



## Predicting selectivity of paracellular pores for biomimetic applications

Journal:	<i>Molecular Systems Design &amp; Engineering</i>
Manuscript ID	ME-ART-12-2019-000177.R1
Article Type:	Paper
Date Submitted by the Author:	12-Jan-2020
Complete List of Authors:	Rajagopal, Nandhini; Syracuse University, Department of Biomedical and Chemical Engineering Durand, Alejandro; Syracuse University, Department of Biomedical and Chemical Engineering Nangia, Shikha; Syracuse University, Department of Biomedical and Chemical Engineering

SCHOLARONE™  
Manuscripts

**Design, System, Application**

Biological systems, through millions of years of evolution, have perfected size-selective transport of solute through cell and tissue barriers. Our focus here is to develop biomimetic ion channels based on paracellular transport through tight junctions. The tight junctions are a network of membrane proteins that undergo association to create pore and barriers in paracellular space. Using advanced molecular simulation techniques and big data analysis of protein-protein interactions, we obtained the tight junction architecture for three different proteins. Based on the nature of the protein, we predicted the charged residues responsible for the cation or anion-selectivity of the tight junctions. Our results are in excellent agreement with experiment data, and we anticipate that this approach will be instrumental in developing on-demand ion channels for ultrafiltration applications.

## Predicting selectivity of paracellular pores for biomimetic applications

Nandhini Rajagopal, Alejandro J Durand, and Shikha Nangia\*

*Department of Biomedical and Chemical Engineering, Syracuse University, Syracuse, NY 13244*

**Abstract:** Biological systems exhibit diverse examples of controlled solute permeability and selectivity in cell and tissue barriers. The epithelial and endothelial cells lining each organ confer selectivity via tight junctions, physical fence-like structures that regulate paracellular transport, whose primary functional component is a claudin. Members of the claudin family of proteins undergo cis and trans assembly to control paracellular selectivity. However, based on the type of claudin and its expression level in a tissue, the tight junction selectivity varies from cationic to anionic or permeability changes from zero to leaky. In vitro and in vivo characterization of tight junction macroassemblies is a challenge, especially when molecular-level precision is essential for using nature's design principle for biomimetic applications, such as ion separation platforms and nanosensors. In the present work, we use a recently developed method, protein association energy landscape (PANEL), to exploit the cis architecture of claudin proteins to explain their paracellular selectivity. Using PANEL, we generated millions of claudin-claudin dimer geometries and analyzed the amino acid residue contacts. We demonstrate that a rigorous analysis of cis architectures can not only predict the critical residues responsible for tight junction selectivity, but the cis structures obtained can also provide putative tight junction pore configurations. A deeper understanding of tight junction architecture at a molecular level has been possible using specially designed computational tools and techniques. This approach has promise in determining the selectivity of tight junction proteins and their subsequent use in biomimetic ultrafiltration devices.

\*Address for correspondence:

Dr. Shikha Nangia

343 Link Hall

Department of Biomedical and Chemical Engineering

Syracuse University, Syracuse, NY 13244, USA

Phone (315) 443 0571 | Email: [snangia@syr.edu](mailto:snangia@syr.edu)

## 1. INTRODUCTION

The design of biomimetic sensors and filtration systems requires fundamental knowledge of features critical to natural biological transport processes.<sup>1-4</sup> Separation methods rely on charge and size-selective channels. Most biomimetics focus on the specialized active transporters responsible for transcellular channels (through the cell).<sup>4-13</sup> There has been limited success in exploring the passive and gradient-dependent paracellular (between cells) channels (Figure 1a), where solutes move through hydrophilic channels between cells and not across the hydrophobic lipid bilayer.<sup>14,15</sup>

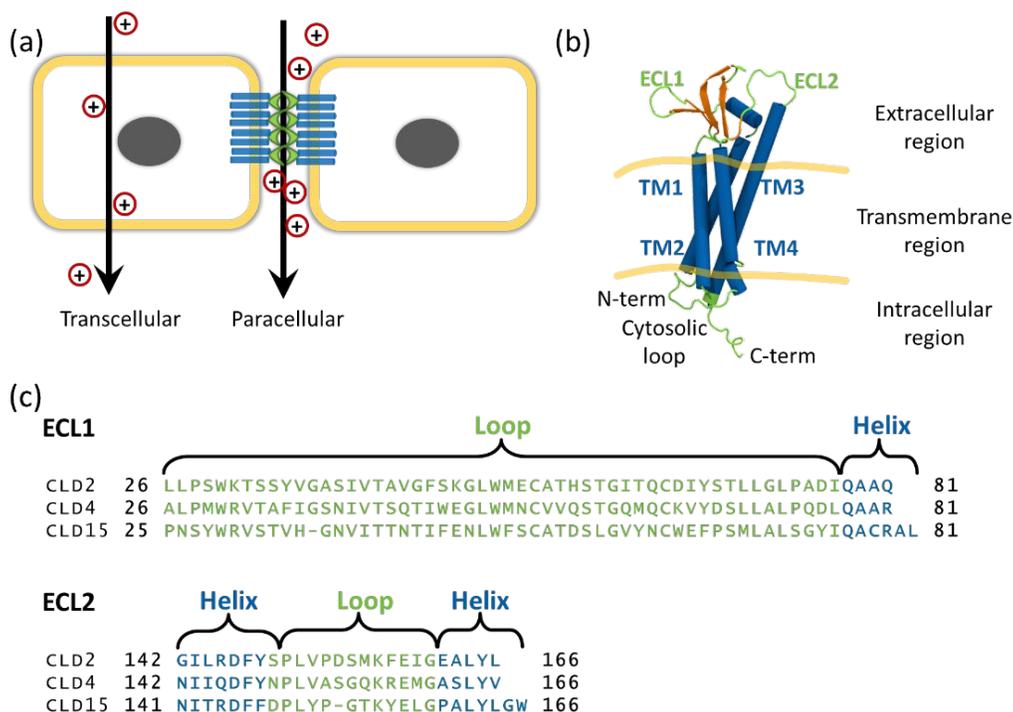


Figure 1. (a) Pathways for solute transport across epithelial and endothelial cells. Paracellular charge-selective ion channels are formed via a network of cis and trans interacting claudin proteins. (b) Structure of claudin protein (cartoon representation) showing four transmembrane helices (TM1-4), extracellular loop (ECL1 and ECL2) domains, cytosolic loop domain, along with N-terminal and C-terminal domains. (c) Sequence alignment of ECL1 and ECL2 regions (green) and TM regions (blue) of human claudin-2, claudin-4, and claudin-15.

The tight junctions are present in epithelia and endothelia that compartmentalize the organs and body cavities in multicellular organisms.<sup>16-18</sup> Besides being the physical and chemical barriers, the epithelial layer allows selective permeability of ions and solutes in a tissue-specific manner.<sup>14,16,19,20</sup> The functional gatekeepers of the paracellular transport are the claudin family of tight junction proteins.<sup>15,18-22</sup> Tight junctions are macroassemblies in which each claudin molecule has a central role in establishing the size- and charge-selectivity of the ion permeation channel.<sup>23,24</sup> Variations in the amino acid sequence, residue type, and residue location in the secondary structure, profoundly influences the macromolecular assembly of the tight junction, and therefore its function.

