Structure *19*

Supplemental Information

The Extracellular Architecture

of Adherens Junctions Revealed

by Crystal Structures of Type I Cadherins

Oliver J. Harrison, Xiangshu Jin, Soonjin Hong, Fabiana Bahna, Goran Ahlsen, Julia Brasch, Yinghao Wu, Jeremie Vendome, Klara Felsovalyi, Cheri M. Hampton, Regina B. Troyanovsky, Avinoam Ben-Shaul, Joachim Frank, Sergey M. Troyanovsky, Lawrence Shapiro, and Barry Honig

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Protein expression and purification

Coding sequences for the ectodomains of mouse E-cadherin (Asp1 to Ala544 of the mature protein) and mouse N-cadherin (Asp1 to Val553) were cloned into the mammalian expression vector pCEP4 (Invitrogen). A sequence encoding the signal peptide of human CD33 was included before the ectodomain coding region to facilitate protein secretion, and a hexahistidine tag was included at the C-terminus. All point mutations were introduced using Quikchange site-directed mutagenesis (Stratagene). To express proteins for structural studies, plasmids were transfected into human embryonic kidney (HEK) 293 GnTI$^-$ cells (Reeves et al., 2002). These cells are deficient in the enzyme *N*-acetylglucosaminyltransferase I and consequently modify proteins at N-linked glycosylation sites with only minimal high mannose type glycosyl moieties ($Man_5GlcNAc_2$). E-cadherin used for functional experiments was expressed instead in normal HEK 293F cells. Transfections were performed using Lipofectamine 2000 (Invitrogen) and positive cells were selected with hygromycin B. Proteins were purified from 4 liters of conditioned media by Ni-NTA affinity chromatography in a batch procedure, and then further purified by anion and cation exchange chromatography and size exclusion chromatography on an Akta FPLC system (GE Healthcare). E-cadherin ectodomains for crystallization were additionally treated with endoglycosidase H (EndoH, New England Biosciences) prior to the size exclusion step in order to remove N-linked sugars, leaving only a single GlcNAc residue at sites of N-linked modification. All proteins were concentrated in 150mM NaCl, 10mM Tris-Cl pH 8.0, 3mM $CaCl_2$ to 5-10mg/ml and were judged to be pure from Coomassie staining of SDS-PAGE gels. Peptide sequencing confirmed that N-termini were correct. Molecular masses determined by mass spectrometry were 65878Da (E-cad, from HEK 293F cells), 65735Da (E-cad, GnTI- cells), 62700Da (E-cad, EndoH treated) and 70260Da (N-cad, GnTI- cells).

The coding sequence for an EC1-EC2 fragment of mouse E-cadherin (Asp1 to Asp213 of the mature protein) was cloned into the bacterial expression vector pSMT3 (Mossessova and Lima, 2000). Point mutations were introduced using the Quikchange method (Stratagene). Proteins were expressed in Rosetta 2 *E.coli* cells (Novagen) and were purified exactly as described previously (Harrison et al., 2010) to yield proteins with

native mature N-termini. These were concentrated to ~10mg/ml in 150mM NaCl, 10mM Tris-Cl pH 8.0, 3mM $CaCl_2$.

Crystallization and Structure determination

All crystals were grown using the vapor diffusion method in hanging drops at 20°C. For E-cadherin EC1-5, 2µl of protein solution at 8mg/ml was mixed with 1µl of reservoir solution containing 11% (w/v) PEG 5000 monomethyl ether (MME), 0.1M Bicine pH 8.5, 10mM CaCl2, 30mM MnCl2, 2% (v/v) dioxane and 4% (v/v) 1-butanol. Crystals were dehydrated for 24hrs over reservoirs containing an additional 11% PEG 5000MME and were cryoprotected with 30% (v/v) ethylene glycol. For N-cadherin EC1-5, the protein at 8mg/ml was mixed with an equal volume of reservoir solution containing 25% PEG8000, 0.1M Tris, pH 8.5; crystals were cryoprotected with 30% glycerol. For E-cadherin EC1-2 V81D, crystallization conditions were: 10% (v/v) PEG 400, 0.1M sodium acetate pH4.6, 0.13M $CaCl_2$. E-cadherin EC1-2 L175D crystallized in similar conditions, except with 25% (v/v) PEG 400 and 0.15M $CaCl_2$. Both were cryoprotected by increasing PEG 400 to 30%.

Data was collected from single crystals at 100K using a wavelength of 0.979Å at the X4A and X4C beamlines of the National Synchrotron Light Source, Brookhaven National Laboratory. Images were processed using HKL (Otwinowski and Minor, 1997). E-cadherin EC1-5 structure was solved by molecular replacement with Molrep (Vagin and Teplyakov, 1997) using the structure of C-cadherin EC1-3 as the search model. After several rounds of refinement, models of EC4 and EC5 domains were built into areas of clear density. N-cadherin EC1-5 structure was solved by molecular replacement with Molrep (Vagin and Teplyakov, 1997) using the E-cadherin EC1-5 structure as the search model. E-cadherin EC1-2 structures were solved by molecular replacement in Phaser (McCoy et al., 2007) using the structure of wild-type mouse E-cadherin EC1-2 (PDB 2QVF, (Harrison et al., 2010)) as a search model. Model refinement was performed using Refmac (Murshudov et al., 1997) and Coot (Emsley and Cowtan, 2004). NCS was not used in the refinements. Data collection and refinement statistics are listed in Table 1.

