Supporting Information

Vendome et al. 10.1073/pnas.1416737111

SI Materials and Methods

Bacterial Protein Production. Coding sequences of mouse N-, E-, P-, R-cadherin extracellular cadherin (EC)1–EC2 (Asp1-Val216, Asp1-Asp213, Glu1-Asp213, and Asp1-Asp215, respectively) and *Xenopus* C-cadherin EC1–EC2 (Asp1-Asp213) were amplified by PCR from cDNA libraries (Clontech) and cloned in frame with an N-terminal hexahistidine-tagged small ubiquitin-like modifier (SUMO) protein into the BamHI/NotI sites of the pSMT3 vector [a modified pET-28b(+) expression plasmid]. Cleavage of SUMO-fusion proteins with Ulp1 (Ubiquitin-like protease 1) after a Gly-Gly motif yields cadherin proteins with native termini. Any extra amino acids occurring after the cleavage site because of cloning were removed by using the QuikChange site-directed mutagenesis kit (Stratagene, Agilent Technologies) to ensure native N termini of all proteins used in our studies, unless specifically altered. We introduced all point mutations using either the QuikChange mutagenesis kit or KOD Hot Start DNA Polymerase kit (Novagen, EMD Chemicals). For surface plasmon resonance (SPR) analysis, a Gly-Gly-Gly-Cys motif was introduced by PCR before the stop codon in wild-type P-, C-, N-cadherin, E-cadherin W2A, N-cadherin W2A, and N-cadherin R14E mutants, yielding a free C-terminal cysteine in these constructs for attachment to thiol-reactive SPR chips. Biotinylated wild-type E- and N-cadherin proteins were produced with a C-terminal Avi tag, as described previously (1), for immobilization on neutravidine-derivatized SPR chips. R-cadherin was fused with biotin carboxyl carrier protein after Asp215 for in vivo biotinylation and subsequent chip immobilization. All PCR primers were obtained from Life Technologies.

For protein production, *Escherichia coli* Rosetta 2DE3 pLysS (Novagen) or BL21-Gold (DE3) cells (Agilent) were transformed with prepared pSMT3 vector and grown at 37 °C until OD_{600} reached 0.6 shaking at 200 rpm. To induce protein expression, we added 100 μM isopropyl-β-D-thiogalactopyranoside (IPTG) and lowered the temperature to 18 °C. After 18 h cells were harvested by centrifugation at 4,000 × *g* for 15 min. Pelleted bacteria were resuspended in lysis buffer (500 mM NaCl, 10–20 mM Tris·HCl pH 8.0, 20 mM Imidazole pH8.0, 3 mM $CaCl_2$) and lysed for 3–6 min by sonication in 15-s intervals with a 45-s rest in between pulsing. Cell debris was spun down at 4 °C and 20,000 × *g* for 30 min to 1 h and His-tagged proteins were extracted from the clear lysate by flowing over nickel charged ion metal affinity chromatography (IMAC) Sepharose 6 Fast Flow resin (GE Healthcare). Beads were subsequently washed with 20- to 40-column volumes of lysis buffer to remove contaminants and His_6-SUMO fusion proteins were eluted with lysis buffer containing 250 mM imidazole. The His_6-SUMO tag of the fusion proteins were cut enzymatically by adding Ulp-1 to a final concentration of 2 μg/mL to the elution. Proteins were dialyzed into a low ionic strength buffer (100 mM NaCl, 10–20 mM Tris·HCl pH 8.0, and 3 mM $CaCl_2$). We removed Ulp1 (which has a His tag), cleaved His_6-SUMO tags, and any remaining uncut fusion protein by batch binding to nickel charged IMAC resin equilibrated in dialysis buffer. The cadherins were further purified by anion exchange chromatography (Mono Q 10/10 HR; GE Healthcare) using a NaCl gradient, and size-exclusion chromatography (HiLoad 26/60 SuperdexTM S75 prep grade; GE Healthcare) in a final buffer of 150 mM NaCl, 10 mM Tris·HCl pH 8.0 and 3 mM $CaCl_2$. Proteins were concentrated to a final concentration of ∼1–10 mg/mL using Amicon Spin concentrators (Millipore) and flash-frozen. Cysteine-containing proteins were purified using the same method, except all buffers were supplemented with 1 mM TCEP to keep cysteines reduced.

