Functional characterization of 67 endocytic accessory proteins using multiparametric quantitative analysis of CCP dynamics

Madhura Bhave[a,1], Rosa E. Mino[a,1], Xinxin Wang[a,b], Jeon Lee[b], Heather M. Grossman[a,b], Ashley M. Lakoduk[a], Gaudenz Danuser[a,b], Sandra L. Schmid[a,2,3], and Marcel Mettlen[a,1,3]

[a]Department of Cell Biology, University of Texas Southwestern Medical Center, Dallas, TX 75390; and [b]Lyda Hill Department of Bioinformatics, University of Texas Southwestern Medical Center, Dallas, TX 75390

This contribution is part of the special series of Inaugural Articles by members of the National Academy of Sciences elected in 2020.

Contributed by Sandra L. Schmid, October 24, 2020 (sent for review September 29, 2020; reviewed by Pietro De Camilli and Mark von Zastrow)

Clathrin-mediated endocytosis (CME) begins with the nucleation of clathrin assembly on the plasma membrane, followed by stabilization and growth/maturation of clathrin-coated pits (CCPs) that eventually pinch off and internalize as clathrin-coated vesicles. This highly regulated process involves a myriad of endocytic accessory proteins (EAPs), many of which are multidomain proteins that encode a wide range of biochemical activities. Although domain-specific activities of EAPs have been extensively studied, their precise stage-specific functions have been identified in only a few cases. Using single-guide RNA (sgRNA)/dCas9 and small interfering RNA (siRNA)-mediated protein knockdown, combined with an image-based analysis pipeline, we have determined the phenotypic signature of 67 EAPs throughout the maturation process of CCPs. Based on these data, we show that EAPs can be partitioned into phenotypic clusters, which differentially affect CCP maturation and dynamics. Importantly, these clusters do not correlate with functional modules based on biochemical activities. Furthermore, we discover a critical role for SNARE proteins and their adaptors during early stages of CCP nucleation and stabilization and highlight the importance of GAK throughout CCP maturation that is consistent with GAK's multifunctional domain architecture. Together, these findings provide systematic, mechanistic insights into the plasticity and robustness of CME.

clathrin-mediated endocytosis | CRISPRi screen | SNAREs | GAK | total internal reflection fluorescence microscopy

Clathrin-mediated endocytosis (CME) regulates the uptake of nutrients, growth factors, adhesion molecules, transmembrane ion channels, transporters, signaling receptors, and other ligand–receptor complexes. Thus, CME plays a crucial role in cell homeostasis by constantly remodeling and controlling the composition of the plasma membrane (PM) in response to various extracellular and intracellular stimuli. Consequently, dysregulated CME has been extensively linked to disease (1, 2). CME is a multistep process involving 1) "priming," i.e., the regulated and localized activation of clathrin assembly proteins, predominantly adaptor protein 2 (AP2) complexes at the PM; 2) initiation of clathrin assemblies; 3) stabilization of clathrin-coated pits (CCPs) in the form of a macromolecular complex; 4) productive CCP growth and maturation, which culminates in 5) fission and the release of newly formed clathrin-coated vesicles (CCVs) into the cytosol (3, 4). Many nascent CCPs fail to complete this multistep process and instead rapidly disassemble as early or late abortive pits (5–7). In addition to the major coat proteins, clathrin and AP2, successful completion of CME requires the activities of a myriad of endocytic accessory proteins (EAPs). These EAPs, many of which are multidomain proteins, encode multiple biochemical activities, including curvature generation and sensing, cargo recruitment, scaffolding, and lipid modification (1, 4, 8).

The activities of EAPs, or of their individual functional domains, were largely identified through in vitro biochemical assays. The in vivo functions of EAPs in CME are frequently measured by cargo uptake, which scores the net accumulation of cargo inside cells, but lacks the temporal resolution and sensitivity to capture early steps or the regulation of CCP growth. Indeed, several publications showed that measurements of cargo uptake alone are unable to reveal alterations in the early kinetics of CCP maturation caused by the absence of one or several EAPs (9–11). An alternative approach to assess stage-specific EAP functions has been to measure the temporal hierarchy of their recruitment to CCPs. Using a pH-sensitive fluorescent cargo to mark scission events, Merrifield and colleagues (7) measured the recruitment profiles of 34 EAPs to CCPs with high temporal resolution, providing insight into their sequential roles in CME. The study also highlighted the nonuniform molecular composition of individual CCPs. However, its major limitation was that, prior to the advent of genome-editing technologies, fluorescently

Significance

Clathrin-mediated endocytosis (CME), the major pathway for uptake into cells, is a multistep process involving a myriad of endocytic accessory proteins (EAPs). Although the biochemical activities of many EAPs have been extensively studied, their stage-specific role(s) during clathrin-coated pit (CCP) initiation, stabilization, and/or maturation are poorly defined. Here, using quantitative total internal reflection fluorescence microscopy and a rigorous experimental and multiparametric analytical pipeline, we study the effect of CRISPR interference (CRISPRi)- or small interfering RNA (siRNA)-mediated knockdown of 67 individual EAPs on CCP dynamics. Our comprehensive analyses combined with unsupervised phenotypic clustering reveal the complex and overlapping roles of EAPs during early, critical stages of CME, providing a valuable resource to spur further research into their function.

Author contributions: M.B., R.E.M., S.L.S., and M.M. designed research; M.B., R.E.M., and M.M. performed research; M.B., R.E.M., X.W., H.M.G., and A.M.L. contributed new reagents/analytic tools; M.B., R.E.M., X.W., J.L., G.D., S.L.S., and M.M. analyzed data; M.B., R.E.M., S.L.S., and M.M. wrote the paper; X.W. and G.D. edited the paper; S.L.S. supervised the research; and M.M. prepared all figures.

Reviewers: P.D.C., HHMI, Yale University; and M.v.Z., University of California, San Francisco.

The authors declare no competing interest.

See Profile on page 31563.

[1]M.B., R.E.M., and M.M. contributed equally to this work.

[2]Present address: Chan Zuckerberg Biohub, San Francisco, CA 94158.

[3]To whom correspondence may be addressed. Email: sandra.schmid@czbiohub.org or marcel.mettlen@utsouthwestern.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2020346117/-/DCSupplemental.

First published November 30, 2020.

labeled EAPs were transiently overexpressed, which is especially problematic given the likely competition arising from widely shared protein interaction domains and binding motifs. Moreover, as individually tracked CCPs, referred to hereafter as intensity traces, were aligned to the terminal fission event, early molecular signatures of these EAPs at CCPs were missed, especially given the heterogeneity of CCP lifetimes (7, 12). Nonetheless, these and similar pioneering studies in yeast (13, 14) have led to the concept that the CME machinery is organized into functional modules that act sequentially, and in a stereotypic manner, during CCP maturation (1, 4).

It has become increasingly evident that CME and cell surface receptor signaling are reciprocally regulated by feedback loops (15–20). This has led to the understanding that CME is not a passive process, but that it can respond and adapt to multiple inputs. Moreover, consistent with the essential role of CME in cellular physiology, the process is robust and exhibits plasticity, in that compensatory mechanisms can restore CME even when individual stages of CCP maturation are significantly perturbed (9, 19, 21). This robustness and plasticity likely derive from the overlapping functions, and thus redundancies, of EAPs regulating clathrin assembly and CCP maturation.

Although the individual activities of many EAPs have been extensively studied, many controversies exist as to how, and at which stage(s), these activities contribute to the overall process of CME (22–29). In part, these controversies may reflect the wide range of cell types, experimental systems, and assays used to study EAP function. A more comprehensive analysis under identical experimental conditions has never been undertaken and could lead to a better understanding of the functional hierarchy and the role played by each EAP during CME. Such a systematic analysis requires a readout that directly measures discrete early stages of CCP nucleation, initiation, and maturation. Confocal and total internal reflection fluorescence microscopy (TIR-FM)-based live-cell imaging (5, 6, 30, 31) paired with unbiased and high-content image analyses (6, 9, 11, 32) are essential for the detection of these alterations in early regulatory stages of CCV formation.

Equipped with live cell TIR-FM and computer vision tools to quantify several stage-specific parameters of CME, including rates of CCP initiation, stabilization, and maturation (9, 12, 33), we have quantified the knockdown (kd) effect (i.e., phenotype) of most known or suspected EAPs on CCP dynamics within a uniform and rigorous experimental framework. Based on previous studies and their biochemically defined activities, the 67 proteins studied can be assigned to functionally distinct modules (1, 4, 8). Here, we have clustered EAPs based on their phenotypic signatures. Interestingly, these phenotypically defined clusters do not overlap with the biochemically defined modules. Our results highlight the functional complexity of protein–protein interactions and EAP activity during CCV formation. The overlapping activities and hence functional redundancies of multidomain EAPs provide a mechanistic basis for the robustness of CME and the

