Dynamical model of the CLC-2 ion channel exhibits a two-step gating mechanism

Keri A. McKiernan[1], Anna K. Koster[1,2], Merritt Maduke[2,5], and Vijay S. Pande[1,3,4,5,*]

[1]Department of Chemistry, Stanford University, Stanford CA 94305, USA
[2]Department of Molecular & Cellular Physiology, Stanford University, Stanford CA 94305, USA
[3]Department of Computer Science, Stanford University, Stanford, CA 94305, USA
[4]Department of Structural Biology, Stanford University, Stanford, CA 94305, USA
[5]Program in Biophysics, Stanford University, Stanford, CA 94305, USA
[*]pande@stanford.edu

ABSTRACT

This work reports a Markov state model of the conformational dynamics involved in regulating the inner and outer gates of the CLC-2 ion channel. This dynamical model, based on 600 microseconds of molecular dynamics simulation, suggests a new mechanism for channel activation. Transition path analysis highlights a dominant pathway between a highly stable closed state (similar to known structures) and several new protein conformations not observed in any previously elucidated CLC structure. Movement from the closed state involves the intracellular rotation of the inner-gate backbone along residues S168-G169-I170. This rotated state is similar to that observed in the recent cryo-EM structure of the bovine CLC-K channel, but the volume of the intracellular (inner) region of the ion conduction pathway is expanded to a greater extent. From this state, which remains closed at the extracellular side, two additional states may be assumed, each involving a unique rotameric flip of the GLU_{ex} (E211) side chain. Both additional states involve conformational changes that orient GLU_{ex} away from the extracellular (outer) region of the ion conduction pathway. In the first additional state, the rotameric flip of GLU_{ex} results in an open, or near-open, channel pore. The equilibrium population of this state is low ($\sim$1%), consistent with the low open probability of CLC-2 in the absence of a membrane potential stimulus (0 mV). In the second additional state, GLU_{ex} rotates to occlude the channel pore. This state, which has a low equilibrium population ($\sim$1%), is only accessible when GLU_{ex} is protonated. Together, these pathways model the opening of both an inner and outer gate within the CLC-2 selectivity filter, as a function of GLU_{ex} protonation. Collectively, our findings build on and are consistent with known properties of CLC-2 and will guide future investigations into CLC-2 channel biophysics.

Introduction

The chloride channel (CLC) protein family plays a wide variety of physiological functions in a diverse array of organisms, ranging from bacteria to humans[1–7]. This family of membrane proteins is composed of both channels and H^+/Cl^- exchange transporters that share a structurally unique homodimeric architecture[8,9]. Each subunit within the homodimer is an independent functional unit[10–12] composed of 17 membrane-embedded alpha helices[13]. These helices coalesce to form a narrow, electropositive ion-conducting pore that is highly selective for Cl^-[13]. Such an architecture is unusual for ion channels and raises many questions about the mechanisms of ion channel opening and closing (gating).

High-resolution CLC structures provide invaluable starting points for understanding the structure-function relationships of CLC protein activation. The first CLC structure solved, a prokaryotic H^+/Cl^- exchange transporter[13], proved relevant for guiding structure-function studies on a wide range of CLCs, including eukaryotic CLC channels[14–17]. This structure represents a closed conformational state in which bound chloride (Cl^-) is occluded by proteinaceous "gates" that block exit to either the outer (extracellular) or inner (intracellular) side of the protein. The outer gate is formed in part by a conserved glutamate residue, GLU_{ex}[18–25]. When mutated to GLN (to resemble the protonated GLU), this side chain rotates outwards and partially unblocks the Cl^- permeation pathway[18]. The inner gate is formed by conserved SER and TYR residues[13,26–28] that, according to solved CLC structures, physically obstruct the Cl^- translocation pathway from the intracellular side. In the recently published structure of the bovine CLC-K channel (the first structure of a CLC channel homolog), the loop containing the SER gate residue is rotated to unblock the intracellular gate[29]. Despite the advances in structural understanding afforded by these discoveries, additional analysis is required to turn these static conformational snapshots into a complete dynamic picture of CLC gating. Experimental biophysicists have shown that these mechanisms are complex, involving interdependent influences of transmembrane voltage, $[Cl^-]$, and $[H^+]$[8,30–34]. To begin to resolve these complexities and reveal the dynamic, atomistic details of the CLC gating machine, it will be critical to pair structural studies with high-resolution conformational modeling[35–44].

In this study, we investigate the molecular dynamics of CLC-2, a chloride ion channel of particular interest due to its high levels of expression in the brain, in both neurons and glia[3,45]. In glia, the importance of CLC-2 was illuminated in studies of knockout mice, which develop white matter vacuolization resembling human leukoencephalopathy[46], and complementary studies[47,48]. Indeed, mutations in human CLC-2 are known to underlie a specific form of the disease[49], although the mechanistic details of causation are still uncertain[50]. In neurons, CLC-2 has been suggested to play an important role in regulation of excitability and/or chloride homeostasis[51–59]. However, the details of these roles are controversial[3,58,60,61]. A precise understanding of the atomic interactions and rearrangements involved in channel activation will be valuable to facilitate understanding of CLC-2 physiological behavior and its role in human disease.

The spatial and temporal resolution afforded by computer simulations allows for the atom-scale examination of protein structure and function. Here we apply molecular dynamics (MD) simulations and Markov State Modeling (MSM) to study the conformational dynamics of CLC-2 gating. MSMs allow for the rigorous statistical characterization of biophysical simulations[62–65]. Prior to constructing the MSM, the simulation data is first transformed into its kinetically slowest representation via the method of time-structure based independent component analysis (tICA)[66–68]. Such a transformation allows us to identify the slowest conformational changes in our dataset, and to build the MSM over these processes. We queried this MSM in order to predict the thermodynamics and kinetics of CLC-2 channel activation. Our model suggests that the top activation pathway begins in the closed state, near the template structure on which we based our model. The slowest timescale conformational change involves the rotation of the backbone along the S168-G169-I170 residues on the intracellular side of the entrance to the ion conduction pathway. The second major motion describes the flip of the GLU_{ex} side chain on the extracellular vestibule. This second motion displays a sensitivity to residue protonation that correlates well with experimental evidence that the conserved glutamate residue, GLU_{ex}, regulates the CLC "fast" gate in CLC-2[69,70]. Furthermore, the model suggests that these "inner" and "outer" gates are nontrivially coupled, providing a novel depiction of CLC conformational change and a new mechanistic hypothesis that describes the stepwise process of channel activation. Together, these data provide a high-resolution mechanistic model of CLC-2 gating.

Results and discussion

Ion channel "gating" involves the conformational rearrangement of the protein, such that transmembrane ion flow through the channel pore is either increased or decreased. These gates are highly regulated so as to avoid dissipation of cellular ion gradients. High-resolution structures of CLCs have revealed two "gates" within each subunit. These gates physically occlude the permeation pathway, one towards the extracellular (outer) side and one towards the intracellular (inner) side. Outer-gate opening is thought to describe the circumstance where the GLU_{ex} residue (E211 for rCLC-2, see Fig. 1 for reference) swings upward, such that it ceases to block the ion conduction pathway. Several lines of experiments support this model[18–25]. This conformational change may be influenced by a variety of factors, including but not limited to, external membrane potential, thermal fluctuation, local ion concentration, or residue protonation[69–77]. The conformational changes involved with inner gate opening are less precisely understood[24,78], though experiments are consistent with the narrowest constriction being localized near the SER/TYR site and not further towards the intracellular side[14,26].

We have applied dynamical modeling in order to identify and piece together the major conformational changes characterizing inner- and outer-gate opening. First, a homology model was built for the rCLC-2 sequence, using the 3ORG structure of the cmCLC transporter[79] as a template structure. This template is 75% identical to CLC-2 within the pore-forming (selectivity filter) regions of the protein. The sequence conservation between the target and template[80,81] is depicted in Fig. 1, bottom right. From this model, two MD datasets were generated, identical with the exception of GLU_{ex} protonation. To improve computational sampling, and because each subunit contains an independent and identical pore[10–12], only one subunit of the homodimer was simulated, as depicted in Fig. 1. These datasets were analyzed with respect to the selectivity filter conformation (Fig. 1 inset).

