PNAS
www.pnas.org

Supplementary Information for

Structural basis for *Clostridium perfringens* enterotoxin targeting of claudins at tight junctions in mammalian gut

Alex J. Vecchio[a,b,1], Sewwandi S. Rathnayake[a], and Robert M. Stroud[c,1]

[a]Department of Biochemistry, University of Nebraska−Lincoln, NE 68588

[b]Nebraska Center for Integrated Biomolecular Communication, University of Nebraska−Lincoln, NE 68588

[c]Department of Biochemistry and Biophysics, University of California, San Francisco, CA 94158

[1]To whom correspondence may be addressed.

Email: stroud@msg.ucsf.edu and avecchio@unl.edu

ORCID: 0000-0002-4222-7874 [Vecchio, AJ], 0000-0003-2083-5665 [Stroud, RM]

This PDF file includes:

Supplementary Information (SI)

SI Results and Discussion

Structural Determination and Analysis of the hCLDN-4/cCpE Complex. hCLDN-4 was recombinantly expressed and isolated as detailed in **SI Methods**. The structure of the hCLDN-4/cCpE complex was determined to 3.37 Å from a single crystal dataset in space group $P2_12_12_1$, with cell dimensions of 70.4, 116.2, and 118.1 Å. Phases for the complex were determined with molecular replacement (MR) by independent placement of each protein using a poly-alanine model of hCLDN-9 and the cCpE from PDB ID 6ov2 (1). Post-MR maps allowed modelling of 87% of full-length hCLDN-4 and 100% of cCpE with no structural breaks in non-terminal regions. The asymmetric unit contains a single hCLDN-4/cCpE complex, like our previous structures of hCLDN-9 bound to cCpE (1). This single complex within the crystal differs from other claudin/cCpE complex structures, which crystallize in non-physiological antiparallel multimeric arrays that influence intermolecular interactions and structural analysis (**Figure S5**) (2-4). Data collection and refinement statistics are summarized in **Table S1**.

Structural analysis of the hCLDN-4/cCpE complex reveals a plethora of intermolecular interactions occur between proteins. In ECS2, polar interactions (distances 2.4-4.0 Å) between five residues make eleven hydrogen bonds with cCpE, six which are mediated by main chain atoms (**Figure S6A**). Additionally, hCLDN-4 Asp146 and cCpE Arg227, and Arg158 of hCLDN-4 and Asp225 of cCpE form two ionic interactions and help bond the two molecules together. Non-polar interactions (distances 3.0-4.0 Å and PDBePISA buried surface areas >50 $Å^2$ (5)) within ECS2 are focused on Leu151 and Ala153, part of the $NPLVA^{153}$ motif. In ECS1, eight residues throughout the segment form polar interactions with cCpE exclusively through nine hydrogen bonds, one coming from a main chain carbonyl (**Figure S6B and S6C**). No ionic interactions exist in ECS1 and only Phe35 makes significant non-polar contacts. Generally, contacts between hCLDN-4 and cCpE are sparsely spread across β1-β4, ECL1, and TM2 in ECS1, while in ECS2 they are concentrated on ECH2 and ECL4.

Contradictory findings using various biochemical and cellular techniques have contended that one of two regions within ECS2 is solely responsible for driving cCpE binding to claudins. Previously, the $NP(V/L)(V/L)(P/A^{153})$ motif (6-8) or an ECS2 basic region that electrostatically interacts with an acidic cleft of cCpE (9) have been shown to be essential for complex association and cytotoxicity. The structure of the hCLDN-4/cCpE complex demonstrates nine residues in ECS1 and eight residues in ECS2 are involved in binding cCpE (**Figure S7**). Interactions between hCLDN-4 and cCpE are concentrated on ECS2 while they are sparsely spread across ECS1. Of the 69 total residues that constitute the two ECS, 48 (70%) reside in ECS1 and 21 (30%) reside in ECS2. Analysis of the intermolecular contacts between hCLDN-4 and cCpE reveals 38% of all ECS2 residues compared to only 18% of all ECS1 side chains interact with cCpE (**Figure S7A**). Despite having 2.3-fold fewer amino acids, ECS2 uses a greater percentage of residues to interact with cCpE compared to ECS1, creating a focused cCpE binding region. This focused group of interactions on the smaller ECS2 domain explain why $NP(V/L)(V/L)(P/A^{153})$ motif mutations alter cCpE binding.

In this report we explain through structure/function analyses how both classes of interactions facilitate hCLDN-4/cCpE complex formation and preservation. Leu151 in the $NPLVA^{153}$ motif functions to slow cCpE dissociation (**Table 1**) by accessing a surface pocket on cCpE and forming non-polar side chain and polar backbone interactions (**Figures S6 and S11**), while ECS2 electrostatics outside of ECL4 facilitate complex association via a hydrogen bond network that comprises the cCpE-binding motif (**Figure 3**). These findings allow us to speculate that both modes of interactions play specific roles in toxin binding—electrostatics drive association and the $NPLVA^{153}$ motif prevents dissociation. This idea would explain how in the absence of one mode—an $NPLVA^{153}$ motif in hCLDN-1 ($DPMTP^{154}$) and hCLDN-19 ($NPSTP^{154}$)—that these subtypes can still bind toxin and render cells susceptible to CpE cytotoxicity (**Figure 4I and** (1)). In combination with our finding that residues that interact with cCpE are extensively spread across the hCLDN-4 ECS (**Figure S7**), a two-mode binding process provides further proof why ECS2 residues alone

cannot distinguish CpE receptors. In summary, our analyses reveal that both the NPLVA^{153} motif and electrostatic interactions outside of it play supportive roles in ECS2 binding to cCpE, that these two binding modes direct distinct binding processes, that both ECS support cCpE binding, and that ECS2 residues alone cannot be used to distinguish CpE receptors from non-receptors. Our results unite two former disparate ideas that appeared conflicting, and resolve a more nuanced amino acid-level understanding of CpE targeting of claudins.

CpE Disruption of Claudin Cis/Trans Assemblies and Paracellular Pores. Between β1 and β2 of cCpE a ten residue α-helix forms from Asn210-Pro219 (**Figures 2B and S6B**) The Pro219 of cCpE comes within 3.2 Å of hCLDN-4 and interacts with Cys64 and Lys65, residues that form the Cys54–64 disulfide bridge and impart anion selectivity to hCLDN-4 (10). These interactions allow cCpE to affect the conformation of the hCLDN-4 β3 to β4 "little finger" that is tethered by the Cys54–64 disulfide bridge. Targeting of this hCLDN-4 region by cCpE is interesting because: i) the Cys54–64 disulfide is conserved in all claudins, making this interaction a potential claudin fold recognition site; and ii) Lys65 is the anion-selective residue in CLDN-4 and equivalent residue 65 side chains in other claudins impart ion-selectivity (10, 11). Pro219 and this helix in cCpE could therefore disrupt claudin paracellular pore formation through interactions in this region. Moreover, Pro219 comes within 3.8 Å of Gln63. This is notable because the equivalent residue in mCLDN-15, Asn61, has been shown to reside at the interface between cis interacting claudin homodimers (12). These observations lead us to conclude that Pro219 and the helix of cCpE could be used for claudin fold recognition in order to disable claudin cis and/or trans assemblies, resulting in losses to ion selectivity via paracellular pore disruption. The cCpE's targeting of hCLDN-4 regions used for these specific functions is worth further investigation in other subtypes.

Structural Comparisons of Claudins Bound to cCpE. To date, nine crystal structures of claudins have been determined, eight of which are in complex with cCpE (1-4, 13). The cCpE bound to hCLDN-4 exhibits no major conformational alterations compared to other structures and has root-mean-square deviations (RMSD) between Cα atoms that range between 0.4–0.7 Å (**Figure S8A**). Compared to other structures, the cCpE bound to hCLDN-4 is most similar to cCpE from the complex with hCLDN-9>full-length CpE>cCpE from the complex with mCLDN-19>cCpE from the complex with mCLDN-3>cCpE from the complex with T4 lysozyme-hCLDN-4. Superposition of the claudins from these structures show cCpE moves as a rigid-body with a rocking motion within the claudin hand (**Figure S8B**). These movements influence the conformations of claudin "fingers" and in turn alter the sizes and shapes of solvent exposed pockets at the interfaces of these complexes (**Figure S9**). We found that the hCLDN-4/cCpE interface has an area of 2663 Å^2 and that a solvent accessible pockets exists there and has a volume of 5281 Å^3 (**Figure 2C**). Solvent pockets are observed at the interfaces of other claudin/cCpE complexes, making them subtype-specific features useful for comparing cCpE interactions with claudins (**Figure S10**) (1). These structural overlays illustrate cCpE can adopt numerous conformational ensembles depending on the claudin it binds and underscores how the unique structures of the solvent accessible pockets at the protein interfaces act as subtype-specific features of claudin/cCpE complexes.

Previously, a crystal structure of a C-terminally truncated form of hCLDN-4 fused on the N-terminus to T4 lysozyme (T4L-hCLDN-4) and bound to cCpE was determined to 3.5 Å (4). This complex had a 1:1 hCLDN-4/cCpE stoichiometry but packed in an antiparallel tetrameric arrangement (**Figure S5**). Comparing T4L-hCLDN-4 and our hCLDN-4, the cCpEs have an RMSD of 0.84 Å between Cα atoms while the Cα RMSD between the CLDN-4s is 1.4 Å, indicating significant differences. Superposition of hCLDN-4 between structures shows conformational deviations are present in the ECS, especially in the loops encompassing ECL1-4 (**Figure S11A**). The crystal packing of T4L-hCLDN-4 reveals ECL3 of chain A interacts with antiparallel oriented claudins from chain C and E; and cCpE makes crystallographic and non-crystallographic symmetry contacts on four sides with T4L (chain C, E, and G) and another cCpE from chain H. It is likely these contacts limit cCpE's conformational sampling, influence its orientation, and the resulting movements of the two ECS of hCLDN-4. This would explain the 0.84 Å Cα deviations we observe between structures and why the intermolecular interactions differ. T4L-hCLDN-4 ECS2 makes three hydrogen bonds and one ionic interaction with cCpE, while ECS1 forms six hydrogen bonds

with cCpE. For the T4L-hCLDN-4/cCpE complex six residues in ECS1 and five residues in ECS2 contribute most to binding cCpE, which corresponds to 24% of all ECS2 residues and 12% in ECS1; much less than in our hCLDN-4/cCpE complex. Furthermore, structural comparison reveals the hydrogen bonding network of the cCpE-binding motif from hCLDN-4/cCpE (**Figure 3A**) is less extensive in T4L-hCLDN-4/cCpE (**Figure S11B**). In T4L-hCLDN-4/cCpE the ECL3 moves to contact neighboring antiparallel claudins; altering TM2, opening the solvent exposed pocket, and creating conformational differences in the bonding network side chains. These changes result in losses to cCpE interactions for Arg158 and Leu531. Arg158 makes no interactions with cCpE residues and Leu531 forms only five non-polar and one polar interaction versus eight non-polar and three potential polar interactions for Leu531 in our hCLDN-4/cCpE complex (**Figure S6A**). The other major difference between these two complexes is that Gln156 and the conformation of ECL4 is significantly altered by cCpE Asp225 in the T4L-hCLDN-4/cCpE complex. This causes cCpE Asp225 to be unable to ionically interact with Arg158, and Gln156 unable to form polar interactions with cCpE Arg227 or Ser313. In summary, the antiparallel crystal packing as well as the presence of T4L in the T4L-hCLDN-4/cCpE complex significantly hinders the conformational freedom of both hCLDN-4 and cCpE, trapping them in a state less conducive to tight association. The structure of the hCLDN-4/cCpE complex reported here reveals more numerous and unique intermolecular interactions that form at a multitude of structural elements, which we hypothesize to be a more accurate representation of *in vitro* and *in vivo* complexes.

