Cell Reports, Volume *23*

Supplemental Information

Homophilic and Heterophilic Interactions of Type II Cadherins Identify Specificity Groups Underlying Cell-Adhesive Behavior

Julia Brasch, Phinikoula S. Katsamba, Oliver J. Harrison, Göran Ahlsén, Regina B. Troyanovsky, Indrajyoti Indra, Anna Kaczynska, Benjamin Kaeser, Sergey Troyanovsky, Barry Honig, and Lawrence Shapiro

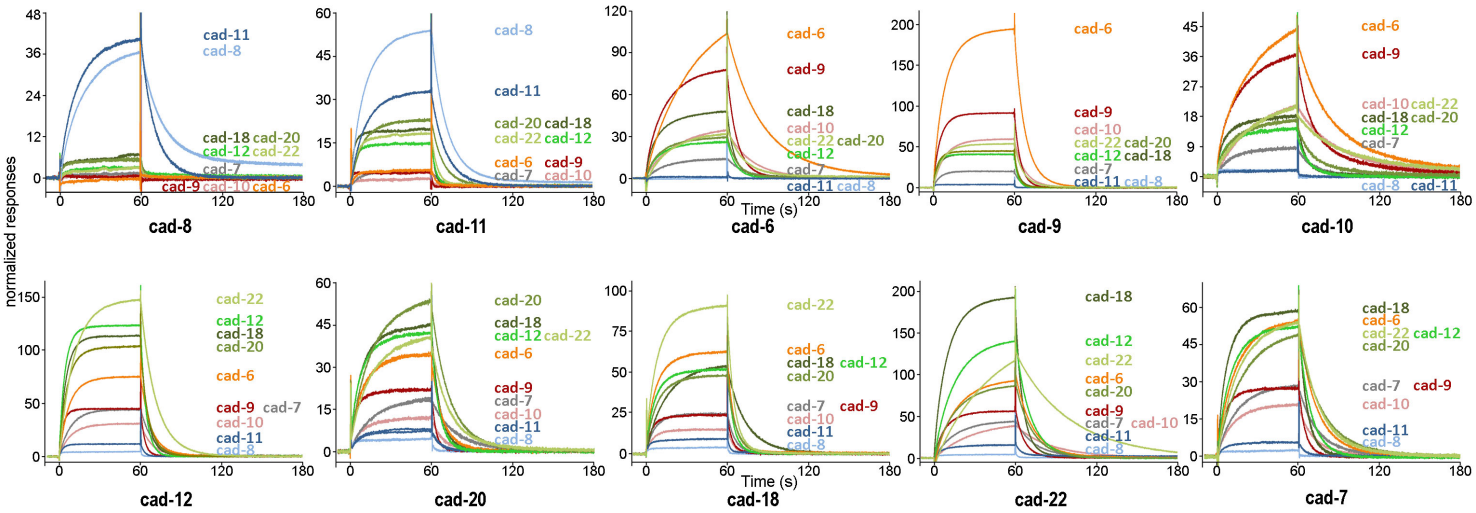
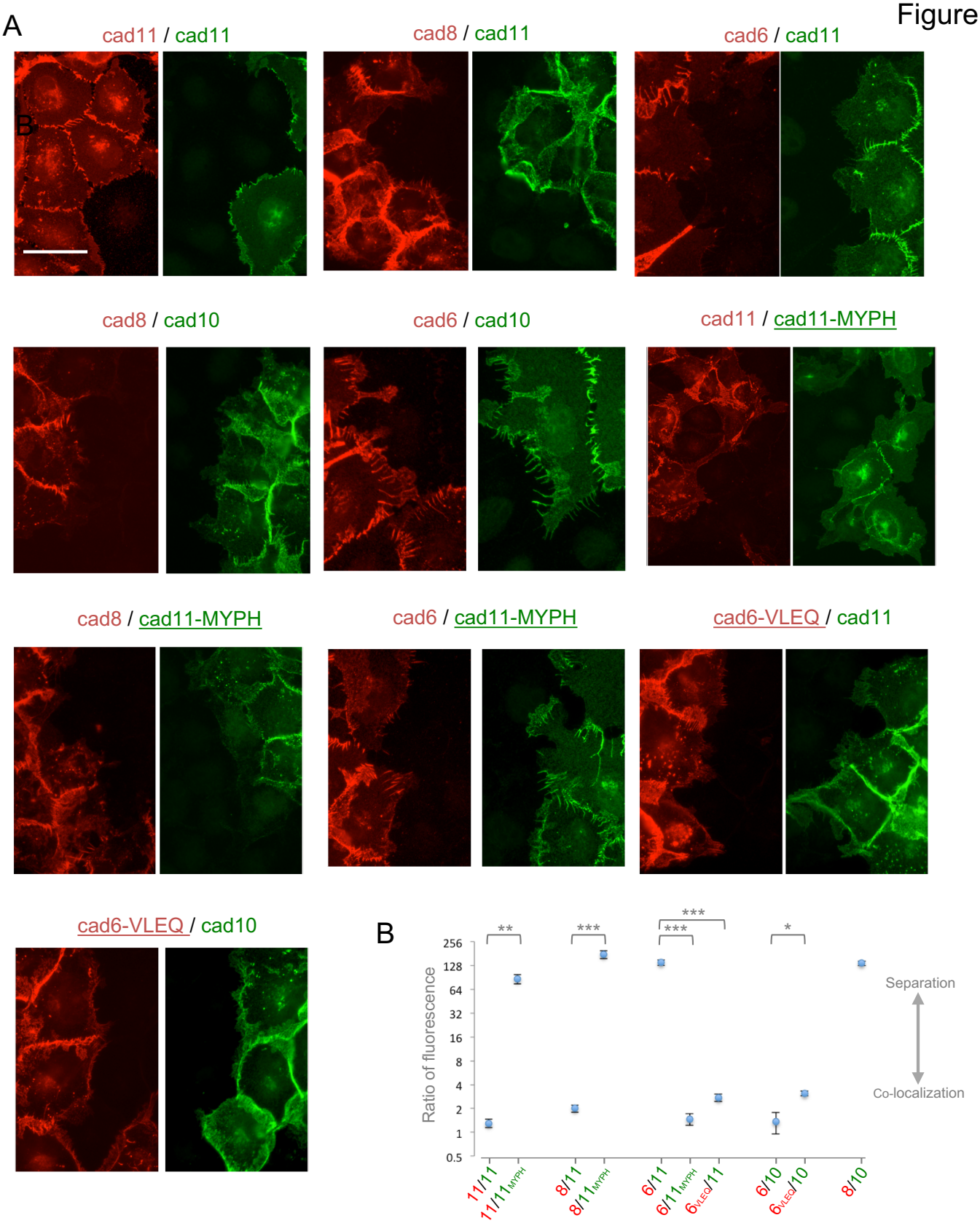


