Supplementary information

Systematic Assessment of the Accuracy of Subunit Counting in Biomolecular Complexes Using Automated Single Molecule Brightness Analysis

John S H Danial[1,2,3,6]*, Yuri Quintana[1,6], Uris Ros[1,4], Raed Shalaby[1,4], Eleonora G Margheritis[5], Sabrina Chumpen Ramirez [5], Christian Ungermann[5], Ana J Garcia-Saez[1,4]* and Katia Cosentino[1,5]*

[1] Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen 72076, Germany

[2] Yusuf Hamied Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, United Kingdom

[3] UK Dementia Research Institute, University of Cambridge, Cambridge CB2 1EW, United Kingdom

[4] Institute for Genetics and Cologne Excellence Cluster on Cellular Stress Responses in Aging-Associated Diseases (CECAD), Cologne 50931, Germany

[5] Department of Biology/Chemistry and Center for Cellular Nanoanalytics (CellNanOs), University of Osnabrück, Osnabrück 49076, Germany

[6] These authors contributed equally.

* To whom correspondence should be addressed: Katia Cosentino (kacosentino@uni-osnabrueck.de), John S H Danial (js2494@cam.ac.uk) and Ana J Garcia-Saez (ana.garcia@uni-koeln.de)

Methods

Simulating ground-truth data

To simulate the ground-truth data, a pre-set number of particles (i.e. complexes) were randomly scattered across images which are 256 by 256 pixels in size (except for those simulated in **Figures 4c-n** which were 1024 by 2024 pixels in size to accommodate for the larger number of complexes in the same field of view without significantly increasing the density) with pixel size of 100 nm. The standard deviation in the PSF of each particle was fixed at 130 nm. Each complex contained a pre-set number of molecules with intensities sampled from a random number generator based on a normal distribution with a mean intensity of I_m and standard deviation of v, where v is the inter complex variation in the intensity. Each molecule had a photobleaching time that was randomly sampled from 1 to the maximum number of frames which in our simulation was set to 500 frames. The intensity of each molecule was set to the sampled intensity before the photobleaching time point and to 0 afterwards. Each molecule was convoluted with a 2D Gaussian Kernel. To account for the noise statistics of Electron-Multiplying Charge Coupled Detectors (EMCCDs), the produced images in counts I_c were modified as follows:

$$I_e = I_c * QE + DC + RO \tag{1}$$

where I_e is the image signal in electrons, QE is the quantum efficiency, DC is the dark current and RO is the readout noise. The quantum efficiency was set at 95%, dark current was set at 0.0002 electrons/second and readout noise was set to 1 electron. These values are typical of commercially available EMCCDs. To simulate noise, I_e was modified as follows:

$$I_e = gamma\left(I_e, G - 1 + \frac{1}{I_e}\right) + O \tag{2}$$

where G is the camera gain (corrected for conversion factor) and O is the camera bias offset. The gain was set to 58.8 (electron gain of 300 and conversion factor of 5.1) and offset to 100. This noise model was chosen to best replicate the electron-multiplication feature in EMCCDs[1]. The produced images were saved in the big TIFF format at 16 bits for further processing using SAS.

Detecting single molecule using a deep convolutional network

Detection of each complex / molecule was carried out as described in[2]. Briefly, we developed and used DeepSinse, a simple, multi-layer Convolutional Neural Network (CNN) architecture to enable fast detection of single molecules using as few parameters as possible. Our neural network is composed of a CNN, a dense layer and a SoftMax (classification) layer. The neural network was first trained to classify simulated ground-

truth datasets of noise and Gaussian bursts in pre-labelled Regions Of Interest (ROIs), then validated on different, unseen, datasets of pre-labelled ROIs. We then tested it on ground-truth generated ROIs. The neural network is finally deployed by feeding an image into a peak-finding algorithm based on identifying regional maxima. The peak-finding algorithm outputs hundreds of noise- and burst- containing ROIs which are then fed into the trained network for classification, thus, resulting in an annotated image. Burst-containing, and pure-noise, images were simulated with the formers' peak burst intensities varying from 50 to 100 counts corresponding to signal-to-noise ratios from 24.85 to 45.81 and 100 images were simulated with pure noise. Pre-annotated ROIs were picked from each of these images, intensity scaled between 0 and 1 to avoid subjective segmentation parameters such as the intensity threshold, shuffled and fed into the neural network for training. To optimize performance, the neural network was trained using different ROI radii and number of ROIs. The lowest false negative rate FNR (61.45%) and false positive rate FPR (0.3%) were achieved at a ROI radius of 5 pixels and 10,000 training ROIs. The trained network was then integrated into SAS for immediate use. Particles that are less than 6 pixels away from the borders of each image, as well as those which are less than 5 pixels apart from each other are rejected. Finally, the detected particles were fed into a least-squares solver to fit a Gaussian function to the PSF of each particle using the cpufit plugin[3] and extract a value to the standard deviation which was compared to a user-specified value (200 nm for the simulated and experimental data as commonly used in single molecule microscopy[4]). Particles exhibiting a standard deviation value smaller than the user-set value are accepted and, otherwise, rejected. This approach allows for discarding multiple particles present in the same ROI.

