Homophilic and heterophilic interactions of type II cadherins identify specificity groups underlying cell-adhesive behavior

Julia Brasch[a], Phinikoula S. Katsamba[a,b], Oliver J. Harrison[a], Göran Ahlsén[a,b], Regina B. Troyanovsky[c], Indrajyoti Indra[c], Anna Kaczynska[d], Benjamin Kaeser[d], Sergey Troyanovsky[c], Barry Honig[a,b,d,e,*], Lawrence Shapiro[a, d, f,*]

[a] Zuckerman Mind Brain Behavior Institute, New York, NY 10027; USA.
[b] Howard Hughes Medical Institute, Columbia University, New York, NY 10032; USA.
[c] Department of Dermatology, The Feinberg School of Medicine, Northwestern University, Chicago, IL 60611; USA.
[d] Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032; USA.
[e] Center for Computational Biology and Bioinformatics, Department of Systems Biology, and Department of Medicine Columbia University, New York, NY 10032; USA
[f] Lead contact.

*Corresponding authors:
Lawrence Shapiro (lss8@columbia.edu)
Barry Honig (bh6@columbia.edu)

Summary

Type II cadherins are cell-cell adhesion proteins critical for tissue patterning and neuronal targeting, but whose molecular binding code remains poorly understood. Here we delineate binding preferences for type II cadherin cell-adhesive regions revealing extensive heterophilic interactions between specific pairs in addition to homophilic interactions. Three distinct specificity groups emerge from our analysis with members that share highly similar heterophilic binding patterns and favor binding to one another. Structures of adhesive fragments from each specificity group confirm near-identical dimer topology conserved throughout the family, allowing interface residues whose conservation corresponds to specificity preferences to be identified. We show that targeted mutation of these residues converts binding preferences between specificity groups in biophysical and co-culture assays. Our results provide a detailed understanding of the type II cadherin interaction map and a basis for defining their role in tissue patterning and the emerging importance of their heterophilic interactions in neural connectivity.

Introduction

Vertebrate classical cadherins are a family of calcium-dependent cell adhesion receptors whose selective interactions are critical for morphogenesis, patterning and maintenance of solid tissues including the CNS, in which they contribute to neural circuit assembly, axon guidance and synapse formation and plasticity (Basu et al., 2017; Hirano and Takeichi, 2012; Redies et al., 2012; Williams et al., 2011). All are single-pass transmembrane proteins with extracellular regions composed of five successive extracellular cadherin (EC) repeats and intracellular regions containing binding sites for adaptor proteins, α-, β- and p120 catenins, which link adhesion mediated by the extracellular regions to the actin cytoskeleton (Brasch et al., 2012; Hirano and Takeichi, 2012). Classical cadherins can be divided into type I cadherins, comprising E-, N-, P-, R- and M-cadherin, and type II cadherins, which comprise a separate subfamily of thirteen members: cadherins-6, -7, -8, -9, -10, -11, -12, -18, -19, -20, -22, -24 and a divergent member, vascular endothelial (VE)-cadherin (Brasch et al., 2011). While the molecular interactions of type I cadherins have been well characterized, the larger type II cadherin subfamily is comparatively less understood.

Individual type II cadherins are differentially expressed in the CNS (Hirano and Takeichi, 2012), often with expression of distinct subsets demarcating specific sub-regions, as observed in the visual system (Duan et al., 2014), hippocampus (Basu et al., 2017; Bekirov et al., 2002) and spinal cord (Demireva et al., 2011; Patel et al., 2006; Price et al., 2002). In functional studies, single and double type II cadherin knock-out mice show a variety of distinct non-lethal phenotypes relating to cell-targeting and synaptic function in the CNS and to morphogenesis in other tissues. These phenotypes include failure of a subset of retinal ganglion cells to innervate their target neurons ($Cdh6^{-/-}$ mice (Osterhout et al., 2011)), reduction of high-magnitude long-term potentiation (LTP) in the

hippocampus (Cdh9$^{-/-}$, Cdh10$^{-/-}$, Cdh6$^{-/-}$;Cdh10$^{-/-}$ (Basu et al., 2017)), impaired targeting of bipolar cells in the retina (Cdh8$^{-/-}$, Cdh9$^{-/-}$ (Duan et al., 2014), impaired synaptic coupling in cold-sensitive sensory neurons (Cdh8$^{-/-}$ (Suzuki et al., 2007)), and, outside the CNS, delayed kidney development (Cdh6$^{-/-}$ (Mah et al., 2000)) and reduction of bone density (Cdh11$^{-/-}$ (Kawaguchi et al., 2001)). In addition, *in vivo* mis-expression studies demonstrate that expression of specific complements of type II cadherins in individual neurons directs their sorting into segregated populations in the developing chicken spinal cord and mouse telencephalon (Inoue et al., 2001; Patel et al., 2006; Price et al., 2002).

The molecular interactions of type II cadherins underlying these complex behaviors are not yet fully defined. Structural studies of cadherin-8, -11 and -20 and the divergent member VE-cadherin have revealed that type II cadherins form 'strand swapped' adhesive dimers between their membrane-distal EC1 domains in which N-terminal β-strands are reciprocally exchanged (Brasch et al., 2011; Patel et al., 2006). This strand exchange is anchored by docking of two conserved tryptophan residues, Trp2 and Trp4, into a hydrophobic pocket in the partner EC domain with additional interactions contributed by a hydrophobic patch at the base of the domain (Patel et al., 2006), except in the case of VE-cadherin which lacks these additional hydrophobic interactions (Brasch et al., 2011). Individual type II cadherins share this canonical interface, but show selectivity in their binding interactions. In *in vitro* cell aggregation assays, type II cadherins mediate both homophilic adhesive interactions between cells expressing identical cadherins, and selective heterophilic interactions between cells expressing different cadherins (Nakagawa and Takeichi, 1995; Patel et al., 2006; Shimoyama et al., 1999; Shimoyama et al., 2000). Type II cadherins frequently show partially overlapping, though distinct, expression patterns *in vivo*, and there is evidence that both types of

interaction contribute to their roles in cell sorting and targeting (Basu et al., 2017; Duan et al., 2014; Inoue et al., 2001; Osterhout et al., 2011; Patel et al., 2006; Price et al., 2002; Williams et al., 2011). Notably, biological roles for type II cadherin heterophilic interactions have emerged from recent *in vivo* studies of cadherins-8, -9 in the mouse retina (Duan et al., 2014) and cadherins-6, -9 and -10 in the mouse hippocampus (Basu et al., 2017). However, the precise molecular binding preferences underlying type II cadherin function remain to be fully determined.

Here we use a comprehensive biophysical approach to quantitatively analyze homophilic and heterophilic binding interactions of type II classical cadherins. We find that heterophilic interactions between different cadherins are highly selective and are frequently preferred over homophilic interactions. Three distinct 'specificity groups' emerge from our analysis, within which closely related cadherins preferentially interact and exhibit highly similar overall patterns of heterophilic binding preferences, unique to each group. Based on these observations, we examine structural and sequence conservation of the adhesive interface between specificity groups, determine crystal structures of adhesive regions of cadherins from previously unrepresented groups, and identify critical specificity residues that can convert binding preferences between groups in biophysical and cell culture experiments.

Results

Homophilic binding affinities of type II cadherin adhesive fragments

We used a bacterial expression system to produce soluble EC1-2 adhesive fragments of mouse classical type II cadherins-6, -7, -9, -10, -11, -12, -18, -20, and -22 and EC1-3 fragments of cadherins-8 and -24, since the latter were unstable as shorter fragments. These represent all members of the classical type II cadherin family except for cadherin-

19, for which a stable protein could not be produced, and the divergent member vascular endothelial (VE) cadherin, whose characterization we reported previously (Brasch et al., 2011). Table 1 lists homophilic binding affinities determined by sedimentation equilibrium analytical ultracentrifugation (AUC) analysis of the eleven type II cadherins produced here. All formed homodimers in solution and fitting of the data to a monomer-dimer equilibrium model yielded K_D values in the low micromolar range, from 3.1 to 42.2 μM (Table 1). Within this range five cadherins show relatively tight binding affinities with K_D values below 10 μM, three cadherins have intermediate affinities in the 10-30 μM range, and three have weak affinities >30 μM (Table 1). Notably, cadherins with similar homodimerization strengths do not share highest sequence identity. For example, cadherins-6 and -10 share the highest amino acid sequence identity (84 % over EC1-2), but lie at each extreme of the K_D range (Table 1), while cadherin-6 and cadherin-22 share only 61 % sequence identity, yet have nearly identical homophilic affinities (Table 1).

Heterophilic interactions identify distinct type II cadherin specificity groups

Type II cadherins have been shown to exhibit both homophilic and heterophilic binding behavior in cell aggregation studies (Nakagawa and Takeichi, 1995; Patel et al., 2006; Shimoyama et al., 1999; Shimoyama et al., 2000) and *in vivo* (Basu et al., 2017; Duan et al., 2014; Inoue et al., 2001; Osterhout et al., 2011; Patel et al., 2006; Price et al., 2002; Williams et al., 2011). However, a comprehensive quantitative analysis of all heterophilic binding interactions in the family has not been reported. We therefore aimed to delineate the heterophilic binding behavior using surface plasmon resonance (SPR) to measure binding for all cognate pairs. In these experiments, each cadherin adhesive fragment was covalently coupled to a sensor chip surface by thiol-coupling of an engineered C-terminal Cysteine-residue (Cys-tag) to present functional EC1 domains in a favorable

orientation for interaction. All Cys-tagged proteins were analyzed by AUC confirming homophilic binding affinities to be broadly similar to those of the untagged proteins (Table S1). Untagged cadherin-6, -7, -8, -9, -10, -11, -12, -18, -20, and -22 adhesive fragments were passed over each surface and homophilic and heterophilic binding responses were recorded for all combinations to provide a comprehensive SPR matrix of all potential interactions (Figure 1). As expected, homophilic binding interactions were observed for each cadherin tested (Figure 1, diagonal) and dissociation rates varied up to six fold between cadherins (Table S2). In addition to these homophilic responses, all cadherin surfaces supported significant levels of heterophilic binding to selective subsets of cadherin family members (Figure 1, rows). Importantly, response levels for the strongest heterophilic interactions for each surface were comparable to or exceeded those of the respective homophilic interactions, suggesting functional significance. Homophilic binding was favored over all other heterophilic interactions for only two cadherins, cadherin-6 and -20 (Figure 1 and Figure S1). Heterophilic binding affinities could not be determined from the SPR data due to competing homodimerization of surface and analyte cadherins (Katsamba et al., 2009), nevertheless, relative binding strengths could be assessed by comparing response levels over the same surface (Figure 1, rows). Based on the precise pattern of binding preferences observed for each cadherin, the type II cadherin family can be divided into three distinct ‘specificity groups’. Within these three groups, members share near identical binding profiles, in that they bind heterophilically to the same set of cadherins (Figure 1, compare rows), and show a preference for interactions within the same group (Figure 1, boxes).