The recently reported 3.6 Å crystal structure of a mCLDN-3/cCpE complex gives a glimpse into the associations between toxin and another high-affinity gut-expressing claudin (**Table 1**) (2). The cCpE and claudin conformations amid mCLDN-3 and our hCLDN-4 have RMSDs between Cα atoms of 0.62 Å and 1.8 Å, respectively. Differences in claudin conformations primarily arise in the loops encompassing ECL1-4. The cCpE binding pose alters the conformations of ECL1-4 in mCLDN-3 so that cCpE moves out of the claudin "palm" completely and pushes toward the tips of the fingers compared to hCLDN-4 (**Figure S11C**). The mCLDN-3 NPLVP^{152} motif accesses the pocket between β4 and β8–β9 of cCpE but does not penetrate it as deeply as the NPLVA^{153} motif of hCLDN-4 (**Figure S11D**). Two prolines in the mCLDN-3 NPLVP^{152} motif limit the flexibility of ECL4 prohibiting it from adapting to the cCpE surface groove when positioned like cCpE in hCLDN-4. These prolines align mCLDN-3 Leu150 with hCLDN-4 Pro150 instead of Leu151, preventing Leu150 from entering the cCpE groove. As hCLDN-3 has a NPVVP^{152} motif we speculate that the lower binding affinity for cCpE that we measured for it is a result of Val150 not being able to penetrate the cCpE surface groove as extensively as Leu150 of mCLDN-3 or Leu151 of hCLDN-4 (**Table 1**). In addition, the cCpE-binding motif hydrogen bonding network is severely disrupted due to a lack of cCpE contacts, the arginine stacking is lost and Arg227 of cCpE makes no ionic contacts to any mCLDN-3 side chain. The resulting cCpE pose decreases the claudin/cCpE interface area and creates a solvent accessible pocket at least twice as large in mCLDN-3 compared to hCLDN-4 (**Figures S10D and S10H**). This cCpE pose may be influenced by contacts with an antiparallel neighboring mCLDN-3/cCpE complex, where the ECS interact with TM1 and TM2 (**Figure S5C**). In structures of hCLDN-9 in complex with cCpE previously determined by our lab we observed two cCpE orientations with one having a larger solvent exposed pocket than the other—highlighting that multiple cCpE poses may exist when bound to claudins although they are difficult to trap individually using crystallography (1). Because hCLDN-4 and mCLDN-3 share 75% residue identity in the cCpE-binding motif we hypothesize that mCLDN-3 can form nearly identical interactions with cCpE, with subtle changes caused by the NPLVP^{152} motif, but that these interactions were not observed due to the steric influence of antiparallel molecules as well as the cCpE interactions with symmetry-related molecules in the crystal packing. The hypothesis that mCLDN-3 and hCLDN-4 can bind cCpE nearly identically will require validation via new structures where complexes do not pack in antiparallel orientations.

Prediction of NPLVA^{153} Structure and Leu^{151}Val Mutant Function in hCLDN-4. Structural comparisons of the mCLDN-3/cCpE and hCLDN-4/cCpE complexes reveal structural information about ECL4's NPLVA^{153} motif that are useful for interpreting our studies. We demonstrate the biophysical function of Leu151 in enterotoxin binding but have not postulated how the more hCLDN-3-like mutant, hCLDN-4$^{L151V/A153P}$, exhibits no significant loss in cCpE binding affinity compared to

wild type hCLDN-4 (**Table 1**) or how this mutant maintains longer complexes with cCpE compared to hCLDN-3 (**Figure 5B**). Sequence alignments show 70% identity between hCLDN-3 and -4 overall and 67% identity in the cCpE-binding motif (**Figures 6 and S3**). In ECS2, hCLDN-3 and -4 share only 63% identity. The two mutations in hCLDN-4$^{L151V/A153P}$ improves this identity to hCLDN-3 11%, to 74%, but there remains 26% divergence. This contrasts to comparisons to wild type hCLDN-4, where ECS identity between hCLDN-4$^{L151V/A153P}$ remains high at 89%. The increased sequence identity of hCLDN-4$^{L151V/A153P}$ to hCLDN-4 (89%) over hCLDN-3 (74%) explains why hCLDN-4$^{L151V/A153P}$ binds cCpE more like hCLDN-4 (**Table 1**). The structural basis of this is hard to surmise without structures of hCLDN-3 or hCLDN-4$^{L151V/A153P}$ in complex with cCpE. But, we hypothesize that the structure and conformation of ECS2, which is important for cCpE binding, differs between hCLDN-3 and -4 due to the 37% sequence divergence in ECS2 residues between subtypes. Indeed, comparison of hCLDN-4 and mCLDN-3 in complex with cCpE reveals the presence of structural perturbations in ECS2 that result in alternate binding modes of cCpE (**Figures S11C and S11D**). The structural orientation of the position three residues, Leu150/151, are misaligned in the one amino acid shorter mCLDN-3 compared to hCLDN-4. As hCLDN-3 is also one amino acid shorter and has a Val150 rather than Leu150 of mCLDN-3, these two changes would disallow the position three side chain of hCLDN-3 from aligning to and penetrating the cCpE surface groove deeply. Therefore, the higher structural homology between hCLDN-4$^{L151V/A153P}$ and hCLDN-4 would mean that the NPVVP^{153} motif of the former would align and potentially somewhat penetrate the cCpE surface groove (**Figure S11D**). We believe that this would explain why hCLDN-4$^{L151V/A153P}$ has only 5-fold lower affinity and 4-fold lower $t_{1/2}$ compared to hCLDN-4 (**Table 1**), and why its complex with cCpE is maintained outside of cells like other high-affinity binding claudins (**Figure 5B**). The loss of Leu151 decreases the binding affinity to cCpE, but residues outside of the NPVVP^{153} motif in the mutant hCLDN-4$^{L151V/A153P}$ can compensate for this loss better compared to hCLDN-3. This exercise shows how residues outside of the NPLVA^{153} and cCpE-binding motifs facilitate toxin binding by structural alteration of ECS2, that the structures of ECS2 likely differ between hCLDN-3 and -4, and further verifies that no residue solely directs claudin/cCpE complex formation.

SI Methods

hCLDN-4 and Enterotoxin Expression, Purification, and Complex Formation. *Homo sapiens* claudin-4 gene was subcloned into the pFastBac1 (ThermoFisher) vector and expressed in Tn5 (*Trichoplusia ni*, High Five, Expression Systems) cells with a C-terminal decahistidine tag preceded by a thrombin cleavage site (hCLDN-4-His$_{10}$). Tn5 cells in suspension were infected with baculovirus particles generated according to the manufacturer's protocol at a MOI of 0.2 and cultured for 48 hours at 27°C. Cells were harvested by centrifugation and the pellet resuspended in Lysis buffer containing 50 mM Tris pH 7.4, 200 mM NaCl, 1 mM PMSF, and EDTA-free SigmaFast protease tablets (Sigma). Cells were sonicated at 4°C, and the crude lysate was resuspended in Lysis buffer supplemented with 1 M NaCl, dounce homogenized, and membranes prepared via ultracentrifugation at 100,000 xg for 1 hour. Membrane pellet was resuspended in Lysis buffer, dounce homogenized, and solubilized with 1% (w/v) n-undecyl-β-D-maltopyranoside (UDM, Anatrace) and 0.04% cholesteryl hemisuccinate (CHS, Anatrace) overnight at 4°C. Insoluble material was removed by ultracentrifugation at 100,000 xg for 30 minutes, and the supernatant spiked with 15 mM imidazole, filtered through a 1.2 µm filter, then incubated with NiNTA resin (Sigma) for 12 hours at 4°C. Resin was captured and washed with Lysis buffer containing 500 mM NaCl, 25 mM imidazole, and 0.29% UDM, followed by a wash with Lysis buffer containing 300 mM NaCl, 50 mM imidazole, and 0.29% UDM. Lysis buffer containing 150 mM NaCl and 0.15% UDM was then used for incubation and protein release by thrombin, which occurred overnight at 4°C. Sample was captured and used for binding studies and cCpE complex formation.

Clostridium perfringens enterotoxin (CpE-His$_{10}$) and its C-terminal domain comprising residues 194-319 (cCpE-His$_{10}$) were subcloned into the pFastBac1 vector and expressed in Tn5 cells with a C-terminal decahistidine tag preceded by a thrombin cleavage site. Tn5 cells in suspension were infected with baculovirus particles at a MOI of 4.0 and cultured for 72 hours at 27°C. Cells were harvested by centrifugation and the pellet resuspended in Lysis buffer, as above. Cells were sonicated at 4°C, and the crude lysate was resuspended in Lysis buffer supplemented with 1 M

NaCl, dounce homogenized, and cell debris removed via ultracentrifugation at 100,000 xg for 1 hour. The supernatant was spiked with 15 mM imidazole, filtered through a 1.2 µm filter, then incubated with NiNTA resin for 12 hours at 4°C. Protein purification of CpE and cCpE was similar to hCLDN-3 apart from the omission of UDM and CHS in the buffers. Both toxins were captured, concentrated, 0.2 µm filtered, then loaded onto a Superdex 200 Increase 10/300 GL (GE Healthcare Life Sciences) equilibrated in SEC buffer (10 mM Hepes pH 7.4, 100 mM NaCl, and 4% glycerol). Peak fractions from size-exclusion chromatography (SEC) were collected, concentrated, then flash-frozen in liquid nitrogen and stored at -80°C until needed for complex formation, and binding and cytotoxicity assays.

Purified hCLDN-4 was mixed with cCpE at a ratio of 1:1.5 (moles:moles) and nutated overnight at 4°C. The protein complex was concentrated, filtered through a 0.2 µm filter and loaded onto a Superdex 200 Increase column equilibrated in SEC buffer with 0.1% UDM. Peak fractions from SEC were collected, concentrated to 15 mg/mL, and used for crystallization.