Extraction of intensity traces

The intensity of the detected and accepted molecules were extracted from the movies by drawing ROI with a user-specified radius around each of these molecules at each time point. The local background at each time point is calculated by extracting the intensity in a region around the ROI, 2 pixels larger than the ROI. Intensity traces are used for both, the selection of monomeric complexes for calibration (see "Selection of monomeric traces for calibration"), and for the extraction of the intensity value for the detected molecule. In this last case, the maximum intensity of each trace is calculated by taking the median of the first 5 time points of the intensity trace, and the background is subtracted by taking the median of the intensity of the last 5 time points of the background trace (which, theoretically, should yield the same result as the first 5 time points).

Selection of monomeric traces for calibration

The traces extracted for particles belonging to the calibration sample are fed into a trace annotator which calculates the absolute gradient of each intensity trace, normalizes the calculated gradient between 0 and 1 and, finally, extracts the number of peaks in the normalized gradient above a threshold value, known as the minimum peak height, which we set at 0.5 (for simulated data), 0.9 (for experimental data). The value of the minimum peak height was optimized once for each dataset by inspection of the calibration curve so that,

ideally, a single peak corresponding to the monomeric traces was dominant. Gradients with a single peak were chosen as monomeric traces (i.e. arising from a complex with a single molecule).

Calculating the proportion of labelled species

After measuring the intensity for each detected and accepted particle (measured as described in "Extraction of intensity traces"), a kernel probability distribution function (pdf) with a bandwidth of 5 photons was calculated from the intensity values of the calibration species and normalized between 0 and 1. A peak finder was subsequently used to find peaks in the normalized kernel pdf with a minimum height of 0.8 (normalized value). The first peak, corresponding to the intensity distribution of monomers, was selected and its mean intensity value (I_m) was used to fit the following Gaussian Mixture Model (GMM) using a non-linear curve fitting solver to the kernel pdf to extract the standard deviation in the intensity values of the population of monomers (σ_m):

$$f(I) = \sum_{n=1}^{n} \quad a_i e^{-\left(\frac{I-I_i}{2\sigma_i}\right)^2} \tag{3}$$

Where $a_i, I_i and\ \sigma_i$ are the amplitude, mean intensity value and standard deviation in intensity value of each Gaussian. n is the number of Gaussians and was set to 2 to account for any dimers present in the sample used for calibration. In this case, $I_m \overset{\text{def}}{=} I_1$ and $\sigma_m \overset{\text{def}}{=} \sigma_1$. Following, a kernel pdf was constructed from the sample of unknown species with a bandwidth of 5 photons which was fit, using a non-linear curve fitting solver, with an idealized GMM:

$$f(I) = \sum_{n=1}^{n_{max}} \quad \frac{p_i}{\sigma_m\sqrt{2\pi n}} e^{-\left(\frac{I-(I_m n)}{2\sigma_m\sqrt{n}}\right)^2} \tag{4}$$

Where p_i is the proportion of each species corresponding to the area below each Gaussian curve and n_{max} is the maximum number of Gaussians. n_{max} was calculated as described in[22] from I_m and σ_m using the following formula:

$$n_{max} = floor\left[\left(\frac{I_m}{\sigma_m}\right)^2\right] \tag{5}$$

Constraining the maximum number of Gaussians as such does not prevent overfitting. To account for overfitting, we performed non-linear curve fitting with n from 1 to n_{max}. In each time we performed a fit, we calculated the root sum squared residual and whenever this was less than 95% of the minimum calculated value, this last was set as the minimum value and the used number of Gaussians as n_{max}. Finally, an optional

refinement step was performed were I_m was scanned in a $\pm$ 10 photons region with 1 photon resolution and modified to where the root sum squared residual was minimized.

The residual error is calculated in an ascending order of Gaussians. The minimum residual error is also calculated in that specific order. For configurations where the proportion of species decreases with increasing the species number (as in figure 3h) a minimum residual error is reached with low species numbers which have higher proportions. This is not the case with configurations where the proportion of species increases with increasing the specie number (as in figure 3i).

Correcting for labelling efficiency to calculate the true proportion of labelled and unlabelled species

Labelling efficiency was corrected for as described in[5]. Briefly, each molecule was assumed to be either labelled (1) or not labelled (0). To uncover the true proportion of species from those measured above, we constructed a binomial probability distribution function of the following form:

$$f(n,p) = \left(\frac{n}{x}\right) p^x q^{n-x} I_{0,1,\ldots,n}(x) \tag{6}$$

Where x is the species number, n is the number of trials, p is the probability of a molecule being labelled (set to the labelling efficiency) and q is the probability of a molecule not being labelled ($= 1 - p$). x and n took values from 1 to the number of measured species. A linear least-squares problem solver was then used to calculate the true proportion of species taking the constructed binomial probability distribution function as the multiplier matrix and the measured proportion of species as the constant vector. The solver was constrained between 0 and 100% across all species.

