Determination of protein stoichiometries via dual-color colocalization with single molecule localization microscopy

Hua Leonhard Tan[a,b], Stefanie Bungert-Plümke[a], Daniel Kortzak[a], Christoph Fahlke[a], and Gabriel Stölting[c,a,1]

[a]Institute of Biological Information Processing, Molecular and Cellular Physiology (IBI-1), Forschungszentrum Jülich, 52428 Jülich, Germany; [b]Department of Biology, University of Maryland College Park, MD 20742, USA; [c]Berlin Institute of Health at Charité – Universitätsmedizin Berlin, Hypertension and Molecular Biology of Endocrine Tumors, Charitéplatz 1, 10117 Berlin, Germany

[1] To whom correspondence should be addressed. gabriel.stoelting@bih-charite.de

The stoichiometry of plasma membrane protein complexes is an important determinant of their function and of interactions between the individual proteins. Most approaches used to address this question rely on extracting these complexes from their native environment, which may disrupt weaker interactions. Therefore, microscopy techniques have been increasingly used in recent years to determine protein stoichiometries in situ. Classical light microscopy suffers from insufficient resolution, but super-resolution methods such as single molecule localization microscopy (SMLM) can circumvent this problem. When using SMLM to determine protein stoichiometries, subunits are labeled with fluorescent proteins that only emit light following activation or conversion at different wavelengths. Typically, individual signals are counted based on a binomial distribution analysis of emission events detected within the same diffraction-limited volume. This strategy requires low background noise, a high detection efficiency for the fluorescent tag and intensive post-imaging data processing. To overcome these limitations, we developed a new method based on SMLM to determine the stoichiometry of plasma membrane proteins. Our dual-color colocalization (DCC) approach allows for accurate in situ counting even with low efficiencies of fluorescent protein detection. In addition, it is robust in the presence of background signals and does not require strong temporal separation of emission events within the same diffraction-limited volume, which greatly simplifies data acquisition and processing. We used DCC-SMLM to resolve the controversy surrounding the stoichiometries of two SLC26 multifunctional anion exchangers and to determine the stoichiometries of four members of the SLC17 family of organic anion transporters.

single molecule localization microscopy | fluorescent protein | protein stoichiometry

Most membrane proteins form oligomeric assemblies, in which individual subunits cooperate to fulfill specialized functions. Some membrane proteins also form larger clusters, in which distinct classes of proteins interact with each other. The number of subunits within such protein structures is a crucial determinant of their function (1). Native gel electrophoresis has been successfully used to determine subunit stoichiometries and the molecular determinants of oligomeric assemblies (2–4). However, this technique relies on isolating protein complexes from their native environment, which can disrupt some protein interactions. To overcome this limitation, *in situ* methods such as fluorescence correlation spectroscopy (5), stepwise photobleaching (6) and single molecule localization microscopy (SMLM) (7, 8) have become increasingly popular in recent years.

Stepwise photobleaching relies on the irreversible and stochastic photobleaching of fluorescent tags linked to the subunit to be counted, and the resulting intensity steps depend on the number of subunits within the complex being tested. Such data can be acquired using comparatively inexpensive microscopes which has undoubtedly led to the widespread adoption of this method (1, 6, 9, 10). However, this approach is only feasible for low expression densities as it is limited by the resolution of light microscopy. Therefore, *Xenopus laevis* oocytes, which are large in diameter and in which ectopic protein expression can be controlled by adjusting the amount of injected RNA, are most often used for this method (6). However, some proteins complexes exhibit a different stoichiometry when expressed in the *Xenopus* system compared with mammalian expression systems (11). For mammalian cells, the resolution limit of light microscopy can be circumvented using super-resolution microscopy methods (8, 12–17). SMLM achieves a higher resolution by reconstructing the precise positions of isolated fluorophore emissions detected over thousands of individual frames (18, 19). The fluorophores used for quantitative SMLM are most often photoactivatable or photoconvertible fluorescent proteins that are activated at wavelengths different from that of the excitation laser (20). Their positions in each frame may then be assigned to individual proteins in post-processing, to produce accurate information about the number and position of proteins subunits (8).

Several problems have to be overcome during image acquisition and processing to determine protein stoichiometries: To increase the lateral resolution in SMLM, the simultaneous activation of multiple fluorophores within the same diffraction-limited volume must be avoided. To assign signals to individual subunits within the same protein complex, the activation of each fluorophore needs even greater separation in time. These criteria can only be achieved if the probability of activating a certain protein is kept minimal through a gradually increased activation laser intensity. Algorithms to optimize laser control have been developed, but they are not intended to temporally separate clusters of emissions from individual proteins (13). During the long recordings required for SMLM, inherent mechanical and thermal fluctuations, as well as sample drift, combine to reduce spatial resolution. The complex photophysics of fluorescent protein blinking may also result in over- or undercounting (7, 13). Another source of miscounting is the background noise, which cannot be easily distinguished from genuine signals at the single molecule level (21, 22). Furthermore, it is impossible to detect all fluorophores in the sample because the probability of having a working fluorescent marker on an individual protein (also called the detection efficiency) is always less than one. This may be due to incomplete protein translation or folding, failure to activate into a fluorescent state, or pre-imaging bleaching (15, 17). Thus, the observed counts are not equal to the stoichiometry of interest but instead follow a binomial distribution that depends on both the real number of subunits per complex and the detection efficiency. Typically, only a detection efficiency of at least 0.5 ensures reasonable resolution between different stoichiometries (6, 15).

Here we developed a new strategy, dual-color colocalization SMLM (DCC-SMLM), which utilizes two spectrally different fluorescent proteins attached to the same protein of interest. DCC-SMLM is very tolerant of background noise and low detection efficiencies and does not require stringent control of the activation laser. We tested the new DCC-SMLM aproach on proteins of known stoichiometry and then used it to study proteins from two plasma membrane transporter families with controversial or previously unknown stoichiometries: The SLC26 and SLC17 families.

The human SLC26 family has 10 members that function as anion transporters or channels as well as the motor protein in cochlear outer hair cells. The physiological importance of these proteins is illustrated by the growth defects, chronic diarrhea or deafness associated with naturally occuring mutations.

