MOLECULAR SIMULATIONS OF VOLTAGE-GATED PROTON-SELECTIVE CHANNEL

Mr. Panisak Boonamnaj

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

A Dissertation Submitted in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy in Chemistry

Department of Chemistry

FACULTY OF SCIENCE

Chulalongkorn University

Academic Year 2018

Copyright of Chulalongkorn University

การจำลองเชิงโมเลกุลของแชนนัลแบบเลือกจำเพาะโปรตอนชนิดโวลเทจเกต

นายปณิศักย์ บุญอำนาจ

MOLECULAR SIMULATIONS OF VOLTAGE-GATED PROTON-SELECTIVE CHANNEL

Mr. Panisak Boonamnaj

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

A Dissertation Submitted in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy in Chemistry

Department of Chemistry

FACULTY OF SCIENCE

Chulalongkorn University

Academic Year 2018

| Dissertation Title | MOLECULAR SIMULATIONS OF VOLTAGE-GATED PROTON-SELECTIVE CHANNEL |
| --- | --- |
| By | Mr. Panisak Boonamnaj |
| Field of Study | Chemistry |
| Thesis Advisor | Professor Pornthep Sompornpisut, Ph.D. |

Accepted by the FACULTY OF SCIENCE, Chulalongkorn University in Partial Fulfillment of the Requirement for the Doctor of Philosophy

---------------------------------------- Dean of the FACULTY OF SCIENCE

(Professor Polkit Sangvanich, Ph.D.)

DISSERTATION COMMITTEE

---------------------------------------- Chairman

(Associate Professor Vudhichai Parasuk, Ph.D.)

---------------------------------------- Advisor

(Professor Pornthep Sompornpisut, Ph.D.)

---------------------------------------- Examiner

(Professor Supot Hannongbua, Dr. rer. nat.)

---------------------------------------- Examiner

(Assistant Professor Somsak Pianwanit, Ph.D.)

---------------------------------------- External Examiner

(Assistant Professor Chalermpol Kanchanawarin, Ph.D.)

ปณิศักย์ บุญอำนาจ : การจำลองเชิงโมเลกุลของแชนนัลแบบเลือกจำเพาะโปรตอนชนิดโวลเทจเกต. (MOLECULAR SIMULATIONS OF VOLTAGE-GATED PROTON-SELECTIVE CHANNEL) อ.ที่ปรึกษาวิทยานิพนธ์หลัก : ศ. ดร.พรเทพ สมพรพิสุทธิ์

แชนนัลแบบเลือกจำเพาะโปรตอนชนิดโวลเทจเกต (Hv1) เป็นโปรตีนฝังในเยื่อหุ้มเซลล์ทำหน้าที่คัดเลือกและขนส่งโปรตอนไอออนข้ามเยื่อหุ้มเซลล์ ซึ่งมีบทบาทสำคัญต่อกระบวนการต่าง ๆ ทางสรีรวิทยา โครงสร้างของโปรตอนแชนนัลถูกพบว่าอยู่ร่วมกันในลักษณะไดเมอร์ (dimer) ในเยื่อหุ้มเซลล์ กลไกการทำงานในระดับโมเลกุลต่าง ๆ ที่เกี่ยวข้องกับการขนส่งโปรตอนเป็นประเด็นที่ถูกวิจัยมาอย่างต่อเนื่องแต่ยังคงไม่เป็นที่แน่ชัด ในงานวิจัยนี้การจำลองพลวัตเชิงโมเลกุล (molecular dynamics simulation) ถูกนำมาใช้ เพื่อให้เข้าใจเกี่ยวกับความสัมพันธ์ระหว่างโครงสร้างและการทำงานได้มากขึ้น ในส่วนแรกเกี่ยวข้องกับการตรวจสอบบทบาทของโดเมนปลายซี (C-terminal domain) ที่เกี่ยวข้องกับการรักษาเสถียรภาพของโครงสร้างแบบไดเมอร์ ข้อมูลจากการจำลองพลวัตเชิงโมเลกุลแสดงให้เห็นว่าเสถียรภาพของโครงสร้างแบบไดเมอร์นั้นขึ้นกับ อันตรกิริยาต่าง ๆ ระหว่างกรดอะมิโนที่อยู่บริเวณโดเมนปลายซี ในส่วนที่สองเป็นการตรวจสอบผลกระทบจากสถานะประจุ (ionization state) ของเรซิดิวต่าง ๆ บนโครงสร้างของโดเมนรับรู้ศักย์ไฟฟ้า (voltage-sensing domain) พบว่าการเปลี่ยนสถานะประจุของเรซิดิวบริเวณช่องขนส่งโปรตอนมีผลต่อการเปลี่ยนแปลงทางโครงสร้างของโดเมนรับรู้ศักย์ไฟฟ้า ทั้งในส่วนของขนาดของช่องขนส่งโปรตอนและจำนวนน้ำ ในการศึกษายังพบว่าการจัดเรียงตัวของน้ำในครึ่งบนและครึ่งล่างของช่องขนส่งโปรตอนนั้นมีทิศทางที่ตรงกันข้ามกัน ซึ่งทิศทางการจัดเรียงตัวนั้นถูกปรับเปลี่ยนไปเมื่อการเปลี่ยนสถานะประจุของเรซิดิว ในส่วนสุดท้ายเป็นการตรวจสอบการเปลี่ยนแปลงทางโครงสร้างของโดเมนปลายซีเนื่องจากอุณหภูมิ ด้วยวิธีการจำลองพลวัตเชิงโมเลกุลและการจำลองสถานการณ์มอนติคาร์โลแบบอะตอมหยาบ (coarse-grained Monte Carlo simulation) พบว่าในการเพิ่มอุณหภูมินั้น โครงสร้างโดเมนปลายซีมีเปลี่ยนแปลงรูปร่างไปเป็นแบบก้อนกลม (globular conformation) ในช่วงแรก และคลายตัวออกในภายหลังเนื่องจากการเสียสภาพ ผลที่ได้จากทั้งสองวิธีการคำนวณให้ผลที่สอดคล้องกันแสดงให้เห็นถึงประสิทธิภาพของเทคนิคการจำลองสถานการณ์มอนติคาร์โลแบบอะตอมหยาบ

| | | |
|---|---|---|
| ภาควิชา | ภาควิชาเคมี | ลายมือชื่อนิสิต ................................................. |
| สาขาวิชา | เคมี | ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก ................ |
| ปีการศึกษา | 2561 | |

# # 5672892223 : DOCTOR OF PHILOSOPHY

VOLTAGE-GATED PROTON-SELECTIVE CHANNEL / C-TERMINAL DOMAIN / VOLTAGE-SENSING DOMIAN / MD SIMULATION / CGMC SIMULATION

Panisak Boonamnaj : MOLECULAR SIMULATIONS OF VOLTAGE-GATED PROTON-SELECTIVE CHANNEL. ADVISOR: Prof. Pornthep Sompornpisut, Ph.D.

Voltage-gated proton-selective channels (Hv1) are a membrane protein that permeates proton across cell membranes. They play an important role in various physiological processes. Hv1 was found as a dimer in membranes. Molecular mechanisms underlying the proton conduction of Hv1 have been a subject of intense research but is still remained unclear. In this study, molecular dynamics (MD) simulations of Hv1 were carried out to gain better understanding of the structure-function relationship. The first part of this thesis involved with an investigation of the role of the C-terminal domain (CTD) in stabilizing the dimer structure of Hv1. MD data showed the stability of the dimer structure depended on intersubunit interactions between amino acids located on the CTD. In the second part, an effect of ionization state of the charged residues on voltage-sensing domain (VSD) was examined. Upon changes in protonation states of conserved acid residues in the pore, conformational changes of the VSD were observed, affecting the size and hydration properties of the aqueous pore in Hv1. This study showed the orientation of water molecules in the pore. It was proposed that water molecules in the upper pore oriented its dipole vector in an opposite direction to those in the lower pore. The water orientation in the upper and lower pore was swapped as a result of changing the protonation state. In the last part, thermal-induced conformational changes of the CTD were investigated using coarse-grained Monte Carlo (CGMC) and all-atom MD simulations. As temperature increased, the CTD maintained its globular conformations in its native phase but it expanded in denatured phase. The results obtained from CGMC were compatible with the MD results demonstarting an efficiency of the CGMC approach.

| | | |
|---|---|---|
| Department: | Department of Chemistry | Student's Signature ...................... |
| Field of Study: | Chemistry | Advisor's Signature ...................... |
| Academic Year: | 2018 | |

# ACKNOWLEDGEMENTS

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# LIST OF ABBREVIATIONS

| | | |
|---|---|---|
| *Ci*-VSP | = | *Ciona intestinalis* voltage-sensitive phosphatase |
| EPR | = | Electron paramagnetic resonance |
| HF | = | Hatree-Fock |
| Hv | = | Voltage-gated proton-selective channel |
| Kv | = | Voltage-gated potassium-selective channel |
| MD | = | Molecular dynamics |
| MC | = | Monte Carlo |
| Nav | = | Voltage-gated sodium-selective channel |
| PaDAR | = | Pseudoatom-driven solvent accessibility refinement |
| PDB | = | Protein Data Bank |
| POPC | = | 1-palmitoyl2-oleoyl-sn-glycero-3-phophochloline |
| RMSD | = | Root-mean square deviation |
| RMSF | = | Root-mean square fluctuation |
| S1-S4 | = | Transmembrane segments 1 to 4 of VSD |
| SF | = | Selectivity filter |
| TIP3P | = | Transferable intermolecular potential three-position |
| TM | = | Transmembrane |
| vdW | = | van der Waals |
| VSD | = | Voltage-sensing domain |
| WT | = | Wild-type |
| X-ray | = | Crystallographic spectroscopy |

# CHAPTER I: INTRODUCTION

## 1.1 RESEARCH RATIONALES AND THEORIES

### 1.1.1 ION CHANNLES

Ion channels are essential for the survival of a cell by control the transport of ions across cell membrane. Ion channels were classified into 3 broad categories according to the opening mechanism: voltage-gated ion channels (VGICs), ligand-gated ion channels (LGICs), and mechanically gated (mechanosensitive, MS) ion channels (MGICs). [1]
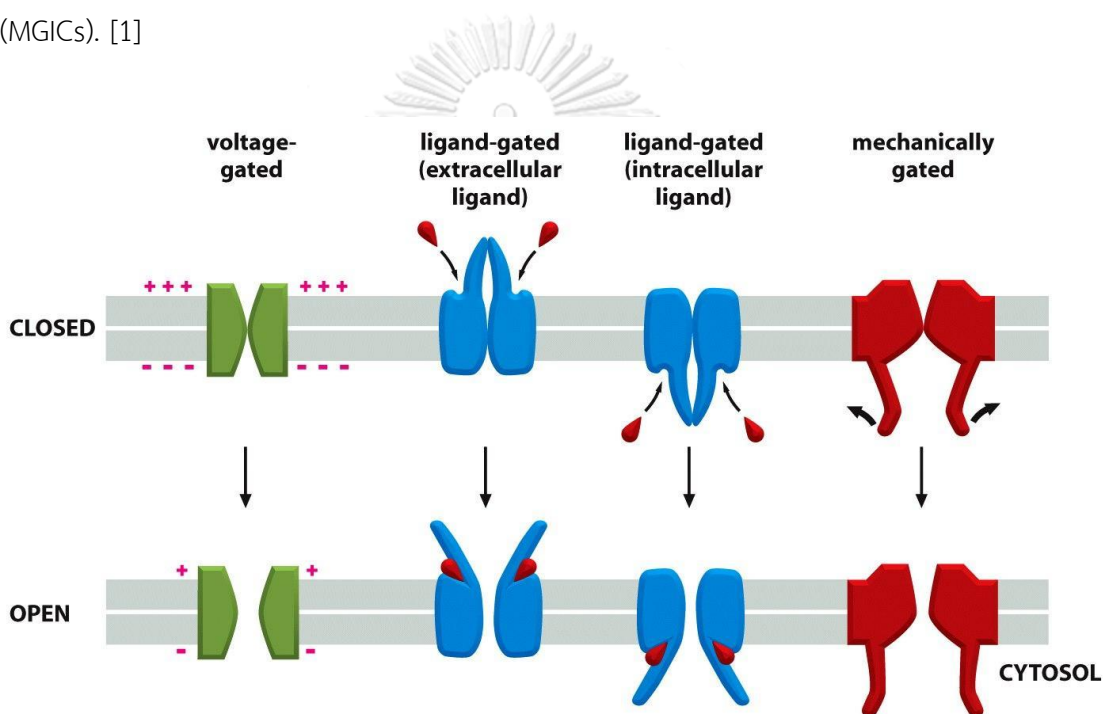


**Figure 1.1**     Different kinds of ion channels; voltage-gated ion channels (VGICs), ligand-gated ion channels (LGICs), and mechanically gated (mechanosensitive, MS) ion channels (MGICs). [1]

VGICs respond to a change in the membrane potential and cause conformational changes resulting in the opening of the pore such as the sodium-selective channel (Nav), potassium-selective channel (Kv), calcium-selective channel (Cav) and proton-selective (Hv) channels. LGICs open in response to binding of a ligand which includes GABA and glycine receptor channels. In case of MGICs, ion flux signals are generated by response to external or internal force. There are two families of MS channels: small conductance mechanosensitive ion channel (MscS) and large conductance mechanosensitive ion channel (MscL). In this thesis, voltage-gated proton-selective channel is focused.

### 1.1.2 VOLTAGE-GATED PROTON-SELECTIVE CHANNEL

Voltage-gated proton-selective (Hv) channel or proton channel is a proton specific ion channels that open and close strongly depend on both membrane potential ($\Delta V$) and the pH-gradient across the membrane ($\Delta pH$). [2] In 1972, the proton channel was first identified in dinoflagellates. [3] It mediates proton extrusion and serves important physiological roles in a variety of cell types such as pH homeostasis, B-lymphocyte signaling, sperm motility and production of reactive oxygen species. [2] A single proton channel can transport $H^+$ up to 104 ions per second under physiological conditions. Furthermore, the proton channel can promote tumor propagation in highly metastatic human breast cancer. Commonly, similar structures of proteins are found in nature and they perform the same or similar biological roles, called "isoform". However, only one gene of voltage-gated proton-selective channel encoding has been identified, namely hydrogen voltage-gated channel 1 (HVCN1). Thus, the voltage-gated proton-selective channel or Hv1 channel was studied in this thesis (H for $H^+$, the conducted ionic specie, subscript V for voltage-gated and 1 for isoform 1).

The canonical voltage-gated ion channels, i.e. the voltage-gated potassium-selective (Kv), sodium-selective (Nav), and calcium-selective channels (Cav), are comprised of 6 transmembrane segments (S1-S6 or 6TM). [2] In the case of Kv, the channels are formed by four identical 6TM protomers, making the channel a four-fold symmetric tetramer. In the cases of eukaryotic Nav and Cav, the four 6TM subunits are present in a tandem single polypeptide. The S1–S4 segments form the voltage-sensing domain (VSD) and the S5-S6 segments constitute pore domain (PD). In contrast, the Hv1 channel contains only a voltage-sensing domain (VSD), lacking a separate pore domain (PD) and hence the alternative names voltage-sensing only protein (VSOP). [2] The VSD of Hv1 channel consists of 4 transmembrane segments (S1-S4 or 4TM), as in the VSDs of other voltage-gated ion channels and is responsible for proton permeation. There are N-terminal domain (NTD) and C-terminal domain (CTD) extending outside of the 4TM in cytoplasmic side. The S4 segment, which contains voltage-sensitive arginines (R1-R3) highly conserved in other voltage-gated ion channels, is directly connected to the CTD. Hv1 has been found that form homodimer molecule in which two VSD subunits are held together by the CTD. The CTD of Hv1 is not only responsible for channel dimerization but also cooperative, and thermosensitive gating. [4] Activation of the Hv1 subunits was controlled by each other. Exclusive proton selectivity is a hallmark of the proton channel. Aspartate (Asp) in the middle of the S1 segment is a crucial part of the selectivity filter in the Hv1 channel. Replacement of Asp, the mutant channel lost proton specificity. [2]
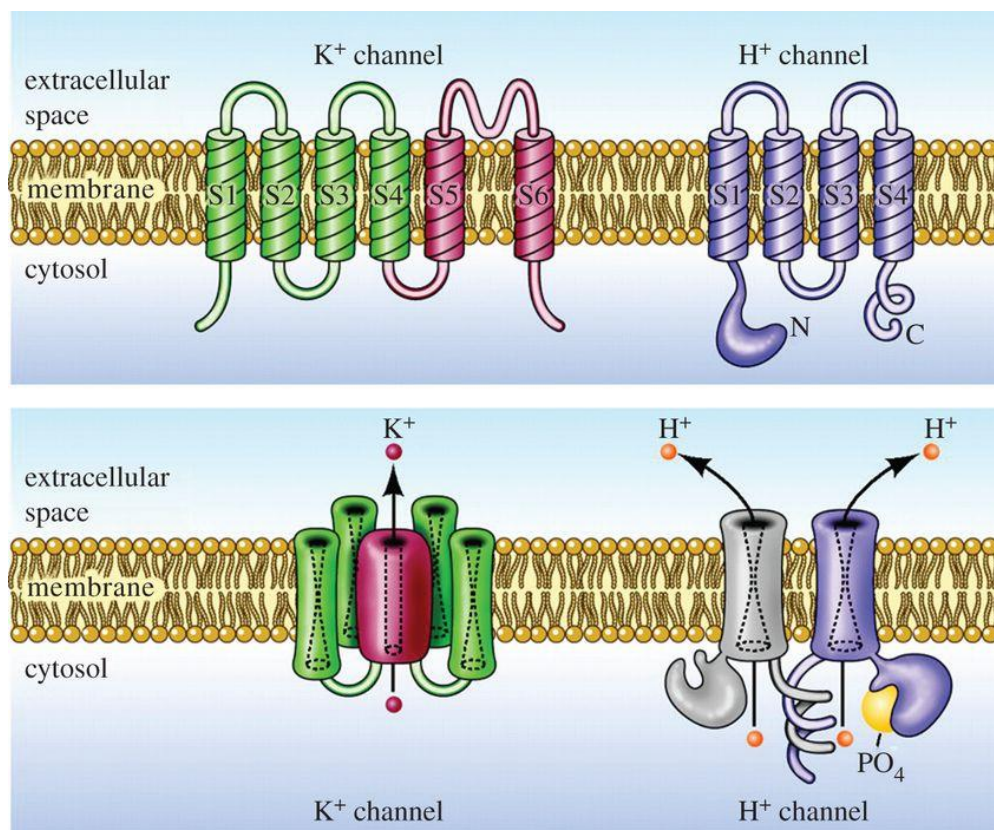
**Figure 1.2** Structures of voltage-gated potassium-selective (Kv) and voltage-gated proton-selective (Hv) channels. The top row shows monomeric subunits of the complete molecule in the lower row. Kv channels comprise four identical subunits with six TM helices per monomer. The S1–S4 form the voltage-sensing domain (VSD) and S5–S6 form the conduction pathway. In case of Hv1 channels, VSD is made with only four TM helices with no explicit pore domain. Hv1 channels form a dimer, largely due to coiled-coil interaction in the C-terminal domain in which each subunit has its own conduction pathway. [2]

### 1.1.3 MOLECULAR DYNAMICS SIMULATION

Molecular dynamics (MD) simulation is a computational method based on molecular mechanics (MM). In MM calculation, all atoms in system are treated as balls and their bonds as springs. An empirical potential energy (PE) function is used to describe how atoms interact with each other based on their relative position. The below mathematical equation to compute the potential energy, $U$, of a system containing atoms and includes terms describing both bonded and nonbonded interactions.

$$U = U_{bonded} + U_{nonbonded} \qquad (1.1)$$

The first term contains the bonded contributions: bond stretching, bond angle bending and bond twisting (dihedral-, rotational- or proper torsions) and improper torsion.

$$U_{bonded} = U_{bond} + U_{angle} + U_{dihedral} + U_{improper} \qquad (1.2)$$

$$U_{bonded}(\vec{r}_1,...,\vec{r}_i,...,\vec{r}_N) = \sum_{i,j} U_{bond}(\vec{r}_i,\vec{r}_j) + \sum_{i,j,k} U_{angle}(\vec{r}_i,\vec{r}_j,\vec{r}_k)$$
$$+ \sum_{i,j,k,l} U_{dihedral}(\vec{r}_i,\vec{r}_j,\vec{r}_k,\vec{r}_l) + \sum_{i,j,k,l} U_{improper}(\vec{r}_i,\vec{r}_j,\vec{r}_k,\vec{r}_l) \qquad (1.3)$$

In a system composed of $N$ particles interacting according to the rules of the force field, the force acting on each particle, $i$, is estimated by the gradient of the potential energy with respect to atomic displacement.

$$\vec{F}_i = -\nabla_i U\left(\vec{r}_1,...,\vec{r}_i,...,\vec{r}_N\right) \qquad (1.5)$$

When the forces have been assigned, Newton's law of motion, $\vec{F} = m\vec{a}$, can be used to numerically solve the motion of the particle $i$.

In the above equation, $m_i$, $\vec{a}_i$, $\vec{v}_i$ represent mass, acceleration, velocity, respectively, of the $i$-th particle at time, $t$. The variable $\vec{r}_i$ represents the position vector of particle $i$ in a cartesian coordinate system. The new position of each particle is computed, the interatomic forces are updated in order to calculate the subsequent positions after another time step. This entire process is repeated to obtain a trajectory of coordinates of the system for a finite time period. This method is widely used to study the molecular behavior. Several program packages have been developed for simulating the atomistic level of biomolecules. Here, we use NAMD which is the parallel molecular dynamics program designed for high-performance simulation of large biomolecular systems. NAMD uses the popular molecular graphics program VMD for simulation setup and trajectory analysis.

### 1.1.4 MONTE CARLO SIMULATION

Monte Carlo (MC) simulation is named after the city of Monte Carlo in Monaco, which is famous for gambling, The MC simulation is a computational method based on repeated random sampling. The goal of a Monte Carlo simulation is to explore a representative set of states based on randomness according to a statistical mechanical ensemble. A system can be in a certain state with a certain probability. This is expressed as

$$\rho_i = \frac{e^{-\beta U(\vec{r}_i)}}{\int\limits_r e^{-\beta U(\vec{r})} dr} \tag{1.6}$$

where $\rho_i$ is the probability that the system is in state $i$ with potential energy $U(\vec{r}_i)$. This is the idea behind Monte Carlo simulation. Metropolis et al. [5] were the first to show the way for sampling with calculation of thermal averages according to the principles of statistical mechanics in a given equilibrium ensemble. A set of random numbers is used to determine the sequence of states through which the system evolves. The probability of moving from one state to another depends only upon the state that the system is currently staying in, no matter how the system got into that current state. A trial state is generated that is nearby the current state of the system upon a trial move calculated by random numbers. After moving, one calculates the potential energy change between two states, $(U(\vec{r}_{new}) - U(\vec{r}_{old}))$, and the new configuration is accepted or rejected based on the acceptance probability.

$$A_{old \to new} = \min[1, e^{-\beta(U(\vec{r}_{new}) - U(\vec{r}_{old}))}] \tag{1.7}$$

If the energy change is negative, which means that the system moved to a lower energy state, then the exponential term is greater than 1 and the move is accepted with probability 1. If the energy change is positive, then the move is accepted with probability $e^{-\beta(U(\vec{r}_{new}) - U(\vec{r}_{old}))}$ and compares with a generated random number in uniformly distributed over the interval (0,1). The new state will be accepted if the generated random number is less than $e^{-\beta(U(\vec{r}_{new}) - U(\vec{r}_{old}))}$, otherwise the system is returned to its original state. The entire process repeatedly occurs to obtain a representative set of states of the system. This is generally known as a Metropolis algorithm. This method is highly used to study motions in biomolecules.

### 1.1.5 COARSE-GRAINED MODEL

The use of coarse grained (CG) models in a variety of simulation techniques has known as a useful tool to probe the time and length scales of systems beyond what is feasible with traditional all atom (AA) models. In all-atom methods, simulations with long time scales are very high cost and impossible in many cases. The CG models are applied to overcome these limitations. The time scales and length of simulations can be extended by making some approximations and integrating out the unnecessary degrees of freedom. Atoms in a molecule are grouped into pseudoatoms referred to as beads. Coarse Graining models were initiated for modeling huge biomolecules. Among different CG models, MARTINI coarse-grained model, developed by Marrink and coworkers [6] is used in this study. The MARTINI model provides suitable level of coarse-graining, as it retains information about the chemistry. In the MARTINI approach 4:1 with mapping scheme, four heavy atoms are grouped together into one bead represented by a single interaction center. There are only four types of representative bead are considered: polar (P), nonpolar (N), apolar (C), and charged (Q). Apart from the 4:1 mapping scheme for standard beads, CG also uses 2-3:1 mapping scheme to represent ring compounds. Each bead type has a number of subtypes, which allow for a more accurate representation of the chemical nature of the underlying atomic structure. Within a main type, subtypes are either distinguished by a letter denoting the hydrogen-bonding capabilities (d = donor, a = acceptor, da = both, 0 = none), or by a number indicating the degree of polarity (from 1, low polarity, to 5, high polarity). In this thesis, the CG models was combined with MC method in studying dynamics of protein molecules.

### 1.1.6 POISSON-BOLTZMANN IMPLICIT SOLVATION MODEL

There are several computational approaches to approximate solvent effects in molecular dynamics. They can be roughly divided into 3 classes: explicit, implicit and mixed solvation models. Such a large system of biomolecules with water molecules, the explicit solvation model is computationally expensive. Important properties such as solvation free energies may often converge slowly and take long time due to water molecules. Thus, the implicit solvation model is most widely used even it is simplistic.

It would be useful to reduce computational effort. Solvent was presented as a continuous medium instead of individual explicit solvent molecules. Estimation ability of solvent behavior with non-computationally expensive cost is acceptable. The most implicit solvent models are commonly based on the Poisson equation.

$$\nabla \varepsilon \cdot \nabla \phi\left(\vec{r}\right) = -4\pi \rho\left(\vec{r}\right) \tag{1.8}$$

where $\varepsilon$ is the dielectric constant at a given point in space, $\phi(\vec{r})$ is the electrostatic potential distribution function and $\rho(\vec{r})$ is the total charge density.

The Poisson equation must be augmented with a distribution of ions surrounding the biomolecules because it is only valid at zero ionic strength. The probability of finding an ion in a particular region, $P(\vec{r})$, of space is related to the Boltzmann factor, $P\left(\vec{r}\right) \propto e^{-\beta q\phi(\vec{r})}$ where $\beta = \dfrac{1}{k_B T}$ in which $T$ is the absolute temperature of the system and $k_B$ is the Boltzmann constant. $q\phi\left(\vec{r}\right)$ is the energy of a point charge in a given electrostatic potential. The probability of finding positive and negative charges of ions must be considered. The equation for calculating the electrostatic potential in a biomolecular system with a given solution ionic strength, termed the Poisson-Boltzmann (PB) equation, is shown in Eq. 1.9.

$$\nabla \varepsilon \cdot \nabla \phi\left(\vec{r}\right) = -4\pi \left(\rho_{mol}\left(\vec{r}\right) + \rho_{ions}\left(\vec{r}\right)\right) \tag{1.9}$$

where total charge density, $\rho\left(\vec{r}\right)$, in the system is equal to the total charge density of biomolecules, $\rho_{mol}\left(\vec{r}\right)$, and the total charge density of ions , $\rho_{ions}\left(\vec{r}\right)$. The total charge density of ions in a particular region with quantities of several ion types, $n_i$, each bearing charge, $q_i$, is given by $\rho_{ions}(\vec{r}) = \sum_{i=1}^{n} q_i n_i$ . Distribution of the $i$ particle with potential energy, $\phi\left(\vec{r}\right)$, based on the Boltzmann distribution law of classical mechanics is written as $n_i(\vec{r}) = n_i^0 e^{-\beta q_i \phi(\vec{r})}$ when $n_i^0$ is bulk value. Thus, the total charge density of ions can be expressed in the form $\rho_{ions}(\vec{r}) = \sum_{i=1}^{n} q_i n_i^0 e^{-\beta q_i \phi(\vec{r})}$ . Hence, the PB equation can be rewritten as

$$\nabla \varepsilon \cdot \nabla \phi\left(\vec{r}\right) = -4\pi \left(\rho_{mol}\left(\vec{r}\right) + \sum_{i=1}^{n} q_i n_i^0 e^{-\beta q_i \phi(\vec{r})}\right) \tag{1.10}$$

To make linearized PB equation, Debye-Huckel theory is applied. Based on the assumption that electrostatic potential of the system, $\phi(\vec{r})$, would be smaller than the thermal energy, $k_B T$, then $q_i\phi(\vec{r}) \ll k_B T$ or $\dfrac{q_i\phi(\vec{r})}{k_B T} \ll 1$. The exponential term $e^{-\frac{qi\phi(\vec{r})}{k_B T}}$ can be expanded using a Taylor series as following, $e^{-\frac{q_i\phi(\vec{r})}{k_B T}} = 1 - \dfrac{q_i\phi(\vec{r})}{k_B T} + \dfrac{1}{2}\left(\dfrac{q_i\phi(\vec{r})}{k_B T}\right)^2 + ...$, and neglect all except the first two terms then the total charge density of ions will be,

$$\rho_{ions}(\vec{r}) = \sum_{i=1}^{n} q_i n_i^0 \left(1 - \frac{q_i\phi(\vec{r})}{k_B T}\right) = \sum_{i=1}^{n} q_i n_i^0 - \sum_i^n \frac{q_i^2 n_i^0 \phi(\vec{r})}{k_B T} \tag{1.11}$$

The first term $\sum_{i=1}^{n} q_i n_i^0$ is the charge of solution that is zero. Hence,

$$\rho_{ions}(\vec{r}) = -\sum_{i=1}^{n} \frac{q_i^2 n_i^0 \phi(\vec{r})}{k_B T} \tag{1.12}$$

The linearized PB equation is obtained

$$\nabla\varepsilon \cdot \nabla\phi(\vec{r}) = -4\pi\rho_{mol}(\vec{r}) + 4\pi\sum_{i=1}^{n} \frac{q_i^2 n_i^0 \phi(\vec{r})}{k_B T} \tag{1.13}$$

or

$$\nabla\varepsilon \cdot \nabla\phi(\vec{r}) = -4\pi\rho_{mol}(\vec{r}) + \left(\frac{4\pi}{k_B T}\sum_{i=1}^{n} q_i^2 n_i^0\right)\phi(\vec{r}) \tag{1.14}$$

The constant in the right-hand parentheses can be lumped together and called as a new constant $\kappa^2$,

$$\kappa^2 = \frac{4\pi}{\varepsilon k_B T}\sum_{i=1}^{n} q_i^2 n_i^0 \tag{1.15}$$

Then, the linearized PB equation in term of $\kappa^2$ is shown,

$$\nabla \varepsilon \cdot \nabla \phi \left( \vec{r} \right) = -4\pi \rho_{mol} \left( \vec{r} \right) + \varepsilon \kappa^2 \phi(\vec{r}) \tag{1.16}$$

At high concentration, however, the Debye-Huckel theory fails. An alternative approach with failure of the Debye-Huckel theory base on linearized PB equation is adding higher terms. For a symmetrical electrolyte, $z_+ = z_- = z$ , with equal concentrations of positive and negative ions, $n_+^0 = n_-^0 = n_i^0$ and $q_i = z_i e_0$ where $z_i$ is charge number and $e_0$ is the elementary charge, one gets

$$\rho_{ions} \left( \vec{r} \right) = \rho_+ \left( \vec{r} \right) + \rho_- \left( \vec{r} \right) = \left( z_+ e_0 \right) n_i^0 e^{-\frac{(z_+ e_0)\phi(\vec{r})}{k_B T}} + \left( z_- e_0 \right) n_i^0 e^{-\frac{(z_- e_0)\phi(\vec{r})}{k_B T}} \tag{1.17}$$

$$\rho_{ions} \left( \vec{r} \right) = \left( z e_0 \right) n_i^0 \left( e^{-\frac{(z_- e_0)\phi(\vec{r})}{k_B T}} - e^{\frac{(z_+ e_0)\phi(\vec{r})}{k_B T}} \right) \tag{1.18}$$

But

$$e^{+x} - e^{-x} = 2\sinh\left( x \right) \tag{1.19}$$

and therefore

$$\rho_{ions} \left( \vec{r} \right) = -2 n_i^0 z e_0 \sinh \left( \frac{(z e_0)\phi\left( \vec{r} \right)}{k_B T} \right) \tag{1.20}$$

The nonlinearized PB equation is obtained

$$\nabla \varepsilon \cdot \nabla \phi \left( \vec{r} \right) = -4\pi \left( \rho_{mol} \left( \vec{r} \right) - 2 n_i^0 z e_0 \sinh \left( \frac{(z e_0)\phi\left( \vec{r} \right)}{k_B T} \right) \right) \tag{1.21}$$

or

$$\nabla \varepsilon \cdot \nabla \phi \left( \vec{r} \right) = -4\pi \rho_{mol} \left( \vec{r} \right) + 8\pi n_i^0 z e_0 \sinh \left( \frac{(z e_0)\phi\left( \vec{r} \right)}{k_B T} \right) \tag{1.22}$$

## 1.2 RESEARCH METHODOLOGY

1.2.1    The structure and dynamics of protein were calculated using the MD and CGMC simulations.

1.2.2    PB implicit solvation model was employed to estimate free energies of total solvation, electrostatic and nonpolar components.

## 1.3 OBJECTIVES

1.3.1    To investigate dynamical properties and important interactions involved with the dimer stabilization by C-terminal domain (CTD).

1.3.2    To study the effect of ionization state of the charged residues on conformational change and hydration profiles in voltage-sensing domain (VSD).

1.3.3    To characterize the thermal-response conformational change of C-terminal domain (CTD) dimer using all-atom MD with coarse-grained MC (CGMC) simulations.

## 1.4 SCOPE OF THIS DISSERTATION

The purposes of this thesis are to investigate the structural functions of the Hv1 channel involved with dimer stabilization by C-terminal domain (CTD), conformational change of monomeric Hv1 channel in response to pH and thermal denaturation of dimeric CTD. In Chapter II, the role of CTD on dimer stability of the Hv1 channels was initially studied. Binding strength between two subunits depended on intersubunit interactions including salt-bridges, hydrophobic interactions, hydrogen bonds, and a disulfide bond across the dimer interface. In Chapter III, influence of ionization states on conformation change was also investigated. The results suggested that the voltage-

sensing domain (VSD) has been shown to undergo a pH-dependent conformational change. Modification of the ionization states of charged residues through the VSD induced reorganization of a hydrogen-bond network affected the packing of VSD core. The hydration profiles and dipole moment orientation of waters also exhibited differently. In Chapter IV, thermal stability of the CTD dimer has been examined using CGMC simulation compared with results obtained from MD simulation. Observations from investigation by both different approaches were highly consistent. The radius of gyration ($R_g$) decreases upon an increasing temperature and switch to higher values of the $R_g$ by further increasing the temperature.

## 1.5 EXPECTED BENEFICIAL OUTCOMES

1.5.1     Understanding the dynamics properties and important interactions involved with the dimer stabilization by C-terminal domain (CTD).

1.5.2     Influence of ionization state on conformational change and hydration profiles in voltage-sensing domain (VSD).

1.5.3     Understanding the denaturation behavior of CTD dimer by temperature obtained from MD simulation with CGMC simulation.

The results could be useful information for better understanding of biological functions of the Hv1 channels. In addition, it may provide the greatest benefit for medical knowledge which can be utilized for a development of drug design in the future.

# CHAPTER II: STRUCTURAL DYNAMICS AND INTERACTIONS OF C-TERMINAL DOMAIN OF THE Hv1 CHANNEL IN DIMER SATABILIZATION

Insight into the Role of the Hv1 C‑Terminal Domain in Dimer Stabilization

Panisak Boonamnaj and Pornthep Sompornpisut

*Department of Chemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand*

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

## 2.1 ABSTRACT

The voltage-gated proton-selective channel (Hv1) conducts protons in response to changes in membrane potential. The Hv1 protein forms dimers in the membrane. Crystal structures of Hv1 channels have revealed that the primary contacts between the two monomers are in the C-terminal domain (CTD), which forms a coiled-coil structure. The role of Hv1-CTD in channel assembly and activity is not fully understood. Here, molecular dynamics (MD) simulations of full-length and truncated CTD models of human and mouse Hv1 channels reveal a strong contribution of the CTD to the packing of the transmembrane domains. Simulations of the CTD models highlight four fundamental interactions of the key residues contributing to dimer stability. These include salt bridges, hydrophobic interactions, hydrogen bonds, and a disulfide bond across the dimer interface. At neutral pH, salt-bridge interactions increase dimer stability and the dimer becomes less stable at acidic pH. Hydrophobic core packing of the heptad pattern is important for stability, as shown by favorable nonpolar binding free energies rather than by electrostatic components. Moreover, free-energy calculations indicate that a more uniform hydrophobic core in the coiled-coil structure of the Hv1-NIN, a channel carrying the triple mutation M234N–N235I–V236N, leads to an increase in dimer stability with respect to the wildtype. A Cys disulfide bond has a strong impact on dimer stability by holding the dimer together and facilitating the interactions described above. These results are consistent with dissociative temperatures and energy barriers of dimer dissociation obtained from the temperature-accelerated MD.

## 2.2 INTRODUCTION

A voltage-gated proton-selective channel (Hv1) serves fundamental roles in physiological processes [2] such as cell signaling for microbial killing in the innate immune [7-10] system and acid extrusion of sperm motility in the male reproductive [11, 12] system by regulation of pH. The Hv1 is made of four transmembrane segments (S1–S4) that comprise a voltage-sensing domain (VSD) [13] similar to that in other voltage-gated ion channels, [14-16] but it lacks a separate pore domain. [17] Recent evidence suggests that native Hv1 functions as a dimer channel in which each channel subunit has its own proton conduction pore. [18-21] The voltage-sensing arginines show remarkably outward movement driven by transmembrane voltage during channel opening. [22] C-terminal domain (CTD), which connects downstream to S4 located on the cytoplasmic side of the membrane, is responsible for dimerization, cooperative gating, and structural stability of Hv1 channels. [18, 19, 21-28] The CTD has a characteristic heptad repeating pattern, and formation of an $\alpha$-helical coiled-coil structure was found in several X-ray crystal structures. [23, 26, 29] Moreover, the stability of the coiled-coil motif is affected by changes not only in temperature [26, 29] but also pH. [30] It has been shown that the CTD could modulate the channel activation kinetics. The monomeric Hv1 channel lacking its CTD (Hv1-$\Delta C$) remains functional and can also be activated significantly faster than the full-length (FL) dimeric Hv1 channel. [18, 21, 24, 26, 27] The S4 movement in the channel opening is facilitated by deletion of the CTD. [22] Moreover, the coiled-coil helices also contain Cys residues which can form a disulfide bond across the dimer interface. The Hv1 channel facilitates the robust production of reactive oxygen species in phagocyte [7-9, 31] that causes a fluctuating environment of oxidation–reduction and alters coiled-coil folding by changing the redox state of Cys residues. Fujiwara et al. [29] demonstrated that disulfide bond, formed under oxidizing condition, could stabilize the coiled-coil structure and increase the thermal stability. Even though previous studies have shown

an important role of the CTD in dimerization and activation of the Hv1 channel, [4, 32] the interactions between these domains for dimer stability remains unclear. Here, we employed molecular dynamics (MD) simulations to identify interactions of the key residues contributing to the stability of the dimer structure. This study provides more structural insight into the fundamental role of the Hv1 C-terminal domain in dimer stabilization.