Crystallization and Structure Determination. Crystals were grown at 4°C using vapor diffusion by mixing 1 µL of complex with 1 µL of a mother liquor containing 100 mM DL-malic acid, 2-(*N*-morpholino)ethanesulfonic acid, and Tris base (1:2:2) pH 6.0 and 25% polyethylene glycol 1500. Rod-shaped crystals appeared at 12 hours and grew to variable lengths approaching 500 µm at the longest edge. Crystals were cryoprotected by removing excess mother liquor using a brief soak in Al's oil (50:50 paraffin/silicon, Hampton Research), then flash-frozen in liquid nitrogen. Data from a single crystal were collected at the Advanced Light Source beamline 8.3.1 at -170°C at a wavelength of 1.11583 Å, and processed using XDS (14) in space group $P2_12_12_1$. The structure was solved by molecular replacement (MR) in PHENIX (15) Phaser-MR (16) using an all poly-alanine version of hCLDN-9 and the cCpE from PDB ID 6ov2 (1), searching sequentially for hCLDN-9 then cCpE using two search components. To unbias the orientation of cCpE, it was manually moved away from poly-alanine hCLDN-9 5000 Å using Coot (17). One solution was found in space group $P2_12_12_1$ using a search for all possible solutions in the same point group, and high-quality maps were obtained post-MR. This initial incomplete model was used for subsequent model building and refinement in Coot (17) and Phenix (18), culminating in the final placement of residues 5-186 (87%) of hCLDN-4 and 191-319 (100%) of cCpE into the electron density. Final R factors for the hCLDN-4/cCpE complex is R_{work}=29.38% and R_{free}=29.57%. All structure figures were generated using PyMOL (19). Structural biology applications used in this project were compiled and configured by SBGrid (20). Data collection and refinement statistics are summarized in **Table S1**.

Claudin Expression, Mutagenesis, and Purification. The genes for human claudin-1, -3, and -4 and mouse claudin-3 and -4 were subcloned into pFastBac1 and expressed in Tn5 cells with a C-terminal decahistidine tag preceded by a thrombin cleavage site (CLDN-His_{10}) or with a C-terminal decahistidine tag preceded by eGFP (CLDN-eGFP-His_{10}). The hCLDN-$3^{V150L/P152A}$ mutant was created using the forward primer (5'-CTTCTACAACCCCCTGGTGGCCGAGGCGCAGAAGCG-3') and equivalent reverse compliment, while the hCLDN-$4^{L151V/A153P}$ mutant used the forward primer (5'- CTTCTACAATCCGGTGGTGCCCTCCGGGCAGAAGCG-3'). Protocols for expression and purification of all CLDN-His_{10} claudins were identical to those of hCLDN-4-His_{10}. Post-IMAC and thrombin digested CLDN-His_{10} samples were captured and used for cCpE and CpE binding studies. CLDN-eGFP-His_{10} were not purified and used for cytotoxicity, microscopy, and post-infection binding assays.

Pocket Detection and Measurements. The PDB coordinates for chain A (CLDN) and B (cCpE) for hCLDN-4 in complex with cCpE (7kp4), 6ov2, 6ov3, 6ake, 5b2g, and 3x29 were uploaded to Fpocket 1.0 (21) and PDBsum (22). For PDBSum the Interface Area and for Fpocket the "Real Volume (approximation)" is reported (**Figure S10H**). For 6ake Fpocket could not define the pocket accurately as it is so solvent exposed the probe likely could not distinguish borders. Comparison of pockets found from Fpocket (21) and MetaPocket 2.0 (23) were shown to be nearly identical. Pockets from MetaPocket 2.0 were read into PyMOL (19) and represented as spheres (**Figure S9 and S10**).

Affinity and Kinetic Measurements Using Bio-layer Interferometry (BLI). Briefly, cCpE-His_{10} and CpE-His_{10} were expressed in Tn5 cells, the cells were lysed and purified via NiNTA as described above. Instead of protease cleavage however the proteins were eluted in 500 mM imidazole, buffer exchanged into no imidazole buffer, concentrated to 1 mg/mL, then flash frozen in liquid nitrogen and stored at -80°C until use. CLDN-His_{10} claudins used in these analyses were purified as described above; post-cleaved protein of >90% purity was flash frozen in liquid nitrogen at 1 mg/mL and stored at -80°C until use.

BLI analyses were performed using an Octet© RED384 System (FortéBIO) with cCpE-His_{10} and CpE-His_{10} immobilized on NiNTA or Anti-His (HIS2) Biosensors (Dip and Read™) and free claudins. Assays were designed and setup using Data Acquisition and analyzed using Data Analysis versions 10.0 from Octet© Software. Experiments were run using 200 µL sample volume in black 96-well flat bottom microplates (Greiner Bio-One). Each experiment consisted of a sensor Wash (60 seconds) followed by toxin Loading (200 seconds), a Baseline measurement (120 seconds), and then an Association and Dissociation (300 seconds each) steps. Wash, Baseline, and Dissociation measurements were performed in SEC buffer with 0.1% UDM, and Baseline and Dissociation were performed in the same wells. Loading and Association measurements were performed in SEC buffer with 0.1% UDM with 10 µg/mL (100 nM) toxin and a range of 0-2000 nM claudins, respectively. A minimum of seven and maximum of eight concentrations were used for these experiments. For high-affinity claudins, a concentration range of 0-250 nM claudin was used, while for low-affinity binders, 0-2000 nM claudin was required. All experiments were run at 25°C using an acquisition rate of 5 Hz averaged by 20, in at least duplicate and quadruplicate for hCLDN-3 and -4. The effect of NaCl salt (5 and 100 mM) and temperature (25 and 37°C) on binding affinities of hCLDN-4 were tested prior to analyses and showed no significant differences.

For high-affinity binders, ligand sensors were subtracted from reference wells, the Y axis was aligned to the baseline for the last 5 seconds of measurement, and an inter-step correction was applied aligned to dissociation, then processed using Savitzky-Golay Filtering. No significant non-specific binding of claudins to naked NiNTA sensors were observed using </=250 nM claudin. For hCLDN-1 and -3, some non-specific binding to the NiNTA was observed due to the higher concentrations required (likely due to the basic claudin C-terminus), in which case a double reference subtraction was applied from reference sensors loaded with no toxin in SEC buffer. For hCLDN-1 we used both NiNTA and Anti-His (HIS2) biosensors and observed no interactions between hCLDN-1 and Anti-His antibody alone; a double reference subtraction was applied for these data. Processed data was analyzed using an association and dissociation step, using a 1:1 model with a global (full) fit applied. and the R_{max} was unlinked by sensor using a window of interest from 0-300 seconds. All curves were fit using this criteria. In instances where low concentrations resulted in negative responses, these wells were excluded from analysis and curve fitting. Data related to these experiments are shown in **Table 1** and **Figure S12**.

Insect Cell Cytotoxicity Assay and Microscopy. Recombinant baculoviruses containing CLDN-eGFP-His_{10} and hCLDN-$3^{V150L/P152A}$-His_{10} were produced using established methods. To each well of a 24-well cell culture plate, 500 µL of $0.6x10^6$ adherent Sf9 (*Spodoptera frugiperda*, Expression Systems) cells were added. After 1 hour at 27°C, virus containing claudins and a non-claudin plasma membrane-expressed protein (PMP) were added at a MOI of 1.0. Cultures were rocked gently then placed at 27°C. After 36 hours, 75 µg purified cCpE (10 µM) and 25 µg CpE (1.4 µM) in SEC buffer were added to the culture medium of two individual wells expressing claudins, PMP, and untreated Sf9 cells and placed back at 27°C. Twelve hours post-incubation with enterotoxins, Sf9 morphological damage was visualized using an EVOS FL Auto (ThermoFisher) microscope with 20x objective. Post microscope visualization, CpE cytotoxicity was quantified using a cell viability analysis by gently removing 50 µL of untreated and treated Sf9 cells then adding an equivalent volume of 0.04% trypan blue. After 5 minutes, 10 µL of stained cells were transferred to a hemacytometer and counted manually in duplicate by two individuals. Viability counts consisted of dividing the total number of live cells (unstained) by the total number of cells. Average cell viability for toxin treated cells were compared to three controls; 1) Sf9 cells with no baculovirus added in 3 wells, 1 well that was not treated with any toxin and 2 wells that were each treated with cCpE or

CpE; 2) Sf9 cells expressing PMP in 3 wells, 1 well that was not treated with any toxin and 2 wells that were each treated with cCpE or CpE; and 3) Sf9 cells expressing claudins in 3 wells, 1 of which was not treated with any toxin.

Post-Infection Binding Assay of Claudin/cCpE Complex. Claudin-expressing cells not treated with any toxin and cells treated with cCpE from the **Cytotoxicity Assay** were used for this analysis. 450 µL of Sf9 cells expressing CLDN-eGFP-His_{10} were harvested from a 24-well plate by gently pipetting and placed in a 1.5 mL microfuge tube (Beckman Coulter). Cells were pelleted at 13,000 rpm on a table top centrifuge for 2 minutes, culture media was removed and then the pellet was resuspended in 1 mL Lysis buffer with 1% (w/v) UDM. The solution was rotated for 1 hour at 4°C. After time, samples were spun via ultracentrifugation at 100,000 xg for 20 minutes using a TLA-45 rotor (Beckman Coulter) to remove insoluble debris. 300 µL of supernatant was loaded onto a Superdex 200 Increase column equilibrated in buffer consisting of 25 mM Tris pH 7.4, 150 mM NaCl, 0.1% UDM and 1% glycerol. Fluorescence of eGFP (λ_{ex}=488 nm and λ_{em}=509 nm) was monitored and the samples were subjected to fluorescence-detection SEC (FSEC) (24). Sf9 cells used in the assay had been infected with baculovirus for 48 hours and treated with 10 µM cCpE for 12 hours before harvesting. Total assay time from harvest to end of FSEC was 2 hours. Control samples representing CLDN-eGFP-His_{10} not treated with toxin from **Cytotoxicity Assay** were incubated with 10 µg cCpE (2 µM), incubated at 20°C for 2 hours, then assayed via FSEC to verify complexes from the cell-based assay represent intact claudin/cCpE complexes.