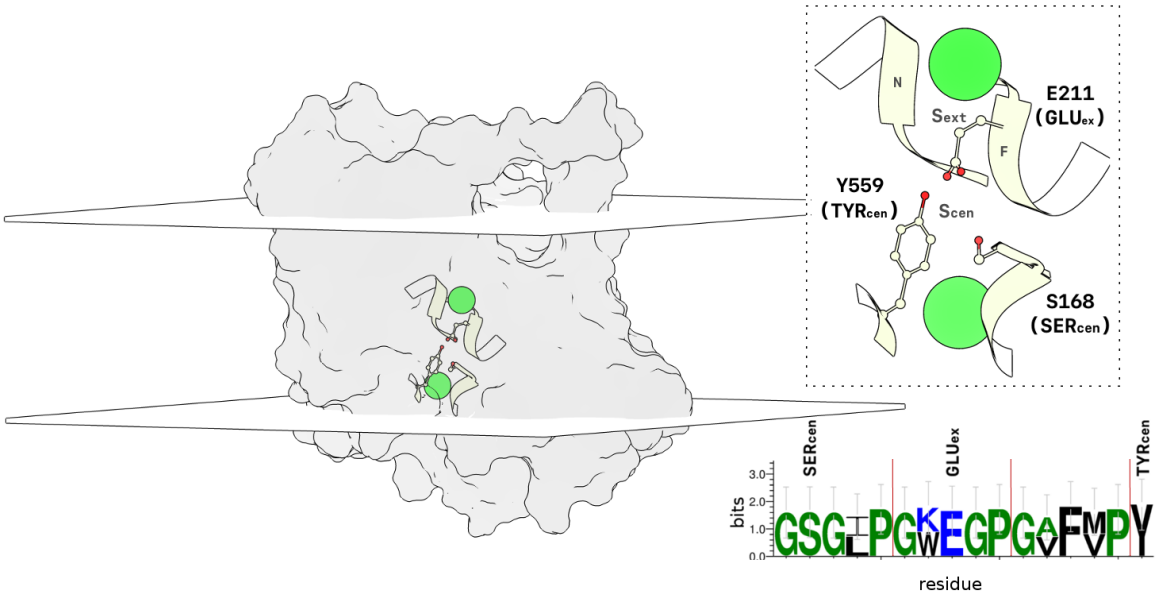

Figure 1. Illustration of the simulated rCLC-2 monomer. The protein is shown as a transparent surface. The planar surfaces represent the system membrane bilayer (planes intersect the average z-coordinate of the lipid phosphorus atoms). The selectivity filter (SF) is highlighted in a ribbon representation. A larger view of the SF is given by the inset with a dashed border. Mechanistically important Cl- binding sites, S_{cen} and S_{ext}, are labeled, as are nearby residues GLU_{ex}, SER_{cen}, and TYR_{cen}. Helices F and N help form the S_{ext} binding site. GLU_{ex}, on Helix F, is positioned at the S_{cen} ion binding site in this conformation. Chloride ions are shown in green. Sequence conservation between the CLC-2 and 3ORG selectivity filter residues is displayed on the bottom right. Hydrophilic residues are highlighted in blue, neutral residues in green, and hydrophobic residues in black. Residue height denotes the relative frequency of the amino acid at a given position. When there is a residue discrepancy, the CLC-2 residue is given on top. Each region of the SF has been segmented by the red lines.

Selectivity filter conformational dynamics

Markov State Models (MSMs) have emerged as a powerful tool for robustly estimating the statistics of biophysical systems from MD datasets[62–65]. This technique first partitions a dataset into a set of discrete conformational states, and then parameterizes a transition matrix describing the probability of transition for each state pair. Prior to constructing the MSM, it is advantageous to featurize the data according to a structural metric, and then transform the featurized dataset using time-structure independent component analysis (tICA)[66]. tICA determines the slowest-decorrelating dynamical degrees of freedom[82]. This improves the MSM statistics, and additionally provides a set of unbiased reaction coordinates that are useful for separating and visualizing important conformational states. These reaction coordinates can be thought of as dynamical modes, and are referred to as "tICs". A tIC is given by a linear combination of input features, and can be interpreted by identifying the features with the highest linear coefficients (referred to as loadings). After the MSM is constructed, the state representation can be simplified via macrostating. This simplification allows for the identification of the most thermodynamically stable conformational states. Here, we employ these techniques to characterize the conformational dynamics sampled in our datasets.

Our analysis identified four major conformational states: C_{oi}, C_o, O, and U. These states were named to reflect whether the channel appears closed (C) or open (O) at the inner (i) or outer (o) gate, with the exception of U, a postulated non-conducting state. A representative conformation for each thermodynamic macrostate is illustrated in Fig. 2. For a detailed labeling of each selectivity filter component, see Fig. 1. Each of these states can be kinetically distinguished from the others by tIC 0, tIC 1, and tIC 3. Note that tIC 2 was found to be redundant with tIC 1 and hence will not be discussed. These tICs were interpreted by analyzing the input features with the highest loadings. In Table 1 we have detailed how each tIC (dynamical mode) separates a particular two-state kinetic process, and which features distinguish each tIC. Greater detail describing the tIC loadings is given in the supplementary information.

Along each major dynamical mode, we observe two-state behavior. tIC 0 involves the rotation of the backbone dihedrals of the S168, G169, and I170 residues. This implies that the bottom strand of the SF can exist in two unique conformations. tICs 1 and 3 both involve flipping the GLU_{ex} χ_1 side chain dihedral, a motion that is coupled to the residues involved in the tIC 0 conformational change. Although tICs 1 and 3 begin in the same initial state (C_o) and describe the rotameric flipping of the same residue, the end states (O and U) differ in conformation. This suggests that the GLU_{ex} rotamer can exist in three possible states.

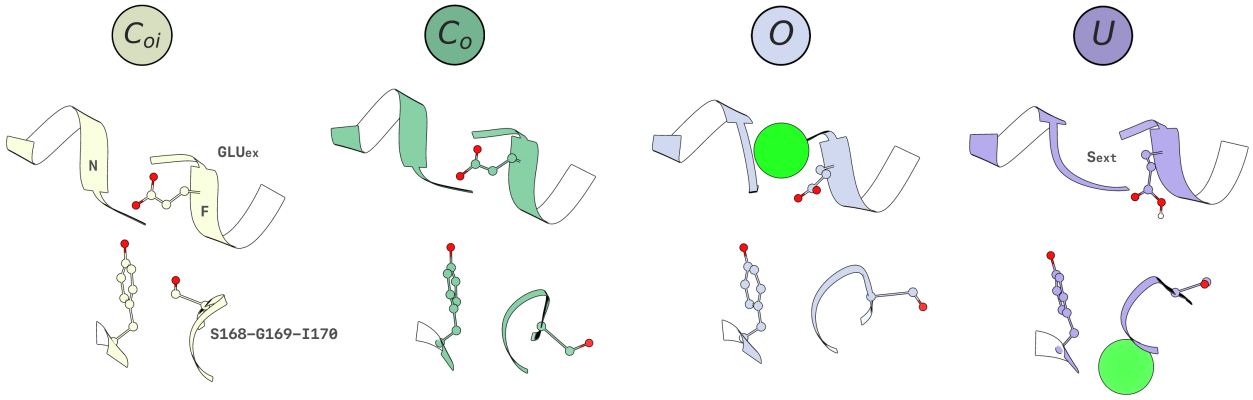

Figure 2. Structural visualization of each CLC-2 selectivity filter macrostate conformation. The S168-G169-I170 residues, located at the inner region of the SF, exist in two orientations, either pointing toward the ion conduction pathway (C_{oi}), or away (C_o, O, U). The GLU_{ex} residue side chain, located at the outer region of the SF, exists in three orientations, either occluding the S_{ext} Cl^- binding site, located in the gap between helices F and N (C_{oi}, C_o), or in one of the two distinct outward rotamers (O and U). Note that in the U state, the geometry of the S_{ext} binding site is distorted by the conformation of helix N.

dynamical mode	conformational change	top residue features	kinetic process
tIC 0	SER_{cen} backbone rotation	ψ(I170,G169), ψ(G169,S168)	$C_{oi} \leftrightarrow C_o$
tIC 1	GLU^p_{ex} side chain flip	χ_1(E211), ϕ(S168,G167)	$C_o \leftrightarrow U$
tIC 3	GLU_{ex} side chain flip	χ_1(E211), ϕ(G169,S168)	$C_o \leftrightarrow O$

Table 1. Dynamical modes of the MD data, as determined by tICA.

