Supporting Information to

Quantifying the Oligomeric States of Membrane Proteins in Cells through Super-resolution Localizations

Xihong Xie[†], Yu-Shan Cheng[†], Meng-Hsuan Wen, Aparna Calindi, Karen Yang, Chi-Wei Chiu, Tai-Yen Chen[*]

Department of Chemistry, University of Houston, Houston, Texas 77204, United States.

[†] Equal contribution. [*] Correspondence: tchen37@central.uh.edu

Supplementary Results

Table of content

1. Schematic overview of experimental strategy

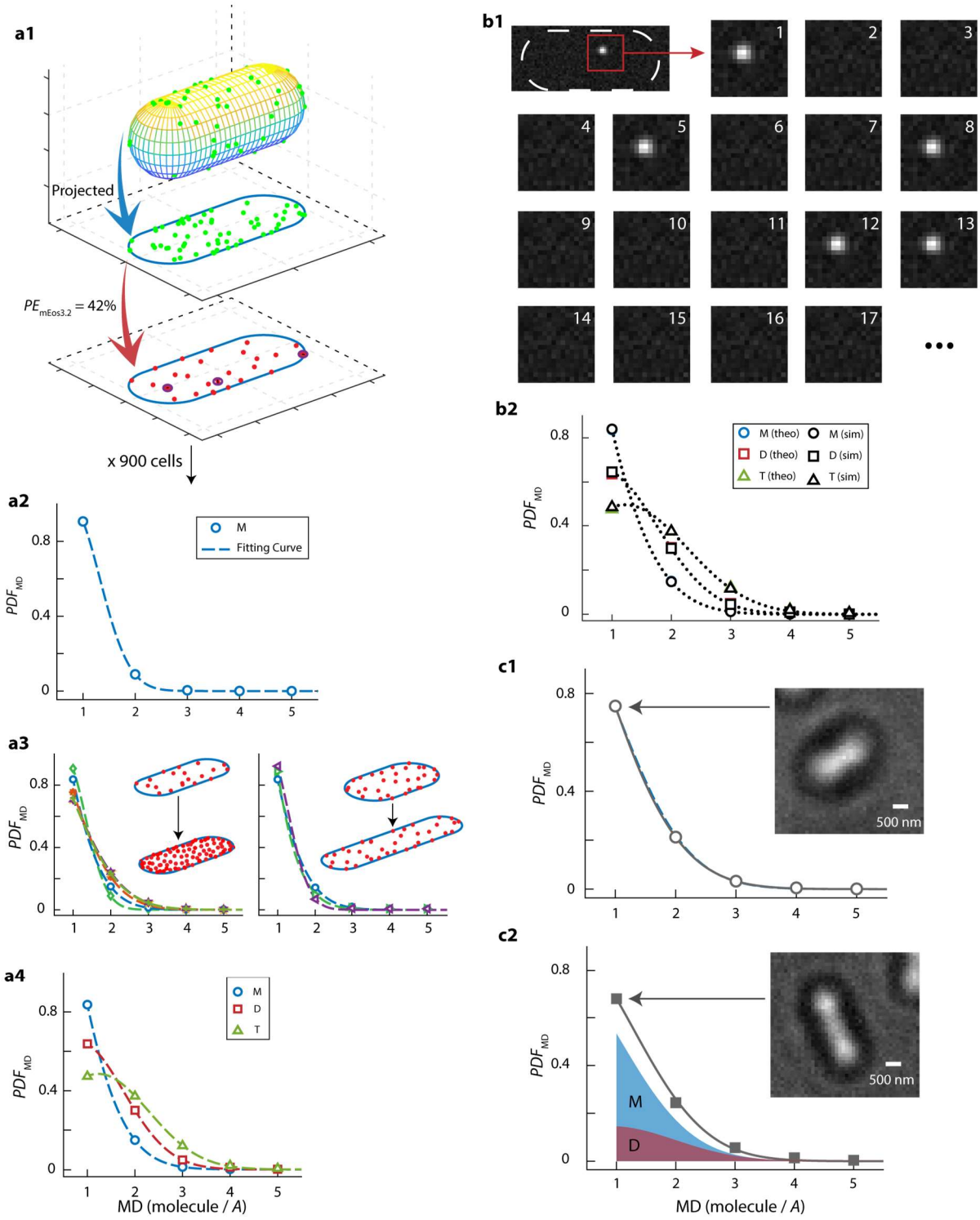


Fig. S1. Schematic overview of quantifying the oligomeric state of membrane proteins in cells. (**a1**) Simulation of localizations of membrane proteins in cells. (**a2**) Estimation of PDF_{MD}. (**a3**) Effect of protein concentration and cellular geometry on PDF_{MD}. (**a4**) Fitting PDF_{MD} with the truncated Gaussian model quantifies different oligomeric states. (**b1**) Simulation of PALM images with photophysics of mEos3.2. (**b2**). Validation of analysis method by comparing the theoretical and simulated PDF_{MD}. (**c1**) Application of PDF_{MD} analysis to monomeric membrane proteins UhpT. (**c2**) Application of PDF_{MD} analysis to dimeric membrane proteins SbmA.