Preparing BAX reference sample and imaging setup used for experimental validation

The sample was prepared as described in[5]. Briefly, egg phosphatidylcholine and cardiolipin (Avanti Polar Lipids, US) were mixed in a 7:3 ratio and dissolved in chloroform that was evaporated under reduced pressure for 3 hours. The lipid film was then resuspended with 150 mM NaCl, 10 mM Hepes (pH7.4) to a final concentration of 1 mg/mL. The lipid solution was subjected to five cycles of freezing and thawing after which they were manually extruded through a polycarbonate membrane with a defined pore size (100 nm) using glass syringes. The formed large unilamellar vesicles (LUVs) were then incubated with 10 nM Atto488 labelled Bax (labelling efficiency 84%) for 1 hour at 43°C to give proteoliposomes which were subsequently diluted 1:10 with untreated liposomes. The supported lipid bilayer (SLB) was formed by incubating the diluted proteoliposomes on a glass slide, previously cleaned with a piranha solution, at 37 °C for 2 minutes in the presence of 3 mM $CaCl_2$ and then washed several times with 150 mM NaCl, 10 mM Hepes (pH7.4) to remove non-fused vesicles. Samples were imaged on the setup described below for a total of 1200 frames under a 35 ms exposure time and 25 ms delay between frames with a power density of ~ 1 kW/cm^2.

The sample was imaged on a Total Internal Reflection Fluorescence (TIRF) microscope. The laser excitation from a 4-wavelengths (405 nm, 488 nm, 561 nm and 647 nm) laser engine (iChrome MLE, Toptica Photonics AG, DE) was coupled into a multi-mode fibre onto a TIRF-alignment module (Laser TIRF 3, Carl Zeiss AG, DE) inserted into the side port of an upright microscope (Axiovert 200, Carl Zeiss AG). Using the TIRF-alignment module, the excitation light was focused onto the back focal plane of a 100x, 1.46 Numerical Aperture (NA) objective (Apo-TIRF, Carl Zeiss AG) after passing through a triple-band clean-up filter (TBP 483 + 564 + 642 (HE), Carl Zeiss AG) and being reflected off a quad-band dichroic filter (TFT 506 + 582 + 659 (HE), Carl Zeiss AG). Image was additionally magnified by 1.6x to obtain a final pixel size of 100 nm. The emission was collected using the same objective, passed through a triple-band emission filter (TBP 526 + 601 + 688 (HE), Carl Zeiss AG) and focused on an EMCCD camera (iXon 88X, Andor, IE) cooled at -70 degC.

Preparing Atg9 reference sample and imaging setup used for experimental validation

Atg9 was overexpressed in yeast as GFP-3xFLAG fusion protein and purified as previously described[6]. For the generation of Atg9 containing liposomes or protein free liposomes, 50% DOPC, 30% DOPE and 20% DOPS (Avanti Polaris) were mixed and dried under vacuum 1h at 37°C. The lipid film was resuspended in buffer A (300 mM KCl, 50 mM HEPES, pH 7.4) to a concentration of 20 mg/ml and sonicated 15 min. Generated SUVs were destabilized by adding 35 mM CHAPS followed by incubation 1h at 23°C. 1.5mg of lipids were mixed with 10 µg of the purified protein or with lysis buffer (for protein free liposomes) and the mixes were incubated 1h at 4°C. Then, samples were diluted 10 times in buffer A to reduce the concentration of detergent below the critical micelle concentration (CMC). Samples were dialyzed using Slide-A-Lyzer dialysis cassettes 20 MWCO (Thermo Scientific) against buffer A plus 0.2 g of Bio-beads SM2 adsorbent Media (BIO-RAD) per liter of buffer. Reconstituted liposomes were freeze-thawed two times. The SLB was formed by incubating 1:10 diluted proteoliposomes on a glass slide, cleaned with a piranha solution, at 37°C for 10min with 3mM $CaCl_2$, and washed 15 times with 150 mM NaCl, 10mM Hepes (pH7.4) buffer to remove non-fused vesicles.

Samples were imaged on a custom-designed TIRF microscope for a total of 2000 frames under a 30 ms exposure time. Laser excitation from a 488 nm laser, max. power 400 mW (Sapphire, Coherent) was coupled into a single mode polarization maintaining fiber to a TIRF module connected to an Olympus IX83 inverted microscope with hardware autofocus system (IX3-ZDC, Olympus) and 100x oil-immersion objective ((UPLAPO100xOHR). Image was additionally magnified by 1.6x (IX3-CAS, Olympus) to obtain a final magnification of 160x and a pixel size of 100 nm. Fluorescence was filtered by a four-line polychroic mirror (zt405/488/561/640rpc, Chroma, 3 mm) and rejection band filter (zet405/488/561/647 TIRF, Chroma), and the emission was focussed on an iXon Ultra EMCCD Camera (Andor Technologies).

