THE UNIVERSITY OF CHICAGO

MULTISCALE SIMULATIONS OF ION TRANSPORT IN MEMBRANE PROTEINS

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF CHEMISTRY

BY

ZHI WANG

CHICAGO, ILLINOIS

AUGUST 2020

To my beloved.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost, I would like to express my greatest gratitude to my advisor, Dr. Gregory A. Voth, who has provided me precious opportunities, excellent guidance, abundant computational resources, and excellent research atmosphere throughout my Ph.D. study. His influence over me can never be overstated, as I have been constantly learned from his vast knowledge, endless passion, dedicated work and rigorous attitude toward science.

I would also like to thank the rest of my thesis committee: Dr. Aaron R. Dinner and Dr. Gregory S. Engel, for their invaluable suggestions and advices on the thesis.

It is hard to overemphasize the influence from Dr. Jessica M. J. Swanson, who has provided me powerful direction and vision at the field of biological simulations as well as invaluable and insightful advices to my research projects. She has also given me countless advices on scientific writing for proposals and journal articles.

I would also like to thank the excellent team of the Voth group including Dr. Yuxing Peng, Dr. Sangyun Lee, Dr. Ruibin Liang, Dr. Zhi Yue (Shane), Dr. Christopher Arntsen, Dr. Heather Mayes, Dr. James Dama, Dr. Rajib Biswas, Dr. Rui Sun, Dr. Chen Chen, Dr. Shule Liu, Dr. Dudu Tong, Dr. Yining Han, Dr. Morris Cohen, Dr. Thomas Dannenhoffer-Lafage, Dr. Jacob Wagner, Dr. Peng He, Dr. Xinyou Ma, Mr. Chenghan Li, Mr. Zhefu Li, Ms. Yu Liu (Grace), Mr. Jaehyeok Jin, Mr. Paul Calio, Ms. Laura Watkins, Mr. Zack Jarin, and Mr. Aleksander Durumeric. It would be hard to overstate the importance of frequent conversations within the group.

Last but not least, I would like to thank my family: my wife Yumeng Si, my father Mingshi Wang, and my mother Jiuying Cao, for their unconditional and endless love in my life.

ABSTRACT

Studying the migration of an excess proton or a fluoride anion is crucial to understanding mechanisms of many physiological processes including transepithelial transport and adenosine triphosphate (ATP) synthesis, as these processes generally involve ion transport across a channel or a transporter. However, the molecular nature of proton transport (PT) is complicated, involving the Grotthuss shuttling mechanism, where the covalent and hydrogen bond topologies are dynamically changing. The accurate modeling of fluoride is also very challenging as the polarization effect can be non-trivial in a complicated protein system.

Quantum mechanical approaches, such as *ab initio* molecular dynamics (AIMD) or hybrid quantum mechanics/molecular mechanics (QM/MM) methods, explicitly treat the reactive nature of proton transport and the electronic polarization effect, but at the cost of their computational expense which could cause insufficient free energy sampling, resulting in large systematic errors. Therefore, we employed the multiscale reactive molecular dynamics (MS-RMD) method for proton transport and the classical Drude oscillator model for the fluoride transport taking both computational accuracy and efficiency into account.

For proton transport, the MS-RMD method defines a linear combination of multiple possible protonation configurations ("states") of water molecules and/or titratable residues, each with a specific bonding topology and protonation state. These diabatic basis states constitute the diagonal entries in a Hamiltonian matrix, while off-diagonal terms capture the coupling between states. Through an on-the-fly diagonalization of this Hamiltonian matrix, a potential energy surface is defined along which the nuclei can evolve according to Newtonian dynamics. In this process, chemical bonds are dynamically broken and reconstructed, and the positive charge defect is delocalized (or shared). For fluoride transport, the classical Drude oscillator model,

which has been widely employed to simulate ion channels and transporters, has been shown to improve the ion transport properties in aqueous bulk, the ion paring properties, and the free energy sampling in ion channels/transporters.

In this dissertation, MS-RMD is employed to systematically investigate the proton solvation and transport mechanism through a Cl^-/H^+ antiporter ClC-ec1, a prototype of the ClC antiporters, proteins that stoichiometrically exchange Cl^- and H^+ ions in opposite directions across a membrane. It has been shown that alternative polyatomic anion flux including NO_3^- and SCN^- or site-specific mutation (*e.g.*, the I109F mutation) can partially or completely block the proton flux. With the help of multiscale computer simulations in which the Grotthuss mechanism of proton transport (PT) is treated explicitly, it is demonstrated how the chemical nature of these anions alters the coupling mechanism and qualitatively explain the shifts in the ion stoichiometry. Multidimensional free energy profiles for PT and the coupled changes in hydration are presented for NO_3^- and SCN^-. The calculated proton flux agrees with the experiments, showing reduced or abolished rate. We suggested that the size of anions and interactions with the protein significantly alter hydration and shift its influence on PT from facilitating to inhibiting, and we revealed that the most relevant coordinate to the PT free energy barrier is the water connectivity along the PT pathway which is significantly affected by the nature of the bound anion. In addition, reactive molecular dynamics simulations of explicit proton transport across the central region in the I109F mutant has also been performed, and a two-dimensional free energy profile has been constructed that is consistent with the experimental transport rates. The importance of a phenylalanine gate formed by F109 and F357 and its influence on hydration connectivity through the central proton transport pathway is revealed. Our

work demonstrates how seemingly subtle changes in local conformational dynamics can dictate hydration changes and thus transport properties.

In the following chapter of the dissertation, we present our work of employing the CHARMM-Drude model to investigate the fluoride transport mechanism through a F^- channel Ec2 in the Fluc family, which is discovered in *Escherichia coli*. The Fluc protein transports fluoride across the membrane at the rate of 10^6–10^7 s^{-1}. It has been shown that conserved residues N41, F80, F83, H106 as well as an interfacial Na^+ are crucial to the F^- transport. Our constant-pH molecular dynamics simulations reveal that F^- is deprotonated while His106 remains protonated. Following replica-exchange umbrella sampling (REUS) simulations using both nonpolarizable force field and polarizable force field indicates the resulting diffusion constant as well as rate constant from polarizable simulations matches the experimental measurements better, implying the necessity of utilizing the polarizable force field. Our simulations suggest several F^- binding states, which reveal the electrostatic attraction between F^- and Na^+, Phe C–H···F hydrogen bonds, a H106A side-chain rotation, and a F83A–H106A gate, explaining aforementioned experimental discoveries. These binding states also provide insights on following mutagenetic experiments on conserved residues R19 and T37. These findings help to explain the F^- transport mechanism in Ec2 as well as other Fluc proteins.

CHAPTER 1

INTRODUCTION

1.1 Simulation of Ion Transport in Proteins

Simulating ion transport in proteins is a challenging task. Due to the charge delocalization and/or polarization effect of an ion, it is usually inaccurate to describe it using standard force field with fixed point charges, especially for protons and anions like fluoride.

1.1.1 Simulating Proton Transport

One of the most commonly transported ions is proton. The "proton" in a hydrated biological system is a unit of net positive charge (1) due to a missing electron that can be transported between and among biomolecules (2). However, the molecular nature of proton transport (PT) is complicated, involving the Grotthuss shuttling mechanism, where the covalent and hydrogen bond topologies are dynamically changing. Studying the migration of an excess proton is crucial to understanding mechanisms of many physiological processes including transepithelial transport and adenosine triphosphate (ATP) synthesis, due to their involvement with PT (3-7). For example, the ClC antiporters regulate membrane excitability, transepithelial transport and cell volume in different tissues (8, 9). They function as exchanger proteins which transport both Cl^- and H^+ in opposite directions (10, 11).

Despite the significance, it is complicated to establish the atomistic details of PT in proteins using experimental techniques due to its microscopic nature. Therefore, computational studies can be important to the investigation of PT mechanisms in proteins, adding insights at

molecular scale as well as increased temporal and spatial resolution to experimental data. However, it can be quite complicated to explicitly model proton transport, even in seemingly simple bulk water solution, since it involves charge delocalization, Grotthuss mechanism of proton hopping, and solvent reorganization. Moreover, the migration of an excess proton in confined spaces (*e.g.* proteins and carbon nanotubes) can be coupled to changes in the hydration level along PT pathway (12-18). Hence, force fields mostly adopted in molecular dynamics (MD) simulations like CHARMM (19, 20) and AMBER (21, 22) cannot describe the reactive nature of the delocalized proton in water and the protonation/deprotonation reactions of some ionizable residues in protein.

Quantum mechanical approaches, such as *ab initio* molecular dynamics (AIMD) or hybrid quantum mechanics/molecular mechanics (QM/MM) MD methods, explicitly treat the reactive nature of amino acid protonation/deprotonation and the Grotthuss hopping mechanism of PT, but at the cost of their computational expense which limits their application in free energy samplings of large biomolecular systems. Their high computational cost could cause insufficient free energy sampling, resulting in large systematic errors. Therefore, it is necessary to develop approximations to *ab initio* QM methods that take both computational accuracy and efficiency into account.

Multiscale reactive molecular dynamics (MS-RMD) method (2, 23, 24) is such a method that combines the advantage of quantum mechanics approaches and that of molecular mechanics approaches. This MS-RMD method defines a linear combination of multiple possible protonation configurations ("states") of water molecules and/or titratable residues, each with a specific bonding topology and protonation state. These diabatic basis states constitute the diagonal entries in a Hamiltonian matrix, while off-diagonal terms capture the coupling between states. Through

an on-the-fly diagonalization of this Hamiltonian matrix, a potential energy surface is defined along which the nuclei can evolve according to Newtonian dynamics. In this process, chemical bonds are dynamically broken and reconstructed, and the positive charge defect is delocalized (or shared).

MS-RMD is "multiscale" in the sense that (1) the region where PT reactions occur is described by the reactive force field allowing for changes in bonding topology and charge delocalization whereas other parts of the simulation system is described by nonreactive force field (*e.g.* CHARMM (19, 20) force field) which significantly increases the computational efficiency; (2) the reactive region is changed and assigned on-the-fly based on the position of excess charge; (3) the quantum information of electronic states is variationally encoded into the RMD model through FitRMD parametrization. As shown in our previous publications, MS-RMD is several orders of magnitude more computationally efficient than QM/MM MD, while still accurately describing the charge delocalization and reactive nature of the PT (25). With these advantages, MS-RMD method was adopted for the study of PT in various biological systems (15-18, 26-28).

In the CHAPTER 2 (17) and CHAPTER 3 (18), MS-RMD is employed to systematically investigate the proton solvation and transport mechanism through a Cl^-/H^+ antiporter ClC-ec1. ClC-ec1 is a prototype of the ClC antiporters, proteins that stoichiometrically exchange Cl^- and H^+ ions in opposite directions across a membrane (10, 29). It has been shown that other polyatomic anions, such as NO_3^- and SCN^-, can also be transported by ClC-ec1, but with partially or completely uncoupled proton flux (30). In the CHAPTER 2 (17), with the help of multiscale computer simulations in which the Grotthuss mechanism of proton transport (PT) is treated explicitly, it is demonstrated how the chemical nature of these anions alters the coupling

mechanism and qualitatively explain the shifts in the ion stoichiometry. Multidimensional free energy profiles for PT and the coupled changes in hydration are presented for NO_3^- and SCN^-. The calculated proton conductances agree with experiment, showing reduced or abolished proton flux. Surprisingly, the proton affinity of the anion is less influential on the PT, while its size and interactions with the protein significantly alter hydration and shift its influence on PT from facilitating to inhibiting. It is found that the hydration of the cavity below the anion is relatively fast, but connecting the water network past the steric hindrance of these polyatomic anions is quite slow. Hence, the most relevant coordinate to the PT free energy barrier is the water connectivity along the PT pathway – but importantly only in the presence of the excess proton – and this coordinate is significantly affected by the nature of the bound anion. This work again demonstrates how PT is intrinsically coupled with protein cavity hydration changes as well as influenced by the protein environment. It additionally suggests ways in which ion exchange can be modulated and exchange stoichiometries altered.

Other than alternative anion flux, experiments have demonstrated that several mutations, including I109F, affect ion transport (31). The I109F mutation has been shown to decrease the Cl^- and H^+ transport rates by an order of magnitude (31). In the CHAPTER 3 (18), reactive molecular dynamics simulations of explicit proton transport across the central region in the I109F mutant has been performed, and a two-dimensional free energy profile has been constructed that is consistent with the experimental transport rates. The importance of a phenylalanine gate formed by F109 and F357 and its influence on hydration connectivity through the central proton transport pathway is revealed. This work demonstrates how seemingly subtle changes in local conformational dynamics can dictate hydration changes and thus transport properties.

1.1.2 Simulating Fluoride Transport

Fluoride is another ion that is very challenging to simulate, because an accurate description of the interactions between ion and ion, ion and water, and ion and biomolecules is required. Several nonpolarizable models for the fluoride ion have been developed (32-35). However, these nonpolarizable models can hardly account for the non-trivial hydrogen bond between fluoride anion and hydrogen donor where induced electronic polarization warrants flexible charge description that is not included in a nonpolarizable force field. Thus, nonpolarizable models tend to overestimate the electrostatic interaction between fluoride and hydrogen-donating molecules (and, similarly, cations), and they are likely to generate higher energy barriers (36). In addition, such effect varies as the chemical environment of F^- changes, making the approach of implicitly including atomic polarization (37, 38) unable to describe the polarization effect uniformly well in a complicated biological system. The influence of induced polarization by ions has also been shown crucial to the study of ion channels (39-41).

Quantum mechanical (QM) approaches explicitly treat the polarization effect by calculating the electronic density and/or wavefunctions. However, their computational cost could cause insufficient statistical sampling, leading to errors in free energy samplings of large biomolecules (42). Considering the balance between accuracy and computational expense, one can adopt a polarizable force field (43-46) as an alternative approach. This approach has been shown to reproduce *ab initio* (MP4/aug-cc-pVTZ) enthalpy of formation (47), experimental hydration free energies (48), and neutron- and x-ray diffraction data (49).

Three different methods have been developed to explicitly model electronic polarization, *i.e.*, induced dipole model (50, 51), classical Drude oscillator model (52-55), and fluctuation

charge model (56, 57). The CHARMM-Drude force field (58-63), which is based on the Drude oscillator model, has been employed to simulate ion channels and transporters (36, 64-67). It has been shown the polarizable Drude force field improves the ion transport properties in aqueous bulk (68), the ion paring properties (69), and the free energy sampling in ion channels/transporters (36), compared with the nonpolarizable force field. The CHARMM-Drude force field also benefits from its computational efficiency. The computational cost is only about 1.2 to 1.8 times greater than that of fixed-charge CHARMM (70). The force field is also implemented in multiple MD programs including CHARMM, NAMD, AMBER, GROMACS, and OpenMM, and it covers multiple types of biomolecules including protein, liquid, ions, DNA, and sugar, making it versatile and applicable to various simulation systems (71).

In the CHAPTER 4, the CHARMM-Drude model is employed to investigate the fluoride transport mechanism through a F^- channel Ec2 in the Fluc family, which is discovered in *Escherichia coli*. The Fluc protein transports fluoride across the membrane at the rate of 10^6–10^7 s^{-1} (72, 73). It has been shown that conserved residues N41, F80, F83, H106 (74) as well as an interfacial Na^+ (75) are crucial to the F^- transport. Our constant-pH molecular dynamics simulations (76, 77) reveal that F^- is deprotonated while His106 remains protonated. Following replica-exchange umbrella sampling (REUS) (78) simulations using both nonpolarizable force field and polarizable force field indicates the resulting diffusion constant as well as rate constant from polarizable simulations matches the experimental measurements (72, 73, 79, 80) better, implying the necessity of utilizing the polarizable force field. Our simulations suggest several F^- binding states, which reveal the electrostatic attraction between F^- and Na^+, Phe C–H···F hydrogen bonds, a H106A side-chain rotation, and a F83A–H106A gate, explaining aforementioned experimental discoveries (74, 75). These binding states also provide insights on

following mutagenetic experiments on conserved residues R19 and T37. These findings help to explain the F^- transport mechanism in Ec2 as well as other Fluc proteins.

1.2 Multiscale Reactive Molecular Dynamics

Standard molecular dynamics approaches are incapable of describing charge delocalization or Grotthuss proton shuttling due to their fixed bonding topologies. To overcome these limitations, the multiscale reactive molecular dynamics (MS-RMD) approach (23, 24) defines a linear combination of multiple possible protonation configurations ("states") of water molecules and/or titratable residues, each with a specific bonding topology and protonation state. These diabatic basis states constitute the diagonal entries in a Hamiltonian matrix, while off-diagonal terms capture the coupling between states. Through an on-the-fly diagonalization of this Hamiltonian matrix (described below), a potential energy surface is defined along which the nuclei can evolve according to Newtonian dynamics making and breaking chemical bonds, and delocalizing (or sharing) the positive charge defect.

The MS-RMD Hamiltonian **H** is defined as

$$\mathbf{H} = \sum_{ij} |i(\mathbf{r})\rangle h_{ij}(\mathbf{r}) \langle j(\mathbf{r})|, \tag{1.1}$$

where **r** are the system nuclear coordinates, h_{ii} is the potential energy surface for the diabatic state $|i\rangle$ described mostly by a classical force field, and h_{ij} is the coupling between states $|i\rangle$ and $|j\rangle$, which usually requires further parametrization. The on-the-fly diagonalization of the matrix representation of the above Hamiltonian provides the energy and the eigenvector of the ground state at each configuration of the nuclei **r** by solving the following eigenequation

$$\mathrm{H}\mathbf{c} = E_0\mathbf{c},$$

1.2)

where H is the matrix representation of the Hamiltonian **H**, **c** denotes the vector of ground state, and E_0 is the ground-state energy. The simulation system is propagated using the force calculated from the Hellmann-Feynman theorem:

$$\mathbf{F} = -\sum_{ij} c_i c_j \boldsymbol{\nabla} h_{ij},$$

1.3)

where c_i is the i-th component of the ground-state vector.

The diagonal elements h_{ii} of the MS-RMD matrix are mostly given by the potential energy function of each basis state i. However, due to the excess proton in the system, additional terms need to be added. In a biological system, either a titratable residue (*e.g.*, glutamate) or water is protonated at each state. The h_{ii} corresponding to the state with protonated glutamate (GLUH) is described as

$$h_{ii}^{GLUH} = V_{surrounding}^{intra} + V_{GLUH}^{intra} + V_{surrounding,GLUH}^{inter} + V_{ii},$$

1.4)

where the first three terms are the inter- and the intra-molecular potentials of protonated glutamate and all other surrounding molecules, such as waters, other protein residues, lipids, and ions in the system. Notably, the O-H bond in the carboxyl (-COOH) group of GLUH is described by a Morse potential, $U^{Morse}(r)$:

$$U^{Morse}(r) = D\left[1 - \exp\left(-\alpha(r - r_0)\right)\right]^2,$$

1.5)

where r is the O-H bond length, and D, α, and r_0 are parameters, which are taken from our previous work (81). Since the zero-point energies of two protonated forms of glutamate and

water do not match under the description of classical force fields, V_{ii} is introduced to compensate the constant energy shift between two states.

In order to correct the overestimated electrostatic interaction between opposite charges on a hydronium and deprotonated glutamate at a short distance (82), two repulsive terms, $V_{OO_k}^{rep}$ and $V_{H_jO_k}^{rep}$, are introduced in h_{ii} corresponding to the state with deprotonated glutamate:

$$V_{OO_\epsilon}^{rep}\left(R_{OO_\epsilon}, R_{H_jO_\epsilon}; j = 1,2,3\right) = B\exp\left[-b\left(R_{OO_\epsilon} - d_{OO}^0\right)\right] \cdot \sum_j^3 \exp\left[-b'\left(R_{H_jO_\epsilon}\right)^2\right], \quad (1.6)$$

$$V_{H_jO_\epsilon}^{rep}\left(R_{HO_\epsilon}\right) = C\exp\left[-c\left(R_{H_jO_\epsilon} - d_{OH}^0\right)\right], \quad (1.7)$$

where R_{OO_ϵ} is the distance between the hydronium oxygen, O, and the carboxyl oxygen of glutamate, O_ϵ (OE1 and OE2 in the PDB), and $R_{H_jO_\epsilon}$ is the distance between each of three hydronium hydrogen atoms, H_j, and the carboxyl oxygen of glutamate. The functional forms for the repulsive terms are the same as those used in our previous model (81). B, b, b', C, and c are fitted parameters, and d_{OO}^0 and d_{OH}^0 are fixed to the same value used in our previous publication (25), which are 2.4 and 1.0 Å, respectively.

The off-diagonal element h_{ij} for the coupling between protonated glutamate and water is given by,

$$h_{ij}^{GLUH} = c_1\exp[-c_2(r_{OH} - c_3)^2], \quad (1.8)$$

where r_{OH} is the distance between the donor oxygen of the carboxyl group of glutamate and the acceptor hydrogen of the adjacent hydronium molecule, c_1, c_2, and c_3 are fitted parameters. The

h_{ij} term for the coupling between hydronium and water is the same as the one used in reference (83).

The MS-RMD formalism also provides a convenient and physically intuitive description of the excess proton, or center of the excess charge (CEC), and its coordinate $\mathbf{r}_{\mathrm{CEC}}$ is defined in our previous work (83, 84) as

$$\mathbf{r}_{\mathrm{CEC}} = \sum_{i=1}^{N} c_i^2 \mathbf{r}_{\mathrm{COC}}^i, \tag{1.9}$$

in which the summation is performed over all N MS-RMD states, and c_i and $\mathbf{r}_{\mathrm{COC}}^i$ correspond to the population and the center of charge (COC) in the i-th MS-RMD state, respectively. The COC of i-th state is given by

$$\mathbf{r}_{\mathrm{COC}}^i = \left(\sum_{k\in\{i\}} |q_k| \mathbf{r}_k\right) / \left(\sum_{k\in\{i\}} |q_k|\right), \tag{1.10}$$

where $\{i\}$ denotes the MS-RMD complex in the i-th state, and q_k and $\mathbf{r}_k$ stand for the atomic partial charge and the coordinate of the k-th atom, respectively.

Assumptions of the method, more detailed explanations of the functionals in the Hamiltonian matrix, and the parameters for the excess proton in water have been extensively discussed in our previous publications (23, 24, 83-87).

The key to MS-RMD modelling is the optimization of the parameters introduced in the Hamiltonian matrix. This is generally done by performing quantum mechanical single point calculations and force-matching (88) for an ensemble. The overall fitting procedure has also been extensively described and justified in previous publications (25, 81).

1.3 References

1. Knight C & Voth GA (2012) The Curious Case of the Hydrated Proton. *Acc. Chem. Res.* 45(1):101–109.

2. Swanson JMJ, *et al.* (2007) Proton Solvation and Transport in Aqueous and Biomolecular Systems: Insights from Computer Simulations. *J. Phys. Chem. B* 111(17):4300–4314.

3. Wikstrom MKF (1977) Proton pump coupled to cytochrome *c* oxidase in mitochondria. *Nature* 266(5599):271–273.

4. Steinmeyer K, Schwappach B, Bens M, Vandewalle A, & Jentsch TJ (1995) Cloning and Functional Expression of Rat CLC-5, a Chloride Channel Related to Kidney Disease. *J. Biol. Chem.* 270(52):31172–31177.

5. Maduke M, Miller C, & Mindell JA (2000) A Decade of CLC Chloride Channels: Structure, Mechanism, and Many Unsettled Questions. *Annu. Rev. Biophys. Biomol. Struct.* 29(1):411–438.

6. Jentsch TJ (2008) CLC Chloride Channels and Transporters: From Genes to Protein Structure, Pathology and Physiology. *Crit. Rev. Biochem. Mol. Biol.* 43(1):3–36.

7. Weinert S, *et al.* (2010) Lysosomal Pathology and Osteopetrosis upon Loss of H^+-Driven Lysosomal Cl^- Accumulation. *Science* 328(5984):1401–1403.

8. Günther W, Lüchow A, Cluzeaud F, Vandewalle A, & Jentsch TJ (1998) ClC-5, the chloride channel mutated in Dent's disease, colocalizes with the proton pump in endocytotically active kidney cells. *Proc. Natl. Acad. Sci. U. S. A.* 95(14):8075–8080.

9. Devuyst O, Courtoy PJ, Christie PT, Thakker RV, & Beauwens R (1999) Intra-renal and subcellular distribution of the human chloride channel, CLC-5, reveals a pathophysiological basis for Dent's disease. *Hum. Mol. Genet.* 8(2):247–257.

10. Accardi A & Miller C (2004) Secondary active transport mediated by a prokaryotic homologue of ClC Cl^- channels. *Nature* 427(6977):803–807.

11. Novarino G, Weinert S, Rickheit G, & Jentsch TJ (2010) Endosomal Chloride-Proton Exchange Rather Than Chloride Conductance Is Crucial for Renal Endocytosis. *Science* 328(5984):1398–1401.

12. Yamashita T & Voth GA (2012) Insights into the Mechanism of Proton Transport in Cytochrome *c* Oxidase. *J. Am. Chem. Soc.* 134(2):1147–1152.

13. Goyal P, Lu J, Yang S, Gunner MR, & Cui Q (2013) Changing hydration level in an internal cavity modulates the proton affinity of a key glutamate in cytochrome *c* oxidase. *Proc. Natl. Acad. Sci. U. S. A.* 110(47):18886–18891.

14. Peng Y, Swanson JMJ, Kang S-g, Zhou R, & Voth GA (2015) Hydrated Excess Protons Can Create Their Own Water Wires. *J. Phys. Chem. B* 119(29):9212–9218.

15. Lee S, Swanson JMJ, & Voth GA (2016) Multiscale Simulations Reveal Key Aspects of the Proton Transport Mechanism in the ClC-ec1 Antiporter. *Biophys. J.* 110(6):1334–1345.

16. Liang R, Swanson JMJ, Peng Y, Wikström M, & Voth GA (2016) Multiscale simulations reveal key features of the proton-pumping mechanism in cytochrome *c* oxidase. *Proc. Natl. Acad. Sci. U. S. A.* 113(27):7420–7425.

17. Wang Z, Swanson JMJ, & Voth GA (2018) Modulating the Chemical Transport Properties of a Transmembrane Antiporter via Alternative Anion Flux. *J. Am. Chem. Soc.* 140(48):16535–16543.

18. Wang Z, Swanson JMJ, & Voth GA (2020) Local Conformational Dynamics Regulating Transport Properties of a Cl^-/H^+ Antiporter. *J. Comput. Chem.* 41(6):513–519.

19. MacKerell AD, Jr., Feig M, & Brooks CL, III. (2004) Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulations. *J. Comput. Chem.* 25(11):1400–1415.

20. Klauda JB, *et al.* (2010) Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *J. Phys. Chem. B* 114(23):7830–7843.

21. Lindorff-Larsen K, *et al.* (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Struct., Funct., Bioinf.* 78(8):1950–1958.

22. Li D-W & Brüschweiler R (2010) NMR-Based Protein Potentials. *Angew. Chem. Int. Ed.* 49(38):6778–6780.

23. Knight C, Lindberg GE, & Voth GA (2012) Multiscale reactive molecular dynamics. *J. Chem. Phys.* 137(22):22A525.

24. Yamashita T, Peng Y, Knight C, & Voth GA (2012) Computationally Efficient Multiconfigurational Reactive Molecular Dynamics. *J. Chem. Theory Comput.* 8(12):4863–4875.

25. Lee S, Liang R, Voth GA, & Swanson JMJ (2016) Computationally Efficient Multiscale Reactive Molecular Dynamics to Describe Amino Acid Deprotonation in Proteins. *J. Chem. Theory Comput.* 12(2):879–891.

26. Lee S, Mayes HB, Swanson JMJ, & Voth GA (2016) The Origin of Coupled Chloride and Proton Transport in a Cl^-/H^+ Antiporter. *J. Am. Chem. Soc.* 138(45):14923–14930.

27. Liang R, Swanson JMJ, Wikström M, & Voth GA (2017) Understanding the essential proton-pumping kinetic gates and decoupling mutations in cytochrome *c* oxidase. *Proc. Natl. Acad. Sci. U. S. A.* 114(23):5924–5929.

28. Parker JL, *et al.* (2017) Proton movement and coupling in the POT family of peptide transporters. *Proc. Natl. Acad. Sci. U. S. A.* 114(50):13182–13187.

29. Basilio D, Noack K, Picollo A, & Accardi A (2014) Conformational changes required for H^+/Cl^- exchange mediated by a CLC transporter. *Nat. Struct. Mol. Biol.* 21(5):456–463.

30. Nguitragool W & Miller C (2006) Uncoupling of a CLC Cl^-/H^+ Exchange Transporter by Polyatomic Anions. *J. Mol. Biol.* 362(4):682–690.

31. Han W, Cheng RC, Maduke MC, & Tajkhorshid E (2014) Water access points and hydration pathways in CLC H^+/Cl^- transporters. *Proc. Natl. Acad. Sci. U. S. A.* 111(5):1819–1824.

32. Senn HM, O'Hagan D, & Thiel W (2005) Insight into Enzymatic C−F Bond Formation from QM and QM/MM Calculations. *J. Am. Chem. Soc.* 127(39):13643–13655.

33. Jensen KP & Jorgensen WL (2006) Halide, Ammonium, and Alkali Metal Ion Parameters for Modeling Aqueous Solutions. *J. Chem. Theory Comput.* 2(6):1499–1509.

34. Joung IS & Cheatham TE, III. (2008) Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. *J. Phys. Chem. B* 112(30):9020–9041.

35. Joung IS & Cheatham TE, III. (2009) Molecular Dynamics Simulations of the Dynamic and Energetic Properties of Alkali and Halide Ions Using Water-Model-Specific Ion Parameters. *J. Phys. Chem. B* 113(40):13279–13290.

36. Sun R-N & Gong H (2017) Simulating the Activation of Voltage Sensing Domain for a Voltage-Gated Sodium Channel Using Polarizable Force Field. *J. Phys. Chem. Lett.* 8(5):901–908.

37. Jämbeck JPM & Lyubartsev AP (2013) Implicit inclusion of atomic polarization in modeling of partitioning between water and lipid bilayers. *Phys. Chem. Chem. Phys.* 15(13):4677–4686.

38. Jia X & Li P (2019) Solvation Free Energy Calculation Using a Fixed-Charge Model: Implicit and Explicit Treatments of the Polarization Effect. *J. Phys. Chem. B* 123(5):1139–1148.

39. Roux B (1993) Non-additivity in cation–peptide interactions. A molecular dynamics and ab initio study of Na^+ in the gramicidin channel. *Chem. Phys. Lett.* 212(3):231–240.

