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]

s

[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, Center of Functional Genomics, 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 a result of the interactions between individual proteins and an important determinant of their function. 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 membraneproteins. 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 temporal clustering of signals from individual proteins 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 (Lussier et al., 2019). Native gel electrophoresis has been successfully used to determine subunit stoichiometries and the molecular determinants of oligomeric assemblies (Detro-Dassen et al., 2008; Gendreau et al., 2004; Schägger & von Jagow, 1991). 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 (Chen et al., 2003), stepwise photobleaching (Ulbrich & Isacoff, 2007) and single molecule localization microscopy (SMLM) (Annibale et al., 2011; Sengupta et al., 2011) 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 (Bartoi et al., 2014; Coste et al., 2012; Lussier et al., 2019; Ulbrich & Isacoff, 2007). 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 (Ulbrich & Isacoff, 2007). However, some protein complexes exhibit a different stoichiometry when expressed in the *Xenopus* system compared with mammalian expression systems (Krashia et al., 2010). For mammalian cells, the resolution limit of light microscopy can be circumvented using super-resolution microscopy methods (Durisic et al., 2014; Hummer et al., 2016; Lee et al., 2012; Nan et al., 2013; Nicovich et al., 2017; Puchner et al., 2013; Sengupta et al., 2011). SMLM achieves a higher resolution by reconstructing the precise positions of isolated fluorophore emissions detected over thousands of individual frames (Betzig et al., 2006; Rust et al., 2006). The fluorophores used for quantitative SMLM

are usually photoactivatable or photoconvertible fluorescent proteins that are activated at wavelengths different from that of the excitation laser (H. Li & Vaughan, 2018). 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 (Sengupta et al., 2011).

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 (Lee et al., 2012). 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 (Annibale et al., 2011; Lee et al., 2012). Another source of miscounting is the background noise, which cannot be easily distinguished from genuine signals at the single molecule level (Shivanandan et al., 2014; Zhao et al., 2018). 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 (Durisic et al., 2014; Nicovich et al., 2017). 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 (Durisic et al., 2014; Ulbrich & Isacoff, 2007).

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 (Detro-Dassen et al., 2008; Geertsma et al., 2015; Walter et al., 2019), whereas other studies have proposed a tetrameric stoichiometry (Hallworth & Nichols, 2011; Sinha et al., 2010; Zheng et al., 2006). 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 flurorescent protein (F) with a detection efficiency of *p* (the probability of being fluorescent and being detected), will be detected with a probability *P*:

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

In theory, *P* is equal to the ratio of all proteins with visible fluorescence (N_F) by the number of labeled proteins in total (*N*):

$$P = \frac{N_F}{N} \quad [2]$$

However, direct determination of *P* from single-color microscopy images is impossible as the total number of proteins, including those not exhibiting fluorescence, is unknown. Therefore, in order to experimentally determine *P*, we additionally labeled each protein subunit with another fluorescent protein (M) as a baseline marker (Figure 1A). As for F, the number of proteins observable by the fluorescence of M (N_M) will be smaller than *N*. Assuming that the fluorescence of F and M is independent, *P* can be determined from the number of proteins showing fluorescence of both proteins, M and F (N_{MF}), and N_M which can be counted from dual-color microscopy images (Figure 1B):

$$P = \frac{N_F}{N} = \frac{N_{MF}}{N_M} \quad [3]$$

The values of *P* obtained for proteins with known stoichiometries were used to solve Equation 1 for *p*, which could then be used to determine unknown stoichiometries (Figure 1C). Independence of both fluorescence signals was guaranteed by using proteins with separate emission and excitation spectra as well as performing sequential rather than simultaneous imaging of color channels.

We chose mVenus (Nagai et al., 2002) as the fluorescent marker protein M because of the very low background signal in its detection channel (Figure 1D-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 (Nagai et al., 2002). Like other GFP derivatives, it blinks by cycling between the fluorescent and dark states under continuous

illumination with the excitation laser (Dickson et al., 1997). PAmCherry was chosen as the indicator because it is quickly bleached by continuous excitation after activation (Durisic et al., 2014). However, this channel had considerably more background signals. PAmCherry fluorescence is stable for several frames after its activation (Durisic et al., 2014; Subach, Patterson, et al., 2009), resulting in small clusters of localizations per fluorescent protein. We partially corrected for the background by requiring a certain number of signals per cluster (see methods) as well as removing clusters with very large numbers of localizations (Figure 1 – figure supplement 1). However, some background signals were indistinguishable from genuine PAmCherry signals, which explains the higher background in that particular color channel (Figure 1D-F). The order of mVenus and PAmCherry tags fused to the protein of interest did not significantly impact the probability of colocalization (Figure 1 – figure supplement 2). To reduce the contribution of the background noise to the final result even more, we corrected P by subtraction of the predicted contribution of colocalization with random noise from the observed colocalization (N_{co}) and the observed mVenus clusters (N_M) (Figure 1F):

$$P=\frac{N_{co}-N_M P_{rd}}{N_M-N_M P_{rd}} \quad [4]$$

The probability for the colocalization of any mVenus cluster with red background noise clusters within a certain distance (d) follows a Poisson distribution. P_{rd} is the probability of finding an mVenus cluster randomly colocalized with red background noise clusters and was calculated as

$$P_{rd}=1-e^{-\pi d^2 D_{rb}} \quad [5]$$

where D_{rb} is the density of red clusters in regions outside transfected cells.

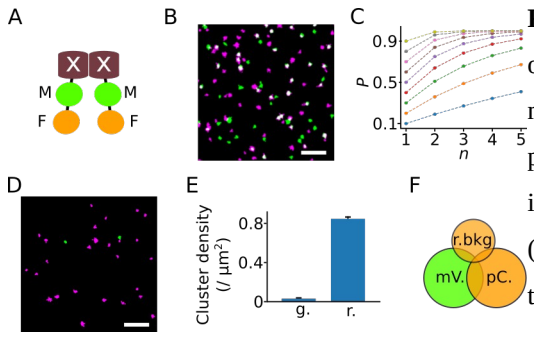

Figure 1. The DCC strategy to determine protein stoichiometry. (*A*) Model of a dimeric fusion protein used for DCC-SMLM. The protein of interest is marked with a white "X", whereas the marker and indicator fluorescent proteins are labeled "M" and "F", respectively. (*B*) A reconstructed SMLM image from 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*) A representative image reconstructed from a recording of a bare coverslip suface showing background signals in the PAmCherry (magenta) and mVenus channels (green). Scale bar: 1 µm. (*E*) 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. *N* = 283 recordings. (*F*) 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).

Correction of Sample Drift, Chromatic Aberrations and Definition of Colocalization. The resolution of all SMLM data is compromised by sample drift and optical aberrations which affect the determination of colocalization. We recorded sample drift using TetraSpeck beads as fiducial markers and correcting for this during post-imaging data processing (Yi et al., 2016). This correction reduced the radius of clustered signals to ~30 nm (Figure 2A 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 (Erdelyi et al., 2013; Kozubek & Matula, 2000). Values from a linear fit of the measured inter-color distances of individual beads to their coordinates were used for correction (Figure 2C). This way, lateral chromatic aberration could be reduced to 13.8 ± 8.9 nm for bead samples (Figure 2D and E). Chromatic aberrations may also be corrected by determining regional averages with similar results (Figure 2 – figure supplement 1) (Churchman & Spudich, 2012; Georgieva et al., 2016). 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 (Figure 2F).

Unlike FPs in cellular samples, beads show an intense and long-lasting fluorescent signal. To determine the improvement in fixed cell samples, and to determine the threshold for colocalization, we applied the correction to recordings of samples expressing monomeric barttin or tetrameric $K_{ir}2.1$, each subunit tagged with mVenus and PAmCherry (Figure 2G). We determined the distances between each individual mVenus cluster and the surrounding PAmCherry clusters.-These include the distances from mVenus to (I) PAmCherry signals 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 (RDF) for PAmCherry signals around each mVenus cluster for two different constructs, the monomeric barttin and the tetrameric $K_{ir}2.1$ (Figure 2H). Prior to the correction of chromatic abberations, the RDF shows high densities centered on a distance of about 100 nm but reaching up to 250 nm. Following correction, the majority of PAmCherry signals are distributed in a radius of below 100 nm around mVenus clusters so this value was chosen as threshold distance for colocalization.

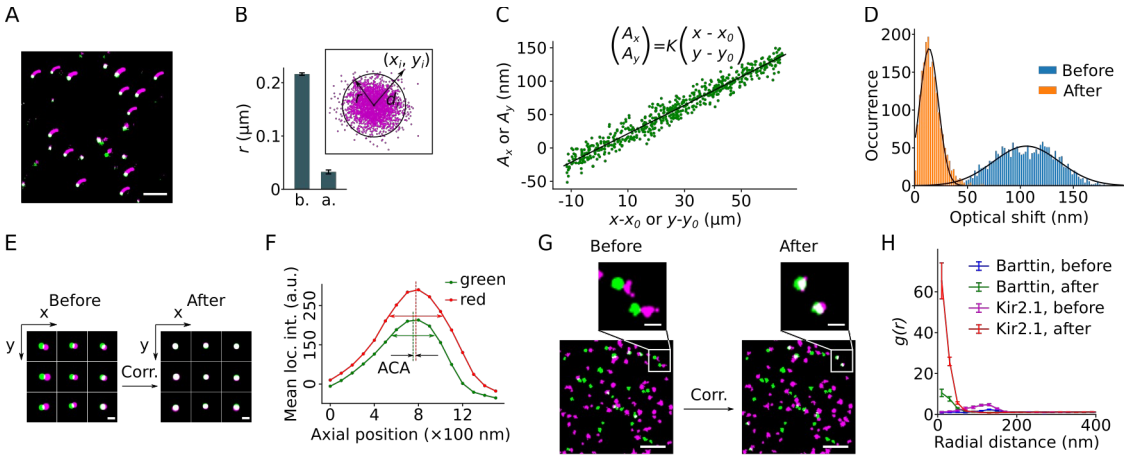

Figure 2. Correction of sample drift and chromatic aberration.

