Systematic spatial mapping of proteins at exocytic and endocytic structures

Ben T. Larson, Kem A. Sochacki, Jonathan M. Kindem, and Justin W. Taraska
Laboratory of Molecular Biophysics, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD 20892

ABSTRACT **Vesicular secretion (exocytosis) involves the release and then compensatory recycling of vesicle components through endocytosis. This fundamental cellular process is controlled by the coordinated assembly and interactions of dozens of proteins at the plasma membrane. Understanding the molecular composition of individual exocytic and endocytic structures and their organization across the plasma membrane is critical to understanding the behavior and regulation of these two cellular processes. Here we develop a high-resolution and high-throughput fluorescence imaging–based approach for the unbiased mapping of 78 proteins at single exocytic vesicles and endocytic structures in neuroendocrine PC12 cells. This analysis uses two-color single-frame images to provide a systems-level map of the steady-state distributions of proteins at individual exocytic and endocytic structures in the cell. Along with this quantitative map, we find that both calcium-regulated exocytic vesicles (dense core vesicles) and endocytic structures (clathrin-coated structures) and the proteins associated with these structures exhibit a random spatial distribution in unstimulated neuroendocrine PC12 cells. This approach is broadly applicable for quantitatively mapping the molecular composition and spatial organization of discrete cellular processes with central molecular hubs.**

Monitoring Editor
Sandra L. Schmid
University of Texas
Southwestern Medical Center

Received: Feb 26, 2014
Revised: Apr 25, 2014
Accepted: May 1, 2014

INTRODUCTION

Modular design is a feature of all living systems (Barabasi and Albert, 1999; Hartwell *et al.*, 1999; Albert *et al.*, 2000). In eukaryotic cells, a group of proteins directs the fusion of vesicles with the plasma membrane, a process called exocytosis, and another group controls the recapture of plasma membrane components—endocytosis (Dittman and Ryan, 2009). The regulation of these two cellular processes relies on the dynamic interplay between these complex, robust, adaptable, and possibly interchangeable protein networks at discrete molecular structures (e.g., vesicles) located across the cell membrane. Although exocytosis and endocytosis use separate molecular modules, these processes need to be balanced for proper cellular function (Houy *et al.*, 2013). Specifically, exocytosis is the primary means by which materials including chemicals, hormones, and membrane proteins are released from cells into the extracellular space or plasma membrane (Jahn and Sudhof, 1999), and endocytosis is required to recapture membrane and proteins added during exocytosis, as well as to internalize extracellular nutrients and signals (Conner and Schmid, 2003). Understanding the coordinated regulation of exocytosis and endocytosis is a central problem for cell biology.

In the most basic sense, identifying which proteins are commonly found on either exocytic or endocytic vesicles would aid in unraveling the core proteins involved in these processes. Similarly, determining which proteins are found rarely at vesicles would point to specialized or regulatory roles for these components. Even proteins that never associate with vesicles could be ruled out and possibly assigned to other cellular roles. At a different level of organization, how proteins and their associated vesicles distribute across the cell is critical for their function (Hartwell *et al.*, 1999). Specifically, altering

This article was published online ahead of print in MBoC in Press (http://www.molbiolcell.org/cgi/doi/10.1091/mbc.E14-02-0771) on May 7, 2014.

J.W.T., B.T.L., and K.A.S. conceived of and designed the study. J.W.T. and K.A.S. performed the TIRF imaging. B.T.L., J.M.K., and J.W.T. analyzed data. B.T.L. and K.A.S. designed and programmed the simulations. J.W.T. and B.T.L. wrote the article. All authors developed analysis software, discussed results, and helped to develop the article.

The authors declare no competing financial interests.

Address correspondence to: Justin W. Taraska (justin.taraska@nih.gov).

Abbreviations used: *c*, Pearson's correlation coefficient; CCS, clathrin-coated structure; CSR, complete spatial randomness; DCV, dense core vesicle; GFP, green fluorescent protein; NA, numerical aperture; RFP, red fluorescent protein; TIRF, total internal reflection fluorescence.

protein distributions would generate different behaviors. This is because the activity of a protein can be influenced by its binding specificity, overall abundance, and cellular locations (Hartwell *et al.*, 1999). For example, some have proposed that endocytic proteins localize specifically around docked exocytic vesicles (Roos and Kelly, 1999; Blanpied *et al.*, 2002; Holroyd *et al.*, 2002; Jaiswal *et al.*, 2009) or that endocytic proteins are preassembled on exocytic vesicles (Holroyd *et al.*, 2002; Jaiswal *et al.*, 2009). This tight physical coupling might accelerate the retrieval of proteins released during triggered exocytosis (Vogel, 2009). In addition, localizing specific adaptors to a subpopulation of clathrin-coated pits could allow for the coordinated retrieval of membrane proteins at subcellular endocytic hot spots (Puthenveedu and von Zastrow, 2006; Liu *et al.*, 2009, 2010; Loerke *et al.*, 2009; Traub, 2009; Nunez *et al.*, 2011). In these situations, signals could be confined to zones within the cell. All of these scenarios would strongly influence the underlying behavior of these cellular systems. Thus developing a spatial, systems-level understanding of the proteins that associate with exocytic and endocytic vesicles across the plasma membrane is key to understanding how these two biological processes function.

Here we begin to fill this gap in understanding by using live-cell, two-color, total internal reflection fluorescence (TIRF) microscopy along with automated image processing and a uniform, high-throughput quantitative analysis framework to map a library of 78 proteins proposed to have roles in membrane trafficking in single images of individual vesicles in the cell. With this framework we determine both the vesicular and cell-wide distributions of a library of proteins in relation to both calcium-regulated exocytic vesicles (dense core vesicles [DCVs]) and clathrin-coated endocytic structures (clathrin-coated structures [CCSs]) at the plasma membrane—a total of 156 pairs of interactions. These results provide a unified systems-level steady-state spatial map of the protein landscape of individual structures and a framework for the unbiased spatial analysis of other molecular systems. With these data we identify two core groups of proteins that associate with endocytic structures or exocytic vesicles. We also observe a broad range in degree of associations. We find few shared components and several new associations. Furthermore, most proteins are distributed among these structures and across the membrane in a well-mixed random pattern with no apparent underlying spatial organization.

RESULTS

Imaging and analysis framework

We aimed to map the association of a library of proteins to individual endocytic structures (CCSs) and exocytic vesicles (DCV) in an unbiased manner across the plasma membrane of cells. To accomplish this, we labeled CCSs with green fluorescent protein–tagged clathrin light chain (GFP-CLTA) and DCVs with neuropeptide Y-GFP (NPY-GFP). Both probes have been used extensively in other studies (Gaidarov *et al.*, 1999; Merrifield *et al.*, 2002; Taraska *et al.*, 2003). These markers and a collection of 76 other proteins were tagged with red fluorescent protein (Supplemental Table S1). The majority of the proteins in this library have been proposed to be involved in exocytosis or endocytosis. Many, however, have unknown or unrelated roles. Furthermore, the relative associations, distributions, and overlap between these molecular components are unknown.

