Determination of oligomeric states of proteins via dual-color colocalization with single molecule localization microscopy

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

[a]Institute of Biological Information Processing, Molecular and Cellular Physiology (IBI-1), Forschungszentrum Jülich, 52428 Jülich, Germany; [b]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 oligomeric state of plasma membrane proteins is the 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 oligomeric states *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 oligomeric states of proteins, subunits are labeled with fluorescent proteins that only emit light following activation or conversion at different wavelengths. Typically, individual molecules 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 recall rate 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 oligomeric state 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 temporal clustering of localizations 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 oligomeric state of two SLC26 multifunctional anion exchangers and to determine the oligomeric state of four members of the SLC17 family of organic anion transporters.

single molecule localization microscopy | fluorescent protein | protein oligomeric state

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 oligomeric states of proteins 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 proteins exhibit a different oligomeric state 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 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 or converted 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 the oligomeric state of proteins: 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 and emission of each fluorophore need to be well separated in time. This criterion can only be fulfilled 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 the 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 fluorescent marker detected during the imaging period (i.e., the recall rate) is always less than one. This likely arises from a combination of incomplete protein translation or misfolding, failure to get activated or converted and pre-imaging bleaching (Durisic et al., 2014; Nicovich et al., 2017). Thus, the observed counts are not equal to the oligomeric state of interest but instead follow a binomial distribution that depends on both the real number of subunits per complex and the recall rate. This binomial strategy demands a recall rate of at least 0.5, to ensure reasonable resolution between different oligomeric states (Durisic et al., 2014; Ulbrich & Isacoff, 2007). Such a demand for high recall rates poses challenges with respect to fluorescent protein selection, sample preparation, data acquisition and post-imaging data processing (Shivanandan et al., 2014). These criteria and drawbacks, as a major setback, have limited the application of SMLM on counting.

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 the background noise and low recall rates of fluorescent proteins, and does not require stringent control of the activation laser. We tested the new DCC-SMLM approach on proteins with known subunit composition, and then used it to study proteins from two plasma membrane transporter families with controversial or unknown oligomeric states: 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 occurring 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 state (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

DCC-SMLM for Determining Oligomeric States. A protein complex of n identical subunits, each labeled with a single indicator fluorescent protein (F) with a recall rate of p, will be detected with a probability P as

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

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

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

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 ratio of the number of proteins showing fluorescence from both M and F (N_{MF}) to that of the proteins showing fluorescence from M (N_M),

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

Both N_{MF} and N_M can be counted from dual-color microscopy images (Figure 1B). The values of P obtained for proteins with known oligomeric states were used to solve Equation 1 for p, which could then be used to determine unknown oligomeric states (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. However, based on our experience with solubilizing ion channels and transporters tagged with fluorescent proteins (Guzman et al., 2022; Stölting et al., 2014; Stölting, Bungert-Plümke, et al., 2015; Tan et al., 2017), we expected that a small fraction of fluorescent proteins might be cleaved off within the linker region or within M, leading to the simultaneous removal of both F and M. We introduced an additional factor, m, to describe the probability of this phenomenon occurring. Notwithstanding that we cannot exclude other possibilities contributing to the loss of both fluorescent proteins, m remains generally as the probability of both M and F simultaneously being dysfunctional. Dysfunction of F alone is still represented by the factor p. With m included, the probability P was hence described as

$$P = \sum_{k=1}^{n} \left\{ \frac{\binom{n}{k}(1-m)^k m^{n-k}}{1-m^n} \left[1-(1-p)^k\right] \right\} \quad [4]$$

To assess how the values of m and p might alter the performance of the DCC strategy, an *in silico* sensitivity analysis based on Equation 4 demonstrated that the optimal separation from monomers to tetramers roughly occurred in the range where $0.2 < p < 0.4$, while m should generally remain below 0.3 (Figure 1 – figure supplement 1).

In SMLM, diffraction-limited video frames, which capture the temporally isolated emission events of individual fluorophores, are fit to yield 2D-coordinates associated with additional attributes such as

intensity and precision of the emission event. These multi-feature coordinates are termed "localizations", and are used to reconstruct super-resolution images after processing (Betzig et al., 2006). The emission from individual fluorophores typically remains active over several frames, dependent on the video acquisition rate and fluorophore kinetics (e.g., for PAmCherry, see Durisic et al., 2014; Subach, Patterson, et al., 2009), resulting in a cluster of localizations per emission event. Ideally, when the sample drift and chromatic aberrations are well corrected, all emission events from a single point emitter (e.g., a fluorescent protein) in a fixed sample fall into the same cluster.

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 high tolerance to cellular acidification and the 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.

To deal with background signals, we partially corrected for the background by setting a minimum number of localizations for clustering (see Methods). However, some background signals were still indistinguishable from genuine PAmCherry signals, which explains the higher background level in that color channel (Figure 1D-F). To reduce the contribution of the background noise to the final result, as illustrated in Figure 1F, we corrected P as

$$P = \frac{N_{MF} - N_M P_{rb}}{N_M - N_M P_{rb}} \quad [5]$$

in which P_{rb} is the probability of the mVenus cluster colocalizing with red background clusters. For any given mVenus cluster, the number (N_{rb}) of background clusters from the red channel located in its radial vicinity d follows the Poisson distribution: $N_{rb} \sim \mathrm{Pois}(\pi d^2 D_{rb})$, in which D_{rb} is the average density of red background clusters measured from the area on the coverslip outside of the cell. Therefore, we have

$$P_{rb} = 1 - e^{-\pi d^2 D_{rb}} \quad [6]$$

185 To prevent miscounting due to protein aggregation, we excluded the small fraction of clusters with large
186 diameters (> 500 nm) (Figure 1 – figure supplement 2).
187
188 **Correction of Sample Drift, Chromatic Aberrations and Definition of Colocalization.** The resolution
189 of SMLM is compromised by sample drift and optical aberrations, which affect the determination of
190 colocalization. We recorded sample drift using TetraSpeck beads as fiducial markers and corrected the
191 drift during post-imaging data processing (Yi et al., 2016). This correction reduced the radius of clustered
192 signals to ~30 nm (Figure 2A and B). Since our approach relies on colocalization of the marker and the
193 indicator in two color channels, we also corrected chromatic aberrations resulting from different
194 refractive indices at different emission light wavelengths. We confirmed earlier reports that lateral
195 chromatic aberration can be described as a set of linear functions centered on the optical axis of the
196 microscope (Erdelyi et al., 2013; Kozubek & Matula, 2000). Values from a linear fit of the measured
197 inter-color distances of individual beads to their coordinates were used for correction (Figure 2C). This
198 way the lateral chromatic aberration could be reduced to 13.8 ± 8.9 nm for bead samples (Figure 2D and
199 E). Chromatic aberrations may also be corrected by determining regional averages with similar results
200 (Figure 2 – figure supplement 1) (Churchman & Spudich, 2012; Georgieva et al., 2016). The axial
201 chromatic aberration determined as the focus offset between colors was less than 100 nm and thus
202 considered negligible for SMLM signal registration in only two dimensions (Figure 2F). Unlike
203 fluorescent proteins in cellular samples, beads show an intense and long-lasting fluorescent signal. To
204 evaluate the correction in fixed cell samples and to determine the threshold for colocalization, we applied
205 the correction to cell samples expressing monomeric barttin (Scholl et al., 2006, Estévez et al., 2001), or
206 tetrameric K_{ir}2.1 (Kcnj2, Doyle et al., 1998), each subunit tagged with a copy of mVenus and PAmCherry
207 (Figure 2G). We calculated the radial distribution function (RDF) of the red clusters around each mVenus
208 cluster, Figure 2H. Prior to the correction of chromatic aberrations, the RDFs from both proteins showed
209 high densities peaked at a distance of about 150 nm. Following correction, highest densities of red
210 clusters were found within a radius of ~100 nm around mVenus clusters for both proteins, representing
211 the achievable two-color resolution of our experimental setup. We therefore chose 100 nm as the
212 threshold distance for colocalization.
213

Fixation Conditions and the Fusion Pattern of the Fluorescent Proteins. 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 we observed a significantly higher signal density, a higher ratio of PAmCherry clusters to mVenus clusters and a higher colocalization ratio than with the aldehyde fixation (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 compared with PFA/GA fixation and, therefore, we chose cold methanol as our routine fixation reagent.

Since proteins of interest (POI) were tagged with both PAmCherry and mVenus, we investigated whether the fusion pattern affected the outcome. We compared the colocalization ratios when PAmCherry and mVenus were tagged to the C-terminus of POIs in two different ways: either PAmCherry or mVenus preceding the other. The results, as shown in Figure 3 – figure supplement 1, indicated that the fusion pattern did not change the outcome significantly.

Calibration of the DCC Model. Given that the recall rate (p) of fluorescent proteins may differ between experimental setups (Durisic et al., 2014; Wang et al., 2014), we first calibrated the DCC model (Equations 4, 5 and 6) using standards with known oligomeric state: the monomeric ClC-K chloride channel β-subunit barttin (Scholl et al., 2006, Estévez et al., 2001), 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 $K_{ir}2.1$ (Kcnj2, Doyle et al., 1998). All proteins were C-terminally tagged with PAmCherry and mVenus via two short peptide linkers (Table 2).

244 As an independent confirmation of the reported oligomeric state from the literature as well as to confirm
245 the function of our constructs, we performed high-resolution clear native electrophoresis (hrCNE) with
246 these proteins (Figure 4 – figure supplement 2). Our results also demonstrate that the use of a two
247 fluorescent proteins linked to the proteins in our analysis did not interfere with their reported quaternary
248 structure.
249
250 After correction for the drift and chromatic aberrations, localizations were assigned to clusters using the
251 DBSCAN algorithm (Ester et al., 1996). Background subtracted colocalization ratios for mVenus and
252 PAmCherry clusters were then calculated (Equations 5 and 6, Figure 4A). We used the known oligomeric
253 states (n) to obtain the recall rate (p) of PAmCherry from Equation 1, or p and the modification factor
254 (m) from Equation 4 of the DCC model. A global fit of bootstrap resampled (10,000 resamples)
255 experimental data from all five proteins with known oligomeric states to Equation 1 suggested a low
256 recall rate (p) of 0.12 (95% confidence interval (CI): 0.12 – 0.13) for PAmCherry. In general, the fit did
257 describe the data reasonably well ($R^2 = 0.95$) but showed a large deviation for the monomeric reference
258 protein as expected (Figure 4A). Fitting to Equation 4 gave values of 0.30 (95% CI: 0.21 – 0.39) for m
259 and 0.17 (95% CI: 0.15 – 0.18) for p. As expected, this model only aligns slightly better with the data (R^2
260 $= 0.97$) compared to the simpler model (Equation 1–3). However, the improvement is especially visible
261 for the monomeric reference protein (Figure 4A), and we therefore chose Equation 4 to analyze proteins
262 with contradictory or unknown oligomeric states.
263
264 **Using DCC-SMLM to Determine the Oligomeric States of SLC26 and SLC17 Proteins.** We next
265 used DCC-SMLM to resolve the oligomeric states of several SLC26 and SLC17 proteins. The
266 colocalization ratios of mVenus and PAmCherry clusters for these proteins were determined (Figure 4B)
267 and compared with those expected for different oligomeric states using the DCC model (Equation 4). To
268 quantify the difference between expected and observed values, we introduced the coefficient of mismatch
269 M^2. A M^2 value of 0 indicates a perfect match, with larger values indicating larger differences. The
270 calculated value is dependent on an accurate determination of p and m in Equation 4 as well as of the
271 colocalization ratios of the proteins of interest. To include the combined error of these measurements, we
272 used bootstrap resampling over all of these parameters to calculate means and 95% confidence intervals.
273 As shown in Figure 4 C-H, M^2 values reached their lowest values for a dimeric state of SLC26 proteins

