

## SUPPORTING INFORMATION FOR:

# Computational Assessment of Different Structural Models for Claudin-5 Complexes in Blood-Brain-Barrier Tight-Junctions

*Alessandro Berselli*<sup>#1,2</sup>, *Giulio Alberini*<sup>#1,3</sup>, *Fabio Benfenati*<sup>\*1,3</sup> *Luca Maragliano*<sup>\*1,4</sup>

<sup>1</sup> *Center for Synaptic Neuroscience and Technology (NSYN@UniGe), Istituto Italiano di Tecnologia, Largo Rosanna Benzi, 10, 16132 Genova, Italy*

<sup>2</sup> *Department of Experimental Medicine, Università degli Studi di Genova, Viale Benedetto XV, 3, 16132 Genova, Italy*

<sup>3</sup> *IRCCS Ospedale Policlinico San Martino, Largo Rosanna Benzi, 10, 16132 Genova, Italy*

<sup>4</sup> *Department of Life and Environmental Sciences, Polytechnic University of Marche, Via Brecce Bianche, 60131, Ancona, Italy*

\*Email: [fabio.benfenati@iit.it](mailto:fabio.benfenati@iit.it), [luca.maragliano@iit.it](mailto:luca.maragliano@iit.it)

### List of Abbreviations

- BBB: Blood-brain barrier
- CLDN: Claudin
- TM: Trans-membrane
- ECL: Extracellular loop
- FE: Free energy
- MD: Molecular dynamics
- RMSD: Root-mean square deviation
- US: Umbrella sampling
- WHAM: Weighted histogram analysis method

### List of Models

#### Pore I - Model 1

Single pore configuration, where the four Cldn5 promoters were homology-modeled based on the equilibrated configuration of the Cldn15 single pore model illustrated in Ref. 1.

#### Pore I - Model 2

Single pore configuration, where each Cldn5 promoter is modeled based on the homology with the crystal structure of the mouse Cldn15 promoter (PDB ID: 4P79)<sup>2</sup>. The configuration of the single pore model is obtained starting from the atomic coordinates of the pdb file provided in Ref. 3.

# CONTENTS

1. ANALYSIS OF STANDARD MOLECULAR DYNAMICS SIMULATIONS OF TWO VERSIONS OF THE PORE I MODEL	S2
1.1 RMSD	S2
1.2 Cross-distances	S2
1.3 Pore Radius	S7
2. BLOCK ERROR ANALYSIS	S8
3. STRUCTURAL ANALYSIS OF THE CLDN5 MONOMER	S10
REFERENCES	S12

## 1. Analysis of Standard Molecular Dynamics Simulations of two Versions of the Pore I model

### 1.1 RMSD

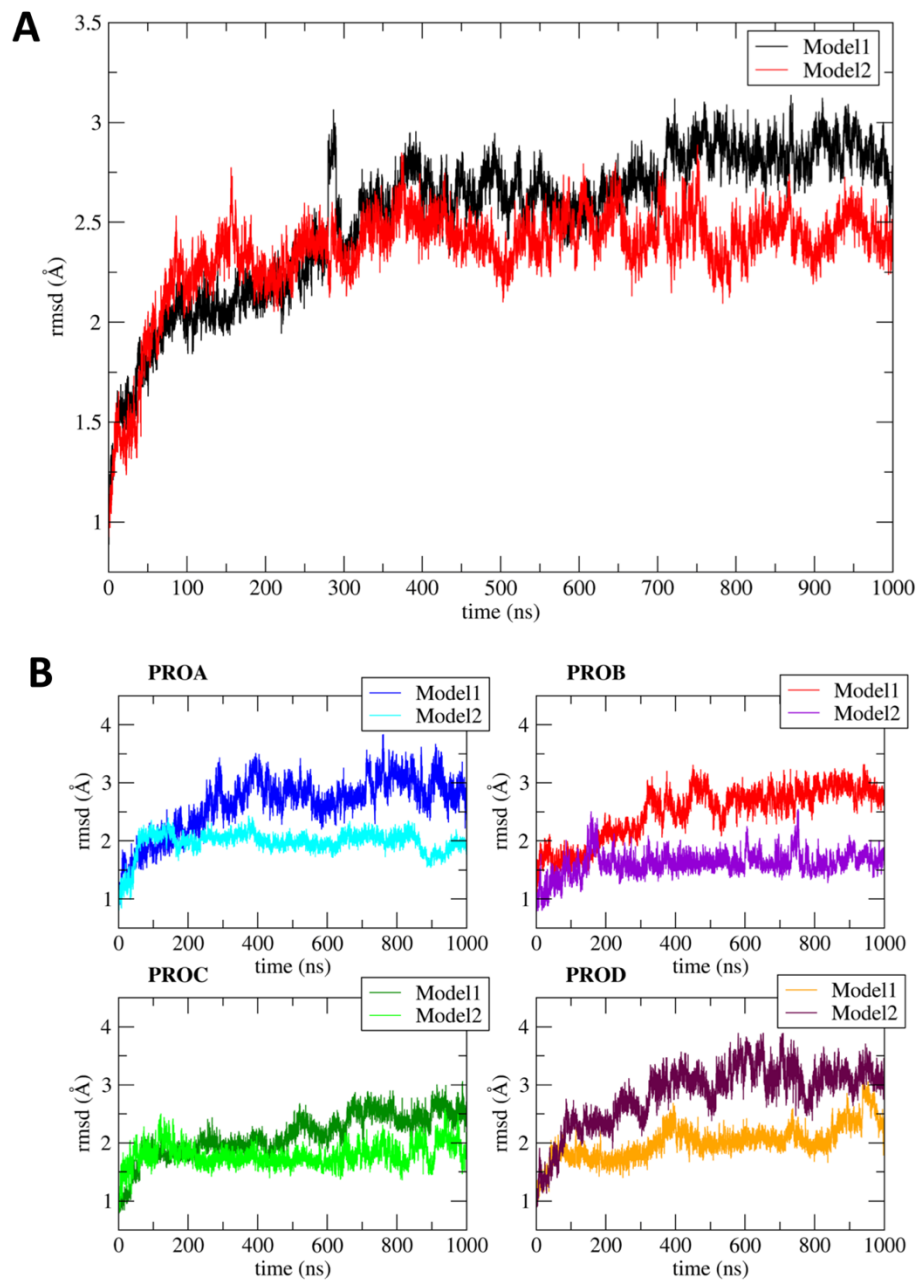
We calculated the root mean squared deviation (RMSD) of the atomic backbone with respect to the starting conformation, using configurations sampled every 100 ps along the trajectories. The RMSD was computed for different domains of the structures (**Figure S1**). In the following, we indicate the four protomers using the PROA/B/C/D labels. For both the models, we observe a plateau between 2.5-3 Å for the tetramer, as illustrated in **Figure S1A**. The RMSD for each protomer and for each ECL was calculated (**Figure S1B, S2**). These results show that the integrity of the  $\beta$ -barrel is preserved, and the dimension of the cavity does not reveal significant variations, as confirmed by the pore radius calculations reported in **Figure S6**.

