Supplementary information for:

Molecular design principles underlying β-strand swapping in the adhesive dimerization of cadherins

Jeremie Vendome [1,2,3,5], Shoshana Posy [1,2,3,5,6], Xiangshu Jin,[1,3] Fabiana Bahna [1,3], Goran Ahlsen [1], Lawrence Shapiro [1,4*] and Barry Honig [1,2,3,*]

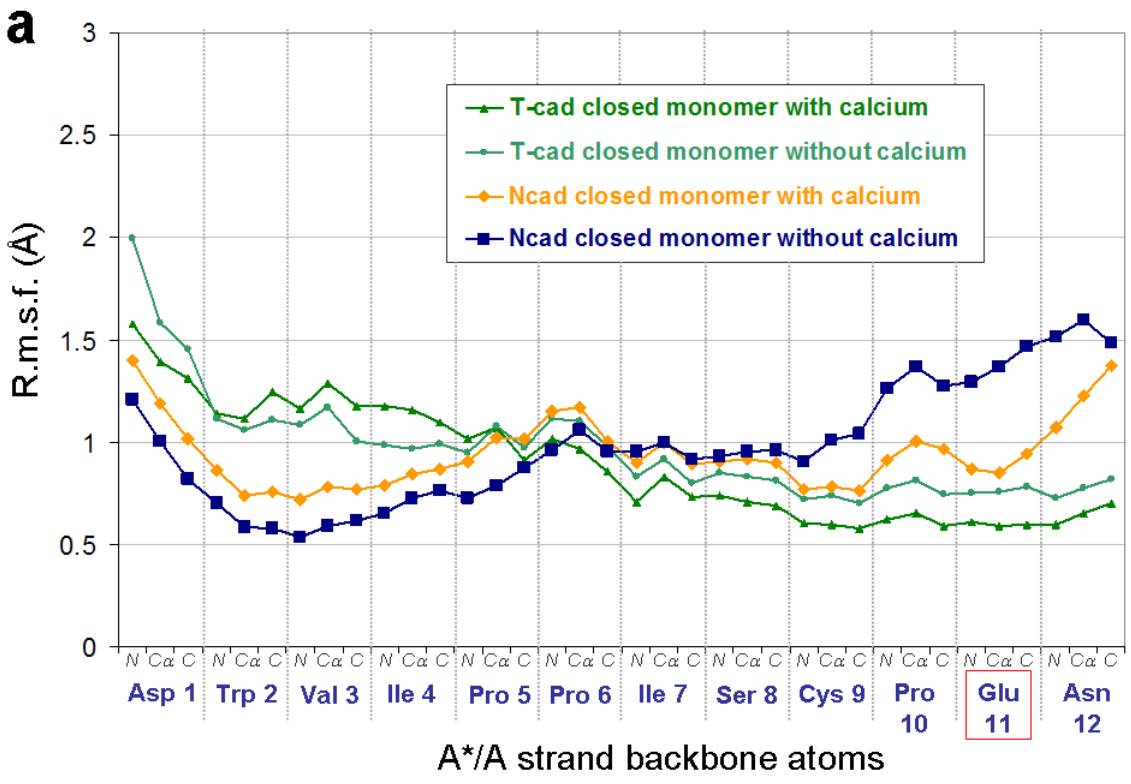
[1]Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032 USA, [2]Center for Computational Biology and Bioinformatics, Columbia University, Room 815, 1130 St. Nicholas Avenue, New York, NY 10032 USA, [3]Howard Hughes Medical Institute, [4]Edward S. Harkness Eye Institute, Columbia University, New York, NY 10032 USA, [5]These authors made equal contributions, [6]Present address: Computer-Assisted Drug Design, Bristol-Myers Squibb Company, Princeton, NJ 08543