2. Construction of plasmids and strains

2.1 Molecular cloning of plasmids

UhpT gene was amplified from *E. coli* strain DH5α by colony PCR with primers UhpT_F and UhpT_R (**Table S1**). First, the PCR fragments containing *UhpT* were inserted into pBADmEos3.2 plasmids (a gift from Dr. Peng Chen) between EcoRI (NEB R3101S) and NheI (NEB R3131S) sites to generate the coding sequence for the UhpT-mEos3.2 fusion protein. The obtained construct was digested by NheI, blunted by Klenow (NEB M0210S) and digested by PstI (NEB R3140S) to restore the Shine-Dalgarno (SD) sequence. The coding region was further subcloned back to the original pBADmEos3.2 plasmid between blunted-EcoRI site and Pst-I site, which resulted in the final construct, pUhpTmE, with a proper protein expression property (**Fig. S2a**). The pSbmAmE construct was created through the same procedures (**Fig. S2b)** starting with primers, SbmA_F and SbmA_R. All constructs were confirmed by DNA sequencing. All the restriction enzymes were purchase from New England Biolabs.

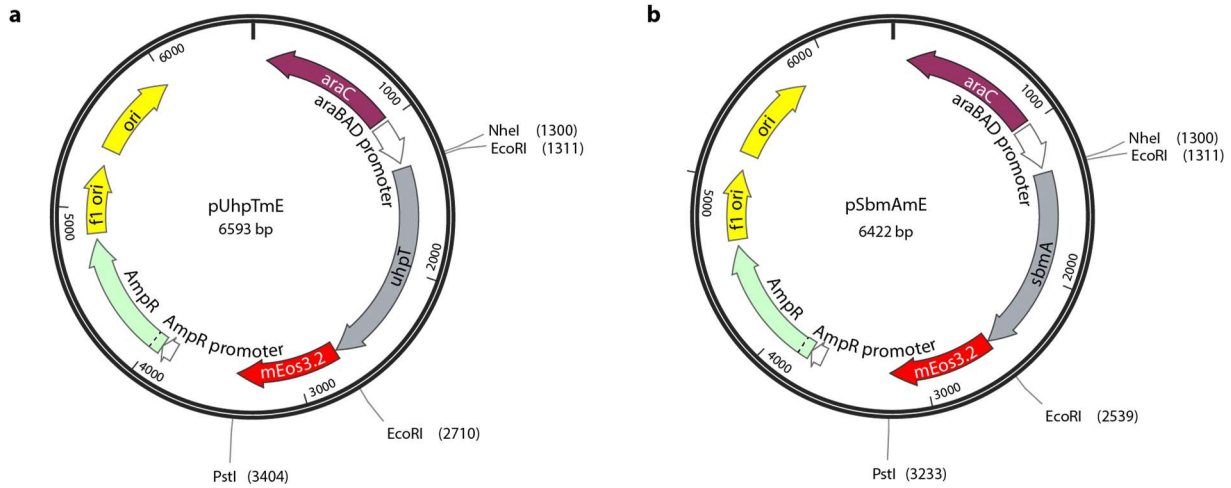


Fig. S2. Map of constructed plasmids. (**a**) This plasmid contains the gene of the target protein UhpT, gene of fluorescent protein mEos3.2, ampicillin resistant gene *AmpR*, and the origin of replication, *ori*, as well as restriction sites NheI, EcoRI, and PstI. (**b**) The same as (**a)** but for target protein SbmA.

Table S1. Primers used for cloning in this study

Primer Name	**Sequence (5'-3')**
1. UhpT_F	CTAGC TAG CAT GCT GGC TTT CTT AAA CCA G
2. UhpT_R	TGA ATT CTG CCA CTG TCA ACT GCT G
3. SbmA_F	CTAGC TAG CAT GTT TAA GTC TTT TTT CCC AAA G
4. SbmA_R	GGA ATT CGC TCA AGG TAT GGG TTA CTT C

Table S2. Plasmids used or constructed in this study

Plasmid Name	**Vector Backbone**	**Gene Insert**	**Resistance**
1. pBADmEos3.2	pBAD24	*mEos3.2*	Amp
2. pUhpTmE	pBAD24	*uhpT-mEos3.2*	Amp
3. pSbmAmE	pBAD24	*sbmA-mEos3.2*	Amp

Table S3. Strains used or constructed in the current study

Strain Name	Plasmid	Chromosomal Gene Tag/Modification	Resistance
1. BW25113	none	*none (experimental wild-type)*	none
2. JW0368-1	none	*ΔsbmA*	Kan
3. U-BW25113	pUhpTmE	*none*	Amp
4. S-JW0368-1	pSbmAmE	*ΔsbmA*	Amp, Kan