### 1.2 Cross-distances

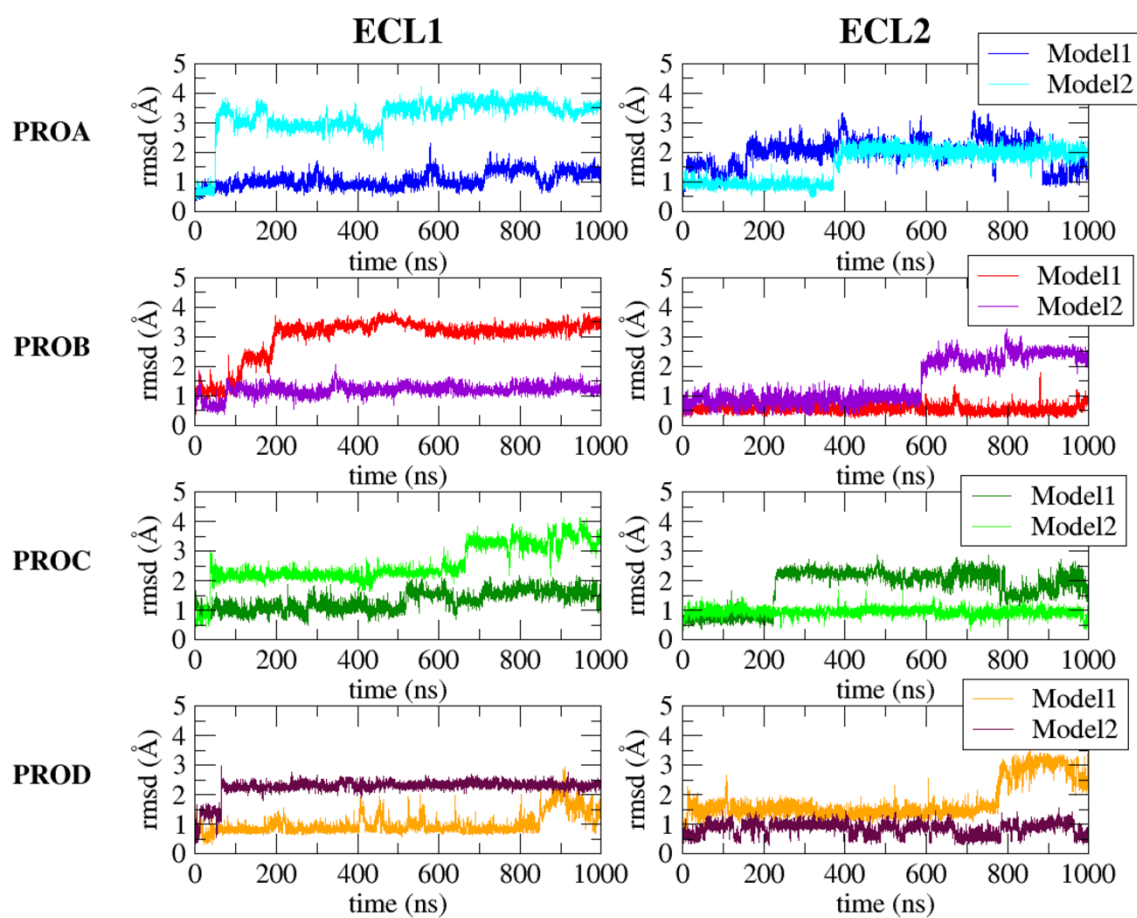
We calculated various cross-distances between pore-lining residues as illustrated in **Figure S3**. In particular, for each pair of residues, we computed the distance between the two C $\alpha$  atoms and between the last C-atom of the sidechains, as summarized in **Table S2**. The Gln57 and Gln63 residues are located in the narrowest region, while Lys65 is next to them (**Figure S3C**). The distance between the residues Val70 and Pro153 accounts for the integrity of the lateral scaffold of the pore while the other between the residues Lys48 describes the fluctuations of the pore diameter at its entrance (**Figure S3D**).

Both the models exhibit stationary fluctuations for each cross-distance, as illustrated in **Figure S4**. In particular, for the Gln63 pairs in Model1 the average value of the CD – CD distance (between the C-atom at the extremities of the sidechains) is  $12.29 \pm 1.15$  Å for the PROA-PROD dimer, and  $8.32 \pm 1.19$  Å for the PROB-PROC dimer (**Figure S4A**). These distances are smaller than those found in the Model2, where values of  $15.89 \pm 1.28$  Å and  $13.69 \pm 1.02$  Å, respectively, are obtained (**Figure S4B**).

The distances between residues at the mouth of the pore are reported in **Figure S4**.