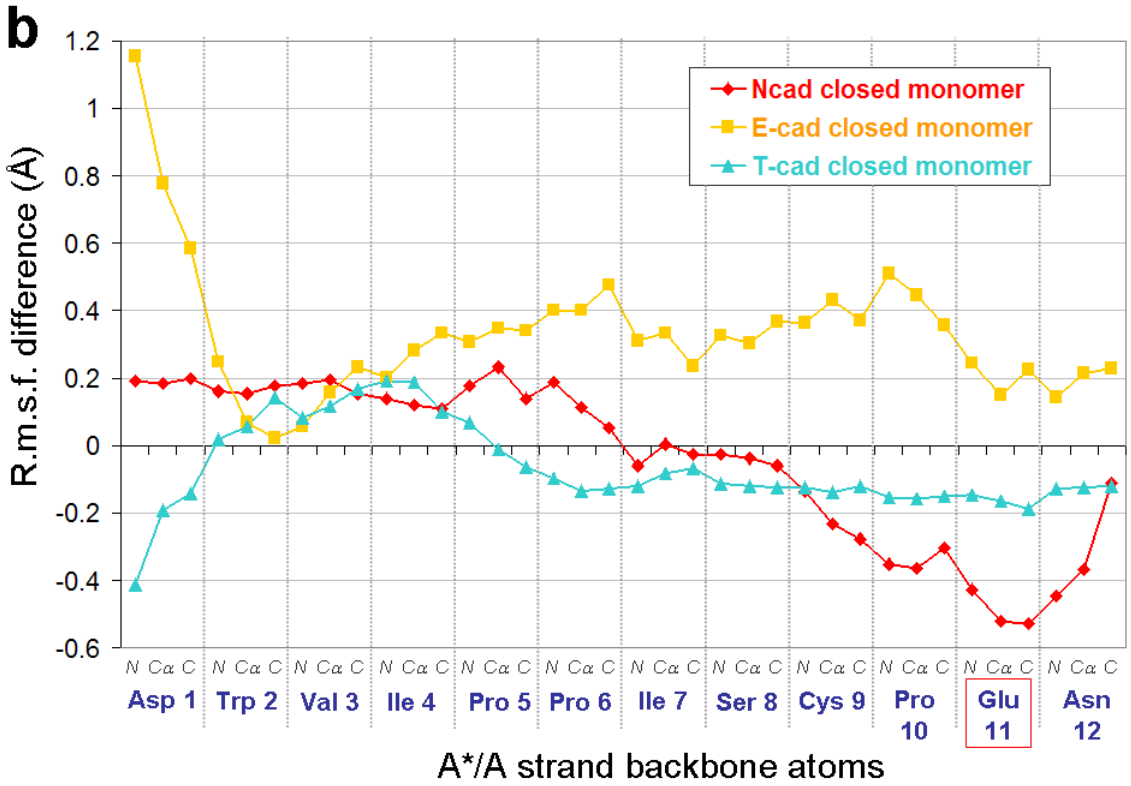
Supplementary Figures 1-7

Supplementary Methods

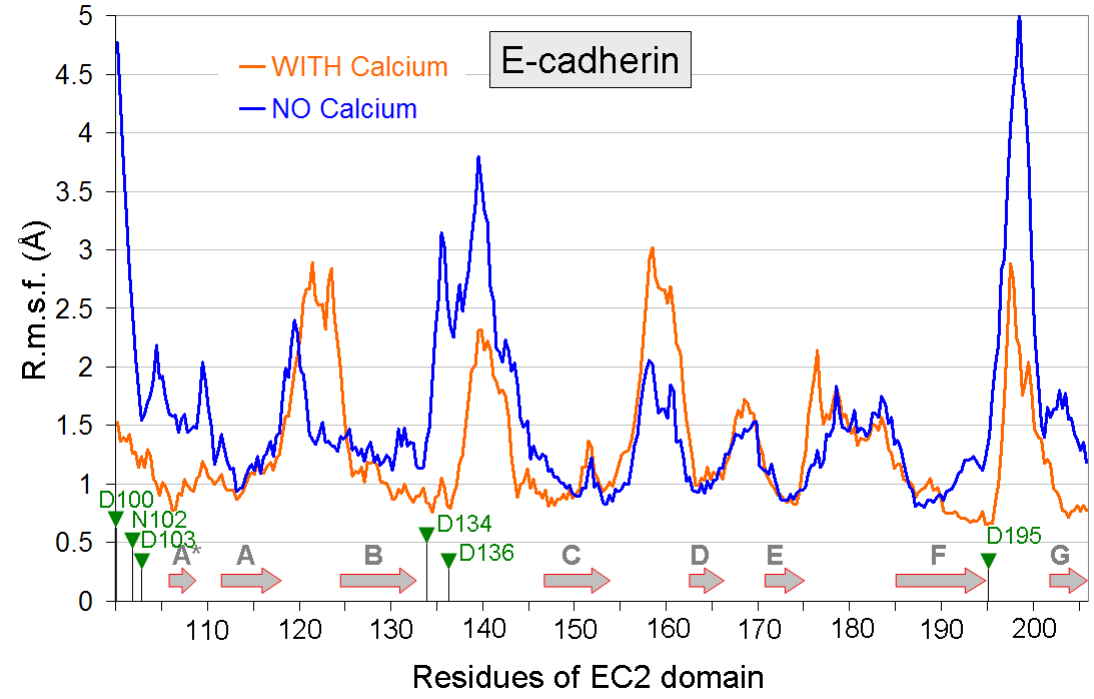
[*] Correspondence: lss8@columbia.edu, bh6@columbia.edu

