Probing Single-Molecule Dynamics in Self-Assembling Viral Nucleocapsids

Thomas Bugea, Roméo Suss, Laetitia Gargowitsch, Charles Truong, Karen Perronet, Guillaume Tresset

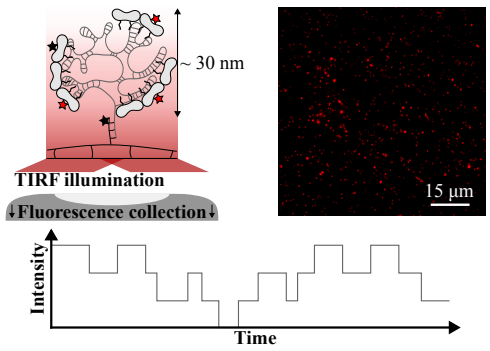
Probing Single-Molecule Dynamics in Self-Assembling Viral Nucleocapsids

Thomas Bugea,[†,‡] Roméo Suss,[†,‡] Laetitia Gargowitsch,[†] Charles Truong,[¶] Karen Perronet,[*,‡] and Guillaume Tresset[*,†]

†*Université Paris-Saclay, CNRS, Laboratoire de Physique des Solides, 91405 Orsay, France*
‡*Université Paris-Saclay, ENS Paris-Saclay, CNRS, CentraleSupélec, LuMIn, 91190 Gif-sur-Yvette, France*
¶*Université Paris-Saclay, Université Paris Cité, ENS Paris-Saclay, CNRS, SSA, INSERM, Centre Borelli, 91190 Gif-sur-Yvette, France*
doi: 10.1021/acs.nanolett.4c04458.

E-mail: karen.perronet@ens-paris-saclay.fr; guillaume.tresset@universite-paris-saclay.fr

Abstract



All viruses on Earth rely on host cell machinery for replication, a process that involves a complex self-assembly mechanism. Our aim here is to scrutinize in real time the growth of icosahedral viral nucleocapsids with single-molecule precision. Using total internal reflection fluorescence microscopy, we probed the binding and unbinding dynamics of fluorescently-labeled capsid subunits on hundreds of immobilized viral RNA molecules simultaneously at each time point. A step-detection algorithm, combined with statistical analysis, allowed us to estimate microscopic quantities such as the equilibrium binding rate and mean residence time, which are otherwise inaccessible through traditional ensemble-averaging techniques. Additionally, we could estimate a set of rate constants modeling the growth kinetics from nonequilibrium measurements, and we observed an acceleration in growth caused by the electrostatic screening effect of monovalent salts. Single-molecule fluorescence imaging will be crucial for elucidating virus self-assembly at the molecular level, particularly in crowded, cell-like environments.

Keywords

Self-assembly, virus, fluorescence microscopy, single-molecule.

The recent pandemic recalls in a vivid manner that viruses can cause significant global health and economic crises. Incapable of replicating on their own, viruses hijack the host cell's machinery to produce large quantities of their components, which then self-assemble into new, highly ordered virus particles. These particles go on to infect other cells, potentially leading to cell death. Viruses are made up of genetic material – either DNA or RNA – encased in one or more protein shells known as capsids, and sometimes enveloped in a lipid membrane. Remarkably, despite the dense and complex nature of the cytoplasmic environment, these newly synthesized molecular components manage to

self-assemble into infectious viral particles with near atomic-scale precision. Furthermore, the survival of a virus depends on its ability to package its own genome, even though the cell's interior is filled with a multitude of nucleic acids essential for cellular functions. Therefore, understanding the complexity of virus self-assembly requires integrated, quantitative approaches that lie at the intersection of virology, statistical physics, and nanotechnology.[1,2]

The simplest spherical viruses co-assemble their capsids, which are made up of multiple copies of the same polypeptide chain, together with their genome, which consists of one or more segments of single-stranded (ss)RNA. An overview of the other mechanisms related to various viral systems can be found elsewhere.[1] Molecular dynamics simulations have revealed two main assembly pathways.[3] In the first pathway, RNA stabilizes interactions between capsid subunits, leading to shell formation through a nucleation and growth process. In the second pathway, subunits rapidly bind to the RNA, creating an amorphous nucleoprotein complex that later anneals into an ordered capsid via a disorder-to-order transition.[4] Various experimental studies using ensemble-averaging techniques, including static light scattering[5–7] and time-resolved small-angle X-ray scattering (TR-SAXS),[8–15] have been conducted to understand the nature of short- and long-lived intermediates during the assembly and disassembly of empty capsids. However, fewer studies have focused on the presence of nucleic acids, even though TR-SAXS has allowed the observation of both of the previously mentioned pathways under varying physicochemical conditions.[16–18]

Ensemble-averaging techniques typically offer high spatiotemporal resolution for probing processes but fall short in revealing the statistical variability of observables across different nano-objects. In contrast, single-molecule and single-particle methods allow for monitoring processes in individual nano-objects, though this often comes with reduced resolution. For instance, mass spectrometry,[19,20] atomic force microscopy,[21] and resistive-pulse sensing[22–24] have showcased the wide range of metastable intermediates associated with icosahedral empty capsids. High-speed atomic force microscopy has resolved the contacts between capsid subunits on a substrate in real time and at the molecular level,[25,26] while optical tweezers[26] have quantitatively measured the forces involved in nucleic acid packaging during co-assembly. Fluorescence correlation spectroscopy[27] has been used to track the growth of nucleocapsids, and cryotransmission electron microscopy has elucidated the nanometer-scale morphology of end-products, highlighting specific defects in capsid symmetry.[28–30] Recently, Garmann and coworkers developed an interferometric scattering microscopy-based[31] single-particle assay that can simultaneously monitor the formation of multiple nucleocapsids in real time.[32,33] They observed a notable lag time in particle growth at weak subunit-RNA interactions, providing direct evidence for a nucleation step. This lag time disappeared when the interactions between subunits and RNA were stronger, consistently with a reduction in the energy barrier.

Our objective was to go beyond this technique by probing the self-assembly dynamics at the level of individual subunits, a resolution that interferometric scattering microscopy could not achieve due to its insufficient signal-to-noise ratio. To accomplish this, we implemented total internal reflection fluorescence microscopy (TIRFM), enabling us to detect in real time the binding and unbinding events of individual fluorescently-labeled subunits simultaneously across multiple RNA molecules.

The selected viral system was the cowpea chlorotic mottle virus (CCMV), a multipartite,[34] non-enveloped, single-stranded (ss)RNA plant virus of $\sim$28 nm in diameter with $T = 3$ icosahedral symmetry.[35] Four genomic RNA segments are distributed across three distinct viral particles in such a way that each one contains approximately 3,000 nucleotides. When purified capsid proteins in their dimeric form – referred to as subunits – are mixed with a genomic RNA segment at physiological pH, amorphous nucleoprotein complexes are initially formed. These complexes can be progressively ordered either by lowering the pH,[32,36]

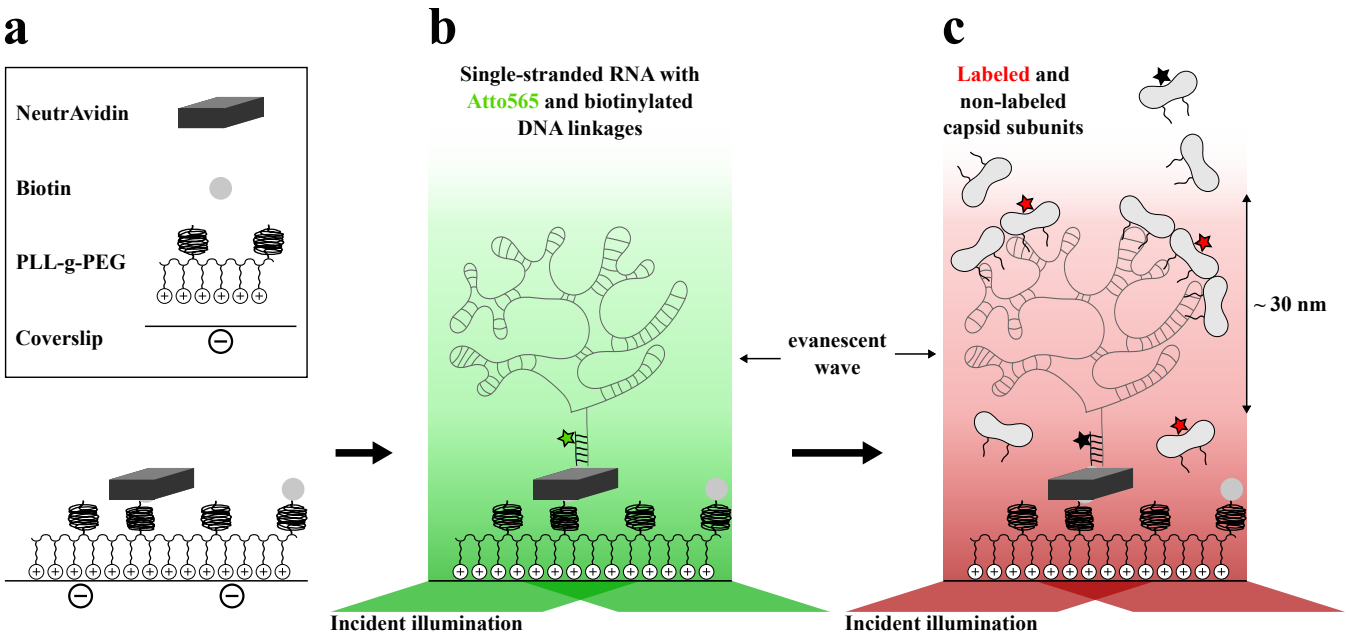


Figure 1: (a) Schematic representation of the coverslip functionalization. (b) Binding of green-labeled single-stranded RNA to the surface and TIRF illumination in the green channel. (c) Injection of non-labeled and red-labeled capsid subunits and TIRF illumination in the red channel.

which decreases the net charge of the subunits,[37] or by increasing the subunit concentration, or equivalently, the chemical potential.[4] In all cases, amorphous nucleoprotein complexes coexist with nearly complete nucleocapsids through a phase separation process reminiscent of disproportionation.[38] All our single-molecule experiments were conducted using CCMV subunits, and we selected the second genomic RNA segment (C2; 2,767 nucleotides) for twofold reasons: (i) it is representative of the genomic content of native CCMV particles in terms of number of nucleotides; and (ii) some of us used it in the past to investigate the role of RNA compactness in packaging process.[28,39] We set pH to 6.1 to optimize surface passivation (see below) and confirmed that at submicromolar concentrations of subunits, particles with masses close to that of native virions were formed, alongside a broad collection of nucleoprotein complexes (see Fig. S1 in the Supporting Information).