Mammalian Protein Production. Mouse M-cadherin EC1-5 (Ala1-Ala547) was expressed in HEK293 cell lines as secreted proteins with a PTPα signal sequence and C-terminal hexa-histidine tag. Conditioned media were harvested 2 d after transient transfection using polyethylenimine buffered to pH 7.0 with 1 M NaOH, and brought to high salt concentration (500 mM NaCl) before purification. The protein was purified using IMAC Sepharose 6 Fast Flow resin charged with nickel and dialyzed against 150 mM NaCl, 10 mM Tris·HCl pH 8.0, and 3 mM $CaCl_2$, followed by two chromatographic steps: ion exchange (MonoQ 10/100GL) and size exclusion (S200 26/60) in a final buffer of 150 mM NaCl, 10 mM Tris pH 8.0, and 3 mM $CaCl_2$.

Crystallization, Data Collection, and Refinement. We expressed and purified mouse P-cadherin EC1–EC2, as described above, and used it for crystallization studies at 2.58 mg/mL. Using the vapor-diffusion method, protein crystals grew at 20 °C after combining 0.6-μL protein with 0.6-μL well solution composed of 38% (wt/vol) PEG 6000, 350 mM calcium chloride, and 100 mM Bis-Tris pH6.5. The crystals grew within 48 h and were flash-frozen in liquid nitrogen after being immersed briefly in cryo protectant [15% (vol/vol) butane-2*R*,3*R*-di-ol, 18% (wt/vol) PEG 6000, 350 mM calcium chloride, 100 mM Bis-Tris pH6.5]. For mouse P-cadherin EC1–EC2, data were collected on a single frozen crystal at the X4C beam line of the National Synchrotron Light Source, Brookhaven National Laboratory at a wavelength of 0.979 Å. We processed the data using the HKL suite (2) and solved the structure by molecular replacement with mouse E-cadherin (PDB ID code 2QVF) as search model using Phaser (3). Refinement was carried out by manual building in Coot (4) followed by automated refinement in Phenix (5). Ramachandran plot statistics for the final model are 96.4% favored, 3.6% allowed and 0% outliers.

Crystals of N-cadherin EC1–EC2 A78S I92M mutant were obtained in 15% (wt/vol) PEG3350, 0.2 M sodium chloride, 0.1 M Tris, pH 8.5. The crystals were cryoprotected with the mother liquor supplemented with 30% (vol/vol) glycerol before flash-cooling in liquid nitrogen. These crystals belong to space group $P2_1$ with cell dimensions a = 59.8 Å, b = 221.2 Å, c = 72.2 Å, and β = 103.9°. Crystals of N-cadherin EC1–EC2 mutant with AA-insertion between residues 2 and 3 were obtained in 0.3 M magnesium formate, 0.1 M Hepes, pH 7.5, and were cryoprotected with the mother liquor supplemented with 30% (vol/vol) glycerol for freezing in liquid nitrogen. These crystals belong to space group $P2_12_12$ with a = 175.5 Å, b = 65.8 Å, and c = 102.5 Å. Crystals of the N-cadherin EC1–EC2 W2F mutant were obtained in 4% (wt/vol) PEG4000, 0.2 M magnesium chloride, 0.1 M MES, pH 6.5. The crystals were cryoprotected with the mother liquor supplemented with 30% (vol/vol) ethylene glycol before flash cooling in liquid nitrogen. These crystals belong to space group $C2$ with cell dimensions a = 116.6 Å, b = 86.2 Å, c = 46.7 Å, and β = 98.5°.

Diffraction data for all mouse N-cadherin EC1–EC2 mutants were collected on single crystals at 100 K at the X4C beam line in the National Synchrotron Light Source at Brookhaven National Laboratory, and processed with the HKL program suite (2).The structures were solved by molecular replacement with PHASER using the structure of wild type N-cadherin EC1–EC2 (PDB ID code 2QVI). Manual rebuilding was done with COOT, and refinement was performed using REFMAC (6) implemented in

the CCP4 program suite (Collaborative Computational Project Number 4, 1994). The statistics of data collection and refinement are summarized in Table S1.