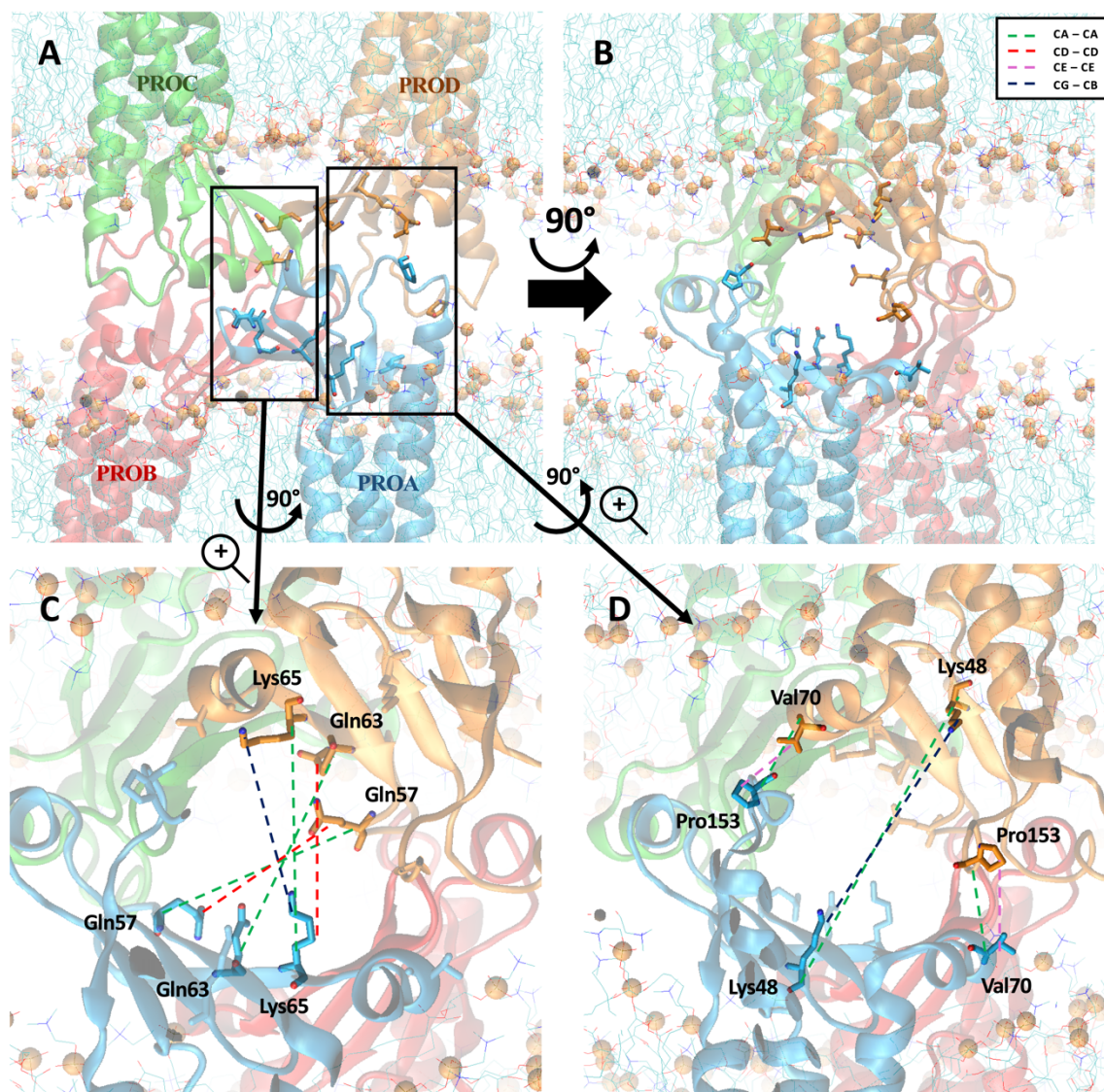
**Figure S1.** Backbone RMSD of Model1 and Model2. A) RMSD for the tetramer. B) RMSD for each protomer.



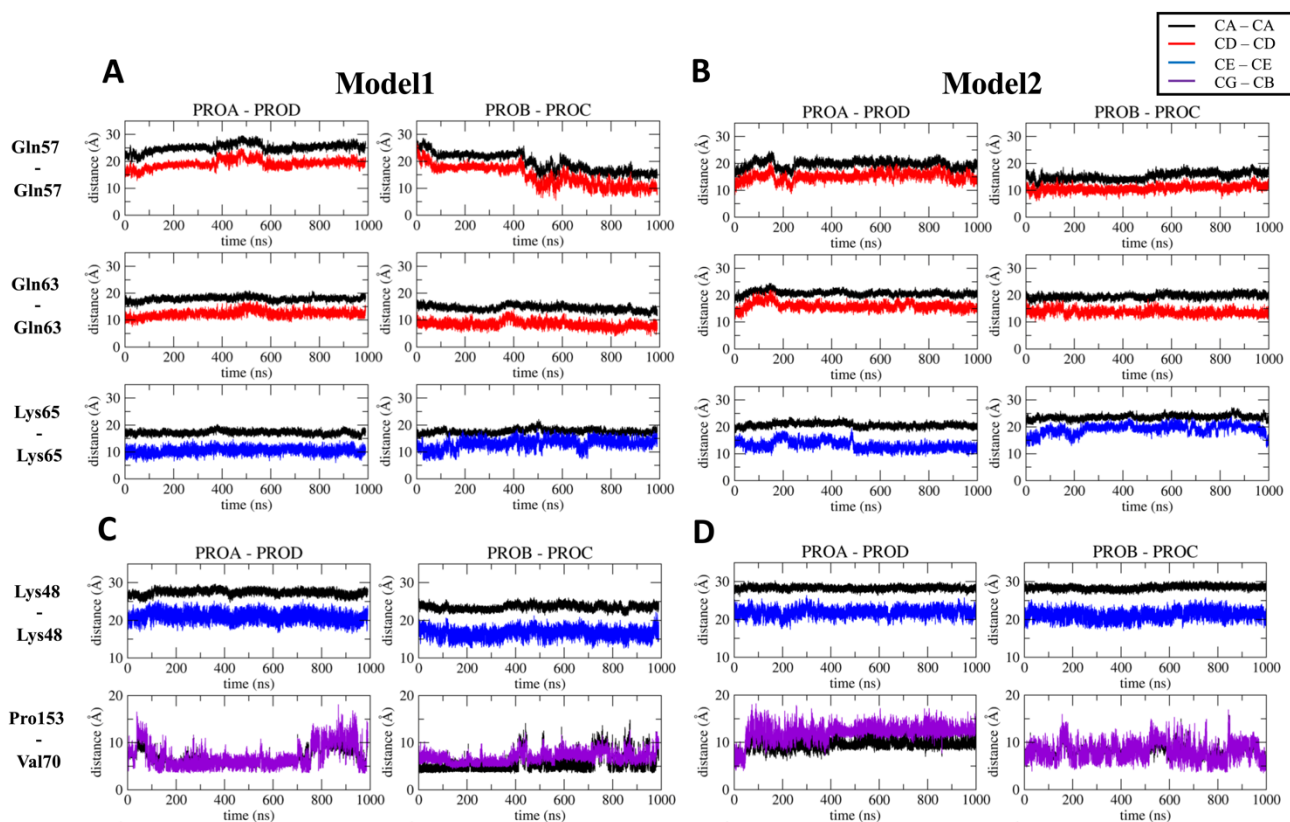
**Figure S2.** RMSD of the Model1 and the Model2 ECLs.

**Table S1.** Cross-distances mapped along the 1- $\mu$ s-long MD simulation of the two Pore I models.

Residue	Atom type	Protomers
<b>Gln57 – Gln57</b>	CD – CD	PROA - PROD
	CA – CA	PROB - PROC
	CD – CD	PROA - PROD
<b>Gln63 – Gln63</b>	CA – CA	PROB - PROC
	CD – CD	PROA - PROD
	CA – CA	PROB - PROC
<b>Lys65 – Lys65</b>	CE – CE	PROA - PROD
	CA – CA	PROB - PROC
	CE – CE	PROA - PROD
<b>Lys48 – Lys48</b>	CA – CA	PROB - PROC
	CE – CE	PROA - PROD
	CA – CA	PROB - PROC
<b>Pro153 – Val70</b>	CG – CB	PROA - PROD
	CA – CA	PROB - PROC
	CG – CB	PROA - PROD
	CA – CA	PROB - PROC