Figure 1A illustrates our basic imaging and analysis pipeline. First, two-color TIRF microscopy was used to acquire single images of seven or more living PC12 cells coexpressing both a green and a red fluorescent marker. Because CCSs and DCVs are approximately 100 nm in diameter, they appear as punctate, diffraction-limited spots (Kasai *et al.*, 1999; Taraska *et al.*, 2003; Sochacki *et al.*, 2012). The red proteins we investigated appeared as diffraction-limited spots, filaments, or diffuse fluorescence of varying intensities (see Supplemental Figure S1 for examples of all red fluorescent protein [RFP] constructs). We then used image processing to automatically identify single green vesicles (CCSs or DCVs) from these individual images and extract small, square regions centered at the brightest pixel of each object. Matched regions from the same cellular location in the red channel image along with an equal number of randomly positioned regions from the red channel image were also extracted, and the correlation between each green and red image pair was calculated using Pearson's correlation coefficient (*c*; Bolte and Cordelieres, 2006). The size of the regions was chosen based on the average nearest-neighbor spacing of CCSs to ensure that the majority of regions contained only one structure (Sochacki *et al.*, 2012). DCVs are roughly half as dense as CCSs in these cells, and thus the same region size could also be used for DCVs (Taraska *et al.*, 2003). The extracted images were normalized to the brightest pixel in each image and grouped into a stack that contained all the images from every cell. These image stacks were averaged to produce an average image of the field (Figure 1B). The relative brightness in the center of the red image reflects the degree of association between the test protein (red) and the reference object (green). We quantified brightness as peak height by averaging radial line scans centered in the images, subtracting the average radial line scan from random regions, and measuring peak height as the maximum intensity difference between the central point and any other point in the scan (Figure 1B). Averaging all of the individual correlations gave us a second measure for degree of colocalization, which we refer to as correlation (*C*). Absolute magnitude of *C* is not sufficient for determination of colocalization, so all values were compared with average correlation between the green and random red regions, as well as between cytoplasmic and empty vector controls. Here the use of two metrics (peak height and *C*) of colocalization allowed us to cross-reference measurements for robustness, reduce false positives, and better determine a protein's association level. Using simulations, we found both peak height and average correlation to be robust to noise, strongly correlated, affected by both the specificity and frequency of association, and sensitive to small differences in the degree of association (Supplemental Figure S2).

CCSs

We first investigated proteins associated with CCSs. Figure 2 shows the distribution of the 78 tested proteins. Each point represents measurements of thousands of vesicles from single time points acquired from at least seven cells. For each protein, the standard errors of the measurements are smaller than the point size. Supplemental Figure S3 shows box plots of all 256,581 individual correlation measurements from each protein, along with the data grouped by individual cell to show the average cell-to-cell variability across the data set of 902 cells analyzed. As a control, we analyzed free cytoplasmic RFP and "empty" samples not expressing red fluorophores. The empty controls established a reference point for no observed correlation, and as expected, yielded a measurement of 0.00 for correlation and 0.01 for peak height. These values set a threshold for specific colocalization and exclusion with CCSs. We scored proteins as associated if they have correlation $C \geq 0.10$ and peak height ≥ 0.10. Proteins with $C \leq 0.01$ and peak height ≤ -0.05 are scored as excluded from CCSs. Exclusion values were set from the cytoplasmic mCherry control, which is absent from dense core vesicles. Proteins with values within these thresholds were considered to have no measurable relation to the reference object.

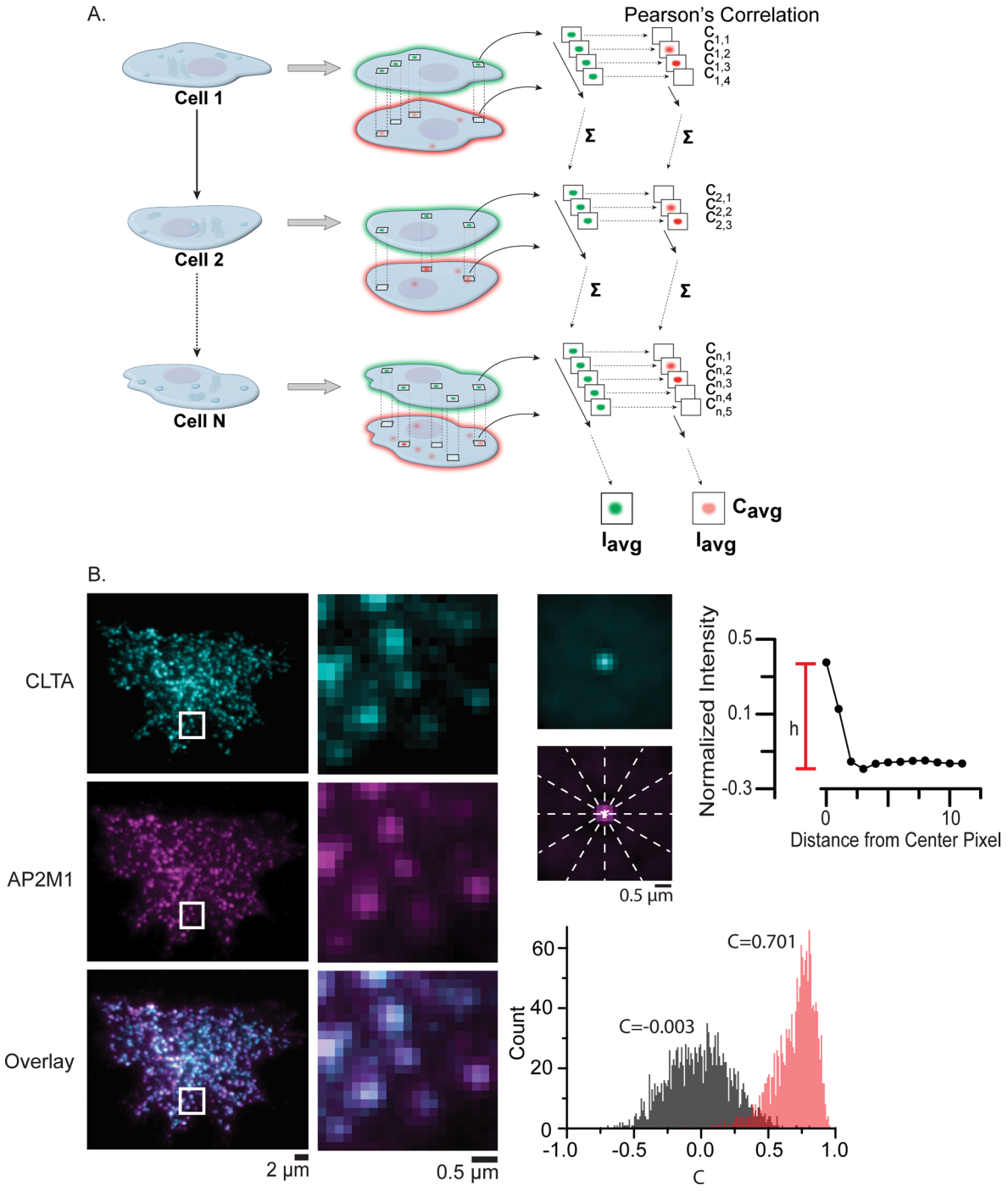

FIGURE 1: Colocalization analysis pipeline. (A) Cells with GFP-tagged structures of interest (CCSs or DCVs) and RFP tags on another protein of interest are imaged in both colors. Small test regions are extracted from the green channel, along with the corresponding region of the cell from the red channel. All regions are then normalized to the brightest pixel. The small regions are averaged together, yielding an image to be used for peak height measurement. Pearson's correlation coefficient (*c*) is measured for each test region pair, and these values are averaged together to give an average correlation value. (B) Example image and illustration of correlation and peak height measurement. Smaller images are average structures resulting from process described in A. Radial line scans are taken of the red average structure (bottom), and the maximum difference from the intensity at the central pixel to any other point in the scan is taken as peak height (graph). The histograms display *c* for all pairs of extracted small regions (red histogram) and for extracted reference regions compared with randomly selected regions from the other channel (gray histogram). The average correlation value is noted above each histogram.