40. Allen TW, Andersen OS, & Roux B (2006) Ion Permeation through a Narrow Channel: Using Gramicidin to Ascertain All-Atom Molecular Dynamics Potential of Mean Force Methodology and Biomolecular Force Fields. *Biophys. J.* 90(10):3447–3468.

41. Bucher D, Guidoni L, Maurer P, & Rothlisberger U (2009) Developing Improved Charge Sets for the Modeling of the KcsA K^+ Channel Using QM/MM Electrostatic Potentials. *J. Chem. Theory Comput.* 5(8):2173–2179.

42. Yu H & Roux B (2009) On the Utilization of Energy Minimization to the Study of Ion Selectivity. *Biophys. J.* 97(8):L15–L17.

43. Steven W. Rick & Stuart SJ (2002) Potentials and Algorithms for Incorporating Polarizability in Computer Simulations. *Rev. Comput. Chem.*, eds Kenny B. Lipkowitz & Boyd DB), pp 89–146.

44. Yu H & van Gunsteren WF (2005) Accounting for polarization in molecular simulation. *Comput. Phys. Commun.* 172(2):69–85.

45. Warshel A, Kato M, & Pisliakov AV (2007) Polarizable Force Fields: History, Test Cases, and Prospects. *J. Chem. Theory Comput.* 3(6):2034–2045.

46. Lopes PEM, Roux B, & MacKerell AD, Jr. (2009) Molecular modeling and dynamics studies with explicit inclusion of electronic polarizability: theory and applications. *Theor. Chem. Acc.* 124(1):11–28.

47. Xantheas SS & Dang LX (1996) Critical Study of Fluoride−Water Interactions. *J. Phys. Chem.* 100(10):3989−3995.

48. Lamoureux G & Roux B (2006) Absolute Hydration Free Energy Scale for Alkali and Halide Ions Established from Simulations with a Polarizable Force Field. *J. Phys. Chem. B* 110(7):3308–3322.

49. Trumm M, *et al.* (2012) Modeling the hydration of mono-atomic anions from the gas phase to the bulk phase: The case of the halide ions F^-, Cl^-, and Br^-. *J. Chem. Phys.* 136(4):044509.

50. Applequist J, Carl JR, & Fung K-K (1972) An Atom Dipole Interaction Model for Molecular Polarizability. Application to Polyatomic Molecules and Determination of Atom Polarizabilities. *J. Am. Chem. Soc.* 94(9):2952–2960.

51. Thole BT (1981) MOLECULAR POLARIZABILITIES CALCULATED WITH A MODIFIED DIPOLE INTERACTION. *Chem. Phys.* 59(3):341–350.

52. Drude P (1902) *The Theory of Optics* (Longmans, Green, and Co., New York); trans C. Riborg Mann & Millikan RA.

53. Pratt LR (1980) Effective field of a dipole in non-polar polarizable fluids. *Mol. Phys.* 40(2):347–360.

54. Høye JS & Stell G (1980) Dielectric theory for polar molecules with fluctuating polarizability. *J. Chem. Phys.* 73(1):461–468.

55. Cao J & Berne BJ (1993) Theory and simulation of polar and nonpolar polarizable fluids. *J. Chem. Phys.* 99(9):6998–7011.

56. Rappe AK & Goddard WA, III. (1991) Charge Equilibration for Molecular Dynamics Simulations. *J. Phys. Chem.* 95(8):3358–3363.

57. Banks JL, *et al.* (1999) Parametrizing a polarizable force field from *ab initio* data. I. The fluctuating point charge model. *J. Chem. Phys.* 110(2):741–754.

58. Lamoureux G, Harder E, Vorobyov IV, Roux B, & MacKerell AD, Jr. (2006) A polarizable model of water for molecular dynamics simulations of biomolecules. *Chem. Phys. Lett.* 418(1):245–249.

59. Yu H, *et al.* (2010) Simulating Monovalent and Divalent Ions in Aqueous Solution Using a Drude Polarizable Force Field. *J. Chem. Theory Comput.* 6(3):774–786.

60. Luo Y, Jiang W, Yu H, MacKerell AD, Jr., & Roux B (2013) Simulation study of ion pairing in concentrated aqueous salt solutions with a polarizable force field. *Faraday Discuss.* 160:135–149.

61. Chowdhary J, *et al.* (2013) A Polarizable Force Field of Dipalmitoylphosphatidylcholine Based on the Classical Drude Model for Molecular Dynamics Simulations of Lipids. *J. Phys. Chem. B* 117(31):9142–9160.

62. Lopes PEM, *et al.* (2013) Polarizable Force Field for Peptides and Proteins Based on the Classical Drude Oscillator. *J. Chem. Theory Comput.* 9(12):5430–5449.

63. Li H, *et al.* (2017) Drude Polarizable Force Field for Molecular Dynamics Simulations of Saturated and Unsaturated Zwitterionic Lipids. *J. Chem. Theory Comput.* 13(9):4535–4552.

64. Kratochvil HT, *et al.* (2016) Instantaneous ion configurations in the K^+ ion channel selectivity filter revealed by 2D IR spectroscopy. *Science* 353(6303):1040−1044.

65. Dhakshnamoorthy B, Rohaim A, Rui H, Blachowicz L, & Roux B (2016) Structural and functional characterization of a calcium-activated cation channel from *Tsukamurella paurometabola*. *Nat. Commun.* 7(1):12753.

66. Vergara-Jaque A, Fong P, & Comer J (2017) Iodide Binding in Sodium-Coupled Cotransporters. *J. Chem. Inf. Model.* 57(12):3043–3055.

67. Mayes HB, Lee S, White AD, Voth GA, & Swanson JMJ (2018) Multiscale Kinetic Modeling Reveals an Ensemble of Cl^-/H^+ Exchange Pathways in ClC-ec1 Antiporter. *J. Am. Chem. Soc.* 140(5):1793–1804.

68. Prajapati JD, Mele C, Aksoyoglu MA, Winterhalter M, & Kleinekathöfer U (2020) Computational Modeling of Ion Transport in Bulk and through a Nanopore Using the Drude Polarizable Force Field. *J. Chem. Inf. Model.* 60(6):3188–3203.

69. Flood E, Boiteux C, Lev B, Vorobyov I, & Allen TW (2019) Atomistic Simulations of Membrane Ion Channel Conduction, Gating, and Modulation. *Chem. Rev.* 119(13):7737–7832.

70. Jiang W, *et al.* (2011) High-Performance Scalable Molecular Dynamics Simulations of a Polarizable Force Field Based on Classical Drude Oscillators in NAMD. *J. Phys. Chem. Lett.* 2(2):87–92.

71. Jing Z, *et al.* (2019) Polarizable Force Fields for Biomolecular Simulations: Recent Advances and Applications. *Annu. Rev. Biophys.* 48(1):371–394.

72. Stockbridge RB, Robertson JL, Kolmakova-Partensky L, & Miller C (2013) A family of fluoride-specific ion channels with dual-topology architecture. *eLife* 2:e01084.

73. Smith KD, *et al.* (2015) Yeast Fex1p Is a Constitutively Expressed Fluoride Channel with Functional Asymmetry of Its Two Homologous Domains. *J. Biol. Chem.* 290(32):19874–19887.

74. Last NB, Sun S, Pham MC, & Miller C (2017) Molecular determinants of permeation in a fluoride-specific ion channel. *eLife* 6:e31259.

75. McIlwain BC, Martin K, Hayter EA, & Stockbridge RB (2020) An Interfacial Sodium Ion is an Essential Structural Feature of Fluc Family Fluoride Channels. *J. Mol. Biol.* 432(4):1098−1108.

76. Lee MS, Salsbury FR, Jr., & Brooks CL, III. (2004) Constant-pH Molecular Dynamics Using Continuous Titration Coordinates. *Proteins: Struct., Funct., Bioinf.* 56(4):738–752.

77. Khandogin J & Brooks CL, III. (2005) Constant pH Molecular Dynamics with Proton Tautomerism. *Biophys. J.* 89(1):141–157.

78. Sugita Y, Kitao A, & Okamoto Y (2000) Multidimensional replica-exchange method for free-energy calculations. *J. Chem. Phys.* 113(15):6042–6051.

79. Wang JH (1954) EFFECT OF IONS ON THE SELF-DIFFUSION AND STRUCTURE OF WATER IN AQUEOUS ELECTROLYTIC SOLUTIONS. *J. Phys. Chem.* 58(9):686–692.

80. Vanýsek P (2006) IONIC CONDUCTIVITY AND DIFFUSION AT INFINITE DILUTION. *CRC Handbook of Chemistry and Physics*, ed Lide DR (Taylor and Francis, Boca Raton, FL), Internet Version 2006 Ed, pp 5:76–78.

81. Nelson JG, Peng Y, Silverstein DW, & Swanson JMJ (2014) Multiscale Reactive Molecular Dynamics for Absolute pK_a Predictions and Amino Acid Deprotonation. *J. Chem. Theory Comput.* 10(7):2729–2737.

82. Akin-Ojo O & Wang F (2009) Improving the Point-Charge Description of Hydrogen Bonds by Adaptive Force Matching. *J. Phys. Chem. B* 113(5):1237–1240.

83. Wu Y, Chen H, Wang F, Paesani F, & Voth GA (2008) An Improved Multistate Empirical Valence Bond Model for Aqueous Proton Solvation and Transport. *J. Phys. Chem. B* 112(2):467–482.

84. Day TJF, Soudackov AV, Čuma M, Schmitt UW, & Voth GA (2002) A second generation multistate empirical valence bond model for proton transport in aqueous systems. *J. Chem. Phys.* 117(12):5839–5849.

85. Schmitt UW & Voth GA (1998) Multistate Empirical Valence Bond Model for Proton Transport in Water. *J. Phys. Chem. B* 102(29):5547–5551.

86. Maupin CM, Wong KF, Soudackov AV, Kim S, & Voth GA (2006) A Multistate Empirical Valence Bond Description of Protonatable Amino Acids. *J. Phys. Chem. A* 110(2):631–639.

87. Biswas R, Tse Y-LS, Tokmakoff A, & Voth GA (2016) Role of Presolvation and Anharmonicity in Aqueous Phase Hydrated Proton Solvation and Transport. *J. Phys. Chem. B* 120(8):1793–1804.

88. Izvekov S, Parrinello M, Burnham CJ, & Voth GA (2004) Effective force fields for condensed phase systems from *ab initio* molecular dynamics simulation: A new method for force-matching. *J. Chem. Phys.* 120(23):10896–10913.

CHAPTER 2

MODULATING THE CHEMICAL TRANSPORT PROPERTIES OF THE CLC ANTIPORTER VIA ALTERNATIVE ANION FLUX

2.1 Introduction

The chloride channel (ClC) proteins constitute a family of transmembrane proteins in species ranging from prokaryotes to human beings that regulate a wide range of physiological processes (1-5). Proteins within this family play a crucial role in the extreme acid resistance of bacteria (ClC-ec1) (5) and their genetic disruption or misregulation has been linked to multiple mammalian diseases, including myotonia congenita (ClC-1) (6, 7), retinal degeneration (ClC-2) (4), Bartter's syndrome type III (ClC-Kb) (2), Dent's disease (ClC-5) (1), and osteopetrosis (ClC-7) (8, 9). ClC proteins were initially assumed to be exclusively channels allowing the passive diffusion of Cl^-, until ClC-ec1 was identified as an antiporter where Cl^- and H^+ ions are exchanged in opposite directions (10). Importantly, this protein can also transport other anions, including NO_3^- and SCN^-, with a similar conductance rate but altered proton coupling (11-18). Due to their prevalence, transport diversity, and functional complexity, the ClC proteins have been aptly described as "a never ending source of surprises" (19).

Among the members of ClC proteins, ClC-ec1 (**Figure 2.1** (A)) from *Escherichia coli* is the most structurally and mechanistically investigated (10, 20-24). It exchanges approximately

two Cl^- ions for one H^+ actively and stoichiometrically under normal metabolic conditions (10, 24). As revealed by the wild-type (WT) crystal structure (20, 21), the cross-linked protein (25), and the monomeric form (26), ClC-ec1 is a homodimer in which two subunits function independently yet similarly. Though ClC-ec1 transports various anions, Miller and co-workers observed in electrophysiological experiments that the coupling of the polyatomic anions like NO_3^- and SCN^- to H^+ transport is weakened, or even extinguished, compared to Cl^-/H^+ coupling (15). The stoichiometry of anion to proton flux is also shifted from (2.2±0.1):1 for Cl^-:H^+ flow (24), to 7–10:1 for NO_3^-:H^+ flow, while proton transport (PT) is completely blocked by thiocyanate (15). The goal of this work is to investigate these altered exchange processes in order to deepen our understanding of the ion exchange mechanism, the origin of chloride selectivity and the chemical facets that regulate exchange coupling and stoichiometry. The insights gained in this model system have broader implications for the driving forces of ion transport in other proteins, especially those that demonstrate partial or complete uncoupling effects, such as ClC-4 and ClC-5 (16, 17, 27).

Unravelling the interaction between the transported ions in ClC-ec1 necessitates first understanding their binding sites and transport pathways. Bound to the crystal structure of ClC-ec1 are four Cl^- anions, two in each monomer (see **Figure 2.1** (A)) (20). One of the two anions bound to each monomer is observed in the central binding site (S_{cen}), located in the central cavity and surrounded by several conserved residues including S107, E148, I356, F357, A358, and Y445 (ClC-ec1 numbering). The S_{cen} anion is stabilized by coordination to the backbone nitrogen atoms of I356 and F357, and to the oxygen atoms of S107/Y445 side chains. The other anion is bound to the internal site (S_{int}) with one side facing the internal bulk, and the other side coordinating the S107 and G108 backbone atoms. Crystal structures and simulations have

suggested that the polyatomic anions occupy the same sites (S_{cen} and S_{int}) during transport, but induce a significant shift in both local environment and proton transport functionality (15, 28). For the nitrate-bound ClC, this phenomenon is most likely attributable to the S_{cen} site for which crystallographic and electrophysiological studies have demonstrated synergy with proton transport in ClC-ec1 (29). Similar synergistic interactions were also revealed in an eukaryotic Cl^-/H^+ exchanger (30). Structural examination indicates that S_{cen} is located at the intersection of the proton pathway and the anion pathway, additionally suggesting its potential effect on PT. This is further corroborated by the isothermal titration calorimetry (ITC) experiments that revealed the stability of Cl^-, Br^-, or NO_3^- bound to S_{cen} in ClC-ec1 (31).

In contrast, the S_{int}-bound halide or nitrate observed in the internal part of the anion transport pathway has a low binding affinity (31). This is not the case for SCN^-, which is thought to have a higher affinity for the internal site based on undetectable electron density in S_{cen} in a crystal structure of ClC-ec1 grown in $SeCN^-$ (a crystallographic analog of SCN^-) (15) and ITC experiments (31). We suppose that one of the reasons for uncoupling of H^+ and SCN^- transport could be the low affinity of SCN^- to S_{cen}, since the S_{cen}-unbound ClC-ec1 was shown to be substantially less proton permeable in our previous publications (32, 33). However, that same work demonstrated how a Cl^- anion at the S_{cen} site has the largest impact on proton conductivity, and thus is central to Cl^-/H^+ coupling. For these reasons, and because SCN^- must pass through S_{cen} with an unknown retention time, we focus herein on the impact of SCN^- at S_{cen} on PT. This also allows us to directly compare the effects of SCN^- and NO_3^- to that of Cl^- (all at S_{cen}).

Though many biophysical studies of the anion transport mechanism have been performed (29, 31), the behavior of the transported proton has been more challenging to characterize experimentally. Site-directed mutagenesis experiments have revealed that two glutamate

residues, E203 and E148, act as crucial proton loading sites for H^+ transport (34, 35). However, between these two residues lies a ~15 Å-long hydrophobic cavity lacking in either ionizable residues or crystallographic water molecules, both of which are fundamental elements that enable PT. It therefore seems impossible that excessive protons can transport across such a dry area, when making judgements based solely on information from the crystal structure.

Despite the absence of a "proton carrier" in the crystal structure, the mystery in the hydrophobic cavity has been explained in a number of previous simulation studies in which the explicit PT is treated via a reactive molecular dynamics methodology that allows for Grotthuss proton shuttling between water molecules (32, 36-38). Moreover, non-reactive simulations have been used to predict that water molecules filled this hydrophobic region transiently depending on the protonation states of certain residues (39, 40). Multiscale reactive molecular dynamics (MS-RMD) simulations went further to demonstrate an important dynamical coupling of the PT with the increased hydration between E203 and E148 (32, 36). This work also revealed that, in contrast to the expectation that PT from E203 to E148 should be rate-limiting and dependent on the presence of Cl^- at S_{cen}, the outward PT (E203 to E148) is actually relatively fast, even in the absence of Cl^-_{cen} (32). Our subsequent work suggested that the role of Cl^- bound to S_{cen} during outward proton flux could instead be the facilitation of PT release from E148 to external bulk (38). Our most recent work suggests that both of these PT steps may be rate-limiting and facilitated by Cl^- occupancy in different contributing pathways to the total outward proton flux (33). In contrast, the rate of proton uptake from external bulk to E148 was shown to be relatively fast at physiological pH (33, 38). This protonation does not require E148 to rotate up and is minimally affected by the presence or absence of an anion at S_{cen}. The slowest rate for proton uptake to E148 was calculated to be $(4\pm1) \times 10^2$ ms^{-1} when Cl^-_{cen} is absent and E148 is rotated

down (38), which is more than two orders of magnitude faster than the Cl^- turnover rate (2.3 ms^{-1}) and PT rate (1.0 ms^{-1}) (23, 24). In the presence of Cl^-_{cen}, the proton uptake rate is slightly increased (less than one order of magnitude) (38). We further revealed that the inward proton flux is rate-limited by transport from E148 to E203, and dependent on the presence of Cl^-_{cen} (33, 38). Collectively, these previous studies explain the mechanism of ion exchange and the origin of the 2.2:1 ratio for Cl^-:H^+ flux. In the present paper, we consider the likely possibility that these rate-limiting steps are not transferable, and that PT within the central cavity may be significantly slowed by the presence of polyatomic anions at S_{cen}. By slowing or blocking PT between E203 and E148, these anions could uncouple PT from anion transport since the requisite protonation of E148 for anion transport would still occur via protonation from the external side of the membrane. Therefore, we focus our attention in this work on PT between E203 and E148 as influenced by polyatomic anions.

Similar to the Cl^--bound system, the ease of hydrating the central hydrophobic region was investigated with standard molecular dynamics (MD) simulations in the NO_3^-- and SCN^--bound systems; Jiang *et al.* observed shortened water networks within the central region of these two PT-suppressing/blocking systems (28). However, analyzing the water network in the absence of an explicit hydrated excess proton (*i.e.*, not including the delocalized protonic charge defect and Grotthuss proton shuttling) can present an incomplete, if not completely incorrect, picture of how hydration influences PT. The hydrated excess proton has been clearly shown to significantly alter water structure and dynamics (32, 41, 42), including within ClC-ec1 (32). Thus, the causality between the change in water structure and the decrease in PT rates warrants further exploration, by explicitly including the dynamic coupled behavior of proton transport and water hydration in these polyatomic anion-bound systems.

The investigation of the explicit proton transport as affected by polyatomic anions is challenging in three distinct ways: (1) as noted earlier the explicit excess proton moves along the water chain through consecutive O–H bond breaking and reforming steps (Grotthuss mechanism), which are not handled correctly in standard non-reactive MD simulations; (2) as was proposed by Jiang *et al.* (28), these polyatomic anions in the central cavity, unlike monatomic Cl^-, show two distinct binding modes (**Figure 2.1** (B) and (C)), which must be treated carefully since the environment within the cavity will clearly influence the PT rate (32); and (3) the polyatomic anions might even interact with excess protons directly in the protein environment during PT between E203 and E148.

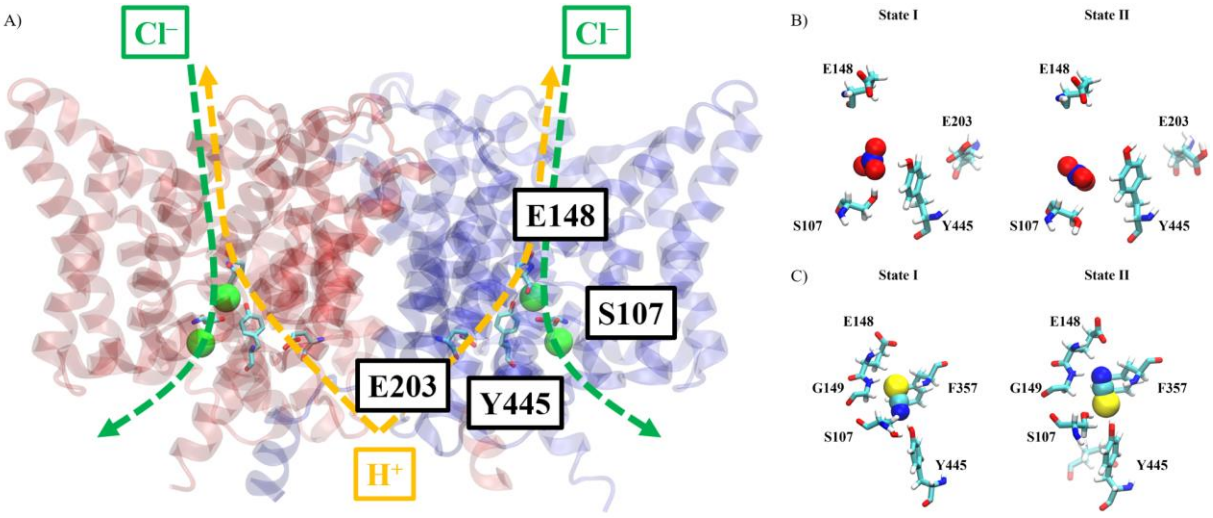

Figure 2.1 The crystal structure of WT ClC-ec1, and two distinct states of S_{cen} anion-binding for nitrate and thiocyanate
(A) The ClC-ec1 homodimer is shown with the key residues rendered and labeled in monomer A (blue). Approximate Cl^- and H^+ transport pathways are indicated by green and orange dashed arrows, respectively. Two bound chloride anions (green spheres) occupy the central (S_{cen}) and internal (S_{int}) binding sites. The anion bound to S_{cen} lies in the intersection of two ion pathways. (B) In the nitrate-bound protein, two states are characterized by the coordination of nitrate anion (van der Waals spheres) to S107 or Y445. (C) In the thiocyanate system, the state is differentiated by the orientation of the anion.

In this work, we demonstrate how ion exchange is modulated by the chemical nature of the anion and the protein environment. We were able to overcome the challenges above by using

MS-RMD simulations and electronic structure calculations to investigate the influence of polyatomic anions on PT. Quantitative free energy calculations of explicit PT and the associated changes in hydration reveal a novel three-phase mechanistic coupling between hydration and PT in the central cavity. As opposed to simply the frequency of water filling the protein cavity, we find that the hydrogen-bond connectivity in narrow regions of the cavity is the critical feature that influences PT and ion coupling. Thus, PT through the central cavity can be either facilitated (Cl^-) or inhibited (polyatomic anions) depending on the anion's size, proton affinity, protein interactions, and influence on hydration. Surprisingly, we find the proton affinity of the anion (*i.e.*, the pK_a of its conjugate acid) has little influence, while its size, stability and influence on hydration are the dominant forces of change. These findings help to explain the coupling mechanism in ClC-ec1, and more broadly suggest ways in which ion exchange can be modified.

2.2 Methods

2.2.1 Overall Methods

The employed method for simulating explicit proton transport, MS-RMD, performs an on-the-fly diagonalization of a Hamiltonian matrix to identify a linear combination of "states", each with a different bonding topology, that delocalizes the excess proton's positive charge and enables Grotthuss shuttling without running much more expensive hybrid quantum mechanics/molecular mechanics (QM/MM) MD simulations. The procedures for setting up the system, a more detailed introduction to MS-RMD, and the parametrization procedure of MS-RMD models (43-46) for residues E148 and E203 are explained in detail in the SI. Initial configurations for the simulations were mutated from a structure obtained from a previous study

of this system (32), which was equilibrated from the ClC-ec1 crystal structure (PDB ID: 1OTS) (21). After equilibration for ~500 ns, the production MS-RMD umbrella sampling runs were performed with the RAPTOR software (44) embedded in the LAMMPS MD package (http://lammps.sandia.gov) (47), with umbrella potentials implemented in the PLUMED package (48).

2.2.2 System Setup

The simulation system was composed of the ClC-ec1 homodimer (PDB ID: 1OTS) (21), 163 POPE lipids, 18 anions ($Cl^-/NO_3^-/SCN^-$), and ~11,000 water molecules in a 92 Å × 92 Å × 79 Å box with periodic boundary conditions. The CHARMM-CMAP force field (49-51) was used for the protein, and CHARMM36 (49-51) force field was used to model lipids. TIP3P was used to simulate water molecules. The force field for anions was the same as the previous study (28). Following previous pK_a calculations performed on the same crystal structure (52) and our previous investigation of proton transport from E113 to E203 (32), E203 and D417 were protonated, while all other residues were in their standard protonation states.

The systems were equilibrated with the NAMD 2.10 software (53). The interval of calculating bonded interactions and short-range nonbonded interactions was 2 fs. The particle mesh Ewald (PME) method (54) was adopted to calculate long-range electrostatic interactions every 4 fs, with a spherical cutoff of 12.0 Å and an accuracy threshold of 10^{-6}. The neighbor list consisting of atom pairs whose interactions were to be evaluated was updated every 20 fs. A switching function was used for van der Waals and short-range electrostatic interactions in the range of 10.0 Å and 12.0 Å. The cutoff of neighbor list was set to 13.5 Å. Langevin dynamics

and the Nosé−Hoover Langevin piston method (55-57) were employed for constant pressure (P = 1 atm) and constant temperature (T = 310 K) (NPT) molecular dynamics simulations.

After ~500 ns of equilibration, the proton on one of the two E203 residues was treated as the excess proton and multiscale reactive molecular dynamics (MS-RMD) (43, 44) simulations were performed. The excess proton and water molecules were described by the MS-EVB 3.2 model (58). Parameters for protonatable residues E148 and E203 (the proton donor/acceptor) were determined with the FitRMD parametrization procedure (see below) (45, 46). All other initial system setup and simulation details were described in our previous work (32).

The weighted histogram analysis method (WHAM) (59, 60) was utilized to combine the potential of mean forces (PMFs) generated from the separate windows of umbrella sampling simulations. The error bars on the PMF calculations were estimated using block averaging by dividing each trajectory into consecutive blocks. PMFs were processed with the help of GNU Parallel (61), and plotted with Matplotlib (62). All PMF energies were zeroed in reference to the minimum free energy. Representative configurations of the trajectories and the crystal structure were rendered in VMD (63).

2.2.3 Simulations of the Anion Motions

The relative stability of the central anions in each binding mode was investigated by calculating the PMF for the rotation or flipping of the anion as described by a single continuous collective variable (CV). In the nitrate-bound system, the CV definition was inspired by the hydrogen-bond interaction between the residues and the anion (28). Among the candidates, the selected CV, smallest distance between the nitrate oxygens and the oxygen on the side chain of S107 residue, shows distinct bimodal distribution within the classical simulations. For the

thiocyanate-bound system, we first defined a reference point located slightly above the anion. The midpoint of the α-carbons of residues R147 and A358 was chosen for stability. Then the anion's flipping motion was characterized by the difference in two distances, from the reference point to either sulfur atom or the nitrogen atom of the thiocyanate at S_{cen}. With the CVs defined above, umbrella sampling for ~10 ns was performed to construct the PMFs.

2.2.4 Parametrization for E148 and E203

In this study, the MS-RMD models for E148 and E203 in ClC-ec1 were parametrized. The parameter set (46) for glutamic residues in the ClC-ec1 protein with Cl^- present in S_{cen} was used as the initial guess. To sample the configurations of deprotonation of glutamic residues, umbrella sampling simulations were performed with the collective variable defined as the shortest distance from center of excess charge (CEC) to the oxygen atoms on the side chain of Glu, where CEC ($\mathbf{r}_{\mathrm{CEC}}$) is defined in section **1.2 Multiscale Reactive Molecular Dynamics**. The umbrella windows ranged from 1.00 to 3.00 Å with a 0.25 Å interval between neighboring windows. The force constant of the harmonic potential was set to be 20 kcal mol^{-1} $Å^{-2}$. Configurations were sampled every ~10 ps, and approximately 800−1000 configurations for fitting were collected along the CV axis. For each configuration, a single point energy calculation was performed with the CP2K software (64, 65) to calculate the atomistic forces in the quantum mechanics (QM) region with the hybrid quantum mechanics/molecular mechanics (QM/MM) method.

In an extensive benchmark of 200 density functional theory (DFT) functionals (66), BLYP-D3 was recognized as the best GGA functional with chemical accuracy for biological and water systems. It was also shown that triple-zeta basis set is sufficient for calculations and that

quadruple-zeta basis set showed no significant improvement. Therefore, in the QM/MM single point energy calculations, the QM atoms were treated with DFT using the BLYP-D3/TZV2P functional and basis set (67-69), including the third generation of the dispersion correction developed by Grimme *et al.* (70, 71), the Gaussian and plane wave (GPW) scheme (72), and Goedecker−Teter−Hutter (GTH) pseudopotentials optimized for BLYP functional (73-75). Note that the diffuse basis set was not adopted, but previous research on the ClC protein (32, 46) and another transmembrane protein (42) involving PT provided results consistent with experimental measurements using the same basis set. The QM atoms consisted of the S_{cen} anion (NO_3^-/SCN^-), the side chains of residues S107, E148, E203, and Y445, and the water molecules within the proton transport channel or within the third solvation shell from E148 and E203. The QM box size was ~18–32 Å in each dimension to ensure that a ~6–8 Å buffer was included between the QM atoms and the boundary of the QM box. The Gaussian expansion of the electrostatic potential (GEEP) scheme was used to treat the QM/MM electrostatic coupling with periodic boundary conditions (PBCs) (76, 77), and the spurious QM/QM periodic image interactions were decoupled according to Blöchl's work (78). The chemical bonds that crossed the QM/MM boundary between the alpha carbon (MM) and the beta carbon (QM) of the residues were capped with hydrogen atoms, on which the forces were calculated following the IMOMM scheme with a scaling factor of 1.50 (79). The particle mesh Ewald method developed by Essmann *et al.* (80) and orbital transformation optimized by VandeVondele *et al.* (81) were also adopted. Fourier transform was performed utilizing FFTW3 (82). Otherwise, the QM setup was consistent with previous work (32, 46).

