Surface delivery quantification reveals distinct trafficking efficiencies among clustered protocadherin isoforms

Elizabeth J. May[1] and Rachelle Gaudet*

Department of Molecular and Cellular Biology, Harvard University, 52 Oxford Street, Cambridge, MA 02138, USA

*Correspondence: Rachelle Gaudet, Harvard University, 52 Oxford St, Cambridge, MA 02138

gaudet@mcb.harvard.edu

[1] Present Address: California Institute for Quantitative Biosciences, University of California, Berkeley, Stanley Hall, Berkeley, CA

Author contributions statement:

E.J.M and R.G. designed the research, wrote the paper, and prepared the figures. E.J.M performed experiments, developed the analysis and wrote the code, and analyzed the data.

Classification:
Major: Biological Sciences
Minor: Cell Biology

Keywords: clustered protocadherins, cell surface trafficking, plasma membrane proteins

This PDF file includes:

Main text
Figures 1 to 6
Table 1

Abstract

Proteins that transmit molecules and signals across the plasma membrane are crucial in cell biology because they enable cells to sense and respond to their surroundings. A major challenge for studying cell-surface proteins is that often they do not fold or traffic properly to the plasma membrane when produced in heterologous cells. We developed a strategy for quantifying surface localization from fluorescence microscopy images of surface-stained cells. Using clustered protocadherins, a protein family important for cell-cell recognition during neuronal development, as a model system, we found that surface delivery levels vary among clustered protocadherin isoforms and between wildtype and engineered variants. Quantifying these differences provides evidence that challenges the generally accepted hypothesis that *cis* dimerization controls surface delivery of clustered protocadherins. This work establishes a generalizable framework for screening proteins and variants of interest for proper cell surface localization.

Significance

Surface proteins allow cells to interact with their environments, and their activities are often regulated by their delivery to and removal from the plasma membrane. We developed a strategy to quantitatively compare surface delivery of proteins based on established epitope tag-based surface staining methods. Using natural and engineered variants of clustered protocadherins, cell-surface proteins essential for neuron development, we show that such quantitative comparisons of surface trafficking facilitate the interpretation of mutational effects and can shed light on key regulatory mechanisms. We find that surface trafficking levels differ between variants and that, contrary to what was previously thought, a domain that inhibits surface delivery in some clustered protocadherins may do so independently of its protein-protein interaction interface.

Introduction

Clustered protocadherins (cPCDHs) are a large family of single-pass transmembrane proteins that function as adhesion and signaling molecules in brain development. They are primarily expressed in the nervous system (1-4) and their genetic deletion in mice is lethal at birth due to extensive neuronal cell death (5-8). At the cellular scale, cPCDH perturbations lead to a range of defects in neuronal cell morphology and connectivity, and their roles in dendrite complexity and self-avoidance are the most studied and best understood (9-14). These essential neuronal activities rely on homophilic cPCDH interactions at cell-cell contacts, so proper localization to the plasma membrane is a prerequisite for cPCDH function.

In mammalian genomes, including mice and humans, more than 50 genes encode a diverse set of paralogous cPCDH isoforms (4, 15, 16). Each isoform contains six extracellular cadherin (EC) domains, a transmembrane helix, and an intracellular region **(Figure 1A)**. The cPCDH genomic region includes clustered arrays of exons encoding isoform sequences from the N-terminal extracellular domains through the first ~100 amino acids of the intracellular region. From there, the sequences differ only by subfamily: α-PCDHs and γ-PCDHs have α- and γ-specific C-terminal sequences encoded by shared exons, while β-PCDHs do not have shared exons and thus have shorter intracellular regions (3, 4, 16-18). At the cell surface, cPCDHs form strictly homophilic *trans* dimers between molecules on juxtaposed membranes and preferentially heterophilic *cis* dimers between isoforms on the same membrane (19-31). Defining these cPCDH interactions has led to a proposed mechanism wherein cPCDH oligomerization via both *cis* and *trans* dimers serves as a self-contact signal that ultimately leads to self-avoidance through an as-yet-unknown signaling pathway (20, 22, 23, 26-30, 32). Cell surface localization and *cis* and *trans* dimerization are all required for this mechanism.

In the brain and in cultured neurons, cPCDHs are not uniformly distributed on the cell surface; instead, they are largely intracellular, have a punctate distribution, and are often enriched at cell-cell contact sites (31, 33-39). Using cell aggregation assays as a functional readout, investigators have noted that when transfected into heterologous cells, α-PCDHs do not reach the cell surface, while β- and γ-PCDHs do (21). This differential trafficking has been attributed to *cis* dimerization (21, 26, 28, 40). The cPCDH *cis* interaction involves EC6 from one subunit and both EC5 and EC6 (EC5-6) from the other subunit forming an asymmetric dimer (28). Several conditions induced α-PCDHs to go to the cell surface: (i) deletion of the EC6 domain, (ii) replacement of the α-PCDH EC6 domain with a β- or γ-PCDH EC6 domain, (iii) co-transfection with a full-length β- or γ-PCDH, or (iv) co-transfection with an EC5-6 fragment (but not an EC6-only fragment) of a β- or γ-PCDH (21). Disruptive mutations introducing charged amino acids into *cis* interface positions in β- and γ-PCDHs also interfered with their surface delivery (26, 28). These observations have led to the hypothesis that *cis* dimerization is required for cPCDH surface delivery (21, 26, 28, 40), although the mechanism of this regulation remains unclear. An additional layer of γ-PCDH trafficking regulation has been mapped to a motif in the variable part of the intracellular region, ubiquitination or phosphorylation of which may impact surface delivery (31, 38, 39, 41-44).

Several previous studies have included observations of cPCDH subcellular localization in heterologous cells (1, 19, 21, 33-35, 41, 45-50), but few have made quantitative surface delivery measurements (31, 49). To identify cPCDH features that influence their surface trafficking, we quantified cell surface localization using antibody detection of an extracellular epitope tag and developed generalizable metrics that summarize multiple aspects of diverse surface delivery phenotypes. We tested mutation-containing and domain-swapped α-PCDH variants to determine which molecular features are sufficient to enable cell surface trafficking. We found that while

replacing the EC6 domain of PCDHα4 with EC6 of a β- or γ-PCDH increased surface localization compared to wildtype α4, introducing *cis* interface residues from γB6 into α4 did not. EC6-swapped α4 constructs also had lower surface delivery levels than wildtype β- and γ-PCDHs. Our results indicate that EC6 contributes to, but does not completely control surface delivery of cPCDHs, and suggest that the mechanism of cPCDH surface trafficking regulation may be independent of *cis* dimerization.

Results

Defining metrics for surface delivery quantification

To investigate the surface trafficking of cPCDHs, we pursued a quantitative approach that would facilitate comparisons of wildtype cPCDH isoforms and engineered variants. Surface staining by antibody detection of extracellular epitope tags is a common method for assessing a protein's surface localization that has previously been applied to cPCDHs in heterologous cell lines and cultured neurons (21, 34, 48-51). We used a derivative of the 293T cell line that lacks the adhesion protein N-cadherin (293T-ΔNC) (52). 293T-ΔNC cells are adherent but detach easily and can be used for cell aggregation assays (52). Using cPCDH constructs containing an extracellular Myc tag and an intracellular fluorescent protein (FP), we added fluorescent anti-Myc antibodies (Ab) to transfected cells and collected fluorescence microscopy images. We first compared two wildtype cPCDH isoforms known to have different surface trafficking behaviors: α4, which does not traffic to the plasma membrane, and γC3, which does (**Figure 1B**) (21). The FP served as a transfection marker and was visible throughout the cells in all samples (**Figure 1C**). Notably, the cellular distribution of the FP signal was indistinguishable between α4 and γC3, although the overall FP intensity was different. There was no visible Ab signal in the α4-transfected sample (**Figure 1C**),

confirming the expected non-surface localization of α4. As expected, the Ab stained the cell periphery in the Myc-tagged γC3 sample (**Figure 1C**).

Although the overall surface delivery behaviors of α4 and γC3 were apparent in our images, we noted some variation in surface delivery between cells within each sample that we wanted to investigate further. For example, some γC3-expressing cells had very little or no measurable surface staining (**Figure 1C, arrowheads**). We therefore devised a method that accounts for the phenotypes of individual cells and evaluates the overall surface delivery behavior of a construct based on the distribution of cellular phenotypes. For each sample, we collected brightfield images and fluorescence images in the FP and Ab channels (**Figure 2A**). We used automated segmentation in brightfield to identify cells and calculated the mean FP and Ab signal per cell (FP_{cell} and Ab_{cell}) from the fluorescence channels. We identified transfected cells using a threshold FP_{cell} value determined from images of untransfected control cells (**Figure 2B**). To visualize the distribution of surface stain signals across cells in each sample, we displayed the Ab_{cell} values for transfected cells as a histogram (**Figure 2C**).

The surface stain histograms illustrate the variation in surface delivery phenotypes present within a sample and can be superimposed to facilitate wholistic comparisons between samples. We compared γC3 to a γC3 control construct lacking the extracellular Myc tag. The γC3 distribution had two peaks, with surface-stained cells making up a broad, right-shifted subpopulation and unstained transfected cells contributing to a peak that overlapped with the untagged γC3 distribution (**Figure 2C**). We calculated the fraction of transfected cells with measurable surface staining, f^+, using a threshold Ab_{cell} value determined from the untagged γC3 distribution (**Figure 2C**, cyan vertical line). For α4, f^+ was close to zero (0.01 ± 0.01, mean ± SD), while for γC3 f^+ was 0.62 ± 0.07 (**Figure 2E**). We also calculated the mean Ab_{cell} value for all transfected cells in

each sample, $\overline{Ab}$, which is proportional to the average cPCDH concentration on the cell surface for a given construct (**Figure 2C**, black vertical line). As expected, f^{+} and $\overline{Ab}$ for γC3 were significantly higher than for α4 ($p = 2\times10^{-12}$ for f^{+} and $p = 5\times10^{-7}$ for Ab using Welch's unequal variances t-test; **Figure 2F**).