2.2 Introduce plasmids into E. coli cells through electroporation

To prepare electrocompetent cells, BW25113 (CGSC7739 Keio Collection, Yale; genotype: F- *Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), λ-, rph-1, Δ(rhaDrhaB)568, hsdR514*) strain and JW0368-1(CGSC8547 Keio Collection, Yale; genotype: F- *Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), ΔsbmA742::kan, λ-, rph-1, Δ(rhaD-rhaB)568, hsdR514*) strain were first grown in 3 mL Luria-Bertani (LB) broth overnight, then subcultured 50 µL of the overnight culture to 5 mL fresh LB media at 37 °C until OD_{600} ~0.6. The cells were pelleted by centrifugation (4 °C, 3,000 g, 5 min). The pellets were washed with chilled 10 mL of 10% glycerol for four times and finally resuspended in 100 µL 10% glycerol. The plasmid pUhpTmE was transformed into the electrocompetent BW25113 strain while pSbmAmE was transformed into electrocompetent the JW0368-1 strain by electroporation (2.5 kV, Gene Pulser XcellTM Total System, BIO-RAD 1652660). After recovery with SOC media (NEB B9020S), the cells were plated on LB agar containing carbenicillin (100 µg/mL), a chemically stable substitute for ampicillin.

2.3 Intactness of mEos3.2-tagged proteins in E. coli

The western blot was performed to verify the expression of UhpT-mEos3.2 and SbmA-mEos3.2 as an intact fusion protein in cells. BW25113, JW0368-1, U-BW25113 and S-JW0368-1 strains were amplified overnight in LB with corresponding antibiotics. The cultures were diluted 1:100 in M9 media ($[Na_2HPO_4]$ = 33.7 mM, $[KH_2PO_4]$ = 22.0 mM, [NaCl] = 8.55 mM, $[NH_4Cl]$ = 9.35 mM) containing 8% amino acids (Fisher Scientific 11-130-051), 4% vitamins (Fisher Scientific 11-120-052) and 0.4% glycerol and incubated at 37 °C with shaking for 4 h. The culture was continued for two h at 22 °C after adding L-arabinose to a final concentration of 1 mM. The volume of 1 mL culture was pelleted (2,000 g, 5 min) and lysed by 1X sample buffer (BIO-RAD 161-0737) containing 1% 2-mercaptoethanol (Sigma M3148). The lysis was heated at 95 °C for 5 min. The TGX gel (BIO-RAD 456-8094), with the samples and the protein marker (BIORAD 1610373) loaded, was run for 80 min under 100 V in 1X Tris/Glycine/SDS buffer (BIO-RAD 161-0732). The proteins in the gel were transferred to a PVDF membrane in a transfer pack (BIO-RAD 1704156) by the Trans-Blot® Turbo™ Transfer System (25 V, 10 min, BIO-RAD 1704150). The membrane was blocked with blocking solution (2% Bovine Serum Albumin (Sigma A7906) in 1X PBS plus 0.05% Tween 20 (Fisher Scientific BP337-100)) for 1 h and incubated with rabbit-derived primary antibody mEOS2 pAb (Badrilla A010-mEOS2) solution for 2 h at 22 °C where the primary antibody was diluted 1:2,000 in the blocking solution. The membrane was rinsed in the washing solution (1X PBS plus 0.05% Tween 20), and further incubated with the secondary antibody, goat-derived Anti-Rabbit IgG (Jackson ImmonoResearch 111-035-045), in a dilution of 1:20,000 in the blocking solution for one h. After rinsing in the washing solution, the membrane was treated with the ECL substrate (BIO-RAD 170-5060) and loaded onto the ChemiDoc Touch Imaging System (BIO-RAD 1708371, **Fig. S3**). To estimate the percent cleavage of UhpT and SbmA, we compared the line profiles of each condition. The contribution of non-specific binding from the parental cell was first removed mathematically. We then estimated the percent cleavage of UhpT and SbmA using the integrated area at the correct molecular weight positions. In both cases, the percent cleavage is less than 10%.

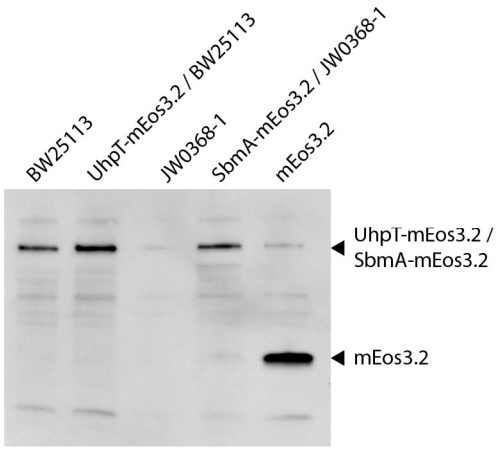


Fig. S3. Western Blot of mEos3.2-tagged UhpT, SbmA shows both tagged proteins are mostly intact in the cell.