The MS-RMD parameters for E148 and E203 were optimized using a genetic algorithm (83) to minimize the residual χ^2 of squared differences between the forces calculated with QM/MM ($\mathbf{F}_{ij}^{\mathrm{ref}}$) and those calculated with MS-RMD ($\mathbf{F}_{ij}^{\mathrm{cal}}$):

$$\chi^2 = \frac{1}{3MN}\sum_{j=1}^{M}\sum_{i=1}^{N}\left|\mathbf{F}_{ij}^{\mathrm{cal}} - \mathbf{F}_{ij}^{\mathrm{ref}}\right|^2, \tag{2.1}$$

where the M and N are the number of configurations and the number atoms included in the fitting in each configuration, respectively. All atoms in the "reactive complex" (*i.e.*, the Glu residue and first three solvation shells), as determined by the MS-RMD state-selection algorithm (84), were included in the fitting. The models for E148 and E203 were developed independently

Table 2.1 The MS-RMD model parameters for E148 and E203 with a polyatomic anion (NO_3^-/SCN^-) present in S_{cen}. The definitions of the parameters are described in our previous work (45, 46, 84). Diagonal parameters for the bonding potential in the protonated Glu state were consistent with our previous publication (46). Off-diagonal parameters defined in our previous publication (84) were set to be zero unless explicitly shown in the table.

	NO_3^-		**SCN^-, state I**		**SCN^-, state II**	
	E148	E203	E148	E203	E148	E203
$\boldsymbol{B}$	0.310278	0.327409	0.070932	0.159072	0.246881	0.191776
$\boldsymbol{b}$	0.647138	0.077244	0.974834	1.313819	0.416484	0.923576
$\boldsymbol{b'}$	0.472166	0.497267	1.416915	1.263502	0.061358	0.115218
$\boldsymbol{d_{OO}^0}$ [a]	2.4	2.4	2.4	2.4	2.4	2.4
$\boldsymbol{C}$	0.430479	0.087287	0.471578	0.138039	0.208488	0.298328
$\boldsymbol{c}$	1.420285	0.208499	1.414776	1.088506	0.899271	1.928760
$\boldsymbol{d_{OH}^0}$ [a]	1.0	1.0	1.0	1.0	1.0	1.0
$\boldsymbol{r_s^l}$ [a]	3.5	3.5	3.5	3.5	3.5	3.5
$\boldsymbol{r_s^h}$ [a]	4.0	4.0	4.0	4.0	4.0	4.0
$\boldsymbol{V_{ii}}$	–142.1753	–152.1397	–154.5561	–149.4336	–146.4417	–155.2515
$\boldsymbol{V_{const}^{ij}}$	–27.15733	–25.66442	–20.51090	–24.35824	–25.71849	–22.49573
$\boldsymbol{\gamma}$	1.145386	2.291855	3.934933	2.525677	1.541663	3.040086
$\boldsymbol{\beta}$ [a]	1000	1000	1000	1000	1000	1000
$\boldsymbol{P'}$ [a]	1.0	1.0	1.0	1.0	1.0	1.0
$\boldsymbol{\alpha}$	0.000196	0.001431	0.002298	0.003671	0.002940	0.000328
$\boldsymbol{r_{OO}^0}$	1.981408	2.223412	3.376918	2.407584	3.193151	3.984573

a These parameters are fixed in the parametrization procedure, while others are optimized.

due to their separation of ~10−16 Å precluding any direct coupling. To capture the deviation of CEC from the midpoint of the two oxygen atoms, the off-diagonal functional characterizing the coupling between the protonated glutamate state and the hydronium state was modified to the original form for proton transport in water (84). The new MS-RMD parameters for E148 and E203 for each system are given in **Table 2.1**. Note that there are variations in parameters between the S_{cen} anions and the anion states, because the water environment and hence the PT behavior is sensitive to the nature of the bound anion. Similar variations were also reported in our previous study in the ClC-ec1 protein (32, 38, 46).

2.2.5 Simulations of the Proton Transport

Four two-dimensional (2D) PMFs (two for the SCN^- cases, one for NO_3^-, and one for WT) were calculated from umbrella sampling simulations. Each of them is described with two CVs, one (ξ_1) representing the extent to which center of excess charge transports, and the other one (ξ_2) indicating the presence of water molecules surrounding the transport channel (32).

The first CV is the ratio of the distance between the CEC and the starting point to the distance between the end points. The advantage of this CV is that it quantitatively captures the movement of the CEC and the relative distances between donor and acceptor simultaneously. Mathematically, the first CV is described as

$$\xi_1 = \frac{\min\limits_{i\in\{1,2\}} (\mathbf{r}_{\mathrm{CEC}} - \mathbf{r}_{\mathrm{E203},i})\cdot\hat{\mathbf{n}}_{\mathrm{PT}}}{\min\limits_{i,j\in\{1,2\}} (\mathbf{r}_{\mathrm{E148},j} - \mathbf{r}_{\mathrm{E203},i})\cdot\hat{\mathbf{n}}_{\mathrm{PT}}}, \tag{2.2}$$

where $\hat{\mathbf{n}}_{\mathrm{PT}}$ is a prescribed unit vector pointing from E203 to E148, $\mathbf{r}_{\mathrm{CEC}}$, $\mathbf{r}_{\mathrm{E203},i}$, and $\mathbf{r}_{\mathrm{E148},j}$ are the coordinates of the CEC, i-th oxygen of E203 side chain, and j-th oxygen of E148 side chain, respectively, with i and j enumerating from 1 to 2. The min function used in ξ_1 is a smoothed

approximation (defined in section **2.2.7 Smooth Minimum/Maximum Function**) of the mathematical minimum function. This CV, different from the curvilinear path CV used in our previous work (32), describes the system in a more systematic and complete manner. See section **2.3.3 Justification of the New CV** for more discussions.

The second CV (ξ_2) is the effective number of water molecules in a box, with the center of the box and the principal axis defined through our published procedure (32). It is described as

$$\xi_2 = \sum_{i=1}^{N_{\mathrm{H_2O}}} N_i, \tag{2.3}$$

where $N_{\mathrm{H_2O}}$ is total number of water molecules and N_i is the occupancy of the i-th water molecule in a defined rectangular box, calculated from

$$N_i = \prod_{\alpha}^{x,y,z} N_{i,\alpha}, \tag{2.4}$$

$$N_{i,\alpha} = \left(1 - R_{i,\alpha}^6\right)/\left(1 - R_{i,\alpha}^{12}\right), \tag{2.5}$$

in which the multiplication is performed over all Cartesian coordinates, and $R_{i,\alpha}$ is given by

$$R_{i,\alpha} = \max\left\{0{,}1 + \left(\left|r_{i,\alpha} - r_{0,\alpha}\right| - b'_\alpha\right)/d_\alpha\right\}, \tag{2.6}$$

where $r_{0,\alpha}$ and $r_{i,\alpha}$ are the α (either x, y, or z) coordinate of the box center, and the oxygen in i-th water molecule, respectively. Parameters b'_x, b'_y, and b'_z characterize the box size and d_α dictates how far $R_{i,\alpha}$ extends beyond the box. Note that the Cartesian coordinate system involved in previous equations is aligned with the water box, instead of the simulation box.

In this study, the z axis of the box was aligned with the vector connecting the center of mass of the carboxyl group of E148 to that of E203. For the thiocyanate simulations, the

distances d_x, d_y, and d_z were chosen to be 5.0 Å, 5.0 Å, and 6.7 Å, respectively. The parameter b'_z was set to be 6.0 Å. Parameters b'_x and b'_y were set to 3.2 Å for state I and 3.7 Å for state II. In the system for NO_3^- present in S_{cen}, the distances d_x, d_y, and d_z were all set to be 5.0 Å, and parameters b'_x, b'_y, and b'_z were set to 3.0 Å, 3.0 Å, and 5.0 Å, respectively. The boundary of the box was determined so that waters outside the central region are not counted.

2.2.6 QM/MM MD Simulation Details

QM/MM MD simulations were also performed with the CP2K software (64, 65). The equilibrated configurations in the MS-RMD simulations were used as the initial structure of the QM/MM simulations. Short equilibration with the CSVR thermostat (85) (time constant = 5) was performed to quickly stabilize the system temperature around the targeted value (300 K), which was followed by additional equilibration with Nosé–Hoover chain (55, 56) as the thermostat (chain length = 4). The total equilibration time was ~5 ps in total. The integration time step was set to be 0.5 fs, and the smallest distance from CEC, calculated from the geometric configurations (86), to the anion was collected every time step for ~30 ps in multiple independent trajectories. To restrain the QM atoms within the QM box, a quadratic confining wall potential was used with the wall skin thickness being 2 Å × 2 Å × 2 Å. A harmonic wall (force constant: 100 kcal mol^{-1} $Å^{-2}$) for CEC starting at the position of 3.6 Å away from the anion was added to make sure that the CEC is close to the anion. The distributed block-compressed sparse row (DBCSR) library (87) was utilized to perform sparse matrix calculation; and the time-reversible always stable predictor-corrector (ASPC) method (88) was adopted for

integration. All other setup was consistent with the aforementioned QM/MM single point energy calculations and previous QM/MM MD simulations for ClC-ec1 (32, 46).

2.2.7 Smooth Minimum/Maximum Function

The smooth minimum/maximum function involved in the CV definition is implemented in PLUMED 2 (48), and defined as

$$\min_i\{x_1, x_2, \dots, x_i, \dots, x_n\} \approx \beta/\log\sum_{i=1}^{n}\exp(\beta/x_i), \tag{2.7}$$

$$\max_i\{x_1, x_2, \dots, x_i, \dots, x_n\} \approx \beta\log\sum_{i=1}^{n}\exp(x_i/\beta), \tag{2.8}$$

where β is a positive number set as 250 Å for the configuration sampling for FitRMD and 500 Å for the nitrate motion PMF construction. For the first CV (ξ_1) of the 2D PMF of proton transport, the above built-in minimum function causes some numerical issues, and the definition is therefore modified to be

$$\min/\max_i\{x_1, x_2, \dots, x_i, \dots, x_n\} \approx \bar{x} + \beta\log\sum_{i=1}^{n}\exp[(x_i - \bar{x})/\beta], \tag{2.9}$$

where $\bar{x}$ is the arithmetic mean of $x_1, x_2, \dots, x_i, \dots, x_n$, and β is either a positive number for max function or a negative number for min function. $|\beta|$ was set as 0.025 Å for definition of CV1 (ξ_1). The modified min/max function was implemented in PLUMED 2 (48) by calling GNU libmatheval library.

2.2.8 Rate Constant Calculations

As is derived from the transition state theory (89, 90), the rate constant k was calculated from

$$k = \sqrt{\frac{k_{\mathrm{B}}T}{2\pi m_{\mathrm{eff}}}} \frac{\int_S \exp(-V(\mathbf{x})/k_{\mathrm{B}}T)\mathrm{d}\mathbf{x}}{\int_A \exp(-V(\mathbf{x})/k_{\mathrm{B}}T)\mathrm{d}\mathbf{x}}, \quad (2.10)$$

in which T represents the system temperature, and the effective mass m_{eff} is estimated from $m_{\mathrm{eff}}\langle v^2\rangle = k_{\mathrm{B}}T$ based on the equipartition theorem where the instantaneous velocity v of the excess proton CEC motion is calculated from the finite-difference method. In the equation above, $V(\mathbf{x})$ stands for system potential energy, and $\mathbf{x}$ is the full set of system variables. The probability (Boltzmann factor) is integrated over the dividing surface S in the numerator and over the reactive basin A in the denominator. The characteristic time was calculated as the inverse of rate constant.

2.2.9 Water-Wire Analysis

A continuous water-wire is defined as a water network starting from E203 to E148, connected through hydrogen bonds, determined by two criteria: (1) the distance between the donor and acceptor heavy atoms is shorter than 3.0 Å; and (2) the angle formed by the donor atom, the central hydrogen, and the acceptor atom is larger than 150 degrees. The breadth-first algorithm was used to search for any possible continuous water network. In the case of non-connecting water-wire, the searching algorithm was performed twice, one from E203 to E148 and the other in the opposite direction, with maximum depth set as 20. The smallest distance between two groups of heavy atoms in the water networks starting from either E203 or E148 is

divided by the smallest distance between two groups of side-chain oxygen atoms of either E203 or E148, and plotted against the minimum free energy path.

2.3 Results and Discussion

2.3.1 Basicity of the Protein-Bound Anion

Thiocyanic acid, the conjugate acid of SCN^-, is a strong acid in water solution ($pK_a = -1.8$) (91), indicating that SCN^- does not readily react with H^+ directly in an aqueous environment. However, the environment of bulk water is quite different from that of a buried protein binding site, as demonstrated by the altered effective dielectric constant (91, 92), which could affect the basicity of the bound anion significantly. Thus, SCN^- could be explicitly protonated during PT in the protein. Similarly, NO_3^-, the conjugate base of a strong acid (HNO_3), might also react with H^+ in the cavity during PT from E203 to E148. If this is the case, we must include SCN^- into the MS-RMD parametrization via our iterative fitting procedure (45). In order to understand the possible change of the anion basicity influenced by the ClC-ec1 environment, extensive QM/MM MD simulations were performed. The excess proton CEC was restrained in the vicinity of the anion in the QM/MM simulations to avoid having the excess charge escape from the central anion due to unfavorable starting configurations.

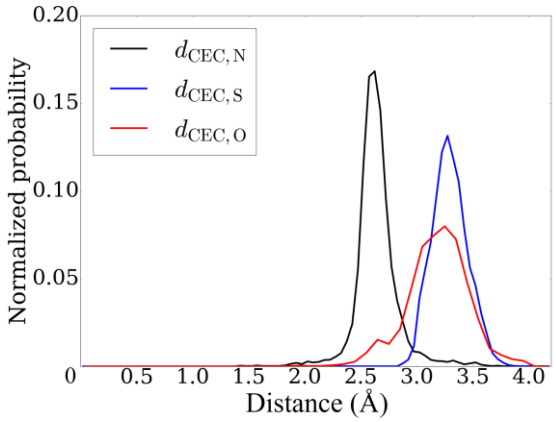

Figure 2.2 The normalized probability of observing a CEC around the anion
Black and blue lines are calculated from the thiocyanate-bound protein, whereas the red line is from the nitrate-bound system. The horizontal axis stands for the distance in Å between CEC and the corresponding atom of the anion, and the vertical axis indicates the normalized probability. The bin size of the histograms is 0.05 Å.

Our QM/MM simulations showed no configurations of proton-bound thiocyanate or nitrate, as is indicated by the disappearing probability at short distances (< 2 Å) between the CEC and the anion (**Figure 2.2**). In the thiocyanate case, the most probable distance is 2.6 Å for the N side of SCN^-, and 3.3 Å for the S side, the former of which corresponds to the length of a typical X···H–O hydrogen bond where X is an electronegative period 2 element and O is the oxygen of a hydronium (93). The latter indicates a weak hydrogen bond of S···H–O, the length of which is typically 3.2–3.5 Å (94, 95). Therefore, the thiocyanate anion does not react with the hydronium, though they are able to form hydrogen bonds. For the nitrate-bound system, the distance shows a bimodal distribution, peaked at 2.6 Å and 3.3 Å. Noticeably, the 2.6-Å peak attributed to the aforementioned hydrogen bond (93) is much lower than the other peak, implying that the nitrate anion interacts with the proton even more weakly. To eliminate the effect of initial configurations, simulations starting from a protonated configuration were also performed. They ended up with the deprotonation of the anion. Therefore, the polyatomic anions were

considered to not react with the PT process directly (chemically) and were regarded as a nonreactive species within the MS-RMD parametrization.

2.3.2 Anion Binding Modes

Since we observed that the anions stayed stable in S_{cen} for ~500 ns of MD simulations, and that they bound to the protein in two different states (**Figure 2.1** (B) and (C)), we next wanted to assess whether or not the two anion binding modes needed to be explicitly sampled in our PT analysis. Thus, we analyzed the relative stability and rate of transitioning between the two modes with MS-RMD umbrella sampling simulations including the explicit H^+ on E203. The potentials of mean force (PMFs, *i.e.*, free energy profiles) for the motion of anions (**Figure 2.3**) confirmed two distinct binding modes for both anions and revealed a larger transition barrier for thiocyanate. Based on **Figure 2.1** (B) and (C), we defined state I as the left minima and state II as the right. In the nitrate case, the anion either coordinated with S107/Y445 (state I) or not (state II). For the SCN^--bound protein, either the S-side (state I) or the N-side (state II) of the anion interacted with G149/F357, while the other side coordinated the lower part of the cavity (S107/Y445). Contrary to previous results reported (28), we found that state II for both nitrate and thiocyanate was thermodynamically favorable. This difference is likely due to the different force fields and methods employed. Jiang *et al.* (28) used a non-reactive TIP3P water model with no enhanced sampling, whereas we performed reactive MD with an excess proton on E203 and the SPC/Fw water model (96) as well as umbrella sampling (see section **2.2 Methods**). Given the sensitivity of water dynamics to the underlying model, the water environment around the anion, and hence the relative enthalpic and entropic stability, could also be affected by the water

force fields employed. Observing that the nitrate should experience less steric hindrance after the S107 side chain rotates down from state I to state II, we suggest that the unexpected relative stability of nitrate state II is due to entropic freedom in a larger free space, compensating for the energetic loss of coordination with the S107 and Y445 side chains. The free energy barrier and rate constants (**Table 2.2**) extracted from the PMFs indicate that the rotation of nitrate is relatively fast (picosecond time scale) and able to be sampled within the production runs (nanosecond time scale), whereas the flipping of the bulkier thiocyanate anion is slower (1.1 ns) and will not be sampled sufficiently. The fast state transition of NO_3^- is due to its relatively small size as compared to the protein cavity, and the relative weakness of both steric and coordination effects. In comparison, the slow state transition of SCN^- is due to its linear and more bulky structure. The surrounding residues squeeze SCN^-, disfavoring intermediate configurations and slowing the flipping motion of SCN^-. Based on these results, we decided to treat the two thiocyanate binding modes separately in the PT analysis, whereas the faster transitions for nitrate were deemed sufficient to be captured in PT simulations without distinguishing or explicitly sampling the two binding modes.

Table 2.2 Information extracted from the PMFs for anion motion with free energy barriers, $\Delta F_{\mathrm{I}}^{\ddagger}$ and $\Delta F_{\mathrm{II}}^{\ddagger}$, for the system to escape from the state I and II, respectively, and the characteristic time t defined as $1/k$ in which k is the TST rate constant for the system to escape from the most stable basin.

System	$\Delta F_{\mathrm{I}}^{\ddagger}$ (kcal/mol)	$\Delta F_{\mathrm{II}}^{\ddagger}$ (kcal/mol)	t (ps)
NO_3^-	1.3 ± 0.2	1.9 ± 0.3	3.7 ± 1.7
SCN^-	3.6 ± 0.2	5.1 ± 0.3	$(1.1 \pm 0.6) \times 10^3$

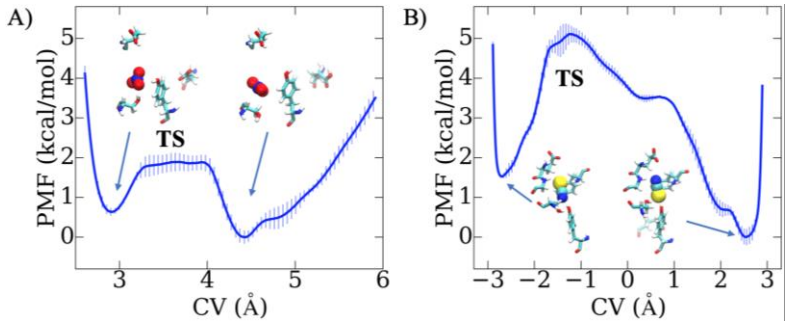

Figure 2.3 PMFs with error bars for the two binding modes of nitrate (A) and thiocyanate (B), within the central cavity of ClC-ec1. "TS" is short for transition state. Screenshots of the representative configurations of the metastable states, *i.e.*, local minima, are depicted. The CV for the PMF is defined as: (A) the smallest distance between the nitrate oxygens and the oxygen on the side chain of S107 residue; (B) $d_{\mathrm{S,ref}} - d_{\mathrm{N,ref}}$, the difference in the two distances from "ref" to either "S" or "N", in which the subscripts "S" and "N" stand for the sulfur atom and the nitrogen atom of the thiocyanate anion in the PT pathway, respectively, and the "ref" represents the midpoint of the α-carbons of the residues R147 and A358.

2.3.3 Justification of the New CV

Our previous investigation (32) into the central region of WT ClC-ec1 indicated that the whole PT process involves three variables: (1) the distance between donor and acceptor (*x*-axis of **Figure 2.4** (B)); (2) the progress of proton transport (*y*-axis of **Figure 2.4** (B) and *x*-axis of **Figure 2.4** (C)); and (3) the water density within the region (*y*-axis of **Figure 2.4** (C)). It would then seem that an accurate and complete inclusion of the three coupled CVs involved in the PT process requires a full 3D free energy sampling, which is computationally infeasible. However, we observed that the change in donor-acceptor distance and the CEC, which is used to track the excess proton, is stepwise (**Figure 2.4** (B)) (32), providing us an opportunity to incorporate them into a single CV. By including them into a ratio, the improved CV (ξ_1, Equation (2.2)), characterizing the extent to which proton is transported, captures both the absolute coordinate of CEC and the relative distance between the donor and the acceptor, removing one dimension of

required sampling. However, it has been suggested that a shift in CV may possibly alter the free energy barrier artificially (97). To exclude this possibility, we therefore ran 2D umbrella sampling simulations for the WT Cl^--bound ClC-ec1 with our new CV to explore consistency between our new PMF and those sampled with the previous CVs (32, 97).

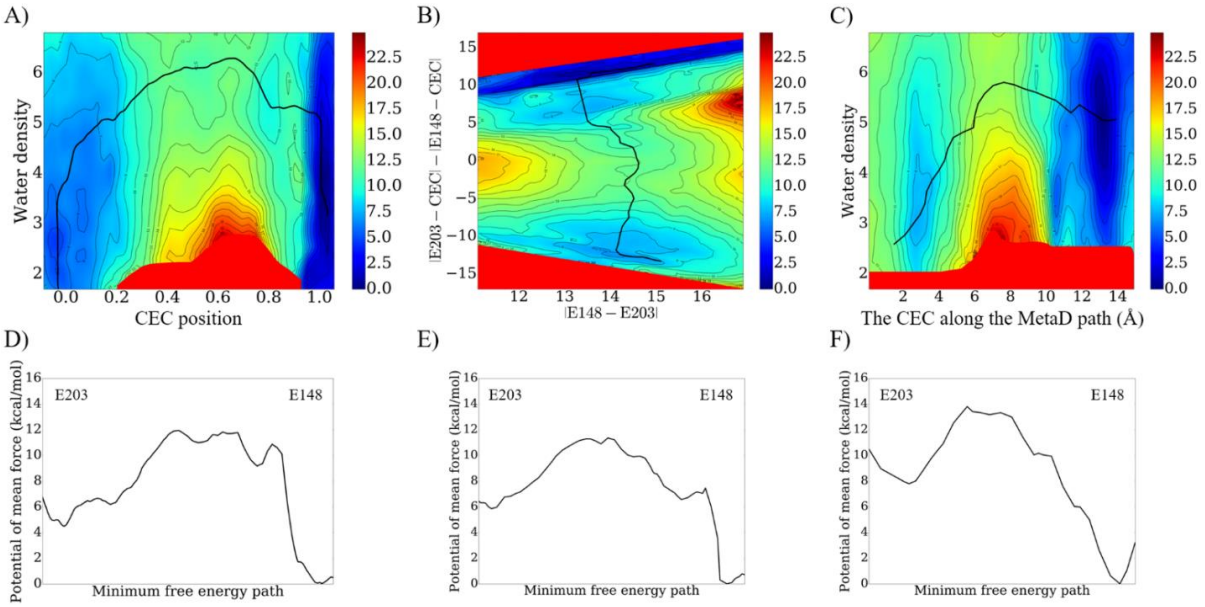

Figure 2.4 Comparison of PMFs
(A) The 2D PT PMF of WT ClC-ec1, with Cl^-_{cen} present, calculated with the new definition of CEC position in Equation (2.2), and (B and C) along with previously published PMFs of the same system for comparison. For (A), the CEC CV (the horizontal axis) is defined in Equation (2.2), and the water density CV (the vertical axis) characterizes the number of water molecules in the central region (41). For (B), the horizontal axis is the distance between E148 and E203, and the vertical axis is the difference in the two distances from the CEC to either E148 or E203. For (C), the *y*-axis is the same as (A), whereas the *x*-axis is the distance of CEC along the proton pathway sampled by the MetaD, as described in the previous paper (32). The line on each 2D PMF is the minimum free energy path for the proton transfer process from E203 to E148 (D, E, and F). The extracted 1D PMFs along the minimum free energy path, from the corresponding 2D PMFs in (A, B and C). The error of the PMFs is ~0.3 kcal/mol (D) or ~0.8 kcal/mol (E, F), estimated from splitting the trajectories and block-averaging.

The traced 1D minimum of the 2D PMF shows good agreement with the previous PMF using relative distance as a CV (compare **Figure 2.4** (D) and (E)), thus validating that the new CV reflects the movement of the CEC and the relative distance between donor and acceptor

simultaneously; while the 2D PMFs verify that the influence of hydration is consistently captured in the second dimension (compare **Figure 2.4** (A) and (C)). Note that the minimum free energy paths of **Figure 2.4** (A) and **Figure 2.4** (B) are incomparable because CVs in the latter PMF are incorporated into a single CV in the former PMF. Thus, the change in CV does not artificially affect the PT barrier, but describes the system in a more systematic and complete manner.

2.3.4 Proton Transport Simulations

Our initial studies described above determined the need for three simulations of PT: two for SCN^- and one for nitrate. Preliminary one-dimensional (1D) sampling of the PT across the central cavity showed a significant drop of the hydration level at the state of protonated glutamic acid, implying the coupling between PT and the water environment. Hence, 2D umbrella sampling simulations were performed to construct PT PMFs. The PMFs and calculated rate constants are summarized in **Figure 2.5** and **Table 2.3**, respectively. The increased free energy barrier and calculated rate constants agree with the experimental results that PT is partly uncoupled to the nitrate transport and completed blocked by the thiocyanate (15). Combining with the experimental measurement where the ratio of anion transport rates of Cl^-:SCN^-:NO_3^- is approximately 3:3:1 (15) and the Cl^- turnover rate of 2.3 ms^{-1} (23, 24), the shift in PT rate is qualitatively consistent with the stoichiometry change from 2.2:1 for Cl^-:H^+ flow (24), to 7–10:1 for NO_3^-:H^+, and to almost infinity for SCN^-:H^+ (15). The consistency justifies our computational results and allows for additional analysis based on these 2D PMFs.

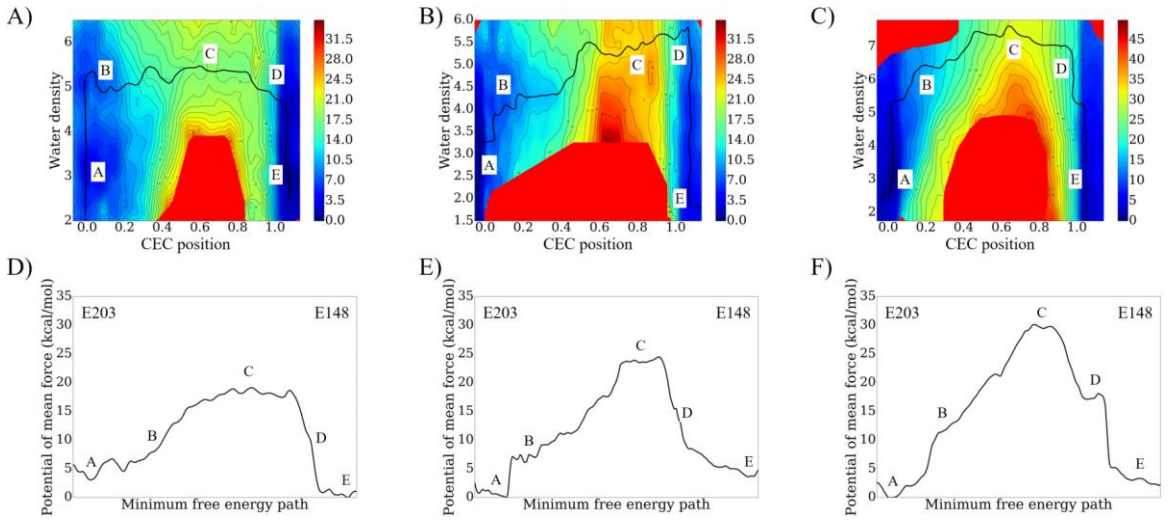

Figure 2.5 Proton transport PMFs
2D PT PMFs for nitrate (A) and thiocyanate in binding mode I (B) and II (C) tracking the improved definition for the ratio-based and unitless CEC position (horizontal axis) as defined in Equation (2.2) and water density (vertical axis) defined in the SI. Irrelevant high-energy areas (red) are not sampled to reduce the cost of computation. The black lines trace the minimum free energy paths for PT from E203 to E148. (D, E, and F) The extracted 1D PMFs, along the minimum free energy path, from the corresponding 2D PMFs above. The error of the PMFs is ~0.2 kcal/mol for the nitrate case and ~1 kcal/mol for the thiocyanate case, estimated from splitting the trajectories and block-averaging. Key points are labeled in each PMF denoting: A) protonated E203; B) the proton in a hydrated channel just above E203; C) the transition state; D) the proton passing through a connected water network to E148; and E) protonated E148 and dehydration of the channel. Note the much larger energy scale of the vertical axis in panels (D, E, and F) versus that of the same lettered panels of **Figure 2.4** for the WT protein.

To understand the PT mechanism, 2D PMFs as well as 1D traces along the minimum free energy paths are inspected. According to the 2D PMFs, the PT process can be categorized into three phases: (1) hydration (path A to B in PMFs), (2) proton transport (path B to D in PMFs), and (3) dehydration (path D to E in PMFs). In the hydration phase, the hydrophobic cavity is solvated by 3-5 water molecules, as shown by the increase in water density CV along the initial part of the minimum free energy paths. For both nitrate and thiocyanate, the initial deprotonation of E203 is facile once the hydration level reaches 3-3.5. However, for nitrate the cavity can easily fill to ~5 waters before or during E203 deprotonation. In contrast, it is harder to increase the solvation to ~5 or more water molecules in the thiocyanate system, requiring more free

energy and coupling to PT (meaning the excess proton transitioning to the water above E203 facilitates more water filling the cavity). In the proton transport phase, the excess proton progresses toward E148. This is an uphill process in free energy, peaking as the proton passes the anion in each case. For thiocyanate, it is coupled to increasing hydration. Finally, E148 is protonated and the cavity returns to the low hydration state due to the absence of an excess proton. These results are consistent with our preliminary 1D sampling.