Claudins are transmembrane proteins that contain 207-305 amino acids (~21-34 kDa) and fold into four helices (TM1-4) with two extracellular loops (ECL1-2), and a cytosolic loop (Figure 1b).<sup>25</sup> The claudin proteins interact laterally (cis assembly) within the membrane of the same cell, followed by head-on (trans assembly) with claudins of the opposing cell to form a network of tight junctions.<sup>26,27</sup> Thus far, 27 members of the claudin family have been identified in mammals.<sup>28</sup> Sequence alignment of a subset of human claudin proteins shows a longer ECL1 loop and a shorter ECL2 loop that have several acidic and basic amino acid residues (Figure 1c).

Based on permeability, claudins are broadly classified as pore-forming (solute permissive) or barrier-forming (solute restrictive).<sup>17,28</sup> However, despite the classification, a claudin can be a pore for an ion or solute, while being a barrier to a different solute given its charge or size.<sup>18,23</sup> Biochemical analyses of claudin proteins have correlated the paracellular ion selectivity to the protein sequence in the ECL1 domain. For example, claudin-2 and -15, which both have a net negative charge in the ECL1 domain have been reported to be cation selective with pore diameters in 0.8-1.0 nm,<sup>19,29-31</sup> whereas claudins with a net positively charged ECL1 are reported as anion selective.<sup>32-35</sup> In addition, the TM domains anchor the claudins in the lipid bilayer and facilitate macromolecular cis assembly that is visible in freeze fracture monographs as a network of claudin strands. In an earlier study by Colegio *et al.*, the authors used claudin-2 and -4 chimeras to conclude that the ECL domains determine paracellular charge selectivity but not tight junction fibril architecture.<sup>36</sup>

Despite the fundamental knowledge that the two ECL domains form cis and trans associations, it is unclear which segments of and what proportion of ECL1 and ECL2 domains participate in cis versus trans associations. In this work, we utilize a large claudin dimer data set obtained from the potential energy landscape to compute the cis binding probability of each amino acid residue. We used normalized contact analysis to categorize the residues as high, medium, and low probability for cis interaction. The normalized contacts also show which of the two ECLs in a claudin is engaged in cis formation. To predict the residues responsible for ion selectivity of the pore, we focused on the charged residues with low cis contacts. These were further refined to eliminate the sterically buried residues and to identify residues that were exposed to the solvent. Using our systematic approach, we present our findings for claudins-2, -4, and -15 that are bona fide pore-forming claudin proteins for small ions.<sup>29-32,37-40</sup> In each case, our predictions are in excellent agreement with the biochemical and computational studies reported in the literature.

## 2. METHODS

**Claudin structures.** The claudin structures were obtained by homology modeling using crystal structures claudin-15 (4P79),<sup>41</sup> claudin-19 (3X29),<sup>42</sup> claudin-4 (5B2G),<sup>43</sup> and claudin-3 (6AKG)<sup>44</sup> as described previously.<sup>45-48</sup> Structures were then minimized and equilibrated using the CHARMM36 force field.<sup>49-51</sup> Each protein was individually embedded in 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC) bilayer and surrounded by water with 0.15 M NaCl using the CHARMM-GUI server. Each system was then allowed to undergo equilibration, first in isochoric-isothermal (*NVT*) ensemble conditions for 100 ns and then second in the isobaric-isothermal (*NPT*) ensemble conditions for 100 ns. Pressure was maintained at 1 bar using Parrinello-Rahman barostat<sup>52</sup> with  $\tau_P = 1$  ps. Temperature was maintained at 310 K using Nose-Hoover thermostat<sup>53</sup> and  $\tau_T = 5$  ps. The equilibrated structures were coarse grained (CG) using standard MARTINI coarse graining<sup>54-56</sup> protocol with EINEDyn network and side-chain fixes.<sup>57,58</sup>

**PANEL.** The coarse grained claudin structures were used as inputs to generate their minimum energy association landscapes using the PANEL (Protein Association Energy Landscape) method, developed by our group.<sup>59</sup> The PANEL method takes the CG protein structure as input to generate multiple dimeric seed geometries (~2000-2500) that are selected randomly from a uniform distribution of the dimeric rotational space.

Table 1. Parameters for PANEL runs

PANEL parameters	Claudin-2	Claudin-15	Claudin-4
$d_s$ (nm)	3.4	3.2	3.1
$N$	2,483	2,352	2,000
$M$ (ns)	500	500	500
Total data set ( $N \times M$ )	1,241,500	1,176,000	1,000,000
Selected data	867,410	879,548	916,327
Coverage (%)	88.5	92.9	95.9
Average grid sampled	6.7	6.8	7.1

Using the automated PANEL scripts, each seed geometry was embedded in DOPC bilayer, energy minimized, and equilibrated for 75 ns. The *NVT* and *NPT* equilibration steps were performed using v-rescale

thermostat and Berendsen barostat,<sup>60</sup> respectively. Temperature was maintained at 310 K and pressure at 1 bar ( $\tau_T = 1$  ps,  $\tau_P = 5$  ps). The production run was performed for 500 ns using the v-rescale thermostat<sup>61</sup> and Parrinello-Rahman barostat<sup>52</sup> ( $\tau_T = 1$  ps,  $\tau_P = 12$  ps). The PANEL input parameters for each claudin system are provided in Table 1. The separation distance ( $d_s$ ) for each claudin was computed before the initializing the PANEL scripts. The number of seed geometries ( $N$ ) for each system was in the 2000-2500 range to maximize the energy landscape coverage. Each seed geometry was simulated for 500 ns, and the protein-protein interaction energies were recorded every 1 ns. The total dataset ( $N \times M$ ) from the production run, was analyzed to remove any spurious non-interacting conformations that do not represent cis interaction. The remaining data were then selected to generate the PANEL plots. These interaction energies of the selected data were projected on the 2D rotational space to generate the protein association energy landscape. The energy landscape grid coverage ranged from 89-96%, which implied that each grid point was sufficiently sampled with an average of  $\sim 7$  samples per grid (Table 1). The PANEL method allows easy identification of stable dimer orientations from the lowest energy regions of the energy landscape, which can then be used to analyze amino acid contacts involved in dimer formation.

**Contact Analysis.** The minimum energy configurations corresponding to the PANEL plots were sorted based on their energies, and the configurations falling in the lowest 33% of the energy scale were selected to represent the stable set of claudin dimer configurations. The selected configurations were then used to obtain pair-wise residue between the interacting claudins. The amino acid residue contacts were determined by computing the distance between the CG backbone beads of every amino acid in either protein. All inter-protein backbone beads within 6 Å distance were assigned to be in contact. Further, the contacts were normalized relative to the largest contacting residue to be in 0 to 1 range. The normalized contacts were then labeled as high ( $>0.66$ ), moderate (0.33–0.66), and low ( $<0.33$ ). The normalized contact analysis was performed using in-house python scripts and MDAnalysis package.<sup>62,63</sup>

**Pore structure and analysis.** The atomistic pore I and II structures of claudin-5 from our earlier work were used as templates to generate putative pore models for the claudins-2, -4, and -15.<sup>45-47</sup> The pore modeling and alignment were performed using YASARA, molecular-graphics, modeling, and simulation software.<sup>64,65</sup> The structures were energy minimized using steepest descent with periodic boundary conditions in YASARA software. PyMol software was used to analyze the final pore structures and to visualize the pore-lining charged residues.<sup>66</sup> The inner diameter along the length of the pore was estimated using the CAVER plugin.<sup>67</sup>

