



Antiparallel protocadherin homodimers use distinct affinity- and specificity-mediating regions in cadherin repeats 1-4

The Harvard community has made this article openly available. [Please share](#) how this access benefits you. Your story matters

Citation	Nicoludis, John M, Bennett E Vogt, Anna G Green, Charlotta PI Schärfe, Debora S Marks, and Rachelle Gaudet. 2016. "Antiparallel protocadherin homodimers use distinct affinity- and specificity-mediating regions in cadherin repeats 1-4." eLife 5 (1): e18449. doi:10.7554/eLife.18449. http://dx.doi.org/10.7554/eLife.18449 .
Published Version	doi:10.7554/eLife.18449
Citable link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:29002614
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Antiparallel protocadherin homodimers use distinct affinity- and specificity-mediating regions in cadherin repeats 1-4

John M Nicoludis^{1,2}, Bennett E Vogt², Anna G Green³, Charlotta PI Schärfe^{3,4}, Debora S Marks³, Rachele Gaudet^{2*}

¹Department of Chemistry and Chemical Biology, Harvard University, Cambridge, United States; ²Department of Molecular and Cellular Biology, Harvard University, Cambridge, United States; ³Department of Systems Biology, Harvard Medical School, Boston, United States; ⁴Applied Bioinformatics, Department of Computer Science, University of Tübingen, Tübingen, Germany

Abstract Protocadherins (Pcdhs) are cell adhesion and signaling proteins used by neurons to develop and maintain neuronal networks, relying on *trans* homophilic interactions between their extracellular cadherin (EC) repeat domains. We present the structure of the antiparallel EC1-4 homodimer of human Pcdh γ B3, a member of the γ subfamily of clustered Pcdhs. Structure and sequence comparisons of α , β , and γ clustered Pcdh isoforms illustrate that subfamilies encode specificity in distinct ways through diversification of loop region structure and composition in EC2 and EC3, which contains isoform-specific conservation of primarily polar residues. In contrast, the EC1/EC4 interface comprises hydrophobic interactions that provide non-selective dimerization affinity. Using sequence coevolution analysis, we found evidence for a similar antiparallel EC1-4 interaction in non-clustered Pcdh families. We thus deduce that the EC1-4 antiparallel homodimer is a general interaction strategy that evolved before the divergence of these distinct protocadherin families.

DOI: [10.7554/eLife.18449.001](https://doi.org/10.7554/eLife.18449.001)

*For correspondence: gaudet@mcb.harvard.edu

Competing interests: The authors declare that no competing interests exist.

Funding: See page 11

Received: 06 June 2016

Accepted: 28 July 2016

Published: 29 July 2016

Reviewing editor: William I Weis, Stanford University, United States

© Copyright Nicoludis et al. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

Introduction

Protocadherins (Pcdhs) encompass about 70% of the cadherin superfamily in mammals, and are involved in cell adhesion in the nervous system of higher animals (*Hulpiau and van Roy, 2011; Hulpiau et al., 2013; Keeler et al., 2015; Sotomayor et al., 2014*). Pcdhs segregate into two groups: clustered (expressed from a large gene cluster) and non-clustered. Clustered Pcdhs, comprising the α , β , and γ subfamilies, mediate neuronal survival, self-avoidance and self/nonself discrimination, and promote dendritic arborization in neuronal cells (*Emond and Jontes, 2008; Garrett et al., 2012; Kostadinov and Sanes, 2015; Ledderose et al., 2013; Lefebvre et al., 2012; Molumby et al., 2016; Suo et al., 2012; Wang et al., 2002; Weiner et al., 2005*). Clustered Pcdhs thus function analogously to insect Dscam isoforms to mediate neuronal identity (*Zipursky and Sanes, 2010*). Clustered Pcdhs are also broadly involved in synapse formation and maintenance, neuronal connectivity and neuronal survival (*Hayashi and Takeichi, 2015; Keeler et al., 2015*). Non-clustered Pcdhs play key roles in neuronal development (*Keeler et al., 2015; Kim et al., 2011*). For example, Pcdh7 is involved in germ layer differentiation (*Rashid et al., 2006; Yoshida, 2003*), and Pcdh1 and Pcdh8 mediate cell sorting and migration during gastrulation (*Kim et al., 1998; Kuroda et al., 2002*). Both clustered and non-clustered Pcdhs control these phenotypes through homophilic interactions of their extracellular cadherin (EC) repeat domains (*Hirano et al., 1999*;

eLife digest As the brain develops, nerve cells or neurons connect with one another to form complex networks. These connections form between branch-like structures, called dendrites, that project from the cell body of each neuron. To prevent unneeded connections from forming, dendrites that belong to the same neuron need a way to recognize and avoid one another. A family of proteins called protocadherins supports this process of self-avoidance.

Protocadherins have three main parts or domains: an extracellular domain that faces outwards away from the cell, a transmembrane domain that sits within the cell's surface membrane and an intracellular domain that faces into the cell's interior. There are two major groups of protocadherins – clustered and non-clustered – and the former are responsible for the self-avoidance behavior between dendrites. Clustered protocadherins in turn comprise three subfamilies, each of which consists of multiple variants with slightly different structures (known as isoforms). The particular set of protocadherin isoforms that a neuron displays on its surface distinguishes that neuron from all others, a little like a barcode.

When two dendrites meet, the protocadherins in their membranes come into contact with one another. If both dendrites come from the same neuron and therefore possess identical sets of protocadherins, then all protocadherins can form two-subunit complexes containing one copy of the same isoform from each dendrite. These complexes are called homodimers and their formation acts as a signal that informs the cell that it has encountered one of its own dendrites and should therefore not establish a connection. By using X-rays to determine the structure of a crystallized protocadherin fragment down to the level of its individual atoms, Nicoludis et al. now reveal exactly how clustered protocadherins form homodimers. The results show that each protocadherin subfamily uses a slightly different type of interaction due to differences in the structure of their extracellular domains.

The next challenge is to identify the signaling cascade that is triggered by the formation of clustered protocadherin homodimers, and to work out how activation of this cascade prevents a permanent connection from forming. In addition, the results of Nicoludis et al. predict that some non-clustered protocadherins form dimers with a similar architecture to that of clustered protocadherins. This possibility should also be tested experimentally.

DOI: [10.7554/eLife.18449.002](https://doi.org/10.7554/eLife.18449.002)

Hoshina et al., 2013; Kim et al., 1998; Kuroda et al., 2002; Schreiner and Weiner, 2010; Thu et al., 2014; Yamagata et al., 1999; Yoshida, 2003).

Clustered Pcdh subfamilies show distinct phenotypes. In zebrafish, α and γ Pcdhs express in overlapping but distinct brain regions (*Biswas et al., 2012*). In mammals, α Pcdhs regulate sorting of olfactory sensory neuron axons into glomeruli, serotonergic axon maturation, and dendritic patterning in CA1 pyramidal neurons (*Hasegawa et al., 2012, 2008; Katori et al., 2009; Suo et al., 2012*). The γ subfamily is important for self/non-self discrimination in retinal starburst amacrine cells and Purkinje neurons (*Kostadinov and Sanes, 2015; Lefebvre et al., 2012*). Thus, available data suggest that the different Pcdh subfamilies may function independently or cooperatively, perhaps depending on the brain region and/or neuronal cell type.

Our recent Pcdh γ A1 and Pcdh γ C3 EC1-3 structures revealed dimer interactions between EC2 and EC3 (*Nicoludis et al., 2015*), consistent with previous biochemical and bioinformatics data (*Schreiner and Weiner, 2010; Wu, 2005*). Using sequence co-evolution analysis, we predicted inter-subunit EC1-EC4 interactions, and proposed that clustered Pcdhs form extended antiparallel homodimers engaging EC1-4. A complementary biochemical and structural study arrived at a very similar docking model (*Rubinstein et al., 2015*), which was recently confirmed for α and β clustered Pcdhs (*Goodman et al., 2016*).

We determined the crystal structure of Pcdh γ B3 EC1-4, the first full antiparallel dimer for a γ isoform. We analyzed the clustered Pcdhs structures in light of biological, biochemical and evolutionary data to further resolve how clustered Pcdhs encode specificity. We describe how structural differences between the α , β and γ subfamilies generate distinct modes of specificity encoding. We also provide evidence that the EC1/EC4 and EC2/EC3 interfaces are functionally different: EC1/EC4

provides nonselective dimerization affinity while EC2/EC3 is generally responsible for enforcing specificity. Finally, we extend our sequence coevolution analysis to the non-clustered Pcdhs and provide evidence that the EC1-4 interaction is broadly used by Pcdhs.