Figure 2A shows examples of three general classes of the steady-state distributions we observed around CCSs from single image pairs. Many proteins (32 proteins tested), such as EPN2, were measured to be associated with CCSs, others (34 proteins), including CAV1, showed no relationship, and some (12 proteins), including STX1A, were excluded (Figure 2). Within the "associated" group we

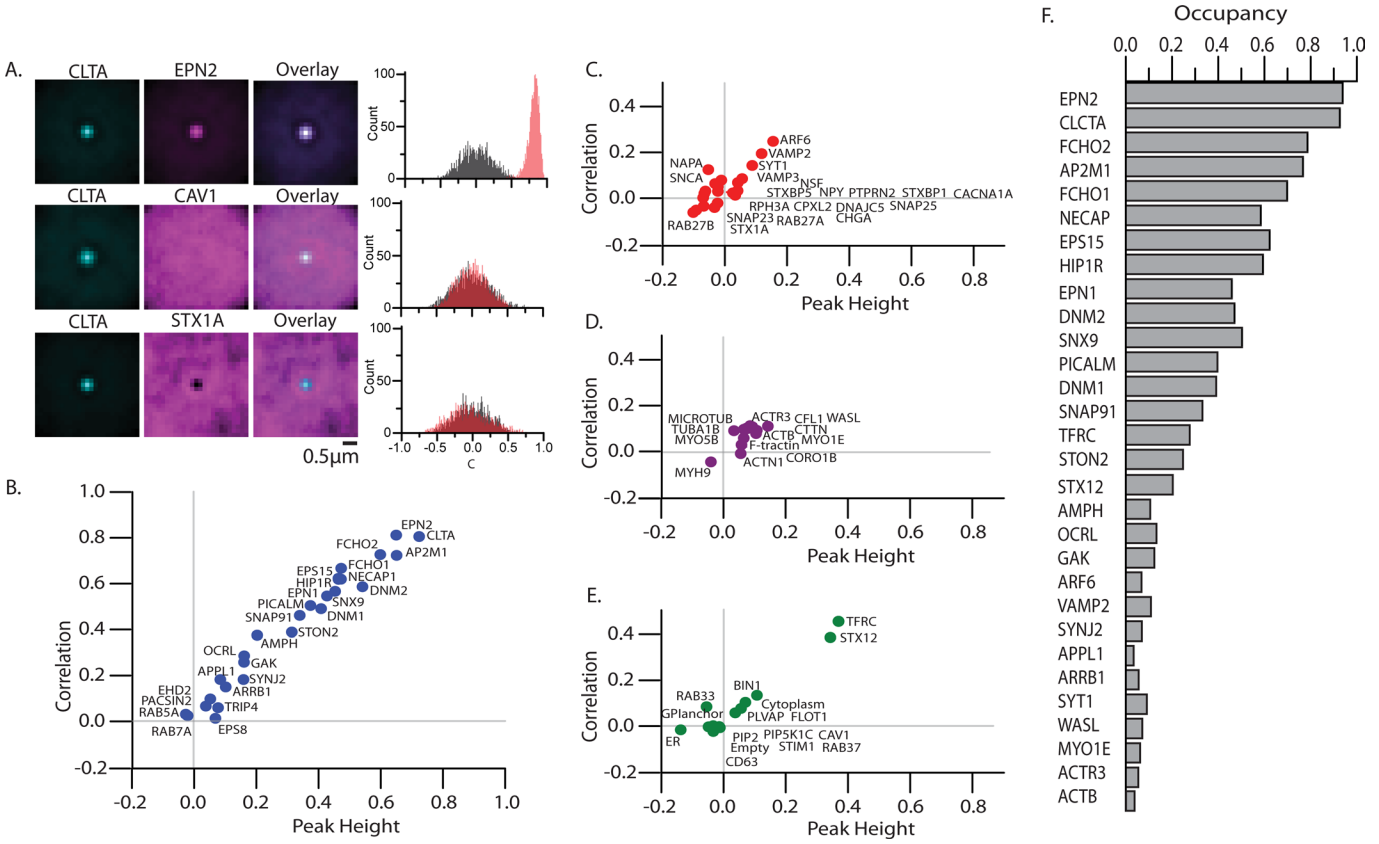

FIGURE 2: Colocalization analysis and occupancy for CCSs. (A) Representative images for range of observed patterns of association along with correlation histograms. Top to bottom: colocalized protein (EPN2), no relationship (CAV1), and exclusion (STX1A). The histograms display *c* for all pairs of extracted small regions (red histograms) and for extracted reference regions compared with randomly selected regions from the other channel (gray histograms). (B–E) Peak height and correlation for the groups of proteins: (B) endocytic, (C) exocytic, (D) cytoskeletal, and (E) "other" or unknown function. Standard error bars are included but smaller than the point size in all cases. (F) Occupancy for all colocalized, punctate proteins ordered by correlation.

see a large range of association values, presumably due to the complex, dynamic, and possibly heterogeneous nature of clathrin-coated pits, as well as to the diversity of proteins surveyed (Taylor *et al.*, 2011). For example, EPN2, FCHO2, and AP2M1 were highly colocalized ($C = 0.81–0.72$), STON2, TFR, and AP180 were moderately colocalized ($C = 0.46–0.38$), and ARRB1, OCRL1, and VAMP2 were only weakly associated over controls ($C = 0.28–0.15$; Figure 2, B–E). We reason that strongly associated proteins are frequently and selectively associated with CCSs. Lower values could represent a combination of rare, less selective, or brief associations. Without inherent dynamic knowledge of the system, we cannot unambiguously distinguish between these behaviors in this analysis. Excluded proteins, such as STX1A, must never associate with CCSs even in a nonspecific manner.

To highlight broad patterns of association, we clustered proteins into four general categories according to their major proposed cellular functions gleaned from the literature: endocytic, exocytic, cytoskeletal, and other. We found most proteins in the endocytic classification associated with CCSs (Figure 2B). However, these proteins exhibit a wide range of values (Figure 2B). For example, the adapter proteins ARRB1, STON2, and AP2M1 have correlation values of 0.15, 0.39, and 0.72 respectively. Several proteins in the exocytic group (VAMP2, SYT1) were weakly colocalized. Most of these, however, were previously identified as endocytic cargo for clathrin-coated pits. Another cargo protein, TFRC, which cycles through the endocytic/exocytic system, was also found to be associated with CCSs, with a correlation of 0.46. Many important exocytic proteins were excluded (Figure 2C). Cytoskeletal proteins exhibited a weak tendency toward colocalization (Figure 2D). This is expected, as the cytoskeleton has been proposed to have an important role in clathrin-mediated endocytosis (Merrifield *et al.*, 2002; Kaksonen *et al.*, 2005). Finally, a small number of proteins in the uncategorized group (other) were associated (Figure 2E). For example, STX12, a protein proposed to be involved in endosome function, was associated with CCSs (Tang *et al.*, 1998).