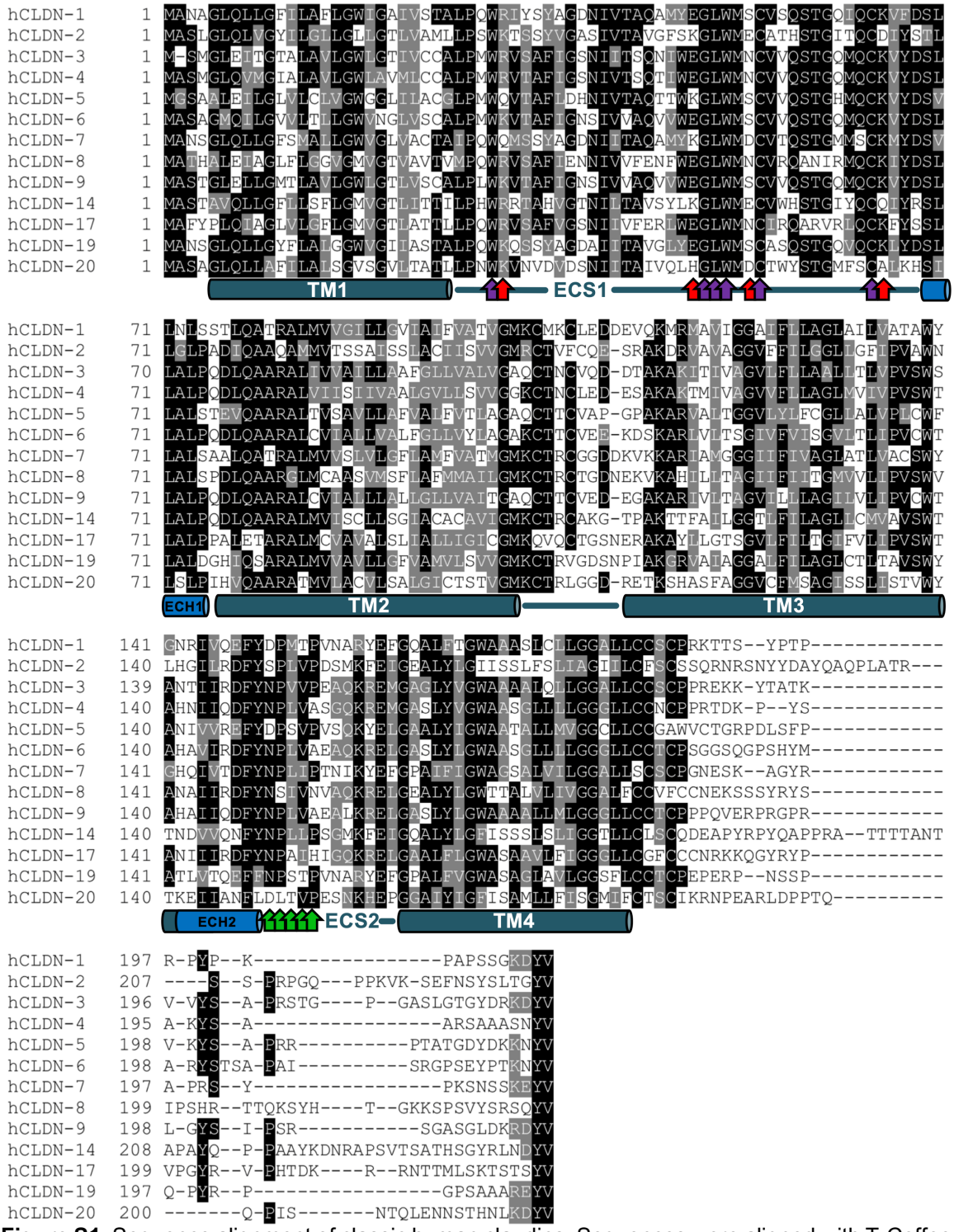


Figure S1. Sequence alignment of classic human claudins. Sequences were aligned with T-Coffee (25). Highlighted in the sequences with a cylinder helix are TM1-4 (teal), the consensus SLLALP[74] for ECH1 and ECH2 (blue); and noted with arrows are the claudin-defining sequence W[30]-GLW[51]-C[54]-C[64] (purple), the purported ion-selective pore lining residues (red), and the NPLVA[153] motif (green).

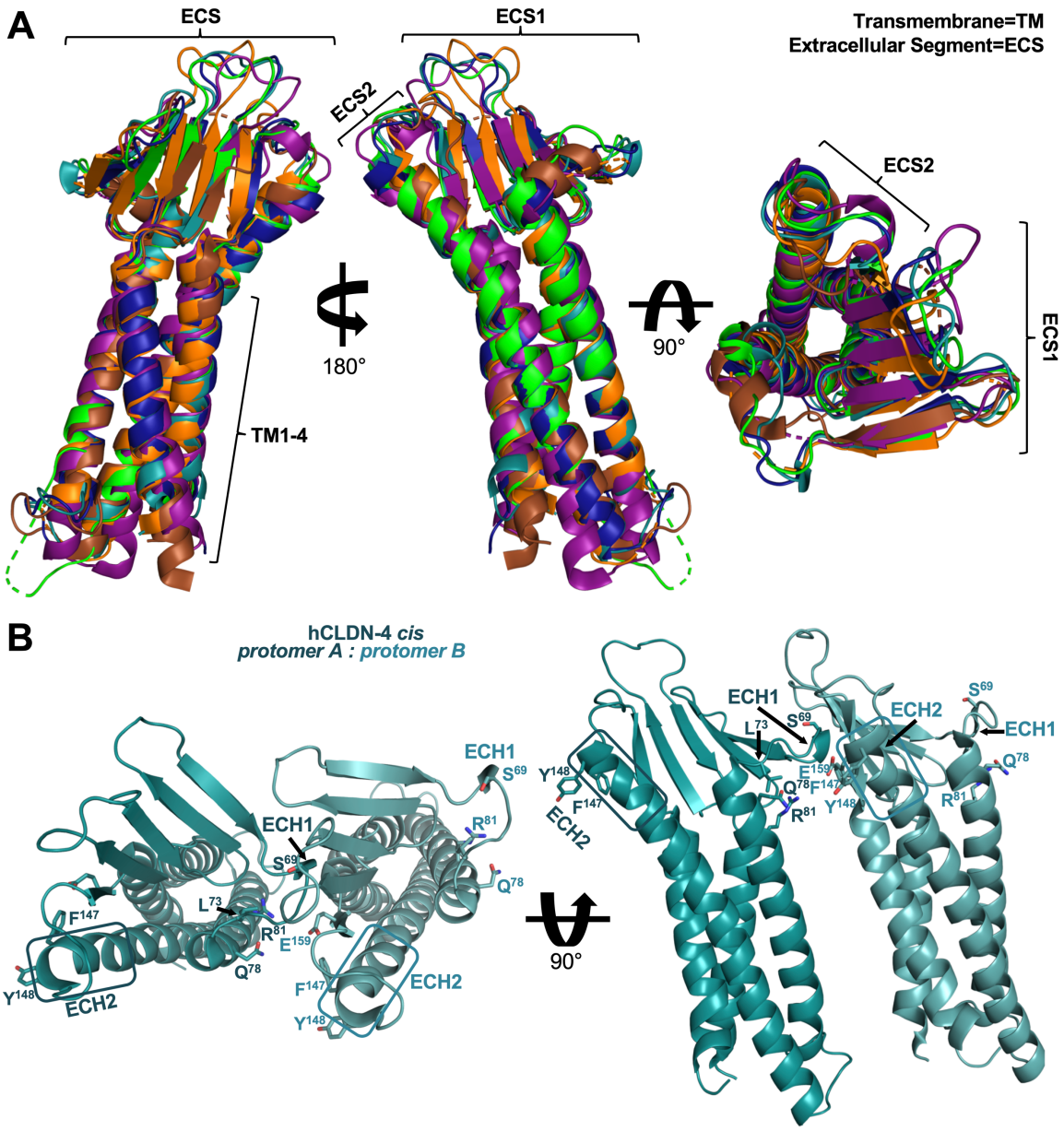


Figure S2. Claudin structural topology and a model for hCLDN-4 *cis* homodimers. (*A*) Superposition of hCLDN-4 (teal), hCLDN-9 (blue, 6ov2), mCLDN-3 (purple, 6ake), T4L-hCLDN-4 (green, 5b2g), mCLDN-19 (orange, 3x29), and mCLDN-15 (brown, 4p79), viewed parallel to the membrane (left), rotated 180° (middle), then rotated 90° viewing top-down (right). cCpE is removed for clarity. (*B*) hCLDN-4 monomers were manually docked based on the Zhao *et al.* (26) model for mCLDN-15 *cis* homodimers. Side chains thought to be involved in *cis* assembly are labeled and colored accordingly, and ECH2 are highlighted with rectangles (teal). Note the loss of hydrogen bonds between Ser69/Arg81/Gln78 and Glu159, and the distances between protomers create a loss of hydrophobic interactions between Leu73 and Phe147/Tyr148 as compared to structures of mCLDN-15 and hCLDN-9 in complex with cCpE (1).

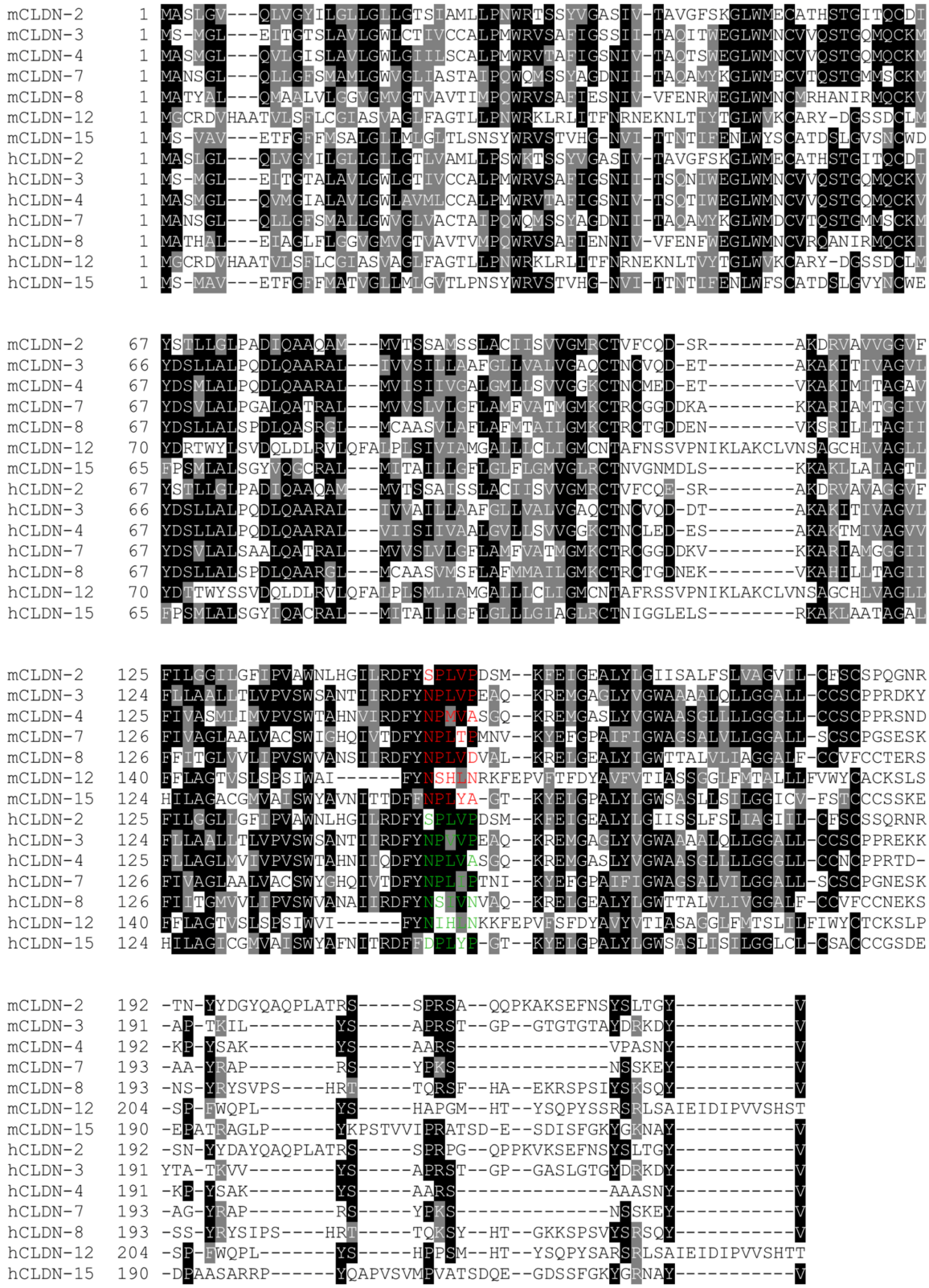


Figure S3. Sequence alignment of gut claudins. Gastrointestinal-expressing claudins from mouse and humans were aligned with T-Coffee (25). Highlighted in the sequences are the NPLVP[152] motif from mouse (red) and NPLVA[153] from humans (green).