Results and discussion

Structure of the Pcdh γ B3 EC1-4 extended antiparallel dimer

The in vitro-refolded recombinantly-expressed Pcdh γ B3 EC1-4 (47 kDa) yielded two peaks on size exclusion chromatography (SEC; **Figure 1—figure supplement 1**). Based on multi-angle light scattering (MALS; **Figure 1—figure supplement 1**), peak 1 was wide and polydisperse (~200–300 kDa). Peak 2 was monodisperse at 80 kDa – consistent with a dimer – and readily yielded a crystal structure (**Figure 1—source data 1**). As expected, each EC forms a seven-stranded Greek key β -sandwich motif (**Figure 1A**), similarly to other clustered Pcdh structures (**Goodman et al., 2016; Nicoludis et al., 2015; Rubinstein et al., 2015**). Notably, EC4 has a unique β -strand arrangement compared to EC1-EC3 (**Figure 1B,C**) and all known cadherin repeat structures. Strand β 1a is extended by 4–5 residues, while β 1b is correspondingly shortened. Additionally, while in EC1-3 strand β 2 splits into β 2a and β 2b, interacting with strands β 5 and β 1a, respectively, in EC4 it forms a continuous strand interacting with both simultaneously. This distinct structural feature contributes to intersubunit EC1/EC4 interactions (see below).

Although the asymmetric unit contains a single Pcdh γ B3 molecule, a crystallographic two-fold axis generates an antiparallel dimer with intersubunit EC1/EC4 and EC2/EC3 interactions (**Figure 1A**). This dimer is consistent with the Pcdh γ A1 EC1-3 crystal structure, validating the previously predicted interface (**Nicoludis et al., 2015; Rubinstein et al., 2015**), and with recent α and β Pcdhs structures (**Goodman et al., 2016**), confirming that this interaction mechanism is conserved among all clustered Pcdh subfamilies (**Figure 1D**). The structures do differ noticeably in overall twist, including subfamily-specific differences in relative EC1/EC4 orientation (**Figure 1E**).

The linear architecture of clustered Pcdhs enables extended antiparallel dimer interfaces. Overall, the tilt and azimuthal angles between adjacent clustered Pcdh repeats are distinct from those of classical cadherins (**Figure 1—figure supplement 2**) (**Nicoludis et al., 2015**). Classical cadherins, which typically dimerize through EC1/EC1 interfaces, exhibit smaller tilt angles and thus an overall curved structure (**Boggon et al., 2002**). Notably, the clustered Pcdh repeat orientation is such that EC1 and EC3 use the same face for intersubunit contacts, as do EC2 and EC4 (**Figure 1—figure supplement 3**), suggesting that longer cadherins could readily form even more extended interfaces.

Clustered protocadherin subfamilies have distinct specificity mechanisms dictated by structural differences

Clustered Pcdh subfamilies control different phenotypes in vivo and have discrete expression patterns (**Biswas et al., 2012; Keeler et al., 2015**), suggesting that they encode specificity using distinct modes, which may relate to subfamily-specific structural features. To investigate this hypothesis, we calculated the isoform conservation ratio (ICR) within individual subfamilies, which quantifies the extent to which individual residue positions are conserved among orthologs (same isoform in different species) and diversified in paralogs (different isoforms in the same species) (**Nicoludis et al., 2015**), resulting in three ICR value sets for the α , β and γ subfamilies, respectively (**Figure 2—figure supplement 1**). To account for subfamily differences in sequence conservation, we normalized the ICR values by dividing by the subfamily average. We then mapped them onto the Pcdh α 7 EC1-5, Pcdh β 8 EC1-4 and Pcdh γ B3 EC1-4 structures (**Figure 2A**). Comparing the structures and isoform-specific conservation in the different subfamilies allowed us to identify key specificity determinant regions for individual subfamilies. We illustrate three examples of how the subfamilies have encoded specificity using unique structural features.

In α isoforms, the EC2 β 4- β 5 loop is enriched in high-ICR and chemically diverse residues, and differs in conformation in the Pcdh α 4 and Pcdh α 7 structures (**Figure 2B**): the Pcdh α 4 EC2 β 4- β 5 loop contacts β 1b of EC3, while the corresponding loop in Pcdh α 7 does not, suggesting variable interactions in other isoforms. In comparison, the EC2 β 4- β 5 loop residues in both β and γ isoforms have lower ICR values, more similar loop structure, and do not contact β 1b of EC3. Thus, this loop may have evolved to generate diversity within α isoforms, but not in other subfamilies.

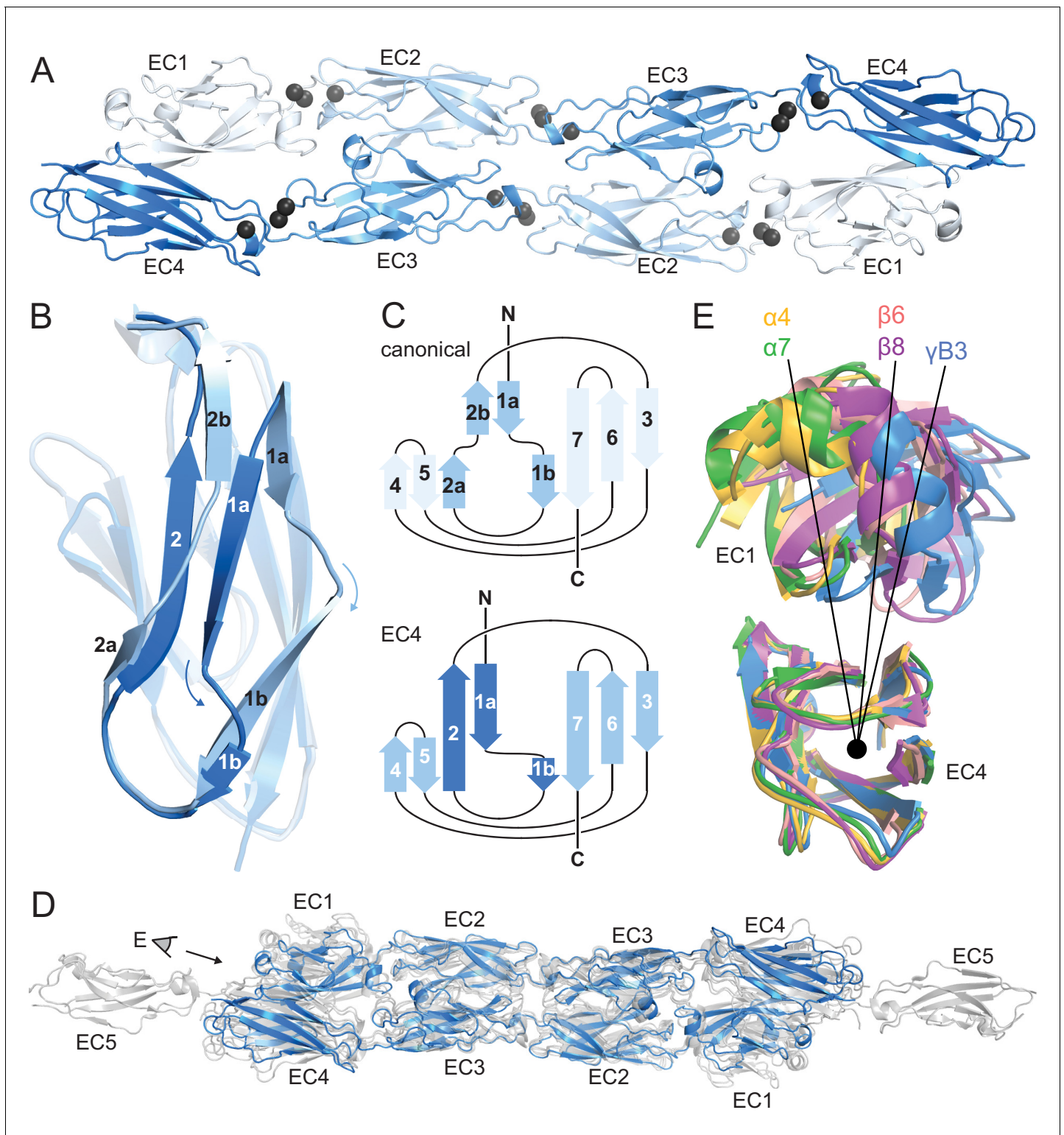


Figure 1. PcdhyB3 EC1-4 extended antiparallel dimer relies on unusual EC4 β -strand arrangement and is similar to other clustered Pcdh dimers. (A) Structure of the PcdhyB3 EC1-4 antiparallel dimer, with each EC a different shade of blue and the Ca²⁺ ions in grey. (B) Superposition of PcdhyB3 EC2 and EC4 highlighting the differences in β -strands 1 and 2. (C) Comparison of the canonical cadherin (top) and EC4 (bottom) β -strand arrangement. (D) The structures of Pcdh dimers $\alpha 4$ EC1-4, $\alpha 7$ EC1-5, $\beta 6$ EC1-4, and $\beta 8$ EC1-4 (grey) were superimposed using the dimeric EC2-3 region onto $\gamma B3$ EC1-4 (blue), illustrating variations in twist/corkscrew. (E) The EC4 domains of clustered Pcdh structures (colored as labeled) were superimposed, highlighting EC1 position differences that correlate with subfamily. Point of view (eye symbol) shown in (D). See **Figure 1—figure supplements 1–4**.