Based on several prior computational results, it was expected that the hydration phase should contribute to the proton transport free energy barrier most significantly (28), and that the proton transport should couple with the hydration (32, 41). Surprisingly, these two CVs do not couple clearly in terms of the minimum free energy path, especially for the nitrate system (**Figure 2.5** (A)), and the hydration phase does not induce much of a barrier (less than 5 kcal/mol in the nitrate case). This is attributable to the fact that the water density CV ξ_2 introduced in our previous paper (41) simply reflects the water count of a specified region, but not necessarily the water wire connectivity with the excess proton in it. In the hydration phase, water molecules cluster below the anion, which is relatively easy due to the proximal ionizable residue (E203). Importantly, cavity hydration is not necessarily always equal to water wire connectivity.

Table 2.3 Summary of PT PMFs with different anions

System	$\Delta F^{\ddagger}$ (kcal/mol)	k (s^{-1})	k_{ref} (s^{-1})
Cl^-	7.4 ± 0.3	$(7 \pm 3) \times 10^7$	4×10^8
NO_3^-	14.6 ± 0.2	14 ± 5	$\sim 7 \times 10^1$
SCN^-, I	21 ± 1	$(2 \pm 3) \times 10^{-4}$	< 1
SCN^-, II	28 ± 1	$(1 \pm 2) \times 10^{-7}$	< 1

a $\Delta F^{\ddagger}$ is the energy barrier for the proton transport between E203 and E148, from the higher-energy state (E203 for Cl^-/ NO_3^- and E148 for SCN^-) to the lower-energy state (E148 for Cl^-/ NO_3^- and E203 for SCN^-), and k stands for the rate constant calculated from transition state theory. The experimentally measured anion rate constant combined with the anion/proton stoichiometry is used to calculate the reference k_{ref} for NO_3^-, and k_{ref} for SCN^- is based on the detection limit of ion transport (15, 23, 24). Since PT in the central region is not rate-limiting in ClC-ec1, our previous result for PT from E203 to E148 is listed as the reference k_{ref} for Cl^- (32).

In order to better understand the origin of the dominant contribution to the PT barrier, we analyzed the correlation between free energy and water connectivity, by plotting the extracted 1D PMF along with either the probability of forming a continuous water-wire (**Figure 2.6** (A)-(C)), or the ratio of the length of the disconnected area to the distance between two glutamates, E203 and E148 (**Figure 2.6** (D)-(F)). These plots indicate a strong correlation between the free energy barrier and the formation of a continuous water-wire, as is highlighted by two facts: (1) the water network is only connected around the dividing surface where the derivative of the free energy over the path is near zero; (2) the free energy and the ratio of disconnected distance to total distance is inversely related. This explains the unexpected free energy barrier in the PT phase and recaptures the correlation between PT and water dynamics. Thus, it is the water connectivity along the PT pathway in the presence of the explicit excess proton that is most relevant to the PT barrier. Interestingly, the probability of forming a continuous water-wire in the state II of thiocyanate system (**Figure 2.6** (C)) is extremely low, indicating that the water-wire actually breaks on one side when it is forming on the other side. According to the correlation between water-wire formation and PT barrier, this result is consistent with the highest PT free energy barrier in the system.

2.3.5 Water-Network Analysis

For each of the polyatomic anion PMFs, the free energy for PT peaks (at the transition state) as the proton passes the bulky anion. This was not the case for Cl^- (see **Figure 2.4**). At the initial stage of the PT phase where free energy escalates, the water network is incomplete (**Figure 2.7** (A) and (D)). Connecting the water network to form a so called “water-wire” past the bulky anions is strongly coupled to, and enabling of, PT past the anions to E148. Hence, the

free energy ceases to increase when the water network around the bound anion is connected (**Figure 2.7** (B) and (E)). The position of center of excess charge (CEC, defined in Equation (2.2)) is exactly next to the anion at this stage. Finally, the system is stabilized after the proton passes the anion, binds to E148, and the water-wire is again disconnected (**Figure 2.7** (C) and (F)). For the entire PT process, only when the excess proton is near the anion does a continuous water-network (*i.e.*, continuous near the central anion) become significantly observable, which is near the dividing surface. Similar proton-induced wetting has been observed in other biological systems and materials (41, 42, 98). The distinct state change of the water-wire near the dividing surface confirms that the water molecules close to the polyatomic anion are energetically unfavorable. Thus, the coupled phenomena of PT and hydration, as well as the challenge of moving a proton through a confined space with limited delocalization, significantly increases the free energy barriers for PT between E203 and E148. Based on this analysis, the dynamic process of the PT phase can be more accurately summarized as: (1) proton transport from E203 to below the anion along with the elongation of the water-network but no connectivity past the anion (state B); (2) the water network connecting past the polyatomic anion, in the presence of a proximal excess proton, enabling PT past the anion (state C); and (3) breaking of the water network with the loss of the excess proton as it binds E148 (state D). The dividing surface lies in the second stage.

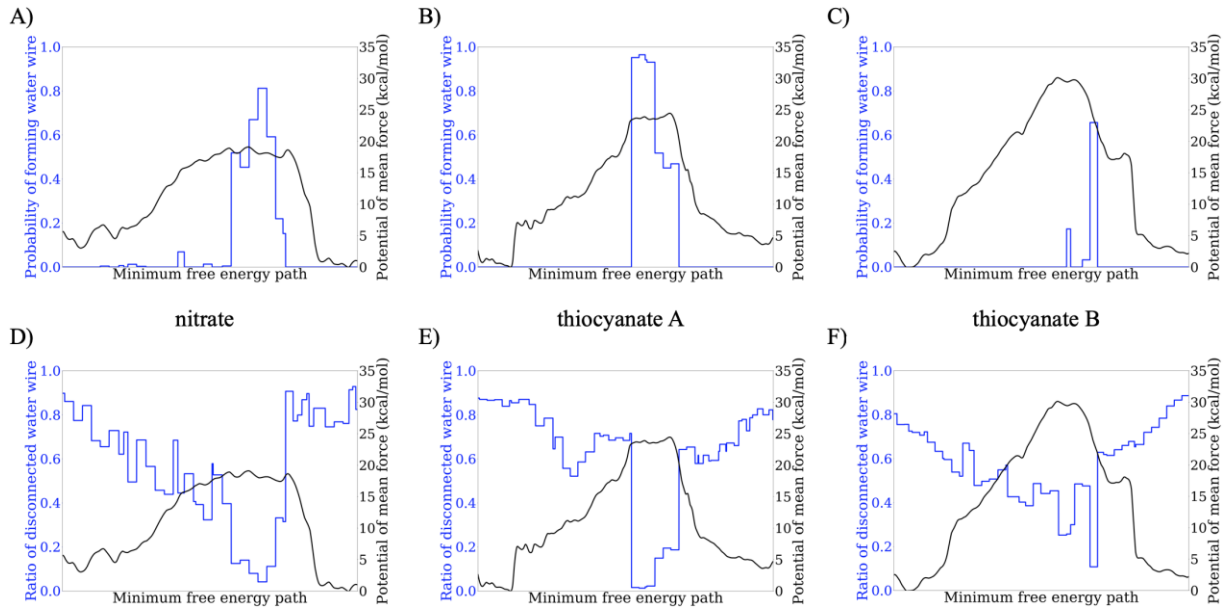

Figure 2.6 Water-network analysis
The overlap of extracted 1D PMF (black) and either the probability of forming a continuous water-wire (A, B and C) or the ratio of the length of the disconnected area to the distance between two glutamates, E203 and E148 (D, E and F) (blue), plotted against the minimum free energy pathway. They are plotted for the nitrate-bound protein (A, D), or the thiocyanate-bound protein in state I (B, E) or state II (C, F).

To further elucidate the molecular basis of the PT/water-induced barrier, we analyzed the anion structure close to the free energy barrier and found significant steric hindrance by the polyatomic anions. Unlike monatomic Cl^-, polyatomic anions like NO_3^- and SCN^- occupy more space within the central region. NO_3^- has a trigonal planar geometry, with the negative charge distributed over three O atoms. The effective radius of NO_3^- is ~1.97 Å, which is larger than that of Cl^- (1.64–1.82 Å) (99, 100). A slightly larger spatial occupation of nitrate anion is enough to slow PT across the central region, but insufficient to completely block the H^+ transport. In contrast, SCN^- has a linear structure which spans 4.77 Å in length (101), making it much bulkier than either Cl^- or NO_3^-. The increased size of SCN^- increases the free energy necessary to connect the water network and transport the excess proton through the transient water chain. The

influence of steric hindrance is consistent with our finding that the spatial effect slightly outcompetes the coordination effect in the thiocyanate motion PMF (**Figure 2.3**).

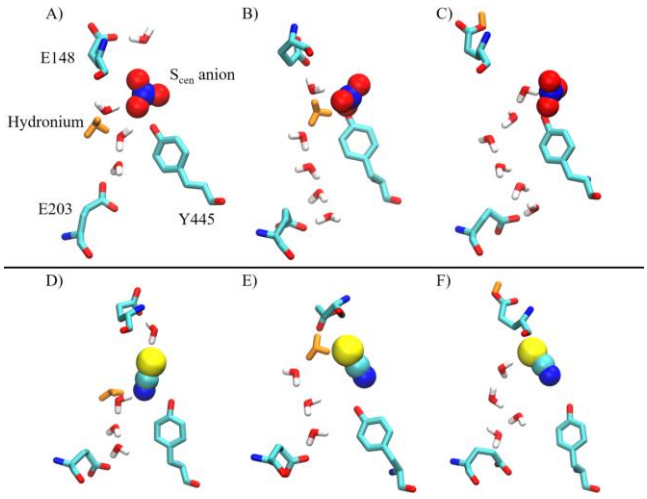

Figure 2.7 Representative configurations of the nitrate-bound (A, B, and C) and thiocyanate-bound (D, E and F) systems in the initial stage of the PT phase (A, D), connected water network (B, E), and final stage of the PT phase (C, F). The S_{cen} anion is shown in VDW representation, and the hydronium/protonated glutamic acid is rendered orange for clarity. Key residues (labeled in panel A) and water molecules are colored according to element.

2.4 Conclusions

We have performed extensive multiscale reactive MD simulations combined with free energy sampling to investigate the influence of polyatomic anions on PT and the related mechanisms of ion coupling and stoichiometric ion exchange in ClC-ec1. In our simulations, PT was treated explicitly, including Grotthuss shuttling and proton charge defect delocalization. We first quantified the relative stability and transition rates between two binding modes of the nitrate and thiocyanate anions in the central binding site of ClC-ec1, S_{cen}. We found that the anions are stable for at least the submicrosecond time scale in the central site. Our PMF calculations reveal that separate treatment of the two thiocyanate binding modes was necessary due to the slow interconversion. In contrast, nitrate rotates relatively quickly, making it possible to sufficiently

sample this transition during a single PT PMF calculation. This difference between nitrate and thiocyanate can be explained by the increased steric restriction of the bulky thiocyanate. We then defined and justified a new reaction coordinate to characterize PT between E203 and E148, which incorporates the excess proton CEC and the distance between the two glutamate residues simultaneously. This CV significantly reduces the computational demands. It not only captures two crucial factors of PT in a single variable, but also introduces no new spurious effects on the PMF based on comparisons to previously published PMFs (32) for PT for Cl^--bound ClC-ec1. Next, we used QM/MM simulations to verify that the polyatomic anions do not react with proton directly during the PT process in the protein, and hence do not require explicit MS-RMD treatment.

After addressing the challenges of studying the NO_3^-/SCN^--bound ClC-ec1, we constructed three 2D PMFs for the PT across the central region, one for NO_3^- and two for SCN^-. The rate constants calculated from the PMFs utilizing transition state theory are consistent with the experimental measurements, in which PT is partly uncoupled to the nitrate transport and completely uncoupled to thiocyanate transport. A previously published hypothesis about the water structure being influenced by the polyatomic anions was confirmed but significantly expanded to a three-phase mechanism of proton shuttling, on the basis of quantitative PT free energy calculations. Interestingly, these PMFs revealed that simply increasing the number of water molecules in the region of PT is not sufficient to enable PT. The critical hydration CV seems to be water network connectivity, which we demonstrate by analyzing the probability of water connectivity and the ratio of the length of the disconnected area to the distance between two glutamates. Both of these measures are highly coupled with PT through the transition state, as indicated by the correlation with the 1D PMFs for all three systems (one for NO_3^- and two for

SCN^-). A continuous water-wire with an explicit excess proton in it is in fact able to exist in the NO_3^-/SCN^--bound protein, albeit transiently and conditionally, contrary to the hypothesis arising from a prior standard empirical MD simulation in which the explicit PT was not treated (28).

In-depth analysis on the PT phase reveals three stages for the process. First, there is an increase in hydration weakly coupled to PT to just below the anion binding site. Second, the transition state is defined by strongly coupled connection of the water network and PT past the anion. The cause of the free energy barrier, and thus decreased proton coupling, is the steric hindrance and unfavorable hydrogen bond connectivity induced by NO_3^- and SCN^-. For the smaller polyatomic anion $NO_3^-$the PT free energy barrier is raised enough to make PT through central cavity clearly rate-limiting and slower than it is with Cl^-, whereas the bulkier SCN^- completely inhibits PT. Third, the water network breaks as E148 is protonated. This mechanism demonstrates that it is not primarily the hydration of the cavity that is rate limiting, but rather forming a connected hydrogen bonded pathway past the steric hindrance of the polyatomic anions. This provides insightful contrast to the common assumption that blocking water influx pathways is an effective way to inhibit PT (or an explanation of inhibited PT). For this system, and we anticipate for many others, water influx is a relatively fast and easy process compared to water connectivity through narrow regions. The latter was in fact recently found to be the cause of the rate limiting step for PT in cytochrome *c* oxidase (42, 102).

Collectively this work shows how seemingly subtle changes in the chemical nature of the anion can flip it from facilitating coupled proton exchange, to inhibiting or even blocking it in both directions. The shift in coupling changes the associated ion exchange stoichiometries, instead of completely blocking anion transport, because the rate and ease (*i.e.*, pK_a) of protonating E148, which is essential for anion flux, from the external bulk is fast and minimally

influenced by the presence of the anions (33, 38). Rather, PT is hindered/blocked in the central cavity between E148 and E203. The demonstrated entanglement between PT, water dynamics, and the local environment is expected to play a significant role in a number of other proton-conducting proteins. Our insights also potentially suggest guidelines for one way of controlling the function of the ClC family of proteins by influencing hydration, either through tuning the spatial constriction of the central cavity via binding, or by adjusting the steric hindrance caused by the channel-forming residues. However, our conclusions for these permeant anions are not necessarily transferrable to the uncoupling of PT to fluoride-bound ClC, due to the impermeant nature (3, 103) and the strong electronegativity of F^-. Therefore, additional research on PT in fluoride-bound ClC proteins, including a ClC^F exporter (104), is needed and will be the focus of future efforts. Additionally, we hope to identify new hydration CVs that directly track water connectivity and to investigate the S_{int}-bound SCN^- system, as well as the complete anion/proton transport cycle employing multiscale kinetic modeling (MKM) (33).

2.5 References

1. Lloyd SE, *et al.* (1996) A common molecular basis for three inherited kidney stone diseases. *Nature* 379(6564):445–449.

2. Simon DB, *et al.* (1997) Mutations in the chloride channel gene, *CLCNKB*, cause Bartter's syndrome type III. *Nat. Genet.* 17(2):171–178.

3. Maduke M, Pheasant DJ, & Miller C (1999) High-Level Expression, Functional Reconstitution, and Quaternary Structure of a Prokaryotic ClC-type Chloride Channel. *J. Gen. Physiol.* 114(5):713–722.

4. Bösl MR, *et al.* (2001) Male germ cells and photoreceptors, both dependent on close cell−cell interactions, degenerate upon ClC-2 Cl^- channel disruption. *EMBO J.* 20(6):1289−1299.

5. Iyer R, Iverson TM, Accardi A, & Miller C (2002) A biological role for prokaryotic ClC chloride channels. *Nature* 419(6908):715–718.

6. Lipicky RJ & Bryant SH (1966) Sodium, Potassium, and Chloride Fluxes in Intercostal Muscle from Normal Goats and Goats with Hereditary Myotonia. *J. Gen. Physiol.* 50(1):89–111.

7. Lipicky RJ, Bryant SH, & Salmon JH (1971) Cable Parameters, Sodium, Potassium, Chloride, and Water Content, and Potassium Efflux in Isolated External Intercostal Muscle of Normal Volunteers and Patients with Myotonia Congenita. *J. Clin. Invest.* 50(10):2091–2103.

8. Kornak U, *et al.* (2001) Loss of the ClC-7 Chloride Channel Leads to Osteopetrosis in Mice and Man. *Cell* 104(2):205–215.

9. Cleiren E, *et al.* (2001) Albers-Schönberg disease (autosomal dominant osteopetrosis, type II) results from mutations in the *CICN7* chloride channel gene. *Hum. Mol. Genet.* 10(25):2861–2867.

10. Accardi A & Miller C (2004) Secondary active transport mediated by a prokaryotic homologue of ClC Cl^- channels. *Nature* 427(6977):803–807.

11. Fahlke C, Dürr C, & George AL, Jr. (1997) Mechanism of Ion Permeation in Skeletal Muscle Chloride Channels. *J. Gen. Physiol.* 110(5):551–564.

12. Fahlke C, Yu HT, Beck CL, Rhodes TH, & George AL, Jr. (1997) Pore-forming segments in voltage-gated chloride channels. *Nature* 390(6659):529–532.

13. Rychkov GY, Pusch M, Roberts ML, Jentsch TJ, & Bretag AH (1998) Permeation and Block of the Skeletal Muscle Chloride Channel, ClC-1, by Foreign Anions. *J. Gen. Physiol.* 111(5):653–665.

14. Hebeisen S, *et al.* (2003) Anion Permeation in Human ClC-4 Channels. *Biophys. J.* 84(4):2306–2318.

15. Nguitragool W & Miller C (2006) Uncoupling of a CLC Cl^-/H^+ Exchange Transporter by Polyatomic Anions. *J. Mol. Biol.* 362(4):682–690.

16. Zdebik AA, *et al.* (2008) Determinants of Anion-Proton Coupling in Mammalian Endosomal CLC Proteins. *J. Biol. Chem.* 283(7):4219–4227.

17. Alekov AK & Fahlke C (2009) Channel-like slippage modes in the human anion/proton exchanger ClC-4. *J. Gen. Physiol.* 133(5):485–496.

18. De Stefano S, Pusch M, & Zifarelli G (2011) Extracellular Determinants of Anion Discrimination of the Cl^-/H^+ Antiporter Protein CLC-5. *J. Biol. Chem.* 286(51):44134–44144.

19. Accardi A (2015) Structure and gating of CLC channels and exchangers. *J. Physiol.* 593(18):4129–4138.

20. Dutzler R, Campbell EB, Cadene M, Chait BT, & MacKinnon R (2002) X-ray structure of a ClC chloride channel at 3.0 Å reveals the molecular basis of anion selectivity. *Nature* 415(6869):287–294.

21. Dutzler R, Campbell EB, & MacKinnon R (2003) Gating the Selectivity Filter in ClC Chloride Channels. *Science* 300(5616):108–112.

22. Walden M, *et al.* (2007) Uncoupling and Turnover in a Cl^-/H^+ Exchange Transporter. *J. Gen. Physiol.* 129(4):317–329.

23. Lim H-H, Shane T, & Miller C (2012) Intracellular Proton Access in a Cl^-/H^+ Antiporter. *PLoS Biol.* 10(12):e1001441.

24. Basilio D, Noack K, Picollo A, & Accardi A (2014) Conformational changes required for H^+/Cl^- exchange mediated by a CLC transporter. *Nat. Struct. Mol. Biol.* 21(5):456–463.

25. Nguitragool W & Miller C (2007) CLC Cl^-/H^+ transporters constrained by covalent cross-linking. *Proc. Natl. Acad. Sci. U. S. A.* 104(52):20659–20665.

26. Robertson JL, Kolmakova-Partensky L, & Miller C (2010) Design, function and structure of a monomeric ClC transporter. *Nature* 468(7325):844–847.

27. Orhan G, Fahlke C, & Alekov AK (2011) Anion- and Proton-Dependent Gating of ClC-4 Anion/Proton Transporter under Uncoupling Conditions. *Biophys. J.* 100(5):1233–1241.

28. Jiang T, Han W, Maduke M, & Tajkhorshid E (2016) Molecular Basis for Differential Anion Binding and Proton Coupling in the Cl^-/H^+ Exchanger ClC-ec1. *J. Am. Chem. Soc.* 138(9):3066–3075.

29. Accardi A, Lobet S, Williams C, Miller C, & Dutzler R (2006) Synergism Between Halide Binding and Proton Transport in a CLC-type Exchanger. *J. Mol. Biol.* 362(4):691–699.

30. Picollo A, Xu Y, Johner N, Bernèche S, & Accardi A (2012) Synergistic substrate binding determines the stoichiometry of transport of a prokaryotic H^+/Cl^- exchanger. *Nat. Struct. Mol. Biol.* 19(5):525–531.

31. Picollo A, Malvezzi M, Houtman JCD, & Accardi A (2009) Basis of substrate binding and conservation of selectivity in the CLC family of channels and transporters. *Nat. Struct. Mol. Biol.* 16(12):1294–1301.

32. Lee S, Swanson JMJ, & Voth GA (2016) Multiscale Simulations Reveal Key Aspects of the Proton Transport Mechanism in the ClC-ec1 Antiporter. *Biophys. J.* 110(6):1334–1345.

33. Mayes HB, Lee S, White AD, Voth GA, & Swanson JMJ (2018) Multiscale Kinetic Modeling Reveals an Ensemble of Cl^-/H^+ Exchange Pathways in ClC-ec1 Antiporter. *J. Am. Chem. Soc.* 140(5):1793–1804.

34. Accardi A, *et al.* (2005) Separate Ion Pathways in a Cl^-/H^+ Exchanger. *J. Gen. Physiol.* 126(6):563–570.

35. Lim H-H & Miller C (2009) Intracellular Proton-Transfer Mutants in a CLC Cl^-/H^+ Exchanger. *J. Gen. Physiol.* 133(2):131–138.

36. Wang D & Voth GA (2009) Proton Transport Pathway in the CIC Cl^-/H^+ Antiporter. *Biophys. J.* 97(1):121–131.

37. Zhang Y & Voth GA (2011) The Coupled Proton Transport in the ClC-ec1 Cl^-/H^+ Antiporter. *Biophys. J.* 101(10):L47–L49.

38. Lee S, Mayes HB, Swanson JMJ, & Voth GA (2016) The Origin of Coupled Chloride and Proton Transport in a Cl^-/H^+ Antiporter. *J. Am. Chem. Soc.* 138(45):14923–14930.

39. Ko YJ & Jo WH (2010) Secondary Water Pore Formation for Proton Transport in a CIC Exchanger Revealed by an Atomistic Molecular-Dynamics Simulation. *Biophys. J.* 98(10):2163–2169.

40. Han W, Cheng RC, Maduke MC, & Tajkhorshid E (2014) Water access points and hydration pathways in CLC H^+/Cl^- transporters. *Proc. Natl. Acad. Sci. U. S. A.* 111(5):1819–1824.

41. Peng Y, Swanson JMJ, Kang S-g, Zhou R, & Voth GA (2015) Hydrated Excess Protons Can Create Their Own Water Wires. *J. Phys. Chem. B* 119(29):9212–9218.

42. Liang R, Swanson JMJ, Peng Y, Wikström M, & Voth GA (2016) Multiscale simulations reveal key features of the proton-pumping mechanism in cytochrome *c* oxidase. *Proc. Natl. Acad. Sci. U. S. A.* 113(27):7420–7425.

43. Knight C, Lindberg GE, & Voth GA (2012) Multiscale reactive molecular dynamics. *J. Chem. Phys.* 137(22):22A525.

44. Yamashita T, Peng Y, Knight C, & Voth GA (2012) Computationally Efficient Multiconfigurational Reactive Molecular Dynamics. *J. Chem. Theory Comput.* 8(12):4863–4875.

45. Nelson JG, Peng Y, Silverstein DW, & Swanson JMJ (2014) Multiscale Reactive Molecular Dynamics for Absolute pK_a Predictions and Amino Acid Deprotonation. *J. Chem. Theory Comput.* 10(7):2729–2737.

46. Lee S, Liang R, Voth GA, & Swanson JMJ (2016) Computationally Efficient Multiscale Reactive Molecular Dynamics to Describe Amino Acid Deprotonation in Proteins. *J. Chem. Theory Comput.* 12(2):879–891.

47. Plimpton S (1995) Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* 117(1):1–19.

48. Tribello GA, Bonomi M, Branduardi D, Camilloni C, & Bussi G (2014) PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* 185(2):604–613.

49. MacKerell AD, Jr., Feig M, & Brooks CL, III. (2004) Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulations. *J. Comput. Chem.* 25(11):1400–1415.

50. Klauda JB, *et al.* (2010) Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *J. Phys. Chem. B* 114(23):7830–7843.

51. Best RB, *et al.* (2012) Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ, ψ and Side-Chain χ_1 and χ_2 Dihedral Angles. *J. Chem. Theory Comput.* 8(9):3257–3273.

52. Faraldo-Gómez JD & Roux B (2004) Electrostatics of Ion Stabilization in a ClC Chloride Channel Homologue from *Escherichia coli*. *J. Mol. Biol.* 339(4):981–1000.

53. Phillips JC, *et al.* (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26(16):1781–1802.

54. Darden T, York D, & Pedersen L (1993) Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* 98(12):10089–10092.

55. Nosé S (1984) A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* 52(2):255–268.

56. Nosé S (1984) A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.* 81(1):511–519.

57. Hoover WG (1985) Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* 31(3):1695–1697.

58. Biswas R, Tse Y-LS, Tokmakoff A, & Voth GA (2016) Role of Presolvation and Anharmonicity in Aqueous Phase Hydrated Proton Solvation and Transport. *J. Phys. Chem. B* 120(8):1793–1804.

59. Kumar S, Bouzida D, Swendsen RH, Kollman PA, & Rosenberg JM (1992) The Weighted Histogram Analysis Method for Free‐Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* 13(8):1011–1021.

60. Grossfield A ("WHAM: the weighted histogram analysis method", version 2.0.9.

61. Tange O (2011) GNU Parallel: The Command-Line Power Tool. *;login: The USENIX Magazine* 36(1):42–47.

62. Hunter JD (2007) Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* 9(3):90–95.

63. Humphrey W, Dalke A, & Schulten K (1996) VMD: Visual molecular dynamics. *J. Mol. Graphics* 14(1):33–38.

64. VandeVondele J, *et al.* (2005) Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Comput. Phys. Commun.* 167(2):103–128.

65. Hutter J, Iannuzzi M, Schiffmann F, & VandeVondele J (2014) cp2k: atomistic simulations of condensed matter systems. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* 4(1):15–25.

66. Mardirossian N & Head-Gordon M (2017) Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals. *Mol. Phys.* 115(19):2315–2372.

67. Lee C, Yang W, & Parr RG (1988) Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* 37(2):785–789.

68. Becke AD (1988) Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* 38(6):3098–3100.

69. Schäfer A, Huber C, & Ahlrichs R (1994) Fully optimized contracted Gaussian basis sets of triple zeta valence quality for atoms Li to Kr. *J. Chem. Phys.* 100(8):5829–5835.

70. Grimme S, Antony J, Ehrlich S, & Krieg H (2010) A consistent and accurate *ab initio* parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* 132(15):154104.

71. Grimme S, Ehrlich S, & Goerigk L (2011) Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* 32(7):1456–1465.

72. Lippert G, Hutter J, & Parrinello M (1997) A hybrid Gaussian and plane wave density functional scheme. *Mol. Phys.* 92(3):477–487.

73. Goedecker S, Teter M, & Hutter J (1996) Separable dual-space Gaussian pseudopotentials. *Phys. Rev. B* 54(3):1703–1710.

74. Hartwigsen C, Goedecker S, & Hutter J (1998) Relativistic separable dual-space Gaussian pseudopotentials from H to Rn. *Phys. Rev. B* 58(7):3641–3662.

75. Krack M (2005) Pseudopotentials for H to Kr optimized for gradient-corrected exchange-correlation functionals. *Theor. Chem. Acc.* 114(1):145–152.

76. Laino T, Mohamed F, Laio A, & Parrinello M (2005) An Efficient Real Space Multigrid QM/MM Electrostatic Coupling. *J. Chem. Theory Comput.* 1(6):1176–1184.

77. Laino T, Mohamed F, Laio A, & Parrinello M (2006) An Efficient Linear-Scaling Electrostatic Coupling for Treating Periodic Boundary Conditions in QM/MM Simulations. *J. Chem. Theory Comput.* 2(5):1370–1378.

78. Blöchl PE (1995) Electrostatic decoupling of periodic images of plane‐wave‐expanded densities and derived atomic point charges. *J. Chem. Phys.* 103(17):7422‐7428.

79. Maseras F & Morokuma K (1995) IMOMM: A New Integrated *ab initio* + Molecular Mechanics Geometry Optimization Scheme of Equilibrium Structures and Transition States. *J. Comput. Chem.* 16(9):1170–1179.

80. Essmann U, *et al.* (1995) A smooth particle mesh Ewald method. *J. Chem. Phys.* 103(19):8577–8593.

81. VandeVondele J & Hutter J (2003) An efficient orbital transformation method for electronic structure calculations. *J. Chem. Phys.* 118(10):4365–4369.

82. Frigo M & Johnson SG (2005) The Design and Implementation of FFTW3. *Proc. IEEE* 93(2):216–231.

83. Mühlenbein H & Schlierkamp-Voosen D (1993) Predictive Models for the Breeder Genetic Algorithm I. Continuous Parameter Optimization. *Evol. Comput.* 1(1):25–49.

84. Wu Y, Chen H, Wang F, Paesani F, & Voth GA (2008) An Improved Multistate Empirical Valence Bond Model for Aqueous Proton Solvation and Transport. *J. Phys. Chem. B* 112(2):467–482.

85. Bussi G, Donadio D, & Parrinello M (2007) Canonical sampling through velocity rescaling. *J. Chem. Phys.* 126(1):014101.

86. König PH, *et al.* (2006) Toward Theoretical Analyis of Long-Range Proton Transfer Kinetics in Biomolecular Pumps. *J. Phys. Chem. A* 110(2):548–563.