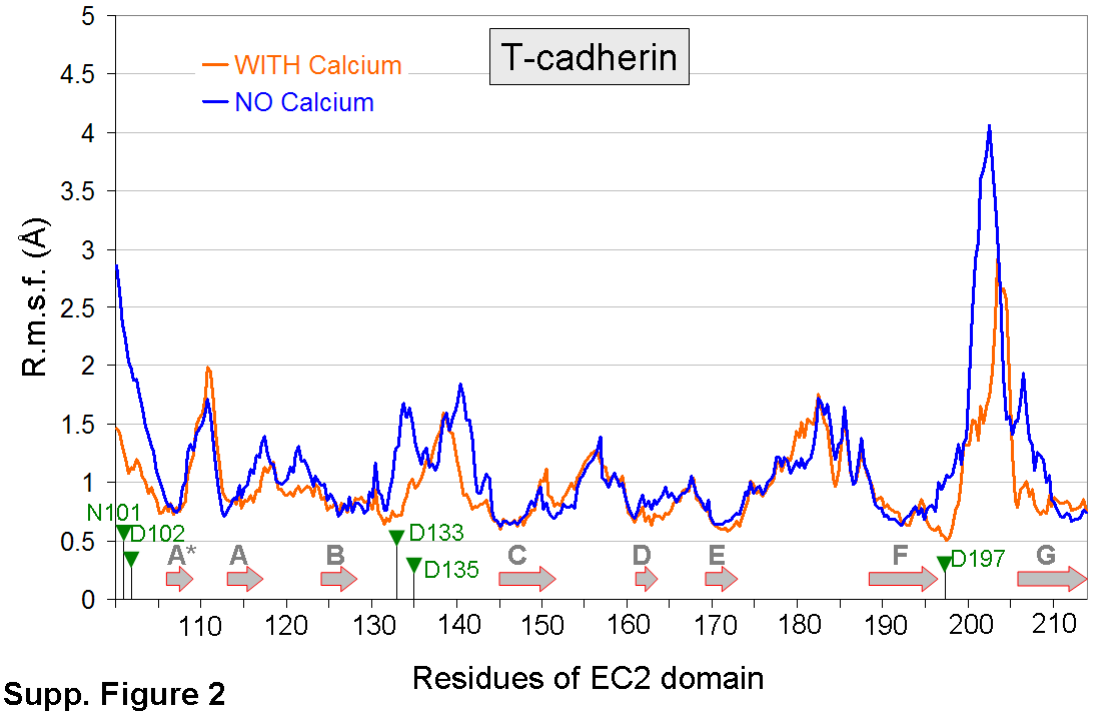
Running title: Design principles of cadherins adhesive swapping


a
R.m.s.f. (Å)
3
2.5
2
1.5
1
0.5
0
T-cad closed monomer with calcium
T-cad closed monomer without calcium
Ncad closed monomer with calcium
Ncad closed monomer without calcium
N Cα C
Asp 1
Trp 2
Val 3
Ile 4
Pro 5
Pro 6
Ile 7
Ser 8
Cys 9
Pro 10
Glu 11
Asn 12
A*/A strand backbone atoms


b
R.m.s.f. difference (Å)
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
Ncad closed monomer
E-cad closed monomer
T-cad closed monomer
N Cα C
Asp 1
Trp 2
Val 3
Ile 4
Pro 5
Pro 6
Ile 7
Ser 8
Cys 9
Pro 10
Glu 11
Asn 12
A*/A strand backbone atoms

Supplementary Figure 1: Root mean-square fluctuations of the A*/A strand backbone atoms of N-cadherin closed monomer during the MD simulations in the presence and absence of Ca^{2+}. (**a**) R.m.s.f. of the A*/A strand backbone atoms (N, Cα and C atoms of residues 1-12) of N-cadherin and T-cadherin in the presence or the absence of Ca^{2+}. The color code is indicated in the panel on the graph. (**b**) Difference between the R.m.s.f. in the presence of Ca^{2+} and in the absence of Ca^{2+} for N-, E- and T-cadherin closed monomers. The color code is indicated in the top left panel in on the graph.