Figure 1 continued on next page

Figure 1 continued

DOI: [10.7554/eLife.18449.003](https://doi.org/10.7554/eLife.18449.003)

The following source data and figure supplements are available for figure 1:

Source data 1. Statistics for Pcdh γ B3 EC1-4 structure.

DOI: [10.7554/eLife.18449.004](https://doi.org/10.7554/eLife.18449.004)

Figure supplement 1. Pcdh γ B3 refolding yields two species, one of which corresponds to monodisperse dimeric protein.

DOI: [10.7554/eLife.18449.005](https://doi.org/10.7554/eLife.18449.005)

Figure supplement 2. Protocadherins and non-classical cadherins have a distribution of orientation between repeat pairs that is distinct from classical cadherins.

DOI: [10.7554/eLife.18449.006](https://doi.org/10.7554/eLife.18449.006)

Figure supplement 3. EC1 and EC3 use the same face for intersubunit contacts, as do EC2 and EC4.

DOI: [10.7554/eLife.18449.007](https://doi.org/10.7554/eLife.18449.007)

Figure supplement 4. HEPES molecule near the EC2/EC3 interface.

DOI: [10.7554/eLife.18449.008](https://doi.org/10.7554/eLife.18449.008)

In β isoforms, the Phe-X₁₀-Phe loop between β 3 and β 4 of EC3 has limited diversity compared to α and γ isoforms and wedges between the EC2 β 4- β 5 loop and β 2b strand (**Figure 2C**). In contrast, the Phe-X₁₀-Phe loop of α and γ isoforms has a helical conformation, and has residues with higher ICR values and greater chemical diversity. Therefore alterations in secondary structure can affect how specificity is encoded within the subfamilies.

The short loop following the extended β 1a strand in EC4 contacts the EC1 β 6- β 7 loop (**Figure 2D**), and there are large structural differences in the EC1/EC4 interaction between subfamilies (**Figure 1E**). α Isoforms have low-ICR residues at this interface, whereas β and γ isoforms have higher ICR value residues. This thus suggests that the large structural differences drive inter-subfamily specificity, on which may be layered additional isoform specificity.

In all cases, sequence regions with high isoform-specific conservation correlate with interface contacts, revealing the interplay between dimer structure and how subfamilies encode specificity. Diversity in the composition and conformation of loop regions provides distinct specificity mechanisms to subfamilies. Phylogenetic analysis indicates that isoforms are more similar within than across subfamilies (*Sotomayor et al., 2014; Wu, 2005*), and the available structures show that the interface architecture is more similar within subfamilies as well (**Figure 1E**) (*Goodman et al., 2016; Nicoludis et al., 2015*). With this insight, the dimer interface seen in the Pcdh γ C3 EC1-3 crystal structure may represent a unique dimer architecture for C-type isoforms (*Nicoludis et al., 2015*), as these isoforms are transcriptionally, functionally and evolutionarily distinct from the subfamilies in which they reside (*Chen et al., 2012; Frank et al., 2005; Kaneko et al., 2006*). Distinct expression of the clustered Pcdh subfamilies in different tissues and at different developmental stages supports the necessity for intra-subfamily specificity (*Biswas et al., 2012; Frank et al., 2005*). Differences in subfamily structure and isoform-specific conservation suggest that homophilic specificity mechanisms emerged independently in each subfamily through diversification of subfamily-specific interface contacts.

The EC1/EC4 interaction provides affinity of dimerization

The EC2/EC3 interaction is integral to clustered Pcdh dimerization specificity, as evidenced by bioinformatics and cell-aggregation assays (*Nicoludis et al., 2015; Rubinstein et al., 2015; Schreiner and Weiner, 2010; Thu et al., 2014; Wu, 2005*). We sought to understand the functional purpose of the EC1/EC4 interaction, and made three observations. First, for all isoforms with available structures, fewer EC1/EC4 interface residues have high isoform-specific conservation compared to the EC2/EC3 interface residues (**Figure 2A, Figure 2—figure supplement 1**). Second, interface residues shared by most isoforms are more hydrophobic in EC1/EC4 than in EC2/EC3 (**Figure 3A**). Third, the Pcdh γ B3 EC1/EC4 interface is much larger (BSA = 976 Å² per protomer) than the EC2/EC3 interface (555 Å² per protomer). The lack of isoform specificity, the hydrophobic nature, and large interface area together suggest that the EC1/EC4 interface promotes binding with little specificity.

To probe this hypothesis, we used the Computational Interface Alanine Scanning Server to assess each interface residue's contribution to the free energy of complex formation ($\Delta\Delta G_{\text{calc}}$) when

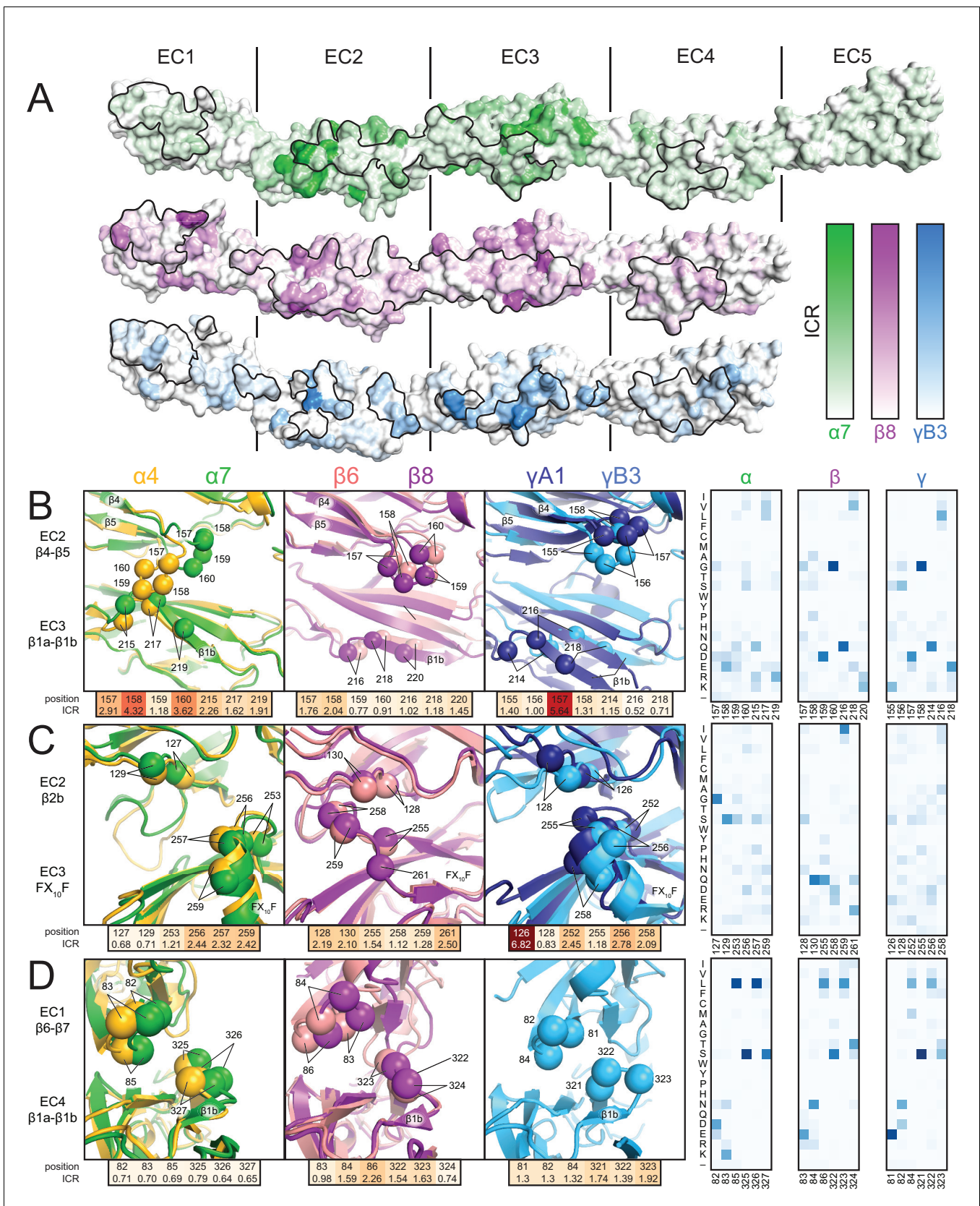


Figure 2. Isoform-specific conservation and structural differences reveal subfamily differences in diversity generation. (A) Subfamily-specific ICR values mapped onto the surfaces of Pcdh $\alpha 7$ (top, green), Pcdh $\beta 8$ (middle, magenta) and Pcdh $\gamma B3$ (bottom, blue). The black outline marks the dimer interface. Figure 2 continued on next page

Figure 2 continued

footprint. (B, C, D) Unique structural features of the α (left), β (center), and γ (right) structures (colored according to Figure 1). ICR values for highlighted residues shown below and normalized amino acid frequencies for these positions shown on the right. See Figure 2—figure supplement 1.