The single-molecule experiments were conducted in custom-built microfluidic cells which included a glass coverslip at the bottom (Fig. S2 in SI). The first step involved developing a surface functionalization strategy that allowed us to immobilize fluorescently-labeled RNA while preventing non-specific interactions between the subunits and the coverslip. We hypothesized that these interactions were primarily electrostatic: the N-terminus of a capsid protein contains ten positively-charged amino acids,[40] while the glass surface is negatively charged at near-neutral pH.[41] We passivated the coverslip with a layer of a random graft copolymer composed of a poly(L-lysine) backbone with poly(ethylene glycol) side chains (PLL-*g*-PEG)[41–44] whom 20 % were modified with biotin to create specific binding sites (Fig. 1a).[45–48] Viral ssRNA was hybridized with two DNA linkers, one of which was biotinylated and strongly attached to the biotinylated PLL-*g*-PEG layer via NeutrAvidin (Fig. S3 in SI). The second DNA linker was labeled with a green fluorophore (ATTO 565). Prior to the addition of subunits, NeutrAvidin and labeled ssRNA were injected, and the positions of the labeled ssRNA were localized using the green channel of the TIRF microscope (Fig. 1b). Subsequently, we introduced into the medium capsid subunits, whom ~10 % were red-labeled (ATTO 647N, see Supporting Information for justification). Finally, a series of

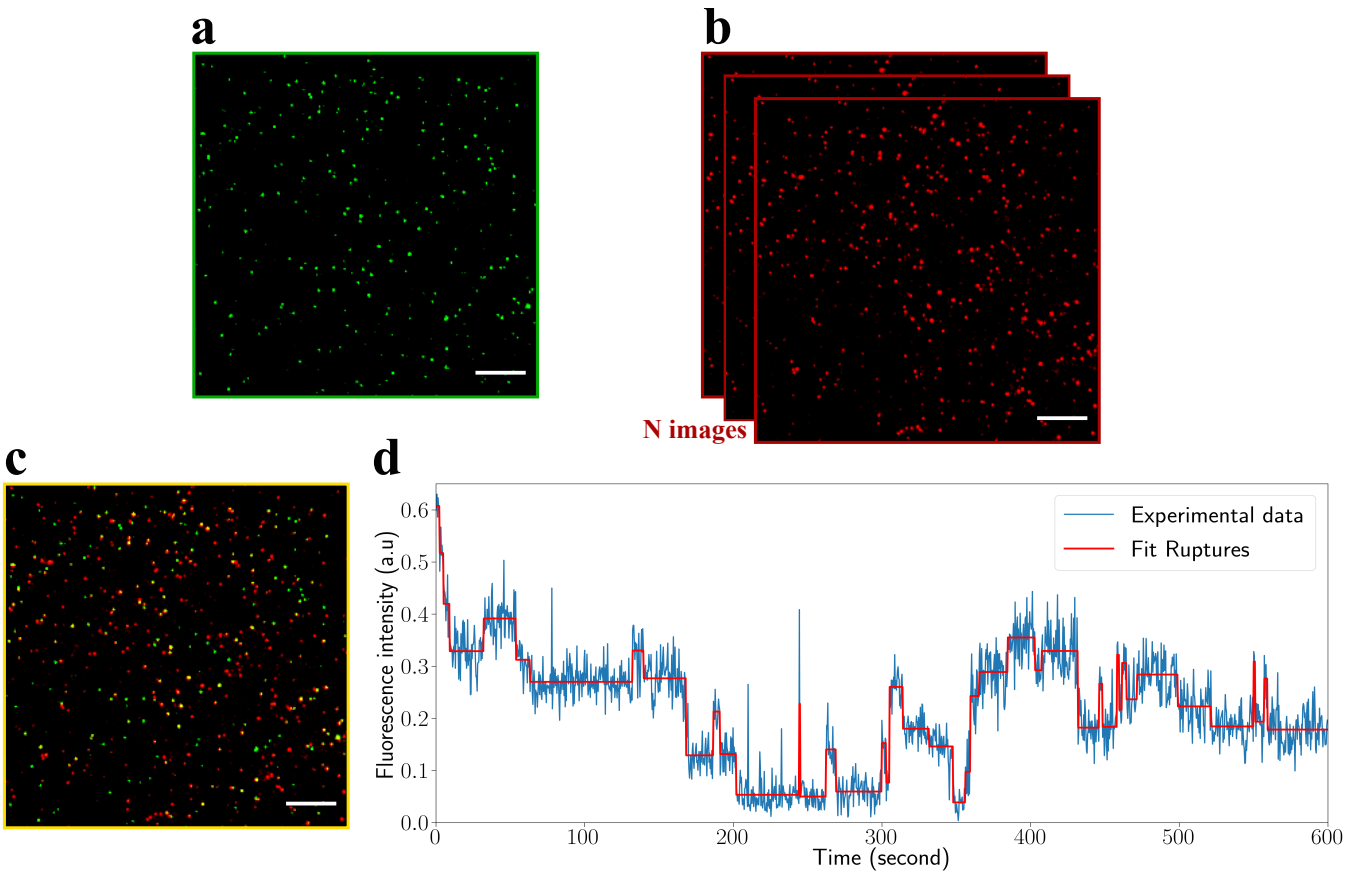


Figure 2: (a) TIRFM image of green-labeled ssRNA. (b) Series of N TIRFM images in the red channel to detect red-labeled capsid subunits. (c) Colocalizations in yellow between labeled RNA and red-labeled subunits. (d) Evolution of the fluorescence intensity emitted by red-labeled capsid subunits colocalized with a RNA molecule. The red trace is a fit with the `Ruptures` algorithm on raw data. Scales bar are 10 µm.

N images were recorded using the red channel (Fig. 1c) with a sampling rate that depends on the type of experiment being conducted.

Examples of fluorescence images showing green-labeled ssRNA and red-labeled subunits are depicted in Figs. 2a and 2b, respectively. The red channel acquisition was recorded after equilibrium was reached – i.e., 15 minutes after subunit injection – during 10 minutes with 0.375 second between each frame, using 25 nM of subunits with a labeling ratio $r = 10$ %. At the beginning, the positions of ssRNA were detected, revealing approximately 1,700 potential assembly sites. Similarly, on the first image in the red channel, we estimated around 3,200 binding sites of subunits. Although some non-specific adsorption of subunits to the glass coverslip was still present, we could accurately identify colocalized spots, highlighted in yellow in Fig. 2c. We followed these subunit-RNA interactions through intensity variations, across the N images recorded in the red channel, which occurred on the RNA positions. An example of the collected temporal traces is shown in Fig. 2d. Control measurements confirmed that the binding and unbinding events observed on temporal traces were predominantly due to interactions between labeled subunits and immobilized ssRNA (see Supporting Information). From the experiment associated to Fig. 2, we observed that approximately 90 % of the RNA spots exhibited at least one binding event of a labeled subunit over the $N = 1{,}600$ acquired images. To detect these spots, we identified intensities in the red step-fit signal that exceeded the mean fluorescence h emitted by a single fluorophore for at least two consecutive images. This reference intensity was deter-

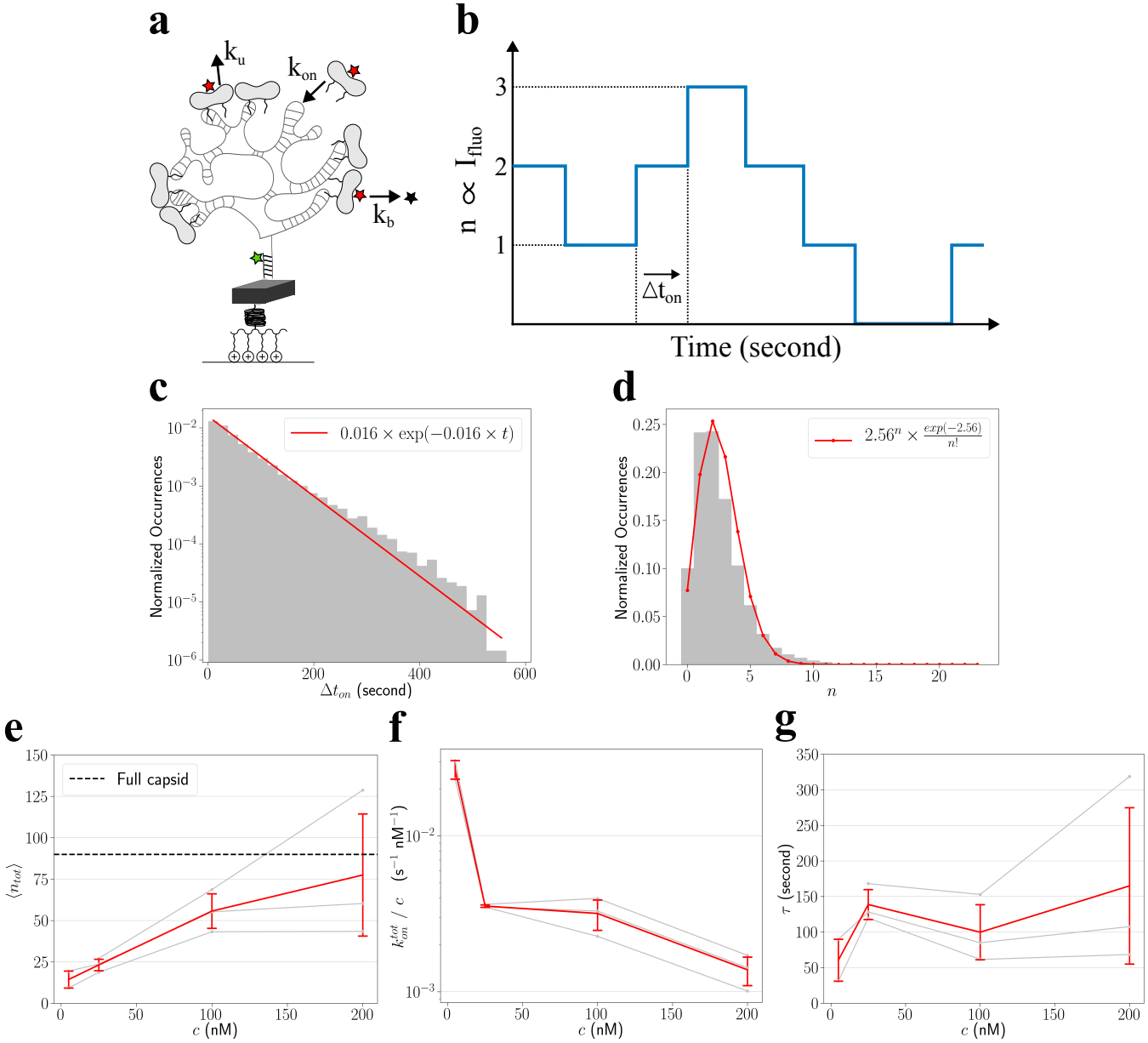


Figure 3: (a) Schematic representation of the single-molecule events occurring within a viral nucleoprotein complex: binding, unbinding and bleaching of fluorescently-labeled subunits. (b) Illustration of the time evolution of the number of bound labeled subunits n on a given RNA molecule. (c) Distribution of Δt_{on} and (d) n with 25 nM of subunits and $r = 9.4$ %. Poisson and exponential laws are plotted respectively in red. (e) Variations of $\langle n_{\text{tot}} \rangle$, (f) $k_{\text{on}}^{\text{tot}}/c$ and (g) τ as a function of subunit concentration. Each grey line was obtained with a different batch of subunits and the red curves represent the average. The labeling ratio is ∼10 % for 5 and 25 nM of subunits, and 2.5 % for 100 and 200 nM.

mined for each experiment by attaching single ATTO 647N molecules to the surface and observing their photobleaching. In this case, we found $h = 0.101$ a.u (Fig. S5a in SI). Finally, stepwise events were identified in noisy fluorescence traces (see e.g. Fig. 2d), by applying a step-detection algorithm called `Ruptures`.[49,50] Detailed explanations on the functionality and performance of the algorithm are provided in the Supporting Information.

Our primary focus was on measuring microscopic quantities that characterize the self-assembly dynamics of viral nucleoprotein complexes but that cannot be captured by

ensemble-averaging methods. As depicted on Fig. 3a, three single-molecule events can occur within a nucleoprotein complex at each time step: binding, unbinding and bleaching of a labeled subunit, with rate constants k_{on}, k_{u}, and k_{b} respectively. Fig. 3b illustrates the time evolution of the number of bound labeled subunits n on a nucleoprotein complex, which can be monitored via fluorescence. Since in our experiments, the binding rate k_{on} was much smaller than the acquisition rate, the time between two consecutive binding events, noted hereafter Δt_{on}, could be readily inferred from step detection. When equilibrium has been reached (Fig. S7 in SI), given that the bulk concentration of subunits remains unchanged, and assuming that k_{on} and k_{u} can be considered as constant, Δt_{on} follows an exponential probability distribution:

$$p(\Delta t_{\mathrm{on}}) = k_{\mathrm{on}} \exp(-k_{\mathrm{on}} \Delta t_{\mathrm{on}}) \qquad (1)$$

with mean value $\langle \Delta t_{\mathrm{on}} \rangle = 1/k_{\mathrm{on}}$ (see Supporting Information for justification). Fig. 3c shows a distribution of Δt_{on} obtained with 25 nM of subunits (r = 9.4 %). By fitting the distribution with an exponential decay function, we estimated Δt_{on} = 63 s and accordingly $k_{\mathrm{on}} = 0.016 \pm 0.001$ s^{-1}. These results were inferred from the analysis of ∼6,600 temporal traces – i.e. RNA spots – recorded on 1,600 images with a time interval of 0.375 second, starting 15 minutes after the injection of subunits. Similarly, the number of labeled subunits obeys a Poisson law (see Supporting Information):

$$p(n) = \frac{\langle n \rangle^n}{n!} e^{-\langle n \rangle} \qquad (2)$$

with $\langle n \rangle = k_{\mathrm{on}}/k_{\mathrm{off}}$, and $k_{\mathrm{off}} = k_{\mathrm{u}} + k_{\mathrm{b}}$ if photobleaching has to be taken into account, $k_{\mathrm{off}} = k_{\mathrm{u}}$ otherwise. Fig. 3d depicts an histogram of the number of labeled subunits on the first recorded image, under the same experimental conditions as in Fig. 3c. Notice that photobleaching did not come into play here since the histogram was built only with the first acquired image. At 25 nM of subunits, $\langle n \rangle$ was estimated to be 2.6 ± 0.2, which led to a mean residence time $\tau = 1/k_{\mathrm{u}}$ of 162 s. In other words, each subunit spent nearly 3 minutes on average attached to a RNA molecule. Moreover, with $r = 9.4 \pm 0.4\%$, we deduced the mean total number of subunits $\langle n_{\mathrm{tot}} \rangle$ per RNA to be 27 ± 2, which is approximately one third of a fully-formed capsid of 90 subunits ($T = 3$). This estimate was consistent with measurements performed by mass photometry at 25 nM of subunits (Fig. S1c in Supporting Information). Finally, we estimated the total binding rate constant of all subunits, noted $k_{\mathrm{on}}^{\mathrm{tot}}$, to be 0.17 ± 0.01 s^{-1} which corresponds to a binding event every 6 s.