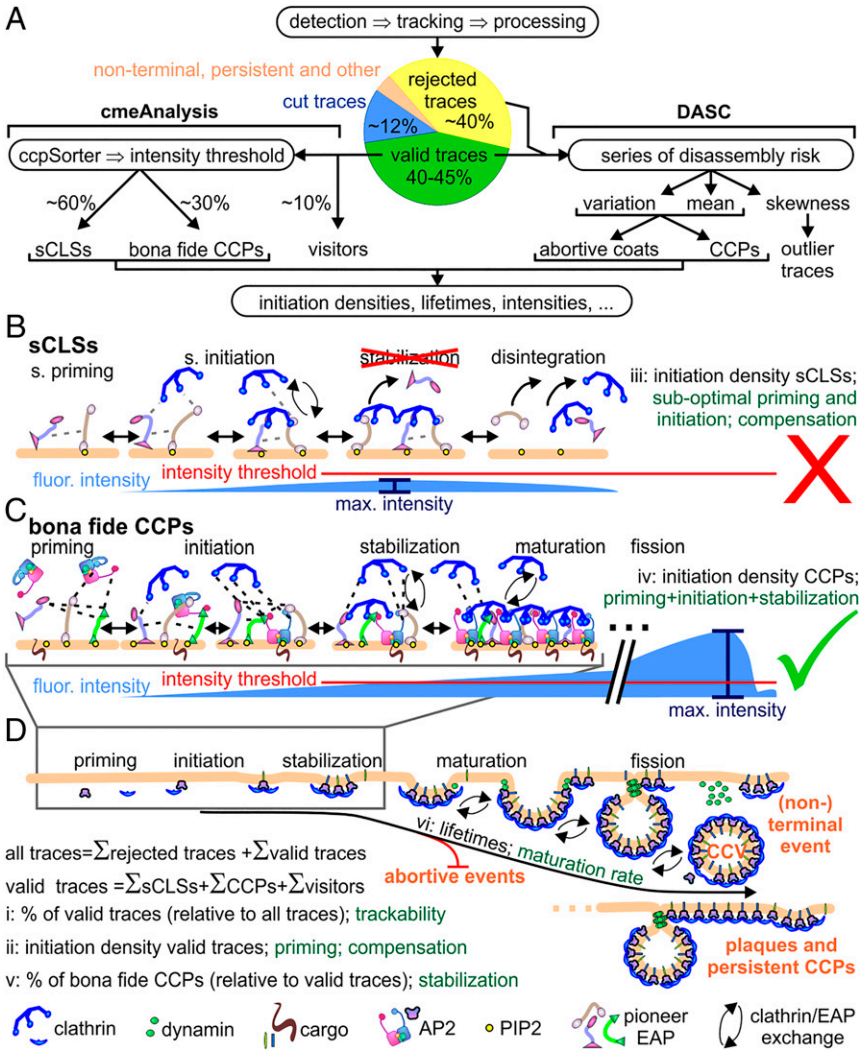

Fig. 1. Schematic representation of workflow and stages leading to CCV formation identified by cmeAnalysis. (*A*) Image series were analyzed via the cmeAnalysis or DASC processing pipelines. Clathrin-coated structures were detected, then tracked and processed. Only valid traces were partitioned by the cmeAnalysis pipeline into sCLSs and bona fide CCPs, based on a user-defined intensity threshold. In contrast, DASC (disassembly asymmetry score classification) takes into account both valid and rejected traces, as defined by cmeAnalysis, to measure total initiation rates of all clathrin-coated structures, and then, from valid traces only, classifies ACs, CCPs, and outlier traces in an intensity threshold-independent manner, instead relying on intensity fluctuations during CCP assembly and maturation (11). (*B*) sCLSs are small clathrin assemblies that never grow beyond a user-defined intensity threshold, because the partial clathrin coats fail to be stabilized and thus rapidly disassemble. (*C*) Nascent clathrin structures that are stabilized at the PM by a multitude of weak protein–protein interactions are defined as bona fide CCPs that can complete their maturation and internalize cargo as productive CCPs. (*D*) Maturation of bona fide CCPs involves their invagination and finally dynamin-catalyzed fission to release cargo-laden clathrin-coated vesicles (CCVs). A subset of stabilized coats also fail to mature and disassembled as abortive pits. Six parameters quantified by cmeAnalysis (i–vi) are indicated in black, and the CCP behaviors they measure are indicated in green.

ability of cells to counter defects in CCP maturation by activation of compensatory mechanisms.

Results

An Experimental Pipeline to Define the Stage-Specific Functions of EAPs. Normal diploid human retinal pigmented epithelial (ARPE-19) cells that stably express eGFP-CLCa as a fiduciary marker to visualize CME were chosen for this study as they are ideally suited for TIR-FM and image-based analysis of CCP dynamics (9, 11). For CRISPR interference (CRISPRi)-mediated kd, dCas9-BFP-KRAB (dCas9 fused to a Krüppel-associated box, KRAB, transcriptional repressor domain and BFP for selection) was stably introduced into these cells. For each EAP target gene, two sequence-specific single-guide RNAs (sgRNAs) (see *Materials and Methods* for details of cell line generation and sgRNA selection) were cloned into lentiviral vectors together with the selectable marker for puromycin resistance (*SI Appendix*, Fig. S1*A*).

To ensure optimal kd efficiency and consistency in treatment, which we confirmed with a subset of EAPs by western blotting (*SI Appendix*, Fig. S2), a strict experimental timeline was followed (*SI Appendix*, Fig. S1*B*). Each experiment included cells transduced with viruses encoding scrambled sgRNA, which served as a positive control for the efficacy of the lentiviral transduction (i.e., a high survival rate of cells after puromycin selection when compared to noninfected cells) and as the control condition and reference for our analysis pipeline. Cells were also infected with virus encoding sgRNA against CALM (clathrin assembly lymphoid myeloid leukemia) because the consistently high efficacy of CALM kd (*SI Appendix*, Fig. S3 *A* and *B*) resulted in a readily identifiable change in cell morphology, i.e., the presence of broad lamellipodia (*SI Appendix*, Fig. S3*C*) and strong effects on CCP dynamics (*SI Appendix*, Fig. S3 *D*–*F*), as previously observed (11). CALM kd thus served as a visual and functional indicator of the CRISPRi activity in our cells. Three days post lentiviral transduction and puromycin selection, cells from each condition were split and seeded on two gelatin-coated cover glasses. Thus, every imaging session included a negative control (sgScrambled), a positive control (sgCALM), and sgRNAs against up to three different EAPs. The negative controls and experimental conditions were imaged in duplicates to average possible variations in cell seeding, attachment, and cover glass coating. To reduce the computational load, each large field of view (up to 130 μm × 120 μm at 108× final magnification) was split in two, yielding a total of 22 to 24 movies per condition (*Materials and Methods*). Each movie contained two to five cells, and at least 250,000 valid traces were analyzed per condition.

After acquisition, the raw data were transferred to a high-performance computational cluster and analyzed using the previously published cmeAnalysis pipeline (9). Briefly, cmeAnalysis comprises three modules (Fig. 1*A*). The detection module uses a model-based particle detector to distinguish CCPs from image noise, making it extremely sensitive and reliable in detecting true clathrin assemblies. The tracking module then links images of CCPs between consecutive frames in the time series (32). Given that CCPs continuously appear and disappear, and the possibility that CCP signals, especially at early stages, may not be strong enough to be robustly detected in every single frame, the tracker uses crucially important "gap closing" algorithms to identify broken intensity traces. Finally, the processing module applies self-learned data quality standards to all traces (6, 32, 33) to identify "valid" traces, which have a minimum length of five frames (i.e., 5 s) and a minimum of at least three detections above noise within these five frames (9). Traces are deemed "invalid" if 1) the underlying detected particles are larger than diffraction limited, 2) they have too many consecutive gaps later in their lifetime, 3) they merge or split (some of the latter could correspond to nonterminal events) (34), 4) they are "cut" at the beginning or end of the movie, and/or 5) they exhibit exceptional fluctuations in background intensity. A separate group of traces, which are either diffraction-limited or larger, are classified as persistent, if they are present throughout the entirety of the 7.5-min movie. Closer inspection of these persistent structures reveals that most of these CCPs are highly dynamic with large intensity fluctuations, indicating that they support CME by multiple CCV fission events either from large, preexisting structures or unresolved adjacent CCPs (*SI Appendix*, Fig. S4 *A* and *B*). Although these CCPs stand out to the observer due to their larger size and temporal and spatial persistence, they represent only a small fraction (<2.5%) of all traces in ARPE cells (*SI Appendix*, Fig. S4 *C* and *D*). Given their dynamic characteristics (e.g., no quantifiable lifetime, multiple CCV fission events, merging, splitting), they cannot be accurately quantified by our pipeline and were therefore not included in the analyses. By analyzing each individual trace, the processing module yields thousands of valid intensity traces per cell (6, 9, 32). Next, cmeAnalysis isolates so-called visitors from all valid traces (an invariant ∼10%), which correspond to structures that are bright, yet short-lived, and are thus most likely preformed clathrin-coated structures that travel through the TIR-FM field (6, 9). The remaining valid traces contain a large number of very dim clathrin assemblies (e.g., partially formed coats or even single triskelia carrying three eGFP-labeled clathrin light chains) that are nonetheless confidently captured by our highly sensitive detection method. Due to the sheer number of these short-lived assemblies, their dynamics dominate the analysis results. To distinguish these structures from larger, more stable clathrin assemblies, cmeAnalysis applies a user-defined intensity threshold (*Materials and Methods*) that splits valid traces based on the maximum intensity they reach into 1) subthreshold clathrin-labeled structures (sCLSs), which are small clathrin assemblies that fail to grow beyond the intensity threshold and are rapidly turned over, and 2) bona fide CCPs that can mature and yield CCVs (Fig. 1). Importantly, as sCLSs must remain at the PM for ≥5 s and given that clathrin does not interact directly with the PM, we ascribe these events to failed attempts at CCP nucleation.

Changes in CCP Dynamics Yield Important Insights into the Stage-Specific Regulation of CME. The effects of EAP kd on CCP dynamics were assessed using six main readouts (Fig. 1 *B*–*D* and *SI Appendix*, Figs. S5–S7). Several additional parameters were quantified, including initiation density of all traces, CCP maximum intensities, and percent persistent structures (see Datasets S1 and S2 for raw data and percent difference, respectively, after EAP kd for all nine parameters measured). Together, the effects of EAP kd on each of these readouts informs us as to their stage-specific function(s). Thus, as illustrated in Fig. 1 *B*–*D*, the percentage of valid traces relative to all detected traces (i), which typically falls between 40 and 45%, controls for consistency between experiments and is indicative of global changes in CCP behaviors that might affect our ability to accurately track CCPs. For example, large changes in this parameter could indicate a switch of CME from diffraction-limited terminal to more nonterminal or plaque-based events. We also measured, as events per second per square micrometer of cell surface area, the rate of initiation of valid traces (ii) and, after thresholding to identify sCLSs and CCPs, the initiation rates of sCLSs (iii) and bona fide CCPs (iv). These initiation rates correspond to the nucleation of clathrin assembly and reflect the activities of clathrin assembly proteins, the most abundant and predominant being AP2. AP2 interacts with the PM via PM-enriched phosphatidylinositol-4,5-bisphosphate (PIP2) and/or sorting motifs in the cytoplasmic domains of cargo molecules (35). To successfully bridge between clathrin and the PM, AP2 must undergo a series of activating conformational changes that regulate CCP initiation (36–38). EAPs that affect these factors (i.e., PIP2 concentration, cargo recruitment, and/or AP2 activation) and thus the "priming" of sites for clathrin assembly would alter the rates of initiation of sCLSs,

CCPs, or both. Interestingly, transient sCLSs make up the majority (60 to 65%) of valid traces, indicating that most clathrin initiation/nucleation events are unsuccessful (herein referred as "suboptimal" [s.] priming and s. initiation; Fig. 1*B*). A selective decrease in the initiation rates of bona fide CCPs would indicate defects in CCP stabilization subsequent to nucleation. Together, these measurements capture early events in CCP priming, nucleation, and stabilization.

Later events in CCP maturation are reflected in (v) the percentage of bona fide CCPs (Fig. 1*C*), which measures (vi) the efficiency of stabilization/maturation and the mean lifetime of bona fide CCPs (i.e., the average length, in seconds, of CCP intensity traces, Fig. 1*D*). CCP lifetimes typically exhibit a broad, Rayleigh-like distribution peaking near ∼30 s (9). Changes in the shape of this distribution curve reflect changes in both the rate and efficiency of CCP maturation (9, 29, 39). In the absence of compensatory mechanisms, the initiation density of bona fide CCPs, their lifetime, and percentage are expected to directly relate to the efficiency of CME.

Because cmeAnalysis relies on a user-defined threshold based on control cells, changes in the relative numbers of sCLSs and CCPs could indicate global changes in CCP intensities rather than, or in addition to, stage-specific effects. Therefore, we recently introduced a complementary, threshold-independent analytic approach. This pipeline, termed DASC (disassembly asymmetry score classification) analysis (11), uses the detection algorithms of cmeAnalysis but considers both valid and rejected traces when determining the initiation rates of clathrin structures (Fig. 1*A* and *SI Appendix*, Fig. S4*C*). Then, analyzing only valid traces, DASC classifies traces into either abortive coats (ACs) or bona fide CCPs based on frame-to-frame intensity fluctuation and disassembly relation among all of the traces (11). Despite significant overlap in the lifetimes and intensities of ACs and CCPs, these functionally and structurally distinct subpopulations are accurately resolved by DASC (11). Together, cmeAnalysis and DASC faithfully capture early events in CCP nucleation, stabilization, and maturation. The DASC analysis was applied to a subset of EAP kd conditions. In most cases, changes in the %CCPs determined by cmeAnalysis correlated well with those determined by DASC (*SI Appendix*, Fig. S4*E*). As expected, the outliers corresponded to EAPs that altered the average intensities of CCPs (*SI Appendix*, Fig. S7*B*).

