Spring 2019

# Computational Studies of Protein-Inhibitor Interactions

Mohammadjavad Mohammadi
*University of New Hampshire, Durham*

# COMPUTATIONAL STUDIES OF PROTEIN-INHIBITOR INTERACTIONS

BY

MOHAMMADJAVAD MOHAMMADI

BS, Chemical Engineering, Amirkabir University of Technology, 2010
MS, Chemical Engineering, University of Southern California, 2013

DISSERTATION

Submitted to the University of New Hampshire
in Partial Fulfillment of
the Requirements for the Degree of

Doctor of Philosophy

in

Chemical Engineering

May 2019

This dissertation has been examined and approved in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Chemical Engineering by:

**Dissertation Director, Dr. Harish Vashisth**
Assistant Professor, Department of Chemical Engineering
University of New Hampshire, Durham, NH

**Dr. Russell Carr**
Professor and Chair, Department of Chemical Engineering
University of New Hampshire, Durham, NH

**Dr. Nivedita Gupta**
Professor, Department of Chemical Engineering
University of New Hampshire, Durham, NH

**Dr. Rick Cote**
Professor and Chair, Department of Molecular, Cellular, and Biomedical Sciences
University of New Hampshire, Durham, NH

**Dr. Richard Neubig**
Professor and Chair, Department of Pharmacology and Toxicology
Michigan State University, East Lansing, MI

on [May 2019].

Original approval signatures are on file with the University of New Hampshire Graduate School.

*This dissertation is dedicated to my family.*

# ACKNOWLEDGMENTS

My first and sincere appreciation goes to my thesis advisor, Dr. Harish Vashisth. I am grateful for his continuous guidance and support in all stages of my research work at UNH. I would also like to thank him for all I have learned from him and for encouraging and helping me to shape my ideas and interests. I also sincerely thank Prof. Russell Carr for continued guidance at various steps of my studies at UNH. I must thank Prof. Richard Neubig for many insightful discussions, and reviewing my work and providing valuable feedback. I also sincerely thank my other committee members Prof. Nivedita Gupta and Prof. Rick Cote for guidance, support, and serving on my dissertation committee. I am grateful to Dr. Luca Maragliano (Istituto Italiano di Tecnologia, Genoa, Italy) for his guidance, insights, and willingness to help in understanding single-sweep methodology. I also thank Vincent Shaw (Michigan State University, East Lansing, MI) for many insightful discussions and cooperation during my thesis work. I also thank my fellow labmates, Dr. Hossein Mohammadiarani and Yong Liu for valuable collaborations.

I also thank my undergraduate advisor, Dr. Cavus Falamaki (Amirkabir University of Technology, Tehran, Iran), who encouraged me to pursue higher studies. Last but not least, I would like to thank my parents for their unconditional love and support, financially and spiritually, throughout my life. Without them I would have never made it to this stage in my career.

# TABLE OF CONTENTS

# NOMENCLATURE

**BC** — Boundary Conditions.

**CV** — Collective Variable.

**BSA** — Buried Surface Area.

**DCC** — Dynamic Cross Correlation.

**DI** — Deuterium Incorporation.

**GDP** — Guanosine Diphosphate.

**GPCR** — G-protein-coupled Receptor.

**GTP** — Guanosine Triphosphate.

**HDX** — Hydrogen-deuterium Exchange.

**MD** — Molecular Dynamics.

**MM/GBSA** — Molecular Mechanics Generalized Born Surface Area.

**MM/PBSA** — Molecular Mechanics Poisson-Boltzmann Surface Area.

**TAMD** — Temperature Accelerated Molecular Dynamics.

**PME** — Particle Mesh Ewald.

**PMF** — Potential of Mean Force.

**RCSB** — Research Collaboratory for Structural Bioinformatics.

**RGS** — Regulator of G-protein Signaling.

**PDE4** — Phosphodiesterase 4.

**RMSD** — Root Mean Squared Deviation.

**RMSF** — Root Mean Squared Fluctuation.

**SASA** — Solvent Accessible Surface Area.

**TDZD** — Thiadiazolidinone.

# LIST OF TABLES

**Table** **Page**

# LIST OF FIGURES

**Figure** **Page**

# Abstract


## Computational Studies of Protein-Inhibitor Interactions

In this thesis work, I conducted large-scale molecular dynamics (MD) simulation studies of interactions of enzymes and signaling proteins with inhibitory ligands. Specifically, I have studied three classes of proteins: the first part of my thesis reports studies on the hydrogen-producing [FeFe]-hydrogenase enzyme, the second part reports on studies of regulatory proteins from the G-protein coupled receptor (GPCR) family, and the third part reports on studies of the phosphodiesterase (PDE) enzyme family.

In the first part, I studied the problem of [FeFe]-hydrogenase sensitivity to the presence of inhibitory gases oxygen ($O_2$) and carbon monoxide (CO) that cause irreversible damage to the active site of this enzyme. Therefore, a detailed knowledge of the diffusion pathways of these inhibitory gases is necessary to develop strategies for designing novel enzymes that are tolerant to these gases. Specifically, I studied the diffusion pathways of $O_2$ and CO in the CpI [FeFe]-hydrogenase from *Clostridium pasteurianum*. I used several enhanced sampling and free-energy simulation methods to reconstruct a three-dimensional free-energy surface for diffusion of each gas which revealed free-energy minima forming an interconnected network of pathways. I discovered multiple pathways of minimal free-energy as diffusion portals for $O_2$ and CO, and observed that the global minimum in the free-energy surface is located in the vicinity of the active site metal cluster, the H-cluster. Among potential residues that I propose as candidates for future mutagenesis studies to increase the tolerance of this enzyme to both inhibitory gases, 11 residues are shared between $O_2$ and CO. I hypothesize that these shared candidate residues are potentially useful for designing new variants of this enzyme that are tolerant to both inhibitory gases.

In the second part, I have studied the interplay of protein conformational dynamics and effects of small-molecule inhibitors in a class of signaling proteins, known as the Regulators of G protein signaling (RGS) proteins, that negatively modulate signaling in GPCRs. Recently discovered thiadiazolidinone(TDZD) compounds that target cysteine residues have shown different levels of specificities and potencies for several known RGS proteins, thereby suggesting intrinsic differences in dynamics of these proteins upon binding of these compounds. I characterized the effect of binding of several small-molecule inhibitors on perturbations and dynamical motions in RGS4.