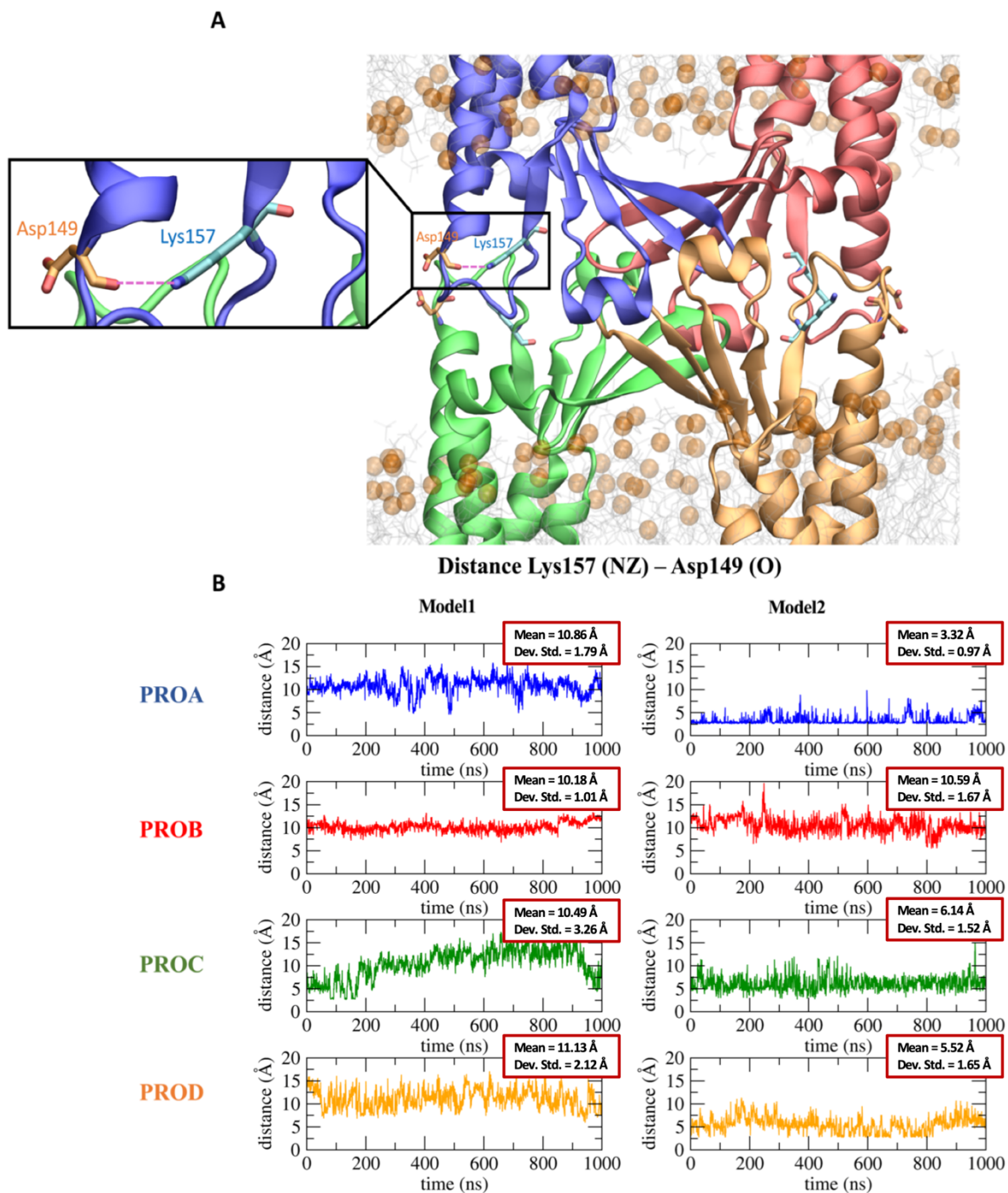


**Figure S3.** Graphical representation of the cross-distances calculated along the trajectories of Model1 and Model2, visualized on Model2. Individual protomers are distinguished by their coloring and referred as PROA (blue), PROB (red), PROC (green), PROD (orange). C-atoms are reported with the CHARMM atom name. A) Lateral side view; residues are displayed only for protomers PROA and PROD. B) Apical view of the pore cavity. C) Close-up view of the distances monitored at the center of the pore. D) Close-up view of the distances monitored at the mouth of the pore.



**Figure S4.** Cross-distances for Model1 and Model2. A) Residue pairs close to the pore center of Model1. B) Residue pairs close to the pore center of Model2. C) Residue pairs at the entrance of the Model1 pore. D) Residue pairs at the entrance of the Model2 pore.