(*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. The fit is representative for a single recording of a bead sample.

(*D*) Lateral chromatic aberrations (blue) from nine recordings of a bead sample with a total of 1980 beads were corrected with the values obtained in *C*, as shown in orange. Black lines show Gaussian fitting. (*E*) Reconstructed images of nine representative beads from different areas before and after drift & chromatic aberration corrections. The green and magenta indicate signals from the mVenus and PAmCherry channels, respectively. Scale bar: 200 nm. (*F*) 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). (*G*) Representative reconstructed images showing the mVenus (green) and PAmCherry clusters (magenta) recorded from HEK293T cells expressing barttin-mVenus-PAmCherry before and after drift & chromatic aberration corrections. Non-clustered signals were omitted from the images. Scale bars indicate 1 μm and 200 nm respectively in the bottom and the zoomed-in images. (*H*) Radial distribution of barttin (17 cells, 1 preparation) and $K_{ir}2.1$ (35 cells, 3 preparations) PAmCherry clusters around mVenus clusters before and after drift & chromatic aberration corrections. Un-clustered signals were excluded from the analysis.

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 (Stanly et al., 2016; Tanaka et al., 2010). To test whether these fixation methods affect mVenus and PAmCherry fluorescence in SMLM we stably expressed the chloride channel *β*-subunit barttin(Estévez et al., 2001) C-terminally tagged with both PAmCherry and mVenus. Both fixation methods led to comparable correlation distances between the two colors (Figure 3A and B). However, upon MeOH fixation a significantly higher signal density, a higher ratio of PAmCherry clusters to mVenus clusters and a higher colocalization ratio were observed (Figure 3 A', B' and C-F). Similar results were obtained for many of the other proteins used in this study (Table 1). 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.

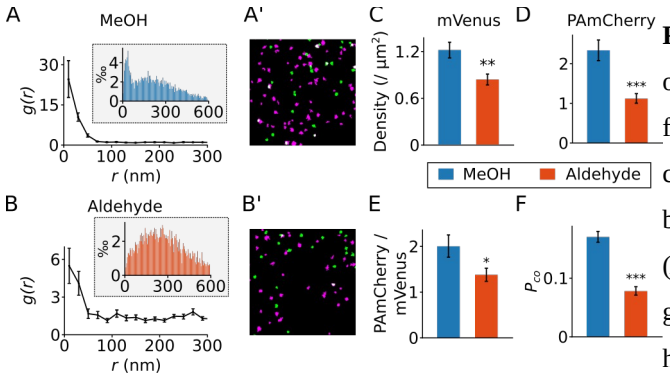

Figure 3. Effect of different fixation methods on the number of clusters. (*A* and *B*) Results from the radial distribution function *g(r)* for PAmCherry clusters around mVenus clusters recorded in Flp-In T-REx 293 cells expressing barttin-mVenus-PAmCherry and fixed with cold methanol (*A*, MeOH, 24 cells) or para-formaldehyde plus glutaraldehyde (*B*, Aldehyde, 29 cells). The inset shows the histograms of the nearest neighbor distances. (*A'* and *B'*) Reconstructed images from recordings of cells expressing barttin-mVenus-PAmCherry fixed with cold methanol (*A'*) or the aldehyde (*B'*), with green indicating mVenus clusters and magenta for PAmCherry clusters. (*C* and *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.

Calibration of the Dual-Color Colocalization Model. Given that the detection efficiency *p* of fluorescent proteins may differ between experimental setups (Durisic et al., 2014; Wang et al., 2014), we first calibrated the DCC model (Equations 1 and 3) using standards of known stoichiometry: the monomeric ClC-K chloride channel *β*-subunit barttin (Scholl et al., 2006), the dimeric chloride channels ClC-2 and ClC-K (Park et al., 2017; Stölting et al., 2014; Thiemann et al., 1992), the trimeric glutamate transporter EAAT2 (Canul-Tec et al., 2017; Nothmann et al., 2011) and the tetrameric potassium channel $K_{ir}2.1$ (Doyle et al., 1998). All proteins were C-terminally tagged with PAmCherry and mVenus via two short peptide linkers (Table 2).

As an independent confirmation of the reported stoichiometries from the literature as well as to confirm the function of our constructs, we used high-resolution clear native electrophoresis (hrCNE) (Figure 4 – figure supplement 1). Our results also demonstrate that the use of a two fluorescent proteins linked to the proteins in our analysis did not interfere with their reported quarternary structure.

After correction for drift and chromatic aberration, signals were assigned to clusters using the DBSCAN algorithm (Ester et al., 1996). Colocalization ratios (*P*) for mVenus and PAmCherry clusters were then calculated (Equation 3, Figure 4A) and used with the known oligomeric states (*n*) to obtain *p*

from Equation 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 Figure 4A), 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, having been detected via mVenus, the probability of it also being detected via a PAmCherry signal is:

$$P(n)=\sum_{k=1}^{n}\left\{\frac{\binom{n}{k}(1-m)^k m^{n-k}}{1-m^n}\left[1-(1-p)^k\right]\right\} \quad [6]$$

Fitting Equation 6 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 Figure 4A), 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 (Figure 4B) and compared with those expected for different oligomeric states using the DCC model (Equation 6). To quantify the difference between expected and observed values, we introduced the coefficient of mismatch M^2 (SI). A M^2 value of 0 indicates a perfect match, with larger values indicating larger differences. As shown in Figure 4 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 (Figure 4 – figure supplement 1). 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 (Figure 4I–N). Moreover, we independently confirmed our results for the SLC26 and SLC17 proteins with high-resolution clear native electrophoresis utilizing the same constructs and expression system as in our DCC-SMLM experiments (Figure 4 – figure supplement 1).

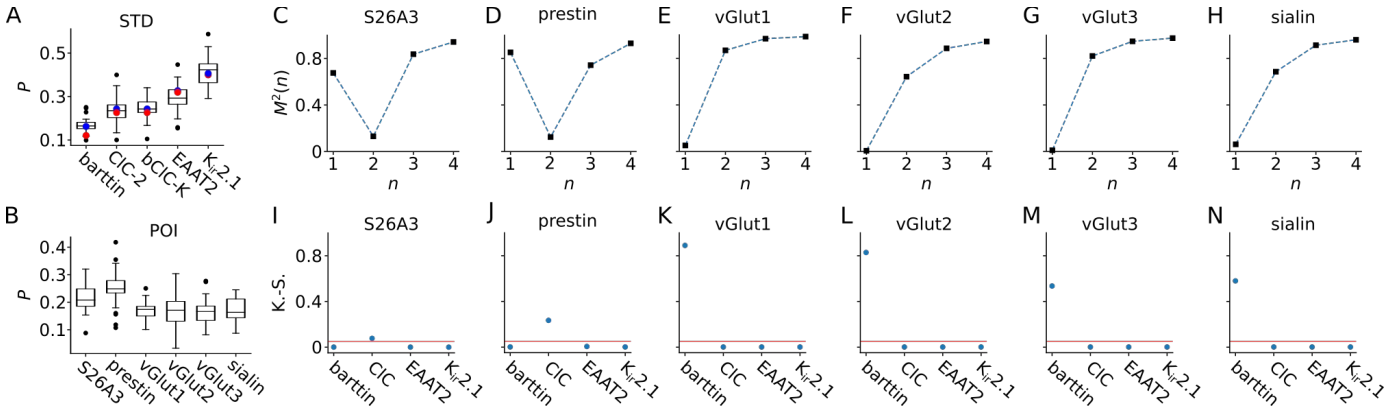

Figure 4. Calibration and determination of stoichiometry through analyzing the colocalization ratios with DCC-SMLM. (*A*) The colocalization ratios (*P*) for mVenus and PAmCherry linked to proteins with known stoichiometries (STD). Red dots show the fit with Equation 1 and blue dots show the fit with Equation 6 including the modification factor *m*. The number of cells included in the analysis are (in brackets): barttin (13), ClC-2 (30), bClC-K (14), EAAT2 (19) and $K_{ir}2.1$ (34). (*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).

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 (Figure 4 – figure supplement 2) overcomes several limitations of methods commonly used to determine protein stoichiometries: unlike for biochemical methods, no solubilization or purification of proteins is necessary as DCC-SMLM works with proteins *in situ*. In comparison to photo-bleaching experiments, DCC-SMLM can be applied reliably even in small mammalian cells where the high densities of protein expression are problematic. In contrast to previously reported quantitative SMLM methods, DCC-SMLM holds several advantages that result in a simplification of the procedure and offers greater reliability of counting. Binomial SMLM counting strategies require clustering of signals in at least three dimensions (x, y and time) to identify individual protein subunits. This demands stringent control of the activation laser and intensive post-imaging data processing (Lee et al., 2012). Another recently published SMLM-based quantification strategy also foregoes the need for temporal clustering as it utilizes the blinking behavior of some fluorescent proteins (Fricke et al., 2015; Hummer et al., 2016). This, however, requires fluorescent proteins that show blinking behavior (multiple on $\leftrightarrow$ off sequences) which is not guaranteed for all fluorescence proteins or dyes (Durisic et al., 2014; Subach, Malashkevich, et al., 2009). Additionally, all previously published SMLM quantification and photobleching step analysis methods require large photoefficiency values ($p \geq 0.5$) as they are otherwise unable do resolve the small oligomeric stoichiometries typically observed for most membrane proteins. In contrast, DCC-SMLM only relies on the relationship between the total detection probability for a protein complex and the number of fluorophores within the complex. This relationship is determined by the detection efficiency of the individual fluorophores (Equations 1 and 6) and is independent of temporal separation or specific fluorophore behavior. In fact, while we use the

x- and y-coordinates provided by SMLM, DCC itself is not even reliant on SMLM and the algorithm may also be used with other microscopy modalities given enough spatial resolution. Examples may include structured illumination (Kner et al., 2009) and super-resolution radial fluctuations (Gustafsson et al., 2016) 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 thereby removing much of the information required by other algorithms.