Specifically, I studied two conformational models of RGS4 in which a buried cysteine residue is solvent-exposed due to side-chain motions or due to flexibility in neighboring helices. I found that TDZD compounds with aromatic functional groups perturb the RGS4 structure more than compounds with aliphatic functional groups. Moreover, small-molecules with aromatic functional groups but lacking sulfur atoms only transiently reside within the protein and spontaneously dissociate to the solvent. I further probed the salt-bridges forming across isoforms of RGS proteins, resulting in a hypothesis that differences in salt-bridges between a pair of helices in RGS proteins are responsible for differences in flexibility and potency among isoforms.

In the final part of my thesis, I evaluated differences in binding interactions of phosphodiesterase 4 (PDE4) inhibitors within the PDE4 catalytic domain. From residues within 5 Å of the ligand binding site, five residues revealed significant differences in non-bonded interaction energies that could account for the differential binding affinities of inhibitors. I found one site (Phe506 in human PDE4; Tyr253 in the *C. elegans* PDE4 catalytic domain) that alters the binding conformation of roflumilast and zardaverine (human PDE 4 inhibitors) into a less energetically favorable state. These results support the feasibility of designing the next-generation of anthelmintics/nematicides that could selectively bind to nematode PDEs. Overall, my thesis has resulted in enhancement of detailed mechanistic insights into several protein and inhibitory ligand interactions that are potentially useful in the developement of novel inhibitors targeting protein/protein and protein/ligand interactions.

# CHAPTER 1

# INTRODUCTION

## 1.1 Protein-Protein Interactions

The past decade has seen many advances in molecular understanding of regulatory systems critical in physiology and pathology, and protein-protein interactions (PPIs) (Fig. 1.1) signify a key objective space for pharmacological intervention. Recent genomics advances found that there are up to 10000 disease related proteins [18], and yet only approximately 400 proteins have been explored for therapeutic development. The majority of these proteins are the G protein-coupled receptors (GPCRs), enzymes, ion channels, and nuclear hormone receptors [19].

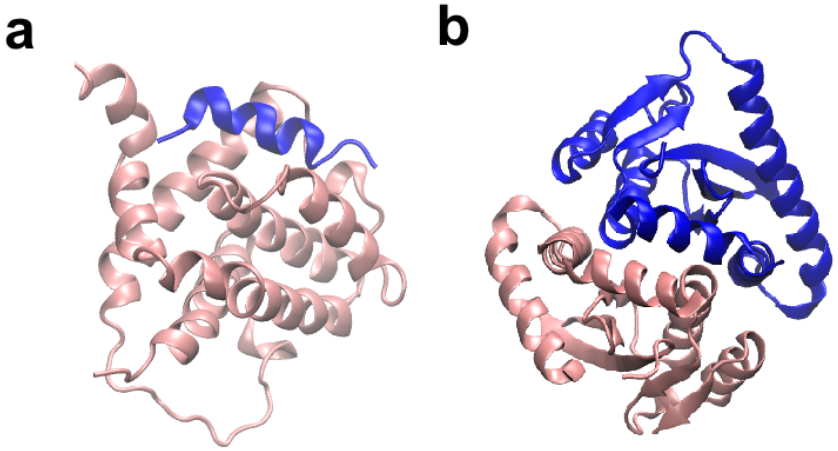



Figure 1.1: The protein-protein interaction (a) between a protein and a peptide (Protein Data Bank; PDB:1BXL), and (b) between two proteins (Protein Data Bank; PDB:2B4J).

Most of these disease related proteins exert their functions via interactions with other proteins (Fig. 1.1), and recent developments demonstrate that PPIs can be modulated with small-molecule drugs and peptides [19]. To enhance development of novel therapeutic compounds, a molecular understanding of the structure and dynamics of interacting proteins is essential.

## 1.2 Protein-Ligand Interactions

Protein functions are usually coupled to binding of small molecules such as substrates or ligands. These ligands often interact with endogenous ligand-binding sites of their target proteins. The modulation of a target protein by a ligand has been extensively explored for a diverse range of objectives in drug discovery [19]. The interaction of a ligand with a protein is a complex biophysical event, as the surface of a protein constitutes a multitude of polar and hydrophobic sites that represent challenges to understanding the interaction mechanism of a ligand with the target protein.

## 1.3 Classification of ligands

Ligands modulate proteins and PPIs via three broad mechanisms ([18]): (1) orthosteric inhibition, (2) allosteric regulation, and (3) interfacial binding. The orthosteric ligands bind target proteins at the overlapping areas in PPIs, directly inhibiting the formation of macromolecular complexes. The allosteric inhibitors bind to their target protein at a site distinct from the macromolecular interface. The interfacial inhibitors bind to a pocket at the internal space or macromolecular interface often transiently formed, and cause a nonproductive conformation of the protein/protein complex. In this thesis, I studied three categories of ligands, small-molecule interfacial inhibitors, orthosteric inhibitors and allosteric inhibitors.

### 1.3.1 Small-Molecules Inhibitors

Small-molecule inhibitors have been extensively explored in modulation of many diseases related to proteins [20]. These inhibitors could be as simple as small gas molecules $O_2$ and CO, that are known to inhibit enzymes with metallic clusters. Most proteins exert their functions through interactions with other proteins, and they often lack obvious druggable pockets for small molecules.

Although PPIs are essential for many aspects of cellular functions and provide vast potential for drug development, targeting PPIs with small molecules is challenging for two main reasons. Firstly, the interacting areas between proteins typically contain numerous polar and hydrophobic interactions distributed across a large interface. Unless the interaction hotspot can be identified, a small molecule with a reduced contact area with a protein has difficulty in achieving effective binding due to lower interactions. Secondly, the protein-protein interfaces are typically featureless. The lack of pockets limits the site of binding to only limited areas of a small molecule; such binding is much less effective than binding with multiple areas in a deeper pocket. Despite these difficulties, there has been a growing interest in identifying small-molecule inhibitors against PPIs [21].

### 1.3.2 Allosteric Inhibitors

These small-molecule ligands bind to target proteins at sites distinct from the macromolecular interface. Ligand binding induces changes in either the static conformation or the dynamic properties of the target protein and hinders the macromolecular interaction in an allosteric (distant) manner. In this thesis, I studied thiadiazolidinone (TDZD) compounds that have been successfully exploited to modulate the activity of their target proteins allosterically.