Additional measurements carried out with various subunit concentrations in the same aqueous buffer pointed out the critical role of concentration in the growth of nucleoprotein complexes towards icosahedral nucleocapsids. Indeed, Fig. 3e reveals an increase of $\langle n_{\mathrm{tot}} \rangle$ from ∼15 subunits at 5 nM to ∼78 subunits at 200 nM. The binding rate constant $k_{\mathrm{on}}^{\mathrm{tot}}$ divided by the subunit concentration c (in order to normalize $k_{\mathrm{on}}^{\mathrm{tot}}$ with the purely diffusion-limited binding rate constant that scales as c) dropped continuously with increasing c (Fig. 3f) most certainly because remaining binding sites became less accessible as each RNA molecule was increasingly populated with bound subunits. Similarly, the mean residence time τ globally increased with c (Fig. 3g) consistently with the fact that bound subunits were more prone to reside in nucleoprotein complexes with a higher $\langle n_{\mathrm{tot}} \rangle$ since they could make more contacts with each other.

In the following experiment, we demonstrate that this single-molecule methodology can also provide valuable insights into nonequilibrium dynamics. Figure 4a shows the time evolution of probability distributions $p_i(n)$ – the index i refers to as the image number –, each of which being inferred from 452 fluorescence spots at each time step. The assembly experiment was carried out with a subunit concentration of 25 nM (r = 10 %) and a time interval between consecutive images Δt of 2 min, which enabled to neglect photobleaching. The inset depicts $\langle n \rangle$ and reveals that it took more than 15 min to reach equilibrium under these conditions.

Returning to the binding/unbinding model illustrated on Fig. 3a, the probability distribution evolves according to a master equation with rate constants depending on n (see Supporting Information for justification) expressed as

$$p_{i+1}(n) - p_i(n) = p_i(n-1)\Delta t k_{\text{on}}(n-1) - p_i(n)\Delta t k_{\text{on}}(n) - n p_i(n)\Delta t k_{\text{off}}(n) + (n+1)p_i(n+1)\Delta t k_{\text{off}}(n+1). \quad (3)$$

Note that $k_{\text{off}}(n) = k_{\text{u}}(n)$ because the low acquisition rate made photobleaching negligible. This equation applied to the experimental $p_i(n)$ (see Fig. 4a) collected between 0 and 20 min allowed us to estimate $k_{\text{on}}(n)$ and $k_{\text{off}}(n)$ by non-negative least-squares minimization. Figs. 4b and c show the rate constants with their uncertainties determined by bootstrapping the residuals. Noticeably, the two rate constants peaked about the equilibrium value of $\langle n \rangle$ ($\sim$3 in this experiment). It is worth stressing that the rate constants introduced earlier on Fig. 3 are estimates valid only near equilibrium. By contrast, $k_{\text{on}}(n)$ and $k_{\text{off}}(n)$ inferred from $p_i(n)$ give a comprehensive mathematical description of the growth kinetics through master equations (Eq. 3), even though they suffer from high uncertainties at large n because $p_i(n)$ vanishes.

Alongside equilibrium experiments, we adapted slightly the acquisition conditions to study the initial steps of capsids assembly. Here, a 25-minute acquisition (1,000 images, time interval of 1.5 seconds) started immediately at the beginning of the subunit injection (32.5 nM, $r = 8$ %). On Fig. 5a, the fluorescence intensity emitted by labeled subunits attached to a single RNA reached its maximum I_{max} around 10 minutes after subunit injection in a buffer with 10 mM NaCl and a pH of 6.1. Then, it decreased until 25 minutes which could be explained either by red-labeled subunits bleaching on the ssRNA or unbinding from the latter. After 28 minutes, subunits were newly injected (black arrow) and the 10-minute acquisitions recorded did not highlight numerous additional binding events which corroborate

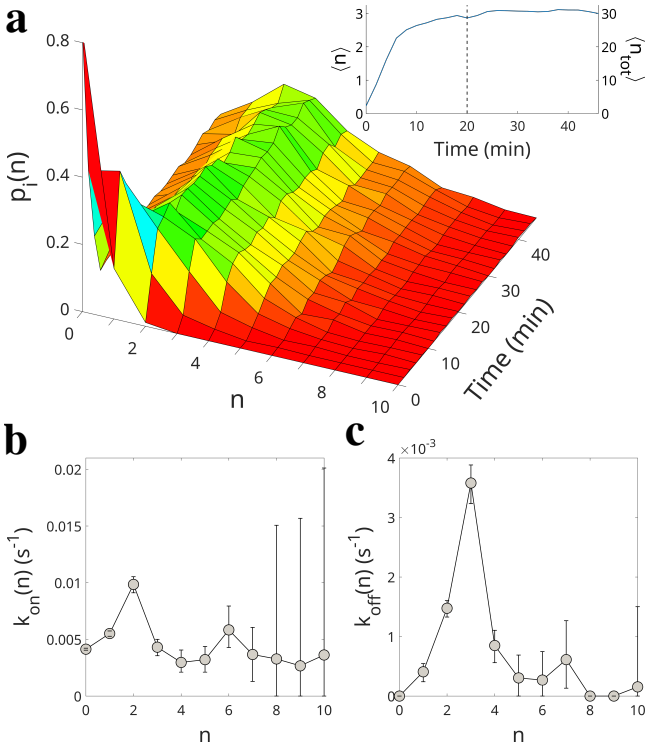


Figure 4: (a) Experimental probability distributions $p_i(n)$ as a function of time for an assembly experiment performed with 25 nM of subunits and a labeling ratio $r = 10$ %. The inset shows the mean value $\langle n \rangle$ versus time as well as the corresponding mean total number of subunits $\langle n_{\text{tot}} \rangle = \langle n \rangle / r$. (b) Binding and (c) unbinding rate constants versus n estimated from $p_i(n)$ between 0 and 20 minutes (vertical dashed line in the inset of (a)). Error bars represent 68 % confidence intervals.

the stability of the nucleoprotein complexes studied previously at equilibrium. This whole behavior was also observed in 100 and 200 mM NaCl but we used to notice a maximum of fluorescence intensity reached sooner (Fig. S8 in SI). We collected the time at which I_{max} was obtained for each RNA considered and a shift towards small time values appeared when the concentration of salt increased (Fig. 5b). We computed the median time t_{median} averaged over several experiments at three different ionic strengths (Fig. 5c). The continuous decrease we observed, was inferred from the analysis of respectively $\sim$1,600 - 1800 and 300 signals.

The results derived from nonequilibrium experiments indicated that binding events on ssRNA were eased with a greater salt concen-

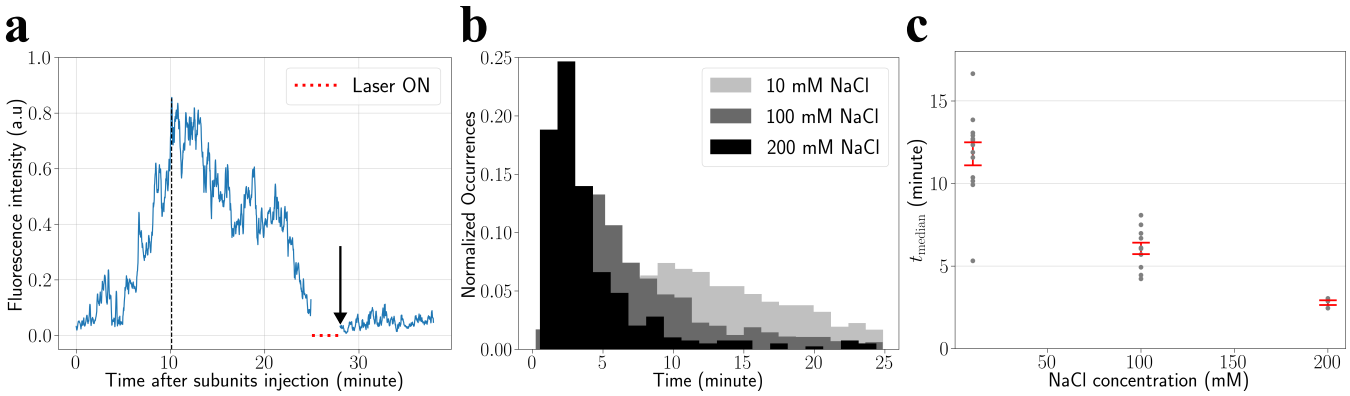


Figure 5: (a) Evolution of fluorescence intensity emitted by labeled subunits from one RNA molecule in a buffer with 10 mM NaCl. The dashed black lines show when I_{max} is reached, the dotted red lines indicate a continuous illumination for 3 minutes and the black arrow correspond to a second injection of CCMV subunits. (b) Distribution of t_{median} for a salt concentration of 10 mM (light grey), 100 mM (dark grey) and 200 mM (black). (c) Evolution of t_{median} with the concentration of salt. Each grey disk was obtained by analyzing the results coming from one microfluidic cell. The subunit concentration is 32.5 nM ($r = 8$ %) and the pH is set to 6.1.

tration. Nevertheless, seeing this evolution for salt concentration higher than the physiological one, raised some questions on how the interactions between the genomic material and subunits were affected. If we consider two simple species, respectively with positive and negative charges, higher ionic strength should reduce their interaction. Such behavior has been demonstrated in brome mosaic virus (BMV), another ssRNA plant virus with $T = 3$ icosahedral symmetry, where the delay before assembly onset around RNA increased at higher salt concentrations for a pH of 6.1.[33] However, despite the 70 % homology between their amino-acid sequence, those two viruses do not carry the same charge. Indeed, while a CCMV subunit has an isoelectric point (pI) of 4.8, the one of BMV is closer to neutrality with 6.5.[51] Therefore, in an aqueous buffer where the pH is 6.1, the capsid protein of BMV is slightly positive when it is negative for CCMV. This difference in charges between CCMV and BMV subunits may have a key role in the way subunit-RNA interactions are affected when ionic strength is modified. Future experiments with different subunit and salt concentrations will be performed to complete the evolution described by Fig. 5c.

In summary, we have shown that TIRFM can accurately probe in real time binding and unbinding events of capsid subunits on viral ssRNA at the single-molecule level. Contrary to ensemble-averaging techniques often used to probe the self-assembly of viral nucleocapsids, we have access to the variability of the number of bound labeled subunits across hundreds to thousands of assembly sites at each time step. After data processing and statistical analysis, microscopic quantities such as binding rate, mean residence time, and rate constants can be deduced, thus shedding light on the equilibrium dynamics and growth kinetics.

The main limitation of our methodology stems from the concentration of labeled subunits. The latter has to be low enough so that the background fluorescence does not outshine the signals of interest. Accordingly, we lowered the labeling ratio for high subunit concentrations at the expense of the resolution on the number of bound labeled subunits. This issue could be lifted by using labeled subunits emitting fluorescence in two or more different colors. Another option could be to reduce further the fluorescence background by imaging the nucleoprotein complexes in zero-mode waveguide nanowells.[52] Besides, combining TIRFM with interferometric scattering microscopy[31] could be a way to refine the size of nucleoprotein complexes estimated by fluorescence when the la-

beling ratio is low.