Calculation of sequence variation scores

We sought independent evidence for physiological relevance of the cadherin *cis* interface observed in crystal structures by analysing sequence conservation patterns along a multiple sequence alignment of type I cadherin EC1-2 domains. Psi-Blast was used to search for type I cadherins in the nonredundant protein sequences (NR) database. 88 full-length type I cadherin sequences were found and aligned with Muscle (Edgar, 2004) and the EC1-2 portion of the multiple sequence alignment served as an input for Rate4site (Pupko et al., 2002). Rate4site was used to calculate a score, normalized to have a mean of 0 and standard deviation of 1, which reflects relative conservation at each position of the multiple sequence alignment. Scores were then compared for surface residues in four categories: swapped dimer interface residues, X-dimer interface residues (Harrison et al., 2010), cis interface residues and exposed residues not part of any known interface. To define interface residues, Surfv (Sridharan et al., 1992) was used to calculate the surface area buried by each residue side chain at the different interfaces observed in type I crystal structures (swapped dimer, X-dimer and putative *cis* interface). Structures used were N-cadherin EC1-5, E-cadherin EC1-5 and PDB structures 1L3W, 2QVI, 1FF5, 1EDH, 1Q1P, 2QVF and 2O72. Residues whose side chain buries more than 5 $Å^2$ in more than half the structures were considered interface residues. Exposed residues were defined

using Surfv as residues with at least 25% of the side chain accessible to solvent in more than half of the crystal structures; remaining residues were considered buried. Calcium-binding residues were excluded from all categories, because of their high conservation due to their structural role. Residues found in both swapped and X-dimer interfaces were included in each category. For statistical comparison of variation scores between the residue categories, the non-parametric Mann-Whitney-U test (Wilcoxon rank-sum test) (Mann and Whitney, 1947) was used, since the distribution of scores does not follow a normal distribution.

Liposome aggregation assays

Liposomes were prepared by a hydration and extrusion method from a 9:1 molar ratio of 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC) and the nickel salt of 1,2-dioleoyl-*sn*-glycero-3-([*N*(5-amino-1-carboxypentyl)iminodiaceticacid]-succinyl) (DOGS-NTA) according to the manufacturer's protocol (Avanti Lipids). Lipids were hydrated with assay buffer (25mM HEPES pH 7.4, 0.1M KCl, 10% (v/v) glycerol, 3mM $CaCl_2$); membranes with a pore size of 100nm were used for extrusion. Liposome aggregation assays were performed essentially as described in (He et al., 2009): liposomes and hexahistidine-tagged E-cadherin ectodomain proteins were mixed at final concentrations of ~10mM lipid and 20μM protein in assay buffer and incubated at room temperature. Aggregation was monitored by measuring light scattering at 650nm in a spectrophotomer at 20 second intervals starting immediately after mixing (time zero) for a total of 40 minutes. Experiments were performed in triplicate; samples containing only buffer and liposomes served as controls.

Cryo-EM of Cadherin Liposomes

Aggregated cadherin-conjugated liposomes were applied to 300 mesh copper TEM grids with R 2/1 Quantifoil carbon film (Quantifoil Micro Tools GmbH, Jena, Germany). Grids were plasma-cleaned using a Solarus plasma cleaner (Gatan Inc., Warrendale, PA) just prior to use. The samples were vitrified by blotting and plunge-freezing into liquid nitrogen-cooled ethane using an automated Vitrobot (FEI Company, Hillsboro, Oregon). Frozen grids were transferred to a Tecnai Polara F30 TEM (FEI) and imaged at 300 kV using a Tietz 4K x 4K CCD camera (Tietz Video and Image Processing Systems GmbH, Gauting, Germany). Images were recorded under low-dose conditions at ~10 μm underfocus using the SerialEM software (Mastronarde, 2005). Microscope magnification was 39,000-59,000x. Images were processed using the IMOD software (Kremer et al., 1996).

Analytical Ultracentrifugation

Analytical ultracentrifugation (AUC) experiments were performed using a Beckman XLA/I ultracentrifuge, equipped with either a Ti50An or a Ti60An rotor. For sedimentation equilibrium experiments, all proteins were dialyzed overnight at 4ºC in 10mM Tris-HCl, pH 8.0, 150mM NaCl, 3mM $CaCl_2$, with 3mM TCEP added for E-cadherin samples. 120 μL of each protein at a concentration of 0.7, 0.46, or 0.24 mg/mL were loaded into a cell with a 12 mm six-channel centerpiece and sapphire windows. Data was collected at 25ºC using UV at 280 nm and interference at 660 nm. Samples of two-domain proteins were spun at 18000 rpm for 20 hours after which four scans (1 per hour) were collected. Speed was increased to 22000 rpm for 10 hours, then to 26000 rpm for 10 hours, with four hourly scans taken after each period. Five-domain proteins were spun using the same protocol except for using 11000, 13000 and 15000 rpm, respectively.

This protocol gives 72 scans for each protein. Buffer density and v-bar were calculated using SednTerp. (Alliance Protein Laboratories, Corte Cancion, Thousand Oaks, CA, USA). The scans were processed and analyzed using HeteroAnalysis 1.1.0.28 (http://www.biotech/uconn.edu/auf/). Data from all concentrations, speeds and both detection systems were globally fitted by non-linear regression to a monomer/dimer equilibrium model to calculate the K_D for each homodimer. All values were determined from at least two independent experiments.

Cell Culture, Antibodies, Plasmids, and DNA Transfections.
Transfection, growth, and immunofluorescence microscopy of human A-431, and A-431D cells were performed as described (Troyanovsky et al., 2007). All cell lines were cultured in DMEM/10% FCS. The plasmid pRc-EcDendra-Δ748-KL encoding human E-cadherin truncated mutant EcadΔ-Dendra was described (Hong et al., 2010). The double point mutation V81D V175D inactivating the *cis*-interface and the point mutation W2A inactivating the strand-swap interface were introduced by site-directed mutagenesis. After transfection and G-418 selection, cells were sorted by flow cytometry. Only cells expressing a moderate level of transgenes were studied. The selected cultures exhibited the same expression level of recombinant Dendra-tagged proteins. For co-culture and co-immunoprecipitation experiments, Dendra tag in the pRc-EcDendra-Δ748-KL plasmid was replaced for mCherry and 6xmyc, respectively.

The following antibodies were used: anti-E-cadherin, clone C20820 (BD Biosciences, San Jose, CA), recognizing only the endogenous cadherin; mouse anti-β-catenin (BD Biosciences); rabbit anti-Dendra2 (Evrogen, Moscow, Russia); and mouse anti-myc, clone 9E10 (Santa Cruz Biotechnology, Santa Cruz, CA).