## 1.4 Background of Major Proteins Studied

In this thesis, I have studied three essential classes of proteins that are affected by ligands: (i) [FeFe]-hydrogenase, (ii) Regulators of G-protein signaling (RGS) proteins, and (iii) Phosphodiesterase 4 (PDE4) enzyme. The hydrogen-producing [FeFe]-hydrogenase is a promising source of renewable $H_2$. RGS proteins are clinically significant drug targets. PDE4 enzymes are drug targets for the development of nematicides. In this thesis, I explored these three classes of proteins and their interactions with inhibitory compounds.

### 1.4.1 FeFe-hydrogenase

Biological $H_2$ production from [FeFe]-hydrogenase, found in photosynthetic microorganisms, is an attractive alternative source of renewable energy in future [22, 23, 24]. Among such microorganisms, green algae and cyanobacteria are potential sources of $H_2$ production because they have a maximum solar to hydrogen (STH) conversion efficiency [11, 25, 26]. The hydrogenase family is broadly classified into [NiFe]-hydrogenase and [FeFe]-hydrogenase, based upon the metal-cluster present in the active-site [27, 28]. In comparison to [NiFe]-hydrogenase, [FeFe]-hydrogenase is primarily responsible for $H_2$ production, but is significantly inhibited by $O_2$ [29, 30]. In this work, I have focused on CpI [FeFe]-hydrogenase from *Clostridium pasteurianum*. Fig. 1.2 shows the crystal structure (PDB code 1FEH) of the CpI [FeFe]-hydrogenase containing 574 amino-acid residues, and six metal clusters: four $[Fe_4S_4]$ metal clusters, three of which are ligated to four cysteines, distal metal cluster that is ligated to three cysteines and one histidine, and one surface cluster $[Fe_2S_2]$ that is ligated to four cysteines.

The central metal cluster, which is known as the H-cluster, consists of [2Fe]-subclusters with ligation to carbonyl and cyanide groups [31]. Hydrogen production in this enzyme takes place at the H-cluster buried inside the protein [32]. In [FeFe]-hydrogenase, the H-cluster is a complex structure consisting of $Fe_2(CO)_3(CN)_2$(dithiomethylamine), which is covalently

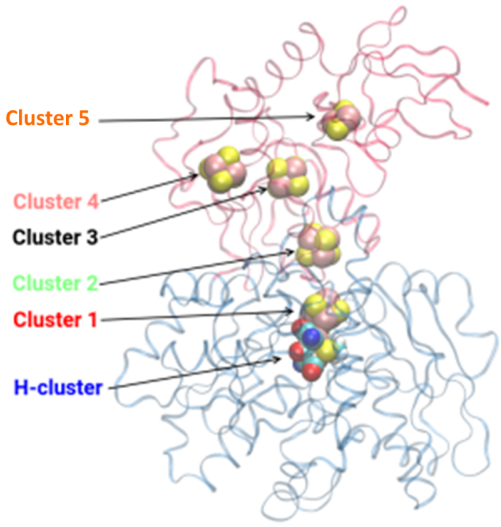



Figure 1.2: Locations of metal clusters in [FeFe]-hydrogenase are highlighted. The protein-backbone is shown as ribbons, and metallic clusters are shown as space-filling. The active-site domain of the protein is shown in blue ribbons, and the non-active-site domain is shown in red ribbons.

bound to an accessory metal cluster [$Fe_4S_4$] [33, 34, 35]. The side of the protein that harbors the H-cluster (blue ribbons in Fig. 1.2), is designated as the active-site domain, and the side of the protein that harbors other metal clusters, is designated as the non-active-site domain (red ribbons in Fig. 1.2).

A major obstacle to achieving the goal of sustained $H_2$ production is the inactivation of this hydrogenase by oxygen [25, 26]. Although both NiFe and FeFe hydrogenases are sensitive to inactivation by $O_2$, [NiFe]-hydrogenases capable of naturally resisting the inactivation by $O_2$ with different metallic clusters were recently found [36, 37]. Recently, success has also been achieved in improving the $O_2$ sensitivity of [NiFe]-hydrogenase [38, 39, 40], but no similar progress has been made on $H_2$-producing [FeFe]-hydrogenase. It is thought that the inactivation of [FeFe]-hydrogenase by $O_2$ consists of diffusion of $O_2$ from the solvent into the enzyme interior, getting to the H-cluster buried in the core of the protein-matrix, and subsequently binding to the H-cluster [5, 41, 42].

Although the exact mechanism of inactivation of the H-cluster after binding of $O_2$ is not well understood, X-ray absorption measurements show that binding of $O_2$ to the distal Fe ($Fe_d$) in the H-cluster can subsequently damage the [$Fe_4S_4$] accessory cluster next to the H-cluster [43, 44, 45, 46, 47]. It has been proposed that possible formation of reactive oxygen upon $O_2$ binding to $Fe_d$ of the H-cluster and its diffusion toward the [$Fe_4S_4$] accessory-cluster damages this cluster (cluster 1 in Fig. 1.2) [43, 48]. However, Swanson et al. [49] describe a rather complex mechanism for degradation of the H-cluster, including the formation of a reversible state before its full degradation. Aside from the H-cluster, Liebgott et al. [50] have shown that by altering the structure of [NiFe]-hydrogenase using a site-directed mutagenesis approach, the diffusion rate of $O_2$ can be slowed by orders of magnitude, thereby slowing the inactivation rate. This is accomplished by replacing the amino acids along known gas diffusion channels with bulky amino acids, likely blocking the possible $O_2$ diffusion path within the protein [50]. Similarly, Nienhaus et al. [51, 52] show that CO diffusion in myoglobin is altered by mutations of amino acids along CO-diffusion paths.