Phenotypic Clustering of EAPs. With these parameters in hand, we then measured the kd effects of 67 EAPs on the dynamic behavior of CCPs. Prior to running cmeAnalysis, the intensity threshold was determined in control cells so that 30 to 35% of all traces were classified as bona fide CCPs, yielding the typical Rayleigh-shaped lifetime distributions, seen in multiple cell lines (5, 9, 18). Then, the same intensity threshold was applied to all experimental conditions captured the same day. To normalize for slight experimental day-to-day variations (*SI Appendix*, Fig. S3 *E* and *F*), the effects of EAP kd were expressed as relative percent differences (%Δ) with respect to the corresponding sgScrambled control for each parameter measured (Dataset S2). We first confirmed that none of the kd conditions led to a strong decrease, i.e., negative %Δ, in the percentage of valid traces (*SI Appendix*, Fig. S5*A*) and therefore did not per se affect the trackability of CCP traces. As will be discussed later, kd of a small subset of EAPs resulted in an increase of valid traces.

The quantitative nature of our multiparametric analysis allowed us to group all studied EAPs into phenotypically similar clusters using unsupervised clustering. The quality of this approach is vulnerable to the number of clusters (k) chosen, which is generally not trivial; therefore, we applied gap statistics (40) to compare the total intracluster variation for different k values to their expected values under the null reference distribution generated by a Monte Carlo sampling process. Using the globalSEmax method (40), a maximum gap statistic for $k = 10$ was determined. Then, *K*-means, the most common clustering algorithm, was used for partitioning all studied EAPs into 10 phenotypic clusters (Fig. 2 *A* and *B*). To validate the *K*-means clustering results by an independent clustering method, affinity propagation clustering (APC) was also applied to the screen data. APC does not require a predetermined k, but rather creates pairwise phenotype similarities based on negative distances (41). As for the gap statistic/globalSEmax method, APC also generated 10 clusters. Not only were the numbers of clusters determined by these two independent methods identical, but the grouping of all 67 EAPs into these clusters was closely matched. Hence, an Alluvial diagram comparing the two methods has only a few and thin crossing stream fields (Fig. 2*A*). The majority of these differences were observed with proteins whose kd effects were weaker (clusters 6 to 10) resulting in an ambiguous phenotypic classification. These included DBNL1 (also known as Hip-55 or mAbp1), the human ortholog of yeast Abp1, Snx9 (whose kd efficiency using CRISPRi was weak; *SI Appendix*, Fig. S2), and PIK3CA. The fourth protein was Eps15, which moved from *K*-means cluster 3 to APC cluster 5. These two clusters differed quantitatively, but not qualitatively in their phenotypic signatures.

Based on the coherence between these two independent clustering methods, we have classified EAP effects into 10 groups. However, given the limitations of comprehensive screens and the fact that several of these groups exhibit only mild phenotypes, we cannot (and do not) infer any significance to the exact number of clusters. Nonetheless, we pooled all conditions cluster-by-cluster and applied a nonparametric, two-sample permutation test for multivariate datasets (42) to obtain the statistical significance of the overall phenotype in each cluster and found that P values in the 10 clusters are all ≤ 0.0001, indicating that the overall phenotype in every cluster is significant. Thus, the observed phenotypic signatures provide insights into the complexity of EAP function during CME, as discussed below.

Phenotypic Signatures of EAP Clusters. The heatmap in Fig. 2*B* depicts the %Δ for each kd relative to its control and illustrates the specific phenotypic signature of each cluster. Clusters 8 to 10 have uniformly weak or no detectable phenotypes, which could reflect either that 1) these proteins do not play a significant role in CME, at least in these cells and under our experimental conditions (most of the actin-related and early endosomal proteins fall into these groups); 2) there exists functional redundancy between closely related isoforms (e.g., Itsn1, NECAP1, and Fcho2); or 3) kd was inefficient. In cases where a phenotype was expected based on previous publications, and/or poor kd efficacy was confirmed by western blotting, we repeated our assays using small interfering RNA (siRNA)-mediated kd (indicated by an asterisk in Fig. 2 *B* and *C*). Note that although kd of SNX9 was inefficient (*SI Appendix*, Fig. S2), its function has been extensively studied using these techniques (29, 43) and therefore was not pursued here. The following describes the phenotypic signatures of clusters 1 to 7 in general terms. Insights gained as to the roles played by specific proteins will be described in more detail later.

Clusters 1 through 5 are characterized by a defect in CCP initiation and stabilization, as indicated by the decreased initiation rate and percentage of bona fide CCPs (Fig. 2*B*, columns iv and v, respectively). However, these clusters differ in their effects on the rates of sCLS initiation/nucleation (column iii). In the absence of cluster 1 and 2 EAPs, the rate of sCLSs initiations increased indicative of an increase in suboptimal priming events. The increased rate of sCLSs initiation could be due to rapid, localized turnover of coat proteins released during unsuccessful stabilization attempts, or compensatory mechanisms induced in an effort to restore CME. In contrast, in the absence of cluster 3 EAPs, the rate of sCLSs initiations decreased. AP2 is among the cluster 1 proteins, suggesting that EAPs in cluster 3 might function upstream of AP2 and/or together with other adaptors

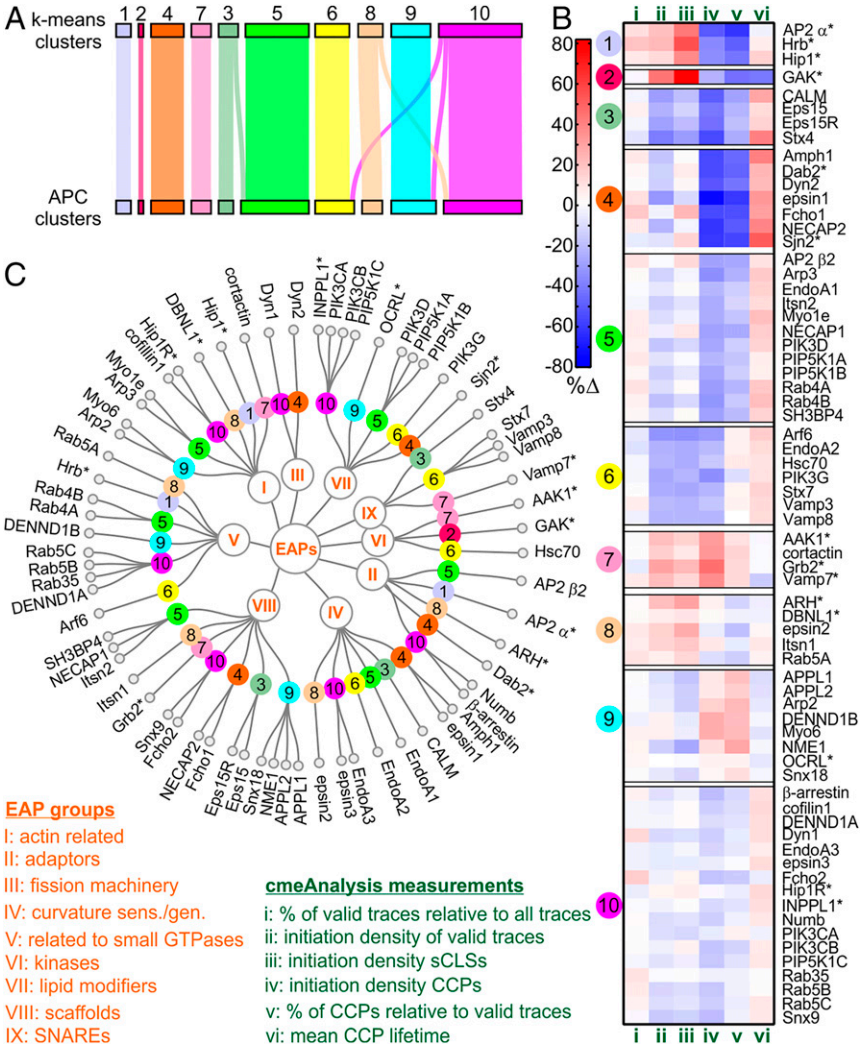

Fig. 2. Cluster analysis and summary of EAP phenotypes. (*A*) Two independent clustering methods (*k*-means and APC) identified 10 phenotypic clusters. Alluvial plot representing the compositional changes in the clusters determined either by the *k*-means or the APC clustering method (*Materials and Methods*). The width of the blocks represents the size of the cluster, and the width of a stream field connecting two blocks represents the number of components differentially assigned by the two methods. As indicated, *k*-means and APC clusters are very similar in composition. (*B*) Heatmap reporting percent differences (%Δ) in the six indicated parameters derived from cmeAnalysis measurements of all studied EAPs, grouped in clusters obtained by the *k*-means clustering method, visualizing the phenotypic signature of each protein. (*C*) Circular dendrogram visualizing the partitioning of each functional EAP group (I–IX) into the various clusters determined by the *k*-means clustering method. EAPs denoted with an asterisk (*) were depleted using siRNA.

and assembly proteins to initiate clathrin recruitment. The lack of a substantial effect of cluster 4 and 5 proteins on sCLS initiation rates suggests that they act downstream of successful nucleation to stabilize nascent CCPs. Cluster 3 to 5 EAPs are further distinguished from clusters 1 and 2 in that the mean lifetimes of CCP are increased (Fig. 2*B*, column vi). The observation that their kd slows the rate of CCP maturation suggests that these proteins continue to function at later stages of CME.

The kd of proteins in cluster 6 proportionally reduce the initiation densities of both sCLSs and CCPs (Fig. 2*B*, columns iii and iv, respectively). Given that the ratio of these two populations (i.e., the %CCPs) does not change (Fig. 2*B*, column v), these effects indicate that cluster 6 EAPs likely function exclusively during initiation/nucleation rather than CCP stabilization. Surprisingly, among these proteins is Hsc70, well characterized for its function during CCV uncoating (44, 45). How precisely Hsc70 affects the early stages of CME remains to be elucidated; however, Hsc70 has been suggested to play a role in chaperoning cytosolic clathrin, and hence may play an early role in regulating clathrin assembly (46).

In contrast to all other clusters, kd of proteins in cluster 7 mainly increased, albeit to differing degrees, the initiation rates of sCLSs (Fig. 2*B*, column iii) and to a greater extent those of bona fide CCPs (column iv), resulting in an increase in the percent of CCPs (column v). This phenotypic signature, which is more difficult to interpret, could reflect activation of compensatory mechanisms that restore CME in perturbed cells (9, 19) or that these factors act as negative regulators of CCP initiation and stabilization.