87. Borštnik U, VandeVondele J, Weber V, & Hutter J (2014) Sparse matrix multiplication: The distributed block-compressed sparse row library. *Parallel Computing* 40(5):47–58.

88. Kolafa J (2004) Time-Reversible Always Stable Predictor–Corrector Method for Molecular Dynamics of Polarizable Molecules. *J. Comput. Chem.* 25(3):335–342.

89. Vineyard GH (1957) FREQUENCY FACTORS AND ISOTOPE EFFECTS IN SOLID STATE RATE PROCESSES. *J. Phys. Chem. Solids* 3(1–2):121–127.

90. Chandler D (1987) *Introduction to Modern Statistical Mechanics*, (Oxford University Press, New York), pp 245–246.

91. Rumble JR ed (2018) *CRC Handbook of Chemistry and Physics* (CRC Press/Taylor & Francis, Boca Raton, FL), 98th Ed.

92. Li L, Li C, Zhang Z, & Alexov E (2013) On the Dielectric "Constant" of Proteins: Smooth Dielectric Function for Macromolecular Modeling and Its Implementation in DelPhi. *J. Chem. Theory Comput.* 9(4):2126–2136.

93. Botti A, Bruni F, Imberti S, Ricci MA, & Soper AK (2004) Ions in water: The microscopic structure of a concentrated HCl solution. *J. Chem. Phys.* 121(16):7840–7848.

94. Bhattacharyya S, Bhattacherjee A, Shirhatti PR, & Wategaonkar S (2013) O–H···S Hydrogen Bonds Conform to the Acid–Base Formalism. *J. Phys. Chem. A* 117(34):8238–8250.

95. Bhattacharyya S, Roy VP, & Wategaonkar S (2016) Acid–Base Formalism Extended to Excited State for O–H···S Hydrogen Bonding Interaction. *J. Phys. Chem. A* 120(35):6902–6916.

96. Wu Y, Tepper HL, & Voth GA (2006) Flexible simple point-charge water model with improved liquid-state properties. *J. Chem. Phys.* 124(2):024503.

97. Zhang Y & Voth GA (2011) Combined Metadynamics and Umbrella Sampling Method for the Calculation of Ion Permeation Free Energy Profiles. *J. Chem. Theory Comput.* 7(7):2277–2283.

98. Duster AW & Lin H (2017) Restrained Proton Indicator in Combined Quantum-Mechanics/Molecular-Mechanics Dynamics Simulations of Proton Transfer through a Carbon Nanotube. *J. Phys. Chem. B* 121(36):8585–8592.

99. Masterton WL, Bolocofsky D, & Lee TP (1971) Ionic Radii from Scaled Particle Theory of the Salt Effect. *J. Phys. Chem.* 75(18):2809–2815.

100. Fawcett WR (2005) Charge distribution effects in the solution chemistry of polyatomic ions. *Condens. Matter Phys.* 8(2):413–424.

101. Iwadate Y, Kawamura K, Igarashi K, & Mochinaga J (1982) Effective Ionic Radii of NO_2^- and SCN^- Estimated in Terms of the Böttcher Equation and the Lorentz–Lorenz Equation. *J. Phys. Chem.* 86(26):5205–5208.

102. Swanson JMJ (2018) Chapter 10 Computational Means of Assessing Proton Pumping in Cytochrome c Oxidase (Complex IV). *Mechanisms of Primary Energy Transduction in Biology*, ed Wikström M (The Royal Society of Chemistry), pp 249–272.

103. Lim H-H, Stockbridge RB, & Miller C (2013) Fluoride-dependent interruption of the transport cycle of a CLC Cl^-/H^+ antiporter. *Nat. Chem. Biol.* 9(11):721−725.

104. Stockbridge RB, *et al.* (2012) Fluoride resistance and transport by riboswitch-controlled CLC antiporters. *Proc. Natl. Acad. Sci. U. S. A.* 109(38):15289–15294.

CHAPTER 3

LOCAL CONFORMATIONAL DYNAMICS REGULATING TRANSPORT PROPERTIES OF THE CLC ANTIPORTER

3.1 Introduction

The chloride channel (ClC) family of transmembrane proteins, composed of Cl^- channels and Cl^-/H^+ antiporters (1), has been discovered in nature across multiple species including many bacteria, some archaea, and almost all eukaryotes (including humans) (2-4). They are responsible for various physiological processes ranging from epithelial salt homeostasis (5) to pH regulation of intracellular compartments (6). Mutations in human ClC genes have proven to be related to pathologies of bone, muscle, kidney, and eye (1), emphasizing their importance in human physiology.

ClC-ec1 (**Figure 3.1** (A)) is a Cl^-/H^+ antiporter from *Escherichia coli*, and also the most structurally, functionally, and mechanistically investigated protein in the ClC family (7-12). Structurally, crystallographic data (7, 8) of the protein has revealed its homodimeric structure as well as three Cl^- binding sites (S_{int}, S_{cen}, and S_{ext}) per monomer. The central binding site (S_{cen}), with distinguishable Cl^- occupancy in the crystal structure (8), is located in the central cavity and surrounded by several conserved residues including S107, E148, I356, F357, A358, and Y445 (ClC-ec1 numbering). Inferred from the crystal structure, the S_{cen} anion is stabilized by the

backbone atoms of I356 and F357 and the side chains S107 and Y445. The other anion sites are located either above or below S_{cen} along the membrane normal axis.

Functionally, the wild-type (WT) ClC-ec1 exchanges Cl^- ions for H^+ at a stoichiometry of (2.2±0.1):1 under normal conditions (9, 12), and therefore should be classified as a ClC antiporter. The normal proton transport rate is 1.0×10^3 s^{-1} (12, 13), but it can be slowed down or even eliminated by alternative polyatomic anions like NO_3^- and SCN^- (14). The anion/proton stoichiometry is also shifted to 7–10:1 or a much higher value by these anions (14). The two subunits of the homodimer have also been shown to work independently by experiments employing cross-linking (15) and by the monomeric form of ClC-ec1 (16).

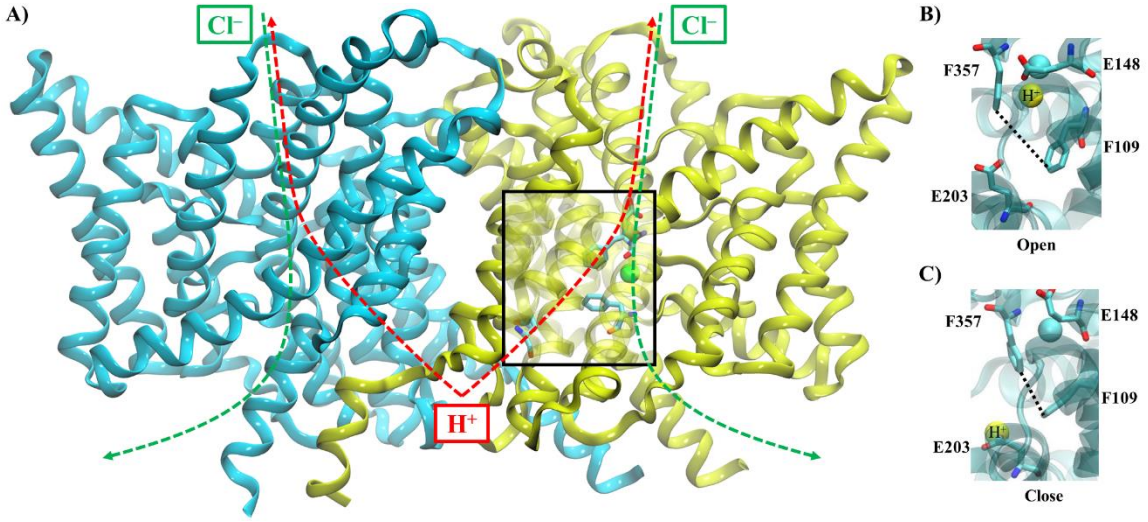

Figure 3.1 Equilibrated structure of the ClC mutant and phenylalanine gate states
(A) Equilibrated homodimeric structure (monomer A in yellow and monomer B in cyan) of ClC-ec1 I109F mutant with Cl^- (green sphere) bound to the S_{cen} site and some key residues shown in monomer A. Approximate ion transport pathways are indicated as red (H^+) and green (Cl^-) dashed arrows. The (B) open and (C) closed state of the phenylalanine gate formed by the mutated residue F109 and F357 in I109F mutant, with residues F109, E148, E203, F357 displayed and labeled. The excess proton and S_{cen} Cl^- are shown as yellow and cyan spheres, respectively.

Mechanistically, the anion and proton transport (PT) pathways have been identified (see **Figure 3.1** (A)) based on crystal structures (7, 8) as well as numerous mutagenesis experiments

(13, 17).These pathways intersect in the central cavity around the S_{cen} site, which is the most stable binding site confirmed by isothermal titration calorimetry (ITC) experiments (18) and computational calculations of the free energy profile (19). Three anion binding sites, S_{int}, S_{cen}, and S_{ext}, were identified in the anion transport pathway, while E148 and E203 were recognized as important proton binding sites (7, 8, 13, 17). Surprisingly, there are no crystallographic waters to date, nor other titratable residues, between these two residues (7, 8), complicating our understanding of the PT mechanism. Such proton carriers are crucial for the Grotthuss shuttling of PT, since the excess proton diffuses via the dynamic formation and concomitant cleavage of covalent O–H bonds among proton-carrying molecules, unlike other hydrated ions.

Based on several molecular dynamics (MD) simulations of the WT protein, it was proposed that a transient water network could be formed to facilitate PT (20-23). Our reactive MD simulations further revealed an important dynamical coupling between PT and an increased hydration between E203 and E148 (20, 23). These simulations also demonstrated how Cl^- at S_{cen} is necessary for PT through the central cavity in one direction, and influential in proton release from E148 in the other (23-25). Our recent work further demonstrated how multiple competing pathways with different intermediates and rate limiting steps cumulatively contribute to the total anion/proton flux (25). This notion, that multiple sequences of transitions between microstates contribute to the total flux (25), is quite different from the commonly proposed single-pathway mechanisms (12, 26), and explains for the first time the origin of the 2.2:1 stoichiometry.

Despite being extensively studied, several aspects of ClC-ec1 remain unanswered or under debate. For example, it is still unclear how the ClC family of proteins are partitioned into channels and antiporters despite sharing very similar structures (27). Similar to other channels, but unlike other active secondary transporters such as LacY and NapA (28), ClC-ec1 does not

seem to use large-scale conformational changes to enable antiporting (29). However, it is thought that conformational dynamics of residues along the transport pathway, including I109 (22), E148 (23, 30), and F357 (30), are important for transport properties. The side chain rotation of I109 was proposed to be important for hydration in the central cavity (22). Similarly, the concerted movement of F357 and E148 side chains was suggested to be an essential component of the transport cycle based on MD simulations of WT ClC-ec1, combined with uptake assay and ITC measurements on the F357A mutant (30).

Several experimental (12, 31-34) and computational (34) studies have also suggested that conformational changes in helical structures further away from the transport pathways could contribute to the ion transport cycle. However, no crystal structure of ClC-ec1 or other homologs has shown such global conformational changes, and the magnitude as well as the impact of these structural rearrangements remains unclear.

In the absence of large conformational changes, it is curious how the functionality of ClC proteins are altered and regulated. Experiments have demonstrated that alternative anions modulate the transport properties of ClC-ec1 (14). We were able to explain this effect of alternative anions in S_{cen} by their ability to disrupt the water network in the central cavity and emphasized the importance of a complete water network for PT (35).

Several ClC mutants including the I109F mutant have also been shown to disrupt ion transport (22). For the I109F mutant, the PT rate is significantly reduced by an order of magnitude to ~1.0×10^2 s^{-1}, and the Cl^- rate is decreased from ~2.3×10^3 s^{-1} to ~2.6×10^2 s^{-1}. In non-reactive MD simulations of the WT protein, side-chain rotation of I109 was observed, and the residue was identified as a gating residue lying within the central cavity (22). However, the simulation did not directly simulate PT, and the I109 rotation was not characterized statistically

or energetically. Interestingly, PT from E203 to E148 is not rate-limiting in one orientation based on a more recent reactive MD study (23), which contradicts the proposed gating effect of I109 between E203 and E148. Nevertheless, the rate-limiting step is not necessarily transferrable to the I109F mutant, and no simulation of PT in the I109F mutant has been performed yet. Thus, further computational investigations of the I109F mutant are necessary to understand the mechanism by which this mutation regulates function.

This work focuses on the mechanistic details of how a mutation influences PT properties. Through comprehensive simulations with the correct description of delocalization of an excess charge and Grotthuss shuttling, we have determined the free energy profile for PT through the central cavity in I109F mutant. We find that an important phenylalanine gate is formed between F109 and F357 that disrupts the hydration network in the central cavity and thus slows down PT by an order of magnitude. We further demonstrate how water connectivity is crucial for PT, and propose a new approach for regulating hydration to alter transport properties. This alternative approach of controlling hydration through local conformational dynamics is expected to influence the PT mechanism of other similar proteins, such as cytochrome *c* oxidase (C*c*O) (36).

3.2 Methods

3.2.1 Equilibration with Nonreactive Molecular Dynamics

The simulation system consisted of the ClC-ec1 homodimer (PDB ID: 1OTS) (8), 163 POPE lipids, 17 Cl^- anions, and ~11,000 TIP3P water molecules in a 92 Å × 92 Å × 79 Å box with periodic boundary conditions. The equilibrated wild-type structure (23) was mutated to

I109F, followed by energy minimization. It was then equilibrated with the GROMACS 5.1 software (37) for ~0.5 μs.

The CHARMM-CMAP and CHARMM36 (38-40) force field was used to model the protein and lipids, respectively. E203 and D417 in monomer A and E113 in monomer B were protonated, while all other residues were in their standard protonation states, consistent with previous pK_a calculations performed on the same crystal structure (41) and our previous investigation of proton transport from E113 to E203 (23).

The interval of leap-frog integration was 2 fs. The neighbor list (cutoff = 12.0 Å) was updated every 40 fs. The particle mesh Ewald (PME) method (42) was adopted to calculate long-range electrostatic interactions, with a spherical cutoff of 12.0 Å, while the short-range electrostatic potential was shifted by a constant. A switching function was used for van der Waals interactions in the range of 10.0 Å and 12.0 Å. Velocity rescaling with a stochastic term (43) was employed for constant volume and constant temperature (T = 303.15 K) (NVT) molecular dynamics simulations.

3.2.2 Parametrization for E148 and E203

To describe charge delocalization and Grotthuss proton shuttling which are not covered in standard molecular dynamics, the multiscale reactive molecular dynamics (MS-RMD) method (44, 45) defines a set of possible protonation states of water molecules and/or titratable residues, each with a specific bonding topology. Each point of the potential energy surface (PES) at a specific atomic configuration is the optimized system energy within the linear space spanned by these protonation basis states. It can be calculated though solving the eigensystem problem of a Hamiltonian matrix, where the diagonal entries are the system energy given by the force field

and the bonding topology of a specific protonation state, and the off-diagonal terms capture the coupling between basis states. Therefore, chemical bonds form and/or break when the pivot state (most probable basis state) changes, and the charge delocalization is a natural result of the system being in a linear combination of the basis states. Assumptions of the method, detailed explanations of the functionals in the Hamiltonian matrix, and the parameters for the excess proton in water have been extensively discussed in our previous publications (44-50).

To model the system with MS-RMD, several parameters introduced in the Hamiltonian matrix must be optimized. Usually this is done by performing quantum mechanical single point calculations and force-matching (51) for an ensemble. The fitting procedure has also been described in detail and justified in previous publications (52, 53).

In this study, the MS-RMD parameters for E148 and E203 in monomer A of ClC-ec1 were parametrized. They were treated differently due to their sensitivity to local environment. The initial guess was chosen to be parameters for glutamic residues in the ClC-ec1 protein with Cl^- present in S_{cen} (53). This set of parameters was also used in our test simulations for the discovery of a second CV. To generate the training set with configurations of deprotonation of glutamic residues, 2D REUS simulations were performed. The collective variables were the same as the production run, but the parameters were the initial guess.

Configurations were sampled every ~20 ps, and approximately 1600−2400 configurations were collected for each fitting. For each configuration, a single point energy calculation was performed with the CP2K 5.1 software (54, 55) to calculate the atomistic forces in the quantum mechanics (QM) region with the hybrid quantum mechanics/molecular mechanics (QM/MM) method.

Consistent with our previous work on the same protein (35), the QM atoms were treated with density functional theory (DFT) using the BLYP-D3/TZV2P functional and basis set (56-58), the third generation of the dispersion correction (59, 60), the Gaussian and plane wave (GPW) scheme (61), and Goedecker−Teter−Hutter (GTH) pseudopotentials optimized for BLYP functional (62-64). The QM region includes the S_{cen} Cl^- anion, the side chains of residues S107, E148, E203, and Y445, and the water molecules within the transport cavity or within the third solvation shell from the fitted residue. The QM box size was ~18–32 Å in each dimension to ensure that a ~6–8 Å buffer was included between the QM atoms and the boundary of the QM box. The Gaussian expansion of the electrostatic potential (GEEP) scheme was used to treat the QM/MM electrostatic coupling with periodic boundary conditions (PBCs) (65, 66), and the spurious QM/QM periodic image interactions were decoupled according to Blöchl's work (67). The chemical bonds that crossed the QM/MM boundary between the alpha carbon (MM) and the beta carbon (QM) of the residues were capped with hydrogen atoms, on which the forces were calculated following the IMOMM scheme with a scaling factor of 1.50 (68). The particle mesh Ewald method developed by Essmann *et al.* (69) and orbital transformation optimized by VandeVondele *et al.* (70) were also adopted. Fourier transform was performed utilizing FFTW3 (71). The distributed block-compressed sparse row (DBCSR) library (72) was utilized to perform sparse matrix calculation. Otherwise, the QM setup was consistent with previous work (35).

In the optimization procedure of MS-RMD parameters for E148 and E203, genetic algorithm (73) was employed to minimize the residual χ^2 of squared differences between the forces calculated with QM/MM ($\mathbf{F}_{ij}^{\mathrm{ref}}$) and those calculated with MS-RMD ($\mathbf{F}_{ij}^{\mathrm{cal}}$):

$$\chi^2 = \frac{1}{3MN}\sum_{j=1}^{M}\sum_{i=1}^{N}\left|\mathbf{F}_{ij}^{\mathrm{cal}} - \mathbf{F}_{ij}^{\mathrm{ref}}\right|^2, \quad (3.1)$$

where the M and N are the number of configurations and the number atoms included in the fitting in each configuration, respectively. The fitted atoms are those in the "reactive complex" (*i.e.*, the Glu residue and first three solvation shells), as determined by the MS-RMD state-searching algorithm (49). The models for E148 and E203 were developed independently due to their separation of ~10−16 Å precluding any direct coupling. To account for the chemical nonequivalence of Glu and water, asymmetric off-diagonal form was adopted (48). The new MS-RMD parameters for E148 and E203 are given in **Table 3.1**.

Table 3.1 The MS-RMD model parameters for E148 and E203 with Cl^- present in S_{cen}.

	E148	**E203**
$\boldsymbol{B}$ **(kcal mol^{-1})**^{a}	3.41	1.50
$\boldsymbol{b}$ **(Å^{-1})**^{a}	0.83	1.79
$\boldsymbol{b'}$ **(Å^{-2})**^{a}	0.40	1.77
$\boldsymbol{C_{\mathrm{diag}}}$ **(kcal mol^{-1})**^{a}	3.89	4.96
$\boldsymbol{c}$ **(Å^{-1})**^{a}	1.78	1.79
$\boldsymbol{V_{ii}}$ **(kcal mol^{-1})**^{a}	–146.1	–135.3
$\boldsymbol{V_{ij}^{\mathrm{const}}}$ **(kcal mol^{-1})**^{b}	–69.9	–46.7
$\boldsymbol{r_{\mathrm{sc}}^{0}}$ ^{b}	1.46	1.21
$\boldsymbol{\alpha}$ **(Å^{-2})**^{b}	0.94	1.19
$\boldsymbol{a_{\mathrm{DA}}}$ **(Å)**^{b}	3.24	3.25
$\boldsymbol{\gamma}$ **(Å^{-2})**^{b}	1.34	0.85

a The definition of diagonal parameters was described in our previous publication (52, 74). Bonding parameters for the protonated Glu state, the switching function, d_{OO}^0, and d_{OH}^0 were fixed and consistent with our previous work (52).
b Off-diagonal parameters were defined in Reference (48) and set to be zero unless explicitly shown in the table, with one exception (C_{offdiag} = 1).

3.2.3 Definition of Collective Variables

To find the most relevant collective variable (CV) coupled to the water dynamics, test simulations were performed. They used the same settings as production simulations, except that water density CV (75) was adopted as an initial guess of the second CV, and that the MS-RMD parameters for E148 and E203 were not reparametrized (*i.e.*, these parameters were retrieved

from our previous publication (53)). The absolute value of covariance between the candidate CV and the water-wire gap ratio (a quantity reflecting the quality of water network, described in the next section) was calculated to determine the CV most correlated with hydration level.

In production simulations, two-dimensional replica-exchange umbrella sampling (REUS) (76) was employed to calculate potentials of mean force (PMFs) with an accelerated convergence rate. The first CV, ξ_1, was the ratio-based center of excess charge (CEC) position defined in our previous work (35). This ratio-based CV describes the distance between donor and acceptor and the progress of proton transport simultaneously, but does not introduce artificial bias (35). The second CV, ξ_2, was set to be the tip distance between the outmost carbon atoms of the phenyl rings in the F109 and F357 side chains, as it has the largest absolute covariance with the gap ratio, and was shown to be correlated to the hydration level in the central cavity (see **Figure 3.2** and **Figure 3.3**).

3.2.4 PT Simulations, PMF Calculation, and Rate Constant Estimation

After the equilibration, the proton on the E203 residue in monomer A was treated as the excess proton. All following discussions and simulations were done for monomer A, since two subunits have been shown to work independently (15, 16).

To simulate PT explicitly, multiscale reactive molecular dynamics (MS-RMD) (45) was employed. To correctly describe charge delocalization and Grotthuss shuttling of proton transport, MS-RMD performs an on-the-fly diagonalization of a Hamiltonian matrix and calculates the most probable mixture of states, each with a different bonding topology.

Simulations were performed with the RAPTOR package (45) embedded in the LAMMPS software (77), with umbrella potentials implemented in the PLUMED 2 package (78).

The excess proton and water molecules were described by the MS-EVB 3.2 model (50). Parameters for protonatable residues, E148 and E203 in monomer A, were determined with the FitRMD parametrization procedure (see section **3.2.2 Parametrization for E148 and E203**) (52, 53).

The time interval for velocity-Verlet integration (79) was adjusted to 1 fs in the production run. Neighbor list was built every 1 fs after at least 2 fs of delay. Particle-particle–particle-mesh (P^3M) solver (80) was used to calculate long-range electrostatic interactions, while the short-range counterpart was multiplied by a switching function in the range of 10.0 Å and 12.0 Å and an additional damping factor. Nosé–Hoover chain (81, 82) (chain length = 3) was employed for NVT (T = 300 K) simulations. To be consistent with our previous work, a very weak harmonic restraint (force constant = 2.4 kcal mol^{-1} $Å^{-2}$) was applied to each of α-carbon atoms 20 Å away from the reaction center to tether them to their initial position. This was essential in previous simulations in which a water density box was explicitly sampled, and is not expected to be important given the lack of significant conformational motions in distal regions in unrestrained simulations (29, 35). All other system setup and simulation details are consistent with our previous work (23, 35).

In the replica-exchange umbrella sampling (REUS) simulations, the first collective variable (CV), ξ_1, was the ratio-based center of excess charge (CEC) position defined (35) as

$$\xi_1 = \frac{\min\limits_{i\in\{1,2\}}\left(\mathbf{d}^{i}_{\mathrm{E203,CEC}}\cdot\hat{\mathbf{n}}\right)}{\min\limits_{i,j\in\{1,2\}}\left(\mathbf{d}^{ij}_{\mathrm{E203,E148}}\cdot\hat{\mathbf{n}}\right)}, \tag{3.2}$$

where $\hat{\mathbf{n}}$ is a prescribed unit vector pointing from E203 to E148, $\mathbf{d}^{i}_{\mathrm{E203,CEC}}$ and $\mathbf{d}^{ij}_{\mathrm{E203,E148}}$ are the displacement vector pointing from the i-th oxygen of E203 side chain to either CEC (47, 49) or the j-th oxygen of E148 side chain, respectively, with i and j enumerating from 1 to 2. The min

function used in ξ_1 is a smoothed approximation of the mathematical minimum function defined in our previous publication (35). Despite being different from the curvilinear path CV used in our previous work (23), this ratio-based CV describes the distance between donor and acceptor and the progress of proton transport simultaneously, but does not introduce artificial bias (35).

The weighted histogram analysis method (WHAM) (83, 84) was utilized to combine the potential of mean forces (PMFs) generated from the separate windows of REUS simulations. The error bars on the PMFs were estimated using block averaging by dividing each trajectory into consecutive blocks. PMFs were processed and plotted with the help of GNU Parallel (85) and Matplotlib (86). The protein structures were rendered in VMD (87).

Rate constants were calculated following the same procedure introduced in our previous work (35) based on the transition state theory (TST) (88, 89) and the equipartition theorem.

3.2.5 Characterization of Water Connectivity

The water-wire gap ratio and K-shortest path (KSP) length were used to quantify water connectivity. They were plotted against the tip distance CV (ξ_2) to identify correlation. Neither of these measures is continuously differentiable, so they cannot be used as a CV for REUS.

The water-wire gap ratio was calculated following our previous procedure (35). The breadth-first algorithm was used to search for any possible water network connected through hydrogen bonds, which were defined by distance criterion and angle criterion: (1) the distance between the donor and acceptor heavy atoms is shorter than 3.5 Å; and (2) the angle formed by the donor atom, the central hydrogen, and the acceptor atom is larger than 150 degrees. With the maximum depth set to 20, the searching algorithm was performed twice, once from E203 to E148 and once in the opposite direction, so as to search for all paths between the donor and the

acceptor. The gap ratio was the gap distance divided by channel length, where "gap distance" was defined as the smallest distance between two disconnected groups of heavy atoms in the water networks starting from either E203 or E148, and "channel length" was given by the smallest distance between two groups of side-chain oxygen atoms of either E203 or E148.

To calculate the K-shortest path length (K = 1), the "distance matrix" of water molecules in the transport channel as well as oxygen atoms of side chains of E148 and E203 was calculated first. The ij-th element d_{ij} was given by

$$d_{ij} = 1 + \left(r_{ij}/r_0\right)^6, \tag{3.3}$$

where r_{ij} represents the distance between the i-th atom and the j-th atom, and r_0 is a scaling factor and set as 3.0 Å. On the undirected graph given by the distance matrix, the length of the shortest path starting from carboxyl oxygen atoms of E203 to those of E148 was calculated using Yen's algorithm (90). The natural logarithm of the resulting quantity was taken for a smaller range and better description of water connectivity. The calculation was implemented in a revised version of PLUMED 2 (78) and Boost Graph Library.

Clustering of data points in **Figure 3.2** and **Figure 3.3** were done using Gaussian mixture model (91), calculated with scikit-learn package (92) implemented in Python 3. Other clustering algorithms (*k*-means++ (93), spectral clustering, *etc.*) gave similar results. The black lines separating four quadrants were placed along the dividing line between two clusters, such that minimal points are in quadrants I and III. The density plot was normalized such that the integral over the area is unity.

3.3 Results and Discussion

3.3.1 Tip Distance of the Phenylalanine Gate Representing Water Connectivity

To study PT in the I109F mutant of ClC-ec1, we employed REUS for enhanced sampling. As with all rare event processes, finding the CVs that define the slowest motions limiting the process of interest is of paramount importance. As shown in our previous work (35), the water density CV (75), which was used to simulate PT for ClC-ec1 with alternative anions in S_{cen}, was not perfect in all situations. At high water densities, it favored water molecules clustering below the S_{cen} anion instead of a continuous water-wire along the PT pathway. As a result, the minimum free energy path was uncoupled (*i.e.*, it had a Pi-shape, with two ~90-degree kinks) for the nitrate system, meaning the PT process did not show a clear dependence on the water density CV, especially through the transition barrier region. Although we recovered their correlation through post-processing, we want to capture this directly through our enhanced sampling simulations. We developed several quantities to characterize water connectivity directly, but they are not continuously differentiable and cannot be used as a CV in REUS. Therefore, we characterized alternative degrees of freedom (DOFs) in the system to find the most relevant CV coupled to the water dynamics and to get an idea of factors influencing water structure.

From the trajectories of the test simulations, we observed some conformational changes of F109 and F357 (**Figure 3.1** (B) and (C)). Given that these two residues were proposed to affect transport properties (22, 30), we supposed that these conformational changes could couple with the amount of water around the S_{cen} anion. We proposed three candidate CVs: (1) the minimum distance between the carbon atoms in two Phe residues, (2) the distance between the center of the phenyl carbons, (3) the tip distance between the outmost carbon atoms of two Phe residues. After analyzing the test simulations, we discovered that the tip distance has the largest

absolute covariance with the water network gap ratio. We believe that the tip distance CV is flexible enough to allow rotation along the principal axis of the phenyl group without being changed, but it is not overly flexible in that it forbids free rotation along an arbitrary axis.