At the single molecule level, noise and genuine signals become hard to distinguish, especially for photoactivatable fluorescent proteins with low brightness such as PAmCherry (Subach, Patterson, et al., 2009). 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 ~0.8 clusters/μm^2 inside cells (Figure 1E) suggests that only 2.6% of all mVenus molecules erroneously colocalize with noise (Equation 5). Compared with the colocalization ratios of > 0.2 for dimers and higher oligomeric forms, this results in a signal-to-noise ratio of > 10. This estimate compares favorably with single-color experiments: transfected cells showed a mean density of ~2.2 PAmCherry clusters/μm^2 consistent with a signal-to-noise ratio of < 4 for this FP alone. 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 (Equation 3), and should be avoided. Our use of the much brighter mVenus tag as the marker allowed us to use a narrow detection window, resulting in very low levels of background noise (Figure 1D and Table 3).

For data processing, filtering of noise signals is important for determining the detection efficiency accurately. 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 (Figure 5). This is confirmed by our observation that signal clusters recorded from background regions showed mostly only small clusters (Figure 5).

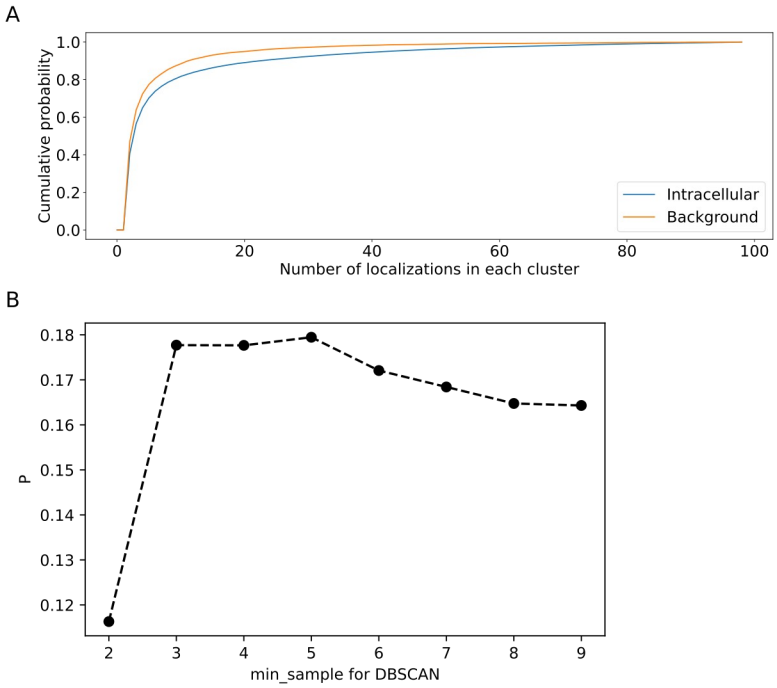

Figure 5. *(A)* The cumulative probability distribution of the number of localizations per cluster in the background and within Flp-In T-Rex 293 cells expressing Barttin-mVenus-PAmCherry. While most clusters in the background only exhibit a small number of localizations, clusters within cells exhibit a higher proportion of larger clusters. *(B)* The dependence of the colocalization ratio P versus the lower cutoff to be considered a true cluster. A cutoff of 2 results in low ratios due to the inclusion of a large number of background red clusters that are not colocalizing with green clusters.

While data on the stoichiometry in large protein clusters and macro-structures can be relevant, DCC-SMLM was developed to only resolve smaller oligomeric structures *in situ* as they are typically the smallest functional units of proteins. We therefore removed larger structures from further analysis. The number of clusters removed, however, was very low with typically more than 98% remaining (Figure 1 – figure supplement 1).

The use of two fluorescent proteins linked sequentially to our reference proteins and proteins of interest did not interfere with their structure as indicated by our hrCNE results (Figure 4 – figure supplement 1). This may not be surprising as similar results have been reported previously for similar constructs in other contexts (Leeman et al., 2018; Liang et al., 2018; Murata et al., 2005; Ranade et al., 2014).

In conventional binomial analyses of SMLM data, a low detection efficiency ($p < 0.5$) is a significant limitation when determining stoichiometries (Durisic et al., 2014). Theoretically, DCC-SMLM can optimally distinguish oligomeric states from monomer to pentamer if p is between 0.2 and 0.4 (Figure 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 (Figure 4). The original DCC model (Equations 1 and 3) suggested a *p* value of 0.12 for PAmCherry, but did not result in an acceptable fit for our experimental data from reference proteins (Figure 4A). 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. Fitting this modified DCC model to our reference data suggested a slightly higher *p* value of 0.16 for PAmCherry which resulted in a fit that aligned well with the experimental data (Figure 4A). Previously reported detection efficiencies for PAmCherry range from 0.04 when expressed in *E. coli* (Wang et al., 2014) to about 0.5 in *Xenopus* oocytes (Durisic et al., 2014). This variability suggests that the expression system influences the functionality of fluorescent proteins, but may also reflect differences in image acquisition and data processing. It should be considered that many fluorescent proteins and dyes are sensitive to their specific microenvironment and it may not be taken for granted that the same values of *p* and *m* can be used in every circumstance. However, in our case, the same values gave reasonable results that matched the stoichiometries as determined by high-resolution clear native electrophoresis (Figure 4 – figure supplement 1).

To achieve the necessary spatial resolution across 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. While our implementation is based on SMLM, it could also be adapted for other types of microscopy as stated above. In our experiments, there was no difference between a linear fit of chromatic aberrations and local average correction (Figure 2 and Figure 2 – figure supplement 1); 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 large number of protein complexes. Therefore, if the protein complex of interest does not exist in a single oligomeric state, DCC-SMLM counting can only report an average stoichiometry. We simulated such cases but there was no clear

pattern that allowed users to determine whether such averaging occurs based on a single experiment only (Figure 4 – figure supplement 3). Although this is a limitation, similar caveats apply to most other methods that rely on the aggregation of information from multiple protein complexes. To exclude an effect of expression density on the stoichiometry of our observed proteins, we used cells which covered a wide range of expression densities. A plot of the colocalization ratio versus the density of expressed clusters (Figure 4 – figure supplement 4) did not reveal any significant correlation between the two values, effectively ruling out aggregation at higher expression densities. Future efforts should aim to explore whether additional information obtained from DCC-SMLM experiments or other experimental designs could be used to increase its usefulness in analyzing proteins that adopt a mixture of oligomeric states.

Using DCC-SMLM, we determined the stoichiometries of proteins from the SLC26 and SLC17 families. Studies using Blue native electrophoresis (Detro-Dassen et al., 2008) and X-ray and cryo-EM (Geertsma et al., 2015; Walter et al., 2019) previously reported a dimeric assembly for SLC26, whereas microscopy approaches supported a tetrameric assembly (Hallworth & Nichols, 2011; Sinha et al., 2010). 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 (Figure 4B-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 (Figure 4 – figure supplement 1). An earlier report suggested dimeric organization for SLC17 transporters based on SDS-PAGE results (Andersen, 2004) but a recent cryo-EM study reported a monomeric structure (F. Li et al., 2020). Functional expression studies have also demonstrated that vesicular glutamate transporters and sialin can function without accessory

subunits (Eriksen et al., 2016; Morin et al., 2004), 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 drawbacks and requirements of previous methods, making stoichiometry determination more reliable and much easier to perform.

Materials and Methods

Cell Culture and DNA Constructs. HEK293T cells and Flp-In T-REx 293 cells were grown in DH10 medium (DMEM supplemented with 50 U/mL penicillin-streptomycin and 10% FBS at 37 °C and 5% CO_2 (Tan et al., 2017)). 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.

Cleaning of Coverslips. Piranha solution was made by slowly adding 1 volume of hydrogen peroxide solution (30%, Merck) to 3 volumes of concentrated sulfuric acid (>95%, Sigma-Aldrich) while stirring. **Caution: Piranha solution reacts violently with organic matter and must be handled and disposed of according to institutional guidelines!** Coverslips (VWR) were placed in a glass container and freshly prepared piranha solution was added. Following incubation for 4-7 days at room temperature, coverslips were rinsed with double distilled water at least 20 times. Afterwards, coverslips were stored in double distilled water at 4 °C for experiments within the next 2 weeks.

Preparation of Cells 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 2-3 h before transfection. Cells were transfected with 0.5-1.0 µg plasmid DNA plus 1.0 µL Lipofectamine 2000 (Life Technologies) according to the manufacturers' instructions 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.

Afterwards, cells were washed with 2 ml 4 °C PBS for 3 times and afterwards fixed in 1 ml of -20 °C cold methanol (gradient grade for liquid chromatography, Merck) for 5 minutes at -20 °C. Methanol was then removed and the cells were washed 5 times with 4 °C cold PBS.

For experiments using aldehyde fixation, 4 g PFA (Merck) was dissolved in 80 ml PBS heated to 60 °C. Once completely dissolved, the solution was filled to 100 ml with PBS. 1 ml aliquots were stored at

-20 °C for long-term storage. Prior to fixation, aliquots were warmed to 37 °C and vortexed to dissolve precipitates. 8 µl of a 25% GA (electron microscopy grade, Agra Scientific) was added to the PFA aliquot for a final concentration of 0.2% (w/v) GA. Cells were washed 3 times with 37 °C warm PBS and fixed in 1 ml of the PFA/GA solution for 30 minutes at 37 °C. Afterwards, cells were washed 3 times with PBS, followed by incubation in 0.5% (w/v in PBS) ice cold $NaBH_4$ (Sigma-Aldrich) solution for 1 h at 4 °C.