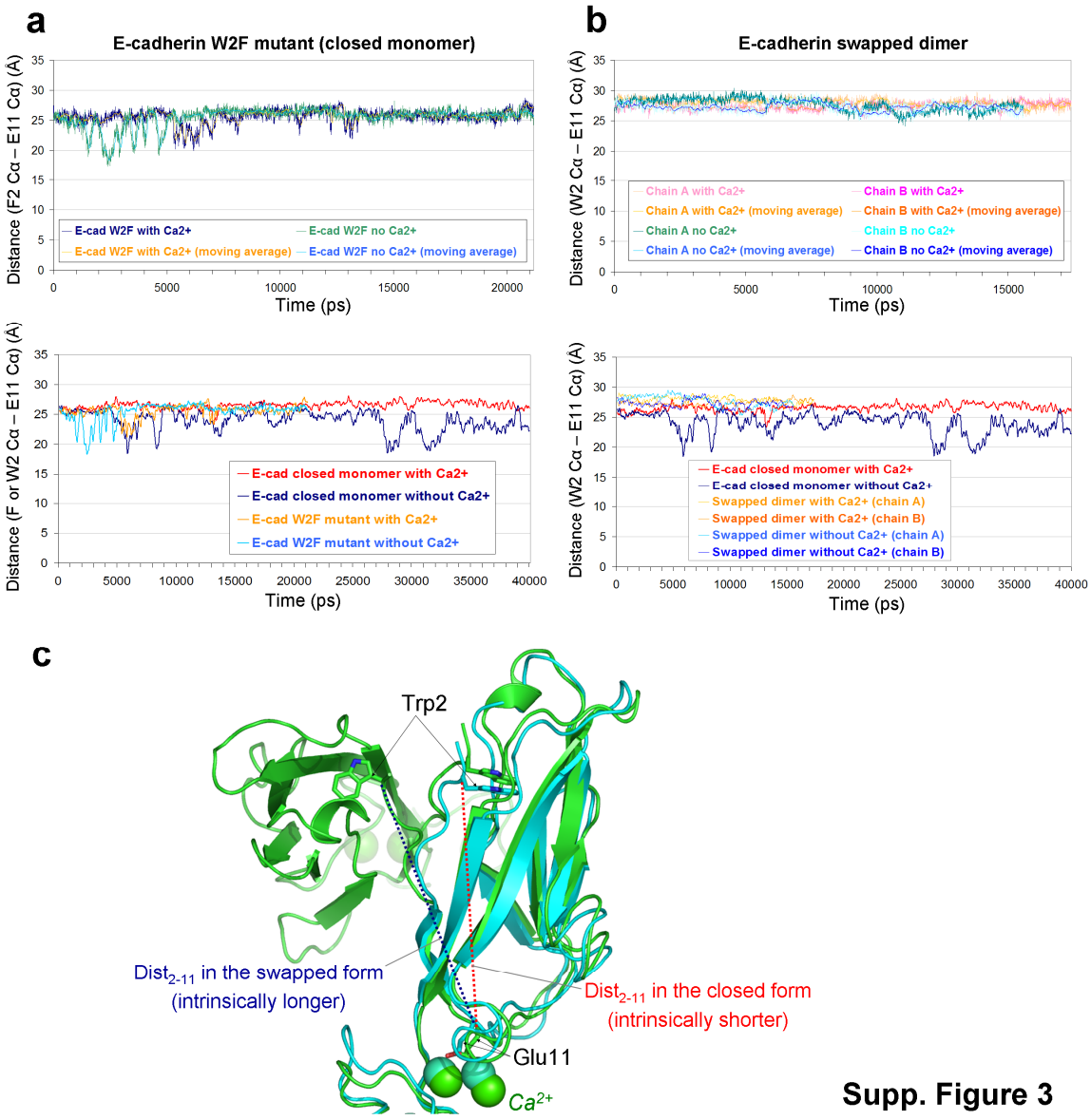




Supp. Figure 2

Supplementary Figure 2: Root mean-square fluctuations of the EC2 domain

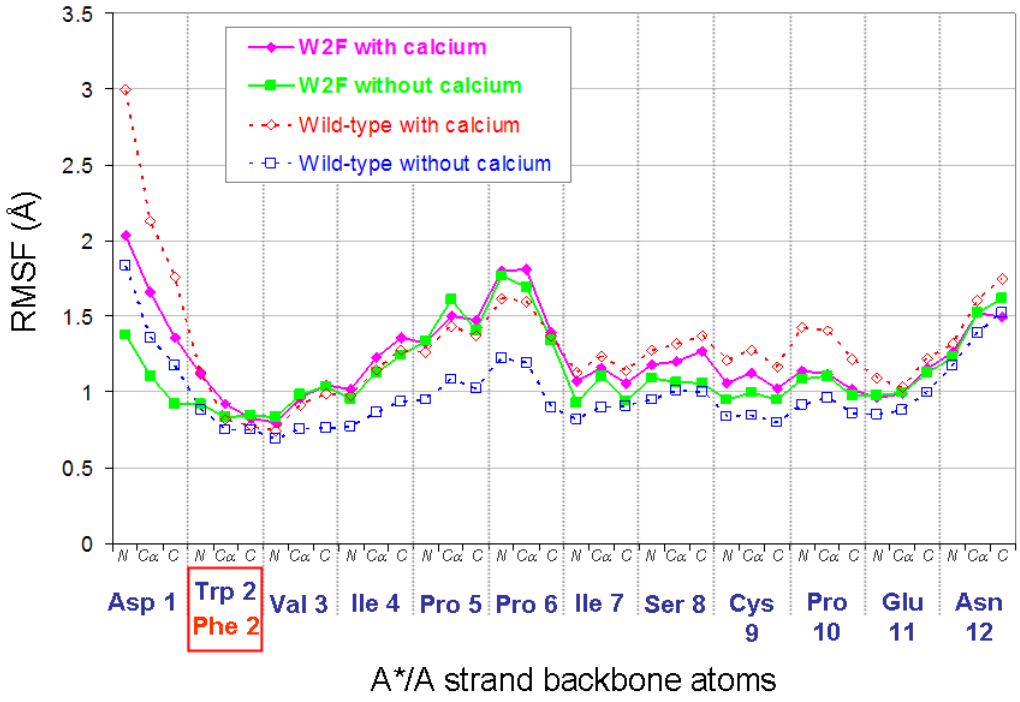
backbone atoms during the equilibrium MD simulations. (Top) R.m.s.f. observed in the closed E-cadherin monomer simulations. (Bottom) R.m.s.f. observed in the closed T-cadherin monomer simulations. In both graphs, the orange trace corresponds to simulations in presence of Ca^{2+}, and the blue one to simulation in the absence of Ca^{2+}. Secondary structure elements are represented by grey arrows and the positions of the Ca^{2+}-binding residues are indicated by green triangles.



Supp. Figure 3

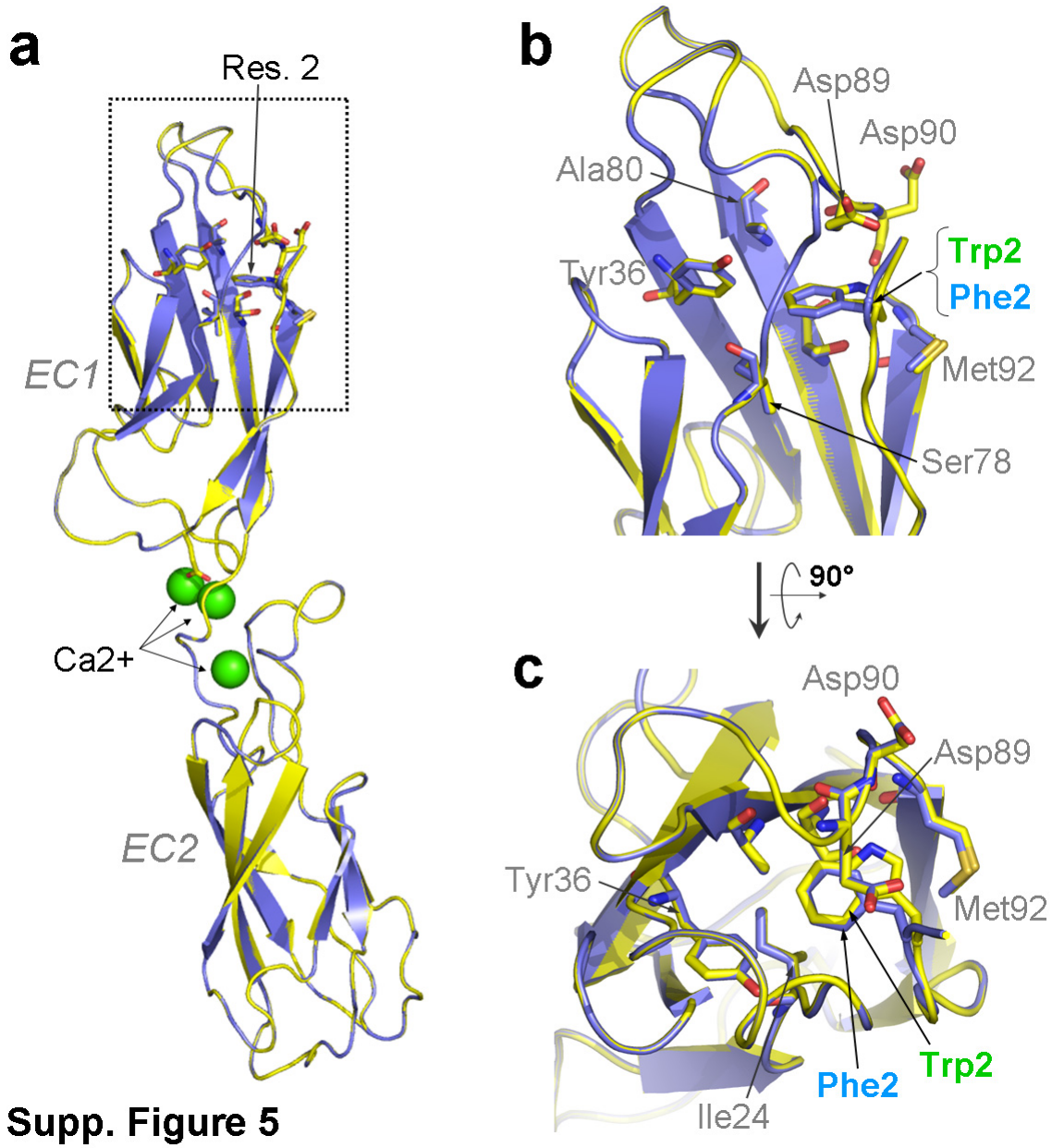
<u>Supplementary Figure 3:</u> Distance between the Cα atoms of residues 2 and Glu11 ($dist_{2-11}$) during the MD simulations. $Dist_{2-11}$ as a function of time during the MD simulations of E-cadherin W2F mutant in its closed conformation (**a**), and E-cadherin swapped dimer (**b**). For panel (**a**) and (**b**), the top graph represents $dist_{2-11}$ in the simulation in the presence and in the absence of Ca^{2+}, together with the moving average (over 20 values) for each curve. The bottom graph of each panel represents the same curves together with the curves obtained for wild type E-cadherin in its closed conformation, for a better comparison. For more clarity, only the moving averages (over 20 values) have been represented in the bottom graphs. The color code is indicated in the panel on each graph.