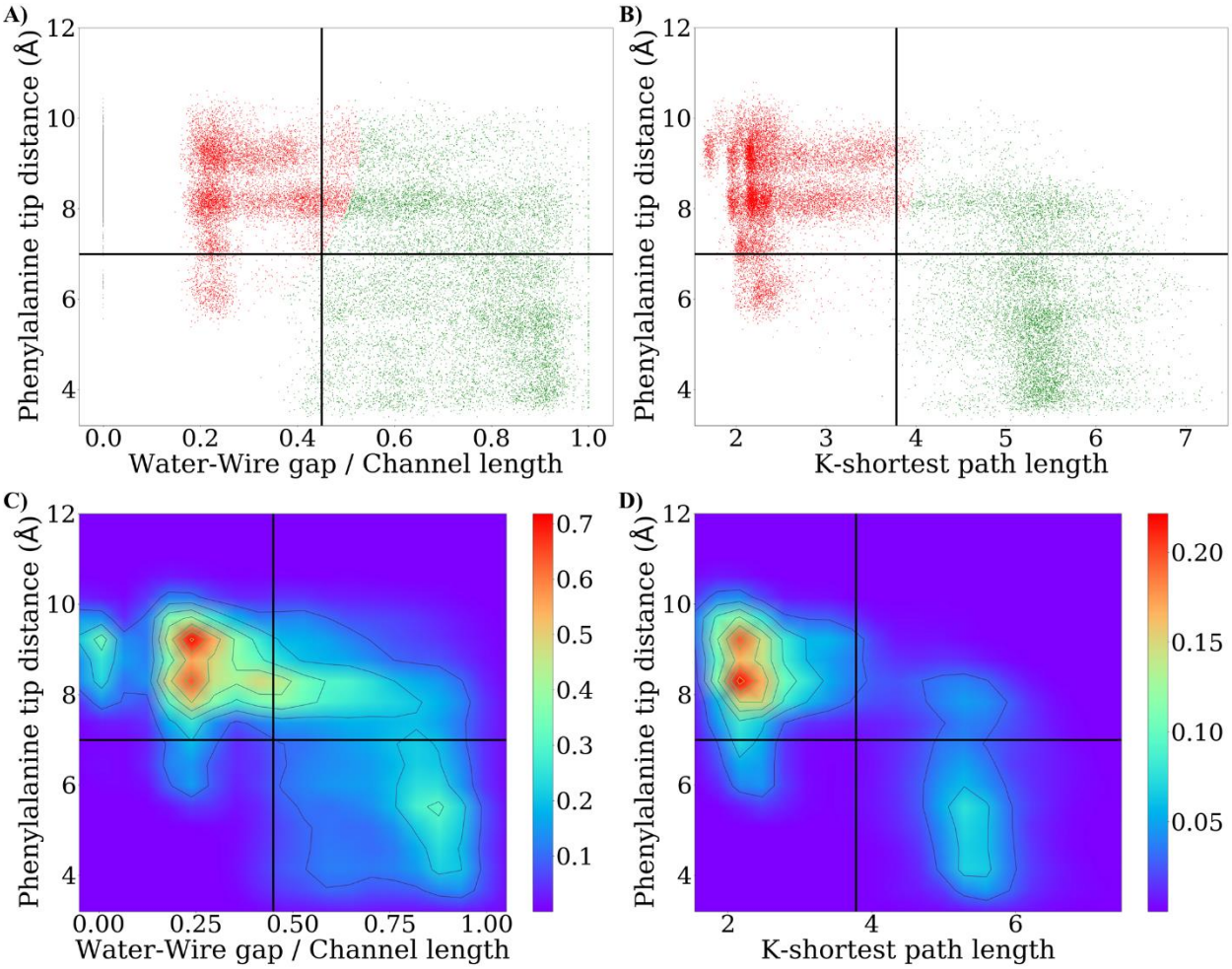

Figure 3.2 Scatter-point plot of (A) gap ratio or (B) KSP length vs tip distance of Phe gate calculated from REUS simulations, and (C, D) the corresponding density plots
Data points were clustered into (A) 3 or (B) 2 components according to Gaussian mixture model (91), with each cluster colored red, green, and gray, respectively. Black lines partition the plot into four quadrants. These lines were placed along the dividing line between red and green clusters, such that minimal points are in quadrants I and III. The density plot was normalized such that the integral over the area is unity.

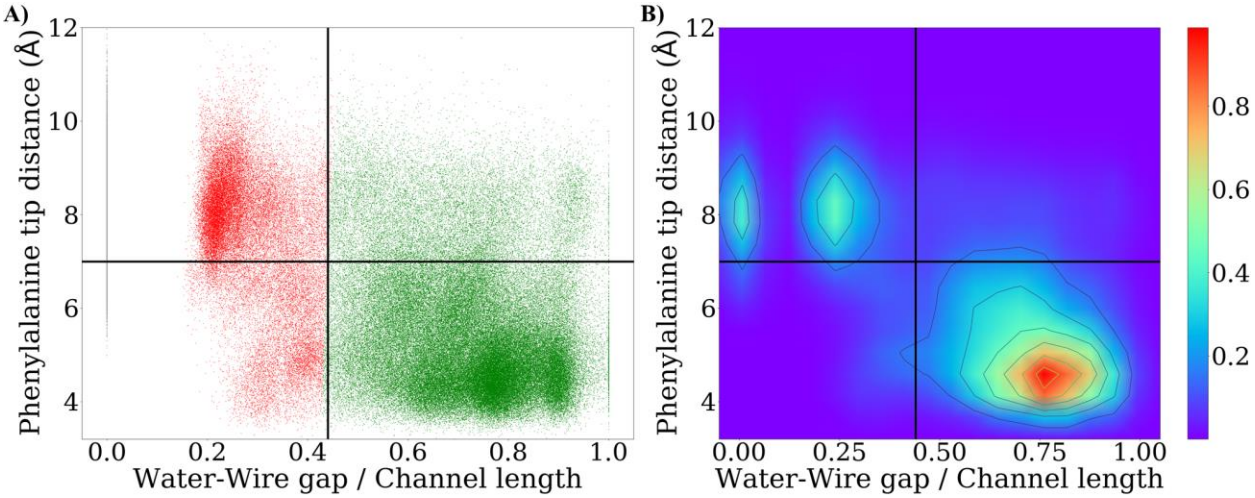

Figure 3.3 (A) Scatter-point plot of gap ratio vs tip distance of Phe gate calculated from test simulations, and (B) the corresponding density plots
Data points were clustered into 3 components according to Gaussian mixture model (91), with each cluster colored red, green, and gray, respectively. Black lines partition the plot into four quadrants, and they were placed along the dividing line between red and green clusters, such that minimal points are in quadrants I and III. The density plot was normalized such that the integral over the area is unity. **Figure 3.3** (B) is quantitatively different from **Figure 3.2** (C), because they were obtained from two ensembles. This figure shows a consistent trend that data points in quadrant II is more than those in quadrant I, and that points in quadrant IV is more than those in quadrant III. All interpretations for **Figure 3.2** (A) mentioned in the main text hold for **Figure 3.3** (A) as well.

The gating effect of the phenylalanine gate formed by F109 and F357 (**Figure 3.1** (B) and (C)) was further quantified by the clustering in the correlation plot between this tip distance and two other quantities reflecting water connectivity: the water network gap ratio and KSP length (**Figure 3.2** and **Figure 3.3**). Although we used test simulations to identify the correlation, we further confirmed this by running additional analyses on REUS simulations. The latter (see **Figure 3.2** (A)) is consistent with results from test simulations (**Figure 3.3**). Therefore, the discussion below is based on newer analyses.

The data in **Figure 3.2** (A) clusters in either the upper-left region, where the tip distance is larger than 7 Å and the gap-ratio lower than ~0.4, or the lower-right region, where the tip distance is shorter and the gap-ratio is larger. Thus, a well-connected water network (gap-ratio

less than ~0.3) is strongly correlated with an open Phe gate. Note that the line at the zero-gap region represents a completely connected water network, based on the hydrogen bond definition with distance criterion and angle criterion. These completely connected states are dominated by the larger tip distances (gap ratio above 6 Å) and do not exist for tip distances below 5 Å. As expected, there is an abrupt jump in values from 0 to ~0.2, because our hydrogen bond definition is discontinuous (*i.e.*, 0/1-based). A small shift in the position of a water oxygen/hydrogen atom in a continuous water chain might disconnect its hydrogen bond to a neighboring water molecule along the chain if this bond is just at the definition boundary based on the discontinuous criteria. Such an abrupt change will suddenly increase the gap distance from 0 to the new O–O distance, which is ~3 Å if the angle criterion is broken and ~3.5 Å if the distance criterion is broken. Given the approximate Glu–Glu distance at ~15 Å, it is reasonable to have a gap in the gap ratio between 0 and ~0.2.

In the correlation plot between tip distance and KSP length (**Figure 3.2** (B)), for which a longer length means worse water connectivity, the data shows a similar pattern, with dividing lines at ~7 Å tip distance and at 4 KSP length. Again, the formation of high-quality water networks (KSP length < 3) requires an open Phe gate (tip distance > ~6 Å), and the best-quality water networks (KSP length < 1.8) only accompany a fully open Phe gate (tip distance > ~8 Å).

These plots (**Figure 3.2**) demonstrate that the conformational change of F109 and F357 is important to water structure. Despite a few exceptions, the Phe–Phe tip distance is correlated with the quality of water network within the transport channel. Based on these results, we decided to directly sample the tip distance to control the hydration of the PT pathway. Unlike the gap ratio and KSP length, the tip distance has the advantage of being continuously differentiable

in sampling regions. Thus, the tip distance was used in REUS for enhanced sampling of the channel hydration.

3.3.2 Phenylalanine Gate Regulating Proton Transport

With two CVs defined, we employed REUS to sample PT between E203 and E148. Since both orientations of the protein in the membrane are possible in the *in vitro* experiments (94), we consider the rate-limiting step to be either from E203 to E148, or vice-versa, depending on which has a lower barrier. The orientation with the higher barrier should have lower transport rate and thus contribute less to the experimentally measured flux. The 2D PMF (**Figure 3.4** (A)) shows a 14.3 ± 0.6 kcal mol^{-1} free energy barrier for PT across the transport channel (E148 to E203). The increased transport barrier compared to WT (by ~6 kcal mol^{-1}) (23) shifts the rate-limiting step to the transport within the central region. Based on transition state theory (TST), the rate constant for this PT step is estimated to be $(1.6 \pm 1.1) \times 10^2$ s^{-1}, which is consistent with the experimental value of $\sim 1.0 \times 10^2$ s^{-1} (22). Note that the protonated E203 state is ~1.5 kcal mol^{-1} more stabilized than the protonated E148 state in the I109F mutant of ClC-ec1. In contrast, our previous studies (23) showed energetically downhill PT from E203 to E148 in WT ClC-ec1, using the same MS-RMD methods and sampling protocols. This should be attributed to the decrease in pK_a of E148, as was suggested by its correlation with decreased Cl^- transport rate (25, 95).

From the 2D PMF, the minimum free energy path and extracted 1D free energy profile (**Figure 3.4** (B)) show an inverse U-shaped path compared to the Pi-shape in our previous work on alternative anions bound to ClC-ec1 (35), suggesting a possibly similar stepwise mechanism of PT but more coupling between the Phe gate (and thus water connectivity) and PT. Thus, the

similarity of paths is consistent with the above discussion, in that the conformation of the Phe gate indeed influences the water dynamics within the channel and thereby PT. With mostly hydrophobic residues in the transport cavity, the system starts in a dehydrated state with closed phenylalanine gate and an excess proton loaded at E203 (**Figure 3.1** (C) and **Figure 3.5** (A)). After the phenylalanine gate is opened (shown as increased tip distance in the 2D PMF), the hydration of the central cavity increases (**Figure 3.5** (B)). This is consistent with the 6–7 Å tip distance region in **Figure 3.2**. With a connected water-wire, the excess proton is transported to the center (**Figure 3.5** (C)) and through the phenylalanine gate (**Figure 3.1** (B) and **Figure 3.5** (D)). The gate then closes due to interactions between the two Phe residues and the channel recovers its original dry state attributable to hydrophobic effect (**Figure 3.5** (E)). The increased transport barrier in the I109F mutant compared to WT (23) indicates that the interaction between F109 and F357 is sufficiently stabilized, compared to the interaction between I109 and F357 in the WT protein, to hinder the formation of a connected water-wire, but not strong enough to completely block water penetration and PT.

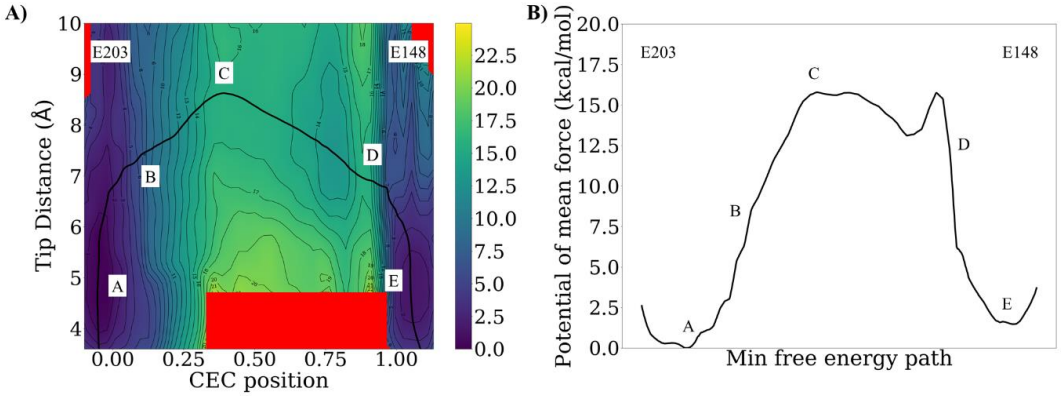

Figure 3.4 Proton transport PMFs for the I109F ClC-ec1
(A) 2D PMF of I109F ClC-ec1, with Cl^-_{cen} present in S_{cen}. Irrelevant high-energy areas (red) are not sampled to reduce the cost of computation. The black line traces the minimum free energy path for PT from E203 to E148. (B) The extracted 1D PMF along the minimum free energy path. The error of the PMF is ~0.6 kcal mol^{-1}, estimated from splitting the trajectories and block-averaging. Five key points along the path are labeled and described in the Results and Discussion section.

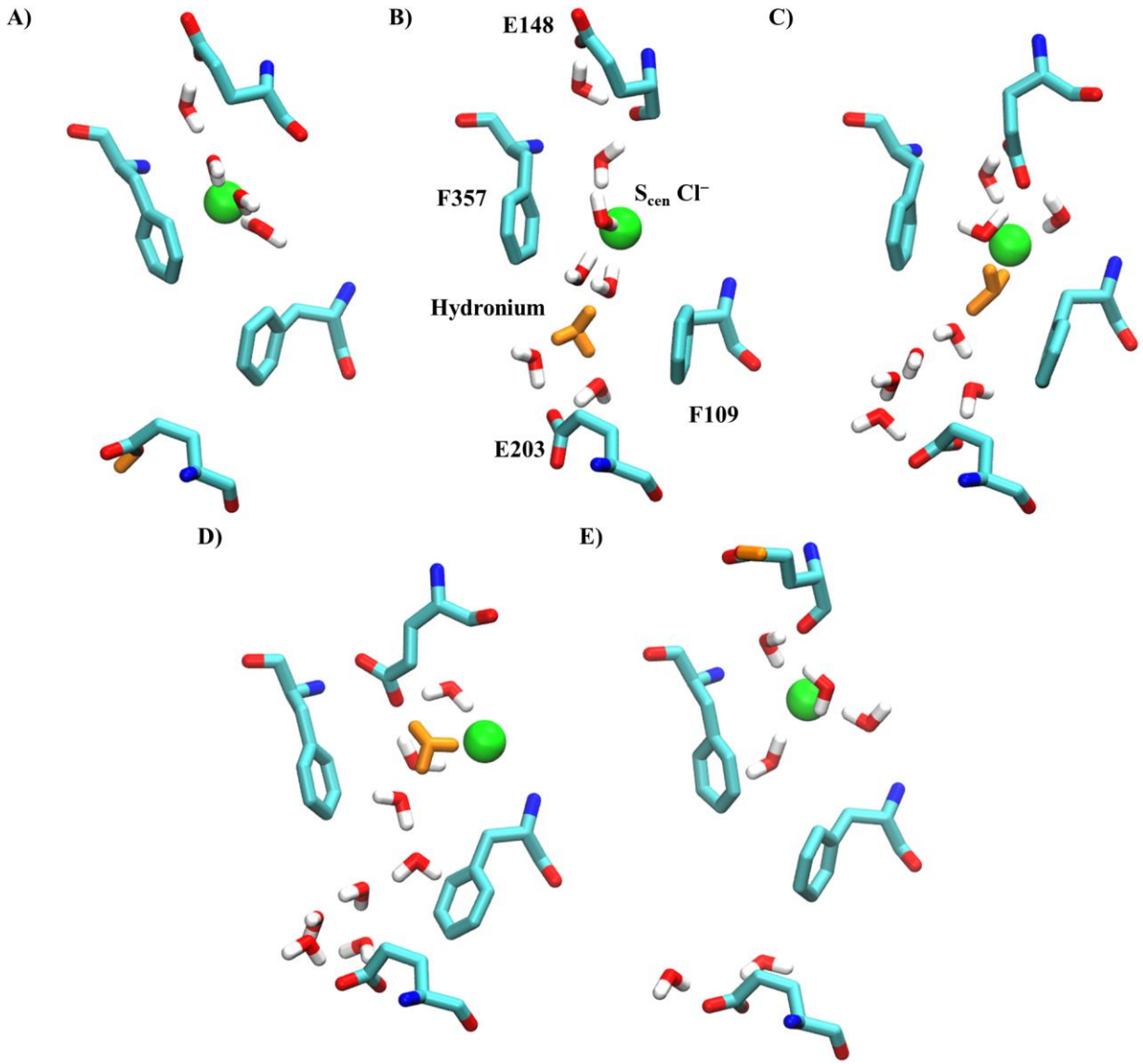

Figure 3.5 Representative configurations of proton transport across the central cavity in the I109F mutant of ClC-ec1 protein.
The panel label is consistent with **Figure 3.4**. The S_{cen} Cl^- anion is shown in VDW representation, and the hydronium/protonated glutamic acid is rendered orange for clarity. Key residues (labeled in panel B) and water molecules are colored according to element.

The observed mechanism of PT re-emphasizes the importance of a connected water network around the excess charge, which is also the reason why alternative anions slow or block PT in WT protein. It also demonstrates that in addition to the chemical nature of alternative

anions (35), the conformational dynamics of the phenylalanine gate, formed by F109 (the mutated residue) and F357, can also regulate water connectivity and hence PT through the central cavity. A similar gating mechanism was reported in C*c*O (36), where the side-chain rotation of N139 controls the asparagine gate formed by N121 and N139, and thus regulates the hydration level of D-channel. Noticeably, these local conformational changes are coupled with the proton movement in the protein. In the ClC mutant, the closed state of the Phe gate is ~2.5 kcal mol^{-1} more stable than the open state when proton is bound to E203, but the open state becomes dominant when proton moves to the center (region C in **Figure 3.4** (A)). A similar coupling effect between PT release from D132 and conformational changes of the asparagine gate in WT C*c*O were also reported (36).

3.4 Conclusions

Multiscale reactive MD simulations with enhanced free energy sampling were performed to investigate the mechanism of I109F mutation slowing down PT in ClC-ec1. With both the mutation and PT treated explicitly, we constructed a 2D PMF for the PT process across the central region past the central Cl^- anion binding site (S_{cen}). Based on transition state theory, the rate constant estimated from the PMF is consistent with the experimental measurements (22), where the PT rate is decreased by an order of magnitude to ~1.0×10^2 s^{-1}. Our result is also consistent with the decreased Cl^- transport rate, since protonated E148, which is essential for Cl^- transport, is destabilized in our simulations.

By comparing our results with previous work on the nitrate/thiocyanate-bound ClC-ec1 (35), we herein confirm a similar hydration–PT–dehydration stepwise mechanism, but in this case the controlling factor is the conformation of the phenylalanine gate formed by F109 and

F357. The correlation plot shows the relationship between the tip distance of the gate and the water connectivity. With these results combined, we conclude that it is the formation and stabilization of the Phe gate that disrupts water connection and hence the proton transport across the central region. This is an alternative approach to regulate PT, compared to the previously reported regulation by the chemical nature of the transported anion. Thus, this work demonstrates how subtle local conformational dynamics can significantly influence the ease of forming hydration networks and thus the rates of charge transport. Such subtle conformational dynamics could be an important factor distinguishing a ClC antiporter from a ClC channel, and are certainly influential regulators of ClC functionality.

Despite quantitively matching experimental results of PT rate, our study does not fully explain the effect of I109 mutation on Cl^- transport. Our future research will include a more complete investigation of the whole Cl^-/H^+ transport cycle of the mutant using multiscale kinetic modeling (MKM) (25).

3.5 References

1. Jentsch TJ (2008) CLC Chloride Channels and Transporters: From Genes to Protein Structure, Pathology and Physiology. *Crit. Rev. Biochem. Mol. Biol.* 43(1):3–36.

2. Pusch M (2004) Structural Insights into Chloride and Proton-Mediated Gating of CLC Chloride Channels. *Biochemistry* 43(5):1135–1144.

3. Chen T-Y (2005) STRUCTURE AND FUNCTION OF CLC CHANNELS. *Annu. Rev. Physiol.* 67(1):809–839.

4. Dutzler R (2007) A structural perspective on ClC channel and transporter function. *FEBS Lett.* 581(15):2839–2844.

5. Uchida S, *et al.* (1993) Molecular Cloning of a Chloride Channel That Is Regulated by Dehydration and Expressed Predominantly in Kidney Medulla. *J. Biol. Chem.* 268(6):3821–3824.

6. Devuyst O, Courtoy PJ, Christie PT, Thakker RV, & Beauwens R (1999) Intra-renal and subcellular distribution of the human chloride channel, CLC-5, reveals a pathophysiological basis for Dent's disease. *Hum. Mol. Genet.* 8(2):247–257.

7. Dutzler R, Campbell EB, Cadene M, Chait BT, & MacKinnon R (2002) X-ray structure of a ClC chloride channel at 3.0 Å reveals the molecular basis of anion selectivity. *Nature* 415(6869):287–294.

8. Dutzler R, Campbell EB, & MacKinnon R (2003) Gating the Selectivity Filter in ClC Chloride Channels. *Science* 300(5616):108–112.

9. Accardi A & Miller C (2004) Secondary active transport mediated by a prokaryotic homologue of ClC Cl^- channels. *Nature* 427(6977):803–807.

10. Walden M, *et al.* (2007) Uncoupling and Turnover in a Cl^-/H^+ Exchange Transporter. *J. Gen. Physiol.* 129(4):317–329.

11. Lim H-H, Shane T, & Miller C (2012) Intracellular Proton Access in a Cl^-/H^+ Antiporter. *PLoS Biol.* 10(12):e1001441.

12. Basilio D, Noack K, Picollo A, & Accardi A (2014) Conformational changes required for H^+/Cl^- exchange mediated by a CLC transporter. *Nat. Struct. Mol. Biol.* 21(5):456–463.

13. Lim H-H & Miller C (2009) Intracellular Proton-Transfer Mutants in a CLC Cl^-/H^+ Exchanger. *J. Gen. Physiol.* 133(2):131–138.

14. Nguitragool W & Miller C (2006) Uncoupling of a CLC Cl^-/H^+ Exchange Transporter by Polyatomic Anions. *J. Mol. Biol.* 362(4):682–690.

15. Nguitragool W & Miller C (2007) CLC Cl^-/H^+ transporters constrained by covalent cross-linking. *Proc. Natl. Acad. Sci. U. S. A.* 104(52):20659–20665.

16. Robertson JL, Kolmakova-Partensky L, & Miller C (2010) Design, function and structure of a monomeric ClC transporter. *Nature* 468(7325):844–847.

17. Accardi A, *et al.* (2005) Separate Ion Pathways in a Cl^-/H^+ Exchanger. *J. Gen. Physiol.* 126(6):563–570.

18. Picollo A, Malvezzi M, Houtman JCD, & Accardi A (2009) Basis of substrate binding and conservation of selectivity in the CLC family of channels and transporters. *Nat. Struct. Mol. Biol.* 16(12):1294–1301.

19. Wang C-H, Duster AW, Aydintug BO, Zarecki MG, & Lin H (2018) Chloride Ion Transport by the *E. coli* CLC Cl^-/H^+ Antiporter: A Combined Quantum-Mechanical and Molecular-Mechanical Study. *Front. Chem.* 6(62):1–17.

20. Wang D & Voth GA (2009) Proton Transport Pathway in the CIC Cl^-/H^+ Antiporter. *Biophys. J.* 97(1):121–131.

21. Ko YJ & Jo WH (2010) Secondary Water Pore Formation for Proton Transport in a CIC Exchanger Revealed by an Atomistic Molecular-Dynamics Simulation. *Biophys. J.* 98(10):2163–2169.

22. Han W, Cheng RC, Maduke MC, & Tajkhorshid E (2014) Water access points and hydration pathways in CLC H^+/Cl^- transporters. *Proc. Natl. Acad. Sci. U. S. A.* 111(5):1819–1824.

23. Lee S, Swanson JMJ, & Voth GA (2016) Multiscale Simulations Reveal Key Aspects of the Proton Transport Mechanism in the ClC-ec1 Antiporter. *Biophys. J.* 110(6):1334–1345.

24. Lee S, Mayes HB, Swanson JMJ, & Voth GA (2016) The Origin of Coupled Chloride and Proton Transport in a Cl^-/H^+ Antiporter. *J. Am. Chem. Soc.* 138(45):14923–14930.

25. Mayes HB, Lee S, White AD, Voth GA, & Swanson JMJ (2018) Multiscale Kinetic Modeling Reveals an Ensemble of Cl^-/H^+ Exchange Pathways in ClC-ec1 Antiporter. *J. Am. Chem. Soc.* 140(5):1793–1804.

26. Feng L, Campbell EB, Hsiung Y, & MacKinnon R (2010) Structure of a Eukaryotic CLC Transporter Defines an Intermediate State in the Transport Cycle. *Philos. Trans. R. Soc., B* 330(6004):635–641.

27. Accardi A (2015) Structure and gating of CLC channels and exchangers. *J. Physiol.* 593(18):4129–4138.

28. Ryan RM & Vandenberg RJ (2016) Elevating the alternating-access model. *Nat. Struct. Mol. Biol.* 23(3):187–189.

29. Jiang T, Han W, Maduke M, & Tajkhorshid E (2016) Molecular Basis for Differential Anion Binding and Proton Coupling in the Cl^-/H^+ Exchanger ClC-ec1. *J. Am. Chem. Soc.* 138(9):3066–3075.

30. Xu Y, Vien M, Liang C, Accardi A, & Bernèche S (2017) Molecular Mechanism of Cl^-/H^+ Coupling in a CLC Antiporter. *Biophys. J.* 112(3):336a.

31. Bell SP, Curran PK, Choi S, & Mindell JA (2006) Site-Directed Fluorescence Studies of a Prokaryotic ClC Antiporter. *Biochemistry* 45(22):6773–6782.

32. Elvington SM, Liu CW, & Maduke MC (2009) Substrate‐driven conformational changes in ClC‐ec1 observed by fluorine NMR. *The EMBO Journal* 28(20):3090‐3102.

33. Abraham SJ, *et al.* (2015) ^{13}C NMR detects conformational change in the 100-kD membrane transporter ClC-ec1. *J. Biomol. NMR* 61(3):209–226.

34. Khantwal CM, *et al.* (2016) Revealing an outward-facing open conformational state in a CLC Cl^-/H^+ exchange transporter. *eLife* 5:e11189.

35. Wang Z, Swanson JMJ, & Voth GA (2018) Modulating the Chemical Transport Properties of a Transmembrane Antiporter via Alternative Anion Flux. *J. Am. Chem. Soc.* 140(48):16535–16543.

36. Liang R, Swanson JMJ, Peng Y, Wikström M, & Voth GA (2016) Multiscale simulations reveal key features of the proton-pumping mechanism in cytochrome *c* oxidase. *Proc. Natl. Acad. Sci. U. S. A.* 113(27):7420–7425.

37. Abraham MJ, *et al.* (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 1–2:19–25.

38. MacKerell AD, Jr., Feig M, & Brooks CL, III. (2004) Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulations. *J. Comput. Chem.* 25(11):1400–1415.

39. Klauda JB, *et al.* (2010) Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *J. Phys. Chem. B* 114(23):7830–7843.

40. Best RB, *et al.* (2012) Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ, ψ and Side-Chain χ_1 and χ_2 Dihedral Angles. *J. Chem. Theory Comput.* 8(9):3257–3273.

41. Faraldo-Gómez JD & Roux B (2004) Electrostatics of Ion Stabilization in a ClC Chloride Channel Homologue from *Escherichia coli*. *J. Mol. Biol.* 339(4):981–1000.

42. Darden T, York D, & Pedersen L (1993) Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* 98(12):10089–10092.

43. Bussi G, Donadio D, & Parrinello M (2007) Canonical sampling through velocity rescaling. *J. Chem. Phys.* 126(1):014101.

44. Knight C, Lindberg GE, & Voth GA (2012) Multiscale reactive molecular dynamics. *J. Chem. Phys.* 137(22):22A525.

45. Yamashita T, Peng Y, Knight C, & Voth GA (2012) Computationally Efficient Multiconfigurational Reactive Molecular Dynamics. *J. Chem. Theory Comput.* 8(12):4863–4875.

46. Schmitt UW & Voth GA (1998) Multistate Empirical Valence Bond Model for Proton Transport in Water. *J. Phys. Chem. B* 102(29):5547–5551.

47. Day TJF, Soudackov AV, Čuma M, Schmitt UW, & Voth GA (2002) A second generation multistate empirical valence bond model for proton transport in aqueous systems. *J. Chem. Phys.* 117(12):5839–5849.

48. Maupin CM, Wong KF, Soudackov AV, Kim S, & Voth GA (2006) A Multistate Empirical Valence Bond Description of Protonatable Amino Acids. *J. Phys. Chem. A* 110(2):631–639.

49. Wu Y, Chen H, Wang F, Paesani F, & Voth GA (2008) An Improved Multistate Empirical Valence Bond Model for Aqueous Proton Solvation and Transport. *J. Phys. Chem. B* 112(2):467–482.

50. Biswas R, Tse Y-LS, Tokmakoff A, & Voth GA (2016) Role of Presolvation and Anharmonicity in Aqueous Phase Hydrated Proton Solvation and Transport. *J. Phys. Chem. B* 120(8):1793–1804.

51. Izvekov S, Parrinello M, Burnham CJ, & Voth GA (2004) Effective force fields for condensed phase systems from *ab initio* molecular dynamics simulation: A new method for force-matching. *J. Chem. Phys.* 120(23):10896–10913.

52. Nelson JG, Peng Y, Silverstein DW, & Swanson JMJ (2014) Multiscale Reactive Molecular Dynamics for Absolute pK_a Predictions and Amino Acid Deprotonation. *J. Chem. Theory Comput.* 10(7):2729–2737.

53. Lee S, Liang R, Voth GA, & Swanson JMJ (2016) Computationally Efficient Multiscale Reactive Molecular Dynamics to Describe Amino Acid Deprotonation in Proteins. *J. Chem. Theory Comput.* 12(2):879–891.

54. VandeVondele J, *et al.* (2005) Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Comput. Phys. Commun.* 167(2):103–128.

55. Hutter J, Iannuzzi M, Schiffmann F, & VandeVondele J (2014) cp2k: atomistic simulations of condensed matter systems. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* 4(1):15–25.