## 2.3 METHODS

### 2.3.1 Structure Modeling

Dimeric Hv1 channels of the X-ray crystal structure of mHv1cc (PDB ID: 3WKV) [13] (Figure S2.1A) and the electron paramagnetic resonance (EPR)-refined homology model of hHv1 provided by Li [33] (Figure S2.1B) were used as the initial state of the Hv1 proton channel. Loops connecting each transmembrane segment of the mHv1cc were modeled with Loopy, [34] and the protein quality was structurally validated by PROCHECK. [35] The combined model of mHv1ccVSD/hHv1-CTD (Figure S2.1C) and the truncated C-terminal domain model in the dimer form were built using visual molecular dynamics (VMD). [36] The favorable position of protein in the lipid bilayer was carefully predicted by a method previously developed, [37] using the adaptive Poisson–Boltzmann solver (APBS) software package, [38] with PARSE partial charges. [39] The hydrogen atoms and atomic charges were assigned with the PDB2PQR program. [40, 41]

2.3.2 Molecular Dynamics Simulations

All simulations were performed with NAMD [42] using the CHARMM27 force field. [43] Simulation systems were initially constructed using VMD. The full-length and truncated C-terminal domain models were inserted in a 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) lipid bilayer solvated by 100 mM NaCl solutions at 298 K. The dimer of isolated Hv1 C-terminal domain systems (mHv1-CTD and hHv1-CTD) was solvated with 100 mM NaCl solutions. Harmonic constraints were also applied to amino-terminal ends on I214 residues for mHv1cc-CTD and I218 residues for hHv1-CTD to compensate effects of the transmembrane VSD region. Multiple independent MD simulations were carried out for 100 ns. For temperature induced dimer dissociation of Hv1-CTD, simulations without constraints on their ends were further performed for 45 ns and they were solvated with 150 mM NaCl. Side-chain ionization states were assigned on the basis of pKa using PROPKA. [44] All simulation systems were run with a time step of 2 fs and surrounded with explicit TIP3P water molecules. [45] A distance cutoff of 12 Å was used for calculating nonbonded interactions, electrostatic interactions with particle mesh Ewald summation via fast Fourier transform, and van der Waals interactions, and a switching distance was set at 10 Å. Langevin dynamics was used to control the temperature of systems, with a damping coefficient of 1 ps$^{-1}$. Pressure was kept constant at 1 atm using the Nose–Hoover Langevin piston method, with a piston period of 200 fs and a damping time of 50 fs. Energy minimization was used to remove hard geometry and contact between atoms. The surrounding environment, including lipids, ions, and waters, was relaxed, followed by constrained proteins. In full-length models, the hydration of the membrane–protein interface during equilibration was prevented by applying forces on water molecules.

### 2.3.3 MD Trajectory Analysis

The MD trajectories were mainly analyzed by modified VMD scripts, including root-mean square deviation (RMSD), helix–helix crossing angle, salt bridges, and interhelical hydrogen bond. The RMSD was calculated using backbone atoms. The crossing angle was computed by determining the angle between two helix vectors. The initial and end points of each vector were selected from the center of mass of three residues. Distribution curves were constructed using bin width ($\Delta\theta$) of 1°. Subinteracting states of salt bridges were classified by measuring the O–N distance on the basis of the following criteria: $\nu$-state, $d_{E-K'} > 3.5$ Å and $d_{E'-K} > 3.5$ Å; $\sigma$-state, $d_{E-K'} \leq 3.5$ Å and $d_{E'-K} > 3.5$ Å; $\sigma^{*}$-state, $d_{E-K'} > 3.5$ Å and $d_{E'-K} \leq 3.5$ Å; $\delta$-state, $d_{E-K'} \leq 3.5$ Å, and $d_{E'-K} \leq 3.5$ Å. The correlation contours of subinteracting states were plotted by mapping representative colors for the subinteracting states of the lower pairs on the matching positions of salt-bridge distance coordinates of the upper pairs.

### 2.3.4 Binding Free Energy

Protein–protein binding free energy was computed based on the interaction energy and the solvation energy changes of the system (Figure S2.2). Binding free energy, $\Delta\Delta G_{bind}$, is defined by [46]

$$\Delta\Delta G_{bind} = (\Delta G_{solv}^{AB} - \Delta G_{solv}^{A} - \Delta G_{solv}^{B}) + \Delta E_{coul,vac} \qquad (2.1)$$

where solvation energy of dimer, $\Delta G_{solv}^{AB}$, is subtracted with solvation energies of each individual subunit, $\Delta G_{solv}^{A}$ and $\Delta G_{solv}^{B}$, respectively, and Coulombic energy in vacuum of binding ($\Delta E_{coul,vac} = E_{coul,vac}^{AB} - E_{coul,vac}^{A} - E_{coul,vac}^{B}$) is the last term. The solvation energy comprises electrostatic and nonpolar solvation energies ($\Delta G_{solv} = \Delta G_{elec} + \Delta G_{np}$). The binding free energy can be separated into electrostatic component ( $\Delta\Delta G_{bind}^{elec} = \Delta G_{elec}^{AB} - \Delta G_{elec}^{A} - \Delta G_{elec}^{B} + \Delta E_{coul,vac}$ ) and nonpolar component (

$\Delta\Delta G_{bind}^{np} = \Delta G_{np}^{AB} - \Delta G_{np}^{A} - \Delta G_{np}^{B}$ ). The PDB2PQR/APBS software package was used to carry out nonlinear Poisson–Boltzmann implicit solvation calculations. For electrostatic solvation energy, dielectric constants of protein, water, and vacuum were 2, 80, and 1, respectively. Protein charges were assigned on the basis of pKa values. The bathing solution had an ionic strength of 0.1 M with Coulomb charges of +1 and −1 and a radius of 2.0 Å. The water probe radius was 1.4 Å. [47] The electrostatic calculation used grid points of 129 Å × 129 Å × 129 Å and multigrid lengths of 67 Å × 67 Å × 67 Å. For nonpolar solvation energy, solvent pressure was applied as 0.150624 kJ·mol$^{-1}$·Å$^{-3}$, bulk solvent density was 0.033428 Å$^{-3}$, solvent molecule radius was 1.4 Å, and the optimized surface tension coefficient was 0.0209 kJ·mol$^{-1}$·Å$^{-2}$. AMBER ff99 was used for protein charges and van der Waals radius. [48] All energy calculations were carried out using >1000 snapshots for each simulation system extracted from MD trajectories by random sampling. The representative snapshots were superimposed on the initial structure to keep a similar arrangement before calculations. Binding free energies were significantly populated according to the Chauvenet's criterion. [49]

### 2.3.5 Fraction of Dimer Contact

The contact fraction of the CTD dimer is used to define levels of dimer dissociation. It was computed by considering only residues facing the hydrophobic core interface. The distances between $C_\beta - C_\beta$ of two subunits were measured and the number of residue pairs within 10 Å counted. Snapshots were used to evaluate the contact fraction of the dimer at a given time, $\alpha(t_i)$, by the following equation

$$\alpha(t_i) = \frac{\sum_{a=1}^{A} \sum_{b=1}^{B} c_{ab}(t_i)}{\sum_{a=1}^{A} \sum_{b=1}^{B} c_{ab}(t_0)} \tag{2.2}$$

where $c_{ab}(t_i)$ and $c_{ab}(t_0)$ are the number of pairs within the cutoff contact distance at time $t_i$ and the initial time, respectively. The contact fraction was represented in the average value, $\langle\alpha\rangle$, by dividing the number of simulation frames, $N$.

$$\langle\alpha\rangle = \frac{\sum_{k=1}^{N}\alpha(t_i)}{N} \tag{2.3}$$

2.3.6 Dimer Dissociation Constant Based on the Two-State Model

In the simplest case, the dimer-to-monomer transition was fitted to the following two-state scheme [50, 51]

$$\left[(CTD)_2\right] \rightleftharpoons 2\left[CTD\right] \tag{2.4}$$
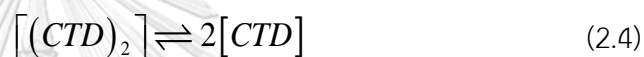
At temperature $T$, the dissociation equilibrium constant, $K_D$, can be determined from the average contact fraction ($\langle\alpha\rangle$) and $P_t$ as follows

$$K_D = \frac{2\left(1-\langle\alpha\rangle^2\right)P_t}{\langle\alpha\rangle} \tag{2.5}$$

where $P_t$ is the total protein concentration in the simulation box and can be obtained from the number of protein molecules, $N_p$, divided by the Avogadro number, $N_A$, and the equilibrated volume of the simulation box, $V$.

$$P_t = \frac{N_p}{N_A V} \tag{2.6}$$

The dissociation constant for the interactions throughout the temperature range can be used to calculate the standard enthalpy and entropy changes of the reaction, $\Delta H^o$ and $\Delta S^o$, respectively, using

$$\ln K_D = -\frac{\Delta H^o}{RT} + \frac{\Delta S^o}{R} \tag{2.7}$$

where $R$ is the ideal gas constant and $T$ is the absolute temperature. Thus, plotting $\ln K_D$ against $\dfrac{1}{T}$ gives the standard enthalpy and entropy changes of the reaction from slope, $-\dfrac{\Delta H^o}{R}$, and y-intercept, $\dfrac{\Delta S^o}{R}$, of the linear fit, respectively.

### 2.3.7 Energy Barrier of Dimer Dissociation

The dissociation process of dimer was kinetically characterized using the Eyring–Polanyi theory, and the equation is shown below

$$\ln\left(\frac{k_{off}}{T}\right) = -\frac{\Delta G^{o\ddagger}}{RT} + \ln\left(\kappa\frac{k_B}{h}\right) \tag{2.8}$$

where $k_{off}$ is the dissociation rate constant, $\Delta G^{o\ddagger}$ is the standard free-energy change of activation, $\kappa$ is transmission coefficient, $h$ is the Planck's constant, and $k_B$ is the Boltzmann constant. The dissociation rate constant was calculated from dissociative half-life $\left(t_{1/2} = \dfrac{0.693}{k_{off}}\right)$, when $\alpha\left(t_{1/2}\right) = 0.5$. The standard free-energy change of activations was obtained from the slope, $-\dfrac{\Delta G^{o\ddagger}}{R}$, and y-intercept, $\ln\left(\kappa\dfrac{k_B}{h}\right)$, of the linear fit, respectively.

## 2.4 RESULTS AND DISCUSSION

### 2.4.1 Structural Models of Hv1 C-Terminal Domain

To gain structural insight into the basis for the specificity of Hv1 dimer stabilization, the resting state of a recently determined X-ray crystal structure of a chimeric protein of the mouse proton channel (mHv1cc) [13] (Figure S2.1A) and a homology model of the human proton channel (hHv1) (Figure S2.1B) developed previously [33] were used in simulations. However, the C-terminal region of the mHv1cc

model (mHv1cc-CTD) was replaced by a leucine zipper motif of the transcriptional activator GCN4 from yeast. [13] Although this makes the CTD length of mHv1cc channel shorter than what is actual, the GCN4 leucine zipper still has a similar common structural blueprint of typical coiled-coil proteins (Figure 2.1A). [52]
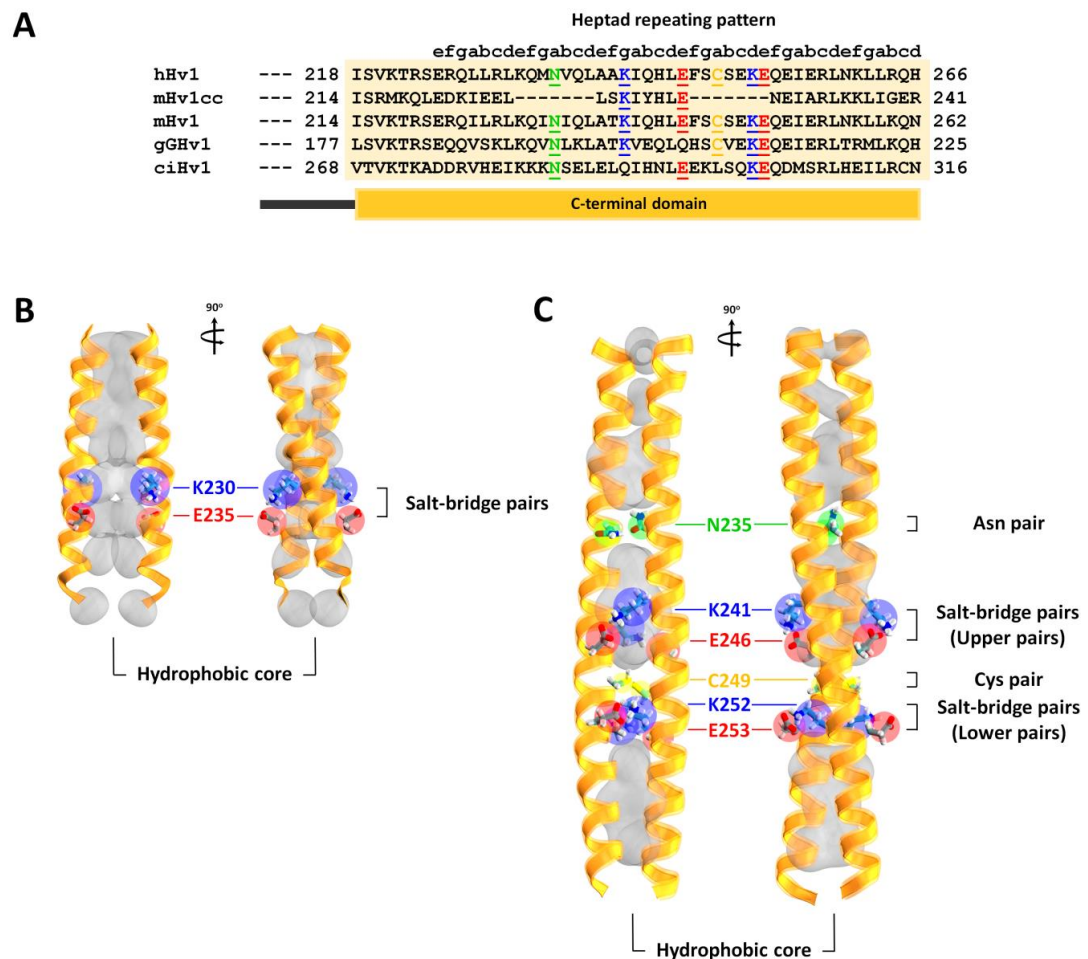


**Figure 2.1**   Structure comparison of CTDs. (A) Relative alignment of C-terminal region from various species with characteristic heptad repeating pattern, $(abcdefg)_n$. Hydrophobic residues predominantly occupy at "a" and "d" positons, and charged residues occupy at the "e" and "g" positions. The C-terminal domain forms a coiled-coil architecture shown in (B) mHv1cc-CTD and (C) hHv1-CTD models, with underpinning residues which might engage in dimer stability highlighted with different shades of sphere color. The hydrophobic core is displayed as a grey quick surface. At dimer interface, the hydrophobic core is disrupted by polar (green sphere) and salt-bridge residues (red and blue sphere).

Additionally, the cytoplasmic region of mHv1cc between the middle of S2 to the middle of S3 was replaced by the corresponding section of voltage sensitive phosphatase from *Ciona intestinalis* (*Ci*VSP). [13] In the case of the hHv1 model, its VSD was built based on the *Ci*VSP dimer and combined with the X-ray crystal structure of isolated hHv1 CTD (hHv1-CTD). [33] Its structure was refined using the structure refinement PaDSAR method that employs electron paramagnetic resonance (EPR) data. [53] Using mHv1cc-VSD and hHv1-CTD of different models, a dimer structure of mHv1cc-VSD/hHv1-CTD was constructed (Figure S1C) where this model makes the actual contributions from CTD more complete for mHv1cc-VSD. The CTD is essential for achieving the dimer structure with two identical monomers forming a coiled-coil architecture via parallel double helical interactions. To explore the dimerization mediated by CTD, available X-ray crystal structures of isolated Hv1-CTD [23, 26, 29] were used to compare with results of the C-terminal region from the models mentioned above. The Hv1-CTD possesses a distinctive heptad repeat of (abcdefg)$_n$, which is highly conserved among different species (Figure 2.1A). Figure 2.1B, C illustrates a well-packed structure of two models in which mHv1cc-CTD is a shorter helical dimer, whereas hHv1-CTD is the longer one, respectively. The CTD interactions were driven from such repeating sequences, enabling the coiled-coil formation in multiple proteins. [54] From the structure and composition of Hv1-CTD, the overall stability of the dimer could involve a variety of side-chain interactions, including salt bridges, hydrogen bonds, hydrophobic interactions, and disulfide bonds. Salt-bridge interactions are engaged by conserved ionizable residues (Glu and Lys pairs) which flank the coiled-coil core. At the middle of the dimeric hHv1-CTD center, polar side chains of the Asn pair, N235–N235′, may contribute to dimer stability by forming hydrogen bonds (Figure 2.1C). Another feature in dimeric hHv1-CTD is the presence of the Cys pair, C249–C249′, (Figure 2.1C) which is known as a redox sensor. [29] Under oxidizing condition, the interhelical disulfide bond will likely be formed and this permits core packing more

tightly than usual. As shown in Figure 2.1A, the hHv1-CTD contains a similar character of these residues in the sequence of mHv1-CTD. Thus, only hHv1-CTD is enough for studying the effect of these residues instead of mHv1-CTD. At this point, the hypothesis on the requirement of the C-terminal region for dimer formation and effect of contributions on the structure will be verified. This study exclusively focuses on inter-subunit interactions involving these underpinning residues marked above.

2.4.2 Hv1 C-Terminal Domain Mediates Transmembrane Domain Interactions

Initially, three independent replications for each equilibrated 100 ns MD trajectory of the dimeric Hv1 channel in full-length (FL) and truncated C-terminal domain ($\Delta C$) embedded with the explicit POPC lipid bilayer were generated to investigate the impact of CTD on interactions between the two VSD units. Here, the simulations were conducted with different FL models, including (1) mHv1cc-FL, (2) hHv1-FL, (3) hHv1-FL with a disulfide bond of C249 pair (hHv1ss-FL), (4) combined mHv1cc-VSD/hHv1-CTD (mHv1cc-VSD/hHv1-CTD) models, and with $\Delta C$ models: (5) mHv1cc-$\Delta C$ and (6) hHv1-$\Delta C$ models. The dimeric Hv1-FL channel contains the two-stranded coiled-coil structure of CTDs. The hydrophobic residue contacts and other potentially inter-subunit interactions on the coiled-coil structure might contribute to the conformational stability and dimerization of the Hv1 channels. [55-57] Figure 2.2A, B clearly shows the effect of the coiled-coil structure on the contact of the two VSD units. The MD results show that the distances between S1, S2, and S4 and their identical transmembrane segments are relatively farther apart in the absence of CTD. The dimeric VSD units of the Hv1-FL channel are more close contacts than that of the $\Delta C$ channel. As expected, the $\Delta C$ models have high RMSD values for the VSD part (Figure 2.2C). In addition, the overall RMSD values of CTD are significantly lower than those of the VSD region, which reflects the stability of this motif. This suggests that the

interactions at the CTD contribute significantly to dimer stabilization. Clearly, the C-terminal region is the important structural motif involved in dimerization.
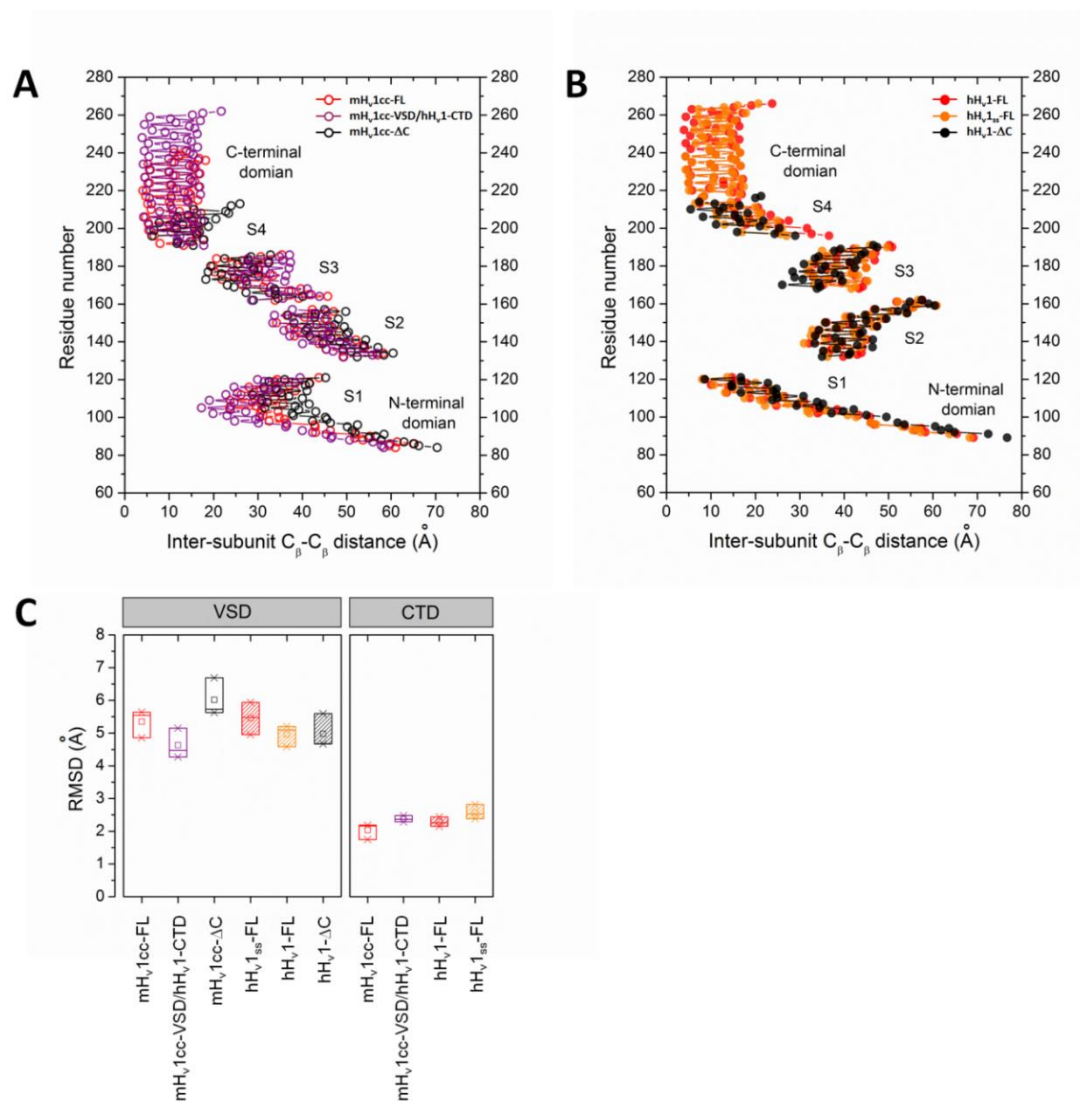


**Figure 2.2**      The average $C_\beta - C_\beta$ inter-subunit distance profiles of (A) mHv1cc and (B) hHv1 models. (C) Statistical box plots of backbone RMSD values for the voltage sensing domain (VSD) and C-terminal domain (CTD). The boxes are 25-75 % with the median as horizontal line; the little squares represent mean values, and the asterisks show the 1–99% of percentile. The data obtained from three independent simulations for each system averaged over 80-100 ns.

2.4.3 Fundamental Interactions of the Key Residues for Dimer Stability

To gain a deep understanding of the dimer stabilization mechanism, 100 ns MD simulations of several dimers of isolated Hv1-CTD were conducted. From Figure 2.2C, the CTD of the full-length models, i.e., mHv1cc-FL and hHv1-FL, exhibits the RMSDs of 2.0–3.0 Å, which falls within the same RMSD range of the only CTD model, as shown in Figure S2.3. This clearly suggests that the simulations with and without the VSD did not significantly alter the structure and dynamics of CTD dimer. The side-chain dynamics and interactions of conserved underpinning amino acids on the coiled-coil motif were then analyzed in more detail. The strategy was to determine the influence of the underpinning amino acids on CTD assembly. We primarily focused on the pairs of oppositely charged residues that participate in the formation of the interhelical salt bridge across the dimer interface, as shown in Figure 2.1B, C. MD simulations at different temperatures and pH conditions were carried out with an attempt to understand the impact of salt-bridge interaction on dimer stability. In the mHv1cc-CTD model, two salt-bridge pairs, E235–K230′ and E235′–K230, wrap around the long hydrophobic core of dimeric mHv1cc-CTD (Figure 2.1B), whereas hHv1-CTD model consists of two "upper" (E246–K241′ and E246′–K241) and two "lower" (E253–K252′ and E253′–K252) salt-bridge pairs (Figure 2.1C) (a prime ′ denotes the residues on the opposing subunit). However, not every charged residue on the C-terminal region can form salt-bridges, depending on the proximity between the salt-bridge partners.
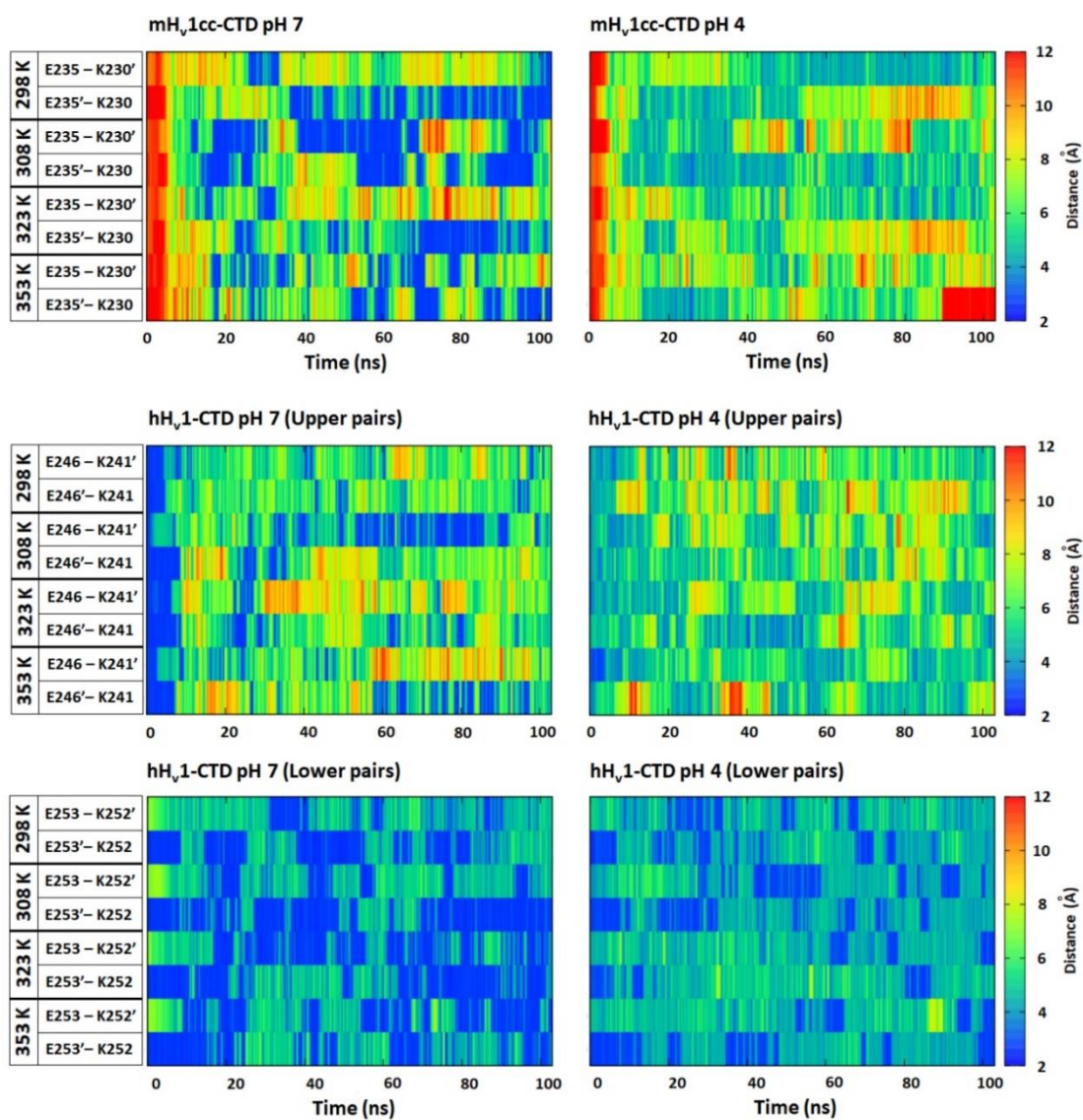
Figure 2.3     Time evolution of distance profiles at different temperatures involved with salt-bridge residues on CTD in mHv1cc-CTD model (E235–K230' and E235'–K230) and hHv1-CTD model for upper pairs (E24 –K241' and E246'–K241) and lower pairs (E253–K252' and E253'–K253). A prime (') denotes the residues on an opposing subunit. The distance between O–N atoms at pH 7 (left) and 4 (right) was measured and plotted in different shades of color scale. Blue color indicates the forming of salt-bridge interactions and changing in appearance of blue color reflects the swapping dynamics of the salt-bridge pairs.

An analysis of the O–N distances revealed that each salt-bridge pair is repeatedly formed and disrupted (Figure 2.3). We also found that the salt-bridges were swapped between each pair throughout the trajectories. The salt-bridge pair swapping of the mHv1cc-CTD and hHv1-CTD models occurs more frequently when the temperature increases (Figure 2.3). The change in salt-bridge interactions has been continuously observed during the course of simulations for all models (Figures S2.4 and S2.5). To determine the frequency of the two salt-bridge pairs being formed or disrupted, we introduced subinteracting states (see Methods) that are $\nu$ for noninteracting, $\sigma$ or $\sigma^*$ for single-interacting, and $\delta$ for double-interacting states (Figure 2.4A).
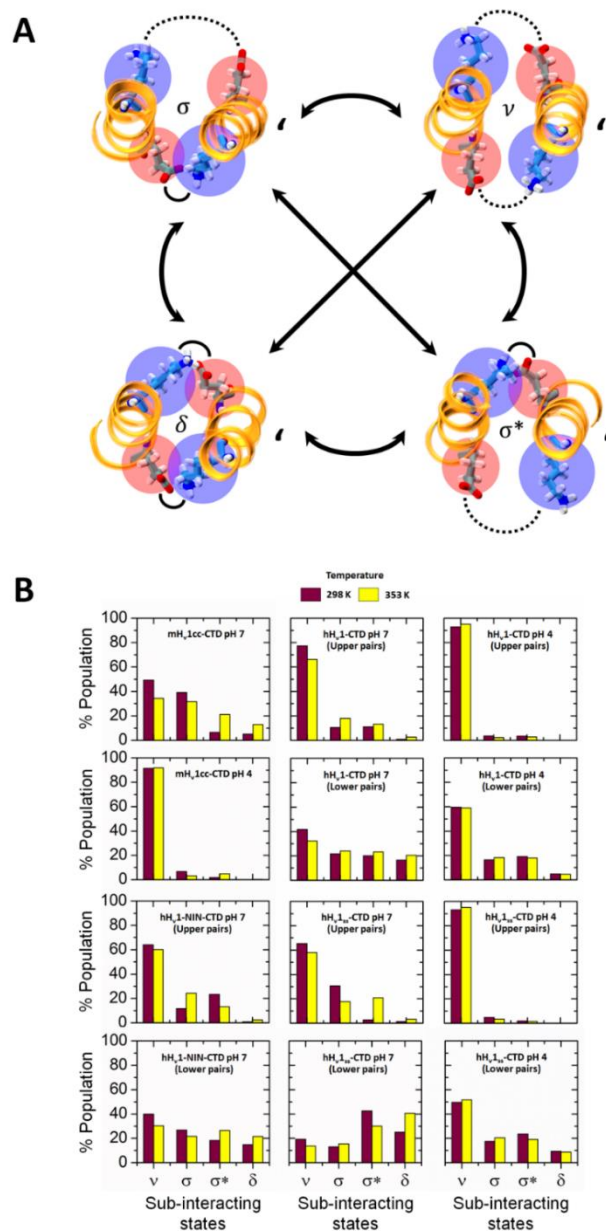
**Figure 2.4**    Dynamics of salt-bridge residues along the MD simulations. (A) observed movement of salt-bridges residues causes by oppositely charged residues where Glu and Lys residues are shaded with red and blue sphere, respectively. A prime (') denotes an opposing subunit. Each conformation independently changes to any sub-interacting state as indicated with arrows. An interacting pair is shown in a solid line with overlapping spheres. (B) Population of sub-interacting states associated with interaction of salt-bridge residues was counted over the entire simulation at 298 K and 353 K.

The frequency of the systems visiting the individual subinteracting state is shown in Figure 2.4B. MD data of all of the Hv1-CTD, except for hHv1ss-CTD pH 7 (lower salt bridges), revealed that the  -state was the dominant conformation, whereas $\sigma$-, $\sigma^*$-, and $\delta$-states were detected with a relatively smaller quantity (Figures 2.4B and S2.6). Especially, the upper salt-bridge pairs of hHv1-CTD at neutral pH are mostly unbound, suggesting that the upper pairs form very weak electrostatic interactions. In contrast, the lower pairs are stronger when compared with the upper pairs at the same temperature and pH. The O–N distances of the lower pairs are, on average, significantly shorter than those of the upper pairs (Figures S2.4 and S2.5), and therefore they made interactions with a greater frequency of $\sigma$-, $\sigma^*$-, and $\delta$-states than those observed from the upper pairs (Figures 2.4B and S2.6). Interestingly, all of the simulated models exhibit a pH-dependent salt-bridge interaction. At acidic pH, the salt-bridge pairs change from interacting to noninteracting states because of a protonation of the Glu residues preventing a salt-bridge formation. Particularly, the subinteracting state of mHv1cc-CTD and of the upper pairs of hHv1-CTD shifts to the $\nu$-state for almost the entire trajectory (Figures 2.4B and S2.6), suggesting that these interactions become weak in low pH medium. The $\Delta\Delta G_{bind}$ bind of mHv1cc-CTD and hHv1-CTD also displayed higher stability at pH 7 than at pH 4 (Table 2.1), implying that an increase of dimer stability is associated with salt-bridge interactions. Apparently, the electrostatic interactions become energetically unfavorable at pH 4 by a significant increase in $\Delta\Delta G_{bind}^{elec}$. Although $\sigma$-, $\sigma^*$-, and $\delta$-states at acidic pH were still minorly observed, interactions of these salt-bridge pairs are nonelectrostatic. In our simulations, each salt-bridge pair exhibits different subinteracting states, whereas the $\nu$-state is more frequently observed in the Hv1-CTD. The results are in good agreement with the mixed subinteracting state observed from dimer-of-dimers in Hv1-CTD crystal structures (Table S2.1). We also found that these residues adopted different conformations consistent with anisotropic atomic fluctuation (*B*-factor plots in Figure S2.7) extracted

from the crystallographic data. In the hHv1-CTD, only lower salt-bridge pairs appear to contribute to the stability of the dimer rather than the upper pairs. At pH 7, the lower pairs of hHv1-CTD display substantial levels of salt-bridge interactions similar to that of mHv1cc-CTD (with about 50% or more of occurrence frequency for all combined interacting states). The upper salt-bridges do not play the main factor for dimer stability. In addition, we did not find any correlation coupling between interactions of the upper and lower pairs in the hHv1-CTD model (Figure S2.8). The subinteracting states of the upper and lower pairs are independent from each other.
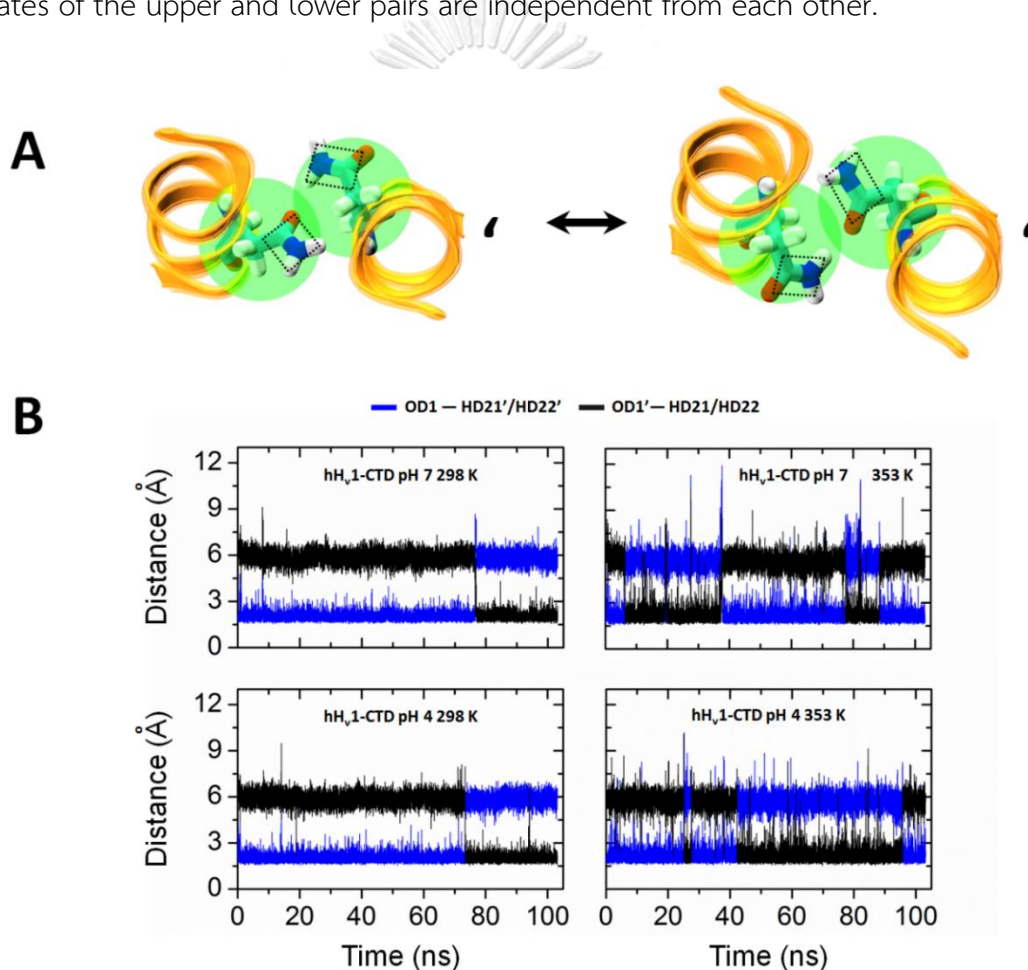


Figure 2.5    Inter-helical hydrogen bond at the coiled-coil core of hHv1-CTD model involving the N235 pair. (A) amide-plane flipping of N235 side chains indicated with dash line causes swapping of hydrogen bond across the dimer interface. A prime (') denotes an opping subunit. (B) Evolution of distances between hydrogen bond acceptor-donor pairs belongs to N235 residues at 298 K and 353 K.