Co-imunoprecipitation Assay and Western Blotting.
To analyze cadherin dimerization, we used a co-immunoprecipitation assay (Troyanovsky et al., 2006). In brief, cells were extracted with 1 ml of 1% Triton X-100-containing immunoprecipitation (IP) lysis buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 2 mM EDTA, and 0.5 mM AEBSF) and then subjected to subsequent incubations, first with anti-Dendra antibody and then with protein A-Sepharose. The resulting immunoprecipitates were analyzed by immunoblotting

Live-cell imaging and data processing.
Experiments were performed as described previously (Hong et al., 2010). All imaging experiments were performed one day after plating in imaging media (L-15 plus 10% FBS). Cells were imaged using an Eclipse Ti-E microscope controlled with Nikon's NIS-Elements software and Perfect Focus System. The microscope was equipped with an incubator chamber, a CoolSNAP HQ2 camera, and two different halogen and mercury light sources. To analyze cadherin junctional turnover, we used Dendra photoactivation assay (Hong et al., 2010). A circular region of interest (ϕ=5μm) was photoactivated by a three-second-long exposure to the 402 nm light using the mercury arc light source and a pinhole insert. Time-lapse images were taken in both FITC and TRITC filter sets using a halogen light source that minimized phototoxicity and photobleaching. In case of co-cultures of Dendra- and mCherry-tagged cadherin-expressing cells, a FRAP assay was used: in the region of interest, Dendra intrinsic green fluorescence was first fully photoconverted by four-second-long exposure to the 402 nm light. Then the recovery of green fluorescence in the region was monitored over the time. All images were saved as Tiff files for the further image analysis.

Tiff stacks were processed using ImageJ software (National Institutes of Health). A circular region of interest (ϕ=0.65μm) was positioned on a single AJ and the mean value of fluorescent intensity in each frame was calculated. In photoactivation and FRAP experiments, the red and green fluorescent intensity, correspondingly, was normalized in such a way that 0 and 1 corresponded to the background and the initial values, respectively. The background value was obtained from the image taken right before the photoactivation. In case of FRAP, the background value was obtained immediately after 4-second long green-to-red photoconversion. The time course of intensity change was produced from 10 sets of independent experiments.

Cell aggregation assays

Confluent monolayers of transfected or parental A431D cells were dissociated with non-enzymatic cell dissociation solution (Mediatech) and resuspended in DMEM/10%FCS with 3mM $CaCl_2$ or 3mM EDTA (for control samples) at a density of 6 x 10^5 viable cells per ml. 0.8ml of each cell suspension was transferred to a 1.5ml tube and allowed to aggregate at 37°C with constant inversion (~12rpm). Cell aggregation was assessed after 30 minutes of aggregation by light microscopy of 20μl samples. 'Time zero' samples taken at the start of the aggregation assay were monodisperse (not shown). Triplicate samples of each cell line were tested in three independent experiments, giving similar results.

SUPPLEMENTAL REFERENCES

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res *32*, 1792-1797.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr *60*, 2126-2132.

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.

He, Y., Jensen, G.J., and Bjorkman, P.J. (2009). Cryo-electron tomography of homophilic adhesion mediated by the neural cell adhesion molecule L1. Structure *17*, 460-471.

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.

Kremer, J.R., Mastronarde, D.N., and McIntosh, J.R. (1996). Computer visualization of three-dimensional image data using IMOD. J Struct Biol *116*, 71-76.

Liwosz, A., Lei, T., and Kukuruzinska, M.A. (2006). N-glycosylation affects the molecular organization and stability of E-cadherin junctions. J Biol Chem *281*, 23138-23149.

Mann, H.B., and Whitney, D.R. (1947). On a Test of Whether One of 2 Random Variables Is Stochastically Larger Than the Other. Ann Math Stat *18*, 50-60.

Mastronarde, D.N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. J Struct Biol *152*, 36-51.

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.

Mossessova, E., and Lima, C.D. (2000). Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. Molecular Cell *5*, 865-876.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr *53*, 240-255.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Macromolecular Crystallography, Pt A *276*, 307-326.

Pupko, T., Bell, R.E., Mayrose, I., Glaser, F., and Ben-Tal, N. (2002). Rate4Site: an algorithmic tool for the identification of functional regions in proteins by surface mapping of evolutionary determinants within their homologues. Bioinformatics *18 Suppl 1*, S71-77.

Reeves, P.J., Callewaert, N., Contreras, R., and Khorana, H.G. (2002). Structure and function in rhodopsin: High-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. Proceedings of the National Academy of Sciences of the United States of America *99*, 13419-13424.

Sridharan, S., Nicholls, A., and Honig, B. (1992). A New Vertex Algorithm to Calculate Solvent Accessible Surface-Areas. Faseb J *6*, A174-A174.

Troyanovsky, R.B., Laur, O., and Troyanovsky, S.M. (2007). Stable and unstable cadherin dimers: mechanisms of formation and roles in cell adhesion. Mol Biol Cell *18*, 4343-4352.

Troyanovsky, R.B., Sokolov, E.P., and Troyanovsky, S.M. (2006). Endocytosis of cadherin from intracellular junctions is the driving force for cadherin adhesive dimer disassembly. Mol Biol Cell *17*, 3484-3493.

Vagin, A., and Teplyakov, A. (1997). MOLREP: an automated program for molecular replacement. J Appl Crystallogr *30*, 1022-1025.