A small subset of EAPs increase the fraction of valid traces relative to all traces (*SI Appendix*, Fig. S5*A*). Strikingly, the α and β subunits of AP2 fall into this group, as do the pioneer EAPs Fcho1/2. Others include Hrb, which clusters with AP2α; epsin1, which clusters with Fcho1; and dynamin-1, which clusters with Fcho2. Given the criteria for determining valid traces (see above), this result could indicate the induction of compensatory mechanisms to restore CME, changes in the early rates of clathrin recruitment, and/or changes in the relative numbers of terminal and nonterminal events. Further analyses would be necessary to determine which of the criteria used to validate traces (i.e., early fluctuations in intensity, CCP splitting and merging events, etc.) are altered before interpreting this observation.

Biochemically Defined Modules Span Multiple Phenotypic Clusters. EAPs can be roughly divided into nine functionally related modules (1, 47, 48) (Fig. 2*C*): (I) regulators of the actin cytoskeleton, (II) adaptor proteins, (III) fission machinery, (IV) curvature sensors and generators, (V) small GTPases and their regulators, (VI) protein kinases and ATPases, (VII) lipid kinases and phosphatases, (VIII) scaffolds, and (IX) SNAREs. These groupings are mainly based on in vitro biochemical assays that define protein activities and are thus closely related to the domain structure of

each individual EAP (1, 2, 49). However, many EAPs encode several functionally distinct domains. Thus, grouping proteins based on the biochemical activity of any single domain may not be predictive of the protein's full functionality in vivo. For instance, epsin1 (Fig. 2*C*, group IV—generally considered a curvature sensor/generator) contains an epsin N-terminal homology (ENTH) domain that binds to PIP2 and can indeed induce membrane curvature (27). However, in addition, epsin1 bears several motifs that bind to the clathrin heavy chain (50) and a multitude of AP2 binding sites (51), which can trigger clathrin assembly onto, and CCV formation from, liposomes (52). Epsin1 also encodes two ubiquitin interacting motifs (UIMs) and can function as a cargo adaptor (28, 53). Given the functional versatility of its domains, epsin1—as for many other EAPs discussed below—cannot be unambiguously assigned to any single functional group.

Given the flexibility and modularity of biochemical activities, which are further complicated by temporal aspects of protein recruitment to CCPs, it is not surprising that the biochemically defined protein groups do not overlap with the phenotypic clusters observed here. Rather, EAPs belonging to the biochemically defined groups split up into the various phenotypically defined clusters (Fig. 2*C*), and thus each phenotypic cluster contains a mixture of proteins encoding diverse activities. Together, these observations underline the concept that the activities of individual domains cannot account for the phenotypic signature of a given EAP, but rather the complex interplay between the various domain activities of an EAP determines its function(s) during CCP maturation. Due to the number of EAPs studied, hereafter we describe only those EAPs whose depletion resulted in the strongest phenotypes.

Cluster 1 EAPs Function in CCP Initiation/Nucleation and Stabilization. Upon depletion of cluster 1 proteins (Fig. 3), the rate of initiation of sCLSs increased (Fig. 3 *A* and *B*), indicative of an increase in failed, suboptimal nucleation events. Correspondingly, the rate of bona fide CCP initiation and the percentage of stabilized CCPs decreased (Fig. 3 *A*, *C*, and *D*). Together, these data suggest that cluster 1 proteins are critical factors in the early stages of CCP initiation/nucleation. Not unexpectedly, among the cluster 1 proteins is the AP2 complex that is well known to nucleate clathrin assembly (21, 24). Its α subunit has two isoforms encoded by separate genes (AP2A1 and AP2A2). Although highly homologous, α1 and α2 are most divergent in their unstructured hinge and appendage domains (*SI Appendix*, Fig. S8*A*), suggesting possible differences in their interactions with EAPs. In our initial screen, we attempted to study the effect of kd of these individual α isoforms by using specific sgRNAs. The kd of α2 resulted in a decrease in CCP initiations, but kd of α1 had no significant effect on any of the parameters tested (*SI Appendix*, Fig. S8 *B–F*). This was surprising, considering that AP2 α1 is expressed at 1.77 ± 0.16-fold higher levels than AP2 α2 in these cells, based on relative abundance of unique peptides as determined by mass spectrometry. As antibodies that distinguish these two isoforms are not available, we could not assess either the efficacy of individual subunit kd or whether kd of one subunit leads to a compensatory up-regulation of the other. For these reasons, we chose to knock down both isoforms by using published siRNA sequences.

As expected, efficient AP2 α kd reduced the initiation rates and numbers of bona fide CCPs (Fig. 2*B*). The kd of the β2-adaptin subunit of AP2 phenocopied that of the α-adaptin subunit, albeit to a lesser extent, which could indicate a less efficient kd of β2 or partial redundancy with the β1-adaptin subunit of AP1 (54). The increased rate of sCLS initiation could also reflect compensatory activities of other EAPs capable of triggering clathrin assembly. Consistent with early functions in CME, neither of the cluster 1 proteins substantially altered the lifetime distribution or mean lifetimes of CCPs (Fig. 3 *E* and *F*, respectively). However, kd of the α (as well as the μ2 subunits; *SI Appendix*, Fig. S8*F*) of AP2 seemed to tighten the distribution (Fig. 3*E*), which peaked more sharply at ∼40 s, indicative of a more homogeneous/synchronous maturation process.

The other two EAPs identified in cluster 1, Hrb (HIV Rev-binding protein) and Hip1 (Huntingtin-interacting protein 1), had not previously been identified as CCP initiation/nucleation factors, but their membership in this cluster can be rationalized based on previous studies. Hrb, also known as AGFG1 (Arf-GAP domain and FG repeating-containing protein), has been shown to bind Eps15 (55), AP2 (56), and SNAREs (57, 58). The latter interaction is mediated by Hrb's binding to longin domains, found in a subset of SNAREs, e.g., VAMP7. This interaction requires Vamp7 to be in an open conformation, which is stabilized when bound to a partner Q-SNARE (58). Thus, Hrb will mediate inclusion of a primed QR heterodimeric SNARE complex into CCVs. Hrb also encodes an ArfGAP domain, although its GAP activity has not yet been demonstrated. Given that ArfGAPs specifically recognize the activated GTP-bound form of an Arf, Hrb might also function as an Arf effector that is recruited to the PM by an as-yet-unidentified, activated Arf to enhance CCP nucleation through its interactions with AP2 and Eps15. Both these properties, i.e., inclusion of SNARE proteins and interactions with Arf GTPases, are key to the stable assembly of COPI and COPII coats (59, 60), and thus might also apply to CCV formation.

Hip1, the third EAP identified in cluster 1, is classified as an actin regulator based on the C-terminal talin-like domain it shares with Hip1R and yeast Sla2p that both bind F-actin. However, Hip1 binds actin with much lower affinity than its close relative Hip1R, and its ability to bind actin remains controversial (61–63). Both proteins also encode potentially curvature-generating N-terminal ANTH domains that interact with PIP2 at the PM, as well as a coiled-coiled domain that supports multiple protein–protein interactions including homodimerization. In addition, Hip1 (61, 64) encodes additional binding sites for clathrin heavy and light chains, as well as for AP2. Consistent with its assignment to cluster 1, Hip1 functions synergistically with AP2 to promote clathrin assembly in vitro (61, 64, 65). Furthermore, it has been reported that short-lived CCPs, which are presumably abortive, fail to recruit Hip1, suggesting an early role in nucleating and stabilizing nascent CCPs (66). Given that Hip1 binds to the N terminus of CLC, which is known to negatively regulate clathrin assembly (61, 67), we speculate that Hip1–CLC interactions might play a role in enhancing the rate of clathrin assembly, which has been shown to correlate with the stabilization of nascent CCPs (11, 12).

GAK, a Multifunctional Protein with a Unique Phenotypic Signature. GAK (Cyclin-G-associated kinase, also known as auxilin2), the lone member of cluster 2 (Fig. 2*B*), was phenotypically unique among the EAPs we tested, perhaps reflecting its hybrid nature. GAK is homologous to both neuronally enriched auxilin and AAK1 (adaptor associated kinase 1) (68, 69). Like auxilin, GAK encodes a DNAJ domain that functions together with Hsc70 in catalyzing clathrin uncoating, as well as a lipid binding PTEN domain. GAK also encodes a kinase domain, which like its homolog AAK1, can phosphorylate μ2 to stabilize activated AP2 complexes for cargo sorting (70, 71). In addition, GAK encodes multiple clathrin binding sites that can mediate clathrin assembly in vitro (72). In a previous study, kd of GAK also stood out by producing a strong phenotype on both late abortive and productive CCPs, and uniquely altering the lifetime distribution of CCPs (73).

In our screen, GAK depletion had a strong phenotypic signature (Fig. 2*B*). Interestingly, there was a marked increase in the initiation density of all valid traces (Fig. 2*B*, column ii, and *SI Appendix*, Fig. S6*A*), due to an almost doubling of the rate of initiation of sCLSs (Fig. 2*B*, column iii and *SI Appendix*, Fig. S6*B*). Given the atypical phenotype observed for GAK in our screen, we repeated this experiment under the more stringent imaging conditions conducive to DASC analysis (11) (*Materials*

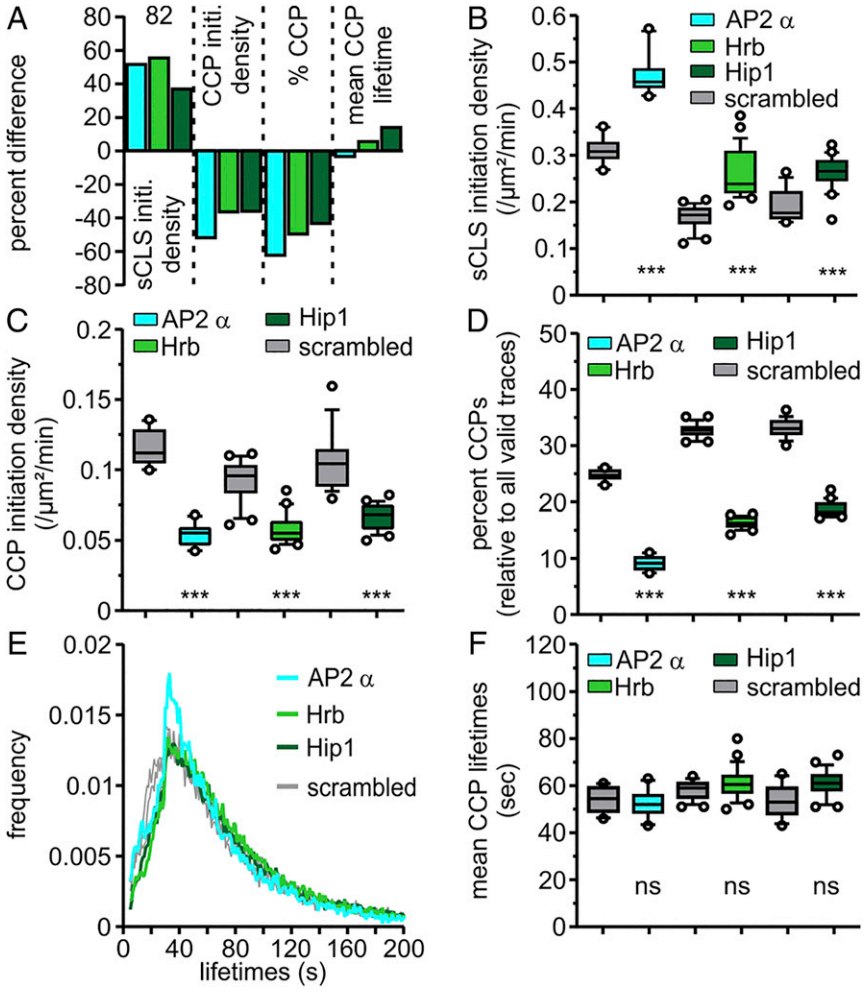

Fig. 3. Effect of cluster 1 EAPs on nucleation and initiation of CCP growth (*A*). Percent difference plot of initiation density of sCLSs, initiation density of bona fide CCPs, percent CCPs, and mean lifetimes of CCPs. Results are expressed as %Δ relative to the experimental control. Initiation densities of (*B*) sCLSs and (*C*) bona fide CCPs. (*D*) Percentage of CCPs relative to all valid traces. (*E*) Lifetime distribution of bona fide CCPs. (*F*) Mean lifetimes of bona fide CCPs. The box-and-whiskers plots in this and all subsequent figures show median and 10th to 90th percentiles. Individual circles correspond to outlier movies ($n \geq 22$ per condition). In this and all subsequent figures, ordinary one-way ANOVA was used to compare control with kd, each performed on the same day. *P* values: ***$P < 0.001$. Bona fide CCP analyzed: controls for AP2 α, Hip1, Hrb: 43,866, 59,824, 86,929, respectively; AP2 α: 23,549; Hip1: 52,047; Hrb: 56,474.