(**c**) Superposition of the structures of the closed (cyan) and swapped (green) conformations of E-cadherin. The distance used to monitor the length of the A*/A strand ($dist_{2\text{-}11}$) is represented for each conformation. It appears that $dist_{2\text{-}11}$ is, for geometrical reasons, intrinsically longer in the swapped conformation, and that this longer distance does not, in that particular case, reflect a more important "stretch" of the A*/A strand relative to the closed conformation.



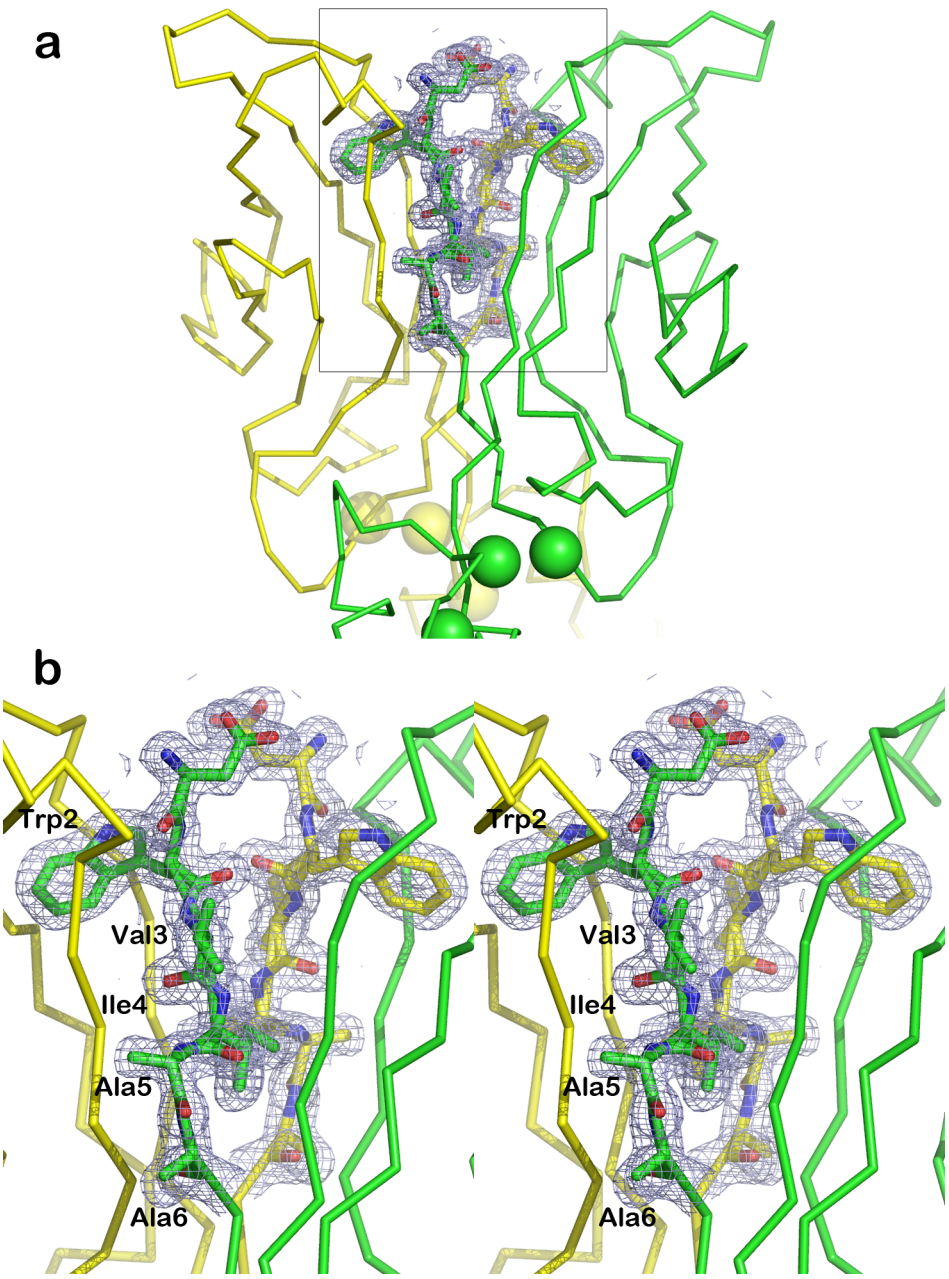
Suppl. Figure 4

Supplementary Figure 4: Effect of the W2F mutation on R.m.s.f. of the A*/A strands during the equilibrium MD simulations. The graph shows the R.m.s.f. of the N-terminal backbone atoms (residues 1-12) during the MD simulations of the W2F mutant closed monomer in presence (magenta) or in absence (green) of calcium. For a better comparison, the results obtained for the wild-type closed monomer with or without calcium are reported in dashed lines (respectively in red and blue).