As a control, we investigated the general effect of overexpression on a subset of proteins that spanned the range of associations to CCSs. We find that expression levels over a 12-fold range in magnitude appeared to have little to no effect on cell-wide average correlation values (Supplemental Figure S4). Furthermore, cell-to-cell variability for all tested proteins in the library was relatively small, with a limited number exhibiting notable variability or extreme outliers (Supplemental Figure S3). In particular, EPN1, F-tractin, and EPS15, showed much larger than average cell-to-cell average correlation variability. APPL1 and SYNJ2 exhibited outliers that indicate a different association pattern than the average *C* for each of these proteins.

Peak height and *C* in our analysis represent how often a protein is found spatially associated with a reference structure at steady state. To further characterize the system, we developed an explicit

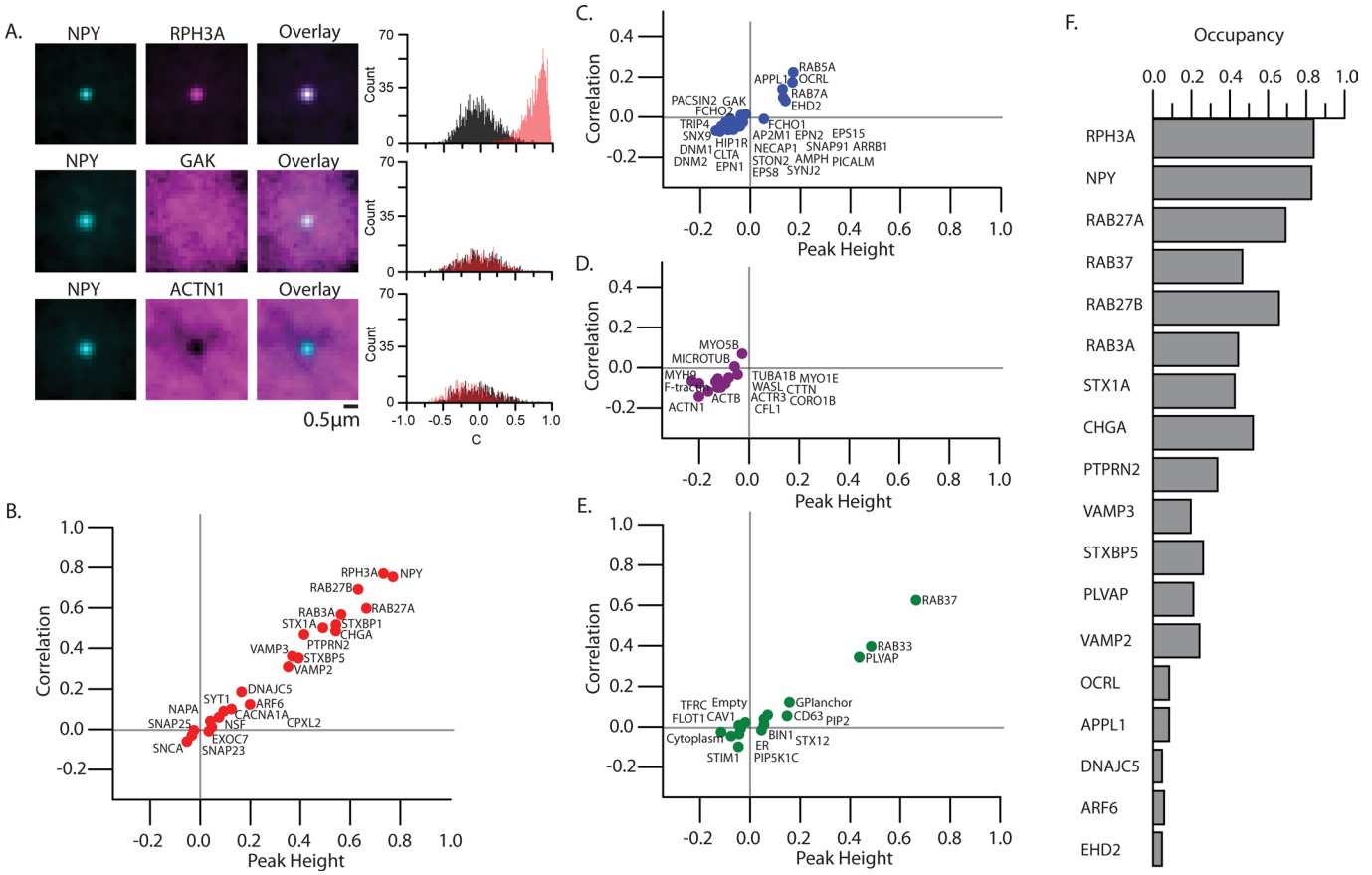

FIGURE 3: Colocalization analysis and occupancy for DCVs. (A) Representative images for range of observed patterns of association along with correlation histograms. Top to bottom: colocalized protein (RPH3A), no relationship (GAK), and exclusion (ACTN1). The histograms display *c* for all pairs of extracted small regions (red histograms) and for extracted reference regions compared with randomly selected regions from the other channel (gray histograms). (B–E) Peak height and correlation for the groups of proteins: (B) exocytic, (C) endocytic, (D) cytoskeletal, and (E) "other" or unknown function. All protein names are from the HUGO Gene Nomenclature Committee. Standard error bars are included but are smaller than the point size in all cases. (F) Occupancy for all colocalized, punctate proteins ordered by correlation.

measure of occupancy. Here we define occupancy as the frequency of colocalization above a control threshold. For analysis, we excluded proteins with diffuse distributions because we define occupancy as a spatially discrete event. Of the 78 red proteins, 57 were classified as punctate. To measure occupancy, we used *c* values for all test images. We set a threshold by finding the correlation value 2 SDs above the mean correlation of the random image control set. Pairs of images with values above were scored as occupied, and occupancy is the percentage of total structures occupied (see Supplemental Figure S5 for additional information about the occupancy calculation). Figure 2F displays the results of our analysis. Occupancy in general matches the trend for degree of colocalization. Of interest, some values were low, at <10%. These data highlight the diversity of the population of individual CCSs at any one time point.

DCVs

We next characterized the associations of the same group of test proteins with DCVs located close to the plasma membrane (see Supplemental Figure S6 for an example image of an NPY-GFP cell). In PC12 cells, DCVs are involved in the release of peptides and chemicals through the process of calcium-triggered exocytosis (Taraska *et al.*, 2003). They are proposed to be functionally linked to CCSs through the compensatory recycling of some vesicle components lost into the membrane during exocytosis (Allersma *et al.*, 2004; Taraska and Almers, 2004; Bittner *et al.*, 2013; Houy *et al.*, 2013). Similar to our analysis of CCs, we see a wide distribution of correlation values to DCVs (Figure 3). Again, each point represents measurements of thousands of DCVs from single images of at least seven cells, and standard errors are smaller than the point size. Our controls set similar bounds for considering proteins to be colocalized ($C \geq 0.08$ and peak height ≥ 0.10), excluded ($C \leq -0.01$ and peak height ≤ -0.05), or not spatially related (neither colocalized nor excluded). Some proteins (27 proteins), such as RPH3A, were colocalized; others, such as GAK, show no relationship (19 proteins); and some, including ACTN1, were excluded (32 proteins; Figure 3). There were, however, fewer proteins with intermediate values compared with CCSs, potentially due to the relatively static nature of DCVs docked at the plasma membrane. Supplemental Figure S7 presents box plots of correlation data from all 61,375 individual DCV measurements and 760 average NPY-GFP cell measurements grouped by protein.