Table S1. Data collection and refinement statistics.

	hCLDN-4/cCpE
Data collection	
Beamline	ALS 8.3.1
Space group	P 2_1 2_1 2_1
Cell dimensions	
a, b, c (Å)	70.36, 116.23, 118.05
α, β, γ (°)	90, 90, 90
Asymmetric unit volume (Å^3)	241,187
Asymmetric unit solvent (%)	81.37
Wavelength (Å)	1.11583
Resolution (Å)[a]	59.03-3.37 (3.45-3.37)
R_{meas} (%)	6.30 (594.00)
Total reflections	75314 (4503)
Unique reflections	26037 (1756)
I/σI	8.11 (0.29)
$CC_{1/2}$	100.00 (13.90[b])
Completeness (%)	98.00 (88.80)
Multiplicity	2.90 (2.56)
Wilson B-factor	156.91
Refinement	
Resolution (Å)	44.81–3.37
Number of reflections[c]	24373 (1908)
R_{work}/R_{free}	0.2938/0.2957
Number of atoms	2405
Protein	2405
Water	0
Other	0
Avg. B-factor	170.92
R.M.S. deviations	
Bond length (Å)	0.003
Bond angles (°)	0.820
Molprobity statistics[d]	
Ramachandran (%)	
Favored	95.83
Allowed	4.17
Outliers	0.00
Poor rotamers (%)	0.00
Cβ deviations (%)	0.00
Clashscore	12.40
Molprobity score	1.89
PDB ID[e]	7KP4

[a]Values in parenthesis are for highest resolution shell.
[b]Percentage of correlation between intensities from random half-datasets. Correlation significant at the 0.1% level ($CC_{1/2}^*$), as calculated by XDS.
[c]Values in parenthesis are number of reflections used for R_{free}.
[d]As determined by http://molprobity.biochem.duke.edu/
[e]Deposition code in the Protein Data Bank.

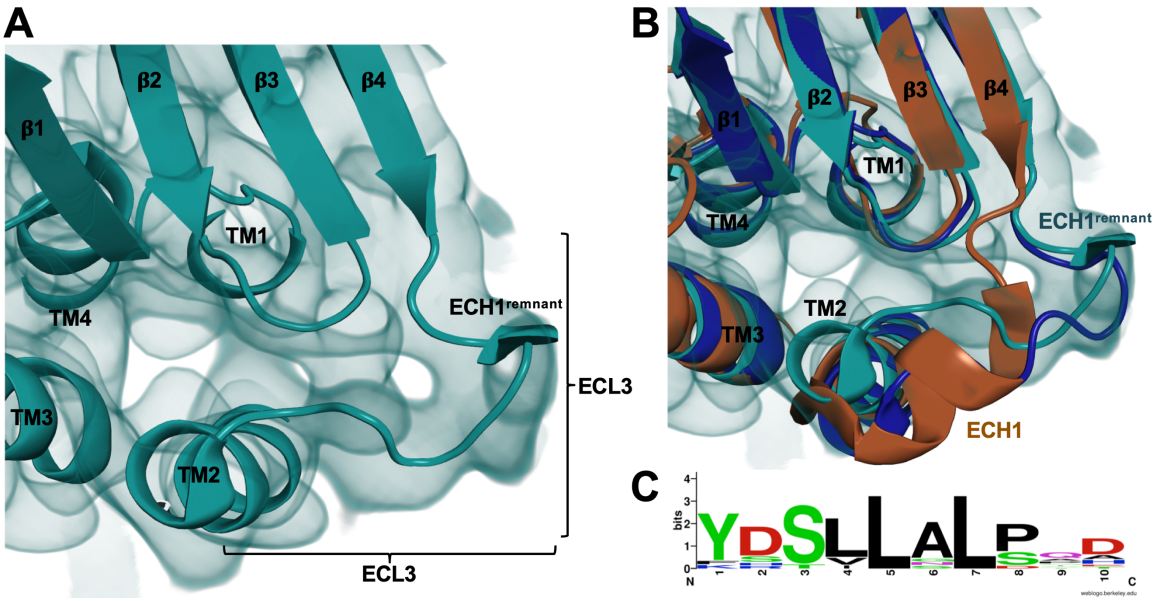


Figure S4. The ECL3 and ECH1 from claudin crystal structures. (*A*) hCLDN-4 structure (teal) and $2F_o$-F_c electron density map (teal, volume) contoured to 1.0σ showing the bulbous density surrounding the ECH1 helical remnant. cCpE is removed for clarity. (*B*) hCLDN-4 structure and electron density map as in *A* (teal) with superposition of hCLDN-9 (blue) and mCLDN-15 (brown) structures. Note the conformational differences in ECL3 and the structural differences in ECH1 between the cCpE-bound hCLDN-4 and -9 complexes and the unbound mCLDN-15. (*C*) Sequence alignments of classic claudins from humans (**Figure S1**) reveals a consensus sequence for ECH1 of SLLALP[74] using WebLogo (27).

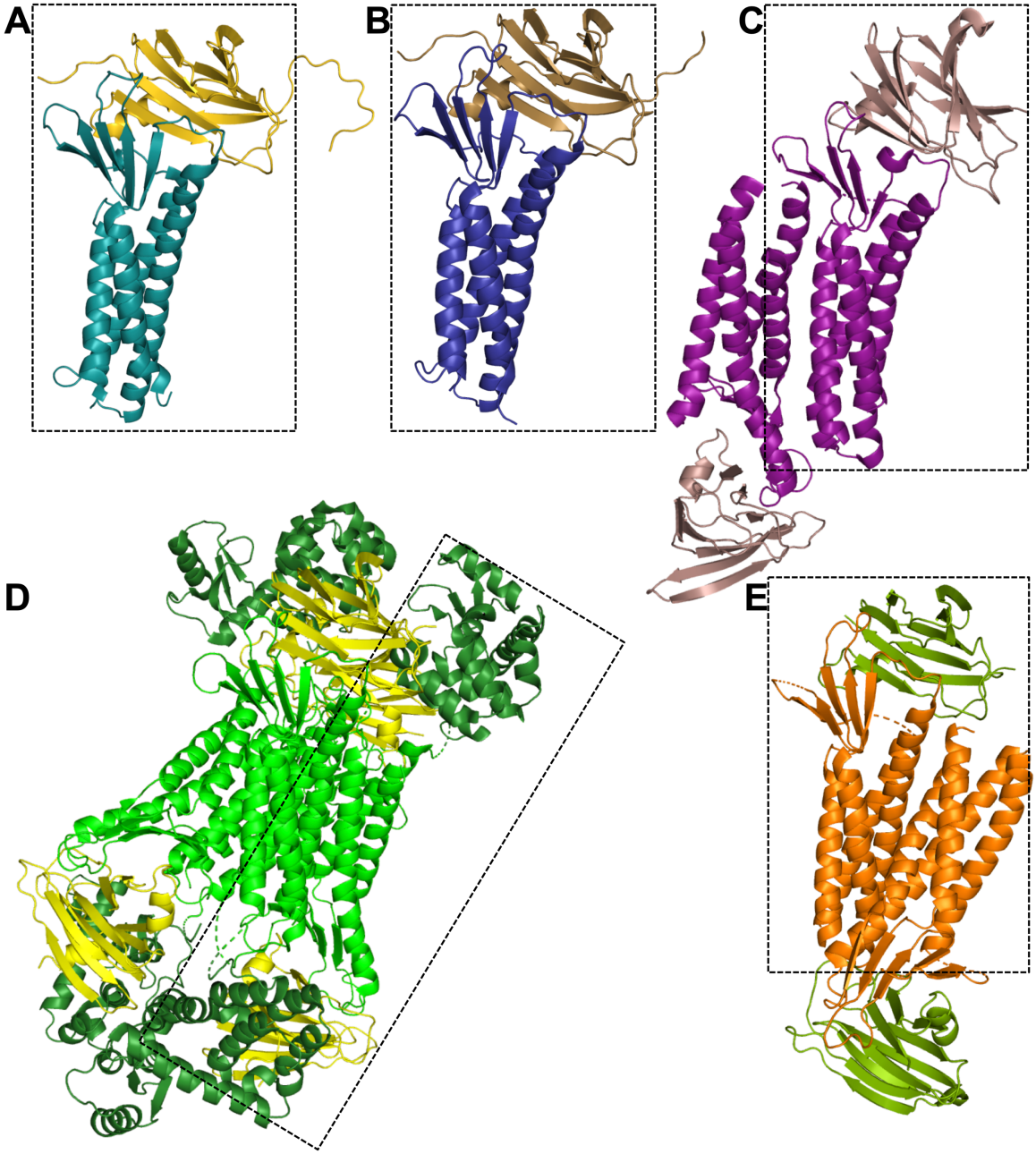


Figure S5. Crystal packing of claudin/cCpE complexes. Monomer A from claudin/cCpE structures were superimposed on hCLDN-4. (*A*) hCLDN-4 (teal) in complex with cCpE . (*B*) hCLDN-9 (blue) in complex with cCpE (copper). (*C*) Antiparallel dimeric mCLDN-3 (purple) in complex with cCpE (light purple). (*D*) Antiparallel tetrameric T4L-hCLDN-4 (T4L forest, hCLDN-4 green) in complex with cCpE (yellow). (*E*) Antiparallel dimeric mCLDN-19 (orange) in complex with cCpE (light green). Dashed rectangles (black) outline a single 1:1 claudin/cCpE complex.