### 3. APPROACH

We have adopted a systematic bottom-up approach to predict the paracellular ion selectivity of tight junction pores. We hypothesized that charged, surface-exposed, pore-lining residues in the extracellular domains of tight junction proteins are key to the charge/ion selectivity of the tight junction pores. To test our hypothesis, we started with the evaluation of the amino acid residue sequence of the protein, and based on the three-dimensional structure of the protein, we identified the charged residues that are exposed to the surface (Figure 2). By focusing only on these surface-exposed residues, we aimed to pinpoint those residue(s) that, despite protein oligomerization (cis and trans), remain available to participate in pore

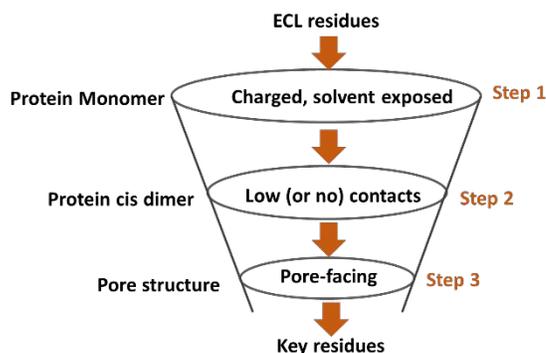


Figure 2. Flowchart of the three-step approach used to predict the key residues that determine the paracellular ion selectivity of tight junction pores.

selectivity. In the second step, we evaluated the participation of the residues in cis interactions to pick out the residues that form the least number of cis contacts. The study of cis contacts was performed using the stable, low-energy dimer conformations obtained from the PANEL method. From the first set of residues, we identified the subset of residues with few or zero cis contacts. Finally, based on the predicted structure of the pore, we further narrowed down the residues from the second step and predicted the pore-facing residues that were involved in the paracellular ion transport.

## 4. RESULTS AND DISCUSSION

### 4.1. Claudin-2 paracellular cation selectivity is determined by D65 and D146 residues

The PANEL minimum energy plot shows the landscape of cis interacting claudin-2 dimers as a function of the rotational angles ( $\theta$ ,  $\theta'$ ). The details of PANEL implementation and the definition of the rotational angles

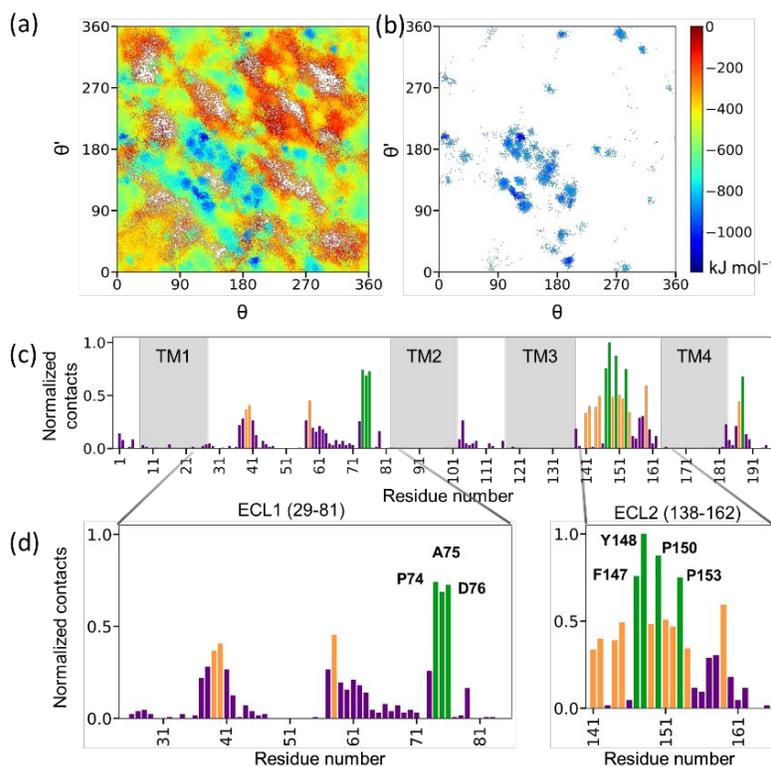


Figure 3. Claudin-2 cis interaction analysis. (a) Minimum energy PANEL plot as a function of ( $\theta$ ,  $\theta'$ ) rotational angles; (b) Regions in the landscape that have interaction energy below  $-797 \text{ kJmol}^{-1}$  cutoff. The energy scale on the plots ranges from 0 (red) to  $-1196 \text{ kJmol}^{-1}$ . (c) Normalized residue-residue contacts for amino acid residues (1-198) and their location in TM1-4 and ECL1-2 domains. (d) Enlarged view of normalized contacts in the ECL1 and ECL2 regions. The contacts are categorized as high (green), medium (orange), and low (purple).

were reported in our earlier work.<sup>59</sup> To generate claudin-2 PANEL plots,  $8.7 \times 10^5$  dimer conformations were sampled (Table 1). The interaction energy of these dimer conformations ranged from 0 to  $-1196 \text{ kJmol}^{-1}$  representing non-interacting to highly stable dimer conformations (Figure 3a). Although each dimer conformation is significant, our goal was to identify low-energy, stable cis interacting claudin-2 partners. Therefore, we focused on conformations that have interaction energies below the  $-797 \text{ kJmol}^{-1}$  cut-off (or lower one-third of the energy range). On the claudin-2 PANEL plot, these low energy conformations are primarily clustered around  $\theta$  and  $\theta'$  ranging from  $90^\circ$  to  $120^\circ$  (Figure 3b). Identifying these stable cis interacting conformations provides unprecedented insights into the multiple ways two claudins-2 proteins can interact with each other. Obtaining such detailed and systematic information is impossible in biochemical assays. Furthermore, analyzing these cis geometries provides a quantitative assessment of the residue-residue contacts that form these stable conformations.

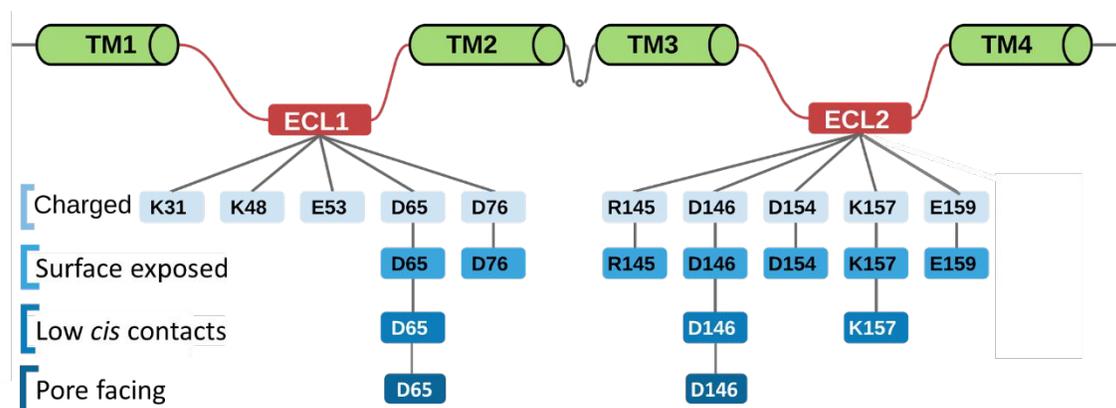


Figure 4. Systematic analysis of the ECL1 and ECL2 domain residues for claudin-2. The method predicts D65 and D146 as the key residues for claudin-2 pore selectivity.