As expected, proteins classified as exocytic were mostly colocalized. Surprisingly, a subset of the endocytic group, including RAB5A, OCRL, APPL1, RAB7A, and EHD2, was weakly associated (Figure 3B). These proteins had not previously been mapped to DCVs.

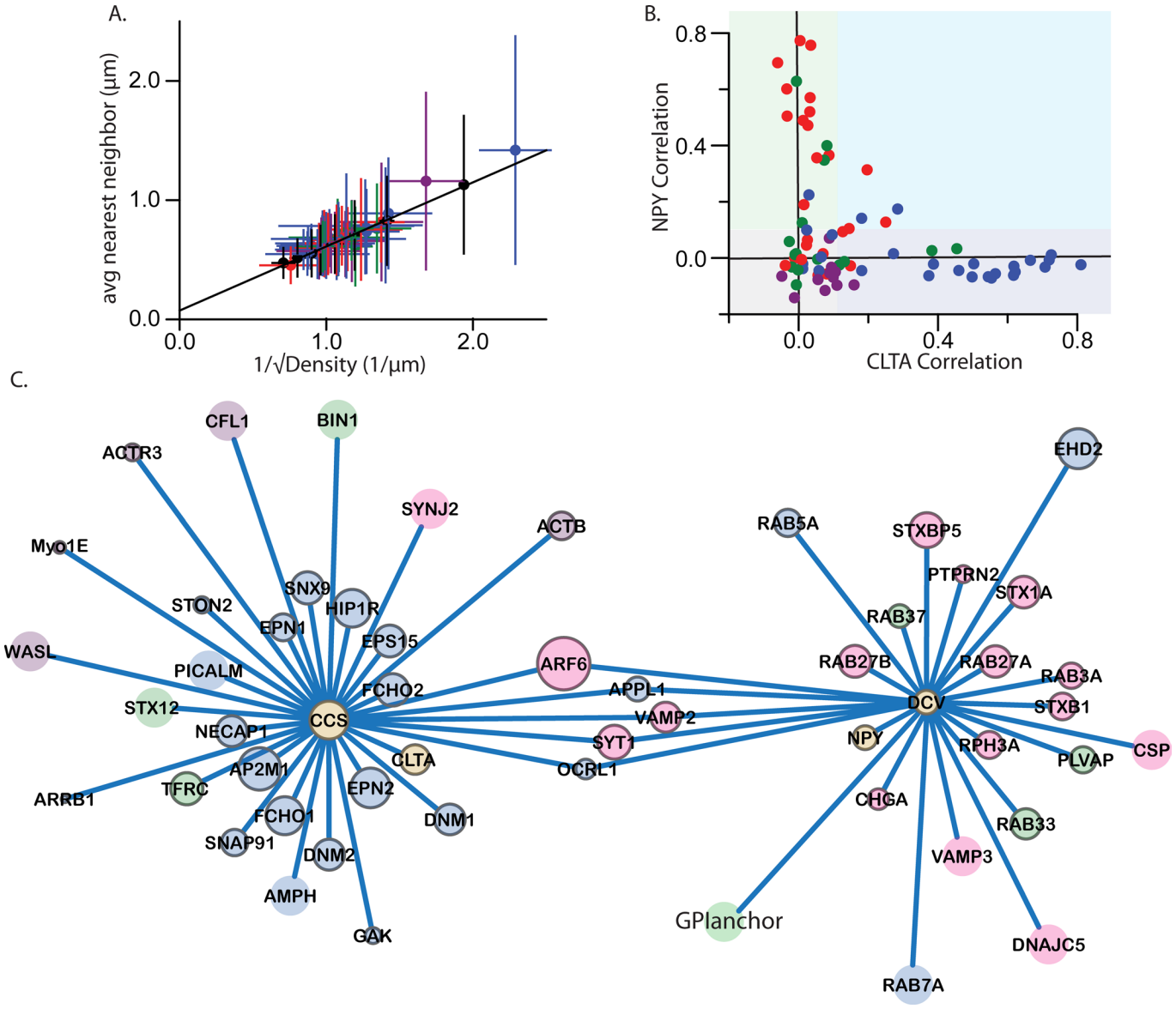

FIGURE 4: Spatial distributions and global analysis. (A) Spatial distributions of proteins compared with complete spatial randomness (CSR) as determined by simulations (black markers and line of best fit). For CSR, we expect a proportional relationship between density and the square of the average nearest-neighbor distance and, hence, a linear relationship between average nearest-neighbor distance and the reciprocal square root of density. (B) CLTA correlation vs. NPY correlation for all proteins. (C) Map of proteins associated to CCS and DCV. Edge lengths are proportional to average correlation value, and node size is proportional to density of protein puncta. Nodes for which background was too high to reliably measure density of puncta have no outline.

Many endocytic proteins were excluded (Figure 3C). This is consistent with previous work showing that calcium-regulated exocytic vesicles tend to fuse with regions of membrane between CCSs (Sochacki *et al.*, 2012). We additionally found that cytoskeletal proteins were largely excluded (Figure 3D). This also confirms previous reports that exocytic vesicles dock in holes of the cytoskeleton in endocrine cells (Wang and Richards, 2011). A small number of proteins in the uncategorized group (other), including RAB37, RAB33, PLVAP, and GPI anchor, were found to be associated (Figure 3E). Similar to our analysis of CCSs, cell-to-cell variability for DCVs was mostly small; however, a few proteins, including CHGA, STXBP1, RAB5A, and VAMP2, had much larger than average cell-to-cell variability. Again, we determined occupancy of the proteins of interest for DCVs (Figure 3F). Occupancy was correlated to degree of colocalization, and low occupancy values indicate a molecularly heterogeneous population of DCVs at any one time point.

Spatial distribution patterns and global network analysis

We next considered the spatial distribution patterns of endocytic and exocytic proteins across the plasma membrane. We hypothesized that there may be cell-wide organization in the spatial distributions of components involved in membrane trafficking. Here we characterized these distributions statistically. Complex spatial point patterns can be described by nearest-neighbor distributions (Clark and Evans, 1954). Thus, for each protein that showed a punctate distribution, we determined the position of all structures and calculated both the density and nearest-neighbor distances (Figure 4A). We measured the density of CCSs to be 1.22/µm^2, consistent

with previous manual measurements from our group (Sochacki *et al.*, 2012), and the density of DCVs to be 0.76/μm^2. The distribution of nearest-neighbor distances for nearly all proteins resembled Rayleigh distributions (see Supplemental Figure S2 for an example) characteristic of spatial randomness and stochastic spatial point processes. Although complete spatial randomness (CSR) is well understood and characterized, there are no general solutions for the basic descriptive statistics for finite spatial fields with complex boundary geometries (i.e., cell outlines; Brewer and Mccann, 1985). On addition, the point-spread function introduces a minimum possible distance between localized objects (the diffraction limit). Therefore, to compare our experimental distributions to CSR, we simulated spatially random distributions of objects of finite radius within cell boundaries drawn from our experimental data. These simulations give us a characteristic mean nearest-neighbor distance and variance for a given density of random distributions within cells. These were then compared with experimental distributions (Figure 4A). In Figure 4A, we plot the reciprocal square root of density versus average nearest-neighbor distance to highlight the expected roughly linear relationship between these two quantities for random distributions of objects with finite radii (Matérn, 1986). We find that, at the optical resolution of this study, nearly all of the distributions of punctate proteins are consistent with CSR (Figure 4A). One notable exception is SYT1, for which the mean nearest-neighbor distance is smaller than expected, indicating clustering. Closer inspection of SYT1 images reveals that this protein tends to be concentrated at the edges of cells (Supplemental Figure S1).