Figure S1. SPR analysis of heterophilic binding interactions of type II cadherins. Overlaid binding responses of each of the ten type II cadherin analytes shown in Fig.1 at 12μM over surfaces of cadherin-8 (top left), cadherin-11, cadherin-6, cadherin-9, cadherin-10, cadherin-12, cadherin-20, cadherin-18, cadherin-22 and cadherin-7 (bottom right).



Supplementary Figure 2. Quantitation of full-length type II cadherin localization at homotypic and heterotypic contact sites between transfected A431D cells in co-culture. (A) Fluorescence images showing separate red and green channels for co-cultures of transfected A431D cells displayed as combined images in Figure 6. Scale bar 50μm. (B) Mean ratios of red fluorescence at homotypic versus heterotypic contact sites calculated from three measurements for each co-culture. Error bars: standard error. *=p<0.05, **=p<0.01, ***=p<0.001 from unpaired Student's t-test.

Table S1: Dissociation constants (K_D) for homodimerization of type II cadherin CYS-tagged wild-type and mutant protein fragments determined by analytical ultracentrifugation. Compare to wild-type measurements presented in Table 1.

Cadherin	K_D[μM]	Description
Cadherin-6 CYS-tag	3.6 ±0.5[a]	Wild-type
Cadherin-7 CYS-tag	19.6 ±0.6	Wild-type
Cadherin-8 CYS-tag	19.6 ±2.4	Wild-type
Cadherin-9 CYS-tag	7.8 ±1.4	Wild-type
Cadherin-10 CYS-tag	40.2 ±5.7	Wild-type
Cadherin-11 CYS-tag	19.2 ±4.6	Wild-type
Cadherin-12 CYS-tag	4.2 ±1.8	Wild-type
Cadherin-18 CYS-tag	16.7 ±1.9	Wild-type
Cadherin-20 CYS-tag	14.4 ±1.3	Wild-type
Cadherin-22 CYS-tag	5.0 ±1.0	Wild-type
Cadherin-6 W4A	321±0.5[b]	Strand-swap mutant
Cadherin-6 M188D	12.6±0.5[b]	X-dimer mutant
Cadherin-6 W4A + M188D	Monomer[b]	Double interface mutant
Cadherin-8 W4A	Monomer	Strand-swap mutant