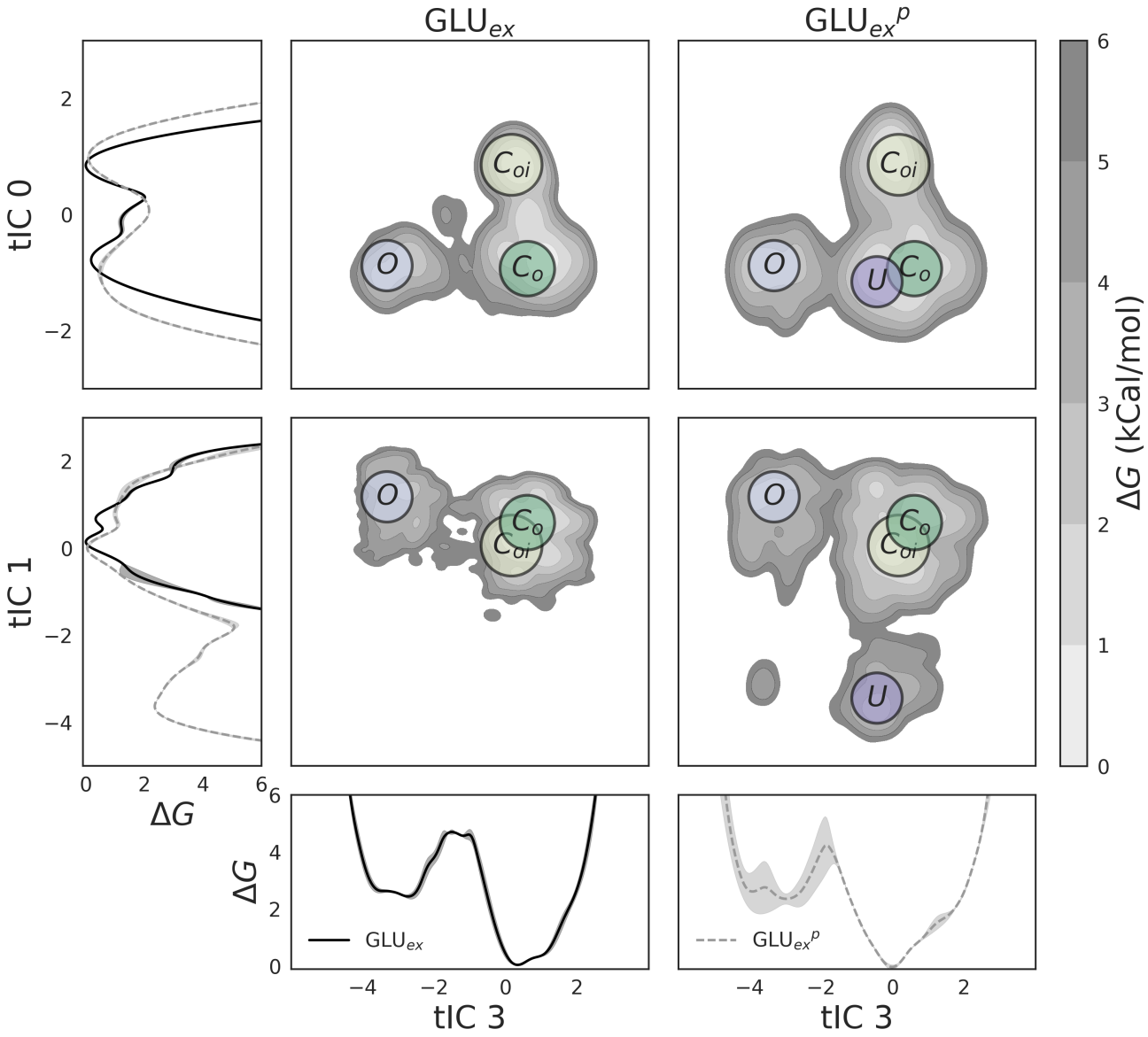

Figure 3. Free energy of major CLC-2 conformational states, projected onto all major dynamical modes (tICs), as a function of GLU$_{ex}$ protonation. Note that tIC 2 is redundant with tIC1 and is not shown. These landscapes are shown in both 1D and 2D, for both the deprotonated dataset (GLU$_{ex}$, black solid lines) and the protonated dataset (GLU$_{ex}{}^{p}$, gray dashed lines). tIC 0 denotes the conformation of the S168-G169-I170 backbone, while tIC 1 and tIC 3 denote the conformation of the GLU$_{ex}$ side chain. In all cases, C_{oi} was found to be the global free energy minimum, and to not vary greatly with respect to GLU$_{ex}$ protonation. Also insensitive to protonation are the basin free energies of C_o and O. The largest difference caused by GLU$_{ex}$ protonation is the existence of U, which occurs only for GLU$_{ex}{}^{p}$ (as isolated by tIC 1), and the structure of the kinetic barriers.

The thermodynamics of each conformational state was further analyzed using the MSM-derived free energies. This analysis is illustrated by Fig. 3, where the data have been separated by the GLU_{ex} protonation state, and projected onto each major dynamical mode. We observe that in all cases, C_{oi} describes the conformational free energy minimum. This state, so named because the channel pore appears closed to both the outer (o) and inner (i) sides of the membrane, is similar to the template (3ORG) structure with the exception that GLU_{ex} has moved upwards out of the S_{cen} site, towards the S_{ext} site. This positioning at S_{ext} is similar to that observed in all other CLC structures found in the PDB. The C_o state is the next most stable state, with a free energy less than 1 kCal/mol greater than the C_{oi} state. This state was labeled C_o because the channel pore is closed to the outer side but open towards the inner side. In this state, the S168, G169, and I170 backbone strand has changed conformation, but the GLU_{ex} residue remains in the same geometry. As can be seen by the landscapes in Fig. 3, GLU_{ex} protonation does not affect the free energies of the C_{oi} and C_o states. The transition barrier height for the $C_{oi} \leftrightarrow C_o$ conformational change is approximately 2 kCal/mol.

tICs 1 and 3 identify two higher energy conformational states, O and U, each connected to state C_o. As discussed previously, these states each represent rotamers of the GLU_{ex} residue. In the O ("open") state, GLU_{ex} appears to move out of the way of the Cl^- permeation pathway, whereas in the U state, GLU_{ex} collapses into the pore, which remains occluded (as discussed below). Both states O and U are between 2 and 3 kCal/mol greater in energy than states C_o and C_{oi}. The O state shows a mild sensitivity to protonation, while the U state exists only when $GLU_{ex}{}^p$ is protonated. This result can be seen by the absence of a U basin in the GLU_{ex} 2D free energy plots of Fig. 3. The transition barrier height of $C_o \rightarrow O$, along tIC 3, is approximately 4-5 kCal/mol for both protonation states, though there is considerable uncertainty for $GLU_{ex}{}^p$ (shown by the error bar along tIC 3). The transition barrier height of $C_o \rightarrow U$, along tIC 1, for $GLU_{ex}{}^p$ is approximately 5 kCal/mol.

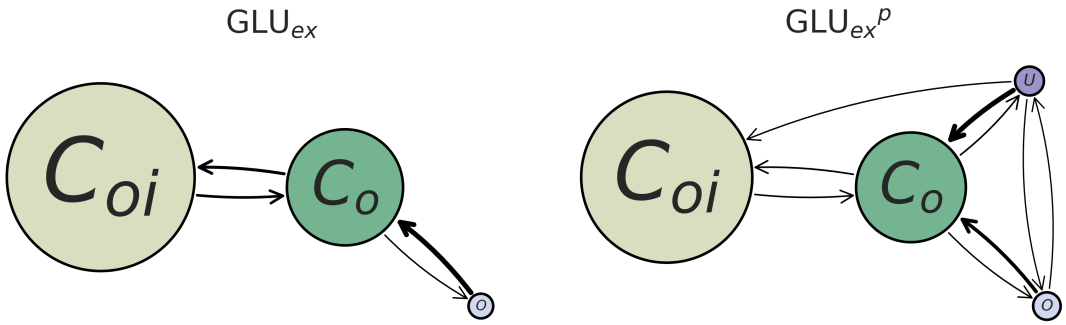

Figure 4. Kinetic rates describing the transition between major CLC-2 conformations, as a function of GLU_{ex} protonation. Note that the size of each node is proportional to its equilibrium population (larger denotes a lower free energy), and the weight of each arrow is proportional to the speed of the kinetic transition (heavier weight denotes a faster rate). Additionally, it is possible that there exist transitions occurring over timescales longer than what was sampled in our datasets (which likely underlies the lack of an observed transition from C_{oi} to U). As identified by the thermodynamic analysis, the U conformational state exists only for the $GLU_{ex}{}^p$ model. Both the GLU_{ex} and $GLU_{ex}{}^p$ models show similar rates describing $C_{oi} \leftrightarrow C_o$ and $C_o \leftrightarrow O$. However, in the case of $GLU_{ex}{}^p$, state U may transition forward to any other state.

To examine kinetics in greater detail, continuous-time rate-matrix MSMs were computed for the macrostate model in order to estimate the relative rate of transition between each major conformational state (as the traditional MSM gives transition probabilities, rather than transition rates). The resultant kinetic network graphs are given in Fig. 4. We observe that the C_{oi} state transitions only to the C_o state, while the C_o state may transition to any of the other states. The forward and backward rate of $C_{oi} \leftrightarrow C_o$ are comparable. For both the O and U states, the reverse rates from these states to the C_o state are faster than the forward rates. While the primary effect of GLU_{ex} protonation is the addition of the U state, kinetic analysis highlights two additional, mechanistically relevant, features of protonation. First, the connectivity of the modeled closed state, C_{oi}, to the modeled open state, O, suggests that protonation of GLU_{ex} is not essential for channel conduction. While this result is in contrast to a computational study of the 3ORG structure, which found that protonation of the GLU_{ex} side chain is required for ion transport to occur[38], it is in harmony with electrophysiological studies of CLC-2, which have demonstrated that protonation of GLU_{ex} does not affect channel opening[69]. The second additional highlight of our kinetic analysis is that the transition rate from O to C_{oi} is slowed when GLU_{ex} is protonated. While this result is uncertain due to the magnitude of the error bars of our calculations, it is worth highlighting given its consistency with experimental results, which demonstrate that the CLC-2 closing rate decreases with increasing extracellular $[\text{H}^+]$[70].

Structural analysis

Above, we computed estimates for the thermodynamic and kinetic behavior of four major CLC-2 conformational states. Now, we delve deeper into the structural analysis of these states. To further investigate the geometric structure of each macrostate, the volume of the ion conduction pathway was analyzed. For each macrostate, several conformations were sampled, and the radius of the ion conduction pathway computed. These results are given with respect to distance from the GLU_{ex} residue (one path to the extracellular region, one path to the intracellular region), and are illustrated in Fig. 5. It was found that GLU_{ex} is the bottleneck of the conduction pathway for all macrostates. States C_{oi}, C_o, and O display a comparable bottleneck of approximately 1.5 Å. In contrast, state U is more congested in this region, reducing to a radius of only 1 Å. It is unlikely that this especially constricted bottleneck would allow for chloride ion permeation.