Structural biology applications were provided by SBGrid (7).

SPR Binding Assays. Biotinylated versions of mouse E-, N-, and R-cadherins EC1–EC2 domains were captured over neutravidin-immobilized surfaces. Mouse P- and N-cadherins EC1–EC2, together with *Xenopus* C-cadherin EC1–EC2 were covalently immobilized via a C-terminal cysteine using a ligand thiol-coupling protocol. The biotinylated and thiol-coupled N-cadherin surfaces yielded similar results suggesting that the tethering method does not influence the binding responses. Each cadherin ligand was tethered to the chip surface at 70.0 μM monomer concentration, which was calculated using the homophilic binding K_Ds, as listed in Table 1.

Neutravidin immobilization and capture of biotinylated cadherin were performed as described previously (1). Thiol-coupled proteins were immobilized in HBS pH 7.4 (10 mM Hepes, 150 mM NaCl, pH 7.4), 3 mM $CaCl_2$ at 25 °C using a flow rate of 20 μL/min. For the immobilization reaction, the carboxyl groups were activated for 2 min using 400 mM EDC [*N*-ethyl-*N*-(3-dimethylaminopropyl)carbodiimide], mixed at 1:1 ratio (vol/vol) 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 (vol/vol), to yield a final concentration of 80 mM PDEA and injected over the same flow cell for 4 min. A sample of the cadherin protein to be immobilized was freshly desalted in 10 mM sodium acetate, pH 4.0 and sequentially injected over the activated surface at 10–50 mg/mL, depending on the protein, until the desired immobilization level was achieved. Any remaining disulfides were blocked using a 4-min injection of 50 mM L-cysteine/1.0 M NaCl in 0.1 M sodium acetate, pH 4.0.

Cadherin binding experiments were performed at 25 °C in a running buffer of 10 mM Tris·HCl, pH 8.0, 150 mM NaCl, 3 mM $CaCl_2$, 0.25 mg/mL BSA, and 0.005% (vol/vol) Tween 20. Soluble cadherin (analytes) were diluted in running buffer to a 12.0 μM monomer concentration, which were calculated using the homophilic K_D values listed in Table 1. Samples were injected for 60 s at 50 μL/min followed by a 60-s dissociation phase. A 1-min buffer wash at the end of the binding cycle minimized sample contamination of the fluidics system of the instrument. Each cadherin was tested at least twice to verify the reproducibility of the assay and confirm stability of the immobilized surfaces. The binding responses were double-referenced (1) and processed using Scrubber 2.0 (BioLogic Software).

1. Katsamba P, et al. (2009) Linking molecular affinity and cellular specificity in cadherin-mediated adhesion. *Proc Natl Acad Sci USA* 106(28):11594–11599.
2. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Macromolecular Crystallography, Part A*, Methods in Enzymology, eds Abelson JN, Simon MI, Carter CW, Jr, Sweet RM (Academic, San Diego), Vol 276, pp 307–326.
3. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Crystallogr* 40(Pt 4): 658–674.
4. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr, Sect D: Biol Crystallogr* 60:2126–2132.
5. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr, Sect D: Biol Crystallogr* 66(Pt 2):213–221.
6. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr, Sect D: Biol Crystallogr* 53(Pt 3): 240–255.
7. Morin A, et al. (2013) Collaboration gets the most out of software. *eLife* 2:e01456.