Supplementary Figure 1

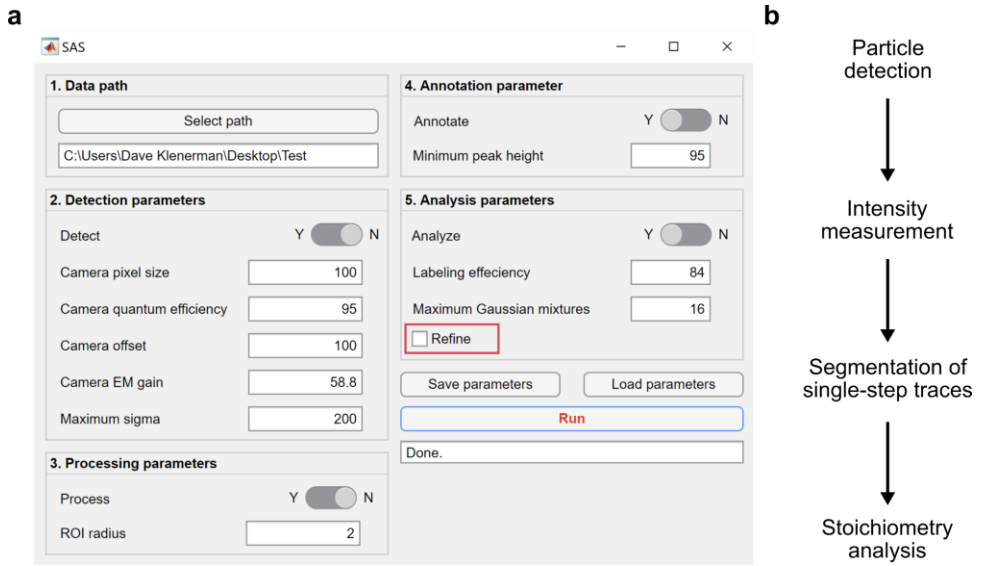

Figure S1. (a) Graphical user interface (GUI) of SAS. The GUI is divided into five panels: 1. Data import; 2. Detection parameters to be fed with the experimental parameters used by the experimenter during data acquisition; 3. Processing parameters that allows the user to select the ROI radius based on the experimental parameters used; 4. Annotation of parameters in a text file and choice of the threshold value (minimum peak height) for appropriate monomeric curves selection; 5. Analysis parameters that allow to introduce the values of the protein labelling efficiency and of the maximum number of Gaussian mixtures the user aims to resolve. Importantly, the refinement step can be selected in this panel (red box). (b) SAS workflow.