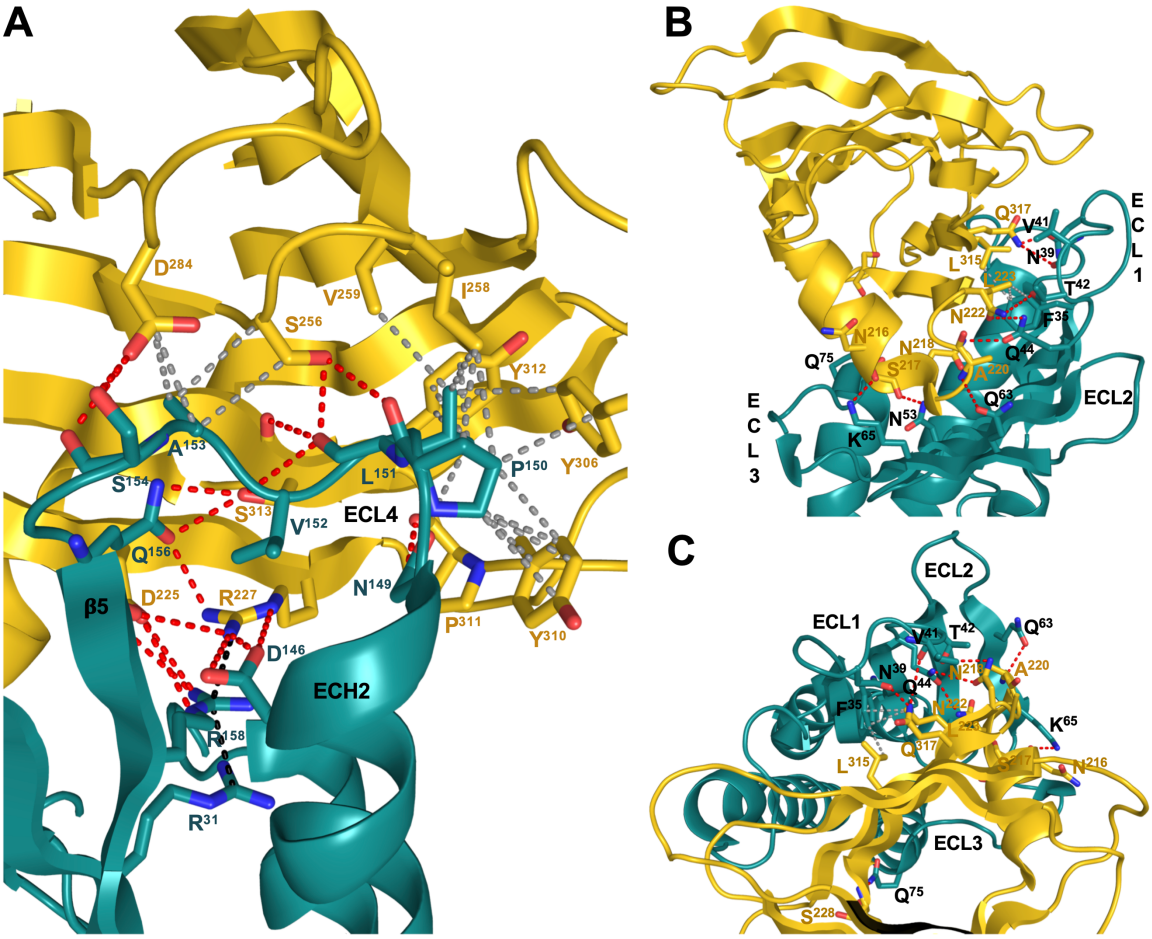


Figure S6. Intermolecular interactions in the hCLDN-4/cCpE complex. (*A*) Bonds between the ECS2 of hCLDN-4 (teal, cartoon) and cCpE (gold, cartoon). Potential polar (red, dash), non-polar (grey, dash), and arginine stacking (black, dash) interactions between side chains. Side chains are represented as sticks with carbon (teal/gold), oxygen (red), and nitrogen (blue) colored, respectively. (*B and C*) Bonds between the ECS1 of hCLDN-4 and cCpE. Proteins, side chains, and bonds are represented and colored as in *A*.

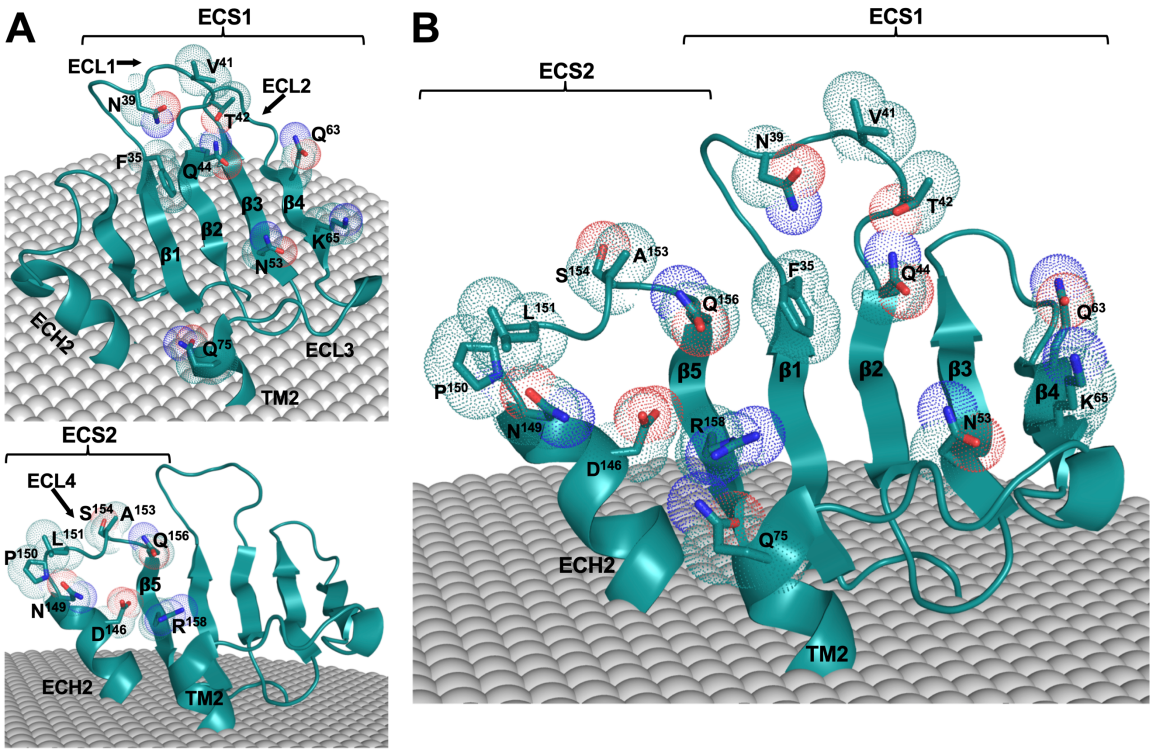


Figure S7. Amino acid determinants of hCLDN-4 interactions with cCpE. (*A*) Cartoon representation depicting the 9 residues of ECS1 (top) and 8 residues of ECS2 (bottom) that contact cCpE through ionic, polar, or non-polar interactions. Residue side chains are shown as sticks with surrounding van der Waals surfaces (dots) with carbon (teal), oxygen (red), and nitrogen (blue) coloring, respectively. The membrane surface (grey, spheres) calculated by the PPM server (28) is shown and cCpE is removed for clarity. (*B*) All 17 cCpE-binding residues of hCLDN-4 represented as in *A*. Note the extent of interacting residues across the extracellular surface of hCLDN-4.

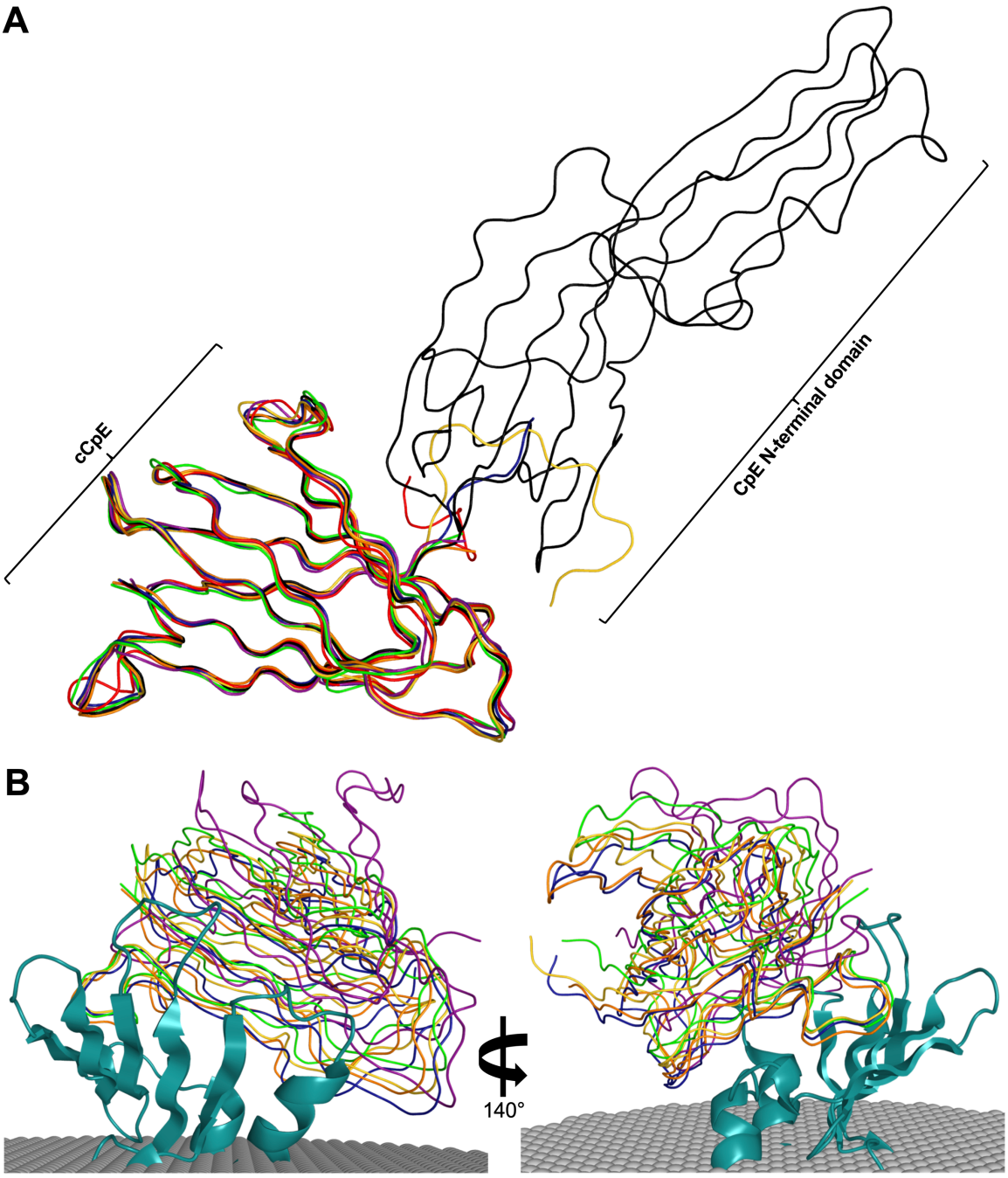


Figure S8. Structures of cCpE alone and from complexes with claudins. (*A*) Superposition of cCpE from hCLDN-4 (gold, 7kp4), hCLDN-9 (blue, 6ov2), mCLDN-3 (purple, 6ake), T4L-hCLDN-4 (green, 5b2g), mCLDN-19 (orange, 3x29), unbound (red, 2quo), and full-length CpE (black, 3am2). Chain B of cCpE from cCpE-bound structures was used for superpositions and claudins are removed for clarity. The N-terminal domain of CpE is shown for reference. (*B*) Superposition of claudins bound to cCpE. cCpE are colored as in *A* with hCLDN-4 (teal, cartoon) and membrane (grey, spheres) shown for reference. Chain A of claudins from cCpE-bound structures was used for superpositions. Note the alternate conformations of cCpE in the claudin hand that "rock" as a rigid-body.