Fixed cells were stored in 2 ml PBS in the dark at 4 °C until imaging within 24 h.

Preparation of Bead Samples for imaging. To prepare bead samples, 5 µl multi-color TetraSpeck fluorescent bead solution (100 nm diameter, Thermo Fisher Scientific) was diluted with 400 µl double distilled water, added to cleaned coverslips and let dry.

SMLM Imaging. Imaging was performed on a custom-built wide field and total internal reflection fluorescence (TIRF) microscope described previously (Stölting et al., 2015; Tang et al., 2015). We used 561 nm and 405 nm diode lasers for imaging and photoactivation of PAmCherry, respectively. Prior to imaging, the sample was washed 3 times with PBS and mounted in the chamber. 1 µl TetraSpeck fluorescent bead solution was diluted in 0.5 ml PBS, added to the coverslip and incubated for 1-2 h in the dark. Afterwards, the solution was removed and 1 ml PBS was added for imaging. To reduce vaporization, the chamber was covered.

The focal plane was adjusted using red color signals of the fluorescent beads on the coverslip surface. A transmission image was taken to record positions of cells and beads. Videos were recorded at an exposure of 85.59 ms per frame, during which the illumination lasted for 50 ms. Background signals were bleached by illumination with the 561 nm laser at full power for 500-1000 frames. Color channels were recorded sequentially (Table 3), starting with mVenus. Recordings were stopped when no signals appeared.

The intensity of the 405 nm laser used for photoactivation of PAmCherry was gradually increased during the recording to keep signals sparsely distributed over time. After recording both color channels,

the focus was re-adjusted using the fluorescent beads, and the displacement of the sample stage was taken as the axial drift. Recordings with an axial drift > 300 nm were discarded.

Extraction of SMLM signals. Raw videos were processed by SNSMIL (Tang et al., 2015) with the following parameters: optical magnification, 200; numerical aperture, 1.49; gain, 300; e/AD count, 11.9; excess noise factor square, 2; pixel diameter, 16 µm; bias offset (AD counts), 100; quality threshold, 1; PSF width tolerance, 3; fitting model, fixed SD 2D gaussian non-integral model. For recordings of the green channel (mVenus and beads), the peak emission wavelength was set to 528 nm and for the red channel (PAmCherry and beads) to 595 nm. Coordinates obtained from SNSMIL were translated from pixels to nm using the known size of 80×80 nm/pixel. Binary reconstructions were plotted using a custom written python script using the package Pillow (v.6.0.0).

Drift correction. The algorithm was adapted from a previously reported application using nanodiamonds (Yi et al., 2016). First, the drift was corrected relative to the first frame of each color. Beads close to the recorded cells were manually selected from the reconstructed image using Fiji (Schindelin et al., 2012) and precise coordinates were extracted from the output of SNSMIL using a custom python script. In a recording with n total frames, the average drift (δ_k) on the kth frame (k = 1,2,3,…,n) was calculated across all m beads (with the index $i = 1...m$) relative to the first frame as:

$$\delta_k = \frac{\sum_{i=1}^{m} \left(x_{i,k} - x_{i,1}, y_{i,k} - y_{i,1} \right)}{m_k} \quad [7]$$

The drift was subtracted from each localization (p_k)to give the drift corrected position for each frame (p'_k):

$$p'_k = p_k - \delta_k \quad [8]$$

As the red channel was recorded after the green channel, the corrected positions of the red channel were further corrected relative to the last frame in the green channel:

$$p''_{Red} = p'_{Red} - \delta_{Green,n} \quad [9]$$

Determination of Axial Chromatic Aberration (ACA). To determine ACA, a bead sample was used. First, the focus was set in the red channel. The stage was moved from 700 nm above to 700 nm below the focal plane in 100 nm steps and 10 frames in each color channel were recorded at each step. The mean intensities of the signals from each step were then analyzed.

Correction of Lateral Chromatic Aberration (LCA). For the determination of LCA, multiple beads (>10 beads) without cells were recorded in wide field mode for 50 frames in the green and red channel, respectively. Data from multiple recordings were concatenated into a single dataset (100-200 beads in total, evenly distributed around the field of view) for each color channel. In this dataset, the coordinates of signals were clustered using the DBSCAN (Density-Based Spatial Clustering of Applications with Noise) implementation of the python package scikit-learn (v.0.21.0) (Pedregosa et al., 2011) to assign localizations to individual beads. The parameter *eps* defines the upper threshold of the euclidean distances between any two localizations for a single bead and was empirically determined as 40 nm. The parameter *min samples* defines the lower threshold of localizations required for each cluster and was set to 50 to ensure that localizations must occur in every frame of the recording. Following clustering in separate color channels, distances to the nearest neighboring cluster in the other color channel were calculated and considered to be from the same bead if within 200 nm.

The second (red) was corrected relative to the first recorded color (green). For every bead in the dataset, the LCA was determined as the shift between the mean coordinates of the clustered localizations in the red (p_{Red}) and green (p_{Green}) channels:

$$A = p_{Red} - p_{Green} \quad [10]$$

In general, correction of LCA was performed relative to the green channel according to yield corrected coordinates for each localization:

$$p'''_{Red} = p''_{Red} - A \quad [11]$$

For the regional LCA correction (rCA) (Churchman & Spudich, 2012; Georgieva et al., 2016), the field of view was divided into 100 squares (each 1049×1049 nm). Within each square with the coordinates *(i,j)*, the mean LCA ($A_{i,j}$) was calculated by averaging the LCA from the individual beads found within this square. For squares without any beads, $A_{i,j}$ was interpolated from neighboring regions. rCA for

samples of interest was performed by subtracting $A_{i,j}$ from the localizations in each square according to Equation 11.

To correct LCA by linear fit (fCA) (Erdelyi et al., 2013; Kozubek & Matula, 2000), the LCA (*Ax*, *Ay*) was determined as described above along the positions of all beads detected in the red channel (*x*, *y*). The LCA varies across the field of view according to:

$$\begin{pmatrix} A_x \\ A_y \end{pmatrix} = K \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} \quad [12]$$

K, x_0 and y_0 were determined by a fit of the LCA and positions of all beads present in the sample.

With these values, the LCA (A_s) of any given localization with position (x_s, y_s) in samples of interest was then calculated by

$$A_S = K \begin{pmatrix} x_s - x_0 \\ y_s - y_0 \end{pmatrix} \quad [13]$$

LCA correction was then performed according to Equation 11.

Cluster analysis of mVenus and PAmCherry localizations. One emission event from either mVenus or PAmCherry lasts for several frames, generating a series of localizations that are temporally and spatially close(Durisic et al., 2014; Nan et al., 2013). Similar to the assignment of localizations to individual beads, we used DBSCAN for a cluster analysis of localizations. The parameter “eps” was empirically determined as 40 nm while “min samples” was set to 10 for mVenus and 6 for PAmCherry (Figure 5). Non-clustered signals were discarded from further analysis.

Colocalization Analysis. Mean coordinates of localization clusters were taken as the detected position of the fluorescent proteins. To determine the cutoff value of the inter-color distance for colocalization, we performed pair correlation analysis across both color channels (Sengupta et al., 2011; Sengupta & Lippincott-Schwartz, 2012). In short, the densities of PAmCherry localization clusters around any given mVenus cluster was calculated according to

$$g_{MF}(r) = \frac{a}{\pi N_M N_F} \sum_{i=1}^{N_M} \frac{d N_{Fi}(r)}{(2r + dr) dr} \quad [14]$$

in which a is the area of the selected ROI; N_M and N_F are the numbers of green and red clusters in the ROI, respectively; $dN_{Fi}(r)$ ($1 <= I <= N_M$) represents the number of red clusters that fall into the annulus around the ith cluster with the inner radius r and the outer radius $r+dr$. To accelerate computing, distances to all PAmCherry clusters within 400 nm from each mVenus cluster were calculated using the spatial.distance.cdist function from the python package SciPy v.1.2.1 (Virtanen et al., 2020). The values were grouped into bins of 20 nm width and the pair correlation calculated for each bin.

For comparison, distances to the nearest and second-nearest neighboring clusters were additionally plotted in a histogram and fit with a Gaussian curve.

Estimation of the Oligomeric State of the Protein of Interest. To estimate the oligomeric state of proteins of interest and the observed colocalization ratios, we defined the coefficient of mismatch, $M^2(n)$, as

$$M^2(n) \equiv 1 - \frac{\sum_i (x_i - \bar{x})^2}{\sum_i [x_i - E(n)]^2} \quad [15]$$

in which $\bar{x}$ is the mean of all observations (x_i) and $E(n)$ is the expected colocalization ratio for the stoichiometry n. The expected value is calculated from Equation 6 with the optimal values of m and p as determined by the fit to the calibration data (Figure 4A).

High-resolution clear native electrophoresis. HEK293T cells were transfected with 3 – 5 µg of plasmid DNA using the calcium phosphate method and treated as described previously (Guzman et al., 2017). In brief, cells were washed with ice cold PBS and incubated for 15 min with a lysis buffer containing 0.1 M sodium phosphate, either 0.5% (w/v) digitonin or 20 mM DDM, protease inhibitors and 20 mM iodoacetamide. The buffer was then transferred into a 1.5 ml Eppendorf tube. After centrifugation at 4 °C, an aliquot of the supernatant containing approximately 10 µg of protein was loaded into a native gel. The 4-14% acrylamide gradient gels were prepared as described(Wittig et al., 2007). The anode buffer contained 25 mM imidazole/HCl, pH 7.0 and the cathode buffer contained 50 mM tricine, 7.5 mM imidazole/HCl, pH 7.0 supplemented with the anionic detergent DOC (0.05%) and

the non-ionic detergent Triton X-100 (0.05%) (Wittig et al., 2007). The gels were run in a cold room (4 °C) and the starting voltage was set to 100 V. After 1 h, the voltage was increased to 150 V for another 2 h. Gels were scanned on a fluorescence gel scanner (Typhoon FLA 9500, GE Healthcare, Freiburg, Germany) at 100 µm resolution. mVenus was excited with the built-in 473 nm laser and the emission was filtered with a 530/20 bandpass filter. Gel images were visualized using Fiji (Schindelin et al., 2012).