We next investigated whether the amount of protein on the cell surface correlated with its expression level. Taking advantage of the variation in expression generated by transient transfection and reported by our transfection marker, we constructed dose-response curves. To group cells by expression level, we binned by FP_{cell} (FP_{bin}) and calculated the mean Ab_{cell} for each bin (Ab_{bin}). The FP_{bin}–Ab_{bin} relationship appears to follow a power law (**Figure S1D)**, so we plotted FP_{bin} and Ab_{bin} on a log-log scale to display surface expression as a function of protein expression for each sample (**Figure 2D**). The slope of a linear fit to these data represents a surface delivery "efficiency" ($E_{surface}$) that is related to the fraction of protein produced that localized to the cell surface. $E_{surface}$ was positive for γC3, indicating that cells making more γC3 protein overall had more γC3 on the surface than cells with lower γC3 expression levels, whereas $E_{surface}$ for α4 was close to zero, showing that there was little to no surface staining of α4-transfected cells at any expression level (**Figure 2G**).

Quantification can distinguish different surface delivery levels

To validate our quantification strategy and investigate the relationship between cPCDH *cis* dimerization and surface delivery, we focused on the *cis* interface domain EC6. We tested an α4 construct lacking EC6 (α4ΔEC6), EC6-swapped α4 constructs (α4-γB6$_{EC6}$, α4-β17$_{EC6}$, and α4-γC3$_{EC6}$), and their wildtype counterparts (γB6, β17, and γC3) for surface delivery in 293T-ΔNC cells (**Figure 3A**). There was no Ab signal for α4, as expected (**Figure S2A**). The EC6-deleted and

EC6-swapped samples had dim, punctate Ab signal, often at contact points between transfected cells (**Figure S2, B-E**). Wildtype β- and γ-PCDHs had more uniform surface staining with bright signal at cell-cell contacts, and β17 and γC3 were brighter than γB6 (**Figure S2, F-H**). These samples thus presented a range of surface delivery levels, and we used f^{+}, $\overline{Ab}$, and $E_{surface}$ to compare and summarize their distinct surface delivery phenotypes.

All the Ab_{cell} histograms at least partially overlapped with the untagged γC3 distribution, showing that every sample contained a subpopulation of transfected cells without any surface stain signal (**Figure S3B**). Cells transfected with α4ΔEC6 or the EC6-swapped constructs had f^{+} between 0.1 and 0.5, while the wildtype β- and γ-PCDHs generally had higher f^{+} ranging from 0.3 to 0.8 (**Figure 3B**). $\overline{Ab}$ for α4ΔEC6 and the EC6-swapped constructs was significantly higher than $\overline{Ab}$ for wildtype α4 ($p \leq 0.03$) and $\overline{Ab}$ was highest for the wildtype β- and γ-PCDH isoforms (**Figure 3C**). The $\overline{Ab}$ metric emphasized the ultra-bright surface staining of some cells in the γC3 and β17 samples, capturing the difference we noticed compared to γB6 and indicating that γC3 and β17 traffic to the cell membrane particularly well for cPCDHs. Compared to α4, $E_{surface}$ was positive and significantly higher for all other constructs, and $E_{surface}$ was highest for the wildtype β- and γ-isoforms (**Figure 3D**).

Overall, our quantification indicates that surface levels of α4ΔEC6 and the EC6-swapped variants are higher than wildtype α4 in 293T-ΔNC cells. These findings agree with a previous study showing that the same α4ΔEC6 and EC6-swapped α4 variants were detected on the cell surface and mediated cell aggregation in K562 cells like wildtype β- and γ-PCDHs (21). Furthermore, our quantification revealed that the constructs with detectable levels of cell surface protein were divided are two groups, with wildtype isoforms (γB6, β17, and γC3) generally having more surface protein than the engineered α4 variants (α4ΔEC6, α4-γB6$_{EC6}$, α4-β17$_{EC6}$, and α4-γC3$_{EC6}$). As

removing or swapping EC6 was insufficient to bring the amount of cell surface α4 to the levels of wildtype β- or γ-PCDHs, cPCDH features beyond EC6 must also contribute to surface delivery.

Surface levels of EC6-swapped α-PCDHs are lower than wildtype β- and γ-PCDHs

To test whether our findings would generalize to all cPCDHs, we generated an expanded set of EC6-swapped α4 constructs using isoforms from the different cPCDH subfamilies (**Table 1; Figure 4A, Figure S4A**). Most of the EC6-swapped α4 constructs behaved similarly: f^+, $\overline{Ab}$, and $E_{surface}$ were significantly higher than wildtype α4 ($p \leq 0.002$ for $E_{surface}$), but not as high as wildtype β- and γ-PCDHs (**Figure 4B-C, Figure S4B-E**). The one exception was αC2, which is part of the α-PCDH genomic locus but is phylogenetically distinct from the α-PCDHs, like the other C-type isoforms. Wildtype αC2 apparently did not traffic to the cell membrane, with f^+, $\overline{Ab}$, and $E_{surface}$ similar to α4, while α4-αC2$_{EC6}$ had higher f^+, $\overline{Ab}$, and $E_{surface}$ than α4, similar to the other EC6-swaps (**Figure 4B-C, Figure S4B-E**).

As cPCDH *cis* dimerization is asymmetric, with EC6 of one subunit binding EC5 and EC6 of the other subunit, we wondered whether α4 constructs with both EC5 and EC6 swapped would traffic as well as their wildtype β- and γ-PCDH counterparts. We generated EC5-6-swapped α4 variants that represented each cPCDH subfamily: α4-γA11$_{EC5-6}$, α4-γB6$_{EC5-6}$, α4-β17$_{EC5-6}$, and α4-γC3$_{EC5-6}$ (**Table 1; Figure 4A, Figure S4A**). Interestingly, swapping EC5 did not bridge the gap in surface delivery between EC6-swapped α4 constructs and wildtype β- and γ-PCDHs (**Figure 4B-C**). The α4-γB6$_{EC5-6}$ construct came the closest, with f^+ and $\overline{Ab}$, but not $E_{surface}$, exceeding α4-γB6$_{EC6}$ and approaching wildtype γB6 levels (**Figure 4C and Figure S4D-E**, light blue bars). The other EC5-

6-swapped α4 constructs showed no improvement over their EC6-swapped counterparts. Therefore, features of EC6, but not EC5, contribute to cPCDH surface delivery.

To explain the EC6-swap rescue phenotype, we and others hypothesized that *cis* homodimerization may be conserved among β- and γ-PCDHs, but not α-PCDHs, and that *cis* dimerization is required for surface delivery (21, 26, 28, 40). A prediction from this hypothesis is that surface delivery should be correlated between wildtype isoforms and their corresponding EC6-swapped α4 variants. For example, we expected that α4-αC2$_{EC6}$ would have the lowest surface abundance of the EC6-swapped α4 constructs, and that α4-β17$_{EC6}$ would have the highest, because they have the same EC6 *cis* interface surfaces as αC2 and β17, the wildtype isoforms with the weakest and strongest surface delivery phenotypes, respectively (**Figure 4C, Figure S4D-E**). Surprisingly, this was not the case; we found no convincing correlation between wildtype and EC6-swapped α4 constructs by any of our surface delivery metrics (**Figure 4D, Figure S4J-K**). Instead, regardless of the identity of the EC6 domain, all the EC6-swapped α4 constructs had f^{+}, $\overline{Ab}$, and $E_{surface}$ values that were similar to α4ΔEC6 (**Figure 4D, Figure S4F-G**). These findings suggest that the surface delivery levels of EC6-swapped α4 variants do not depend on the isoform identity of EC6 (or its interactions), but rather on the absence of the α4 EC6 domain.

The γ-PCDH *cis* interface is insufficient to enable α-PCDH surface delivery

Our domain-swap results hinted that while EC6 plays a role in cPCDH surface delivery, its role may not depend on *cis* dimerization. This was surprising because a previous study used a loss-of-function approach—introducing disruptive mutations into the *cis* interface of representative β- and γ-PCDHs—to show that some, but not all, *cis* interface mutations prevented surface delivery in K562 cells (26, 28). This finding supported the conclusion that *cis* dimerization is necessary for

cPCDH surface delivery (21, 26, 28, 40). To clarify the role of cPCDH *cis* dimerization in surface delivery, we took a complementary gain-of-function approach—swapping β- and γ-PCDH amino acids into an α-PCDH at differentially conserved *cis* interface positions—to test whether *cis* interface residues were sufficient to increase surface delivery.

To identify differentially conserved *cis* interface positions, we aligned all wildtype mouse cPCDH isoforms using their extracellular sequences (EC1-6) and divided the sequences into two alignments according to their cell aggregation behavior: cell surface trafficking isoforms (β- and γ-PCDHs, αC2, γC3, and γC5) and non-trafficking isoforms (α-PCDHs, αC1, and γC4) (21). We calculated the amino acid frequencies at each position in EC6 separately for the two alignments and compared the resulting distributions using the Kullback-Leibler divergence (KL_{div}; **Figure 5A**). KL_{div} scores are high for positions that are conserved within each alignment but differ between the two alignments, and low for positions that are either variable or conserved across both alignments.

We mapped the KL_{div} scores of EC6-side *cis* interface residues onto the γB7 *cis* dimer structure (28) and the predicted mouse α4 model from the AlphaFold Protein Structure Database (53) (**Figure 5B**). In general, the residues in the top and center of the *cis* interface are strictly conserved among all cPCDHs, while differentially conserved positions are at the sides and bottom of the interacting surface, as previously noted (28). Several of the most differentially conserved positions in the *cis* interface have different sidechain properties (e.g., R595 in α4 versus V560 in γB7, or Y627 in α4 versus R590 in γB7; **Figure 5B**). To test whether these differences underlie subfamily-specific surface delivery, we introduced conserved β- and γ-PCDH *cis* interface residues into the corresponding positions of an α-PCDH.