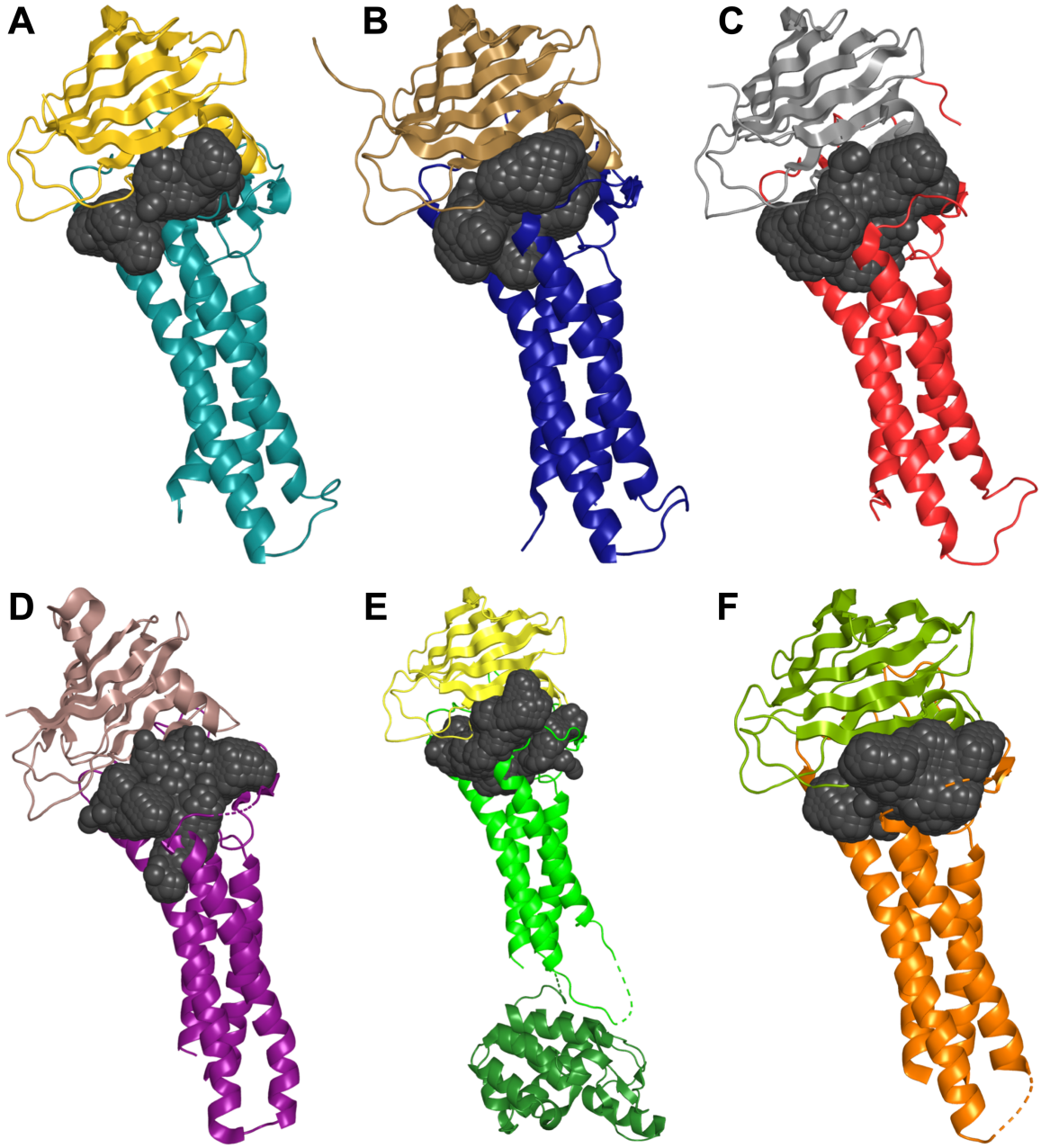
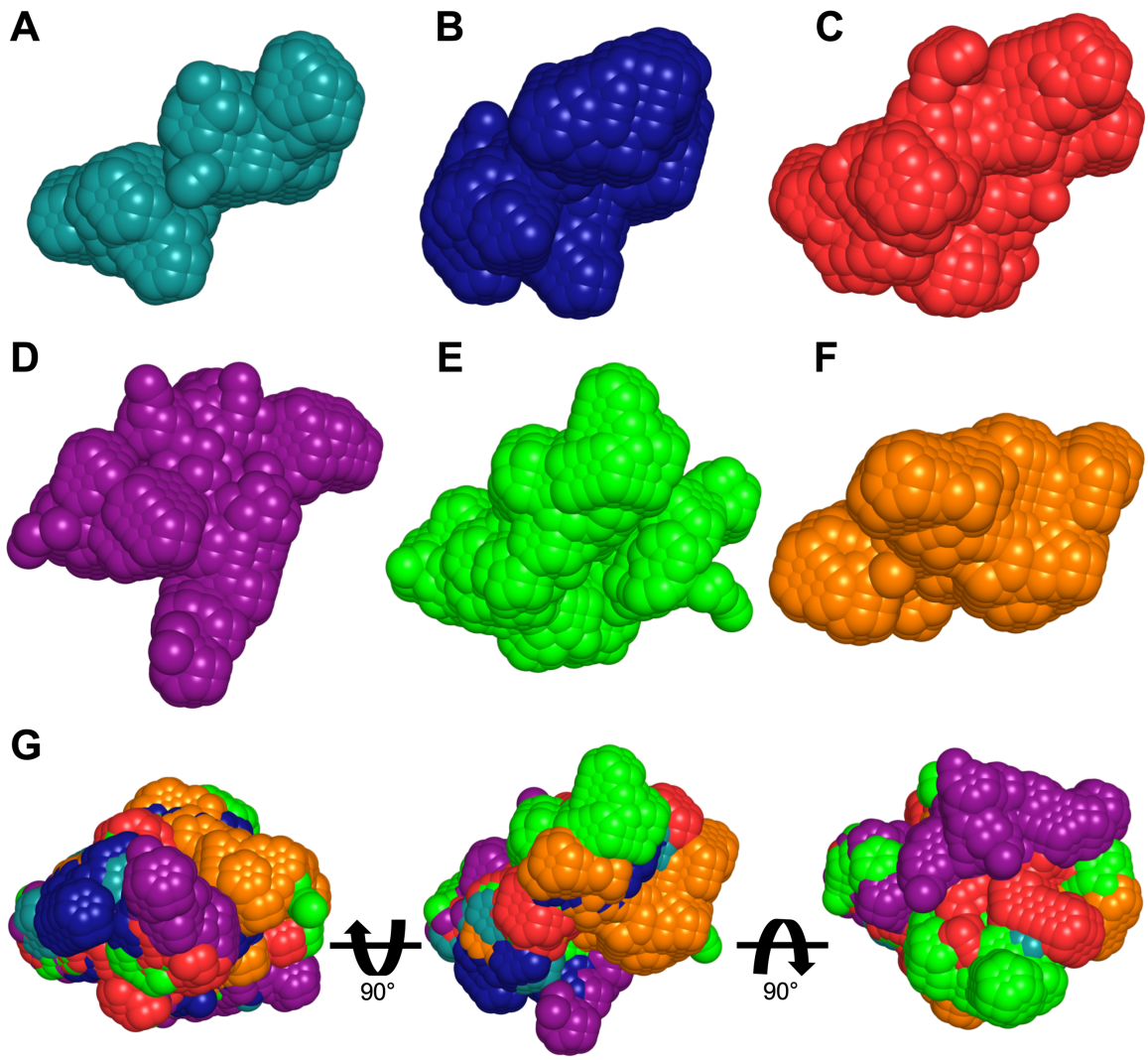


Figure S9. Solvent exposed pockets at claudin/cCpE interfaces. (*A*) hCLDN-4/cCpE (teal/gold, 7kp4). (*B*) hCLDN-9 closed/cCpE (blue/sand, 6ov2). (*C*) hCLDN-9 open/cCpE (red/grey, 6ov3). (*D*) mCLDN-3/cCpE (purple/light purple, 6ake). (*E*) T4L-hCLDN-4/cCpE (dark green and green/yellow, 5b2g). (*F*) mCLDN-19/cCpE (orange/light green, 3x29). Claudin/cCpE complexes (cartoons) are shown with pocket volumes (dark grey, spheres). Pockets were calculated with MetaPocket 2.0 (23) using chain A/B (claudin/cCpE) from all structures.



H

Complex	Interface Area (Å^2)	Pocket Volume (Å^3)	Relative Interface Area / Pocket Volume (%)
hCLDN-4	2663	5281	100 / 100
hCLDN-9 closed	2585	7491	97 / 142
hCLDN-9 open	2304	10125	87 / 192
mCLDN-3	1929	9869*	72 / 187*
T4L-hCLDN-4	2312	9552	87 / 181
mCLDN-19	2143	8169	81 / 155

Figure S10. Comparison of interface pockets in claudin/cCpE complexes. Proteins were removed and the solvent accessible pockets from **Figure S9**: (*A*) hCLDN-4/cCpE (teal), (*B*) hCLDN-9 closed/cCpE (blue), (*C*) hCLDN-9 open/cCpE (red), (*D*) mCLDN-3/cCpE (purple), (*E*) T4L-hCLDN-4/cCpE (green), and (*F*) mCLDN-19/cCpE (orange) are shown as spheres. (*G*) Pockets from *A*-*F* were superposed and viewed parallel to the membrane (middle), rotated +90° (right, view down to membrane surface) and -90° (left, view up through membrane) to demonstrate their unique shapes. (*H*) The areas of the claudin/cCpE interfaces and the pocket volumes as calculated by PDBSum and Fpocket (21, 22) are shown relative to hCLDN-4. Note that hCLDN-4 has the highest interface area and smallest pocket volume. *The low interface area of mCLDN-3 in complex with cCpE produces a pocket so large that the borders are difficult to define with a probe and thus are not measured accurately.