Cadherins-8 and -11 comprise one such specificity group and display clear preference for heterophilic binding to each other over all other cadherins (Figure 1, top rows). The cadherin-8 surface supported heterophilic binding of cadherin-11 as the strongest

observed interaction, followed by homophilic binding of the cadherin-8 analyte (Figure 1, top row). All other analytes bound at very low levels: cadherins-20 and -18 bound very weakly while cadherins-6, -9, -10 and cadherins-7, -12 and -22 did not show binding above background levels. A similar binding pattern was observed for the cadherin-11 surface with heterophilic binding of cadherin-8 comprising the strongest response, followed by homophilic binding and comparatively lower binding to all other analytes including modest levels of binding to cadherins-12, -20, -18 and -22 (Figure 1, second row).

Cadherins-6, -9, and -10 define a second specificity group whose members share closely similar binding profiles (Figure 1). On surfaces coated with cadherin-6, -9 or -10, binding of cadherin-6 analyte showed the strongest response, followed by binding of cadherins-9 and -10. No binding of cadherins-8 and -11 was observed on any of the three surfaces, while the remaining cadherins, cadherin-12, -20, -18, -22 and -7, showed intermediate binding comparable to that of the cadherin-10 analyte.

The third specificity group comprises cadherins-12, -18, -20, -22, and -7 (Figure 1). These cadherins showed generally strong binding responses to each other (Figure 1, green box), intermediate binding to the cadherin-6/9/10 specificity group and no strong binding responses to cadherins-8 and -11. Cadherin-7 binding responses as an analyte were lower overall than those of other members in this group (Figure 1, right column), but the binding profile of cadherin-7 was closely similar to those of cadherin-12, -18, -20, and -22 (Figure 1, bottom row).

To determine if heterophilic interactions between type II cadherins form through the same strand-swap binding mechanism shown to mediate homophilic interactions (Patel

et al., 2006) and see Figure 4 below), we tested the effects of ablating this interface in SPR (Figure 2). As expected, alanine substitution of Trp4, which anchors the strand-swap dimer, ablated homophilic binding of cadherins-6, -8 and -11 in SPR and AUC (Figure 2A-C, W4A-mutants; Table S1), and, importantly, also abolished heterophilic binding between all cognate pairs (Figure 2A-C). In addition, mutation of the X-dimer interface (M188D) in cadherin-6, which was previously shown to be a necessary binding intermediate for homophilic interactions (Harrison et al., 2010), also resulted in severely diminished heterophilic binding of cadherin-6 M188D mutant to cadherins-9 and -10 (Figure 2A). Residual heterophilic binding between cadherin-6 M188D and cadherin-9 was abolished in a double mutant, in which both strand-swap and X-dimer interfaces were mutated (Figure 2A, W4A M188D-mutant). Together these data confirm that homophilic and heterophilic binding interactions in the type II cadherin family form through the same adhesive interfaces.

Specificity groups correspond to branches of the type II cadherin phylogenetic tree

To compare the relative strengths of heterophilic interactions across the type II cadherin family, values derived from normalized responses for each interaction in our SPR matrix were used to weight a force-directed interaction network (Figure 3A). Cadherins linked closely in the network reflect strong binding interactions (e.g. cadherins-18 and -22), while cadherins linked distantly in the network reflect weak or background level binding interactions (e.g. cadherins-8 and -9). As expected, three discrete clusters emerge from the binding network which are consistent with the specificity groups we identified above based on qualitative comparison of our SPR traces (see Figure 1).

We next investigated if cadherins belonging to the three specificity groups in SPR share other characteristics. Remarkably, the groups derived purely from the binding data correspond closely to the phylogenetic grouping generated from alignment of sequences of adhesive EC1-2 regions (Figure 3B). This suggests that type II cadherins bind to, and share binding preferences with, phylogenetically closely related family members. Cadherins-8 and -11, in particular, are more distantly related to the other members of the type II cadherin family (Figure 3B), which is reflected in their binding preferences. Cadherin-24, which was excluded from SPR analyses due to solubility problems (see Methods), is also likely to share binding preferences with cadherin-8 and -11 since it belongs to the same phylogenetic branch. Interestingly, the divergence of the cadherin-6/9/10 and cadherin-8/11 groups also extends to the sequence of their cytoplasmic domains (Nollet et al., 2000; Sotomayor et al., 2014), suggesting the possibility of differences in downstream events upon binding.

The clustering of cadherins into specificity groups is also partially reflected in their chromosomal locations. Genes encoding the closely related cadherins-8 and -11 are located in close proximity in mouse chromosome 8, human chromosome 16, and chicken chromosome 11 (Nollet et al., 2000). Similarly, cadherins-6, -9 and -10 are clustered on the same chromosome in mouse (chromosome 15), human (chromosome 5) and chicken (chromosome 2). Together with the phylogenetic data, these observations suggest that type II cadherins sharing similar binding preferences are likely to have diverged most recently during evolution.

Crystal structures of type II cadherins show highly conserved dimer topology across all specificity groups

To investigate the molecular basis of grouped specificity behavior observed in SPR, we set out to compare homodimer structures of representative type II cadherins. We have previously reported crystal structures of adhesive fragments of cadherin-8, -11 and -20 representing two members of the cadherin-8/11 group and a single member of the cadherin-7/12/18/20/22 group (Figure 3B, (Patel et al., 2006)). To extend structural coverage to include multiple representatives for all specificity groups, we have now determined crystal structures of adhesive EC1-2 fragments of cadherins-6 and -10 belonging to the previously uncharacterized cadherin-6/9/10 group and cadherins-7 and -22 from additional sub-branches of the cadherin-7/12/18/20/22 specificity group (Figures 3B and 4).

Crystals of cadherins-6, -10, -7 and -22 diffracted to between 1.7 and 2.7 Å resolution (Table S3) and structures were solved by molecular replacement. Each cadherin structure adopted an extended conformation rigidified by coordination of three calcium ions in the inter-domain linker regions and formed strand-swapped homodimers (Figure 4A), as was observed for cadherins-8, -11 and -20 in the past (Patel et al., 2006). In each dimer, reciprocal exchange of A-strands between EC1 domains is anchored by docking of Trp2 and Trp4 residues into a hydrophobic pocket of the partner molecule, bringing strands A, B and G into intermolecular contact (Figure 4B). Further buried surface area is also contributed by the BC-loop of domain EC2, which packs against strands B and E of the partner EC1 domain in all four dimer structures (Figure 4A, arrows).

Superposition of these new structures with previously determined type II cadherin homodimer structures reveals that the overall dimer topology, including the angle between partner EC1 domains, is essentially identical across members of all three specificity groups (Figure 4B, r.m.s.d. < 0.9 Å between 145-185 aligned Cα-atoms per dimer). Owing to this shared topology, identical regions of the EC1 A, B, and G strands contact each other in all type II cadherin dimers. While the interface is dominated by docking of Trp2 and Trp4 into the partner hydrophobic pocket formed by residues from the B, C, F and G strands, additional mostly hydrophobic contacts form between paired A-strands and between strands B and A or G, extending the dimer interface over the whole face of the domain in all structures.

Specificity determinants in the type II cadherin adhesive interface

Structural conservation of the type II cadherin adhesive interface across all branches of the family suggests that subtype-specific differences in structurally equivalent interfacial residues could govern binding preferences. Since we observe binding within specificity groups that is favored over binding between groups, residues in the interface with group-specific conservation may be particularly important. We aligned amino acid sequences of type II cadherin EC1 domains from mouse, human, and chicken, and examined sequence conservation both within and across specificity groups. Figure 4C shows a sequence logo representation calculated separately for each specificity group. Interfacial residues derived from all available crystal structures were mapped onto the sequences to identify conserved and variable regions of the interface (Figure 4C, magenta bars). The majority of interface residues are fully conserved, or conserved in consensus, across the type II cadherin family and comprise the core of the strand swap interface including residues of the exchanged A-strands, the acceptor pocket and most of the hydrophobic patch towards the base of the domain (Fig. 4C-E, grey shading). Positions

of variable residue identity, where consensus residues differ between at least two specificity groups, are restricted to ten residues at the periphery of the interface and define two distinct regions (Figure 4C-E, highlights). First, the lower part of the interface is encircled by variable residues contributed by the base of the A-strand (L/V9, L/V/I10); the base of the B-strand (Q/L/V/I19, Y/L20, K/R23); the E-strand (D/N56); and the base of the G-strand (I/V96, H/Q97). Second, the upper periphery of the interface contains two group-specific residues, namely M/V3 in the A-strand and E/P89 at the top of the G-strand.