The residue contact analysis shows that some residues are more heavily involved in cis interactions than others. In terms of the claudin-2's secondary structure, the ECL domains form appreciably higher contacts than the membrane-embedded TM domains (Figure 3c). In particular, the ECL2 residues F147, Y148, P150, and P153 show high normalized contacts ( $>0.66$ ) along with other residues with a moderate number of contacts. Interestingly, the majority of the ECL1 residues have low contact, except P74, A75, and D76. The TM4 domain has one or two residues with high contacts. The overall contact analysis shows that ECL2 is the primary cis binding domain for claudin-2.

The non-cis binding residues can potentially engage in other structural and functional roles in the tight junctions trans interactions, pore-lining residues, remain solvent exposed, or participate in the channel's ion selectivity (if charged). Narrowing down to surface exposed ECL residues, reveals D65 in ECL1 and D146 and K157 in the ECL2 domains (Figure 4). The location of D65 and D146 strongly suggests their availability for dictating the charge selectivity or a potential involvement in trans interactions.

Using claudin-5 pore I and II models as templates from our previous work, we generated putative pore models for claudin-2 (Figure 5).<sup>46</sup> In the pores, the D65 and D146 residues line the pore cavity, whereas the positively charged side chains of K157 face outwards suggesting their non-involvement in the pore selectivity. In pore I, D65 is in the constricted neck region of the pore, and D146 lies near the wider opening of the pore, with both providing selectivity for cation transport. Conversely, in the pore II structure, D146 lies in vicinity of the pore neck, while D65 guards the mouth of the pore. This presents the possibility of a

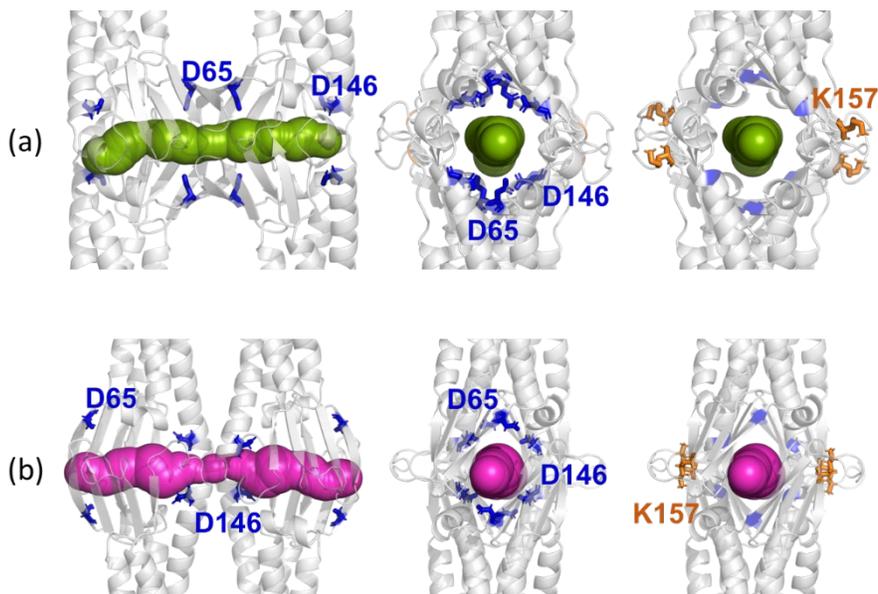


Figure 5. Claudin-2 pore models showing predicted residues for paracellular charge selectivity. The longitudinal and transverse views of (a) pore I channel (green) and (b) Pore II channel (pink) are shown along with pore-lining residues D65 and D146 (blue) and non-pore lining K157 (orange) residues.

continuous conduction of cations through channels with variable morphologies, yet similar charge selectivity in complementary positions.

Our results based on data analysis of claudin-2 cis interactions are in excellent agreement with the experimental findings. In a series of molecular studies of claudin-2 and -4 chimeras, swapping the ECL1 domains altered the paracellular charge selectivity, demonstrating that ECL1 domain plays a pivotal role in ion permeation.<sup>36</sup> Similar studies involving ECL2 chimeras showed little difference in ion permeation. Furthermore, mutating D65 with a positively charged residue markedly reduced its cation selectivity.<sup>24,68,69</sup> Mutation of D65N exhibits threefold lower conductance and Na<sup>+</sup> permeability compared to wildtype, and no change in Cl<sup>-</sup> permeability relative to the wildtype claudin-2.<sup>69</sup> Despite experimental studies indicating non-involvement of ECL2 residues in pore selectivity, our work predicts possible involvement of D146 in pore selectivity. It is possible that the considerable involvement of ECL2 in cis interactions screens D146 to a certain extent; however, further experimental work needs to be performed to compare the effects of D146 mutations.

### 3.2. Claudin-15 paracellular cation selectivity is determined by D55 and D148 residues

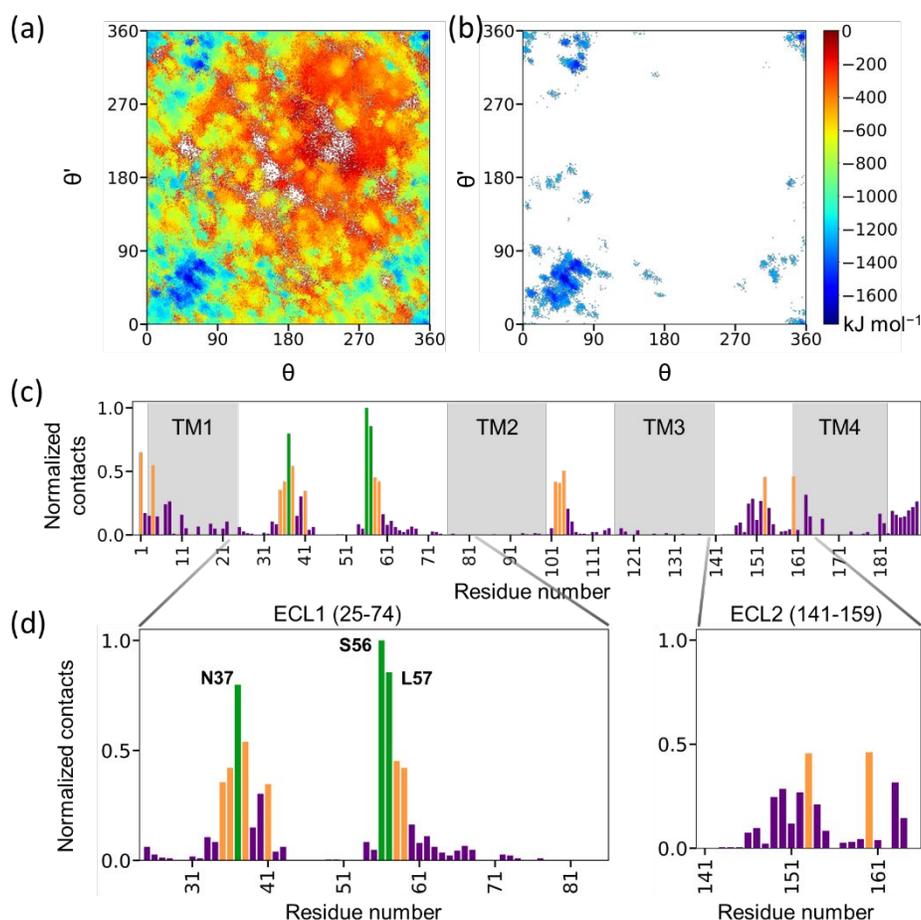


Figure 6. Claudin-15 cis interaction analysis. (a) Minimum energy PANEL plot as a function of ( $\theta$ ,  $\theta'$ ) rotational angles; (b) Regions in the landscape that have interaction energy below  $-1184.5 \text{ kJmol}^{-1}$  cutoff. The energy scale on the plots ranges from 0 (red) to blue ( $-1776.82 \text{ kJmol}^{-1}$ ). (c) Normalized residue-residue contacts for amino acid residues (1-190) and their location in TM1-4 and ECL1-2 domains. (d) Enlarged view of normalized contacts in the ECL1 and ECL2 regions. The contacts are categorized as high (green), medium (orange), and low (purple).