Biochemical and recent structural studies indicated that the proteins exist as dimers (4, 23, 24), whereas other studies have proposed a tetrameric stoichiometry (25–27). The SLC17 family includes proton-driven vesicular glutamate, aspartate, sialic acid, and nucleotide transporters, as well as Na^+-coupled phosphate transporters. Using DCC-SMLM, we determined the oligomeric states for two SLC26 proteins, the human epithelial chloride-bicarbonate exchanger SLC26A3 and the rat motor protein prestin (Slc26a5), and for four members of the SLC17 family: mouse vGlut1, 2, and 3 (Slc17a7, Slc17a6, and Slc17a8, respectively) and human sialin (SLC17A5).

Results

Dual-Color Colocalization SMLM for Determining Stoichiometry. A protein complex of *n* identical subunits, each labeled with a single indicator fluorescent protein with a detection efficiency of *p* (the probability of being fluorescent and being detected), will be detected with a probability *P*:

$$P = 1 - (1 - p)^n \quad [1]$$

The direct determination of *P* from single-color microscopy images is impossible because the total number of proteins is unknown. Therefore, in order to experimentally determine *P*, we labeled each protein subunit with an additional fluorescent protein, *M*, as a baseline marker and determined *P* by colocalization of the marker with the indicator fluorescent protein (*F*) (Fig. 1 A and B). This approach allows us to measure the probability *P* by

$$P = \frac{N_{MF}}{N_M} \quad [2]$$

where N_M and N_{MF} are the number of protein complexes detected via the marker and via both the marker and indicator, respectively. The values of *P* obtained for proteins with known oligomeric states (*n*) were used to solve Eq. 1 for *p*, which could then be used to determine unknown stoichiometries (Fig. 1C).

We chose mVenus (28) as the fluorescent marker protein *M* because of the very low background signal in its detection channel (Fig. 1 D and E). mVenus is derived from yellow fluorescent protein with several improvements such as faster maturation, a monomeric structure, and a higher tolerance for cellular acidification and the surrounding halide concentration (28). Like other GFP derivatives, it blinks by cycling between the fluorescent and dark states under continuous illumination with the excitation laser (29). PAmCherry was chosen as the indicator because it is quickly bleached by continuous excitation after activation (15). However, this channel had a considerably higher background signal. PAmCherry fluorescence is stable for several frames after its activation (15, 30), resulting in small clusters of coordinates per fluorescent protein. We partially corrected for the background by filtering out clusters with only a few signals (Supplementary information, Fig. S1). However, some background signals were indistinguishable from genuine PAmCherry signals, which explains the higher background in that particular color channel (Fig. 1D). The order of mVenus and

PAmCherry tags did not significantly impact the probability of colocalization (Supplementary information, Fig. S2).

Correction of Sample Drift and Chromatic Aberrations. The resolution of all SMLM data is compromised by sample movement and optical aberrations. We recorded sample drift using TetraSpeck beads as fiducial markers and correcting for this during post-imaging data processing (Supplementary information) (31). This correction reduced the radius of clustered signals to ~30 nm for a single protein complex (Fig. 2 A and B). Since our approach relies on colocalization of the marker and indicator in two color channels, we also corrected chromatic aberrations resulting from different refractive indices at different emission light wavelengths. We confirmed earlier reports that lateral chromatic aberration can be described as a set of linear functions centered on the optical axis of the microscope (32, 33). Values from a linear fit of the measured inter-color distances of individual beads to their coordinates were used for correction (Fig. 2C). This way, lateral chromatic aberration could be reduced to 13.8 ± 8.9 nm (Fig. 2D). Chromatic aberrations may also be corrected by determining regional averages (34, 35). A simple algorithm that calculates and corrects the mean color shift in 10 × 10 squares over the field of view gave similar results to the linear fit (13.4 ± 8.5 nm; Supplementary information, Fig. S3). The axial chromatic aberration as determined by the focus offset between colors was less than 100 nm and thus considered irrelevant for SMLM signal registration in only two dimensions (Fig. 2E).

Definition of Colocalization. In SMLM, colocalization is defined as an inter-color distance below a certain threshold value. To determine this threshold, we calculated the distances between each individual mVenus cluster and its surrounding PAmCherry clusters from recordings of cells in which both fluorescent proteins were linked to the C-terminus of $K_{ir}2.1$ potassium channels (35 cells from three independent experiments). These include the distances from mVenus to (I) PAmCherry within the same protein complex, (ii) neighboring clusters in the absence of any active PAmCherry in the same complex and (iii) background signals (which are random and widely spread). We then calculated radial distributions for PAmCherry signals around each mVenus cluster and estimated a colocalization threshold of 100 nm (Fig. 3A). We also plotted a histogram of the distances from each mVenus cluster

to the nearest and second nearest PAmCherry cluster and fitted the shorter component with a Gaussian function (black line in Fig. 3B), obtaining values of 25.8 nm for μ and 19.0 nm for σ. As both approaches gave similar results, we chose 100 nm as the colocalization threshold in the following experiments.

Fixation Conditions. A critical prerequisite for DCC studies is to sufficiently immobilize the studied proteins while preserving the function of fluorescent proteins. A large proportion of membrane proteins remain mobile after fixation with paraformaldehyde (PFA) alone, requiring the addition of glutaraldehyde (GA) or the use of cold methanol (MeOH) for adequate immobilization (36, 37). To test whether these fixation methods affect mVenus and PAmCherry fluorescence in SMLM we stably expressed the chloride channel *β*-subunit barttin (38) C-terminally tagged with both PAmCherry and mVenus. Both fixation methods led to comparable correlation distances between the two colors (Fig. 4 A and B). However, upon MeOH fixation a significantly higher signal density was observed (Fig. 4 C and D) and a higher ratio of PAmCherry clusters to mVenus clusters was observed (Fig. 4E). We also observed a significantly lower colocalization ratio with PFA/GA than MeOH fixation (Fig. 4F). Similar results were obtained for many of the other proteins used in this study (Supplementary information, Table S1). Therefore, we concluded that fixation with cold methanol better preserved the function of the fluorescent proteins used in this study compared with PFA/GA fixation and, therefore, chose cold methanol as our routine fixation method.