Supplementary Figure 2

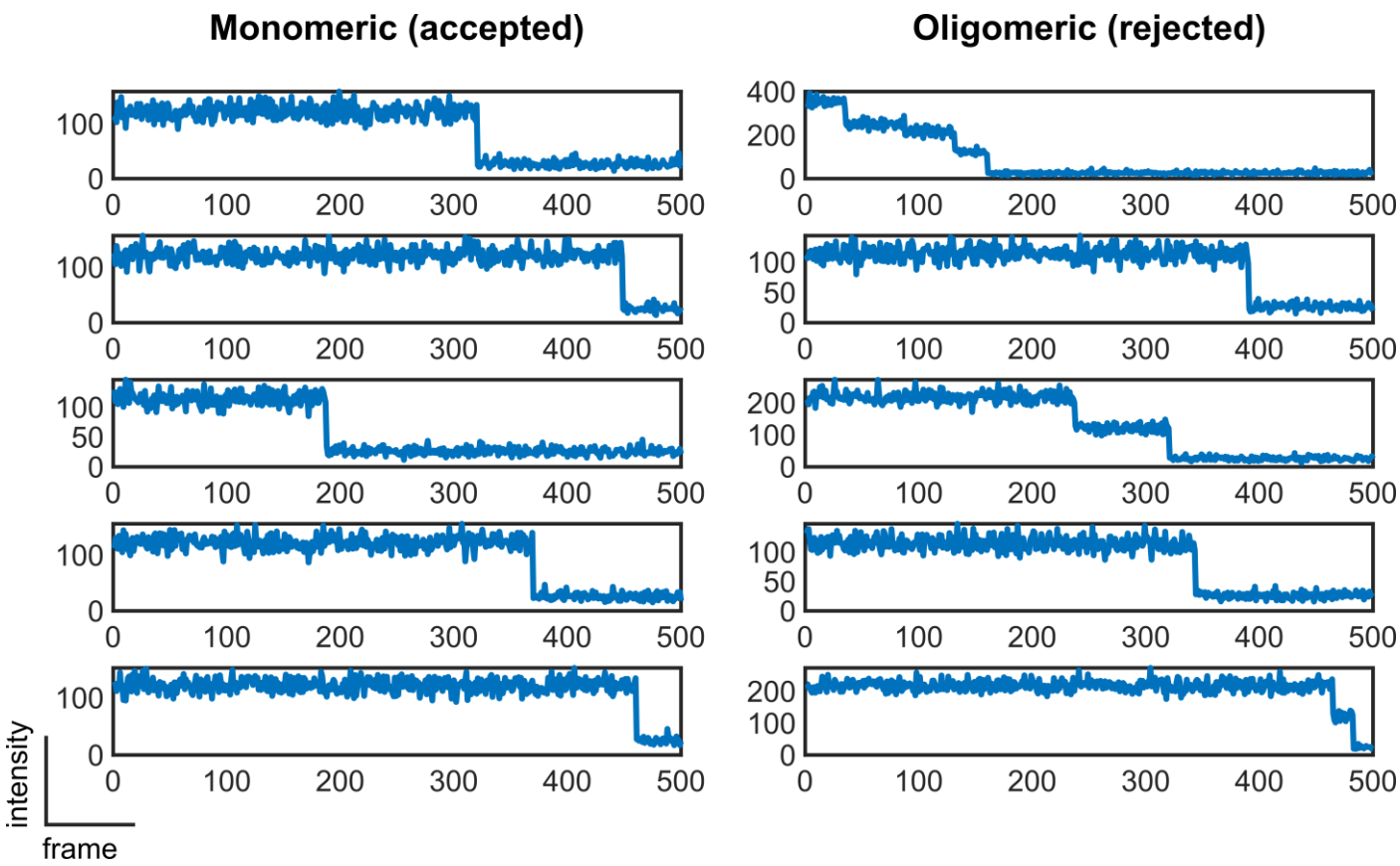

Figure S2. Gallery of randomly accepted (monomeric) and rejected (oligomeric) traces selected by the trace annotator tool of SAS.

Supplementary Figure 3

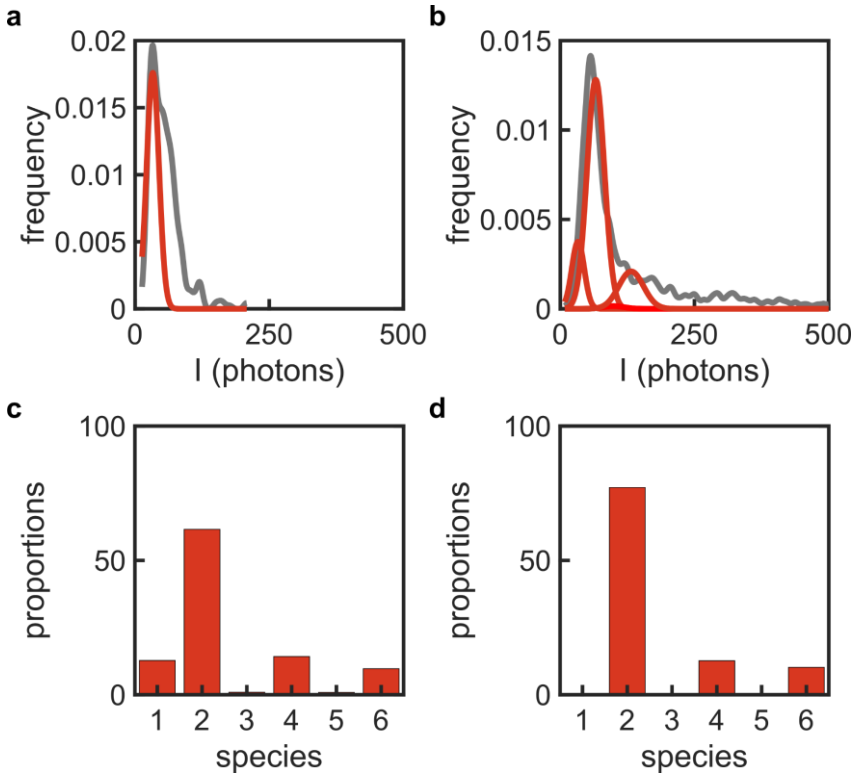

Figure S3. Graphs of the (a) calibration, and (b) unknown ~~(c)~~ intensity distributions obtained from the experimental data to measure the stoichiometry of BAX complexes. Percentage of occurrence of Bax species (c) before~~,~~ and (d) after labelling correction (labelling efficiency 84%). Illustrated data are without refinement. Results after refinement are shown in figure 5.