Within the ten variable interface residues, five are not fully conserved within each respective specificity group and thus are less likely to underlie shared group binding behavior (Figure 4C-E, green highlighting). The remaining five residues show perfect conservation within each specificity group whilst differing between at least two groups (Figure 4C-E, blue highlighting), strengthening their candidacy as specificity determinants. Most strikingly, residues Y/L20 and H/Q97 form three distinct group-specific interaction pairs in the adhesive homodimer structures (Figure 4F). In cadherin-6 and -10 structures, Tyr20 and His97 are closely apposed and engage in near-parallel π-π-stacking interactions at approximately 3.3 Å distance (Figure 4F, left). In the cadherin-7, -20 and -22 homodimers, Tyr20 and Gln97 are also apposed, and likely engage in van der Waals interactions between the amino group of Gln97 and the tyrosine ring (Burley and Petsko, 1986). In the cadherin-8/11 group, Leu20 and Gln97 are in proximity, but do not interact closely. These subgroup-specific pairwise interactions appear likely to contribute to the restricted binding preferences observed in SPR. The remaining residues showing strict group specific conservation, M/V3 and L/V9 in the A-strand and E/P89 at the top of the G-strand forming the rim of the pocket, do not form such interacting pairs with each other or with any other potential specificity determining

residues, but instead contact fully conserved residues (Figure 4C-E). Nonetheless, they could contribute to specificity indirectly through effects on neighboring residues or by overall effects on the shape of the interaction surface.

Targeted mutation of group-specific interface residues converts binding preferences in SPR

We investigated whether these group-specific residues governed binding preferences in SPR. We first introduced point mutations at positions 20 and 97, which form distinct interacting pairs at the interface in each specificity group (see Figure 4F). For these experiments we chose cadherin-6 from the cadherin-6/9/10 group and cadherin-11 from the cadherin-8/11 group, since these groups show clearly distinguishable binding preferences in SPR (Figure 1). Tyr20 and His97 in cadherin-6 were substituted with the corresponding Leu20 and Gln97 residues from cadherin-11 (cad-6 LQ mutant) and the reverse substitutions were made in cadherin-11 (cad-11 YH mutant). Binding of mutants was tested against wild-type proteins in SPR (Figure 5A). As described above, cadherin-6 binds selectively to cadherin-6, -9 and -10 surfaces, but does not bind to cadherin-8 and -11 surfaces (Figure 5A, left). Over the same set of surfaces cad-6 LQ mutant displayed decreased binding to the cad-6, -9 and -10 surfaces, but concomitantly increased binding to cadherin-8 and -11 surfaces (Figure 5A), consistent with conversion of overall specificity. Corresponding behavior was observed for the cad-11 YH mutant, which, compared to wild-type cadherin-11, showed dramatically decreased binding to members of the same specificity group (cadherin-8, -11) with concomitantly increased binding to cadherins-6, -9 and -10 (Figure 5A). Our results are consistent with a decisive role for residues 20 and 97 as specificity determinants in the cadherin-6/9/10 and -8/11 specificity groups.

Since binding responses of the cad-6 LQ and cad-11 YH mutants were lower than those of the target wild-type proteins, we tested if conversion of additional group-specific residues enhanced changes in binding preference. We mutated the group-specifically conserved surface residues M/V3 and E/P89 in cad-6 LQ and cad-11 YH mutants to generate quadruple mutants cad-6 VLEQ and cad-11 MYPH. In SPR, cad-11 MYPH mutant showed a similar conversion of binding preferences to the double mutant, but with additional enhancement of responses to cadherins-6, -9 and -10 (Figure 5A). Similarly, cad-6 VLEQ also showed enhanced binding to the opposite specificity group in comparison to the cad-6 LQ double mutant, though in this case binding was also increased over same specificity group surfaces (Figure 5A). Since both quadruple mutants enhanced binding to the opposite specificity group, our data suggest that M/V3 and E/P89 contribute modestly to specificity, at least in combination with residues L/Y20 and H/Q97. In order to interconvert all 37 non-identical EC1 residues between cadherin-6 and -11 we also prepared chimeric proteins cad-6_{EC1}cad-11_{EC2} (cad-6/11) and cad-11_{EC1}cad-6_{EC2} (cad-11/6). We confirmed the fidelity of these chimeric proteins in AUC (Table 1) and by structural analysis of cad-11/6 (Figure 4A, right panel). As expected both formed homophilic dimers, and the structure confirmed the cad-11/6 homodimer arrangement to be near-identical to that of wild-type cadherin-11. In SPR experiments these chimeric proteins near-perfectly mimicked the binding behavior of the wild-type proteins from which their EC1 domains derived (Figure 5B) confirming that complete inter-conversion of binding behavior can be observed when sufficient residues are changed and that all specificity determinants are restricted to EC1. Thus our data suggest that residues 20 and 97, and to a lesser extent residues 3 and 89, are major specificity determinants with other variable interface residues in EC1 likely contributing indirectly to precise binding specificities.

Localization of full-length type II cadherins to heterotypic cell contacts depends on specificity determinants.

To relate our biophysical observations for adhesive fragments and their mutants to the behavior of full-length proteins at cell-cell contact sites, we examined localization of fluorescently labeled type II cadherins in transfected A431D cells (Figure 6 and S2). Full-length cadherins-8 and -11, and cadherins-6 and -10, representing pairs from distinct specificity groups, were labeled at their C-termini with either red-fluorescent mCherry, or green fluorescent Dendra2. These were transfected singly into A431D cell lines that were then co-cultured to allow formation of homotypic cell contacts between cells of the same cell line and heterotypic contacts between cells from different cell lines. When combinations of cadherins from the same specificity group, cadherins-8 and 11, or -6 and -10, were co-cultured, both cadherins co-localized at heterotypic contact sites, in addition to the respective homotypic sites, consistent with heterophilic binding observed for these pairs in SPR (Figure 6B, E; arrow heads). In marked contrast, co-culture of cells expressing mismatched cadherins-6 and -11, or cadherins-8 and -10 from different specificity groups produced heterotypic cell contact sites devoid of cadherins, which accumulated only at homotypic sites (Figure 6C and E). Lastly, a homophilic pair, cadherin-11-dendra and cadherin-11-cherry localized equally to homotypic and heterotypic contacts (Figure 6A, arrowheads) reflecting uniform homophilic interactions. The cell adhesive behavior of these full-length wild-type proteins thus closely mirrors binding preferences observed in biophysics.

We also tested our quadruple specificity mutants in co-culture assays with full-length cadherin-6 and -11 (cad-6 VLEQ, cad-11 MYPH). Compared to wild-type protein, localization of cad-11 MYPH mutant to heterotypic cell-cell contacts with cadherin-8 (compare Figure 6G and B) or cadherin-11 (compare Figure 6F to A) from the same

specificity group was ablated or dramatically reduced (Figure S2). At the same time, cad-11 MYPH mutant now strikingly co-localized with cadherin-6 from the opposite specificity group (compare Figure 6H to C), showing a significant shift in overall binding preference (Figure 6 and S2). Similar behavior was observed for mutant cad-6 VLEQ, which acquired strong co-localization with cadherin-11 from the opposite specificity group (compare Figure 6I to C), while co-localization with cadherin-10 was concomitantly decreased (compare Figure 6J to E). Behavior of these mutants shows that conversion of four residues between specificity groups produces a shift of binding preferences sufficient to convert adhesive specificity between cells.

Discussion

Type II cadherins represent a large family of adhesion proteins with overlapping differential expression patterns and diverse functional roles, presenting a challenge in relating their molecular and functional properties. The comprehensive matrix of binding interactions determined here reveals that all type II cadherins tested participate both in homophilic and heterophilic interactions that are of comparable strengths. Heterophilic interactions are not promiscuous: each member of the family binds to a specific subset of other members with characteristic relative response levels. Importantly, these subsets and responses are shared between multiple cadherins, giving rise to three distinct specificity groups: 6/9/10, 8/11 and 7/12/18/20/22. Heterophilic interactions form between all members within each specificity group, and weaker heterophilic interactions form between the 6/9/10 and 7/12/18/20/22 groups, while the 8/11 group is more isolated (Figure 1). Within groups, individual cadherins are nonetheless distinguished by widely differing homodimerization affinities (Table 1) and dissociation rates (Table S2).

We observe consistent specificity behavior between SPR experiments with purified adhesive fragments and cellular co-culture experiments with full-length proteins (Figures 1 and 6). Our results are also in agreement with previous co-culture experiments testing other pairings from the 6/9/10 specificity group (Basu et al., 2017) and with cell-cell aggregation assays using transfected human and chicken type II cadherins (Basu et al., 2017; Nakagawa and Takeichi, 1995; Patel et al., 2006; Shimoyama et al., 1999; Shimoyama et al., 2000). In a broad cell aggregation study using a range of human type II cadherins, selective co-aggregation of a number of specific cadherin pairs was observed (Shimoyama et al., 2000), representing a subset of the heterophilic binding pairs we observe in SPR. Heterophilic pairs only observed in SPR were either not tested in the co-aggregation study (all interactions involving cadherins-20 and -22, Figure 1) or did not produce detectable co-aggregation, likely due to differences in assay sensitivity and variable cadherin expression levels in transfected cells. Nonetheless, observation of closely similar specificity preferences across multiple assay systems and species supports the biological relevance of the type II cadherin adhesive interaction map determined here.

By sequence analysis we identified potential N-linked glycosylation sites at Asn202 for cadherins-6, -9, -10, -12, -18 and -20, at Asn127 for cadherin-8, and no sites in cadherins-7, -11, -22 and -24. Mapping the sites onto type II cadherin structures reveals Asn127 and Asn202 to be located on the B- and G-strands of EC1, respectively, distal from adhesive sites. Together with close correlation of binding preferences observed between SPR and cell-based experiments this suggests that absence of glycosylation in the bacterially produced adhesive fragments is unlikely to affect interaction behavior of type II cadherins, as we have also observed previously for adhesive regions of VE-cadherin (Brasch et al., 2011).

Type II family members cadherins-19, -24 and VE-cadherin were excluded from our SPR experiments. Of these, cadherin-24 is predicted to be part of the 8/11 specificity group, based on its position in the phylogenetic tree and conservation of potential specificity residues Val3, Leu20, Pro89 and Gln97 (Figures 3 and 4). Cadherin-19 can not be assigned to one of the specificity groups identified here based on sequence information, since it contains unique residues at putative specificity sites 20 and 97 and lacks conservation of invariant residues Arg5 and Gln6 in the swapped A-strand. Based on these observations, we predict that cadherin-19 is unlikely to engage in heterophilic binding with other members of the type II cadherin family, though this remains to be tested. As reported previously, the structure and interface characteristics of the VE-cadherin homodimer are divergent from, and very likely incompatible with, those of other members of the type II cadherin family, consistent with its specialized biological role (Brasch et al., 2011).