The presence of gas channels in [FeFe]-hydrogenase for $O_2$ and $H_2$ has been investigated both computationally and experimentally. The existence of two main gas channels in the active-site domain of the [FeFe]-hydrogenase has been proposed [5, 11, 53, 25]. Kubas et al. [10] investigated the rate of $O_2$ diffusion along these two main channels [10]. A major limitation of these investigations has been a narrow understanding of $O_2$ diffusion pathways within the other parts of [FeFe]-hydrogenase, and interactions of these diffusion channels with each other. Moreover, computational studies for gaining insights into ligand diffusion mechanisms, such as $O_2$ diffusion in [FeFe]-hydrogenase, is still an active area of research [54]. In fact, Bingham et al. [55] have reported that the mutations in an area away from the two main gas channels affect $O_2$ diffusion noticeably, indicating the potential presence of other gas channels in [FeFe]-hydrogenase or an allosteric communication between $O_2$ binding sites. Although, both Lautier et al.'s [6] and Ghirardi et al.'s [11] mutagenesis investigations suggest that $O_2$ is possibly using the two main channels, the likelihood of the presence of other pathways

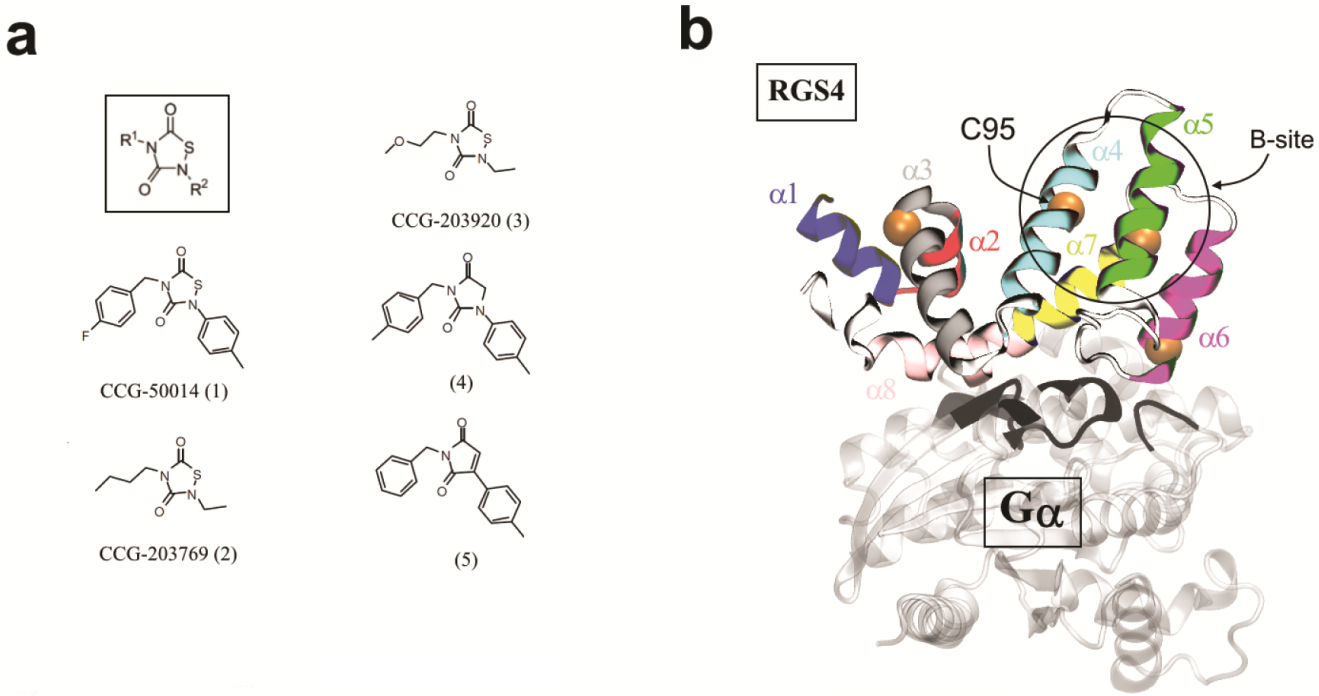
for $O_2$ within the protein was highlighted [6]. Indeed, a network of interconnected diffusion pathways have been found for ligand diffusion in various proteins [56, 57, 58], clearly showing the presence of key residues or reaction sites along the predicted pathways, and elucidating the interconnected nature of diffusion pathways of ligands in proteins. Therefore, a better understanding of gas diffusion in [FeFe]-hydrogenase is needed to develop strategies for its functional modification. Thus, a detailed mapping of diffusion pathways of CO and O2 in the FeFe-hydrogenase is not only needed for resolving these questions, but also for developing approaches to enhance the tolerance of the FeFe-hydrogenase to inhibitory gases.

### 1.4.2 Regulators of G-protein signaling (RGS)

GPCRs are membrane proteins of profound clinical relevance [59, 60, 61, 62, 63], as they mediate key roles in many cellular reactions to neurotransmitters and extracellular ligands [64]. Signaling by GPCRs is negatively modulated by a family of proteins known as RGS proteins that serve as a critical node in controlling various cellular responses [65, 66, 67, 68]. The mechanism of action of RGS proteins is to bind to activated (GTP-bound) G protein $\alpha$-subunits (G$\alpha$) and accelerate the rate of GTP-hydrolysis, resulting in conversion of GTP to GDP and deactivation of G$\alpha$-subunits. By this mechanism, RGS proteins rapidly dampen signaling by GPCRs, and therefore inhibitor candidates targeting RGS/G$\alpha$ protein-protein interaction are potentially useful to enhance signaling by GPCRs [69]. Inhibiting protein-protein interactions, such as the one between an RGS protein and a G$\alpha$ subunit, is particularly challenging [70, 71, 72, 73], because of the lack of suitable binding pockets for small-molecules at the RGS/G$\alpha$ protein-protein interface. Although progress has been made recently in the development of small molecules as covalent allosteric protein modifiers [74, 75, 76, 77], application of small-molecule ligands for the inhibition of RGS proteins and protein-protein interactions is an active area of research [78, 65, 79, 80, 81, 82, 70, 83, 84, 85, 4].

Specifically, high-throughput screens have revealed several classes of small molecule inhibitors acting through covalent modification of cysteine thiols on the RGS4 protein [85, 84, 86, 87], which is a well-characterized member of the R4 subfamily [66]. Among them, the thiadiazolidinone (TDZD) inhibitor CCG-50014 and its congeners (Fig. 1.3a) have shown high potency (measure of the concentrations of a drug at which it is effective) and specificity (the degree to which the effects of a drug are due to the one pharmacological action) [4]. CCG-203769 is one of the congeners of CCG-50014 with two aliphatic side-chains, but with substantially improved solubility [4]. RGS4 is highly expressed in the striatum [88, 89], regulating synaptic plasticity in response to dopamine signaling [90, 91], and has been associated with Parkinson's disease [92]. The effectiveness of CCG-203769 on RGS4 has been demonstrated through in-vivo studies, by reducing bradykinesia in a raclopride model of certain Parkinson's-like motor deficits in mice [93]. Furthermore, RGS4 has also been implicated in cancer [94, 95, 96], heart rate control in the sinoatrial node [97], suppression of the anticonvulsant action of adenosine [98], and opioid receptor signaling [99].