Data availability. The latest version of the raw data and python scripts for analysis are available for download via https://doi.org/10.5281/zenodo.6000530.

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 Dr. 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.

Author contributions: *H.L.T., G.S. and Ch.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.

Andersen, S. S. L. (2004). Expression and purification of recombinant vesicular glutamate transporter VGLUT1 using PC12 cells and High Five insect cells. *Biological Procedures Online*, *6*(1), 105–112. https://doi.org/10/dc6m9v

Annibale, P., Vanni, S., Scarselli, M., Rothlisberger, U., & Radenovic, A. (2011). Quantitative Photo Activated Localization Microscopy: Unraveling the Effects of Photoblinking. *PLoS ONE*, *6*(7), e22678. https://doi.org/10/cxf64n

Bartoi, T., Augustinowski, K., Polleichtner, G., Gründer, S., & Ulbrich, M. H. (2014). Acid-sensing ion channel (ASIC) 1a/2a heteromers have a flexible 2:1/1:2 stoichiometry. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(22), 8281–8286. https://doi.org/10/f5449g

Betzig, E., Patterson, G. H., Sougrat, R., Lindwasser, O. W., Olenych, S., Bonifacino, J. S., Davidson, M. W., Lippincott-Schwartz, J., & Hess, H. F. (2006). Imaging Intracellular Fluorescent Proteins at Nanometer Resolution. *Science*, *313*(5793), 1642–1645. https://doi.org/10/dd8dkd

Canul-Tec, J. C., Assal, R., Cirri, E., Legrand, P., Brier, S., Chamot-Rooke, J., & Reyes, N. (2017). Structure and allosteric inhibition of excitatory amino acid transporter 1. *Nature*, *544*(7651), 446–451. https://doi.org/10/f93fz5

Chen, Y., Wei, L.-N., & Müller, J. D. (2003). Probing protein oligomerization in living cells with fluorescence fluctuation spectroscopy. *Proceedings of the National Academy of Sciences*, *100*(26), 15492–15497. https://doi.org/10/bqzw3c

Churchman, L. S., & Spudich, J. A. (2012). Single-Molecule High-Resolution Colocalization of Single Probes. *Cold Spring Harbor Protocols*, *2012*(2), 242–245. https://doi.org/10/fx5xs4

Coste, B., Xiao, B., Santos, J. S., Syeda, R., Grandl, J., Spencer, K. S., Kim, S. E., Schmidt, M., Mathur, J., Dubin, A. E., Montal, M., & Patapoutian, A. (2012). Piezo proteins are pore-forming

subunits of mechanically activated channels. *Nature*, *483*(7388), 176–181. https://doi.org/10/f4h7sd

Detro-Dassen, S., Schänzler, M., Lauks, H., Martin, I., zu Berstenhorst, S. M., Nothmann, D., Torres-Salazar, D., Hidalgo, P., Schmalzing, G., & Fahlke, C. (2008). Conserved dimeric subunit stoichiometry of SLC26 multifunctional anion exchangers. *The Journal of Biological Chemistry*, *283*(7), 4177–4188. https://doi.org/10/ccmgmt

Dickson, R. M., Cubitt, A. B., Tsien, R. Y., & Moerner, W. E. (1997). On/off blinking and switching behaviour of single molecules of green fluorescent protein. *Nature*, *388*(6640), 355–358. https://doi.org/10/d8ngjm

Doyle, D. A., Cabral, J. M., Pfuetzner, R. A., Kuo, A., Gulbis, J. M., Cohen, S. L., Chait, B. T., & MacKinnon, R. (1998). The Structure of the Potassium Channel: Molecular Basis of K+ Conduction and Selectivity. *Science*, *280*(5360), 69–77. https://doi.org/10/ff8wqd

Durisic, N., Laparra-Cuervo, L., Sandoval-Álvarez, Á., Borbely, J. S., & Lakadamyali, M. (2014). Single-molecule evaluation of fluorescent protein photoactivation efficiency using an *in vivo* nanotemplate. *Nature Methods*, *11*(2), 156–162. https://doi.org/10/f5rs9w

Erdelyi, M., Rees, E., Metcalf, D., Schierle, G. S. K., Dudas, L., Sinko, J., Knight, A. E., & Kaminski, C. F. (2013). Correcting chromatic offset in multicolor super-resolution localization microscopy. *Optics Express*, *21*(9), 10978–10988. https://doi.org/10/gf2n7g

Eriksen, J., Chang, R., McGregor, M., Silm, K., Suzuki, T., & Edwards, R. H. (2016). Protons Regulate Vesicular Glutamate Transporters through an Allosteric Mechanism. *Neuron*, *90*(4), 768–780. https://doi.org/10/f8mkkr

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96)*, 226–231.

Estévez, R., Boettger, T., Stein, V., Birkenhäger, R., Otto, E., Hildebrandt, F., & Jentsch, T. J. (2001). Barttin is a Cl- channel beta-subunit crucial for renal Cl- reabsorption and inner ear K+ secretion. *Nature*, *414*(6863), 558–561. https://doi.org/10.1038/35107099

Fricke, F., Beaudouin, J., Eils, R., & Heilemann, M. (2015). One, two or three? Probing the stoichiometry of membrane proteins by single-molecule localization microscopy. *Scientific Reports*, *5*(1), 1–8. https://doi.org/10/ggktjz

Geertsma, E. R., Chang, Y.-N., Shaik, F. R., Neldner, Y., Pardon, E., Steyaert, J., & Dutzler, R. (2015). Structure of a prokaryotic fumarate transporter reveals the architecture of the SLC26 family. *Nature Structural & Molecular Biology*, *22*(10), 803–808. https://doi.org/10/f7tpdt

Gendreau, S., Voswinkel, S., Torres-Salazar, D., Lang, N., Heidtmann, H., Detro-Dassen, S., Schmalzing, G., Hidalgo, P., & Fahlke, C. (2004). A Trimeric Quaternary Structure Is Conserved in Bacterial and Human Glutamate Transporters. *Journal of Biological Chemistry*, *279*(38), 39505–39512. https://doi.org/10/b6dfgq

Georgieva, M., Cattoni, D. I., Fiche, J.-B., Mutin, T., Chamousset, D., & Nollmann, M. (2016). Nanometer resolved single-molecule colocalization of nuclear factors by two-color super resolution microscopy imaging. *Methods*, *105*, 44–55. https://doi.org/10/f8w5jw

Gustafsson, N., Culley, S., Ashdown, G., Owen, D. M., Pereira, P. M., & Henriques, R. (2016). Fast live-cell conventional fluorophore nanoscopy with ImageJ through super-resolution radial fluctuations. *Nature Communications*, *7*, 12471. https://doi.org/10/f82gtk

Guzman, R. E., Bungert-Plümke, S., Franzen, A., & Fahlke, C. (2017). Preferential Association with ClC-3 Permits Sorting of ClC-4 into Endosomal Compartments. *Journal of Biological Chemistry*, jbc.M117.801951. https://doi.org/10/gcmnz7

Hallworth, R., & Nichols, M. G. (2011). Prestin in HEK cells is an obligate tetramer. *Journal of Neurophysiology*, *107*(1), 5–11. https://doi.org/10/dp85fx

Hummer, G., Fricke, F., & Heilemann, M. (2016). Model-independent counting of molecules in single-molecule localization microscopy. *Molecular Biology of the Cell*, *27*(22), 3637–3644. https://doi.org/10/f895sx

Kner, P., Chhun, B. B., Griffis, E. R., Winoto, L., & Gustafsson, M. G. L. (2009). Super-resolution video microscopy of live cells by structured illumination. *Nature Methods*, *6*(5), 339–342. https://doi.org/10/cckqwz

Kozubek, M., & Matula, P. (2000). An efficient algorithm for measurement and correction of chromatic aberrations in fluorescence microscopy. *Journal of Microscopy*, *200*(3), 206–217. https://doi.org/10/dn5zdr

Krashia, P., Moroni, M., Broadbent, S., Hofmann, G., Kracun, S., Beato, M., Groot-Kormelink, P. J., & Sivilotti, L. G. (2010). Human α3β4 Neuronal Nicotinic Receptors Show Different Stoichiometry if They Are Expressed in Xenopus Oocytes or Mammalian HEK293 Cells. *PLOS ONE*, *5*(10), e13611. https://doi.org/10/ds2h2h

Lee, S.-H., Shin, J. Y., Lee, A., & Bustamante, C. (2012). Counting single photoactivatable fluorescent molecules by photoactivated localization microscopy (PALM). *Proceedings of the National Academy of Sciences*, *109*(43), 17436–17441. https://doi.org/10/f4dj3c

Leeman, D. S., Hebestreit, K., Ruetz, T., Webb, A. E., McKay, A., Pollina, E. A., Dulken, B. W., Zhao, X., Yeo, R. W., Ho, T. T., Mahmoudi, S., Devarajan, K., Passegué, E., Rando, T. A., Frydman, J., & Brunet, A. (2018). Lysosome activation clears aggregates and enhances quiescent neural stem cell activation during aging. *Science (New York, N.Y.)*, *359*(6381), 1277–1283. https://doi.org/10/gc77km