Single-molecule fluorescence imaging can certainly help decipher the lifecycle of many viruses with unprecedented resolution. We envision that this powerful technique will be essential for future *in vitro* investigations aiming to reconstitute the assembly and disassembly processes occurring within the intracellular milieu, or to assess packaging affinity between competing RNA templates, and for which the viral components will have to be distinguished from a vast majority of other biomolecules. Deep insights into these processes are not only crucial to combat viral infections, but can also be useful to design synthetic, highly-symmetric nanocages with antiviral properties or for delivery applications.[53–55]

Acknowledgement This work received a financial support from Investissements d'Avenir through LabEx PALM (ANR-10-LABX-0039-PALM) and LabEx NANOSACLAY (ANR-10-LABX-0035-NANOSACLAY) and has been supported as part of France 2030 program "ANR-11-IDEX-0003".

Supporting Information Available

Details of sample preparation, assembly experiments, TIRF microscopy, image processing and data analysis.

References

(1) Bruinsma, R. F.; Wuite, G. J. L.; Roos, W. H. Physics of viral dynamics. *Nat. Rev. Phys.* **2021**, *3*, 76–91.

(2) Zandi, R.; Dragnea, B.; Travesset, A.; Podgornik, R. On virus growth and form. *Phys. Rep.* **2020**, *847*, 1–102.

(3) Elrad, O. M.; Hagan, M. F. Encapsulation of a polymer by an icosahedral virus. *Phys. Biol.* **2010**, *7*, 045003.

(4) Panahandeh, S.; Li, S.; Marichal, L.; Leite Rubim, R.; Tresset, G.; Zandi, R. How a Virus Circumvents Energy Barriers to Form Symmetric Shells. *ACS Nano* **2020**, *14*, 3170–3180.

(5) Zlotnick, A.; Johnson, J. M.; Wingfield, P. W.; Stahl, S. J.; Endres, D. A theoretical model successfully identifies features of hepatitis B virus capsid assembly. *Biochemistry* **1999**, *38*, 14644–14652.

(6) Zlotnick, A.; Aldrich, R.; Johnson, J. M.; Ceres, P.; Young, M. J. Mechanism of capsid assembly for an icosahedral plant virus. *Virology* **2000**, *277*, 450–456.

(7) Johnson, J. M.; Tang, J.; Nyame, Y.; Willits, D.; Young, M. J.; Zlotnick, A. Regulating self-assembly of spherical oligomers. *Nano Lett.* **2005**, *5*, 765–770.

(8) Tuma, R.; Tsuruta, H.; French, K. H.; Prevelige, P. E. Detection of intermediates and kinetic control during assembly of bacteriophage P22 procapsid. *J. Mol. Biol.* **2008**, *381*, 1395–1406.

(9) Tresset, G.; Le Coeur, C.; Bryche, J.-F.; Tatou, M.; Zeghal, M.; Charpilienne, A.; Poncet, D.; Constantin, D.; Bressanelli, S. Norovirus capsid proteins self-assemble through biphasic kinetics via long-lived stave-like intermediates. *J. Am. Chem. Soc.* **2013**, *135*, 15373–15381.

(10) Law-Hine, D.; Sahoo, A. K.; Bailleux, V.; Mehdi Zeghal; Prevost, S.; Maiti, P. K.; Bressanelli, S.; Constantin, D.; Tresset, G. Reconstruction of the disassembly pathway of an icosahedral viral capsid and shape determination of two successive intermediates. *J. Phys. Chem. Lett.* **2015**, *6*, 3471–3476.

(11) Law-Hine, D.; Zeghal, M.; Bressanelli, S.; Constantin, D.; Tresset, G. Identification of a major intermediate along the self-assembly pathway of an icosahedral viral capsid by using an analytical model of a spherical patch. *Soft Matter* **2016**, *12*, 6728–6736.

(12) Asor, R.; Selzer, L.; Schlicksup, C. J.; Zhao, Z.; Zlotnick, A.; Raviv, U. Assembly Reactions of Hepatitis B Capsid Protein into Capsid Nanoparticles Follow a Narrow Path through a Complex Reaction Landscape. *ACS Nano* **2019**, *13*, 7610–7626.

(13) Asor, R.; Schlicksup, C. J.; Zhao, Z.; Zlotnick, A.; Raviv, U. Rapidly Forming Early Intermediate Structures Dictate the Pathway of Capsid Assembly. *J. Am. Chem. Soc.* **2020**, *142*, 7868–7882.

(14) Chevreuil, M.; Lecoq, L.; Wang, S.; Gargowitsch, L.; Nhiri, N.; Jacquet, E.; Zinn, T.; Fieulaine, S.; Bressanelli, S.; Tresset, G. Nonsymmetrical dynamics of the HBV capsid assembly and disassembly evidenced by their transient species. *J. Phys. Chem. B* **2020**, *124*, 9987–9995.

(15) Kra, K.; Li, S.; Gargowitsch, L.; Degrouard, J.; Pérez, J.; Zandi, R.; Bressanelli, S.; Tresset, G. Energetics and Kinetic Assembly Pathways of Hepatitis B Virus Capsids in the Presence of Antivirals. *ACS Nano* **2023**, *17*, 12723–12733.

(16) Kler, S.; Asor, R.; Li, C.; Ginsburg, A.; Harries, D.; Oppenheim, A.; Zlotnick, A.; Raviv, U. RNA encapsidation by SV40-derived nanoparticles follows a rapid two-state mechanism. *J. Am. Chem. Soc.* **2012**, *134*, 8823–8830.

(17) Chevreuil, M.; Law-Hine, D.; Chen, J.; Bressanelli, S.; Combet, S.; Constantin, D.; Degrouard, J.; Möller, J.; Zeghal, M.; Tresset, G. Nonequilibrium self-assembly dynamics of icosahedral viral capsids packaging genome or polyelectrolyte. *Nat. Commun.* **2018**, *9*, 3071.

(18) Oliver, R. C.; Potrzebowski, W.; Najibi, S. M.; Pedersen, M. N.; Arleth, L.; Mahmoudi, N.; André, I. Assembly of Capsids from Hepatitis B Virus Core Protein Progresses through Highly Populated Intermediates in the Presence and Absence of RNA. *ACS Nano* **2020**, *14*, 10226–10238.

(19) Uetrecht, C.; Barbu, I. M.; Shoemaker, G. K.; van Duijn, E.; Heck, A. J. R. Interrogating viral capsid assembly with ion mobility-mass spectrometry. *Nat. Chem.* **2011**, *3*, 126–132.

(20) Lutomski, C. A.; Lyktey, N. A.; Pierson, E. E.; Zhao, Z.; Zlotnick, A.; Jarrold, M. F. Multiple pathways in capsid assembly. *J. Am. Chem. Soc.* **2018**, *140*, 5784–5790.

(21) Medrano, M.; Fuertes, M. A.; Valbuena, A.; Carrillo, P. J. P.; Rodriguez-Huete, A.; Mateu, M. G. Imaging and quantitation of a succession of transient intermediates reveal the reversible self-assembly pathway of a simple icosahedral virus capsid. *J. Am. Chem. Soc.* **2016**, *138*, 15385–15396.

(22) Harms, Z. D.; Selzer, L.; Zlotnick, A.; Jacobson, S. C. Monitoring Assembly of Virus Capsids with Nanofluidic Devices. *ACS Nano* **2015**, *9*, 9087–9096.

(23) Kondylis, P.; Zhou, J.; Harms, Z. D.; Kneller, A. R.; Lee, L. S.; Zlotnick, A.; Jacobson, S. C. Nanofluidic Devices with 8 Pores in Series for Real-Time, Resistive-Pulse Analysis of Hepatitis B Virus Capsid Assembly. *Anal. Chem.* **2017**, *89*, 4855–4862.

(24) Starr, C. A.; Nair, S.; Huang, S.-Y.; Hagan, M. F.; Jacobson, S. C.; Zlotnick, A. Engineering Metastability into a Virus-like Particle to Enable Triggered Dissociation. *J. Am. Chem. Soc.* **2023**, *145*, 2322–2331.

(25) Valbuena, A.; Maity, S.; Mateu, M. G.; Roos, W. H. Visualization of Single Molecules Building a Viral Capsid Protein Lattice through Stochastic Pathways. *ACS Nano* **2020**, *14*, 8724–8734.

(26) Buzón, P.; Maity, S.; Christodoulis, P.; Wiertsema, M. J.; Dunkelbarger, S.; Kim, C.; Wuite, G. J.; Zlotnick, A.; Roos, W. H. Virus self-assembly proceeds through contact-rich energy minima. *Sci. Adv.* **2021**, *7*, eabg0811.

(27) Patel, N.; Dykeman, E. C.; Coutts, R. H. A.; Lomonossoff, G. P.; Rowlands, D. J.; Phillips, S. E. V.; Ranson, N.; Twarock, R.; Tuma, R.; Stockley, P. G. Revealing the density of encoded functions in a viral RNA. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 2227–2232.

(28) Marichal, L.; Gargowitsch, L.; Rubim, R. L.; Sizun, C.; Kra, K.; Bressanelli, S.; Dong, Y.; Panahandeh, S.; Zandi, R.; Tresset, G. Relationships between RNA topology and nucleocapsid structure in a model icosahedral virus. *Biophys. J.* **2021**, *120*, 3925–3936.

(29) Bond, K.; Tsvetkova, I. B.; Wang, J. C.-Y.; Jarrold, M. F.; Dragnea, B. Virus Assembly Pathways: Straying Away but Not Too Far. *Small* **2020**, *16*, 2004475.

(30) Ruszkowski, M.; Strugala, A.; Indyka, P.; Tresset, G.; Figlerowicz, M.; Urbanowicz, A. Cryo-EM reconstructions of BMV-derived virus-like particles reveal assembly defects in the icosahedral lattice structure. *Nanoscale* **2022**, *14*, 3224–3233.

(31) Ortega Arroyo, J.; Cole, D.; Kukura, P. Interferometric scattering microscopy and its combination with single-molecule fluorescence imaging. *Nat. Protoc.* **2016**, *11*, 617–633.

(32) Garmann, R. F.; Goldfain, A. M.; Manoharan, V. N. Measurements of the self-assembly kinetics of individual viral capsids around their RNA genome. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116*, 22485–22490.

(33) Garmann, R. F.; Goldfain, A. M.; Tanimoto, C. R.; Beren, C. E.; Vasquez, F. F.; Villarreal, D. A.; Knobler, C. M.; Gelbart, W. M.; Manoharan, V. N. Single-particle studies of the effects of RNA–protein interactions on the self-assembly of RNA virus particles. *Proc. Natl. Acad. Sci. U.S.A.* **2022**, *119*, e2206292119.

(34) van de Waterbeemd, M.; Snijder, J.; Tsvetkova, I. B.; Dragnea, B. G.; Cornelissen, J. J.; Heck, A. J. R. Examining the Heterogeneous Genome Content of Multipartite Viruses BMV and CCMV by Native Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2016**, *27*, 1000–1009.

(35) Speir, J. A.; Munshi, S.; Wang, G.; Baker, T. S.; Johnson, J. E. Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by X-ray crystallography and cryo-electron microscopy. *Structure* **1995**, *3*, 63–78.

(36) Garmann, R. F.; Comas-Garcia, M.; Gopal, A.; Knobler, C. M.; Gelbart, W. M. The assembly pathway of an icosahedral single-stranded RNA virus depends on the strength of inter-subunit attractions. *J. Mol. Biol.* **2014**, *426*, 1050–1060.

(37) Tresset, G.; Chen, J.; Chevreuil, M.; Nhiri, N.; Jacquet, E.; Lansac, Y. Two-Dimensional Phase Transition of Viral Capsid Gives Insights into Subunit Interactions. *Phys. Rev. Applied* **2017**, *7*, 014005.

(38) Bruinsma, R. F.; Comas-Garcia, M.; Garmann, R. F.; Grosberg, A. Y. Equilibrium self-assembly of small RNA viruses. *Phys. Rev. E* **2016**, *93*, 032405.

(39) Tresset, G.; Li, S.; Gargowitsch, L.; Matthews, L.; Pérez, J.; Zandi, R. Glass-like Relaxation Dynamics during the Disorder-to-Order Transition of Viral Nucleocapsids. *J. Phys. Chem. Lett.* **2024**, *15*, 10210–10218.

(40) Vriend, G.; Verduin, B. J. M.; Hemminga, M. A. Role of the N-terminal part of the coat protein in the assembly of

cowpea chlorotic mottle virus: A 500 MHz proton nuclear magnetic resonance study and structural calculations. *Journal of Molecular Biology* **1986**, *191*, 453–460.