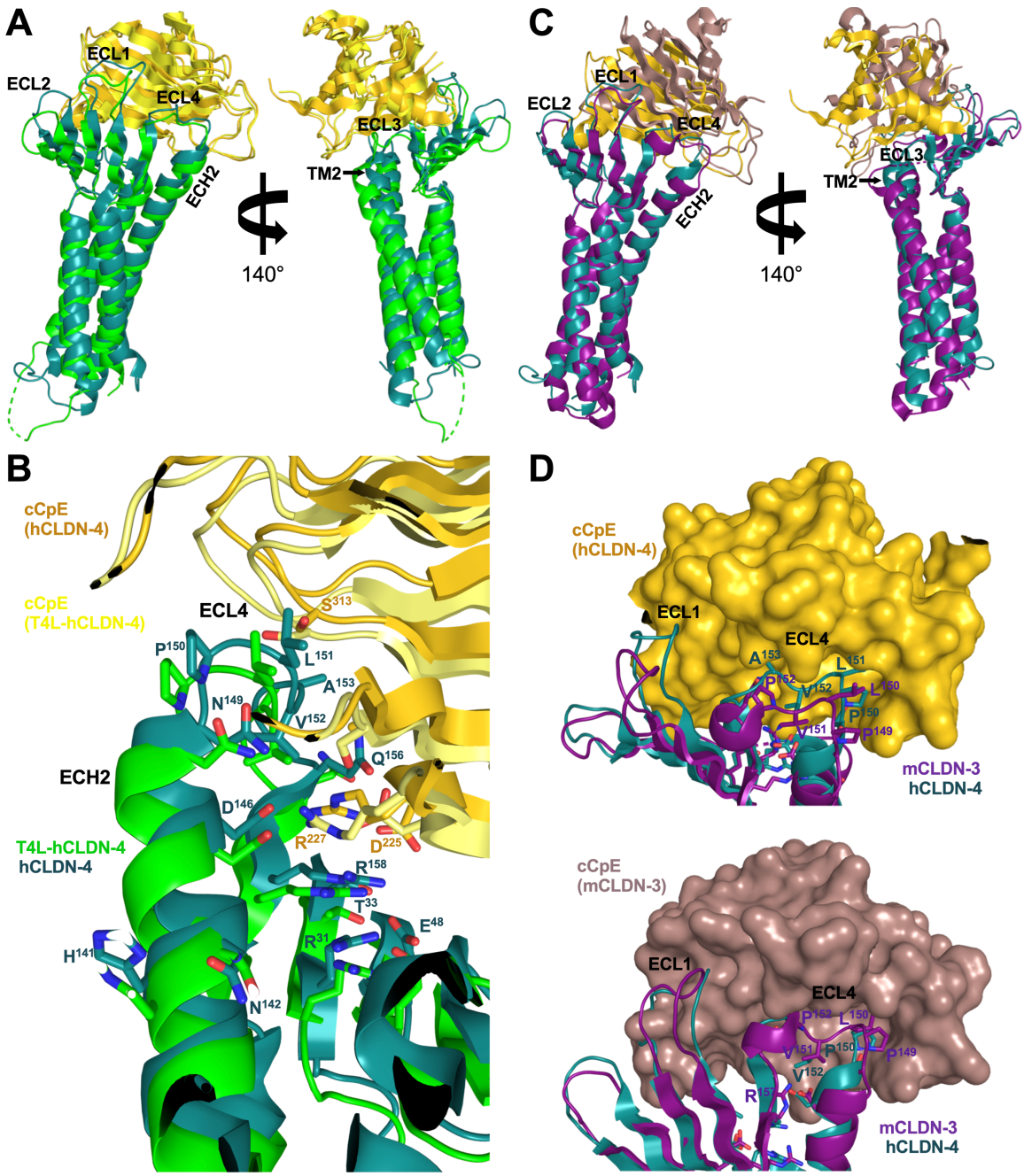


Figure S11. Structural alterations between high-affinity cCpE claudin receptors. (*A*) Superposition of hCLDN-4/cCpE (teal/gold) and T4L-hCLDN-4/cCpE (green/yellow). (*B*) Side chain alterations in the cCpE-binding motif between hCLDN-4/cCpE and T4L-hCLDN-4/cCpE. Complexes and side chains are represented and colored as in *A*. (*C*) Superposition of hCLDN-4/cCpE (teal/gold) and mCLDN-3/cCpE (purple/light purple). (*D*) ECL4 conformations between hCLDN-4 (teal) and mCLDN-3 (purple) with cCpE (gold, surface) from hCLDN-4 (top); and with cCpE (light purple, surface) from mCLDN-3 (bottom). Complexes are colored as in *C*. Note the altered ECL4 structure and resulting cCpE conformational changes between structures.

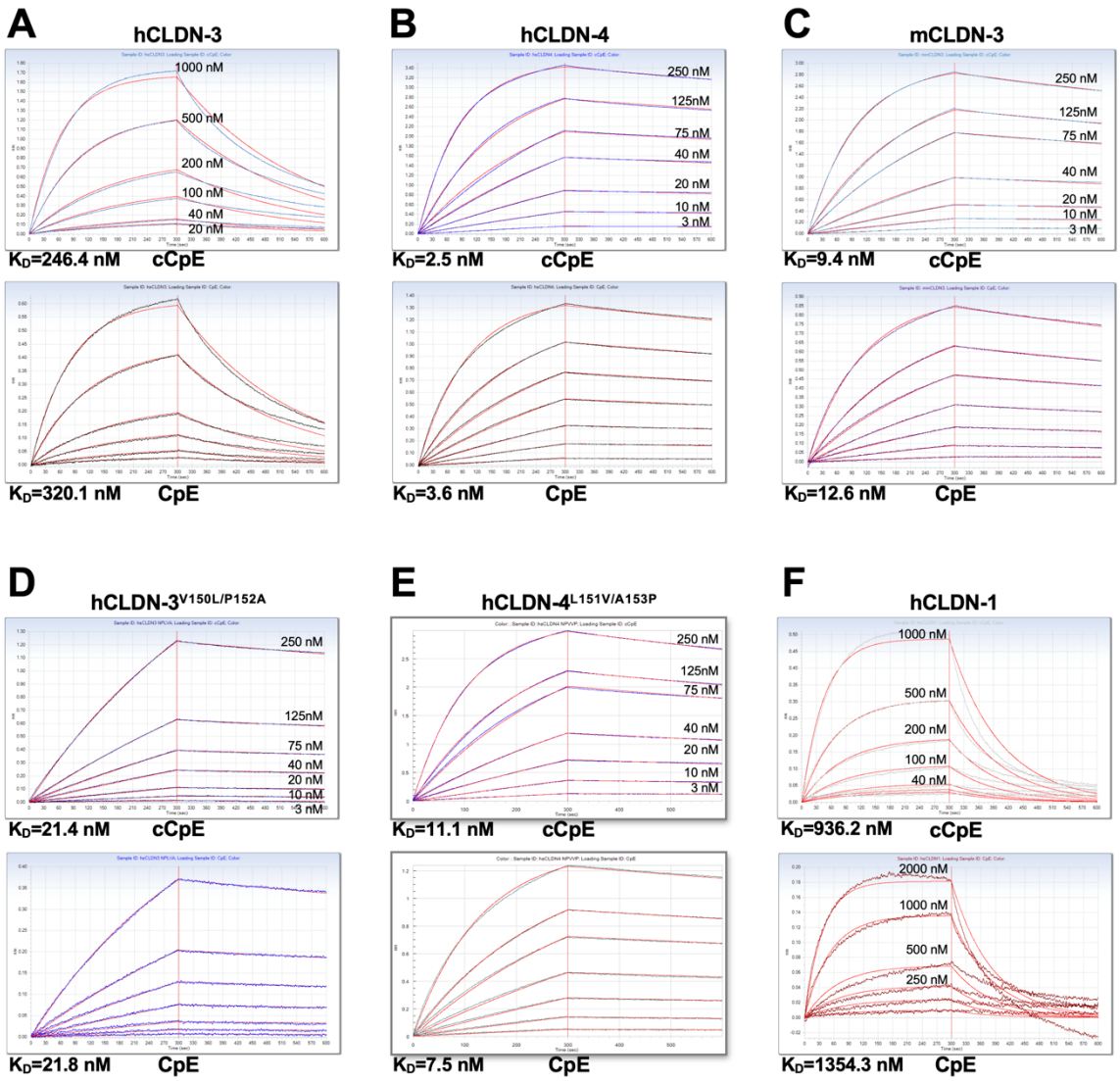


Figure S12. Measurements of claudin binding to enterotoxins. Sensograms from BLI using immobilized cCpE and CpE in SEC buffer with UDM detergent binding to: (*A*) hCLDN-3, (*B*) hCLDN-4, (*C*) mCLDN-3, (*D*) hCLDN-3$^{V150L/P152A}$, (*E*) hCLDN-4$^{L151V/A153P}$, and (*F*) hCLDN-1. Raw data (various colors) are overlaid with best fit lines (red). The vertical red line at 300 seconds indicates start of dissociation recording. Note rapid dissociation of hCLDN-1 and -3 from enterotoxin-bound biosensors. The sensograms represent one sample set from at least duplicate measurements.

SI References

1. Vecchio AJ & Stroud RM (2019) Claudin-9 structures reveal mechanism for toxin-induced gut barrier breakdown. *Proc Natl Acad Sci U S A* 116(36):17817-17824.
2. Nakamura S*, et al.* (2019) Morphologic determinant of tight junctions revealed by claudin-3 structures. *Nat Commun* 10(1):816.
3. Saitoh Y*, et al.* (2015) Structural insight into tight junction disassembly by Clostridium perfringens enterotoxin. *Science* 347(6223):775-778.
4. Shinoda T*, et al.* (2016) Structural basis for disruption of claudin assembly in tight junctions by an enterotoxin. *Sci Rep* 6:33632.
5. Krissinel E & Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372(3):774-797.
6. Robertson SL, Smedley JG, 3rd, & McClane BA (2010) Identification of a claudin-4 residue important for mediating the host cell binding and action of Clostridium perfringens enterotoxin. *Infect Immun* 78(1):505-517.
7. Veshnyakova A*, et al.* (2012) Mechanism of Clostridium perfringens enterotoxin interaction with claudin-3/-4 protein suggests structural modifications of the toxin to target specific claudins. *J Biol Chem* 287(3):1698-1708.
8. Winkler L*, et al.* (2009) Molecular determinants of the interaction between Clostridium perfringens enterotoxin fragments and claudin-3. *J Biol Chem* 284(28):18863-18872.
9. Kimura J*, et al.* (2010) Clostridium perfringens enterotoxin interacts with claudins via electrostatic attraction. *J Biol Chem* 285(1):401-408.
10. Hou J, Renigunta A, Yang J, & Waldegger S (2010) Claudin-4 forms paracellular chloride channel in the kidney and requires claudin-8 for tight junction localization. *Proc Natl Acad Sci U S A* 107(42):18010-18015.
11. Yu AS*, et al.* (2009) Molecular basis for cation selectivity in claudin-2-based paracellular pores: identification of an electrostatic interaction site. *J Gen Physiol* 133(1):111-127.
12. Suzuki H, Tani K, Tamura A, Tsukita S, & Fujiyoshi Y (2015) Model for the architecture of claudin-based paracellular ion channels through tight junctions. *J Mol Biol* 427(2):291-297.
13. Suzuki H*, et al.* (2014) Crystal structure of a claudin provides insight into the architecture of tight junctions. *Science* 344(6181):304-307.
14. Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):125-132.
15. Adams PD*, et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213-221.
16. Bunkoczi G*, et al.* (2013) Phaser.MRage: automated molecular replacement. *Acta Crystallogr D Biol Crystallogr* 69(Pt 11):2276-2286.
17. Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486-501.
18. Afonine PV*, et al.* (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* 68(Pt 4):352-367.
19. Schrödinger (The PyMOL Molecular Graphics System Version 2.2.3).
20. Morin A*, et al.* (2013) Collaboration gets the most out of software. *Elife* 2:e01456.
21. Schmidtke P, Le Guilloux V, Maupetit J, & Tuffery P (2010) fpocket: online tools for protein ensemble pocket detection and tracking. *Nucleic Acids Res* 38:W582-589.
22. Laskowski RA (2001) PDBsum: summaries and analyses of PDB structures. *Nucleic Acids Res* 29(1):221-222.
23. Huang B (2009) MetaPocket: a meta approach to improve protein ligand binding site prediction. *OMICS* 13(4):325-330.
24. Kawate T & Gouaux E (2006) Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins. *Structure* 14(4):673-681.
25. Notredame C, Higgins DG, & Heringa J (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* 302(1):205-217.
26. Zhao J*, et al.* (2018) Multiple claudin-claudin cis interfaces are required for tight junction strand formation and inherent flexibility. *Commun Biol* 1:50.
27. Crooks GE, Hon G, Chandonia JM, & Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Res* 14(6):1188-1190.

28. Lomize MA, Pogozheva ID, Joo H, Mosberg HI, & Lomize AL (2012) OPM database and PPM web server: resources for positioning of proteins in membranes. *Nucleic Acids Res* 40:D370-376.