To summarize our results, we plotted the correlation values of each protein to both CCSs and DCVs (Figure 4B). Most proteins are specific for only one cellular organelle. Only six components show positive correlation values with both vesicle types. These data highlight the distinct identity of these cellular organelles. Finally, we used our analysis to generate a systems-level network map (Figure 4C). Most proteins cluster around only one cellular node. The proteins closest to the central nodes are those found most often at each organelle as determined by correlation value. The six shared components were new associations or those that cycle through the endocytic/exocytic pathway as cargo. Node size is scaled by density, and edge length is scaled by correlation. Note that proteins have a diverse mixture of colocalization values and densities.

DISCUSSION

DCVs and CCSs are discrete multicomponent molecular assemblies at the plasma membrane. Here we develop an approach to produce a systems-level quantitative molecular map of individual exocytic and endocytic vesicles at the plasma membrane of cells. We identify proteins found often at vesicles, those that populate vesicles at lower densities, and proteins that do not associate with either one or both of the structures. The steady-state picture that emerges from our analysis is that of a decoupled, well-mixed system. Furthermore, we identify several novel components of DCVs and CCSs. We conclude that this general approach can be used to interrogate and quantitate the molecular composition of complex molecular systems in cells.

From our analysis we observe a wide range of association values for proteins and CCSs. We see many proteins that are either excluded from CCSs (STX1A, RAB3A, RAB27B) or have no relationship to the structures (CAV1, RAB5A, STIM1). We do, however, see a number of components at very high occupancies. These include EPN2, FCHO2, FCHO1, EPS15, HIP1R, and NECAP1. We reason that these proteins are likely important structural or functional components of the CCS maturation. Indeed, some of these proteins, such as AP2M1, are essential for clathrin-mediated endocytosis (Dittman and Ryan, 2009). Proteins with lower colocalization values (i.e., less frequently encountered), such as GAK, APPL1, and SYNJ2, likely represent transient molecular associations or long-lived but very rare associations. Indeed, many of these proteins, including GAK and SYNJ2, have been shown in live-cell imaging studies to transiently associate with CCSs soon before or soon after the vesicle separates from the plasma membrane (Taylor *et al.*, 2011). Of interest, even within the same subclass of proteins, we observed widely different levels of occupancy. For example, the clathrin adapter proteins EPN2, ARRB1, STON2, SNAP91, and AP2M1 are found at drastically different occupancies (Owen *et al.*, 2004; Traub, 2009). Adapter proteins bind clathrin and unique cargo destined for endocytosis (Sorkin, 2004; Traub, 2009; Kelly and Owen, 2011). A change in the occupancy of a given adapter at CCSs could regulate rates of cargo internalization across the population of existing CCSs (von Zastrow and Kobilka, 1994). This could represent a tunable, distributed sensing and response network for endocytosis.

Similar to CCSs, we observed high occupancy rates for a subset of DCV-associated proteins. These proteins, such as RPH3A, RAB3A, RAB27B, and RAB27A, are likely crucial for proper DCV function. Several proteins found to be strongly associated with DCVs, including STX1A, STXBP1, and STXBP5, likely constitute part of the membrane complex responsible for docking DCVs at the plasma membrane (Dittman and Ryan, 2009). Also similarly to CCSs, we observe a range of degrees of association. Because DCVs are believed to be more static than CCSs and persist at the membrane during the resting state, the observed protein heterogeneity of DCVs was unexpected. Exocytosis, however, of unique populations of DCVs may be triggered by different signals. This would be important if individual DCVs package unique cargoes (Taraska *et al.*, 2003). This diversity could be maintained and adjusted by a dynamic maturation process or by differential assembly of exocytic vesicles. Again, like CCSs, the density of a particular component could tune the activity of the underlying network of total available vesicles.

Our method also revealed several new partners for both CCSs and DCVs. We found that STX12 is associated with CCSs, and that RAB5A, OCRL1, APPL1, RAB7A, EHD2, and PLVAP associate with DCVs. The roles of these proteins in exocytosis and endocytosis are unknown. Many of these associations are relatively weak and may have been missed by other techniques. Unlike standard correlation-based colocalization analyses, our approach also allows us to measure exclusion. Many proteins, including most cytoskeletal components, were strongly excluded from DCVs. These include ATCB, CORO1B, and WASL. Vesicles appear to dock in distinctive cytoskeletal cages at the membrane (Wang and Richards, 2011). For CCSs, the protein SXT1A was strongly excluded, and this was a novel observation generated from our analysis. These findings demonstrate the utility of our approach in discovering new protein relationships. Of interest, we find relatively few shared components. Given that CCSs and DCVs are reciprocal nodes in the membrane-recycling pathway and previous results suggested that some endocytic proteins are preassembled on exocytic structures (Holroyd *et al.*, 2002; Jaiswal *et al.*, 2009), this was unexpected. Some shared proteins are membrane proteins internalized by clathrin-mediated endocytosis (VAMP2 and SYT1; Dittman and Ryan, 2009). OCRL1 and APPL1, which are also shared components, however, have unknown roles in both processes.

At the cellular level we found that endocytic and exocytic structures and their associated molecules are randomly distributed across the plasma membrane. This pattern is consistent with an underlying Poisson point-process of structure formation (or docking) and/or

maturation for both vesicles types. In general, there are twice as many CCSs as DCVs in the cell. We propose that CCSs nucleate and mature through the self-organizing, autocatalyzing, and stochastic nucleation of diffuse cellular components. DCVs tend to dock randomly in regions between CCSs. Future work with electron and superresolution microscopy will reveal whether local patterns exist at the scale of tens of nanometers (Sochacki *et al*., 2014).

Although our analysis is robust and unbiased, it does have some specific limitations. For example, underexpression of a protein may lead to an artificially low occupancy, whereas overexpression may lead to artificially high occupancy. Other factors, including state of the cell cycle, competition with endogenous proteins, effect of neighboring cells, or effects of the tags, could change our results. These issues are problems in all imaging studies in which proteins are tagged and exogenously expressed. We tried to account for these factors by measuring at least seven cells for every construct, which spanned the range of intensities encountered. Our detailed analysis of expression level indicated that cells of widely different expression levels showed very little difference in average correlation value. Furthermore, our systematic analysis of cell-to-cell variability indicated that the majority of the proteins in this study have very low such variability, with a few exceptions. The underlying causes of the cell-to-cell variability in these cases are not known and would provide an interesting avenue to further investigation. Here we present the average steady-state association values of a single protein with both CCS and DCVs. These representative images from single cells should produce representative mean population correlation and intensity values that reflect the average values for the average cell at a single time point. Thus, along with the findings presented here, our map serves as a baseline for probing conditions or perturbations for which patterns of association might deviate from these averages. Because our data were generated from single snapshots of the cell in the resting state, additional dynamic information gleaned from live-cell time-lapse data would provide for a more complete picture of these systems (Kaksonen *et al*., 2005; Henne *et al*., 2010; Loerke *et al*., 2011; Taylor *et al*., 2011; Cocucci *et al*., 2012). Specifically, combining the global spatial information generated from our analysis with an understanding of the system's underlying steady-state behavior could be used to mathematically derive complex information on the system's dynamic properties (Kafri *et al*., 2013).