Calibration of the Dual-Color Colocalization Model. Given that the detection efficiency *p* of fluorescent proteins may differ between experimental setups (15, 39), we first calibrated the DCC model (Eq. 1 and 2) using standards of known stoichiometry: the monomeric ClC-K chloride channel *β*-subunit barttin (38), the dimeric chloride channels ClC-2 and ClC-K (40–42), the trimeric glutamate transporter EAAT2 (43, 44) and the tetrameric potassium channel $K_{ir}2.1$ (45). All proteins were C-terminally tagged with PAmCherry and mVenus via two short peptide linkers (SI Appendix, Table S2). The reported stoichiometries were confirmed using high-resolution clear native electrophoresis (SI Appendix, Fig. S4).

After correction for drift and chromatic aberration, signals were assigned to clusters using the DBSCAN algorithm (46). Colocalization ratios (P) for mVenus and PAmCherry clusters were then calculated (Eq. 2, Fig. 5A) and used with the known oligomeric states (n) to obtain p from Eq. 1 of the DCC model. The fit suggested a low detection efficiency (p) of 0.12 for PAmCherry. However, the fit did not describe the data very well (R^2 = 0.68; red dots in Fig. 5A), especially for the monomer and the trimer.

We suspected that misfolding or truncation of both the marker and indicator in fusion proteins might have led to a loss of fluorescence for some protein subunits. This would especially affect proteins with smaller stoichiometries where the relative contribution of fluorescence is greater for each subunit. Therefore, we introduced an additional factor, m, to describe the probability of this phenomenon occuring. For a protein complex comprising n identical subunits, each C-terminally labeled with mVenus and PAmCherry with at least one functional mVenus, the probability of it also being detected via a PAmCherry signal is:

$$P(n)=\sum_{k=1}^{n}\left\{\frac{\binom{n}{k}(1-m)^k m^{n-k}}{1-m^n}\left[1-(1-p)^k\right]\right\} \quad [3]$$

Fitting Eq. 3 to the observed colocalization ratios of mVenus and PAmCherry clusters for the protein standards, gave values of 0.24 for m and 0.16 for p. This new model better aligns with the data (R^2 = 0.87) compared with the original DCC model, especially for the monomer (blue dots in Fig. 5A), and was therefore used to analyze proteins with contradictory or unknown stoichiometries.

Using DCC-SMLM to Determine the Subunit Stoichiometry of SLC26 and SLC17 Proteins. We next used DCC-SMLM to resolve the subunit stoichiometries of several SLC26 and SLC17 transporters. The colocalization ratios of mVenus and PAmCherry clusters for these proteins were determined (Fig. 5B) and compared with those expected for different oligomeric states using the DCC model (Eq. 3). To quantify the difference between expected and observed values, we introduced the coefficient of mismatch M^2 (SI Appendix). A M^2 value of 0 indicates a perfect match, with larger values indicating larger differences. As shown in Fig. 5 C-H, M^2 values are lowest for a dimeric

assembly of SLC26A3 and prestin and a monomeric state for all of the SLC17 transporters. These results are also consistent with the results of native gel electrophoresis (SI Appendix, Fig. S4). Our M^2 analysis was further validated by a two-sample Kolmogorov-Smirnov (KS) test to compare the distribution of values from each protein of interest to those of the protein standards (Fig. 5 I–N).

Discussion

We have developed a novel DCC-SMLM technique to determine the stoichiometries of plasma membrane protein complexes in situ in mammalian cells. Using mVenus as the marker fluorescent protein and PAmCherry as the indicator fluorescent protein for counting, we established a dimeric state for two representatives of the SLC26 family, the epithelial anion exchanger SLC26A3 and the motor protein prestin (Slc26a5); and a monomeric state for four representatives of the SLC17 family, the vesicular glutamate transporters vGlut1, 2, and 3 and the lysosomal sialic acid transporter sialin.

Our DCC-SMLM approach overcomes several limitations of methods commonly used to determine protein stoichiometries: no solubilization or purification of proteins is necessary, and the method works even in small mammalian cells. In contrast to previously reported quantitative single-color SMLM methods, DCC-SMLM has a simpler procedure and greater reliability of counting. For the binomial counting strategy, spatially clustered signals from protein complexes need to be separated into individual temporal clusters representing individual protein subunits. This process demands stringent control of the activation laser and intensive post-imaging data processing (13). In contrast, DCC-SMLM relies on the relationship between the total detection probability for a protein complex and the number of fluorophores within the complex. This relationship is mainly determined by the detection efficiency of individual fluorophores (Eq. 1) and is independent of temporal separation.

To achieve the necessary spatial resolution in two colors, we implemented and compared simple methods of cell fixation of cells and post-imaging correction of common artifacts such as drift and chromatic aberrations. Our implementation is based on SMLM but could also be adapted for other types of microscopy. In our experiments, there was no difference between a linear fit of chromatic aberrations and local average correction (Fig. 2 and SI Appendix, Fig. S3); however, this may depend on the quality of the optical setup and should be carefully evaluated for every instrument.

It should be emphasized that DCC-SMLM does not consider the behavior of individual protein complexes; instead, it is based on statistical information from a larger number of protein complexes. Therefore, if the protein complex of interest does not exist in a single oligomeric state, then DCC-

SMLM counting can only indicate the average stoichiometry. Although this is a limitation, similar caveats apply to most other methods based on fluorescent proteins. Future efforts should aim to explore whether additional information obtained from DCC-SMLM experiments could be used to increase its usefulness in analyzing proteins that adopt a mixture of oligomeric states.

At the single molecule level, noise and genuine signals become hard to distinguish, especially for photoactivatable fluorescent proteins with low brightness such as PAmCherry (30). Since DCC considers the overlap of two colors, the effect of background signals on the indicator is greatly reduced. The probability of noise signals falling within the colocalization threshold radius of a given mVenus cluster follows a poisson distribution. Our measured background signal density in the PAmCherry channel of ~0.6 clusters/µm² inside cells (Fig. 1D) suggests that only 2% of all mVenus molecules erroneously colocalize with noise (SI Appendix, Eq. S13). Compared with the colocalization ratios of >0.2 for dimers and higher oligomeric forms, this results in a signal-to-noise ratio of more than 10. This estimate compares favorably with single-color experiments: transfected cells showed a density of ~2.2 PAmCherry clusters/µm² consistent with a signal-to-noise ratio of only <4. Therefore, background noise in the indicator color channel impacts DCC-SMLM much less compared with SMLM. However, higher background noise levels in the marker detection channel would lead to the over-counting of marker clusters, resulting in a decreased colocalization ratio (Eq. 2), and should be avoided. Our use of the much brighter mVenus tag as the marker allowed us to use a narrow detection window of 526 to 546 nm, resulting in very low levels of background noise (Fig. 1D and SI Appendix, Table S3).