Li, F., Eriksen, J., Finer-Moore, J., Chang, R., Nguyen, P., Bowen, A., Myasnikov, A., Yu, Z., Bulkley, D., Cheng, Y., Edwards, R. H., & Stroud, R. M. (2020). Ion transport and regulation in a synaptic vesicle glutamate transporter. *Science*, *368*(6493), 893–897. https://doi.org/10/ghtc7d

Li, H., & Vaughan, J. C. (2018). Switchable Fluorophores for Single-Molecule Localization Microscopy. *Chemical Reviews*, *118*(18), 9412–9454. https://doi.org/10/gffnrw

Liang, J. R., Lingeman, E., Ahmed, S., & Corn, J. E. (2018). Atlastins remodel the endoplasmic reticulum for selective autophagy. *The Journal of Cell Biology*, *217*(10), 3354–3367. https://doi.org/10/gd3w4g

Lussier, Y., Fürst, O., Fortea, E., Leclerc, M., Priolo, D., Moeller, L., Bichet, D. G., Blunck, R., & D'Avanzo, N. (2019). Disease-linked mutations alter the stoichiometries of HCN-KCNE2 complexes. *Scientific Reports*, *9*(1), 9113. https://doi.org/10/ghxvhr

Morin, P., Sagné, C., & Gasnier, B. (2004). Functional characterization of wild-type and mutant human sialin. *The EMBO Journal*, *23*(23), 4560–4570. https://doi.org/10/b7mfz7

Murata, Y., Iwasaki, H., Sasaki, M., Inaba, K., & Okamura, Y. (2005). Phosphoinositide phosphatase activity coupled to an intrinsic voltage sensor. *Nature*, *435*(7046), 1239–1243. https://doi.org/10/fwvsxn

Nagai, T., Ibata, K., Park, E. S., Kubota, M., Mikoshiba, K., & Miyawaki, A. (2002). A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nature Biotechnology*, *20*(1), 87–90. https://doi.org/10/cm5fpj

Nan, X., Collisson, E. A., Lewis, S., Huang, J., Tamgüney, T. M., Liphardt, J. T., McCormick, F., Gray, J. W., & Chu, S. (2013). Single-molecule superresolution imaging allows quantitative analysis of RAF multimer formation and signaling. *Proceedings of the National Academy of Sciences*. https://doi.org/10/f5gxxb

Nicovich, P. R., Owen, D. M., & Gaus, K. (2017). Turning single-molecule localization microscopy into a quantitative bioanalytical tool. *Nature Protocols*, *12*(3), 453–460. https://doi.org/10/gffnrk

Nothmann, D., Leinenweber, A., Torres-Salazar, D., Kovermann, P., Hotzy, J., Gameiro, A., Grewer, C., & Fahlke, C. (2011). Hetero-oligomerization of Neuronal Glutamate Transporters. *Journal of Biological Chemistry*, *286*(5), 3935–3943. https://doi.org/10/cdsrh4

Park, E., Campbell, E. B., & MacKinnon, R. (2017). Structure of a CLC chloride ion channel by cryo-electron microscopy. *Nature*, *541*(7638), 500–505. https://doi.org/10/f9kb9j

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*(85), 2825–2830.

Puchner, E. M., Walter, J. M., Kasper, R., Huang, B., & Lim, W. A. (2013). Counting molecules in single organelles with superresolution microscopy allows tracking of the endosome maturation trajectory. *Proceedings of the National Academy of Sciences*, *110*(40), 16015–16020. https://doi.org/10/f5bz39

Ranade, S. S., Qiu, Z., Woo, S.-H., Hur, S. S., Murthy, S. E., Cahalan, S. M., Xu, J., Mathur, J., Bandell, M., Coste, B., Li, Y.-S. J., Chien, S., & Patapoutian, A. (2014). Piezo1, a mechanically activated ion channel, is required for vascular development in mice. *Proceedings of the National Academy of Sciences*, *111*(28), 10347–10352. https://doi.org/10/f59t35

Rust, M. J., Bates, M., & Zhuang, X. (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods*, *3*(10), 793–796. https://doi.org/10/fwrhq3

Schägger, H., & von Jagow, G. (1991). Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Analytical Biochemistry*, *199*(2), 223–231. https://doi.org/10/dkdw6j

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: An open-source platform for biological-image analysis. *Nature Methods*, *9*(7), 676–682. https://doi.org/10.1038/nmeth.2019

Scholl, U., Hebeisen, S., Janssen, A. G. H., Müller-Newen, G., Alekov, A., & Fahlke, C. (2006). Barttin modulates trafficking and function of ClC-K channels. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(30), 11411–11416. https://doi.org/10.1073/pnas.0601631103

Sengupta, P., Jovanovic-Talisman, T., Skoko, D., Renz, M., Veatch, S. L., & Lippincott-Schwartz, J. (2011). Probing protein heterogeneity in the plasma membrane using PALM and pair correlation analysis. *Nature Methods*, *8*(11), 969–975. https://doi.org/10/dbm6dt

Sengupta, P., & Lippincott-Schwartz, J. (2012). Quantitative analysis of photoactivated localization microscopy (PALM) datasets using pair-correlation analysis. *BioEssays*, *34*(5), 396–405. https://doi.org/10/fz4r5b

Shivanandan, A., Deschout, H., Scarselli, M., & Radenovic, A. (2014). Challenges in quantitative single molecule localization microscopy. *FEBS Letters*, *588*(19), 3595–3602. https://doi.org/10/f6hz45

Sinha, G. P., Sabri, F., Dimitriadis, E. K., & Iwasa, K. H. (2010). Organization of membrane motor in outer hair cells: An atomic force microscopic study. *Pflügers Archiv - European Journal of Physiology*, *459*(3), 427–439. https://doi.org/10/fbb24f

Stanly, T. A., Fritzsche, M., Banerji, S., García, E., Serna, J. B. de la, Jackson, D. G., & Eggeling, C. (2016). Critical importance of appropriate fixation conditions for faithful imaging of receptor microclusters. *Biology Open*, *5*(9), 1343–1350. https://doi.org/10/ghx6bm

Stölting, G., Fischer, M., & Fahlke, C. (2014). ClC-1 and ClC-2 form hetero-dimeric channels with novel protopore functions. *Pflügers Archiv - European Journal of Physiology*, *466*(12), 2191–2204. https://doi.org/10/f6pp3r

Stölting, G., Oliveira, R. C. de, Guzman, R. E., Miranda-Laferte, E., Conrad, R., Jordan, N., Schmidt, S., Hendriks, J., Gensch, T., & Hidalgo, P. (2015). Direct Interaction of CaVβ with Actin Up-regulates L-type Calcium Currents in HL-1 Cardiomyocytes. *Journal of Biological Chemistry*, *290*(8), 4561–4572. https://doi.org/10/f63x6s

Subach, F. V., Malashkevich, V. N., Zencheck, W. D., Xiao, H., Filonov, G. S., Almo, S. C., & Verkhusha, V. V. (2009). Photoactivation mechanism of PAmCherry based on crystal structures of the protein in the dark and fluorescent states. *Proceedings of the National Academy of Sciences*, *106*(50), 21097–21102. https://doi.org/10/fkf33f

Subach, F. V., Patterson, G. H., Manley, S., Gillette, J. M., Lippincott-Schwartz, J., & Verkhusha, V. V. (2009). Photoactivatable mCherry for high-resolution two-color fluorescence microscopy. *Nature Methods*, *6*(2), 153–159. https://doi.org/10/fhxdwk

Tan, H., Bungert-Plümke, S., Fahlke, C., & Stölting, G. (2017). Reduced Membrane Insertion of CLC-K by V33L Barttin Results in Loss of Hearing, but Leaves Kidney Function Intact. *Frontiers in Physiology*, *8*. https://doi.org/10/ghj439

Tanaka, K. A. K., Suzuki, K. G. N., Shirai, Y. M., Shibutani, S. T., Miyahara, M. S. H., Tsuboi, H., Yahara, M., Yoshimura, A., Mayor, S., Fujiwara, T. K., & Kusumi, A. (2010). Membrane molecules mobile even after chemical fixation. *Nature Methods*, *7*(11), 865–866. https://doi.org/10/fw3zrx

Tang, Y., Dai, L., Zhang, X., Li, J., Hendriks, J., Fan, X., Gruteser, N., Meisenberg, A., Baumann, A., Katranidis, A., & Gensch, T. (2015). SNSMIL, a real-time single molecule identification and

localization algorithm for super-resolution fluorescence microscopy. *Scientific Reports*, *5*(1), 11073. https://doi.org/10/gg329r

Thiemann, A., Gründer, S., Pusch, M., & Jentsch, T. J. (1992). A chloride channel widely expressed in epithelial and non-epithelial cells. *Nature*, *356*(6364), 57–60. https://doi.org/10/b9j62m

Ulbrich, M. H., & Isacoff, E. Y. (2007). Subunit counting in membrane-bound proteins. *Nat Meth*, *4*(4), 319–321. https://doi.org/10/bvx6n2

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … van Mulbregt, P. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*(3), 261–272. https://doi.org/10.1038/s41592-019-0686-2

Walter, J. D., Sawicka, M., & Dutzler, R. (2019). Cryo-EM structures and functional characterization of murine Slc26a9 reveal mechanism of uncoupled chloride transport. *ELife*, *8*, e46986. https://doi.org/10/ghxvmm

Wang, S., Moffitt, J. R., Dempsey, G. T., Xie, X. S., & Zhuang, X. (2014). Characterization and development of photoactivatable fluorescent proteins for single-molecule–based superresolution imaging. *Proceedings of the National Academy of Sciences*, *111*(23), 8452–8457. https://doi.org/10/f56hgs