Our results illustrate the ability of fluorescence imaging and quantitative analysis to interrogate the underlying structure of a complex spatial system and screen for new associations and patterns. The massive application of this analysis to a library of all the proteins in a mammalian cell could allow for the complete quantitative mapping of the global protein landscape of a cell. The random distribution we observe for almost all proteins studied suggests an overall structural model for how the plasma membrane is organized in relation to the endocytic and exocytic structures in neuroendocrine cells. CCSs are very dense in PC12 cells, from 2 to 11 times as dense as reported in other cell types (Loerke *et al*., 2009; Liu *et al*., 2010; Nunez *et al*., 2011). This dense network of CCSs may allow for robust, rapid sorting and capture of cargo and possible enhancement in signal transduction. Measurements in the time domain and use of targeted perturbations could be used to probe the molecular diversity of these structures and will be needed to fully characterize the system (Kim and Ryan, 2009; Taylor *et al*., 2011; Willox and Royle, 2012). The static network map we generate through imaging can be used as a reference for comparison between different cell types, states, and perturbations. Investigating individual functional units of interacting molecules in their cellular context will deepen our understanding of the self-organization, complexity, spatial organization, and control of biological systems.

MATERIALS AND METHODS

Plasmids, solutions, and cells

A complete list of the 80 plasmids used in this study, their construction, and their original sources are included in the supplemental information (Supplemental Table S1). PC12-GR5 cell stocks were cultured as described previously. Briefly, cells were plated onto poly(L-lysine)-coated glass coverslips. Cells were transfected with 1 µg of plasmid DNAs using Lipofectamine 2000 (Invitrogen, Grand Island, NY) according to the manufacturer's instructions. For imaging, cells were maintained in imaging buffer (in mM: 130 NaCl, 2.8 KCl, 5 $CaCl_2$, 1 $MgCl_2$, 10 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid, 10 glucose, pH 7.4). Cells were imaged 24–48 h after transfection. Experiments were carried out at 28°C.

TIRF microscopy

Cells were imaged with an inverted microscope (IX-81; Olympus, Center Valley, PA) configured for TIRF and equipped with a 100×/1.45 numerical aperture (NA) objective (Olympus) and an image splitter (Dual View; Photometrics, Tucson, AZ) for simultaneous imaging of red and green fluorescence. Fluorescence was excited by a laser at 488 (Coherent; Sapphire) and 561 nm (LD-561-20A; Melles Griot, Carlsbad, CA). Lasers were combined and then controlled with an acousto-optic tunable filter (Andor, Belfast, UK). Optical filters were bright-line, full multiband LF405/488/561/635 filters (Semrock, Rochester, NY). The resulting emission was then divided by the image splitter's dichroic (565DCXR) and projected side by side through 525Q/50 and 605Q/55 emission filters onto the chip of a back-illuminated electron-multiplying charge-coupled device camera (DU 897; Andor). Images were acquired using IQ software (Andor). For precise alignment of red and green images, we imaged 100-nm fluorescent beads (Invitrogen) that emitted light in both the red and green channels before each experiment. The centroids of six beads across the imaging region were mapped from the red to the green side of the image using a projective image transformation. The same transform coordinates were used to overlay red and green live-cell data. Pixel size was 160 nm. Frames were acquired in time-lapse recordings with alternating 488- and 561-nm excitation with 100-ms exposures given at 2 Hz. Each image used for analysis is the average of five exposures.

Image analysis

All image analysis was done with custom MATLAB (MathWorks, Natick, MA) software. All of the local maxima above a user-defined threshold and >2 µm away from the edge of the cell boundary were identified in each whole-cell green image (NPY-GFP for exocytic vesicles and GFP-CLTA for endocytic vesicles). Each local maximum was centered in a 4 µm × 4 µm cropped region. These cropped regions from several cells were each normalized to the brightest pixel and averaged together. The regions corresponding to the green local maxima were also excised from the red images, normalized to the brightest pixel, and averaged together. In each average red image, 12 radial line scans were taken and averaged together. Randomly selected regions equal in number to the excised regions were processed in the same way. The average random line scan was then subtracted from the average red image line scan to account for background intensity statistics. Peak height for each construct was determined by measuring the maximum intensity difference between the first point (center of the image) and all other points in the averaged scan.

Smaller 1.0 μm × 1.0 μm cropped, vesicle-centered red/green regions were used to calculate Pearson's *c* for each green peak:

$$c = \frac{\sum_{i=1}^{n}(G_i - \overline{G})(R_i - \overline{R})}{\sqrt{\sum_{i=1}^{n}(G_i - \overline{G})}\sqrt{\sum_{i=1}^{n}(R_i - \overline{R})}}$$

where G_i and R_i represent pixel intensities in the two channels and the overbar indicates average pixel intensity. The region size was chosen based on the spacing of CCSs in an attempt to capture only one structure per correlation region on average. Correlations were then averaged together, yielding an average *C* for each construct. Correlation values were calculated between peak-centered green excised regions and randomly selected red regions as well. The threshold for statistically significant correlation values for calculation of occupancy was set as 2 SDs above the mean of the green region to random red region correlation values (Supplemental Figure S5).

Simulations

Simulations were performed using MATLAB. For all simulations, the field was populated with randomly distributed points at a given density. This field was then convolved with the Gaussian approximation of the theoretical point-spread function for the microscope used for this study (NA = 1.45, λ = 500) and pixelated with 160-nm pixels. For determination of the expected spatial distributions under complete spatial randomness (Supplemental Figure S2, D and E), the field was then cropped using a mask made from the outline of a real cell to introduce realistic boundary geometry. Structures (diffraction-limited point-spread functions) were then located using the same custom automated program used for image analysis, and density and nearest- neighbor statistics were calculated from these locations. For each density, 10 simulations using 10 different cell outlines were analyzed. For simulations investigating the robustness and sensitivity of our analysis (Supplemental Figure S2, A–C), 4 μm × 4 μm reference regions were created with a centered object and randomly distributed objects within the region at a density of 1/μm^2. Corresponding regions with appropriate patterns of localization (Supplemental Figure S2, B and C) were created for all reference regions. Region pairs were analyzed as described in the *Image analysis* section. For simulations with added noise, a background pedestal of 80% of the maximum object intensity was added, along with an additional 10% shot noise.

ACKNOWLEDGMENTS

We thank Keir Neuman, Jonathan Silver, John Hammer, and members of the Taraska lab for critical reading of the manuscript. We also thank Paul Whittredge and Shiqin Judy Yu for help with cloning. Additionally, we thank the many labs who contributed plasmids to this study and Alan Hoofring of NIH Medical Arts for design work on Figure 1. J.W.T. is supported by the Intramural Research Program of the National Heart, Lung, and Blood Institute, National Institutes of Health.

REFERENCES

Albert R, Jeong H, Barabasi AL (2000). Error and attack tolerance of complex networks. Nature 406, 378–382.

Allersma MW, Wang L, Axelrod D, Holz RW (2004). Visualization of regulated exocytosis with a granule-membrane probe using total internal reflection microscopy. Mol Biol Cell 15, 4658–4668.

Barabasi AL, Albert R (1999). Emergence of scaling in random networks. Science 286, 509–512.