[a] Errors indicate data range from two or more experiments.
[b] previously published in Harrison et. al. (2010)

Table S2: Dissociation rates (k_d) for homodimerization of type II cadherins derived from homophilic binding responses shown in Figure 1 and Figure S1.

protein	k_d (s^{-1})
Cadherin-6	0.0415
Cadherin-9	0.215
Cadherin-10	0.0444
Cadherin-8	0.116
Cadherin-11	0.106
Cadherin-7	0.0657
Cadherin-12	0.235
Cadherin-18	0.0620
Cadherin-20	0.0675
Cadherin-22	0.0236

Table S3: Crystallographic data collection and refinement statistics for crystal structures presented and analyzed in Figure 4.

	Cadherin-6 EC1-2 mouse	**Cadherin-7 EC1-2** mouse	**Cadherin-10 EC1-2** mouse	**Cadherin-22 EC1-2** mouse	**Cadherin-11$_{EC1}$/6$_{EC2}$ chimera** mouse
Data collection					
Space group	$C222_1$	$P2_12_12_1$	$P4_12_12$	$P2_1$	$I2_12_12_1$
Cell dimensions					
a, *b*, *c* (Å)	114.35, 141.65, 142.20	58.304, 82.431, 93.546	87.38, 87.38, 67.68	50.26, 45.077, 128.054	53.62, 81.08, 166.29
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 92.253, 90	90, 90, 90
Resolution (Å)	40-1.90 (1.97-1.90)	40-1.7 (1.78-1.7)	45-2.70 (2.83-2.70)	30-2.70 (2.80-2.70)	30-3.00 (3.18-3.00)
R_{sym} or R_{merge}	0.066 (0.553)	0.07 (0.656)	0.123 (1.041)	0.046 (0.496)	0.124 (1.766)
$I/\sigma I$	23.8 (2.6)	27.33 (2.9)	11.7 (2.1)	22.7 (1.8)	14.8 (1.9)
Completeness (%)	100 (99.9)	99.8 (99.6)	99.9 (100.0)	98.3 (94.0)	99.9 (99.9)
Redundancy	5.2 (4.7)	7.4 (7.1)	8.5 (8.6)	4.1 (3.4)	7.7 (7.8)
Refinement					
Resolution (Å)	20-1.90	20-1.7	20-2.70	20-2.70	20-3.00
No. reflections	90099	48271	7545	15707	7540
R_{work}/ R_{free}	0.1611, 0.2009	0.1598, 0.1991	0.2358, 0.2853	0.2261, 0.2772	0.2703, 0.3005
No. atoms					
Protein	6617	3383	1632	3211	1574
Ligand/ion	14	18	18	6	19
Water	1231	680	26	11	2
B-factors					
Protein	41.25	23.91	73.83	84.72	121.81
Ligand/ion	34.57	34.73	66.32	59.02	115.36
Water	47.69	33.28	52.39	56.94	112.45
R.m.s deviations					
Bond lengths (Å)	0.003	0.009	0.002	0.003	0.004
Bond angles (°)	0.652	1.300	0.487	0.525	0.589
Ramachandran					
Favored (%)	96.7	98.9	96.6	97.3	91
Allowed (%)	3.3	1.1	3.4	2.7	9
Outliers (%)	0	0	0	0	0
PDB Accession Code	6CGU	6CGS	6CG6	6CG7	6CGB

*Highest resolution shell is shown in parenthesis.

Supplemental Experimental Procedures

Protein production in bacteria

Coding sequences of EC1-2 fragments of mouse cadherin-6, -7, -8, -9, -10, -11, -12, -18, -20, -22 and EC1-3 fragments of cadherin-8 and -24 encompassing residues 1-207 (EC1-2) or 1-322 (EC1-3) of the mature proteins were amplified by PCR from cDNA libraries (Clontech). Sequences were cloned in frame with an N-terminal hexa histidine-tagged SUMO protein into the *BamHI/NotI* sites of the vector pSMT3. Cleavage of SUMO-fusion-proteins with Ulp1 (Ubiquitin-like protease 1) after a Gly-Gly motif yields cadherin proteins with native N-termini. Extra amino acids occurring due to cloning after the cleavage site were removed by using the QuikChange site directed mutagenesis kit (Stratagene) to ensure native N-termini of all proteins used in our studies. We introduced all point mutations in cadherin-6, -8 (EC1-2) and -11 (EC1-2) using the QuikChange method.