While states C_{oi}, C_o, and O display bottlenecks of approximately 1.5 Å in radius (narrower than the $\sim$1.8 Å radius of a Cl^- ion), the extent of these bottlenecks along the length of the channel pore varies significantly between the conformations. For C_{oi} and C_o, the pore is constricted to less than 1.8 Å for a $\sim$19 Å stretch, whereas for state O, the pore is constricted for less than half that distance (Fig. 5). Comparing C_{oi} and C_o, the most striking difference is along the intracellular pathway, which widens sufficiently to accommodate chloride in C_o. This finding aligns with our initial observation that these states are distinguished by the conformational state of the inner gate, while the outer gate remains closed. Across both the extracellular and intracellular region, the O state is the most voluminous. This finding aligns with our observation that both inner- and outer-gate residues (GLU_{ex} and SER_{cen}) have rotated away from S_{cen}. The fact that a narrow constriction persists within the putative "open" state (the most open of the set of macrostates) suggests that further conformational change may be involved in reaching a fully open state. Such an additional conformational change (to widen the pore following the rotation of GLU_{ex}) has been proposed based on experimental studies[37,83,84].

For both extracellular and intracellular pathways, the U state is the most constricted, despite the fact that the inner-gate SER residue is rotated to the "open" position. The GLU_{ex} residue is in a unique rotameric state, rotated away from SER_{cen}, but this rotation does not increase the volume of the extracellular region of the channel, in part due to the distortion of helix N in this state (see Fig. 2). The constricted nature of the U state indicates that it is a non-conductive (closed) channel. That the U state is only accessible when GLU_{ex} is protonated (neutral) is in agreement with experimental observation that low extracellular pH induces channel closure[19,85,86] (though it is known that an extracellular histidine is also involved[87]).

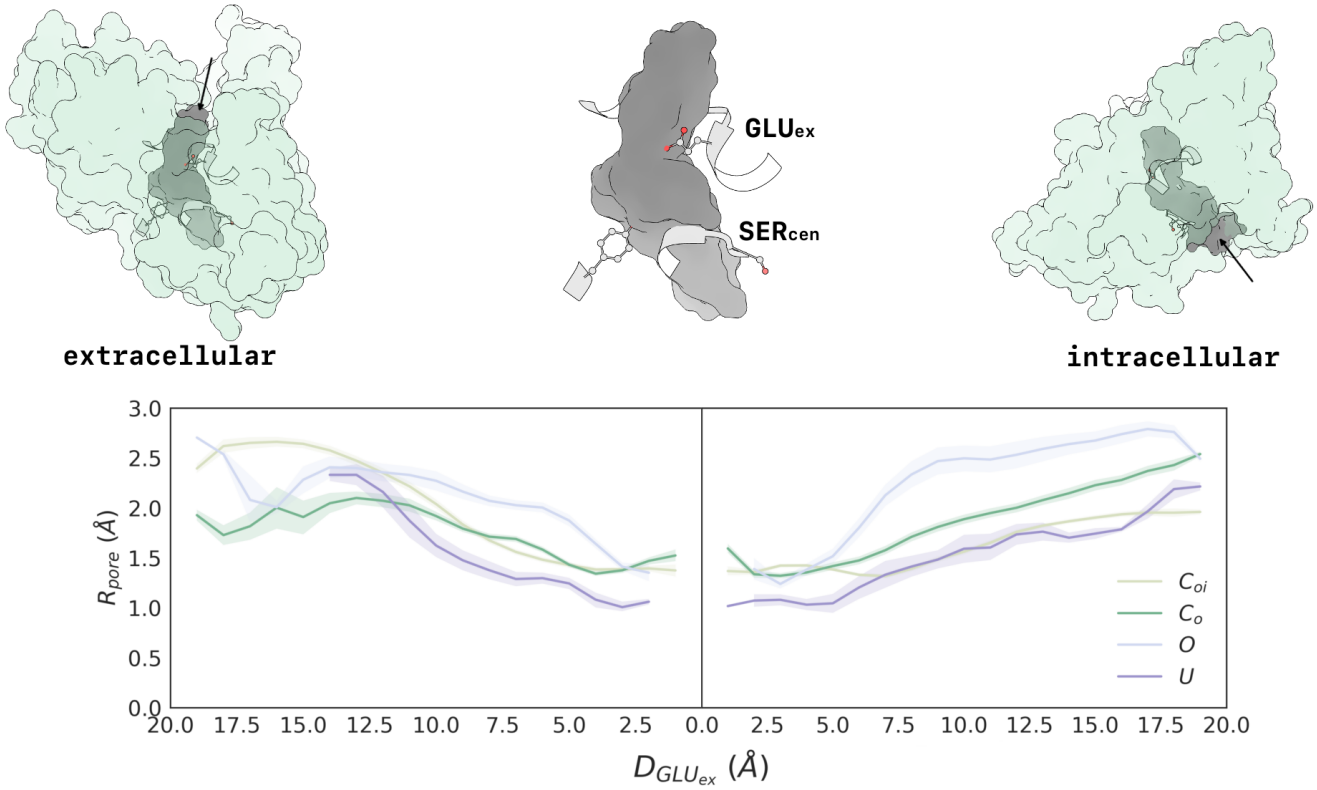

Figure 5. The radius of the ion conduction pathway demonstrates a strong dependence on macrostate. Distance along the pathway (x axis) is given with respect to the GLU_{ex} residue. The plot on the left illustrates the pore radius from the extracellular side of the protein, descending into the pore to GLU_{ex}, while the plot on the right illustrates the pore radius from GLU_{ex} to the intracellular side of the protein. As an example, the top structures illustrate the volume of the C_o ion conduction pathway from multiple perspectives. All macrostates are relatively open near the extracellular mouth of the protein, but narrow in distinct ways to the bottleneck located at GLU_{ex}. The O state is the most open in the region just above GLU_{ex}, and U is the most narrow. Additionally, the minimum radius of the bottleneck is approximately the same for all states except U, which is highly constricted. Moving toward the intracellular opening from GLU_{ex}, O is once again the most voluminous, while U is the most constricted.

Now, we further interpret our model (Figs. 3, 4) structurally by visualizing representative conformations from each macrostate. Superpositions of these macrostate conformations are depicted in Fig. 6A. Observe that the inner residues, S168-G169-I170, assume two major conformational states, while the outer residues, GLU_{ex} and those surrounding the S_{ext} binding site, assume three major conformational states.

In Fig. 6B, we compare these structures to a set of relevant structures from the PDB. In the known structures, the GLU_{ex} side chain has been observed to adopt three conformations, referred to as "down", "middle", and "up"[26]. The "down" conformation, with GLU_{ex} pointing into S_{cen}, has been seen only in 3ORG, our template structure. In the CLC-2 model based on this structure, we found that GLU_{ex} quickly equilibrated away from this conformation to a rotamer where it occupied the S_{ext} binding site (the "middle" position). This positioning is similar to that observed in most CLCs in the PDB (overlay shown for representative structure 1OTS). The GLU_{ex} "up" position, which has been observed when GLU_{ex} is mutated to GLN (PDB 1OTU[18]) has been considered to represent a more open pore conformation; however this conformation is more occluded than the *O* macrostate[37]. The *O* state demonstrates a SF-conformation similar to that of the 5TQQ structure, especially with respect to the intracellular region.