We continued using α4 as a representative non-trafficking α-PCDH isoform. To make α4 more likely to *cis* homodimerize, we introduced residues from a γB-subfamily isoform because surface plasmon resonance measurements of cPCDH *cis* heterodimers suggested that γB-subfamily cPCDHs are the most amenable to participating as either the EC6-only or EC5-6 sides of the *cis* interface (29). We chose γB6 because it is highly similar to γB7, the only isoform for which a *cis* dimer structure is available (28), with 97% sequence identity in EC6. Of the 28 EC6 interface residues, 14 differ between α4 and γB6. We mutated the top five (α4-5$x_{\gamma B6}$), top ten (α4-10$x_{\gamma B6}$), and top twelve (α4-12$x_{\gamma B6}$) most differentially conserved amino acids in α4 to the corresponding γB6 amino acid and quantified the surface delivery of the resulting chimeric constructs (**Figure 5C**). Contrary to our hypothesis, none of the chimeras showed improved surface delivery compared to wildtype α4 (**Figure 5D-F**). To verify that the mutations did not destabilize the protein, we co-expressed α4 and the chimeric constructs with untagged γA11; all the variants were detected on the surface to similar extents (**Figure S5A-D**). Therefore, changing differentially conserved *cis* interface positions of α4 to match γB6 was insufficient to increase surface delivery compared to wildtype α4. This result aligns with our other EC6 and EC5-6 domain swap findings to suggest that *cis* dimerization does not control surface delivery.

We also generated α4 variants in which all the different EC6 interface residues (α4-14$x_{\gamma B6}$) and all the different EC5-6 interface residues (α4-22$x_{\gamma B6}$) were changed to the corresponding γB6 amino acid (**Figure 5C**). Neither of these constructs showed increased surface delivery compared to wildtype α4 (**Figure 5D-F**), but they also failed to reach the cell surface upon co-transfection with untagged γA11 (**Figure S5B-D**), despite being expressed similarly to surface trafficking constructs such as α4-γ$B6_{EC6}$ (**Figure S5E**). It is possible that these constructs do not fold properly, and that at least some α4 EC6 *cis* interface positions are important for the protein's stability.

Surface delivery is independent of *cis* dimerization affinity

Finally, we investigated the relationship between *cis* binding affinity and surface delivery in wildtype cPCDH isoforms. As our previous experiments relied on engineered α-PCDH variants, it remained possible that *cis* dimerization is important for surface trafficking of β- and γ-PCDHs. To test this, we compared the surface delivery of wildtype β- and γ-PCDHs to published *cis* homodimerization affinity values (or that of the closest paralog for which an affinity has been reported (22, 26-29); **Figure 6A)**. If *cis* dimerization enables surface delivery, we expect isoforms that bind with higher affinity to traffic better than isoforms with lower affinity, leading to an inverse correlation on a plot of K_D versus surface delivery. We did not observe a convincing correlation between K_D and any of our surface delivery metrics (**Figure 6B-D)**. Surprisingly, the surface delivery of the strongest *cis*-binding isoform, αC2, was not significantly different from α4, suggesting that along with α4, αC2 does not traffic to the cell membrane despite its strong *cis* homodimerization affinity (7-9 μM (22, 29); **Figure 6E-G**). Another isoform, γC3, did not *cis* homodimerize in solution (26) and yet had high levels of surface protein (**Figure 6B-D, triangles**). As a counterexample, the *cis* interacting fragment of a β17 paralog (β9) forms a moderately strong dimer in solution (35 μM) (29), and β17 had strong cell surface trafficking (**Figure 6B-D, blue crosses**). The remaining isoforms we tested, γB6, γB2, γA3, and γA11, have affinities that range from unmeasurably low to moderate affinities of around 60 μM (26, 29), but trafficked at similar levels (**Figure 6B-D**).

Overall, we found *cis* dimerization affinity to be a poor predictor of surface delivery behavior for wildtype cPCDH isoforms. By comparing our different quantification metrics, we were able to rank the wildtype cPCDH isoforms by their overall surface delivery behaviors. While the γA and

γB isoforms had similar f^{+}, for $\overline{Ab}$ γB2 and γA11 tended to be slightly higher than γB6 and γA3, suggesting that there was usually more protein on the surface of γA11- and γB2-transfected cells (**Figure 6E-F**). For the isoforms with the strongest surface delivery, γC3 and β17, $\overline{Ab}$ was overlapping and highly variable between experiments (**Figure 6E-G**). The quantification metrics thus enabled nuanced comparisons between cPCDH isoforms because they capture distinct aspects of surface delivery, which may be particularly useful for specific applications. For screening constructs to maximize secreted protein levels for purification, high f^{+} and $\overline{Ab}$ are the most important considerations. For assessing how well different variants traverse the secretory pathway, $E_{surface}$ is more informative.

Discussion

In heterologous expression systems, surface delivery is a useful metric to validate proper surface localization, to account for differences in surface protein levels, and to identify protein modifications that disrupt trafficking. We developed a robust method to quantify the surface delivery of plasma membrane proteins using epitope-tag-based surface staining and fluorescence microscopy. Our metrics describe three axes of variation common to any surface staining experiment: the fraction of cells with detectable levels of surface protein, relative surface protein abundance, and relative surface trafficking efficiency. These metrics are compatible with data acquired either by microscopy or flow cytometry, and our scripts are available for researchers to modify or apply directly to their own datasets. To validate our approach, we applied it to cPCDHs and found that the quantification recapitulated known subfamily-specific trafficking trends: β- and γ-PCDHs reach the plasma membrane, while α-PCDHs do not (21, 34, 45). We also found that while removing EC6 from PCDHα4 enabled its surface localization, replacing α4's EC6 or EC5-

6 with β- or γ-PCDH domains imparted no additional increase in surface trafficking. Our measurements further showed how trafficking levels varied among wildtype and engineered cPCDH variants, enabling analyses that tested the relationship between cPCDH *cis* dimerization and surface delivery.

We focused on *cis* dimerization because *cis* dimer-disrupting mutations in γB6 and γB7, including mutations introducing α-specific residues, prevented surface trafficking of those variants (26, 28). We were therefore surprised to find that the inverse experiment, incorporating γB6 *cis* interface residues into α4, was insufficient to enable surface delivery. Revisiting the previous work, we noted that of eight γB6 *cis* interface residues independently mutated to aspartic acid, three interfered with cell aggregation (L558D, V563D, and R598D) while five had no effect (Y535D, M548D, K561D, W574D, R593D) (26), either because the mutations did not disrupt *cis* dimerization or because *cis* dimerization is not required for surface delivery. Two other *cis* interface mutations, V563R in γB6 and A570R in γB7, disrupted *cis* dimerization in solution but did not impede cell aggregation (28, 29), either because *cis* dimerization was maintained in the two-dimensional membrane context or because *cis* dimerization is not required for surface delivery. Conversely, the V560R mutation in γB7, which is equivalent to V563R in γB6, did impede cell aggregation, as did the double mutations V563R/Y535G and V563R/R598S, but not V563R/R593Y, in γB6 and their equivalents in β17 (28). Interpreting the effects of these loss-of-function mutations is also challenging because a lack of cell aggregation can be due to effects other than impeded surface delivery, such as altered protein expression or misfolding. Here, we found that the identity of the *cis* dimerization regions and their dimerization properties—EC6 identity for domain-swapped α-PCDH constructs and binding affinity for wildtype isoforms—were not

correlated with direct surface trafficking measurements, supporting the interpretation that surface trafficking does not depend on *cis* dimerization.

If *cis* dimerization does not control cPCDH surface delivery, then how is it regulated? Previous research with α4ΔEC6 led to the proposal that the α-PCDH EC6 domain inhibits surface delivery (21), and our finding that replacing EC6 in α4 with EC6 from a β-, γ-, or C-type cPCDH did not improve surface delivery compared to α4ΔEC6 substantiates that idea. Investigations into the cellular mechanisms of cPCDH trafficking suggest potential alternative models for how surface delivery might be controlled independently of *cis* dimerization. The EC domains of cPCDHs are O-mannosylated, an unusual type of glycosylation in metazoans catalyzed by four glycosyltransferases dedicated to cadherin superfamily proteins, including cPCDHs (54). As protein glycosylation and folding are tightly linked quality control mechanisms in the endoplasmic reticulum, differences in glycosylation site motifs or interactions with these glycosyltransferases could represent a rate-limiting step that determines the trafficking efficiency of a particular cPCDH isoform. We observed that even in the absence of the α-PCDH EC6 domain, EC6- and EC5-6-swapped α4 constructs trafficked less efficiently than wildtype β- and γ-PCDHs, suggesting that differences affecting surface delivery remain in the unchanged parts of the chimeric variants, including potential O-mannosylation sites in EC1-4. Although we focused on extracellular interactions here, a motif in the variable intracellular region of γ-PCDHs has been shown to regulate their intracellular trafficking in a ubiquitinylation- and phosphorylation-dependent manner in 293T cells (38, 39, 43, 44). Finally, correlative light and electron microscopy studies have identified specialized intracellular compartments that appear when γ-PCDHs, but not α-PCDHs, are overexpressed in 293T cells, raising the possibility that α- and γ-PCDHs may be trafficked by distinct intracellular pathways (39, 41). Quantification of surface delivery with the

tools presented here will facilitate future investigations of the complex cellular regulation of cPCDH surface delivery.

In the cPCDH field, cell aggregation assays have played a crucial role in deciphering the molecular rules of cPCDH-mediated cell recognition (20-22, 26, 28). Perhaps most importantly, sorting occurs in cell aggregation assays, with cells preferring to bind others with completely matching sets of cPCDH isoforms, the same scenario that leads to self-recognition and avoidance in neurons (21, 22). In these experiments, the extent of cell aggregation depended on relative isoform expression levels (20, 21). Direct measurements of surface trafficking could provide insight into the relative surface concentrations of different isoforms in co-transfection experiments and distinguish how each isoform contributes to the energetics of competitive binding.