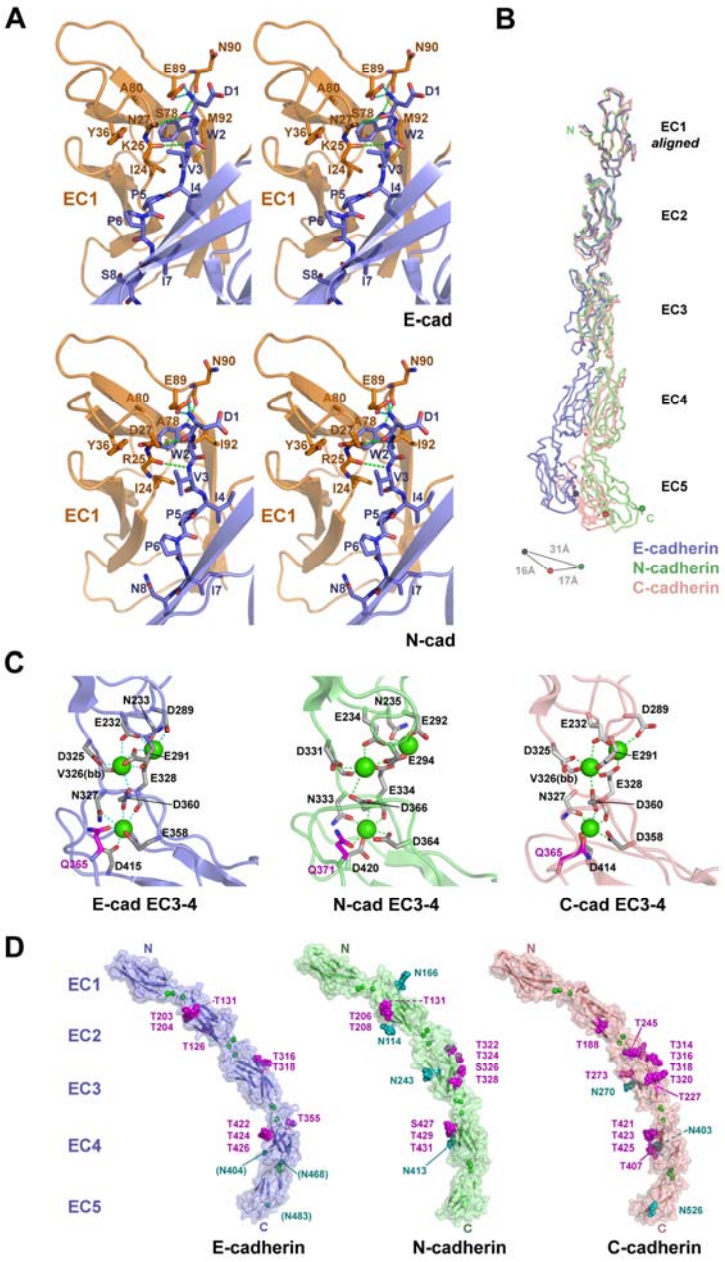

Figure S1. Structural details of type I cadherin ectodomains (related to Figure 1) **(A)** Strand swapped *trans* dimer interfaces of mouse E- and N-cadherin. Interactions between the swapped A* strand and the partner EC1 domain are shown for E-cadherin (upper panel) and for N-cadherin (lower panel). One docked strand is shown, identical interactions are observed for the docked strand of the partner molecule. Protomers are colored blue and orange. Stick representation is used for the swapped strands and for the side chains of EC1 residues forming dimer contacts. Blue and green dashed lines represent ionic bonds and hydrogen bonds, respectively. **(B)** Ectodomain curvature of E-, N- and C-cadherin. Whole ectodomain structures of E-, N- and C-cadherin (1L3W)

superposed over the EC1 domain. Distances between equivalent residues near the C-termini (Cys532, Cys542 and Cys530 in E-, N- and C-cadherin, circles) due to curvature differences are indicated. Note the divergent interdomain angles between EC2-3, EC3-4 and EC4-5. **(C)** Calcium ion coordination in E-, N-, and C-cadherin. One of the three calcium ions in the EC3-4 linker is coordinated by the side chain carboxyl group of Gln365 from EC4, in place of a backbone carbonyl group that occupies this coordination position in the EC1-2, EC2-3 and EC4-5 regions. This results in a difference in the angle between EC3 and EC4 in comparison to other pairs of successive domains, contributing to the overall curvature of the ectodomain. Side chains or backbone atoms of residues involved in calcium co-ordination are shown as sticks, calcium ions are shown as green spheres. Dashed lines indicate coordinating interactions. Gln365 is highlighted in magenta. **(D)** Glycosylation of the E-, N- and C-cadherin ectodomains. Electron density for ten O-linked carbohydrate modifications is observed in domains EC2 to EC4 of both E-cadherin and N-cadherin. Four N-linked sites were found in N-cadherin, also in EC2 to EC4. N-linked glycosylation could not be determined in the E-cadherin structure because the protein was treated with endoglycosidase H to aid crystallization, removing ~ 3kDa of N-linked carbohydrate; single GlcNAc units remaining after cleavage could not be visualized in the structure due to possible disorder and/or insufficient map quality. Three sites of N-linked glycosylation in E-cadherin (labelled in parentheses) have previously been identified (Liwosz et al., 2006). O-linked glycosylation of a T-x-T-x-T motif in the G strands of EC3 and EC4 is conserved in E-, N- and C-cadherin, while modification of the equivalent region of EC2 occurs additionally in E- and N-cadherin. Sequence conservation suggests that these regions are likely to be O-glycosylated in most type I classical cadherins. Ectodomain structures are shown in ribbon form, with a molecular surface. N-linked glycosyl moieties observed in the electron density maps are shown as blue spheres; O-linked sugars are colored magenta. Green spheres represent calcium ions.