56. Lee C, Yang W, & Parr RG (1988) Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* 37(2):785–789.

57. Becke AD (1988) Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* 38(6):3098–3100.

58. Schäfer A, Huber C, & Ahlrichs R (1994) Fully optimized contracted Gaussian basis sets of triple zeta valence quality for atoms Li to Kr. *J. Chem. Phys.* 100(8):5829–5835.

59. Grimme S, Antony J, Ehrlich S, & Krieg H (2010) A consistent and accurate *ab initio* parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* 132(15):154104.

60. Grimme S, Ehrlich S, & Goerigk L (2011) Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* 32(7):1456–1465.

61. Lippert G, Hutter J, & Parrinello M (1997) A hybrid Gaussian and plane wave density functional scheme. *Mol. Phys.* 92(3):477–487.

62. Goedecker S, Teter M, & Hutter J (1996) Separable dual-space Gaussian pseudopotentials. *Phys. Rev. B* 54(3):1703–1710.

63. Hartwigsen C, Goedecker S, & Hutter J (1998) Relativistic separable dual-space Gaussian pseudopotentials from H to Rn. *Phys. Rev. B* 58(7):3641–3662.

64. Krack M (2005) Pseudopotentials for H to Kr optimized for gradient-corrected exchange-correlation functionals. *Theor. Chem. Acc.* 114(1):145–152.

65. Laino T, Mohamed F, Laio A, & Parrinello M (2005) An Efficient Real Space Multigrid QM/MM Electrostatic Coupling. *J. Chem. Theory Comput.* 1(6):1176–1184.

66. Laino T, Mohamed F, Laio A, & Parrinello M (2006) An Efficient Linear-Scaling Electrostatic Coupling for Treating Periodic Boundary Conditions in QM/MM Simulations. *J. Chem. Theory Comput.* 2(5):1370–1378.

67. Blöchl PE (1995) Electrostatic decoupling of periodic images of plane ‐ wave ‐ expanded densities and derived atomic point charges. *J. Chem. Phys.* 103(17):7422–7428.

68. Maseras F & Morokuma K (1995) IMOMM: A New Integrated *ab initio* + Molecular Mechanics Geometry Optimization Scheme of Equilibrium Structures and Transition States. *J. Comput. Chem.* 16(9):1170–1179.

69. Essmann U, *et al.* (1995) A smooth particle mesh Ewald method. *J. Chem. Phys.* 103(19):8577–8593.

70. VandeVondele J & Hutter J (2003) An efficient orbital transformation method for electronic structure calculations. *J. Chem. Phys.* 118(10):4365–4369.

71. Frigo M & Johnson SG (2005) The Design and Implementation of FFTW3. *Proc. IEEE* 93(2):216–231.

72. Borštnik U, VandeVondele J, Weber V, & Hutter J (2014) Sparse matrix multiplication: The distributed block-compressed sparse row library. *Parallel Computing* 40(5):47–58.

73. Mühlenbein H & Schlierkamp-Voosen D (1993) Predictive Models for the Breeder Genetic Algorithm I. Continuous Parameter Optimization. *Evol. Comput.* 1(1):25–49.

74. Wu Y, Chen H, Wang F, Paesani F, & Voth GA (2008) An Improved Multistate Empirical Valence Bond Model for Aqueous Proton Solvation and Transport. *J. Phys. Chem. B* 112(23):7146–7146.

75. Peng Y, Swanson JMJ, Kang S-g, Zhou R, & Voth GA (2015) Hydrated Excess Protons Can Create Their Own Water Wires. *J. Phys. Chem. B* 119(29):9212–9218.

76. Sugita Y, Kitao A, & Okamoto Y (2000) Multidimensional replica-exchange method for free-energy calculations. *J. Chem. Phys.* 113(15):6042–6051.

77. Plimpton S (1995) Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* 117(1):1–19.

78. Tribello GA, Bonomi M, Branduardi D, Camilloni C, & Bussi G (2014) PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* 185(2):604–613.

79. Swope WC, Andersen HC, Berens PH, & Wilson KR (1982) A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *J. Chem. Phys.* 76(1):637–649.

80. Hockney RW & Eastwood JW (1988) Particle-Particle–Particle-Mesh (P^3M) Algorithms. *Computer Simulation Using Particles*, (CRC Press, New York), 1st Ed, pp 267–304.

81. Nosé S (1984) A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.* 81(1):511–519.

82. Hoover WG (1985) Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* 31(3):1695–1697.

83. Kumar S, Bouzida D, Swendsen RH, Kollman PA, & Rosenberg JM (1992) The Weighted Histogram Analysis Method for Free‐Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* 13(8):1011‐1021.

84. Grossfield A ("WHAM: the weighted histogram analysis method", version 2.0.9.

85. Tange O (2011) GNU Parallel: The Command-Line Power Tool. *;login: The USENIX Magazine* 36(1):42–47.

86. Hunter JD (2007) Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* 9(3):90–95.

87. Humphrey W, Dalke A, & Schulten K (1996) VMD: Visual molecular dynamics. *J. Mol. Graphics* 14(1):33–38.

88. Vineyard GH (1957) FREQUENCY FACTORS AND ISOTOPE EFFECTS IN SOLID STATE RATE PROCESSES. *J. Phys. Chem. Solids* 3(1–2):121–127.

89. Chandler D (1987) *Introduction to Modern Statistical Mechanics*, (Oxford University Press, New York), pp 245–246.

90. Yen JY (1971) Finding the *K* Shortest Loopless Paths in a Network. *Management Science* 17(11):712–716.

91. Yu G, Sapiro G, & Mallat S (2012) Solving Inverse Problems With Piecewise Linear Estimators: From Gaussian Mixture Models to Structured Sparsity. *IEEE Trans. Image Process.* 21(5):2481–2499.

92. Pedregosa F, *et al.* (2011) Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* 12(85):2825–2830.

93. Arthur D & Vassilvitskii S (2007) k-means++: The Advantages of Careful Seeding. *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, ed Gabow H (Society for Industrial and Applied Mathematics, Pennsylvania), pp 1027–1035.

94. Miller C & Nguitragool W (2009) A provisional transport mechanism for a chloride channel-type Cl^-/H^+ exchanger. *Philos. Trans. R. Soc., B* 364(1514):175–180.

95. Picollo A, Xu Y, Johner N, Bernèche S, & Accardi A (2012) Synergistic substrate binding determines the stoichiometry of transport of a prokaryotic H^+/Cl^- exchanger. *Nat. Struct. Mol. Biol.* 19(5):525–531.

CHAPTER 4

ION TRANSPORT MECHANISM OF A FLUORIDE CHANNEL

4.1 Introduction

The fluoride ion is ubiquitous in the aqueous biosphere, appearing in groundwater, sea, and soil typically at 10–100 μM levels (1). Given its toxicity (2, 3), many organisms utilize F^- exporter proteins to reduce the cytoplasmic F^- concentration, so that its toxic effect is minimized (4, 5). Fluoride exporters are divided into two subclasses: the CLC^F proteins, as F^-/H^+ antiporters found in eubacteria (5), and the Fluc proteins, as F^- channels mostly discovered in unicellular organisms (4, 6). Being more broadly distributed nature, the Fluc proteins have been investigated structurally, functionally, and mechanistically since their discovery (4, 6-12).

Crystallographic data (10) of two Fluc proteins, Ec2 and Bpe, has revealed its antiparallel homodimeric structure as well as a F^- binding site for each monomer (see **Figure 4.1**). This binding site discovered in crystal structures of both proteins (PDB ID: 5A43 for Ec2 and 5NKQ for Bpe) (10), is located in the center of the protein and surrounded by conserved residues N41 and F80 (Ec2 numbering). In the crystal structure of Ec2, the F^- anion is in proximity to the phenyl carbon of F80 and the amide group of N41, with a distance of 3.7 Å and 4.7 Å, respectively, implying possible interactions between them. A F83–H106 pair is also observed in the crystal structure, with a minimum distance of 3.4 Å between the phenyl group and the imidazole group. Note that a Na^+ cation is also detected in the interfacial region of two subunits.

Functionally, the Fluc channels have been identified as F^- channels with a turnover rate as high as 10^6–10^7 s^{-1} and >10^4-fold F^- over Cl^- selectivity (6, 9). The two subunits of the homodimer have also been shown to work independently by single-pore mutation experiments in

which each transport pathway can be individually inactivated without blocking the F^- transport through the other pathway (11). The allows us to focus the study of F^- transport on one monomer (monomer A in this work). Since one transport pathway is composed of residues from both monomers, we add an alphabet A or B after residue indices when necessary, with A being the same monomer surrounding the F^- ion and B being the other one.

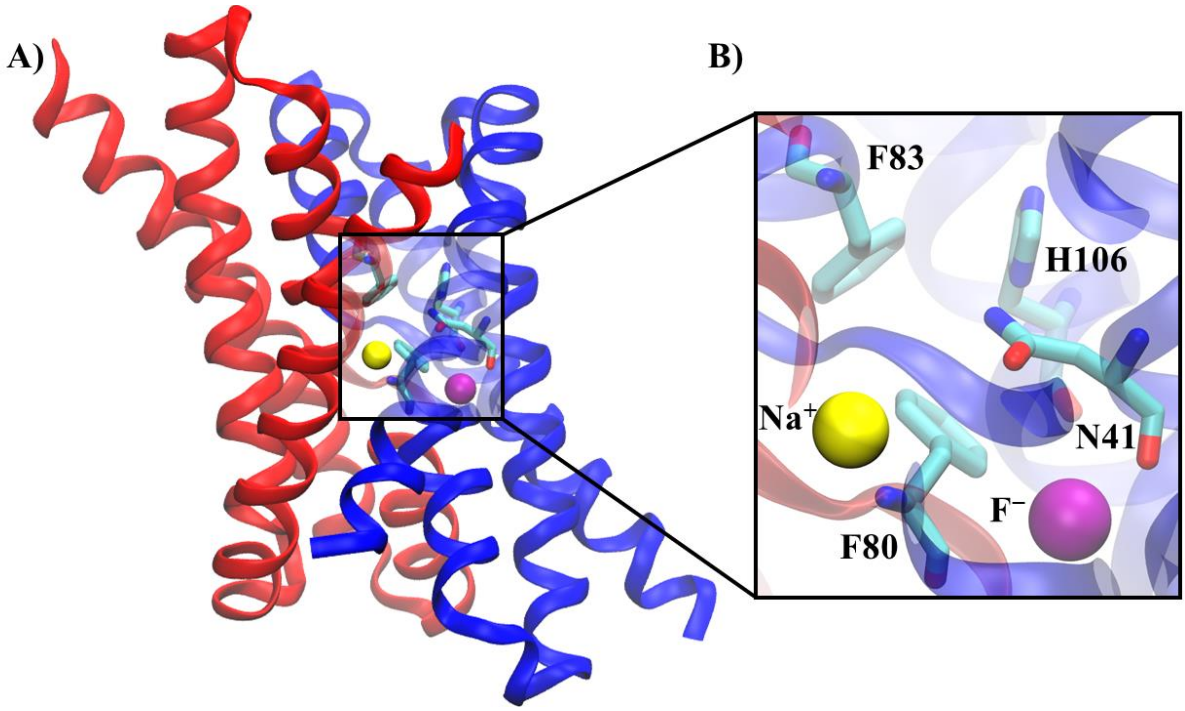

Figure 4.1 The crystal structure of Ec2 protein showing several mutation-sensitive residues (A) The crystal structure (PDB ID: 5A43) (10) of Ec2 (monomer A in blue and monomer B in red) with Na^+ (yellow sphere) and F^- (purple sphere) in monomer A and some key residues crucial to the F^- transport. (B) The zoomed in crystal structure of Ec2 with residues N41, F83, and H106 in monomer A and F80 in monomer B displayed and labeled. The protein-bound Na^+ and F^- are shown as yellow and purple spheres, respectively.

Mechanistically, several mutagenetic experiments have revealed the N41, F80, F83, and H106 residues being essential to the F^- transport (12). It is also revealed that the interfacial Na^+ is also a crucial component necessary for the F^- activity despite not being transported itself (13).

However, several questions regarding the transport mechanism remain unresolved. First of all, what is the form of fluoride in the protein? As a F83–H106 pair is depicted in the crystal structure (10), it is possible that the H106 will take a protonated state under the physiological pH condition due to the cation-π interaction. If this is the case, F^- will interact with the protonated H106 electrostatically and possibly be trapped in the protein due to the strong attraction, in which case F^- must transiently receive a proton to form a neutral HF and deprotonate afterwards to complete the transport cycle. Moreover, it is still unclear how the Na^+ and conserved residues affect the F^- transport at a molecular level. In addition, many uninvestigated conserved residues warrant further exploration to identify new residues crucial to the F^- transport.

Therefore, we aim at solving these unanswered questions by performing molecular dynamics (MD) simulations, but this can only be done after settling several issues. Firstly, an appropriate force field must be chosen carefully, because nonpolarizable force fields can hardly model the hydrogen bond between fluoride anion and hydrogen donor as well as the electrostatic interaction between F^- and Na^+, as the fixed charge description in these models does not account for induced electronic polarization. Thus, nonpolarizable models tend to overestimate the electrostatic interaction between fluoride and hydrogen-donating moieties and cations, and hence energy barriers are likely to be exaggerated (14). Several research has proved the importance of inclusion of induced polarization into the force field to model ion channels accurately (15-17). Secondly, several titratable residues must be designated with correct protonation states. Some empirical methods including the PROPKA method (18, 19) can predict the protonation states of these residues with almost no computational expense, but at the cost of accuracy. Such approach could generate erroneous results for essential residues including H106, and in turn mislead our understanding of the F^- protonation state and its transport mechanism.

In this work, we employed the Drude force field (20-25) with the cation-π correction (26) to model F^- transport in the Ec2 protein considering the balance between accuracy and computational expense. The Drude force field (20-25) has been widely utilized to simulate ion channels and transporters (14, 27-30), and it has been shown to improve the ion transport properties in aqueous bulk (31), the ion paring properties (32), and the free energy sampling in ion channels/transporters (14), without sacrificing computational efficiency (33). We then evaluated the effect of including polarization effect by comparing results from simulations using Drude force field and those employing additive nonpolarizable force fields.

We also employed the constant-pH molecular dynamics (CpHMD) simulations (34, 35) to identify protonation states of ionizable residues as well as the fluoride species. Surprisingly, F is found to be deprotonated along the channel without being transiently protonated, despite H106 in its charged state. We also demonstrate several F^- binding states involving essential interactions between F^- and residues N41, F80, F83, and H106, explaining their mutagenetic sensitivity. These findings help to explain the F^- transport mechanism in Ec2 as well as other Fluc proteins.

4.2 Methods

4.2.1 Equilibration with Nonpolarizable Molecular Dynamics

Generated from CHARMM-GUI *Membrane Builder* (36-42), the simulation system consisted of the Ec2 protein (PDB ID: 5A43) (10), 135 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine (POPE) lipids, 17 K^+ anions, 25 Cl^- anions, and ~7,000 TIP3P water molecules in a 71 Å × 71 Å × 82 Å box with periodic boundary conditions. Because the force

field parameters for F^- were not included in the standard CHARMM force field (43-46), two F^- ions identified in the crystal structure (10) were replaced with Cl^-. The structure was then equilibrated with the GROMACS 2019.3 software (47) for 100 ns.

The CHARMM-CMAP and CHARMM36m (43-46) force field was used to model the protein and lipids, respectively. Based on pK_a calculations performed by PROPKA3.1 (18, 19), E86 in both monomers were protonated, while all other residues were in their standard protonation states. The interval of leap-frog integration was 2 fs. LINear Constraint Solver (LINCS) (48) was used to restrain bonds involving hydrogen atoms to allow for a 2-fs step length. The neighbor list (cutoff = 12.0 Å) was updated every 40 fs. The particle mesh Ewald (PME) method (49, 50) was adopted to calculate long-range electrostatic interactions, with a spherical cutoff of 12.0 Å, while the short-range electrostatic potential was shifted by a constant. A switching function was used for van der Waals interactions in the range of 10.0 Å and 12.0 Å. All heavy atoms of the protein and two bound ions were restrained using an umbrella potential (force constant = 1.00 kcal mol^{-1} $Å^{-2}$). Other simulation details are consistent with our previous work (51, 52).

4.2.2 Constant-pH Molecular Dynamics Simulation Settings

Simulations were performed using CHARMM (38) (version c42b2) or GROMACS (47) (version 2019.3) simulation packages. All simulations used the CHARMM27/CMAP (43, 53), CHARMM36 (44), and CHARMM-modified (54) TIP3P (55) models to represent protein, lipids, and waters, respectively. The parameters for F^- and HF were reported by Senn *et al.* (56) and Laage *et al.* (57), respectively. The constant-pH molecular dynamics (CpHMD) parameters for HF were derived using the protocol described by Lee *et al*. (34). The leapfrog integrator was

used to propagate spatial coordinates. A switching function was applied to the van der Waals force from 8 to 12 Å. Electrostatic interactions were computed using the particle mesh Ewald (49, 50) method with a real-space cutoff of 12 Å. SHAKE (58) in CHARMM or LINCS (48) in GROMACS were used to restrain bonds involving hydrogen atoms to allow for a 2-fs step length. All simulations were conducted with periodic boundary conditions at constant temperature of 310 K by the Nosé-Hoover thermostat (59, 60), and constant pressure of 1 atm by the Langevin piston pressure-coupling algorithm (61) in CHARMM or the Parrinello-Rahman barostat (62) in GROMACS.

In the membrane-enabled (63) hybrid-solvent (64) continuous constant-pH molecular dynamics (CpHMD) simulations (34, 35), the conformational dynamics was propagated using explicit solvent and lipid molecules, while the electrostatic hydration forces on the titration coordinates were computed using implicit solvent and membrane described by the generalized-Born (GB) model GBSW (65, 66) with optimized GB input radii by Chen *et al.* (67). The implicit membrane was represented by an infinite low-dielectric-constant slab harboring a high-dielectric-constant exclusion cylinder that aligns with the Z-axis to mimic the protein region (66). The implicit-membrane thickness was 35 Å and the radius of the exclusion cylinder was 15 Å. The dielectric constant was switched from 2 to 80 within 2.5 Å for both leaflets. The ionic strength in the GB calculation was set to 0.150 M. Following Wallace *et al.* (64), the fictitious particles with a mass of 10 atomic mass unit were propagated using the Langevin algorithm with a collision frequency of 5 ps^{-1}, the titration coordinates were updated every 10 MD steps to allow for water relaxation. All Asp, Glu, His, Arg19 residues, and two HF ligands were allowed to titrate. When the pH-based replica exchange (pH-REX) (64) was applied, exchanges between neighboring replicas were attempted every 500 MD steps.

4.2.3 Constant-pH Molecular Dynamics Simulation Protocol

Simulations were conducted based on a protocol previously established (68, 69). The last frame from the nonpolarizable simulation was taken as the initial structure, with all Cl^- ions replaced back to F^- and a crystallographic Na^+ docked into the protein. Following the CHARMM-GUI protocol (36, 40), the system was minimized and relaxed in a stepwise fashion using GROMACS until the POPE bilayer was equilibrated. The last snapshot was then retrieved as the starting configuration for further equilibration at pH 7 using the CpHMD in CHARMM. During the CpHMD equilibration, a spherical restraint was applied via the MMFP utility in CHARMM (38) to the center-of-mass (COM) of the protein heavy atoms with a force constant of 50 kcal mol^{-1} $Å^{-2}$. Ions were excluded from the hydrophobic membrane region (–17.5 Å < Z < 17.5 Å) by a planar restraint of 1 kcal mol^{-1} $Å^{-2}$. The HF in chain B was restrained around its initial location by a spherical restraint of 2 kcal mol^{-1} $Å^{-2}$. A planar restraint of 10 kcal mol^{-1} $Å^{-2}$ was applied to the Z-component of the COM of the HF in chain A (ZCOM) to relocate it to the designated location along the membrane normal. 21 replicates of the starting configuration were thus generated by constraining the ZCOM from –15 Å to 15 Å with a 1.5 Å interval. The CpHMD equilibrations ran for 50 ns until the backbone RMSD of the Ec2 protein plateaued.

4.2.4 Replica-Exchange Umbrella Sampling Simulations

21 replicates of the Ec2 structure generated from the CpHMD equilibrations were retrieved and assigned protonation states (protonated His106, deprotonated F^-, and all other residues in their standard states; see section **4.3.1 Protonation States of Titratable Residues**)

using CHARMM (38). The number of cations is adjusted accordingly. With these replicates used as initial structures, replica-exchange umbrella sampling (REUS) (70) simulations were conducted to calculate potentials of mean force (PMFs) of F^- transport, using the NAMD simulation package (71).

In simulations using polarizable force field, the CHARMM Drude-2013 polarizable force field (20-25) was employed to model the protein (24), lipids (23, 25), ions (21, 22), and SWM4-NDP waters (20). Additional cation-π correction (26) was added to the standard CHARMM-Drude force field for better description of interactions between protonated His106 and Phe83. In nonpolarizable REUS simulations, the CHARMM-CMAP and CHARMM36m (43-46) force field was used to model the protein and lipids, respectively, and the CHARMM-modified (54) TIP3P (55) water model was adopted.

The MD time step was 0.25 fs for polarizable simulations and 2 fs for nonpolarizable simulations. All interactions, including bonded interactions, short-range nonbonded interactions, and long-range electrostatic interactions, were calculated each MD step. The particle mesh Ewald (PME) method (49) was adopted to calculate long-range electrostatic interactions, with a spherical cutoff of 12.0 Å and an accuracy threshold of 10^{-6}. The neighbor list consisting of atom pairs whose interactions were to be evaluated was updated every 10 MD step. A switching function was used for van der Waals and short-range electrostatic interactions in the range of 10.0 Å and 12.0 Å. The cutoff of neighbor list was set to 16.0 Å. The SETTLE algorithm (72) was used to restrain bonds involving hydrogen atoms. In the bulk regions, cylindrical restraints along the x–y plane were applied to allow the fluoride anion to explore a similar area to that within the protein, similar to our previous work (73). Other simulation details are consistent with our previous work (51, 52).

In the replica-exchange umbrella sampling (REUS) simulations, the collective variable (CV) was the z-component of the vector from the center of mass of α-carbon atoms of the protein to the F^- ion. The weighted histogram analysis method (WHAM) (74, 75) was utilized to combine the potential of mean forces (PMFs) generated from the separate windows of REUS simulations. The error bars on the PMFs were estimated by dividing each trajectory, generating PMFs separately, and calculating the standard deviation. PMFs were processed and plotted with the help of GNU Parallel (76) and Matplotlib (77). The protein structures were rendered in VMD (78).

4.2.5 Calculation of Fluoride Diffusion Constant and Conductance

The effective rate constant k for the F^- transport was calculated from the PMF $F(z)$ based on the mean first passage time (79):

$$k = \left\{ \int_{z_a}^{z_b} dz' D(z')^{-1} \exp[F(z')/k_B T] \int_{z_a}^{z'} dz'' \exp[-F(z'')/k_B T] \right\}^{-1}, \quad (4.1)$$

where $z_a = -20$ Å and $z_b = 20$ Å are the entrance and the exit of the protein, k_B is the Boltzmann constant, T is the temperature, D is the diffusion constant given by the Woolf-Roux equation (80):

$$D(z_i) = \lim_{s \to 0} \frac{-\hat{C}(s;z_i)\langle \delta z^2 \rangle_{(i)} \langle \dot{z}^2 \rangle_{(i)}}{\hat{C}(s;z_i)[s\langle \delta z^2 \rangle_{(i)} + \langle \dot{z}^2 \rangle_{(i)}/s] - \langle \delta z^2 \rangle_{(i)} \langle \dot{z}^2 \rangle_{(i)}}, \quad (4.2)$$

where $\delta z = z - z_i$ is the deviation from the center, $C(t; z_i) = \langle \dot{z}(t)\dot{z}(0) \rangle_i$ is the position-dependent velocity autocorrelation function for the F^- transport along z-axis, calculated based on the data from window *i*, z_i is the reference point for the harmonic restraint potential in window *i*,

and $\hat{C}(s;z_i) = \int_0^\infty e^{-st} C(t;z_i)dt$ is the Laplace transform of this function. To estimate the value of the limit as $s \to 0$, we linearly extrapolate from the range 0.2 (MD step)$^{-1} \leq s \leq$ 0.3 (MD step)$^{-1}$. The error of the diffusion constant of F^- in bulk was estimated by averaging multiple windows in bulk region and calculating the standard deviation. The rate constant error was estimated from the standard deviation of 6 rate constants given by divided trajectories.

4.3 Results and Discussion

4.3.1 Protonation States of Titratable Residues and Ions

We performed a 20-ns pH-REX CpHMD simulation to determine the charge states of the relevant ionizable residues and ligands in the Ec2 protein. Asp, Glu, His residues and the two HF ligands were made titratable in the simulation. Two Arg19 residues were also allowed to ionize because they are chelated with the positively charged protein-bound Na^+. The resulting predicted pK_a's are listed in **Table 4.1**. The table indicates around physiological pH condition, HFs are deprotonated, *i.e.*, F^-, when they stay inside the PDB binding pocket. The ionizable gating residues His106s have pK_a's at least 1.5 pH units above the physiologic pH and thus less likely to deprotonate during F^- transfer. Hence, the following REUS simulations were performed with protonated His106, deprotonated F^-, and all other residues in their standard protonation states.

Table 4.1 The pK_a values predicted by pH-REX CpHMD
pH from 1.0–10.5, 20 ns per replica

Residue	pK_a	
	Chain A	Chain B
Arg19	Stay protonated	
His60	7.3	7.4
Asp62	3.4	3.7
Glu86	3.9	3.8
His106	8.6	9.6
HF	Stay deprotonated	

4.3.2 Comparison between Polarizable and Nonpolarizable Simulations

With protonation states determined by CpHMD simulations, we employed REUS to sample F^- transport across the Ec2 protein, using either polarizable CHARMM-Drude force field (20-25) with the cation-π correction (26) or nonpolarizable CHARMM force field (43-46, 54).

Table 4.2 The diffusion constant D of F^- in bulk and the rate constant k of F^- transport calculated from REUS simulations

Model	D (10^{-5} cm^2 s^{-1})	k (s^{-1})
CHARMM-Drude	1.4 ± 0.2	(4 ± 1) × 10^6
Additive CHARMM36m	2.1 ± 0.1	0.29 ± 0.26
Experiment	1.4 ± 0.1	10^6–10^7

Based on the Woolf-Roux equation (80), the position-dependent diffusion constant of F^- was calculated from our REUS simulations. The diffusion constant in the bulk region is estimated to be (1.4 ± 0.2) × 10^{-5} cm^2 s^{-1} from the polarizable simulations and (2.1 ± 0.1) × 10^{-5} cm^2 s^{-1} from nonpolarizable simulations (**Table 4.2**). The experimental F^- diffusion constant in infinite dilute aqueous solution is 1.475 × 10^{-5} cm^2 s^{-1} (81), and its diffusion constant in aqueous bulk is estimated to be (1.4 ± 0.1) × 10^{-5} cm^2 s^{-1} considering the finite concentration (82). Therefore, our results indicate that polarizable simulations provide the correct F^- diffusivity while the nonpolarizable simulations overestimate the diffusion constant, which is consistent with Prajapati *et al.*'s research on calculating ion conductivity for aqueous solutions lower than 1 M using either polarizable or nonpolarizable force field (31).

The PMFs (**Figure 4.2**) were then constructed from both simulations. They have similar shapes but differ significantly in energetics. The PMF from polarizable simulations shows a 7.4 ± 0.3 kcal mol^{-1} free energy barrier for F^- transport from the global minimum to the external bulk, while the one from nonpolarizable simulations indicates that the barrier is 17.7 ± 0.8 kcal

mol^{-1}. Based on the mean first passage time (MFPT) model (79), the rate constant for the F^- transport is estimated to be $(4 \pm 1) \times 10^6$ s^{-1} from polarizable simulations, and (0.29 ± 0.26) s^{-1} from nonpolarizable simulations, the former of which is consistent with the experimental F^- turnover rate of 10^6–10^7 s^{-1} (6, 9) (see **Table 4.2**). Hence, both diffusion constant and rate constant indicate that the polarizable Drude model better describes F^- transport in the Ec2 protein, and our further discussions will solely rely on the Drude simulations.

4.3.3 Fluoride Transport Mechanism and the Fluoride Binding States

The F^- transport PMF (**Figure 4.2**) reveals several metastable F^- binding states (**Figure 4.3**), which are (with increasing z coordinate) (A) F^- bound with Na^+ and stabilized by R19B, (B and C) F^- bound with F80B side chain and alternatively bound with N41A/H106A side chains, (D) F^- shuttled above through F80B side-chain rotation and located at the bottom of the F83A–H106A gate, and (E) F^- passing through the gate and stabilized by R19A and T37A. These binding states involve several mutagenetic sensitive residues, including N41A, F80B, F83A, and H106A (12), and explain how these residues are crucial to the F^- transport. Notably, the state B is similar to the binding state in the crystal structure (**Figure 4.1**) (10), which features the interaction between F^- and N41A/F80B.

The transition from state A to state E explains the F^- transport mechanism, featuring a strong F^-–Na^+ interaction, a phenyl C–H···F hydrogen bond, a H106A side-chain rotation, and a F83A–H106A gate. In the binding state A (**Figure 4.3** (A)), the F^- interacts with the protein-bound Na^+, R19B, and F80B backbone. In both PMFs (**Figure 4.2**), this state is shown as a global energy minimum, which should be attributable to the strong electrostatic interaction between F^- and these electropositive moieties. This also explains the experimental results

implying that the interfacial Na^+ is an essential structural component of the Fluc channels necessary for the $F^–$ transport despite not itself transported (13). Notably, the energy stabilization is overestimated in the nonpolarizable PMF, leading to a much deeper energy well compared with the polarizable one. Since both Na^+ and $F^–$ are monoatomic ions and modeled as point charges in the nonpolarizable model, the polarization effect is completely ignored, and thus the electrostatic attraction is overestimated. It further illustrates the necessity of employing a polarizable model to accurately model $F^–$ transport in a Na^+-bound protein.

In state B, C, and D (**Figure 4.3** (B)-(D)), the interaction between F80B and $F^–$ is revealed. The representative configurations show a phenyl C–H···F hydrogen bond stabilizing the binding of $F^–$ to the protein. This hydrogen bond explains the experimental measurement of the missing $F^–$ transport activity when F80 is mutated (11, 12). Such interaction is captured in the crystal structure as well (10), but our simulations provide dynamic information of the hydrogen bond and indicate that it persists in a wide CV range (CV = –3 – 2 Å).