274 and for a monomeric state of all SLC17 proteins. As predicted for a monomer, accurate determination
275 with Equation 1 alone (only *p*, no *m*) was not feasible as seen for the SLC17 proteins (Figure 4 – figure
276 supplement 1).
277
278 Even using Equation 4, the confidence interval at the dimeric state for SLC26A3 overlapped with that of
279 the monomeric state (Figure 4C), preventing the confident designation as a dimer based on the M^2 values
280 alone. To resolve these uncertainties, we additionally used a two-sample Kolmogorov-Smirnov (KS) test
281 (Figure 4I–N), which allows for the determination of the most likely oligomeric state by comparing the
282 colocalization ratios from the protein of interest with those from proteins with known stoichiometry,
283 thereby foregoing the need to determine *p* and *m*. With this test, we confirmed the results of the M^2 values
284 and also demonstrated that SLC26A3 is most likely a dimeric protein as it showed a significantly different
285 distribution compared to reference proteins with other oligomeric states (Figure 4I). Moreover, we
286 independently confirmed our results for the oligomeric states of SLC26 and SLC17 proteins with high-
287 resolution clear native electrophoresis utilizing the same constructs and expression system as in our
288 DCC-SMLM experiments (Figure 4 – figure supplement 2).
289
290 **Discussion**
291 We have developed a novel DCC-SMLM technique to determine the oligomeric state of plasma
292 membrane protein complexes *in situ* in mammalian cells. Using mVenus as the marker fluorescent protein
293 and PAmCherry as the indicator fluorescent protein for counting, we established a dimeric state for two
294 representatives of the SLC26 family, the epithelial anion exchanger SLC26A3 and the motor protein
295 prestin (Slc26a5); and a monomeric state for four representatives of the SLC17 family, the vesicular
296 glutamate transporters vGlut1, 2, and 3 and the lysosomal sialic acid transporter sialin.
297
298 Our DCC-SMLM approach (Figure 4 – figure supplement 3) overcomes several limitations of the
299 methods commonly used to determine the oligomeric state of proteins. Unlike for biochemical methods,
300 no solubilization or purification of proteins is necessary as DCC-SMLM works with proteins *in situ*. In
301 comparison to photo-bleaching experiments, DCC-SMLM can be applied reliably even in small
302 mammalian cells where the high densities of protein expression are usually problematic. In contrast to
303 previously reported quantitative SMLM methods, DCC-SMLM simplifies the procedure and offers

304 greater reliability of counting. Binomial SMLM counting strategies require clustering of signals in at
305 least three dimensions (x, y and time) to identify individual protein subunits. This demands stringent
306 control of the activation laser and intensive post-imaging data processing (Lee et al., 2012). Another
307 recently published SMLM-based quantification strategy also foregoes the need for temporal clustering
308 as it utilizes the blinking behavior of some fluorescent proteins (Fricke et al., 2015; Hummer et al., 2016).
309 This, however, requires fluorescent proteins that show blinking behavior (multiple on ↔ off sequences)
310 which is not guaranteed for all fluorescence proteins or dyes (Durisic et al., 2014; Subach, Malashkevich,
311 et al., 2009). Additionally, all previously published SMLM quantification and photobleching step
312 analysis methods require large recall rates ($p \geq 0.5$) as they are otherwise unable to resolve the small
313 oligomers typically observed for most membrane proteins. In contrast, DCC-SMLM only relies on the
314 relationship between the total detection probability for a protein complex and the number of fluorophores
315 within the complex, as displayed in Equations 1 and 4. This relationship is independent of temporal
316 separation or demanding fluorophore properties, such as high recall rates, thereby simplifying the data
317 acquisition and post processing.
318
319 At the single molecule level, noise and genuine signals become hard to distinguish, especially for
320 photoactivatable fluorescent proteins with low brightness such as PAmCherry (Subach, Patterson, et al.,
321 2009). Since DCC considers the overlap of two colors, the effect of background signals on the indicator
322 is greatly reduced. The probability of noise signals falling within the colocalization threshold distance of
323 a given mVenus cluster follows a Poisson distribution. Our measured background signal density in the
324 PAmCherry channel of ~0.8 clusters/µm² inside cells (Figure 1E) suggests that only 2.6% of all mVenus
325 molecules erroneously colocalize with the background clusters from the red channel (Equation 6). Given
326 the colocalization ratios greater than 0.2 (for dimers and higher oligomeric forms; Figure 4), this noise
327 level results in a signal-to-noise ratio greater than 10. This estimate compares favorably with single-color
328 experiments: transfected cells showed a mean density of ~2.2 PAmCherry clusters/µm² consistent with a
329 signal-to-noise ratio less than 4 for this fluorescent protein alone. Higher background noise levels in the
330 marker detection channel would lead to the over-counting of marker clusters, resulting in a falsified
331 colocalization ratio and should be avoided. Our use of the bright mVenus tag as the marker allowed us to
332 use a narrow detection window, resulting in very low levels of background noise (Figure 1D and Table
333 3).

334

335 For data processing, filtering of noise signals is important for determining the recall rate accurately. The
336 higher tolerance to background noise enabled us to use only a simple filter based on an empirical
337 threshold of the minimal number of localizations for a single PAmCherry cluster (Figure 5). This was
338 based on our observation that signal clusters recorded from background regions showed mostly only
339 small clusters with few localizations (Figure 5). Fusion of two fluorescent proteins linked sequentially
340 did not interfere with their structure as indicated by our hrCNE results (Figure 4 – figure supplement 2).
341 This may not be surprising as similar results have been reported previously for similar constructs in other
342 contexts (Leeman et al., 2018; Liang et al., 2018; Murata et al., 2005; Ranade et al., 2014).
343

344 In conventional binomial analyses of SMLM data, a low recall rate ($p < 0.5$) is a significant limitation
345 when determining oligomeric states of proteins (Durisic et al., 2014). Theoretically, DCC-SMLM can
346 optimally distinguish oligomeric states from monomer to pentamer if p is between 0.2 and 0.4 (Figure
347 1C; Figure 1 – figure supplement 1), dependent on the oligomeric state of interest. This makes
348 PAmCherry and other fluorescent proteins with low recall rates suitable for determining oligomeric states
349 with DCC-SMLM (Figure 4). The original DCC model (Equations 1–3) suggested a recall rate of only
350 0.12 for PAmCherry, but did not result in an optimal fit for our experimental data from reference proteins
351 (Figure 4A). We reasoned that the function of the two fluorescent proteins may somehow become
352 impaired in the context of fusion proteins. Therefore, we introduced a modification factor, m, to describe
353 the probability of both fluorescent proteins becoming dysfunctional. Fitting this modified DCC model to
354 our reference data suggested a slightly higher recall rate of 0.17 for PAmCherry, and aligned well with
355 the experimental data (Figure 4A). Previously reported recall rates for PAmCherry range from 0.04 when
356 expressed in *E. coli* (Wang et al., 2014) to about 0.5 in *Xenopus* oocytes (Durisic et al., 2014). This
357 variability implies that the expression system may influence the functionality of fluorescent proteins.
358 Nevertheless, it may also result from differences in image acquisition and data processing. The software,
359 SNSMIL (Tang et al., 2015), which was used in our study to extract localizations, is no longer optimal,
360 and newer software such as SMAP (Ries, 2020) may be more sensitive in localization extraction and
361 yield different recall rates. In comparison with SNSMIL, however, SMAP did not significantly improve
362 the colocalization ratios because of its higher sensitivity to background signals (Figure 5 – figure
363 supplement 1). However, this may not be representative as SNSMIL was developed on our microscope

364 and may thus perform better than expected and we recommend to use more up-to-date software. It should
365 also be considered that many fluorescent proteins and dyes are sensitive to their specific
366 microenvironment and it may not be taken for granted that the same values of *p* and *m* can be used in
367 every circumstance. However, in our case, the same values gave reasonable results that matched the
368 oligomeric states as determined by high-resolution clear native electrophoresis (Figure 4 – figure
369 supplement 2).
370
371 Introduction of the modification factor *m* improved our original DCC model. Different from the factor *p*,
372 which describes the recall rate of PAmCherry alone, *m* additionally regards those behaviors that lead to
373 loss of fluorescence of both mVenus and PAmCherry simultaneously. This takes away the independence
374 of the two signals and invalidates the use of *p* alone. Based on our experience with fluorescent protein-
375 linked proteins, we assume that the most likely cause for such behavior is truncation within or proximal
376 to mVenus. We often observe such events when solubilizing proteins containing fluorescent proteins that
377 show up as smaller, additional bands on polyacrylamide gels such as seen in hrCNE (Figure 4 – figure
378 supplement 1). We currently suspect that this is largely dependent on the linker sequence used. Therefore
379 we recommend to use the same linker for calibration proteins and proteins of interest.
380
381 Another reason to carefully consider is the expression of endogenous proteins that can form oligomers
382 with the transfected proteins of interest. This would also reduce the apparent colocalization ratios due to
383 the inclusion of subunits without fluorescent proteins. We used a publicly available database of protein
384 expression (www.proteinatlas.org) to make sure that our calibration proteins and proteins of interest are
385 not expressed in the HEK293T cells that we used. The only protein with meaningful expression in this
386 cell line was ClC-2 which is why we also used the bovine ClC-K channel as another dimer control.
387 Colocalization ratios obtained for ClC-2, however, were similar to those of bClC-K (Figure 4A),
388 indicating that dimerization with endogenous channels was negligible.
389
390 If the interest is only in determining higher oligomeric states (larger than dimers), the factor *m* may be
391 omitted. However, as theoretically expected, the determination of a monomeric state becomes almost
392 impossible without *m* (Figure 4 C-H) when taking the uncertainty of all values into account. Alternatively,
393 the determination of the most likely oligomeric state can be performed without the knowledge of either

394 *p* or *m* using a comparison to the recorded colocalization ratios of reference proteins. We used the two-
395 sample Kolmogorov-Smirnov test with Bonferroni-adjusted significance thresholds (Figure 4I-N). This
396 has the potential to further simplify the determination of oligomeric states but is dependent on very
397 accurate recordings of reference protein values.
398
399 To achieve the necessary spatial resolution across two colors, we implemented and compared simple
400 methods of cell fixation of cells and post-imaging correction of common artifacts such as drift and
401 chromatic aberrations. While our implementation is based on SMLM, it could also be adapted for other
402 types of microscopy as stated above. In our experiments, there was no difference between a linear fit of
403 chromatic aberrations and local average correction (Figure 2 and Figure 2 – figure supplement 1);
404 however, this may depend on the quality of the optical setup and should be carefully evaluated for every
405 instrument.
406
407 It should be emphasized that DCC-SMLM does not consider the behavior of individual protein
408 complexes; instead, it is based on statistical information from a large number of protein complexes.
409 Therefore, if the protein complex of interest does not exist in a single oligomeric state, DCC-SMLM
410 counting can only report an average oligomeric state (simulated in Figure 4 – figure supplement 4).
411 Future efforts should aim to explore whether additional information obtained from DCC-SMLM
412 experiments or other experimental designs could be used to increase its usefulness in analyzing proteins
413 that adopt a mixture of oligomeric states.
414
415 To exclude an effect of expression density on the oligomeric state of our observed proteins, we used cells
416 which covered a wide range of expression densities. A plot of the colocalization ratio versus the density
417 of expressed clusters (Figure 4 – figure supplement 5) did not reveal any significant correlation between
418 the two values, effectively ruling out aggregation over the expression densities achieved in our samples.
419 It would also be expected that higher expression densities may be used to decrease the variability of the
420 results as the inclusion of more clusters per recordings would help the mean colocalization ratio per
421 recording to fall closer to the actual mean.
422