Beyond cPCDHs, our method provides a general-purpose, quantitative way to screen plasma membrane proteins and variants of interest for surface delivery or secretion. The data we extract from our microscopy images is analogous to flow cytometry data, and indeed, the methods are complementary. For attached cells, tissues, and adhesion protein studies, microscopy is preferable to avoid disrupting surface proteins when detaching cells from the culture vessel and each other. For suspended cells, flow cytometry has the advantage of being extremely high-throughput and therefore can detect rare events. Our surface delivery summary metrics are compatible with either approach. Cell surface proteins including ion channels, G-protein coupled receptors, and transporters represent highly important drug targets, and their binding pockets, gating and regulatory mechanisms are often mapped through structure determination and mutational scans (55-62). Membrane protein instability and mis-localization are common in heterologous expression systems (60, 63), and few screening studies validate loss-of-function hits by evaluating surface delivery (56, 58, 59, 62). Implementing quantitative surface delivery analysis in

combination with activity screens will facilitate clearer interpretation of these experiments and improve our investigations of surface protein functions.

Materials and Methods

Differential conservation sequence analysis

The sequences of all mouse cPCDHs were manually gathered from UniProt (release 2023_05) (64), aligned using MUSCLE v.5.1 (65), and truncated to EC6 (Geneious Prime 2023.1.2 [http://www.geneious.com]). The subsequent alignment manipulations were performed in Python 3 using bioviper v.0.20 (66). The alignment was separated into two sets of sequences: α-PCDHs, αC1, and γC4 in one set, and β-PCDHs, γ-PCDHs, αC2, γC3, and γC5 in the other set. For each sub-alignment, the amino acid frequencies were calculated and the frequency distributions at each position were compared between the two alignments using the entropy function in scipy v.1.11.4 (67) to calculate the Kullback-Leibler divergence, which we refer to as the KL_{div} score.

Molecular cloning

The α4, α4ΔEC6, β17, αC2, γC3, and untagged γC3 plasmids were provided by Tom Maniatis's lab (21). All other plasmids were cloned into the pmax-mCherry backbone of the α4 plasmid using Gibson assembly (Gibson Assembly Master Mix, NEB E2611S). Inserts were generated by DNA synthesis (IDT gBlocks or eBlocks) or PCR (Q5 polymerase, NEB M0491S). The complete plasmid sequence for α4 and the full coding region sequences for all other plasmids are provided in .fasta format on Zenodo (https://doi.org/10.5281/zenodo.13345292). For all plasmids generated, DNA from single clones was purified (E.Z.N.A. Plasmid DNA Mini Kit I, Omega Bio-Tek D6942)

and sequenced (Whole Plasmid Sequencing, Plasmidsaurus) to verify the integrity of the constructs.

Cell culture and transfection

The 293T N-cadherin knock-out cell line (293T-ΔNC) (52) was provided by Joshua Sanes's lab. Cells were cultured in a 1:1 mixture of Dulbecco's Modified Eagle's Medium and Ham's F-12 medium containing L-glutamine and without phenol red (DMEM/F-12; Corning, 16-405-CV) supplemented with 10% fetal bovine serum (FBS; Corning, 35-011-CV) and 1X non-essential amino acids (NEAA; Lonza, 13-114E). Cells were maintained in exponential growth by passaging every 3-4 days.

A day before transfection, confluent 293T-ΔNC cells were treated with a solution of 0.5 g/L trypsin and 0.2 g/L ethylene diaminetetraacetic acid (EDTA) without calcium or magnesium (Lonza; 17-161E) at 37°C to detach and dissociate the cells. After five minutes, the trypsinization was stopped by adding an excess of fresh medium and the cells were centrifuged for five minutes at 225 x g. The cell pellet was resuspended in fresh medium, and cells were seeded at 100,000 cells/well in a total volume of 400 µL in glass-bottom culture chambers (Cellvis, C8-1.5-N). The cells adhered to the culture chambers overnight in an incubator maintained at 37°C with 5% CO_2.

Stocks of DNA for transfections were prepared by diluting plasmids to 50 ng/µL with DNA elution buffer (10 mM Tris-HCl pH 8.5; Omega Bio-Tek, D6942-02) so that each transfection reaction mix contained the same amount of DNA and salts. Transfection was carried out using LipoFectMax 3000 Transfection Reagent (ABP Biosciences, FP318). For each sample, one well of cells was transfected with 0.3 µg DNA in 25 µL total volume of transfection mixture containing a 1:3:2 mass ratio of DNA, Component A, and Component B in Opti-MEM I Reduced Serum Medium (Opti-MEM; Gibco, 31985-70) per transfection reaction. The transfection reaction mixtures were

incubated at room temperature for 10-15 minutes, then added directly to the medium of the cells in the culture chambers. Transfection proceeded overnight in a 37°C, 5% CO_2 incubator.

Sample preparation and staining

Live cells were stained 20-24 hours post-transfection. A FITC-conjugated antibody recognizing the Myc epitope tag (Miltenyi Biotec, 130-116-485) was diluted 1:50 in fresh medium. Three hundred and fifty microliters of culture medium were removed from each well and replaced with 50 µL diluted antibody solution (final antibody working concentration: 1:100 dilution). Unstained samples were treated identically, except that the spent culture medium was replaced with 50 µL fresh medium instead of antibody solution. Samples were incubated at 37°C, 5% CO_2 for one hour, then 300 µL fresh medium were added to each well to dilute the antibody. Two sets of three washes were subsequently performed, with a 30-minute incubation at 37°C, 5% CO_2 between the two sets of washes. For each wash, 200 µL medium were removed from each well and replaced with 200 µL fresh medium (antibody diluted 2X per wash; final antibody dilution 256X after all washes). Cells were imaged live immediately following the final wash.

Microscopy

Samples were imaged using a Zeiss LSM 880 laser scanning confocal microscope using a Plan-Apochromat 10x/0.45 NA air objective. The sample was brought into focus and the pinhole was set to correspond to a 16 µm optical section, or approximately the height of a cell. Three fields of view were manually defined using only the brightfield channel for each sample well, and the fields of view were acquired sequentially using automated stage positioning with hardware autofocus to maintain the same focal plane at each position. Images were acquired in bidirectional scanning mode with a 1 µsec pixel dwell time and 4x line-by-line averaging. The two fluorescence channels were acquired sequentially on a PMT detector, and the brightfield channel was acquired in

transmission mode. For the transfection marker mCherry (FP) channel, we used a 561 nm excitation laser and collected emitted light of wavelengths 578-696 nm. For the FITC-conjugated Myc antibody (Ab) channel, we used a 488 nm excitation laser and collected emitted light of wavelengths 493-592 nm.

Image processing and quantification

Image processing was performed in Python 3 using czifile v.2019.7.2 to load the images and metadata. Cells were segmented in the brightfield channel with cellpose v.2.2.3 (68) using a GPU on the FASRC Cannon cluster supported by the FAS Division of Science Research Computing Group at Harvard University. For both fluorescence channels, the background was estimated for each individual image as the median intensity of all non-cell pixels and subtracted. FP_{cell} and Ab_{cell} were calculated with the single_cell.average function in microutil v.0.4.0 (69). Cells containing any saturated pixels were excluded from downstream analysis. The transfection threshold was set as the 0.995^{th} quantile of the pooled FP_{cell} values from the mock-transfected samples in all biological replicate experiments. Thus, the probability that a cell that we designated as transfected (having an FP_{cell} value greater than the transfection threshold) was actually not transfected was less than 0.005. The surface stain threshold was calculated similarly, as the 0.995^{th} quantile of pooled Ab_{cell} values from samples transfected with untagged γC3 or untagged γA11. Histograms were calculated and displayed after applying a logicle transform (70) using FlowKit v.1.0.1 (71) and seaborn v.0.13.0 (72). For each sample, f^+ was calculated as the fraction of transfected cells with Ab_{cell} values higher than the surface stain threshold, and $\overline{Ab}$ was calculated as the average of all transfected cell Ab_{cell} values. To calculate $E_{surface}$, log-spaced bins were defined using the minimum and maximum FP_{cell} values for all transfected cells (across biological replicate experiments and samples). Cells were assigned to bins based on their FP_{cell} values, and FP_{bin} and

Ab_{bin} were then calculated individually for each sample as the mean FP_{cell} and Ab_{cell} values per bin, respectively. $E_{surface}$ was calculated as the slope of the least-squares linear regression of the log of FP_{bin} and Ab_{bin} using the linregress function in scipy (67), excluding bins with negative Ab_{bin} values (this only occurred for a few bins in untagged isoform samples). P-values were calculated using one- or two-tailed Welch's unequal variances t-test or one-way ANOVA, as indicated in the text or figure legends.

Acknowledgements

The authors thank Barry Honig, Tom Maniatis, and Erin Flaherty for suppling plasmids, Josh Sanes for supplying 293T-ΔNC cells, and John Russell, José Velilla, and Sam Berry for helpful feedback on the manuscript. We also thank Brenda Chiang and Nhu Dang for early work on the project. R.G. acknowledges funding from the Harvard Brain Science Initiative Bipolar Disorder Seed Grant. E.J.M. thanks the Aramont Fund for Emerging Science Research and the Simmons family for generous funding through the Simmons Awards. We also thank the Harvard Physics of Living Systems grant for funding and the Harvard Center for Biological Imaging (RRID:SCR_018673) for infrastructure and support.

Data availability:

Microscopy datasets, plasmid sequences, alignment files, and raw data files containing values for the figures in the manuscript are available on Zenodo (https://doi.org/10.5281/zenodo.13345292). Python scripts are available on GitHub (https://github.com/emay2022/surface-trafficking). Plasmids are available from the authors upon reasonable request.