In the hHv1-CTD, two helices in a coiled-coil motif are mainly held together by hydrophobic interactions. The binding free energies show that for interactions at the CTD, nonpolar contributions ( $\Delta\Delta G_{bind}^{np}$ ) are energetically favorable, whereas the electrostatic contributions are less favorable (Table 2.1). Therefore, interactions between hydrophobic residues in the CTD are important for dimer stability. Nevertheless, the wildtype hHv1-CTD, is composed of hydrophilic N235 residue (Figure 2.1C) instead of hydrophobic residue at the a position in the heptad repeating pattern (Figure 2.1A). This unusual hydrophilic Asn residue, which is conserved among Hv1 channel species, interacts to another Asn of the opposite chain via a hydrogen bond (Figure 2.5A). The MD results show short and long OD1–ND2 side-chain distances of the two N235 residues, suggesting that only one hydrogen bond forms between the hHv1-CTD (Figure 2.5B). Nevertheless, the hydrogen bond can be alternated between the two N235 side chains by amide-plane flipping (Figure 5A). The hydrogen bond switching process becomes faster with increasing temperature (Figures 2.5B and S2.9). The results observed from the simulations are highly consistent with those found in the crystal structures of both human and mouse Hv1-CTD (Table S2.1). It has been proposed that N235 and its neighboring isoleucine residues could destabilize the CTD packing by disrupting the continuity of the periodic hydrophobicity in the coiled-coil structure along the dimer interface. Fujiwara et al. demonstrated that the triple mHv1-NIN mutant which adds a better continuity to the periodic hydrophobicity in the structure enhances dimer stability. To validate this idea, a model of triple NIN mutation of hHv1-CTD (hHv1-NINCTD) was built by replacing M234–N235–V236 with N234–I235–N236. The molecular surface analysis showed the hHv1-NIN-CTD mutant has a more complete hydrophobic core along the dimer interface (Figure S2.10). From our simulations, the hHv1-NIN-CTD mutant has an effect on dimer interaction, as shown

by a decrease in the $\nu$-state of the salt-bridge (Figures 2.4B and S2.6). In addition, the $\Delta G_{bind}$ of hHv1-NIN-CTD mutant is energetically more stable than that of hHv1-CTD (Table 2.1). A uniformly hydrophobic core between the dimer interface increases the stability of dimer by a significant increase in $\Delta \Delta G_{bind}^{np}$ and a decrease in $\Delta \Delta G_{bind}^{elec}$. The results support the hypothesis of "hydrophobic continuity" by Fujiwara et al. [26]

**Table 2.1**$^a$    Average binding free energies with separated components (kcal·mol$^{-1}$), mid-point dissociative temperatures (K) and thermodynamic parameters (kcal·mol$^{-1}$) of mHv1cc-CTD and hHv1-CTD models.

| Model | pH | $\Delta \Delta G_{bind}$ | $\Delta \Delta G_{bind}^{elec}$ | $\Delta \Delta G_{bind}^{np}$ | $T_m$ | $\Delta H^o$ | $T \Delta S^o$ | $\Delta G^o$ | $\Delta G^{o\ddagger}$ |
|---|---|---|---|---|---|---|---|---|---|
| mHv1cc-CTD | 7 | -23.4 | -3.4 | -20.4 | 427 | 12.5 | 4.8 | 7.7 | 17.9 |
| | 4 | 6.4 | 24.6 | -18.1 | 413 | 16.0 | 7.7 | 8.4 | 12.8 |
| hHv1-CTD | 7 | -2.9 | 23.0 | -25.9 | 399 | 16.3 | 9.0 | 7.3 | 7.2 |
| | 4 | 7.1 | 31.9 | -24.8 | 378 | 20.3 | 12.5 | 7.8 | 5.4 |
| hHv1-NIN-CTD | 7 | -11.0 | 17.7 | -28.7 | 401 | 15.8 | 8.4 | 7.4 | 8.2 |
| hHv1ss-CTD | 7 | -18.6 | 6.8 | -25.4 | 416 | 14.0 | 6.5 | 7.5 | 17.2 |
| | 4 | -19.4 | 5.0 | -24.4 | 415 | 14.8 | 7.3 | 7.5 | 10.7 |

$^a$Average binding free energies of dimers of isolated CTD, $\Delta \Delta G_{bind}$, are calculated from non-linear Poisson-Boltzmann equation with representative MD snapshots in 10–100 ns at 298 K. $\Delta \Delta G_{bind}^{elec}$ and $\Delta \Delta G_{bind}^{np}$ are electrostatic and non-polar components of the average binding free energies. $T_m$ is the mid-point dissociative temperature obtained from fitting a sigmoidal curve of contact fraction. $\Delta H^o$, $T \Delta S^o$ and $\Delta G^o$ are the thermodynamic equilibrium associated with dimer dissociation calculated from eq 2.7. equation. $\Delta G^{o\ddagger}$ are the standard free energy barrier of dimer dissociation obtained from Eyring equation.

In the wild-type, the hydrogen bond between the polar side chains of N235 disrupts the continuity of hydrophobic packing of the heptad pattern. This disruption could be involved with the channel activity by increasing the mobility of the coiled-coil structure. It has been reported that the activation kinetics of mHv1-NIN mutant was slowed down and a higher temperature was required when compared with wild-type. [26] Therefore, the presence of the unusual hydrophilic Asn could affect activation kinetics of the channel by modulating the stability or mobility of the coiled-coil structure. Taken together with simulation results, they imply that packing of two subunits not only contributes to structural stability but also relates to the activation process.

Another extraordinary feature of the hHv1-CTD model is the presence of a disulfide bond formed by the two C249 residues at the core of the coiled-coil structure (Figure 2.1C). The Hv1 channel commonly expresses in immunocytes, [2] and it is actively involved in the production of reactive oxygen species that causes fluctuation of the redox state. The C249 has been considered as a redox sensor of the Hv1 channel and has participated in dimer stabilization by shortening two neighboring subunits with a disulfide bond. [29] This crucial finding shed light on the function of the C249. The simulations of the hHv1-CTD model containing a Cys disulfide bond (the hHv1ss-CTD) show the greatest negative values of $\Delta\Delta G_{bind}$, suggesting tight interactions in the dimer (Table 2.1). We found that the $\Delta\Delta G_{bind}^{np}$ of the hHv1ss-CTD is about the same as that of the hHv1-CTD but remarkably lower when compared to that of the $\Delta\Delta G_{bind}^{elec}$. Unfavorable electrostatic interactions decrease due to a strong disulfide bond promoting salt-bridge interactions (more $\sigma$-, $\sigma^*$-, and $\nu$-states at pH 7) (Figures 2.4B and S2.6) or making the hydrophobic core to minimize its exposure to water. It should be noted that the $\Delta\Delta G_{bind}$ obtained at neutral and low pH are not significantly different in the presence of the Cys disulfide bond. This suggests that the disulfide

bond between the two C249 residues has a major impact on the dimer stability, consistent with experiments reported previously. [29]

An analysis of helix–helix crossing angles shows a broader distribution for the shorter helix in the mHv1cc-CTD when compared with the longer helix of the hHv1-CTD (Figure S2.11). However, the crossing angles of mHv1cc-CTD, hHv1-CTD, hHv1-NIN-CTD, and hHv1ss-CTD are not altered between the two pH conditions. This suggests that the helix packing geometry of Hv1-CTD is largely unaffected by pH. This should be clearly explained by the primary role of CTD for dimerization. We speculate that upon pH activation, the S4 movement does not significantly induce conformational changes in the CTD region.

2.4.4 Thermal-Induced Dissociation of Dimer

The next focus is on the temperature-induced dimer dissociation of Hv1-CTD. Temperature-accelerated MD simulations were carried out to explore loss of dimer contact on computationally accessible time scales. [58, 59] Each simulation was performed at different temperatures. A thermal dissociation profile was constructed from trajectories using contact fraction analysis that is derived from inter-subunit contact (see Methods). Using a similar strategy as described previously, the dissociation process was observed with the various factors affecting a thermal stability of the dimer structure. The thermal dissociation profiles are shown in Figure 2.6A, B, respectively. Fitting a sigmoidal curve of the data yields midpoint dissociative temperatures ($T_m$) of the dimer-to-monomer transition, providing the information on the thermal stability. There are significant differences in the $T_m$ values, as listed in Table 2.1. It should be noted that $T_m$ obtained in this study does not reflect the actual value because the

dimer dissociation simulations were accelerated by unnaturally high temperatures. We therefore use the relative values for comparison. The dissociation behavior of the mHv1cc-CTD and hHv1-CTD models appears to be dependent upon the pH value. Contact fractions rapidly dropped as a function of temperature, leading to the monomer state becoming decrement in pH (Figure 2.6A, B).
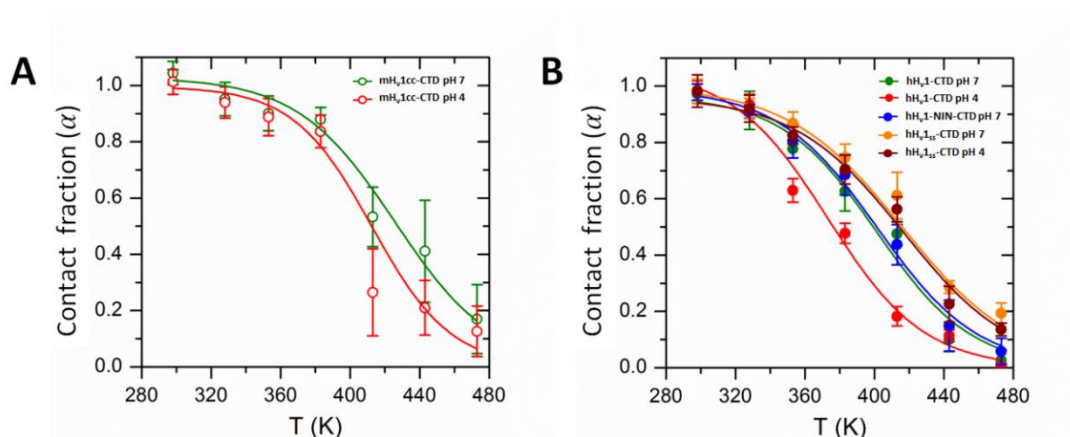


**Figure 2.6** Thermal stability of CTD under different conditions. Thermal dissociation profiles of (A) mHv1cc-CTD and (B) hHv1-CTD models are plotted with contact fraction ($\alpha$) calculated from inter-subunit contacts against temperature. Each point represents the average value ± SD of contact fraction obtained from MD trajectories.

The $T_m$ values at neutral pH are markedly higher than the others at acidic pH and exhibit a similar trend of the $\Delta\Delta G_{bind}$, as shown in Table 2.1. Conversely, Zhao et al. [30] proposed that the thermal stability in coiled-coil unfolding from the CD spectra decreased with an elevating pH. However, it must be noted that our approach used a different definition in the determination of the dimer character from the inter-subunit contacts rather than the quantity of the unfolded structure. The $T_m$ could also change upon hydrophobicity of core packing. The NIN mutation shifted $T_m$ to a higher temperature approximately 5 K away from the hHv1-CTD at pH 7 [26] (Table 2.1). Interestingly, hHv1ss-CTD containing the disulfide bond is thermally very stable, [29]

especially at pH 7, in combination with interhelical salt bridges giving the highest $T_m$ (Table 2.1). To thermodynamically substantiate the effect of interactions on dissociation, a two-state model was introduced to approximate the thermodynamic equilibrium of dimer dissociation. [50, 51] Subsequently, thermodynamic parameters were determined using the eq 2.7 on the basis of the two-state scheme (see Methods). As a result of the plot, the dimer dissociation leads to an endothermic process with a favorable entropy, as shown by positive $\Delta H^o$ and $\Delta S^o$ respectively, (Table 1). The dissociation of dimeric CTDs is apparently a nonspontaneous process due to positive $\Delta G^o$. On the basis of current evidence, perhaps interplaying interactions among these residues is exhibited as a key to regulate the degree of dimer dissociation upon the environmental conditions.

### 2.4.5 Dissociation Kinetics of C-Terminal Domains

The thermodynamics and kinetics of dimer dissociation were further characterized. The standard free energy of activation ($\Delta G^{o\ddagger}$) was obtained from the temperature dependence of the dissociation rate constant using the Eyring equation (see Methods). [51, 60] As summarized in Table 2.1, $\Delta G^{o\ddagger}$ estimated from slope of graphs fell within the range of 8.2–17.9 kcal·mol$^{-1}$. A sequence of the calculated activation energy at pH 7 for the hHv1-CTD model is in the following order: hHv1ssCTD > hHv1-NIN-CTD > hHv1-CTD. Notably, the $\Delta G^{o\ddagger}$ for dissociation of human Hv1-CTD with a disulfide bond is about two times greater than that without disulfide bond. This suggests that the cysteine disulfide bond has a remarkable impact on the CTD packing. Considering the pH effect on dimer stability, the structure is more difficult to break at neutral pH than at acidic pH. with $\Delta G^{o\ddagger}$ of about 1.8–6.5 kcal·mol$^{-1}$ higher for both the mHv1cc-CTD and hHv1-CTD models. The NIN triple mutation raises the free-energy barrier when compared with hHv1-CTD. It is clear that a more hydrophobic character

of the core packing allows a stronger stability of the dimer. [26] Additionally, the $\Delta G^{o\ddagger}$ value was strongly dependent on the contribution of a disulfide bond (hHv1ss-CTD). The free-energy barrier of dimer dissociation dramatically shifted upward by a disulfide bond when compared with hHv1-CTD and hHv1-NIN-CTD reflecting an increase in coiled-coil stability [29] (Table 2.1). With such a positive activation free energy, these contributions on CTD absolutely disfavor the dissociation process. The transition of dimer to monomer seems to be mainly confined by the interhelical interactions. A large impact of contributions involving these residues on CTD could potentially modulate the dimer dissociation. As supported by several pieces of information, the assumption on the critical role of the key residues relating with dimer stability could be confirmed. This should be a useful piece of evidence for providing a more structurally informative pathway of kinetic dissociation.

## 2.5 CONCLUSION

The simulation results reveal the behaviors of the key residues influenced on binding strength between two subunits. Evidently, the salt-bridge residues on CTD can modulate the packing stability of the dimer structure upon pH condition. In hHv1-CTD model, the hydrogen bond switching of side-chain N235 pair clearly destabilizes the dimer structure by disruption of hydrophobic packing. In contrast, stability of NIN mutant with a more uniformly hydrophobic core is enhanced. The CTD with disulfide bond shows the highest stability, which emphasizes that C249 is a major contributor for dimer stabilization. Thus, our study provides insight into interactions of key residues of the Hv1 C-terminal region, emphasizing the ultimate importance of this part of the structure and its links to channel activity.

## 2.6 SUPPLEMENTARY INFORMATION
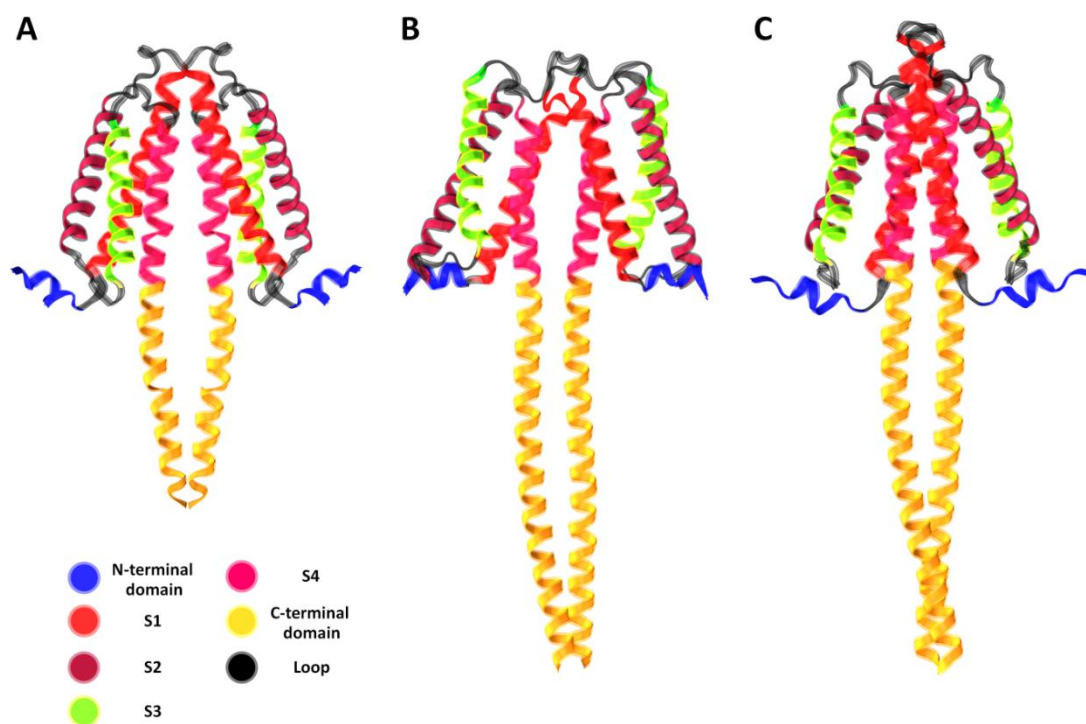


**Figure S2.1**    Model comparison of dimeric Hv1 channels. (A) X-ray crystal structure of a chimeric protein of the mouse Hv1 channel (mHv1cc) (B) EPR refined homology model of the hHv1 (C) combined model of mHv1cc-VSD/hHv1-CTD was built using VSD (S1-S4) and N-terminal domain of mHv1cc connected with CTD of hHv1. Each segment N-terminal domain (blue), S1 (red), S2 (brown), S3 (green), S4 (magenta) and C-terminal domain (yellow) are shown in helix linked with loop (black).
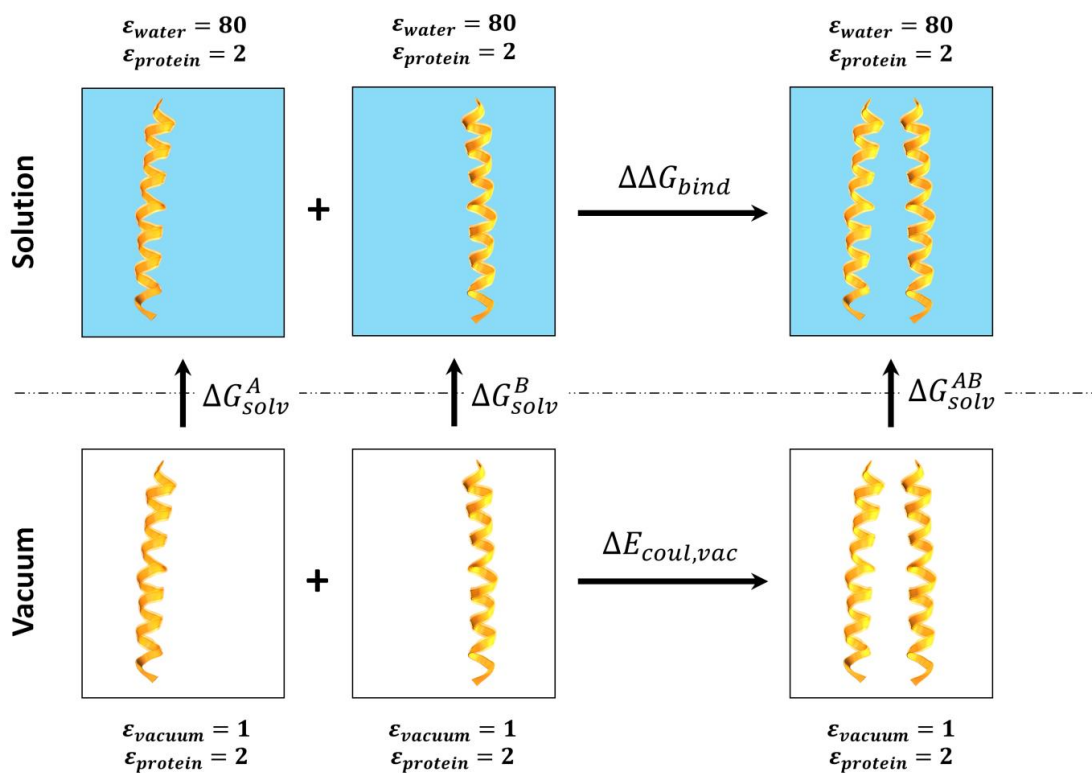
**Figure S2.2** Schematic diagram of thermodynamic cycle for protein-protein binding free energy. The implicit solvation model was used to calculate the binding free energy of dimeric Hv1-CTD. Surrounding dielectric medium of protein is shaded in different colors. Water medium is blue and vacuum is white.
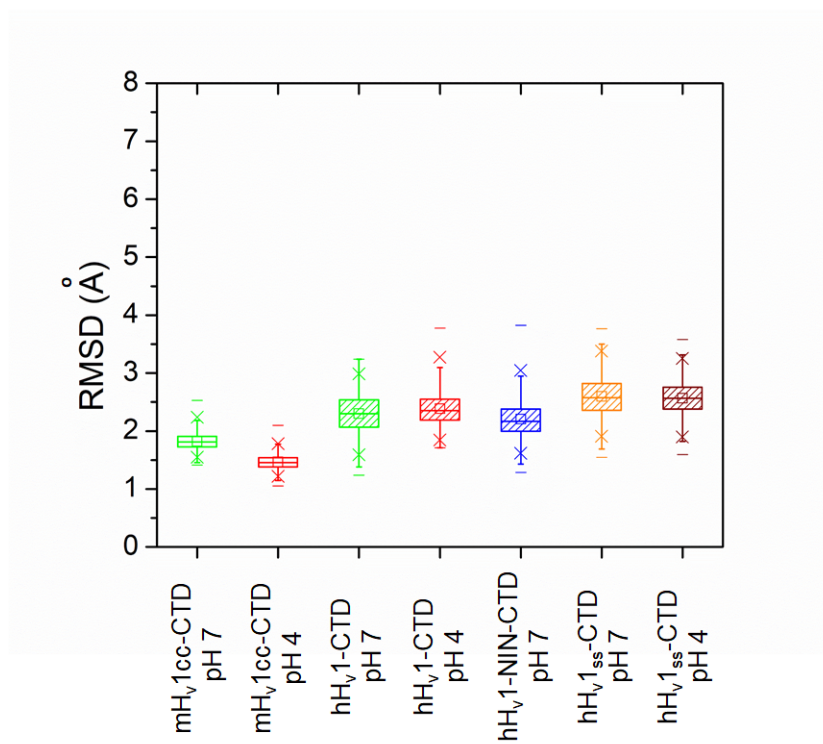
**Figure S2.3**    Statistical box plots of backbone RMSD values for dimers of isolated Hv1-CTD simulation systems at 298 K in different pH conditions. The mHv1cc-CTD and several hHv1-CTD models were used. The hHv1-NIN-CTD is mutant of the hHv1-CTD and the hHv1ss-CTD is hHv1-CTD with disulfide bond. The boxes are 25-75 % with the median as horizontal line; the little squares represent mean values, and the asterisks show the 1–99% of percentile. The data were averaged over 80-100 ns.
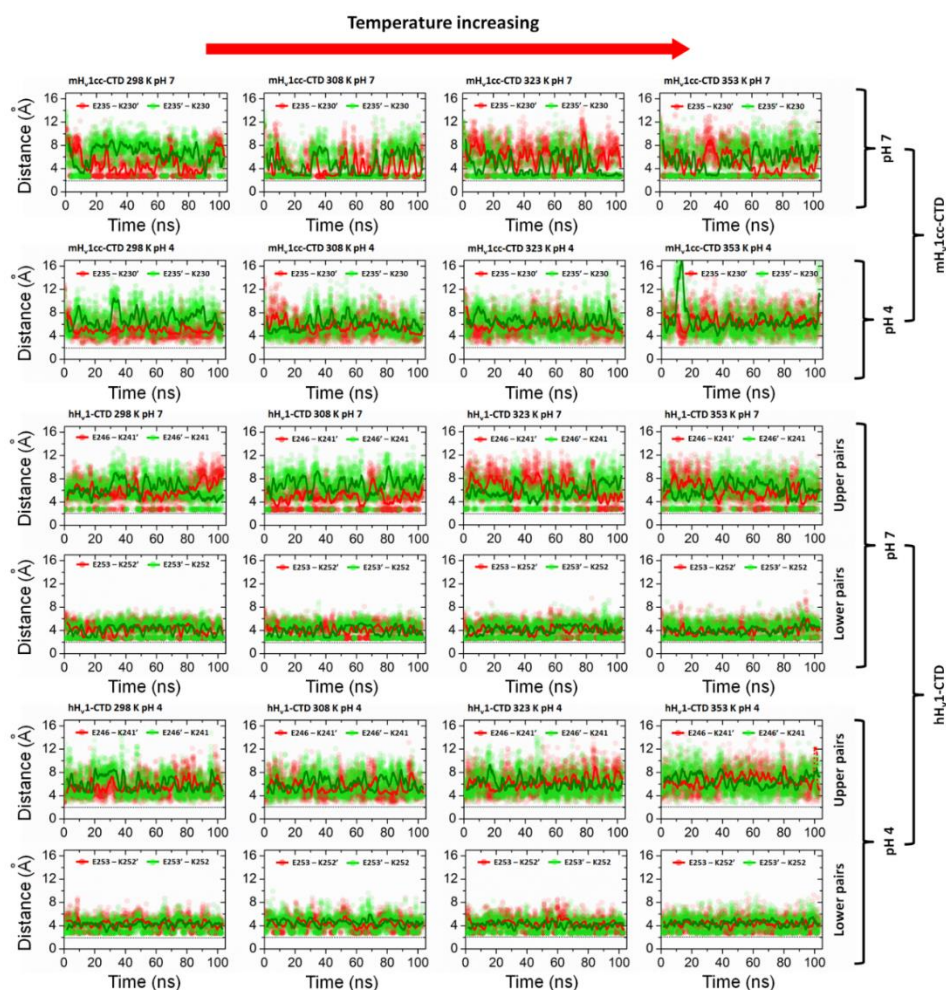
**Figure S2.4**     Distance profiles of salt-bridges residues with average values. Evolution of O—N distance in alternative view was plotted as a dot shown in red and green colors for different pairs of salt-bridge residues, respectively. The solid red and green lines are the moving average value of the distances for each pair. Horizontal dash line is a lower limit distance of interaction. A prime (') denotes the residues on an opposing subunit. The distance profiles of mHv1cc-CTD model (E235–K230' and E235'–K230) are similar to the upper pairs (E246–K241' and E246'–K241) of hHv1-CTD model with a large fluctuation. In the case of the lower pairs (E253–K252' and E253'–K253) of hHv1-CTD model, the distance profiles of salt-bridge residues fluctuate in a smaller range throughout the trajectories because each residue causes the closer interaction with its partner. Thus, the lower pairs of salt-bridge residues should bind tighter than the upper pairs.
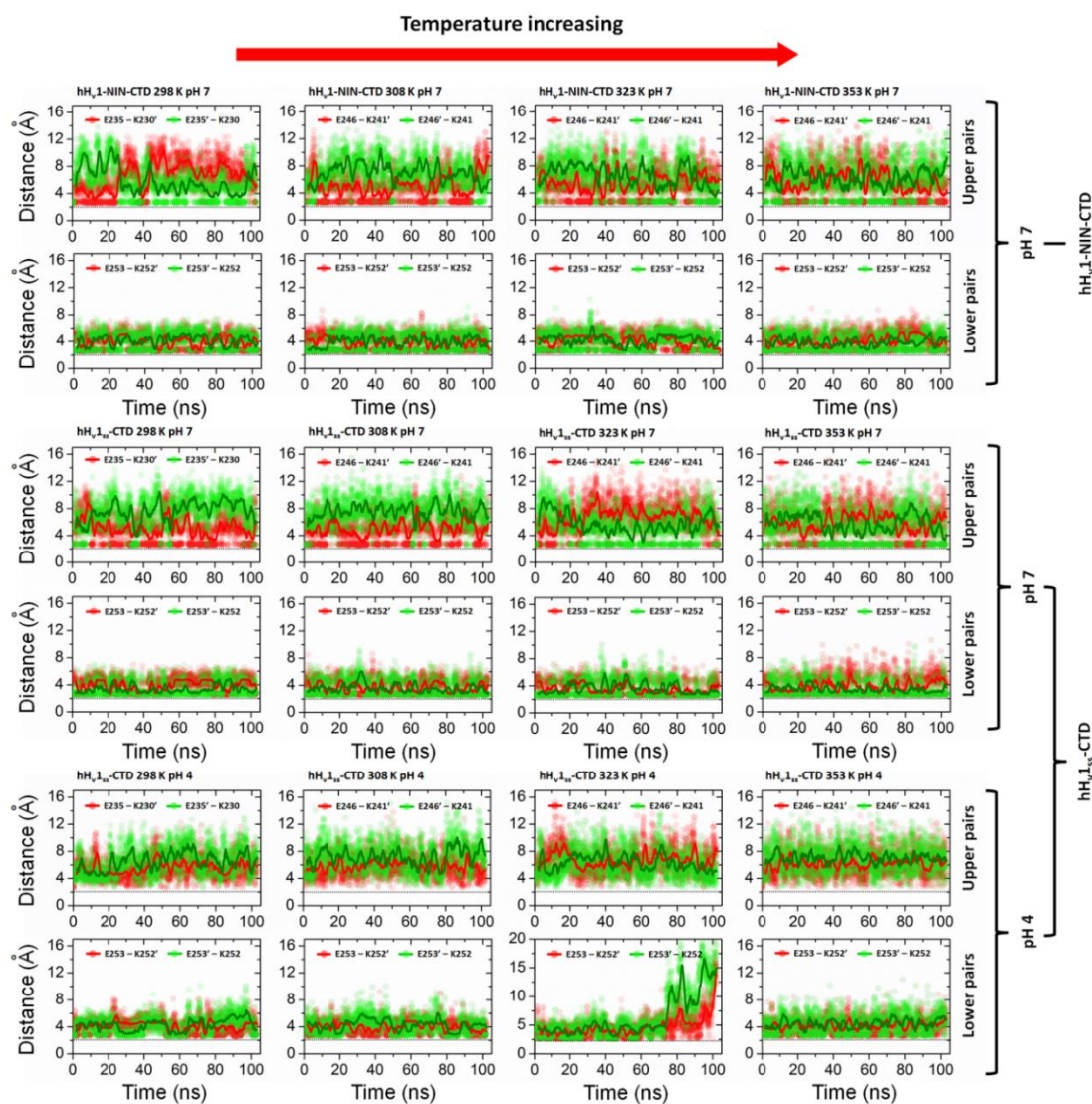
**Figure S2.5**    Distance profiles of salt-bridges residues with average values (continued). Evolution of O—N distance in alternative view was plotted as a dot shown in red and green colors for different pairs of salt-bridge residues, respectively. The solid red and green lines are the moving average value of the distances for each pair. Horizontal dash line is a lower limit distance of interaction. A prime (') denotes the residues on an opposing subunit. The hHv1-NIN-CTD is mutant of the hHv1-CTD by replacement of M234–N235–V236 with N234–I235–N236 and the hHv1ss-CTD is hHv1-CTD with disulfide bond.
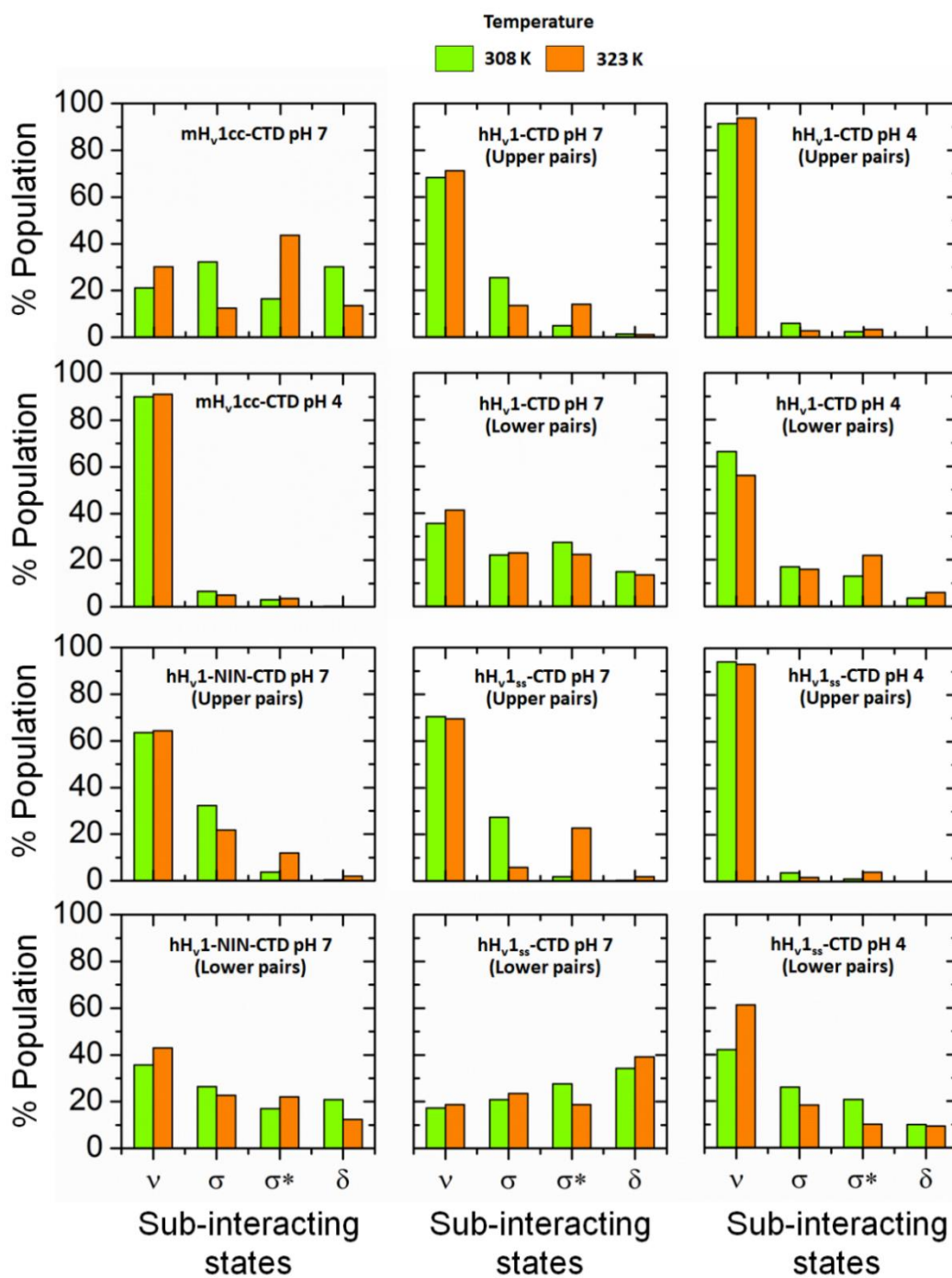
Figure S2.6    Population of sub-interacting states associated with interaction of salt-bridge residues counted over the entire simulation at 308 K and 323 K. The $\nu$-state is highly populated among other sub-interacting states.
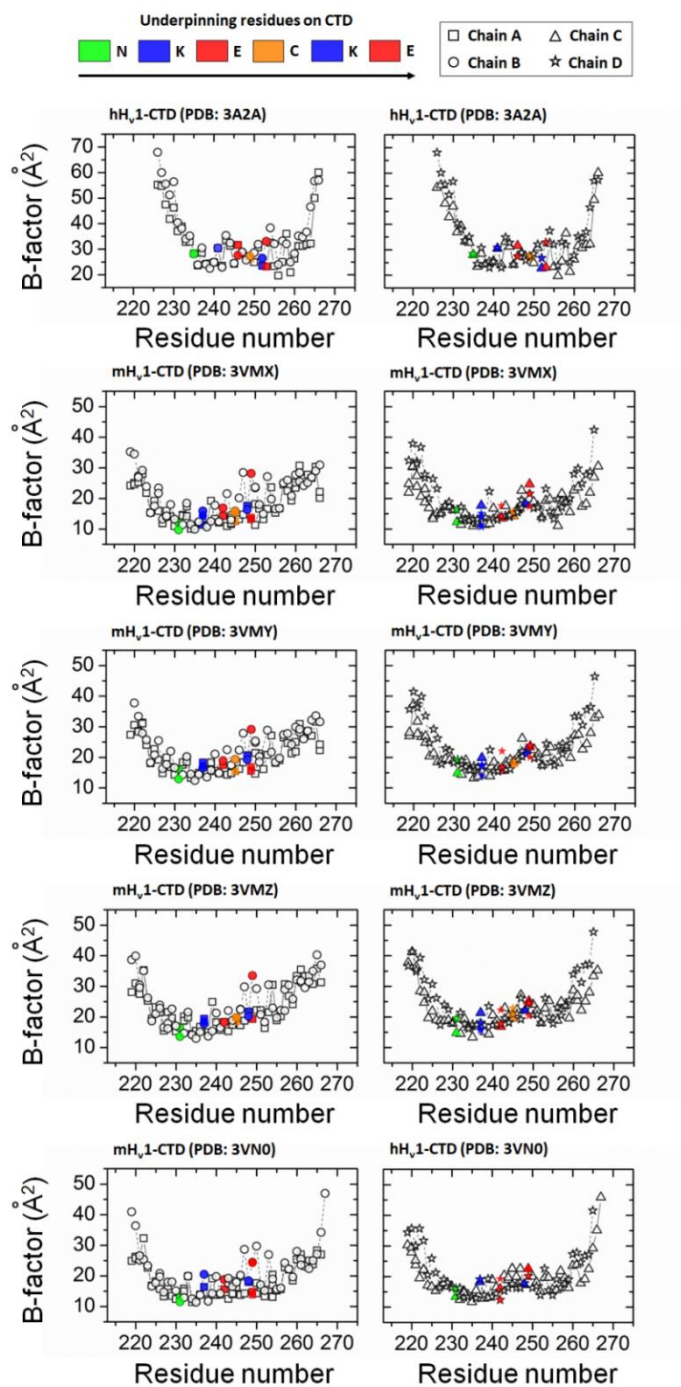
**Figure S2.7** *B*-factor plots extracted from previous X-ray crystal structures of isolated Hv1-CTD. The underpinning residues on CTD are labeled with different colors. Some residues might contain more than one value of *B*-factor as seen in plots due to diffraction anisotropy reflecting the high fluctuation.
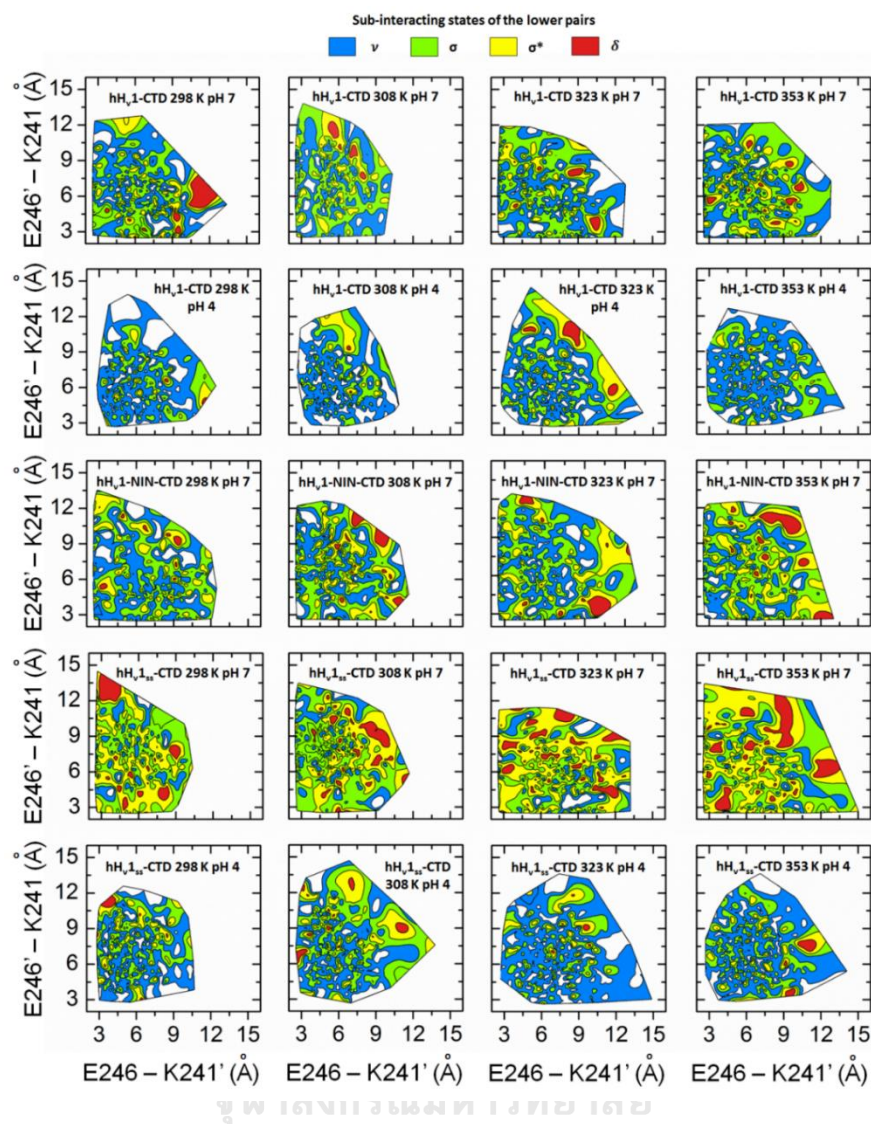
**Figure S2.8** Correlation of sub-interacting states between the upper and lower pairs associated with interaction of salt-bridge residues. Distances of the upper pairs (E246–K241' and E246'–K241) are plotted as a coordinate in graph within data boundary. Each region in graph is corresponding to the sub-interacting states of upper pairs with respect to the distance criteria (see Materials and Methods). The sub-interacting states of lower pairs, matching with distance coordinates of the upper pairs, are represented with different colors. The sub-interacting states of upper and lower pairs independently change each other. For a given sub interacting state of the upper pairs (considering in some region of a graph), the lower pairs is particularly not either $\nu$-, $\sigma$-, $\sigma^*$- or $\delta$-states (showing in different shades of color) as seen in graphs.
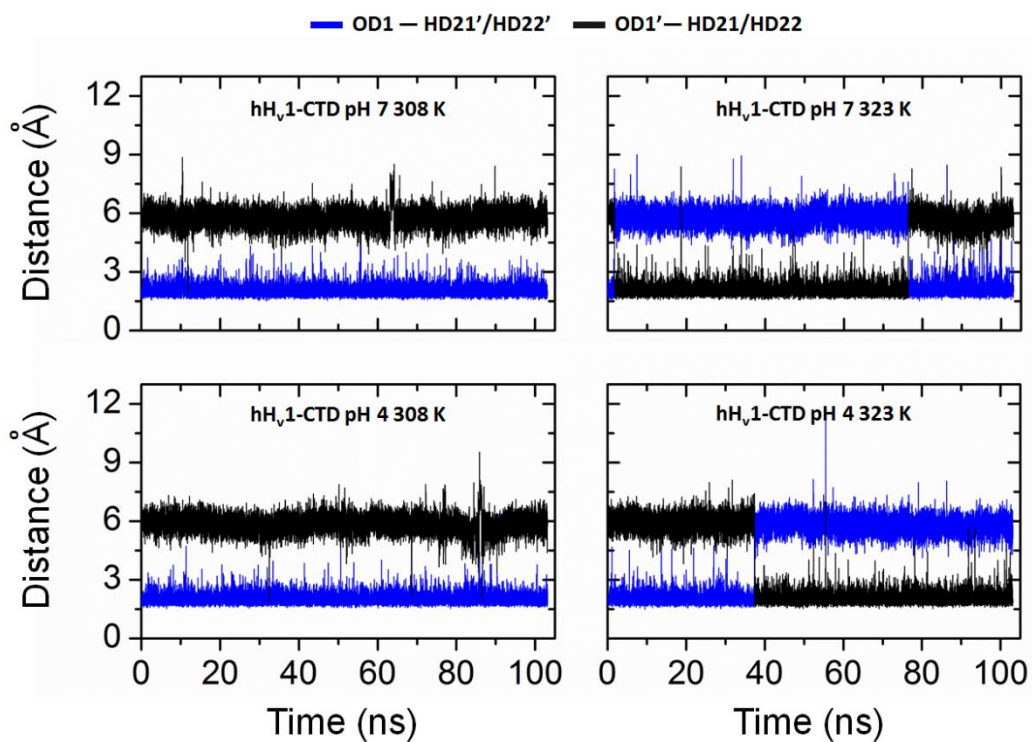
**Figure S2.9**     Inter-helical hydrogen bond of N235 pair. Evolution of distances between O–H pairs belong to N235 residues at 308 K and 323 K.