Structurally, RGS proteins are defined by a conserved ∼120 residue-long box-domain with nine $\alpha$-helices [100]. The crystal structure of RGS4 bound to $G\alpha_{i1}$ (PDB code 1AGR) is shown in Fig. 1.3b [101], highlighting the RGS domain with each helix distinctly colored. The crystal structure of RGS4 reveals the location of four cysteine residues (highlighted by orange spheres in Fig. 1.3b), three of which are solvent-exposed and one cysteine residue (labeled C95 in Fig. 1.3b) is buried. The C95-residue is also a conserved residue among many RGS proteins, thereby making it an attractive target for covalent modification. Although other cysteines in RGS4 play an important role in potent inhibition of RGS4 by covalent inhibitors [84], these residues are not conserved among all inhibited RGS proteins. Therefore, an understanding of the actions of TDZDs at C95 will be broadly applicable to other RGS proteins. Importantly, TDZD compounds can allosterically inhibit the RGS/$G\alpha$ protein-protein interaction via the conserved and buried cysteine residue C95 [84, 85, 102]. Moreover, the activity of RGS4 is regulated allosterically by endogenous lipid modulators




c

```
RGS4      1   VSQEEVKKWAESLENLIN-EAGLAAFKAFLKSEYSEENIDFWISCEEYKKIKSPSKLSPK  59
RGS8      1   -STEEATRWADSFDVLLSUKYGVAAFRAFLKTEFSEENLEFWLACEEFKKTRSTAKLVSK  59
RGS19     1   PSPEEVQSWAQSFDKLMUSPAGRSVFRAFLRTEYSEENMLFWLACEELKAEANQUVVDEK  60
                * **.  **:*:: *:    * :.*:***::*:****: **::*** *   .   : *

RGS4     60   AKKIYNEFISVQATKEVNLDSATREETSRNMLEPTITAFDEAQKKIFNLMEKDSYRRFLK  119
RGS8     60   AURIFEEEFVDVQAPREVNIDFQTREATRKNLQEPSLTAFDQAQGKVUSLMEKDSYPRFLR  119
RGS19    61   ARLIYEDYVSILSPKEVSLDSRVREGINKKMQEPSAUTFDDAQLQIYTLMURDSYPRFLS  120
               *  *::::::.: : :**.:*  .**   ::: **:  :**:** :: .** :*** ***

RGS4    120   SRFYLDLT----                                                  127
RGS8    120   SKMYLDLLSQSQ                                                  131
RGS19   121   SPTYRALL----                                                  128
               *  *  *
```

**Figure 1.3.** (a) Small-molecule structure with $R^1$ and $R^2$ functional groups (inset in box) along with the chemical structures of all small-molecules studied. [4] (b) A cartoon representation of the RGS4-G$\alpha_{i1}$ complex (PDB code 1AGR) is shown. Each of the 9 $\alpha$-helices of RGS4 is colored and labeled, and the location of four cysteine residues are shown by their $C_\alpha$-atoms as orange spheres with the C95 residue labeled. The G$\alpha_{i1}$-subunit is shown in transparent white ribbons, and the loops of G$\alpha_{i1}$ in the proximity of RGS4 are highlighted in black ribbons. (c) Sequence alignment of RGS4, RGS8, and RGS19.

at a site far from the G$\alpha$ interaction interface but near the C95 residue, also known as the "B-site" (the region marked by a circle in Fig. 1.3b) [65]. This allosteric site is also a promising drug target as many physiological regulators of RGS4 directly interact with the B-site [103, 104].

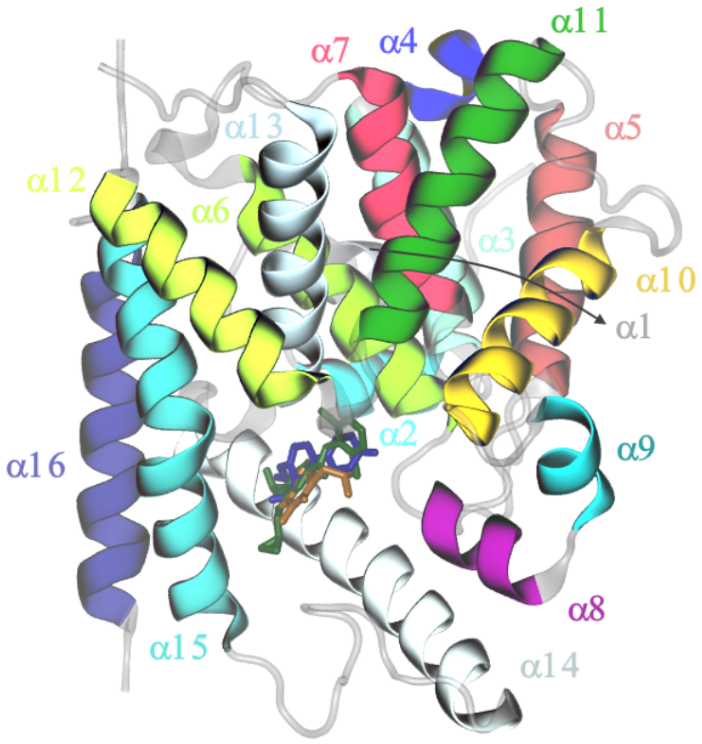



**Figure 1.4.** The Structure of Phosphodiesterase 4 (PDE4), and studied inhibitors (roflumilast, zardaverine, and IBMX are shown in green, yellow, and blue, respectively) overlayed in the catalytic domain of PDE4. The helices of PDE4 are labeled and colored.

### 1.4.3 Phosphodiesterase 4 (PDE4)

The efficacy of and resistance to anthelmintic/nematicidal compounds for controlling parasitic nematodes is a growing concern in the fields of medicine, veterinary medicine, and agriculture [105, 106]. I have studied the human and nematode PDE4 enzyme for targeting

nematode PDEs for the development of nematicides. The class I PDE superfamily in vertebrates consists of eleven PDE families that share a conserved signature motif in the catalytic domain [107]. The crystal structures of the catalytic domains of almost all the PDE families have been solved, providing atomic-level details on the enzymatic and pharmacological properties of enzymes in this superfamily [108].

The catalytic domain is made up of ∼330 amino acids whose secondary structure consists of 16 $\alpha$-helices (Fig. 1.4). These $\alpha$-helices create three subdomains [109] which form a deep catalytic pocket at their center. The active site is composed of two sub-pockets, which bind two divalent metal ions and the substrate, respectively [108]. Zinc and magnesium ions are stabilized by conserved His and Asp residues in the metal binding pocket [108]. For example, the structure of the human PDE4 catalytic domain is stabilized by interactions with divalent cations and Asp and His residues of the metal binding pocket as well as with conserved Gln and Phe residues in the hydrophobic pocket. The invariant Gln residue of PDEs have been shown to be critical for substrate and inhibitor binding [108].