3. Imaging setup, conditions, and analysis

3.1 Imaging sample preparation

The strains contained target plasmids were grown overnight in LB broth with carbenicillin as described above. The culture was subcultured 1:100 into M9 media supplemented with 8% amino acids, 4% vitamins and 0.4% glycerol and incubated at 37 °C with shaking for four h. To express the proteins, we cultured bacteria with one mM L-arabinose at 22 °C. After two h, two mL of cells were pelleted by centrifugation at 3,500 g for 5 min and washed twice with one mL 1X PBS (VWR 97062-948). The cells were fixed by incubating in 200 µL 4% paraformaldehyde PBS solution at 22 °C for 40 min. After washed twice with 1X PBS, the cell pellets were resuspended in 80 µL 1X PBS.[1]

Slides and coverslips used for imaging were thoroughly cleaned and dried by the N_2 gas.[2] The center of the slide was cast with a 3% agarose pad. A volume of 3 µL sample was loaded on the agarose pad and covered with a coverslip which was pre-coated with 10 µL of 20-fold diluted gold nanoparticle (TED PELLA, 15708-9) ethanol solution. The edges of coverslip were sealed with double-sided tapes and epoxy glues.

3.2 Imaging conditions and analysis

PALM images were collected using an Olympus IX83 inverted microscope equipped with a scientific CMOS camera (Photometric Prime 95B) and a 100X TIRF oil immersion objective (Olympus UApoN 100X TIRF) (**Fig. S4**). The images were magnified by 100X. The final image pixel size was 110 nm calibrated by a 0.01mm Microscope Stage Calibration Slide (AmScope MR095). In the excitation path, two laser beams were spatially overlapped using dichroic filters (Chroma 315126) and passed a quarter waveplate (Thorlabs AQWP05M-600) to ensure circularly polarized excitation. All laser beams were expanded five times by an achromatic lens pair and focused (400 mm lens) at the back focal plane of the objective before being reflected toward the objective by a three-band dichroic filter (Chroma C180399) inside the Olympus filter cube. Objective-collimated lasers continuously illuminated the cell in a pseudo-TIRF mode whose beam size at the sample plane is ~30 µm × 40 µm (FWHM). A 405 nm laser (Coherent 1284371) with a power density of 0.0115 W/cm^2 and a 561 nm laser (Coherent 1280720) with a power density of 6.5 kW/cm^2 (or k_b = 35 ± 5 s^{-1}) activated and excited the fluorescent proteins, respectively.

In the detection path, the red-fluorescence of single mEos3.2-tagged molecules passed through the three-band dichroic filter and the red band-pass filter (Semrock FF01-630/92-25) before entering the

camera. A region of the camera with ~ 300 × 300-pixel was used during data acquisition with 30 ms exposure time for 20,000 cycles. The synchronization between camera and lasers was through the Real-Time Controller (Olympus U-RTCE), and the imaging protocol was controlled via the cellSens v.1.14 software. The acquired images were processed using a home-built software, iqPALM, as described in our previous work to extract the protein localizations with ~20 nm spatial resolution.[1]

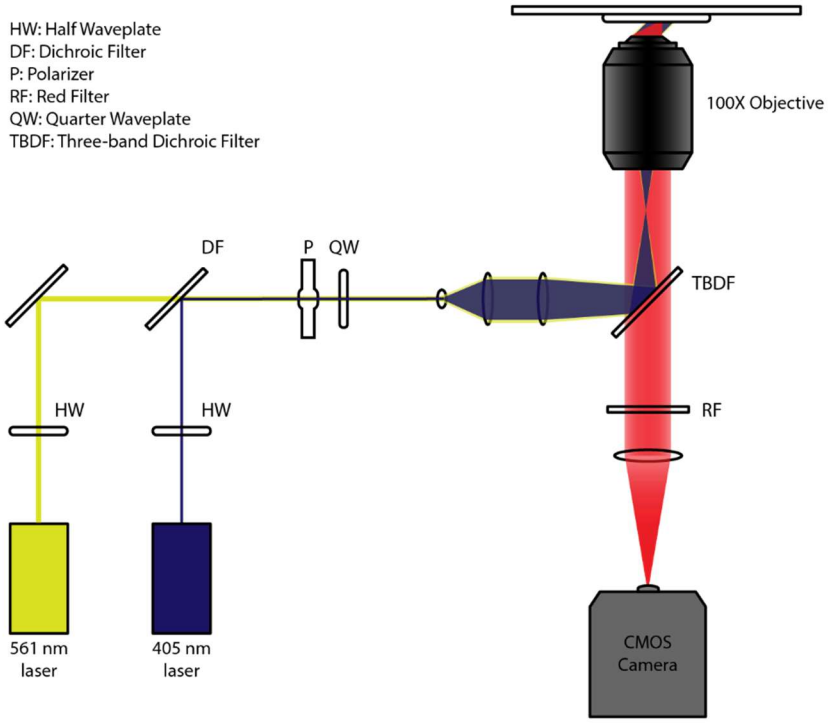


Fig. S4. Schematic diagram of microscope setup. 405 nm and 561 nm lasers were spatially overlapped by using a dichroic filter and passed through quarter waveplate to ensure circularly polarized excitation. Half-wave plates and polarizer controls the incident power densities. The laser beams were focused at the back focal plane of the objective before reflected by a three-band dichroic filter. In the detection path, the red-fluorescence from single mEos3.2-tagged molecules passed through the three-band dichroic filter and the red band-pass filter before entering the camera.