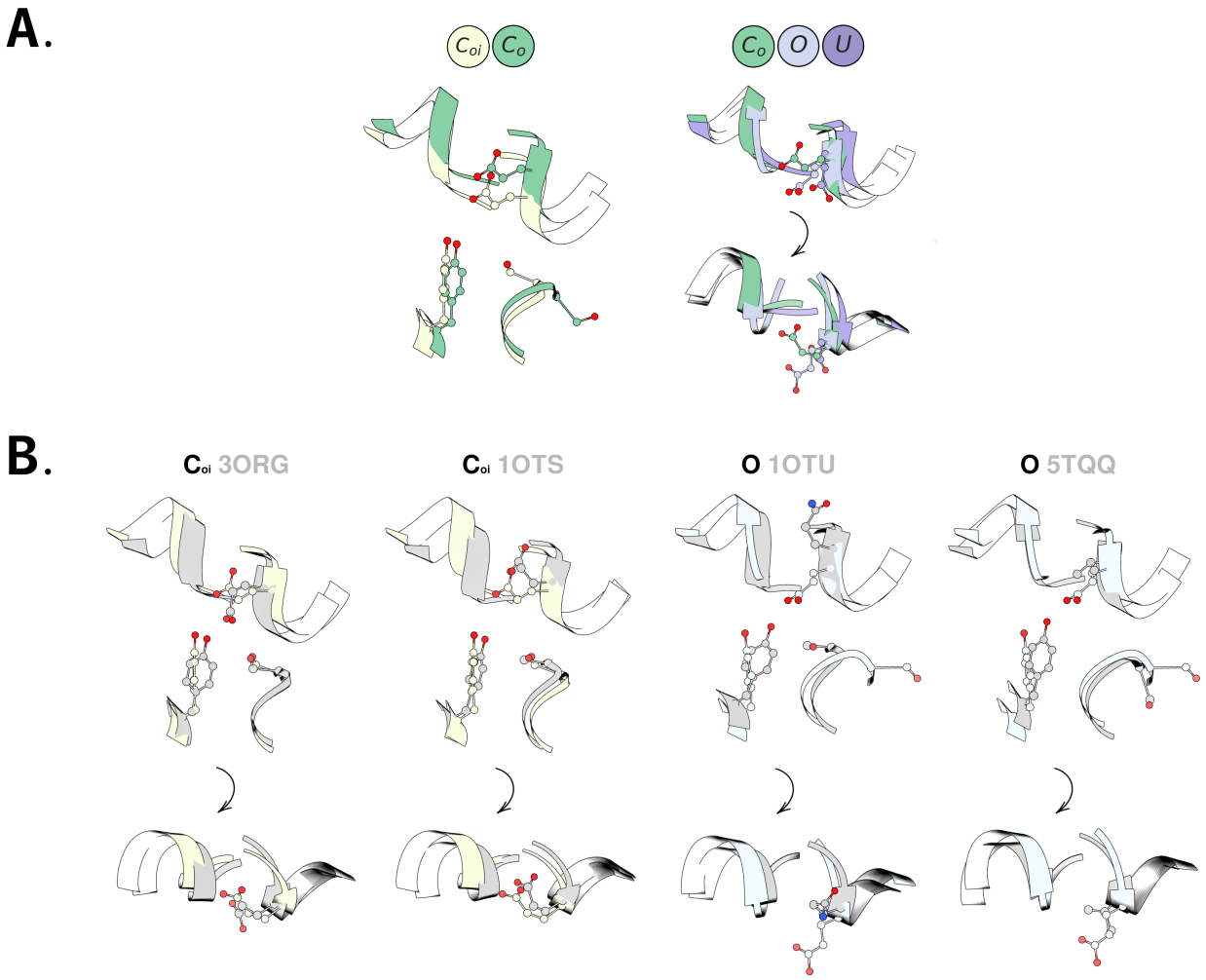

Figure 6. Structural overlays of selectivity filter (SF) conformations. **A.** Overlay of the four macrostates observed in this study. Left: The S168-G169-I170 backbone conformation distinguishes the C_{oi} and C_o states. Right: The GLU_{ex} side chain conformation distinguishes the C_o, O, and U states. These rotamers are depicted from the plane of the membrane (upper), and also from the extracellular plane (lower). The rotation of SF orientation is signified by the black arrow. **B.** Comparison to relevant structures in the PDB. As in A (right panel), these comparisons are shown from the plane of the membrane (upper), and also from the extracellular plane (lower). The stable C_{oi} state adopts a GLU_{ex} conformation different from that in the template structure 3ORG and more similar to that observed in most CLC structures in the PDB, as represented by 1OTS. The O state adopts a GLU_{ex} conformation not seen in any known structure, including 1OTU, where GLU_{ex} is mutated to GLN. At the intracellular side, the conformation of the S168-G169-I170 backbone of the O state resembles that of the 5TQQ structure.

Next, we analyze these conformations along each major dynamical mode (tIC). This is done by examining the structures along each tIC projection, as shown in Fig. 7. Here, the free energy surfaces for each two-state process are projected onto the major reaction coordinates (tICs, or dynamical modes) described earlier, and separated by GLU_{ex} protonation state. The shading surrounding the free energy curves represents the error (1 standard deviation) calculated from statistical bootstrapping of the MSM. The gray structure along each tIC represents the process transition state.

By examining the most significant tIC 0 feature loadings, we determine that the backbone of the S168, G169, and I170 residues determine the conformational states of the $C_o \leftrightarrow C_{oi}$ process. This is illustrated in Fig. 7, top. The C_{oi} state (yellow) positions the S168 side chain toward the ion conduction pathway. The C_o state positions the S168 side chain to the cytosol. In the tIC 0 transition state, the S168 side chain is oriented halfway between the C_{oi} and C_o macrostate structures. Hence, we observe a full 180 degree intracellular rotation of the S168-G169-I170 backbone. Recently, a cryo-EM structure of the bovine CLC-K channel (5TQQ) was found to uniquely (compared to previous CLC structures) display a cytosolic facing SER_{cen} residue[29], similar to the conformation found here. The GLU_{ex} residue adopts the same rotameric state in C_{oi} and C_o, positioned directly in the way of the ion conduction pathway, as is found commonly in known CLC structures. This observation suggests that the outer gate is "closed" for the C_{oi} and C_o states.

Together, tIC 1 and tIC 3 separate two unique rotameric states of the GLU_{ex} residue. Note that along these tICs, the S167-G169-I170 backbone is "out", as can be seen by the outward orientation of the S168 side chain. As revealed by our model, accessibility to the U rotameric state is possible only when the GLU_{ex} residue is protonated. As visualized in Fig. 7 middle and bottom, states O and U describe rotamers of GLU_{ex} where the side chain is no longer resting in the exterior ion binding site (this ion binding site is occupied by an ion for the O state). The structures of the O and U states therefore represent two unique outer gate conformations.

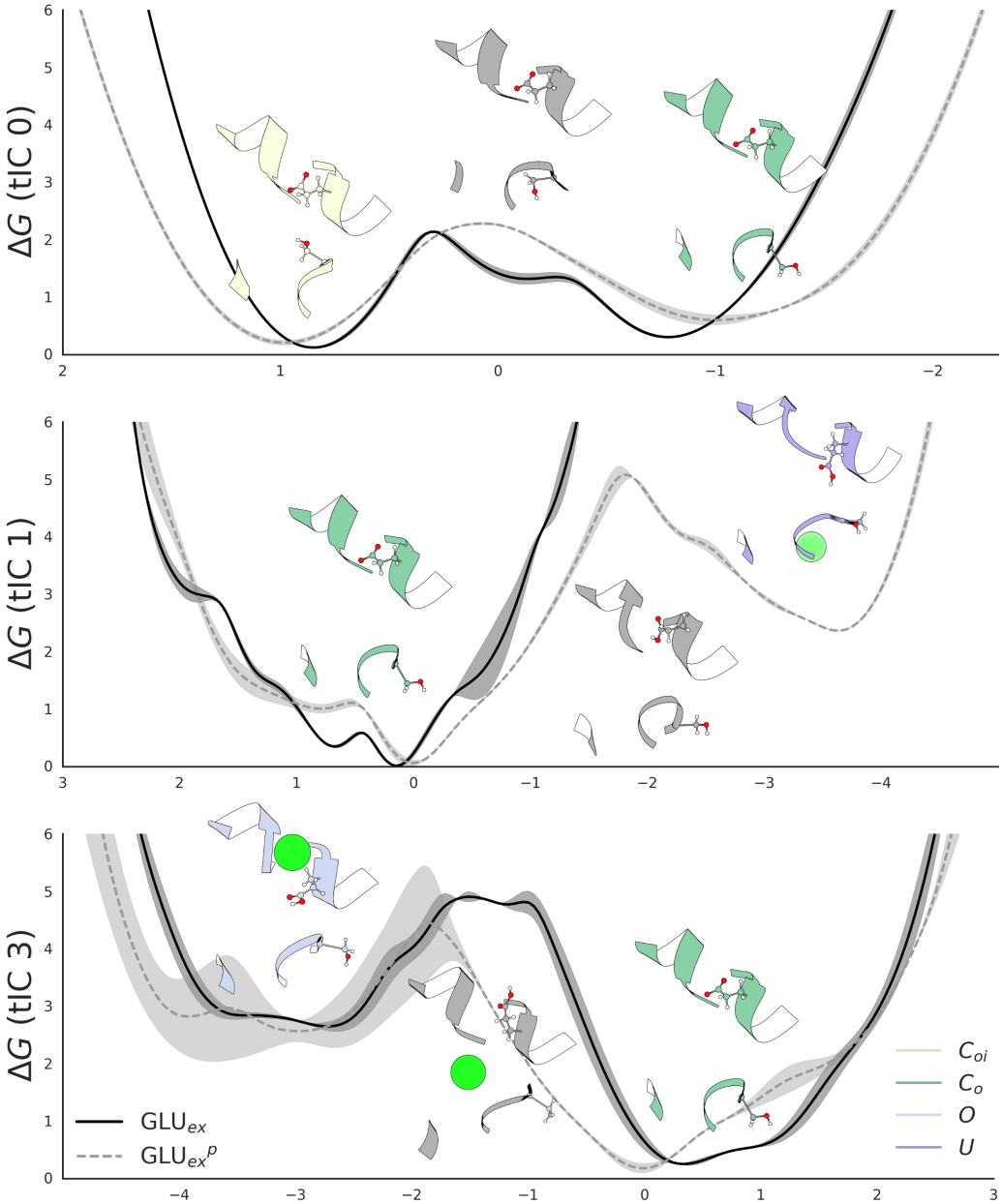

Figure 7. Structural analysis of the major free energy basins along the top reaction coordinates illustrates the conformational changes involved in CLC-2 activation. For each coordinate, the transition state conformation is shown in gray. The conformational changes involve rotation of the S168, G169, and I170 residue backbone (tIC 0), as well as the rotameric flip of the GLU$_{ex}$ residue (tIC 1, tIC 3).