Previous methods of quantifying SMLM rely on the fluorophores having predictable behavior: Fluorescent proteins should either blink only once before being irreversibly bleached (PAmCherry: (15, 47)) or blink multiple times in a predictable pattern (e.g. Eos: (16)). The DCC strategy does not require such behavior, enabling the use of other fluorescent proteins or dyes with more complex photophysical behavior.

In conventional binomial analyses of SMLM data, low detection efficiency ($p < 0.5$) is a significant limitation when determining stoichiometries (14). Theoretically, DCC-SMLM can optimally

distinguish oligomeric states from monomer to pentamer if *p* is between 0.2 and 0.4 (Fig. 1C). However, even lower values of p may be tolerated, as seen in our study. This makes PAmCherry and other fluorescent proteins with low detection efficiencies viable for determining stoichiometries with DCC-SMLM (Fig. 5). The original DCC model (Eq. 1 and 2) suggested a *p* value of 0.12 for PAmCherry, but did not result in an acceptable fit for our experimental data from reference proteins (Fig. 5A). We reasoned that the function of the two fluorescent proteins may somehow become impaired in the context of fusion proteins. Therefore, we introduced a modification factor, *m*, to describe the probability of both fluorescent proteins becoming dysfunctional. Although protein misfolding or premature truncation is likely to be the underlying cause, we cannot exclude other explanations. Therefore, the need for a modification factor should be critically evaluated for every new pair of marker and indicator proteins used for DCC experiments. Fitting this modified DCC model to our reference data suggested a slightly higher *p* value of 0.16 for PAmCherry (Fig. 5A). Previously reported detection efficiencies for PAmCherry range from 0.04 when expressed in *E. coli (39)* to about 0.5 in *Xenopus* oocytes (14). This variability suggests that the expression system influences the functionality of fluorescent proteins, but may also reflect differences in image acquisition and data processing. In data processing, filtering noise signals is important for determining the detection efficiency. The higher tolerance to background noise enabled us to use only a simple filter based on an empirical threshold of the minimal number of signals for a single PAmCherry protein (Supplementary information, Fig. S1A). Consistent with this, signals recorded from background regions showed mostly small clusters (Supplementary information, Fig. S1A). We are, therefore, convinced that our estimates accurately reflect the detection efficiency of PAmCherry under our experimental conditions.

Using DCC-SMLM, we determined the stoichiometries of proteins from the SLC26 and SLC17 families. Studies using Blue native electrophoresis (4) and X-ray and cryo-EM (23, 24) previously reported a dimeric assembly for SLC26, whereas microscopy approaches supported a tetrameric assembly (26, 27). This discrepancy might indicate that SLC26 forms stable dimers that survive protein purification and solubilization as well as weakly associated “dimer-of-dimer” tetramers in native membranes that may only be visible in intact cells. Although we cannot exclude the possibility that a

minor fraction of prestin forms tetrameric structures, our results show that both prestin and SLC26A3 predominantly form dimers in intact HEK293T cells (Fig. 5 B-D).

We also studied the oligomeric state of vGlut1, 2, and 3 (Slc17a7, Slc17a6 and Slc17a8, respectively), and sialin (SLC17A5) and found that they are likely to be monomeric. Again, these results are consistent with native gel electrophoresis (Supplementary information, Fig. S4). An earlier report suggested dimeric organization for SLC17 transporters based on SDS-PAGE results (48), but a recent cryo-EM study reported a monomeric structure (49). Functional expression studies have also demonstrated that vesicular glutamate transporters and sialin can function without accessory subunits (50, 51), which, when considered alongside our results, confirms that SLC17 proteins are functional monomers.

Our study provides a new approach to determine the stoichiometries of membrane proteins *in situ*. DCC-SMLM overcomes many of the inherent drawbacks and requirements of previous methods, making stoichiometry determination more reliable and much easier to perform. Although we established the DCC strategy based on SMLM, the principle of our approach could also be applied to other types of microscopy techniques, provided that the spatial resolution is sufficient to separate the proteins of interest. Examples may include structured illumination (52) and super-resolution radial fluctuations (53) which are otherwise difficult to quantify at the single molecule level because they rely on the mathematical processing of multiple raw images to yield super-resolution images.

Materials and Methods

Cell Culture and DNA Constructs. HEK293T cells and Flp-In T-REx 293 cells were grown in DMEM supplemented with 50 U/mL penicillin-streptomycin and 10% FBS at 37 °C and 5% CO_2 (54). For Flp-In T-REx 293 cells, the medium was additionally supplemented with 10 µg/mL blasticidin S HCl and 100 µg/mL hygromycin B. cDNAs encoding the fusion proteins were inserted into the cloning site of the pcDNA5/FRT/TO or pcDNA3 vector plasmid via overlap extension PCR, restriction digestion and DNA ligation.

Sample Preparation for SMLM Imaging. For SMLM imaging, HEK293T cells were sparsely seeded onto clean 25-mm coverslips in 1.5 mL DH10 in 3.5 cm petri dishes 3 h before transfection. Cells were transfected with 0.5-1.0 µg plasmid DNA plus 1.0 µL Lipofectamine 2000 (Life Technologies) and cultured overnight (12-16 h) before fixation. Flp-In T-REx 293 cells were seeded on coverslips in the same way and 0.01 µg/mL tetracycline was added to the culture medium 14.5 h before fixation. For details, see *Supplementary information*.

SMLM Imaging and Data Processing. SMLM imaging started 16–24 h after fixation. All SMLM imaging was performed on a custom-built wide field and TIRF microscope (55). The fixed cells were imaged in PBS and diluted TetraSpeck beads were added to track sample drift. Raw images were first loaded into SNSMIL (56) to extract the precise location of fluorescent signals. The resulting datasets were processed with custom python scripts to correct of lateral chromatic aberrations and sample drift and for DCC analysis based on python v3.6.8. For details, see *Supplementary information*.