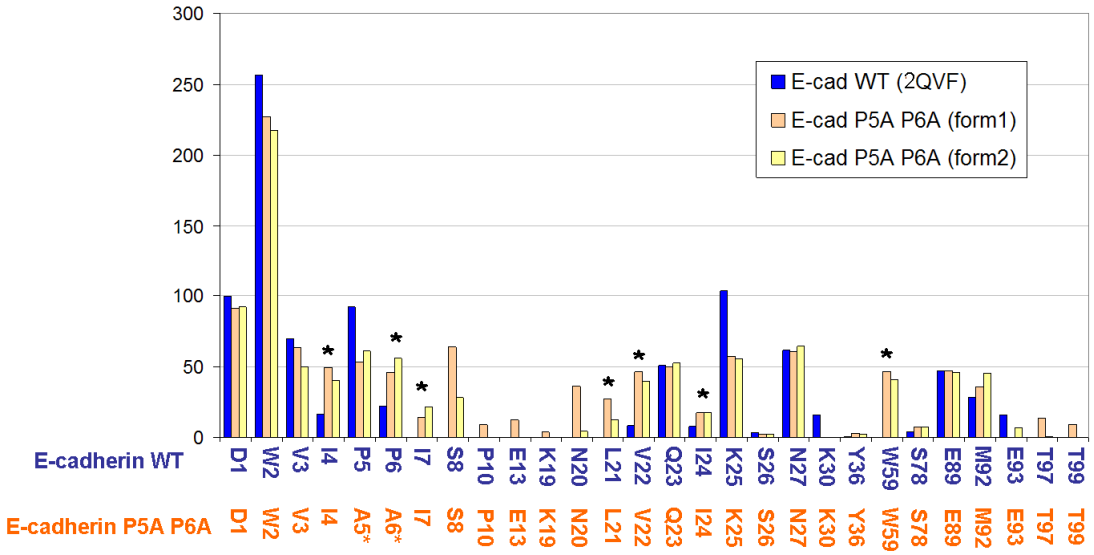


Supp. Figure 5

<u>Supplementary Figure 5</u>: Superposition of the wild type (yellow) and the W2F mutant (blue) E-cadherin closed monomer structures that were used as starting conformations in the MD simulations. Residues Trp2 (in the wild type) and Phe2 (in the W2F mutant) are represented, together with the pocket residues with which they are in contact (Ile24, Tyr36, Ser78, Asp90 and Met92). (**a**) Global view of the EC1-2 domains. (**b**) and (**c**) Zoom into the residue 2 docking pocket, at the top of the EC1 domain. (**b**) Detailed view from the side. (**c**) Detailed view from the top.



Supplementary Figure 6: Simulated annealing Fo-Fc electron density map of the E-cadherin P5A P6A mutant dimer, with residues 1-6 omitted in calculation, contoured at 3σ. (**a**) Overview of the dimerized EC1 domains, with the two protomers represented in yellow and green Cα trace, and the electron density shown for residues 1-6. (**b**) Close up stereoview of the electron density map of residues 1-6.