The structures identified for the major conformational states were arranged into a schematic network mechanism, given in Fig. 8. This network diagram summarizes the results presented in Fig. 7. We observe that there is a dominant motion from the C_{oi} to C_o state, which involves the rotation of the S168-G169-I170 backbone. Following this motion, the GLU_{ex} side chain may rotate away from the ion conduction pathway. This GLU_{ex} flip has a low equilibrium population under these conditions (zero millivolts), such that return to the resting GLU_{ex} conformation is favored. The O rotamer possibly activates the channel to the full conducting state, while the protonation-dependent U rotamer constricts the conduction pathway to a non-conducting state. This hypothesized non-conducting state is consistent with experiments demonstrating the state-dependent effect of pH on the CLC-2 ion channel. Under a constant hyperpolarizing stimulus of -100 mV (to open the channel), application of pH 5.5 external buffer (to favor GLU_{ex} protonation) followed by a depolarizing pulse to close the channel (while GLU_{ex} remains protonated) renders the channel unable to re-open upon application of a second -100 mV stimulus[86]. In contrast, the same experiment performed under basic conditions (pH 8.0, favoring GLU_{ex} deprotonation) reduces current overall but does not prevent the channel from re-opening, suggesting that GLU_{ex} protonation locks the channel in a unique non-conducting conformation that prevents immediate reactivation[86]. This observation is consistent with our model in which the non-conducting U conformational state of CLC-2 is only accessible when GLU_{ex} is protonated.

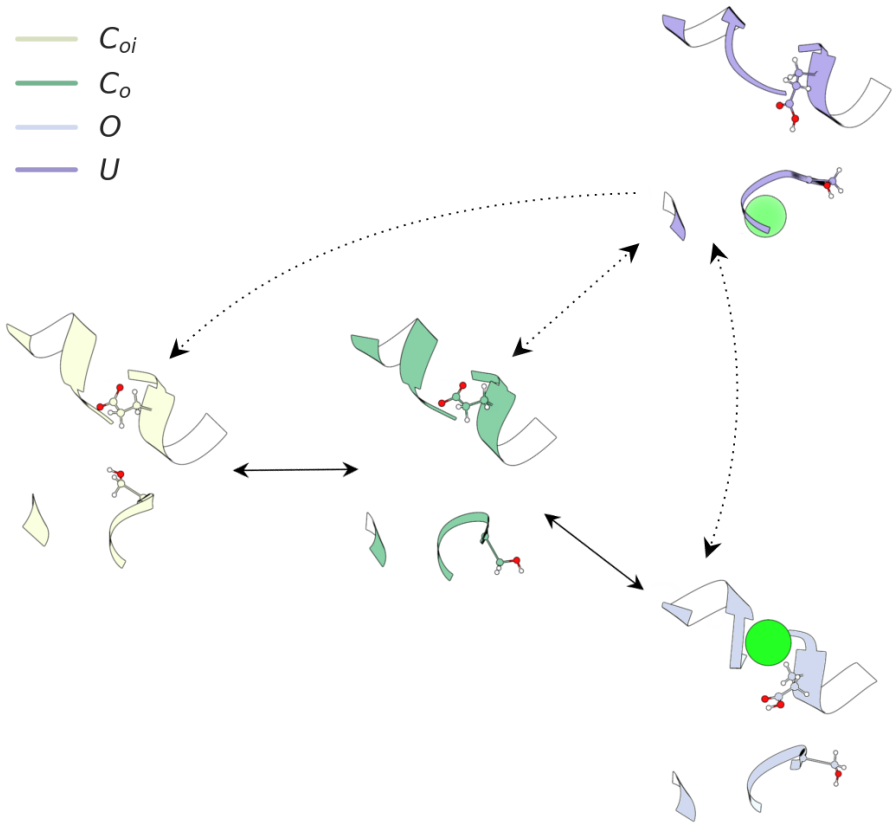

Figure 8. Model of CLC-2 activation. Dashed rates signify transitions only accessible when GLU_{ex} is protonated. The S168-G169-I170 backbone may exist in two possible conformational states. One points toward the ion conduction pathway, the other has the S168 side chain intracellularly solvent-exposed. The GLU_{ex} residue side chain may exist in three possible conformational states. In the closed state, GLU_{ex} rests directly in the S_{ext} ion binding site, obstructing the ion conduction pathway. When GLU_{ex} is protonated, or an ion permeates close to the residue, it flips outward. In the O (putative open) state, GLU_{ex} moves to the side, opening up space for a chloride ion. When GLU_{ex} is protonated, the selectivity filter can adopt a third conformation (the U state), which constricts the permeation pathway, to a non-conducting state.

Summary and conclusions

We have applied computational modeling methodologies to predict the conformational dynamics of CLC-2 activation. Four major conformational states were identified and characterized with respect to thermodynamics, kinetics, and structure. Of the major states, two were found to be highly stable, and two to be less stable but still detectably populated the simulation conditions (zero millivolts). Progression through these states was found to follow a particular sequence. First, from the C_{oi}, or closed state, there is flip of the S168-G169-I170 backbone. This flip occurs via an intracellular rotation and results in a volume increase in the inner regions of the ion conduction pathway. Once this occurs, the GLU_{ex} side chain is able to flip. Our model suggests that intracellular gate opening (mediated by the intracellular rotation of SER_{cen}) must occur prior to opening of the extracellular gate (rotation of GLU_{ex}). Such a mechanism is consistent with the picture emerging from mechanistic electrophysiological studies, which indicate that intracellular Cl^- is the main factor driving CLC-2 channel opening[69]. The intracellular rotation of the S168-G169-I170 backbone that we observe would allow intracellular chloride to access the inner channel pore and facilitate opening of the outer glutamate gate by forcing GLU_{ex} into its open conformation as chloride traverses the membrane.

Our studies of GLU_{ex} protonation reveal consistencies with the existing literature, and suggest potential areas of new investigation. Our observation that protonation of GLU_{ex} is not required for opening is consistent with functional studies of CLC-2[69], as is our observation that protonation of GLU_{ex} decreases the rate of transition from open to closed conformations[70]. Our observation that protonation of GLU_{ex} introduces the ability to adopt an alternative rotamer orientation suggests a potential link between CLC-2 inhibition by extracellular protons (which also involves an extracellular histidine residue[87]). Given the complex relationship between H^+, Cl^-, and voltage in regulating CLC-2 gating, many questions remain. This CLC-2 activation model will complement current and future efforts toward furthering the biophysical understanding of CLC proteins.

Methods

Molecular dynamics

The initial structures consisted of a CLC-2 homology model built using the software Modeller[88]. The rat CLC-2 monomer sequence was mapped onto a monomer of the 3ORG structure. The membrane system was assembled via CHARMMGUI[89]. The TIP3P[90], lipid14[91] and AMBER14SB[92] forcefields were used. Equilibration was done using AMBER 14*, with restraints on the protein and lipids, which were slowly released as the system was heated gradually from 0 to 300 K. An additional 100 ns simulation was run, from which initial system configurations were sampled. The AMBER prmtop and inpcrd files of these structures were used to set up OpenMM[93] simulations in the NPT ensemble. From these initial structures, 100 unique velocity distributions were assigned. The distributed computing platform folding@home[94] was then used to perform 600 μs of aggregate MD simulation. The UCSF Chimera program[95] was used to visualize the MD trajectories and generate figures.

Markov state modeling

MSM analysis was accomplished using MSMBuilder 3.7[96]. The MD datasets were featurized using the signed ϕ, ψ, χ_1, and χ_2 dihedral angles for all residues in the CLC-2 selectivity filter. The SF residue numbers are as follows (using the rCLC-2 primary sequence numbering): 167, 168, 169, 170, 171, 209, 210, 211, 212, 213, 463, 464, 465, 466, 467, 468, 559. The featurized datasets were transformed using kinetic-mapping of all tICA tICs. The transformed data was clustered into 324 microstates via the mini-batch kmeans method. The same tICA model and state decomposition was used for both datasets. 100 bootstrap MSMs were constructed (with a lag-time of 28.8 ns), from which the MLE model has been highlighted. Error bars for the free energy estimates were calculated using all MSM models. Macrostates were determined using the PCCA+ algorithm. A continuous time rate-matrix MSM was used to estimate the kinetic rates of macrostate conversion. The Caver 3.0 software package[97] was used to calculate the pore radii for each macrostate.

References

1. Jentsch, T. J. Discovery of CLC transport proteins: cloning, structure, function and pathophysiology. *The J. physiology* **593**, 4091–4109 (2015).
2. Stauber, T., Weinert, S. & Jentsch, T. J. Cell biology and physiology of CLC chloride channels and transporters. *Compr. Physiol.* (2012).
3. Stölting, G., Fischer, M. & Fahlke, C. Clc channel function and dysfunction in health and disease. *Front. physiology* **5** (2014).
4. Poroca, D. R., Pelis, R. M. & Chappe, V. M. Clc channels and transporters: Structure, physiological functions, and implications in human chloride channelopathies. *Front. pharmacology* **8** (2017).
5. De Angeli, A. *et al.* Clc-mediated anion transport in plant cells. *Philos. Transactions Royal Soc. Lond. B: Biol. Sci.* **364**, 195–201 (2009).