Acknowledgements This research was supported by the Deutsche Forschungsgemeinschaft (FA 301/15-1 to Ch.F.) as part of Research Unit FOR 5046, project P4. We are grateful to Iman Abdollahzadeh, Dr. Johnny Hendriks and Dr. Thomas Gensch for their generous support with the imaging setup. Many thanks also go to Cora Hannack, Dr. Raul E. Guzman and Arne Franzen for providing plasmids encoding vGlut and sialin. We also thank Ashley Craig of Alchemy Editorial Services Ltd for language editing.

Significance Statement

Determination of the stoichiometry of protein complexes in situ in the plasma membrane of within mammalian cells is nontrivial. Conventional strategies for quantifying single molecule localization microscopy (SMLM) data are highly sensitive to the activation laser intensity and background signals, and require high detection efficiency for the fluorophore. The dual-color colocalization (DCC) approach overcame these limitations and significantly improved the efficiency and fidelity of protein stoichiometry determination with SMLM. Using DCC-SMLM, we efficiently determined the stoichiometry of vGlut and SLC26A proteins.

Author contributions: H.L.T., G.S. and C.F. designed the project. H.L.T. was responsible for molecular cloning, the mathematical model and SMLM experiments. S.B.-P. performed the clear native PAGE and the relevant data processing. H.L.T., G.S. and D.K. were responsible for processing and analyzing SMLM data. H.L.T. wrote the first draft and all authors contributed to revising this manuscript.

The authors declare no conflict of interest.

1. Y. Lussier, *et al.*, Disease-linked mutations alter the stoichiometries of HCN-KCNE2 complexes. *Sci. Rep.* **9**, 9113 (2019).

2. H. Schägger, G. von Jagow, Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal. Biochem.* **199**, 223–231 (1991).

3. S. Gendreau, *et al.*, A Trimeric Quaternary Structure Is Conserved in Bacterial and Human Glutamate Transporters. *J. Biol. Chem.* **279**, 39505–39512 (2004).

4. S. Detro-Dassen, *et al.*, Conserved dimeric subunit stoichiometry of SLC26 multifunctional anion exchangers. *J. Biol. Chem.* **283**, 4177–4188 (2008).

5. Y. Chen, L.-N. Wei, J. D. Müller, Probing protein oligomerization in living cells with fluorescence fluctuation spectroscopy. *Proc. Natl. Acad. Sci.* **100**, 15492–15497 (2003).

6. M. H. Ulbrich, E. Y. Isacoff, Subunit counting in membrane-bound proteins. *Nat Meth* **4**, 319–321 (2007).

7. P. Annibale, S. Vanni, M. Scarselli, U. Rothlisberger, A. Radenovic, Quantitative Photo Activated Localization Microscopy: Unraveling the Effects of Photoblinking. *PLoS ONE* **6**, e22678 (2011).

8. P. Sengupta, *et al.*, Probing protein heterogeneity in the plasma membrane using PALM and pair correlation analysis. *Nat. Methods* **8**, 969–975 (2011).

9. B. Coste, *et al.*, Piezo proteins are pore-forming subunits of mechanically activated channels. *Nature* **483**, 176–181 (2012).

10. T. Bartoi, K. Augustinowski, G. Polleichtner, S. Gründer, M. H. Ulbrich, Acid-sensing ion channel (ASIC) 1a/2a heteromers have a flexible 2:1/1:2 stoichiometry. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 8281–8286 (2014).

11. P. Krashia, *et al.*, Human α3β4 Neuronal Nicotinic Receptors Show Different Stoichiometry if They Are Expressed in Xenopus Oocytes or Mammalian HEK293 Cells. *PLOS ONE* **5**, e13611 (2010).

12. X. Nan, *et al.*, Single-molecule super-resolution imaging allows quantitative analysis of RAF multimer formation and signaling. *Proc. Natl. Acad. Sci.* (2013) https:/doi.org/10/f5gxxb (February 4, 2021).

13. S.-H. Lee, J. Y. Shin, A. Lee, C. Bustamante, Counting single photoactivatable fluorescent molecules by photoactivated localization microscopy (PALM). *Proc. Natl. Acad. Sci.* **109**, 17436–17441 (2012).

14. E. M. Puchner, J. M. Walter, R. Kasper, B. Huang, W. A. Lim, Counting molecules in single organelles with super-resolution microscopy allows tracking of the endosome maturation trajectory. *Proc. Natl. Acad. Sci.* **110**, 16015–16020 (2013).

15. N. Durisic, L. Laparra-Cuervo, Á. Sandoval-Álvarez, J. S. Borbely, M. Lakadamyali, Single-molecule evaluation of fluorescent protein photoactivation efficiency using an *in vivo* nanotemplate. *Nat. Methods* **11**, 156–162 (2014).

16. G. Hummer, F. Fricke, M. Heilemann, Model-independent counting of molecules in single-molecule localization microscopy. *Mol. Biol. Cell* **27**, 3637–3644 (2016).

17. P. R. Nicovich, D. M. Owen, K. Gaus, Turning single-molecule localization microscopy into a quantitative bioanalytical tool. *Nat. Protoc.* **12**, 453–460 (2017).

18. E. Betzig, *et al.*, Imaging Intracellular Fluorescent Proteins at Nanometer Resolution. *Science* **313**, 1642–1645 (2006).

19. M. J. Rust, M. Bates, X. Zhuang, Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat. Methods* **3**, 793–796 (2006).

20. H. Li, J. C. Vaughan, Switchable Fluorophores for Single-Molecule Localization Microscopy. *Chem. Rev.* **118**, 9412–9454 (2018).

21. , Ultralow- and Low-Background Surfaces for Single-Molecule Localization Microscopy of Multistep Biointerfaces for Single-Molecule Sensing | Langmuir (February 3, 2021).

22. A. Shivanandan, H. Deschout, M. Scarselli, A. Radenovic, Challenges in quantitative single molecule localization microscopy. *FEBS Lett.* **588**, 3595–3602 (2014).

23. E. R. Geertsma, *et al.*, Structure of a prokaryotic fumarate transporter reveals the architecture of the SLC26 family. *Nat. Struct. Mol. Biol.* **22**, 803–808 (2015).

24. J. D. Walter, M. Sawicka, R. Dutzler, Cryo-EM structures and functional characterization of murine Slc26a9 reveal mechanism of uncoupled chloride transport. *eLife* **8**, e46986 (2019).