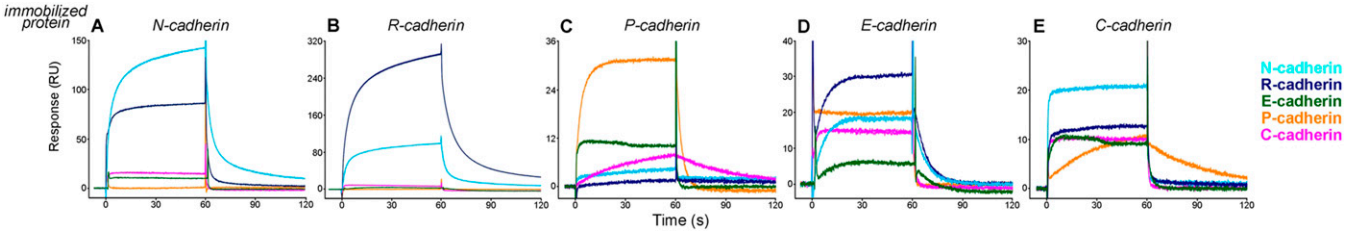

Fig. S1. (*A–E*) SPR analysis of the heterophilic binding of type I cadherins. Each of the five type I cadherins, C-, E-, N-, P-, and R-cadherin analytes was injected at 12-μM monomer concentration over individual sensor chip surfaces immobilized with each of the five type I cadherin at 70-μM monomer concentration. The responses are color-coded as indicated by the legend. The immobilized molecule for each panel protein is shown in italics at the top.

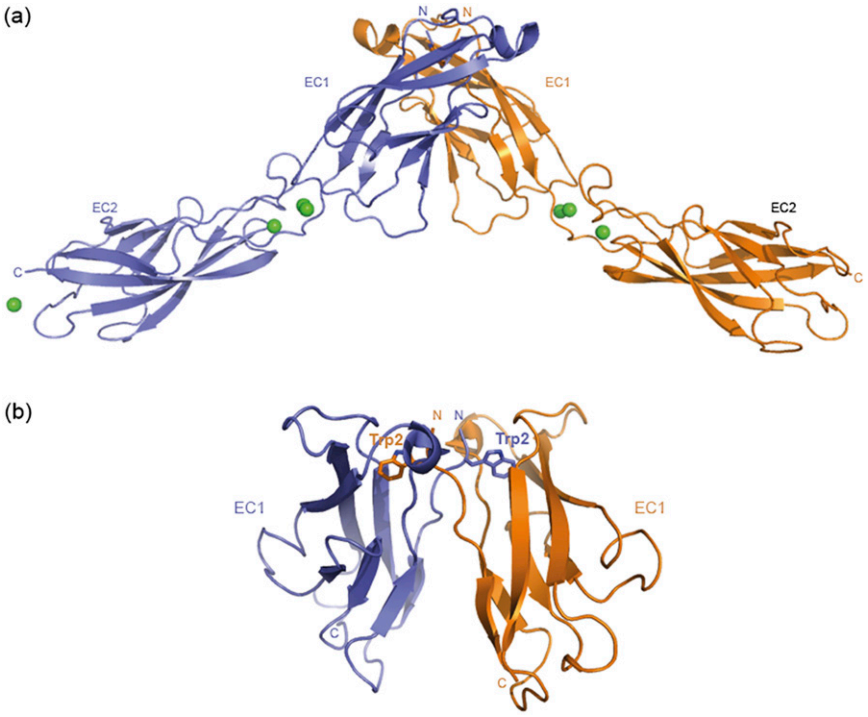

Fig. S2. Crystal structure of P-cadherin adhesive fragment EC1–EC2 reveals an adhesive strand-swapped dimer. (*A*) Strand-swapped dimer formed by wild-type P-cadherin observed in the crystal structure. One protomer shown in blue ribbon representation; the binding partner shown in orange; green spheres represent calcium (II) ions. (*B*) Close-up of EC1 domains showing exchanged A*-strands and docking of Trp2 residues (stick representation). Dimer formed between chains A and B (PDB ID code 4NQQ) is depicted in the figure. Chains C and D also adopt a strand swapped dimer configuration, however, EC1 of chain D is poorly ordered.