and Methods), using single- and double-round siRNA treatments (Fig. 4 and *SI Appendix*, Fig. S9), which resulted in >90% and >95% depletion, respectively (*SI Appendix*, Fig. S9*A*), and dose-dependent effects on CCP dynamics (*SI Appendix*, Fig. S9 *B–E*). At the higher kd efficiency, both cmeAnalysis and DASC confirmed the results obtained in our screen. Thus, the rate of initiation of sCLSs strongly increased, leading to an increase of the initiation density of all clathrin structures (Fig. 4 *A* and *D*). Concomitantly, the initiation rates of bona fide CCPs decreased (Fig. 4 *B* and *E*), leading to a highly significant decrease in the percentage of bona fide CCPs (Fig. 4 *C* and *F*). Furthermore, highly efficient GAK kd reproducibly altered the lifetime distribution of CCPs shifting the normal Rayleigh-like distribution to a quasi-exponentially decreasing lifetime distribution (Fig. 4*G*), which in the past has been interpreted as an increase in abortive events (19). The lifetime distribution of DASC-identified CCPs, in which early abortive APs/ACs are computationally identified and removed, confirms this interpretation (Fig. 4*H*), while its leftward shift also suggests an increase in late abortive events. Combined, these findings point to critical roles for GAK during the entire process of CCV formation—CCP initiation, stabilization, and maturation—consistent with reported defects in the clathrin-mediated internalization of transferrin upon GAK depletion (46, 71), which we reproduce here (*SI Appendix*, Fig. S9 *F* and *G*).

Pioneer Proteins Partition into Two Clusters Suggesting Sequential and Overlapping Functions. In addition to clathrin, adaptors and cargo, the assembly and stabilization of nascent CCPs requires the function of so-called pioneer EAPs (11, 23, 24, 74) (Fig. 1 *B–D*). With their multiple protein–protein and protein–lipid binding sites, these pioneer EAPs support recruitment, activation, and/or binding of clathrin, AP2, and other EAPs onto the PM (74). As we have recently studied the effect of pioneer EAP kd on CCP dynamics using these approaches (11), they will not be discussed in detail here. However, we note that most of the known or suspected pioneer proteins are found in clusters 3, 4, and 5 (Fig. 2*B*). The main distinguishing feature of cluster 3 and 4 EAPs was their effects on the initiation rate of sCLSs: kd of cluster 3 EAPs (e.g., Eps15/Eps15R) decreased the rates of sCLS initiation (Fig. 2*B*, column iii), whereas cluster 4 EAPs (e.g., Fcho1, NECAP 2) did not. We interpret this difference to reflect an earlier function of cluster 3 EAPs in CCP nucleation, relative to cluster 4 EAPs, whose kd primarily affects stabilization of nascent CCPs (Fig. 2*B*, column iv). The kd of cluster 4 EAPs also resulted in more pronounced effects on the median lifetime of CCPs (Fig. 2*B*, column vi), suggesting a sustained role for these proteins in determining the rates of CCP maturation. The phenotypic signature of cluster 5 EAPs qualitatively mirrors that of cluster 4 but was weaker, likely reflecting functional redundancies among isoforms (e.g., NECAP 1/2; Itsn 1/2).

A Role for SNARE Proteins and Their Adaptors during Early Stages of CCP Formation. Nascent CCVs undergo multiple rounds of homotypic fusion, as well as heterotypic fusion with early endosomes that allow for cargo sorting (75). Thus, the ultimate functionality of CCVs requires the incorporation of a sufficient

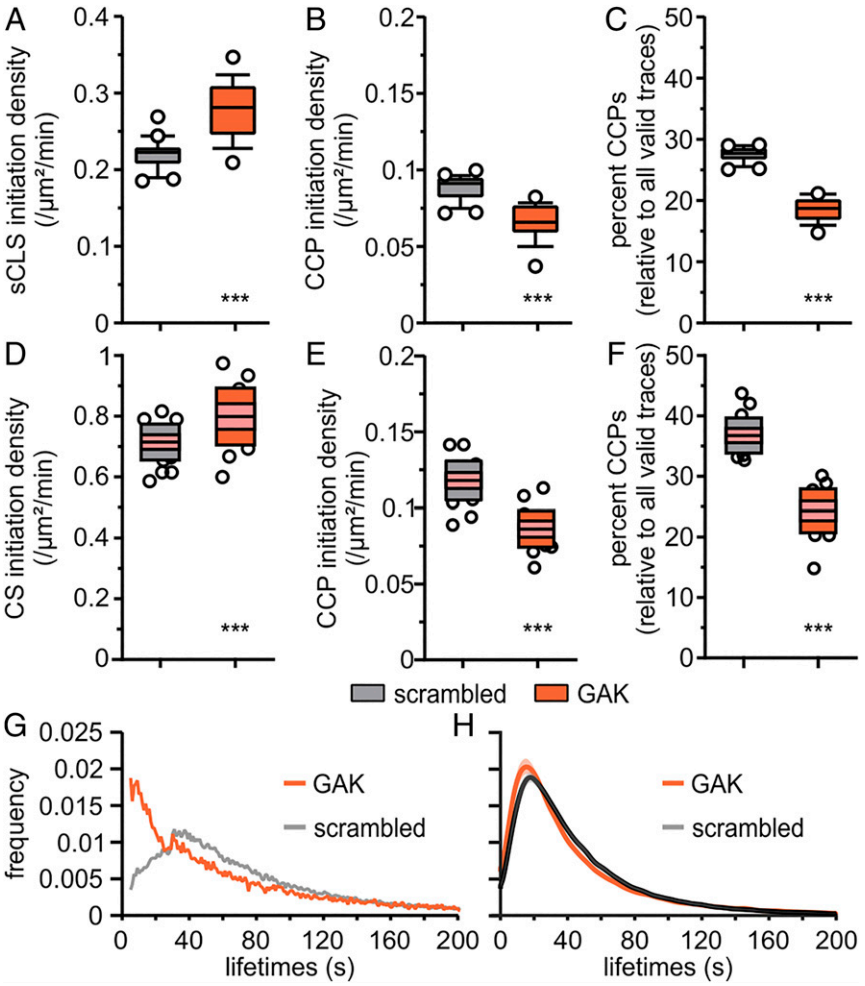

Fig. 4. Effect of GAK kd on CCP dynamics measured by cmeAnalysis (*A*, *C*, and *G*) and DASC (*D*, *F*, and *H*). Initiation densities of (*A*) sCLSs and (*B* and *E*) bona fide CCPs. (*D*) Initiation density of clathrin structures calculated by DASC, considering valid traces and others. (*C* and *F*) Percentage of CCPs relative to all valid traces. (*G*) Lifetime distribution of bona fide (i.e., superthreshold) CCPs identified by cmeAnalysis. (*H*) Lifetime distribution of (i.e., nonabortive) CCPs identified by DASC. Bona fide CCP analyzed: scramble: 212,004; GAK: 191,542 from three biologically independent experiments. (*D*–*F*) Wilcoxon rank-sum test was used to compare control with kd.

number of SNARE proteins to mediate these downstream fusion events. Although the precise SNAREs involved in nascent endocytic vesicle fusion remain poorly defined, we examined the effects of kd of several SNARE proteins known to be present at the PM and in endosomal compartments (76), including the Q-SNAREs syntaxin 4 and 7 (Fig. 5) and their interacting R-SNARE partners, Vamp 3, 7, and 8 (*SI Appendix*, Fig. S10). Consistent with a critical role for the inclusion of SNARE proteins into CCVs, four of five SNARE proteins we studied have clear roles in early stages of CME and cluster along with group 3 (Stx4) and 6 (Stx7, Vamp3, Vamp8) EAPs. In particular, kd of Stx4 resulted in significant effects on every parameter of CCP dynamics we measured (Fig. 5*A*), including the rates of initiation of sCLSs and CCPs (Fig. 5 *B* and *C*), as well as the percent of bona fide CCPs and their median lifetimes (Fig. 5 *D*–*F*). Interestingly, these pleiotropic effects of Stx4 phenocopied those of the second known SNARE adaptor, CALM (Fig. 5 *A*–*F*), which is capable of binding several SNARE proteins, including Vamp 3, 7, and 8 (77–79). While not affecting the percent of CCPs, nor their median lifetimes, kd of Stx7, Vamp 8, and Vamp3 also significantly affected early stages of CME, including rates of initiation of sCLSs and CCPs (Fig. 5 *A*–*F* and *SI Appendix*, Fig. S10). Whether the stronger phenotypes seen with Stx4 kd reflect differences in kd efficiencies, degree of functional redundancy, or distinct functional properties of this Q-SNARE remains to be determined. Together, these results point to a critical role for the recruitment of SNAREs by their respective adaptor proteins for CCP initiation and stabilization.

Analysis of CCP Dynamics Reveals Early Roles for "Late" Module Proteins in CME. Previous live-cell imaging studies that have tracked the recruitment of EAPs to CCPs have identified amphiphysin1 (Amph1), dynamin-2 (Dyn2), and synaptojanin-2 (Sjn2) as late appearing proteins (5, 7, 30, 80) and part of the fission/uncoating module of CME (1). Epsin has also been proposed to mediate membrane fission and vesicle release (81). Consistent with this, kd of either of these EAPs results in a rightward shift in lifetime distributions, especially of the longer-lived tail, and a pronounced increase in median CCP lifetimes (Fig. 6 *A*, *E*, and *F*). However, these proteins clustered in group 4, together with Fcho1 and NECAP2 (Fig. 2*C*), and exhibited decreases in initiation rates and percent of bona fide CCPs (Fig. 6 *A*, *C*, and *D*). Thus, as has been extensively documented for Dyn2 using the same analytical methods (39), these late acting proteins also appear to have early roles in CME. Whether these defects in CCP stabilization relate to: 1) the existence of an endocytic checkpoint responding to defects resulting from the loss of these proteins, as has been suggested for Dyn2 (6, 9) and more recently for epsin1 kd (11), 2) indirect effects caused by downstream inhibition of CME, or 3) direct roles for these proteins during early stages of CCP stabilization, remains to be established.