**90°**

K241 — Salt-bridge pairs (Upper pairs)
E246 —

C249 — Cys pair

K252 — Salt-bridge pairs (Lower pairs)
E253 —
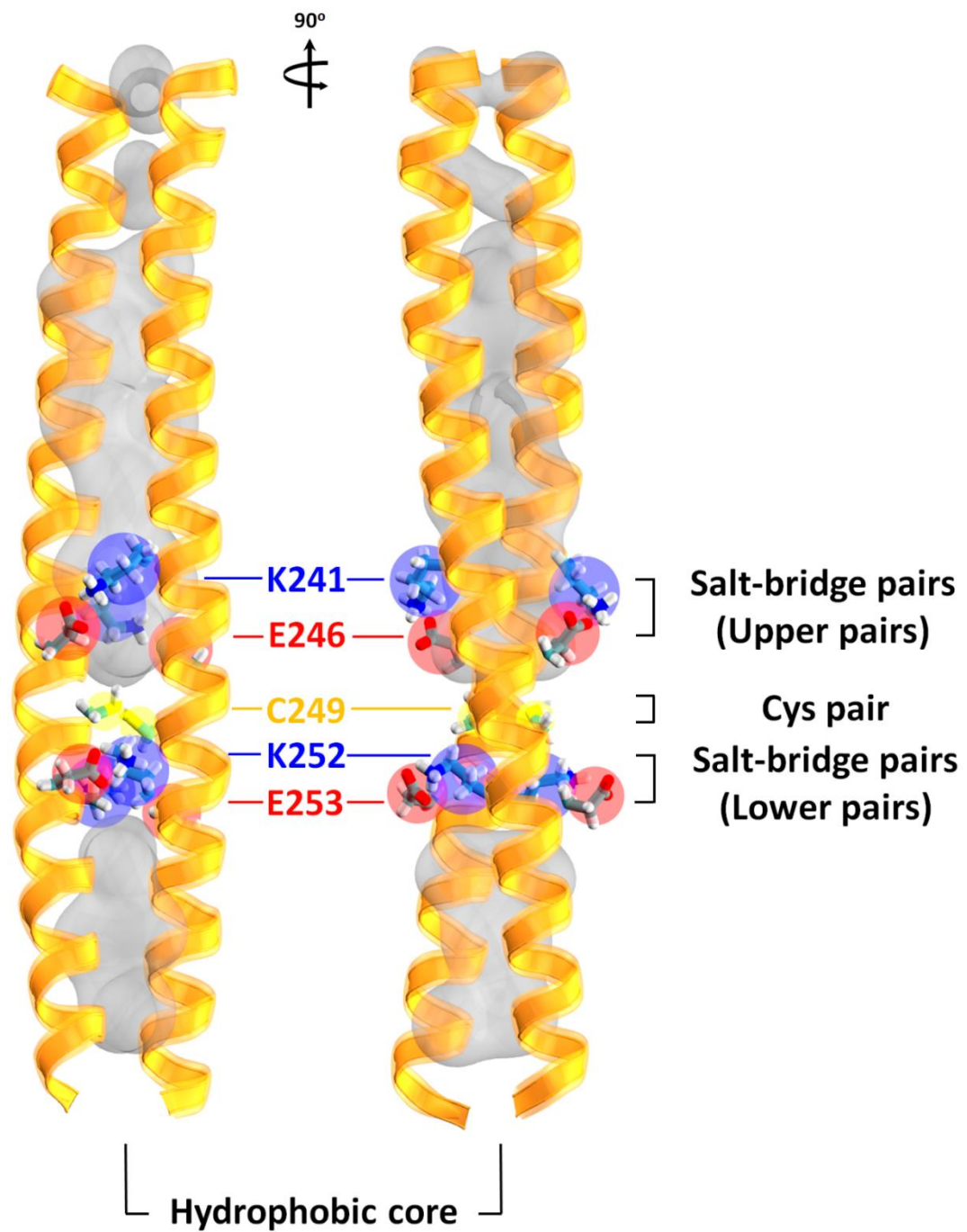
Hydrophobic core

**Figure S2.10**    Structure of hHv1H-NIN-CTD. Residues of M234—N235—V236 on coiled-coil structure were replaced by N234—I235—N236. The hydrophobicity of coiled-coil core is more complete as shown a grey quick surface in above figure.
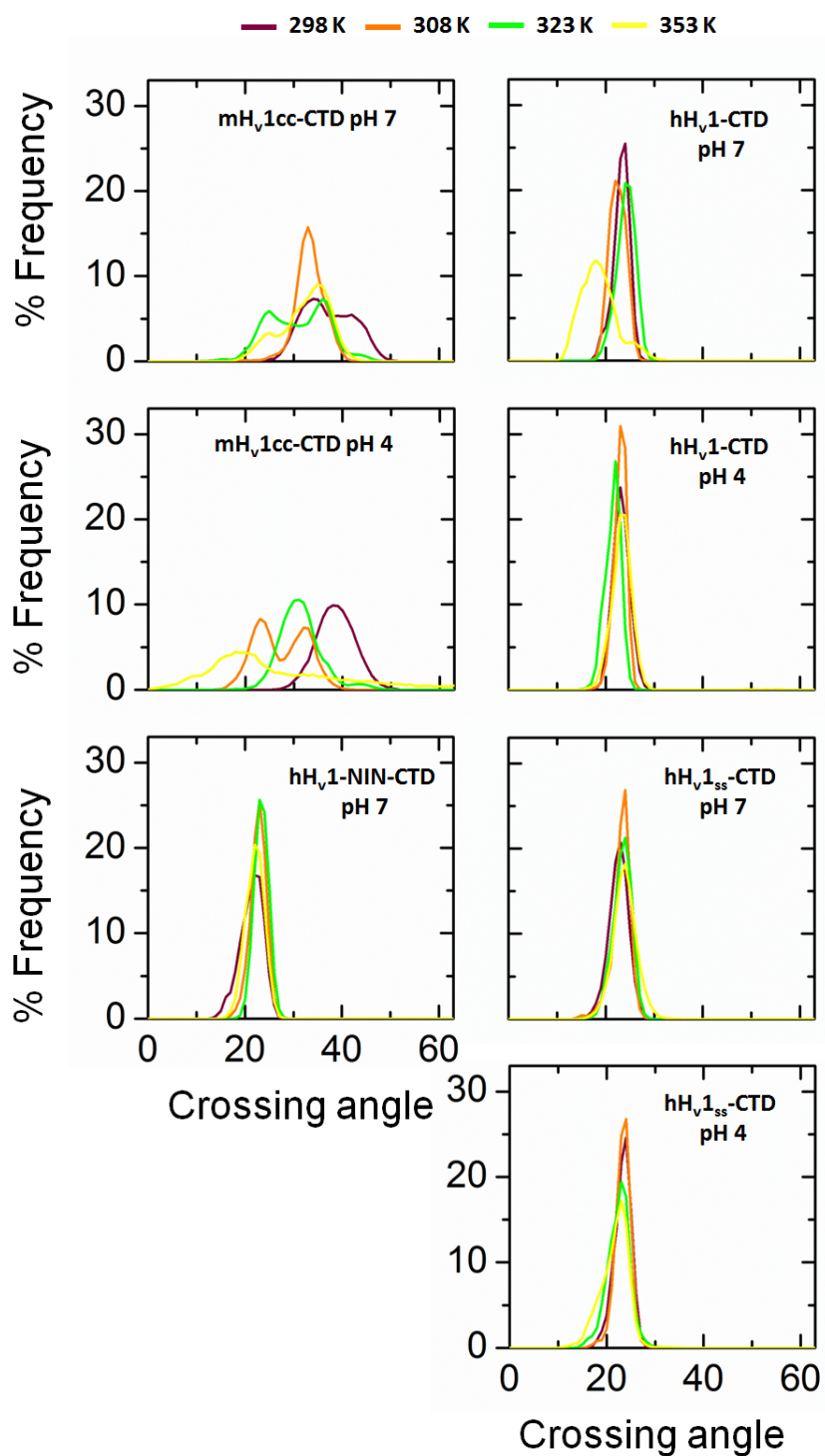
**Figure S2.11**    Distribution of helix-helix crossing angle. The data was extracted from the several dimers of isolated Hv1-CTD.

**Table S2.1**     The conformations of salt-bridge residues, inter-helical hydrogen bond and disulfide bond from previous X-ray crystal structures of isolated CTD.

| Model | hHv1-CTD | | Model | mHv1-CTD | |
|---|---|---|---|---|---|
| | 3A2A [23] | | | 3VMX [26] | |
| | 1st dimer | 2nd dimer | | 1st dimer | 2nd dimer |
| Structure |  |  | Structure |  |  |
| Chain | A and B (B denotes ') | C and D (D denotes ') | Chain | A and B (B denotes ') | C and D (D denotes ') |
| Subinteracting state of salt-bridge residues (Upper pairs) | $\sigma$ | $\sigma$ | Subinteracting state of salt-bridge residues (Upper pairs) | $\nu$ | $\nu$ |
| Distance of E246—K241' | 3.09 (OE2—NZ') | 3.07 (OE2—NZ') | Distance of E242—K237' | 2.94*/6.14* (OE2—NZ$^{1st'}$/ OE2—NZ$^{2nd'}$) | 4.32*/5.34*/ 6.67*/8.33* (OE2 — NZ$^{1st'}$/ OE2 — NZ$^{2nd'}$/ OE2 — NZ$^{3rd'}$/ OE2 — NZ$^{4th'}$) |
| Distance of E246'—K241 | 4.85 (OE2'—NZ) | 4.87 (OE2'—NZ) | Distance of E242'—K237 | 3.11*/4.04* (OE2'—NZ$^{1st}$/ OE2'—NZ$^{2nd}$) | 3.21*/5.88* (OE2$^{1st'}$—NZ/ OE2$^{2nd'}$—NZ) |
| Subinteracting state of salt-bridge residues (Lower pairs) | $\nu$ | $\nu$ | Subinteracting state of salt-bridge residues (Lower pairs) | $\sigma$ | $\nu$ |
| Distance of E253—K252' | 4.36 (OE1—NZ') | 4.33 (OE1—NZ') | Distance of E249—K248' | 2.49*/3.27* (OE2—NZ'/ OE1—NZ') | 4.58 (OE2'—NZ) |
| Distance of E253' — K252 | 4.19 (OE1'—NZ) | 4.21 (OE1'—NZ) | Distance of E249'—K248 | 3.92 (OE2'—NZ) | 4.20*/4.59* (OE1$^{1st'}$—NZ/ OE1$^{2nd'}$—NZ) |
| Conformation of N235 pair | OD1—ND2' | OD1—ND2' | Conformation of N231 pair | OD1'—ND2 | OD1 — ND2' |
| OD1—ND2' | 2.59 | 2.60 | OD1—ND2' | 6.30 | 2.88 |
| OD1'—ND2 | 6.19 | 6.16 | OD1'—ND2 | 3.00 | 6.26 |
| Disulfide bond of C249 pair | Yes | Yes | Disulfide bond of C245 pair | No | No |
| SG—SG' | 2.01 (SG—SG') | 2.04 (SG—SG') | SG—SG' | 2.89*/4.48* (SG—SG$^{1st'}$/ SG—SG$^{2nd'}$) | 2.72*/4.54*/ 4.82*/6.33* (SG$^{1st}$—SG$^{1st'}$/ SG$^{1st}$—SG$^{2nd'}$/ SG$^{2nd}$—SG$^{1st'}$/ SG$^{2nd}$—SG$^{2nd'}$) |

**Table S2.2** (Continued) The conformations of salt-bridge residues, inter-helical hydrogen bond and disulfide bond from previous X-ray crystal structures of isolated CTD.
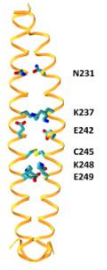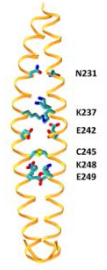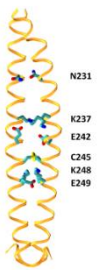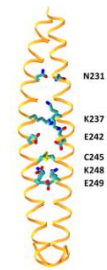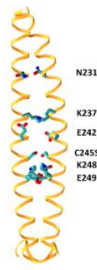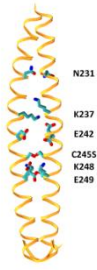
| Model | mHv1-CTD | | Model | mHv1-CTD | |
|---|---|---|---|---|---|
| | 3VMY [29] | | | 3VMZ [29] | |
| | 1st dimer | 2nd dimer | | 1st dimer | 2nd dimer |
| Structure |  |  | Structure |  |  |
| Chain | A and B (B denotes ') | C and D (D denotes ') | Chain | A and B (B denotes ') | C and D (D denotes ') |
| Subinteracting state of salt-bridge residues (Upper pairs) | $\sigma/\nu$ | $\sigma^*/\nu$ | Subinteracting state of salt-bridge residues (Upper pairs) | $\sigma/\nu$ | $\sigma^*/\nu$ |
| Distance of E242—K237' | 2.87*/5.94* (OE2—NZ$^{1st}$'/ OE2—NZ$^{2nd}$') | 4.23*/6.92*/ 8.29* (OE2—NZ$^{1st}$'/ OE2—NZ$^{2nd}$'/ OE2—NZ$^{3rd}$') | Distance of E242—K237' | 2.88*/6.11* (OE2—NZ$^{1st}$'/ OE2—NZ$^{2nd}$') | 4.71*/5.59*/ 8.48* (OE2—NZ$^{1st}$'/ OE2—NZ$^{2nd}$'/ OE2—NZ$^{3rd}$') |
| Distance of E242'—K237 | 4.30 (OE2'—NZ) | 2.92*/4.55* (OE2$^{1st}$'—NZ/ OE2$^{2nd}$'—NZ) | Distance of E242'—K237 | 4.35 (OE2'—NZ) | 2.96*/5.89* (OE2$^{1st}$'—NZ/ OE2$^{2nd}$'—NZ) |
| Subinteracting state of salt-bridge residues (Lower pairs) | $\sigma$ | $\nu$ | Subinteracting state of salt-bridge residues (Lower pairs) | $\sigma$ | $\nu$ |
| Distance of E249—K248' | 2.52*/3.41* (OE2—NZ'/ OE1—NZ') | 4.13*/4.62* (OE2$^{1st}$—NZ'/ OE2$^{2nd}$—NZ') | Distance of E249—K248' | 2.95 (OE1—NZ') | 4.15*/4.38* (OE2$^{1st}$—NZ'/ OE2$^{2nd}$—NZ') |
| Distance of E249'—K248 | 4.00 (OE2'—NZ) | 3.99*/5.36* (OE1$^{1st}$'—NZ/ OE1$^{2nd}$'—NZ) | Distance of E249'—K248 | 3.63 (OE2'—NZ) | 4.20*/4.63* (OE1$^{1st}$'—NZ/ OE1$^{2nd}$'—NZ) |
| Conformation of N231 pair | OD1'—ND2 | OD1 — ND2' | Conformation of N231 pair | OD1'—ND2 | OD1—ND2' |
| OD1—ND2' | 6.31 | 2.95 | OD1—ND2' | 6.40 | 3.00 |
| OD1'—ND2 | 3.03 | 6.27 | OD1'—ND2 | 3.06 | 6.29 |
| Disulfide bond of C245 pair | No | No | Disulfide bond of C245 pair | Yes | Yes |
| SG—SG' | 4.40 (SG—SG') | 2.72*/4.45*/ 4.64*/6.07* (SG$^{1st}$—SG$^{1st}$'/ SG$^{1st}$—SG$^{2nd}$'/ SG$^{2nd}$—SG$^{1st}$'/ SG$^{2nd}$—SG$^{2nd}$') | SG—SG' | 1.98 (SG—SG') | 1.99/6.35* (SG$^{1st}$—SG$^{1st}$'/ SG$^{2nd}$—SG$^{2nd}$') |

**Table S2.3** (Continued) The conformations of salt-bridge residues, inter-helical hydrogen bond and disulfide bond from previous X-ray crystal structures of isolated CTD.

| | mHv1-CTD | |
|---|---|---|
| Model | 3VN0 [29] | |
| | 1st dimer | 2nd dimer |
| Structure |  |  |
| Chain | A and B (B denotes ') | C and D (D denotes ') |
| Subinteracting state of salt-bridge residues (Upper pairs) | $\nu$ | $\sigma^*/\nu$ |
| Distance of E242—K237' | 3.52 (OE2—NZ') | 8.31 (OE2—NZ') |
| Distance of E242'—K237 | 4.15 (OE2'—NZ) | 3.05*/6.27* (OE2$^{1st'}$—NZ/ OE2$^{2nd'}$—NZ) |
| Subinteracting state of salt-bridge residues (Lower pairs) | $\delta/\sigma$ | $\nu$ |
| Distance of E249—K248' | 2.54*/3.39* (OE2$^{1st}$—NZ'/ OE1$^{2nd}$—NZ') | 3.94*/3.98* (OE2$^{1st}$—NZ'/ OE2$^{2nd}$—NZ') |
| Distance of E249'—K248 | 2.64*/3.61* (OE2$^{1st'}$—NZ/ OE2$^{2nd'}$—NZ) | 4.55 (OE1'—NZ) |
| Conformation of N231 pair | OD1'—ND2 | OD1—ND2' |
| OD1—ND2' | 6.29 | 2.97 |
| OD1'—ND2 | 2.99 | 6.28 |
| Disulfide bond of C245 pair | Mutation C245S | Mutation C245S |

Some PDB files contain more than one conformation of these side chains due to diffraction anisotropy and they are marked as 1st, 2nd and 3rd for atoms on different conformations. The possible distances are reported denoted as (*). Sub-interacting states of upper and lower salt-bridge pairs are assigned based on the distance criteria (see Methods). Chain A is equivalent with chain C and chain B is equivalent with chain D, respectively, which are determined using STAMP program [61] in MultiSeq [62] VMD plugin.

# CHAPTHER III: ANALYSIS OF IONIZATION STATE EFFECT ON VOLTAGE SENSOR DYNAMICS IN RESTING STATE OF THE Hv1 CHANNEL

Effect of Ionization State on Voltage Sensor Structure in Resting State of the Hv1 Channel

Panisak Boonamnaj and Pornthep Sompornpisut

*Department of Chemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand*

## 3.1 ABSTRACT

Voltage-gated proton-selective channels (Hv1) mediates proton extrusion during intracellular acidification. Hv1 is gated by the proton electrochemical gradient. The molecular mechanism underlying Hv1 activation remains to be elucidated. It has been proposed that ionizable residues in the permeation pore could serve as a proton-binding site during proton transfer. Here, molecular dynamics (MD) simulations were performed to investigate an effect of ionization states of charged residues on the x-ray structure of Hv1. Modification of protonation state of acidic residues affected the resting conformation of Hv1 by disrupting salt bridges between S4 and the other segments. Upon protonation, conformational changes enabled the displacement of S4 arginines toward the extracellular side and increased a mobility of hydrophobic residues at the gate. The aqueous crevice was considerably wider with the increased hydration in the pore. Solvation free energies of the pore residues were low at the extra- and intracellular entrances whereas the narrowest region exhibited the energy barrier for water translocation. Our MD data showed that water molecules in the upper and lower pore oriented differently. In neutral pH, the pore water oriented its dipole pointing away from the VSD center whereas the opposite direction of the water dipole was observed in acidic pH.

## 3.2 INRODUCTION

Voltage-gated proton selective (Hv1) channels constitute a large family of voltage-gated ion channels. [2] Hv1 channels have been found to be involved in many physiological processes such as immune response of phagocytes, [7-10] sperm motility, [11, 12] and cancer progression. [63] The key role of Hv1 channels is to extrude proton from the intracellular to the extracellular sides. [2] The Hv1 channels are exclusively selective for proton with a unique structure among voltage-gated ion channels. It

comprises a canonical voltage-sensing domain (VSD), like to the first four transmembrane segments (S1-S4 helices) in Nav, Kv and Cav channels. [2] It lacks the pore domain (S5 and S6) typical of voltage-gated ion channels. [17] Instead, the S4 helix is connected to the C-terminal coiled-coil domain (CTD). Thus, the proton transfer occurs through a pore within the Hv1 VSD. The Hv1 channel is known to form a dimer in the membrane. [18, 19] Each subunit is functional independence [18, 20] but is cooperative to another subunit during activation gating. [20, 25, 28] The transmembrane S4 helix of the VSD carries a repeating triplet arginine residue (R1, R2, R3) possessing a sequence motif, XXRXXRXXR, where X are hydrophobic residues.

Extra- and intracellular pH of cells have an important role in protein function and stability. Since the Hv1 channel plays dual roles of voltage and pH sensing, the activation of the channels is controlled by membrane potential ( $\Delta V$ ) and transmembrane pH gradient ( $\Delta pH$ ). [64-66] The efficiency of proton transfer is strongly modulated by a certain range of pH between the extracellular (pH$_o$) and intracellular (pH$_i$) pH. It was shown that pH difference between the outer and inner side of cell membrane creates a pH gradient ( $\Delta pH$ = pH$_o$ – pH$_i$) across the membranes, and a more acidic in the intracellular side promotes channel openning. [17, 64, 65] Without transmembrane pH gradient, Hv1 channels opens at positive voltage. An increase in $\Delta pH$ by increasing pH$_o$ or lowering pH$_i$ results in a shift in voltage activation threshold (V$_{threshold}$) towards more negative potential with higher outward proton currents [17, 64, 65] and faster activation time. [67]

Like other voltage-gated ion channels, the Hv1 channel contains the conserved gating charged arginines in S4 forming a salt-bridge network with aspartate and glutamate residues from the other transmembrane segments. It was shown that at neutral pH these S4 arginines form salt-bridges with those acid residues in S1-S3 through electrostatic interactions, and therefore help to maintain the Hv1 channel in

the resting state. [68, 69] Moreover, there are hydrophobic residues forming a hydrophobic plug at the center of the VSD pore. In the resting state, the salt-bridge network and hydrophobic plug serve as a gate for blocking the permeation pathway, and thus water molecules are unable to pass through the VSD pore. [70] DeCoursey suggested that the mechanism of proton conduction in Hv1 channels differs from gramicidin. The transfer of proton could be in a more complicated mechanism than through the water wire mechanism alone. [71] It was showed that the charged residues aligned along the proton permeation pathway could serve as a candidate for proton-binding site. Changes of protonation states of titratable residues in the pore influenced kinetics of the proton conduction in Hv1.

Despite the importance of pH modulation of proton transport in Hv1 channels, the molecular mechanisms of pH-dependent gating in the Hv1 channel are not completely understood. Several MD studies have not yet reported an effect of the protonation state of the charged residues on the structure and function of Hv1 channels. In this study, we performed molecular dynamics (MD) simulations of a Hv1 channel embedded in phospholipid bilayers to explore the structure and interactions within the VSD under different ionization forms of charged residues of Hv1 in the absence of electric membrane potentials. The study has focused an effect of modification of protonation state of charged residues on structural features associated with function of the Hv1 channel. We used the published crystal structure of a chimera mouse Hv1 (mHv1cc) solved in a resting state as a starting structure. [13] A full-length monomeric mHv1cc was considered to avoid possible interference from the other chain such as the cooperativity effect between two Hv1 subunits. [20, 25, 28] To date, there is no relevant information regarding the impact of His residues on activation of the Hv1 channel. [72] In addition, they are mostly located at the TM loops of the mHv1cc structure. The contributions of protonated/non-protonated histidine at

different pH conditions based on the prediction of p$K_a$ on transition of the VSD core are discarded in the study.

## 3.3 METHODS

### 3.3.1 Structure Models and Positioning Protein in Membrane

The initial structure of the Hv1 channel was taken from the X-ray crystal structure of mHv1cc (PDB: ID 3WKV). [13] Structure of Hv1 monomer in the full-length was used. A multiple sequence alignment of transmembrane segments was constructed using WebLogo3. [73] The structure of the missing loops that connect between TM segments was modeled and energy minimized using the program Loopy. [34] The protein structure quality was subsequently validated by the program PROCHECK. [35] Prior to insertion of the protein into an explicit membrane environment, the optimal spatial position of the protein in membranes was determined using the calculation of the electrostatic solvation free energy approach previously described. [37] In brief, all the missing hydrogen atoms were added to the protein structure using the program PDB2PQR. [40, 41] Partial atomic charges and radii of the protein were taken from the PARSE parameter sets. [39] The electrostatic solvation free energy of the protein was computed with a thickness of 30 Å for the hydrophobic slab, the solvent dielectric of 2 for protein and membrane region, and 80 for non-membrane region. The sequential focusing multigrid algorithm was used with three resolution maps, $300 \times 300 \times 300$ Å$^3$ for coarse, $200 \times 200 \times 200$ Å$^3$ for medium, and $100 \times 100 \times 100$ Å$^3$ for fine resolution. The calculation of the electrostatic solvation free energy was conducted using the program APBS and APBSmem. [38, 74] The optimal position of the protein in membrane is identified by the minimum solvation energy.

3.3.2 Molecular Dynamics Simulations

Protonation states of charged residues in the studied model systems were presumably defined on the basis of five pH conditions, i.e. neutral pH (NpH), intracellular pH gradient or $\Delta pH$ (defined as $pH_o$ = neutral pH and $pH_i$ = acidic pH), acidic pH (ApH) and two basic pH conditions with charged arginine (BpH-R[+]) and neutral arginine (BpH-R). For the two basic pH conditions, lysine residues were deprotonated.

**Table 3.1**   Amino acid charged states present in the five model systems of Hv1. Positive (+), negative (-) and neutral (0) charges of ionizable residues are given based on pH conditions.

| | Ionizable | Model system | | | | |
|---|---|---|---|---|---|---|
| | Residue | BpH-R | BpH-R[+] | NpH | $\Delta pH$ [b] | ApH |
| **Extracellular** | Glu | - | - | - | - | 0 |
| | Asp | - | - | - | - | 0 |
| | Lys | 0 | 0 | + | + | + |
| | Arg | 0 | + | + | + | + |
| | | | | | | |
| **Intracellular** | Glu | - | - | - | 0 | 0 |
| | Asp | - | - | - | 0 | 0 |
| | Lys | 0 | 0 | + | + | + |
| | Arg | 0 | + | + | + | + |

[a]D108 was chosen as the center of the VSD pore with respect to the z-axis. The residues in the extra- and intracellular were relative to the D108 position.  [b]In $\Delta pH$ system, D149 and D170 (located at intracellular) as well as D108 were neutral charged (protonated) whereas E115 and D181 were negatively charged (deprotonated).

The protonation states of ionizable residues were assigned by comparing between the pH values set to the system and the $pK_a$ of amino acids predicted using PROPKA. [44] For instance, if a $pK_a$ value of a given ionizable residue was greater than the pH of the system, then this residue was protonated. The pH-dependent simulation systems containing charged residues with different protonation states are summarized in Table 3.1. In addition, the charges on the deprotonated Tyr and neutral Arg were

taken from previous work [75] and their Lennard-Jones parameters were used as the same with the neutral Tyr and charged Arg, respectively. The protein models were inserted in a 1-Palmitoyl-2-Oleoyl-sn-Glycero-3-Phosphocholine (POPC) lipid solvated with TIP3P water models. [45] $Na^+$ and $Cl^-$ ions at 0.1 M were added to neutralize charges on the system using VMD's Autoionize plugin. All simulation systems were run with a time step of 2 fs and surrounded with explicit TIP3P water molecules. The simulated system had a periodic box with dimensions of $104 \times 102 \times 131$ Å$^3$. Periodic boundary conditions were employed. A distance cutoff of 12 Å was used for calculating nonbonded interactions, electrostatic interactions with particle mesh Ewald summation via fast Fourier transform, and van der Waals interactions, and a switching distance was set at 10 Å. Langevin dynamics at constant temperature of 298 K was used with a damping coefficient of 1 ps$^{-1}$. Pressure was kept constant at 1 atm using the Nose´– Hoover Langevin piston method, with a piston period of 200 fs and a damping time of 50 fs. Energy minimization was performed to remove bad contacts between atoms. Then restrained MD simulations were employed to relax structural strains of the model systems. In the first stage, the restrained MD was conducted with protein and lipid head group atoms kept fixed to their initial positions. The subsequent run was introduced to allow the whole system (waters, lipids and counterions) except for protein to be relaxed. Finally, the equilibration and production runs were performed without any positional restraints. The atomistic MD simulations were performed for 100 ns with ten independent runs for statistics. This gives rise to a total of more than 1 µs of simulation time per system. The CHARMM force field was applied for protein and lipid atoms. The simulation systems were constructed using TCL scripts in VMD. All atom molecular dynamics simulations were performed with the program NAMD. [42]

### 3.3.3 MD Trajectory Analysis

Analyses of the MD trajectories, including hydrogen bond distance root-mean square deviation (RMSD), the root-mean square fluctuation (RMSF), number of pore waters, displacement of residue side chains, water orientation in the pore and transmembrane motions, were carried out using TCL scripts in VMD. Unless otherwise specified, MD trajectories during 80-100ns of the simulations were used for the analysis. The RMSD and RMSF were calculated on the basis of backbone atoms. Analyses of hydrogen-bond network were plotted using python scripts and gnuplot. Positions of the charged and hydrophobic residues in the VSD were extracted using center of mass. The pore radius of Hv1 along the z-axis was determined using the program HOLE. [76] The water orientation in the VSD pore was characterized using $\left\langle \cos\left(\theta\right)_{z} \right\rangle$ where $\theta$ is an angle of the net dipole moment vector of water molecule normal to the z-axis. [77, 78] The positions of the water in the pore were determined based on the z-coordinates of oxygen atoms. All the profiles plotted against the z-axis of the pore generally used a bin width of 1 Å.

### 3.3.4 Principal Component Analysis

A principal component analysis (PCA) was performed on MD trajectories of the systems to identify essential dynamics how ionization states induces changes in mHv1cc conformation PCA was performed for $C_\alpha$ atoms, whose structure from the trajectory was aligned to its initial structure to filter out all the trivial translations and rotations. Last 50 ns (50–100 ns) of the simulation trajectories was selected for each system. The two principal components of motion (PC1 and PC2) which correspond to the first and second Eigenvectors of the covariance matrix, were calculated using the WORDOM software. [79] The representations of collective motions respected to the PC1 were created using VMD.

### 3.3.5 Solvation Energy Calculation

Free energy of solvation was performed using a nonlinear Poisson-Boltzmann continuum model. Solvation free energy, $\Delta G_{solv}$, composes of electrostatic and nonpolar components ($\Delta G_{solv} = \Delta G_{elec} + \Delta G_{np}$). Only residues facing inside the pore were selected for the free energy calculation (61 residues from S1, S2, S3 and S4). Parameters for evaluating $\Delta G_{solv}$ were used according to previous work [80] but multigrid lengths were determined from the protein dimensions in which each was multiplied with 1.5 factor. Protein charges and van der Waals radii were assigned based on AMBER ff99. The free energy calculation was carried out using the program PDB2PQR/APBS.