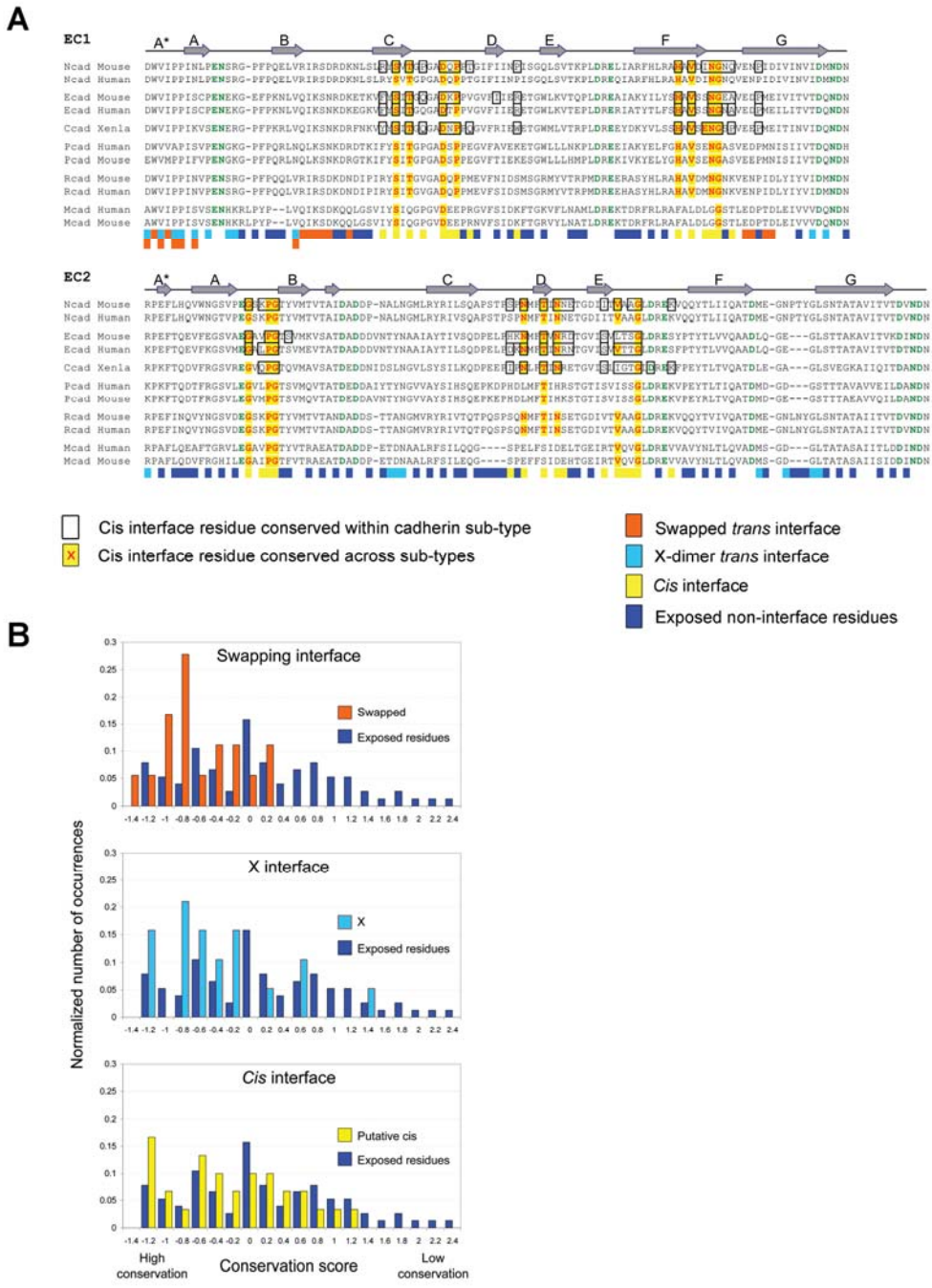

Figure S2. Sequence conservation of interface residues in type I classical cadherins (related to Figure 2)

(A) Multiple sequence alignment of the EC1-2 domains of type I classical cadherins from human, mouse and *Xenopus laevis*. Residues that comprise the *cis* interface in crystal structures of E, N- and C-cadherin (1L3W) are highlighted according to their conservation: those conserved over at least half of the cadherin subtypes are yellow; those conserved only within sub-types are boxed. Shading below the alignment shows the assignment of residues to four categories for conservation analysis: those forming the

cis, swapped *trans* and X-dimer interfaces and surface exposed residues contributing to none of these (see Extended Experimental Procedures and panel B). Buried residues (unshaded) and calcium binding residues (green text) were not used in the analysis. **(B)** Distribution of the Rate4site variation scores of residues at the different interfaces (from top to bottom, the swapping interface, X-interface and the *cis* interface) compared to the score distribution of the non-interface solvent exposed residues (dark blue bars, see Extended Experimental Procedures). Lower scores indicate better conservation. Relative to non-interface exposed residues, those in the swapped, X-dimer or *cis* interfaces all show an enrichment in highly conserved residues (low scores), and an absence of highly variable residues (high scores), as expected for biologically relevant interfaces. A Mann-Whitney-U test confirms the difference in variation between interface and non-interface exposed residues to be significant for the swapped, X-dimer and cis interfaces, with one-sided p-values of 0.0003, 0.0036 and 0.014, respectively.