Discussion

Quantitative multiparametric TIR-FM and a rigorous experimental pipeline has allowed us to characterize the effects of kd of 67 known or suspected EAPs on CCP dynamics under uniform

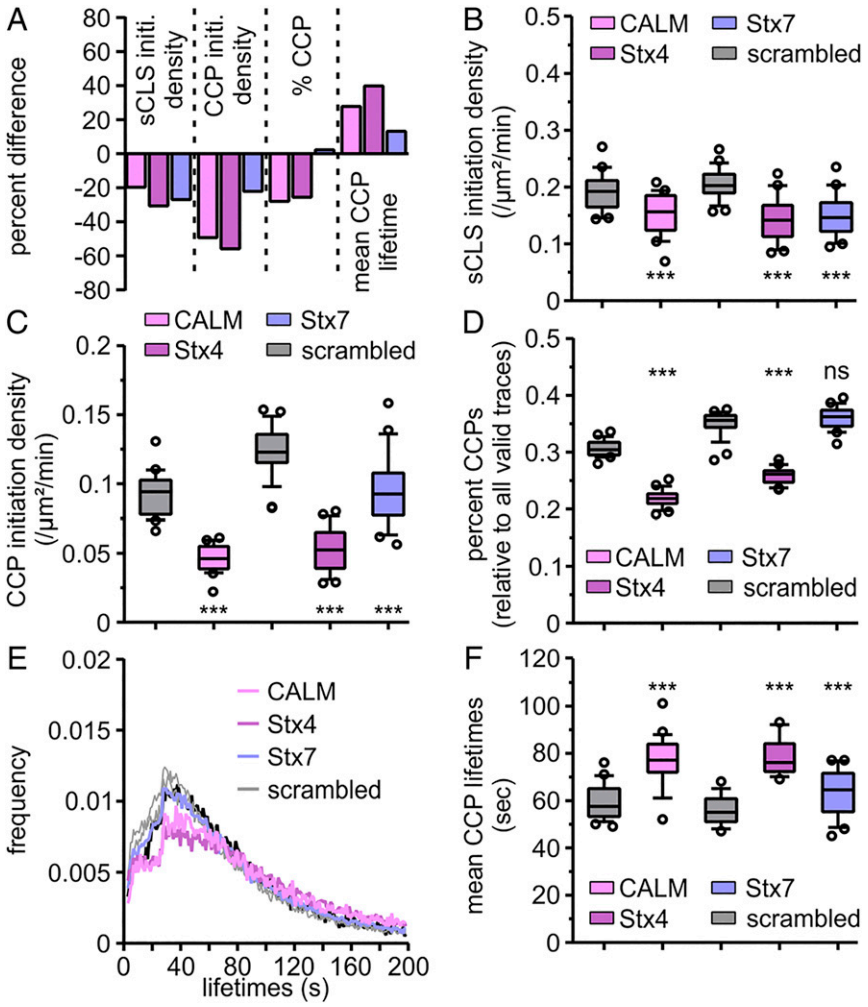

Fig. 5. A role for SNARE proteins and their adaptors during early stages of CCP formation. (*A*). Percent difference plot of initiation density of sCLSs, initiation density of bona fide CCPs, percent CCPs, and mean lifetimes of CCPs. Results are expressed as %Δ relative to the experimental control. Initiation densities of (*B*) sCLSs and (*C*) bona fide CCPs. (*D*) Percentage of CCPs relative to all valid traces. (*E*) Lifetime distribution of bona fide CCPs. (*F*) Mean lifetimes of bona fide CCPs. Bona fide CCP analyzed: controls for CALM, (Stx4, Stx7), Vamp3, and Vamp8: 87,754, (136,457), 95,384, and 104,785, respectively; CALM: 32,921; Stx4: 53,777; Stx7: 98,647; Vamp3: 79,070; Vamp8: 80,769.

experimental conditions. While many of these proteins have been studied individually, in many cases a consensus regarding their functions in CME has failed to emerge, perhaps as a result of differences in experimental conditions, assays, and cell types. Moreover, the sensitivity of our assays allows us to detect phenotypes even when individual isoforms and/or partially functionally redundant EAPs are depleted, and when the effects on CME, as measured by cargo uptake, are mild (9, 11). Unsupervised clustering of our multiparametric data identified phenotypically similar clusters that allowed us to assess the functional hierarchy of EAP activities during the multistep process of CME.

Previously identified functional modules of EAPs, defined based on their predominant biochemical activities and the average kinetics of their peak recruitment to CCPs (1, 7), were dispersed among the phenotypically defined clusters. This suggests that activities such as curvature generation, cargo recognition, scaffolding, AP2 and clathrin interactions are required at multiple stages of CME, including initiation, stabilization, maturation, and fission. The lack of correlation between phenotypic modules of EAPs and their previously ascribed biochemical activities is perhaps not surprising, given that many EAPs are multidomain and hence multifunctional proteins. Moreover, the recruitment of EAPs to CCPs is highly heterogeneous and clear peaks emerge only after averaging (7, 12). Thus, EAP function at CCPs need not be synchronous with their peak recruitment levels. Importantly, motifs involved in protein interactions at CCPs are shared among many EAPs; hence they will likely compete with each other for recruitment. Consequently, cell type differences in relative expression, as well as both overexpression and depletion of one EAP, could alter the recruitment of others. Our results, therefore, provide a snapshot of the hierarchy of EAP function in one cell type, and the effects of kd of one EAP on CCP dynamics could reflect downstream events that alter the integration of other EAP activities.

While many of our findings regarding individual EAP functions can be readily reconciled with those of others, some unexpected results were obtained. Among these was the general role for SNARE proteins and their adaptors during early stages of CME. Several studies have suggested a role for cargo molecules, especially abundant tyrosine-based cargoes, in stabilizing nascent CCPs against early and late abortive events (5, 6, 21). Others have shown that kd of the SNARE adaptors Hrb and, in some cases, CALM strongly inhibits CME (11, 57). Here, we find that several PM-localized SNARE proteins may be critical early regulators of CCP nucleation and stabilization. As has been suggested for the formation of COPI- and COPII-coated vesicles (59, 60), linking incorporation of SNARE proteins to coat assembly is a means to ensure that the newly formed vesicles are capable of fusing with their targets (60). While their fusion activities are highly specific, functional redundancy between SNAREs with regard to their roles in CCP nucleation is likely; hence it will be interesting to examine the effects of dual kd of SNAREs and their adaptors on CME.

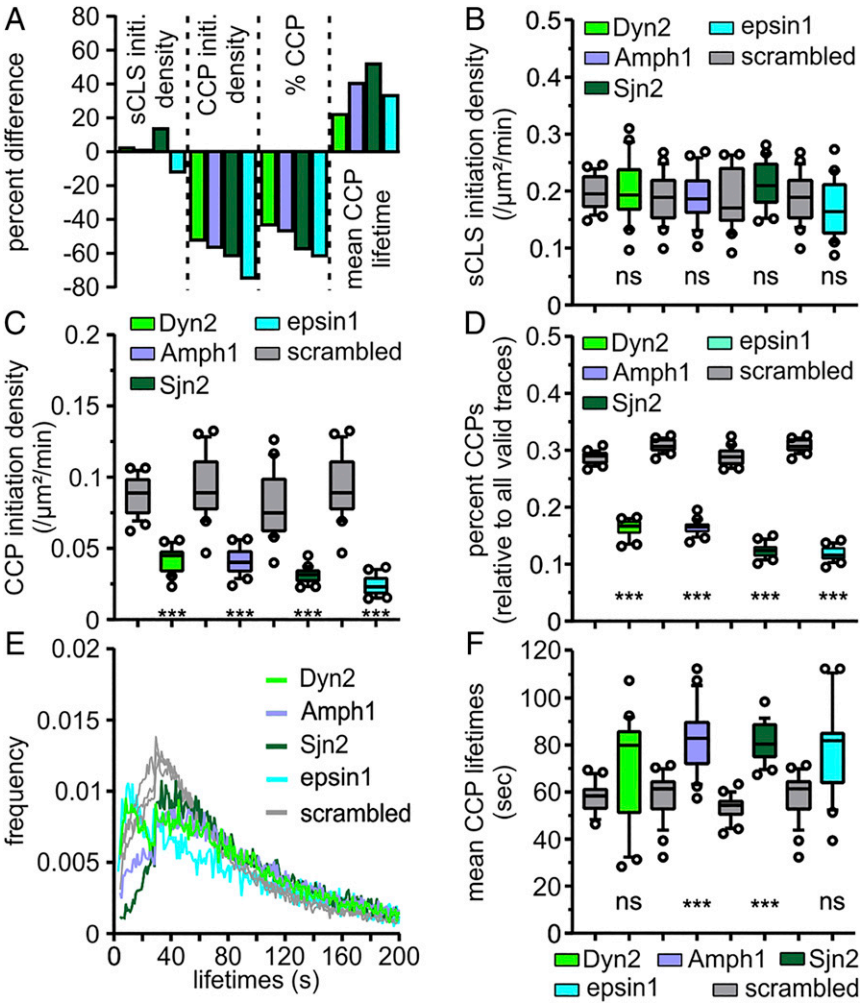

Fig. 6. Analysis of CCP dynamics reveals early roles for "late" module proteins in CME (*A*). Percent difference plot of initiation density of sCLSs, initiation density of bona fide CCPs, percent CCPs, and mean lifetimes of CCPs. Results are expressed as %Δ relative to the experimental control. Initiation densities of (*B*) sCLSs and (*C*) bona fide CCPs. (*D*) Percentage of CCPs relative to all valid traces. (*E*) Lifetime distribution of bona fide CCPs. (*F*) Mean lifetimes of bona fide CCPs. Bona fide CCP analyzed: controls for Dyn2, (Amph1, epsin1), and Sjn2: 73,373, (56,740), and 48,450, respectively; Dyn2: 23,067; Amph1: 26,976; epsin1: 16,140; Sjn2: 22,667.

Several functional studies have pointed to early roles for GAK during CME (46, 82); however, others have shown that GAK is recruited very late to CCPs, concomitant with, or just subsequent to vesicle scission, either when exogenously overexpressed (7, 45, 83) or endogenously tagged (84). We observed both early and late roles for GAK in CME. Most notable was a shift in CCP lifetimes to short-lived species that exhibit a quasi-exponential distribution. That these short-lived CCPs likely correspond to early and late abortive events is consistent with the lifetime distribution of DASC-identified CCPs, as this analysis more accurately resolves productive from abortive structures. Our observed effects on CCP stabilization are consistent with a recent study suggesting that GAK, together with Hsc70, provides an early proofreading mechanism for cargo sorting by coupling coat destabilization with the degree of cargo loading. A role for GAK in CCP maturation is consistent with studies suggesting that GAK-catalyzed clathrin exchange is required for coat rearrangements that drive curvature generation (46), and its potential role in stabilizing AP2-cargo interactions by µ2 phosphorylation. Last, our observed late effects agree with a well-established role for GAK in the uncoating reaction that follows CCV formation (68, 85). These pleiotropic effects of GAK are also expected given that it is essentially a hybrid protein homologous to both AAK1 and auxilin (21). Thus, although the predominant wave of recruitment of GAK to CCPs coincides with CCV formation, our data suggest that lower, catalytic levels of GAK function early during CME. We have similarly reported effects of Dyn1 kd on early CCP dynamics, even when not detectable at CCPs (39, 43).