**3.4 RESULTS AND DISCUSSION**

    3.4.1 Pore Residues and Protonation Sites

       The sequence logo of Hv1 channels from various species shows high conservation of charged residues in S1-S4 as indicated by the stacks of the letters with its height approaching the maximum value of sequence conservation. (Figure 3.1A). The crystal structure shows the arrangement of S1-S4 segments in which the conserved charged and hydrophobic residues are oriented in the central core of the VSD (Figure 3.1B). These residues including the S4 arginines (R201, R204 and R207), the negatively charged residues (D108, E115, D149, D170 and D181) and the conserved hydrophobic residues (V112, F146, F178, L185 and L197) have been highlighted in the study since they were proposed to play an essential role in the proton translocation across membranes. [33, 70] The VSD pore of the mHv1cc x-ray structure contains hydrophobic residues of which the side chains are arranged roughly parallel to bilayer plane (Figure 3.1B). On the basis of the VSD center, the hydrophobic layer of the Hv1 channel can be separately defined as the intracellular layer ($HL_{in}$) and the extracellular layer ($HL_{ex}$). The $HL_{in}$ and $HL_{ex}$ residues act as a hydrophobic gate by creating constriction regions to control water accessibility in the aqueous filled crevice of the VSD. [81] It was proposed that salt-bridge interactions between the basic S4 arginines and acidic residues help to stabilize the VSD in membranes while the hydrophobic residues are clustered together at the VSD center to block water penetration. [13, 70] Assuming that side-chain of acidic residues is the proton binding site required to open the gate. It is therefore interesting to understand a modulation of the protonation state of the charged residues affects structure, salt-bridge interactions and the gate of the voltage sensor in the resting-state conformation.
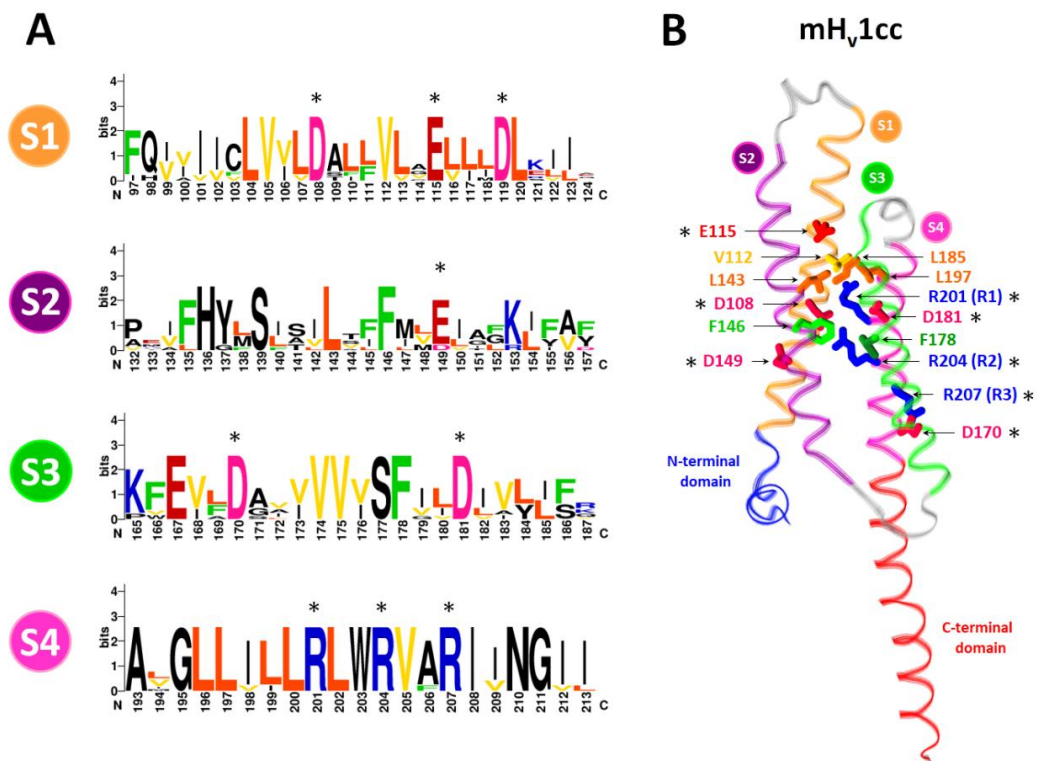
**Figure 3.1** (A) Sequence logos of the TM segments (S1-S4) derived from the primary sequence alignment of voltage-gated proton channel with various members of the Hv1 family. (B) The crystal structure of mHv1cc monomer containing four TM segments: S1 (orange), S2 (purple), S3 (green), S4 (magenta), and N- (blue) and C-terminal helices (red) are shown as ribbon linked with the loops (ribbon grey). The gating charge arginine (R1–R3), conserved acidic and hydrophobic residues are presented as licorice.

3.4.2 Effect on Structure Stability in the Resting State

Global backbone RMSD profiles demonstrated a well-behaved MD simulation for the studied systems (Figure 3.2). Substantial changes were observed for the systems with protonated acidic residues, where RMSDs increase during the first 20 ns for the $\Delta pH$ system and during the first 70 ns for the ApH system. For the neutral pH (NpH) system, no significant change in RMSD was found after equilibration. In the last 20 ns simulation, all systems are considerably stable. The average RMSD values obtained during the last 20 ns simulation time were 3.8 ± 0.5, 5.1 ± 0.5 and 6.9 ± 0.7 Å for NpH, $\Delta pH$ and ApH, respectively.

RMSD per segment from the NpH system were on average lower compared to the corresponding segments of the $\Delta pH$ and ApH systems. For NpH, RMSDs of S1 and S4 were slightly higher than those of S2 and S3. However, MD simulations under acidic pH ($\Delta pH$ and ApH), also showed that structural deviation of S2 and S3 considerably increased compared to the NpH system (Figure 3.2). Since the protonated acid residues were in the S2 and S3 segments, an effect of the protonation on structural stability becomes more apparent. Particularly, a relatively high RMSD in overall for the protonation of all acidic residues (ApH) suggested that under this extreme condition the Hv1 structure deviates dramatically from the initial resting-state conformation (Figure 3.2). In addition, a wide range of deviation in RMSD of S1 and S4 implies that conformations of these segments adopt a relatively large structural deviation. Notes that the S1 segment has three acidic residues whereas the S4 segment has three arginine residues. Therefore, the modification of protonation states of these acidic residues could disrupt all salt-bridges and as a consequence made a stronger impact on the structure stability.
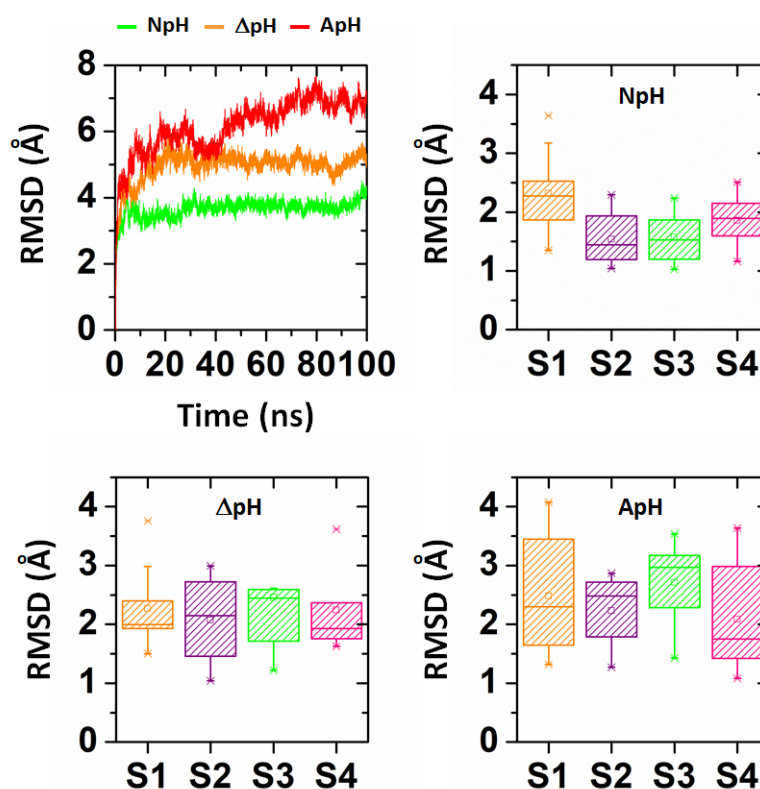
**Figure 3.2**     The upper left panel is the global RMSD profile with respect to the initial structure during the 100 ns MD simulations. The model systems denoted as NpH, $\Delta pH$ and ApH were defined according to pH conditions as described in Method. The remaining panels highlighted averaged RMSD per segment shown as statistical box plots. The boxes were 25–75% with the median as a horizontal line. Mean values were indicated by a square symbol. The asterisks showed 1–99 percentiles.

### 3.4.3 Effect on Salt-bridge Interactions of S4 Arginines

From MD trajectories, we measured hydrogen-bond distances representing salt-bridges between S4 arginine residues (R201, R204 and R207 with a generic name R1, R2 and R3, respectively) and five acidic residues (E115, D108, D181, D149 and D170). At neutral pH (NpH), R1 and R2 form salt-bridges with D108 and/or D181 while R3 interacts with D170 (Figure 3.3).
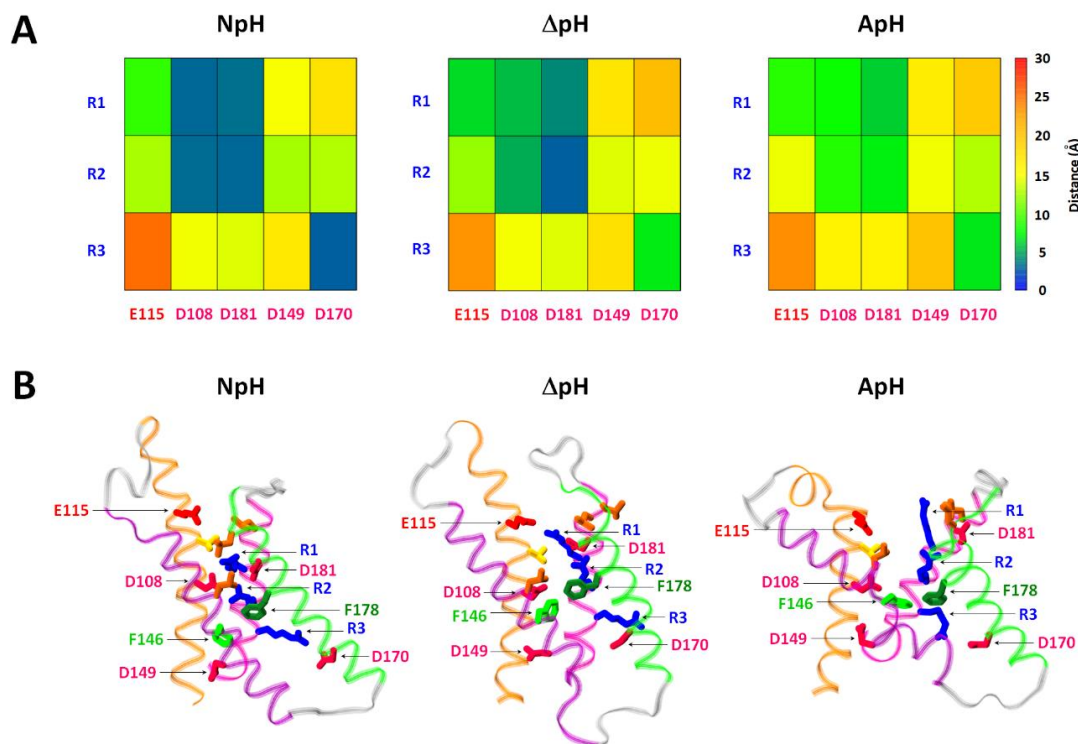
**Figure 3.3** (A) H-donor-acceptor distance patterns of S4 arginines (R1, R2 and R3) with the conserved acidic residues of the Hv1 VSD obtained from the NpH, $\Delta pH$ and ApH systems. With geometric criteria for hydrogen bonding, a blue box is qualified as forming a salt-bridge (B) MD snapshots taken from the equilibration of each system. Representation style of the models was the same as those described in Figure 3.1B.

The obtained salt-bridge pattern at this condition is similar to previous MD results of the Hv1 channel in the resting state. [81] For the $\Delta pH$ system, a significant change of hydrogen-bonding pairs was observed as pHi became acidic but pHo remains neutral pH. As expected, a hydrogen-bond between R3 and D170 was eliminated because of a protonation of the D170 residues preventing the salt-bridge formation. The protonation of D108 gave a similar effect on the R1–D108 and R2–D108 salt-bridge interactions, that are considerably weaken. The only salt-bridge of R2-D181 remains in the system because D181 is not protonated. The ApH system showed a large impact

on salt-bridge disruption associated with S4 arginines. Protonation of all acidic residues has led to conformational rearrangement of the salt-bridges residues. As a result, the original salt-bridge pairs observed from the neutral pH are completely disrupted (Figure 3.3). All salt-bridge pairs are no longer formed. The results of salt-bridge interactions from the simulations support the structural stability analysis or RMSD shown in Figure 3.2.

Our results demonstrate that the acidic condition affects the stability of the channel in the resting conformation through electrostatic exclusion. The pH-induced protonation of the acidic residues disrupted the salt-bridge network and altered the resting-state conformation. The 100 ns simulation time-scale may not be long enough to identify the channel undergoing denaturation or activation under the studied conditions. However, ten-repeat MD runs of the three systems showed that the NpH system produced the most stable structure compared to the initial configuration. On the other hand, structural collapse of the Hv1 was observed in several MD runs of the ApH system, implying that the structure configuration under the protonation of entire acidic residues may not be stable. MD runs of the $\Delta pH$ system produced interesting data which to some extent should not be negligible. According to DeCoursey proposal, one of plausible permeation mechanisms in Hv1 is via a protonation at the carboxyl group of acidic residues. [71, 82] Our MD data with $\Delta pH$ condition may be sufficient to see a significant degree of structural changes associated with a state between the resting state and the conductive state. It might capture a structure which adopts an intermediate position. To support the idea with strong evidences, we have further elucidated other properties inside the VSD core of the obtained structures such as water crevices, solvation free energy and water reorientation. The results will be described in the next section.

We have demonstrated the protonation of acidic residues yielding a significant perturbation of the Hv1 conformation in the resting state. To explore an opposite pH effect, we have further conducted two simulation systems under non-physiological basic conditions. The results showed that modulation of ionization states of basic residues (Lys and Arg) has less effect on the resting state conformation than that of acidic residues (Figure S3.1A). For the BpH-R$^+$ system (containing neutral Lys and charged Arg) there was no significant change in both of the salt-bridge pattern (Figure S3.1B) and the corresponding RMSDs (Figures S3.1C) with respect to the NpH system. Since Arg residues are still charged, the interactions involving S4 arginines are not perturbed significantly. Therefore, the influence of neutral Lys on the stability of the VSD in the resting state conformation is considerably small. In contrary, the effect is more apparent for the system containing neutral Arg and Lys (BpH-R). The salt-bridge network was completely disrupted similar to the ApH system. However, RMSD per segment of the BpH-R system is not as large as that of the ApH system. Instead, the magnitude of TM motions is somewhat similar to the NpH system. Again, notice that both BpH-R$^+$ and BpH-R systems present a non-actual physiological situation for the proton conduction process, as Hv1 is generally activated under intracellular acidification.

### 3.4.4 Effect on Hydration in the Permeation Pathway

The permeation pathway in the Hv1 channel has an hourglass shape that can be accessible by water molecules. It has been suggested that the presence of pore water facilitates the transfer of protons. In the resting state, the salt-bridges and hydrophobic plug at the center of the pathway block the passage of water and thus separate the aqueous crevices into extracellular and intracellular sides. Our MD data

showed that water molecules can access the crevices at different depth depending on the pH conditions (Figure 3.4).
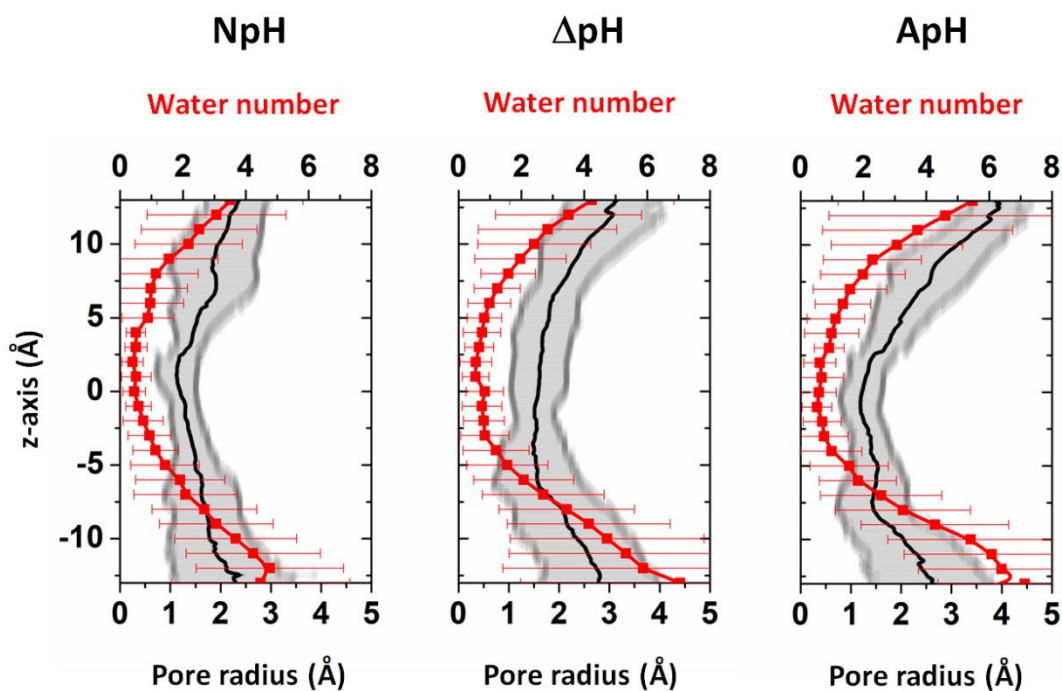


**Figure 3.4**    Averaged pore radii profiles along the z-axis and number of water molecules occupied in the pore of NpH, $\Delta pH$ and ApH systems were shown in black and red curves, respectively. Standard deviation (SD) were shown in grey areas and red error bars.

Amounts of pore waters were also different between the extracellular and intracellular crevices. From the NpH trajectories, the pore radius profile showed that the extracellular and intracellular mouths were wide with an average pore radius of 2.0 Å and 1.8 Å, respectively (Table 3.2).

**Table 3.2**     The pore radius along z-axis for different systems

| Systems | Radius (Å) | | |
| --- | --- | --- | --- |
| | Extracellular side (at z = 10) | Intracellular side (at z = -10) | Narrowest region |
| NpH | 2.0 ± 0.8 | 1.8 ± 0.8 | 1.1 ± 0.4 (0 ≤ z ≤ 3) |
| $\Delta pH$ | 2.5 ± 0.9 | 2.4 ± 1.1 | 1.5 ± 0.4 (-2 ≤ z ≤ 2) |
| ApH | 3.1 ± 0.8 | 2.1 ± 1.0 | 1.2 ± 0.4 (-3 ≤ z ≤ 1) |

The pore radius at the narrowest constriction region fluctuates between 0.6 and 1.5 Å during the simulations with an average radius of ~1.1 Å. We found that side chains of residues involved in the formation of salt-bridges and hydrophobic plug obstruct water penetration at the central region in the VSD core. However, water penetration at the narrowest region of the channel was rarely detected, but still observed in some MD snapshots of the NpH system. It should be noted that the water passage at the same region occurs in the $\Delta pH$ and ApH systems more frequently than it does in NpH. MD simulations under acidic conditions ($\Delta pH$ and ApH) showed a considerable widening of the pathway and the increased hydration in the pore. The pore radius at the extracellular and intracellular entries was on average of 2.5 and 2.4 Å, respectively for $\Delta pH$, and of 3.1 and 2.1 Å, respectively for ApH. At the narrowest region, the pore radius of the $\Delta pH$ system significantly increases to 1.5 Å but a minor increase in the radius (1.2 Å) was observed for the ApH system. The increased radius is because of conformational changes of the salt-bridge residues which was accompanied by the loss of interactions between the S4 arginines and protonated acidic residues. Changes in side chain conformations of S4 arginine, acidic and hydrophobic residues had made more space at the bottleneck region. However, a distribution of water molecules in the permeation pathway of the $\Delta pH$ and ApH systems are slightly greater than that of NpH. This may seem a subtle difference but it could be significant in modulating

the energy of the conduction. As have been shown that a small difference in amounts of water pore can effectively facilitate the ion conduction of a calcium channel by reshaping local electric field. [83]

### 3.4.5 Position of Charged and Hydrophobic Residues in the Pore

Position of side chains relative to the z-axis of the pore was analyzed to demonstrate conformational rearrangement of the charged and hydrophobic residues associated with the protonation state of acidic residues. At the neutral pH, side chains of R1 and R2 were strongly localized around the pore center as shown by the corresponding single narrow peak with its maximum at z = 2 and -2, respectively (Figure 3.5A). Localized side chain of R3 which appeared as a single peak with its maximum at z = -9 (near the intracellular side) was also observed. For most of the cases, the five acidic residues showed restricted mobility of its side chain conformations, especially those (D181, D108, and D107) forming salt-bridges with R1, R2 and R3.
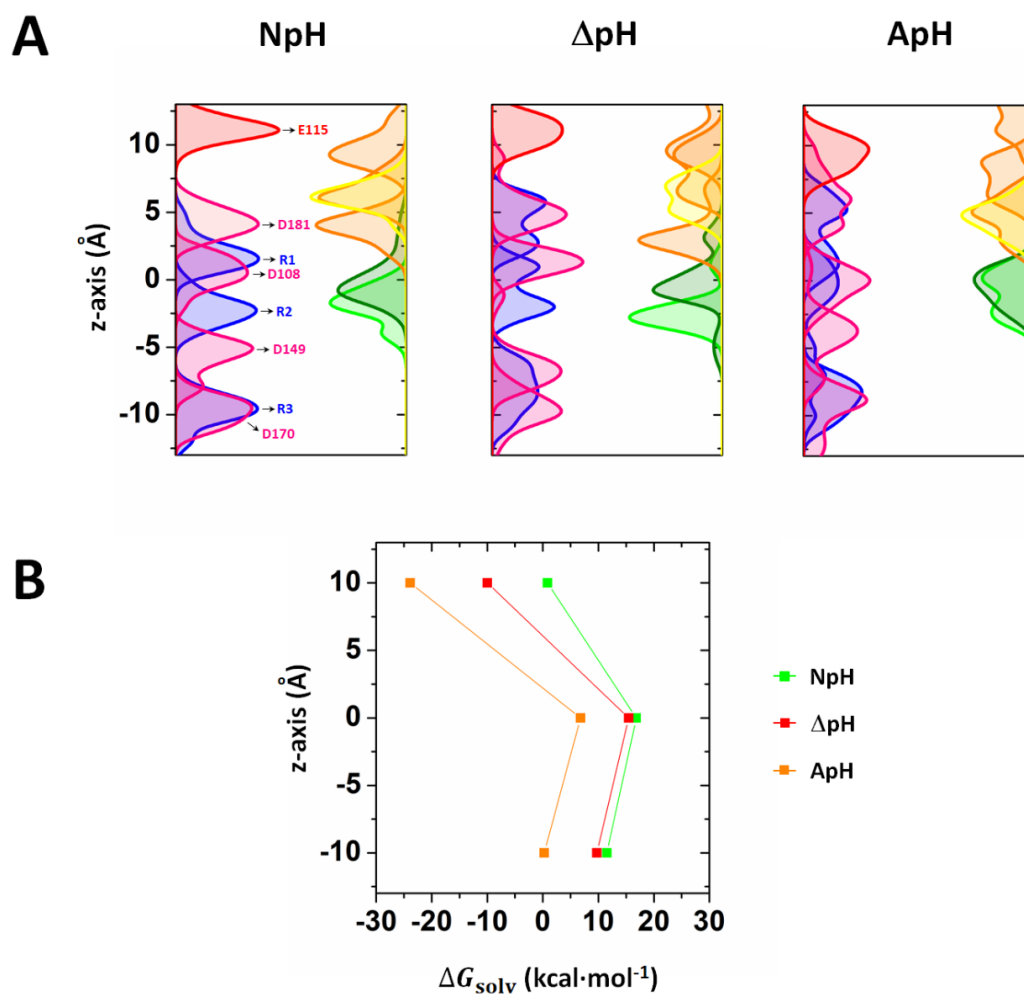
**Figure 3.5** (A) Distributed positions of the charged and hydrophobic residues relative to the z-axis. Distribution curves computed based on the center of mass of the corresponding residue side chain were shown in blue (Arg), red (Glu), pink (Asp) yellow (Val), orange (Leu), light green (F146) and forest green (F178). (B) Solvation free energies at the extracellular, narrowest and intracellular regions of the pore. Each $\Delta G_{solv}$ obtained in the graph was averaged over residues located in the defined regions as described in the main text.

Side chain conformational changes upon acidic residue protonation were clearly observed. For $\Delta pH$, position distributions of R1 and R2 were split into two overlap peaks with unequal height while that of R3 became substantially broader. This suggested that side chains of R1 and R2 switched between two different conformations

with one more favorable than the other. On the other hand, a very broad distribution of R3 position implies no orientational preference due to high conformational mobility of the side chain. The distribution curves of acidic residues exhibited similar shape compared to those observed from the NpH system. Compared to S4 arginines, the effect of conformational changes was less pronounced. The strongest impact was observed in the ApH system. Distribution curves of both basic and acidic residues were relatively broad and overlap each other (Figure 3.5A). This indicated the side chains were highly mobile. Another key finding of this study is that the position of S4 arginines was about 2-3 Å shifted toward the extracellular side. This result is consistent with the idea that the positive gating charge is translocated to the extraplasmic surface during membrane depolarization for the activation of the channels.

Distribution curves of hydrophobic residues in the $HL_{ex}$ (orange and yellow) and $HL_{in}$ (greens) along the z-axis also demonstrated the effect of conformational changes due to acidic residue protonation. The $HL_{in}$ comprises F146 in S2 and F178 in S3 that are located near the intracellular half of the VSD. The $HL_{ex}$ consists of four hydrophobic residues including V112 in S1, L143 in S2, L185 in S3 and L197 in S4 that are on the extracellular half of the VSD. From the simulations, broad overlaps between two distributed peaks of the $HL_{ex}$ residues were mostly observed for the $\Delta pH$ and ApH systems while a single distributed peak was observed for the NpH system (Figure 3.5A). This indicates that the four $HL_{ex}$ residues in the acidic conditions were more displaced and more flexible whereas those in the neutral pH tend to remain relatively stable. However, the position distributions of the $HL_{in}$ residues in the NpH system are slightly different from the $\Delta pH$ and ApH systems; the ApH system showed to some extent a broader distribution than the NpH system. From the results, the hydrophobic plug inside the Hv1 VSD in the resting state conformation is better maintained at the neutral

pH. As a result of protonation of acidic residues, the loss of salt-bridge interactions in the VSD had eventually led to loosely packed hydrophobic cluster at the gate.

We identified the increased hydration in the water crevices associated with changes in the ionization states of charged residues of the channel. MD results showed a low water accessibility in the narrowest region of the channel. To explore the energetic costs, Free energy of solvation ($\Delta G_{solv}$) of the residues facing the pore were computed. Based on the position of the residues relative to the z-axis of the pore, we computed a sum of the $\Delta G_{solv}$ over residues in three regions that are pore residues in the extracellular entry (15 ≤ z ≤ 5), the narrowest region (5 ≤ z ≤ -5) and the intracellular entry (-5 ≤ z ≤ -15). The results showed that for all the three systems the $\Delta G_{solv}$ values were relatively low at the extra- and intracellular entrances of the pore compared with the VSD core. The narrowest region exhibited the highest solvation energy of which the NpH has the most energetically unfavorable (Figure 3.5B). This is consistent with the zero or lowest number of water molecules observed in this region. For all the three regions, the $\Delta G_{solv}$ of $\Delta pH$ and ApH are relatively lower than that of NpH. This indicated that the increased hydration of residues in the pore observed from the simulations was associated with a reduction of energy barrier of solvation. The $\Delta G_{solv}$ together with the electrostatic component ($\Delta G_{elec}$) and non-polar component ($\Delta G_{np}$) of the residues in the entire pore region (15 ≤ z ≤ -15) were shown in Table 3.3. We found that the overall solvation of the pore residues was energetically unfavorable for NpH and $\Delta pH$ with the $\Delta G_{solv}$ of 7.9 and 4.2 kcal·mol$^{-1}$, respectively whereas the favorable $\Delta G_{solv}$ was observed for ApH (-6.6 kcal·mol$^{-1}$). As described previously, the poor solvation of NpH and $\Delta pH$ was associated with the pore residues in the narrowest region. Their side chains came in close contact together, and thus became less solvent-exposed. On the other hands, side chains of the pore residues of ApH were more solvent-exposed than the other two systems. Therefore, they were

more favorably solvated than those of NpH and $\Delta pH$. Among the three systems, the electrostatic free energy of solvation ($\Delta G_{elec}$) was considerably different while the non-polar solvation ($\Delta G_{np}$) was similar (Table 3.3).

**Table 3.3**    The sum of free energy of solvation overall all selected residues

| $\Delta G$ (kcal·mol⁻¹) | NpH | $\Delta pH$ | ApH |
|---|---|---|---|
| $\Delta G_{solv}$ | 7.9 | 4.2 | -6.6 |
| $\Delta G_{elec}$ | -8.7 | -12.3 | -23.0 |
| $\Delta G_{np}$ | 16.7 | 16.5 | 16.3 |

3.4.6 Structure Dynamics of Subdomains

To better understand large conformational changes of the Hv1 VSD channel, the principal component analysis (PCA) was carried out. MD trajectories of the last 50 ns were used to extract the dominant collective protein motions. Here, we illustrated relevant motions of the transmembrane helices on the basis of the corresponding motion mode of the first and second eigenvectors (Figure S3.2). A superposition of the principal motions of the backbone $C_\alpha$ atoms projected along the first eigenvector with 30 frames of the entire trajectory was shown in Figure 3.6A.
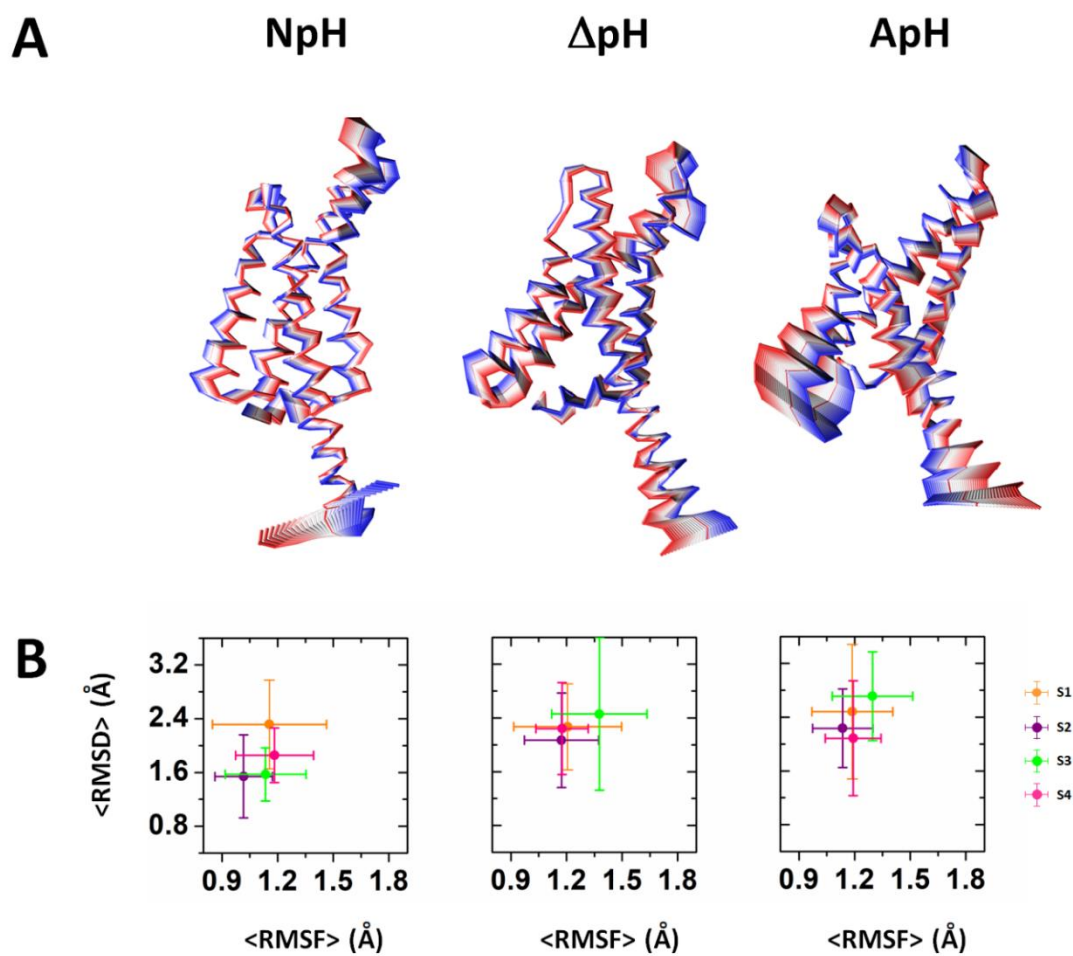
Figure 3.6    (A) Subdomain motions of Hv1 taken from the first eigenvector. Color scale from red–white–blue highlighted from low to high atomic displacement. (B) <RMSD> vs <RMSF> plots of each TM segment.

Note that the major subdomain motion was apparently taken place at the C-terminal helix for all the three simulation systems. From PCA, it appears that the four transmembrane helices of the NpH system were relatively rigid compared to those of $\Delta pH$ and ApH. For $\Delta pH$, the S1, S3 and S4 helices slightly rotated around their own helical axis while the S2 helix tiled toward the membrane plane. A rotation of the N- and C-terminal helices around its helical axis was also observed. In addition, the S2-S3 loop and the C-terminal helix underwent a concerted outward movement in a way to expand the intracellular entrance of the VSD pore. However, concerted hinge motions of the region encompassing the intracellular end of S2 and S3 helices and the C-terminal helix were found to be an opposite and inward movements in the ApH system. Furthermore, we did not observe any apparent rotation of individual transmembrane helix.

Plots of RMSD vs RMSF per segment shown in Figure 3.6B illustrated the magnitude of the TM displacement and flexibility. Both RMSD and RMSF of all TM segments increased as the pH of the solution shifting from NpH to $\Delta pH$ and to ApH. For ApH, the S3 exhibited the highest displacement and flexibility (with an RMSD of 2.7 Å and an RMSF of 1.3 Å) compared to the S1, S2 and S4 segments. Notably, all TM segments were better maintained in the neutral pH while they were more displaced in the acidic conditions. The plots supported the PCA described above. In conclusion, these results emphasized the important motion of subdomains that may play a significant role in the activation dynamics of Hv1 channels.

3.4.7 Water Orientations in the Pore

Atomic details of the water dipole reorientation found in aquaporin water channels provided insight into the mechanism of water conduction. [77, 78, 84] This has been brought to our attention to investigate this issue in the Hv1 since the mutated aquaporin has the ability to conduct proton. [85] MD trajectories were used to compute the net water orientations determined by a sum of dipole moment vectors of the pore waters selected within a defined region. Figure 3.7A showed the net water orientation (or dipole moments) as a function of the coordinate in the z-axis. Interestingly, the net water dipole profiles along the z-axis exhibited a roughly curvature of which the inflection point is approximately located near the VSD center for both NpH and $\Delta pH$ systems. Especially, the water dipole curves in a region of 5 ≤ z ≤ -5 showed a change from $\vec{\mu}_{up}$ ($\left\langle \cos(\theta) \right\rangle_z > 0$) to $\vec{\mu}_{down}$ ($\left\langle \cos(\theta) \right\rangle_z < 0$) for NpH or vice versa for $\Delta pH$. It seems that water molecules in the upper pore (0 < z < 5) tend to orient its dipole in an opposite direction to those in the lower pores (-5 < z < 0). This observation was the so-called bipolar orientation, which has been reported in aquaporin. [77, 78, 84] In addition, the two opposite curves between the NpH and $\Delta pH$ systems indicated that the pore water oriented differently in the crevices (Figure 3.7B). In the neutral pH, the upper- and lower pore waters oriented its dipole pointing away from the VSD center. In contrast to $\Delta pH$, the upper- and lower pore waters reoriented in such a way that their dipole moments point towards the VSD center. The orientation preference of the pore water in the channel is still highly speculative, but to some extent associated with the electric potential at the pore surface. The change in water orientation should be in response to a change in polarity of the channel interior.
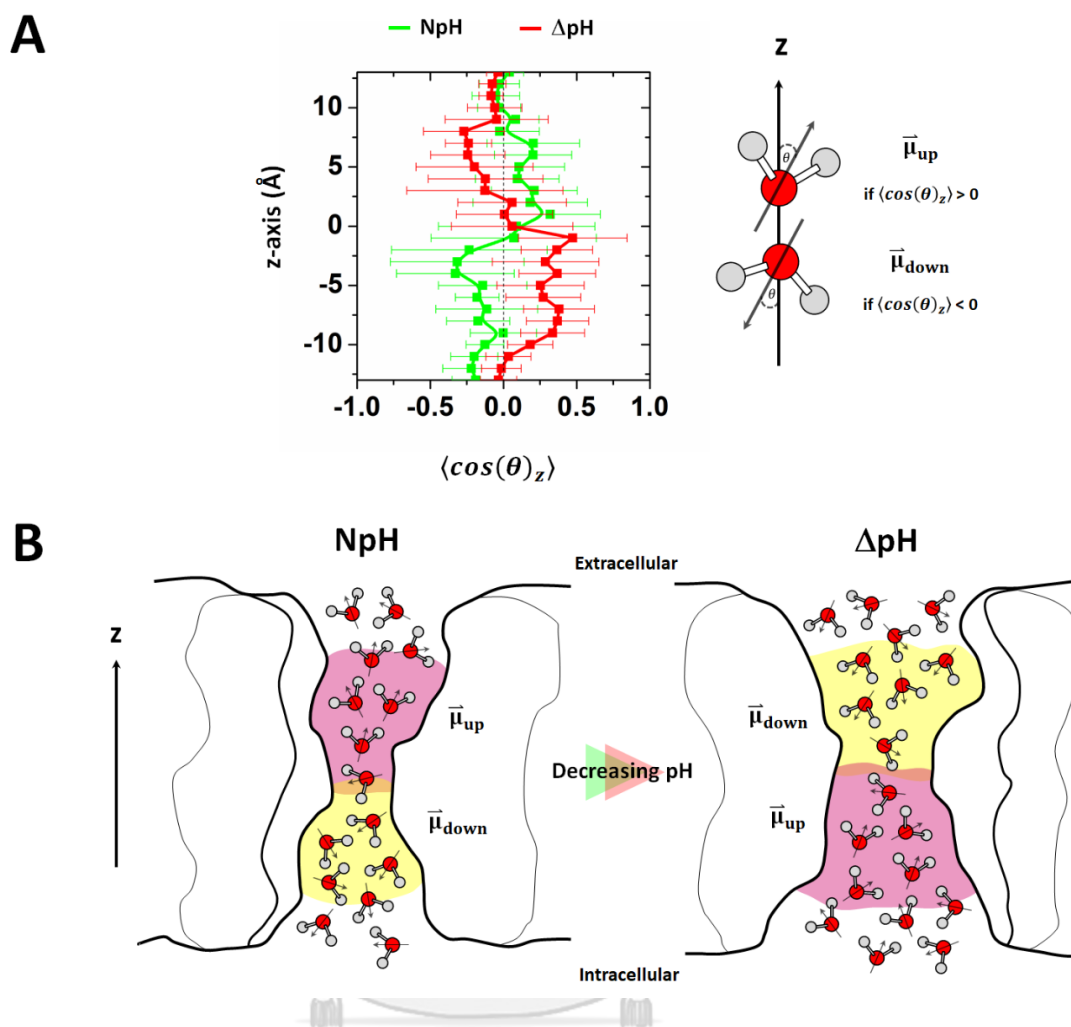
Figure 3.7      (A) Average orientations of water in the Hv1 pore depicted by $\langle \cos(\theta)_z \rangle$ with SD bars. $\theta$ is an angle of the net dipole moment vector of waters ($\vec{\mu}$) relative to the z-axis of the pore. A positive $\langle \cos(\theta)_z \rangle$ corresponds to the water dipole with upward arrow pointing to the extracellular membrane surface ($\vec{\mu}_{up}$) whereas a negative $\langle \cos(\theta)_z \rangle$ refers to the water dipole in downward direction ($\vec{\mu}_{down}$). The data are averaged over 80–100 ns calculated from ten independent simulations for each system. (B) Cartoon illustrating a proposed favorable orientation of waters near the central pore in neutral and $\Delta pH$ conditions.