## 1.5 Thesis Outline

In Chapter 2, I introduce details of all computational methods and software packages used in my thesis work. These include the basics of molecular dynamics (MD) simulations, and other related computational techniques. Chapter 3 describes studies of inhibitory gases $O_2$ and CO in FeFe-hydrogenase. Chapter 4 contains the results of studies of allosteric inhibitors for the RGS4 protein. Chapter 5 contains the results of studies on coupling of protein dynamics and salt-bridging interactions among RGS proteins. Chapter 6 provides structural studies of inhibitors on human and nematode PDE4. Each chapter begins with a brief introduction describing the background information. The introduction of each chapter is followed by results and discussion. The usage of "we" in this thesis refers to work done in collaboration

with other researchers and their contributions have been acknowledged throughout the thesis. Appendices A through D contain samples of computer scripts and codes used in this thesis, appendix E through H provides supporting information for chapters 3 to 6, respectively, and appendix I highlights a cover image related to the work described in chapter 4. My brief curriculum vitae is included in Appendix J.

# CHAPTER 2

# METHODS

## 2.1 Introduction

In this chapter, I describe specific methods in detail including classical MD and enhanced sampling MD techniques. The software used are: VMD (Visual Molecular Dynamics) [110], NAMD (NAnoscale Molecular Dynamics) [111], MD-TASK [112], ICM [113], and CHARMMGUI (CHARMM is an acronym for Chemistry at HARvard Molecular Mechanics) [114]. The analysis scripts were implemented in MATLAB (stands for MATrix LABoratory), TCL (Tool Command Language), and Linux bash scripting. The computations were performed on local supercomputing resources at UNH (Premise and Trillian), and Comet (San Diego Supercomputer Center).

## 2.2 Molecular Dynamics Simulation

Statistical mechanics bridges macroscopic observations with microscopic details using statistical theories combined with classical physics concepts rooted in thermodynamics and Newton's laws [115, 116, 117]. In this context, an ensemble is the group of all microscopic states that correspond to one macroscopic state. In other words, macroscopic properties are statistical averages over multiple microstates of the system, also known as ensemble averages. In statistical mechanics, different kinds of ensembles exist which are characterized by fixed values of thermodynamic variables such as total number of particles $N$, volume $V$, temperature $T$, total energy $E$, or chemical potential $\mu$ [118]. In my thesis, all simulations

were carried out either in the NVT ensemble or in the NPT ensemble. These two ensembles are more commonly compared to experimental settings.

In my thesis work, I used MD simulations as a main tool to study complex interactions between biomolecules and small-molecule inhibitors. Given the expensive nature of quantum mechanical calculations, MD simulations provide a more practical way to study complex biomolecules while retaining classical molecular models based on quantum calculations [119, 120, 121].

MD simulations are rooted in numerical techniques for solving equations of motion for a large number of atoms. The basics of underlying equations of motion originate from classical physics. The source of a force calculation in classical physics is a potential energy function that plays a key role in describing the behaviour of a many-body system. The potential functions account for both bonded and non-bonded interactions. The typical form of the potential energy function is given by the following equation [122]:

$$\begin{aligned} U = & \sum_{bonds} k_b(b-b_0)^2 + \sum_{angles} k_\theta(\theta-\theta_0)^2 + \sum_{dihedrals} k_\phi(1+cos(n\phi-\delta)) \\ & + \sum_{impropers} k_\omega(\omega-\omega_0)^2 + \sum_{Urey-Bradley} k_u(u-u_0)^2 \\ & + \sum_i \sum_{j>i} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \sum_i \sum_{j>i} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \end{aligned} \tag{2.1}$$

The first term in the energy function accounts for bond stretches where $k_b$ is the bond force constant and $(b-b_0)$ is the net displacement from equilibrium. The second term in the equation accounts for angles where $k_\theta$ is the angle force-constant and $(\theta-\theta_0)$ is the angular deviation from equilibrium for three bonded atoms. The third term is for dihedrals (also known as torsion angles) where $k_\phi$ is the dihedral force-constant, $n$ is the multiplicity of the function, $\phi$ is the dihedral angle and $\delta$ is the phase shift. The fourth term accounts for impropers, that is out of plane bending, where $k_\omega$ is the force constant and $(\omega-\omega_0)$ is the out of plane angle deviation. The Urey-Bradley component (cross-term accounting for

angle bending using 1,3 nonbonded interactions) comprises the fifth term, where $k_u$ is the respective force constant and $u - u_0$ is the distance between the 1,3 atoms in the harmonic potential.

Nonbonded interactions between pairs of atoms $(i, j)$ are represented by the last two terms. The electrostatic interactions are represented by a Coulombic potential where $q_i$ represents the charge of an atom, $\epsilon_0$ is the permittivity of free space, and $r_{ij}$ is the distance between the centers of mass of two atoms. The van der Waals forces are represented by a Lennard-Jones potential function. In this equation, $\epsilon_{ij}$ represents the energy well depth, $\sigma_{ij}$ is the hard sphere radius [122]. There are no special terms to represent hydrogen bonding, as they are treated as a sum of van der Waals and electrostatic interactions. For a potential function used in an MD simulation, the above parameters are computed based on quantum-mechanical calculations and optimized against experimental data [111].

In this thesis, I have used the CHARMM force-field with the CMAP correction [122, 123] for all the MD simulations. CHARMM is the most common and a widely used force-field for proteins, peptides, lipids, and small molecule ligands [124]. All small-molecules used in this thesis and their force-fields were parameterized using the Multipurpose Atom-Typer for CHARMM (MATCH) tool [125].

### 2.2.1 Components of an MD Simulation

#### 2.2.1.1 Langevin Equation

To improve the dynamical stability of the system in an MD simulation, the equation of motion is modified to a partially stochastic Langevin equation, which is given by:

$$\begin{aligned} M\dot{v} &= F(r) - \gamma v + \sqrt{\frac{2\gamma k_B T}{M}} R(t) \\ F(r) &= -\nabla U(r) \end{aligned} \tag{2.2}$$

where $M$ is the mass of a particle, $v = \dot{r}$ is the velocity, $\dot{v}$ is the acceleration, $F$ is the force, $U$ is the potential energy, $r$ is the position vector, $\gamma$ is the friction coefficient, $k_B$ is the Boltzmann constant, $T$ is the temperature, and $R(t)$ is a univariate Gaussian random process. The typical values of friction coefficient are 5 $ps^{-1}$ or 10 $ps^{-1}$. The second term is for dissipative forces, and the last term adds fluctuations for stochasticity which crudely mimics molecular collisions and viscosity in the realistic cellular environment.