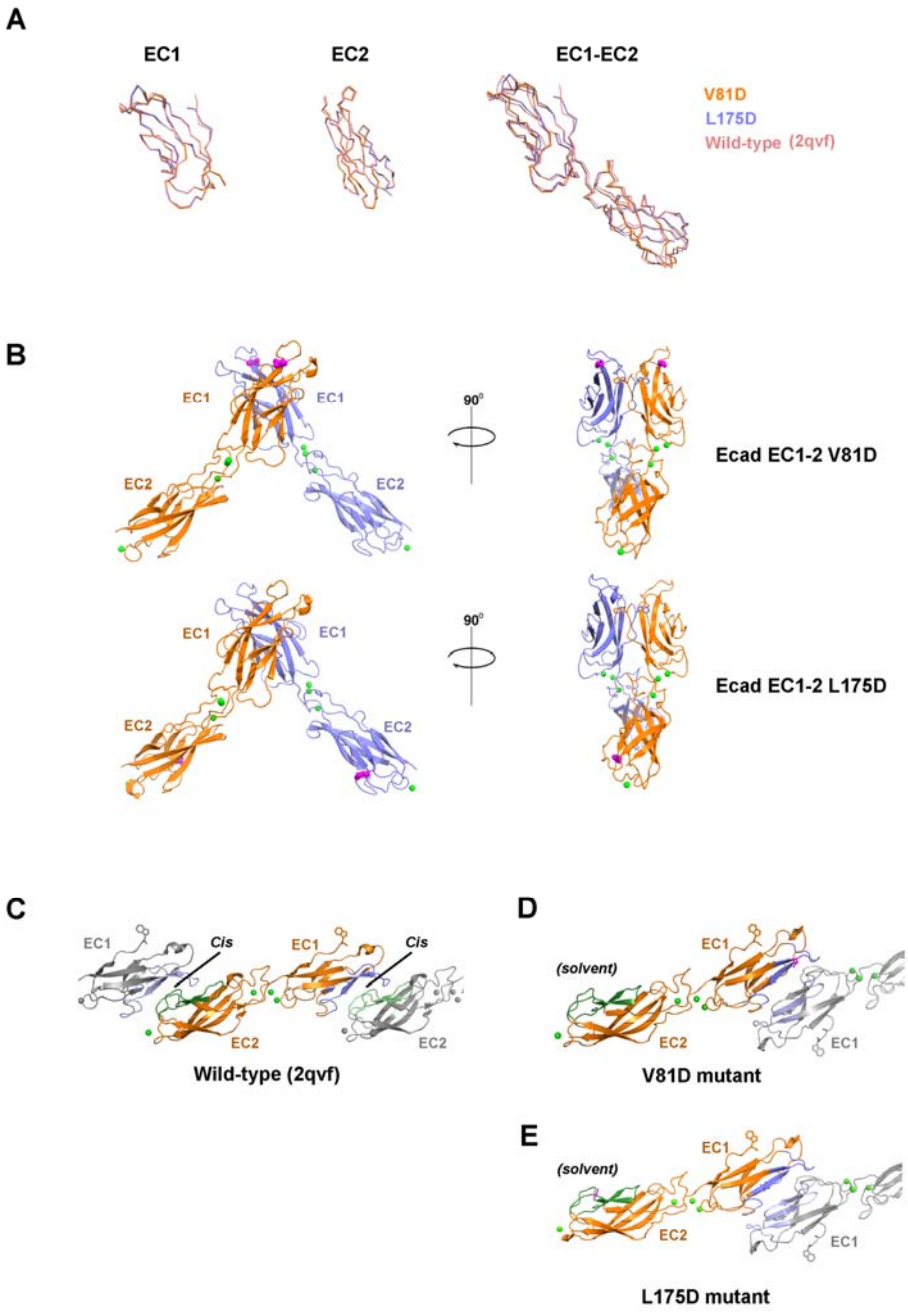

Figure S3. Crystal structures of E-cadherin EC1-2 fragments containing *cis* interface mutations (related to Figure 5)

(A) Structural comparison of *cis* interface mutant structures V81D and L175D with wild-type E-cadherin EC1-2 (2QVF). Superpositions of EC1 (residues 1 to 100), EC2 (residues 101 to 213) and EC1-2 domains are shown. Maximum pairwise rms deviations are: 0.451Å over 90 aligned Cα atoms (EC1); 0.396Å over 110 atoms (EC2); 1.416Å over 197 atoms (EC1-2). **(B)** Strand-swapped *trans* dimer structures of E-cadherin EC1-2 V81D and L175D mutants. Individual protomers are colored orange and blue, calcium

ions are represented by green spheres. Side chains of mutated residues Asp81 and Asp175 are highlighted in magenta. **(C)** *Cis* interface observed in the structure of wild-type E-cadherin EC1-2 (2QVF). Two *cis* interactions, between symmetry related protomers, are shown in the ribbon diagram. The concave (EC1) side of the interface is highlighted in light blue; the convex (EC2) side is highlighted in green. Trp2 side chains are shown to aid orientation, calcium ions are shown as green spheres. **(D, E)** *Cis* interactions are not observed in the structures of *cis* mutants V81D and L175D. Regions of the crystal lattice are displayed as in panel C. Mutated residues Asp81 and Asp175 are highlighted in pink.

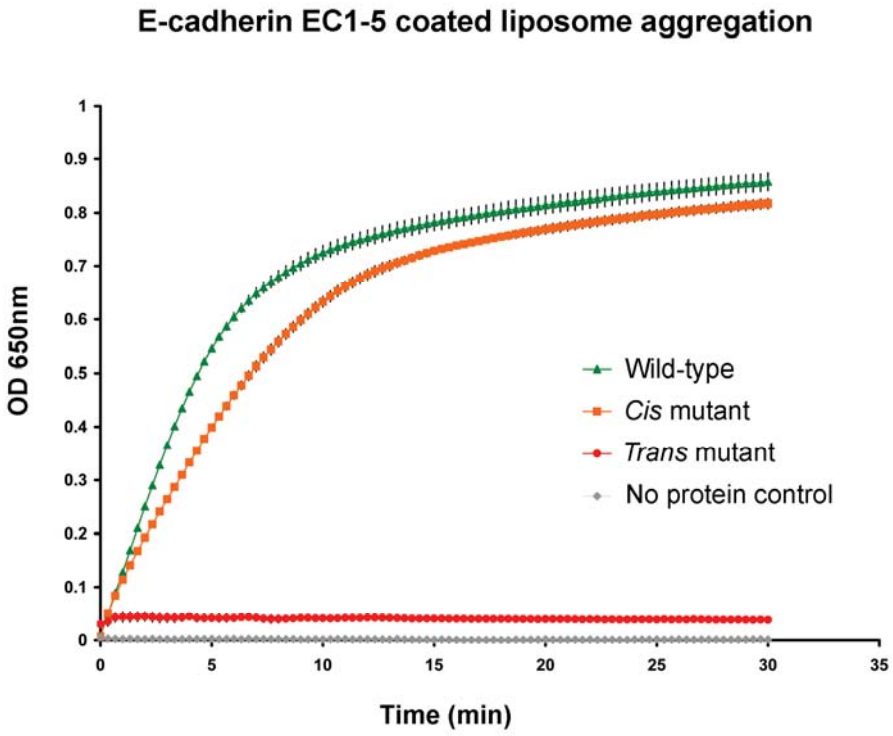

Figure S4. Aggregation of E-cadherin coated liposomes (related to Figure 4)
Liposomes coated with E-cadherin wild-type, *cis* interface mutant (V81D L175D) or *trans* interface mutant (W2A K14E), or with no cadherin (No protein control) were allowed to aggregate for 30 minutes. Increases in optical density at 650nm, corresponding to increases in aggregation of the liposomes, are plotted against time. Error bars represent standard error values from three independent experiments.