In evaluating the paracellular ion selectivity of claudin-15 channels, our aim was to predict the residue(s) responsible for the high permeability of cations while being highly resistant barriers to anions. Using the steps outlined above, we generated the PANEL minimum energy plot that for claudin-15 dimers. A total of

$8.8 \times 10^5$  dimer conformations were sampled to generate the PANEL plot as function of rotational angles ( $\theta$ ,  $\theta'$ ). The interaction energy of the dimer conformations ranged from 0 to  $-1776.82 \text{ kJmol}^{-1}$  (Figure 6a). The claudin-15 PANEL shows a low energy region (below  $-1184.5 \text{ kJmol}^{-1}$  cutoff) of cis interactions concentrated in the range of  $0^\circ$ - $90^\circ$  for both  $\theta$  and  $\theta'$  rotational angles (Figure 6b). Additional low energy regions are scattered over other parts of the landscape.

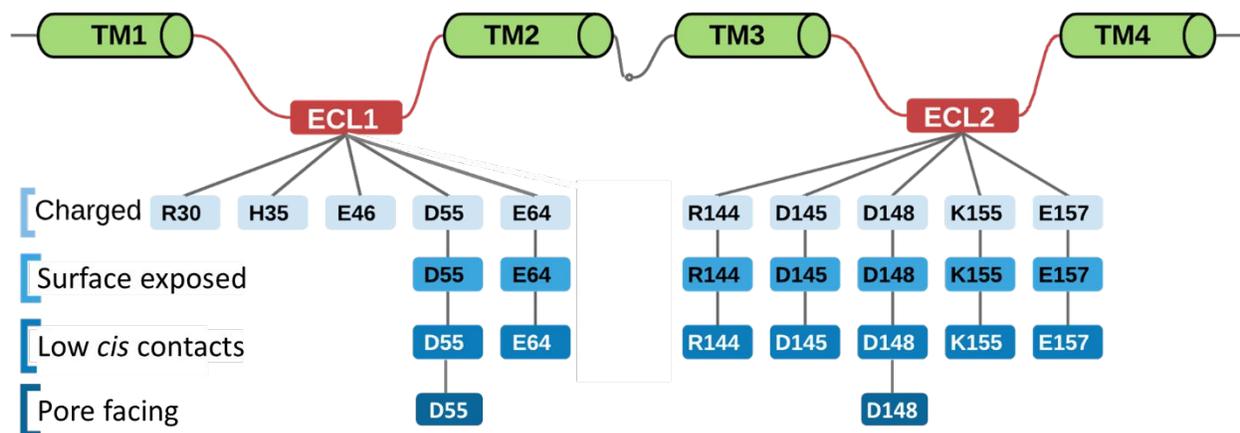


Figure 7. Systematic analysis of the ECL1 and ECL2 domain residues for claudin-15. The method predicts D55 and D148 as the key residues for claudin-15 pore selectivity.

These regions correspond to strong interactions mediated by TM1 and ECL1 with some contribution from the TM4 helix. The ECL1 residues N37, S56, and L57 have the highest contacts. Of note, the highest contributors to cis interactions were from ECL1 domain compared to ECL2 (Figure 6c). This observation challenges the notion that claudin ECL2 domains are primarily engaged in cis assembly. Multiple *in vitro* transepithelial electrical resistance (TEER) measurements have shown that changes in ECL1 domain residues alone can explain changes in the paracellular charge selectivity without any support from the ECL2 domains.<sup>36</sup> There is, however, no direct biochemical evidence to support that the primary function of ECL2 is to participate in formation of cis strands.

The ECL1 domain, in addition to having cis contacts, has multiple surface exposed, charged residues: D55, E64, R144, D145, D148, K155 and E157 (Figure 7). To identify residues contributing to ion selectivity, we generated the pore I and pore II structures based on our previous work.<sup>46</sup> The pore II structure, however, was occluded with no clear channel; it was not analyzed further. In pore I, D55, R144, D145, and D148 were pore facing among the charged residue. The remaining charged residues E64, K155, and K157 face away from the pore, and therefore, do not participate in paracellular selectivity (Figure 8). Further, the side chains of oppositely charged R144 and D145 residues (within 4 Å radius) form a salt bridge interaction

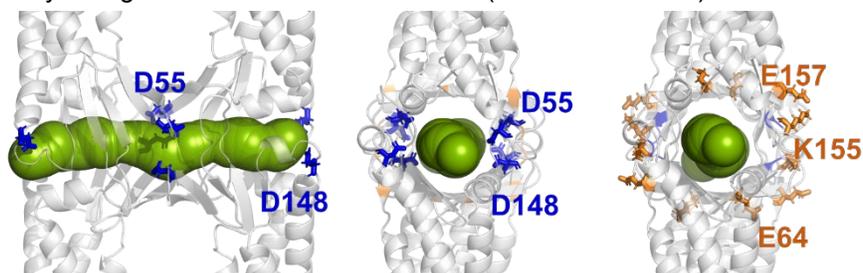


Figure 8. Claudin-15 pore models showing predicted residues for paracellular charge selectivity. The longitudinal and transverse views of pore I channel (green) along with pore-lining residues D55 and D148 (blue) and outward facing E64, K155, and E157 (orange) residues.

rendering themselves ineffective in influencing the pore selectivity. In terms of location, D55 lies in the middle of the permeation path, while D148 is close to the pore entrance. The symmetry of the pore places four D55 residues in the center of the ion permeation pathway (Figure 8), and a pair of D148 residues on

either side of the pore entrance. Together these aspartic acid residues provide a net negative charge in the pore cavity resulting in cation selectivity.

In a recent study, the key role of the D55 residue was highlighted in in vitro and computational studies of ion conductance through claudin-15 channels.<sup>70,71</sup> Using multiple mutant systems (E46K/D55K/E64K) Samanta *et al.* demonstrated that charge reversal of D55 reduced cation permeability, while mutation of E46 and E64 has little influence. In another report, Alberini *et al.* confirmed that E64 side chain is external to the pore cavity.<sup>72</sup>

Although the PANEL plot shows high stability in the cis dimeric regions corresponding to pore I model, the pore II forming cis dimer regions were found to be highly unstable. Consequently, we found that the docked pore I structure shows a significant opening of the pore pathway while pore II docking did not result in forming an open pore but had a block at the neck region. This suggests that the inferences made from PANEL plot regarding stable cis dimer regions may be extended to give a perspective about the possible trans interactions and nature of pore morphologies in claudins. Further, all the predicted acidic residues show a favorable arrangement for cation transport by forming a continuous lining along the length of the pore.