Consistent with the extensive heterophilic interactions observed for type II cadherins, structure and topology of the adhesive homodimers are near identical and residues mediating the core interactions of strand swapping are conserved (Figure 4). The observations that heterophilic and homophilic interactions are formed through the same interface (Figure 2), and that phylogenetically related cadherins maintain heterophilic recognition (Figure 3), suggest that divergence of a common binding mechanism gave rise to the selective heterophilic interactions we observe. In the most extreme cases, the interfaces of cadherins-8 and -11 appear to have diverged sufficiently to become incompatible with all other members of the family, particularly the 6/9/10 group (Figure 1). Variable residues are restricted to the periphery of the interface and only a small number of these show group-specific conservation. In a previous study we postulated

that some of these variable residues could be specificity determinants (Patel et al., 2006). However, identification of specificity groups and determination of structures of multiple members of each have now allowed us to identify the variable residues most likely to contribute to overall specificity preferences (Figure 4). Targeted mutations identified positions 20 and 97, a group-specific apposed pair in the strand-swapped dimer, as major determinants of the incompatibility between the 6/9/10 and 8/11 groups (Figure 5, 6). These same residues could in principle also explain the distinct specificity of the 7/12/18/20/22 group, however, since they are partially shared with the 6/9/10 and 8/11 groups, it is likely that other variable residues also contribute. The residues tested by mutation also only partially account for differences in homophilic binding affinity and subtle differences in heterophilic binding responses of individual members in each group, which can be converted more substantially by exchange of EC1 domains in chimeric proteins (Figure 5 and (Patel et al., 2006)).

Binding interactions have been comprehensively determined for a limited number of other adhesion protein families within the cadherin and immunoglobulin (Ig) superfamilies. Close correspondence has been consistently observed between molecular binding affinities, behavior in cell aggregation assays and phenotypes *in vivo*, and has revealed a range of overall binding characteristics that differ from those observed here for type II cadherins. Like type II cadherins, clustered protocadherin and Dscams families each share a canonical adhesive interface, however, they uniformly favor homophilic interactions in cell aggregation experiments, likely reflecting selection pressure for homophilic binding by their biological roles in neuronal identity and self-recognition (Goodman et al., 2016; Hattori et al., 2008; Rubinstein et al., 2017; Thu et al., 2014). Nectins and SynCAMs, families of vertebrate Ig-like proteins, are able to form both homophilic and heterophilic binding interactions through canonical interfaces as

observed for type II cadherins (Fogel et al., 2007; Harrison et al., 2012; Narita et al., 2011). However, in both cases, heterophilic interactions are strongly preferred, correlating with primary biological functions in heterotypic adhesion (Takai et al., 2008). Desmosomal cadherins also show a strong preference for heterophilic binding (Harrison et al., 2016), although the biological role of this preference remains to be determined. Binding characteristics of type II cadherins are partially reminiscent of those of their close relatives type I cadherins. These display a mixture of homophilic and specific heterophilic binding in SPR experiments (Katsamba et al., 2009; Vendome et al., 2014), but in standard cell aggregation assays type I cadherin heterophilic pairs form only partially mixed aggregates (Katsamba et al., 2009; Shan et al., 2000) and do not fully intermix as was observed for a number of type II cadherin pairs (Patel et al., 2006; Shan et al., 2000; Shimoyama et al., 1999; Shimoyama et al., 2000). The extensive selective heterophilic binding of type II cadherins in combination with their differential homophilic affinities (Table 1), where neither type of interaction is dominantly preferred, may allow them to encode subtle differences in adhesiveness to drive fine sorting events within generally cohesive tissues. Notably, for heterophilic pairs that co-localize to heterotypic junctions in our co-culture assays, homophilic junctions in the same cells do not appear depleted of the respective cadherins (Figure 6). Therefore, in tissues co-expressing multiple type II cadherins, we would expect both types of interaction to contribute to adhesive identity of individual cells.

Biological roles for homophilic interactions of type II cadherins -6, -7 and -20 have been suggested by mis-expression studies in the developing chicken spinal cord, chicken optic tectum and the mouse telencephalon (Inoue et al., 2001; Patel et al., 2006; Price et al., 2002; Treubert-Zimmermann et al., 2002). Experimental equalization of type II cadherin complements of normally segregated tissues led to mis-sorting or mis-targeting

of neurons into compartments expressing matching cadherins. Similarly, homophilic binding of cadherins-6 and -9 appear to be required for targeting of retinal ganglion cells in the mouse visual system (Osterhout et al., 2011) and for formation and differentiation of mossy fiber synapses in the hippocampus (Williams et al., 2011). While these studies do not exclude the potential involvement of heterophilic interactions, homophilic adhesion provides a parsimonious explanation for the observed phenotypes.

Biological roles for heterophilic binding of type II cadherins have emerged only recently. In mouse retina, targeting of Cdh8-expressing type 2 bipolar cells (BC2) and Cdh9-expressing type 5 bipolar cells (BC5) to distinct sublaminae of the inner plexiform layer depended critically on their respective cadherin identities (Duan et al., 2014). However, surrounding cells of the target sublaminae do not express these cadherins and targeting of BC2s and BC5s expressing ectopic cadherin-8 or cadherin-9 is maintained even in cadherin knockout backgrounds, ruling out dependence on homophilic interactions (Duan et al., 2014). These data suggest heterophilic interactions of cadherins 8 and 9 with partner cadherins in the target layers, though the exact complement of type II cadherins in the IPL remains to be determined. The mouse hippocampus provides a clearer example of *in vivo* function driven by heterophilic interactions (Basu et al., 2017). High-magnitude long-term potentiation (LTP) in the stratum oriens, where hippocampal CA3 neurons synapse with basal dendrites of CA1 neurons, depends on presynaptic cadherin-9 expressed only in CA3 and on postsynaptic cadherins-6 and -10 expressed only in CA1 neurons. Knock out of either cadherin-9 alone or both cadherins-6 and -10 produces identical phenotypes, in which high-magnitude LTP is impaired (Basu et al., 2017). Since each cadherin is restricted to one side of the synapse, these findings implicate heterophilic interactions within the 6/9/10 specificity group in functioning of this

neural circuit. Further functional roles for the broad range of selective heterophilic interactions we identify here await determination.

Notably, specific functions of individual type II cadherins may be masked by functional redundancy due in part to their frequently overlapping expression patterns and in part to the shared binding preferences we observe for members of each specificity group. An example of this type of redundancy was observed in the mouse hippocampal circuit described above: single knock-out of cadherin-10 alone maintained wild-type levels of high-magnitude LTP, which was impaired only when cadherins-6 and -10 were knocked out together (Basu et al., 2017). This is likely because cadherin-6 could substitute for cadherin-10 in the single knock out to bind heterophilically to cadherin-9. Surprisingly subtle phenotypes have been observed for numerous single type II cadherin knock outs (Basu et al., 2017; Duan et al., 2014; Kawaguchi et al., 2001; Mah et al., 2000; Osterhout et al., 2011; Saarimaki-Vire et al., 2011; Suzuki et al., 2007), despite their strong expression in the CNS and these could be similarly explained by functional substitution by other cadherins belonging to the same specificity groups. Knock out of complete specificity groups should reveal additional functional roles for type II cadherins. Our analysis of specific binding patterns of type II cadherins and their segregation into distinct specificity groups provide a basis for further investigation of their biological roles.

Experimental Procedures

Protein expression and purification

Recombinant type II cadherin ectodomain fragments were expressed in bacteria and purified from lysate by nickel affinity chromatography, ion exchange chromatography and gel filtration.

Biophysical analyses

AUC experiments were performed at 25 °C in a Beckman XL-A/I analytical ultracentrifuge with UV detection. For SPR, Cysteine-tagged proteins were captured by thiol coupling to CM4 sensor chips, and binding of analytes was assessed in a Biacore T100 biosensor at 25 °C.

Structure determination

Protein crystals were grown by vapor diffusion in hanging drops in the conditions listed in the Supplemental Experimental Methods. X-ray diffraction data were collected from single crystals at 100 K, using a wavelength of 0.979 Å at the X4 beamlines at Brookhaven National Laboratory. Structures were solved by molecular replacement and refined using phenix (Adams et al., 2010).

Co-culture assays

Full-length cadherins with C-terminal dendra2-Myc-or mCherry-Flag tags in the vector pRc/CMV were transfected into human A-431D cells, and co-cultures were analyzed by fluorescence microscopy as described previously (Hong et al., 2010).

Accession numbers

The atomic coordinates of mouse cadherin-6 EC12, cadherin-7 EC12, cadherin-10 EC12, cadherin-22 EC12 and chimera cadherin-1 EC1 cadherin-6 EC2 were deposited with the Protein Data Bank with accession codes 6CGU, 6CGS, 6CG6, 6CG7 and 6CGB, respectively.

Acknowledgements

This work was supported by US National Institutes of Health grant R01 GM062270 (L.S.), and R01 AR044016 (S.T.) and US National Science Foundation grant MCB-1412472 (B.H.). X-ray data were acquired at the X4A and X4C beamlines of the National Synchrotron light source, Brookhaven National Laboratory; the beamlines are operated

by the New York Structural Biology Center. We thank J. Schwanof, R. Abramowitz and Qun Liu at BNL for support with synchrotron data collection.

Author Contributions

JB designed experiments, prepared proteins, crystallized proteins and determined structures; PSK designed and performed the SPR experiments; OJH determined structures; GA performed analytical ultracentrifugation experiments; BK crystallized chimeric protein; AK crystallized cadherin-7 protein; RBT, II and ST performed all co-culture experiments; JB, OJH, PSK, ST, BH, and LS analyzed data and wrote the manuscript.

Declaration of Interests

The authors declare no competing interests.

References

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W.*, et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr *66*, 213-221.

Basu, R., Duan, X., Taylor, M.R., Martin, E.A., Muralidhar, S., Wang, Y., Gangi-Wellman, L., Das, S.C., Yamagata, M., West, P.J.*, et al.* (2017). Heterophilic Type II Cadherins Are Required for High-Magnitude Synaptic Potentiation in the Hippocampus. Neuron *96*, 160-176 e168.