DOI: 10.7554/eLife.18449.009

The following figure supplement is available for figure 2:

Figure supplement 1. Clustered Pcdh subfamilies have distinct patterns of isoform-specific conservation.

DOI: 10.7554/eLife.18449.010

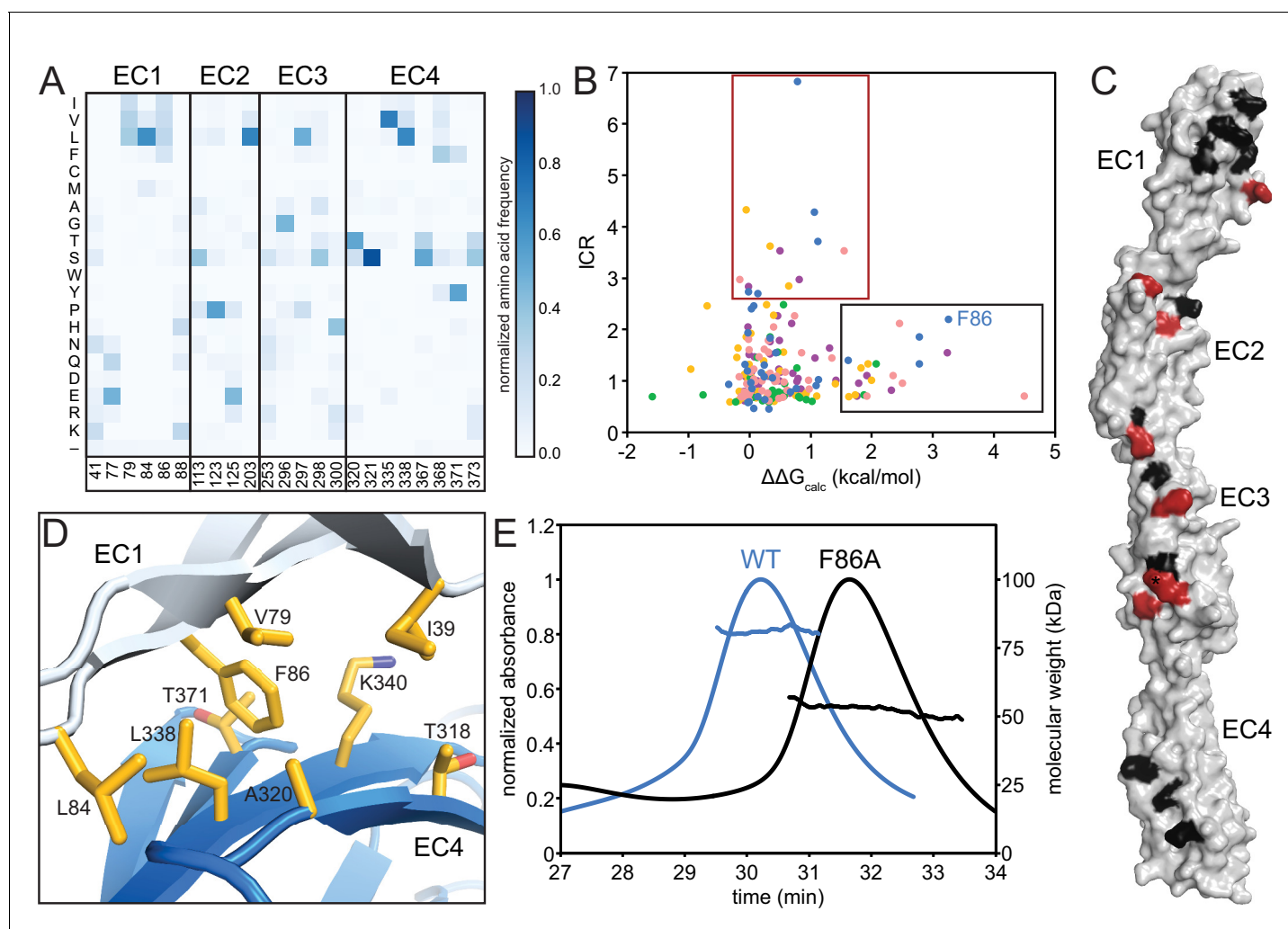


Figure 3. The EC1/EC4 interface is enriched in affinity-driving hydrophobic residues, while the EC2/EC3 interface contains high-ICR residues driving specificity. (A) Amino acid frequencies in clustered Pcdhs of conserved interface residues (see Figure 1—figure supplement 2). (B) Plot of ICR value and $\Delta\Delta G_{\text{calc}}$ of interface residues of Pcdh α EC1-4 (yellow), α 7 EC1-5 (green), β 6 EC1-4 (salmon), β 8 EC1-4 (magenta), γ B3 EC1-4 (blue). Two subsets of interface residues segregate from the main cluster: high- $\Delta\Delta G_{\text{calc}}$ and low-ICR residues ('affinity'; black box) and low- $\Delta\Delta G_{\text{calc}}$ and high-ICR residues ('specificity'; crimson box). Residue F86 from PcdhyB3 EC1-4 is labeled. (C) High- $\Delta\Delta G_{\text{calc}}$ and low-ICR residues (black) map primarily to EC1 and EC4, while low- $\Delta\Delta G_{\text{calc}}$ and high-ICR residues (crimson) primarily map to EC2 and EC3. N253 (*) is found in the 'specificity' region for γ B3 and in the 'affinity' region for β 6 and β 8. (D) The EC1/EC4 interface features a hydrophobic cluster, with EC1 F86 near its center. (E) SEC-MALS profiles of WT PcdhyB3 EC1-4 (blue; molecular weight 82 kDa) and F86A (black; molecular weight 52 kDa) run on a Superdex S200 10/300 column, are consistent with dimeric and monomeric proteins, respectively.

DOI: 10.7554/eLife.18449.011

computationally mutated to alanine (Kortemme and Baker, 2002; Kortemme et al., 2004). Using the five available EC1-4 interfaces, two residue groups emerged from comparing the ICR values to $\Delta\Delta G_{\text{calc}}$: one group with low ICR values and high $\Delta\Delta G_{\text{calc}}$, the other with high ICR values and low $\Delta\Delta G_{\text{calc}}$ (Figure 3B). These residue groups can be regarded as contributing to the affinity and specificity of the complex, respectively. When mapped on Pcdh γ B3, predicted affinity residues concentrated on EC1 and EC4, and predicted specificity residues on EC2 and EC3, corroborating the distinction between EC1/EC4 and EC2/EC3 interactions (Figure 3C).

In the Pcdh γ B3 EC1/EC4 interface, F86, one of the predicted affinity-driving residues from EC1, wedges into a cavity created by hydrophobic EC4 residues (Figure 3D). A Pcdh γ B3 EC1-4 F86A mutant indeed disrupted dimerization, resulting in a monomeric protein as measured by MALS (Figure 3E). Thus, the hydrophobic interactions between EC1 and EC4 are crucial to dimerization. Analogously, purified EC1-3 constructs failed to dimerize in vitro whereas EC1-4 constructs did (Nicoludis et al., 2015; Rubinstein et al., 2015), and K562 cells expressing Δ EC1 or Δ EC4-6 constructs did not aggregate while cells expressing chimeras in which EC1 and EC4 derived from different paralogs did (Schreiner and Weiner, 2010; Thu et al., 2014). Together, these results indicate that the EC1/EC4 interaction is not strictly required for the specificity of dimerization but it drives dimerization affinity through non-specific hydrophobic interactions.