Bittner MA, Aikman RL, Holz RW (2013). A nibbling mechanism for clathrin-mediated retrieval of secretory granule membrane after exocytosis. J Biol Chem 288, 9177–9188.

Blanpied TA, Scott DB, Ehlers MD (2002). Dynamics and regulation of clathrin coats at specialized endocytic zones of dendrites and spines. Neuron 36, 435–449.

Bolte S, Cordelieres FP (2006). A guided tour into subcellular colocalization analysis in light microscopy. J Microsc 224, 213–232.

Brewer R, Mccann MT (1985). Spacing in acorn woodpeckers. Ecology 66, 307–308.

Clark PJ, Evans FC (1954). Distance to nearest neighbor as a measure of spatial relationships in populations. Ecology 35, 445–453.

Cocucci E, Aguet F, Boulant S, Kirchhausen T (2012). The first five seconds in the life of a clathrin-coated pit. Cell 150, 495–507.

Conner SD, Schmid SL (2003). Regulated portals of entry into the cell. Nature 422, 37–44.

Dittman J, Ryan TA (2009). Molecular circuitry of endocytosis at nerve terminals. Annu Rev Cell Dev Biol 25, 133–160.

Gaidarov I, Santini F, Warren RA, Keen JH (1999). Spatial control of coated-pit dynamics in living cells. Nat Cell Biol 1, 1–7.

Hartwell LH, Hopfield JJ, Leibler S, Murray AW (1999). From molecular to modular cell biology. Nature 402, C47–52.

Henne WM, Boucrot E, Meinecke M, Evergren E, Vallis Y, Mittal R, McMahon HT (2010). FCHo proteins are nucleators of clathrin-mediated endocytosis. Science 328, 1281–1284.

Holroyd P, Lang T, Wenzel D, De Camilli P, Jahn R (2002). Imaging direct, dynamin-dependent recapture of fusing secretory granules on plasma membrane lawns from PC12 cells. Proc Natl Acad Sci USA 99, 16806–16811.

Houy S, Croise P, Gubar O, Chasserot-Golaz S, Tryoen-Toth P, Bailly Y, Ory S, Bader MF, Gasman S (2013). Exocytosis and endocytosis in neuroendocrine cells: inseparable membranes! Front Endocrinol (Lausanne) 4, 135.

Jahn R, Sudhof TC (1999). Membrane fusion and exocytosis. Annu Rev Biochem 68, 863–911.

Jaiswal JK, Rivera VM, Simon SM (2009). Exocytosis of post-Golgi vesicles is regulated by components of the endocytic machinery. Cell 137, 1308–1319.

Kafri R, Levy J, Ginzberg MB, Oh S, Lahav G, Kirschner MW (2013). Dynamics extracted from fixed cells reveal feedback linking cell growth to cell cycle. Nature 494, 480–483.

Kaksonen M, Toret CP, Drubin DG (2005). A modular design for the clathrin- and actin-mediated endocytosis machinery. Cell 123, 305–320.

Kasai H, Kishimoto T, Liu TT, Miyashita Y, Podini P, Grohovaz F, Meldolesi J (1999). Multiple and diverse forms of regulated exocytosis in wild-type and defective PC12 cells. Proc Natl Acad Sci USA 96, 945–949.

Kelly BT, Owen DJ (2011). Endocytic sorting of transmembrane protein cargo. Curr Opin Cell Biol 23, 404–412.

Kim SH, Ryan TA (2009). Synaptic vesicle recycling at CNS synapses without AP-2. J Neurosci 29, 3865–3874.

Liu AP, Aguet F, Danuser G, Schmid SL (2010). Local clustering of transferrin receptors promotes clathrin-coated pit initiation. J Cell Biol 191, 1381–1393.

Liu AP, Loerke D, Schmid SL, Danuser G (2009). Global and local regulation of clathrin-coated pit dynamics detected on patterned substrates. Biophys J 97, 1038–1047.

Loerke D, Mettlen M, Schmid SL, Danuser G (2011). Measuring the hierarchy of molecular events during clathrin-mediated endocytosis. Traffic 12, 815–825.

Loerke D, Mettlen M, Yarar D, Jaqaman K, Jaqaman H, Danuser G, Schmid SL (2009). Cargo and dynamin regulate clathrin-coated pit maturation. PLoS Biol 7, e57.

Matérn B (1986). Spatial Variation, 2nd ed., Berlin: Springer-Verlag.

Merrifield CJ, Feldman ME, Wan L, Almers W (2002). Imaging actin and dynamin recruitment during invagination of single clathrin-coated pits. Nat Cell Biol 4, 691–698.

Nunez D, Antonescu C, Mettlen M, Liu A, Schmid SL, Loerke D, Danuser G (2011). Hotspots organize clathrin-mediated endocytosis by efficient recruitment and retention of nucleating resources. Traffic 12, 1868–1878.

Owen DJ, Collins BM, Evans PR (2004). Adaptors for clathrin coats: structure and function. Annu Rev Cell Dev Biol 20, 153–191.

Puthenveedu MA, von Zastrow M (2006). Cargo regulates clathrin-coated pit dynamics. Cell 127, 113–124.

Roos J, Kelly RB (1999). The endocytic machinery in nerve terminals surrounds sites of exocytosis. Curr Biol 9, 1411–1414.

Sochacki KA, Larson BT, Sengupta DC, Daniels MP, Shtengel G, Hess HF, Taraska JW (2012). Imaging the post-fusion release and capture of a vesicle membrane protein. Nat Commun 3, 1154.

Sochacki KA, Shtengel G, van Engelenburg SB, Hess HF, Taraska JW (2014). Correlative super-resolution fluorescence and metal-replica transmission electron microscopy. Nat Methods 11, 305–308.
Sorkin A (2004). Cargo recognition during clathrin-mediated endocytosis: a team effort. Curr Opin Cell Biol 16, 392–399.
Tang BL, Tan AE, Lim LK, Lee SS, Low DY, Hong W (1998). Syntaxin 12, a member of the syntaxin family localized to the endosome. J Biol Chem 273, 6944–6950.
Taraska JW, Almers W (2004). Bilayers merge even when exocytosis is transient. Proc Natl Acad Sci USA 101, 8780–8785.
Taraska JW, Perrais D, Ohara-Imaizumi M, Nagamatsu S, Almers W (2003). Secretory granules are recaptured largely intact after stimulated exocytosis in cultured endocrine cells. Proc Natl Acad Sci USA 100, 2070–2075.
Taylor MJ, Perrais D, Merrifield CJ (2011). A high precision survey of the molecular dynamics of mammalian clathrin-mediated endocytosis. PLoS Biol 9, e1000604.
Traub LM (2009). Tickets to ride: selecting cargo for clathrin-regulated internalization. Nat Rev Mol Cell Biol 10, 583–596.
Vogel SS (2009). Channeling calcium: a shared mechanism for exocytosis-endocytosis coupling. Sci Signal 2, pe80.
von Zastrow M, Kobilka BK (1994). Antagonist-dependent and -independent steps in the mechanism of adrenergic receptor internalization. J Biol Chem 269, 18448–18452.
Wang J, Richards DA (2011). Spatial regulation of exocytic site and vesicle mobilization by the actin cytoskeleton. PLoS One 6, e29162.
Willox AK, Royle SJ (2012). Stonin 2 is a major adaptor protein for clathrin-mediated synaptic vesicle retrieval. Curr Biol 22, 1435–1439.