While we have endeavored to be rigorous in our experimental pipeline and analyses, the comprehensive nature of our screen imposes a number of limitations. First and foremost, as is the case for most screens, we did not routinely measure the efficiency of kd. Thus, quantitative, and even qualitative differences in phenotype could reflect the degree of depletion rather than functional differences between EAPs. Second, while we analyzed multiple cells seeded on duplicate coverslips, with the exception of our analysis of CALM (11) and GAK (Fig. 4 and *SI Appendix*, Fig. S9), our findings have not been replicated by biologically independent experiments. Moreover, as the DASC methodology was being developed while our screen was ongoing, we later discovered that some of our data collection lacked sufficient fluorescent signal-to-noise to accurately apply DASC analysis. Adding DASC analyses and additional biological repeats to all EAPs would provide more statistical power and features to strengthen our clustering results. Finally, although we carefully selected the sgRNAs to pass stringent off-target filtering, off-target effects cannot be completely ruled out. Nonetheless, our comprehensive analysis reveals more complex functional relationships between EAPs and their overlapping roles in early, critical stages of CME. Moreover, it provides a valuable resource to spur further research into EAP function. Building on our findings, future studies involving kd and reconstitution experiments using wt and mutant EAPs will be critical for dissecting

the potentially stage-specific roles of individual domain activities, while double kd experiments will help define functional redundancies driving CCV formation.

In sum, our analysis suggests a more integrated and flexible model for the complex process of CME than the previously proposed stereotypic, modular organization (1, 4). Our studies suggest that functional redundancies between the activities of multidomain EAPs can, in potentially stochastic and variable combinations, drive discrete stages of CCP maturation. As recently proposed (86), this flexibility provides a mechanistic basis for the observed robustness and plasticity of CME (9, 11). Indeed, numerous whole-genome screens for components of the endocytic machinery based on changes in cargo uptake have largely failed to identify EAPs (87–90). The flexible and partially redundant roles of EAPs may also explain apparent discrepancies in the literature, as the extent of involvement of a specific EAP may vary between cell types.

Materials and Methods

Detailed methods are provided in *SI Appendix, Materials and Methods*. In brief, the ARPE-19 cells used in this study were obtained from ATCC and stably express the CCP fiduciary marker eGFP-CLCa (9). These were subsequently transfected with a mammalian lentiviral expression vector encoding the dCas9-BFP-KRAB fusion protein downstream of a SFFV promoter and stable cell lines sorted for BFP fluorescence. For CRISPRi kd of EAPs, cells were infected with lentiviruses encoding two sgRNAs (91) per target gene together with a puromycin selectable marker. The day after infection, cells were subjected to puromycin selection for 2 d, split and seeded onto gelatin-coated coverslips, and CCP dynamics were imaged by TIR-FM. CCP dynamics were measured using our custom-developed cmeAnalysis (6, 9, 32) and DASC (11) analysis pipelines, which are freely available in Github at https://github.com/DanuserLab/cmeAnalysis. The script used for the EAP phenotypic clustering is available in Github at https://github.com/bioinformatics-jeonlee/EAPs_Phenotypic_Clustering.

Data Availability. All cmeAnalysis results are available in Datasets S1 and S2. Due to the amount of data, totaling ~5.2 Tb, all raw imaging data (or subsets of it) will be made available upon request. However, a full dataset and 24 movies each for control and CALM kd, a representative EAP, has been deposited in the NIH FigShare Archive (https://doi.org/10.6084/m9.figshare.13203704.v7 and https://doi.org/10.6084/m9.figshare.13203725.v3, respectively).

ACKNOWLEDGMENTS. We are grateful for early discussions with Jonathan Weissman (University of California, San Francisco) regarding use of CRISPRi technology. We thank Aparna Mohanakrishnan for modifying the original sgRNA cloning vector and Zhiming Chen for help with mass spectrometry. We thank all members of the S.L.S. laboratory and Dinah Loerke (University of Denver) for critical scientific discussions. J.L. was supported by a Cancer Prevention Research Institute of Texas grant (RP150596) that supports the Bioinformatic Core Facility at the University of Texas Southwestern. This work was supported by NIH Grants MH61345 (to S.L.S.) and GM73165 (to S.L.S., G.D., and M.M.).

1. H. T. McMahon, E. Boucrot, Molecular mechanism and physiological functions of clathrin-mediated endocytosis. *Nat. Rev. Mol. Cell Biol.* **12**, 517–533 (2011).
2. D. Azarnia Tehran, T. López-Hernández, T. Maritzen, Endocytic adaptor proteins in health and disease: Lessons from model organisms and human mutations. *Cells* **8**, 1345 (2019).
3. M. Mettlen, P. H. Chen, S. Srinivasan, G. Danuser, S. L. Schmid, Regulation of clathrin-mediated endocytosis. *Annu. Rev. Biochem.* **87**, 871–896 (2018).
4. M. Kaksonen, A. Roux, Mechanisms of clathrin-mediated endocytosis. *Nat. Rev. Mol. Cell Biol.* **19**, 313–326 (2018).
5. M. Ehrlich *et al*., Endocytosis by random initiation and stabilization of clathrin-coated pits. *Cell* **118**, 591–605 (2004).
6. D. Loerke *et al*., Cargo and dynamin regulate clathrin-coated pit maturation. *PLoS Biol.* **7**, e57 (2009).
7. M. J. Taylor, D. Perrais, C. J. Merrifield, A high precision survey of the molecular dynamics of mammalian clathrin-mediated endocytosis. *PLoS Biol.* **9**, e1000604 (2011).
8. C. J. Merrifield, M. Kaksonen, Endocytic accessory factors and regulation of clathrin-mediated endocytosis. *Cold Spring Harb. Perspect. Biol.* **6**, a016733 (2014).
9. F. Aguet, C. N. Antonescu, M. Mettlen, S. L. Schmid, G. Danuser, Advances in analysis of low signal-to-noise images link dynamin and AP2 to the functions of an endocytic checkpoint. *Dev. Cell* **26**, 279–291 (2013).
10. A. M. Motley *et al*., Functional analysis of AP-2 alpha and mu2 subunits. *Mol. Biol. Cell* **17**, 5298–5308 (2006).
11. X. Wang *et al*., DASC, a sensitive classifier for measuring discrete early stages in clathrin-mediated endocytosis. *eLife* **9**, 353686 (2020).
12. D. Loerke, M. Mettlen, S. L. Schmid, G. Danuser, Measuring the hierarchy of molecular events during clathrin-mediated endocytosis. *Traffic* **12**, 815–825 (2011).
13. M. Kaksonen, Y. Sun, D. G. Drubin, A pathway for association of receptors, adaptors, and actin during endocytic internalization. *Cell* **115**, 475–487 (2003).
14. M. Kaksonen, C. P. Toret, D. G. Drubin, A modular design for the clathrin- and actin-mediated endocytosis machinery. *Cell* **123**, 305–320 (2005).
15. R. C. Delos Santos *et al*., Selective regulation of clathrin-mediated epidermal growth factor receptor signaling and endocytosis by phospholipase C and calcium. *Mol. Biol. Cell* **28**, 2802–2818 (2017).
16. K. Eichel, D. Jullié, M. von Zastrow, β-Arrestin drives MAP kinase signalling from clathrin-coated structures after GPCR dissociation. *Nat. Cell Biol.* **18**, 303–310 (2016).
17. P. Liberali, B. Snijder, L. Pelkmans, A hierarchical map of regulatory genetic interactions in membrane trafficking. *Cell* **157**, 1473–1487 (2014).
18. P. H. Chen *et al*., Crosstalk between CLCb/Dyn1-mediated adaptive clathrin-mediated endocytosis and epidermal growth factor receptor signaling increases metastasis. *Dev. Cell* **40**, 278–288.e5 (2017).
19. C. R. Reis *et al*., Crosstalk between Akt/GSK3β signaling and dynamin-1 regulates clathrin-mediated endocytosis. *EMBO J.* **34**, 2132–2146 (2015).
20. S. L. Schmid, Reciprocal regulation of signaling and endocytosis: Implications for the evolving cancer cell. *J. Cell Biol.* **216**, 2623–2632 (2017).
21. Z. Kadlecova *et al*., Regulation of clathrin-mediated endocytosis by hierarchical allosteric activation of AP2. *J. Cell Biol.* **216**, 167–179 (2017).
22. G. M. Beacham, E. A. Partlow, J. J. Lange, G. Hollopeter, NECAPs are negative regulators of the AP2 clathrin adaptor complex. *eLife* **7**, e32242 (2018).
23. B. Ritter *et al*., NECAP 1 regulates AP-2 interactions to control vesicle size, number, and cargo during clathrin-mediated endocytosis. *PLoS Biol.* **11**, e1001670 (2013).
24. E. Cocucci, F. Aguet, S. Boulant, T. Kirchhausen, The first five seconds in the life of a clathrin-coated pit. *Cell* **150**, 495–507 (2012).
25. W. M. Henne *et al*., FCHo proteins are nucleators of clathrin-mediated endocytosis. *Science* **328**, 1281–1284 (2010).
26. P. K. Umasankar *et al*., Distinct and separable activities of the endocytic clathrin-coat components Fcho1/2 and AP-2 in developmental patterning. *Nat. Cell Biol.* **14**, 488–501 (2012).
27. M. G. Ford *et al*., Curvature of clathrin-coated pits driven by epsin. *Nature* **419**, 361–366 (2002).
28. M. J. Hawryluk *et al*., Epsin 1 is a polyubiquitin-selective clathrin-associated sorting protein. *Traffic* **7**, 262–281 (2006).
29. Z. Chen *et al*., Wbox2: A clathrin terminal domain-derived peptide inhibitor of clathrin-mediated endocytosis. *J. Cell Biol.* **219**, e201908189 (2020).
30. C. J. Merrifield, M. E. Feldman, L. Wan, W. Almers, Imaging actin and dynamin recruitment during invagination of single clathrin-coated pits. *Nat. Cell Biol.* **4**, 691–698 (2002).
31. M. A. Puthenveedu, M. von Zastrow, Cargo regulates clathrin-coated pit dynamics. *Cell* **127**, 113–124 (2006).
32. K. Jaqaman *et al*., Robust single-particle tracking in live-cell time-lapse sequences. *Nat. Methods* **5**, 695–702 (2008).
33. M. Mettlen, G. Danuser, Imaging and modeling the dynamics of clathrin-mediated endocytosis. *Cold Spring Harb. Perspect. Biol.* **6**, a017038 (2014).
34. D. Perrais, C. J. Merrifield, Dynamics of endocytic vesicle creation. *Dev. Cell* **9**, 581–592 (2005).
35. D. J. Owen, B. M. Collins, P. R. Evans, Adaptors for clathrin coats: Structure and function. *Annu. Rev. Cell Dev. Biol.* **20**, 153–191 (2004).
36. L. P. Jackson *et al*., A large-scale conformational change couples membrane recruitment to cargo binding in the AP2 clathrin adaptor complex. *Cell* **141**, 1220–1229 (2010).
37. B. T. Kelly *et al*., Clathrin adaptors. AP2 controls clathrin polymerization with a membrane-activated switch. *Science* **345**, 459–463 (2014).
38. A. G. Wrobel *et al*., Temporal ordering in endocytic clathrin-coated vesicle formation via AP2 phosphorylation. *Dev. Cell* **50**, 494–508.e11 (2019).
39. M. Bhave, M. Mettlen, X. Wang, S. L. Schmid, Early and nonredundant functions of dynamin isoforms in clathrin-mediated endocytosis. *Mol. Biol. Cell* **31**, 2035–2047 (2020).
40. R. Tibshirani, G. Walther, T. Hastie, Estimating the number of clusters in a data set via the gap statistic. *J. R. Stat. Soc.* **63**, 411–423 (2001).
41. B. J. Frey, D. Dueck, Clustering by passing messages between data points. *Science* **315**, 972–976 (2007).
42. J. Pierezan, Non-parametric Multivariate Permutation Test (Version 1.0.0, MATLAB Central File Exchange, MathWorks, 2020).
43. S. Srinivasan *et al*., A noncanonical role for dynamin-1 in regulating early stages of clathrin-mediated endocytosis in non-neuronal cells. *PLoS Biol.* **16**, e2005377 (2018).
44. E. Eisenberg, L. E. Greene, Multiple roles of auxilin and hsc70 in clathrin-mediated endocytosis. *Traffic* **8**, 640–646 (2007).
45. R. H. Massol, W. Boll, A. M. Griffin, T. Kirchhausen, A burst of auxilin recruitment determines the onset of clathrin-coated vesicle uncoating. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 10265–10270 (2006).
46. D. W. Lee, X. Zhao, F. Zhang, E. Eisenberg, L. E. Greene, Depletion of GAK/auxilin 2 inhibits receptor-mediated endocytosis and recruitment of both clathrin and clathrin adaptors. *J. Cell Sci.* **118**, 4311–4321 (2005).
47. E. M. Schmid, H. T. McMahon, Integrating molecular and network biology to decode endocytosis. *Nature* **448**, 883–888 (2007).