For protein expression, E. coli BL21 DE3 pLysS (NEB) were transformed with each construct and grown at 37°C shaking at 200 rpm until OD_{600} reached 0.6. To induce protein expression, we added 100 µM IPTG and lowered the temperature to 16°C. After 18h bacteria were harvested by centrifugation at 4,000xg for 15 min. Pelleted bacteria were resuspended in lysis buffer (500 mM NaCl, 20 mM Tris-Cl pH 8.0, 20 mM Imidazole pH 8.0, 3 mM $CaCl_2$) and lysed for 6 minutes by sonication in 15 second intervals with 45 seconds rest in between pulsing. Cell debris was pelleted at 4°C and 20,000xg for 1 hour and His-tagged proteins were extracted from cleared lysate by nickel affinity chromatography using 5 mL nickel charged IMAC Sepharose 6 Fast Flow resin (GE Healthcare). Beads were subsequently washed with 40 column volumes of lysis buffer to remove contaminants and His-SUMO-fusion proteins were eluted with lysis buffer containing 250 mM imidazole. The His-SUMO tag of fusion proteins was

cleaved enzymatically by adding Ulp1 enzyme to a final concentration of 2 μg/mL to the elution. Proteins were then dialyzed into a low ionic strength buffer (75 mM NaCl, 20 mM Tris-Cl pH 8.0 and 3 mM $CaCl_2$). We removed cleaved His-SUMO tags and uncut fusion protein on nickel charged IMAC resin, equilibrated in dialysis buffer. The cadherins were further purified by anion exchange chromatography (Mono Q 10/10 HR, GE Healthcare) and size exclusion chromatography (HiLoad 26/60 Superdex 75 for EC1-2 fragments, Superdex 200 for EC1-3 fragments (GE Healthcare)) leaving them in a final buffer of 150 mM NaCl, 20 mM Tris-Cl pH 8.0 and 3 mM $CaCl_2$. Proteins were concentrated to a final concentration of approximately 5-10 mg/mL using Amicon Spin concentrators (Millipore) and flash frozen. Production of mouse cadherins-6 EC1-2, -8 EC1-3, -9 EC1-2, -10 EC1-2, and -11 EC1-2 and cadherin-6 EC1-2 mutants W4A, M188D and W4A M188D was also described previously (Brasch et al., 2011; Harrison et al., 2010).

Analytical Ultracentrifugation

Equilibrium analytical ultracentrifugation experiments were performed using a Beckman XLA/I ultracentrifuge, with a Ti50An or Ti60An rotor. Prior to each experiment, all proteins were diluted with buffer (150 mM NaCl, 10 mM Tris-Cl pH 8.0, 3 mM $CaCl_2$) and dialyzed for 16 hours at 4°C in the same buffer. 120 μL of proteins at three different concentrations 0.65 mg/mL, 0.43 mg/mL and 0.23 mg/mL were loaded into six-channel equilibrium cells with parallel sides and sapphire windows. We performed all experiments at 25°C and collected UV data at 280 nm, using dialysis buffer as blank. Three-domain proteins were spun for 20 hours at 8740xg and four scans (1 per hour) were collected, speed was increased to 14160xg for 10 hours and four scans (1 per hour) were collected, speed was increased to 20880xg for 10 hours and four scans (1 per hour) were taken, and finally speed was increased to 28910xg for 10 hours and four scans (1 per hour) were collected. This yielded 48 scans per sample. Two-domain

proteins were analyzed using the same protocol, except that 16260xg, 26090xg, 38230xg and 52680xg were used, respectively. RCF's are given at the measuring cell center at a radius of 65 mm. We calculated the buffer density and protein v-bars using the program SednTerp (Alliance Protein Laboratories), and analyzed the retrieved data using HeteroAnalysis 1.1.44 (http://www.biotech.uconn.edu/auf). We fitted data from all concentrations and speeds globally by nonlinear regression to either a monomer-dimer equilibrium model or an ideal monomer model. All experiments were performed at least in duplicate.