6. Iyer, R., Iverson, T. M., Accardi, A. & Miller, C. A biological role for prokaryotic CLC chloride channels. *Nat.* **419**, 715–718 (2002).
7. Fong, P. Twenty-five years of CLC chloride transport proteins. *The J. physiology* **593**, 4083–4084 (2015).
8. Miller, C. CLC chloride channels viewed through a transporter lens. *Nat.* **440**, 484–489 (2006).
9. Accardi, A. Structure and gating of CLC channels and exchangers. *The J. physiology* **593**, 4129–4138 (2015).
10. Middleton, R. E., Pheasant, D. J. & Miller, C. Homodimeric architecture of a cic-type chloride ion channel. *Nat.* **383**, 337 (1996).
11. Ludewig, U., Pusch, M. & Jentsch, T. J. Two physically distinct pores in the dimeric clc-0 chloride channel. *Nat.* **383**, 340–343 (1996).
12. Maduke, M., Miller, C. & Mindell, J. A. A decade of CLC chloride channels: structure, mechanism, and many unsettled questions. *Annu. review biophysics biomolecular structure* **29**, 411–438 (2000).
13. Dutzler, R., Campbell, E. B., Cadene, M., Chait, B. T. & MacKinnon, R. X-ray structure of a CLC chloride channel at 3.0 å reveals the molecular basis of anion selectivity. *Nat.* **415**, 287–294 (2002).
14. Lin, C.-W. & Chen, T.-Y. Probing the pore of clc-0 by substituted cysteine accessibility method using methane thiosulfonate reagents. *The J. general physiology* **122**, 147–159 (2003).
15. Engh, A. M. & Maduke, M. Cysteine accessibility in clc-0 supports conservation of the CLC intracellular vestibule. *The J. general physiology* **125**, 601–617 (2005).
16. Matulef, K. & Maduke, M. The CLC 'chloride channel' family: revelations from prokaryotes. *Mol. membrane biology* **24**, 342–350 (2007).
17. Estévez, R., Schroeder, B. C., Accardi, A., Jentsch, T. J. & Pusch, M. Conservation of chloride channel structure revealed by an inhibitor binding site in clc-1. *Neuron* **38**, 47–59 (2003).
18. Dutzler, R., Campbell, E. B. & MacKinnon, R. Gating the selectivity filter in CLC chloride channels. *Sci.* **300**, 108–112 (2003).
19. Niemeyer, M. I., Cid, L. P., Zúñiga, L., Catalán, M. & Sepúlveda, F. V. A conserved pore-lining glutamate as a voltage-and chloride-dependent gate in the clc-2 chloride channel. *The J. physiology* **553**, 873–879 (2003).
20. Vien, M., Basilio, D., Leisle, L. & Accardi, A. Probing the conformation of a conserved glutamic acid within the cl- pathway of a CLC h+/cl- exchanger. *The J. Gen. Physiol.* jgp–201611682 (2017).
21. Phillips, S. *et al.* Surprises from an unusual CLC homolog. *Biophys. journal* **103**, L44–L46 (2012).
22. Picollo, A., Malvezzi, M. & Accardi, A. Proton block of the clc-5 cl-/h+ exchanger. *The J. general physiology* **135**, 653–659 (2010).
23. Costa, A. *et al.* The arabidopsis central vacuole as an expression system for intracellular transporters: functional characterization of the cl-/h+ exchanger clc-7. *The J. physiology* **590**, 3421–3430 (2012).
24. Bell, S. P., Curran, P. K., Choi, S. & Mindell, J. A. Site-directed fluorescence studies of a prokaryotic clc antiporter. *Biochem.* **45**, 6773–6782 (2006).
25. de Santiago, J. A., Nehrke, K. & Arreola, J. Quantitative analysis of the voltage-dependent gating of mouse parotid clc-2 chloride channel. *The J. general physiology* **126**, 591–603 (2005).
26. Basilio, D., Noack, K., Picollo, A. & Accardi, A. Conformational changes required for h^+/cl^- exchange mediated by a CLC transporter. *Nat. structural & molecular biology* **21**, 456–463 (2014).
27. Jayaram, H., Accardi, A., Wu, F., Williams, C. & Miller, C. Ion permeation through a cl^- selective channel designed from a CLC cl^-/h^+ exchanger. *Proc. Natl. Acad. Sci.* **105**, 11194–11199 (2008).
28. Elvington, S. M., Liu, C. W. & Maduke, M. C. Substrate-driven conformational changes in clc-ec1 observed by fluorine nmr. *The EMBO journal* **28**, 3090–3102 (2009).
29. Park, E., Campbell, E. B. & MacKinnon, R. Structure of a CLC chloride ion channel by cryo-electron microscopy. *Nat.* **541**, 500–505 (2017).
30. Pusch, M., Ludewig, U., Rehfeldt, A. & Jentsch, T. J. Gating of the voltage-dependent chloride channel clc-0 by the permeant anion. *Nat.* **373**, 527–531 (1995).
31. Chen, T.-Y. & Miller, C. Nonequilibrium gating and voltage dependence of the clc-0 cl^- channel. *The J. Gen. Physiol.* **108**, 237–250 (1996).

32. Lísal, J. & Maduke, M. Proton-coupled gating in chloride channels. *Philos. Transactions Royal Soc. Lond. B: Biol. Sci.* **364**, 181–187 (2009).

33. Accardi, A. & Picollo, A. Clc channels and transporters: proteins with borderline personalities. *Biochimica et Biophys. Acta (BBA)-Biomembranes* **1798**, 1457–1464 (2010).

34. Pusch, M. Structural insights into chloride and proton-mediated gating of CLC chloride channels. *Biochem.* **43**, 1135–1144 (2004).

35. Lee, S., Mayes, H. B., Swanson, J. M. & Voth, G. A. The origin of coupled chloride and proton transport in a cl^-/h^+ antiporter. *J. Am. Chem. Soc* **138**, 14923–14930 (2016).

36. Jiang, T., Han, W., Maduke, M. & Tajkhorshid, E. Molecular basis for differential anion binding and proton coupling in the cl^-/h^+ exchanger clc-ec1. *J. Am. Chem. Soc.* **138**, 3066–3075 (2016).

37. Khantwal, C. M. *et al.* Revealing an outward-facing open conformational state in a CLC cl^-/h^+ exchange transporter. *eLife* **5**, e11189 (2016).

38. Cheng, M. H. & Coalson, R. D. Molecular dynamics investigation of cl^- and water transport through a eukaryotic CLC transporter. *Biophys. journal* **102**, 1363–1371 (2012).

39. Ko, Y. J. & Jo, W. H. Chloride ion conduction without water coordination in the pore of clc protein. *J. computational chemistry* **31**, 603–611 (2010).

40. Bisset, D., Corry, B. & Chung, S.-H. The fast gating mechanism in clc-0 channels. *Biophys. journal* **89**, 179–186 (2005).

41. Bostick, D. L. & Berkowitz, M. L. Exterior site occupancy infers chloride-induced proton gating in a prokaryotic homolog of the clc chloride channel. *Biophys. journal* **87**, 1686–1696 (2004).

42. Cohen, J. & Schulten, K. Mechanism of anionic conduction across clc. *Biophys. journal* **86**, 836–845 (2004).

43. Picollo, A., Xu, Y., Johner, N., Bernèche, S. & Accardi, A. Synergistic substrate binding determines the stoichiometry of transport of a prokaryotic h+/cl- exchanger. *Nat. structural & molecular biology* **19**, 525–531 (2012).

44. Lee, S., Swanson, J. M. & Voth, G. A. Multiscale simulations reveal key aspects of the proton transport mechanism in the clc-ec1 antiporter. *Biophys. journal* **110**, 1334–1345 (2016).

45. Thiemann, A., Grunder, S. *et al.* A chloride channel widely expressed in epithelial and non-epithelial cells. *Nat.* **356**, 57 (1992).

46. Blanz, J. *et al.* Leukoencephalopathy upon disruption of the chloride channel clc-2. *J. Neurosci.* **27**, 6581–6589 (2007).

47. Jeworutzki, E. *et al.* Glialcam, a protein defective in a leukodystrophy, serves as a clc-2 cl- channel auxiliary subunit. *Neuron* **73**, 951–961 (2012).

48. Hoegg-Beiler, M. B. *et al.* Disrupting mlc1 and glialcam and clc-2 interactions in leukodystrophy entails glial chloride channel dysfunction. *Nat. communications* **5**, 3475 (2014).

49. Depienne, C. *et al.* Brain white matter oedema due to CLC-2 chloride channel deficiency: an observational analytical study. *The Lancet Neurol.* **12**, 659–668 (2013).

50. Jentsch, T. J. From mice to man: chloride transport in leukoencephalopathy. *The Lancet Neurol.* **12**, 626–628 (2013).

51. Ge, Y.-X., Liu, Y., Tang, H.-Y., Liu, X.-G. & Wang, X. Clc-2 contributes to tonic inhibition mediated by $\alpha 5$ subunit-containing gaba a receptor in experimental temporal lobe epilepsy. *Neurosci.* **186**, 120–127 (2011).

52. Madison, D. V., Malenka, R. C. & Nicoll, R. A. Phorbol esters block a voltage-sensitive chloride current in hippocampal pyramidal cells. *Nat.* **321**, 695–697 (1986).

53. Földy, C., Lee, S.-H., Morgan, R. J. & Soltesz, I. Regulation of fast-spiking basket cell synapses by the chloride channel clc-2. *Nat. neuroscience* **13**, 1047–1049 (2010).