Wittig, I., Karas, M., & Schägger, H. (2007). High resolution clear native electrophoresis for in-gel functional assays and fluorescence studies of membrane protein complexes. *Molecular & Cellular Proteomics: MCP*, *6*(7), 1215–1225. https://doi.org/10/cjqpkm

Yi, J., Manna, A., Barr, V. A., Hong, J., Neuman, K. C., & Samelson, L. E. (2016). madSTORM: A superresolution technique for large-scale multiplexing at single-molecule accuracy. *Molecular Biology of the Cell*, *27*(22), 3591–3600. https://doi.org/10/f8945f

Zhao, M., Nicovich, P. R., Janco, M., Deng, Q., Yang, Z., Ma, Y., Böcking, T., Gaus, K., & Gooding, J. J. (2018). Ultralow- and Low-Background Surfaces for Single-Molecule Localization Microscopy of Multistep Biointerfaces for Single-Molecule Sensing. *Langmuir*, *34*(34), 10012–10018. https://doi.org/10/gd788t

Zheng, J., Du, G.-G., Anderson, C. T., Keller, J. P., Orem, A., Dallos, P., & Cheatham, M. (2006). Analysis of the Oligomeric Structure of the Motor Protein Prestin *. *Journal of Biological Chemistry*, *281*(29), 19916–19924. https://doi.org/10/d5h9pg

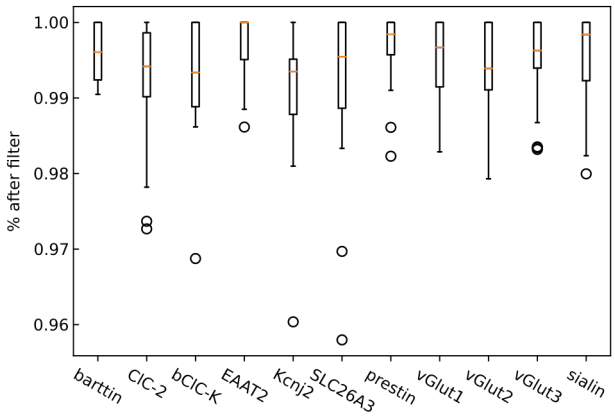

Figure 1 – figure supplement 1. Fraction of proteins remaining after the removal of large clusters.

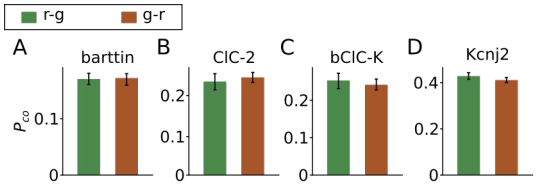

Figure 1 – figure supplement 2. *(A - E)* The colocalization ratios of mVenus with PAmCherry fused to the C-terminus of five different proteins transiently transfected in HEK293T cells fixed with cold methanol. The two fusion patterns POI-PAmCherry-mVenus (r-g) and POI-mVenus-PAmCherry (g-r) are compared for each protein.

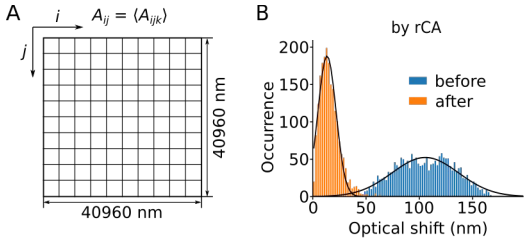

Figure 2 – figure supplement 1. Regional averaging for the correction of lateral chromatic abberation. *(A)* The whole microscopic field of view was divided into 100 squares and the mean lateral chromatic aberration (A_{ijk}) measured in each square was taken as the offset (A_{ij}) for correction. *(B)* Distribution of the optical shift before (blue) and after (orange) correction by rCA. Data are from the same recordings as in Figure 2D. Black lines show the result of a Gaussian fit.

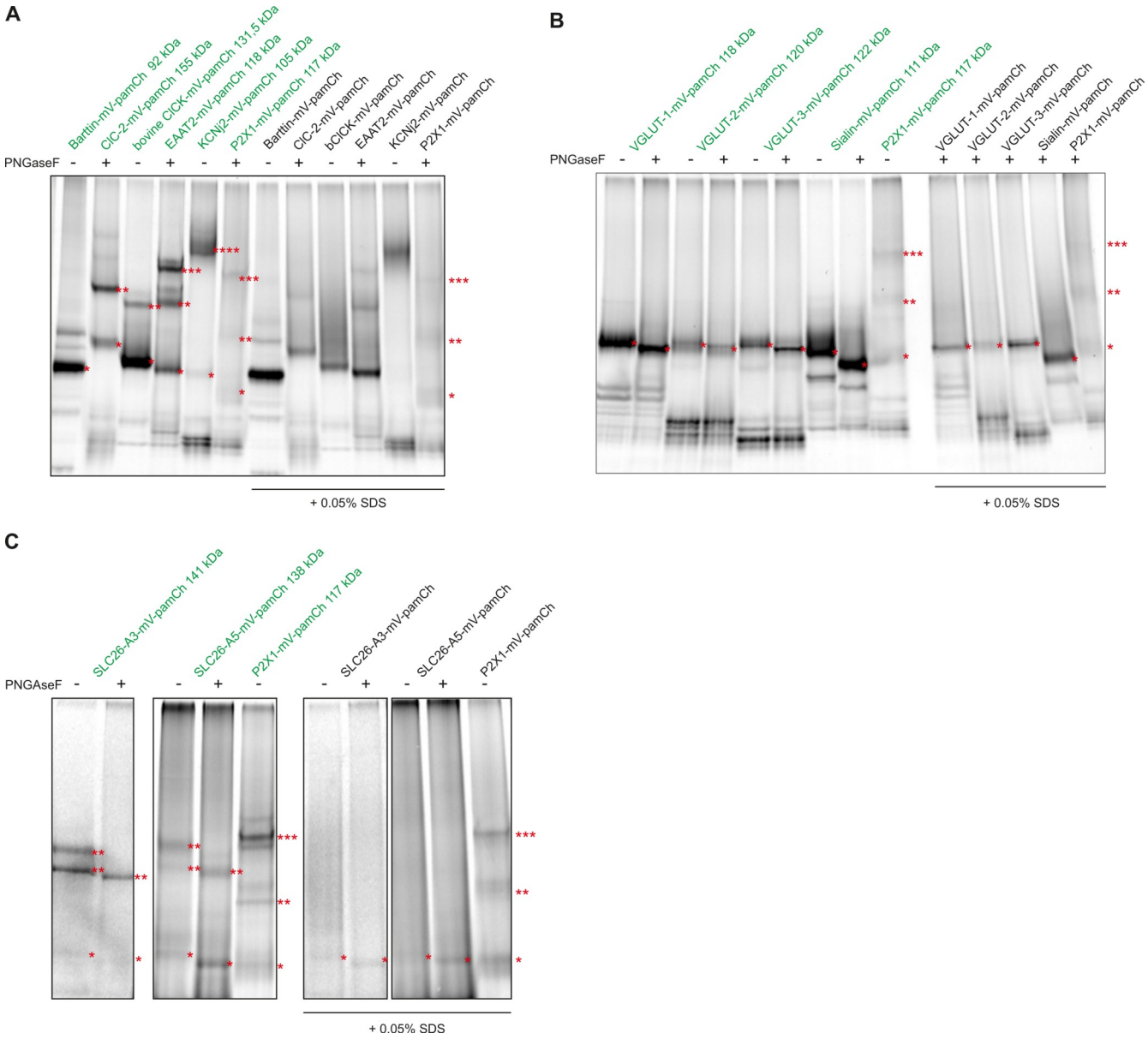

Figure 4 – figure supplement 1. High-resolution clear native gel electrophoresis (hrCNE) of reference proteins and proteins of interest. *(A)* hrCNE of standard proteins with known stoichiometry. Depending on their oligomeric state, protein bands are indicated with an increasing number of red stars (monomer,

*; dimer, **, timer, ***; tetramer, ****). *(B)* hrCNE of VGLUTs and sialin. To minimize artifacts due to protein glycosylation, proteins were treated with PNGaseF at 34 °C for 30 min. For comparison, enzyme treated (+) and untreated (-) samples were loaded. *(C)* hrCNE of SLC26A3 and SLC26A5/prestin. Both proteins are natively glycosylated, therefore dimers exhibit two bands in non-PNGaseF treated samples. Proteins were also treated with a final concentration of 0.05% SDS for 15 min at 37 °C to dissociate oligomers (right half of the gel). The calculated molecular masses of protein monomers are shown in green.