### **3.3. Claudin-4 paracellular anion selectivity is influenced by K65 and R158 residues**

The PANEL minimum energy plot shows the landscape of cis interacting claudin-4 dimers as function of the rotational angles ( $\theta$ ,  $\theta'$ ). A total of  $9.1 \times 10^5$  dimer conformations were sampled to generate the PANEL

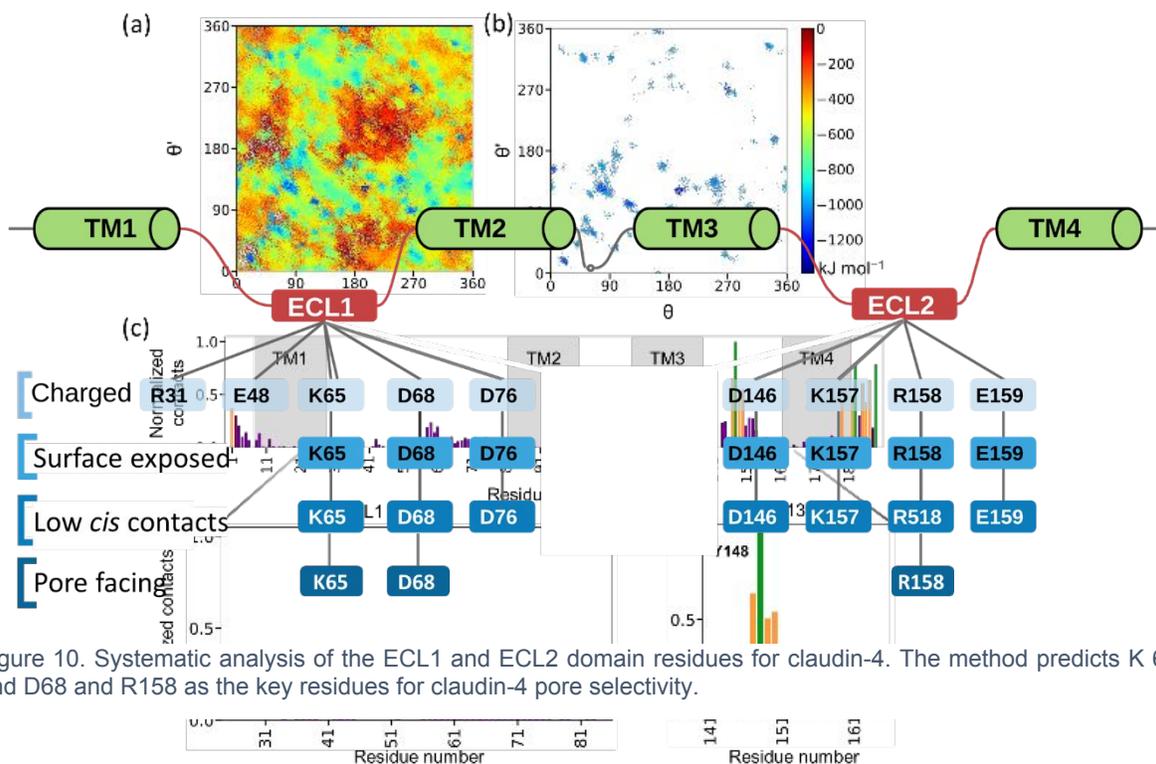


Figure 10. Systematic analysis of the ECL1 and ECL2 domain residues for claudin-4. The method predicts K 65 and D68 and R158 as the key residues for claudin-4 pore selectivity.

plots. The interaction energy of the dimer conformations ranged from 0 to  $-1391 \text{ kJmol}^{-1}$  (Figure 9a). The Figure 9. Claudin-4 cis interaction analysis. (a) Minimum energy PANEL plot as a function of ( $\theta$ ,  $\theta'$ ) rotational angles; (b) Regions in the landscape that have interaction energy below  $-927 \text{ kJmol}^{-1}$  cutoff. The energy scale on the plots ranges from 0 (red) to  $-1391 \text{ kJmol}^{-1}$ . (c) Normalized residue-residue contacts for amino acid residues (1-189) and their location in TM1-4 and ECL1-2 domains. (d) Enlarged view of normalized contacts in the ECL1 and ECL2 regions. The contacts are categorized as high (green), medium (orange), and low (purple).

low energy regions (below the  $-927 \text{ kJmol}^{-1}$  cutoff) were observed in multiple locations over the rotational space (Figure 9b). The low-energy conformations show preferred cis interactions through TM1, TM3 and TM4, in addition to ECL domains (Figure 9c). Like claudin-2, ECL2 shows larger involvement in contact formation in stable cis dimers with highest contacts made by Y148 (Figure 9d). High frequency contacts were also observed in the TM4 region. The amino acid residues least involved in cis dimerization included all of the exposed charged residues in the ECL region. Of these, D76, R157 and E159 were found to be facing away from the pore. The pore-facing residues were K65, D68 and R158, which are available for influencing the nature of the pore (Figure 10).

Like claudin-2 pore morphologies and the predicted key residues, claudin-4 also presents two basic residues K65 and R158, in the pore opening and neck regions of the pores. Their positions are swapped in pore I and pore II. Although there is an inclusion of the acidic residue, D68, it is buried deeper in the ECL

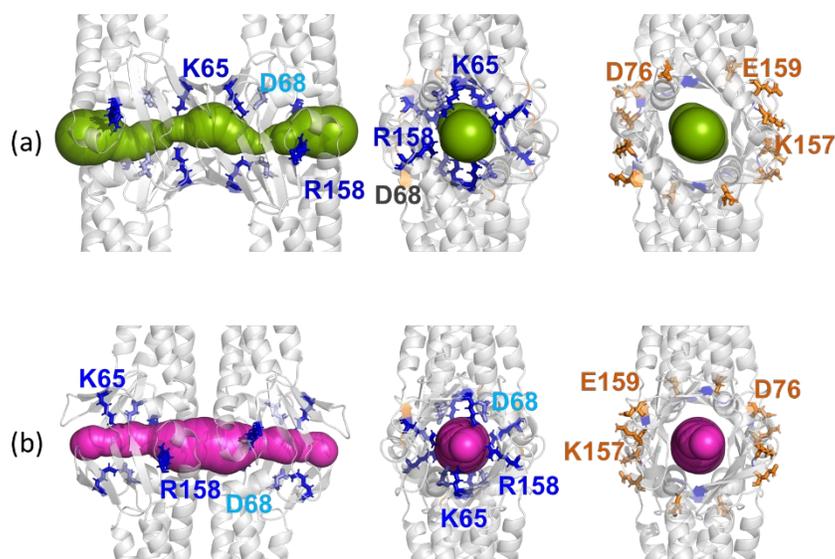


Figure 11. Claudin-4 pore models showing predicted residues for paracellular charge selectivity. The longitudinal and transverse views of (a) pore I channel (green) and (b) Pore II channel (pink) are shown along with pore-lining residues K65 and R158 (blue), D68 (light blue), and non-pore lining D76 and K157 (orange) residues.

region and may be shielded by other nonpolar side chains that restrict the pore size (Figure 11). The presence of stable cis dimers that are precursors to pore I and pore II in claudin-4 PANEL presents the possibility of different pore pathways in claudin-4 (i.e., a basic pore environment supporting an anion selective channel or a cation barrier property).