(41) Perry, S. S.; Yan, X.; Limpoco, F. T.; Lee, S.; Müller, M.; Spencer, N. D. Tribological Properties of Poly(l-lysine)-graft-poly(ethylene glycol) Films: Influence of Polymer Architecture and Adsorbed Conformation. *ACS Appl. Mater. Interfaces* **2009**, *1*, 1224–1230.

(42) Huang, N.-P.; Michel, R.; Voros, J.; Textor, M.; Hofer, R.; Rossi, A.; Elbert, D. L.; Hubbell, J. A.; Spencer, N. D. Poly(l-lysine)-g-poly(ethylene glycol) Layers on Metal Oxide Surfaces: Surface-Analytical Characterization and Resistance to Serum and Fibrinogen Adsorption. *Langmuir* **2001**, *17*, 489–498.

(43) Pasche, S.; De Paul, S. M.; Vörös, J.; Spencer, N. D.; Textor, M. Poly(l-lysine)-graft-poly(ethylene glycol) Assembled Monolayers on Niobium Oxide Surfaces: A Quantitative Study of the Influence of Polymer Interfacial Architecture on Resistance to Protein Adsorption by ToF-SIMS and in Situ OWLS. *Langmuir* **2003**, *19*, 9216–9225.

(44) Michel, R.; Pasche, S.; Textor, M.; Castner, D. G. Influence of PEG Architecture on Protein Adsorption and Conformation. *Langmuir* **2005**, *21*, 12327–12332.

(45) Huang, N.-P.; Vörös, J.; De Paul, S. M.; Textor, M.; Spencer, N. D. Biotin-Derivatized Poly(l-lysine)-g-poly(ethylene glycol): A Novel Polymeric Interface for Bioaffinity Sensing. *Langmuir* **2002**, *18*, 220–230.

(46) Städler, B.; Falconnet, D.; Pfeiffer, I.; Höök, F.; Vörös, J. Micropatterning of DNA-Tagged Vesicles. *Langmuir* **2004**, *20*, 11348–11354.

(47) Morgenthaler, S.; Zink, C.; Städler, B.; Vörös, J.; Lee, S.; Spencer, N. D.; Tosatti, S. G. P. Poly(l-lysine)-grafted-poly(ethylene glycol)-based surface-chemical gradients. Preparation, characterization, and first applications. *Biointerphases* **2006**, *1*, 156–165.

(48) Gunnarsson, A.; Jönsson, P.; Marie, R.; Tegenfeldt, J. O.; Höök, F. Single-Molecule Detection and Mismatch Discrimination of Unlabeled DNA Targets. *Nano Lett.* **2008**, *8*, 183–188.

(49) Truong, C.; Oudre, L.; Vayatis, N. Selective review of offline change point detection methods. *Signal Processing* **2020**, *167*, 107299.

(50) Truong, C.; Bugea, T.; Bouhet, B.; Tresset, G.; Marquier, F.; Perronet, K. Photobleaching Step Counting and Localization for Fluorescence Microscopy. 2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2024.

(51) Duran-Meza, A. L.; Villagrana-Escareño, M. V.; Ruiz-García, J.; Knobler, C. M.; Gelbart, W. M. Controlling the surface charge of simple viruses. *PLOS ONE* **2021**, *16*, e0255820.

(52) Levene, M. J.; Korlach, J.; Turner, S. W.; Foquet, M.; Craighead, H. G.; Webb, W. W. Zero-Mode Waveguides for Single-Molecule Analysis at High Concentrations. *Science* **2003**, *299*, 682–686.

(53) Cannon, K. A.; Ochoa, J. M.; Yeates, T. O. High-symmetry protein assemblies: patterns and emerging applications. *Curr. Opin. Struct. Biol.* **2019**, *55*, 77–84.

(54) Sigl, C.; Willner, E. M.; Engelen, W.; Kretzmann, J. A.; Sachenbacher, K.; Liedl, A.; Kolbe, F.; Wilsch, F.; Aghvami, S. A.; Protzer, U.; Hagan, M. F.; Fraden, S.; Dietz, H. Programmable icosahedral shell system for virus trapping. *Nat. Mater.* **2021**, *20*, 1281–1289.

(55) Tetter, S.; Terasaka, N.; Steinauer, A.; Bingham, R. J.; Clark, S.; Scott, A. J. P.; Patel, N.; Leibundgut, M.; Wroblewski, E.; Ban, N.; Stockley, P. G.; Twarock, R.; Hilvert, D. Evolution of a virus-like architecture and packaging mechanism in a repurposed bacterial protein. *Science* **2021**, 1220–1224.

Supporting Information: Probing Single-Molecule Dynamics in Self-Assembling Viral Nucleocapsids

Thomas Bugea,[†,‡] Roméo Suss,[†,‡] Laetitia Gargowitsch,[†] Charles Truong,[¶] Karen Perronet,[*,‡] and Guillaume Tresset[*,†]

†*Université Paris-Saclay, CNRS, Laboratoire de Physique des Solides, 91405 Orsay, France*
‡*Université Paris-Saclay, ENS Paris-Saclay, CNRS, CentraleSupélec, LuMIn, 91190 Gif-sur-Yvette, France*
¶*Université Paris-Saclay, Université Paris Cité, ENS Paris-Saclay, CNRS, SSA, INSERM, Centre Borelli, 91190 Gif-sur-Yvette, France*

E-mail: karen.perronet@ens-paris-saclay.fr; guillaume.tresset@universite-paris-saclay.fr

Virion purification

The purification of CCMV virions starts from infected leaves of black-eye cowpea and mainly follows the protocol described by Ali *et al.*[S1] Approximately 140 g of infected leaves are first mixed within 300 mL of 0.15 M sodium acetate at pH 4.8, and the whole mixture is filtered. Then, 300 mL of iced chloroform are added to the solution before being agitated during 1 h. A first step of centrifugation is performed at 10,000×g during 10 min, from which only the supernatant is collected and stirred for 1 h after incorporation of 20 mM NaCl and poly(etylene glycol) at 8 % (molecular weight of 8,000 Da). A second step of centrifugation is executed at 10,000×g for 10 min, and then the pellet is directly dissolved in 21 mL of 50 mM sodium acetate at pH 4.8 and put through agitation for 1 h. Another step of centrifugation is realized at 9,000×g for 10 min and the aqueous upper phase is collected and follows a final step of centrifugation for 2 h at 150,000×g after 20 % (w/v) sucrose cushion is added. The purity of the samples is assessed when the ratio $A_{260}/A_{280} > 1.60$.

Capsid protein purification and labeling

First, 5 mg of purified virions are centrifuged for 10 min at 14,000×g, three times in a row in a 0.5 mL centrifugal filter to change the buffer: 100 mM HEPES and 5 mM $MgCl_2$ at pH 8.2. Then, a specific volume of 2.37 mM ATTO 647N NHS-Ester (*ATTO-TEC*, Germany) is added to the solution and incubated at room temperature for 2 h, to target a labeling ratio r for purified proteins of 25 %. The solution is then dialyzed for 6 to 8 h against the previous buffer to get rid of free fluorescent dyes. Labeled virions are dissociated through another dialysis step of 24 h against 50 mM HEPES, 500 mM $CaCl_2$, 1 mM ethylenediamine tetraacetic acid (EDTA), 1 mM dithiothreithol (DTT) and 0.5 mM phenylmethylsulfonyl fluoride (PMSF) at pH 7.5. After 18 h of centrifugation at 150,000×g, several 1-mL aliquots taken from the supernatant are collected and conserved if $A_{260}/A_{280} > 0.80$. The selected aliquots are centrifuged two to three times for 15 min at 4,000×g to obtain a total volume between 1 and 2 mL. A final dialysis is performed overnight against 50 mM HEPES, 500 mM NaCl, 1 mM EDTA, 1 mM DTT and

0.5 mM PMSF at pH 7.5 and then the concentration of capsid proteins C_p and the labeling ratio r are estimated with the following equations:

$$C_p = \frac{A_{280} - A_{647} \times CF_{280}}{\epsilon_p(280)}$$
$$r = \frac{A_{647}}{\epsilon_{d,max} \times C_p}$$

where $CF_{280} = \epsilon_{280}/\epsilon_{d,max}$ stands for a correction coefficient, $\epsilon_{d,max}$ corresponds to the molar extinction coefficient at the longest-wavelength absorption maximum (here 647 nm) and ϵ_{280} is the molar extinction coefficient for CCMV capsid protein at 280 nm. At the end of the purification, we usually have a final labeling ratio of ∼10 %.

For the non-labeled subunits, the purification starts directly with the dissociation dialysis step of 24 h.

Mass photometry experiments

To assess the possibility of forming full nucleocapsids with a mass close to the one of native CCMV virions, mass photometry experiments were performed at Institut Pasteur on the `TwoMP` instrument of *Refeyn* (UK). This technique is based on interferometric scattering (iSCAT) microscopy, in which interferences between a reference field, here reflected by a glass coverslip, and the one scattered by nanoparticles are detected.[S2] The signal of interest here is the contrast of the interferences which is directly proportional to the volume of the particles. Moreover, the linear relationship between the contrast and the volume of the detected objects means that the signal obtained in mass photometry is also proportional to their molar mass, provided that the refractive index is similar for the different particles considered.[S3] Therefore, to characterize the mass of unknown particles, it is important to calibrate the device with known particles. In our case, we detected the first four oligomers of Bovine Serum Albumin – i.e. BSA– and urease.

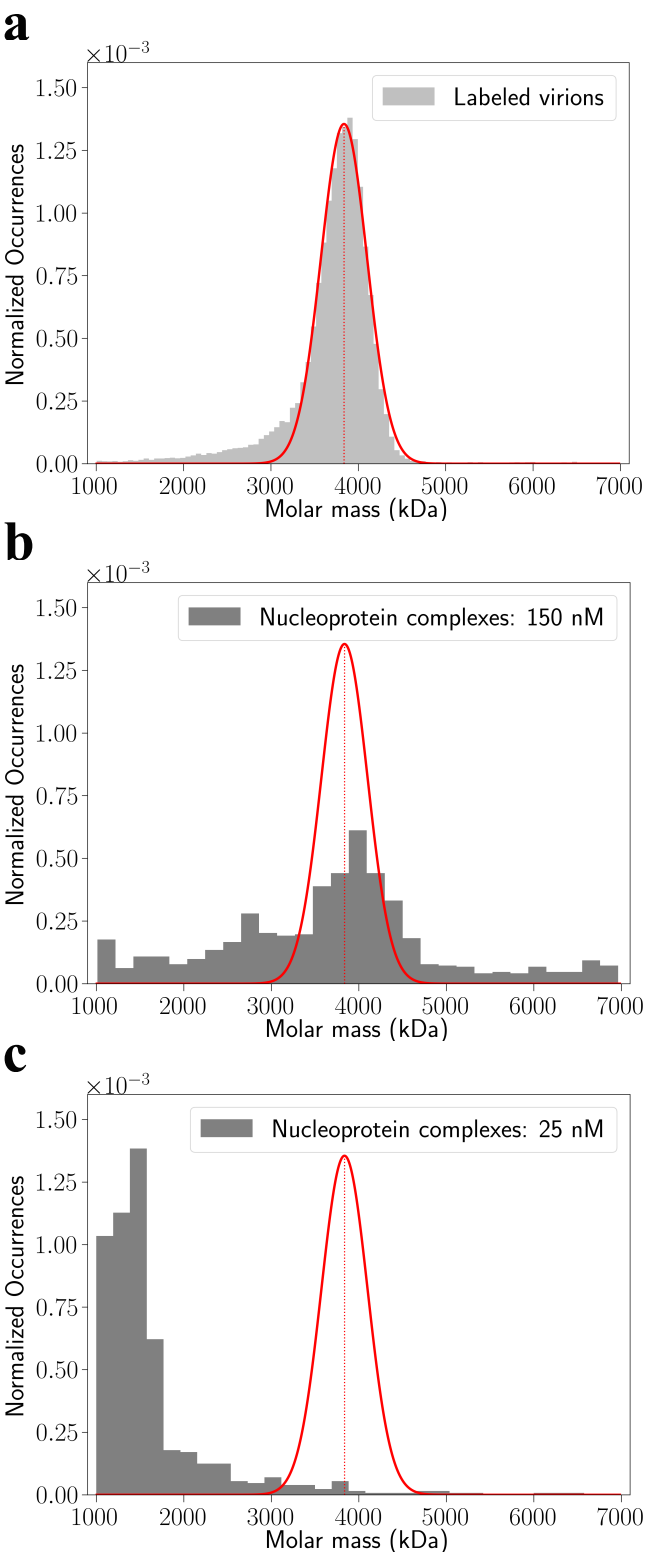


Figure S1: (a) Mass distribution of labeled CCMV virions and (b-c) nucleoprotein complexes. For (b) 150 nM of subunits ($r = 14.2$ %) were mixed with 1 nM of RNA C2 whereas in (c) 25 nM of subunits ($r = 14.2$ %) and 5 nM of RNA were associated together. The curves plotted in red correspond to a Gaussian fit.