423 Using DCC-SMLM, we determined the oligomeric state of proteins from the SLC26 and SLC17 families.
424 Studies using Blue native electrophoresis (Detro-Dassen et al., 2008) and X-ray and cryo-EM (Geertsma
425 et al., 2015; Walter et al., 2019) previously reported a dimeric assembly for SLC26, whereas microscopy
426 approaches supported a tetrameric assembly (Hallworth & Nichols, 2011; Sinha et al., 2010). This
427 discrepancy might indicate that SLC26 forms stable dimers that survive protein purification and
428 solubilization as well as weakly associated "dimer-of-dimer" tetramers in native membranes that may
429 only be visible in intact cells. Although we cannot exclude the possibility that a minor fraction of prestin
430 forms tetrameric structures, our results show that both prestin and SLC26A3 predominantly form dimers
431 in intact HEK293T cells (Figure 4B-D).
432
433 We also studied the oligomeric state of vGlut1, 2, and 3 (Slc17a7, Slc17a6 and Slc17a8, respectively),
434 and sialin (SLC17A5) and found that they are likely to be monomeric. Again, these results are consistent
435 with native gel electrophoresis (Figure 4 – figure supplement 2). An earlier report suggested dimeric
436 organization for SLC17 transporters based on SDS-PAGE results (Andersen, 2004) but a recent cryo-EM
437 study reported a monomeric structure (F. Li et al., 2020). Functional expression studies have also
438 demonstrated that vesicular glutamate transporters and sialin can function without accessory subunits
439 (Eriksen et al., 2016; Morin et al., 2004), which, when considered alongside our results, confirms that
440 SLC17 proteins are functional monomers.
441
442 Our study provides a new approach to determine the oligomeric state of membrane proteins *in situ*. DCC-
443 SMLM overcomes many of the drawbacks and requirements of previous methods, making determination
444 of oligomeric states more reliable and much easier to perform. In fact, while we use the x- and y-
445 coordinates provided by SMLM, the DCC principle is not reliant on SMLM and the algorithm may also
446 be used with other microscopy modalities given enough spatial resolution to separate individual protein
447 complexes. Examples may include structured illumination (Kner et al., 2009) and super-resolution radial
448 fluctuations (Gustafsson et al., 2016) which are otherwise difficult to quantify at the single molecule
449 level because they rely on the mathematical processing of multiple raw images to yield super-resolution
450 images removing much of the information required by other algorithms.
451
452 **Materials and Methods**

453

Key Resources Table				
Reagent type (species) or resource	**Designation**	**Source or reference**	**Identifiers**	**Additional information**
cell line (H. sapiens)	tsa201 (HEK293T)	ECACC via Sigma Aldrich	Cat# 96121229-1VL	
cell line (H. sapiens)	Flp-In T-Rex 293	ThermoFisher	Cat# R78007	
transfected construct (H. sapiens)	SLC26A3	10.1074/jbc.M704924200	GenBank:NM_000111	
transfected construct (R. norvegicus)	Slc26a5 (Prestin)	10.1074/jbc.M704924200	GenBank:NM_030840	
transfected construct (H. sapiens)	ClC-2	10.1093/hmg/4.3.407	Swiss-Prot:P51788	With Y17H and R210H inserted to conform to 360 control chromosomes (Haug et al. 2003)
transfected construct (B. taurus)	bClC-K	This Paper		Cloned from bovine kidney tissue
transfected construct (H. sapiens)	EAAT2	10.1074/jbc.M110.187492		
transfected construct (M. musculus)	vGlut1	other		From R. Guzman, FZ Jülich. Carries a mutation of the di-leucin motif to alanine to facilitate plasma membrane insertion (see Eriksen et al. 2016)
transfected construct (M.	vGlut2	other		From R. Guzman, FZ Jülich. Carries a mutation

musculus)				of the di-leucin motif to alanine to facilitate plasma membrane insertion (see Eriksen et al. 2016)
transfected construct (M. musculus)	vGlut3	other		From R. Guzman, FZ Jülich. Carries a mutation of the di-leucin motif to alanine to facilitate plasma membrane insertion (see Eriksen et al. 2016)
transfected construct (M. musculus)	Sialin	other		From R. Guzman, FZ Jülich. Carries a mutation of the di-leucin motif to alanine to facilitate plasma membrane insertion (see Eriksen et al. 2016)
other	TetraSpeck fluorescent beads (100 nm)	ThermoFisher	Cat# T7279	Fluorescent microspheres used for the correction of drift and chromatic aberration, as well as for the general calibration of the microscope

454

455 **Cell Culture and DNA Constructs.** HEK293T cells and Flp-In T-REx 293 cells were grown in DH10
456 medium (DMEM supplemented with 50 U/mL penicillin-streptomycin and 10% FBS at 37 °C and 5%
457 CO_2 (Tan et al., 2017). For Flp-In T-REx 293 cells, the medium was additionally supplemented with 10
458 µg/mL blasticidin S HCl and 100 µg/mL hygromycin B. None of the cell lines used were authenticated.
459 However, no commonly misidentified lines were used. Cell lines were tested for mycoplasma
460 contamination twice per year. cDNAs encoding the fusion proteins were inserted into the cloning site of

461 the pcDNA5/FRT/TO or pcDNA3 vector plasmid via overlap extension PCR, restriction digestion and
462 DNA ligation.
463
464 **Cleaning of Coverslips.** Piranha solution was made by slowly adding 1 volume of hydrogen peroxide
465 solution (30%, Merck) into 3 volumes of concentrated sulfuric acid (>95%, Sigma-Aldrich) while stirring.
466 **Caution: Piranha solution reacts violently with organic matter and must be handled and disposed**
467 **of according to institutional guidelines!** Coverslips (VWR) were placed in a glass container and freshly
468 prepared piranha solution was added. Following incubation for 4-7 days at room temperature, coverslips
469 were rinsed with double distilled water at least 20 times. Afterwards, coverslips were stored in double
470 distilled water at 4 °C for experiments within the next 2 weeks.
471
472 **Preparation of Cells for SMLM Imaging.** For SMLM imaging, HEK293T cells were sparsely seeded
473 onto clean 25-mm coverslips in 1.5 mL DH10 in 3.5 cm petri dishes 2-3 h before transfection. Cells were
474 transfected with 0.5-1.0 µg plasmid DNA plus 1.0 µL Lipofectamine 2000 (Life Technologies) according
475 to the manufacturers' instructions and cultured overnight (12-16 h) before fixation. Flp-In T-REx 293
476 cells were seeded on coverslips in the same way and 0.01 µg/mL tetracycline was added to the culture
477 medium 14.5 h before fixation.
478
479 Afterwards, cells were washed with 2 ml 4 °C PBS for 3 times and afterwards fixed in 1 ml of -20 °C
480 cold methanol (Gradient grade for liquid chromatography, Merck) for 5 minutes at -20 °C. Methanol was
481 then removed and the cells were washed 5 times with 4 °C cold PBS.
482
483 For experiments using aldehyde fixation, 4 g PFA (Merck) was dissolved in 80 ml PBS heated to 60 °C.
484 Once completely dissolved, the solution was filled to 100 ml with PBS. 1 ml aliquots were stored at -
485 20 °C for long-term storage. Prior to fixation, aliquots were warmed to 37 °C and vortexed to dissolve
486 precipitates. 8 µl of a 25% GA (Electron microscopy grade, Agra Scientific) was added to the PFA aliquot
487 for a final concentration of 0.2% (w/v) GA. Cells were washed 3 times with 37 °C warm PBS and fixed
488 in 1 ml of the PFA/GA solution for 30 minutes at 37 °C. Afterwards, cells were washed 3 times with PBS,
489 followed by incubation in 0.5% (w/v in PBS) ice cold $NaBH_4$ (Sigma-Aldrich) solution for 1 h at 4 °C.
490 Fixed cells were stored in 2 ml PBS in the dark at 4 °C until imaging within 24 h.

491

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

495

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

503

504 The focal plane was adjusted using red color signals of the fluorescent beads on the coverslip surface. A
505 transmission image was taken to record positions of cells and beads. Videos were recorded at an exposure
506 of 85.59 ms per frame, during which the illumination lasted for 50 ms. Background signals were bleached
507 by illumination with the 561 nm laser at full power for 500-1000 frames before the recording started.
508 Color channels were recorded sequentially, starting with the green channel (for mVenus), followed by
509 the red channel (for PAmCherry). See Table 3 for details about the laser power and filter settings. The
510 intensity of the 405 nm laser used for photoactivation of PAmCherry was gradually increased during the
511 recording to keep signals sparsely distributed over time. The recording was stopped when no signals
512 appeared. After recording both color channels, the focus was re-adjusted using the fluorescent beads, and
513 the displacement of the sample stage was taken as the axial drift. Recordings with an axial drift > 300
514 nm were discarded.

515

516 **Extraction of SMLM signals.** Raw videos were processed by SNSMIL (Tang et al., 2015) with the
517 following parameters: optical magnification, 200; numerical aperture, 1.49; gain, 300; e/AD count, 11.9;
518 excess noise factor square, 2; pixel diameter, 16 µm; bias offset (AD counts), 100; quality threshold, 1;
519 PSF width tolerance, 3; fitting model, fixed SD 2D gaussian non-integral model. For recordings of the
520 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). The same parameter settings were used when SMAP was used for data extraction.

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 ($k = 1,2,3,\ldots,n$) was calculated across all m beads (with the index $i = 1...m$) relative to the first frame as:

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

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

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

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:

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

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 toward 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:

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

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

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

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:

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

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

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

579 LCA correction was then performed according to Equation 11.

580

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

587

588 **Pair correlation analysis and computation of colocalization ratios.** Mean coordinates of localization
589 clusters for each color channel were used to represent the positions of the fluorescent proteins. To
590 determine the cutoff value of the inter-color distance for colocalization, we performed pair correlation
591 analysis across both color channels (Sengupta et al., 2011; Sengupta & Lippincott-Schwartz, 2012). In
592 short, the densities of PAmCherry localization clusters around any given mVenus cluster was calculated
593 according to

594 $$g_{MF}(r) = \frac{a}{\pi N_M N_F} \sum_{i=1}^{N_M} \frac{dN_{Fi}(r)}{(2r+dr)dr} [14]$$

595 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
596 ROI, respectively; $dN_{Fi}(r)$ ($1 <= I <= N_M$) represents the number of red clusters that fall into the annulus
597 around the ith cluster with the inner radius r and the outer radius $r+dr$. To accelerate computing, distances
598 to all PAmCherry clusters within 400 nm from each mVenus cluster were calculated using the
599 spatial.distance.cdist function from the python package SciPy v.1.2.1 (Virtanen et al., 2020). The values
600 were grouped into bins of 20 nm width and the pair correlation calculated for each bin. The correlation
601 length from the pair correlation analysis was used as the maximum colocalization distance between
602 clusters (100 nm; Figure 2H). Colocalization was registered for an mVenus cluster if a single red cluster
603 was found in a distance shorter than this maximum colocalization distance. The numbers of clusters and
604 colocalizations were counted in cellular and background regions and applied to Equations 5 and 6 to
605 calculate colocalization ratios.