25. J. Zheng, *et al.*, Analysis of the Oligomeric Structure of the Motor Protein Prestin *. *J. Biol. Chem.* **281**, 19916–19924 (2006).

26. G. P. Sinha, F. Sabri, E. K. Dimitriadis, K. H. Iwasa, Organization of membrane motor in outer hair cells: an atomic force microscopic study. *Pflüg. Arch. - Eur. J. Physiol.* **459**, 427–439 (2010).

27. R. Hallworth, M. G. Nichols, Prestin in HEK cells is an obligate tetramer. *J. Neurophysiol.* **107**, 5–11 (2011).

28. T. Nagai, *et al.*, A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat. Biotechnol.* **20**, 87–90 (2002).

29. R. M. Dickson, A. B. Cubitt, R. Y. Tsien, W. E. Moerner, On/off blinking and switching behaviour of single molecules of green fluorescent protein. *Nature* **388**, 355–358 (1997).

30. F. V. Subach, *et al.*, Photoactivatable mCherry for high-resolution two-color fluorescence microscopy. *Nat. Methods* **6**, 153–159 (2009).

31. J. Yi, *et al.*, madSTORM: a super-resolution technique for large-scale multiplexing at single-molecule accuracy. *Mol. Biol. Cell* **27**, 3591–3600 (2016).

32. M. Kozubek, P. Matula, An efficient algorithm for measurement and correction of chromatic aberrations in fluorescence microscopy. *J. Microsc.* **200**, 206–217 (2000).

33. M. Erdelyi, *et al.*, Correcting chromatic offset in multicolor super-resolution localization microscopy. *Opt. Express* **21**, 10978–10988 (2013).

34. L. S. Churchman, J. A. Spudich, Single-Molecule High-Resolution Colocalization of Single Probes. *Cold Spring Harb. Protoc.* **2012**, 242–245 (2012).

35. M. Georgieva, *et al.*, Nanometer resolved single-molecule colocalization of nuclear factors by two-color super-resolution microscopy imaging. *Methods* **105**, 44–55 (2016).

36. K. A. K. Tanaka, *et al.*, Membrane molecules mobile even after chemical fixation. *Nat. Methods* **7**, 865–866 (2010).

37. T. A. Stanly, *et al.*, Critical importance of appropriate fixation conditions for faithful imaging of receptor microclusters. *Biol. Open* **5**, 1343–1350 (2016).

38. U. Scholl, *et al.*, Barttin modulates trafficking and function of ClC-K channels. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 11411–11416 (2006).

39. S. Wang, J. R. Moffitt, G. T. Dempsey, X. S. Xie, X. Zhuang, Characterization and development of photoactivatable fluorescent proteins for single-molecule–based super-resolution imaging. *Proc. Natl. Acad. Sci.* **111**, 8452–8457 (2014).

40. A. Thiemann, S. Gründer, M. Pusch, T. J. Jentsch, A chloride channel widely expressed in epithelial and non-epithelial cells. *Nature* **356**, 57–60 (1992).

41. G. Stölting, M. Fischer, C. Fahlke, ClC-1 and ClC-2 form hetero-dimeric channels with novel protopore functions. *Pflüg. Arch. - Eur. J. Physiol.* **466**, 2191–2204 (2014).

42. E. Park, E. B. Campbell, R. MacKinnon, Structure of a CLC chloride ion channel by cryo-electron microscopy. *Nature* **541**, 500–505 (2017).

43. D. Nothmann, *et al.*, Hetero-oligomerization of Neuronal Glutamate Transporters. *J. Biol. Chem.* **286**, 3935–3943 (2011).

44. J. C. Canul-Tec, *et al.*, Structure and allosteric inhibition of excitatory amino acid transporter 1. *Nature* **544**, 446–451 (2017).

45. D. A. Doyle, *et al.*, The Structure of the Potassium Channel: Molecular Basis of K+ Conduction and Selectivity. *Science* **280**, 69–77 (1998).

46. M. Ester, H.-P. Kriegel, J. Sander, X. Xu, A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *Proc. 2nd Int. Conf. Knowl. Discov. Data Min. KDD-96*, 226–231 (1996).

47. F. V. Subach, *et al.*, Photoactivation mechanism of PAmCherry based on crystal structures of the protein in the dark and fluorescent states. *Proc. Natl. Acad. Sci.* **106**, 21097–21102 (2009).

48. S. S. L. Andersen, Expression and purification of recombinant vesicular glutamate transporter VGLUT1 using PC12 cells and High Five insect cells. *Biol. Proced. Online* **6**, 105–112 (2004).

49. F. Li, *et al.*, Ion transport and regulation in a synaptic vesicle glutamate transporter. *Science* **368**, 893–897 (2020).

50. P. Morin, C. Sagné, B. Gasnier, Functional characterization of wild-type and mutant human sialin. *EMBO J.* **23**, 4560–4570 (2004).

51. J. Eriksen, *et al.*, Protons Regulate Vesicular Glutamate Transporters through an Allosteric Mechanism. *Neuron* **90**, 768–780 (2016).

52. P. Kner, B. B. Chhun, E. R. Griffis, L. Winoto, M. G. L. Gustafsson, Super-resolution video microscopy of live cells by structured illumination. *Nat. Methods* **6**, 339–342 (2009).

53. N. Gustafsson, *et al.*, Fast live-cell conventional fluorophore nanoscopy with ImageJ through super-resolution radial fluctuations. *Nat. Commun.* **7**, 12471 (2016).

54. H. Tan, S. Bungert-Plümke, C. Fahlke, G. Stölting, Reduced Membrane Insertion of CLC-K by V33L Barttin Results in Loss of Hearing, but Leaves Kidney Function Intact. *Front. Physiol.* **8** (2017).

55. G. Stölting, *et al.*, Direct Interaction of CaVβ with Actin Up-regulates L-type Calcium Currents in HL-1 Cardiomyocytes. *J. Biol. Chem.* **290**, 4561–4572 (2015).

56. Y. Tang, *et al.*, SNSMIL, a real-time single molecule identification and localization algorithm for super-resolution fluorescence microscopy. *Sci. Rep.* **5**, 11073 (2015).