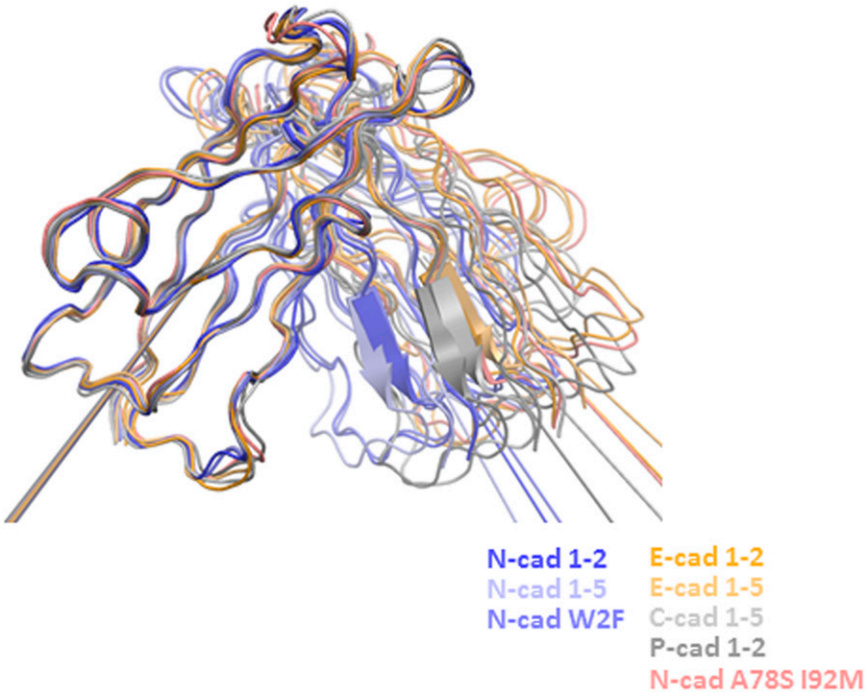

Fig. S3. Ribbon diagram representation of superposed type I classical cadherin strand-swapped dimers. EC1 domains in crystal structures of E-cadherin EC1–EC2 (PDB ID code 2QVF), E-cadherin EC1–EC5 (PDB ID code 3Q2V), C-cadherin EC1–EC5 (PDB ID code 1L3W), P-cadherin (PDB ID code 4NQQ), N-cadherin EC1–EC2 (PDB ID code 2QVI), N-cadherin EC1–EC5 (PDB ID code 3Q2W), and the N-cadherin double-mutant A78S I92M (PDB ID code 4NUM) are shown. The color code is indicated on the figure. For all dimer structures, only the protomer in the foreground has been superposed. The A-strand of the other protomer highlights the difference in the relative orientation of the two protomers within each swapped dimer. The long axis of each EC1 domain used to calculate the dimer angles (*Materials and Methods*) is indicated.