References

1. K. Sano *et al.*, Protocadherins: a large family of cadherin-related molecules in central nervous system. *EMBO J* **12**, 2249-2256 (1993).
2. N. Kohmura *et al.*, Diversity revealed by a novel family of cadherins expressed in neurons at a synaptic complex. *Neuron* **20**, 1137-1151 (1998).
3. S. Obata *et al.*, A common protocadherin tail: multiple protocadherins share the same sequence in their cytoplasmic domains and are expressed in different regions of brain. *Cell Adhes Commun* **6**, 323-333 (1998).
4. Q. Wu, T. Maniatis, A striking organization of a large family of human neural cadherin-like cell adhesion genes. *Cell* **97**, 779-790 (1999).
5. X. Wang *et al.*, Gamma protocadherins are required for survival of spinal interneurons. *Neuron* **36**, 843-854 (2002).
6. W. R. Mancia Leon *et al.*, Clustered gamma-protocadherins regulate cortical interneuron programmed cell death. *Elife* **9**, e55374 (2020).
7. J. A. Weiner, X. Wang, J. C. Tapia, J. R. Sanes, Gamma protocadherins are required for synaptic development in the spinal cord. *Proc Natl Acad Sci U S A* **102**, 8-14 (2005).
8. J. L. Lefebvre, Y. Zhang, M. Meister, X. Wang, J. R. Sanes, gamma-Protocadherins regulate neuronal survival but are dispensable for circuit formation in retina. *Development* **135**, 4141-4151 (2008).
9. J. L. Lefebvre, D. Kostadinov, W. V. Chen, T. Maniatis, J. R. Sanes, Protocadherins mediate dendritic self-avoidance in the mammalian nervous system. *Nature* **488**, 517-521 (2012).
10. M. J. Molumby, A. B. Keeler, J. A. Weiner, Homophilic Protocadherin Cell-Cell Interactions Promote Dendrite Complexity. *Cell Rep* **15**, 1037-1050 (2016).
11. G. Mountoufaris *et al.*, Multicluster Pcdh diversity is required for mouse olfactory neural circuit assembly. *Science* **356**, 411-414 (2017).
12. D. M. Steffen *et al.*, A Unique Role for Protocadherin gammaC3 in Promoting Dendrite Arborization through an Axin1-Dependent Mechanism. *J Neurosci* **43**, 918-935 (2023).
13. A. M. Garrett, D. Schreiner, M. A. Lobas, J. A. Weiner, gamma-protocadherins control cortical dendrite arborization by regulating the activity of a FAK/PKC/MARCKS signaling pathway. *Neuron* **74**, 269-276 (2012).
14. A. B. Keeler, D. Schreiner, J. A. Weiner, Protein Kinase C Phosphorylation of a gamma-Protocadherin C-terminal Lipid Binding Domain Regulates Focal Adhesion Kinase Inhibition and Dendrite Arborization. *J Biol Chem* **290**, 20674-20686 (2015).
15. H. Sugino *et al.*, Genomic organization of the family of CNR cadherin genes in mice and humans. *Genomics* **63**, 75-87 (2000).
16. Q. Wu *et al.*, Comparative DNA sequence analysis of mouse and human protocadherin gene clusters. *Genome Res* **11**, 389-404 (2001).
17. Q. Wu, T. Maniatis, Large exons encoding multiple ectodomains are a characteristic feature of protocadherin genes. *Proc Natl Acad Sci U S A* **97**, 3124-3129 (2000).
18. B. Tasic *et al.*, Promoter choice determines splice site selection in protocadherin alpha and gamma pre-mRNA splicing. *Mol Cell* **10**, 21-33 (2002).
19. S. Obata *et al.*, Protocadherin Pcdh2 shows properties similar to, but distinct from, those of classical cadherins. *J Cell Sci* **108 (Pt 12)**, 3765-3773 (1995).
20. D. Schreiner, J. A. Weiner, Combinatorial homophilic interaction between gamma-protocadherin multimers greatly expands the molecular diversity of cell adhesion. *Proc Natl Acad Sci U S A* **107**, 14893-14898 (2010).

21. C. A. Thu *et al.*, Single-cell identity generated by combinatorial homophilic interactions between alpha, beta, and gamma protocadherins. *Cell* **158**, 1045-1059 (2014).
22. R. Rubinstein *et al.*, Molecular logic of neuronal self-recognition through protocadherin domain interactions. *Cell* **163**, 629-642 (2015).
23. J. M. Nicoludis *et al.*, Structure and Sequence Analyses of Clustered Protocadherins Reveal Antiparallel Interactions that Mediate Homophilic Specificity. *Structure* **23**, 2087-2098 (2015).
24. J. M. Nicoludis *et al.*, Antiparallel protocadherin homodimers use distinct affinity- and specificity-mediating regions in cadherin repeats 1-4. *Elife* **5** (2016).
25. J. M. Nicoludis *et al.*, Interaction specificity of clustered protocadherins inferred from sequence covariation and structural analysis. *Proc Natl Acad Sci U S A* **116**, 17825-17830 (2019).
26. K. M. Goodman *et al.*, gamma-Protocadherin structural diversity and functional implications. *Elife* **5** (2016).
27. K. M. Goodman *et al.*, Structural Basis of Diverse Homophilic Recognition by Clustered alpha- and beta-Protocadherins. *Neuron* **90**, 709-723 (2016).
28. K. M. Goodman *et al.*, Protocadherin cis-dimer architecture and recognition unit diversity. *Proc Natl Acad Sci U S A* **114**, E9829-E9837 (2017).
29. K. M. Goodman *et al.*, How clustered protocadherin binding specificity is tuned for neuronal self-/nonself-recognition. *Elife* **11**, e72416 (2022).
30. J. Brasch *et al.*, Visualization of clustered protocadherin neuronal self-recognition complexes. *Nature* **569**, 280-283 (2019).
31. M. Fernandez-Monreal, S. Kang, G. R. Phillips, Gamma-protocadherin homophilic interaction and intracellular trafficking is controlled by the cytoplasmic domain in neurons. *Mol Cell Neurosci* **40**, 344-353 (2009).
32. N. Boni, L. Shapiro, B. Honig, Y. Wu, R. Rubinstein, On the formation of ordered protein assemblies in cell-cell interfaces. *Proc Natl Acad Sci U S A* **119**, e2206175119 (2022).
33. G. R. Phillips *et al.*, Gamma-protocadherins are targeted to subsets of synapses and intracellular organelles in neurons. *J Neurosci* **23**, 5096-5104 (2003).
34. Y. Murata, S. Hamada, H. Morishita, T. Mutoh, T. Yagi, Interaction with protocadherin-gamma regulates the cell surface expression of protocadherin-alpha. *J Biol Chem* **279**, 49508-49516 (2004).
35. M. Frank *et al.*, Differential expression of individual gamma-protocadherins during mouse brain development. *Mol Cell Neurosci* **29**, 603-616 (2005).
36. S. Kallenbach *et al.*, Changes in subcellular distribution of protocadherin gamma proteins accompany maturation of spinal neurons. *J Neurosci Res* **72**, 549-556 (2003).
37. M. Fernandez-Monreal *et al.*, gamma-protocadherins are enriched and transported in specialized vesicles associated with the secretory pathway in neurons. *Eur J Neurosci* **32**, 921-931 (2010).
38. R. O'Leary *et al.*, A variable cytoplasmic domain segment is necessary for gamma-protocadherin trafficking and tubulation in the endosome/lysosome pathway. *Mol Biol Cell* **22**, 4362-4372 (2011).
39. G. R. Phillips, N. LaMassa, Y. M. Nie, Clustered protocadherin trafficking. *Semin Cell Dev Biol* **69**, 131-139 (2017).

40. R. Rubinstein, K. M. Goodman, T. Maniatis, L. Shapiro, B. Honig, Structural origins of clustered protocadherin-mediated neuronal barcoding. *Semin Cell Dev Biol* **69**, 140-150 (2017).
41. H. H. Hanson *et al.*, LC3-dependent intracellular membrane tubules induced by gamma-protocadherins A3 and B2: a role for intraluminal interactions. *J Biol Chem* **285**, 20982-20992 (2010).
42. S. S. Schalm, B. A. Ballif, S. M. Buchanan, G. R. Phillips, T. Maniatis, Phosphorylation of protocadherin proteins by the receptor tyrosine kinase Ret. *Proc Natl Acad Sci U S A* **107**, 13894-13899 (2010).
43. A. Shonubi, C. Roman, G. R. Phillips, The clustered protocadherin endolysosomal trafficking motif mediates cytoplasmic association. *BMC Cell Biol* **16**, 28 (2015).
44. A. Ptashnik *et al.*, Ubiquitination of the protocadherin-gammaA3 variable cytoplasmic domain modulates cell-cell interaction. *Front Cell Dev Biol* **11**, 1261048 (2023).
45. T. Mutoh, S. Hamada, K. Senzaki, Y. Murata, T. Yagi, Cadherin-related neuronal receptor 1 (CNR1) has cell adhesion activity with beta1 integrin mediated through the RGD site of CNR1. *Exp Cell Res* **294**, 494-508 (2004).
46. H. Sago *et al.*, Cloning, expression, and chromosomal localization of a novel cadherin-related protein, protocadherin-3. *Genomics* **29**, 631-640 (1995).
47. M. Blank, G. B. Triana-Baltzer, C. S. Richards, D. K. Berg, Alpha-protocadherins are presynaptic and axonal in nicotinic pathways. *Mol Cell Neurosci* **26**, 530-543 (2004).
48. G. B. Triana-Baltzer, M. Blank, Cytoplasmic domain of protocadherin-alpha enhances homophilic interactions and recognizes cytoskeletal elements. *J Neurobiol* **66**, 393-407 (2006).
49. S. Bonn, P. H. Seeburg, M. K. Schwarz, Combinatorial expression of alpha- and gamma-protocadherins alters their presenilin-dependent processing. *Mol Cell Biol* **27**, 4121-4132 (2007).
50. I. G. Haas, M. Frank, N. Veron, R. Kemler, Presenilin-dependent processing and nuclear function of gamma-protocadherins. *J Biol Chem* **280**, 9313-9319 (2005).
51. B. Hambsch, V. Grinevich, P. H. Seeburg, M. K. Schwarz, gamma-Protocadherins, presenilin-mediated release of C-terminal fragment promotes locus expression. *J Biol Chem* **280**, 15888-15897 (2005).
52. K. M. Goodman *et al.*, Molecular basis of sidekick-mediated cell-cell adhesion and specificity. *Elife* **5** (2016).
53. J. Jumper *et al.*, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).
54. I. S. B. Larsen *et al.*, Discovery of an O-mannosylation pathway selectively serving cadherins and protocadherins. *Proc Natl Acad Sci U S A* **114**, 11163-11168 (2017).
55. C. B. Billesbolle *et al.*, Structural basis of odorant recognition by a human odorant receptor. *Nature* **615**, 742-749 (2023).
56. H. J. Ellis *et al.*, Deep Mutagenesis of a Transporter for Uptake of a Non-Native Substrate Identifies Conformationally Dynamic Regions [Preprint]. *bioRxiv* 10.1101/2021.04.19.440442, 2021.2004.2019.440442 (2024).
57. A. M. Glazer *et al.*, Deep Mutational Scan of an SCN5A Voltage Sensor. *Circ Genom Precis Med* **13**, e002786 (2020).
58. J. D. Heredia *et al.*, Mapping Interaction Sites on Human Chemokine Receptors by Deep Mutational Scanning. *J Immunol* **200**, 3825-3839 (2018).