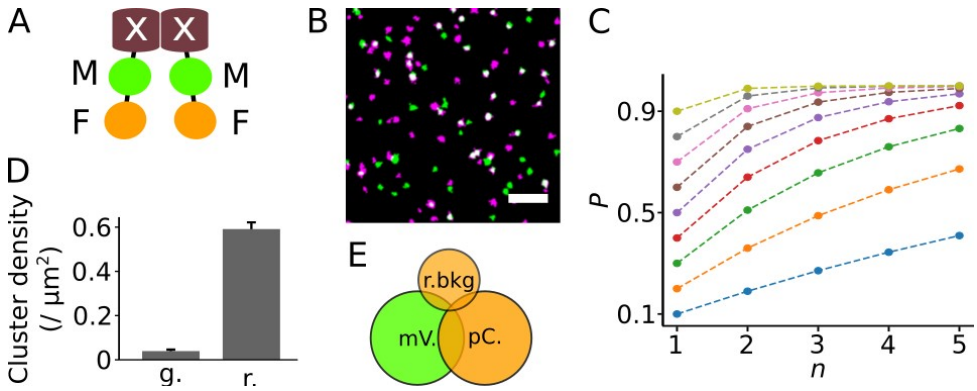

Fig. 1. The DCC strategy to determine protein stoichiometry. (*A*) Model of a dimeric fusion protein used for DCC-SMLM. The protein of interest is marked with a white "X", whereas the marker and indicator fluorescent proteins are labeled "M" and "F", respectively. (*B*) A reconstructed SMLM image from an HEK293T cell expressing K_{ir} 2.1 labeled with both mVenus (green) and PAmCherry (magenta). White indicates colocalization of the two color signals. Scale bar: 1 µm. (*C*) The probability *P* of detecting a protein complex changes with the number (*n*) and detection efficiency (*p*) of the fluorescent proteins it contains. Each color indicates a different *p* value from 0.1 to 0.9 in steps of 0.1. (*D*) The mean background cluster densities in the mVenus (green, g.) and the PAmCherry (red, r.) recording channels. A cluster comprises a series of fluorescent signals that are spatially close and clustered with the DBSCAN algorithm. (*E*) As the background signals in the green channel are negligible, only three types of signals are considered: signals from mVenus (mV.), signals from PAmCherry (pC.) and background signals in the red channel (r.bkg).

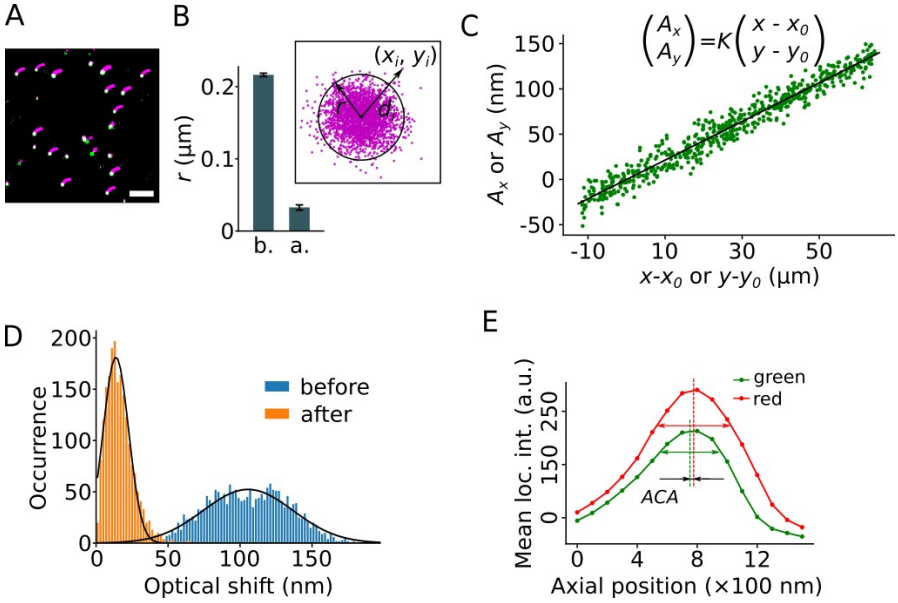

Fig. 2. Correction of sample drift and chromatic aberration. (*A*) An SMLM image of a bead sample reconstructed from 13,000 frames recorded over ~18.5 minutes. The small magenta smears indicate sample drift. The corrected positions are displayed in green or, when overlapping, in white. Scale bar: 1 µm. (*B*) The mean cluster radius of positions extracted from 41 beads before (b.) and after (a.) drift correction, with error bars indicating the SEM. The cluster radius is defined as $\mu + 2\sigma$ of the Euclidean distances (d_i) of all signals (with coordinates (x_i, y_i)) to the center of the cluster, as shown in the inset. (*C*) Plot of the lateral chromatic aberration between the red and green channels, (A_x, A_y), versus the distance of the beads from a center point (x_0, y_0). Data from both dimensions are included in the graph. Linear regression (black line) yielded the values of the slope K, x_0 and y_0. (*D*) Data from nine recordings of a bead sample (1980 beads) were corrected with the values obtained in (*C*). Black lines show Gaussian fitting. (*E*) The mean intensities of the beads recorded at the green and the red channels change along the axial position of the sample. The axial distance between the two peaks indicates the axial chromatic aberration (ACA).

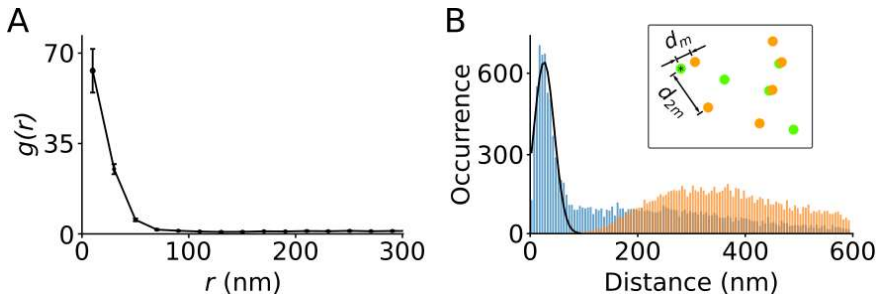

Fig. 3. Definition of colocalization in SMLM. (*A*) The radial distribution *g(r)* of PAmCherry clusters around mVenus clusters (total of 12,132 clusters). mVenus and PAmCherry were both fused to the C-terminus of $K_{ir}2.1$, expressed in HEK293T cells and recorded in SMLM. (*B*) Distribution of the distance from mVenus clusters to the nearest (blue, d_m in inset) and second nearest (orange, d_{2m} in inset) neighboring PAmCherry clusters. The black curve shows a Gaussian fit of the nearest distances. Data for both panels are pooled from 35 cells from three independent experiments.