Bekirov, I.H., Needleman, L.A., Zhang, W., and Benson, D.L. (2002). Identification and localization of multiple classic cadherins in developing rat limbic system. Neuroscience *115*, 213-227.

Brasch, J., Harrison, O.J., Ahlsen, G., Carnally, S.M., Henderson, R.M., Honig, B., and Shapiro, L. (2011). Structure and binding mechanism of vascular endothelial cadherin: a divergent classical cadherin. J Mol Biol *408*, 57-73.

Brasch, J., Harrison, O.J., Honig, B., and Shapiro, L. (2012). Thinking outside the cell: how cadherins drive adhesion. Trends Cell Biol *22*, 299-310.

Burley, S.K., and Petsko, G.A. (1986). Amino-aromatic interactions in proteins. FEBS Lett *203*, 139-143.

Demireva, E.Y., Shapiro, L.S., Jessell, T.M., and Zampieri, N. (2011). Motor neuron position and topographic order imposed by beta- and gamma-catenin activities. Cell *147*, 641-652.

Duan, X., Krishnaswamy, A., De la Huerta, I., and Sanes, J.R. (2014). Type II cadherins guide assembly of a direction-selective retinal circuit. Cell *158*, 793-807.

Fogel, A.I., Akins, M.R., Krupp, A.J., Stagi, M., Stein, V., and Biederer, T. (2007). SynCAMs organize synapses through heterophilic adhesion. J Neurosci *27*, 12516-12530.

Goodman, K.M., Rubinstein, R., Thu, C.A., Bahna, F., Mannepalli, S., Ahlsen, G., Rittenhouse, C., Maniatis, T., Honig, B., and Shapiro, L. (2016). Structural Basis of Diverse Homophilic Recognition by Clustered alpha- and beta-Protocadherins. Neuron *90*, 709-723.

Harrison, O.J., Bahna, F., Katsamba, P.S., Jin, X., Brasch, J., Vendome, J., Ahlsen, G., Carroll, K.J., Price, S.R., Honig, B.*, et al.* (2010). Two-step adhesive binding by classical cadherins. Nat Struct Mol Biol *17*, 348-357.

Harrison, O.J., Brasch, J., Lasso, G., Katsamba, P.S., Ahlsen, G., Honig, B., and Shapiro, L. (2016). Structural basis of adhesive binding by desmocollins and desmogleins. Proc Natl Acad Sci U S A *113*, 7160-7165.

Harrison, O.J., Vendome, J., Brasch, J., Jin, X., Hong, S., Katsamba, P.S., Ahlsen, G., Troyanovsky, R.B., Troyanovsky, S.M., Honig, B.*, et al.* (2012). Nectin ectodomain structures reveal a canonical adhesive interface. Nat Struct Mol Biol *19*, 906-915.

Hattori, D., Millard, S.S., Wojtowicz, W.M., and Zipursky, S.L. (2008). Dscam-mediated cell recognition regulates neural circuit formation. Annu Rev Cell Dev Biol *24*, 597-620.

Hirano, S., and Takeichi, M. (2012). Cadherins in brain morphogenesis and wiring. Physiol Rev *92*, 597-634.

Hong, S., Troyanovsky, R.B., and Troyanovsky, S.M. (2010). Spontaneous assembly and active disassembly balance adherens junction homeostasis. Proc Natl Acad Sci U S A *107*, 3528-3533.

Inoue, T., Tanaka, T., Takeichi, M., Chisaka, O., Nakamura, S., and Osumi, N. (2001). Role of cadherins in maintaining the compartment boundary between the cortex and striatum during development. Development *128*, 561-569.

Katsamba, P., Carroll, K., Ahlsen, G., Bahna, F., Vendome, J., Posy, S., Rajebhosale, M., Price, S., Jessell, T.M., Ben-Shaul, A.*, et al.* (2009). Linking molecular affinity and cellular specificity in cadherin-mediated adhesion. Proc Natl Acad Sci U S A *106*, 11594-11599.

Kawaguchi, J., Azuma, Y., Hoshi, K., Kii, I., Takeshita, S., Ohta, T., Ozawa, H., Takeichi, M., Chisaka, O., and Kudo, A. (2001). Targeted disruption of cadherin-11 leads to a reduction in bone density in calvaria and long bone metaphyses. J Bone Miner Res *16*, 1265-1271.

Mah, S.P., Saueressig, H., Goulding, M., Kintner, C., and Dressler, G.R. (2000). Kidney development in cadherin-6 mutants: delayed mesenchyme-to-epithelial conversion and loss of nephrons. Dev Biol *223*, 38-53.

Nakagawa, S., and Takeichi, M. (1995). Neural crest cell-cell adhesion controlled by sequential and subpopulation-specific expression of novel cadherins. Development *121*, 1321-1332.

Narita, H., Yamamoto, Y., Suzuki, M., Miyazaki, N., Yoshida, A., Kawai, K., Iwasaki, K., Nakagawa, A., Takai, Y., and Sakisaka, T. (2011). Crystal Structure of the cis-Dimer

of Nectin-1: implications for the architecture of cell-cell junctions. J Biol Chem *286*, 12659-12669.

Nollet, F., Kools, P., and van Roy, F. (2000). Phylogenetic analysis of the cadherin superfamily allows identification of six major subfamilies besides several solitary members. J Mol Biol *299*, 551-572.

Osterhout, J.A., Josten, N., Yamada, J., Pan, F., Wu, S.W., Nguyen, P.L., Panagiotakos, G., Inoue, Y.U., Egusa, S.F., Volgyi, B.*, et al.* (2011). Cadherin-6 mediates axon-target matching in a non-image-forming visual circuit. Neuron *71*, 632-639.

Patel, S.D., Ciatto, C., Chen, C.P., Bahna, F., Rajebhosale, M., Arkus, N., Schieren, I., Jessell, T.M., Honig, B., Price, S.R.*, et al.* (2006). Type II cadherin ectodomain structures: implications for classical cadherin specificity. Cell *124*, 1255-1268.

Price, S.R., De Marco Garcia, N.V., Ranscht, B., and Jessell, T.M. (2002). Regulation of motor neuron pool sorting by differential expression of type II cadherins. Cell *109*, 205-216.

Redies, C., Hertel, N., and Hubner, C.A. (2012). Cadherins and neuropsychiatric disorders. Brain Res *1470*, 130-144.

Rubinstein, R., Goodman, K.M., Maniatis, T., Shapiro, L., and Honig, B. (2017). Structural origins of clustered protocadherin-mediated neuronal barcoding. Semin Cell Dev Biol *69*, 140-150.

Saarimaki-Vire, J., Alitalo, A., and Partanen, J. (2011). Analysis of Cdh22 expression and function in the developing mouse brain. Dev Dyn *240*, 1989-2001.

Shan, W., Tanaka, H., Phillips, G., Arndt, K., Yoshida, M., Colman, D., and Shapiro, L. (2000). Functional cis-heterodimers of N- and R-cadherins. J Cell Biol *148*, 579-590.

Shimoyama, Y., Takeda, H., Yoshihara, S., Kitajima, M., and Hirohashi, S. (1999). Biochemical characterization and functional analysis of two type II classic cadherins, cadherin-6 and -14, and comparison with E-cadherin. J Biol Chem *274*, 11987-11994.

Shimoyama, Y., Tsujimoto, G., Kitajima, M., and Natori, M. (2000). Identification of three human type-II classic cadherins and frequent heterophilic interactions between different subclasses of type-II classic cadherins. Biochem J *349*, 159-167.

Sotomayor, M., Gaudet, R., and Corey, D.P. (2014). Sorting out a promiscuous superfamily: towards cadherin connectomics. Trends Cell Biol *24*, 524-536.

Suzuki, S.C., Furue, H., Koga, K., Jiang, N., Nohmi, M., Shimazaki, Y., Katoh-Fukui, Y., Yokoyama, M., Yoshimura, M., and Takeichi, M. (2007). Cadherin-8 is required for the first relay synapses to receive functional inputs from primary sensory afferents for cold sensation. J Neurosci *27*, 3466-3476.

Takai, Y., Miyoshi, J., Ikeda, W., and Ogita, H. (2008). Nectins and nectin-like molecules: roles in contact inhibition of cell movement and proliferation. Nat Rev Mol Cell Biol *9*, 603-615.

Thu, C.A., Chen, W.V., Rubinstein, R., Chevee, M., Wolcott, H.N., Felsovalyi, K.O., Tapia, J.C., Shapiro, L., Honig, B., and Maniatis, T. (2014). Single-cell identity generated by combinatorial homophilic interactions between alpha, beta, and gamma protocadherins. Cell *158*, 1045-1059.

Treubert-Zimmermann, U., Heyers, D., and Redies, C. (2002). Targeting axons to specific fiber tracts in vivo by altering cadherin expression. J Neurosci *22*, 7617-7626.

Vendome, J., Felsovalyi, K., Song, H., Yang, Z., Jin, X., Brasch, J., Harrison, O.J., Ahlsen, G., Bahna, F., Kaczynska, A.*, et al.* (2014). Structural and energetic determinants of adhesive binding specificity in type I cadherins. Proc Natl Acad Sci U S A *111*, E4175-4184.

Williams, M.E., Wilke, S.A., Daggett, A., Davis, E., Otto, S., Ravi, D., Ripley, B., Bushong, E.A., Ellisman, M.H., Klein, G.*, et al.* (2011). Cadherin-9 regulates synapse-specific differentiation in the developing hippocampus. Neuron *71*, 640-655.

Figure Legends

Figure 1. SPR analysis of heterophilic interactions of type II cadherins.