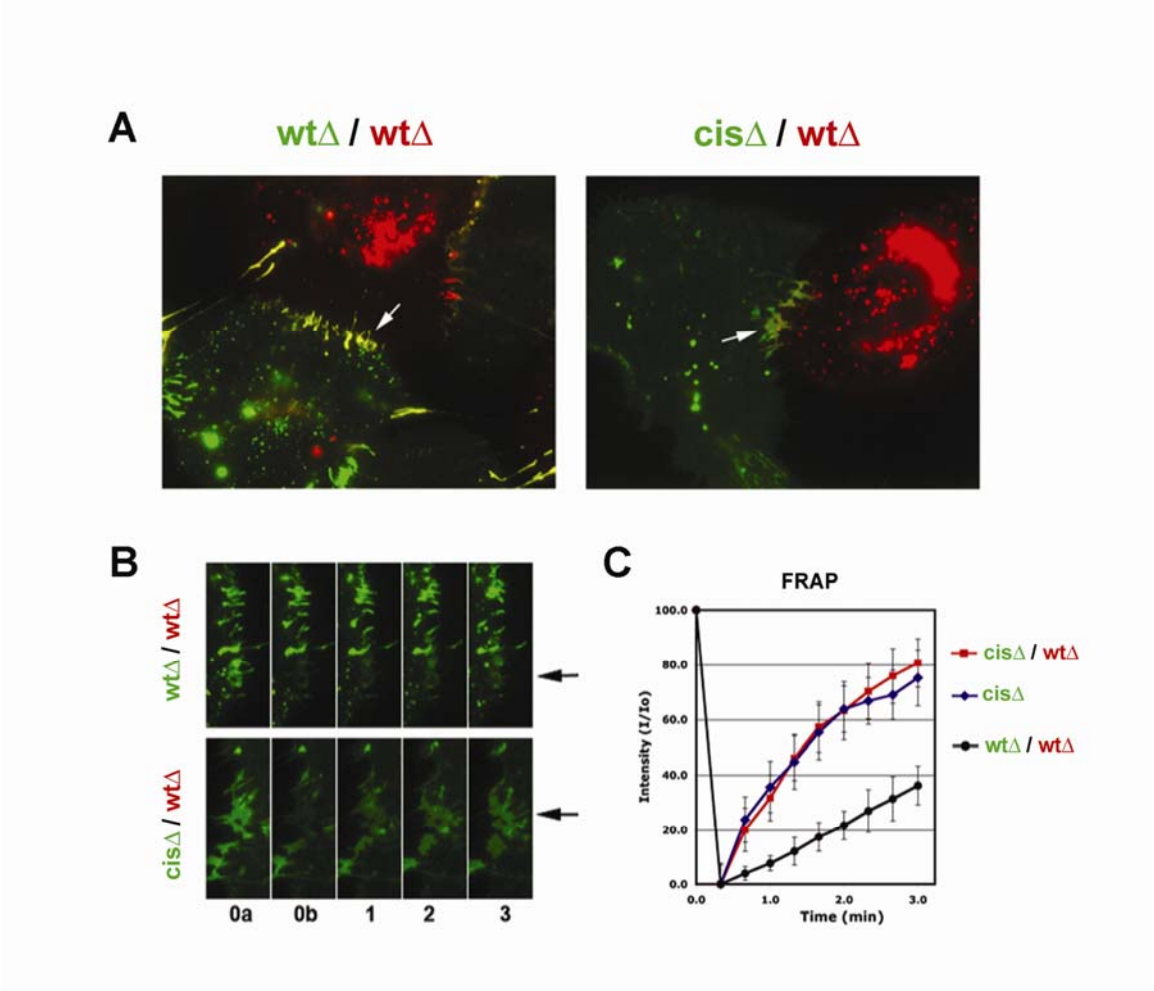

Figure S5. Mixed cell junctions between wild-type and *cis* mutant A431 transfectants (related to Figure 6)

(A) To determine whether adherens junctions can be stabilized by cadherin *cis* interactions in only one of the contacting cells, we co-cultured A-431 cells expressing Ecad^{wt}Δ–Dendra or Ecad^{cis}Δ-Dendra with cells expressing an equivalent construct tagged instead with the red fluorescent protein mCherry (Ecad^{wt}Δ-Cherry). Live imaging of co-cultures showed that the *cis* interface mutant could form mixed contacts (yellow) in *trans*

with its wild-type counterpart on neighboring cells (CisΔ/WtΔ), but that cadherins in these junctions appeared more dispersed than in contacts between cells expressing wild-type constructs (WtΔ/ WtΔ). **(B)** Time-lapse showing green Dendra fluorescence recovery (FRAP) in the mixed junctions. 5μm regions of mixed junctions (arrows in panel A) were photobleached for FRAP experiments to measure turnover of the transfected cadherin. Numbers show minutes after photobleaching. Recovery is more rapid in the mixed CisΔ/WtΔ junctions**. (C)** Graph of FRAP in mixed junctions (WtΔ/WtΔ and CisΔ/WtΔ) and in homotypic junctions between *cis* mutants (CisΔ). Mixed CisΔ/WtΔ junctions show similarly rapid dynamics to homogeneous cisΔ junctions.