Claudin-4 is shown to permeate anions and act as barrier to cation transport in several studies.<sup>32,73,74</sup> A K65D mutation resulted in significant reduction in cation barrier.<sup>24,74</sup> In a later study involving paracellular reabsorption of chloride ions in collecting ducts of the kidneys, claudin-4 was shown to form anion channels facilitating the reabsorption of chloride ions.<sup>32</sup> Further emphasizing the significance of K65 is a study in which, claudin-4 lost its anion selectivity upon K65T mutation.<sup>32</sup>

#### 4. CONCLUSIONS

The claudins create a paracellular seal with well-defined pores for transport of ions and solutes. The permeability of tight junctions occurs when specific members of claudin protein family associate via cis and form charge-selective pores. The selectivity of the pore is mediated by charged residues that remain solvent exposed in the pore walls, without becoming engaged in cis claudin macroassembly. We used a systematic approach to predict the key residues responsible for charge-selectivity in three members of the claudin family. The systematic approach involved generating over million geometries of claudin-claudin interaction and evaluating the residue contacts for low energy dimeric structures. Further analysis of non-cis forming charged residues in the putative pore models helped in precisely pinpointing the residues engaged in ion selectivity. We identified D65 and D146 residues in claudin-2, D55 and D148 in claudin-15, and K65 and R158 in claudin-4 to bestow charge-selective properties to the ion permeation channel. The work is in excellent agreement with biochemical studies performed in prior in vitro and computational work. We also show that in the claudin secondary structure only the ECL domains are responsible for ion-selectivity and not the TM domains. Additionally, mutating or abolishing key residues in the ECL domains offer precise control to reverse or regulate the paracellular ion permeation.

## 5. ACKNOWLEDGMENTS

This work is supported by the CAREER CBET-1453312 grant from National Science Foundation. Computational resources were provided by Information and Technology Services at Syracuse University (Eric Sedore, Larne Pekowsky, and Michael R. Brady) as well as the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1053575.

## Conflicts of interest

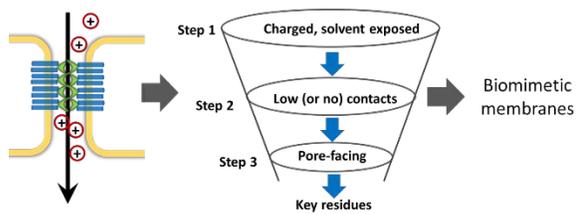
There are no conflicts to declare.

## 6. REFERENCES

- 1 Nielsen, C. H. *Analytical and Bioanalytical Chemistry*, 2009, **395**.
- 2 Shen, Y.-x.; Saboe, P. O.; Sines, I. T.; Erbakan, M.; Kumar, M. *Journal of Membrane Science*, 2014, **454**.
- 3 Chimisso, V.; Maffei, V.; Hürlimann, D.; Palivan, C. G.; Meier, W. *Macromolecular Bioscience*, 2019, **n/a**.
- 4 Hasirci, V.; Vrana, E.; Zorlutuna, P.; Ndreu, A.; Yilgor, P.; Basmanav, F. B.; Aydin, E. *Journal of Biomaterials Science, Polymer Edition*, 2006, **17**.
- 5 Sun, G.; Chung, T.-S.; Jeyaseelan, K.; Armugam, A. *Colloids and Surfaces B: Biointerfaces*, 2013, **102**.
- 6 Xie, W.; He, F.; Wang, B.; Chung, T.-S.; Jeyaseelan, K.; Armugam, A.; Tong, Y. W. *Journal of Materials Chemistry A*, 2013, **1**.
- 7 Zhong, P. S.; Chung, T.-S.; Jeyaseelan, K.; Armugam, A. *Journal of Membrane Science*, 2012, **407-408**.
- 8 Fornasiero, F.; Park, H. G.; Holt, J. K.; Stadermann, M.; Grigoropoulos, C. P.; Noy, A.; Bakajin, O. *Proceedings of the National Academy of Sciences*, 2008, **105**.
- 9 Hilder, T. A.; Gordon, D.; Chung, S.-H. *Nanomedicine: Nanotechnology, Biology and Medicine*, 2011, **7**.
- 10 Hou, X.; Yang, F.; Li, L.; Song, Y.; Jiang, L.; Zhu, D. *Journal of the American Chemical Society*, 2010, **132**.
- 11 Li, X.; Wang, R.; Tang, C.; Vararattanavech, A.; Zhao, Y.; Torres, J.; Fane, T. *Colloids and Surfaces B: Biointerfaces*, 2012, **94**.
- 12 Tang, C. Y.; Zhao, Y.; Wang, R.; Hélix-Nielsen, C.; Fane, A. G. *Desalination*, 2013, **308**.
- 13 Wang, H. L.; Chung, T.-S.; Tong, Y. W.; Jeyaseelan, K.; Armugam, A.; Duong, H. H. P.; Fu, F.; Seah, H.; Yang, J.; Hong, M. *Journal of Membrane Science*, 2013, **434**.
- 14 Madara, J. L. *Annual Review of Physiology*, 1998, **60**.
- 15 Van Itallie, C. M.; Anderson, J. M. *Annual Review of Physiology*, 2006, **68**.
- 16 Krug, S. M.; Schulzke, J. D.; Fromm, M. *Seminars in Cell & Developmental Biology*, 2014, **36**.
- 17 Shen, L.; Weber, C. R.; Raleigh, D. R.; Yu, D.; Turner, J. R. *Annual Review of Physiology*, 2011, **73**.
- 18 Van Itallie, C. M.; Anderson, J. M. *Physiology*, 2004, **19**.
- 19 Krug, S. M.; Günzel, D.; Conrad, M. P.; Lee, I.-F. M.; Amasheh, S.; Fromm, M.; Yu, A. S. L. *Annals of the New York Academy of Sciences*, 2012, **1257**.
- 20 Markov, A. G.; Aschenbach, J. R.; Amasheh, S. *IUBMB Life*, 2015, **67**.
- 21 Angelow, S.; Yu, A. S. L. *Current Opinion in Nephrology and Hypertension*, 2007, **16**.
- 22 Suzuki, H.; Tani, K.; Tamura, A.; Tsukita, S.; Fujiyoshi, Y. *Journal of Molecular Biology*, 2015, **427**.
- 23 Van Itallie, C. M.; Anderson, J. M. *Proceedings of the American Thoracic Society*, 2004, **1**.
- 24 Colegio, O. R.; Van Itallie, C. M.; McCrea, H. J.; Rahner, C.; Anderson, J. M. *American Journal of Physiology-Cell Physiology*, 2002, **283**.
- 25 Morita, K.; Furuse, M.; Fujimoto, K.; Tsukita, S. *Proceedings of the National Academy of Sciences*, 1999, **96**.