59. W. Coyote-Maestas, D. Nedrud, Y. He, D. Schmidt, Determinants of trafficking, conduction, and disease within a K(+) channel revealed through multiparametric deep mutational scanning. *Elife* **11** (2022).
60. S. W. Yee *et al.*, The full spectrum of SLC22 OCT1 mutations illuminates the bridge between drug transporter biophysics and pharmacogenomics. *Mol Cell* **84**, 1932-1947 e1910 (2024).
61. E. M. Jones *et al.*, Structural and functional characterization of G protein-coupled receptors with deep mutational scanning. *Elife* **9** (2020).
62. M. K. Howard *et al.*, Molecular basis of proton-sensing by G protein-coupled receptors [Preprint]. *bioRxiv* 10.1101/2024.04.17.590000 (2024).
63. K. Ikegami *et al.*, Structural instability and divergence from conserved residues underlie intracellular retention of mammalian odorant receptors. *Proc Natl Acad Sci U S A* **117**, 2957-2967 (2020).
64. C. UniProt, UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res* **51**, D523-D531 (2023).
65. R. C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792-1797 (2004).
66. S. Berry, S. Ganesh (2023) bioviper. (Zenodo, https://zenodo.org/doi/10.5281/zenodo.12549380).
67. P. Virtanen *et al.*, SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* **17**, 261-272 (2020).
68. C. Stringer, T. Wang, M. Michaelos, M. Pachitariu, Cellpose: a generalist algorithm for cellular segmentation. *Nat Methods* **18**, 100-106 (2021).
69. J. Russell, I. Hunt-Isaak, D. Hekstra (2021) microutil. (GitHub, https://github.com/Hekstra-Lab/microutil).
70. W. A. Moore, D. R. Parks, Update for the logicle data scale including operational code implementations. *Cytometry A* **81**, 273-277 (2012).
71. S. White *et al.*, FlowKit: A Python Toolkit for Integrated Manual and Automated Cytometry Analysis Workflows. *Front Immunol* **12**, 768541 (2021).
72. M. L. Waskom, seaborn: statistical data visualization. *Journal of Open Source Software* **6**, 3021 (2021).

Table 1: Isoforms and variants used and the cPCDH subfamilies they represent

Subfamily	Isoform	EC6-swapped α4 variant	EC56-swapped α4 variant
α	α4*	α4ΔEC6†	-
β	β17†	α4-β17$_{EC6}$†	α4-β17$_{EC5-6}$
γA	γA3‡	α4-γA3$_{EC6}$	-
	γA11‡	α4-γA11$_{EC6}$	α4-γA11$_{EC5-6}$
γB	γB2‡	α4-γB2$_{EC6}$	-
	γB6†	α4-γB6$_{EC6}$†	α4-γB6$_{EC5-6}$‡
C-type isoforms (polyphyletic group)	αC2‡	α4-αC2$_{EC6}$	-
	γC3*	α4-γC3$_{EC6}$†	α4-γC3$_{EC5-6}$

* When included in subsequent figures, data reproduced from **Figures 2 and S1**
† When included in subsequent figures, data reproduced from **Figure 3**
‡ When included in subsequent figures, data reproduced from **Figures 4 and S4**

Main Text Figures and Legends

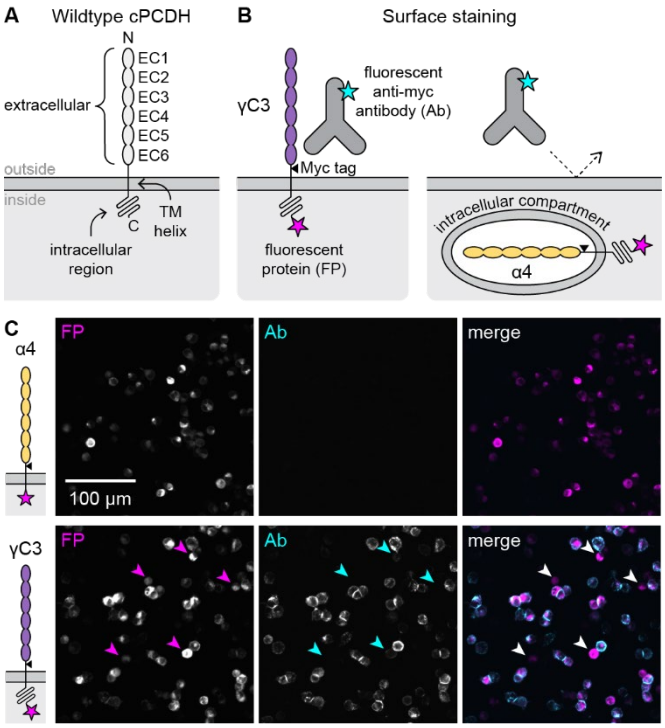

Figure 1. Surface staining by extracellular epitope detection of cPCDHs in 293T-ΔNC cells is the basis of surface delivery quantification. (A) Schematic showing the domain arrangement and topology of a generic cPCDH isoform. From N to C terminus: six extracellular cadherin (EC) domains are tethered to the membrane by a transmembrane helix (TM) followed by an unstructured intracellular region. (B) Diagram of experimental approach. cPCDH isoforms modified to contain an extracellular Myc epitope tag (triangle) and a C-terminal fluorescent protein (magenta star) are expressed in 293T-ΔNC cells. A fluorescent (cyan star) anti-Myc antibody recognizes the Myc tag to detect proteins that reach the cell surface, like γC3. The antibody does not bind proteins that are retained intracellularly, like α4, because it does not cross the plasma membrane. (C) Microscopy images of surface-stained control samples. Top row: An α-PCDH isoform, α4, is produced in transfected 293-ΔNC cells (signal in FP channel) but does not reach the cell surface (no signal in Ab channel). Bottom row: A C-type isoform, γC3, is produced in transfected 293T-ΔNC cells (signal in FP channel), and the anti-Myc antibody stains the periphery of cells with plasma membrane-localized protein (Ab channel). Arrowheads in γC3 images indicate examples where transfected cells have little to no Ab staining, highlighting the heterogeneity of surface trafficking.