54. Clark, S., Jordt, S.-E., Jentsch, T. J. & Mathie, A. Characterization of the hyperpolarization-activated chloride current in dissociated rat sympathetic neurons. *The J. Physiol.* **506**, 665–678 (1998).

55. Staley, K., Smith, R., Schaack, J., Wilcox, C. & Jentsch, T. J. Alteration of gaba a receptor function following gene transfer of the clc-2 chloride channel. *Neuron* **17**, 543–551 (1996).

56. Rinke, I., Artmann, J. & Stein, V. Clc-2 voltage-gated channels constitute part of the background conductance and assist chloride extrusion. *J. Neurosci.* **30**, 4776–4786 (2010).

57. Kleefuss-Lie, A. *et al.* Clcn2 variants in idiopathic generalized epilepsy. *Nat. genetics* **41**, 954–955 (2009).

58. D'agostino, D. *et al.* Mutations and polymorphisms of the clcn2 gene in idiopathic epilepsy. *Neurol.* **63**, 1500–1502 (2004).

59. Cortez, M. *et al.* Disruption of clc-2 expression is associated with progressive neurodegeneration in aging mice. *Neurosci.* **167**, 154–162 (2010).

60. Niemeyer, M. I. *et al.* No evidence for a role of clcn2 variants in idiopathic generalized epilepsy. *Nat. genetics* **42**, 3–3 (2010).

61. Ratté, S. & Prescott, S. A. Clc-2 channels regulate neuronal excitability, not intracellular chloride levels. *J. Neurosci.* **31**, 15838–15843 (2011).

62. Pande, V. S., Beauchamp, K. & Bowman, G. R. Everything you wanted to know about markov state models but were afraid to ask. *Methods* **52**, 99 – 105 (2010). DOI 10.1016/j.ymeth.2010.06.002.

63. Schwantes, C. R., McGibbon, R. T. & Pande, V. S. Perspective: Markov models for long-timescale biomolecular dynamics. *J. Chem. Phys.* **141**, 090901 (2014). DOI 10.1063/1.4895044.

64. Chodera, J. D. & Noé, F. Markov state models of biomolecular conformational dynamics. *Curr. Opin. Struct. Biol.* **25**, 135–144 (2014). DOI 10.1016/j.sbi.2014.04.002.

65. Bowman, G. R., Pande, V. S. & Noé, F. *An Introduction to Markov State Models and Their Application to Long Timescale Molecular Simulation*, vol. 797 (Springer, 2014).

66. Schwantes, C. R. & Pande, V. S. Improvements in markov state model construction reveal many non-native interactions in the folding of ntl9. *J. chemical theory computation* **9**, 2000–2009 (2013).

67. Pérez-Hernández, G., Paul, F., Giorgino, T., De Fabritiis, G. & Noé, F. Identification of slow molecular order parameters for markov model construction. *The J. chemical physics* **139**, 07B604_1 (2013).

68. Naritomi, Y. & Fuchigami, S. Slow dynamics in protein fluctuations revealed by time-structure based independent component analysis: The case of domain motions. *The J. chemical physics* **134**, 02B617 (2011).

69. De Jesús-Pérez, J. J. *et al.* Gating the glutamate gate of clc-2 chloride channel by pore occupancy. *The J. general physiology* **147**, 25–37 (2016).

70. Sánchez-Rodríguez, J. E. *et al.* Sequential interaction of chloride and proton ions with the fast gate steer the voltage-dependent gating in clc-2 chloride channels. *The J. physiology* **590**, 4239–4253 (2012).

71. Traverso, S., Zifarelli, G., Aiello, R. & Pusch, M. Proton sensing of clc-0 mutant e166d. *The J. general physiology* **127**, 51–66 (2006).

72. Zifarelli, G. & Pusch, M. The role of protons in fast and slow gating of the torpedo chloride channel clc-0. *Eur. Biophys. J.* **39**, 869–875 (2010).

73. Chen, T.-Y. Coupling gating with ion permeation in clc channels. *Sci. Signal.* **2003**, pe23–pe23 (2003).

74. Chen, T.-Y., Chen, M.-F. & Lin, C.-W. Electrostatic control and chloride regulation of the fast gating of clc-0 chloride channels. *The J. general physiology* **122**, 641–651 (2003).

75. Sánchez-Rodríguez, J. E., Santiago-Castillo, D., José, A. & Arreola, J. Permeant anions contribute to voltage dependence of clc-2 chloride channel by interacting with the protopore gate. *The J. physiology* **588**, 2545–2556 (2010).

76. Pusch, M., Jordt, S.-E., Stein, V. & Jentsch, T. J. Chloride dependence of hyperpolarization-activated chloride channel gates. *The J. Physiol.* **515**, 341–353 (1999).

77. Zifarelli, G., Murgia, A. R., Soliani, P. & Pusch, M. Intracellular proton regulation of clc-0. *The J. general physiology* **132**, 185–198 (2008).

78. Abraham, S. J. *et al.* 13c nmr detects conformational change in the 100-kd membrane transporter clc-ec1. *J. biomolecular NMR* **61**, 209–226 (2015).

79. Feng, L., Campbell, E. B., Hsiung, Y. & MacKinnon, R. Structure of a eukaryotic CLC transporter defines an intermediate state in the transport cycle. *Sci.* **330**, 635–641 (2010).

80. Crooks, G. E., Hon, G., Chandonia, J.-M. & Brenner, S. E. Weblogo: a sequence logo generator. *Genome research* **14**, 1188–1190 (2004).

81. Schneider, T. D. & Stephens, R. M. Sequence logos: a new way to display consensus sequences. *Nucleic acids research* **18**, 6097–6100 (1990).

82. Schwantes, C. R. & Pande, V. S. Modeling molecular kinetics with tica and the kernel trick. *J. chemical theory computation* **11**, 600–608 (2015).

83. Traverso, S., Elia, L. & Pusch, M. Gating competence of constitutively open clc-0 mutants revealed by the interaction with a small organic inhibitor. *The J. general physiology* **122**, 295–306 (2003).

84. Accardi, A. & Pusch, M. Conformational changes in the pore of clc-0. *The J. general physiology* **122**, 277–294 (2003).

85. Jordt, S.-E. & Jentsch, T. J. Molecular dissection of gating in the CLC-2 chloride channel. *The EMBO J.* **16**, 1582–1592 (1997).

86. Arreola, J., Begenisich, T. & Melvin, J. E. Conformation-dependent regulation of inward rectifier chloride channel gating by extracellular protons. *The J. physiology* **541**, 103–112 (2002).

87. Niemeyer, M. I., Cid, L. P., Yusef, Y. R., Briones, R. & Sepúlveda, F. V. Voltage-dependent and-independent titration of specific residues accounts for complex gating of a clc chloride channel by extracellular protons. *The J. physiology* **587**, 1387–1400 (2009).

88. Šali, A. & Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. molecular biology* **234**, 779–815 (1993).

89. Lee, J. *et al.* CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *J. Chem. Theory Comput.* **12**, 405–413 (2016). DOI 10.1021/acs.jctc.5b00935.

90. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The J. chemical physics* **79**, 926–935 (1983).

91. Dickson, C. J. *et al.* Lipid14: The amber lipid force field. *J. Chem. Theory Comput.* **10**, 865–879 (2014). DOI 10.1021/ct4010307.

92. Maier, J. A. *et al.* ff14sb: improving the accuracy of protein side chain and backbone parameters from ff99sb. *J. chemical theory computation* **11**, 3696–3713 (2015).

93. Eastman, P. & Pande, V. Openmm: A hardware-independent framework for molecular simulations. *Comput. Sci. & Eng.* **12**, 34–39 (2010).

94. Shirts, M., Pande, V. S. *et al.* Screen savers of the world unite. *Comput.* **10**, 43 (2006).

95. Pettersen, E. F. *et al.* Ucsf chimera—a visualization system for exploratory research and analysis. *J. computational chemistry* **25**, 1605–1612 (2004).

96. Harrigan, M. P. *et al.* Msmbuilder: Statistical models for biomolecular dynamics. *Biophys. journal* **112**, 10–15 (2017).

97. Chovancova, E. *et al.* Caver 3.0: a tool for the analysis of transport pathways in dynamic protein structures. *PLoS computational biology* **8**, e1002708 (2012).

Acknowledgements

We acknowledge Professor Justin Du Bois for helpful discussion. We acknowledge funding from NIH grants U19 AI109662 and 2R01GM062868 (VSP), from a Stanford Bio-X seed grant (MM, and VSP), and support from the Center for Molecular Analysis and Design (CMAD). In addition to CMAD, AK acknowledges support from a Stanford Interdisciplinary Graduate Fellowship (SIGF) through the Stanford ChEM-H Institute (Chemistry, Engineering, & Medicine for Human Health). We acknowledge the Folding@home donors who contributed to this project (PROJ9625-9643).

Author contributions statement

All authors designed the experiments. K.A.M. performed the research. M.M. and V.S.P. supervised the research. All authors edited and reviewed the manuscript.

Additional information

VSP is a consultant and SAB member of Schrodinger, LLC and Globavir, sits on the Board of Directors of Apeel Inc, Freenome Inc, Omada Health, Patient Ping, Rigetti Computing, and is a General Partner at Andreessen Horowitz. Other authors declare no competing financial interest.