Although changes in conformations and electrostatic environment of the pore residues were observed in the study, yet, we were not able to identify enough evidences to draw clear conclusions. Nevertheless, our result provides new findings which is a challenge in gaining a fundamental understanding of the proton conduction mechanism in voltage-gated proton channels.

## 3.5 CONCLUSION

Structure, properties and interactions of Hv1 can be modulated by varying protonation state of the charged residues within the voltage sensor domain. Changes in the charged state of these residues eliminated intramolecular electrostatic interactions of the salt-bridge network involved with S4 arginines, leading to the structural instability in the resting state. This affects the hydrophobic packing in the VSD core and allows reshaping of water-filled crevice. A subtle increased hydration in the water crevice could significantly reduce solvation energy barrier in the pore. Furthermore, the pore of the channel accommodated the bipolar orientation of water molecules with an inversion of the water dipole in response to conformational changes. The pH induced conformational changes may be accompanied by a polarity change of the VSD pore. This study should gain further understanding about the structural properties underlying the functional mechanism of the Hv1 channels.

## 3.6 SUPPLEMENTARY INFORMATION
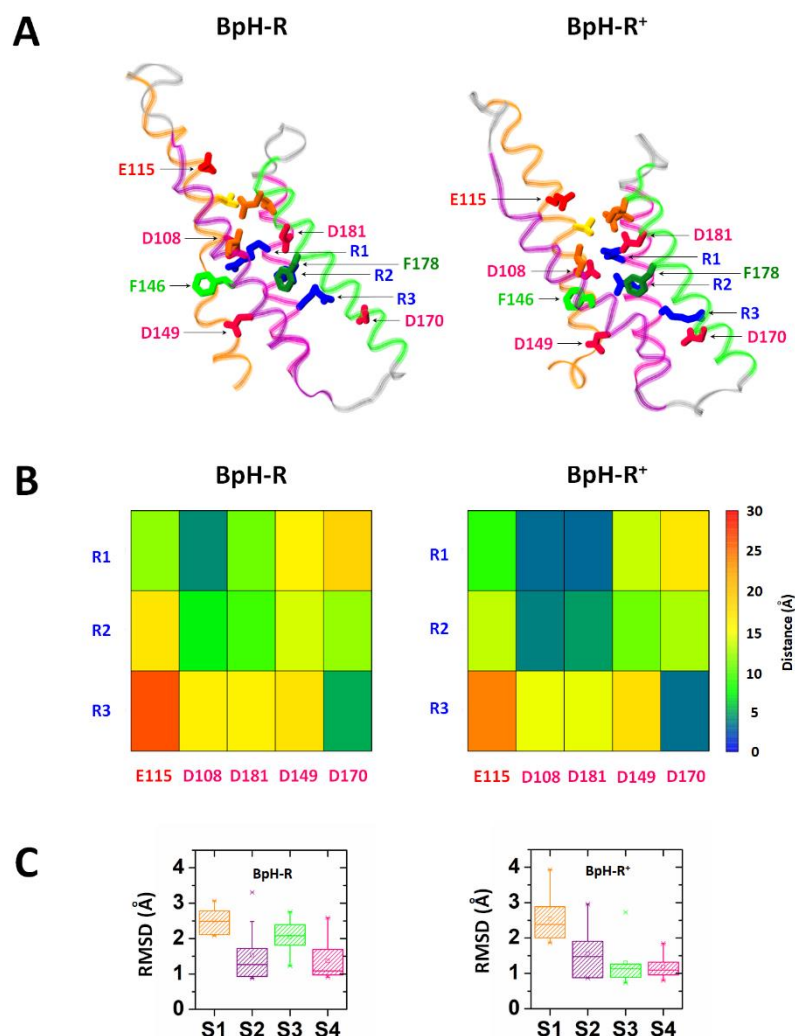


**Figure S3.1** (A) Equilibrated representative conformations showing only the VSD region obtained from different MD trajectories for BpH-R and BpH-R⁺ systems, respectively. The models are colored and labeled as similar in Figure 3.1B. (B) Patterns of hydrogen-bond network of highly conserved charged residues located in the VSD for each system. (C) Statistical box plots of comparing RMSD values for backbone TM segments are also extracted which are highlighted in orange (S1), purple (S2), green (S3) and magenta (S4). The boxes are 25–75% with the median as a horizontal line. Mean values are represented with the little squares. The asterisks show 1–99 percentiles. The data obtained from ten independent simulations for each system averaged over 80–100 ns.
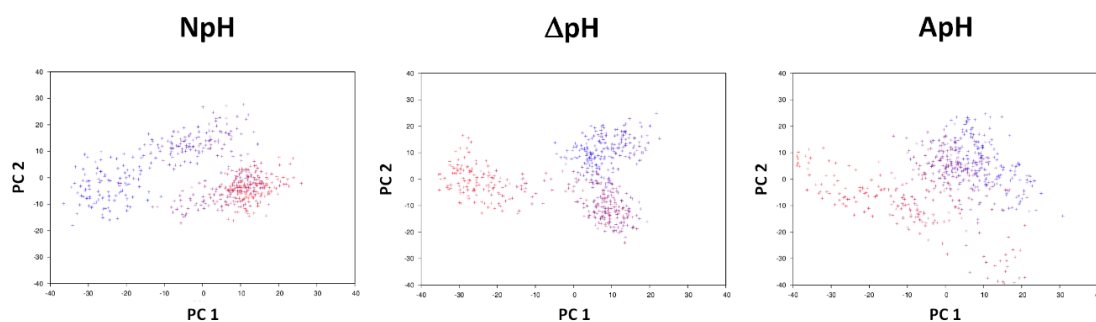
**Figure S3.2**    PCA scatter plot of structural snapshots along the two principal components with a color scale running from red (at 50 ns) to blue (the end at 100 ns).

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

# CHAPTER IV: INVESTIGATION OF THERMAL-RESPONSE OF hHv1 PROTEIN BY COARSE-GRAINED MC AND FINE-GRAINED MD COMPUTER SIMULATIONS

Thermal-response of a protein (hHv1) by a coarse-grained MC and fine-grained MD computer simulations

Panisak Boonamnaj[1], Sunita Subedi Paudel[2], Warin Jetsadawisut[1], Sunan Kitjaruwankul[3], Pornthep Sompornpisut[1], R.B Pandey[2]

[1]*Department of Chemistry, Chulalongkorn University, Bangkok 10330, Thailand*

[2]*Department of Physics and Astronomy, University of Southern Mississippi, Hattiesburg, MS 39406, USA*

[3]*Faculty of Science at Sriracha, Kasetsart University Sriracha Campus, Chonburi 20230, Thailand*

## 4.1 ABSTRACT

A Monte Carlo simulation with a coarse-grained (CGMC) model of protein involving a phenomenological residue-residue interaction potential, augmented by a Molecular Dynamics simulation is used to investigate local and global structures of a protein (hHv1) as a function of temperature. Data from both all-atom MD and a coarse-grained MC simulation show that the radius of gyration of the protein (the cytoplasmic domain of hHv1 monomer and dimer) decreases on increasing the temperature, i.e. it becomes more globular on heating in its native phase in contrast to its thermal expansion in denatured phase. The globularization of the protein is quantified by analyzing the scaling of the structure factor. The crossover from globular to random-coil structure involving segmental reorganizations is found to depend on the length scale and the temperature.

## 4.2 INTRODUCTION

The structures of proteins have been investigated extensively for decades with a range of models involving atomic scale details to coarse-grained representations, first principle force-fields to phenomenological interactions, structure-based constraints to knowledge-based contacts, etc. Because of an explosive growth of literature on the subject, we limit our citations [86-118] for some historical perspective and relevance to this work. Despite enormous interest, fundamental understanding of protein folding remains an open problem. [119] The prolific yet specific characteristics of a protein chain is described by the number and types of the constitutive amino acids, their sequence, and the underlying embedding environment. Statistical average over a huge conformational ensemble leads to specific structural response. Fundamental understanding of unique structural accessibility offers enormous potential for developing insight into a large array of applications in designing drugs, bio-

functionalized nano-materials, etc. As pointed above, exploring the pathways of 'protein folding' in a vast ensemble of protein conformations has been a subject of immense interest for a long time. [86-120] Even though some success has been achieved in developing a better insight into the energetics of protein models, a complete understanding of the fundamental issue of how proteins fold to its native structure still remains an open question from a first principle approach. [119]

Snapshots of protein structures in the Protein Data Bank (PDB), generated from X-ray, IR, and NMR measurements, have played a critical role in developing models, calibrating parameters, designing constraints and effective interactions, extracting local and global characteristics of proteins, etc. [86, 87, 89, 91, 94, 97, 99, 103, 105, 111] Utility of PDB protein structures depends on the prospective and scope of the modeling. For example, the ensemble of protein structures may be used as an input to devise a contact potential (as in this study), direct use of the structure in all-atom molecular dynamics simulations, etc. Massaging the model parameters are frequently used to achieve appropriate characteristics (e.g., native structure by reducing the structural fluctuation around a standard conformation) in latter to identify the residue contacts and associated force fields. For example, the characteristic of structural constraints (e.g., native state) are used to devise an interaction scheme (i.e., favorable interaction between residues with native contacts and unfavorable interaction with non-native contacts) in the original Go model [88] and its variants to study the energy landscape, folding pathways, and binding mechanism. This is a useful approach to explore the conformations around a well-defined structure as the PDB structures provide a valuable glimpse into few elements of a huge conformational phase space. In thermodynamic equilibrium, each microstate should be accessible from any initial state, not just jiggling around a specific snapshot which may be the practice in many computational modeling particularly those with all-atom details.

In developing models of such a complex system as a protein, some degree of approximation is unavoidable in almost all approaches involving all-atom details to minimalist coarse-grained descriptions in order to carry out large-scale computer simulations, gain useful insight, and draw meaningful conclusions. Approximations and coarse-graining procedures may include devising interaction potentials, exploring the phase space selectively, resorting to efficient and effective methods, etc. There can be many coarse-grained mechanisms in protein modeling. [116, 118, 120-124] However, we use an efficient and effective Monte Carlo computer simulation scheme on a discrete lattice with ample degrees of freedom for each constitutive residue to move and interact with its surrounding and peptide-bonds to vary. [125] Our coarse-grained model involves a phenomenological potential with the input from a knowledge-based residue-residue interaction to investigate the structure of a membrane protein hHv1. To verify the results and possibly calibrate the phenomenological scales of the Monte Carlo data, an all-atom MD simulation is also carried out to investigate the same protein. We have observed interesting structural response in native and denatured phases of the protein as a function of temperature.

We focus on a voltage-gated human proton channel hHv1 [33, 126, 127] which is essential in a range of physiological functions such as controlling the motility of sperm, [11, 128] killing pathogens, [9, 129-132] pH regulation, [133] reactive oxygen production, [7] activation of B-cells [8] etc. Conformational transformations of the hHv1 channel are essential in controlling the selective proton conduction in cells. It is believed that the gating mechanism [33, 126, 127] of the proton transport is controlled by the cooperative interaction between the two monomeric units i.e. the dimer of the C-terminal domain (see the schematic Figure S4.1). Fujiwara et al. [26] investigated the cooperative gating in tandem dimers and found that the temperature affects the activation kinetics of proton channel. [33, 126, 127] How does the conformational

changes occur remains an open question. [26] It would, therefore, be interesting to examine the conformations of both, the monomer hHv1 of the C-terminal domain and its tandem dimer tD-hHv1 depicted in Figure S4.2 with an all-atom Molecular Dynamics and Monte Carlo simulations of a coarse-grained phenomenological model as the residue-residue interactions compete with the thermal noise. In general, thermal agitations lead to spreading of protein conformation, we find the opposite in native phase of hHv1, i.e. the protein becomes more globular on heating.

We would like to point out that the conformational fluctuations of proteins, [134] particularly in its native states can be very large. [135] A large ensemble of protein structures obtained from nuclear magnetic resonance has been examined by Tang et al. [135] who suggest that the protein's native state is critical. They have estimated the critical exponent $\nu$ for scaling of the radius of gyration $R_g$ of the proteins with its mass i.e. the number $N$ of amino acids, $R_g \sim N^\nu$ with an universal exponent $\nu \sim \frac{1}{3}$. From a scaling analysis of the structure factor, we find that our results for the conformation of hHv1 including the exponent $\nu$ are not only consistent with the finding of Tang et al. [135] but also it exhibits a negative thermal response in its native phase. In section 2, we briefly describe the residue-residue interaction, phenomenological potential, utility of a lattice host space and bond-fluctuation mechanism and the implementation of the Monte Carlo method in our coarse-grained approach. The technical detail of the standard all-atom Molecular Dynamics method and results on the dependence of the radius of gyration of the protein on the temperature is presented in the following section 3. Data from the MC simulations are presented in section 4 for the thermal response of the local and global physical quantities including the radius of gyration of the protein. The conclusions are provided in section 5.

**4.3 METHODS**

4.3.1 Coarse-Grained Monte Carlo (CGMC) Approach

A protein is described by a set of amino acids tethered together via peptide bonds in a specific sequence. Each of the 20 different amino acids (residues) has its unique characteristics resulting from constitutive atomic and molecular units. In our coarse-grained representation of the protein, each residue is treated as a monomer or node. The structural details at atomic scale is thus ignored but the specificity of each residue is captured via its unique interactions. It is worth pointing out that such a coarse-grained approach [136] has been extensively used in investigating the structural and dynamical properties of polymers (homo, hetero and block copolymers) for over half a century. The power and pitfalls of such coarse-grained methods (e.g., hierarchical scaling) are relatively well understood. Since proteins are a special kind of hetero polymers of amino acids, it is logical to represent a residue as a node i.e. a monomeric unit in polymer terminology and use a viable method to capture its specific characteristics.

*4.3.1.1 Residue-Residue Interactions*

How to find/estimate interactions among the amino acids? The growing ensemble of frozen protein structures in PDB provides an excellent source to devise a knowledge-based contact matrix. Distribution of the amino acids in an ensemble of protein structures can be used to identify residue-residue pair correlations in order to construct a contact map among 20 amino acids. Such a knowledge-based residue-residue contact matrix (20 × 20) consists of 210 independent interaction pairs. [86, 87, 89, 94, 97, 99, 103, 105, 111] Note that the ensemble of residue distribution in proteins structures in PDB involves diverse data sets which includes size and make-up of the protein, the embedding matrix, experimental constraints etc. Various assumptions and

approximations are made in deriving the residue-residue contact matrices. A knowledge-based interaction potential was proposed decades ago by Tanaka and Scheraga [87] and was further developed by Miyazawa and Jernigan (MJ) [89, 97] in an effective medium approximation. Over the years a number of knowledge-based contact [86, 87, 89, 94, 97, 99, 103, 105, 111] potentials have been examined, developed, refined and used extensively in investigation of protein folding. For example, in order to achieve hydrophobicities compatible with the experiment, Betancourt and Thirumalai (BT) [99] refined the MJ contact matrix by selecting an appropriate reference solvent (Thr) within the Miyazawa and Jernigan scheme. [89, 97] We have used some of these contact potentials, not only to verify the resulting data but also to assess the convergence. In this work we will use BT contact matrix as input in our phenomenological potential.

Each residue interacts [118, 125] with its neighboring residues within a range ( $r_c$ ) of interaction with a generalized Lennard-Jones potential,

$$U_{ij} = \left[ \varepsilon_{ij} \left( \frac{\sigma}{r_{ij}} \right)^{12} + \varepsilon_{ij} \left( \frac{\sigma}{r_{ij}} \right)^{6} \right] \tag{4.1}$$

where $r_{ij}$ is the distance between the residues at site $i$ and $j$ and $r_c$ is the cut-off range. The potential strength, $\varepsilon_{ij}$, is unique for each interaction pair with appropriate positive (repulsive) and negative (attractive) values used from the knowledge-based contact interactions. [99] For example, [99] for Leu-Leu pair $\varepsilon_{ij} = -0.81$, for Leu-Lys pair $\varepsilon_{ij} = 0.16$, Ala-Lys pair $\varepsilon_{ij} = 0.11$, etc. in unit of $RT$ where $R$ is the gas constant [89] and  is the temperature. The potential is relatively simple, but it captures the specificity of the underlying residues and provides results that are consistent qualitatively with the results from a typical fine-grained model used frequently with all-atom MD simulations.

*4.3.1.2 Lattice Host Space*

We consider a cubic box for simplicity. Many readers may immediately assume, that we are going to deal with the minimalist models with severely limited degrees of freedom and specificity, for example, a HP model of protein on a lattice where (i) the bond length between consecutive nodes is fixed to a lattice constant, (ii) H (hydrophobic) and P (polar) nodes of the chain occupy one lattice node each and (iii) only very limited and somewhat arbitrary (e.g. end-moves, crank-shaft, reptation) moves [136] of each node is allowed along the lattice topology. These early 'toy' models, still actively in-use, [137] have provided some insight (e.g. energy landscape) into the thermodynamic stability of protein structures. However, in our approach, we are not restricted by severe shortcomings of limited degrees of freedom (movements of both nodes and segments), lack of elastic nature of covalent bond, or specificity of residue nodes, etc. as in the minimalist protein model. In fact, we exploit the advantage of discrete nature of the lattice in coarsening the host space to cover large-scale (temporal and spatial) while incorporating ample degrees of freedom, variability of peptide bonds, specificity of residues, etc. with option to refine via fine-graining (top-down approach).

It is worth reminding the readers about the insight gained from the benefit and pitfall of computational modeling in polymer physics [136] where the structure and dynamics of polymer chains has been studied in depth by both MC and MD simulations. In a homo-polymer chain, the short (Rouse) and asymptotic long time (diffusion) dynamics of a chain segment can be easily verified by both MC and MD methods in dilute solution (single chain limit). Coarsening of the host space via discrete lattice and implementing a bond-fluctuating mechanism [136] is a viable tool to reach large-scale dynamics in many complex systems. The overhead in calculations of the force-field via linked-list of the lattice-cell often used with a continuum host space is

avoided on a discrete lattice. Even though we deal with a relatively simplified model system here, our coarse-grained approach can be easily extended to more complex systems on larger scales than the current MD approaches used in protein modeling. The bond-fluctuation approach used here is one of the most efficient and effective methods to investigate large-scale complex problems [136] in polymers, soft matter and proteins. Moreover, one can easily enhance the degrees of freedom by fine-graining. [138, 139] We would like to emphasize that the popular MD approaches involving atomistic scales details with readily available opensource software are valuable particularly in probing the fine structures and short time scale dynamics. The utility of both, an all-atom MD and a coarse-grained MC, provides better insight into the local and global characteristics of proteins.

Our all-residue Monte Carlo simulation involves coarse-graining the structural details of the constituents, interactions, and the underlying host space. [125] The monomer hHv1 and the tandem dimer tD-hHv1 of the voltage gated protein are represented by 49 and 104 residues, as depicted in Figure S4.2, tethered together by peptide bonds on a cubic lattice. A residue occupies a cube of size $2a^3$ where $a$ is the lattice constant. [136] The covalent bond length between consecutive nodes varies between 2 and $\sqrt{10}$ in unit of the lattice constant. The fluctuation in bond length caused by the movements of each node (residue) is purely stochastic and constrained by the lower and upper limits on the bond length. There is no explicit potential used to control the size of the bond-fluctuation except the hard core wall at these limits. The protein chain is randomly placed in the simulation box in a random configuration initially. Note that there are ample degrees of freedom for each node (with 26 adjacent cells) to move and covalent bonds to fluctuate unlike the minimalist protein chain models on a discrete lattice with constant bond length. [137]

### 4.3.1.3 Metropolis MC Scheme

We use the Metropolis algorithm to move each tethered residue randomly as follows. A residue at a site say $i$ (cubic cell) is selected randomly to move to a neighboring lattice site $j$ (cubic cell) also selected randomly. As long as the excluded volume constraints and the limitations on changes in the covalent bond length (described above) are satisfied, the residue is moved from site $i$ to site $j$ with the Boltzmann probability $e^{-\frac{\Delta E_{ij}}{T}}$, where $\Delta E_{ij}$ is the change in energy between its new ($E_j$) and old ($E_i$) configuration $\Delta E_{ij} = E_j - E_i$ and $T$ is the temperature in reduced unit of the Boltzmann constant. Attempts to move each residue once defines a unit Monte Carlo step (MCS).

We examine a number of local and global physical quantities i.e., energy of each residue, its mobility, contact map and density profile, mean square displacement of the center of mass of the protein, radius of gyration and its structure factor during the course of simulation. Simulations are performed for a sufficiently long time, 107 time steps, at each temperature with many, 10 to 500, independent samples to estimate the average values of the local and global physical quantities. Most of our data are generated on a $150^3$ sample for monomer and $340^3$ sample for a tandem dimer. We have also used different lattice sizes to make sure that our qualitative findings are independent of the lattice size. It should be noted that the potential considered here is phenomenological and all physical quantities are measured in arbitrary scales i.e. in reduced unit. A crude connection with the realistic units for the order of magnitude can however be made by calibrating the results of coarse-grained MC approach with the data from all-atom MD simulations e.g. by comparing the data of the same quantities (see below). A typical residue of size 0.8 nm occupies a cubic cell which implies that the lattice constant a = 0.4 nm. Estimates of the protein size in unit of lattice constant in our coarse-grained model seem consistent with our data

from MD simulations. It is worth pointing that anyone interested in such simulations can develop their own computer code, test our data, and/or improve it further.

### 4.3.2 Molecular Dynamics (MD)

All-atom Molecular dynamics simulations were performed using NAMD [42] with CHARMM22 force field on a $73 \times 73 \times 103$ Å$^3$ box with periodic boundary conditions. The model for the protein was created using structural refinement with PaDSAR method [140] of an X-ray structure of hHv1-CTD taken from Li et al. [33] ; the sequence is shown in Figures S4.1 and S4.2. Two monomers were embedded in a 150 mM NaCl solution with TIP3P for water. [45] Both electrostatic interactions with particle mesh Ewald method and van der Waals interactions were used with a cut-off range of 12 Å. Harmonic constraints were also applied to a central residue to maintain the dimer structure. Each simulation was performed for 45 ns with a time step of 2 fs to generate the NVT ensemble where first 1.2 ns (6000 steps) were only used for relaxation at each temperature with Nose-Hoover Langevin piston method using a piston period of 200 fs and a damping time of 50 fs. Langevin dynamics was implemented to enforce the target temperature with a damping coefficient of 1 ps$^{-1}$. A relatively wide temperature range, $T$ = 298–633 K, was used to examine the conformational transformations of the protein within a feasible time scales. [59]

## 4.4 RESULTS AND DISCUSSION

### 4.4.1 Results on MD
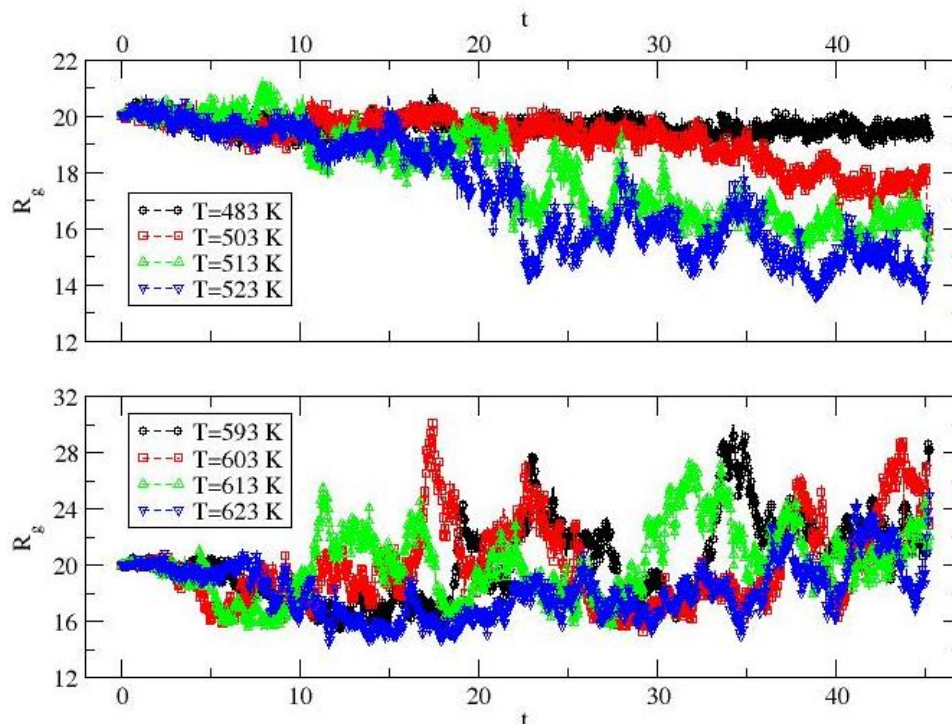


**Figure 4.1**     AAMD data, $R_g$ versus t for hHv1 dimer at representative temperatures in low (upper) and high (lower) temperature regimes.

Figure 4.1 shows the temporal variation of the radius of gyration ( $R_g$ ) at representative temperatures. Despite fluctuations, $R_g$ seems to converge towards its steady-state equilibrium within our observation time. Conformational trajectory of last 20 ns is used to estimate average value of $R_g$ and its standard deviation. The trends in data shows that protein chain has reached its steady state in low temperature regime. Conformational fluctuation is relatively larger at high temperatures where average estimate of the radius of gyration is necessary to identify the trend as shown in Figure 4.2.
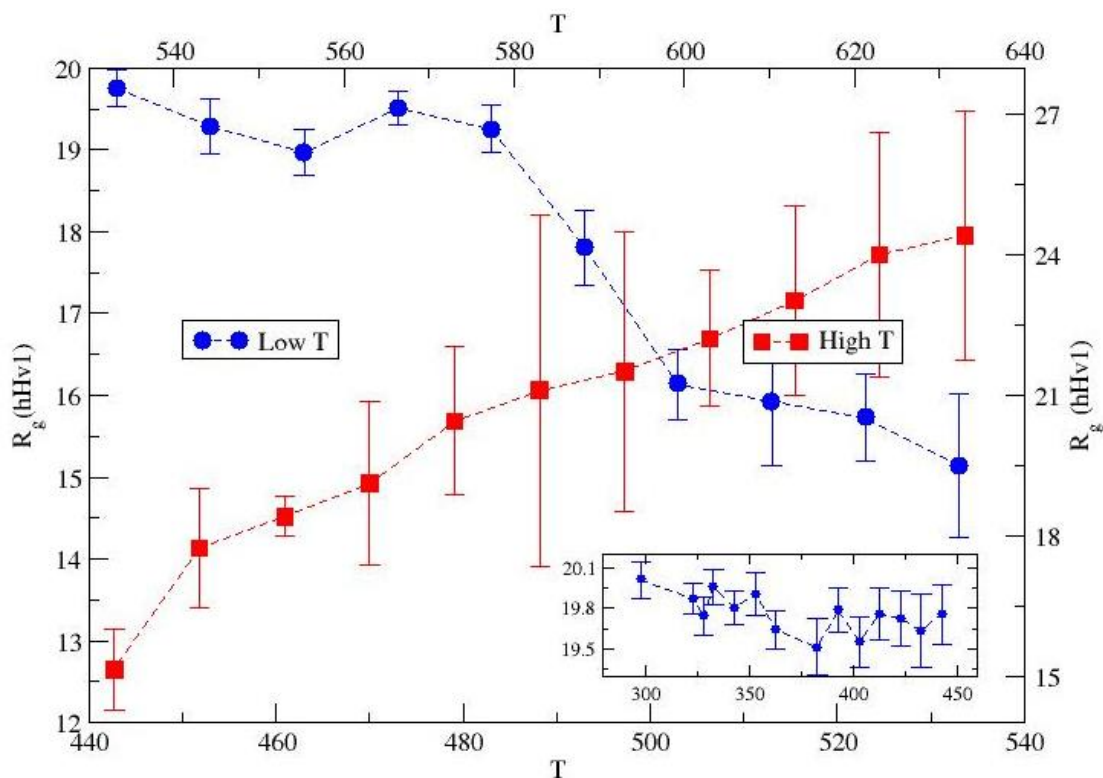
**Figure 4.2**     AAMD data, variation of the radius of gyration $R_g$ of hHv1-dimer with the temperature. Lower x-scale represents low temperature regime with the corresponding $R_g$ on left y-axis. Upper x-scale is for the high temperature regime with the corresponding $R_g$ on right y-axis. Inset figure shows the variation of $R_g$ of hHv1-dimer with the temperature at lower temperatures.

The variation of the average radius of gyration, $R_g$, with the temperature is presented in Figure 4.2. We see that (i.) $R_g$ remains almost constant at low temperatures $T$ = 298–440 K, (ii.) decreases on increasing the temperature $T$ = 440–530 K and iii. expands on raising the temperature further ($T$ = 530–630 K). One may notice immediately that the temperature scales are very high even in the low temperature range. Why the temperature range is so bizarre? The answer lies in the calibration of the temperature in all-atom MD simulation software and the approximate nature of the model systems which can be a subject of worthy debate but we degress for now. We focus on the reality of our investigations of an un-real

model system [141, 142] with the raw data presented here to avoid exaggeration and hype and for interested readers to reproduce, test, and develop it further. One should consider these temperatures as a guide rather than its quantitative measure for low to high range in order to assess the response of the protein structure. It is worth pointing out that Day et al. [59] have observed that 'increasing temperature accelerates protein folding' and that high temperatures are needed to 'elucidate the details of protein unfolding at a minimum of computational cost'.

As mentioned above, one would expect that enhancing the thermal agitations may open up the protein structure, but we find the opposite. The protein becomes more compact on heating in its native phase leading to an interaction-driven thermal-induced conformational response where both monomeric units conform to its dimer configuration. These findings may help hypothesizing in interpretation of experimental observations [26] that the dimeric domain plays a critical role in the proton transport. In the denatured (high temperatures) phase, both monomeric units move apart, the dimer configuration disappears (see Figure 4.3), and the radius of gyration increases with the temperature as seen in Figure 4.2 as one would expect. We speculate that the proton conduction along the hHv1 pathway will decrease in the denatured phase due to underlying tenuous protein conformations and the loss of coordinated dimer conformation. The activation kinetics of proton channel will also respond accordingly to temperatures as pointed out by Fujiwara et al.. [26] In order to verify the trend and complement our finding, we further investigate the thermal response of the protein by a coarse-grained approach described in section 2 with results presented in the following section 4.
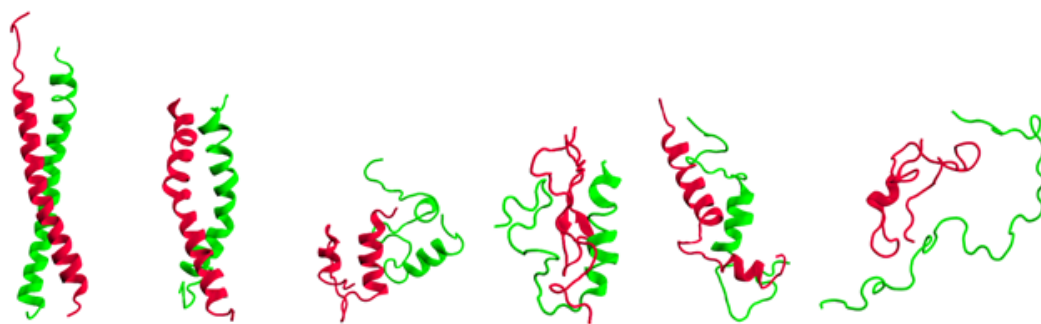
**Figure 4.3**     AAMD data, snapshots of dimer (monomers in green and red) from the MD simulations at temperature $T$ = 443 K, 493 K, 513 K, 533 K, and 603 K respectively from left to right.

### 4.4.2 Results on CGMC

In this section, we present results on the local and global physical properties along with the visual analysis of the snapshots of the monomer hHv1 and its tandem dimer t-hHv1. A large set of data for the thermal response are presented in this section as it is easier to keep track of some of the physical quantities in our coarse-grained approach. Results on the response of the radius of gyration of the protein to temperature from all-atom MD simulation (section 3) are useful not only for assessing and verifying our findings but also for calibrating the scales of data from CGMC simulations.

### 4.4.2.1 Snapshots and Contact Maps

Let us begin by looking at the snapshots and examining the underlying contact maps. Typical snapshots of the tandem dimer at a range of temperatures ($T$ = 0.021–0.032) are presented in Figure 4.4. Although a snapshot is not enough to identify the complete trend, it does provide a glimpse of some structures. For example, the size of the protein seems to contract on raising the temperature ($T$ = 0.021–0.028) and expand on further increasing the temperature ($T$ = 0.028–0.032). Note that, on

increasing the temperature, the separation between the monomeric units appears to decrease in low temperature (native phase) regime and increase in high temperature (denatured phase) regime. The residue interactions are critical in orchestrating the structure of the protein in native phase with relatively lower thermal agitation at low temperatures where the protein conformations are representative of its native structures. The conformations controlled by the thermal-noise at higher temperatures are reflective of the protein structures in its denatured phase. Note that the snapshots in Figure 4.4 looks different from that in Figure 4.3 of the MD simulations. However, the it does capture the essence of contact mechanism of the monomers in the dimer. It should be pointed out that Figure 4.3 is somewhat distilled representation of the structure of the protein and may not be accurate snapshots of the residues.
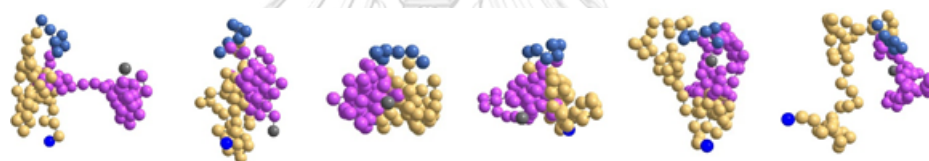


**Figure 4.4**     CGMC data, snapshots of tD-hHv1 protein towards the end of 107 MCS time steps. grey sphere–residue $^1$H, bright blue-residue $^{104}$H, Pink-residue $^2$Q–$^{49}$I, dirty-blue residue $^{50}$A–$^{55}$A, gold residue $^{56}$I–$^{104}$H at the temperature $T$ = 0.021, 0.023, 0.025, 0.028, 0.030, 0.032 from left to right.

Like the snapshots (Figure 4.4), a contact map provides a glimpse of residues in its local proximity i.e. within the range of interaction which may change by changing the temperature. Few representative contact maps of the dimer tD-hHv1 at representative low and high temperatures are presented in Figures 4.5 and S4.3. To understand the Figures, it would be worth reminding that the sequence of residues in two monomers (Figure S4.2) of the tandem dimer is in reverse order i.e. $^1$H–$^{49}$I and $^{56}$I–$^{104}$H. At the temperature $T$ = 0.020 (Figure 4.5), we see some localized assembly around the same segments $^1$H–$^{22}$L and $^{84}$E–$^{104}$H in both monomers almost like a mirror

image. There is also looping via interaction between segments far apart (sequential geometry) i.e. segments $^{35}$K–$^{45}$T of the first monomer and $^{90}$K–$^{100}$L of the second. The map changes on raising the temperature with expanding the contacts among residues. However, the mirror symmetry around the tandem link $^{50}$A–$^{55}$A seems to persist ($T$ = 0.020–0.022) with some fluctuations ($T$ = 0.023). Larger loops (due to contact between residues far apart along the sequence) appear at high temperatures ($T$ = 0.028, 0.029) which eventually vanish on raising the temperature further ($T$ = 0.030, 0.031) (see Figure S4.3).
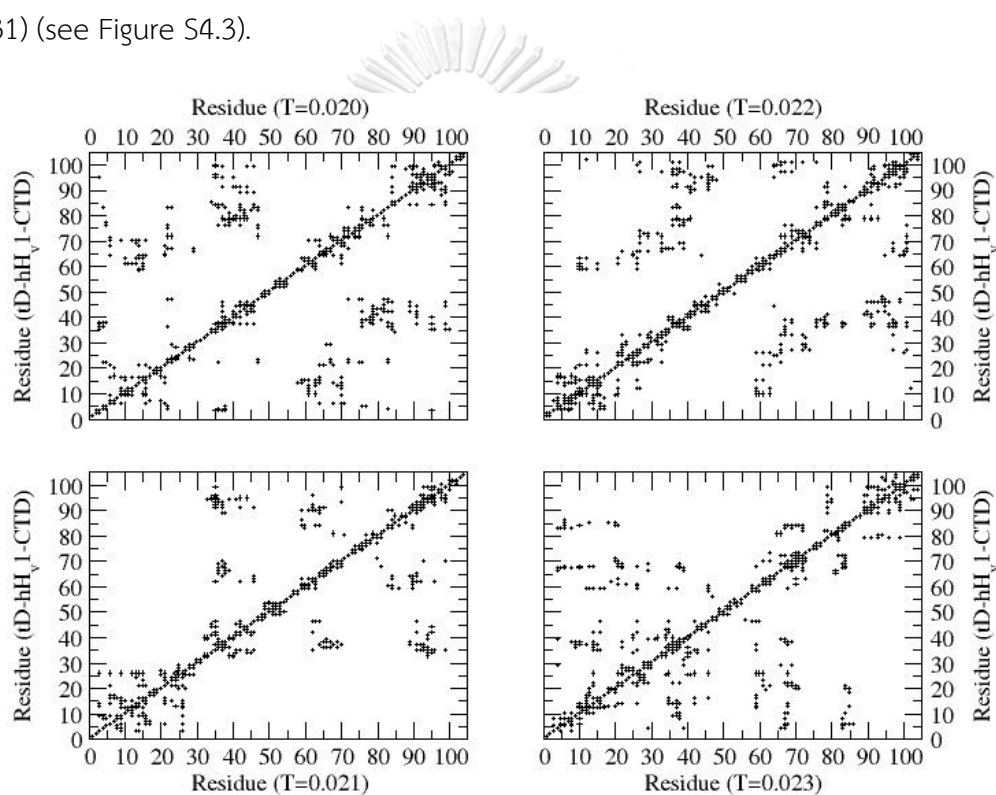


**Figure 4.5**   CGMC data, contact map of tD-hHv1, i.e., its location of residues within the range of their interaction at temperatures $T$ = 0.020–0.023.