4. Statistically saturated PDF_{MD} requires averaging >200 cells

Since the PDF_{MD} was estimated from MD of multiple cells, it was important to ensure the PDF_{MD} was statistically saturated. We first generated PDF_{MD} with different numbers of cells (N_c = 10 to 900) and different total numbers of molecules (N_t = 150 to 4,800). In most conditions, the PDF_{MD} reached a stable distribution when N_c = 900. We then used PDF_{MD} (N_c = 900) as our saturation reference and calculated the percentage error, $\frac{PDF_{MD}(N_c = 1{:}900) - PDF_{MD}(N_c = 900)}{PDF_{MD}(N_c = 900)} \times 100\%$, to probe the minimum cell number requirements for statistically saturated PDF_{MD}. When the percentage error was less than 5%, PDF_{MD} was considered to reach its statistical saturation.

We examined how the percent error of PDF_{MD} correlated with the N_t and N_c among proteins with different oligomeric states (**Fig. S5a**). The result showed that the statistical saturation condition of PDF_{MD} depended on both N_t and N_c. In general, the error decreased when averaging more cells and when N_t is higher. Under similar N_t and N_c, proteins with higher oligomeric state gives larger error than the one with lower oligomeric states. In this paper, the experimental results were derived from N_t ~1,200. To figure out the saturation requirement to determine the oligomeric state, we analyzed error of PDF_{MD} at N_t = 1,200. (**Fig. S5b**). Under this condition, oligomeric states of monomer to hexamer can reach saturated condition when N_c is more than 200. For samples with lower N_t, monomeric, dimeric and trimeric proteins still can

be saturated with more than 200 cells, but protein with higher oligomeric state requires more cells to reach the same condition, that is, more than ~300 for tetrameric proteins, and over 700 cells for pentameric and hexameric proteins (**Fig. S5c**).

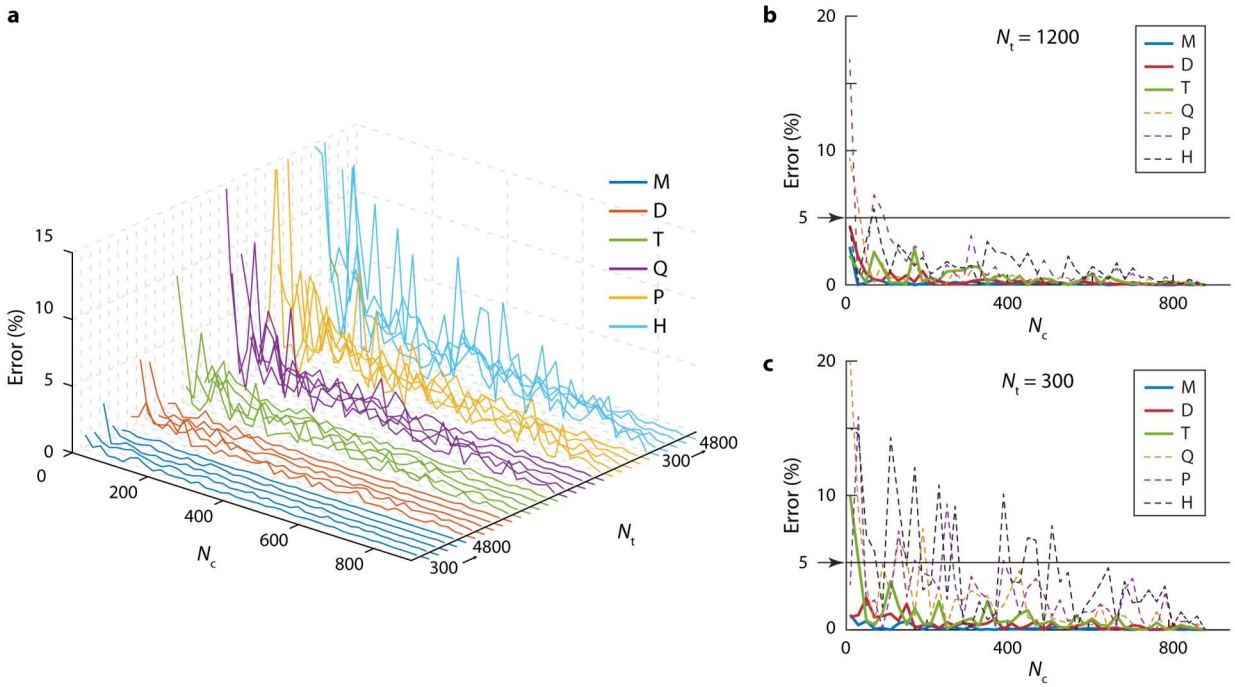


Fig. S5. Statistical saturation requirement analysis. (**a**) Percentage error as a function of N_c at different N_t of 300 to 4800. (**b**) Percentage error as a function of N_c at N_t = 1200 for monomeric to hexameric proteins. (**c**) Same as **b** but N_t = 300.