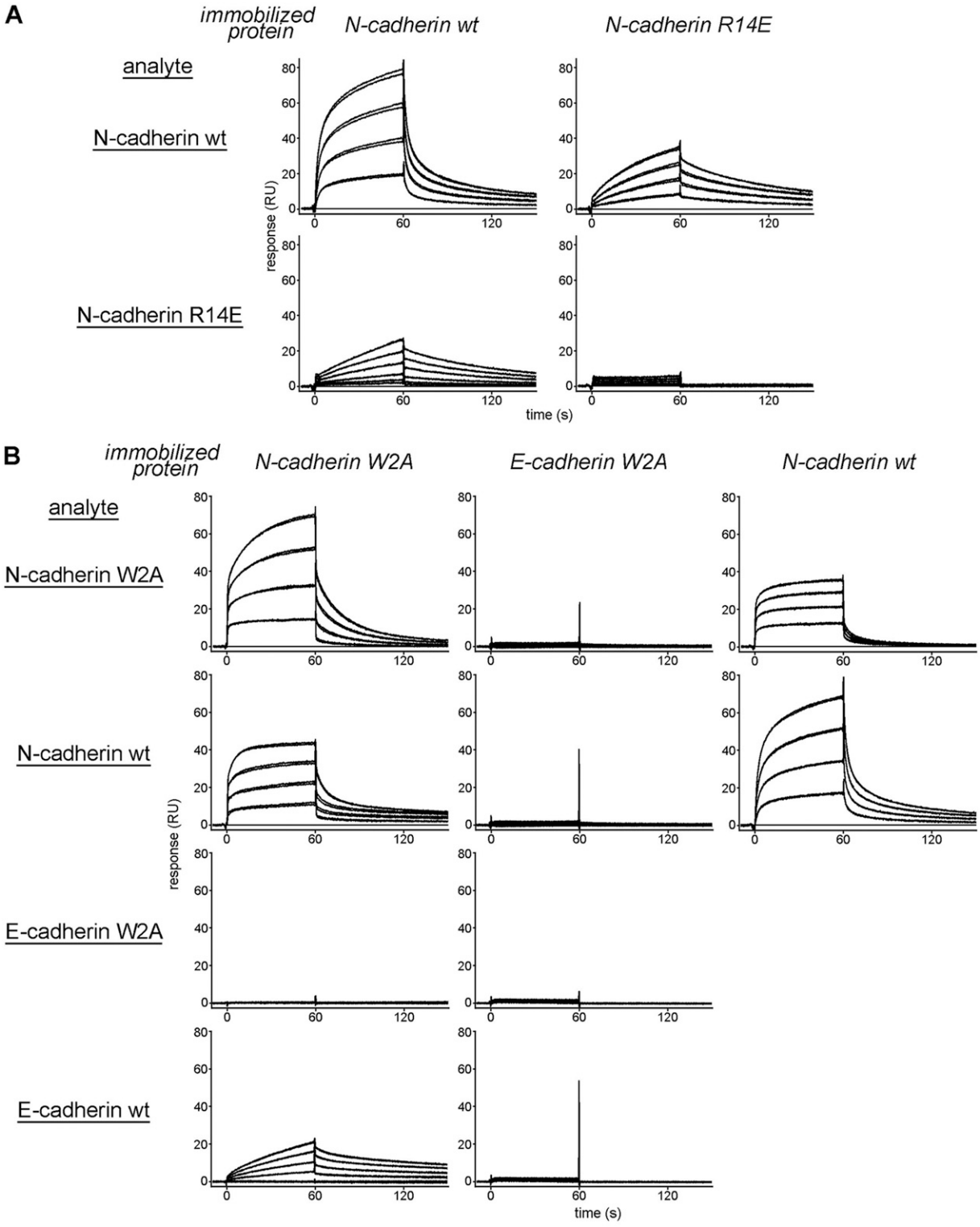

Fig. S4. SPR analysis of W2A strand-swap mutant and R14E X-dimer mutants of N-cadherin. For each panel, the immobilized proteins are indicated in italics at the top and the injected analytes are shown underlined at the left side. All panels are depicted on the same scale. Proteins were immobilized at concentrations corresponding to 100 μM monomer and analyte binding was tested at concentrations corresponding to 12-, 9-, 6-, and 3-μM monomer. (*A*) Binding of N-cadherin wild-type and its R14E X-dimer mutant, to surfaces immobilized with N-cadherin wild-type and N-cadherin R14E. (*B*) N-cadherin W2A and E-cadherin W2A strand-swap mutants were tested for binding over surfaces immobilized with N-cadherin W2A, E-cadherin W2A, and N-cadherin wild-type proteins. The binding of the wild-type N- and E-cadherins was included as a control.

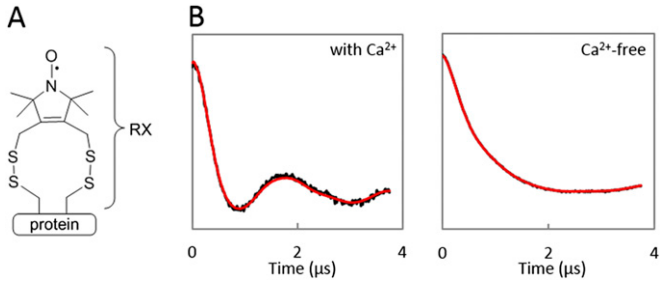

Fig. S5. Double electron-electron resonance (DEER) analysis of RX-labeled E-cadherin. (*A*) Chemical structure of the RX side-chains. (*B*) Background-subtracted dipolar evolutions of E-cadherin 73/75RX 114/116RX in the presence (*Left*) and absence (*Right*) of Ca^{2+}. Both the experimental form factor (black trace) and the best fit (red trace) from the program LongDistances (1) are shown.

1. López CJ, Yang Z, Altenbach C, Hubbell WL (2013) Conformational selection and adaptation to ligand binding in T4 lysozyme cavity mutants. *Proc Natl Acad Sci USA* 110(46): E4306–E4315.