SPR binding assays

Binding assays were performed using a Biacore T100 biosensor equipped with a Series S CM4 chip sensor chip (GE Healthcare). Type II cadherins were covalently immobilized over individual chip surfaces using ligand thiol-coupling chemistry of a C-terminal cysteine in HBS pH 7.4 (10 mM HEPES, 150 mM NaCl), 3 mM $CaCl_2$, at 32°C using a flow rate of 20 µL/min. During the immobilization reaction, the carboxyl groups on the sensor chip surface were activated for 2 minutes using 400 mM EDC (N-ethyl-N_-(3-dimethylaminopropyl)carbodiimide), mixed at 1:1 ratio (v/v) with 100 mM NHS (N-hydroxysuccinimide). Subsequently, a solution of 120 mM PDEA, was mixed with 0.1 M sodium borate pH 8.5 at 2:1 ratio (v/v), to yield a final concentration of 80 mM PDEA and injected over the same flow cell for 4 minutes. The cadherin to be immobilized was freshly desalted in 10 mM sodium acetate, pH 4.0 using the Zeba spin desalting columns (Thermo Scientific) and sequentially injected over the activated surface at 5-25 µg/mL until the desired immobilization level was achieved. Any remaining disulfides were blocked using a four-minute injection of 50 mM L-cysteine/1.0 M NaCl in 0.1M sodium acetate, pH 4.0. Each cys-tagged cadherin ligand was tethered over the dextran layer at the following densities: 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. An unmodified surface was used as a reference flow cell to subtract bulk refractive index changes.

Cadherin binding was tested at 25°C in a running buffer of 10 mM Tris-HCl, pH 8.0, 150 mM NaCl, 3mM $CaCl_2$, 0.25 mg/mL BSA and 0.005% (v/v) Tween 20. Cadherin analytes (injected over the immobilized cadherin surfaces) were diluted in running buffer at 3, 6 and 12 μM monomer concentration, which were calculated using the homophilic K_D values listed in Table 1. Analytes for each concentration series were injected in order of increasing concentration at 50 μL/min for 60 s followed by a 120 s dissociation phase and a 60 s buffer wash at the end of the binding cycle. Each series was tested in duplicate to verify the reproducibility of the assay. Following three sequential cadherin binding cycles a buffer injection replaced the analyte to double reference the responses thus removing systematic noise and instrument drift. The binding responses were processed using Scrubber 2.0 (BioLogic Software). All signals were normalized to account for molecular weight differences between the three-EC domain cadherin-8 and the remaining two-EC domain cadherins. Responses for each type II cadherin analyte were divided by its own molecular weight and multiplied by a constant (23,000).

Binding network

Heterophilic binding data obtained for each SPR surface was normalized to the highest recorded binding response on the surface at equilibrium such that the highest binding interaction received a score of one. Each interaction was recorded twice: once between cadherin ‘A’ on the surface and cadherin ‘B’ as analyte and second between cadherin ‘B’ on the surface and cadherin ‘A’ as analyte; both binding scores for each interaction were

summed. The combined scores were then used to weight a force-directed network using the program Cytoscape (Shannon et al., 2003).

Phylogenetic Tree

The amino acid sequences encompassing EC1-2 of all mouse type II cadherins were used to produce a multiple sequence alignment with the program Muscle. A phylogenetic tree was generated from this alignment using the maximum likelihood method with the program PhML, and the tree was rendered using the program TreeDyn via the phylogeny.fr server (Dereeper et al., 2008).

Crystallization and Structure Determination

Crystals of EC1-2 adhesive fragments of mouse cadherin-6, -7, -10, -22 and cadherin-$11_{EC1}6_{EC2}$ chimera were grown using the vapor diffusion method with 1-2 µL hanging drops at 4ºC (cadherin-7) or 22ºC (all others). Crystallization conditions were: 18.5 % (w/v) polyethylene glycol (PEG) 3,350, 0.2 M sodium acetate, 30 % (v/v) ethylene glycol cryo-protectant for cadherin-6 EC12; 21.5 % (w/v) PEG 6,000, 0.1 M MES pH 5.6, 1 M lithium chloride, 30 % (v/v) glycerole cryo-protectant for cadherin-7; 6 % (v/v) 2-propanol, 0.1 M MES pH 6.0, 0.2 M calcium acetate, 30 % (v/v) ethylene glycol cryo-protectant for cadherin-10; 17 % (v/v) PEG6000, 0.1 M HEPES pH 7.0, 5% ethylene glycol, 30 % (v/v) ethylene-glycol cryo-protectant for cadherin-22; and 4M sodium chloride, 0.1 M sodium acetate pH 5.5, 30 % (v/v) glycerol cryo-protectant for cadherin-$11_{EC1}6_{EC2}$ chimera. Cryo-protection was performed by brief immersion of the crystal prior to flash freezing in well solution, supplemented with the indicated cryo-protectant.