Profiles of type II cadherin analytes (shown in columns) binding over individual surfaces of cadherin-8 (top row), cadherin-11, cadherin-6, cadherin-9, cadherin-10, cadherin-12, cadherin-20, cadherin-18, cadherin-22 and cadherin-7 (bottom row). Analytes were tested at 12, 6 and 3 μM monomer concentrations over each surface as shown in each panel. Responses were normalized to account for the molecular weight variations of the different cadherin analytes. The normalized responses in each row (corresponding to the responses over a surface) are scaled independently allowing quantitative comparison across rows only. Specificity groups identified based on binding preferences (see text) are boxed in blue for the cadherin-8 and cadherin-11 specificity group, in red for the cadherin-6, cadherin-9 and cadherin-10 specificity group and in green for the cadherin-12, cadherin-18, cadherin-20, cadherin-22 and cadherin-7 specificity group. Cys-tagged cadherins were immobilized at a free monomer concentration of 60 μM, corresponding to 4,673 RU for cadherin-6, 945 RU for cadherin-7, 1,006 RU for cadherin-8, 2,174 RU for cadherin-9, 546 RU for cadherin-10, 990 RU for cadherin-11, 3,784 RU for cadherin-12, 1,112 RU for cadherin-18, 1,283 for cadherin-20 and 3,651 RU for cadherin-22. See also Figure S1 and Table S2.

Figure 2. Effects of binding interface mutations on homophilic and heterophilic binding.

(A) SPR binding responses of wild-type cadherin-6 and its respective strand swap mutant (W4A), x-interface mutant (M188D) and a double mutant containing both mutations injected over wild-type cadherin-6 (top row), cadherin-9 (middle row), and

cadherin-10 surfaces (bottom row). **(B)** Wild-type cadherin-8 and strand-swap mutant cadherin-8 W4A and **(C)** cadherin-11 and strand-swap mutant cadherin-11 W4A were injected over wild-type cadherin-8 (top row) and cadherin-11 surfaces (bottom row). Note that responses for each surface are scaled independently.

Figure 3. Biophysically identified heterophilic binding specificity groups correspond to branches of the phylogenetic tree.

(A) Force-directed binding network of type II cadherin heterophilic interactions weighted by binding responses derived from SPR experiments (see methods). Nodes represent individual cadherins colored by specificity group; edges represent heterophilic binding interactions with length inversely proportional to binding strength. **(B)** Phylogram of the type II cadherin family computed from alignment of amino acid sequences of adhesive EC1-2 domain regions using a maximum likelihood method. Branches of the phylogenetic tree are colored according to specificity group. Symbols indicate cadherins for which structures of the adhesive interface are reported in this work (*), or in Patel et. al. 2006 ($^{+}$) and Brasch et al. 2011 (†).

Figure 4. Crystal structures of type II cadherin homodimers and analysis of specificity determinants.

(**A**) Ribbon representation of strand-swapped EC1-2 homodimer structures of cadherin-6, -10, -7, -22 and chimera cad-$11_{EC1}6_{EC2}$. Three Ca^{2+} ions (green spheres) are coordinated by interdomain linker regions of each protomer. (**B**) Superposition of EC1 homodimers reported here and in (Patel et al. 2006) shown as carbon-α traces superposed over left protomer. Docked strand-swap residues Trp2 and Trp4 are shown in stick representation. (**C**) Sequence-logos of aligned EC1 regions of type II cadherins

from human, mouse and chicken (see Methods) separated into specificity groups. Positions containing, or flanked by, interface residues (marked by magenta bars above alignment) are shown in bold and colored according to conservation. Cyan: Residues fully conserved within but different between specificity groups. Green: Residues differing between specificity groups, but not fully conserved within each. Grey: Positions with identical conserved consensus residues across all specificity groups. Secondary structure elements of cadherin-6 shown above logos. Pocket residues are indicated with 'P'. (**D**) Surface representation of EC1 with interface residues colored according to (C). Partner protomer in ribbon representation (salmon). (**E**) Interface residues colored according to (C) shown as sticks in a superposition of EC1 domain structures. Main chain ribbon shown only for cadherin-6. (**F**) Potential specificity determinants (magenta) are shown as sticks on superposed EC1 domain homodimers for members of each specificity group. Trp2 and Trp4 shown for orientation. See also Table S3.

Figure 5: Mutational analysis of type II cadherin specificity using SPR.

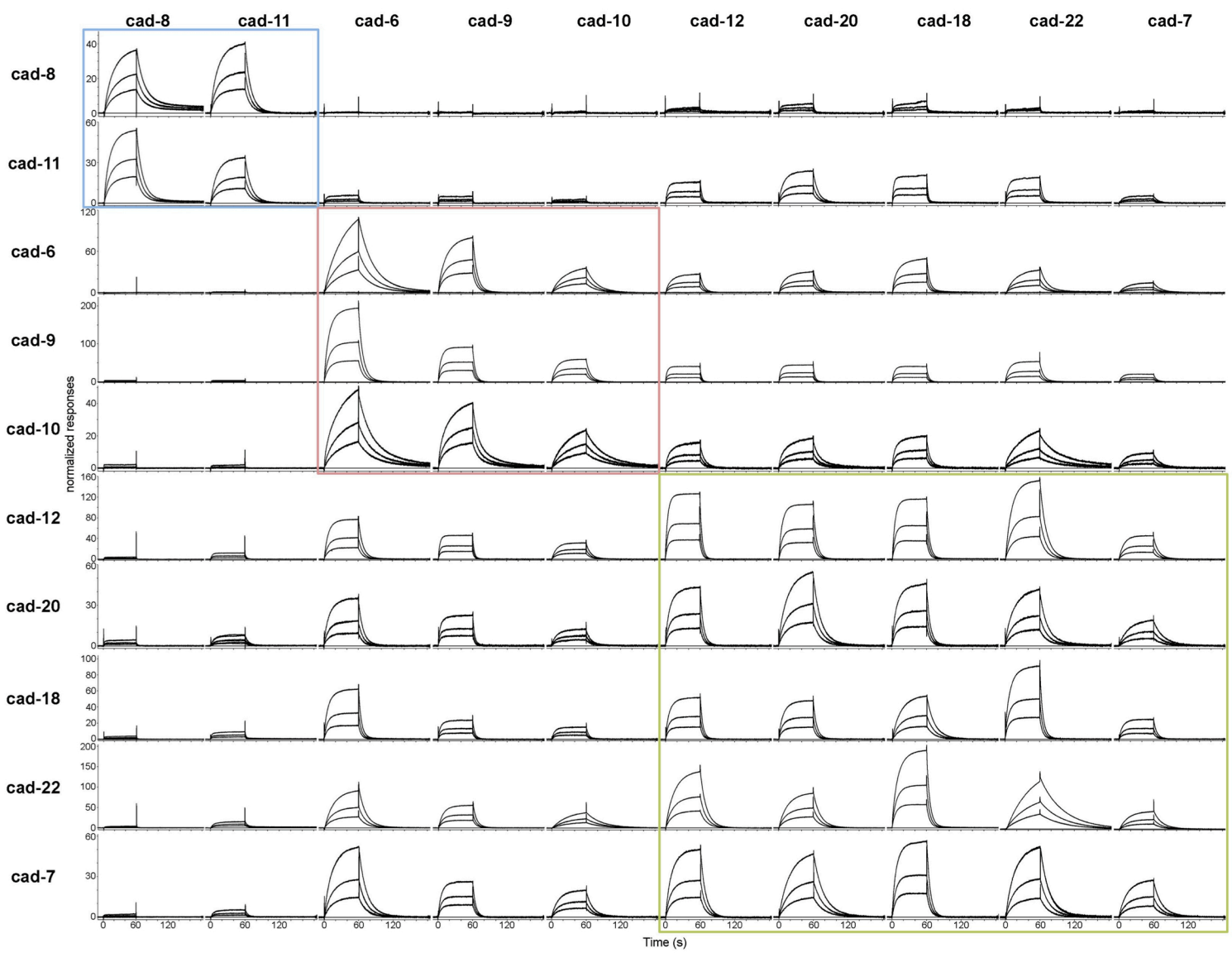
(**A**) Wild-type or specificity mutants of cad-6 or cad-11 (columns) passed over wild type cad-6, -9, -10 and cad-8, -11 surfaces (rows). (**B**) Chimeric proteins cad-$6_{EC1}11_{EC2}$ and cad-$11_{EC1}6_{EC2}$ passed over the same surfaces reproduce the binding characteristics of the wild-type proteins with corresponding EC1 domains. All mutant and chimeric proteins retained homophilic binding in AUC (Table 1).

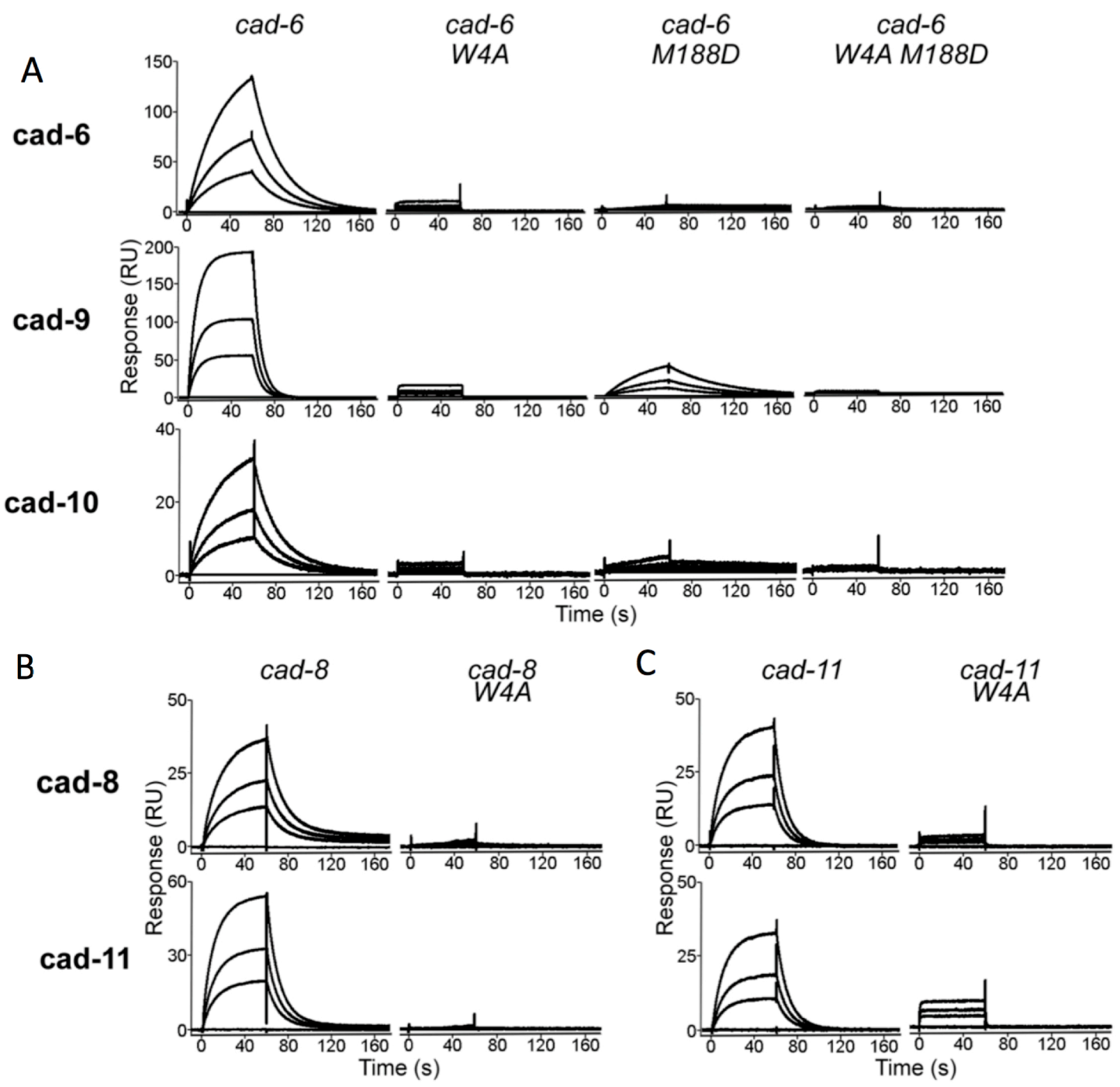
Figure 6. Full-length type II cadherin localization at homotypic and heterotypic contact sites between transfected A431D cells in co-culture.