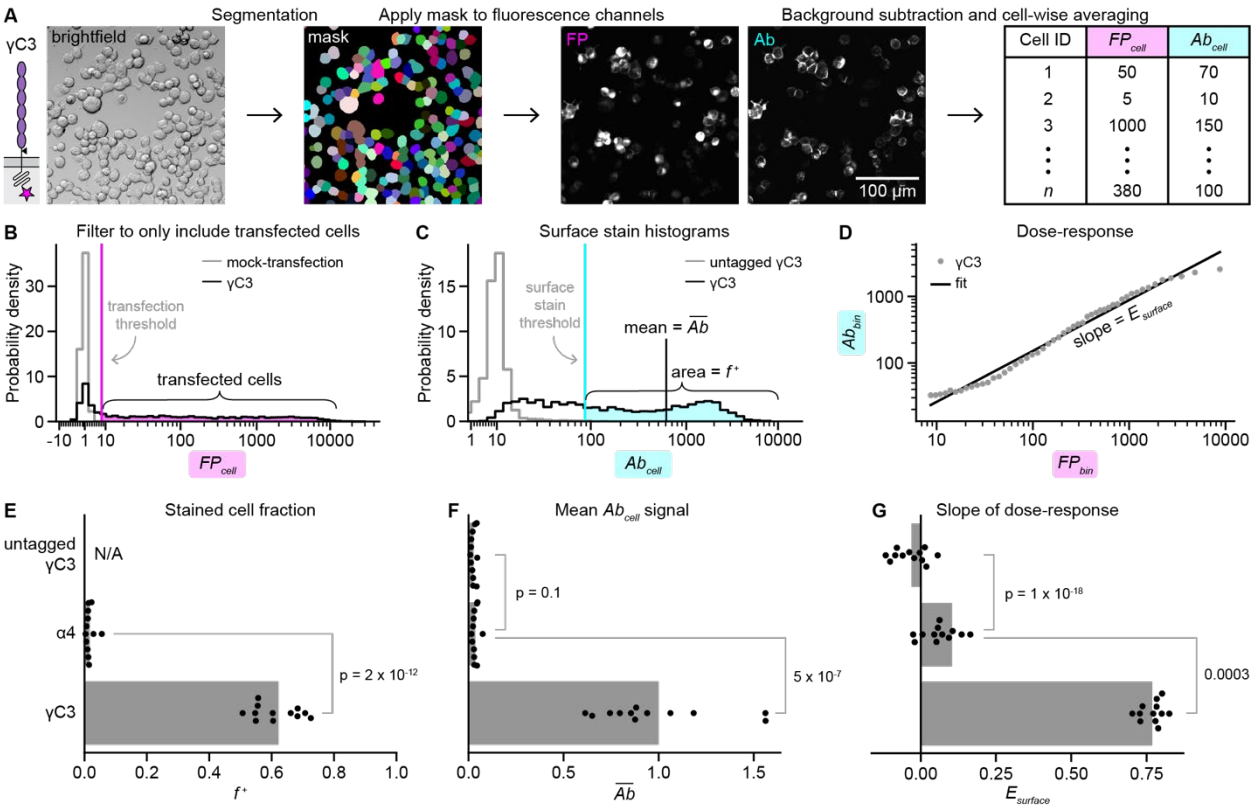

Figure 2. Quantification reveals the heterogeneity of surface trafficking across cells within a single sample and facilitates comparisons between samples. (A) Initial processing steps using images from an example sample, γC3. Automated cell segmentation using the brightfield channel from a particular field of view produces a mask that assigns each pixel in the field of view to a cell (or background). To perform background subtraction, we applied the mask to the FP and Ab images from the same field of view to assign background pixels, and we subtracted the average background pixel values across the image. Cell-wise averaging of background-subtracted pixel intensities produces FP_{cell} and Ab_{cell} values for each of n cells in the field of view. (B) A histogram of FP_{cell} values, where FP_{cell} is a measure of cellular protein levels of the transfected cPCDH. The data are filtered to only include transfected cells (magenta highlight) using a transfection threshold (magenta vertical line) determined using a sample of mock-transfected cells. (C) A histogram of the filtered Ab_{cell} values, where Ab_{cell} a measure of surface abundance of the transfected cPCDH. The mean of the entire Ab_{cell} distribution, $\overline{Ab}$ (black vertical line), corresponds to the average surface stain signal for all transfected cells in the sample. The fraction of Ab-stained cells, f^{+} (cyan highlight), is calculated using a threshold Ab_{cell} value (cyan vertical line) determined using cells transfected with an untagged version of γC3. (D) Dose response plot of the γC3 data. The data are binned by FP_{cell}, and the means of each bin, Ab_{bin} and FP_{bin}, are plotted against each other (gray points). The bins are scaled to maintain the same number of cells in each bin, which results in approximately log-scaled bins. We define $E_{surface}$ as the slope of a linear fit to the data (black line). (E-G) Graphs of f^{+} (E), $\overline{Ab}$ (F), and $E_{surface}$ (G) for untagged γC3, α4, and γC3. Untagged γC3 is excluded from the f^{+} graph (E) because it is used to define the threshold Ab_{cell} value for determining the fraction of stained cells. In panel F, $\overline{Ab}$ values are normalized to the mean $\overline{Ab}$ of Myc-tagged γC3. Points are values for biological replicates (n = 12 for each sample), and bars show the mean values across all experiments. P-values were calculated using a one-tailed Welch's t-test.

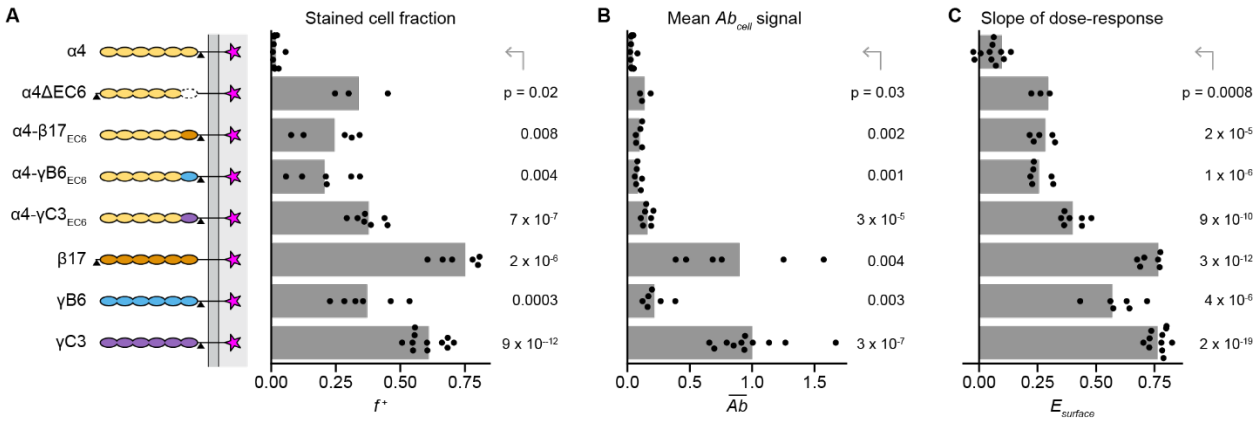

Figure 3. Quantification can distinguish samples that have intermediate surface delivery phenotypes. In construct schematics, domains are color-coded by originating isoform. (A-C) Quantification of f^+ (A), $\overline{Ab}$ (B), and $E_{surface}$ (C) for the constructs shown in A. In B, $\overline{Ab}$ values are normalized to the mean $\overline{Ab}$ of the positive control construct, Myc-tagged γC3. Points are biological replicates (n = 11 for α4 and γC3; 3 for α4ΔEC6; 5 for α4-β17$_{EC6}$; 6 for α4-γB6$_{EC6}$, β17, and γB6; and 7 for α4-γC3$_{EC6}$), and bars are mean values across replicates. P-values for each sample compared to α4 (top row) were calculated using a one-tailed Welch's t-test.

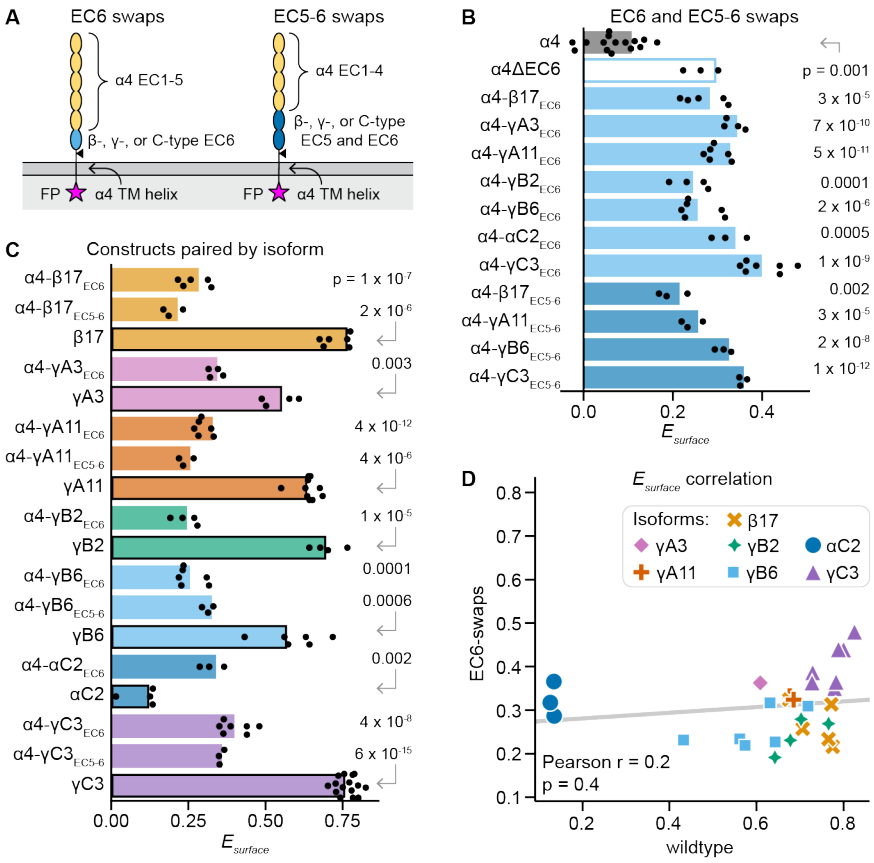

Figure 4. Removing or replacing EC6 of α4 increases its surface delivery, but not to wildtype β- and γ-PCDH levels. (A) Schematics of the types of constructs compared in B-D. EC6-swapped variants have EC1-5 and the transmembrane helix of α4, and EC6 of a different isoform. EC5-6 swaps have EC1-4 and the transmembrane helix of α4, and EC6 of a different isoform. (B) Graph comparing $E_{surface}$ for wildtype α4 (gray bar) to α4 variants with EC6 deleted (unfilled light blue bar), EC6 swapped (lighter blue bars), and EC5-6 swapped (darker blue bars). Points are biological replicates (n = 15 for α4 and γC3; 3 for α4ΔEC6, α4-αC2$_{EC6}$, α4-β17$_{EC5-6}$, α4-γA11$_{EC5-6}$, α4-γB6$_{EC5-6}$, and γC3$_{EC5-6}$; 4 for α4-γA3$_{EC6}$ and α4-γB2$_{EC6}$; 5 for α4-β17$_{EC6}$; 6 for α4-γA11$_{EC6}$ and α4-γB6$_{EC6}$; and 7 for α4-γC3$_{EC6}$), and bars are replicate means. P-values for each sample compared to α4 were calculated using a one-tailed Welch's t-test. (C) Graph of $E_{surface}$ comparing constructs paired by isoform. Bars are color-coded by isoform. Bars without borders are α4 variants (EC6 or EC5-6 swaps). Bars with black borders are wildtype cPCDHs. Points are biological replicates (n = 15 for α4 and γC3; 3 for α4-αC2$_{EC6}$, α4-β17$_{EC5-6}$, α4-γA11$_{EC5-6}$, γB6, α4-γB6$_{EC5-6}$, and γC3$_{EC5-6}$; 4 for αC2, γA3, α4-γA3$_{EC6}$, γB2, and α4-γB2$_{EC6}$; 5 for α4-β17$_{EC6}$; 6 for β17, α4-γA11$_{EC6}$, γB6, and α4-γB6$_{EC6}$; 7 for α4-γC3$_{EC6}$; and 10 for γA11), and bars are replicate means. P-values for each α4 variant compared to the corresponding wildtype isoform were calculated using a two-tailed Welch's t-test. (D) Scatter plot of wildtype cPCDH surface delivery versus the corresponding EC6-swapped α4 variant, as measured by $E_{surface}$. Each point represents an experiment where the wildtype and EC6-swapped variant were tested side by side. Points are shape- and color-coded by isoform. The Pearson r statistic and its associated p-value indicate the degree of correlation.