5. Simulation of the single-molecule imaging data with photophysics of mEos3.2 integrated

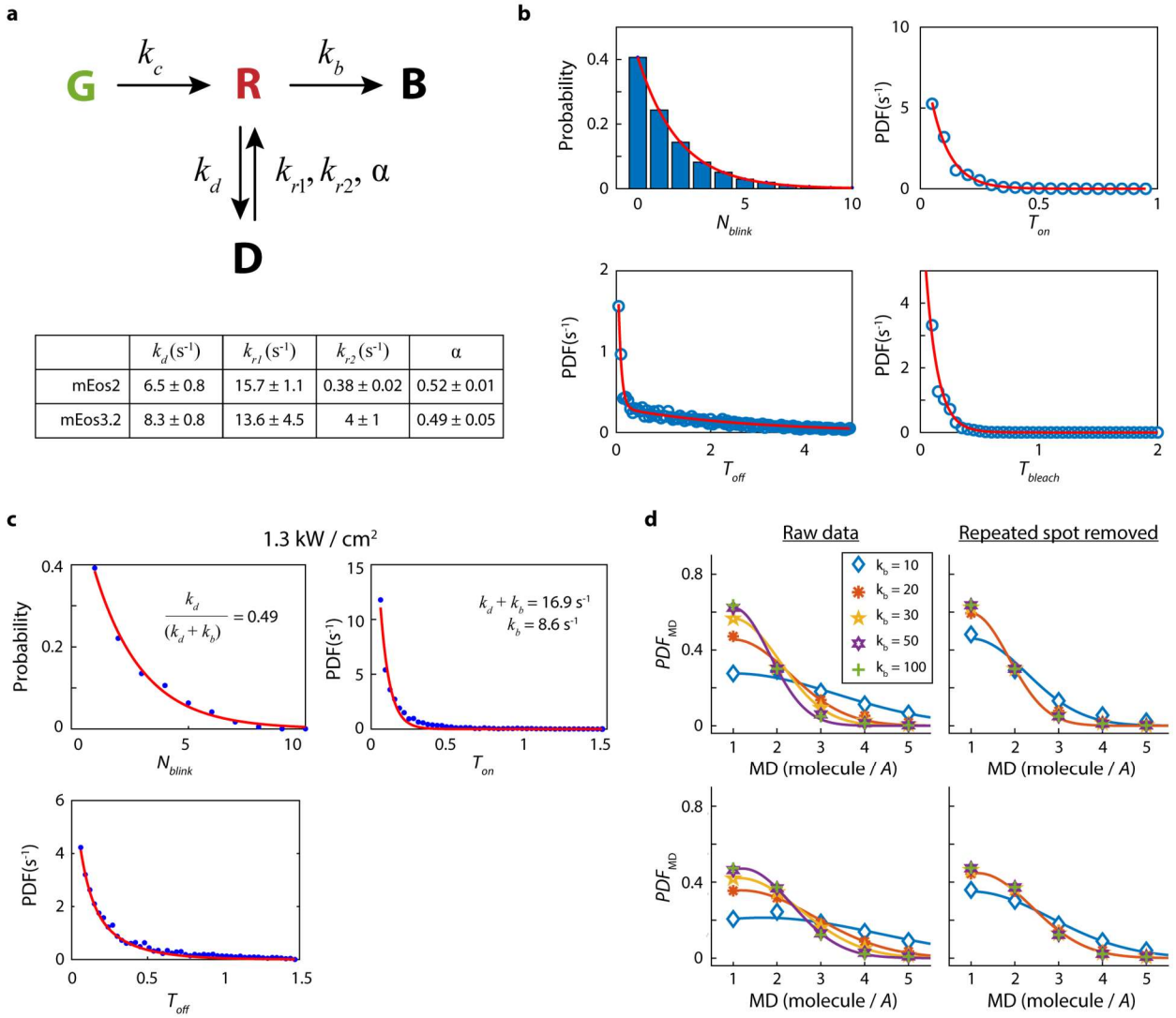
5.1 Photophysics of mEos3.2

Assuming the photo-conversion of mEos3.2 is a first-order chemical reaction with conversion rate constant k_c, we can perform a probability experiment at every observation time interval, ΔT, with a probability seeing a successful conversion of $k_c\Delta T$. For example, after generating a random number $0 \leq x < 1$, one can mark the conversion when $x < k_c\Delta T$ or otherwise moved to the next time interval for next trial. This step was repeated until achieving the targeted number of conversion. For each conversion event, a localization of the membrane monomer was randomly assigned to complete the generation of the sequence of photo-conversion event for labeled membrane proteins.