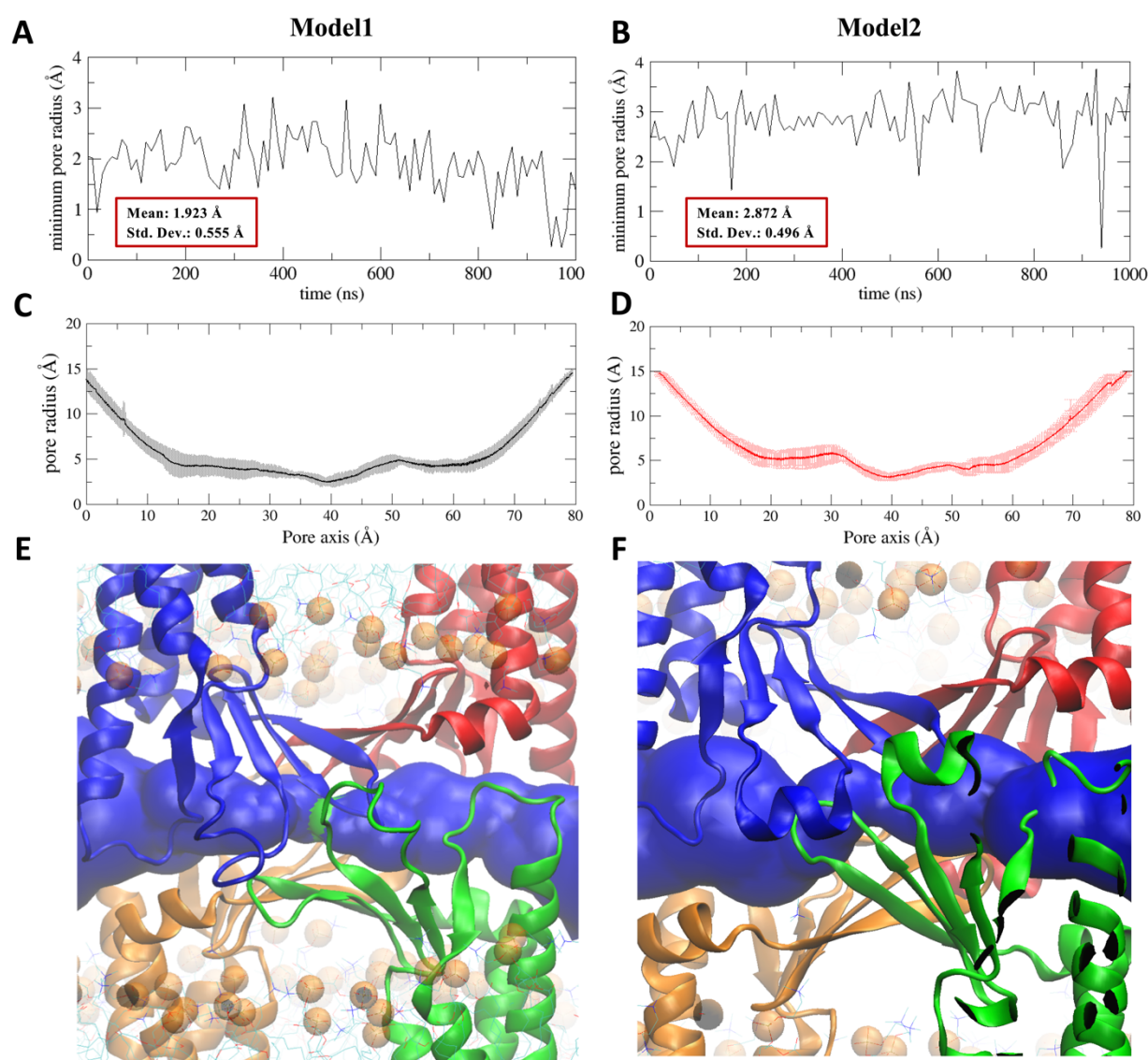
In Ref. 4, the authors reported the presence of the hydrogen bond (HB) between the sidechain of Lys155 and the backbone of Asn148 as pivotal for the preservation of the ECL2 rigidity in the structure of the mouse Cldn15 (PDB ID: 4P79)<sup>2</sup>. Consistently, in the present work, we identified the corresponding interaction in the Cldn5-based models between the residues Lys157 and Asp149 (**Figure S5A**). In each model, we monitored the value of the distance between the N-atom of the Lys157 sidechain and the O-atom of the Asp149 backbone. The interaction is not preserved in any of the four protomers belonging to Model1. On the other hand, in Model2, the distances correlate with a HB for three (**Figure S5B**).



**Figure S5.** Interaction between Lys157 and Asp149 in the four protomers of each model. A) Graphical representation of the two residues at the extremities of the beta-turn in the ECL2. B) Values of distance measured along the trajectory for the individual protomers of each model.

### 1.3 Pore Radius

We used the HOLE program<sup>5,6</sup> to calculate the size of the cavity along the pore axis (Figure S6). In particular, the narrowest value of the pore radius is localized at the center of the cavity (Figure S6E,F).



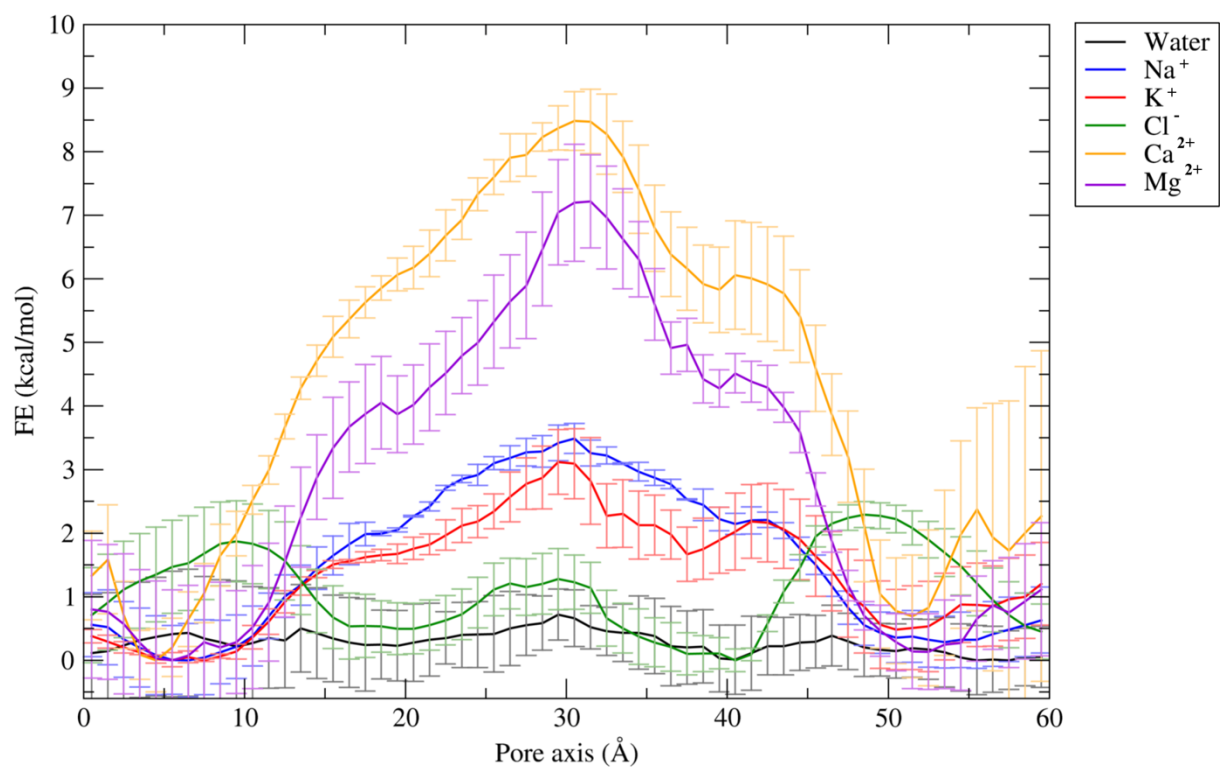
**Figure S6.** Pore radius profiles of Model1 and Model2. A) Minimal pore radius for Model1. B) Minimal pore radius for Model2. C) Average profile along the pore axis for Model1. The grey region identifies the standard deviation. D) Average profile along the pore axis for Model2. The light red region identifies the standard deviation. E) Surface representation of the pore size for Model1. F) Surface representation of the pore size for Model2. The blue-colored areas represent the regions where the radius is at least twice the radius of a single water molecule, while the green-colored section define regions fitting no more than one water molecule.

In Model2, we observed the preservation of the hydrogen bond between Lys157 and Asp149 for three protomers<sup>4</sup>. Based on this result, we chose Model2 as the best representative structure of the Pore I arrangement, to be used for FE calculations.