Supplementary Table 1

Figure	**Stoichiometric configuration** **Experimental conditions (if applicable)** **Reference in supplementary data / reference dataset**
Figure 2a	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [varied from 20 to 320 particles per movie] Unknown: 1-mer (50%), 2-mer (50%) [varied from 20 to 320 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 10, variation in photon count = 0%, lateral sigma = 130 nm, frame size = 256 pixels x 256 pixels
	Columns A to O / dataset # 2 to 5
Figure 2b	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 1-mer (50%), 2-mer (50%) [80 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 2 to 50, variation in photon count = 0%, lateral sigma = 130 nm, frame size = 256 pixels x 256 pixels
	Columns P to AI / dataset # 6 to 10
Figure 2c	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: varied from 1-mer (50%), 2-mer (50%) [80 particles per movie] to 1-to-16-mers (3.125% each)
	# frames = 500, # movies = 10, max. photon count = 10, variation in photon count = 0%, lateral sigma = 130 nm, frame size = 256 pixels x 256 pixels
	Columns AJ to BC / dataset # 11 to 14
Figure 2d	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 1-mer (50%), 2-mer (50%) [80 particles per movie]
	# frames = 500, # movies = varied from 5 to 40, max. photon count = 10, variation in photon count = 0%, lateral sigma = 130 nm, frame size = 256 pixels x 256 pixels
	Columns BD to BW / dataset # 15 to 18
Figure 2e	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 1-mer (50%), 2-mer (50%) [80 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 10, variation in photon count = 0% to 50%, lateral sigma = 130 nm, frame size = 256 pixels x 256 pixels
	Columns BX to CL / dataset # 19 to 22
Figure 2f	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 1-mer (50%), 2-mer (50%) [80 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 10, variation in photon count = 0%, lateral sigma = 130 nm, frame size = 256 pixels x 256 pixels. Analysis bin sized varied from 1 to 20.
	Columns HP to IR / dataset # 45 to 49
Figure 2g	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 1-mer (50%), 2-mer (50%) [80 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 10, variation in photon count = 0%, lateral sigma = varied from 100 nm to 200 nm, frame size = 256 pixels x 256 pixels
	Columns CM to DA / dataset # 23 to 25
Figure 2h	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 1-mer (50%), 2-mer (50%) [80 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 10, variation in photon count = 0%, lateral sigma = 130 nm, frame size = 256 pixels x 256 pixels, pixel size varied from 100 nm to 160 nm.
	Columns IT to JJ / dataset # 50 to 52
Figure 3a	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 1-to-12-mers (8.33% each) [96 particles per movie]

	# frames = 500, # movies = 10, max. photon count = 10, variation in photon count = 0%, lateral sigma = 130 nm, frame size = 256 pixels x 256 pixels
	Columns DB to DF / dataset # 26
Figure 3b	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 1,3,5,7,9 and 11-mers (16.67% each) [96 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 10, variation in photon count = 0%, lateral sigma = 130 nm, frame size = 256 pixels x 256 pixels
	Columns DG to DK / dataset # 27
Figure 3c	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 1,4,7 and 10-mers (25% each) [96 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 10, variation in photon count = 0%, lateral sigma = 130 nm, frame size = 256 pixels x 256 pixels
	Columns DL to DP / dataset # 28
Figure 3d	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 1,5 and 9-mers (33.33% each) [96 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 10, variation in photon count = 0%, lateral sigma = 130 nm, frame size = 256 pixels x 256 pixels
	Columns DQ to DU / dataset # 29
Figure 3e	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 1 and 7-mers (50% each) [96 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 10, variation in photon count = 0%, lateral sigma = 130 nm, frame size = 256 pixels x 256 pixels
	Columns DV to DZ / dataset # 30
Figure 3f	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 1-mer (15.38%), 2-mer (14.10%), 3-mer (12.82%), 4-mer (11.54%), 5-mer (10.26%), 6-mer (8.97%), 7-mer (7.69%), 8-mer (6.41%), 9-mer (5.13%), 10-mer (3.85%), 11-mers (2.56% each) and 12-mer (1.28%) [312 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 10, variation in photon count = 0%, lateral sigma = 130 nm, frame size = 256 pixels x 256 pixels
	Columns EA to EE / dataset # 31
Figure 3g	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 1-mer (1.28%), 2-mer (2.56%), 3-mer (3.85%), 4-mer (5.13%), 5-mer (6.41%), 6-mer (7.69%), 7-mer (8.97%), 8-mer (10.26%), 9-mer (11.54%), 10-mer (12.82%), 11-mers (14.10% each) and 12-mer (15.38%) [312 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 10, variation in photon count = 0%, lateral sigma = 130 nm, frame size = 256 pixels x 256 pixels
	Columns EF to EJ / dataset # 32
Figure 3h	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 1-mer (28.57%), 3-mer (23.81%), 5-mer (19.10%), 7-mer (14.29%), 9-mer (9.52%) and 11-mer (4.76%) [168 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 10, variation in photon count = 0%, lateral sigma = 130 nm, frame size = 256 pixels x 256 pixels
	Columns EK to EO / dataset # 33
Figure 3i	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 2-mer (4.76%), 4-mer (9.52%), 6-mer (14.29%), 8-mer (19.05%), 10-mer (23.81%) and 12-mer (28.57%) [168 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 10, variation in photon count = 0%, lateral sigma = 130 nm, frame size = 256 pixels x 256 pixels
	Columns EP to ET / dataset # 34
Figure 4a	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie]