48. L. M. Traub, Regarding the amazing choreography of clathrin coats. *PLoS Biol.* **9**, e1001037 (2011).
49. V. Legendre-Guillemin, S. Wasiak, N. K. Hussain, A. Angers, P. S. McPherson, ENTH/ANTH proteins and clathrin-mediated membrane budding. *J. Cell Sci.* **117**, 9–18 (2004).
50. M. T. Drake, M. A. Downs, L. M. Traub, Epsin binds to clathrin by associating directly with the clathrin-terminal domain. Evidence for cooperative binding through two discrete sites. *J. Biol. Chem.* **275**, 6479–6489 (2000).
51. H. Chen *et al*., Epsin is an EH-domain-binding protein implicated in clathrin-mediated endocytosis. *Nature* **394**, 793–797 (1998).
52. P. N. Dannhauser, E. J. Ungewickell, Reconstitution of clathrin-coated bud and vesicle formation with minimal components. *Nat. Cell Biol.* **14**, 634–639 (2012).
53. M. Szymanska *et al*., Interaction with epsin 1 regulates the constitutive clathrin-dependent internalization of ErbB3. *Biochim. Biophys. Acta* **1863**, 1179–1188 (2016).
54. L. J. Page, M. S. Robinson, Targeting signals and subunit interactions in coated vesicle adaptor complexes. *J. Cell Biol.* **131**, 619–630 (1995).
55. M. Doria *et al*., The eps15 homology (EH) domain-based interaction between eps15 and hrb connects the molecular machinery of endocytosis to that of nucleocytosolic transport. *J. Cell Biol.* **147**, 1379–1384 (1999).
56. E. M. Schmid *et al*., Role of the AP2 beta-appendage hub in recruiting partners for clathrin-coated vesicle assembly. *PLoS Biol.* **4**, e262 (2006).
57. M. Chaineau, L. Danglot, V. Proux-Gillardeaux, T. Galli, Role of HRB in clathrin-dependent endocytosis. *J. Biol. Chem.* **283**, 34365–34373 (2008).
58. P. R. Pryor *et al*., Molecular basis for the sorting of the SNARE VAMP7 into endocytic clathrin-coated vesicles by the ArfGAP Hrb. *Cell* **134**, 817–827 (2008).
59. U. Rein, U. Andag, R. Duden, H. D. Schmitt, A. Spang, ARF-GAP-mediated interaction between the ER-Golgi v-SNAREs and the COPI coat. *J. Cell Biol.* **157**, 395–404 (2002).
60. S. Springer, A. Spang, R. Schekman, A primer on vesicle budding. *Cell* **97**, 145–148 (1999).
61. V. Legendre-Guillemin *et al*., HIP1 and HIP12 display differential binding to F-actin, AP2, and clathrin. Identification of a novel interaction with clathrin light chain. *J. Biol. Chem.* **277**, 19897–19904 (2002).
62. M. A. Senetar, S. J. Foster, R. O. McCann, Intrasteric inhibition mediates the interaction of the I/LWEQ module proteins Talin1, Talin2, Hip1, and Hip12 with actin. *Biochemistry* **43**, 15418–15428 (2004).
63. J. D. Wilbur *et al*., Actin binding by Hip1 (huntingtin-interacting protein 1) and Hip1R (Hip1-related protein) is regulated by clathrin light chain. *J. Biol. Chem.* **283**, 32870–32879 (2008).
64. S. K. Mishra *et al*., Clathrin- and AP-2-binding sites in HIP1 uncover a general assembly role for endocytic accessory proteins. *J. Biol. Chem.* **276**, 46230–46236 (2001).
65. M. Metzler *et al*., HIP1 functions in clathrin-mediated endocytosis through binding to clathrin and adaptor protein 2. *J. Biol. Chem.* **276**, 39271–39276 (2001).
66. I. Gottfried, M. Ehrlich, U. Ashery, HIP1 exhibits an early recruitment and a late stage function in the maturation of coated pits. *Cell. Mol. Life Sci.* **66**, 2897–2911 (2009).
67. C. Y. Chen, F. M. Brodsky, Huntingtin-interacting protein 1 (Hip1) and Hip1-related protein (Hip1R) bind the conserved sequence of clathrin light chains and thereby influence clathrin assembly in vitro and actin distribution in vivo. *J. Biol. Chem.* **280**, 6109–6117 (2005).
68. A. Umeda, A. Meyerholz, E. Ungewickell, Identification of the universal cofactor (auxilin 2) in clathrin coat dissociation. *Eur. J. Cell Biol.* **79**, 336–342 (2000).
69. S. D. Conner, S. L. Schmid, Identification of an adaptor-associated kinase, AAK1, as a regulator of clathrin-mediated endocytosis. *J. Cell Biol.* **156**, 921–929 (2002).
70. D. Ricotta, S. D. Conner, S. L. Schmid, K. von Figura, S. Honing, Phosphorylation of the AP2 mu subunit by AAK1 mediates high affinity binding to membrane protein sorting signals. *J. Cell Biol.* **156**, 791–795 (2002).
71. C. X. Zhang *et al*., Multiple roles for cyclin G-associated kinase in clathrin-mediated sorting events. *Traffic* **6**, 1103–1113 (2005).
72. A. E. Engqvist-Goldstein *et al*., The actin-binding protein Hip1R associates with clathrin during early stages of endocytosis and promotes clathrin assembly in vitro. *J. Cell Biol.* **154**, 1209–1223 (2001).
73. M. Mettlen *et al*., Endocytic accessory proteins are functionally distinguished by their differential effects on the maturation of clathrin-coated pits. *Mol. Biol. Cell* **20**, 3251–3260 (2009).
74. L. Ma *et al*., Transient Fcho1/2·Eps15/R·AP-2 nanoclusters prime the AP-2 clathrin adaptor for cargo binding. *Dev. Cell* **37**, 428–443 (2016).
75. F. R. Maxfield, T. E. McGraw, Endocytic recycling. *Nat. Rev. Mol. Cell Biol.* **5**, 121–132 (2004).
76. R. Jahn, R. H. Scheller, SNAREs—engines for membrane fusion. *Nat. Rev. Mol. Cell Biol.* **7**, 631–643 (2006).
77. A. Harel, F. Wu, M. P. Mattson, C. M. Morris, P. J. Yao, Evidence for CALM in directing VAMP2 trafficking. *Traffic* **9**, 417–429 (2008).
78. S. J. Koo *et al*., SNARE motif-mediated sorting of synaptobrevin by the endocytic adaptors clathrin assembly lymphoid myeloid leukemia (CALM) and AP180 at synapses. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 13540–13545 (2011).
79. S. E. Miller *et al*., The molecular basis for the endocytosis of small R-SNAREs by the clathrin adaptor CALM. *Cell* **147**, 1118–1131 (2011).
80. R. M. Perera, R. Zoncu, L. Lucast, P. De Camilli, D. Toomre, Two synaptojanin 1 isoforms are recruited to clathrin-coated pits at different stages. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 19332–19337 (2006).
81. E. Boucrot *et al*., Membrane fission is promoted by insertion of amphipathic helices and is restricted by crescent BAR domains. *Cell* **149**, 124–136 (2012).
82. Y. Chen *et al*., Dynamic instability of clathrin assembly provides proofreading control for endocytosis. *J. Cell Biol.* **218**, 3200–3211 (2019).
83. D. W. Lee, X. Wu, E. Eisenberg, L. E. Greene, Recruitment dynamics of GAK and auxilin to clathrin-coated pits during endocytosis. *J. Cell Sci.* **119**, 3502–3512 (2006).
84. K. He *et al*., Dynamics of Auxilin 1 and GAK in clathrin-mediated traffic. *J. Cell Biol.* **219**, e201908142 (2020).
85. T. Greener, X. Zhao, H. Nojima, E. Eisenberg, L. E. Greene, Role of cyclin G-associated kinase in uncoating clathrin-coated vesicles from non-neuronal cells. *J. Biol. Chem.* **275**, 1365–1370 (2000).
86. Z. Chen, S. L. Schmid, Evolving models for assembling and shaping clathrin-coated pits. *J. Cell Biol.* **219**, e202005126 (2020).
87. M. C. Bassik *et al*., A systematic mammalian genetic interaction map reveals pathways underlying ricin susceptibility. *Cell* **152**, 909–922 (2013).
88. C. Collinet *et al*., Systems survey of endocytosis by multiparametric image analysis. *Nature* **464**, 243–249 (2010).
89. D. R. Gulbranson *et al*., AAGAB controls AP2 adaptor assembly in clathrin-mediated endocytosis. *Dev. Cell* **50**, 436–446.e5 (2019).
90. P. Kozik *et al*., A human genome-wide screen for regulators of clathrin-coated vesicle formation reveals an unexpected role for the V-ATPase. *Nat. Cell Biol.* **15**, 50–60 (2013).
91. M. A. Horlbeck *et al*., Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. *eLife* **5**, e19760 (2016).