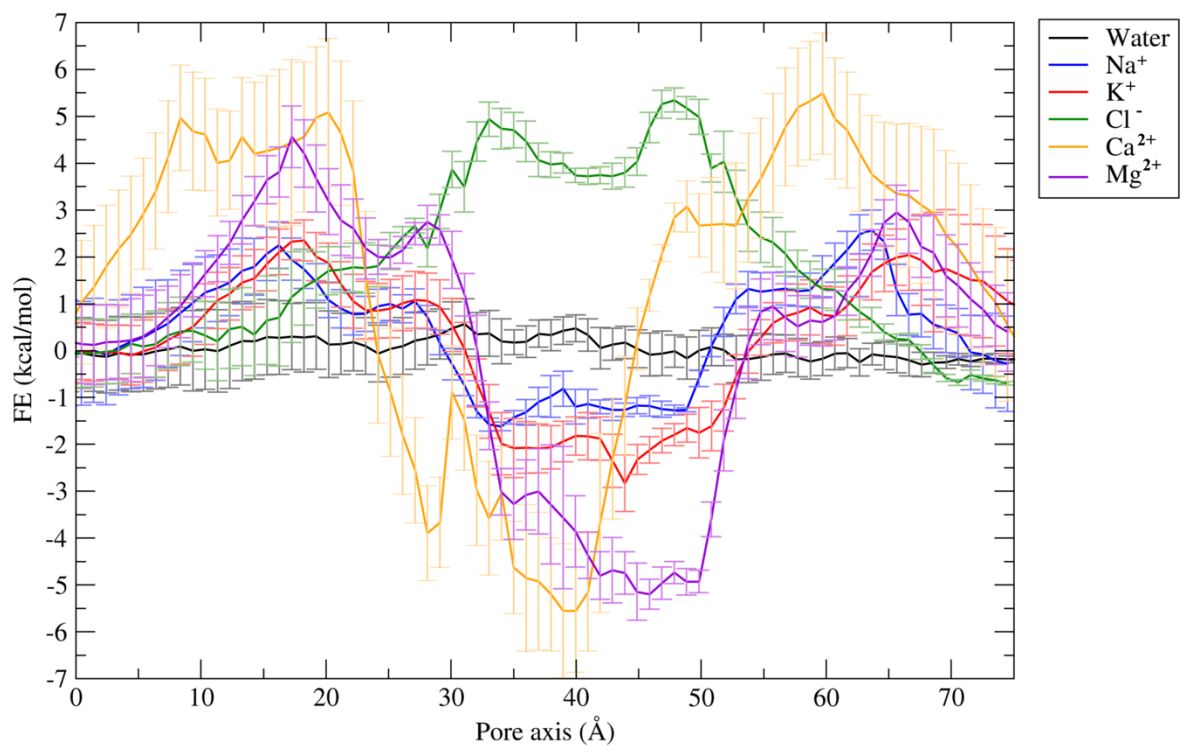
## 2. Block Error Analysis

The FE profiles were calculated with the Umbrella Sampling (US) method<sup>7</sup> in combination with the *weighted histogram analysis method* (WHAM)<sup>8-10</sup>, available at <http://membrane.urmc.rochester.edu/content/wham>. This software implements the bootstrap error analysis to evaluate the uncertainty during the calculation of the FE landscape, which is reported to be underestimated<sup>8,9</sup>. For this reason, we also computed the error estimates using the block analysis, by dividing the trajectories in four blocks<sup>11</sup>, as reported in **Figure S7** and **Figure S8**, respectively.





**Figure S7.** Free energy profiles of water and ions permeating the Pore I. The error, estimated with the block analysis, is represented with bars.



**Figure S8.** Free energy profiles of water and ions permeating the Pore II. The error, estimated with the block analysis, is represented with bars.

### 3. Structural Analysis of the Cldn5 monomer

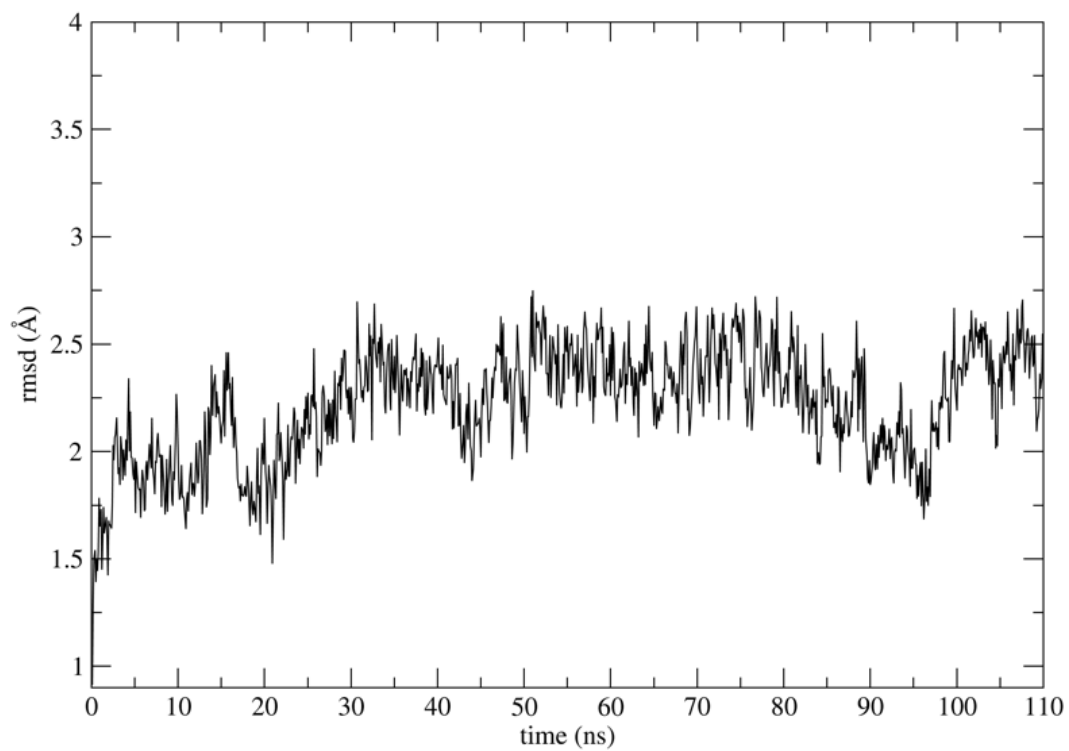
The Cldn5 monomer was homology modelled using the Cldn15 structure as template (PDB ID: 4P79)<sup>2</sup>. In order to assess the structural stability of the model, we performed 110 ns of all-atom MD simulation of the protein embedded in a POPC membrane and solvated with explicit water and counterions. The RMSD of the protein backbone is stationary around 2.25 Å (**Figure S9**). Additionally, we analyzed the network of intramolecular interactions responsible for the stabilization of the tertiary structure<sup>12</sup>. The distances between selected residues are reported in **Table S2** and illustrated in **Figure 10A**. In the inner part of the TM domains, hydrophobic interactions are defined by Trp18, Phe92 and Phe96 of the TM1 and TM2 helices.