	Unknown: 1-mer (15.38%), 2-mer (14.10%), 3-mer (12.82%), 4-mer (11.54%), 5-mer (10.26%), 6-mer (8.97%), 7-mer (7.69%), 8-mer (6.41%), 9-mer (5.13%), 10-mer (3.85%), 11-mers (2.56% each) and 12-mer (1.28%) [312 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 5, variation in photon count = 20%, lateral sigma = 130 nm, frame size = 1,024 pixels x 1,024 pixels
	Columns EV to EZ (rows 1 to 14) / dataset # 31 copy
Figure 4b	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 1-mer (28.57%), 3-mer (23.81%), 5-mer (19.10%), 7-mer (14.29%), 9-mer (9.52%) and 11-mer (4.76%) [168 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 5, variation in photon count = 20%, lateral sigma = 130 nm, frame size = 1,024 pixels x 1,024 pixels
	Columns FB to FF (rows 1 to 14) / dataset # 33 copy
Figure 4c	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 1-mer (15.38%), 2-mer (14.10%), 3-mer (12.82%), 4-mer (11.54%), 5-mer (10.26%), 6-mer (8.97%), 7-mer (7.69%), 8-mer (6.41%), 9-mer (5.13%), 10-mer (3.85%), 11-mers (2.56% each) and 12-mer (1.28%) [312 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 5, variation in photon count = 20%, lateral sigma = 130 nm, frame size = 1,024 pixels x 1,024 pixels
	Columns EV to EZ (rows 22 to 34) / dataset # 31 copy (2)
Figure 4d	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 1-mer (11.76%), 2-mer (11.03%), 3-mer (10.29%), 4-mer (9.56%), 5-mer (8.82%), 6-mer (8.09%), 7-mer (7.35%), 8-mer (6.62%), 9-mer (5.88%), 10-mer (5.15%), 11-mers (4.41%), 12-mer (3.68%), 13-mer (2.94%), 14-mer (2.21%), 15-mer (1.47%) and 16-mer (0.74%) [544 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 5, variation in photon count = 20%, lateral sigma = 130 nm, frame size = 1,024 pixels x 1,024 pixels
	Columns FH to FL / dataset # 35
Figure 4e	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 1-mer (9.52%), 2-mer (9.05%), 3-mer (8.57%), 4-mer (8.10%), 5-mer (7.62%), 6-mer (7.14%), 7-mer (6.67%), 8-mer (6.19%), 9-mer (5.71%), 10-mer (5.24%), 11-mers (4.76%), 12-mer (4.29%), 13-mer (3.81%), 14-mer (3.33%), 15-mer (2.86%), 16-mer (2.38%), 17-mer (1.90%), 18-mer (1.43%), 19-mer (0.95%), 20-mer (0.47%) [840 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 5, variation in photon count = 20%, lateral sigma = 130 nm, frame size = 1,024 pixels x 1,024 pixels
	Columns FN to FR / dataset # 36
Figure 4f	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 1-mer (8%), 2-mer (7.67%), 3-mer (7.33%), 4-mer (7%), 5-mer (6.67%), 6-mer (6.33%), 7-mer (6%), 8-mer (5.67%), 9-mer (5.33%), 10-mer (5%), 11-mers (4.67%), 12-mer (4.33%), 13-mer (4%), 14-mer (3.67%), 15-mer (3.33%), 16-mer (3%), 17-mer (2.67%), 18-mer (2.33%), 19-mer (2%), 20-mer (1.67%), 21-mer (1.33%), 22-mer (1%), 23-mer (0.67%), 24-mer (0.33%) [1200 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 5, variation in photon count = 20%, lateral sigma = 130 nm, frame size = 1,024 pixels x 1,024 pixels
	Columns FT to FX / dataset # 37
Figure 4g	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 1-mer (28.57%), 3-mer (23.81%), 5-mer (19.10%), 7-mer (14.29%), 9-mer (9.52%) and 11-mer (4.76%) [168 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 5, variation in photon count = 20%, lateral sigma = 130 nm, frame size = 1,024 pixels x 1,024 pixels
	Columns EV to EZ (rows 22 to 34) / dataset # 33 copy (2)
Figure 4h	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie]