#### 2.2.1.2 Numerical Integration

For evolving the system in time, the equation of motion for each atom is numerically integrated. The complexity and stochastic nature of biological systems work against the convergence and stability of an integration algorithm.

The time-step ($\Delta$t$\sim$1-2 fs) also increases the cost of numerical integration significantly. The numerical integration algorithm also needs to maintain the accuracy of the system properties, e.g., temperature and pressure, while being efficient even with the existence of fluctuations in the system [111]. The Brünger-Brooks-Karplus (BBK) method [126], the extension of velocity-Verlet algorithm [121], is the common scheme used in MD simulation packages. The BBK method obtains the position and velocity of each particle through the following recurrence relation:

$$r_{n+1} = r_n + \frac{1-\gamma\Delta t/2}{1+\gamma\Delta t/2}(r_n - r_{n-1}) + \frac{1}{1+\gamma\Delta t/2}\Delta t^2[M^{-1}F(r_n) + \sqrt{\frac{2\gamma k_B T}{\Delta M}}Z_n] \qquad (2.3)$$

where $r_n$ is the position vector, $\gamma$ is the friction coefficient, $\Delta$t is the time-step, $M$ is the mass of a particle, $F$ is the force, $k_B$ is the Boltzmann constant, $T$ is the temperature, and $Z_n$ is a group of Gaussian random variables with a mean of zero and a variance of one. In the BBK method, only one random variable is needed for each degree of freedom. This method has a global error proportional to $\Delta t^2$ [127].

#### 2.2.1.3 Choice of integration time-step $\Delta$t

The time step is an important parameter in an MD simulation for numerical integration of differential equations for progress of the system. It is important because it determines the extent of accuracy and convergence in MD simulations. On one hand, larger values of $\Delta$t will lead to instabilities in the integration scheme, but on the other hand, smaller values of $\Delta$t will lead to limited exploration of the phase space and will increase the computational cost. The choice of time-step needs to be adjusted to satisfy the numerical stability of simulation, and convergence of the numerical integration scheme [128]. A suitable value for $\Delta$t in MD simulations of proteins for which the bonds to hydrogen atoms are fixed is 2 fs, and if the bonds to hydrogen atoms are not fixed, the suitable value for $\Delta$t is 1 fs. The implementation of a multiple-time-stepping algorithm could double the computational efficiency [111, 128]. Multiple-time-stepping algorithms compute slow-varying-forces such as long range electrostatics forces less frequently than fast-varying-forces such as Lennard-Jones and short range electrostatic [128]. The limitation of this algorithm is in conserving energy within a simulation system, which makes this scheme inapplicable to the NVE ensemble. In this thesis, I have not used the NVE ensemble and used only the NPT or NVT ensembles in MD simulations.

#### 2.2.1.4 Initial conditions, minimization, temperature control, and pressure control

In an MD simulation, the initial coordinates of each atom in the system must be specified. As in a finite time-scale MD simulation, trajectories are highly dependent on initial configurations. As expected, the initial configuration of proteins and other biomolecules should be near a stable state of the system. The main repository for the initial configurations of different proteins used in this thesis was the Protein Data Bank (PDB)[1], which is

[1]http://www.rcsb.org/pdb/home/home.do

the data repository for 3D structures of biomolecules determined by X-ray crystallography, Nuclear Magnetic Resonace (NMR), and other approaches. In addition to initial coordinates of the system, for solving the differential equations of motion, the initial values of all velocities should also be assigned. In an MD simulation, initial velocities to each atom from a Maxwell-Boltzmann distribution given by:

$$P(v) = \sqrt{(\frac{m}{2\pi k_B T})^3} 4\pi v^2 e^{-\frac{mv^2}{2kT}} \tag{2.4}$$

where $v$ is the velocity, $m$ is the particle mass and $k_B T$ is the product of Boltzmann's constant and the temperature. Based on the ergodic hypothesis in MD runs.

In MD simulations, boundary conditions of physical systems could be defined in three possible ways: $(i)$ vacuum, $(ii)$ a reflecting wall, and $(iii)$ periodic boundary conditions (PBC). The simplest case for BC is vacuum, but the global properties computed will not be representative of the condensed phase [129]. In the case of reflecting wall BC, for keeping the atomic particles inside the simulation box, solid boundaries with a potential are applied [130]. In PBC, a common practice in MD simulations and which is used in this thesis, particles are isolated in a simulation box that is infinitly replicated by periodic translations in all dimensions [129]. In two-dimensions, each unit cell has eight nearest neighbors, while in three-dimensions each unit cell would have 26 nearest neighbors. An atom or a particle can leave the simulation box from one side, and consequently, it is replaced by an image of the particle which enters from the opposite side of the unit cell, in doing so conserving the number of particles in a simulation box, which is requirment of the NVT and NPT ensembles used in this thesis.

Prior to an MD simulation, the initial configuration of the system obtained from PDB may need preparation to add often missing hydrogen atoms or missing atoms in residues. In most cases, the preparation of the system could include incorporating mutations, solvation,

and ionization of the system. Upon preparation of the desired system for an MD simulation, by using steepest descent or conjugate gradient schemes [131] the system arrangement in space energy-minimized. In this thesis, for all pre-simulation systems, I used the conjugate gradient method for minimization.

The temperature control for an MD simulation can be applied by a Langevin thermostat where additional damping and random forces are introduced in the system. In this method, the temperature of the system is maintained at the chosen value by frequently adjusting the momenta of all atoms in the system. Other techniques used in MD schemes are the Andersen thermostat [132] and the Berendsen thermostat. For pressure control, the Nosé-Hoover barostat algorithms were used [133, 134, 135].