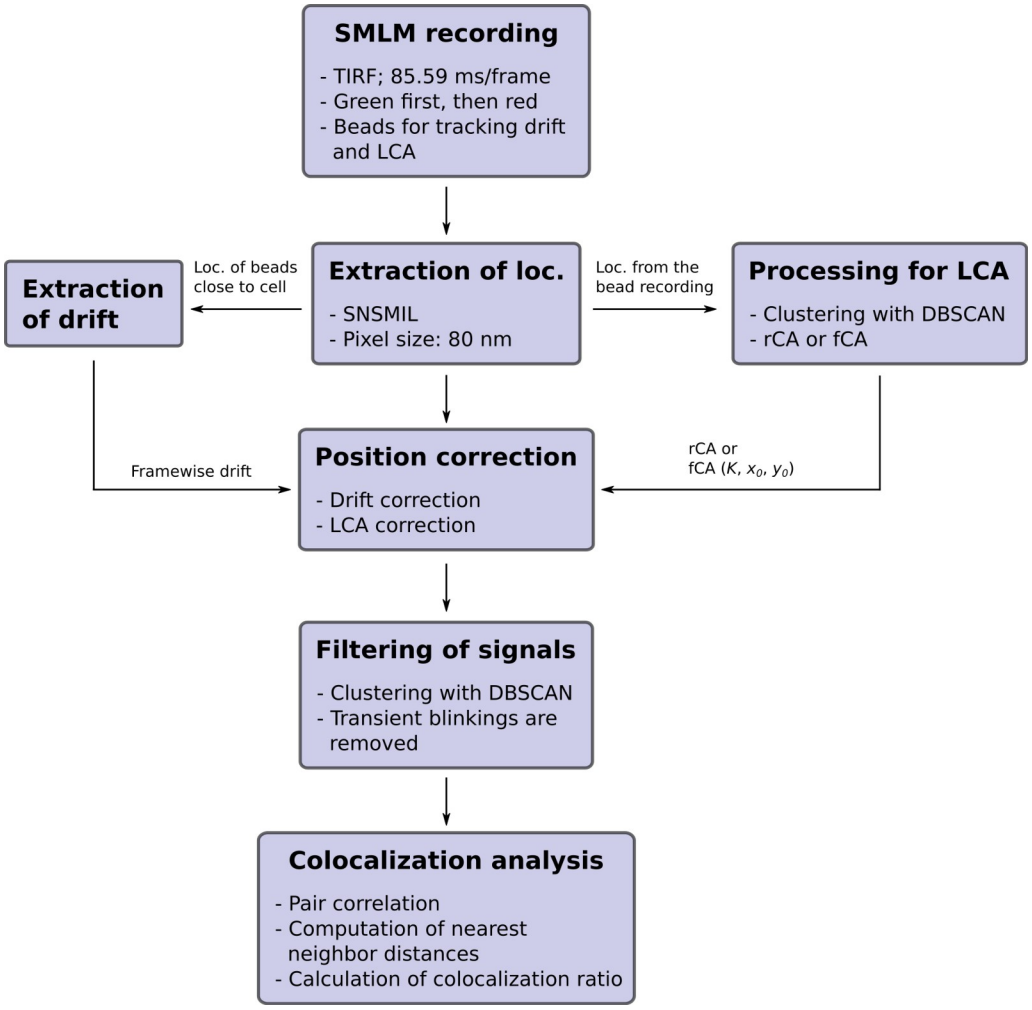

Figure 4 – figure supplement 2. The working paradigm of DCC-SMLM.

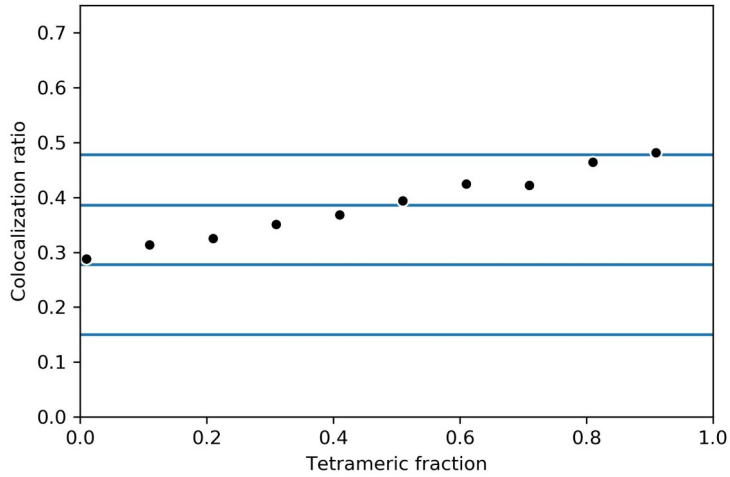

Figure 4 – figure supplement 3. Summary of 1000 simulations per data point of a mixed population of dimeric and tetrameric proteins. Proteins were simulated with p = 0.15, m = 1 at a density of 5 clusters per µm². Protein coordinates were set randomly and DBSCAN clustering was applied as stated in the methods. Blue lines represent expected values for monomeric through tetrameric values. The standard error of the mean is smaller than the symbols.

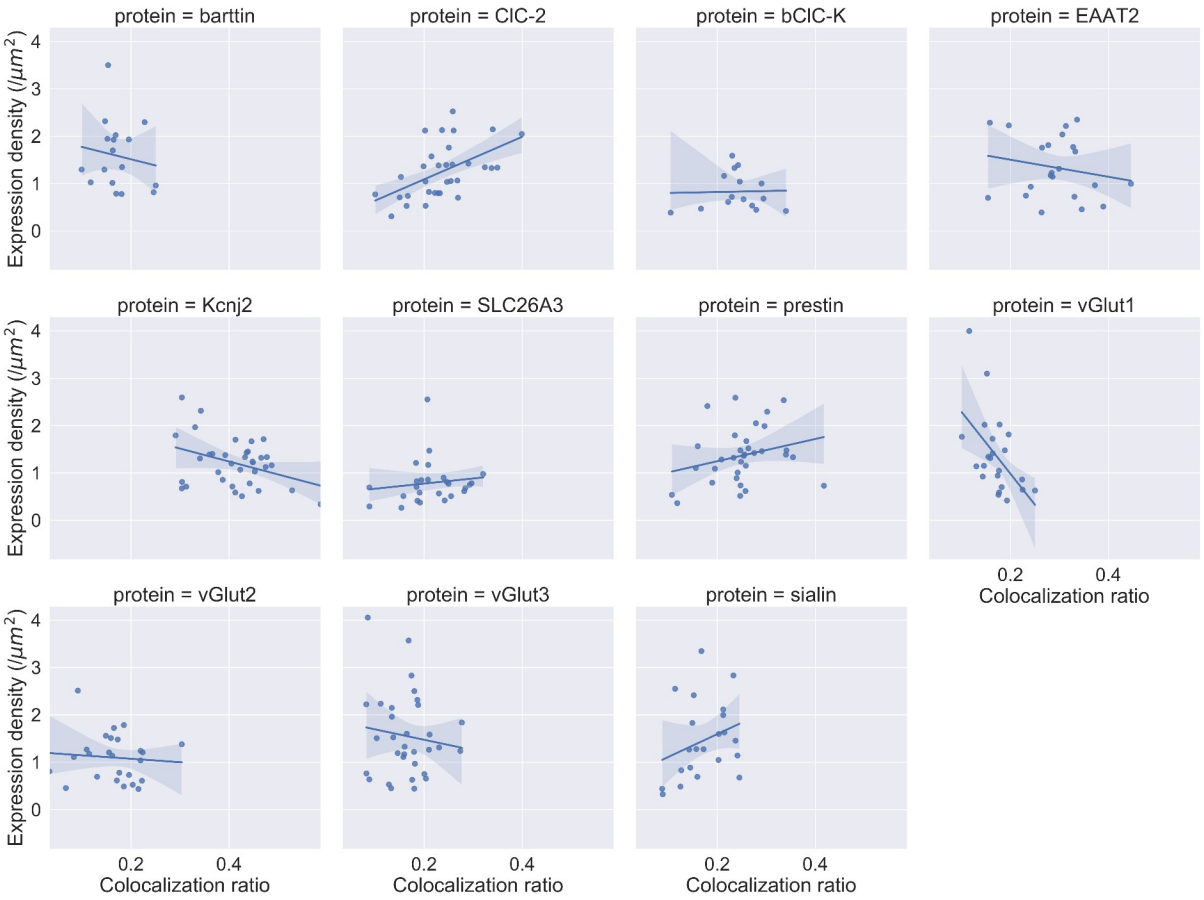

Figure 4 – figure supplement 4. Plots of the experimentally observed expression densities versus the colocalization ratio do not reveal a consistent correlation between these parameters. Blue lines represent a fit of a linear function to the data and the shaded areas represent the 95% CI of this fit as determined by 10,000 bootstrap resamples each.

Table 1. Colocalization ratio (*P*) of mVenus clusters with PAmCherry clusters of fusion proteins fixed with PFA+GA or cold methanol.

	PFA + GA		**Cold methanol**	
Protein	***P***	***N***	***P***	***N***
ClC-2	0.215	19	0.237	32
bClC-K	n/a	n/a	0.241	15
EAAT2	n/a	n/a	0.293	22
Kcnj2	0.248	23	0.411	35
SLC26A3	n/a	n/a	0.214	26
Prestin	0.176	27	0.251	33
vGlut1	0.119	25	0.170	23
vGlut2	0.111	25	0.165	25
vGlut3	0.131	26	0.163	32
sialin	n/a	n/a	0.172	22
barttin	0.078	27	0.172	17

Table 2. The linkers used in the fusion of proteins of interest (POI) to FP. In this project two fusion patterns were tested: POI-Linker #1-PAmCherry-Linker #2-mVenus (r-g) and POI-Linker #3-mVenus-Linker #4-PAmCherry (g-r) (both written from the N- to C-terminus) with the following linker sequences:

Name	Amino acid sequence	Length (a.a.)
Linker #1	GGSGG PSGLR SGSGG GSASG GSGS	24
Linker #2	PPVGT ELGST	10
Linker #3	GGSGG PGGSG GSPVG TELGS T	21
Linker #4	GSGSG GGSAS GGSGS	15

Table 3. The settings used for SMLM imaging. The dichroic mirror Di01-R442/514/561-25x36 was used for imaging of both FP. All filters and dichroic mirrors were acquired from Semrock. The laser power was measured at the sample plane and before the rear entrance of the objective. The power of the 405 nm activation laser for PAmCherry was varied between 4.4 µW and 10.2 mW near the rear entrance, corresponding to 3.0 µW to 4.8 mW in the sample plane.

Channel	FP	Excitation laser	Emission filter	Collection range
Green	mVenus	514 nm @ 6.1/4.4 mW	FF01-485/537/627-25 + FF03-525/50-25	526 – 546 nm
Red	PAmCherry	561 nm @ 6.5/5.4 mW	FF01-609/57-25	580 – 537 nm