The interaction between $F^–$ and His106A is also one of the stabilization effects crucial to the $F^–$ binding, as is shown in both state C and D (**Figure 4.3** (C) and (D)). It does not only include a N–H···F hydrogen bond, but also involves electrostatic attractions between opposite charges as the histidine is protonated. As the $F^–$ is transported along the transport pathway, the His106A residue must be reoriented, indicated by the χ_1 dihedral angel shift from (86 ± 10)° (state C) to (–78 ± 8)° (state D), averaged from the trajectories.

In the binding state D (**Figure 4.3** (D)), $F^–$ coordinates with two phenyl C–H groups (F80B and F83A), one imidazolium N–H group (H106A), and an amide N–H group (N41A). These interactions make the binding state D merely 0.6 kcal $mol^{–1}$ less stable than the global minimum state A. Among these residues, the F83A and H106A form a gate through a cation-π

interaction, which causes a barrier of 3.7 kcal mol^{-1} for F^- to pass through. Such cation-π interaction is more accurately modeled in the polarizable force field thanks to Lin and MacKerell's recent work (26).

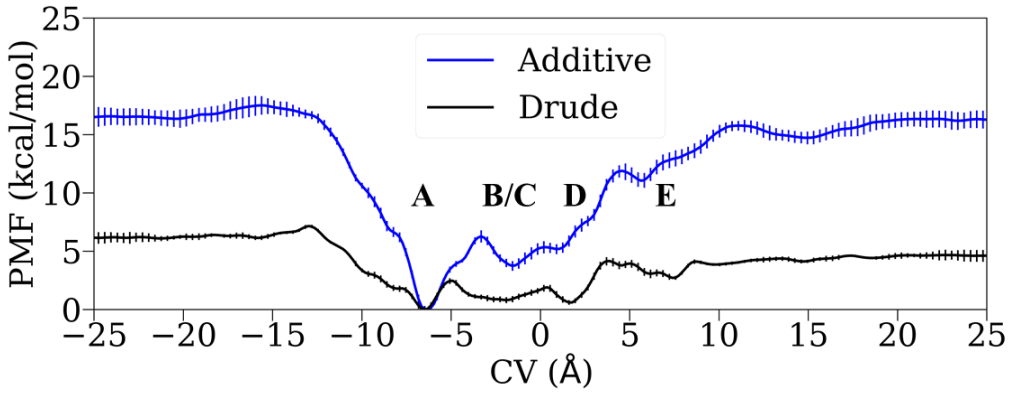

Figure 4.2 The PMF for F^- transport from Drude/Additive REUS simulations
The CV is defined as the z-component of the vector from the center of mass of α-carbon atoms of the protein to the F^- ion. The alphabets A-E indicates F^- binding states illustrated in **Figure 4.3**.

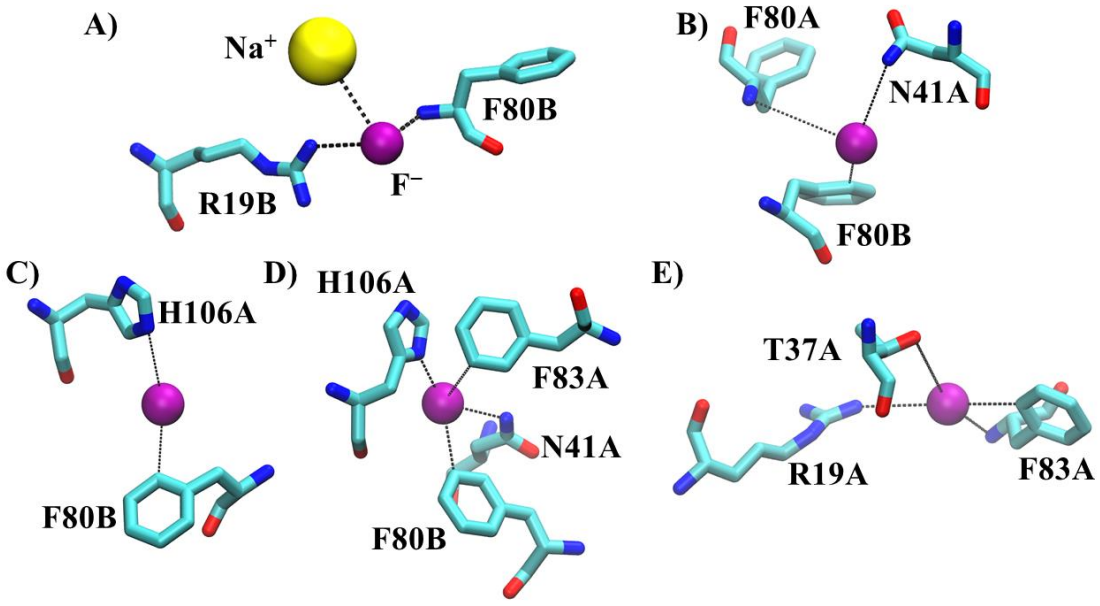

Figure 4.3 Representative configurations of F^- transport across the Ec2 protein.
The panel label is consistent with **Figure 4.2**. The protein-bound F^- and Na^+ are shown as purple and yellow spheres, respectively. Key residues coordinating with F^- are also indicated. The letter following the residue number indicates the monomer to which the residue belongs to, with A being the monomer surrounding the F^- shown in the figure, and B being the other monomer.

Overall, the calculated PMF and observed F^- binding states emphasize the interactions between F^- and various residues as well as Na^+ through electrostatic interactions and hydrogen bonds. Our results show the importance of employing polarizable force field to model F^- channels. More importantly, our Drude REUS simulations demonstrate that the protonated His106 and the Na^+ cation do not trap the electronegative F^- in the binding pockets, and that it is feasible to transfer F completely in its charged state, as F^- in lieu of HF (*i.e.*, no intermediate HF is observed), which is consistent with the pK_a value estimated from the pH-REX CpHMD simulation.

4.4 Conclusions

We performed constant-pH molecular dynamics simulations and polarizable replica-exchange umbrella sampling simulations to investigate the fluoride transport mechanism of Ec2 fluoride channel. The pH-REX CpHMD simulation was first conducted to calculate pK_a values of ionizable residues and ions, confirming that His106 stays protonated, F^- is in its charged form, and all other titratable residues are in their standard states. We followed up with REUS simulations using both nonpolarizable force field and polarizable force field. The resulting diffusion constant as well as rate constant from polarizable simulations matches the experimental measurements (6, 9, 81, 82) better, implying the necessity of utilizing the polarizable force field, which should be attributable to the overestimated interaction between F^- and electropositive moieties under the nonpolarizable description, implied from the much deeper well in the nonpolarizable PMF.

We then discovered and discussed on the F^- binding states, showing key elements in the F^- transport mechanism. These binding states underline the electrostatic attraction between F^- and Na^+, the Phe C–H···F hydrogen bonds, the H106A rotamer switch, and the F83A–H106A gate. These interactions do not only explain mutagenetic experiments on residues N41, F80, F83, and H106 (12) as well as the drop in F^- transport activity in the absence of a binding Na^+ (13), but also provide new insights on following mutagenetic experiments on conserved residues R19 and T37. Notably, the residue R19 in a yeast fluoride channel Fex1p (numbered as R33) was shown to be essential to the fluoride activity (9), but no experiments on this residue have so far been performed to my knowledge for the Ec2 Fluc protein. In addition, our simulations provide dynamic information of the hydrogen bonds discovered in the crystal structure and indicate that they persist as F^- is transported. As a final point, our Drude REUS simulations eliminate the possibility of F^- being transiently protonated by showing the feasibility of F^- transport in its charged state.

Despite providing atomistic details of mutation-sensitive residues, our study does not directly tackle the mutations. Our future research will include a more complete investigation of the mutants of the protein using a polarizable force field. Moreover, the extreme F^-/Cl^- selectivity (6) is yet to be investigated, which is also one of our next tasks.

4.5 References

1. Weinstein L & Davison A (2004) *Fluorides in the Environment: Effects on Plants and Animals* (CABI Publishing, Cambridge, MA).
2. Marquis RE, Clock SA, & Mota-Meira M (2003) Fluoride and organic weak acids as modulators of microbial physiology. *FEMS Microbiol. Rev.* 26(5):493–510.

3. Samygina VR, *et al.* (2007) Reversible Inhibition of *Escherichia coli* Inorganic Pyrophosphatase by Fluoride: Trapped Catalytic Intermediates in Cryo-crystallographic Studies. *J. Mol. Biol.* 366(4):1305–1317.

4. Baker JL, *et al.* (2012) Widespread Genetic Switches and Toxicity Resistance Proteins for Fluoride. *Science* 335(6065):233–235.

5. Stockbridge RB, *et al.* (2012) Fluoride resistance and transport by riboswitch-controlled CLC antiporters. *Proc. Natl. Acad. Sci. U. S. A.* 109(38):15289–15294.

6. Stockbridge RB, Robertson JL, Kolmakova-Partensky L, & Miller C (2013) A family of fluoride-specific ion channels with dual-topology architecture. *eLife* 2:e01084.

7. Stockbridge RB, Koide A, Miller C, & Koide S (2014) Proof of dual-topology architecture of Fluc F^- channels with monobody blockers. *Nat. Commun.* 5(1):5120.

8. Turman DL, Nathanson JT, Stockbridge RB, Street TO, & Miller C (2015) Two-sided block of a dual-topology F^- channel. *Proc. Natl. Acad. Sci. U. S. A.* 112(18):5697−5701.

9. Smith KD, *et al.* (2015) Yeast Fex1p Is a Constitutively Expressed Fluoride Channel with Functional Asymmetry of Its Two Homologous Domains. *J. Biol. Chem.* 290(32):19874–19887.

10. Stockbridge RB, *et al.* (2015) Crystal structures of a double-barrelled fluoride ion channel. *Nature* 525(7570):548−551.

11. Last NB, Kolmakova-Partensky L, Shane T, & Miller C (2016) Mechanistic signs of double-barreled structure in a fluoride ion channel. *eLife* 5:e18767.

12. Last NB, Sun S, Pham MC, & Miller C (2017) Molecular determinants of permeation in a fluoride-specific ion channel. *eLife* 6:e31259.

13. McIlwain BC, Martin K, Hayter EA, & Stockbridge RB (2020) An Interfacial Sodium Ion is an Essential Structural Feature of Fluc Family Fluoride Channels. *J. Mol. Biol.* 432(4):1098−1108.

14. Sun R-N & Gong H (2017) Simulating the Activation of Voltage Sensing Domain for a Voltage-Gated Sodium Channel Using Polarizable Force Field. *J. Phys. Chem. Lett.* 8(5):901–908.

15. Roux B (1993) Non-additivity in cation–peptide interactions. A molecular dynamics and ab initio study of Na^+ in the gramicidin channel. *Chem. Phys. Lett.* 212(3):231–240.

16. Allen TW, Andersen OS, & Roux B (2006) Ion Permeation through a Narrow Channel: Using Gramicidin to Ascertain All-Atom Molecular Dynamics Potential of Mean Force Methodology and Biomolecular Force Fields. *Biophys. J.* 90(10):3447–3468.

17. Bucher D, Guidoni L, Maurer P, & Rothlisberger U (2009) Developing Improved Charge Sets for the Modeling of the KcsA K^+ Channel Using QM/MM Electrostatic Potentials. *J. Chem. Theory Comput.* 5(8):2173–2179.

18. Olsson MHM, Søndergaard CR, Rostkowski M, & Jensen JH (2011) PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pK_a Predictions. *J. Chem. Theory Comput.* 7(2):525–537.

19. Søndergaard CR, Olsson MHM, Rostkowski M, & Jensen JH (2011) Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pK_a Values. *J. Chem. Theory Comput.* 7(7):2284–2295.

20. Lamoureux G, Harder E, Vorobyov IV, Roux B, & MacKerell AD, Jr. (2006) A polarizable model of water for molecular dynamics simulations of biomolecules. *Chem. Phys. Lett.* 418(1):245–249.

21. Yu H, *et al.* (2010) Simulating Monovalent and Divalent Ions in Aqueous Solution Using a Drude Polarizable Force Field. *J. Chem. Theory Comput.* 6(3):774–786.

22. Luo Y, Jiang W, Yu H, MacKerell AD, Jr., & Roux B (2013) Simulation study of ion pairing in concentrated aqueous salt solutions with a polarizable force field. *Faraday Discuss.* 160:135–149.

23. Chowdhary J, *et al.* (2013) A Polarizable Force Field of Dipalmitoylphosphatidylcholine Based on the Classical Drude Model for Molecular Dynamics Simulations of Lipids. *J. Phys. Chem. B* 117(31):9142–9160.

24. Lopes PEM, *et al.* (2013) Polarizable Force Field for Peptides and Proteins Based on the Classical Drude Oscillator. *J. Chem. Theory Comput.* 9(12):5430–5449.

25. Li H, *et al.* (2017) Drude Polarizable Force Field for Molecular Dynamics Simulations of Saturated and Unsaturated Zwitterionic Lipids. *J. Chem. Theory Comput.* 13(9):4535–4552.

26. Lin F-Y & MacKerell AD, Jr. (2020) Improved Modeling of Cation-π and Anion-Ring Interactions Using the Drude Polarizable Empirical Force Field for Proteins. *J. Comput. Chem.* 41(5):439–448.

27. Kratochvil HT, *et al.* (2016) Instantaneous ion configurations in the K^+ ion channel selectivity filter revealed by 2D IR spectroscopy. *Science* 353(6303):1040−1044.

28. Dhakshnamoorthy B, Rohaim A, Rui H, Blachowicz L, & Roux B (2016) Structural and functional characterization of a calcium-activated cation channel from *Tsukamurella paurometabola*. *Nat. Commun.* 7(1):12753.

29. Vergara-Jaque A, Fong P, & Comer J (2017) Iodide Binding in Sodium-Coupled Cotransporters. *J. Chem. Inf. Model.* 57(12):3043–3055.

30. Mayes HB, Lee S, White AD, Voth GA, & Swanson JMJ (2018) Multiscale Kinetic Modeling Reveals an Ensemble of Cl^-/H^+ Exchange Pathways in ClC-ec1 Antiporter. *J. Am. Chem. Soc.* 140(5):1793–1804.

31. Prajapati JD, Mele C, Aksoyoglu MA, Winterhalter M, & Kleinekathöfer U (2020) Computational Modeling of Ion Transport in Bulk and through a Nanopore Using the Drude Polarizable Force Field. *J. Chem. Inf. Model.* 60(6):3188–3203.

32. Flood E, Boiteux C, Lev B, Vorobyov I, & Allen TW (2019) Atomistic Simulations of Membrane Ion Channel Conduction, Gating, and Modulation. *Chem. Rev.* 119(13):7737–7832.

33. Jiang W, *et al.* (2011) High-Performance Scalable Molecular Dynamics Simulations of a Polarizable Force Field Based on Classical Drude Oscillators in NAMD. *J. Phys. Chem. Lett.* 2(2):87–92.

34. Lee MS, Salsbury FR, Jr., & Brooks CL, III. (2004) Constant-pH Molecular Dynamics Using Continuous Titration Coordinates. *Proteins: Struct., Funct., Bioinf.* 56(4):738–752.

35. Khandogin J & Brooks CL, III. (2005) Constant pH Molecular Dynamics with Proton Tautomerism. *Biophys. J.* 89(1):141–157.

36. Jo S, Kim T, & Im W (2007) Automated Builder and Database of Protein/Membrane Complexes for Molecular Dynamics Simulations. *PLoS One* 2(9):e880.

37. Jo S, Kim T, Iyer VG, & Im W (2008) CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM. *J. Comput. Chem.* 29(11):1859–1865.

38. Brooks BR, *et al.* (2009) CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* 30(10):1545–1614.

39. Jo S, Lim JB, Klauda JB, & Im W (2009) CHARMM-GUI Membrane Builder for Mixed Bilayers and Its Application to Yeast Membranes. *Biophys. J.* 97(1):50–58.

40. Wu EL, *et al.* (2014) CHARMM-GUI *Membrane Builder* toward realistic biological membrane simulations. *J. Comput. Chem.* 35(27):1997–2004.

41. Lee J, *et al.* (2016) CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* 12(1):405–413.

42. Lee J, *et al.* (2019) CHARMM-GUI *Membrane Builder* for Complex Biological Membrane Simulations with Glycolipids and Lipoglycans. *J. Chem. Theory Comput.* 15(1):775–786.

43. MacKerell AD, Jr., Feig M, & Brooks CL, III. (2004) Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum

Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulations. *J. Comput. Chem.* 25(11):1400–1415.

44. Klauda JB, *et al.* (2010) Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *J. Phys. Chem. B* 114(23):7830–7843.

45. Best RB, *et al.* (2012) Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ, ψ and Side-Chain χ_1 and χ_2 Dihedral Angles. *J. Chem. Theory Comput.* 8(9):3257–3273.

46. Huang J, *et al.* (2017) CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* 14(1):71–73.

47. Abraham MJ, *et al.* (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 1–2:19–25.

48. Hess B, Bekker H, Berendsen HJC, & Fraaije JGEM (1997) LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* 18(12):1463–1472.

49. Darden T, York D, & Pedersen L (1993) Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* 98(12):10089–10092.

50. Essmann U, *et al.* (1995) A smooth particle mesh Ewald method. *J. Chem. Phys.* 103(19):8577–8593.

51. Wang Z, Swanson JMJ, & Voth GA (2018) Modulating the Chemical Transport Properties of a Transmembrane Antiporter via Alternative Anion Flux. *J. Am. Chem. Soc.* 140(48):16535–16543.

52. Wang Z, Swanson JMJ, & Voth GA (2020) Local Conformational Dynamics Regulating Transport Properties of a Cl^-/H^+ Antiporter. *J. Comput. Chem.* 41(6):513–519.

53. MacKerell AD, Jr., *et al.* (1998) All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* 102(18):3586–3616.

54. Durell SR, Brooks BR, & Ben-Naim A (1994) Solvent-Induced Forces between Two Hydrophilic Groups. *J. Phys. Chem.* 98(8):2198–2202.

55. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, & Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79(2):926–935.

56. Senn HM, O'Hagan D, & Thiel W (2005) Insight into Enzymatic C−F Bond Formation from QM and QM/MM Calculations. *J. Am. Chem. Soc.* 127(39):13643–13655.

57. Laage D, Demirdjian H, & Hynes JT (2005) Intermolecular vibration–vibration energy transfer in solution: Hydrogen fluoride in water. *Chem. Phys. Lett.* 405(4):453-458.

58. Ryckaert J-P, Ciccotti G, & Berendsen HJC (1977) Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of *n*-Alkanes. *J. Comput. Phys.* 23(3):327–341.

59. Nosé S (1984) A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* 52(2):255–268.

60. Hoover WG (1985) Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* 31(3):1695–1697.

61. Feller SE, Zhang Y, Pastor RW, & Brooks BR (1995) Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys.* 103(11):4613–4621.

62. Parrinello M & Rahman A (1981) Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* 52(12):7182–7190.

63. Chen W, Huang Y, & Shen J (2016) Conformational Activation of a Transmembrane Proton Channel from Constant pH Molecular Dynamics. *J. Phys. Chem. Lett.* 7(19):3961–3966.

64. Wallace JA & Shen JK (2011) Continuous Constant pH Molecular Dynamics in Explicit Solvent with pH-Based Replica Exchange. *J. Chem. Theory Comput.* 7(8):2617–2629.

65. Im W, Lee MS, & Brooks CL, III. (2003) Generalized Born Model with a Simple Smoothing Function. *J. Comput. Chem.* 24(14):1691–1702.

66. Im W, Feig M, & Brooks CL, III. (2003) An Implicit Membrane Generalized Born Theory for the Study of Structure, Stability, and Interactions of Membrane Proteins. *Biophys. J.* 85(5):2900–2918.

67. Chen J, Im W, & Brooks CL, III. (2006) Balancing Solvation and Intramolecular Interactions: Toward a Consistent Generalized Born Force Field. *J. Am. Chem. Soc.* 128(11):3728–3736.

68. Yue Z, Chen W, Zgurskaya HI, & Shen J (2017) Constant pH Molecular Dynamics Reveals How Proton Release Drives the Conformational Transition of a Transmembrane Efflux Pump. *J. Chem. Theory Comput.* 13(12):6405–6414.

69. Yue Z, Li C, Voth GA, & Swanson JMJ (2019) Dynamic Protonation Dramatically Affects the Membrane Permeability of Drug-like Molecules. *J. Am. Chem. Soc.* 141(34):13421–13433.

70. Sugita Y, Kitao A, & Okamoto Y (2000) Multidimensional replica-exchange method for free-energy calculations. *J. Chem. Phys.* 113(15):6042–6051.

71. Phillips JC, *et al.* (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26(16):1781–1802.

72. Miyamoto S & Kollman PA (1992) SETTLE: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models. *J. Comput. Chem.* 13(8):952–962.

73. Liang R, Li H, Swanson JMJ, & Voth GA (2014) Multiscale simulation reveals a multifaceted mechanism of proton permeation through the influenza A M2 proton channel. *Proc. Natl. Acad. Sci. U. S. A.* 111(26):9396–9401.

74. Kumar S, Bouzida D, Swendsen RH, Kollman PA, & Rosenberg JM (1992) The Weighted Histogram Analysis Method for Free‐Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* 13(8):1011–1021.

75. Grossfield A ("WHAM: the weighted histogram analysis method", version 2.0.9.

76. Tange O (2011) GNU Parallel: The Command-Line Power Tool. *;login: The USENIX Magazine* 36(1):42–47.

77. Hunter JD (2007) Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* 9(3):90–95.

78. Humphrey W, Dalke A, & Schulten K (1996) VMD: Visual molecular dynamics. *J. Mol. Graphics* 14(1):33–38.

79. Szabo A, Schulten K, & Schulten Z (1980) First passage time approach to diffusion controlled reactions. *J. Chem. Phys.* 72(8):4350–4357.

80. Woolf TB & Roux B (1994) Conformational Flexibility of *o*-Phosphorylcholine and *o*-Phosphorylethanolamine: A Molecular Dynamics Study of Solvation Effects. *J. Am. Chem. Soc.* 116(13):5916–5926.

81. Vanýsek P (2006) IONIC CONDUCTIVITY AND DIFFUSION AT INFINITE DILUTION. *CRC Handbook of Chemistry and Physics*, ed Lide DR (Taylor and Francis, Boca Raton, FL), Internet Version 2006 Ed, pp 5:76–78.

82. Wang JH (1954) EFFECT OF IONS ON THE SELF-DIFFUSION AND STRUCTURE OF WATER IN AQUEOUS ELECTROLYTIC SOLUTIONS. *J. Phys. Chem.* 58(9):686–692.

CHAPTER 5

CONCLUDING REMARK

A series of work have been done to investigate the ion transport mechanism in the ClC-ec1 Cl^-/H^+ exchanger as well as the Ec2 fluoride channel at a molecular level.

In CHAPTER 2, extensive free energy sampling using multiscale reactive molecular dynamics (MS-RMD) method has been performed to investigate the influence of polyatomic anions on proton transport (PT) and the related mechanisms of ion coupling and stoichiometric ion exchange in ClC-ec1. Firstly, the relative stability and transition rates between two binding modes of the S_{cen} NO_3^- and SCN^- anions were quantified. It was found that the anions are stable for at least the submicrosecond time scale in the central site, and that separate treatment of the two thiocyanate binding modes was necessary due to the slow interconversion while nitrate rotates relatively quickly making it possible to sufficiently sample this transition during a single PT potential of mean force (PMF) calculation. This difference between NO_3^- and SCN^- is explained by the increased steric restriction of the bulky thiocyanate. A new reaction coordinate to characterize PT between E203 and E148 was then defined, which incorporates the excess proton center of excess charge (CEC) and the distance between the two glutamate residues simultaneously. This new collective variable (CV) significantly reduces the computational demands but does not introduce new spurious effects on the PMF. Following that, hybrid quantum mechanics/molecular mechanics (QM/MM) simulations verified that the polyatomic anions do not react with proton directly during the PT process in the protein, and hence do not require explicit MS-RMD treatment. 2D PMFs for the PT across the central region were then constructed, one for NO_3^- and two for SCN^-. The rate constants calculated from the PMFs

utilizing transition state theory are consistent with the experimental measurements, in which PT is partly uncoupled to the nitrate transport and completely uncoupled to thiocyanate transport. A previously published hypothesis about the water structure being influenced by the polyatomic anions was confirmed but significantly expanded to a three-phase mechanism of proton shuttling, on the basis of quantitative PT free energy calculations. Interestingly, these PMFs revealed that simply increasing the number of water molecules in the region of PT is not sufficient to enable PT. The critical hydration CV seems to be water network connectivity, which is demonstrated by analyzing the probability of water connectivity and the ratio of the length of the disconnected area to the distance between two glutamates. Both of these measures are highly coupled with PT through the transition state, as indicated by the correlation with the 1D PMFs for all three systems (one for NO_3^- and two for SCN^-). A continuous water-wire with an explicit excess proton in it is in fact able to exist in the NO_3^-/SCN^--bound protein, albeit transiently and conditionally, contrary to the hypothesis arising from a prior standard empirical MD simulation in which the explicit PT was not treated. In-depth analysis on the PT phase reveals three stages for the process. This mechanism demonstrates that it is not primarily the hydration of the cavity that is rate limiting, but rather forming a connected hydrogen bonded pathway past the steric hindrance of the polyatomic anions. This provides insightful contrast to the common assumption that blocking water influx pathways is an effective way to inhibit PT (or an explanation of inhibited PT). Collectively this work shows how seemingly subtle changes in the chemical nature of the anion can flip it from facilitating coupled proton exchange, to inhibiting or even blocking it in both directions. The shift in coupling changes the associated ion exchange stoichiometries, instead of completely blocking anion transport, because the rate and ease (*i.e.*, pK_a) of protonating E148, which is essential for anion flux, from the external bulk is fast and minimally

influenced by the presence of the anions. Rather, PT is hindered/blocked in the central cavity between E148 and E203. The demonstrated entanglement between PT, water dynamics, and the local environment is expected to play a significant role in a number of other proton-conducting proteins. These insights also potentially suggest guidelines for one way of controlling the function of the ClC family of proteins by influencing hydration, either through tuning the spatial constriction of the central cavity via binding, or by adjusting the steric hindrance caused by the channel-forming residues.

In CHAPTER 3, multiscale reactive MD simulations with enhanced free energy sampling were performed to investigate the mechanism of I109F mutation slowing down PT in ClC-ec1. With both the mutation and PT treated explicitly, a 2D PMF for the PT process across the central region past the central Cl^- anion binding site (S_{cen}) was constructed. Based on transition state theory, the rate constant estimated from the PMF is consistent with the experimental measurements, where the PT rate is decreased by an order of magnitude to ~1.0×10^2 s^{-1}. Our result is also consistent with the decreased Cl^- transport rate, since protonated E148, which is essential for Cl^- transport, is destabilized in our simulations. By comparing results in this work with previous work on the nitrate/thiocyanate-bound ClC-ec1 (CHAPTER 2), a similar hydration–PT–dehydration stepwise mechanism is confirmed, but in this case the controlling factor is the conformation of the phenylalanine gate formed by F109 and F357. The correlation plot shows the relationship between the tip distance of the gate and the water connectivity. With these results combined, it is concluded that the formation and stabilization of the Phe gate disrupt water connection and hence the proton transport across the central region. This is an alternative approach to regulate PT, compared to the previously reported regulation by the chemical nature of the transported anion. Thus, this work demonstrates how subtle local conformational

dynamics can significantly influence the ease of forming hydration networks and thus the rates of charge transport. Such subtle conformational dynamics could be an important factor distinguishing a ClC antiporter from a ClC channel, and are certainly influential regulators of ClC functionality.

In CHAPTER 4, constant-pH molecular dynamics (CpHMD) simulations and polarizable replica-exchange umbrella sampling (REUS) simulations were performed to investigate the fluoride transport mechanism of Ec2 fluoride channel. The CpHMD simulation was first conducted to calculate pK_a values of ionizable residues and ions, confirming that His106 stays protonated, F^- is in its charged form, and all other titratable residues are in their standard states. Following up, REUS simulations using both nonpolarizable force field and polarizable force field were performed. Simulations with the polarizable force field result in diffusion constant as well as rate constant consistent the experimental measurements while ones with the nonpolarizable force field fail to match these experimental results, implying the necessity of utilizing the polarizable force field, which should be attributable to the overestimated interaction between F^- and electropositive moieties under the nonpolarizable description, implied from the much deeper well in the nonpolarizable PMF. The F^- binding states are then discovered, showing key elements in the F^- transport mechanism. These binding states underline the electrostatic attraction between F^- and Na^+, the Phe C–H···F hydrogen bonds, the H106A rotamer switch, and the F83A–H106A gate. These interactions do not only explain mutagenetic experiments on residues N41, F80, F83, and H106 as well as the drop in F^- transport activity in the absence of a binding Na^+, but also provide new insights on following mutagenetic experiments on conserved residues R19 and T37. In addition, these simulations provide dynamic information of the hydrogen bonds discovered in the crystal structure and indicate that they persist as F^- is

transported. More importantly, the Drude REUS simulations eliminate the possibility of F^- being transiently protonated by showing the feasibility of F^- transport in its charged state. These findings collectively provide insights into the F^- transport mechanism in the Ec2 as well as other Fluc proteins.

Looking forward, there are many projects that I wanted to work on. For the ClC protein, I have investigated the proton transport mechanism under alternative anion flux or the I109F mutation, but a more complete mechanism for the anion/proton transport cycle should be studied employing the multiscale kinetic modeling (MKM) which integrates both ion transport flux and combines stepwise mechanism.

For the Fluc channel, I have discovered the importance of conserved residues R19 and T37. To look forward, I believe collaborating with experimentalist to confirm this discovery is essential. With the experimental data in hand, direct simulations on these mutants can then be performed to further explain the molecular details of these residues regulating the fluoride transport activity. Another aspect of the Fluc protein to be studied is the extreme F^-/Cl^- selectivity which can be investigated by simulating the Fluc protein across which a chloride instead of fluoride shuttle and by comparing the Cl^- transport mechanism with the F^- transport mechanism proposed here.

In addition, a more general scientific question that I wanted to answer is what makes the difference between a channel and an antiporter. For example, the ClC-1 protein discovered in mammalian species is a Cl^- channel whereas a homolog of this protein, ClC-ec1, is found to be a Cl^-/H^+ antiporter. Having studied the ion exchange mechanism through PMF calculations and MKM analyses for the ClC-ec1, we should reveal the mechanism of proton inactivity in the ClC-

1 channel at the next step. A comparison between the transport mechanism of these two proteins should provide insights into the distinction between channels and antiporters.