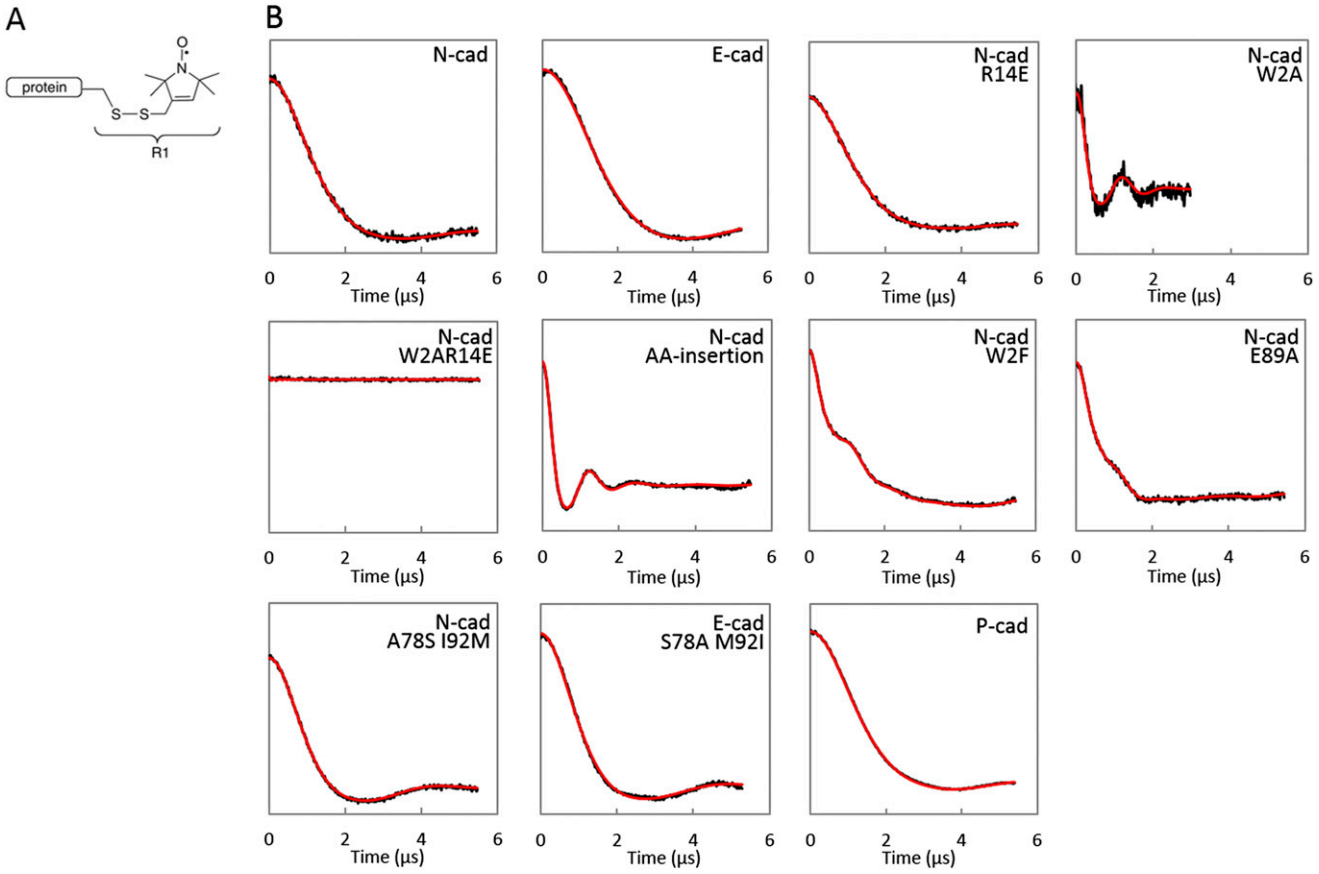

Fig. S6. DEER analysis of R1-labeled E-, N-, P-cadherin and mutants. (*A*) Chemical structure of the R1 side chain. (*B*) Background-subtracted dipolar evolutions of E-, N-, P-cadherin and indicated mutants, all labeled with R1 at position 135. Constructs for E-cadherin include an additional C9S mutation to remove the native cysteine. Both the experimental form factor (black trace) and the best fit (red trace) from the program LongDistances (1) are shown.

1. López CJ, Yang Z, Altenbach C, Hubbell WL (2013) Conformational selection and adaptation to ligand binding in T4 lysozyme cavity mutants. *Proc Natl Acad Sci USA* 110(46): E4306–E4315.