	Unknown: 1-mer (22.22%), 3-mer (19.44%), 5-mer (16.67%), 7-mer (13.89%) 9-mer (11.11%), 11-mer (8.33%), 13-mer (5.56%) and 15-mer (2.78%) [288 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 5, variation in photon count = 20%, lateral sigma = 130 nm, frame size = 1,024 pixels x 1,024 pixels
	Columns FZ to GD / dataset # 38
Figure 4i	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 1-mer (18.18%), 3-mer (16.36%), 5-mer (15.55%), 7-mer (12.73%) 9-mer (10.91%), 11-mer (9.09%), 13-mer (7.27%), 15-mer (5.45%), 17-mer (3.64%) and 19-mer (1.82%) [440 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 5, variation in photon count = 20%, lateral sigma = 130 nm, frame size = 1,024 pixels x 1,024 pixels
	Columns GF to GJ / dataset # 39
Figure 4j	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 1-mer (15.38%), 3-mer (14.10%), 5-mer (12.82%), 7-mer (11.54%) 9-mer (10.26%), 11-mer (8.97%), 13-mer (7.69%), 15-mer (6.41%), 17-mer (5.13%), 19-mer (3.84%), 21-mer (2.56%) and 23-mer (1.28%) [624 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 5, variation in photon count = 20%, lateral sigma = 130 nm, frame size = 1,024 pixels x 1,024 pixels
	Columns GL to GP / dataset # 40
Figure 4k	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 1-mer (40%), 4-mer (30%), 7-mer (20%) and 10-mer (10%) [120 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 5, variation in photon count = 20%, lateral sigma = 130 nm, frame size = 1,024 pixels x 1,024 pixels
	Columns GR to GV / dataset # 41
Figure 4l	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 1-mer (31.37%), 4-mer (25.49%), 7-mer (19.61%), 10-mer (13.73%), 13-mer (7.84%) and 16-mer (1.96%) [204 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 5, variation in photon count = 20%, lateral sigma = 130 nm, frame size = 1,024 pixels x 1,024 pixels
	Columns GX to HB / dataset # 42
Figure 4m	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 1-mer (25.97%), 4-mer (22.08%), 7-mer (18.18%), 10-mer (14.29%), 13-mer (10.39%), 16-mer (6.49%) and 19-mer (2.60%) [308 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 5, variation in photon count = 20%, lateral sigma = 130 nm, frame size = 1,024 pixels x 1,024 pixels
	Columns HD to HH / dataset # 43
Figure 4n	Calibration: 1-mer (50%), 2-mer (25%), 3-mer (25%) [80 particles per movie] Unknown: 1-mer (22.22%), 4-mer (19.44%), 7-mer (16.67%), 10-mer (13.89%), 13-mer (11.11%), 16-mer (8.33%), 19-mer (5.56%), 22-mer (2.78%) [432 particles per movie]
	# frames = 500, # movies = 10, max. photon count = 5, variation in photon count = 20%, lateral sigma = 130 nm, frame size = 1,024 pixels x 1,024 pixels
	Columns HJ to HN / dataset # 44

Supplementary Table S1 Information on the stoichiometric configurations and experimental conditions simulated in this study.

References

(1) Hirsch, M.; Wareham, R. J.; Martin-Fernandez, M. L.; Hobson, M. P.; Rolfe, D. J. A Stochastic Model for Electron Multiplication Charge-Coupled Devices – From Theory to Practice. *PLOS ONE* **2013**, *8* (1), e53671. DOI: 10.1371/journal.pone.0053671.
(2) Danial, J. S. H.; Shalaby, R.; Cosentino, K.; Mahmoud, M. M.; Medhat, F.; Klenerman, D.; Garcia Saez, A. J. DeepSinse: deep learning-based detection of single molecules. *Bioinformatics* **2021**. DOI: 10.1093/bioinformatics/btab352 (acccessed 7/17/2021).
(3) Przybylski, A.; Thiel, B.; Keller-Findeisen, J.; Stock, B.; Bates, M. Gpufit: An open-source toolkit for GPU-accelerated curve fitting. *Scientific Reports* **2017**, *7* (1), 15722. DOI: 10.1038/s41598-017-15313-9.
(4) Thevathasan, J. V.; Kahnwald, M.; Cieśliński, K.; Hoess, P.; Peneti, S. K.; Reitberger, M.; Heid, D.; Kasuba, K. C.; Hoerner, S. J.; Li, Y.; et al. Nuclear pores as versatile reference standards for quantitative superresolution microscopy. *bioRxiv* **2019**, 582668. DOI: 10.1101/582668.
(5) Subburaj, Y.; Cosentino, K.; Axmann, M.; Pedrueza-Villalmanzo, E.; Hermann, E.; Bleicken, S.; Spatz, J.; García-Sáez, A. J. Bax monomers form dimer units in the membrane that further self-assemble into multiple oligomeric species. *Nature Communications* **2015**, *6* (1), 8042. DOI: 10.1038/ncomms9042.
(6) Gómez-Sánchez, R.; Rose, J.; Guimarães, R.; Mari, M.; Papinski, D.; Rieter, E.; Geerts, W. J.; Hardenberg, R.; Kraft, C.; Ungermann, C.; et al. Atg9 establishes Atg2-dependent contact sites between the endoplasmic reticulum and phagophores. *The Journal of cell biology* **2018**, *217* (8), 2743-2763. DOI: 10.1083/jcb.201710116 PubMed.