Antiparallel EC1-4 interaction is predicted in non-clustered Pcdhs

The antiparallel EC1-4 interaction architecture can encode diverse specificities within the clustered Pcdh family. Is this architecture unique to clustered Pcdhs or is it ancestral, and thus also found in non-clustered Pcdhs? These include the δ -1 (Pcdh1, Pcdh7, Pcdh9, Pcdh11) and δ -2 (Pcdh8, Pcdh10, Pcdh17, Pcdh18, Pcdh19) families that are integral to the development and maintenance of the nervous system (Keeler et al., 2015; Kim et al., 2011). We used sequence coevolution analysis, which successfully predicted the clustered Pcdh interface (Nicoludis et al., 2015) (Figure 4—figure supplement 1), to look for evidence of an antiparallel interface in non-clustered Pcdhs (Figure 4A). As in clustered Pcdhs, most covarying residue pairs in non-clustered Pcdhs were intra-domain structural contacts of the well-conserved cadherin fold. Additionally, several covarying pairs are found between EC2 and EC3, or EC1 and EC4, similar to those observed for the clustered Pcdhs. When mapped onto the Pcdh γ B3 dimer, these covarying pairs are somewhat further apart than true interface contacts (Figure 4B), which could be due to differences in dimerization interfaces, as we observe between the clustered Pcdh families, or in the δ -1 or δ -2 Pcdhs secondary structure, for which there are no available structures. This analysis thus predicts an antiparallel EC1-4 interaction in members of the non-clustered Pcdhs. Notably, we cannot determine whether all members or only a subset – and if so, which – likely use this architecture. However maximum parsimony suggests that the ancestral Pcdh used the antiparallel EC1-4 dimer interaction, and Pcdh members which do not show this interaction mechanism either diverged before it evolved or lost it subsequently.

Finally, we looked at the composition of a predicted non-clustered Pcdh interface, by selecting residues homologous to those found at clustered Pcdh interfaces. The predicted EC1/EC4 interface residues are predominantly hydrophobic, while EC2/EC3 residues have more polar and ionic character (Figure 4C). Notably, positions 41 and 77 in EC1 and 320, 321, 371 and 373 in EC4 are more hydrophobic in non-clustered than clustered Pcdhs, indicating that these may form contacts in some non-clustered Pcdhs. Thus, like in the clustered Pcdhs, the EC1/EC4 interaction may promote dimer affinity while the EC2/EC3 interaction provides specificity.

Conclusions

Recently, we and others predicted that clustered Pcdhs form homophilic antiparallel EC1-EC4 complexes based on crystal structures, mutagenesis and bioinformatics (Nicoludis et al., 2015; Rubinstein et al., 2015). Our structure of Pcdh γ B3 EC1-4 confirms our sequence coevolution analysis, demonstrating the robustness of the analysis and revealing the molecular details of this interaction. Here we extended this prediction to other non-clustered Pcdhs using sequence coevolution analysis.

Analysis of the Pcdh γ B3 EC1-4 structure in comparison to α or β subfamily dimers revealed structural differences that correlated with differences in isoform-specific conservation, indicating distinct specificity mechanisms. Unlike the Dscams, where isoforms vary at specific alternatively-spliced

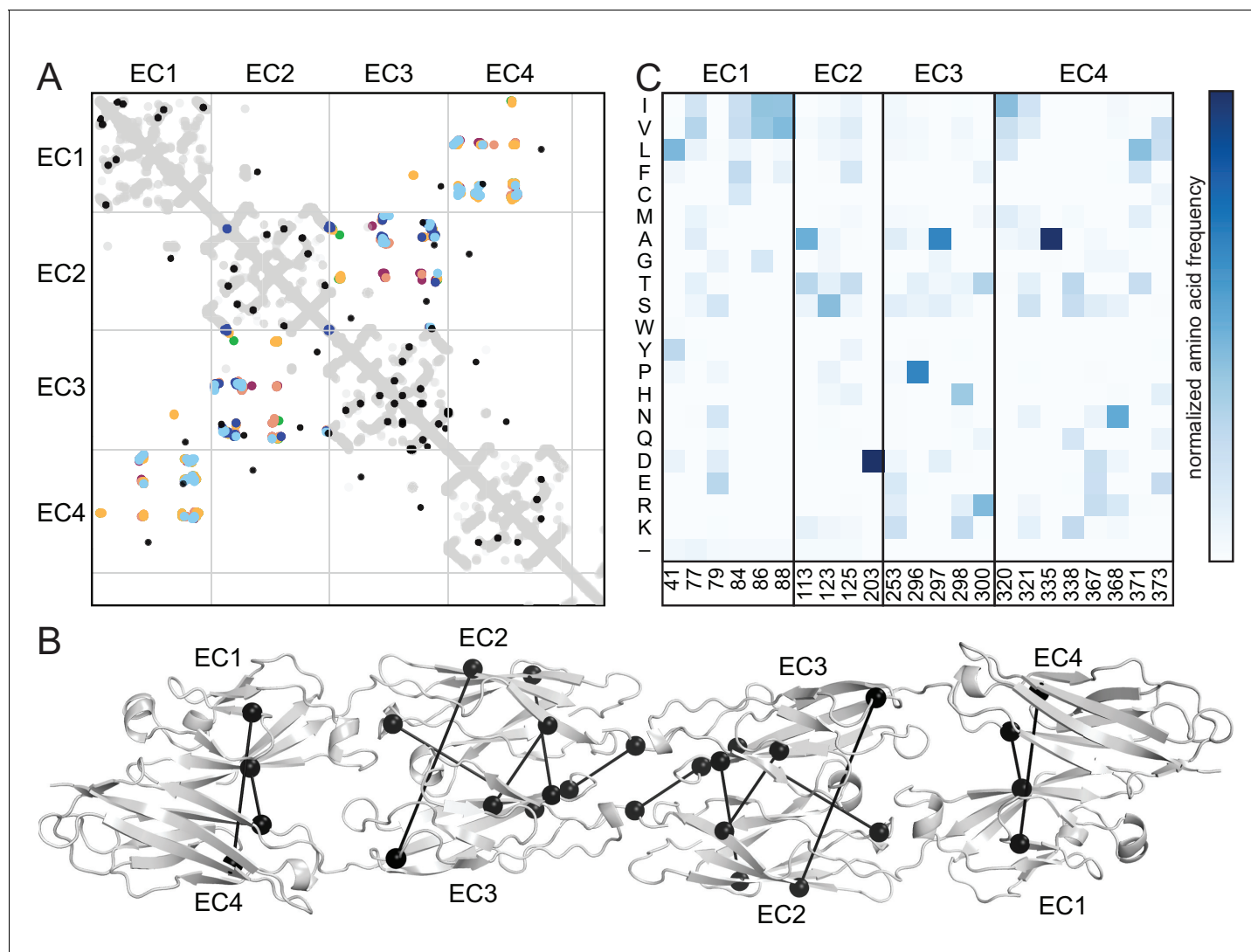


Figure 4. Evolutionary couplings in non-clustered Pcdhs predict an antiparallel interface engaging EC1-EC4. (A) The top 38 covarying pairs are shown in black, and include a number of EC1-EC4 and EC2-EC3 covarying residue pairs. The intramolecular contact maps of Pcdh γ B3 EC1-4, Pcdh α 4 EC1-4, Pcdh α 7 EC1-4, Pcdh β 6 EC1-4, Pcdh β 8 EC1-4 and Pcdh γ A1 EC1-3 are in gray for reference. The observed interface contact residues are also mapped (α 4, yellow; α 7, green; β 6, salmon; β 8, magenta; γ B3, blue; γ A1, dark blue). (B) Covarying residue pairs across EC1-EC4 or EC2-EC3 are mapped onto the Pcdh γ B3 EC1-4 structure with a line between coupled residues. Alignments and evolutionary couplings in **Figure 4—source data 1** and **2**. (C) Amino acid frequencies at non-clustered Pcdh alignment positions corresponding to the conserved interface residue positions identified in clustered Pcdhs (**Figure 1—figure supplement 2**). See **Figure 4—figure supplements 1** and **2**. DOI: [10.7554/eLife.18449.012](https://doi.org/10.7554/eLife.18449.012)

The following source data and figure supplements are available for figure 4:

Source data 1. Alignment of non-clustered Pcdhs EC1-4.

DOI: [10.7554/eLife.18449.013](https://doi.org/10.7554/eLife.18449.013)

Source data 2. Evolutionary couplings from the non-clustered Pcdh alignment.

DOI: [10.7554/eLife.18449.014](https://doi.org/10.7554/eLife.18449.014)

Source data 3. Alignment of clustered Pcdhs EC1-4.

DOI: [10.7554/eLife.18449.015](https://doi.org/10.7554/eLife.18449.015)

Source data 4. Evolutionary couplings from the clustered Pcdh alignment.

DOI: [10.7554/eLife.18449.016](https://doi.org/10.7554/eLife.18449.016)

Figure supplement 1. Evolutionary couplings in clustered Pcdhs are consistent with all available EC1-EC4 antiparallel homodimeric interfaces.

DOI: [10.7554/eLife.18449.017](https://doi.org/10.7554/eLife.18449.017)

Figure supplement 2. Phylogenetic tree distinguishes clustered from non-clustered Pcdhs.