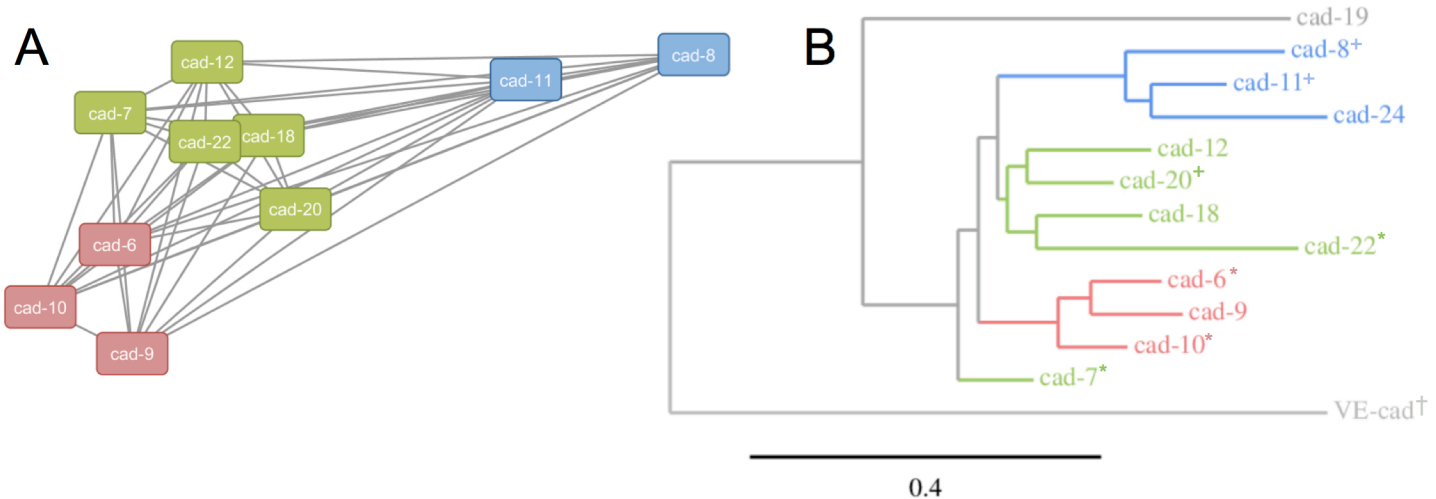
Flourescence images of co-cultures of A431D cells transfected with full length type II cadherins tagged with either cherry (red) or dendra (green) in the combinations indicated. (**A**)-(**E**) Wild-type cadherins localize to homotypic and heterotypic contact sites

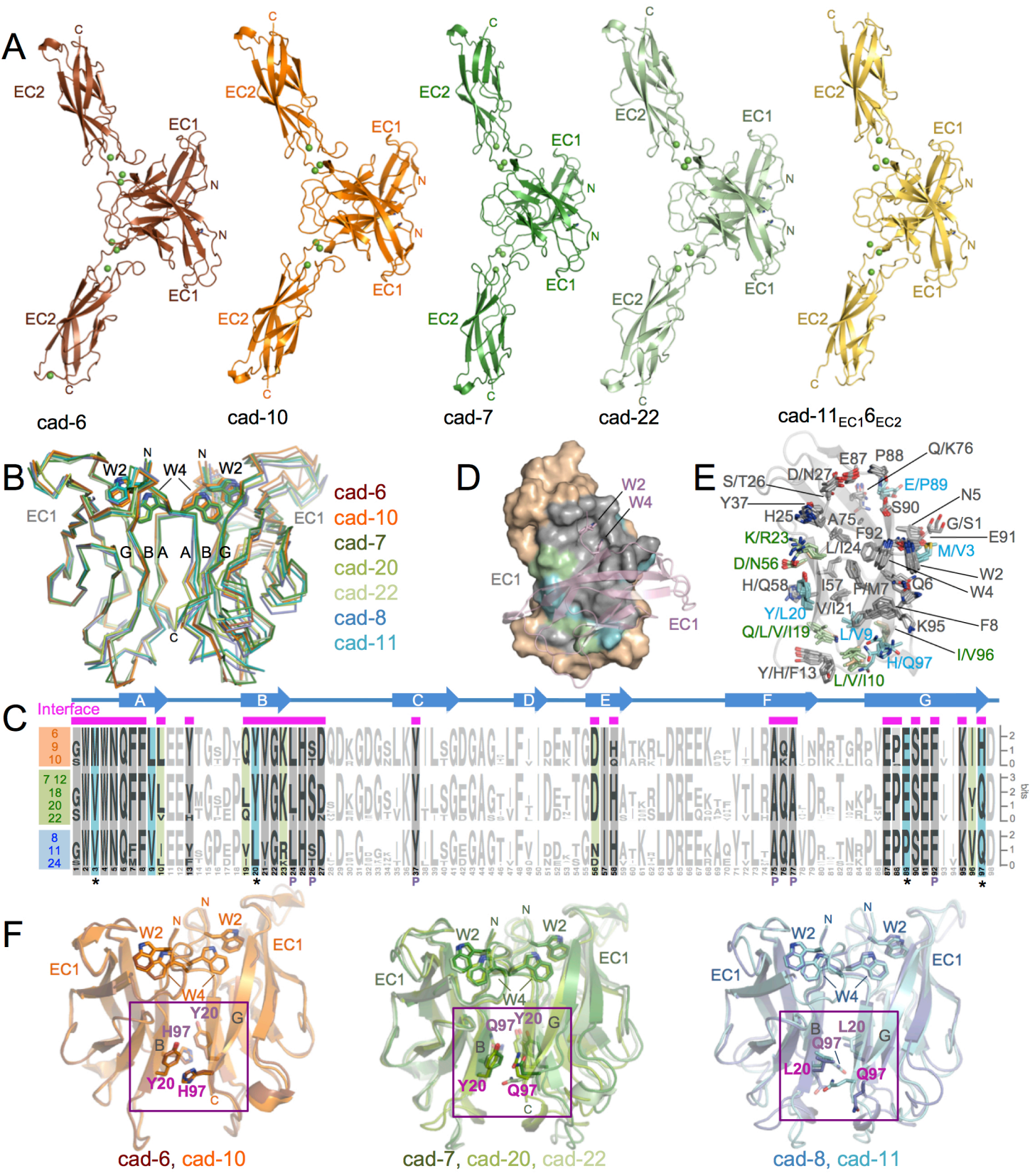
according to their binding preferences. (**F**)-(**J**) Mutations targeting specificity determinants in cad-6 and -11 alter their localization (compare panels linked with arrows). Heterotypic contact sites are delimited by arrowheads at top and bottom. Scale bar 50 μm. See also Figure S2.