Supplementary Figure 7: Buried surface are by residue at the swapped interface of E-cadherin wild type (in blue) and in the two forms of E-cadherin P5A P6A mutant (in orange and yellow). The stars denote the positions where the mutant buries more hydrophobic surface than the wild type.

Supplementary Methods:

Structures used as starting point for the different MD simulations

E-cadherin swapped dimer conformation

For this conformation, the crystal structure of mouse E-cadherin EC1-2 dimer (residues 1-213) with swapped A-strands was used (2QVF.pdb [1]). This structure is very similar to the human E-cadherin EC1-2 swapped dimer structure (2072.pdb [2]) (RMSD of 0.68Å for the backbone atoms) and was chosen for purposes of consistency with the structure of the E-cadherin closed monomer conformation (1FF5.pdb, see below) used for our simulations and with all the E-cadherin EC1-2 constructs used for the AUC experiments, which are all from mouse.

E-cadherin closed monomer conformation

This refers to a structure derived from the crystal structure of mouse E-cadherin EC1-2 closed monomer (1FF5.pdb [3]). We corrected the bias introduced by the presence of an extra N-terminal methionine in the crystal structure (which prevents the formation of the crucial NH3-term./Glu89 salt-bridge [4]) by removing this residue and carrying out a local minimization with a constraint on the distance between the newly generated NH3-terminal group and the carboxyl group of Glu89 (harmonic constraint with a minimum at 5.0 Å between the N atom of the NH3-terminal group and the CG atom of Glu89). We verified that a proper salt-bridge was formed between these two groups in our final structure. Moreover, the nonessential residues 214-218 have also been removed from closed monomer crystal structure (1FF5.pdb) in order to be consistent with the swapped dimer structure which contains only residues 1-213 (see above).

N-cadherin closed monomer conformation

A homology model based on the E-cadherin closed monomer structure described just above was made for N-cadherin EC1-2 domains (residues 1-215), using Nest[5] (57% sequence identity between the two EC1-2 fragments). Both loops and regions with secondary structure were refined using Nest's refining mode #4. The model is accessible online at http://luna.bioc.columbia.edu/~yw2286/pdb/ncad_closed_confo_model.pdb.

T-cadherin closed monomer conformation

The crystal structure of Chicken T-cadherin EC1-2 domains (3K5S.pdb [6]) was used.

E-cadherin W2F mutant closed monomer conformation

The side chain of residue Trp2 was changed to a Phe and a local energy minimization was carried out so as to generate a starting conformation for MD simulations. The local minimization consisted in two rounds of 350 minimization steps (50 steepest descent steps followed by 300 conjugate gradient steps) where only the atoms within 4 Å or 6 Å from residue 2 were respectively left without constraint in the first and second round.

Closed conformations without calcium

The structures of the three closed monomer conformations described above (E-cadherin wild-type and W2F mutant and wild-type T-cadherin) were used to generate the respective closed conformation structures without Ca^{2+} ions. In each structure the crystallographically resolved Ca^{2+} ions were removed, and a local energy minimization

was performed. Three rounds of 350 minimization steps (50 steepest descent steps followed by 300 conjugate gradient steps) were performed. Only atoms within 4 Å, 6 Å or 8Å from any Ca^{2+} coordinating residue were respectively left without constraints during the first, second and third minimization rounds.

Supplementary references

1. Harrison, O.J. et al. Two-step adhesive binding by classical cadherins. *Nat Struct Mol Biol* **17**, 348-57 (2010).
2. Parisini, E., Higgins, J.M., Liu, J.H., Brenner, M.B. & Wang, J.H. The crystal structure of human E-cadherin domains 1 and 2, and comparison with other cadherins in the context of adhesion mechanism. *J Mol Biol* **373**, 401-11 (2007).
3. Pertz, O. et al. A new crystal structure, Ca2+ dependence and mutational analysis reveal molecular details of E-cadherin homoassociation. *EMBO J* **18**, 1738-47 (1999).
4. Harrison, O.J., Corps, E.M. & Kilshaw, P.J. Cadherin adhesion depends on a salt bridge at the N-terminus. *J Cell Sci* **118**, 4123-30 (2005).
5. Petrey, D. et al. Using multiple structure alignments, fast model building, and energetic analysis in fold recognition and homology modeling. *Proteins* **53 Suppl 6**, 430-5 (2003).
6. Ciatto, C. et al. T-cadherin structures reveal a novel adhesive binding mechanism. *Nat Struct Mol Biol* **17**, 339-47 (2010).