DOI: [10.7554/eLife.18449.018](https://doi.org/10.7554/eLife.18449.018)

regions (Li et al., 2016; Meijers et al., 2007; Sawaya et al., 2008; Wojtowicz et al., 2007), the clustered Pcdh subfamilies are structurally diverse, and thus can encode specificity in different ways.

We identified a hydrophobic interaction between EC1 and EC4 that contributes to dimerization affinity, whereas its conservation among clustered Pcdh isoforms suggests that this interaction is not a driver of specificity. Overall, our data support a general role for conserved hydrophobic EC1/EC4 interactions in affinity, and for highly diversified polar EC2/EC3 contacts in specificity, and sequence analyses suggest that this is conserved in at least some non-clustered Pcdhs.

Materials and methods

Expression, purification and crystallization of Pcdh γ B3 EC1-4

Human Pcdh γ B3 EC1-4 (residues 1–414, not counting the signal peptide) was cloned into pET21 with a C-terminal hexahistidine tag, expressed in BL21 Gold (DE3) *Escherichia coli* cells in terrific broth. Cells were induced at OD₆₀₀ = 0.8 with 0.5 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) at 37°C for 4 hr, harvested and lysed by sonication in 8 M guanadinium hydrochloride (GuHCl), 50 mM HEPES pH 7.5, 2 mM CaCl₂, and 20 mM imidazole. Cell lysates were diluted to 5 M GuHCl and loaded onto Ni-Sepharose, washed with 50 mM HEPES pH 7.5, 250 mM NaCl, 10 mM CaCl₂, and 25 mM imidazole and eluted with 250 mM imidazole. Eluted protein was refolded at 1 mg/mL in 12 hr dialysis steps reducing the GuHCl concentration from 2.5 M to 1.25 M and finally 0 M in refolding buffer (100 mM Tris pH 8.5, 10 mM CaCl₂, 1 mM EDTA, 5 mM dithiothreitol (DTT), and 0.5 M L-arginine). Concentrated refolded protein was purified by size-exclusion chromatography (SEC) on a Superdex 200 16/60 column (GE Healthcare, Pittsburgh, PA) in 20 mM Tris pH 8.5, 200 mM NaCl, and 2 mM CaCl₂ (SEC buffer). Two peaks were isolated and each peak was run again separately by SEC before being concentrated for crystallization.

Multi-angle light scattering (MALS)

Approximate molecular mass of Pcdh γ B3 EC1-4 protein (WT or F86A mutant) was determined using a Superdex S200 10/300 column (GE Healthcare, Pittsburgh, PA) with in-line Wyatt Dawn Heleos II and Optilab T-rx refractive index detectors. Protein (100 μ L at 4 mg/mL) was injected and run at 0.4 mL/min in SEC buffer. Signals were aligned, normalized and band-broadened using bovine serum albumin as a standard.

Crystallization, data collection, and structure determination and analysis

Crystals were obtained by vapor diffusion at room temperature in 0.1 M HEPES pH 7, 4% ethylene glycol, and 5% polyethylene glycol monomethyl ether 500 in a 0.3 μ L protein (12 mg/mL) to 0.3 μ L reservoir drop, then cryoprotected with reservoir with 20% glycerol before flash cooling in liquid N₂. Diffraction data (Figure 1—source data 1) were processed in HKL2000 (Otwinowski and Minor, 1997). The Pcdh γ B3 EC1-4 structure was determined by an iterative molecular replacement search with single domains of the Pcdh γ A1 EC1-3 structure (PDBID 4zi9) in PHENIX (Adams et al., 2010). Model building was done in Coot (Emsley and Cowtan, 2004) and refinement in PHENIX (Adams et al., 2010). We analyzed the physicochemical properties of the dimer interface using PISA (Krissinel and Henrick, 2007). In the structure, we found a HEPES molecule near the EC2/EC3 interface that forms a salt bridge with N253 and N155 (Figure 1—figure supplement 4). MALS data were collected with Tris as the buffer (Figure 1—figure supplement 1), indicating that HEPES is not required for dimerization.

ICR value calculations

Overall percent identity of the most common residue at each position was used to calculate ICR values, dividing the percent identity across subfamily orthologs by the percent identity across subfamily paralogs. ICR values were then normalized by dividing by the whole sequence ICR average within each subfamily. The alignment and identity data are provided here (Nicoludis et al., 2015).

Computational Interface Alanine Scanning Server

Pcdh α 4 EC1-4 (5dzw), Pcdh α 7 EC1-5 (5dzv), Pcdh β 6 EC1-4 (5dzx), Pcdh β 8 EC1-4 (5dzy), and Pcdh γ B3 EC1-4 (5k8r) dimer structures were submitted to the Computational Interface Alanine Scanning Server using default settings (Kortemme and Baker, 2002; Kortemme et al., 2004).

Covariation analyses

Previously, we generated an alignment of clustered Pcdhs using mouse Pcdh γ C3 and manually filtered by phylogeny, using FastTree 2.1 (Price et al., 2010), to eliminate non-clustered Pcdhs (Figure 4—figure supplement 2) (Nicoludis et al., 2015). Both this clustered Pcdh and the now separated non-clustered Pcdh alignment were filtered to remove sequences with more than 50% gaps and columns with more than 30% gaps, and trimmed to contain only EC1-EC4. The clustered and non-clustered Pcdh alignments had 2660 and 405.5 effective non-redundant sequences, respectively; sequences were considered redundant and downweighted when more than 90% identical over their full length (Hopf et al., 2014). Evolutionary couplings (Hopf et al., 2014; Marks et al., 2011) were computed using the updated 'PLMC' algorithm (Weinreb et al., 2015) available on <https://github.com/debbiemarkslab/plmc>, which uses a pseudo maximum likelihood approximation (Balakrishnan et al., 2011; Ekeberg et al., 2013; Kamisetty et al., 2013). Alignments and all EC scores are provided (Figure 4—source data 1–4).

We used the precision of the intra-domain evolutionary couplings to determine whether the inter-domain evolutionary couplings are likely to be true. For the clustered Pcdh alignment, 83 non-local (more than five residues apart) contacts fall above a threshold of 80% precision of the intra-domain evolutionary couplings. Intra-domain evolutionary couplings are considered true if they correspond to structural contact (minimum atom distance < 8 Å) in any structure (Figure 4—figure supplement 1). Based on the same 80% precision threshold, the top 38 non-local ECs are significant in the non-clustered Pcdh alignment. We exclude couplings between residues greater than 400 in this analysis due to the false signal from gaps in this region.

Acknowledgements

We thank Kelly Arnett, director of the Center for Macromolecular Interactions, Harvard Medical School) for help in collecting SEC-MALS data. We thank the beamline staff of NE-CAT at the Advanced Photon Source (Argonne, IL, USA) for help with data collection. NE-CAT is funded by NIH (P41 GM103403 and S10 RR029205) and the Advanced Photon Source by the US Department of Energy (DE-AC02-06CH11357). Evolutionary couplings analysis was conducted on the Orchestra High Performance Compute Cluster at Harvard Medical School, which is funded by the NIH (NCRR 1S10RR028832-01). Financial support (to JMN) was provided by the National Defense Science and Engineering Graduate Fellowship. AGG was supported by the National Science Foundation Graduate Research Fellowship (DGE1144152). DSM was supported by National Institutes of Health (GM106303).

Additional information

Funding

Funder	Grant reference number	Author
National Science Foundation	DGE1144152	Anna G Green
National Institutes of Health	NCRR 1S10RR028832-01	Debora S Marks
National Institute of General Medical Sciences	GM106303	Debora S Marks
National Defense Science and Engineering Graduate Fellowship		John M Nicoludis
National Institute of General Medical Sciences	P41 GM103403	Rachelle Gaudet
National Institutes of Health	S10 RR029205	Rachelle Gaudet

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

JMN, AGG, Conception and design, Acquisition of data, Analysis and interpretation of data, Drafting or revising the article; BEV, Conception and design, Acquisition of data, Drafting or revising the article; CPIS, Drafting or revising the article, Contributed unpublished essential data or reagents; DSM, RG, Conception and design, Analysis and interpretation of data, Drafting or revising the article

Author ORCIDs

Rachelle Gaudet,  <http://orcid.org/0000-0002-9177-054X>

Additional files

Major datasets

The following datasets were generated:

Author(s)	Year	Dataset title	Dataset URL	Database, license, and accessibility information
Nicoludis JM, Vogt BE, Gaudet R	2016	Structure of human clustered protocadherin gamma B3 EC1-4	http://www.rcsb.org/pdb/explore.do?structureId=5K8R	Publicly available at the RCSB Protein Data Bank (accession no: 5K8R)
Nicoludis JM, Vogt BE, Gaudet R	2016	Structure of human clustered protocadherin gamma B3 EC1-4	https://data.sbggrid.org/dataset/325	Publicly available at Structural Biology Data Grid (accession no: 325)