606

607 In addition, to prevent miscounting due to protein aggregation, those clusters with a diameter larger than
608 500 nm were excluded from counting. The radius of a cluster was determined as $\mu + 2\sigma$ of the Euclidean
609 distances of all localizations to the center (mean coordinates) of the cluster, as shown Figure 2B.
610
611 **Estimation of the Oligomeric State of the Protein of Interest.** To estimate the oligomeric state of
612 proteins of interest and the observed colocalization ratios, we defined the coefficient of mismatch, $M^2(n)$
613 as
614
615 $$M^2(n) \equiv 1 - \frac{\sum_i (x_i - \overline{x})^2}{\sum_i [x_i - E(n)]^2} [15]$$
616 in which x is the mean of all observations (x_i) and $E(n)$ is the expected colocalization ratio for the
617 oligomeric state n. The expected value was calculated from Equation 4 with the optimal values of m and
618 p, or from Equation 1 with the optimal p alone, as determined by the respective fitting to the calibration
619 data (Figure 4A). For the bootstrapping analysis to determine 95% confidence intervals of the mean, M^2
620 was calculated from resamples of the colocalization ratios measured from each protein of interest and
621 the reference proteins. The Python module scipy.stats.ks_2samp was used to perform the two-sample
622 Kolmogorov-Smirnov test.
623
624 **Fitting of the DCC Model to the Experimental Data.** m and p were determined as the fit parameters
625 of Equation 4 to the experimental data. The value pair (m, p), with m ranging from 0 to 1 and p from 0
626 to 0.4, both with a 0.01 increment, was used to calculate the theoretical colocalization ratios for
627 oligomeric states $n = 1, 2, 3$ and 4. The calculated ratios were then compared with the experimental values
628 obtained from all the proteins with known oligomeric states to calculate the coefficient of determination
629 (R^2). In the case when m was not included, the fitting method was similar, but with m left out.
630
631 **High-resolution clear native electrophoresis.** HEK293T cells were transfected with 3 – 5 µg of plasmid
632 DNA using the calcium phosphate method and treated as described previously (Guzman et al., 2017). In
633 brief, cells were washed with ice cold PBS and incubated for 15 min with a lysis buffer containing 0.1
634 M sodium phosphate, either 0.5% (w/v) digitonin or 20 mM DDM, protease inhibitors and 20 mM
635 iodoacetamide. The buffer was then transferred into a 1.5 ml Eppendorf tube. After centrifugation at 4 °C,

636 an aliquot of the supernatant containing approximately 10 µg of protein was loaded into a native gel. The
637 4-14% acrylamide gradient gels were prepared as described(Wittig et al., 2007). The anode buffer
638 contained 25 mM imidazole/HCl, pH 7.0 and the cathode buffer contained 50 mM tricine, 7.5 mM
639 imidazole/HCl, pH 7.0 supplemented with the anionic detergent DOC (0.05%) and the non-ionic
640 detergent Triton X-100 (0.05%) (Wittig et al., 2007). The gels were run in a cold room (4 °C) and the
641 starting voltage was set to 100 V. After 1 h, the voltage was increased to 150 V for another 2 h. Gels were
642 scanned on a fluorescence gel scanner (Typhoon FLA 9500, GE Healthcare, Freiburg, Germany) at 100
643 µm resolution. mVenus was excited with the built-in 473 nm laser and the emission was filtered with a
644 530/20 bandpass filter. Gel images were visualized using Fiji (Schindelin et al., 2012).
645
646 **Data availability.** The datasets including the extracted localization are available for download via
647 https://doi.org/10.5281/zenodo.6012450. Raw video files can only be provided upon request due to their
648 large file sizes. We implemented a version of DCC-SMLM as a python library on GitHub
649 (https://www.github.com/GabStoelting/DCC-SMLM).
650
651 **Acknowledgements.** This research was supported by the Deutsche Forschungsgemeinschaft (FA 301/15-
652 1 to Ch.F.) as part of Research Unit FOR 5046, project P4. We are grateful to Dr. Iman Abdollahzadeh,
653 Dr. Johnny Hendriks and Dr. Thomas Gensch for their generous support with the imaging setup. Many
654 thanks also go to Cora Hannack, Dr. Raul E. Guzman and Arne Franzen for providing plasmids encoding
655 vGlut and sialin. We also thank Ashley Craig of Alchemy Editorial Services Ltd for language editing.
656
657
658

References

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

Guzman, R. E., Sierra-Marquez, J., Bungert-Plümke, S., Franzen, A., & Fahlke, C. (2022). Functional Characterization of CLCN4 Variants Associated With X-Linked Intellectual Disability and

Epilepsy. *Frontiers in Molecular Neuroscience*, *15*. https://www.frontiersin.org/article/10.3389/fnmol.2022.872407

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

Ries, J. (2020). SMAP: A modular super-resolution microscopy analysis platform for SMLM data. *Nature Methods*, *17*(9), 870–872. https://doi.org/10.1038/s41592-020-0938-1

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., Bungert-Plümke, S., Franzen, A., & Fahlke, C. (2015). Carboxyl-terminal Truncations of ClC-Kb Abolish Channel Activation by Barttin Via Modified Common Gating and Trafficking. *Journal of Biological Chemistry*, *290*(51), 30406–30416. https://doi.org/10/gbf73v

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*, 269. https://doi.org/10.3389/fphys.2017.00269

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

659 **Table 1.** Colocalization ratio (*P*) of mVenus clusters with PAmCherry clusters of fusion proteins fixed
660 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 (K_{ir}2.1)	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

661

662 **Table 2.** The linkers used in the fusion of proteins of interest (POI) to the fluorescent protein. In this
663 project two fusion patterns were tested: POI-Linker #1-PAmCherry-Linker #2-mVenus (r-g) and POI-
664 Linker #3-mVenus-Linker #4-PAmCherry (g-r) (both written from the N- to C-terminus) with the
665 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

666

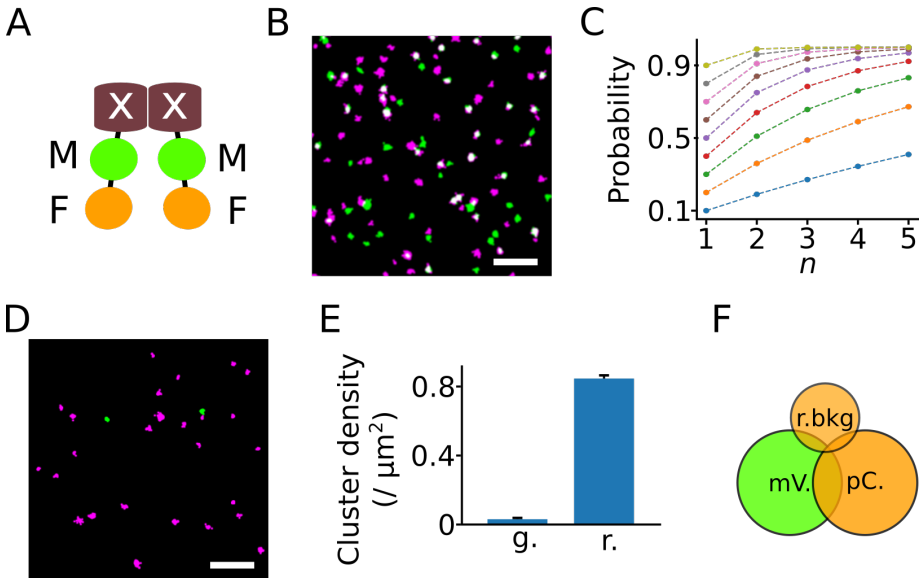
667 **Table 3.** The settings used for SMLM imaging. The dichroic mirror Di01-R442/514/561-25x36 was used
668 for imaging of both fluorescent proteins. All filters and dichroic mirrors were acquired from Semrock.
669 The laser power was measured at the sample plane and before the rear entrance of the objective. The
670 power of the 405 nm activation laser for PAmCherry was varied between 4.4 µW and 10.2 mW near the
671 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

672

673 **Figures**

674



675

676 **Figure 1.** The DCC strategy to determine the oligomeric state of proteins. (*A*) Model of a dimeric fusion
677 protein used for DCC-SMLM. The protein of interest is marked with a white "X", whereas the marker
678 and indicator fluorescent proteins are labeled "M" and "F", respectively. (*B*) A reconstructed SMLM
679 image from an HEK293T cell expressing $K_{ir}2.1$ labeled with both mVenus (green) and PAmCherry
680 (magenta). White indicates colocalization of the two color signals. Scale bar: 1 µm. (*C*) The probability
681 of detecting a protein complex changes with the number (n) and recall rate (p) of the fluorescent proteins
682 it contains. Each color indicates a different recall rate from 0.1 to 0.9 in steps of 0.1. (*D*) A representative
683 image reconstructed from a recording of a bare coverslip surface showing background signals in the
684 PAmCherry (magenta) and mVenus channels (green). Scale bar: 1 µm. (*E*) The mean background cluster
685 densities in the mVenus (green, g.) and the PAmCherry (red, r.) recording channels. A cluster comprises
686 a series of fluorescent signals that are spatially close and clustered with the DBSCAN algorithm. N =
687 283 recordings. (*F*) As the background signals in the green channel are negligible, only three types of
688 signals are considered: signals from mVenus (mV.), signals from PAmCherry (pC.) and background
689 signals in the red channel (r.bkg).