We first performed a mass photometry experiment on labeled native CCMV virions. The corresponding distribution is illustrated on Fig. S1a and the mean molar mass is close to 3.9 MDa. One of the reasons that could explain the slight discrepancy with the expected value of 4.6 MDa may come from the calibration carried out solely with proteins and not with RNA, which is in the core of native virions. Nevertheless, this measurement gave us an estimate of the molar mass to reach if complete nucleocapsids were formed. On Fig. S1b, we can observe the masses of the nucleoprotein complexes formed by mixing 150 nM of subunits with a labeling ratio of 14.2 % and 1 nM of RNA in assembly buffer (50 mM 2-(N-morpholino)ethanesulfonic acid (MES) and 100 mM NaCl at pH 6.1). We see that the characteristic peak of native virions, plotted in red, is superimposed with a part of the detected population, meaning that complete nucleocapsids could be obtained at sub-micromolar concentration of subunits. Smaller objects corresponding to partially-formed nucleocapsids and bigger objects highlighting the presence of some aggregates were also observed. Finally, Fig. S1c shows that 25 nM of subunits associated with 5 nM of RNA led mostly to partial nucleocapsids.

Microfluidic cell and surface chemistry

All experiments were performed into homemade microfluidic cells composed of three layers: a 170-µm-thick glass coverslip, a thin channel of 1-mm width and 50-mm length laser-cut in double-sided adhesive tape and a microscope slide with two holes for inlet and outlet. A container is glued to the inlet whereas the outlet is linked to a tube connected to a syringe (Fig. S2).

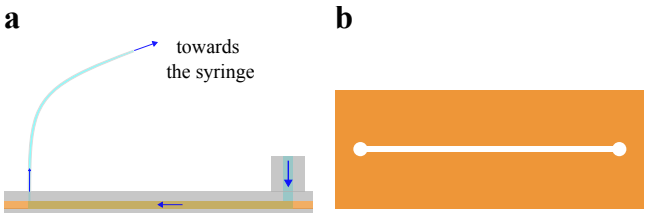


Figure S2: (a) Side view of a microfluidic cell and (b) top view of the tape layer. The blue arrows show the flow direction.

At the beginning of an experiment, the coverslips were activated inside an oxygen-plasma cleaner during 10 min. Then a solution of 1.5 g/L poly(L-lysine)-graft-poly(ethylene glycol) copolymer (PLL(20)-g[3.5]-PEG(2)) and 10^{-4} g/L PLL(20)-g[3.5]-PEG(2)/PEG(3.4)biotin(20 %) in H2 buffer (10 mM HEPES and 150 mM NaCl at pH 7.4) was incubated for 35 min in each microfluidic cell. Once the incubation is done, the cells were rinsed with H2 buffer. Both copolymers were purchased from *SuSoS* (Switzerland). Finally, ∼100 nM NeutrAvidin (*Thermo Scientific*, USA) were incubated for 10 min into each microfluidic cell just before the injection of RNA. The buffer used to prepare NeutrAvidine was the same in which the subunits were prepared (see section Assembly experiments).

Double-hybridization on RNA segments

Assembly experiments were performed with the second single-stranded genomic RNA segment (C2), containing 2,767 nucleotides. C2 5' end was hybridized with a complementary sequence (Oligo-Biot) of 53 bases with a biotin modification on the 3' extremity (*Eurogentec*, Belgium; 5' - ATG-GAG-ACC-CGT-GAG-TCG-ACA-AGG-GAT-TGA-ACC-TCG-CTC-TCG-TGG-ATT-ACA-AA-biotinTEG - 3'). A second hybridization occured, just 5 nucleotides downstream, with a 51-base complementary sequence (Oligo-ATTO) with an ATTO 565 modification at the 5' end (*Eurogentec*; 5' - ATTO 565-AAA-AAG-TAA-AGA-AGA-AAA-ATA-TCA-AGA-AGA-AAA-TGT-ATA-AAC-TGT-TTT-CAA - 3'). The whole architecture is represented on Fig. S3.

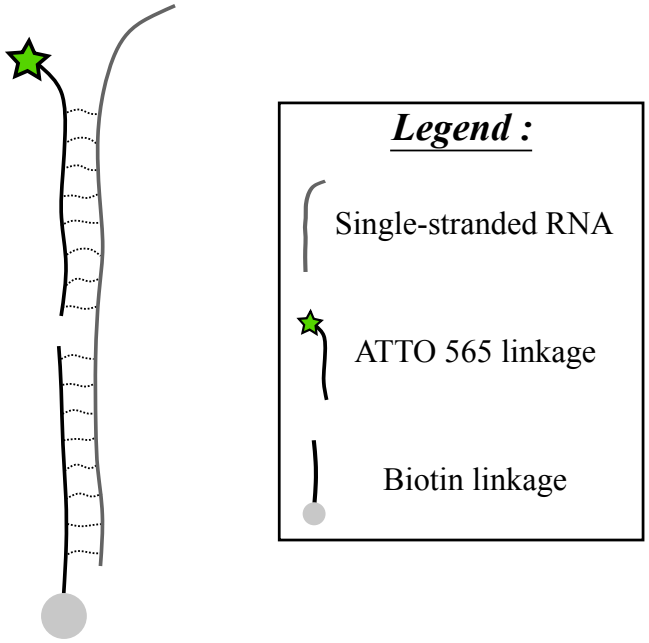


Figure S3: Schematic representation of the double hybridization of RNA C2.

Before each experiment, a 10-µL solution was prepared with the following final composition: 1 µM C2, 3 µM Oligo-ATTO and 1 µM Oligo-Biot in an hybridization buffer (10 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA at pH 7.5). To perform the double-hybridization, the solution was heated for 5 min at 68 °C, for another 5 min at 55 °C and then stored at 4 °C until use.

Total internal reflection fluorescence microscopy (TIRFM)

The TIRFM experiments were performed with a home-made setup described in a slightly different configuration elsewhere.[S4] Briefly, the optical setup can be decomposed into two different channels: red and green according to the two excitation lasers used (638 nm, 06-MLD, *Cobolt*, Sweden, and 561 nm, LCX-561, *Oxxius*, France). Both laser beams are excessively enlarged ($\sim \times 40$), and focused off-axis in the back focal plane of an oil-immersion microscope objective (UPlanApo 60×, NA = 1.50, *Olympus*, Japan), to reach total internal reflexion with a homogeneous illumination of the field of view. The fluorescence emitted by red or green dyes is collected on an Electron-Multiplying Charge-Coupled Device (EMCCD) camera (Ixon 897, *Andor*, Ireland). Each laser beam can be blocked by a digitally controlled shutter that prevents us from exposing the samples continuously. The illumination sequence depends on the purpose of the experiment.

A laboratory-built microscope-based focus drift correction was used to keep the focus during the whole experiment.

Assembly experiments

100 pM hybridized and thus ATTO 565 labeled C2 segments were injected in a functionalized microfluidic cell and imaged in the green channel. Following this first acquisition, the protocol depends on the type of experiment:

- Equilibrium experiment: subunits incubated in the microfluidic cell for 15 minutes. After 14 minutes, a second green image was acquired and then, one minute later, the sequence using the red laser beam started: 1,600 images, 375 ms time interval, and 100 ms exposure time. At the end of this 10-minute film, a third image in the green channel was acquired. For these experiments, we used 5 and 25 nM of subunits ($r \sim$10 %) but also 100 and 200 nM of subunits ($r = 2.5$ %). Here, from the injection of NeutrAvidin to the incubation of CCMV subunits, each step was performed in the following assembly buffer: 50 mM MES and 100 mM NaCl at pH 6.1.

- Towards equilibrium experiments: to highlight thermodynamic equilibrium, subunits were injected in the medium and a 46-minute acquisition (24 images, 2 min time interval, 100 ms exposure time) began in the red channel. At the end, a second movie in the green detection channel was acquired. Here, three concentrations of subunits were used: 25 nM ($r = 10$ %) but also 100 and 200 nM ($r = 2.5$ %). These experiments were performed in the same assembly buffer (50 mM MES and 100 mM NaCl at pH 6.1).

- Nonequilibrium experiments: 32.5 nM of subunits (r = 8 %) were injected and a 25-minute acquisition (1,000 images, 1.5 s time interval, 100 ms exposure time) in the red detection channel started at the same time. Then, the field of view was illuminated continuously for 3 minutes to photobleach each labeled subunit remaining. Finally, a 10-minute movie was acquired in the red channel (400 images, 1.5 s time interval, 100 ms exposure time). Here, three buffers were used with a pH set to 6.1 and a variable salt concentration: 10 - 100 and 200 mM NaCl.

From the injection of NeutrAvidin, the same protocol was repeated for each microfluidic cell used during one experiment.

Image processing and particles detection

The movies acquired after each experiment were analyzed with a custom `Python` algorithm. First, a blurring function was applied to get rid of background noise and increase the signal to noise ratio. Then a top-hat function was used to remove fluorescent spots that were larger than diffraction-limited spots consistent with single emitters. After that, a wavelet function removed the low spatial frequencies present on each image. Finally, another step of top-hat was applied at the end of the algorithm.

Once the movies were processed, we used the function `locate` of the `Python` package called `trackpy`[S5] to detect the fluorescent particles, and especially the labeled RNA spots, visible on an image.

Step-detection algorithm

The assembly experiments enabled us to study signals characterized by the time-evolution of fluorescence intensity emitted by red-labeled subunits from individual RNA spot. In order to detect step events in these temporal traces for equilibrium experiments, we used the detection algorithm `Ruptures`.[S6] The latter needs to be trained on a subset of N signals from which the positions of the steps were manually annotated. To detect steps, `Ruptures` applies a parameter called *penalty* which directly depends on the number of steps, the length of the signal and its noise. During this stage, the detection algorithm is going to use hundreds of *penalties* on the training dataset to predict steps and calculate the cost – i.e. the differences – between its predictions and the given positions. For each *penalty*, the cost is averaged on the entire training dataset. The chosen *penalty* corresponds to the one that minimize the cost function as illustrated by Fig. S4.

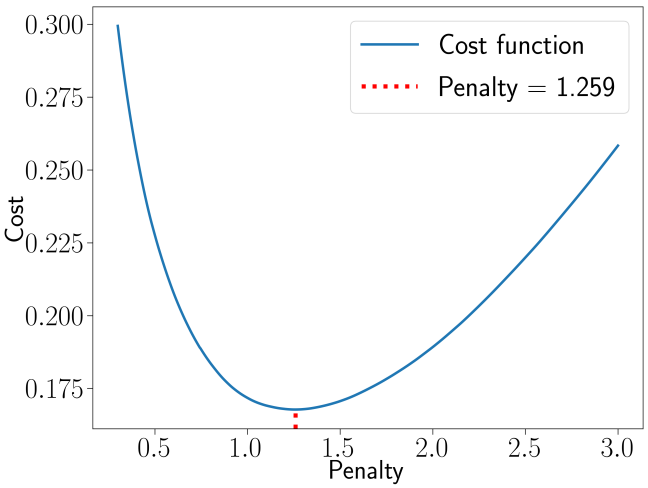


Figure S4: Cost function obtained from the training of `Ruptures` on 50 experimental signals acquired at equilibrium.