References

- Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D Biological Crystallography* **66**:213–221. doi: [10.1107/S0907444909052925](https://doi.org/10.1107/S0907444909052925)
- Balakrishnan S, Kamisetty H, Carbonell JG, Lee SI, Langmead CJ. 2011. Learning generative models for protein fold families. *Proteins* **79**:1061–1078. doi: [10.1002/prot.22934](https://doi.org/10.1002/prot.22934)
- Biswas S, Emond MR, Jontes JD. 2012. The clustered protocadherins Pcdh α and Pcdh γ form a heteromeric complex in zebrafish. *Neuroscience* **219**:280–289. doi: [10.1016/j.neuroscience.2012.05.058](https://doi.org/10.1016/j.neuroscience.2012.05.058)
- Boggon TJ, Murray J, Chappuis-Flament S, Wong E, Gumbiner BM, Shapiro L. 2002. C-cadherin ectodomain structure and implications for cell adhesion mechanisms. *Science* **296**:1308–1313. doi: [10.1126/science.1071559](https://doi.org/10.1126/science.1071559)
- Chen WV, Alvarez FJ, Lefebvre JL, Friedman B, Nwakeze C, Geiman E, Smith C, Thu CA, Tapia JC, Tasic B, Sanes JR, Maniatis T. 2012. Functional significance of isoform diversification in the protocadherin gamma gene cluster. *Neuron* **75**:402–409. doi: [10.1016/j.neuron.2012.06.039](https://doi.org/10.1016/j.neuron.2012.06.039)
- Ekeberg M, Lökvist C, Lan Y, Weigt M, Aurell E. 2013. Improved contact prediction in proteins: Using pseudolikelihoods to infer Potts models. *Physical Review E* **87**:1–16. doi: [10.1103/PhysRevE.87.012707](https://doi.org/10.1103/PhysRevE.87.012707)
- Emond MR, Jontes JD. 2008. Inhibition of protocadherin-alpha function results in neuronal death in the developing zebrafish. *Developmental Biology* **321**:175–187. doi: [10.1016/j.ydbio.2008.06.011](https://doi.org/10.1016/j.ydbio.2008.06.011)
- Emsley P, Cowtan K. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D Biological Crystallography* **60**:2126–2132. doi: [10.1107/S0907444904019158](https://doi.org/10.1107/S0907444904019158)
- Frank M, Ebert M, Shan W, Phillips GR, Arndt K, Colman DR, Kemler R. 2005. Differential expression of individual gamma-protocadherins during mouse brain development. *Molecular and Cellular Neuroscience* **29**:603–616. doi: [10.1016/j.mcn.2005.05.001](https://doi.org/10.1016/j.mcn.2005.05.001)
- Garrett AM, Schreiner D, Lobas MA, Weiner JA. 2012. γ -protocadherins control cortical dendrite arborization by regulating the activity of a FAK/PKC/MARCKS signaling pathway. *Neuron* **74**:269–276. doi: [10.1016/j.neuron.2012.01.028](https://doi.org/10.1016/j.neuron.2012.01.028)
- Goodman KM, Rubinstein R, Thu CA, Bahna F, Mannepli S, Ahlsén G, Rittenhouse C, Maniatis T, Honig B, Shapiro L. 2016. Structural basis of diverse homophilic recognition by clustered α - and β -protocadherins. *Neuron* **90**:709–723. doi: [10.1016/j.neuron.2016.04.004](https://doi.org/10.1016/j.neuron.2016.04.004)

- Hasegawa S**, Hamada S, Kumode Y, Esumi S, Katori S, Fukuda E, Uchiyama Y, Hirabayashi T, Mombaerts P, Yagi T. 2008. The protocadherin-alpha family is involved in axonal coalescence of olfactory sensory neurons into glomeruli of the olfactory bulb in mouse. *Molecular and Cellular Neuroscience* **38**:66–79. doi: [10.1016/j.mcn.2008.01.016](https://doi.org/10.1016/j.mcn.2008.01.016)
- Hasegawa S**, Hirabayashi T, Kondo T, Inoue K, Esumi S, Okayama A, Hamada S, Yagi T. 2012. Constitutively expressed Protocadherin- α regulates the coalescence and elimination of homotypic olfactory axons through its cytoplasmic region. *Frontiers in Molecular Neuroscience* **5**:97. doi: [10.3389/fnmol.2012.00097](https://doi.org/10.3389/fnmol.2012.00097)
- Hayashi S**, Takeichi M. 2015. Emerging roles of protocadherins: from self-avoidance to enhancement of motility. *Journal of Cell Science* **128**:1–10. doi: [10.1242/jcs.166306](https://doi.org/10.1242/jcs.166306)
- Hirano S**, Yan Q, Suzuki ST. 1999. Expression of a novel protocadherin, OL-protocadherin, in a subset of functional systems of the developing mouse brain. *Journal of Neuroscience* **19**:995–1005.
- Hopf TA**, Schärfe CP, Rodrigues JP, Green AG, Kohlbacher O, Sander C, Bonvin AM, Marks DS. 2014. Sequence co-evolution gives 3D contacts and structures of protein complexes. *eLife* **3**:e03430. doi: [10.7554/eLife.03430](https://doi.org/10.7554/eLife.03430)
- Hoshina N**, Tanimura A, Yamasaki M, Inoue T, Fukabori R, Kuroda T, Yokoyama K, Tezuka T, Sagara H, Hirano S, Kiyonari H, Takada M, Kobayashi K, Watanabe M, Kano M, Nakazawa T, Yamamoto T. 2013. Protocadherin 17 regulates presynaptic assembly in topographic corticobasal Ganglia circuits. *Neuron* **78**:839–854. doi: [10.1016/j.neuron.2013.03.031](https://doi.org/10.1016/j.neuron.2013.03.031)
- Hulpiau P**, Gul SI, van Roy F. 2013. New insights into the evolution of metazoan cadherins and catenins. In: Roy Van F (Ed). *The Molecular Biology of Cadherins*. San Diego, CA: Elsevier. p 71–94. doi: [10.1016/b978-0-12-394311-8.00004-2](https://doi.org/10.1016/b978-0-12-394311-8.00004-2)
- Hulpiau P**, van Roy F. 2011. New insights into the evolution of metazoan cadherins. *Molecular Biology and Evolution* **28**:647–657. doi: [10.1093/molbev/msq233](https://doi.org/10.1093/molbev/msq233)
- Kamisetty H**, Ovchinnikov S, Baker D. 2013. Assessing the utility of coevolution-based residue-residue contact predictions in a sequence- and structure-rich era. *PNAS* **110**:15674–15679. doi: [10.1073/pnas.1314045110](https://doi.org/10.1073/pnas.1314045110)
- Kaneko R**, Kato H, Kawamura Y, Esumi S, Hirayama T, Hirabayashi T, Yagi T. 2006. Allelic gene regulation of Pcdh-alpha and Pcdh-gamma clusters involving both monoallelic and biallelic expression in single Purkinje cells. *Journal of Biological Chemistry* **281**:30551–30560. doi: [10.1074/jbc.M605677200](https://doi.org/10.1074/jbc.M605677200)
- Katori S**, Hamada S, Noguchi Y, Fukuda E, Yamamoto T, Yamamoto H, Hasegawa S, Yagi T. 2009. Protocadherin-alpha family is required for serotonergic projections to appropriately innervate target brain areas. *Journal of Neuroscience* **29**:9137–9147. doi: [10.1523/JNEUROSCI.5478-08.2009](https://doi.org/10.1523/JNEUROSCI.5478-08.2009)
- Keeler AB**, Molumby MJ, Weiner JA. 2015. Protocadherins branch out: Multiple roles in dendrite development. *Cell Adhesion & Migration* **9**:214–226. doi: [10.1080/19336918.2014.1000069](https://doi.org/10.1080/19336918.2014.1000069)
- Kim SY**, Yasuda S, Tanaka H, Yamagata K, Kim H. 2011. Non-clustered protocadherin. *Cell Adhesion & Migration* **5**:97–105. doi: [10.4161/cam.5.2.14374](https://doi.org/10.4161/cam.5.2.14374)
- Kim SH**, Yamamoto A, Bouwmeester T, Agius E, Robertis EM. 1998. The role of paraxial protocadherin in selective adhesion and cell movements of the mesoderm during *Xenopus* gastrulation. *Development* **125**:4681–4690.
- Kortemme T**, Baker D. 2002. A simple physical model for binding energy hot spots in protein-protein complexes. *PNAS* **99**:14116–14121. doi: [10.1073/pnas.202485799](https://doi.org/10.1073/pnas.202485799)
- Kortemme T**, Kim DE, Baker D. 2004. Computational alanine scanning of protein-protein interfaces. *Science Signaling* **2004**:pl2. doi: [10.1126/stke.2192004pl2](https://doi.org/10.1126/stke.2192004pl2)
- Kostadinov D**, Sanes JR. 2015. Protocadherin-dependent dendritic self-avoidance regulates neural connectivity and circuit function. *eLife* **4**:e08964. doi: [10.7554/eLife.08964](https://doi.org/10.7554/eLife.08964)
- Krissinel E**, Henrick K. 2007. Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology* **372**:774–797. doi: [10.1016/j.jmb.2007.05.022](https://doi.org/10.1016/j.jmb.2007.05.022)
- Kuroda H**, Inui M, Sugimoto K, Hayata T, Asashima M. 2002. Axial protocadherin is a mediator of prenotochord cell sorting in *Xenopus*. *Developmental Biology* **244**:267–277. doi: [10.1006/dbio.2002.0589](https://doi.org/10.1006/dbio.2002.0589)
- Ledderose J**, Dieter S, Schwarz MK. 2013. Maturation of postnatally generated olfactory bulb granule cells depends on functional γ -protocadherin expression. *Scientific Reports* **3**:1514. doi: [10.1038/srep01514](https://doi.org/10.1038/srep01514)
- Lefebvre JL**, Kostadinov D, Chen WV, Maniatis T, Sanes JR. 2012. Protocadherins mediate dendritic self-avoidance in the mammalian nervous system. *Nature* **488**:517–521. doi: [10.1038/nature11305](https://doi.org/10.1038/nature11305)
- Li S-A**, Cheng L, Yu Y, Chen Q, Li S. 2016. Structural basis of Dscam1 homodimerization: Insights into context constraint for protein recognition. *Science Advances* **2**:e1501118–1501119. doi: [10.1126/sciadv.1501118](https://doi.org/10.1126/sciadv.1501118)
- Marks DS**, Colwell LJ, Sheridan R, Hopf TA, Pagnani A, Zecchina R, Sander C. 2011. Protein 3D structure computed from evolutionary sequence variation. *PLoS One* **6**:e28766. doi: [10.1371/journal.pone.0028766](https://doi.org/10.1371/journal.pone.0028766)
- Meijers R**, Puettmann-Holgado R, Skiniotis G, Liu JH, Walz T, Wang JH, Schmucker D. 2007. Structural basis of Dscam isoform specificity. *Nature* **449**:487–491. doi: [10.1038/nature06147](https://doi.org/10.1038/nature06147)
- Molumby MJ**, Keeler AB, Weiner JA. 2016. Homophilic protocadherin cell-cell interactions promote dendrite complexity. *Cell Reports* **15**:1037–1050. doi: [10.1016/j.celrep.2016.03.093](https://doi.org/10.1016/j.celrep.2016.03.093)
- Nicoludis JM**, Lau SY, Schärfe CP, Marks DS, Weihofen WA, Gaudet R. 2015. Structure and sequence analyses of clustered protocadherins reveal antiparallel interactions that mediate homophilic specificity. *Structure* **23**:2087–2098. doi: [10.1016/j.str.2015.09.005](https://doi.org/10.1016/j.str.2015.09.005)
- Otwinowski Z**, Minor W. 1997. Processing of x-ray diffraction data collected in oscillation mode. In: Carter JR CW, Sweet RM (Eds). *Methods in Enzymology, Volume 276: Macromolecular Crystallography, Part A*. New York: Academic Press. p 307–326. doi: [10.1016/s0076-6879\(97\)76066-x](https://doi.org/10.1016/s0076-6879(97)76066-x)
- Price MN**, Dehal PS, Arkin AP. 2010. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* **5**:e9490. doi: [10.1371/journal.pone.0009490](https://doi.org/10.1371/journal.pone.0009490)