690

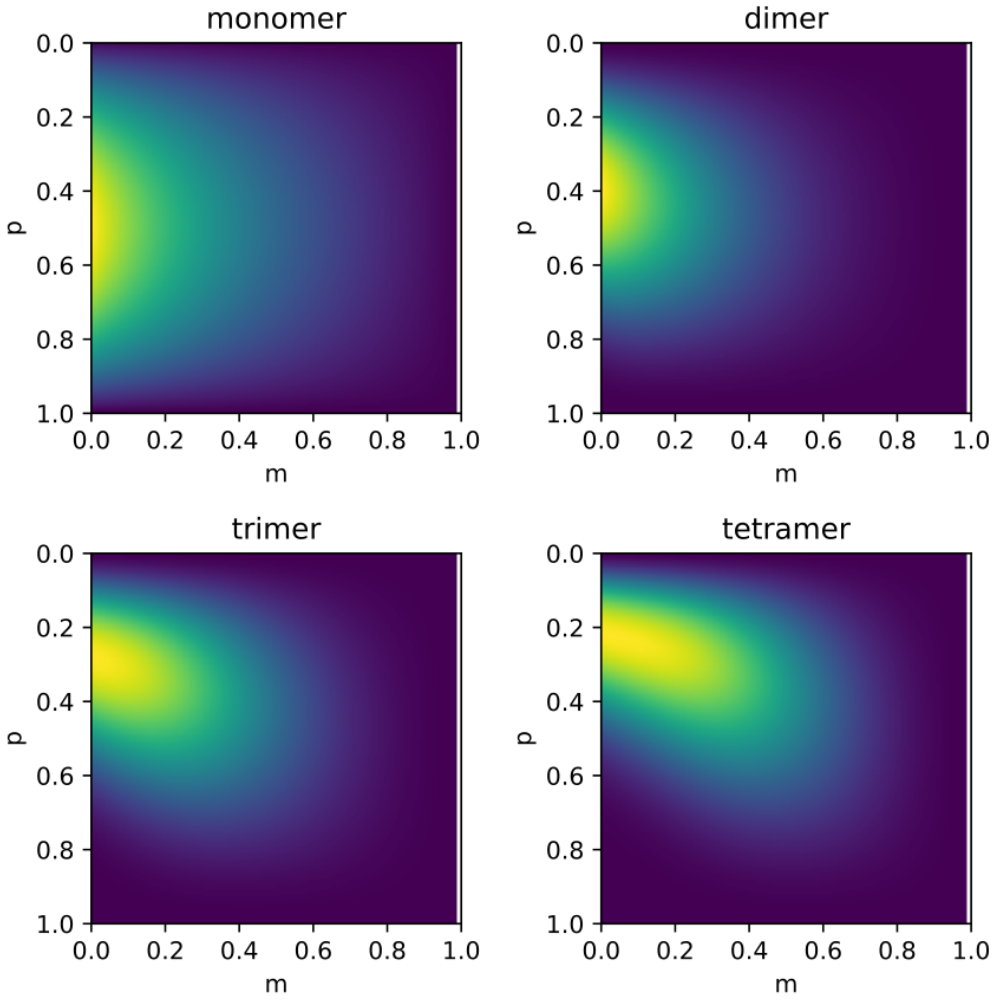


Figure 1 – figure supplement 1. Sensitivity analysis of Equation 1 and Equation 4 for different oligomeric states as indicated in the titles. The results of Equation 4 were calculated using the indicated values of p and m for the indicated as well as the next smaller and larger oligomeric state. Differentiation was then calculated either as the difference from monomer to dimers (for the detection of monomers) or by the product of the differences to the next smaller and larger oligomeric state. Values were normalized by the maximum difference obtained for each panel and plotted using a color code from yellow (maximum) to dark blue (worst). The sensitivity of Equation 1 can be seen where m = 0.

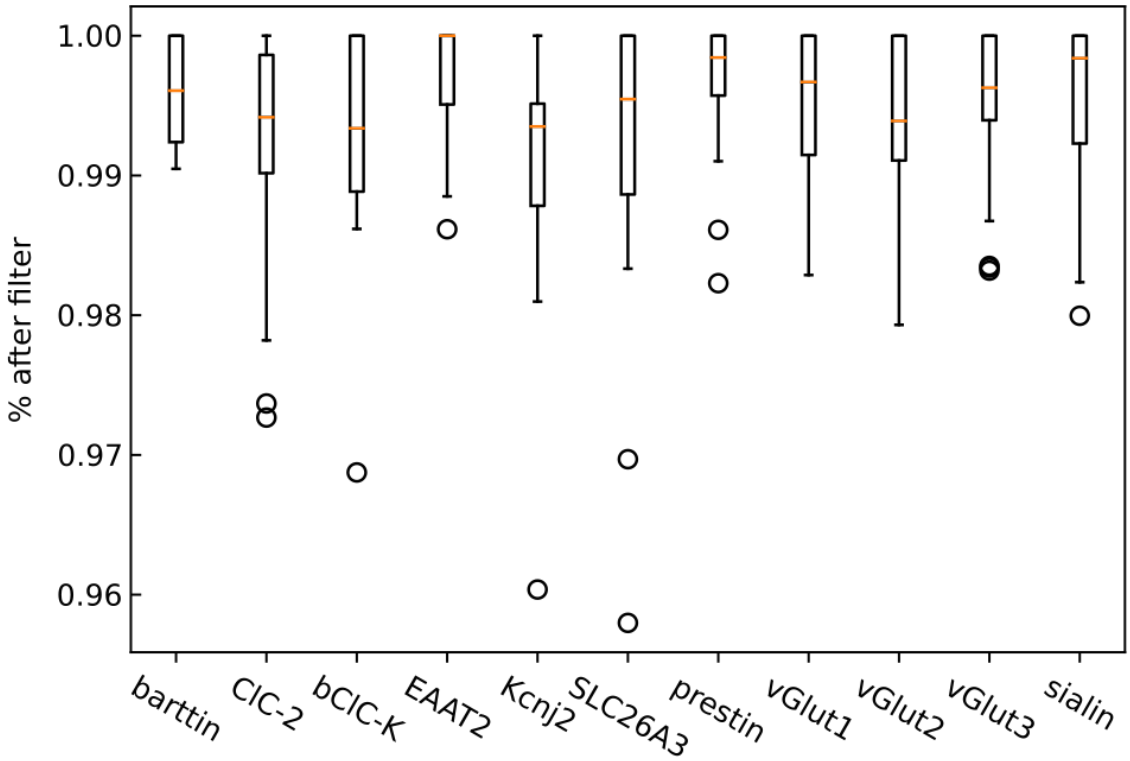


Figure 1 – figure supplement 2. Fraction of clusters remaining after the removal of large clusters (% after filter) by our analysis pipeline for each protein of interest in our study.

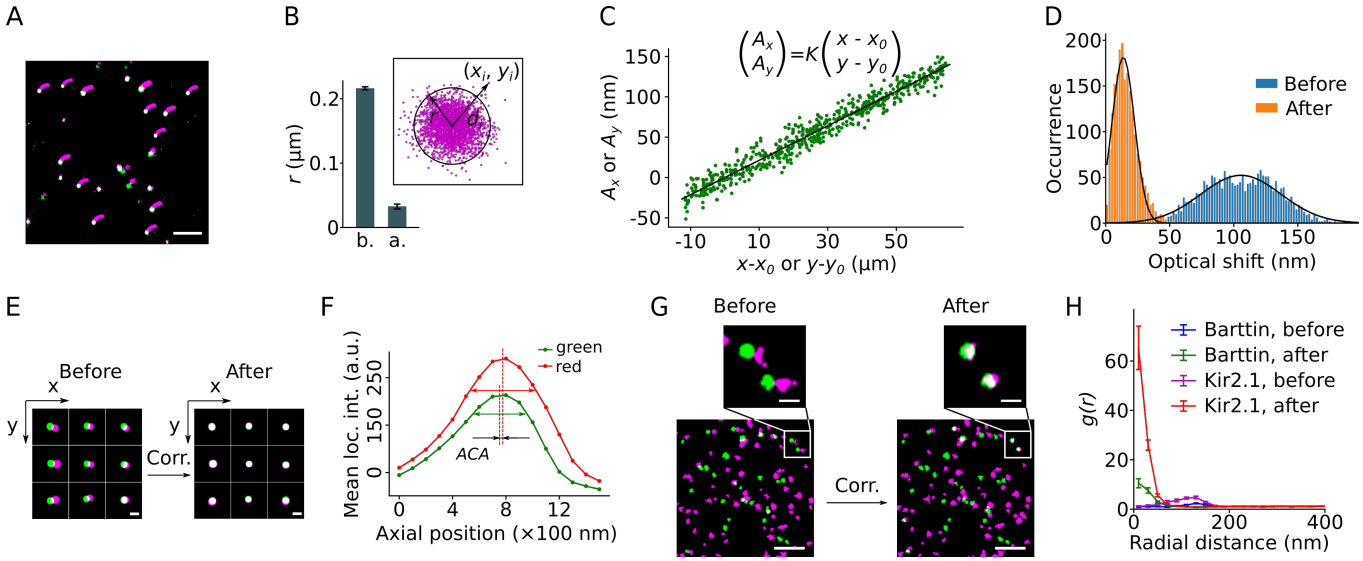


Figure 2. Correction of sample drift and chromatic aberration. (*A*) An SMLM image of a bead sample reconstructed from 13,000 frames recorded over ∼18.5 minutes. The small magenta smears indicate the trajectory of the sample drift. The corrected positions are displayed in green or, when overlapping, 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 localizations (with coordinates (x_i, y_i)) to the center (mean coordinates) 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 from the red coordinates to an arbitrary 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 a representative from a single recording of a bead sample. (*D*) Lateral chromatic aberrations from nine recordings of a bead sample with a total of 1980 beads were corrected with the values obtained in (*C*), with blue and orange indicating before and after the correction, respectively. Black lines show the 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 changed 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 localizations were omitted from the images. Scale

729 bars indicate 1 μm and 200 nm respectively in the bottom and the zoomed-in images. (*H*) Radial
730 distribution of barttin (17 cells, 1 preparation) and $K_{ir}2.1$ (35 cells, 3 preparations) PAmCherry clusters
731 around mVenus clusters before and after drift & chromatic aberration corrections. Non-clustered
732 localizations were excluded from the analysis.
733

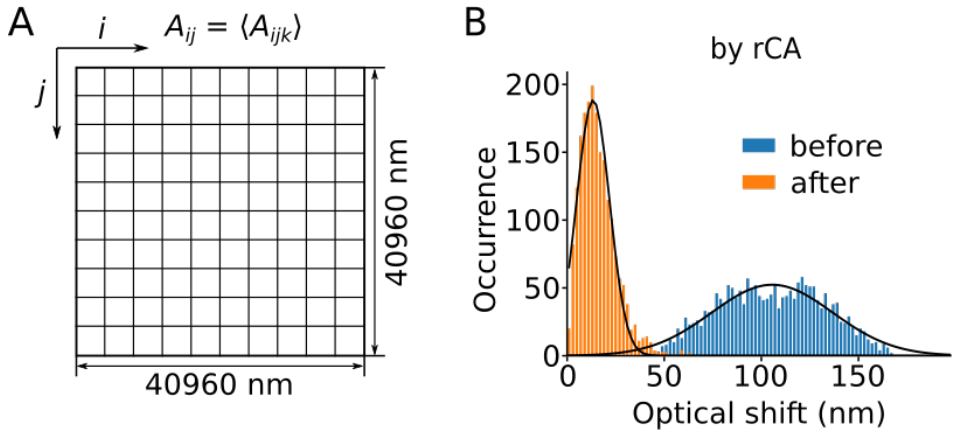
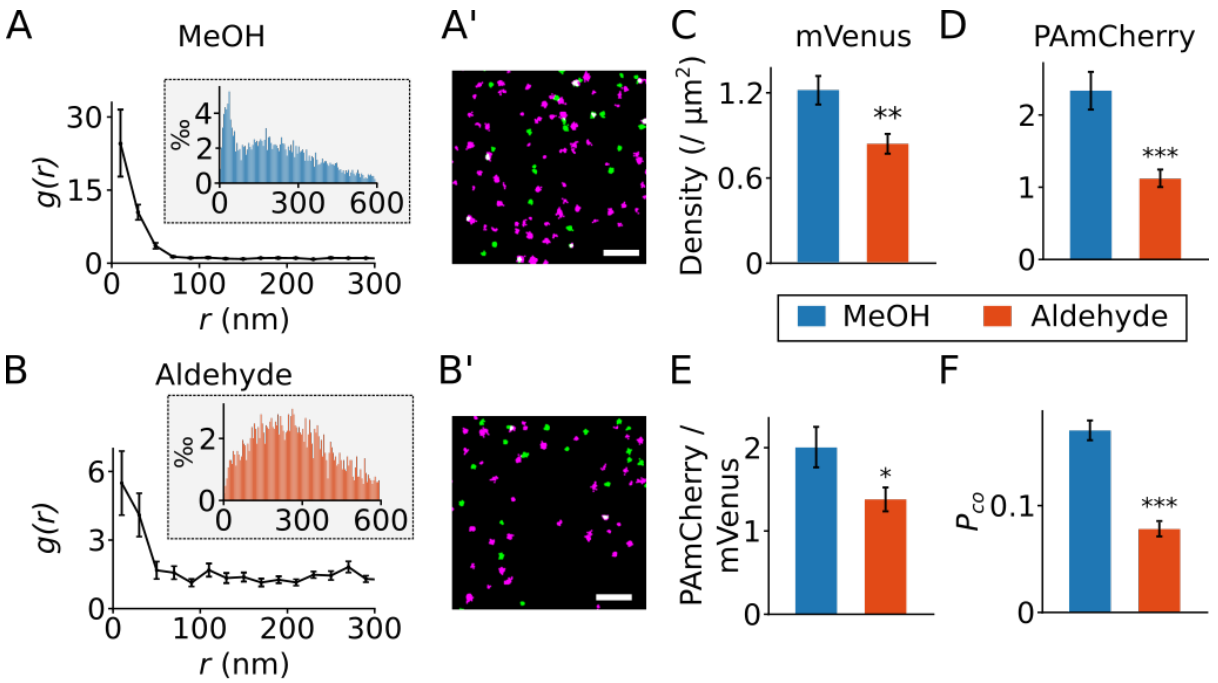


Figure 2 – figure supplement 1. Regional averaging for the correction of lateral chromatic abberation. *(A)* The whole microscopic field of view was divided into 100 squares and the mean lateral chromatic aberration (A_{ijk}) measured in each square was taken as the offset (A_{ij}) for correction. *(B)* 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.