The performances of `Ruptures` were assessed on simulated data with added noise, for which the steps were unambiguously identified. By comparing the breakpoints identified with `Ruptures` and the ground-truth signal, we computed three quantities: (i) the number of true positive events (TP) corresponding to breakpoints correctly identified; (ii) the number of false negative events (FN) representing actual events not detected by `Ruptures`; and (iii) the number of false positive events (FP) wrongly recognized by `Ruptures`. The performances were then further quantified through three metrics, namely, $Recall = TP/(TP + FN)$ providing the fraction of missed breakpoints, $Precision = TP/(TP + FP)$ indicating the fraction of falsely detected breakpoints, and $F_{\text{score}} = 2 \times Recall \times Precision/(Recall +$

Precision) which can be viewed as a combination of the two previous metrics. These metrics were estimated on different series of datasets analyzed by `Ruptures` as well as by `quickPBSA`, another step-detection algorithm designed for identification of photobleaching events.[S7] Table S1 shows that `Ruptures` performed better than `quickPBSA` on overall detection and more particularly, `Ruptures` was more trustworthy over the detected breakpoints than `quickPBSA` as evidenced by *Precision*. Theses performances were obtained by using both `Ruptures` and `quickPBSA` on 1,500 simulated signals. Before, `Ruptures` was trained on 50 signals simulated the same way. By measuring the F_{score} on the training dataset, we found that the *penalty* that minimized the cost function was lower than the one maximizing the F_{score} by a factor that could go up to 2.3. For every simulations analysis, we chose the *penalty* that maximized the F_{score}.

Table S1: Comparison of `Ruptures` and `quickPBSA` performances on steps detection with simulated data.

	Recall (%)	*Precision* (%)	F_{score} (%)
`Ruptures`	72	82	76
`quickPBSA`	70	69	69

For experimental data, `Ruptures` was trained on a dataset of 50 signals from which the positions of steps were manually annotated (Fig. S4). The *penalty* value chosen for the signals collected at equilibrium equals 2.469.

Photophysics study

The analysis of both equilibrium and nonequilibrium experiments were based on photophysical parameters of the ATTO 647N dye. Its properties were studied during control experiments on biotin-modified ATTO 647N attached to the coverslip. A movie was acquired in the red channel with the illumination sequence and parameters adapted to the assembly experiments. During analysis, the same algorithm of image processing and spot detection was used before `Ruptures` was applied on the ATTO 647N sites collected. Once again, the step-detection algorithm was trained on experimental data but here the *penalty* value equals 25.1 because, in theory, only one step down (i.e. the photobleaching event) was expected from each red dye. By using the fitted signals, we made sure to analyze only single dye with clear individual step and collected two parameters: the time τ_b when photobleaching occurred and the average collected fluorescence intensity h before photobleaching. Both distributions are illustrated on Fig. S5.

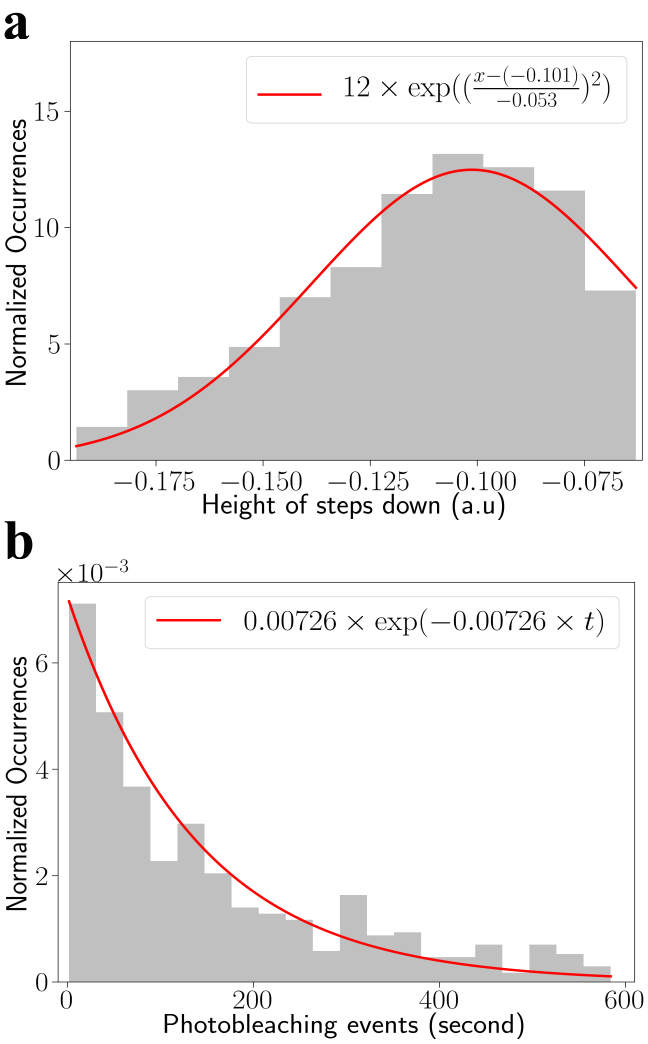


Figure S5: (a) Distribution of h and (b) τ_b. Respectively Gaussian and exponential laws are plotted in red.

In this example, we observe a Gaussian behaviour for h distribution which leads to a mean value of 0.101 a.u and an exponential decay for τ_b with a mean value of ∼138 s.

RNA spots selection for analysis

For assembly experiments performed at equilibrium, three images were acquired in the green channel. Whereas the first one was recorded to have a visual confirmation that RNA molecules were bound to the surface after their incubation, the second and third were analyzed more thoroughly. Indeed, with a `Python` function, we looked here for particles colocalized between those two acquisitions in order to be sure to work with RNA molecules that stayed on the surface during the entire acquisition. In addition, for the analysis, we considered only RNA molecules with at least one binding event of labeled subunits. This type of interaction was taken into account if the intensity fitted by `Ruptures` was above the sum between the fluorescence background I_{bkd} and the mean value of h for at least two consecutive images. The mean intensity $\langle h \rangle$ emitted by a single red dye had been measured for each experiment led at equilibrium (see section Photophysics Study).

Regarding nonequilibrium experiments, the RNA detected in the first green image were the spots of interest. Among them, we considered RNA molecules with at least two binding events of labeled subunits, by selecting the red fluorescent intensity traces recorded at a RNA location which are above $I_{bkd} + 2\langle h \rangle$ for at least $\langle \tau_b \rangle$ seconds without interruption. Indeed, in these experiments, we were observing large increases of fluorescence intensity characterizing early steps of assembly.

Control upon non-specific binding events

Several tests had been performed to evaluate quantitatively the non-specific binding events due to the biological architecture used in our experiments:

- **Non-specific interactions between ATTO 565 and labeled subunits**: Non-specific interactions were studied between 25 nM of subunits ($r = 9.4$ %) and ATTO 565 attached on the surface thanks to a biotin modification. These were compared to specific interactions between labeled subunits and RNA C2 bound to the coverslip that occurred during several assembly experiments: 25 nM of subunits ($r = 9.4$ %), 100 and 200 nM of subunits ($r = 2.5$ %). In all cases, the protocol for equilibrium experiments was followed and an interaction was taken into account if the intensity fitted by `Ruptures` was above $I_{bkd} + \langle h \rangle$ on at least two consecutive images. Fig. S6 shows the percentage of RNA or ATTO 565 colocalized with labeled subunits and it appears clearly that the specific interactions are between 4- to 6-fold higher than the non-specific ones. In other words, labeled subunits did not seem to bind to NeutrAvidin nor a ATTO 565 dye without RNA.

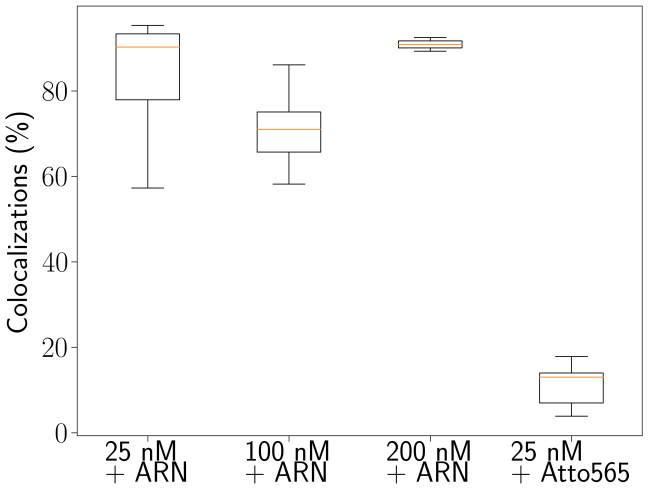


Figure S6: Comparison between specific interactions between labeled subunits and RNA against non-specific interactions between labeled subunits and ATTO 565 dye attached onto a NeutrAvidin. For 25 nM of subunits, the labeling ratio is 9.4 % whereas it is 2.5 % for 100 and 200 nM.

- **Non-specific interactions between Oligo-ATTO and the coverslip**: 300 pM Oligo-ATTO had been directly injected on top of surface chemistry with NeutrAvidin before 100 pM of hybridized RNA. Both samples were incubated 2 min before the beginning of the green acquisition. The goal was to see if the spots observed after an acquisition with

the green laser beam really came from hybridized RNA. By taking the results from three different microfluidic cells, we observed ∼253 ± 123 green spots after the injection of Oligo-ATTO, and then ∼1,924 ± 595 particles when RNA molecules were added to the medium. This observation highlights the fact that single Oligo-ATTO barely bind to the surface when they are not hybridized to the RNA C2.

- **Non-specific interactions between C2 and ATTO 647N NHS-Ester**: Finally, on top of the architecture surface chemistry - NeutrAvidin - hybridized RNA, 5 nM ATTO 647N-NHS-Ester were injected and incubated for 15 minutes to see if binding events can occur with the green-labeled single-stranded RNA. From four microfluidic cells, ∼591 ± 219 RNA spots were analyzed and on average, ∼21 ± 9 ATTO 647N were detected on the first image of the red acquisition. More importantly, we found that only ∼0.1 ± 0.1 % of the genetic material showed binding events with a red dye.

Combined together, these three controls highlighted the fact that the evolution of fluorescence intensity observed on particles detected in the green channel were specific interactions between labeled subunits and single-stranded RNA C2. Each test was performed in the same assembly buffer: 50 mM MES, 100 mM NaCl at pH 6.1.

Computation of $n(t)$ and proof of thermodynamic equilibrium

In equilibrium experiments, the fluorescence intensity emitted by labeled subunits was monitored between 15 and 25 minutes after their injection. To ensure that steady state was established before 15 minutes, control measurements were performed in which we started a 46-minute acquisition immediately after injecting the subunits into the medium. In addition, photobleaching was minimized by reducing the time between two frames to 2 minutes. For each temporal trace, the red fluorescence intensity collected at a time t was divided by $\langle h \rangle$ and the number of labeled subunits n was estimated by taking the closest integer value. In equilibrium experiments this calculation was used only on the first image of each signal with the intensity fitted by `Ruptures` and led to the histogram of Fig. 3d.

To verify when thermodynamic equilibrium was reached, we computed the mean number of labeled subunits $\langle n \rangle$ averaged at each frame over the whole population of RNA considered. Here, it corresponded to the spots colocalized between those detected before and after subunit injection. In Fig. S7, $\langle n \rangle$ was estimated for three concentrations of subunits (25, 100 and 200 nM) and normalized by $\langle n \rangle_{\text{inf}}$, namely:

$$\frac{\langle n(t) \rangle}{\langle n \rangle_{\text{inf}}} = 1 - A \exp\left(-\frac{t}{\tau}\right)$$

For the three concentrations, we saw that $\langle n \rangle$ reached a plateau long before the end of the acquisitions but not exactly at the same time (Fig. S7). These measurements were carried out on several experiments, and we were able to obtain an average estimate of the parameter $3 \times \tau$, which corresponds to the time when $\langle n \rangle$ reaches 95 % of its equilibrium value: 13.3 ± 2.0 min for 25 nM, 10.2 ± 1.6 min for 100 nM and 8.2 ± 2.0 min for 200 nM. Therefore, these results confirmed that between 15 and 25 minutes after the injection of subunits, a steady state was established for 25, 100 and 200 nM.