Photophysics (including photoblinking and bleaching) of mEos3.2 was implanted to each conversion event using the two-state model of mEos2 (a close-related mEos variant of mEos3.2, Error! Reference source not found.**a**) as described in the previous work by Lee *et al*.[8] In short, once photoconverted, mEos3.2 can move to the dark or bleaching state with rate constants of k_d and k_b, respectively. If at the dark states, mEos3.2 can return to the bright state through two different pathways of rate constants (k_{r1} + αk_{r2}), where α indicates the relative contribution between return pathways. By looping through the sequence of conversion event, we implanted the photophysics of mEos3.2 to the simulated data and filled in the original localization information for all blinking events.

Using the literature reported k_b, k_d, k_{r1}, k_{r2}, and α of mEos2 (an earlier variant of mEos3.2), we reproduced distributions of the number of times the molecule blinks (N_{blink}), the fluorescence-on time (T_{on}),

the fluorescence-off time (T_{off}), and the photobleaching time (T_{bleach}). Fitting these distributions recovered the k_b, k_d, k_{r1}, k_{r2} and α within 5% error, confirming that the simulated images faithfully reproduced the photophysics of mEos2.



	k_d(s^{-1})	k_{r1}(s^{-1})	k_{r2}(s^{-1})	α
mEos2	6.5 ± 0.8	15.7 ± 1.1	0.38 ± 0.02	0.52 ± 0.01
mEos3.2	8.3 ± 0.8	13.6 ± 4.5	4 ± 1	0.49 ± 0.05

Fig. S6. Photophysics of mEos3.2. (**a**) The kinetic model for mEos3.2 protein in different states. The green (G) form of mEos3.2 can be converted to the red (R) form of mEos3.2 with rate k_c when under 405 nm laser illumination. Under 561 nm laser, red form of mEos3.2 emits light and can move to the dark or bleaching state with rate constants of k_d and k_b respectively. If at the dark states, mEos3.2 can return to the bright state through two different pathways of rate constants (k_{r1} + αk_{r2}), where α indicates the relative contribution between return pathways. Bottom: summary of the photophysical rate constants extracted from **b** (mEos2) and **c** (mEos3.2) at 561 nm laser power density of 1.3 kW/cm^2. (**b**) Distributions of the number of times the molecule blinks (N_{blink}), the fluorescence-on time (T_{on}), the fluorescence-off time (T_{off}), and the photobleaching time (T_{bleach}) of mEos2. (**c**) Experimentally results of distributions of N_{blink}, T_{on}, and T_{off} of mEos3.2 at 561 nm laser power density of 1.3 kW/cm^2. (**d**) Over-counting effect on the PDF_{MD} of dimeric (top row) and trimeric proteins (bottom row).

Using the same method as Lee *et al.*, we have extracted the k_d, k_b, k_{r1}, k_{r2}, and α for mEos3.2 experimentally (**Fig. S6c**) under 561 nm laser power density of 1.3 kW/cm^2. Under 1.3 kW/cm^2 excitation density, the fitted rate constants are comparable to the reported rate constants of mEos2. Based on the k_d + k_b = 16.9 s^{-1} from the T_{on} curve and $k_d/(k_d + k_b)$ = 0.49 from the N_{blink} curve, we estimated the k_b of 8.6 s^{-1}, which is comparable to the value of mEos2 at similar excitation power density. We then performed the

same experiments under five times higher laser excitation power density of 6.5 kW/cm^2. The extracted k_b is 35 ± 5 s^{-1}, which is slightly lower than the k_b (i.e., 8.6 × 5 = 43 s^{-1}) linearly extrapolated from that of 1.5 kW/cm^2.

We then examined the over-counting effect on the PDF_{MD} of monomeric, dimeric, and trimeric proteins using the simulated data. A more substantial deviation of PDF_{MD} from the theoretical prediction was observed for proteins with higher oligomeric states under low k_b (i.e., k_b = 10 s^{-1}, **Fig. S6d**). However, as the k_b > 50 or 20 (with repeated spots removed) s^{-1}, PDF_{MD} got back to its theoretical form regardless of the oligomeric states. All the experimental determined super-resolution images in this study are obtained under laser excitation power density of 6.5 kW/cm^2 (i.e., k_b = 35 s^{-1}) to avoid the over counting error.

5.2 Regardless of the cluster radius r value, PDF_{MD} is effective to differentiate membrane proteins with different oligomeric states

To probe how the threshold r might change the feasibility of using PDF_{MD} to differentiate different oligomeric states of membrane proteins in the cell, we thresholded the locations of molecules with r = 20, 40, and 60 nm and determined the PDF_{MD} for monomeric, dimeric, and trimeric membrane proteins. **Fig. S7** shows the PDF_{MD} using the fitting results of simulated single-molecule movies from 900 cells with membrane proteins existing in pure monomeric, dimeric, and trimeric states and the number of molecule of 300. Regardless of r value, the peak of PDF_{MD} shifted to higher molecule density as membrane proteins with the higher oligomeric state, suggesting that the effect of r value can be neglected while decoding the oligomeric state of membrane protein in cells.