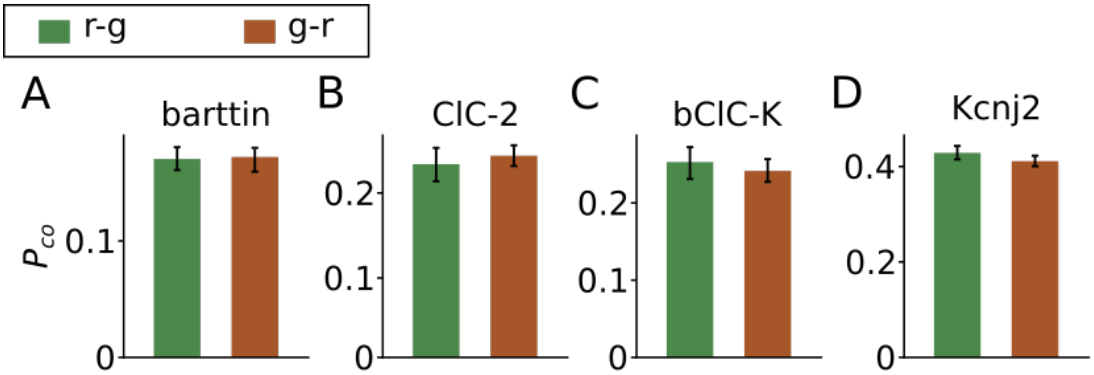


741

742 **Figure 3.** Effect of different fixation methods. (*A* and *B*) Results from the radial distribution function *g(r)*
743 for PAmCherry clusters around mVenus clusters recorded in Flp-In T-REx 293 cells expressing barttin-
744 mVenus-PAmCherry and fixed with cold methanol (*A*, MeOH, 24 cells) or para-formaldehyde plus
745 glutaraldehyde (*B*, Aldehyde, 29 cells). The inset shows the histograms of the nearest neighbor distances.
746 (*A'* and *B'*) Reconstructed images from recordings of cells expressing barttin-mVenus-PAmCherry fixed
747 with cold methanol (*A'*) or the aldehyde (*B'*), with green indicating mVenus clusters and magenta for
748 PAmCherry clusters. Scale bar: 1 µm. (*C* and *D*) The densities of mVenus (*C*, $p = 0.0028$) and PAmCherry
749 (*D*, $p = 4.9\times10^{-5}$) clusters in samples prepared using the two fixation methods. (*E*) The relative density
750 of PAmCherry (pC.) clusters to mVenus (mV.) clusters, $p = 0.02$. (*F*) The measured colocalization ratio
751 of mVenus with PAmCherry, $p = 1.6\times10^{-10}$. *: $0.05 > p > 0.01$, **: $0.01 > p > 0.001$, ***: $p < 0.001$,
752 Student t-test.

753

754



755

756 **Figure 3 – figure supplement 1.** *(A - E)* The colocalization ratios of mVenus with PAmCherry fused to
757 the C-terminus of five different proteins transiently transfected in HEK293T cells fixed with cold
758 methanol. The two fusion patterns POI-PAmCherry-mVenus (r-g) and POI-mVenus-PAmCherry (g-r) are
759 compared for each protein.

760

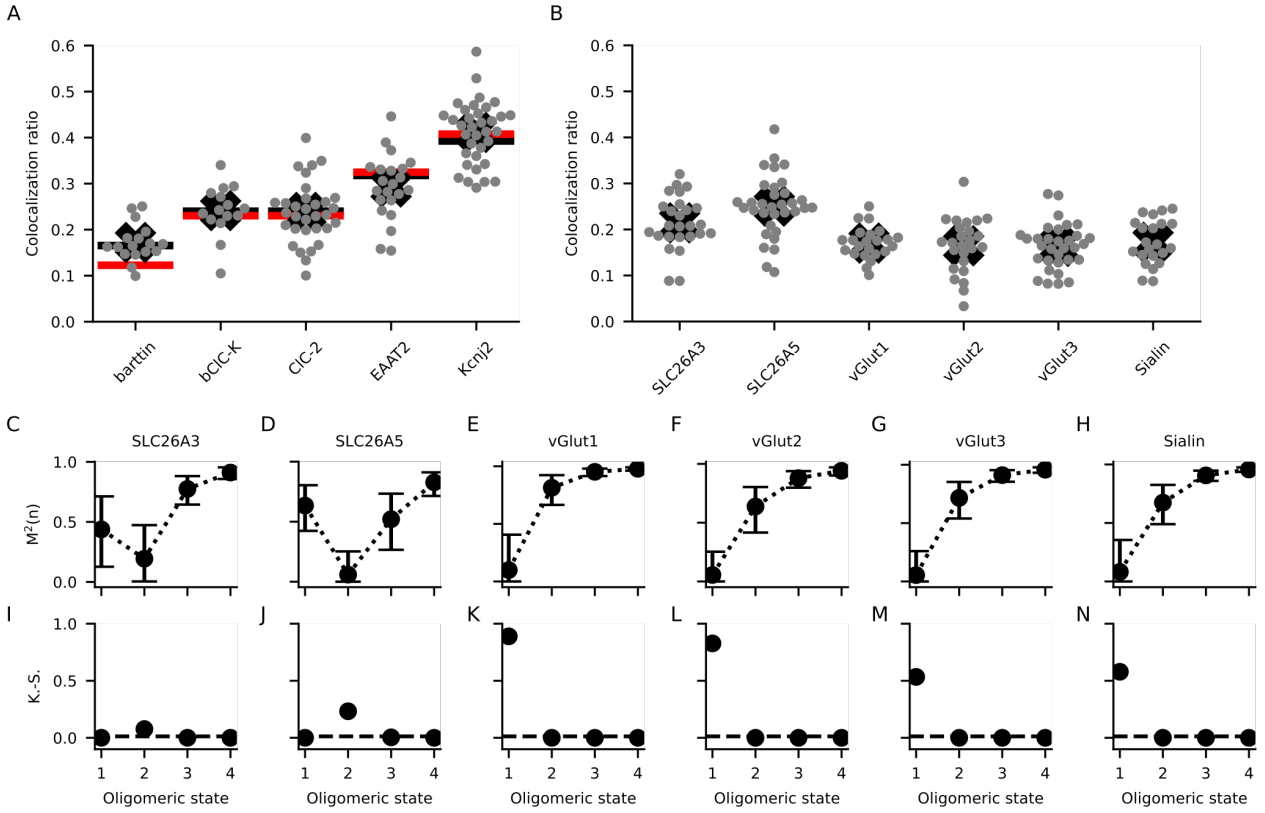


Figure 4. Calibration and determination of oligomeric states through analyzing the colocalization ratios with DCC-SMLM. (*A*) The colocalization ratios for mVenus and PAmCherry linked to proteins with known oligomeric states for calibration. Red lines show the fit with Equation 1 and black lines show the fit with Equation 4 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). Individual values are shown as gray circles and the mean value is shown as a black cross (partially hidden behind data points) (*B*) Colocalization ratios for mVenus and PAmCherry linked to the indicated proteins of interest. 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 oligomeric states of monomer (1), dimer (2), trimer (3) and tetramer (4). Error bars correspond to the range of the 95% confidence intervals as determined by a global bootstrap resampling (10,000 resamples). (*I–N*): Two-sample Kolmogorov-Smirnov (K.-S.) tests of the observed colocalization ratios from each protein of interest compared with

775 the protein standards in (*A*). The dashed horizontal line indicates the Bonferroni-corrected threshold for
776 the α-level (0.0125).
777

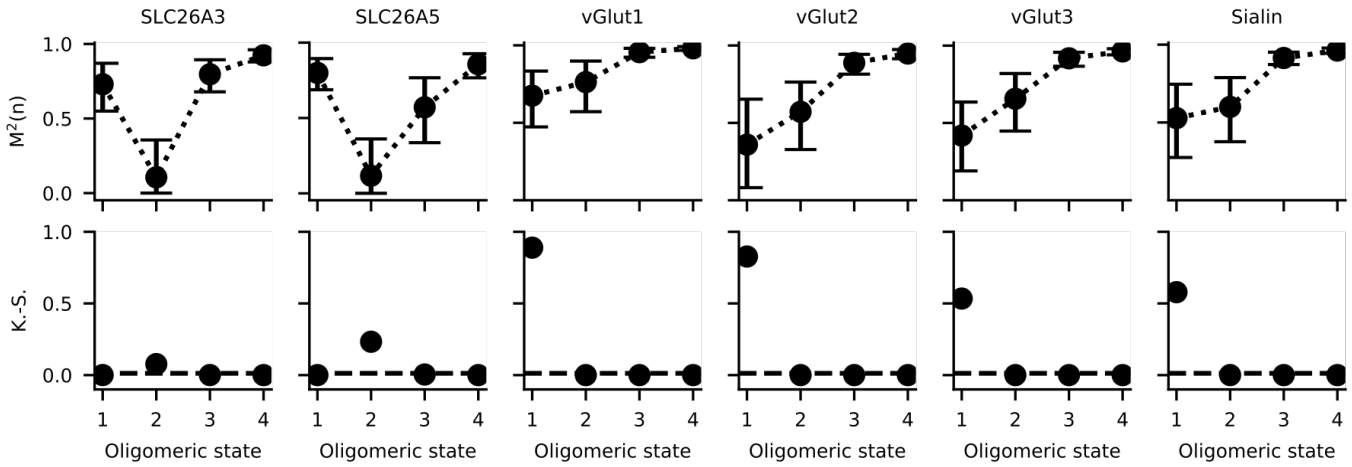


Figure 4 – figure supplement 1. Determination of the most likely oligomeric state of SLC26 and SLC17 proteins using only Equation 1. As expected, the determination of the monomeric state of SLC17 (vGlut and Sialin) proteins is not possible using the COM based only on Equation 1 (top row). Determination of dimers as seen for the SLC26 proteins is much less affected (top row). Error bars correspond to the range of the 95% confidence intervals as determined by a global bootstrap resampling (10,000 resamples). The comparison of the observed colocalization ratios with those of reference proteins using the Kolmogorov-Smirnov test is unaffected (bottom row).