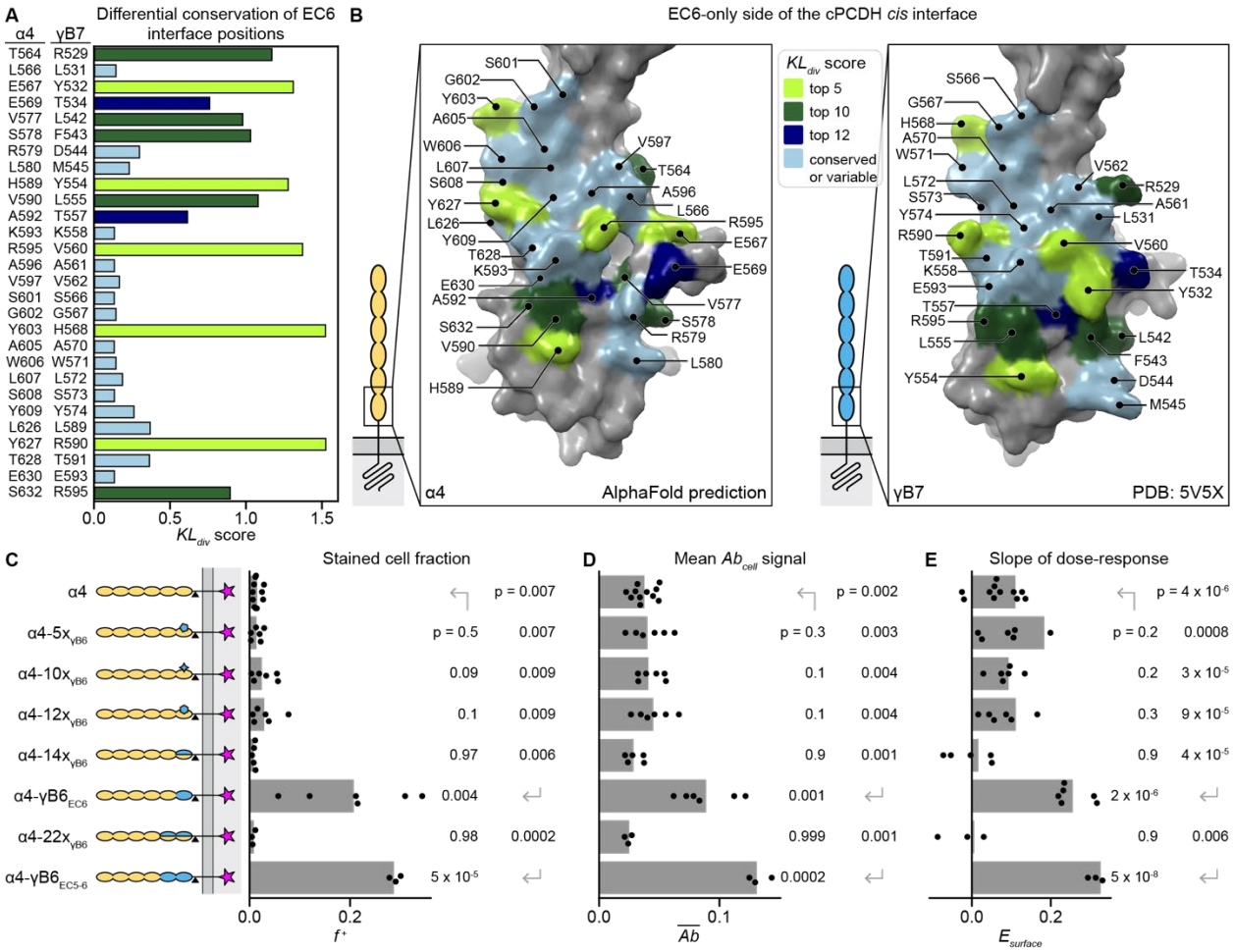

Figure 5. The EC6 *cis* interface of γB6 is insufficient for surface delivery of α4. (A) Graph of KL_{div} scores for all EC6 *cis* interface positions. The color of the bar indicates how positions were grouped into categories by KL_{div} score. (B) Surface representations of the EC6 *cis* interface for α4 (left; structure prediction from AlphaFold protein structure database (53)) and γB7 (right; experimentally-determined structure [PDB: 5V5X] (28)). *Cis* interface residues are colored according to their Kullback-Leibler divergence (KL_{div}) score categories (legend top left). (C) Schematics of the chimeric α4 constructs quantified in panels C-E. (C-E) Quantification of f^{+} (E), $\overline{Ab}$ (D), and $E_{surface}$ (E) for the constructs shown in panel C. In D, $\overline{Ab}$ values are normalized to the mean $\overline{Ab}$ of the positive control construct α4-γB6$_{EC6}$. Points are biological replicates (n = 11 for α4; 3 for α4-22x$_{\gamma B6}$ and α4-γB6$_{EC5-6}$; 5 for α4-14x$_{\gamma B6}$; and 6 for α4-5x$_{\gamma B6}$, α4-10x$_{\gamma B6}$, α4-12x$_{\gamma B6}$, and α4-γB6$_{EC6}$), and bars are replicate means. P-values for each sample compared to α4 were calculated using a one-tailed Welch's t-test. P-values for each sample compared to α4-γB6$_{EC6}$ or α4-γB6$_{EC5-6}$ were calculated using a two-tailed Welch's t-test.

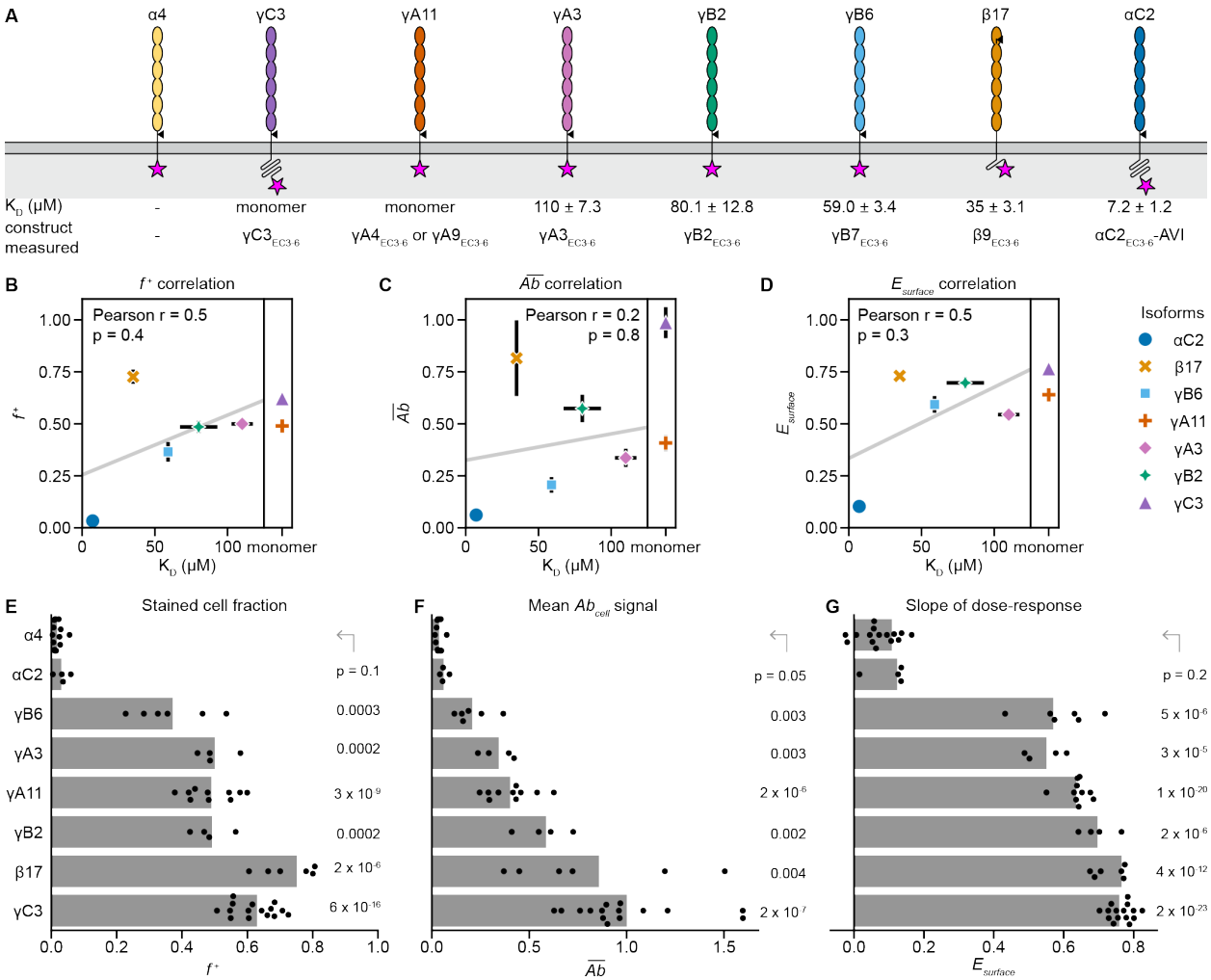

Figure 6. Surface delivery of C-type, β-, and γ-PCDHs does not depend on *cis* dimerization affinity. (A) Schematics of wildtype cPCDHs and reported solution affinity measurements of purified *cis*-interacting fragments (or their closest paralog) (22, 26, 28, 29). Binding affinity is reported as K_D, which is high for weak interactions and low for strong interactions. (B-D) Scatter plots of reported *cis* homodimerization affinity versus the surface trafficking metrics f^+ (B), $\overline{Ab}$ (C), and $E_{surface}$ (D). For each isoform, affinity is plotted as reported K_D ± reported error and surface trafficking is reported as the mean ± standard error of the mean of biological replicates (n = 15 for α4 and γC3; 4 for αC2, γA3, and γB2; 6 for β17 and γB6; and 10 for γA11). Points are shape- and color-coded by isoform. The Pearson r statistic and its associated p-value indicate the degree of correlation. (E-G) Quantification of f^+ (E), $\overline{Ab}$ (F), and $E_{surface}$ (G) for the constructs shown in panel A. In F, $\overline{Ab}$ values are normalized to the mean $\overline{Ab}$ of the positive control construct γC3. Points are biological replicates, and bars are replicate means. P-values for each sample compared to α4 were calculated using a one-tailed Welch's t-test.