### 4.4.2.2 Convergence towards Equilibrium

One of the fundamental ideas behind the Monte Carlo simulations and equilibrium thermodynamics in general is the Markovian accessibility of a microstate from other microstate in a huge ensemble phase space. It can be difficult to reach

equilibrium sometimes in complex systems, particularly those that may involve some degree of frustration. Nevertheless, it is very important to verify the convergence of such quantities as the energy among other physical quantities including radius of gyration (here) to assess the quality of equilibrated data. During the simulation we keep track of various physical quantities as alluded before. Figure 4.6 shows the variation of the radius of gyration ($R_g$) and total energy ($E_c$) of the protein with the MCS time step for a range of temperature ($T$ = 0.021–0.033). We see the slow decay of $R_g$ with the time step at low temperature ($T$ = 0.021) and continue to decay with almost a constant value at long time steps. The convergence towards a near-constant value become faster on raising the temperature ($T$ = 0.023–0.027) as the spread of protein structures consolidates. Although the radius of gyration continues to decay from its random initial configuration on further raising the temperature ($T$ = 0.029), its overall size increases in comparison to lower values at low temperatures ($T$ = 0.023–0.027). Increasing the temperature eventually leads to increase in its radius of gyration with somewhat larger fluctuations ($T$ = 0.021, 0.033). Such a non-monotonic convergence and thermal response require further analysis of the average radius of gyration along with other physical quantities (follows). The energy of the protein (resulting from the residue-residue interaction) seem to fluctuate around zero which is not surprising given the type of residues and their distribution. The energy profile of the residue in equilibrium with the temperature can provide some insight into the segmental structural response and associated energy (see below).
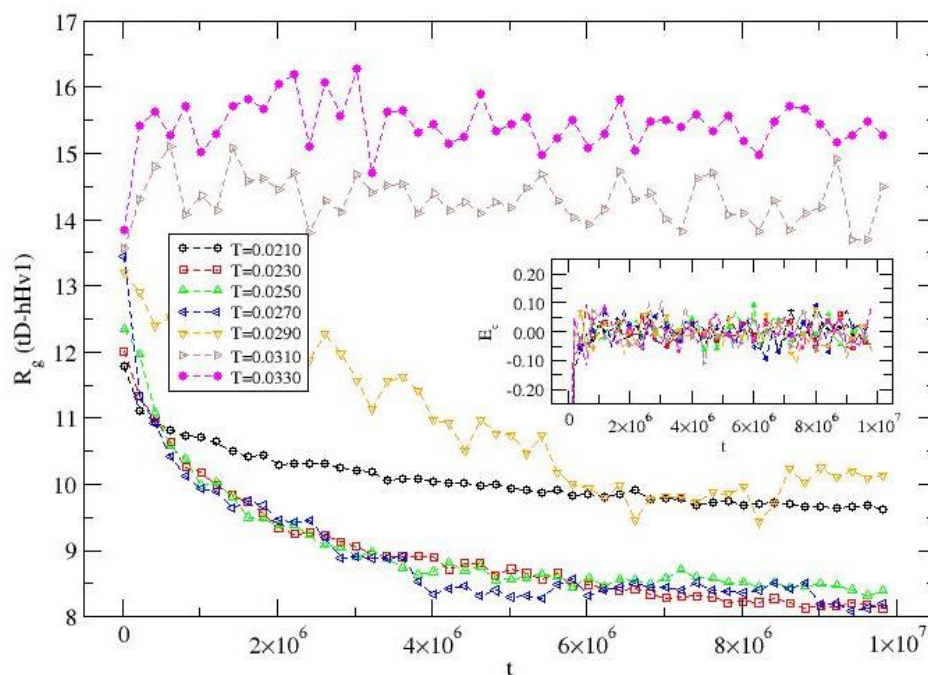
**Figure 4.6**       CGMC data, evolution of the radius of gyration with the time steps for a range of temperature ($T$ = 0.021–0.033). Inset figure shows the variation of the energy of the protein with the time step.

### *4.4.2.3 Radius of Gyration*

Variations of the radius of gyration of the tandem dimer, tD-hHv1, with the temperature are presented in Figures 4.5 covering low to high temperature regimes. In low temperature regime, $R_g$ decays with the temperature, consistent with the results from all-atom MD data of Figure 4.1. In high temperature regime, $R_g$ increases on raising the temperature before saturating to a constant value which is also consistent with the results from all-atom MD simulations. The residue-residue interactions dominate over the thermal noise at low temperatures. The protein structures are almost frozen at extreme low temperatures (i.e. like the inset of Figure 4.2). Thermal agitation stirs the self-organizing residues to adopt to more compact and stable configurations despite reduced entropy. Decay of $R_g$ with $T$ continues until a characteristic value where the thermal energy and residue-residue interactions become comparable and the protein settles into a smallest morphology.

Thermal agitation takes over the residue-residue interactions on increasing the temperature further. Self-organized structure of the protein opens up, the radius of gyration continues to increase until it reaches a constant value when the residue-residue interactions become irrelevant. Variations of the radius of gyration of the monomer hHv1 with the temperature presented in Figure 4.7 insets show a similar trend as that of the tandem dimer tD-hHv1 in low and high temperature regimes on a shifted temperature scale with respect to their size. An eyeball comparison of Figures 4.2 and 4.7 suggests, $T$ = 0.015–0.025 in CGMC corresponds to $T$ = 443–533 K of MD, and similarly, $T$ = 0.027–0.032 in CGMD is equivalent to $T$ = 540–633 K of MD scale. Note that our coarse-grained approach involves phenomenological interaction potential while the MD simulations involve fine-grained force-fields. Regardless of the differences in two approaches, a remarkable similarity in the global response of the protein presented in Figures 4.2 and 4.7 shows that the basic characteristics of the protein can be captured by such an efficient and effective coarse-grained model as the one used extensively in computational soft-matter.
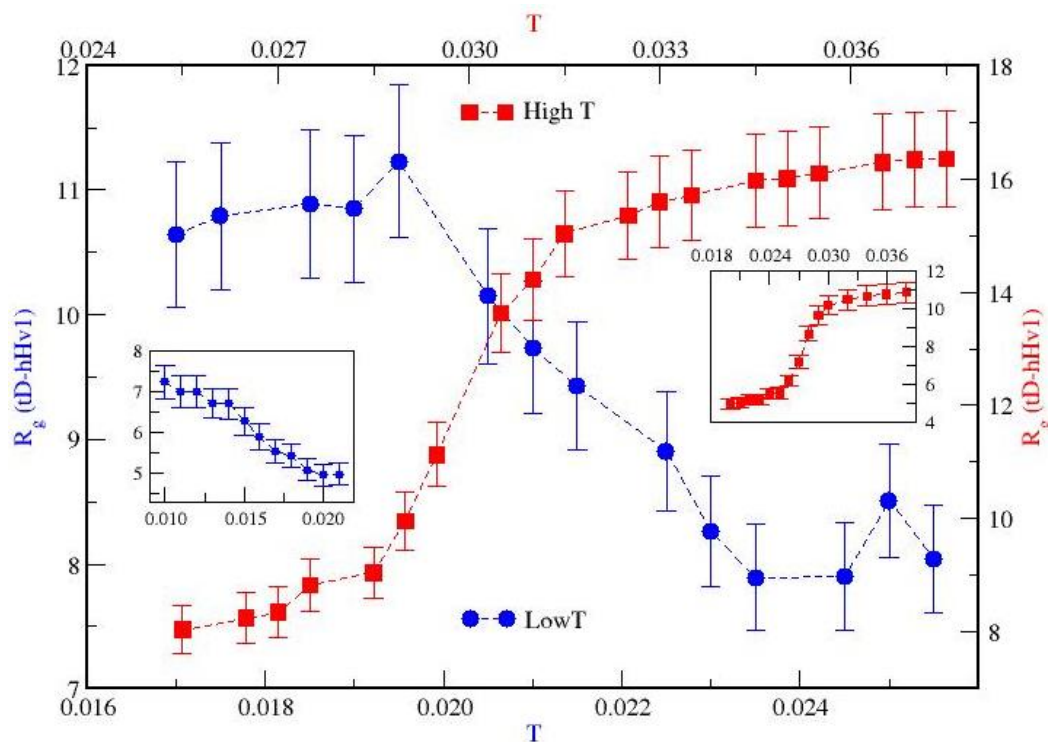
**Figure 4.7**    CGMC data, variation of the radius of gyration, $R_g$ , of the tandem dimer, tD-hHv1 with the temperature, $T$ in low (circle) and high (square) temperature regime. Corresponding data for the monomer hHv1 is presented in the insets. Simulations are performed for 107 time steps, with 10 – 200 independent samples for hHv1 on a $150^3$ lattice and tD-hHv1 on a $340^3$ lattice. Data in last $5 \times 106$ steps in each independent sample are used in estimating the average $R_g$ and its standard deviation shown as error bar. Corresponding variation of $R_g$ of the monomer is included in the inset figures.

*4.4.2.4 Segmental Mobility, Structural, and Energy Profiles*

In order to understand the transport of protons via the connecting pathways along the structure of the protein, it would be informative to examine the local mobility and structural profiles. Since the structural evolution of the hHv1monomer appears to be preserved in its tandem dimer as seen above, we would like to concentrate on tD-hHv1. We define the mobility ($M_n$) of a residue as the probability of its successful moves per unit time step in our CGMC simulation. The mobility profiles

of the residues of tD-hHv1 at representative low and high temperatures are presented in Figures 4.8 and S4.4 respectively. Apart from the residues at the ends, the residues ($^{50}$A–$^{55}$A) in the tandem link and its surrounding are highly mobile even at low temperatures (Figure 4.8). Further, there are localized segments with relatively high (i.e., $^{17}$S–$^{20}$F, $^{85}$F–$^{88}$S) and low mobility ($^{7}$N–$^{15}$K, $^{90}$K–$^{98}$N) with a mirror symmetry around the tandem link. Symmetry in mobility profile seems consistent with the hypothesis used in interpretation of the experimental observations [26, 126, 127] that both monomers provide coordinated but similar pathways for protons transport. Raising the temperature leads to higher mobility (Figure S4.4) while preserving the profile characteristics; the distinction in mobility profiles of the local segments nevertheless becomes more pronounced at lower temperatures (Figure 4.8).
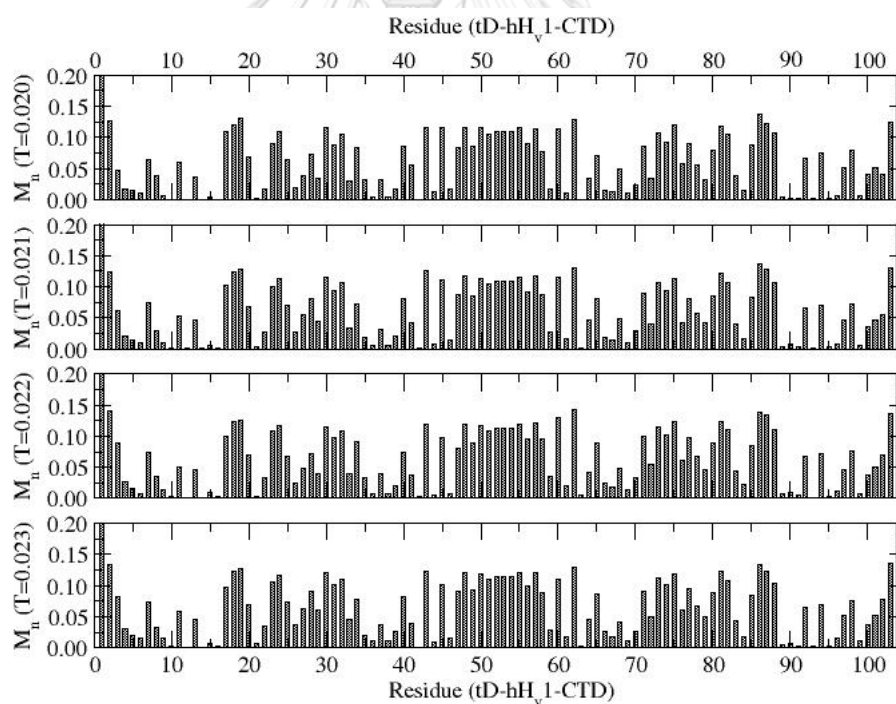


**Figure 4.8**      CGMC data, mobility (successful hops per unit MCS) profile of tD-hHv1-CTD at $T$ = 0.020–0.023.

As described above, the snapshots and contact maps provide a glimpse into the instantaneous structural evolution of the protein. Estimates of the average number ($N_n$) of residues within the range of interaction, a measure of the average contact map, in steady state may provide some insight into the segmental morphology of the protein. The segmental structural profiles of the protein at low and high temperatures (corresponding to mobility profiles in Figures 4.8 and S4.4) are presented in Figures 4.9 and S4.5 respectively. We see that the self-assembled segmental residue profiles (Figures 4.9, S4.5) are complementary to corresponding mobility profiles. Segments with larger number ($N_n$) of interacting residues have lower mobility. Obviously, the segment containing the tandem link has lowest number ($N_n$) of surrounding residue and high mobility, i.e. highly flexible. Segments with the higher value of $N_n$ have more connecting pathways for proton transport. Distribution of the segmental connectivity pathways is symmetric to tandem link. For example, residues $^3$R–$^{16}$E and $^{89}$E–$^{102}$R have relatively high values of $N_n$ (high contact density with local globular morphologies) with a mirror symmetry around the tandem link. Note that the distinction in distribution of segmental globular structures decreases on increasing the temperature (Figure S4.5) with low values of $N_n$ at high temperatures where the residue-residue interaction becomes irrelevant and the structure of the protein conforms to a random-coil conformation.
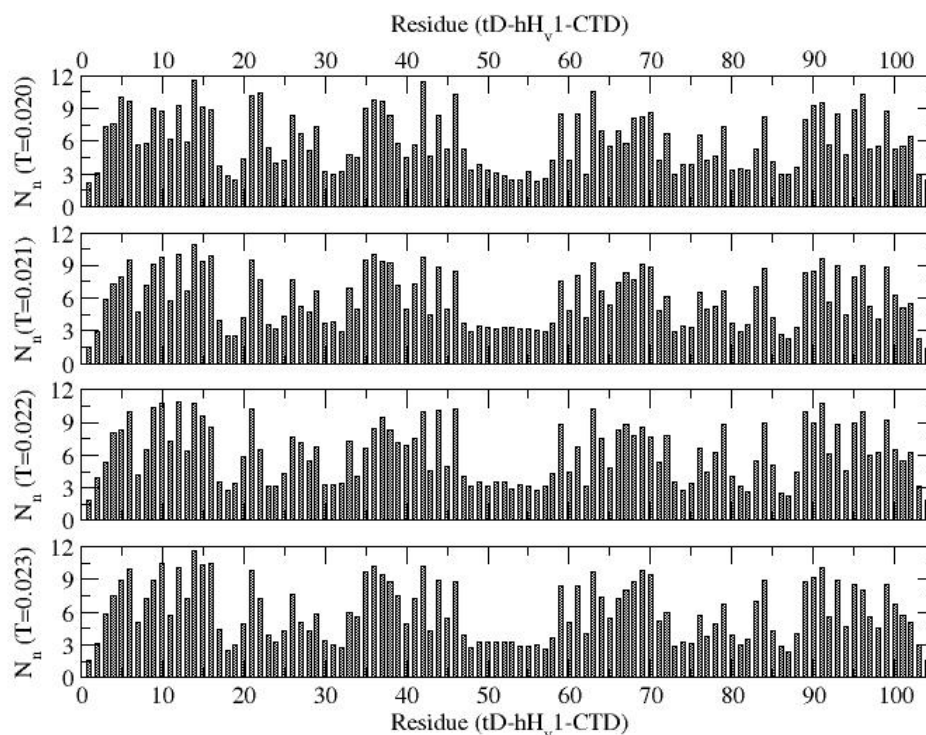
**Figure 4.9**    CGMC data, average number $N_n$ of residues within the range of interaction (a measure of the contact density profile) at temperature $T$ = 0.020–0.023. Statistics is the same as Figure 4.5.

The energy profile of the residue at the range of temperature is presented in Figure 4.10. Distribution of segmental energy appears to remain almost the same with the temperature, except the magnitude of energy which increases with the temperature. Further, note that energy profile is almost symmetric around the tandem link $^{50}$A–$^{55}$A.
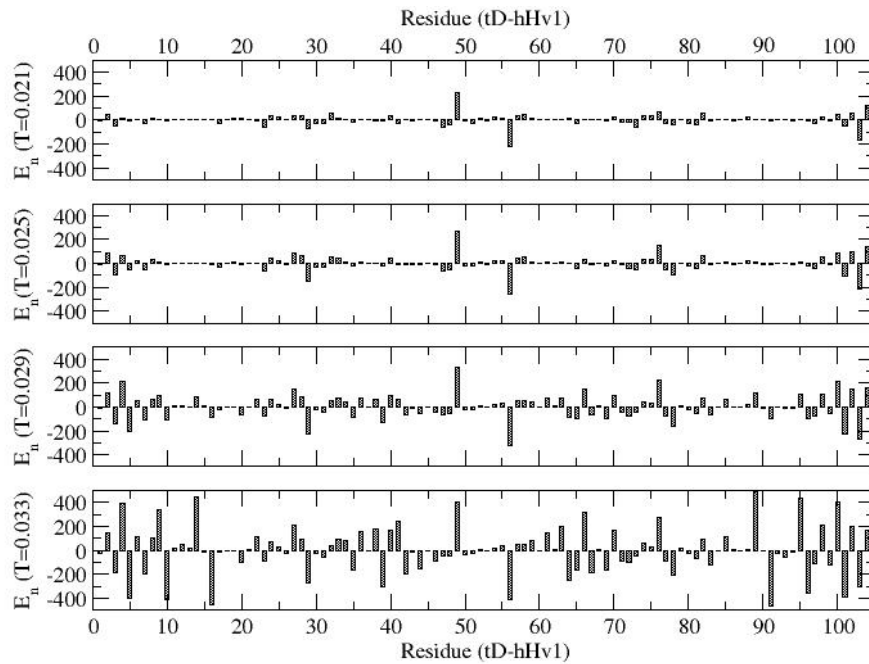
**Figure 4.10**    CGMC data, average energy of each residue for a range of temperature ($T$ = 0.021–0.033)

### 4.4.2.5 Structure Factor

The overall spatial distribution of residues in conformational evolution of the protein may vary with the length scale and temperature. Structural variations over the multiple length scales can be examined by analyzing the structure factor $S(q)$ which is defined as,

$$S(q) = \left\langle \frac{1}{N} \left| \sum_{j=1}^{N} e^{-i\vec{q}\cdot\vec{r}_j} \right|^2 \right\rangle_{|\vec{q}|}$$

(4.2)

where $r_j$ is the position of each residue $j$ of the protein and $|q| = \frac{2\pi}{\lambda}$ is the wave vector with wavelength $\lambda$. One may view the structure factor as a measure of the volume and therefore the number of residue or mass and the wave-length as a linear scale ($r$) for the spread of the protein. Variation of the structure factor with the wavelength or linear length at low and high temperature regimes are presented in Figures 4.11 and 4.12. We see that the variation of the structure factor with the

temperature is not as much at the low temperatures (Figure 4.11) as it is at the high temperatures (Figure 4.12). One may be able to quantify the spread of residues by scaling of the structure factor with the wave-length comparable to radius of gyration of the protein, i.e. $S(q) \propto r^D$, where $D$ is an effective dimension. The slope of $S(q)$ versus $r$ on a log-log scale (Figures 4.11, 4.12) are the estimates of $D$.
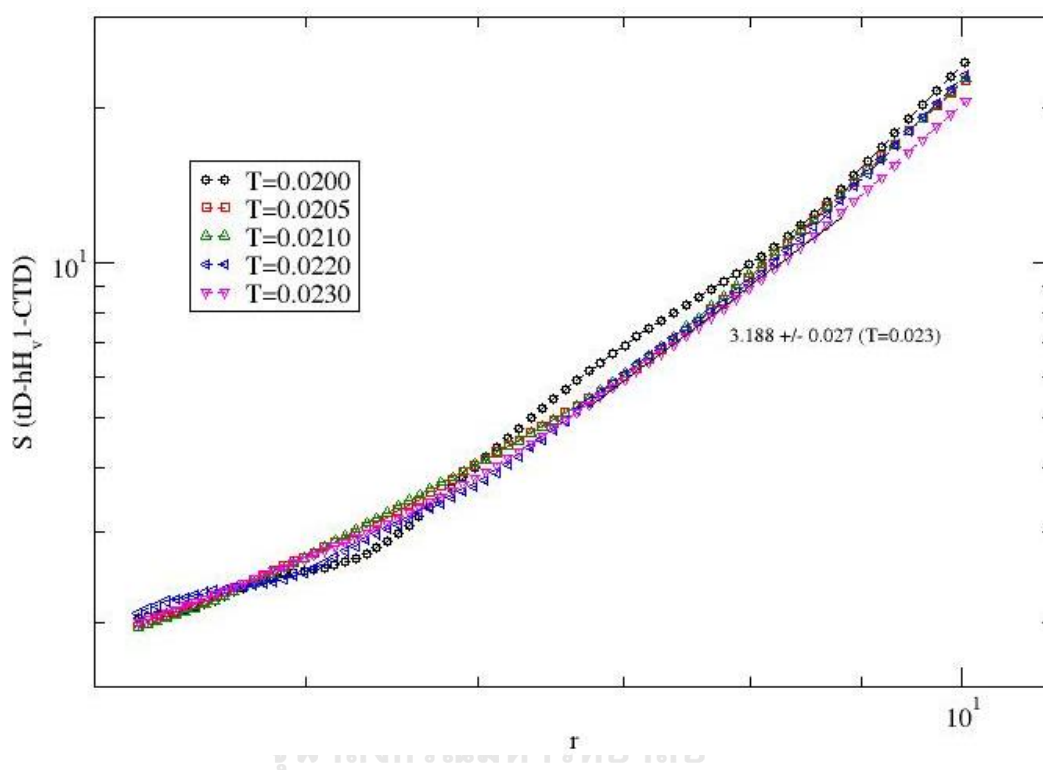


**Figure 4.11**    CGMC data, structure factor $S(q)$ versus wave length $r$ at temperature $T$ = 0.020–0.023 on a log-log scale. Slopes of a set of representative data points at $T$ = 0.023 over length scales comparable to its radius of gyration is an estimate of the effective dimension of the residue apread. Scale of x-axis ranges from 4 to 22.

The radius of gyration of the protein (tD-hHv1) is in the range of 8 – 10 where $D \sim 3$ at $T$ = 0.020, i.e. the protein conforms to a globular (solid) morphology. Comparatively, there is a larger variation in $S(q)$ and radius of gyration ($R_g \sim 9$–16) of the protein at higher temperatures ($T$ = 0.029–0.034). At $T$ = 0.029, the protein remains relatively compact ($D \sim 3$) and becomes random-coil on raising the

temperature further (i.e $D \sim 2$, $T = 0.032$). On a larger length scale, the structure of the protein appears to be linear ($D \sim 1.3$) at high temperature (Figure 4.9). The distribution of residues over the length scale is critical in connectivity for the proton transport along its conformation. In general, smaller the spread (i.e. lower $R_g$, higher $D$), the larger is the connectivity with more pathways in a compact morphology. Thus, the diversity in structural variations in the high temperature regime suggests that proton transport along the protein conformations will show appreciable variations with the temperature.
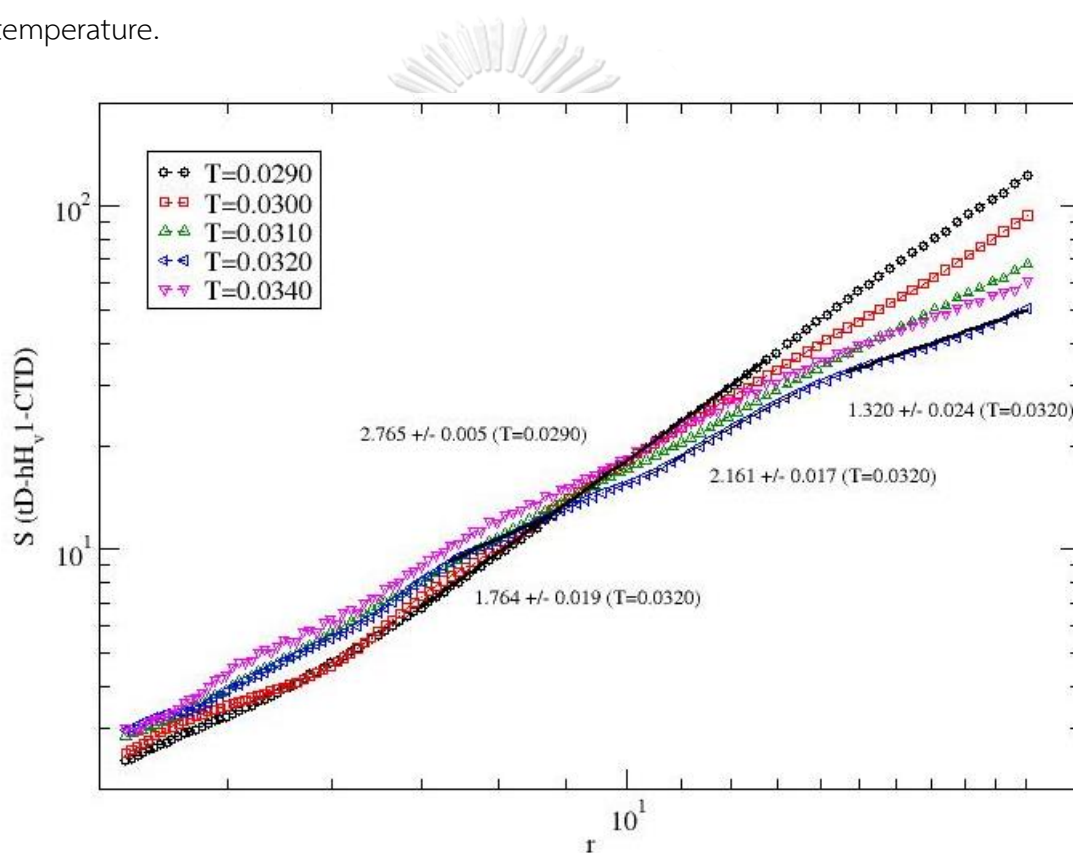


**Figure 4.12**    CGMC data, structure factor $S(q)$ versus wave vector $q$ $\left( |q| = \frac{2\pi}{\lambda} \right)$ at temperature $T = 0.029$–$0.034$ on a log-log scale. Slopes of some representative data points over relevant length scales (included) are estimates of effective dimension of the residue spread. Scale of x-axis ranges from 4 to 22.

**4.5 CONCLUSION**

Conformational response of a dimer (hHv1) and its tandem dimer (tD-hHv1) of C-terminal domain (CTD) of hHv1 channel to temperature are studied by two very different but complementary computational methods: all-atom MD and all-residue coarse-grained MC. All-atom MD simulations incorporate atomic scale structural details while all-residue CGMC covers length scales spanning beyond the size of a residue and involves a knowledge-based residue-residue interaction. Because of the ease and efficiency of implementing the CGMC approach, we are able to analyze a range of local and global physical quantities with a coarse-grained phenomenological interaction potential. A physical quantity such as radius of gyration can be readily calculated in both all-atom as well as all-residue approaches which provides a mean to verify results, i.e., the qualitative trend in variation of $R_g$ with the temperature $T$. Fine (atomic) resolution of all-atom MD simulations data involving trajectories in real space can be used for calibrating the scales (reduced units) for variables used in CG simulations as seen above. Based on data from two approaches (all-atom MD, all-residue MC), we find that both monomer (hHv1) and dimer (tD-hHv1) exhibit similar thermal responses but on different scales (e.g. the range of temperature is based on the size of the proteins) which seem consistent with the hypotheses used in interpreting the laboratory observations.

Visual analysis of the snapshots shows a general trend in variation of the overall conformational spread and structure variability (i.e. distribution of self-organizing fibrous and globular segments in protein) with the temperature. We are able to identify two different temperature regimes, low and high, by visual inspection of the spread (size) of the protein. In the low temperature regime, we find that the size (spread) of the protein decreases on increasing the temperature, an unexpected observation since adding thermal energy by increasing the temperature generally enhances the spread.

Decrease in size of the protein implies increase in compactness and therefore the connected pathways which may result in increase in proton transport on increasing the temperature. In the high temperature regime, on the other hand, we do observe the expected trend, i.e. the conformation of the protein spreads out more on increasing the temperature. Contact maps help in synthesizing the distribution of loops (concentration and size) along the contour of the protein. Decays in loops, i.e., loss in secondary/tertiary structures occur at high temperatures. The analysis of the contact map seems consistent with the visualization. Contact maps of the tandem dimer (tD-hHv1) exhibits the appearance of a mirror symmetry in structural variability around the tandem link which leads us to believe that the two monomeric units ($^1$H–$^{49}$I and $^{56}$I–$^{104}$H) respond somewhat similarly. These observations based on the segmental location of residues and its consequences on the global physical properties are quantified by analyzing the radius of gyration of the protein, its structure factor, and mobility and contact profiles.

Residues continue to perform their stochastic motion to organize in stable local structures as the protein explores its conformational phase space. We have analyzed mobility profile and the interacting contact profiles (average contact map) of the residues to assess the local segmental structures and its stability. Mobility profile of the tandem dimer shows some degree of mirror symmetry in distribution of segments with high and low mobility. For example, the localized segments with low ($^7$N–$^{15}$K, $^{90}$K–$^{98}$N) and high ($^{17}$S–$^{20}$F, $^{85}$F–$^{88}$S) mobility are symmetrically distributed around the high mobile tandem linker in low temperature regime. Segmental residue contact profiles at low and high temperatures are complementary to corresponding mobility profiles. For example, the self-assembled segments with larger number of interacting residues show lower mobility.

How does the radius of gyration $R_g$ (a global physical quantity to estimate the overall size) of the protein (both monomer (hHv1) and tandem dimer (tD-hHv1) depend on temperature? From both all-atom MD and CGMC simulations, we find that the radius of gyration of the protein M-hHv1 decreases on increasing the temperature which is not a typical thermal response, i.e., the protein becomes more compact with reduced entropy on increasing the temperature in low temperature regime. In native phase, the residue-residue interaction dominates over thermal agitation – an interaction-controlled thermal-induced structural response. On further increasing the temperature, we find that the radius of gyration increases on increasing the temperature before reaching a steady-state value. Thermal agitation wins over the residue-residue interaction, the protein expands through its random-coil structures until it conforms to a linear configuration – the thermal-controlled structural response in its denatured phase.

The spread of residues over the length scales spanning over the radius of gyration can be quantified by analyzing the structure factor. Scaling of the structure factor $S(q)$ of the protein with the wave length ($r$), $S(q) \propto r^D$ provides an estimate of the effective dimension $D$ for the distribution of residues. We find $D \sim 3$ at low temperatures (native phase) where the protein conforms to a globular (solid) morphology, i.e., the interacting residues along the protein backbone are connected via multiple pathways (not just the peptide bonds). In high temperature regime, the structures of the protein undergo various structural transformation, i.e., a globular conformation ($D \sim 3$) to random-coil ($D \sim 2$) on raising the temperature and become very tenuous ($D \sim 1.3$) on large scales. In general, smaller the spread (higher $D$), the larger is the connectivity with more pathways in a compact morphology. Thus, the variations in conformations may lead to corresponding variations for the proton transport along the protein structure.

In summary, we have made a number of interesting observations from the investigation of a protein (hHv1) by an all-atom MD and a coarse-grained MC simulation. i. We find that the radius of gyration of the protein decreases on increasing the temperature, a negative thermal response, in its native phase. ii. This is followed by a positive thermal response, in contrast, i.e. expansion in its denatured phase on further heating. iii. The scaling of the structure factor provides an estimate of the universal exponent $v = \frac{1}{D}$ in both native and denatured phases where $v = \frac{1}{3}$ in native phase consistent with the recent findings [135] involving the finite-size scaling at the critical point of the ensemble of NMR structures. iv. We are not only able to quantify the extent of segmental globularity but also to identify residues that contribute to their structural stability. v. All atom MD data helps in calibrating the scales for the coarse-grained MC data and verification of the structural evolution of the protein. vi. Similarity in findings from MD and CGMC data shows that the specific characteristics the protein resulting from an all-atom MD simulation with fine-grained force-field can be captured by the phenomenological interaction potentials of an effective and efficient coarse-grained approach. The structural responses in the native and denatured phases (in low and high temperature regimes) predicted here suggest that proton transport along pathways provided by the protein conformations will show corresponding variations with the temperature which may help in hypothesizing the interpretation of the experimental observations qualitatively. Many important factors such as the crowded cellular environments are yet to be considered, one of the future goals in our on-going systematic investigations.

**4.6 SUPPLEMENTARY INFORMATION**
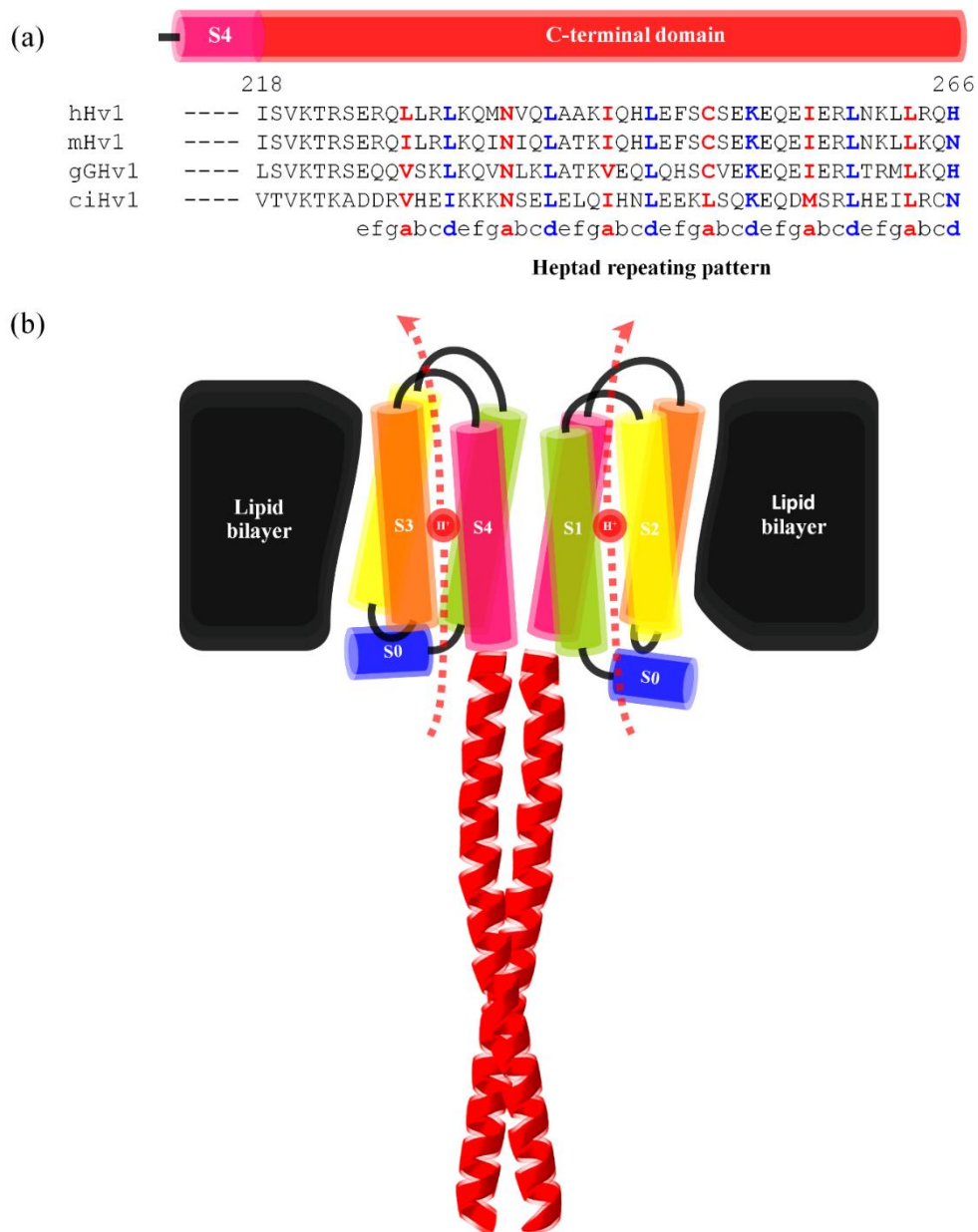




**Figure S4.1** (a) Sequence alignment of the C-terminal domain of Hv1 channel form various species with heptad repeat (abcdefg)$_n$ labeling. Position 'a' and 'd' are hydrophobic positions highlighted in blue and red, respectively, and 'e' and 'g' are charged positions. (b) A dimeric hHv1 model of C-terminal domain based on experimental data and X-ray crystal structure.
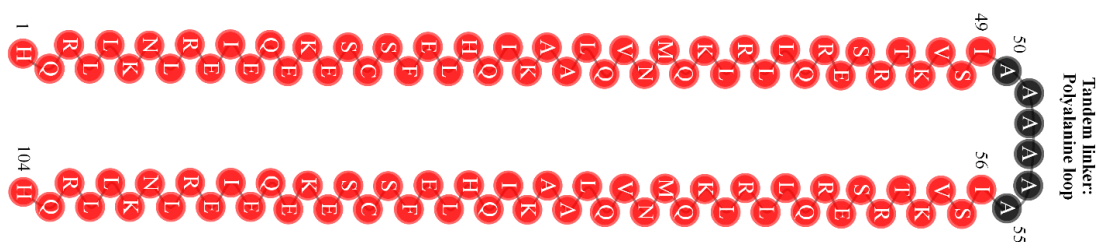
**Figure S4.2**    Coarse-grain model of monomer hHv1 and its tandem dimer tD-hHv1-CTD used in MC simulations
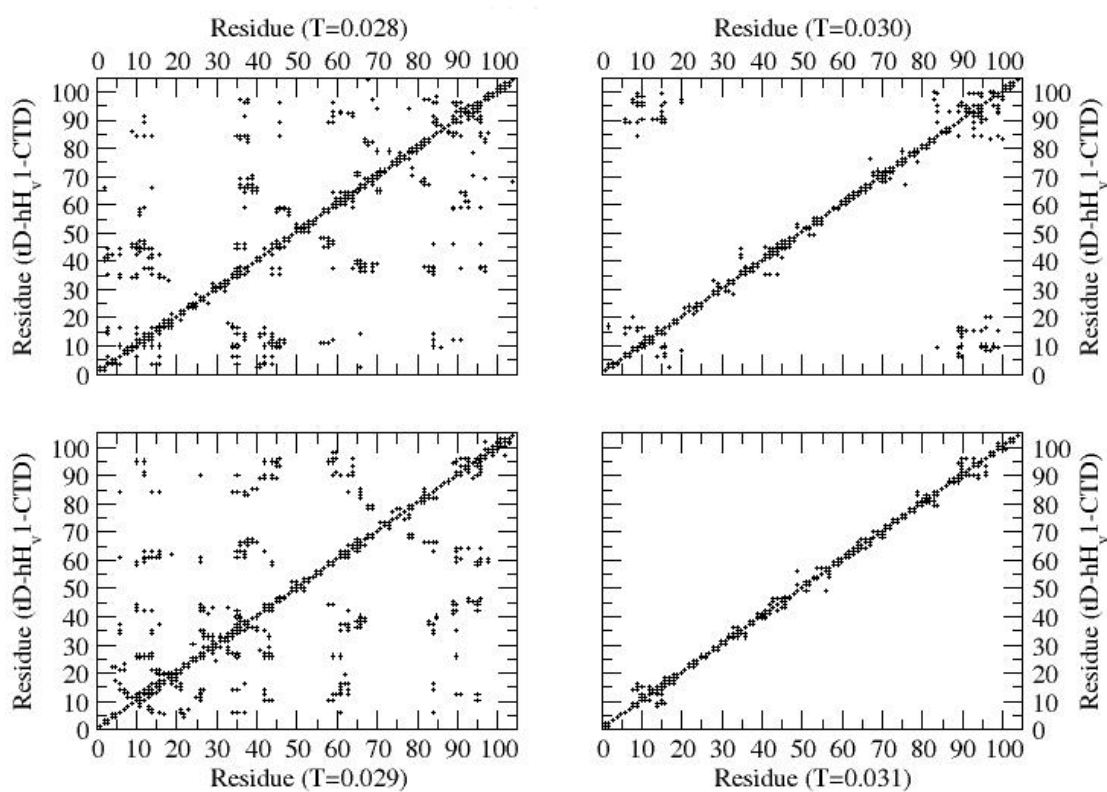