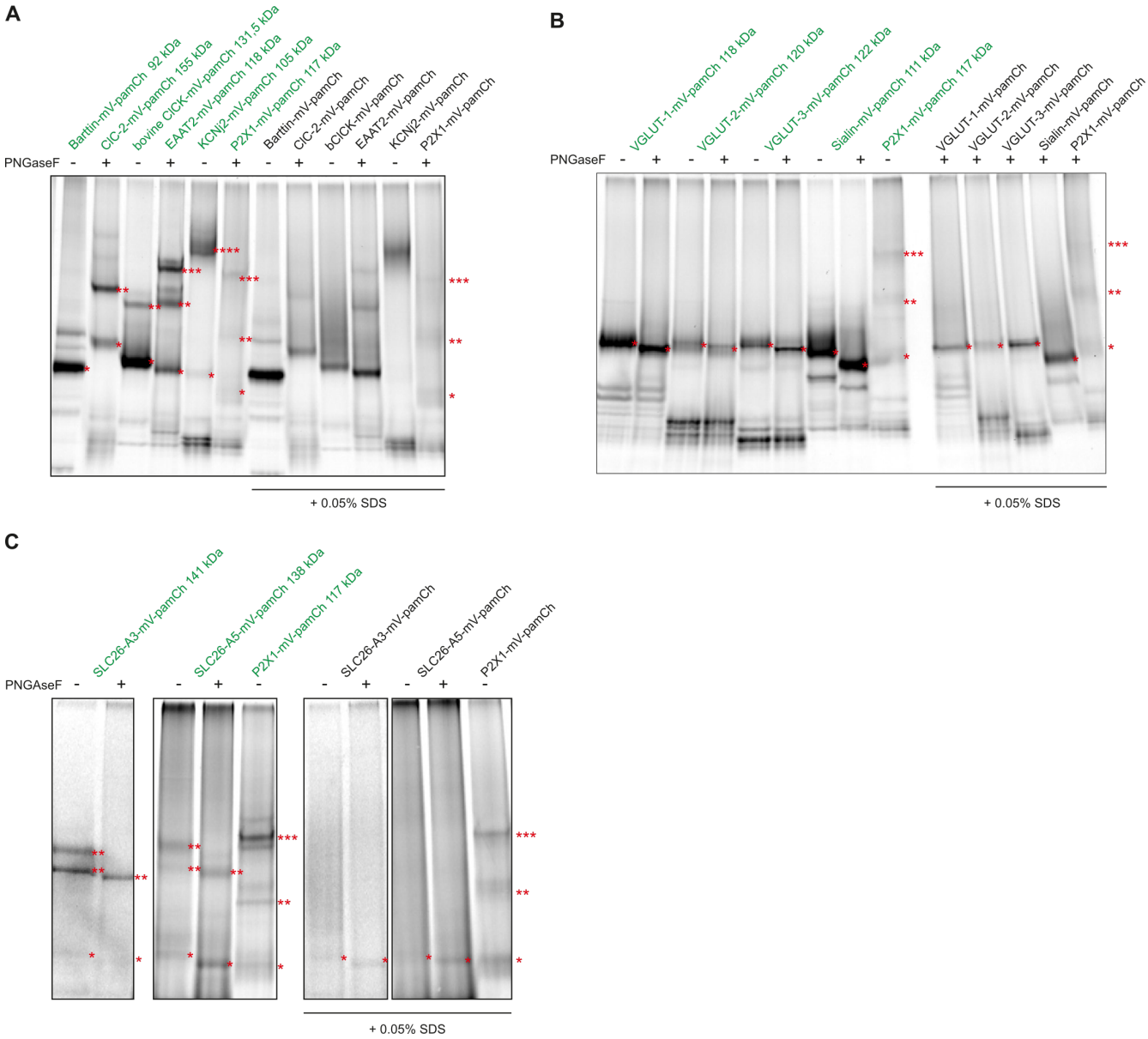


Figure 4 – figure supplement 2. High-resolution clear native gel electrophoresis (hrCNE) of reference proteins and proteins of interest. *(A)* hrCNE of standard proteins with known oligomeric state. Depending on their oligomeric state, protein bands are indicated with an increasing number of red stars (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.

796 Proteins were also treated with a final concentration of 0.05% SDS for 15 min at 37 °C to dissociate
797 oligomers (right half of the gel). The calculated molecular masses of protein monomers are shown in
798 green.
799

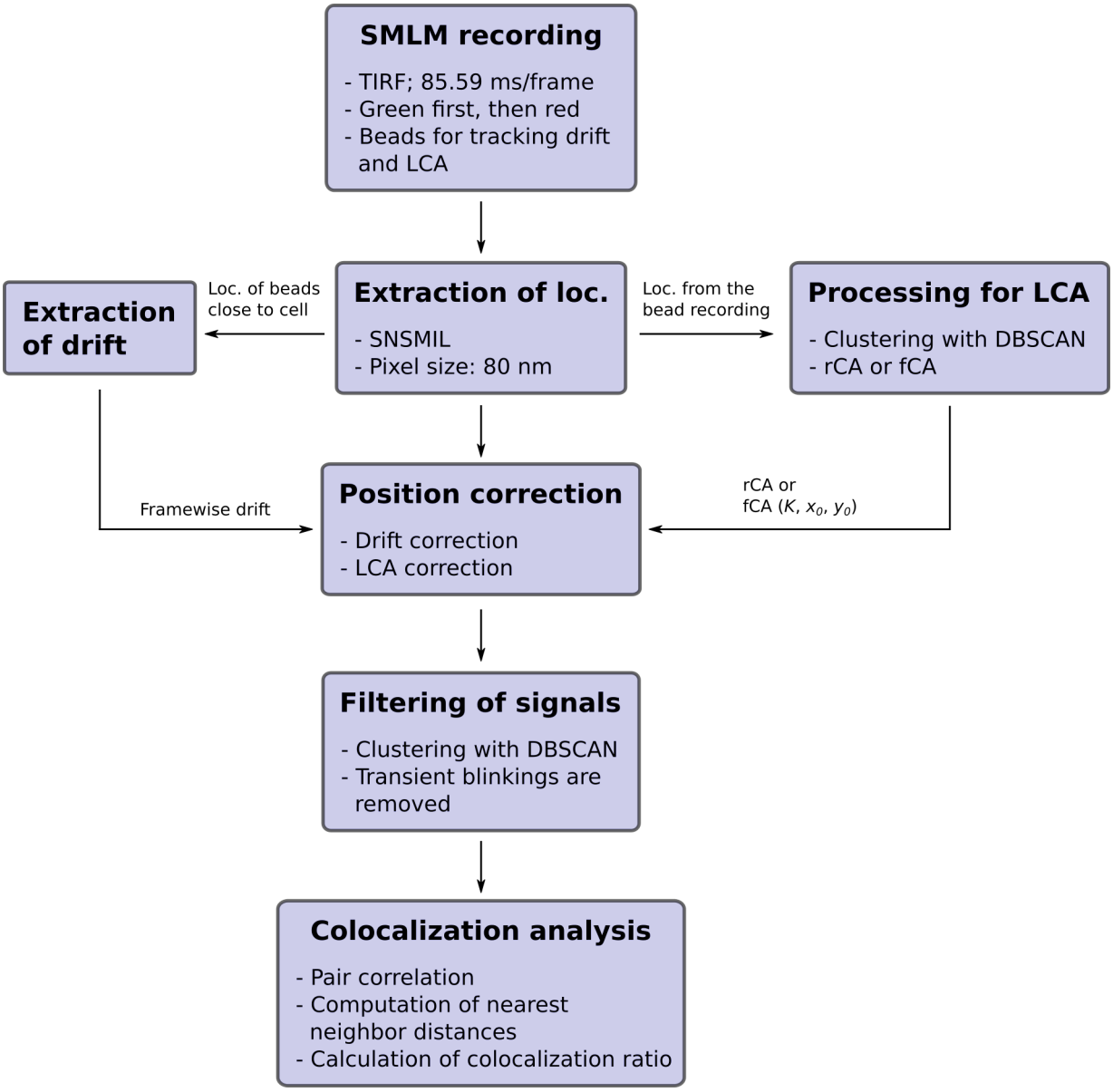


Figure 4 – figure supplement 3. The working paradigm of DCC-SMLM.

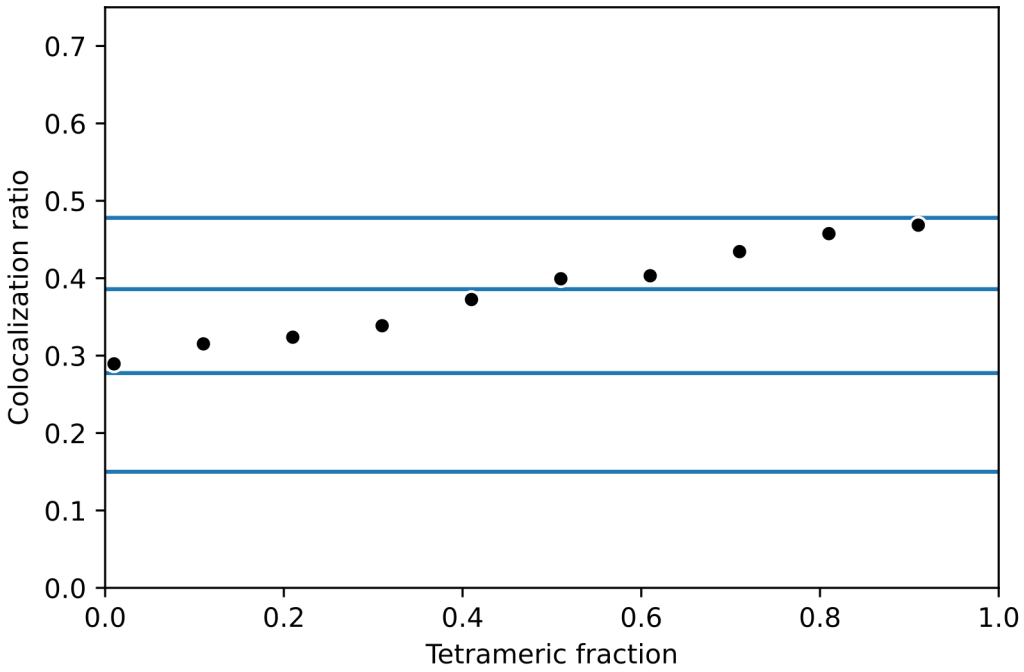


Figure 4 – figure supplement 4. Summary of 1000 simulations per data point of a mixed population of dimeric and tetrameric proteins. Proteins were simulated with p = 0.15, m = 1 at a density of 5 clusters per µm². Protein coordinates were set randomly and DBSCAN clustering was applied as stated in the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