- 26 Krause, G.; Winkler, L.; Mueller, S. L.; Haseloff, R. F.; Piontek, J.; Blasig, I. E. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 2008, **1778**.
- 27 Tsukita, S.; Furuse, M.; Itoh, M. *Nature Reviews Molecular Cell Biology*, 2001, **2**.
- 28 Günzel, D.; Yu, A. S. L. *Physiological Reviews*, 2013, **93**.
- 29 Amasheh, S.; Meiri, N.; Gitter, A. H.; Schöneberg, T.; Mankertz, J.; Schulzke, J. D.; Fromm, M. *Journal of Cell Science*, 2002, **115**.
- 30 Tamura, A.; Hayashi, H.; Imasato, M.; Yamazaki, Y.; Hagiwara, A.; Wada, M.; Noda, T.; Watanabe, M.; Suzuki, Y.; Tsukita, S. *Gastroenterology*, 2011, **140**.
- 31 Wada, M.; Tamura, A.; Takahashi, N.; Tsukita, S. *Gastroenterology*, 2013, **144**.
- 32 Hou, J.; Renigunta, A.; Yang, J.; Waldegger, S. *Proceedings of the National Academy of Sciences*, 2010, **107**.
- 33 Krug, S. M.; Günzel, D.; Conrad, M. P.; Rosenthal, R.; Fromm, A.; Amasheh, S.; Schulzke, J. D.; Fromm, M. *Cellular and Molecular Life Sciences*, 2012, **69**.
- 34 Conrad, M. P.; Piontek, J.; Günzel, D.; Fromm, M.; Krug, S. M. *Cellular and Molecular Life Sciences*, 2016, **73**.
- 35 Milatz, S.; Breiderhoff, T. *Pflügers Archiv - European Journal of Physiology*, 2017, **469**.
- 36 Colegio, O. R.; Itallie, C. V.; Rahner, C.; Anderson, J. M. *American Journal of Physiology-Cell Physiology*, 2003, **284**.
- 37 Luettig, J.; Rosenthal, R.; Barmeyer, C.; Schulzke, J. D. *Tissue Barriers*, 2015, **3**.
- 38 Rosenthal, R.; Milatz, S.; Krug, S. M.; Oelrich, B.; Schulzke, J.-D.; Amasheh, S.; Günzel, D.; Fromm, M. *Journal of Cell Science*, 2010, **123**.
- 39 Muto, S.; Hata, M.; Taniguchi, J.; Tsuruoka, S.; Moriwaki, K.; Saitou, M.; Furuse, K.; Sasaki, H.; Fujimura, A.; Imai, M. et al. *Proceedings of the National Academy of Sciences*, 2010, **107**.
- 40 Rosenthal, R.; Günzel, D.; Krug, S. M.; Schulzke, J. D.; Fromm, M.; Yu, A. S. L. *Acta Physiologica*, 2017, **219**.
- 41 Suzuki, H.; Nishizawa, T.; Tani, K.; Yamazaki, Y.; Tamura, A.; Ishitani, R.; Dohmae, N.; Tsukita, S.; Nureki, O.; Fujiyoshi, Y. *Science*, 2014, **344**.
- 42 Saitoh, Y.; Suzuki, H.; Tani, K.; Nishikawa, K.; Irie, K.; Ogura, Y.; Tamura, A.; Tsukita, S.; Fujiyoshi, Y. *Science*, 2015, **347**.
- 43 Shinoda, T.; Shinya, N.; Ito, K.; Ohsawa, N.; Terada, T.; Hirata, K.; Kawano, Y.; Yamamoto, M.; Kimura-Someya, T.; Yokoyama, S. et al. *Scientific Reports*, 2016, **6**.
- 44 Nakamura, S.; Irie, K.; Tanaka, H.; Nishikawa, K.; Suzuki, H.; Saitoh, Y.; Tamura, A.; Tsukita, S.; Fujiyoshi, Y. *Nature Communications*, 2019, **10**.
- 45 Irudayanathan, F. J.; Wang, X.; Wang, N.; Willsey, S. R.; Seddon, I. A.; Nangia, S. *The Journal of Physical Chemistry B*, 2018, **122**.
- 46 Irudayanathan, F. J.; Wang, N.; Wang, X.; Nangia, S. *Annals of the New York Academy of Sciences*, 2017, **1405**.
- 47 Rajagopal, N.; Irudayanathan, J. F.; Nangia, S. *International Journal of Molecular Sciences*, 2019, **20**.
- 48 Irudayanathan, F. J.; Trasatti, J. P.; Karande, P.; Nangia, S. *The Journal of Physical Chemistry B*, 2016, **120**.
- 49 Huang, J.; MacKerell Jr, A. D. *Journal of Computational Chemistry*, 2013, **34**.
- 50 Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D. *Nature Methods*, 2017, **14**.
- 51 Brooks, B. R.; Brooks lii, C. L.; Mackerell Jr, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S. et al. *BMC Biol*, 2009, **30**.
- 52 Parrinello, M.; Rahman, A. *Journal of Applied Physics*, 1981, **52**.
- 53 Evans, D. J.; Holian, B. L. *The Journal of Chemical Physics*, 1985, **83**.
- 54 Periole, X.; Marrink, S. J.: The martini coarse-grained force field. In *Methods in Molecular Biology*, 2013; Vol. 924; pp 533-565.
- 55 Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S.-J. *Journal of Chemical Theory and Computation*, 2008, **4**.
- 56 Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; De Vries, A. H. *Journal of Physical Chemistry B*, 2007, **111**.
- 57 de Jong, D. H.; Singh, G.; Bennett, W. F. D.; Arnarez, C.; Wassenaar, T. A.; Schäfer, L. V.; Periole, X.; Tieleman, D. P.; Marrink, S. J. *Journal of Chemical Theory and Computation*, 2013, **9**.
- 58 Herzog, F. A.; Braun, L.; Schoen, I.; Vogel, V. *Journal of Chemical Theory and Computation*, 2016, **12**.

- 59 Rajagopal, N.; Nangia, S. *Journal of Chemical Theory and Computation*, 2019, **15**.
- 60 Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *The Journal of Chemical Physics*, 1984, **81**.
- 61 Bussi, G.; Donadio, D.; Parrinello, M. *The Journal of Chemical Physics*, 2007, **126**.
- 62 R. J. Gowers, M. L., J. Barnoud, T. J. E. Reddy, M. N. Melo, S. L. Seyler, D. L. Dotson, J. Domanski, S. Buchoux, I. M. Kenney, and O. Beckstein. Austin, TX2016; SciPy.
- 63 Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. *Journal of Computational Chemistry*, 2011, **32**.
- 64 Krieger, E.; Vriend, G. *Bioinformatics*, 2014, **30**.
- 65 Land, H.; Humble, M. S.; Bornscheuer, U. T., Höhne, M., Eds.; Springer New York: New York, NY, 2018.
- 66 2015.
- 67 Petřek, M.; Otyepka, M.; Banáš, P.; Košinová, P.; Koča, J.; Damborský, J. *BMC Bioinformatics*, 2006, **7**.
- 68 Yu, A. S. L.; Cheng, M. H.; Angelow, S.; Günzel, D.; Kanzawa, S. A.; Schneeberger, E. E.; Fromm, M.; Coalson, R. D. *The Journal of General Physiology*, 2008, **133**.
- 69 Yu, A. S. L. *Annals of the New York Academy of Sciences*, 2009, **1165**.
- 70 Samanta, P.; Wang, Y.; Fuladi, S.; Zou, J.; Li, Y.; Shen, L.; Weber, C.; Khalili-Araghi, F. *The Journal of General Physiology*, 2018, **150**.
- 71 Alberini, G.; Benfenati, F.; Maragliano, L. *PLOS ONE*, 2017, **12**.
- 72 Alberini, G.; Benfenati, F.; Maragliano, L. *The Journal of Physical Chemistry B*, 2018, **122**.
- 73 Borovac, J.; Barker, R. S.; Rievaj, J.; Rasmussen, A.; Pan, W.; Wevrick, R.; Alexander, R. T. *American Journal of Physiology-Cell Physiology*, 2012, **303**.
- 74 Van Itallie, C.; Rahner, C.; Anderson, J. M. *The Journal of Clinical Investigation*, 2001, **107**.



Systematic approach to predicting selectivity of paracellular pores for biomimetic applications