In order to assess the conservation of residues involved in the aforementioned interactions, we performed a multiple sequence alignment with Clustal Omega<sup>13,14</sup>, using the sequences of the human Cldn5, the mouse Cldn15 (mCldn15, PDB ID: 4P79)<sup>2</sup> and the human Cldn4 (PDB ID: 5B2G)<sup>15</sup>. While the Cldn15 forms barrier to the passage of anions and channels for the cations<sup>11,16</sup>, the Cldn4 is mainly supposed to provide an opposite selectivity to the ion permeation<sup>17</sup>. The final alignment is shown in **Figure S10B**. The aromatic Trp30, Trp51, Phe147 and the charged Arg81 residues are conserved among the three homologs, providing an interacting interface between the TM and the ECL domains. Moreover, an analogue cation -  $\pi$  was reported in the Cldn15 between the Arg79 and the Phe65 residues<sup>2,12</sup>. In the Cldn4 structure (PDB ID: 5B2G)<sup>15</sup>, the electrostatic network formed by the acidic residues Glu48, Asp76 and Asp146 stabilizes the stack conformation of Arg31 and Arg158. Conversely, in the Cldn5 sequence, the two Arg residues are converted in the neutral Asn31 and Tyr158 and the Glu48 negatively charged in Cldn4 becomes the basic Lys48 in the Cldn5 primary sequence. Indeed, the salt bridge (SB) between the Lys48 and the Glu76 was identified, together with the interactions between Lys48 and the Tyr67 and Tyr158 sidechains (**Figure S10A**). In addition, the Gln57 residue of Cldn5 and Cldn4 becomes the acidic Asp55 in the mCldn15, which is pivotal for the cation selectivity of the Cldn15-based TJs<sup>11,16</sup>.

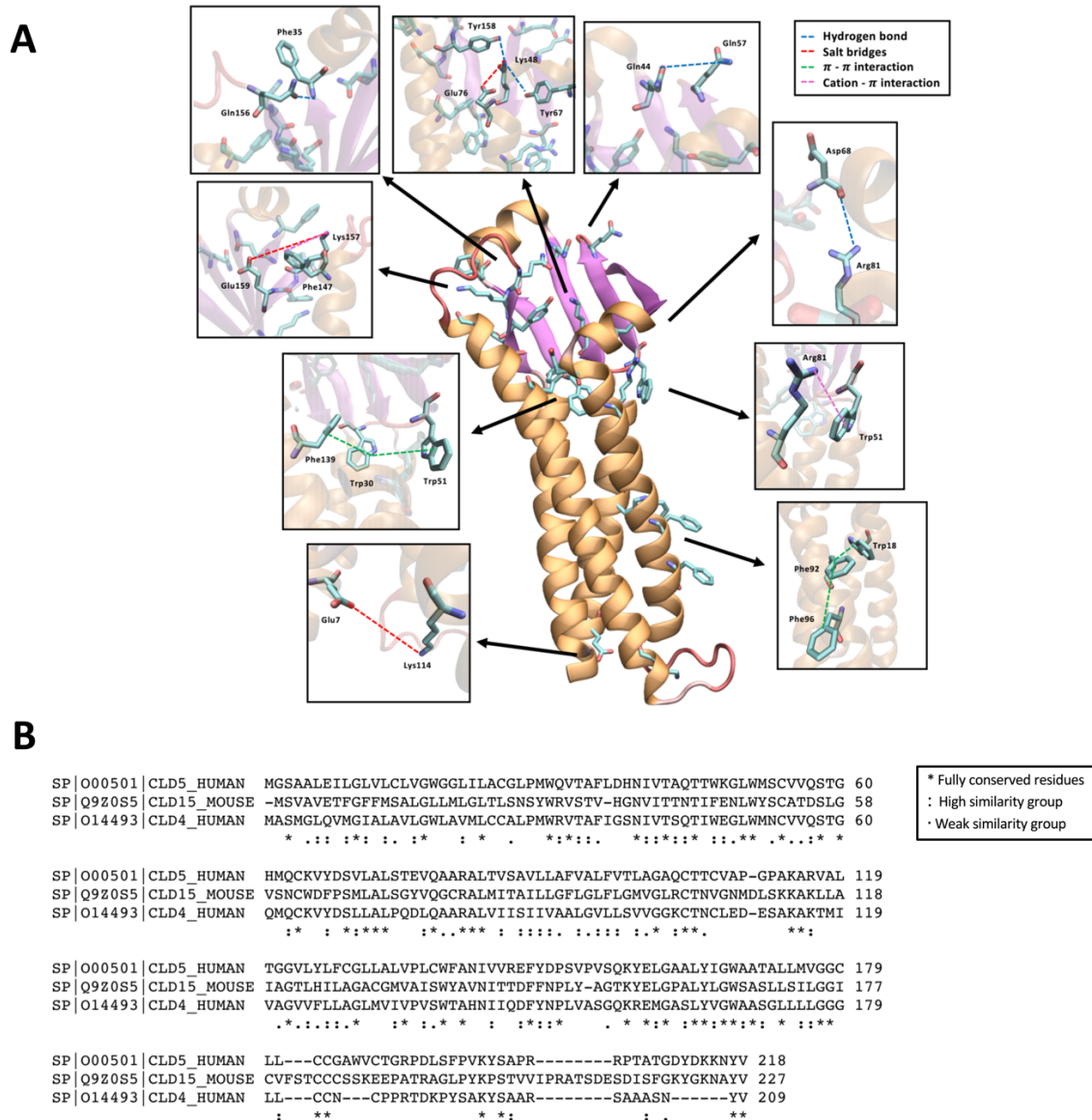
Moreover, other key residues in the ECL domain vary among the three Cldns to differentiate the physiological tasks exerted by the proteins<sup>11,16,17</sup>.