**Figure S4.3**    Contact map of tD-hHv1-CTD, i.e., its location of residues within the range of their interaction at temperatures $T$ = 0.028–0.031.
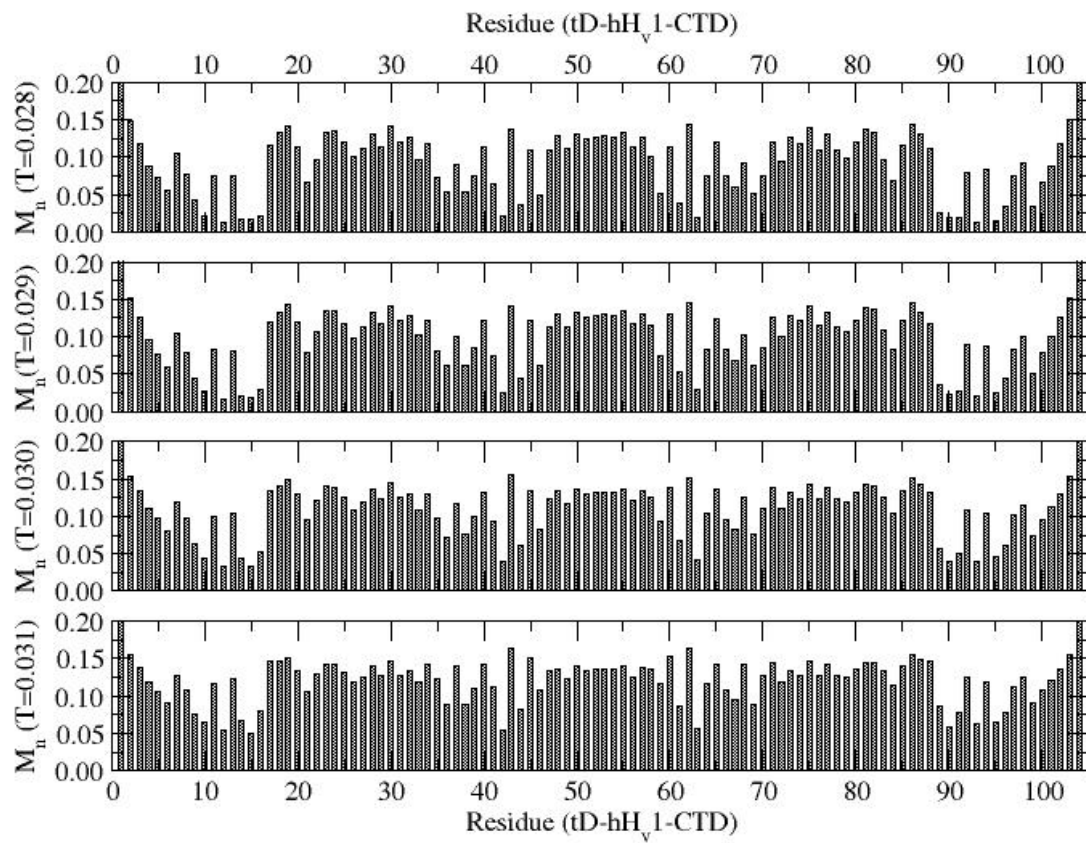
**Figure S4.4**     Mobility (successful hops per unit MCS) profile of tD-hHv1-CTD at $T$ = 0.028–0.031. Statistics is the same as in Figure 4.5.
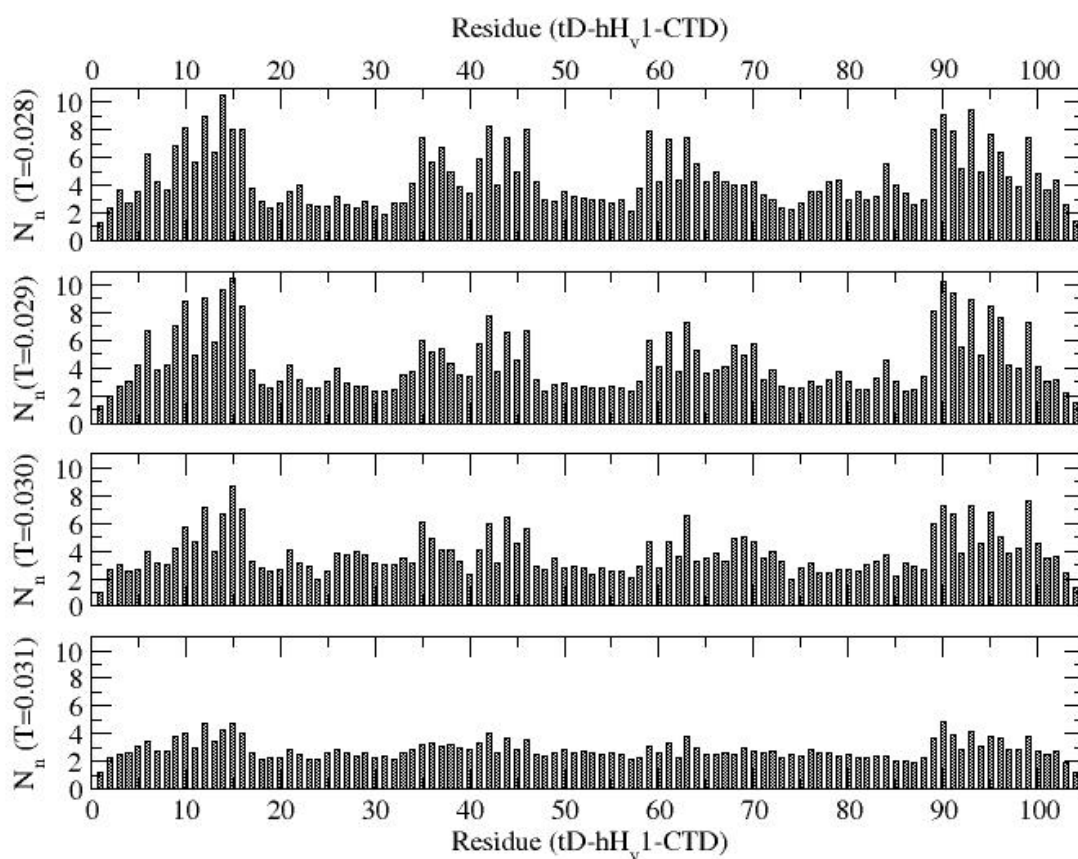
**Figure S4.5**   Average number   of residues within the range of interaction (a measure of the contact density profile) at temperature $T$ = 0.028–0.031. Statistics is the same as Figure 4.5.

จุฬาลงกรณ์มหาวิทยาลัย

**CHULALONGKORN UNIVERSITY**

# CHAPTER V: CONCLUSIONS

## 5.1 CONCLUSIONS

Over past decades, the basic functions of the Hv1 channel has not yet been clearly answered although it involved in various physiological processes. Through the dissertation work, each part of the Hv1 channel structure associated with its function was characterized. Several computational techniques were applied to investigate the dynamic and thermodynamic properties.

In first section, the molecular interactions of C-terminal domain in dimer stabilization were studies by means of MD simulation. There were four types of key interactions contributing the dimer stability. Salt-bridges significantly affected structural packing upon pH condition. Side-chain dynamics of N235 pairs, forming hydrogen bond, remarkably destabilized the dimer structure by disruption of hydrophobic core while uniformly hydrophobic character exhibited a higher stability. Dimer structure of the CTD showed the highest stability in presence of disulfide bond of C249 pair indicated that it was the major contributor for. This work provides the knowledge about the role of CTD on dimer stabilization which could improve the understanding of channel activation and cooperative gating via interfacial interactions of the Hv1 channels.

The Hv1 channel is known that it responds to the pH change. The mechanism of pH sensing could involve with titratable side chains of particular amino acids. The second part of this thesis discussed about the effect of ionization state of charged residues on conformational change. MD results showed that arrangement of salt-bridge network in the VSD core involving with S4 arginines (R1, R2 and R3) and their counter charges was impacted by modification of protonation state, especially at more acidic condition. The resting conformation of Hv1 channel was altered by disrupting the salt-bridge network, leading wider pore radius. Extended pore size significantly reduced solvation energy barrier resulting in a higher hydration of water-filled crevice.

Furthermore, dipole orientation of water molecules in the upper half appeared to have an opposite direction to those in the lower half of pore. The water orientation in the upper and lower regions was swapped as a result of changing the protonation state.

In the last section, conformational response of C-terminal domain of hHv1 channel (hHv1-CTD) to temperature was studied. Two complementary methods were used which were all-atom MD (AAMD) and a coarse-grained MC (CGMC) simulations. Denaturation behavior of the hHv1-CTD was divided into low and high temperature regimes. Decreasing of radius of gyration ($R_g$) was observed upon heating in the low temperature regime, indicating compactness of the structure. Then, the $R_g$ value switched to be higher in the high temperature regime. Conformation of the protein spread out more on increasing the temperature, indicating a fully denatured state. The results obtained from CGMC were also highly consistent with the MD results.

## 5.2 RESEARCH LIMITATIONS

The main reasons are computational cost and resources. Currently available computer power is limited. Then, large-scale simulation is not possible due to time and disk space consuming.

# REFERENCES

[1] Boyle, J., *Molecular biology of the cell, 5th edition by B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter.* Biochemistry and Molecular Biology Education, 2008. **36**(4): p. 317-318.

[2] DeCoursey, T.E., *Voltage-gated proton channels: molecular biology, physiology, and pathophysiology of the Hv family.* Physiological Reviews, 2013. **93**(2): p. 599-652.

[3] Fogel, M. and J.W. Hastings, *Bioluminescence: Mechanism and Mode of Control of Scintillon Activity.* Proceedings of the National Academy of Sciences of the United States of America, 1972. **69**(3): p. 690-693.

[4] Rebolledo, S., F. Qiu, and H. Peter Larsson, *Molecular structure and function of Hv1 channels.* Wiley Interdisciplinary Reviews: Membrane Transport and Signaling, 2012. **1**(6): p. 763-777.

[5] Metropolis, N., et al., *Equation of State Calculations by Fast Computing Machines.* The Journal of Chemical Physics, 1953. **21**(6): p. 1087-1092.

[6] Periole, X. and S.-J. Marrink, *The Martini Coarse-Grained Force Field*, in *Biomolecular Simulations: Methods in Molecular Biology*, L. Monticelli and E. Salonen, Editors. 2013, Humana Press: Totowa, NJ. p. 533-565.

[7] Ramsey, I.S., et al., *Hv1 proton channels are required for high-level NADPH oxidase-dependent superoxide production during the phagocyte respiratory burst.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(18): p. 7642-7647.

[8] Capasso, M., et al., *HVCN1 modulates BCR signal strength via regulation of BCR-dependent generation of reactive oxygen species.* Nature Immunology, 2010. **11**(3): p. 265-272.

[9] DeCoursey, T.E., *Voltage-gated proton channels find their dream job managing the respiratory burst in phagocytes.* Physiology, 2010. **25**(1): p. 27-40.

[10] Capasso, M., *Regulation of immune responses by proton channels.* Immunology, 2014. **143**(2): p. 131-137.

[11]    Lishko, P.V., et al., *Acid extrusion from human spermatozoa is mediated by flagellar voltage-gated proton channel.* Cell, 2010. **140**(3): p. 327-337.

[12]    Lishko, P.V., et al., *The control of male fertility by spermatozoan ion channels.* Annual Review of Physiology, 2012. **74**: p. 453-475.

[13]    Takeshita, K., et al., *X-ray crystal structure of voltage-gated proton channel.* Nature Structural & Molecular Biology, 2014. **21**: p. 352-357.

[14]    Doyle, D.A., et al., *The structure of the potassium channel: molecular basis of K⁺ conduction and selectivity.* Science, 1998. **280**(5360): p. 69-77.

[15]    Long, S.B., E.B. Campbell, and R. MacKinnon, *Crystal structure of a mammalian voltage-dependent Shaker family K⁺ channel.* Science, 2005. **309**(5736): p. 897-903.

[16]    Payandeh, J., et al., *The crystal structure of a voltage-gated sodium channel.* Nature, 2011. **475**(7356): p. 353-358.

[17]    Ramsey, I.S., et al., *A voltage-gated proton-selective channel lacking the pore domain.* Nature, 2006. **440**: p. 1213-1216.

[18]    Koch, H.P., et al., *Multimeric nature of voltage-gated proton channels.* Proceedings of the National Academy of Sciences of the United States of America, 2008. **105**(26): p. 9111-9116.

[19]    Lee, S.-Y., J.A. Letts, and R. MacKinnon, *Dimeric subunit stoichiometry of the human voltage-dependent proton channel Hv1.* Proceedings of the National Academy of Sciences of the United States of America, 2008. **105**(22): p. 7692-7695.

[20]    Tombola, F., M.H. Ulbrich, and E.Y. Isacoff, *The voltage-gated proton channel Hv1 has two pores each controlled by one voltage sensor.* Neuron, 2008. **58**(4): p. 546-556.

[21]    Musset, B., et al., *Oligomerization of the voltage-gated proton channel.* Channels, 2010. **4**(4): p. 260-265.

[22]    Gonzalez, C., et al., *Strong cooperativity between subunits in voltage-gated proton channels.* Nature Structural & Molecular Biology, 2010. **17**(1): p. 51-56.

[23]     Li, S.J., et al., *The role and structure of the carboxyl-terminal domain of the human voltage-gated proton channel Hv1.* The Journal of Biological Chemistry, 2010. **285**(16): p. 12047-12054.

[24]     Musset, B., et al., *Zinc inhibition of monomeric and dimeric proton channels suggests cooperative gating.* The Journal of Physiology, 2010. **588**(9): p. 1435-1449.

[25]     Tombola, F., et al., *The opening of the two pores of the Hv1 voltage-gated proton channel is tuned by cooperativity.* Nature Structural & Molecular Biology, 2010. **17**(1): p. 44-50.

[26]     Fujiwara, Y., et al., *The cytoplasmic coiled-coil mediates cooperative gating temperature sensitivity in the voltage-gated $H^+$ channel Hv1.* Nature Communications, 2012. **3**: p. 816.

[27]     Fujiwara, Y., et al., *Gating of the designed trimeric/tetrameric voltage-gated H+ channel.* The Journal of Physiology, 2013. **591**(3): p. 627-640.

[28]     Qiu, F., et al., *Subunit interactions during cooperative opening of voltage-gated proton channels.* Neuron, 2013. **77**(2): p. 288-298.

[29]     Fujiwara, Y., et al., *Structural characteristics of the redox-sensing coiled coil in the voltage-gated $H^+$ channel.* The Journal of Biological Chemistry, 2013. **288**(25): p. 17968-17975.

[30]     Zhao, Q., C. Li, and S.J. Li, *The pH-sensitive structure of the C-terminal domain of voltage-gated proton channel and the thermodynamic characteristics of Zn2+ binding to this domain.* Biochemical and Biophysical Research Communications, 2015. **456**(1): p. 207-212.

[31]     Okochi, Y., et al., *Voltage-gated proton channel is expressed on phagosomes.* Biochemical and Biophysical Research Communications, 2009. **382**(2): p. 274-279.

[32]     Okamura, Y., Y. Fujiwara, and S. Sakata, *Gating mechanisms of voltage-gated proton channels.* Annual Review of Biochemistry, 2015. **84**(1): p. 685-709.

[33]     Li, Q., et al., *Resting state of the human proton channel dimer in a lipid bilayer.* Proceedings of the National Academy of Sciences of the United States of America, 2015. **112**(44): p. E5926-E5935.

[34]  Xiang, Z., C.S. Soto, and B. Honig, *Evaluating conformational free energies: the colony energy and its application to the problem of loop prediction.* Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(11): p. 7432.

[35]  Laskowski, R.A., et al., *PROCHECK: a program to check the stereochemical quality of protein structures.* Journal of Applied Crystallography, 1993. **26**(2): p. 283-291.

[36]  Humphrey, W., A. Dalke, and K. Schulten, *VMD: visual molecular dynamics.* Journal of Molecular Graphics, 1996. **14**(1): p. 33-38.

[37]  Supunyabut, C., S. Fuklang, and P. Sompornpisut, *Continuum electrostatic approach for evaluating positions and interactions of proteins in a bilayer membrane.* Journal of Molecular Graphics and Modelling, 2015. **59**: p. 81-91.

[38]  Baker, N.A., et al., *Electrostatics of nanosystems: application to microtubules and the ribosome.* Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(18): p. 10037-10041.

[39]  Sitkoff, D., K.A. Sharp, and B. Honig, *Accurate calculation of hydration free energies using macroscopic solvent models.* The Journal of Physical Chemistry, 1994. **98**(7): p. 1978-1988.

[40]  Dolinsky, T.J., et al., *PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations.* Nucleic Acids Research, 2004. **32**(suppl_2): p. W665-W667.

[41]  Dolinsky, T.J., et al., *PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations.* Nucleic Acids Research, 2007. **35**(Web Server issue): p. W522-W525.

[42]  Phillips, J.C., et al., *Scalable molecular dynamics with NAMD.* Journal of Computational Chemistry, 2005. **26**(16): p. 1781-1802.

[43]  MacKerell Jr, A.D., et al., *All-atom empirical potential for molecular modeling and dynamics studies of proteins.* The Journal of Physical Chemistry B, 1998. **102**(18): p. 3586-3616.

[44]     Li, H., A.D. Robertson, and J.H. Jensen, *Very fast empirical prediction and rationalization of protein pKa values.* Proteins: Structure, Function, and Bioinformatics, 2005. **61**(4): p. 704-721.

[45]     Jorgensen, W.L., et al., *Comparison of simple potential functions for simulating liquid water.* The Journal of Chemical Physics, 1983. **79**(2): p. 926-935.

[46]     Izadi, S., B. Aguilar, and A.V. Onufriev, *Protein–ligand electrostatic binding free energies from explicit and implicit solvation.* Journal of Chemical Theory and Computation, 2015. **11**(9): p. 4450-4459.

[47]     Meulemans, A., et al., *Defining the pathogenesis of the Human Atp12p W94R mutation using a Saccharomyces cerevisiae yeast model.* Journal of Biological Chemistry, 2010. **285**(6): p. 4099-4109.

[48]     Wang, J., P. Cieplak, and P.A. Kollman, *How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules?* Journal of Computational Chemistry, 2000. **21**(12): p. 1049-1074.

[49]     Chauvenet, W., *A Manual of spherical and practical astronomy-Vol. 1: Spherical astronomy; Vol. 2: Theory and use of astronomical instruments. Method of least squares.* New York: Dover Publication, 1960, 5th ed., revised and corr., 1960.

[50]     Mason, J.M., U.B. Hagemann, and K.M. Arndt, *Improved stability of the Jun-Fos activator protein-1 coiled coil motif: a stopped-flow circular dichroism kinetic analysis.* Journal of Biological Chemistry, 2007. **282**(32): p. 23015-23024.

[51]     Chen, Z., et al., *Intradimer/intermolecular interactions suggest autoinhibition mechanism in endophilin A1.* Journal of the American Chemical Society, 2014. **136**(12): p. 4557-4564.

[52]     Steinmetz, M.O., et al., *Molecular basis of coiled-coil formation.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(17): p. 7062-7067.

[53]     Li, Q., et al., *Structural basis of lipid-driven conformational transitions in the KvAP voltage-sensing domain.* Nature Structural & Molecular Biology, 2014. **21**(2): p. 160-166.

[54]   Truebestein, L. and T.A. Leonard, *Coiled-coils: The long and short of it.* BioEssays, 2016. **38**(9): p. 903-916.

[55]   Zhou, N.E., C.M. Kay, and R.S. Hodges, *Disulfide bond contribution to protein stability: positional effects of substitution in the hydrophobic core of the two-stranded a-helical coiled-coil.* Biochemistry, 1993. **32**(12): p. 3178-3187.

[56]   Lavigne, P., et al., *Interhelical salt bridges, coiled-coil stability, and specificity of dimerization.* Science, 1996. **271**(5252): p. 1136-1138.

[57]   Mason, J.M. and K.M. Arndt, *Coiled coil domains: stability, specificity, and biological implications.* ChemBioChem, 2004. **5**(2): p. 170-176.

[58]   Mayor, U., et al., *Protein folding and unfolding in microseconds to nanoseconds by experiment and simulation.* Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(25): p. 13518-13522.

[59]   Day, R., et al., *Increasing temperature accelerates protein unfolding without changing the pathway of unfolding.* Journal of Molecular Biology, 2002. **322**(1): p. 189-203.

[60]   Zhou, H.-X., *Rate theories for biologists.* Quarterly reviews of biophysics, 2010. **43**(2): p. 219-293.

[61]   Russell, R.B. and G.J. Barton, *Multiple protein sequence alignment from tertiary structure comparison: assignment of global and residue confidence levels.* Proteins, 1992. **14**(2): p. 309-323.

[62]   Roberts, E., et al., *MultiSeq: unifying sequence and structure data for evolutionary analysis.* BMC Bioinformatics, 2006. **7**(1): p. 382.

[63]   Wang, Y., et al., *Clinicopathological and biological significance of human voltage-gated proton channel Hv1 protein overexpression in breast cancer.* The Journal of Biological Chemistry, 2012. **287**(17): p. 13877-13888.

[64]   Vladimir, V.C., V.S. Markin, and E.D. Thomas, *The voltage-activated hydrogen ion conductance in rat alveolar epithelial cells is determined by the pH gradient.* The Journal of General Physiology, 1995. **105**(6): p. 861-896.

[65]   Sasaki, M., M. Takagi, and Y. Okamura, *A voltage sensor-domain protein is a voltage-gated proton channel.* Science, 2006. **312**(5773): p. 589-592.

[66] Thomas, E.D. and V.C. Vladimir, *Pharmacology of voltage-gated proton channels.* Current Pharmaceutical Design, 2007. **13**(23): p. 2406-2420.

[67] Rodriguez, J.D., et al., *Identification of a vacuolar proton channel that triggers the bioluminescent flash in dinoflagellates.* PLOS ONE, 2017. **12**(2): p. e0171594.

[68] Chakrapani, S., et al., *The activated state of a sodium channel voltage sensor in a membrane environment.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(12): p. 5435-5440.

[69] Delemotte, L., et al., *Intermediate states of the Kv1.2 voltage sensor from atomistic molecular dynamics simulations.* Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(15): p. 6109-6114.

[70] Chamberlin, A., et al., *Hydrophobic plug functions as a gate in voltage-gated proton channels.* Proceedings of the National Academy of Sciences of the United States of America, 2014. **111**(2): p. E273-E282.

[71] DeCoursey, T.E., *CrossTalk proposal: Proton permeation through HV1 requires transient protonation of a conserved aspartate in the S1 transmembrane helix.* The Journal of Physiology, 2017. **595**(22): p. 6793-6795.

[72] DeCoursey, T.E., et al., *Insights into the structure and function of Hv1 from a meta-analysis of mutation studies.* The Journal of General Physiology, 2016. **148**(2): p. 97-118.

[73] Crooks, G.E., et al., *WebLogo: a sequence logo generator.* Genome Research, 2004. **14**(6): p. 1188-1190.

[74] Callenberg, K.M., et al., *APBSmem: a graphical interface for electrostatic calculations at the membrane.* PLOS ONE, 2010. **5**(9): p. e12722.

[75] Bardhan, J.P. and M.G. Knepley, *Communication: modeling charge-sign asymmetric solvation free energies with nonlinear boundary conditions.* The Journal of Chemical Physics, 2014. **141**(13): p. 131103.

[76] Smart, O.S., et al., *HOLE: a program for the analysis of the pore dimensions of ion channel structural models.* Journal of Molecular Graphics, 1996. **14**(6): p. 354-360.

[77] Tajkhorshid, E., et al., *Control of the selectivity of the aquaporin water channel family by global orientational tuning.* Science, 2002. **296**(5567): p. 525-530.

[78]     Jensen, M.Ø., E. Tajkhorshid, and K. Schulten, *Electrostatic tuning of permeation and selectivity in aquaporin water channels.* Biophysical Journal, 2003. **85**(5): p. 2884-2899.

[79]     Seeber, M., et al., *Wordom: a program for efficient analysis of molecular dynamics simulations.* Bioinformatics, 2007. **23**(19): p. 2625-2627.

[80]     Boonamnaj, P. and P. Sompornpisut, *Insight into the role of the Hv1 C-terminal domain in dimer stabilization.* The Journal of Physical Chemistry B, 2018. **122**(3): p. 1037-1048.

[81]     Gianti, E., et al., *On the role of water density fluctuations in the inhibition of a proton channel.* Proceedings of the National Academy of Sciences of the United States of America, 2016. **113**(52): p. E8359-E8368.

[82]     DeCoursey, T.E., *Voltage and pH sensing by the voltage-gated proton channel, Hv1.* Journal of The Royal Society Interface, 2018. **15**(141).

[83]     Dong, H., et al., *Pore waters regulate ion permeation in a calcium release-activated calcium channel.* Proceedings of the National Academy of Sciences of the United States of America, 2013. **110**(43): p. 17332.

[84]     de Groot, B.L. and H. Grubmüller, *Water permeation across biological membranes: mechanism and dynamics of aquaporin-1 and GlpF.* Science, 2001. **294**(5550): p. 2353.

[85]     Chen, H., et al., *Charge Delocalization in Proton Channels, I: The Aquaporin Channels and Proton Blockage.* Biophysical Journal, 2007. **92**(1): p. 46-60.

[86]     *Web site for residue-residue interaction tables: http://gor.bb.iastate.edu/potential/, courtesy of the research group of R.L. Jernigan.*

[87]     Tanaka, S. and H.A. Scheraga, *Medium- and Long-Range Interaction Parameters between Amino Acids for Predicting Three-Dimensional Structures of Proteins.* Macromolecules, 1976. **9**(6): p. 945-950.

[88]     Ueda, Y., H. Taketomi, and N. GŌ, *Studies on protein folding, unfolding, and fluctuations by computer simulation. II. a. three-dimensional lattice model of lysozyme.* Biopolymers, 1978. **17**(6): p. 1531-1548.

[89]    Miyazawa, S. and R.L. Jernigan, *Estimation of effective interresidue contact energies from protein crystal structures: quasi-chemical approximation.* Macromolecules, 1985. **18**(3): p. 534-552.

[90]    Hinds, D.A. and M. Levitt, *A lattice model for protein structure prediction at low resolution.* Proceedings of the National Academy of Sciences of the United States of America, 1992. **89**(7): p. 2536-2540.

[91]    Maiorov, V.N. and G.M. Grippen, *Contact potential that recognizes the correct folding of globular proteins.* Journal of Molecular Biology, 1992. **227**(3): p. 876-888.

[92]    Bryant, S.H. and C.E. Lawrence, *An empirical energy function for threading protein sequence through the folding motif.* Proteins: Structure, Function, and Bioinformatics, 1993. **16**(1): p. 92-112.

[93]    Fiebig, K.M. and K.A. Dill, *Protein core assembly processes.* The Journal of Chemical Physics, 1993. **98**(4): p. 3475-3487.

[94]    Liwo, A., et al., *Calculation of protein backbone geometry from $\alpha$-carbon coordinates based on peptide-group dipole alignment.* Protein Science, 1993. **2**(10): p. 1697-1714.

[95]    Wolynes, P.G., J.N. Onuchic, and D. Thirumalai, *Navigating the folding routes.* Science, 1995. **267**(5204): p. 1619-1620.

[96]    Godzik, A., *Knowledge-based potentials for protein folding: what can we learn from known protein structures?* Structure, 1996. **4**(4): p. 363-366.

[97]    Miyazawa, S. and R.L. Jernigan, *Residue–residue potentials with a favorable contact pair term and an unfavorable high packing density term, for simulation and threading.* Journal of Molecular Biology, 1996. **256**(3): p. 623-644.

[98]    Gerstman, B. and Y. Garbourg, *Structural information content and Lyapunov exponent calculation in protein unfolding.* Journal of Polymer Science Part B: Polymer Physics, 1998. **36**(15): p. 2761-2769.

[99]    Betancourt, M.R. and D. Thirumalai, *Pair potentials for protein folding: choice of reference states and sensitivity of predicted native states to variations in the*

*interaction schemes.* Protein Science : A Publication of the Protein Society, 1999. **8**(2): p. 361-369.

[100] Zhou, Y. and M. Karplus, *Interpreting the folding kinetics of helical proteins.* Nature, 1999. **401**: p. 400-403.

[101] Tobi, D., et al., *On the design and analysis of protein folding potentials.* Proteins: Structure, Function, and Bioinformatics, 2000. **40**(1): p. 71-85.

[102] Vendruscolo, M., R. Najmanovich, and E. Domany, *Can a pairwise contact potential stabilize native protein folds against decoys obtained by threading?* Proteins: Structure, Function, and Bioinformatics, 2000. **38**(2): p. 134-148.

[103] Bastolla, U., et al., *How to guarantee optimal stability for most representative structures in the protein data bank.* Proteins: Structure, Function, and Bioinformatics, 2001. **44**(2): p. 79-96.

[104] Shen, M.-y. and K.F. Freed, *All-atom fast protein folding simulations: the villin headpiece.* Proteins: Structure, Function, and Bioinformatics, 2002. **49**(4): p. 439-445.

[105] Bagci, Z., et al., *The origin and extent of coarse-grained regularities in protein internal packing.* Proteins: Structure, Function, and Bioinformatics, 2003. **53**(1): p. 56-67.

[106] Banavar, J.R. and A. Maritan, *Colloquium: geometrical approach to protein folding: a tube picture.* Reviews of Modern Physics, 2003. **75**(1): p. 23-34.

[107] Kuhlman, B., et al., *Design of a novel globular protein fold with atomic-level accuracy.* Science, 2003. **302**(5649): p. 1364-1368.

[108] Gillespie, B. and K.W. Plaxco, *Using protein folding rates to test protein folding theories.* Annual Review of Biochemistry, 2004. **73**(1): p. 837-859.

[109] Peng, S., et al., *Discrete molecular dynamics simulations of peptide aggregation.* Physical Review E, 2004. **69**(4): p. 041908.

[110] Herges, T. and W. Wenzel, *In silico folding of a three helix protein and characterization of Its free-energy landscape in an all-atom force field.* Physical Review Letters, 2005. **94**(1): p. 018101.

[111] Pokarowski, P., et al., *Inferring ideal amino acid interaction forms from statistical protein contact potentials.* Proteins, 2005. **59**(1): p. 49-57.

[112]    Sorin, E.J. and V.S. Pande, *Exploring the helix-coil transition via all-atom equilibrium ensemble simulations.* Biophysical Journal, 2005. **88**(4): p. 2472-2493.

[113]    Behringer, H., A. Degenhard, and F. Schmid, *Coarse-grained lattice model for molecular recognition.* Physical Review Letters, 2006. **97**(12): p. 128101.

[114]    Skepö, M., P. Linse, and T. Arnebrant, *Coarse-grained modeling of proline rich protein 1 (PRP-1) in bulk solution and adsorbed to a negatively charged surface.* The Journal of Physical Chemistry B, 2006. **110**(24): p. 12141-12148.

[115]    Skolnick, J., et al., *Derivation and testing of pair potentials for protein folding. When is the quasichemical approximation correct?* Protein Science, 2008. **6**(3): p. 676-688.

[116]    Pandey, R.B. and B.L. Farmer, *Globular structure of a human immunodeficiency virus-1 protease (1DIFA dimer) in an effective solvent medium by a Monte Carlo simulation.* The Journal of Chemical Physics, 2010. **132**(12): p. 125101.

[117]    Sharma, S., B.J. Berne, and S.K. Kumar, *Thermal and structural stability of adsorbed proteins.* Biophysical Journal, 2010. **99**(4): p. 1157-1165.

[118]    Fritsche, M., et al., *Conformational temperature-dependent behavior of a histone H2AX: a coarse-grained Monte Carlo approach via knowledge-based interaction potentials.* PLOS ONE, 2012. **7**(3): p. e32075.

[119]    Freddolino, P.L., et al., *Challenges in protein-folding simulations.* Nature Physics, 2010. **6**: p. 751-758.

[120]    Haliloglu, T. and I. Bahar, *Coarse-grained simulations of conformational dynamics of proteins: application to apomyoglobin.* Proteins: Structure, Function, and Bioinformatics, 1998. **31**(3): p. 271-281.

[121]    van Giessen, A.E. and J.E. Straub, *Monte Carlo simulations of polyalanine using a reduced model and statistics-based interaction potentials.* The Journal of Chemical Physics, 2004. **122**(2): p. 024904.

[122]    Zhou, J., et al., *Coarse-grained peptide modeling using a systematic multiscale approach.* Biophysical Journal, 2007. **92**(12): p. 4289-4303.

[123]    de Jong, D.H., et al., *Improved parameters for the Martini coarse-grained protein force field.* Journal of Chemical Theory and Computation, 2013. **9**(1): p. 687-697.

[124]  Noid, W.G., *Perspective: coarse-grained models for biomolecular systems.* The Journal of Chemical Physics, 2013. **139**(9): p. 090901.

[125]  Kitjaruwankul, S., et al., *Asymmetry in structural response of inner and outer transmembrane segments of CorA protein by a coarse-grain model.* The Journal of Chemical Physics, 2016. **145**(13): p. 135101.

[126]  Smith, S.M.E. and T.E. DeCoursey, *Consequences of dimerization of the voltage-gated proton channel*, in *Progress in Molecular Biology and Translational Science*, J. Giraldo and F. Ciruela, Editors. 2013, Academic Press. p. 335-360.

[127]  DeCoursey, T.E., *Structural revelations of the human proton channel.* Proceedings of the National Academy of Sciences of the United States of America, 2015. **112**(44): p. 13430-13431.

[128]  Musset, B., et al., *NOX5 in human spermatozoa: expression, function, and regulation.* Journal of Biological Chemistry, 2012. **287**(12): p. 9376-9388.

[129]  Henderson, L.M. and J.B. Chappell, *NADPH oxidase of neutrophils.* Biochimica et Biophysica Acta (BBA) - Bioenergetics, 1996. **1273**(2): p. 87-107.

[130]  DeCoursey, T.E., D. Morgan, and V.V. Cherny, *The voltage dependence of NADPH oxidase reveals why phagocytes need proton channels.* Nature, 2003. **422**: p. 531-534.

[131]  Morgan, D., et al., *Voltage-gated proton channels maintain pH in human neutrophils during phagocytosis.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(42): p. 18022.

[132]  Demaurex, N., *Functions of proton channels in phagocytes.* 2012. **1**: p. 3-15.

[133]  Fischer, H., *Function of proton channels in lung epithelia.* Wiley interdisciplinary reviews. Membrane transport and signaling, 2012. **1**(3): p. 247-258.

[134]  Michaels, T.C.T., et al., *Fluctuations in the kinetics of linear protein self-assembly.* Physical Review Letters, 2016. **116**(25): p. 258103.

[135]  Tang, Q.-Y., et al., *Critical fluctuations in the native state of proteins.* Physical Review Letters, 2017. **118**(8): p. 088102.

[136]  K., B., *Monte Carlo and Molecular Dynamics Simulations in Polymer Science.* Monte Carlo and Molecular Dynamics Simulations in Polymer Science, 1995.

[137]    van Dijk, E., et al., *Consistent Treatment of Hydrophobicity in Protein Lattice Models Accounts for Cold Denaturation.* Physical Review Letters, 2016. **116**(7): p. 078101.

[138]    Hissam, R.S., B.L. Farmer, and R.B. Pandey, *Scaffolding of an antimicrobial peptide (KSL) by a scale-down coarse-grained approach.* Physical Chemistry Chemical Physics, 2011. **13**(48): p. 21262-21272.

[139]    Pandey, R.B. and B.L. Farmer, *Aggregation and network formation in self-assembly of protein (H3.1) by a coarse-grained Monte Carlo simulation.* The Journal of Chemical Physics, 2014. **141**(17): p. 175103.

[140]    Sompornpisut, P., B. Roux, and E. Perozo, *Structural refinement of membrane proteins by restrained molecular dynamics and solvent accessibility data.* Biophysical Journal, 2008. **95**(11): p. 5349-5361.

[141]    N., P.P. and D. A., *Uncertainty quantification for Molecular Dynamics, arXiv:1801.02483v1 [physics.comp-ph].* 2018.

[142]    Yoo, J. and A. Aksimentiev, *New tricks for old dogs: improving the accuracy of biomolecular force fields by pair-specific corrections to non-bonded interactions.* Physical Chemistry Chemical Physics, 2018. **20**(13): p. 8432-8449.

# VITA

| | |
|---|---|
| **NAME** | Panisak Boonamnaj |
| **DATE OF BIRTH** | 14 March 1991 |
| **PLACE OF BIRTH** | Bangkok |
| **INSTITUTIONS ATTENDED** | 2009 – 2012 B.Sc. in Chemistry (1st honor), Department of Chemistry, Faculty of Science, Srinakharinwirot University, Bangkok, Thailand |
| | 2013 – 2018 Ph.D. student in Physical Chemistry, Faculty of Science, Chulalongkorn University, Bangkok, Thailand |
| **HOME ADDRESS** | 190 Soi 21, Taksin Rd. Samre Thonburi, Bangkok 10600, Thailand |
| **PUBLICATION** | Boonamnaj, P., and Sompornpisut, P. Insight into the role of the Hv1 C-terminal domain in dimer stabilization. J. Phys. Chem. B 122 (2018): 1037-1048. |
| **AWARD RECEIVED** | 2016 ASEA-Uninet grant for short-term research at University of Vienna. |
| | 2013 100th Anniversary Chulalongkorn University Fund for Doctoral Scholarship. |