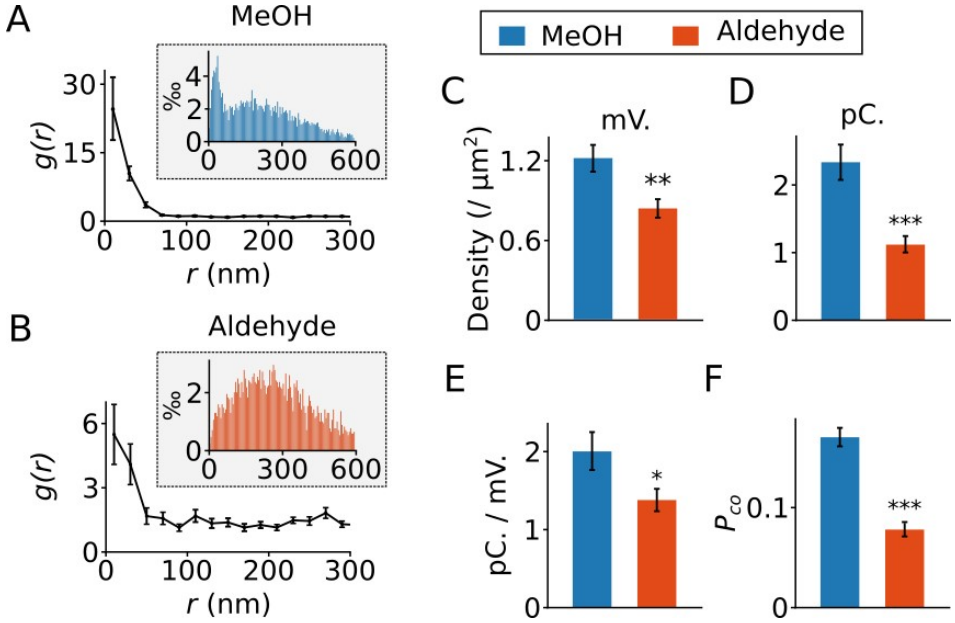

Fig. 4. Effect of different fixation methods on the number of clusters. (*A*, *B*) Results from the radial distribution function *g(r)* for PAmCherry clusters around mVenus clusters recorded in Flp-In T-REx 293 cells expressing barttin-mVenus-PAmCherry and fixed with cold methanol (*A*, MeOH, 24 cells) or para-formaldehyde plus glutaraldehyde (*B*, Aldehyde, 29 cells). The inset shows the histograms of the nearest neighbor distances. (*C*, *D*) The densities of mVenus (*C*, $p = 0.0028$) and PAmCherry (*D*, $p = 4.9\times10^{-5}$) clusters in samples prepared using the two fixation methods. (*E*) The relative density of PAmCherry (pC.) clusters to mVenus (mV.) clusters, $p = 0.02$. (*F*) The measured colocalization ratio of mVenus with PAmCherry, $p = 1.6\times10^{-10}$. *: $0.05 > p > 0.01$, **: $0.01 > p > 0.001$, ***: $p < 0.001$, Student t-test.

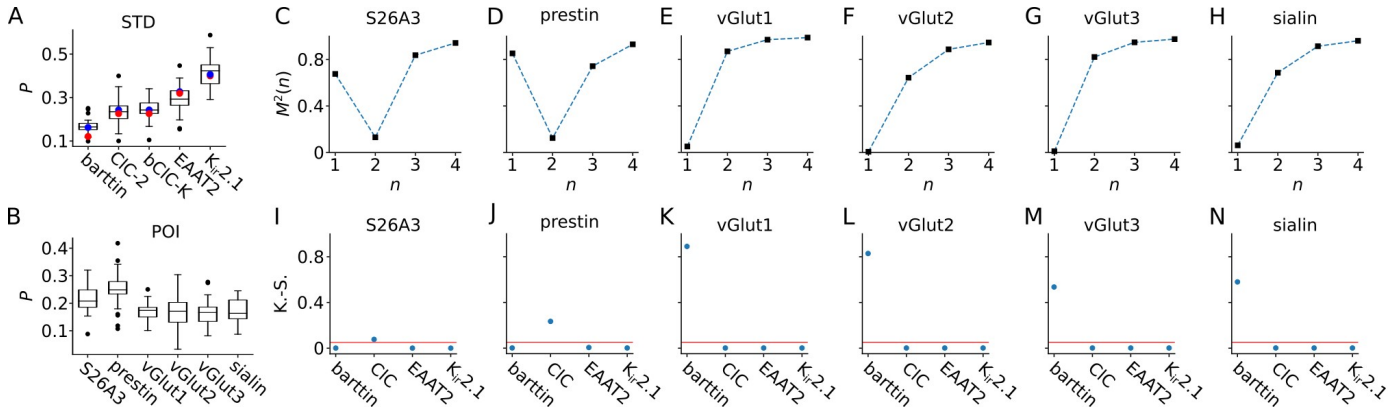

Fig. 5. Calibration and determination of stoichiometry through analyzing the colocalization ratios with DCC-SMLM. (*A*) The colocalization ratios (*P*) for mVenus and PAmCherry linked to proteins with known stoichiometries (STD). Red dots show the fit with Eq. 1 and blue dots show the fit with Eq. 3 including the modification factor *m*. The number of cells included are 13, 30, 14, 19 and 34, respectively, for barttin, ClC-2, bClC-K, EAAT2 and $K_{ir}2.1$. (*B*) Colocalization ratios for mVenus and PAmCherry linked to the indicated proteins of interest (POI). The number of cells included are 24, 27, 22, 25, 30 and 22, respectively. (*C-H*) The coefficient of mismatch (M^2) calculated for all the proteins of interest at assumed stoichiometries (*n)* of 1, 2, 3 and 4. (*I–N*): Two-sample Kolmogorov-Smirnov tests of the observed colocalization ratios from each protein of interest compared with the protein standards in (*A*). The orange horizontal line indicates the Bonferroni-corrected threshold for the α-level (0.0125).