- Rashid D**, Newell K, Shama L, Bradley R. 2006. A requirement for NF-protocadherin and TAF1/Set in cell adhesion and neural tube formation. *Developmental Biology* **291**:170–181. doi: [10.1016/j.ydbio.2005.12.027](https://doi.org/10.1016/j.ydbio.2005.12.027)
- Rubinstein R**, Thu CA, Goodman KM, Wolcott HN, Bahna F, Mannepalli S, Ahlsen G, Chevee M, Halim A, Clausen H, Maniatis T, Shapiro L, Honig B. 2015. Molecular logic of neuronal self-recognition through protocadherin domain interactions. *Cell* **163**:629–642. doi: [10.1016/j.cell.2015.09.026](https://doi.org/10.1016/j.cell.2015.09.026)
- Sawaya MR**, Wojtowicz WM, Andre I, Qian B, Wu W, Baker D, Eisenberg D, Zipursky SL. 2008. A double S shape provides the structural basis for the extraordinary binding specificity of Dscam isoforms. *Cell* **134**:1007–1018. doi: [10.1016/j.cell.2008.07.042](https://doi.org/10.1016/j.cell.2008.07.042)
- Schreiner D**, Weiner JA. 2010. Combinatorial homophilic interaction between gamma-protocadherin multimers greatly expands the molecular diversity of cell adhesion. *PNAS* **107**:14893–14898. doi: [10.1073/pnas.1004526107](https://doi.org/10.1073/pnas.1004526107)
- Sotomayor M**, Gaudet R, Corey DP. 2014. Sorting out a promiscuous superfamily: towards cadherin connectomics. *Trends in Cell Biology* **24**:524–536. doi: [10.1016/j.tcb.2014.03.007](https://doi.org/10.1016/j.tcb.2014.03.007)
- Suo L**, Lu H, Ying G, Capecchi MR, Wu Q. 2012. Protocadherin clusters and cell adhesion kinase regulate dendrite complexity through Rho GTPase. *Journal of Molecular Cell Biology* **4**:362–376. doi: [10.1093/jmcb/mjs034](https://doi.org/10.1093/jmcb/mjs034)
- Thu CA**, Chen WV, Rubinstein R, Chevee M, Wolcott HN, Felsovalyi KO, Tapia JC, Shapiro L, Honig B, Maniatis T. 2014. Single-cell identity generated by combinatorial homophilic interactions between α , β , and γ protocadherins. *Cell* **158**:1045–1059. doi: [10.1016/j.cell.2014.07.012](https://doi.org/10.1016/j.cell.2014.07.012)
- Wang X**, Weiner JA, Levi S, Craig AM, Bradley A, Sanes JR, a. L, Craig S. 2002. Gamma protocadherins are required for survival of spinal interneurons. *Neuron* **36**:843–854. doi: [10.1016/S0896-6273\(02\)01090-5](https://doi.org/10.1016/S0896-6273(02)01090-5)
- Weiner JA**, Wang X, Tapia JC, Sanes JR. 2005. Gamma protocadherins are required for synaptic development in the spinal cord. *PNAS* **102**:8–14. doi: [10.1073/pnas.0407931101](https://doi.org/10.1073/pnas.0407931101)
- Weinreb C**, Riesselman AJ, Ingraham JB, Gross T, Sander C, Marks DS. 2016. 3D RNA and functional interactions from evolutionary couplings. *Cell* **165**:963–975. doi: [10.1016/j.cell.2016.03.030](https://doi.org/10.1016/j.cell.2016.03.030)
- Wojtowicz WM**, Wu W, Andre I, Qian B, Baker D, Zipursky SL. 2007. A vast repertoire of Dscam binding specificities arises from modular interactions of variable Ig domains. *Cell* **130**:1134–1145. doi: [10.1016/j.cell.2007.08.026](https://doi.org/10.1016/j.cell.2007.08.026)
- Wu Q**. 2005. Comparative genomics and diversifying selection of the clustered vertebrate protocadherin genes. *Genetics* **169**:2179–2188. doi: [10.1534/genetics.104.037606](https://doi.org/10.1534/genetics.104.037606)
- Yamagata K**, Andreasson KI, Sugiura H, Maru E, Dominique M, Irie Y, Miki N, Hayashi Y, Yoshioka M, Kaneko K, Kato H, Worley PF. 1999. Arcadlin is a neural activity-regulated cadherin involved in long term potentiation. *Journal of Biological Chemistry* **274**:19473–19479. doi: [10.1074/jbc.274.27.19473](https://doi.org/10.1074/jbc.274.27.19473)
- Yoshida K**. 2003. Fibroblast cell shape and adhesion in vitro is altered by overexpression of the 7a and 7b isoforms of protocadherin 7, but not the 7c isoform. *Cellular & Molecular Biology Letters* **8**:735–741.
- Zipursky SL**, Sanes JR. 2010. Chemoaffinity revisited: dscams, protocadherins, and neural circuit assembly. *Cell* **143**:343–353. doi: [10.1016/j.cell.2010.10.009](https://doi.org/10.1016/j.cell.2010.10.009)