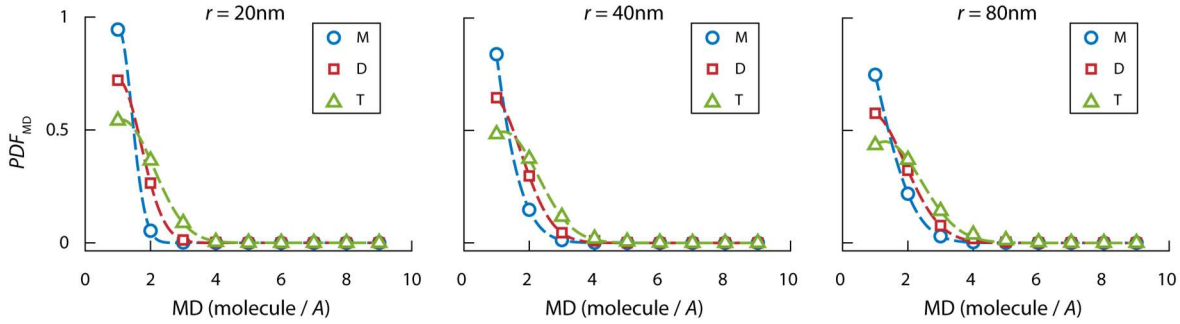


Fig. S7. Oligomeric-state dependent PDF_{MD} with different r. The same trend is observed for r varies from 20 to 60 nm, suggesting that PDF_{MD} is effective to differentiate membrane proteins with different oligomeric states.

6. Selecting *E. coli* cells with membrane proteins uniformly distributed on the membrane surface using randomness index *R*

The quantification of the spatial distribution of discrete objects played an essential role in quantifying the oligomeric state of membrane proteins in this study. Here we adapted the approach[9] developed by Dry *et al.* to estimate the randomness of our data. In short, for a set of n locations, the mean nearest neighbor distance was first calculated through $r_0 = \frac{1}{n}\sum_{i\neq j}^{n} min\{u_{ij}\}$ where u_{ij} is the distance between the i^{th} and j^{th} points. The complete spatial randomness for n points in an area A was typically described by the Poisson process, and the mean distance between randomly distributed points can be defined through $r_E = 0.5\sqrt{A/n}$. We can then define the randomness index R through $R = \frac{r_0}{r_E}$.

After extracting locations from experimental fluorescent images, we first selected cells with similar cellular geometry and protein concentrations. We then used the averaged cellular geometry and protein

concentration to generate the corresponding simulated data describing molecules randomly distributed over the cell surface. The simulated results were then used to estimate the proper randomness index *R*. The *R* was applied to remove cells showing heterogeneous molecule distributions. Finally, the selected cells were combined to generate the final PDF_{MD} for the determination of protein oligomeric state.

7. References

1. Chen, T.-Y.; Santiago, A. G.; Jung, W.; Krzemiński, Ł.; Yang, F.; Martell, D. J.; Helmann, J. D.; Chen, P., Concentration-and chromosome-organization-dependent regulator unbinding from DNA for transcription regulation in living cells. *Nature communications* **2015,** *6*, 7445.
2. Ha, T., Single-molecule fluorescence resonance energy transfer. *Methods* **2001,** *25* (1), 78-86.
3. Chen, T.-Y.; Jung, W.; Santiago, A. G.; Yang, F.; Krzemiński, Ł.; Chen, P., Quantifying multistate cytoplasmic molecular diffusion in bacterial cells via inverse transform of confined displacement distribution. *J. Phys. Chem. B* **2015,** *119* (45), 14451-14459.
4. Javer, A.; Long, Z.; Nugent, E.; Grisi, M.; Siriwatwetchakul, K.; Dorfman, K. D.; Cicuta, P.; Lagomarsino, M. C., Short-time movement of E. coli chromosomal loci depends on coordinate and subcellular localization. *Nat. Commun.* **2013,** *4*, 3003.
5. Durisic, N.; Laparra-Cuervo, L.; Sandoval-Álvarez, Á.; Borbely, J. S.; Lakadamyali, M., Single-molecule evaluation of fluorescent protein photoactivation efficiency using an in vivo nanotemplate. *Nat. Methods* **2014,** *11* (2), 156.
6. Annibale, P.; Scarselli, M.; Greco, M.; Radenovic, A., Identification of the factors affecting co-localization precision for quantitative multicolor localization microscopy. *Optical Nanoscopy* **2012,** *1* (1), 9.
7. Frost, N. A.; Lu, H. E.; Blanpied, T. A., Optimization of cell morphology measurement via single-molecule tracking PALM. *PLoS One* **2012,** *7* (5), e36751.
8. Lee, S.-H.; Shin, J. Y.; Lee, A.; Bustamante, C., Counting single photoactivatable fluorescent molecules by photoactivated localization microscopy (PALM). *Proc. Natl. Acad. Sci. U.S.A.* **2012,** *109* (43), 17436-17441.
9. Dry, M. J.; Preiss, K.; Wagemanns, J., Clustering, randomness and regularity: Spatial distributions and human performance on the traveling salesperson problem and minimum spanning tree problem. **2012**.