Table S1. Data collection and refinement statistics

Data collection and refinement	P-cadherin wild-type	N-cadherin A78S I92M	N-cadherin W2F	N-cadherin AA insertion
Data collection				
Space group	$P2_1\ 2_1\ 2$	$P2_1$	$C2$	$P2_12_12$
Cell dimensions: *a*, *b*, *c* (Å)	123.1, 188.7, 53.2	59.8, 221.2, 72.2	116.6, 86.2, 46.7	175.5, 65.8,102.5
α, β, γ (°)	90, 90, 90	90, 103.9, 90	90, 98.5, 90	90, 90, 90
Molecules per asymmetric unit	4	4	1	3
Resolution limit (Å)	40–3.2	20–3.2	20–2.1	20–2.7
Unique reflections	21,058	25,726	25,187	33,781
Redundancy (highest resolution shell)	6.3 (6.1)	3.5 (1.9)	3.4 (3.0)	5.3 (3.7)
Completeness (%) (highest resolution shell)	99.3 (94.2)	94.7 (68.2)	97.7 (85.7)	99.7 (96.9)
Average I/σ(I) (highest resolution shell)	11.6 (3.1)	9.5 (2.1)	34.3 (3.0)	13.6 (2.0)
R_{merge} (%) (highest resolution shell)	11.6 (58.9)	12.0 (44.5)	11.2 (46.1)	11.6 (44.4)
Refinement				
Resolution limit (Å)	20–3.2	20–3.2	20–2.1	20–2.7
R_{work} (%)	22.8	21.9	21.1	17.2
R_{free} (%)	26.9	25.3	25.3	22.9
Rmsd bonds (Å)	0.01	0.006	0.008	0.009
Rmsd angles (°)	0.9	1.07	1.08	1.26
Protein atoms	6,552	6,668	1,659	5,042
Ligand/ion atoms	23	12	3	10
Water molecules	76	0	159	576
Average B (Å^2) protein atoms	71.4	68.5	53.5	23.0
Average B (Å^2) ligand/ion atoms	63.8	60.7	48.9	26.4
Average B (Å^2) water molecules	24.1	N/A	57.6	23.9

Table S2. Total and hydrophobic surface area buried at the swapped interfaces as determined from crystal structures of N-, E-, C-, P-cadherins and the N-cadherin mutant A78S I92M

Cadherin	Residues 78 and 92	Total BSA (Å^2)	Hydrophobic BSA (Å^2, by residue)	EC1-EC1 dimer angle (°)
N-cadherin wild-type EC1-EC2 (2QVI)	A78, I92	1,839	1,202	75.7
N-cadherin wild-type EC1-EC5 (3Q2W)	A78, I92	1,900	1,218	73.4
N-cadherin W2F (PDB ID code 4NUQ)	A78, I92	1,796	1,149	77.6
N-cadherin A78S I92M (PDB ID code 4NUM form 1)	S78, M92	1,747	998	84.2
N-cadherin A78S I92M (PDB ID code 4NUM form 2)	S78, M92	1,751	1,027	83.9
E-cadherin wild-type mouse (PDB ID code 2QVF)	S78, M92	1,834	1,008	85.3
E-cadherin wild-type mouse EC1-5 (PDB ID code 3Q2V)	S78, M92	1,725	962	89.2
E-cadherin wild-type human (PDB ID code 2O72)	S78, M92	1,800	1,031	88.1
C-cadherin wild-type (PDB ID code 1L3W)	S78, M92	1,840	970	88.7
P-cadherin wild-type (PDB ID code 4NQQ)	G78, M92	1,850	996	83.4

The EC1–EC1 dimer angle is indicated for each crystal structure.

Table S3. Dissociation constants (K_D) for wild type E- and N-cadherins and their pocket mutants

Mutation	K_D (μM)		Mutant description
	E-cadherin	N-cadherin	
M92I	54.6 ± 3.8	—	E-cadherin with N-like Trp2 pocket-lining residue 92
S78A	85.7 ± 3.9	—	E-cadherin with N-like Trp2 pocket-lining residue 78
I92M	—	2.6 ± 0.2	N-cadherin with E-like Trp2 pocket-lining residue 92
A78S	—	34.4 ± 2.0	N-cadherin with E-like Trp2 pocket-lining residue 78

Data are given as mean ± SD.