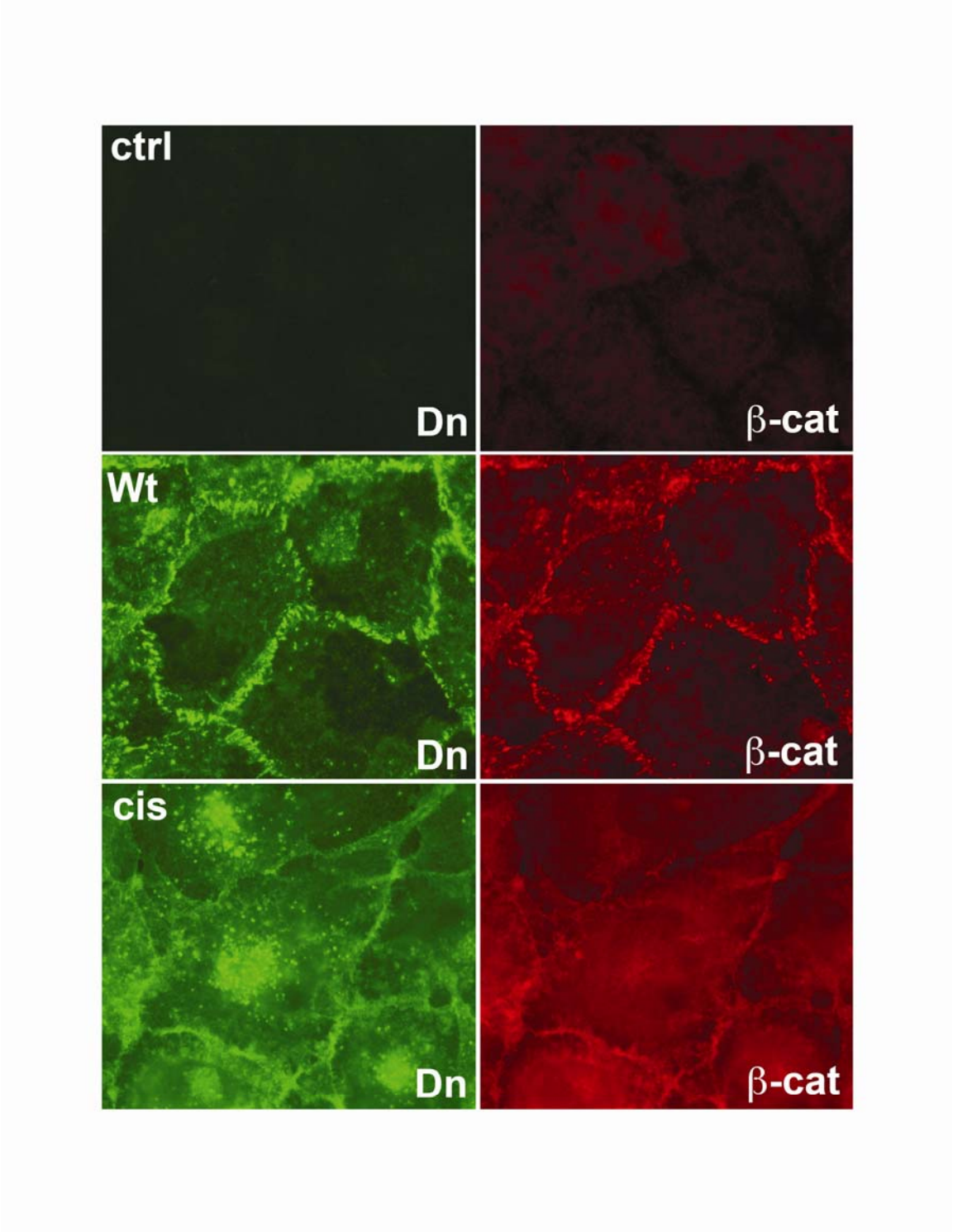

Figure S6. Effect of *cis* mutation on full-length E-cadherin clustering (related to Figure 7)

Parental, cadherin deficient A-431D cells (Ctrl) or their derivates expressing full-length Dendra2-tagged human E-cadherin (Wt) or its V81D V175D *cis* mutant (Cis) were double stained using rabbit anti-Dendra2 antibody against the recombinant cadherins (Dn, green) and mAB against β-catenin (β-cat, red). Note the complete absence of plasma membrane anti-β-catenin staining in parental cells indicating a lack of endogenous

cadherin in these cells. Note also that inactivation of the cis interface nearly completely abolishes cadherin clustering into adherens junctions.

Table S1. **Shared interfaces in crystal lattices of E, N and C-cadherin (related to Figure 2)**

		E-cadherin	**N-cadherin**	**C-cadherin**[1]
***Trans* interface**	Symmetry Operation	x, y, z	x, -y, -z	-x, y, -z
	Buried Surface Area (Å^2)	1612	1723	1690
	Number of Residues	42	46	46
***Cis* interface**	Symmetry Operation	x-½, y-½, z (A)[2] x-½, y+½, z (B)	x-½, y+½, z	x-½, y-½, z
	Buried Surface Area (Å^2)	1356 (A) 1186 (B)	1120	1153
	Number of Residues	48 (A) 45 (B)	39	41

1. Boggon *et al* (2002), PDB 1L3W
2. Interfaces for chains A and B in the E-cadherin asymmetric unit listed separately

Table S2. Dissociation constants (K_D) for homodimerization of cadherin ectodomain fragments determined by equilibrium analytical ultracentrifugation (related to Figure 4)

Protein	**Description**	**K_D (μM)**
N-cadherin EC1-5		
Wild-type (WT)	Wild-type	7.8 ± 0.3
E-cadherin EC1-5		
Wild-type (WT)	Wild-type	109 ± 6
V81D	*Cis*-interface mutant	137 ± 2

L175D	*Cis*-interface mutant	91.5 ± 1
V81D+L175D	*Cis*-interface mutant	118 ± 3.5
W2A+K14E	*Trans*-interface mutant	(Monomer)
E-cadherin EC1-2		
Wild-type (WT)	Wild-type	96.5 ± 10.6[b]
V81D	*Cis*-interface mutant	141 ± 8
L175D	*Cis*-interface mutant	119 ± 10
V81D+L175D	*Cis*-interface mutant	150.5 ± 7.5

[a] Errors indicate data range from two or more experiments.
[b] reported in Katsamba *et al* 2009