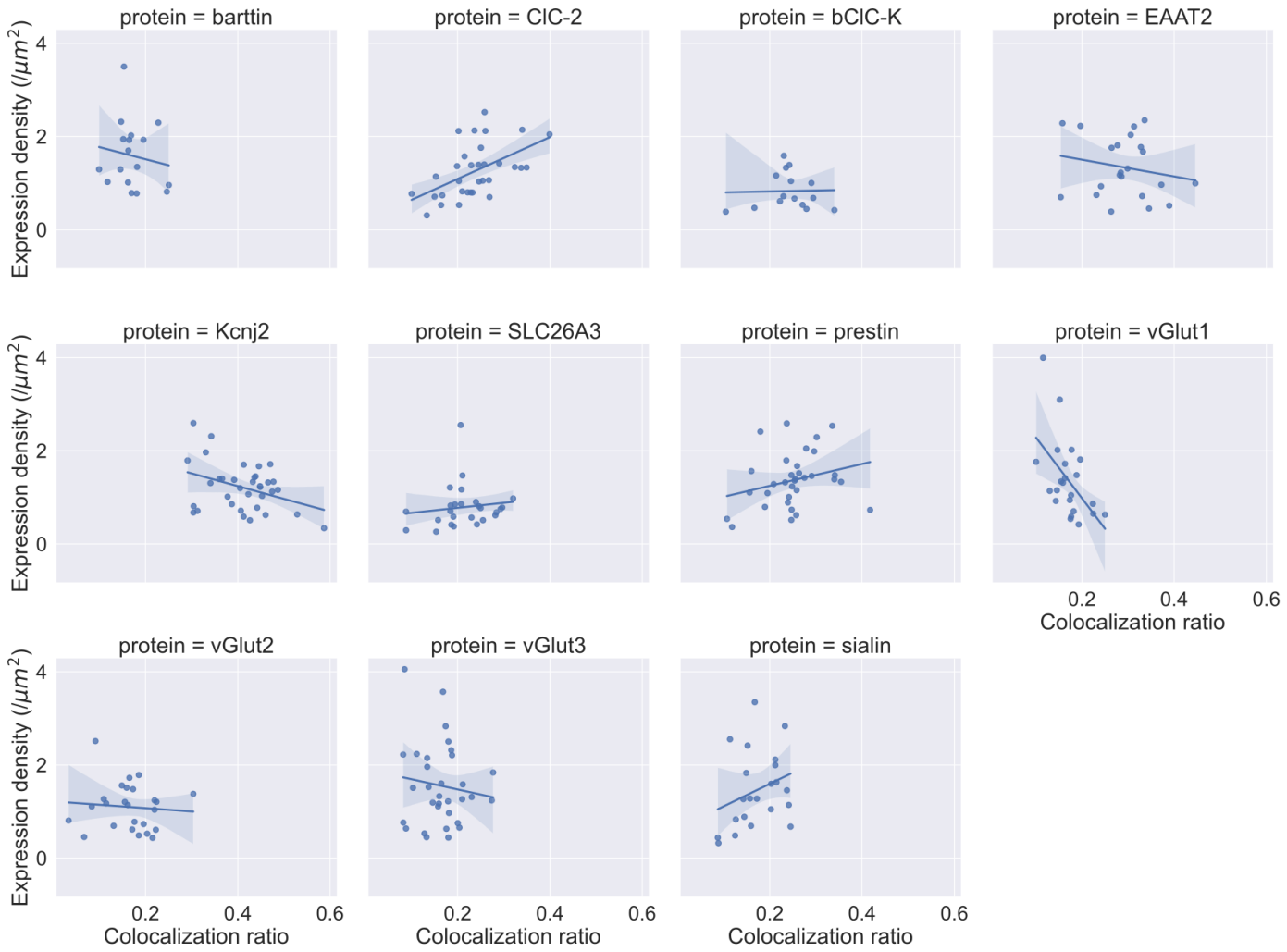


Figure 4 – figure supplement 5. Plots of the experimentally observed expression densities versus the colocalization ratio do not reveal a consistent correlation between these parameters. Blue lines represent a fit of a linear function to the data and the shaded areas represent the 95% CI of this fit as determined by 10,000 bootstrap resamples each.

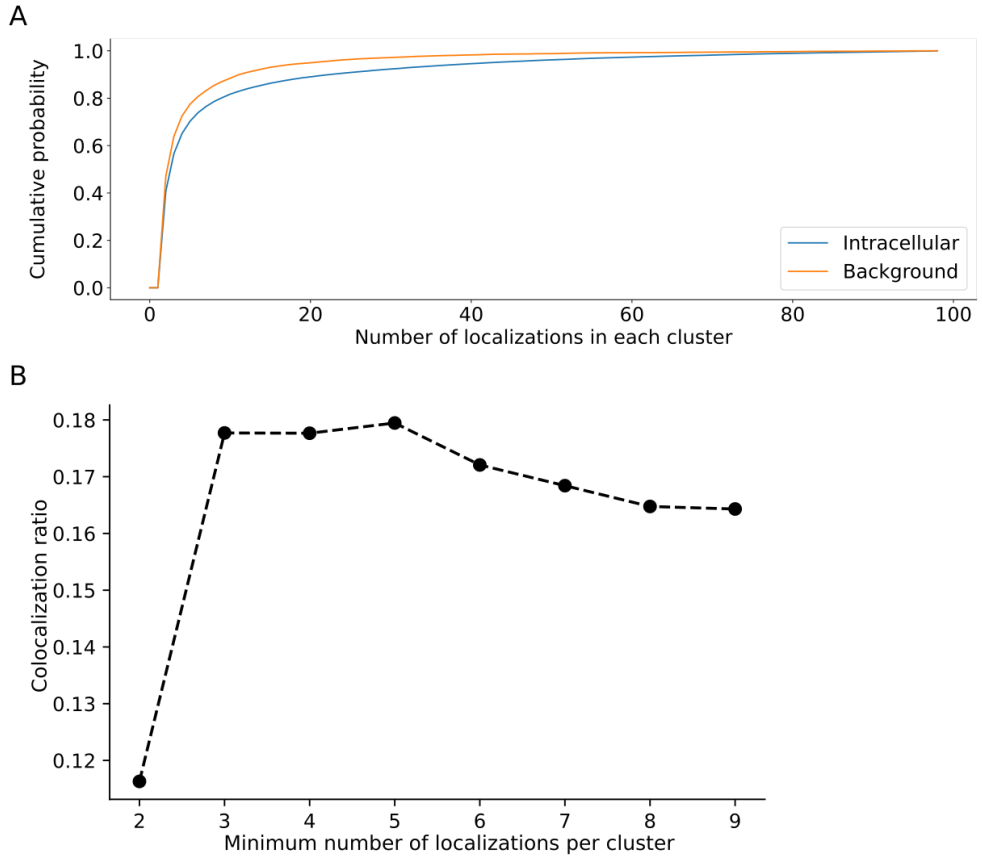


Figure 5. *(A)* The cumulative probability distribution of the number of localizations per cluster in the background and within Flp-In T-Rex 293 cells expressing Barttin-mVenus-PAmCherry. While most clusters in the background only exhibit a small number of localizations, clusters within cells exhibit a higher proportion of larger clusters. *(B)* The dependence of the colocalization ratio *P* versus the lower cutoff to be considered a genuine cluster. A cutoff of 2 results in low ratios due to the inclusion of a large number of background red clusters, which interfere with the computation algorithm by excluding a large amount of colocalizations of the genuine 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).

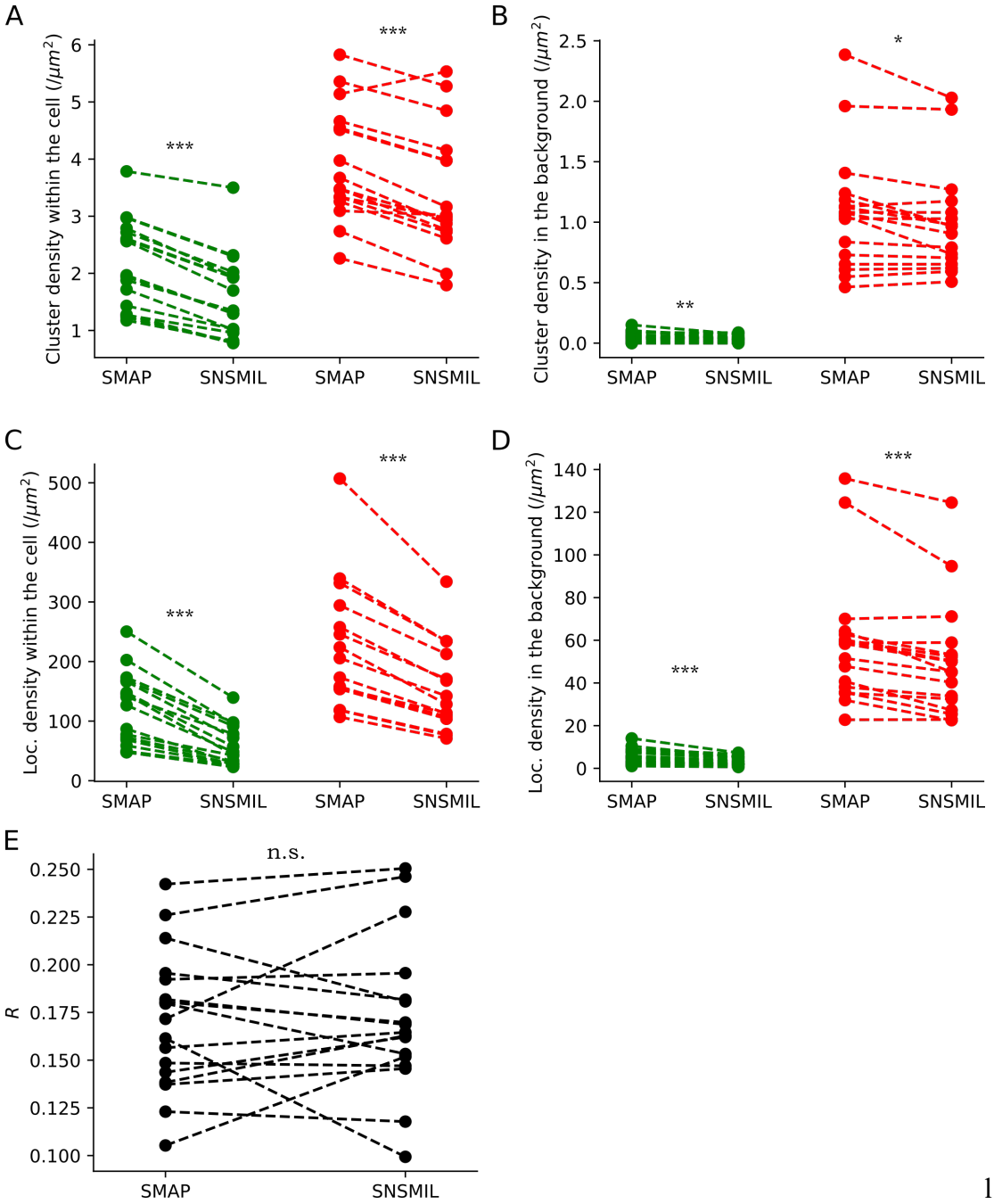


Figure 5 – figure supplement 1. Comparison of localization extraction using two different softwares: SMAP and SNSMIL. SMLM was recorded for mVenus and PAmCherry in HEK293T cells expressing barttin-mVenus-PAmCherry. $N = 17$ recordings. Dashed lines connect the same recordings. (A, B) The cluster densities collected within the cell (A) and in the background (B), with green indicating the mVenus channel and red the PAmCherry channel. (C, D) The localization densities collected within the

835 cell (C) and in the background (D). (E) Comparison of colocalization ratios (*P*) of mVenus and
836 PAmCherry clusters collected within the cell ($p = 0.82$). Paired Student t-Tests were used to compare
837 SMAP and SNSMIL extractions, with *** indicating $p < 0.001$, ** $0.001 <= p < 0.01$, * $0.01 <= p < 0.05$,
838 n.s. $p >= 0.05$.