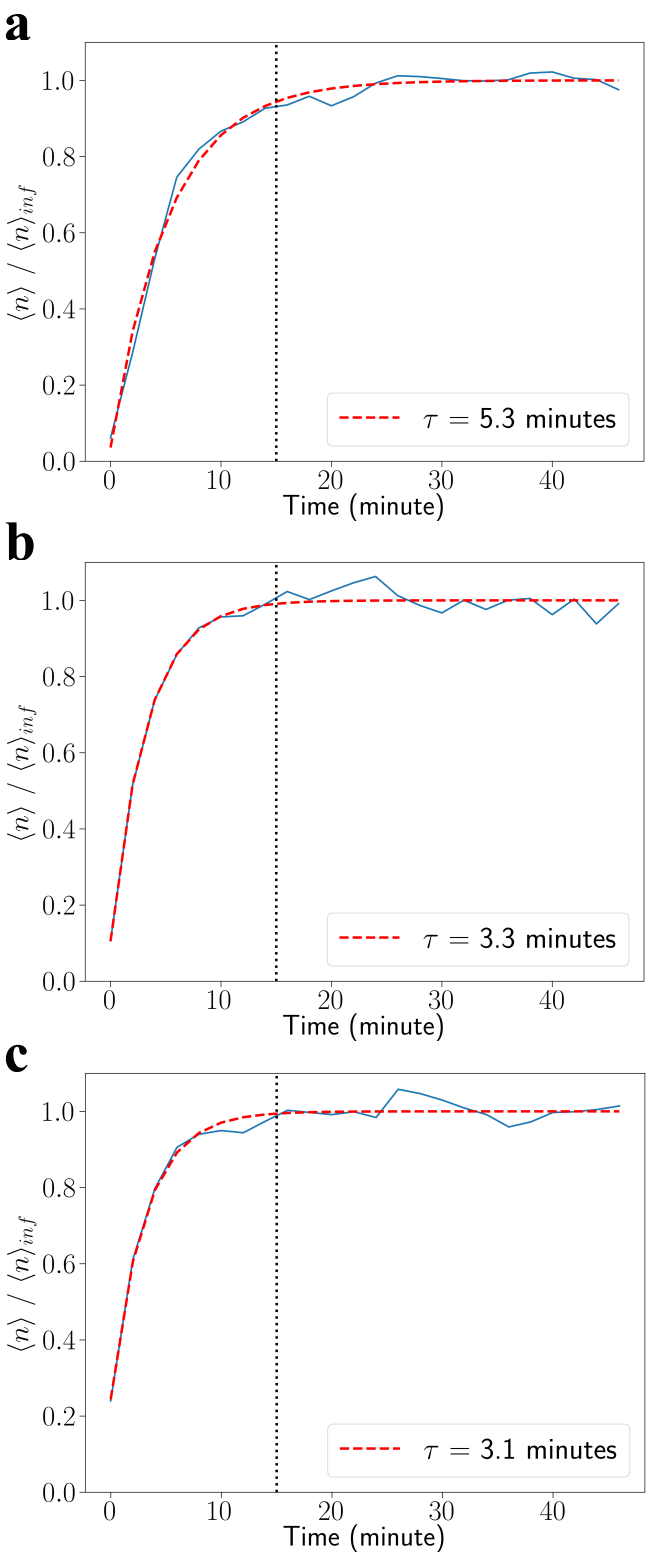


Figure S7: Evolution of the mean number of labeled subunits on ssRNA for subunit concentration of (a) 25 nM (r = 10 %), (b) 100 nM (r = 2.5 %) and (c) 200 nM (r = 2.5 %). Exponential decay evolution is plotted in red dashed lines. Black dotted lines indicate the beginning of the acquisition for equilibrium experiments.

Examples of temporal traces for nonequilibrium experiments

As explained in the main text, the temporal evolution of fluorescence intensity on ssRNA molecules was studied at three different ionic strengths. In Fig. 5a, we observed a characteristic evolution at 10 mM NaCl where the maximum intensity I_{max} was reached around 10 minutes before slowly decreasing due to the combination of unbinding events and photobleaching of bound labeled subunits. For higher salt concentrations, the same general behaviour was obtained as illustrated by Fig. S8. At 100 mM, I_{max} was obtained approximately 5 minutes after subunit injection whereas it occurred around 3 minutes for 200 mM. In these two examples, the values of I_{max} were roughly the same. For both buffers, the concentration of subunits was still set to 32.5 nM (r = 8 %).

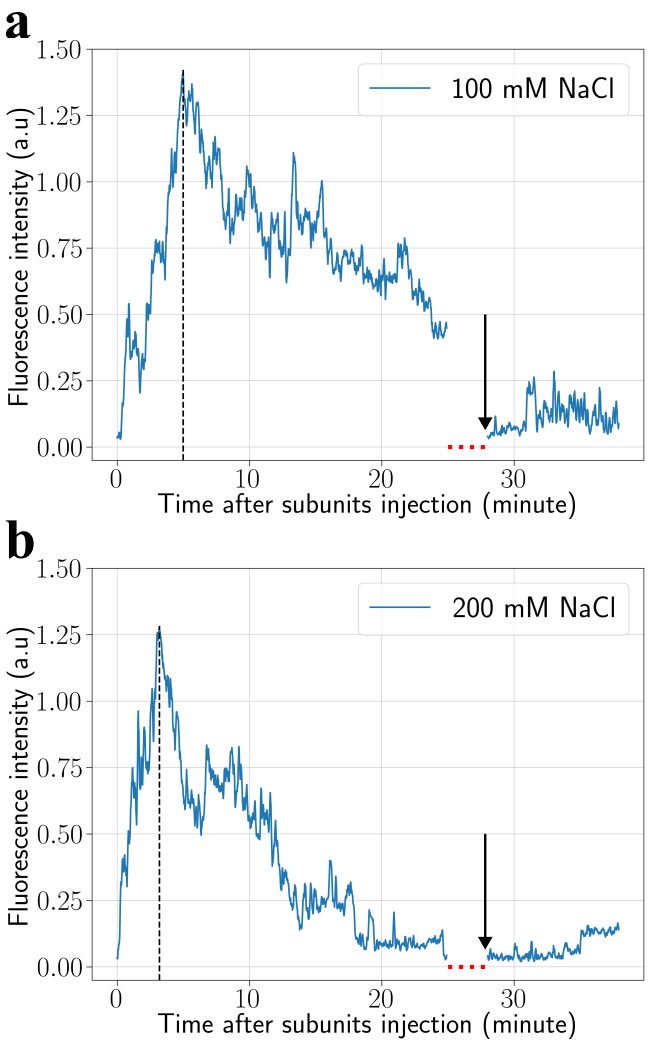


Figure S8: Evolution of fluorescence intensity emitted by labeled subunits from one RNA molecule in a buffer with (a) 100 mM NaCl and (b) 200 mM NaCl. The black dashed lines show when I_{max} is reached, the dotted red lines indicate a continuous illumination for 3 minutes and the black arrows correspond to a second injection of CCMV subunits.

Statistical modeling

Time is divided into frames, each of which has a duration Δt. Assume that at time t a given RNA molecule has n bound labeled subunits. Within the next time interval, three events can occur: (i) a new labeled subunit binds with a rate constant $k_{on}(n)$ and a probability $k_{on}(n)\Delta t$; (ii) a bound labeled subunit disappears, whether by unbinding or by photobleaching, with a rate constant $k_{off}(n)$ and a probability $nk_{off}(n)\Delta t$; (iii) nothing happens with a probability $1 - k_{on}(n)\Delta t - nk_{off}(n)\Delta t$. The probability of having n bound labeled subunits at time $t + \Delta t$ is then expressed by:

$$\begin{aligned} p(n, t+\Delta t) = {} & k_{on}(n-1)\Delta t p(n-1,t) + \\ & (n+1)k_{off}(n+1)\Delta t p(n+1,t) + \\ & + \{1 - k_{on}(n)\Delta t - nk_{off}(n)\Delta t\}\, p(n,t) \end{aligned} \quad (1)$$

In first approximation, near equilibrium, each labeled subunit sees a roughly identical environment, or equivalently the same mean field, when interacting with a RNA molecule, in such a way that k_{on} and k_{off} can be considered as constant depending solely on $\langle n \rangle$. In steady state, $p(n, t+\Delta t) = p(n,t) = p(n)$, and since $p(-1) = 0$, we arrive at

$$np(n) = \mu p(n-1) \quad (2)$$

where the equilibrium constant is $\mu = k_{on}/k_{off}$. By induction and given that normalization imposes $\sum_n p(n) = 1$, we obtain

$$p(n) = \frac{\mu^n}{n!} e^{-\mu} \quad (3)$$

which is the well-known Poisson distribution verifying $\mu = \langle n \rangle$. If k_u is the unbinding rate constant and k_b the photobleaching rate constant, $k_{off} = k_u + k_b$. After determining k_b separately, k_u can be estimated experimentally by computing $\langle n \rangle$ and determining k_{on} from a histogram of binding events (see main text).

In reality, both labeled and unlabeled subunits bind on RNA. We note n_{tot} the total number of subunits bound on a RNA molecule and r the bulk labeling ratio. The number of bound labeled subunits follows the binomial distribution, i.e.,

$$p(n/n_{tot}) = \binom{n_{tot}}{n} r^n (1-r)^{n_{tot}-n}. \quad (4)$$

Eq. 2 still applies with n_{tot} but the equilibrium constant μ_{tot} must be computed with k_{on} for all subunits. Therefore, $p(n_{tot}) = (\mu_{tot}^{n_{tot}}/n_{tot}!)p(0)$ and $p(0)$ is determined from the normalization condition. In principle, n_{tot} cannot exceed a maximal number n_{max} corresponding to the

number of subunits in a complete nucleocapsid, namely here 90. As a result,

$$p(0) = \left(\sum_{k=0}^{n_{\max}} \frac{\mu_{\text{tot}}^k}{k!} \right)^{-1}. \quad (5)$$

Since

$$p(n) = \sum_{k=0}^{n_{\max}} p(n/k)p(k) \quad (6)$$

and by injecting Eq. 4, we arrive at a rather complex distribution for the number of bound labeled subunits:

$$p(n) = \frac{(r\mu_{\text{tot}})^n}{n!} \frac{\sum_{k=0}^{n_{\max}-n} \frac{\mu_{\text{tot}}^k}{k!}(1-r)^k}{\sum_{k=0}^{n_{\max}} \frac{\mu_{\text{tot}}^k}{k!}}. \quad (7)$$

For $n_{\max}$ large enough, we recover the Poisson distribution of Eq. 3,

$$p(n) \approx \frac{(r\mu_{\text{tot}})^n}{n!} e^{-r\mu_{\text{tot}}}. \quad (8)$$

Numerically, given that $n_{\max} = 90$ or higher, this approximation is quite verified for $r \leq 10\%$ and $\mu_{\text{tot}} \leq 70$, which was mostly the case throughout this study. As a result, we could apply the simplified relationship: $\langle n \rangle \approx r\mu_{\text{tot}} \approx r\langle n_{\text{tot}} \rangle$.

References

(S1) Ali, A.; Roossinck, M. J. Rapid and efficient purification of cowpea chlorotic mottle virus by sucrose cushion ultracentrifugation. *J. Virol. Methods* **2007**, *141*, 84–86.

(S2) Piliarik, M.; Sandoghdar, V. Direct optical sensing of single unlabelled proteins and super-resolution imaging of their binding sites. *Nat Commun* **2014**, *5*, 4495.

(S3) Young, G.; Hundt, N.; Cole, D.; Fineberg, A.; Andrecka, J.; Tyler, A.; Olerinyova, A.; Ansari, A.; Marklund, E. G.; Collier, M. P.; Chandler, S. A.; Tkachenko, O.; Allen, J.; Crispin, M.; Billington, N.; Takagi, Y.; Sellers, J. R.; Eichmann, C.; Selenko, P.; Frey, L.; Riek, R.; Galpin, M. R.; Struwe, W. B.; Benesch, J. L. P.; Kukura, P. Quantitative mass imaging of single biological macromolecules. *Science* **2018**, *360*, 423–427.

(S4) Bugaud, O.; Barbier, N.; Chommy, H.; Fiszman, N.; Le Gall, A.; Dulin, D.; Saguy, M.; Westbrook, N.; Perronet, K.; Namy, O. Kinetics of CrPV and HCV IRES-mediated eukaryotic translation using single-molecule fluorescence microscopy. *RNA* **2017**, *23*, 1626–1635.

(S5) Allan, D. B.; Caswell, T.; Keim, N. C.; van der Wel, C. M. soft-matter/trackpy: Trackpy v0.4.2. 2019.

(S6) Truong, C.; Oudre, L.; Vayatis, N. Selective review of offline change point detection methods. *Signal Processing* **2020**, *167*, 107299.

(S7) Hummert, J.; Yserentant, K.; Fink, T.; Euchner, J.; Ho, Y. X.; Tashev, S. A.; Herten, D.-P. Photobleaching step analysis for robust determination of protein complex stoichiometries. *MBoC* **2021**, *32*, ar35.