Data were collected from single frozen crystals at 100K using a wavelength of 0.979 Å at the X4A and X4C beamlines of the National Synchrotron Light Source, Brookhaven

National Laboratory. Data were processed using XDS (Kabsch, 2010) and merged with the aimless program (Evans and Murshudov, 2013) of the ccp4-suite (Winn et al., 2011). Structures were solved by molecular replacement using Phaser (McCoy et al., 2007) within the Phenix suite (Adams et al., 2010). Mouse cadherin-6 EC1-2 W4A (3LND) was used as a search model for cadherin-6 EC1-2 and cadherin-6 was then used as a search model for all other structures. Separate search models were used to solve the cadherin-$11_{EC1}6_{EC2}$ chimera (EC1 from 2A4C, EC2 from cadherin-6 EC1-2 wt). Structure refinement was performed by manual model building in Coot (Emsley et al., 2010) with automated refinement using phenix.refine (Afonine et al., 2012). Non-crystallographic symmetry restraints were used for the first 5-8 rounds of refinement for structures containing more than one molecule per asymmetric unit (cadherin-6, -7, -22). Translation-libration-screw (TLS) parameters were refined for cadherin-6 and cadherin-$11_{EC1}6_{EC2}$ chimera with EC1 and EC2 domains defined as separate TLS groups in phenix.refine. Data collection and refinement statistics are summarized in Table S3.

Sequence Analysis

Amino acid sequences of human, mouse and chicken cadherin-6, -9, -10, -11, -12, -18, -20 and -22, and human and mouse cadherin-24 were obtained from uniprot and aligned using MultAlin (Corpet, 1988) over their mature EC1 domains (residues 1-98). Sequence Logo representations were produced separately for each specificity group from the sequence alignment using the WebLogo server (Crooks et al., 2004). To define interface regions in all type II cadherin crystal structures, residues with at least 5 % of their accessible surface area buried in each dimer were identified using PISA (Krissinel and Henrick, 2007). Residue positions buried in at least half of the available biological dimer structures were defined as interface in the sequence alignments.

Co-culture assays

Full-length sequences encoding mouse cadherins-6, -8, -10 and -11 were cloned in frame with C-terminal dendra2-Myc-or mCherry-Flag tags into the Geneticin-resistant mammalian expression vector pRc/CMV (Invitrogen). Transfection, growth, and immunofluorescence microscopy of transfected human A-431D cells were performed as described (Hong et al., 2010). After selection for Geneticin-resistance, cells were sorted for transgene expression by FACS, and only moderate-expressing cells were used in co-culture experiments. For quantitation, the co-cultures grown on coverslips were fixed and red-fluorescence signals of cells with both homotypic and heterotypic contacts were measured independently for each type of cell-cell contact using regions of 4 x 10 μm placed along the cell-cell contact line. The background fluorescence taken from the area adjacent to the cell-cell contact zone of the same cell was subtracted. Resulting values for homotypic contacts were divided by two, since red-fluorescent cadherins are equally contributed by both contacting cells. Ratios of homotypic over heterotypic red-fluorescence are displayed in Figure S2. Three measurements were performed for each co-culture and the most representative cells were selected.

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.

Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., and Adams, P.D. (2012). Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr D Biol Crystallogr *68*, 352-367.d

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.

Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res *16*, 10881-10890.

Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004). WebLogo: a sequence logo generator. Genome Res *14*, 1188-1190.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.F., Guindon, S., Lefort, V., Lescot, M.*, et al.* (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res *36*, W465-469.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr *66*, 486-501.

Evans, P.R., and Murshudov, G.N. (2013). How good are my data and what is the resolution? Acta Crystallogr D Biol Crystallogr *69*, 1204-1214.

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.

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.

Kabsch, W. (2010). Xds. Acta Crystallogr D Biol Crystallogr *66*, 125-132.

Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. J Mol Biol *372*, 774-797.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J Appl Crystallogr *40*, 658-674.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res *13*, 2498-2504.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A.*, et al.* (2011). Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr *67*, 235-242.