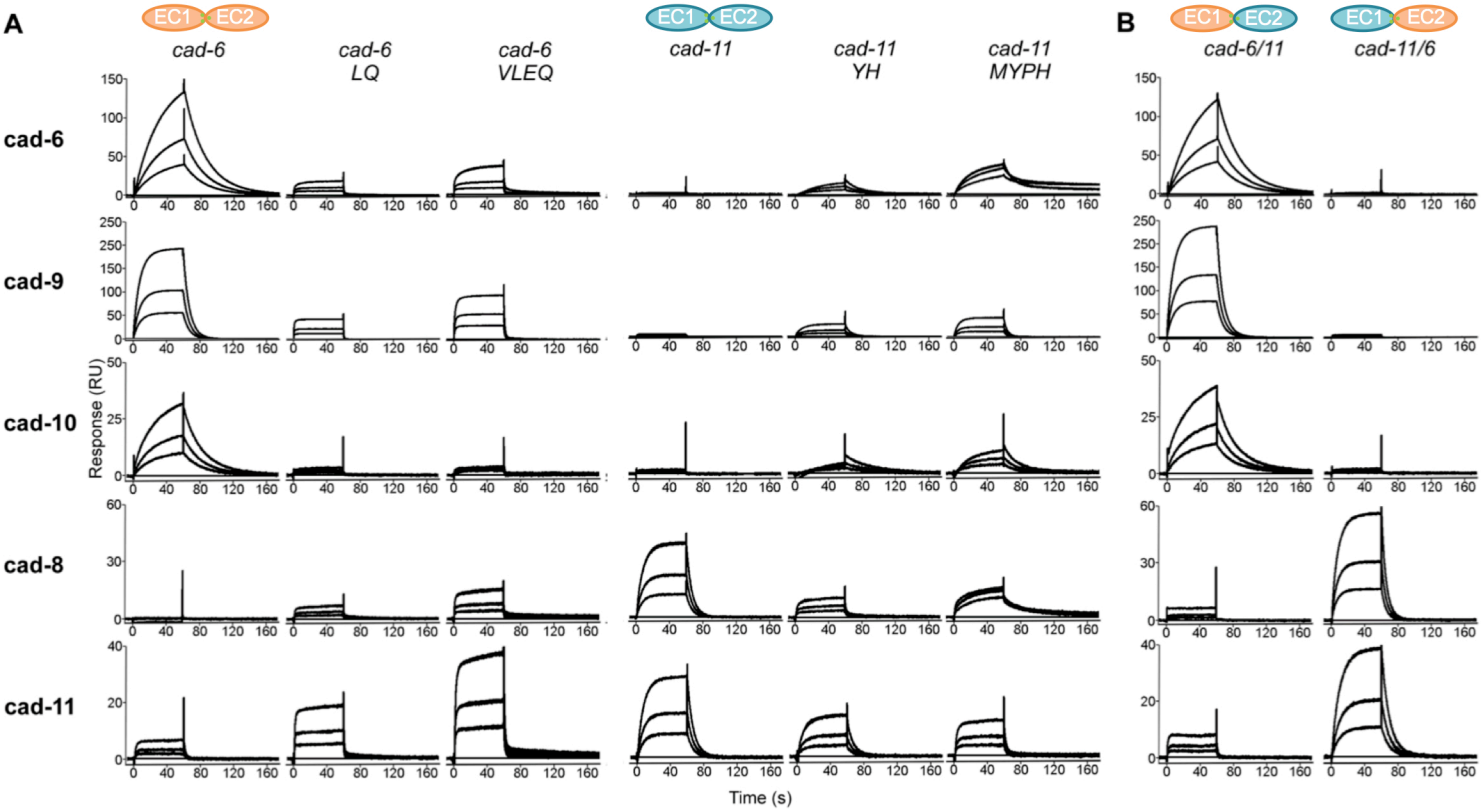
Table 1. Dissociation constants (K_D) for homodimerization of type II cadherin EC1-2 wild-type, chimera and mutant protein fragments determined by analytical ultracentrifugation. See also Table S2.


cad-8
cad-11
cad-6
cad-9
cad-10
cad-12
cad-20
cad-18
cad-22
cad-7
normalized responses
Time (s)


A
cad-6
cad-6 W4A
cad-6 M188D
cad-6 W4A M188D
cad-6
cad-9
cad-10
Response (RU)
Time (s)
B
cad-8
cad-8 W4A
cad-8
cad-11
Response (RU)
Time (s)
C
cad-11
cad-11 W4A
Response (RU)
Time (s)


A
cad-12
cad-8
cad-11
cad-7
cad-22
cad-18
cad-20
cad-6
cad-10
cad-9
B
cad-19
cad-8+
cad-11+
cad-24
cad-12
cad-20+
cad-18
cad-22*
cad-6*
cad-9
cad-10*
cad-7*
VE-cad†
0.4


A
EC2
EC1
N
C
cad-6
cad-10
cad-7
cad-22
cad-11EC16EC2
B
W2
W4
G B A A B G
cad-6
cad-10
cad-7
cad-20
cad-22
cad-8
cad-11
D
E
E87
P88
Q/K76
S/T26
D/N27
E/P89
N5
Y37
S90
H25
A75
G/S1
K/R23
F92
E91
L/I24
M/V3
D/N56
W2
H/Q58
I57
F/M7
Q6
W4
Y/L20
V/I21
F8
Q/L/V/I19
L/V9
K95
I/V96
H/Q97
Y/H/F13
L/V/I10
C
Interface
A
B
C
D
E
F
G
bits
F
Y20
H97
Q97
L20
cad-6, cad-10
cad-7, cad-20, cad-22
cad-8, cad-11


A
B
EC1
EC2
cad-6
cad-6 LQ
cad-6 VLEQ
cad-11
cad-11 YH
cad-11 MYPH
cad-6/11
cad-11/6
cad-6
cad-9
cad-10
cad-8
cad-11
Response (RU)
Time (s)

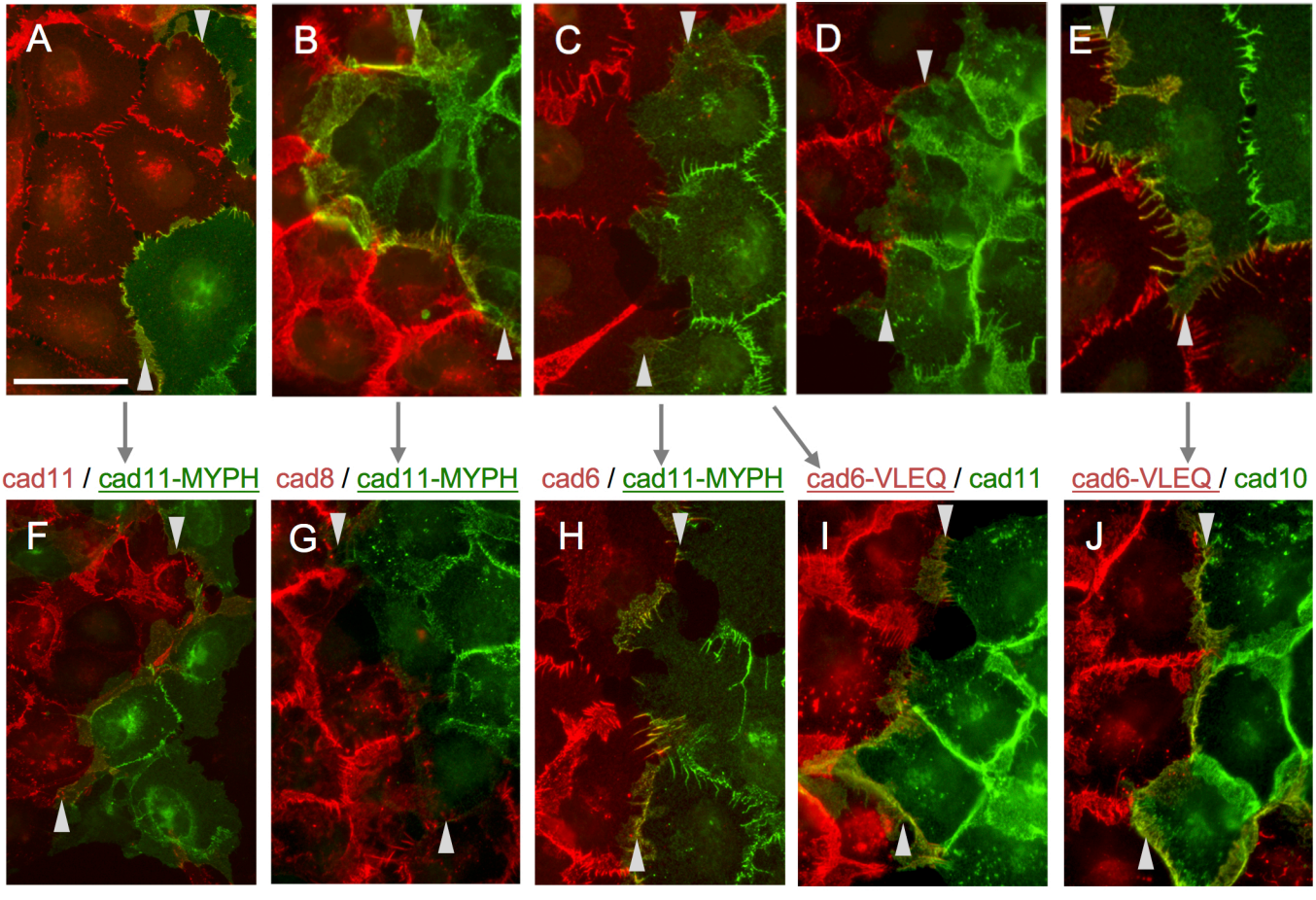

cad11 / cad11
cad8 / cad11
cad6 / cad11
cad8 / cad10
cad6 / cad10
A
B
C
D
E
cad11 / cad11-MYPH
cad8 / cad11-MYPH
cad6 / cad11-MYPH
cad6-VLEQ / cad11
cad6-VLEQ / cad10
F
G
H
I
J

Table 1: Dissociation constants (K_D) for homodimerization of type II cadherin EC1-2 wild-type, chimera and mutant protein fragments determined by analytical ultracentrifugation. See also Table S2.

Cadherin	K_D[μM]	Description
Cadherin-6	3.1 ±0.1[a,b]	Wild-type
Cadherin-9	17.0 ±1.1[c]	Wild-type
Cadherin-10	42.2 ±2.7[c]	Wild-type
Cadherin-8	15.0 ±0.4[c]	Wild-type
Cadherin-11	33.8 ±0.2[c]	Wild-type
Cadherin-24	8.2 ±0.3	Wild-type
Cadherin-7	32.2 ±0.8	Wild-type
Cadherin-12	8.3 ±1.6	Wild-type
Cadherin-18	16.8 ±0.2	Wild-type
Cadherin-20	9.3 ±0.6	Wild-type
Cadherin-22	3.9 ±0.2	Wild-type
Cad-6/11	5.6±0.2	Chimera cad$6_{EC1}11_{EC2}$
Cad-11/6	15.6±0.9	Chimera cad$11_{EC1}6_{EC2}$
Cad-6 Y20L H97Q (LQ)	9.63±1.3	Specificity mutant
Cad-6 M3V Y20L H97Q E89P (VLQP)	6.73±0.8	Specificity mutant

Cad-11 L20Y Q97H (YH)	11.1 ±2.1	Specificity mutant
Cad-11 V3M L20Y Q97H P89E (MYHE)	8.2±0.8	Specificity mutant

[a] Errors standard diviation from two or more experiments,
[b] previously reported in Harrison et al. (2010)
[c] previously reported in Brasch et al. (2011)