**Table S2.** Residue-residue interactions calculated along the 110-ns-long trajectory of the Cldn5 monomer. Distances are calculated between the heavy atoms at the extremities of the sidechains.

residues	interaction	Cldn5 region	distance (Å)
Phe92 - Trp18	$\pi$ - $\pi$	TM2 - TM1	5.33 $\pm$ 0.48
Phe96 - Phe92	$\pi$ - $\pi$	TM2 - TM2	6.49 $\pm$ 0.58
Phe139 - Trp30	$\pi$ - $\pi$	ECL2 - ECL2	6.40 $\pm$ 0.31
Trp30 - Trp51	$\pi$ - $\pi$	ECL1 - ECL1	8.32 $\pm$ 0.41
Arg81 - Trp51	cation - $\pi$	TM2 - ECL1	5.70 $\pm$ 0.63
Lys157 - Phe147	cation - $\pi$	ECL2 - TM3	5.51 $\pm$ 1.48
Lys48 - Glu76	SB	ECL1 - TM2	6.94 $\pm$ 1.59
Lys157 - Glu159	SB	ECL2 - ECL2	5.79 $\pm$ 2.16
Glu7 - Lys114	SB	TM1 - TM3	7.88 $\pm$ 0.45
Glu44 - Glu57	HB	ECL1 - ECL1	5.87 $\pm$ 1.86
Lys48 - Tyr67	HB	ECL1 - ECL1	4.91 $\pm$ 0.90
Lys48 - Tyr158	HB	ECL1 - ECL1	4.41 $\pm$ 0.89
Asp68 - Arg81	HB	ECL2 - TM2	6.84 $\pm$ 1.82
Phe35 - Gln156	HB	ECL1 - ECL2	3.44 $\pm$ 0.51



**Figure S9.** RMSD of the Cldn5 backbone.



**Figure S10.** A) Representation of the *Cldn5* structure with enlarged views of the residues involved in the aforementioned interactions. B) Multiple sequence alignment among *hCldn5*, *mCldn15* and *hCldn4*.

## References

- (1) Alberini, G.; Benfenati, F.; Maragliano, L. A Refined Model of Claudin-15 Tight Junction Paracellular Architecture by Molecular Dynamics Simulations. *PLOS ONE* **2017**, *12* (9), e0184190. <https://doi.org/10.1371/journal.pone.0184190>.
- (2) Suzuki, H.; Nishizawa, T.; Tani, K.; Yamazaki, Y.; Tamura, A.; Ishitani, R.; Dohmae, N.; Tsukita, S.; Nureki, O.; Fujiyoshi, Y. Crystal Structure of a Claudin Provides Insight into the Architecture of Tight Junctions. *Science* **2014**, *344* (6181), 304–307. <https://doi.org/10.1126/science.1248571>.
- (3) Suzuki, H.; Tani, K.; Tamura, A.; Tsukita, S.; Fujiyoshi, Y. Model for the Architecture of Claudin-Based Paracellular Ion Channels through Tight Junctions. *J. Mol. Biol.* **2015**, *427* (2), 291–297. <https://doi.org/10.1016/j.jmb.2014.10.020>.
- (4) Krause, G.; Protze, J.; Piontek, J. Assembly and Function of Claudins: Structure-Function Relationships Based on Homology Models and Crystal Structures. *Semin. Cell Dev. Biol.* **2015**, *42*, 3–12. <https://doi.org/10.1016/j.semcdb.2015.04.010>.

- (5) Smart, O. S.; Breed, J.; Smith, G. R.; Sansom, M. S. A Novel Method for Structure-Based Prediction of Ion Channel Conductance Properties. *Biophys. J.* **1997**, *72* (3), 1109–1126. [https://doi.org/10.1016/S0006-3495\(97\)78760-5](https://doi.org/10.1016/S0006-3495(97)78760-5).
- (6) Smart, O. S.; Neduvilil, J. G.; Wang, X.; Wallace, B. A.; Sansom, M. S. P. HOLE: A Program for the Analysis of the Pore Dimensions of Ion Channel Structural Models. *J. Mol. Graph.* **1996**, *14* (6), 354–360. [https://doi.org/10.1016/S0263-7855\(97\)00009-X](https://doi.org/10.1016/S0263-7855(97)00009-X).
- (7) Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys.* **1977**, *23* (2), 187–199. [https://doi.org/10.1016/0021-9991\(77\)90121-8](https://doi.org/10.1016/0021-9991(77)90121-8).
- (8) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* **1992**, *13* (8), 1011–1021. <https://doi.org/10.1002/jcc.540130812>.
- (9) Souaille, M.; Roux, B. Extension to the Weighted Histogram Analysis Method: Combining Umbrella Sampling with Free Energy Calculations. *Comput. Phys. Commun.* **2001**, *135* (1), 40–57. [https://doi.org/10.1016/S0010-4655\(00\)00215-0](https://doi.org/10.1016/S0010-4655(00)00215-0).
- (10) Smith, L. G.; Tan, Z.; Spasic, A.; Dutta, D.; Salas-Estrada, L. A.; Grossfield, A.; Mathews, D. H. Chemically Accurate Relative Folding Stability of RNA Hairpins from Molecular Simulations. *J. Chem. Theory Comput.* **2018**, *14* (12), 6598–6612. <https://doi.org/10.1021/acs.jctc.8b00633>.
- (11) Alberini, G.; Benfenati, F.; Maragliano, L. Molecular Dynamics Simulations of Ion Selectivity in a Claudin-15 Paracellular Channel. *J. Phys. Chem. B* **2018**, *122* (48), 10783–10792. <https://doi.org/10.1021/acs.jpcc.8b06484>.
- (12) Irudayanathan, F. J.; Trasatti, J. P.; Karande, P.; Nangia, S. Molecular Architecture of the Blood Brain Barrier Tight Junction Proteins--A Synergistic Computational and In Vitro Approach. *J. Phys. Chem. B* **2016**, *120* (1), 77–88. <https://doi.org/10.1021/acs.jpcc.5b09977>.
- (13) Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Weizhong, L.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J. D.; Higgins, D. G. Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7* (1), 539. <https://doi.org/10.1038/msb.2011.75>.
- (14) Sievers, F.; Higgins, D. G. Clustal Omega for Making Accurate Alignments of Many Protein Sequences. *Protein Sci. Publ. Protein Soc.* **2018**, *27* (1), 135–145. <https://doi.org/10.1002/pro.3290>.
- (15) Shinoda, T.; Shinya, N.; Ito, K.; Ohsawa, N.; Terada, T.; Hirata, K.; Kawano, Y.; Yamamoto, M.; Kimura-Someya, T.; Yokoyama, S.; Shirouzu, M. Structural Basis for Disruption of Claudin Assembly in Tight Junctions by an Enterotoxin. *Sci. Rep.* **2016**, *6*, 33632. <https://doi.org/10.1038/srep33632>.
- (16) Samanta, P.; Wang, Y.; Fuladi, S.; Zou, J.; Li, Y.; Shen, L.; Weber, C.; Khalili-Araghi, F. Molecular Determination of Claudin-15 Organization and Channel Selectivity. *J. Gen. Physiol.* **2018**, *150* (7), 949–968. <https://doi.org/10.1085/jgp.201711868>.
- (17) Hou, J.; Renigunta, A.; Yang, J.; Waldegger, S. Claudin-4 Forms Paracellular Chloride Channel in the Kidney and Requires Claudin-8 for Tight Junction Localization. *Proc. Natl. Acad. Sci.* **2010**, *107* (42), 18010–18015. <https://doi.org/10.1073/pnas.1009399107>.