## 2.3 Software Packages for Biomolecular Simulations

### 2.3.1 MD Simulation

In my thesis, I have used NAMD (versions 2.11-2.12) as the primary software tool to carry out MD simulations. NAMD is a parallel MD simulation code designed for large-scale simulations of biomolecular systems [111]. NAMD is capable of effectively using hundreds of processors on computational nodes in supercomputing clusters, and can be further deployed on GPU architectures with a significant reduction in computation time. MD simulations of very large systems of up to ∼64 million atoms are possible with NAMD [136]. The NAMD package contains several efficient techniques to accurately control temperature and pressure, and numerically integrating Newton's equations of motion. NAMD is an open source code package that is available[2] to academic researchers, and its features can be modified by Tcl/Tk scripting (Tk is the standard GUI for Tcl).

[2]http://www.ks.uiuc.edu/Research/namd/

### 2.3.2 Modeling and Analysis

In this thesis, I primarily used NAMD's companion software VMD (version 1.9) [110] for visualization and analysis. VMD was used for preparing input files for NAMD, and for extracting different properties from MD trajectories. VMD is a versatile software that can run from a simple laptop up to GPU clusters, and it is able to perform parallel computations for computationally intensive tasks.

In this thesis, I used MATLAB (versions 2016-2018) for various calculations, optimizations, and creating data plots. The numerous built-in functions, various toolboxes, and libraries make MATLAB a unique tool for computational science researchers. I have provided all MATLAB codes in this thesis in Appendix A.

I used the MD-TASK package [112] for calculating dynamic cross correlation (DCC) maps. I calculated the DCC maps of each system based on the $C_\alpha$ atoms of residues using the MD-TASK package[3] [112]. These maps are reported in Fig. 5.4 (chapter 5), and Fig. 6.5 and Fig. 6.6 (chapter 6). Each cell value ($C_{ij}$) in the matrix of the DCC map was calculated using the following formula:

$$C_{ij} = \frac{< \Delta r_i \Delta r_j >}{\sqrt{< \Delta {r_i}^2 >}\sqrt{< \Delta {r_j}^2 >}} \tag{2.5}$$

Where $\Delta r_i$ represents the displacement from the mean position of atom $i$, and <> denotes the time average over the whole trajectory. The positive values of $C_{ij}$ show correlated motion between residues $i$ and $j$, moving in the same direction, whereas negative values of $C_{ij}$ show anti-correlated motion between residues $i$ and $j$, moving in the opposite direction.

[3]https://md-task.readthedocs.io/en/latest/home.html

I also used Linux-based operating systems (openSUSE and Ubuntu) for working with all local software packages, managing and deploying simulations on clusters, shell scripting for analyzing data, and scripting. In Linux, bash shell is the terminal language for various command line operations from simple tasks to analyzing data and for parallel calculations. I have provided all useful bash codes in Appendix B and C. I used MobaXterm (version 9.4) and Cygwin64 software for remotely accessed Linux-based workstations and supercomputing clusters.

I used CHARMM-GUI[4] and the MATCH web server[5] to produce both the topology and parameter files for small molecules. An important utility program termed catdcd (version 4.0) which is a built-in feature in VMD, was used for concatenating trajectory files and selecting the desired segments of trajectories (Appendix D).

## 2.4 Advanced Sampling Methods

In a conventional MD simulation, due to a large number of degrees of freedom, the explored conformational space of biomolecules in limited [137]. Therefore, several enhanced sampling methods have been developed to overcome the limitations of conventional MD simulations. Enhanced sampling methods such as temperature accelerated MD (TAMD) and metadynamics are methods to accelerate sampling and overcome free-energy barriers for enhancing conformational exploration [111, 138, 139, 137]. In this thesis, I used metadynamics and TAMD for enhanced sampling MD simulations in conjunction with a suite of methods, called the single-sweep, for the free-energy reconstruction, and for finding the interconnected network of migration pathways of ligands within proteins [140, 141, 137]. The details of these

[4]http://www.charmm-gui.org/

[5]http://brooks.chem.lsa.umich.edu/index.php?matchserver=submit

methods are described in the following.

### 2.4.1 TAMD Algorithm

TAMD is a method of enhanced sampling to explore the physical free-energy landscape [140], and has been successfully applied to a number of systems [142, 143, 144, 15]. TAMD algorithm begins by defining an extended version of the physical system in which new variables, $z$, are coupled to the physical variables via collective variables (CV), $\theta(x)$[145, 140]. The new total potential ($U$) contains, in addition to the physical energy ($V(x)$), a term that couples CVs ($\theta(x)$) to new $z$-variables, thus

$$U_\kappa(x,z) = V(x) + \frac{\kappa}{2}\sum_{j=1}^{m}(\theta_j(x) - z_j)^2 \tag{2.6}$$

Here, m is the number of CVs, $\kappa$ is a spring constant and is an adjustable parameter. The new variables $z$ are assigned a fictitious mass ($M$) and a fictitious temperature ($\bar{T}$) which is different from that of the physical system. The evolution of the extended system is governed by the following dynamical equations:

$$M\ddot{x} = -\nabla_x V(x) - \kappa\sum_{j=1}^{m}(\theta_j(x) - z_j)\nabla_x\theta_j(x) + \text{thermostat at } \beta^{-1} \tag{2.7}$$

$$\bar{\gamma}\dot{z} = \kappa(\theta(x) - z) + \sqrt{2\bar{\beta}^{-1}\bar{\gamma}}\eta^z \tag{2.8}$$

where $\bar{\gamma}$ is the artificial friction coefficient, $\bar{\beta}$ is the inverse of the artificial temperature ($\bar{\beta} = 1/(k_B\bar{T})$, where $k_B$ is the Boltzmann's constant), $\beta$ is the inverse of the physical temperature ($\beta = 1/(k_BT)$), and $\eta^z$ is the white-noise associated with the Langevin evolution of the new variables.

#### 2.4.1.1 Free-energy Reconstruction

The exploration of the protein interior space via TAMD results in a dense cloud of ligand positions (Fig. 2.1). In this thesis, I have used the single-sweep method [141] for

reconstructing the free-energy landscape in CVs that are TAMD-explored positions of the center-of-mass of $O_2$ and CO. This method has been successfully used as an efficient way of computing multidimensional free-energies [141]. Out of a large number of explored positions, it is necessary to choose a unique set of positions, namely centers, covering various areas of the protein. For estimating the mean-force required to keep the ligand at those positions, these centers are chosen along the trajectory based on a distance criterion: beginning with the value $z_1$ as the first center, a new point is added to the set when its distance from all other members of the set exceeds a prescribed cut-off distance. In the end, a large number of centers (∼300-700) are harvested (Fig. 2.1a). Each chosen location is referred to as a "center", and the $k^{th}$ center is indexed as $z_k$.

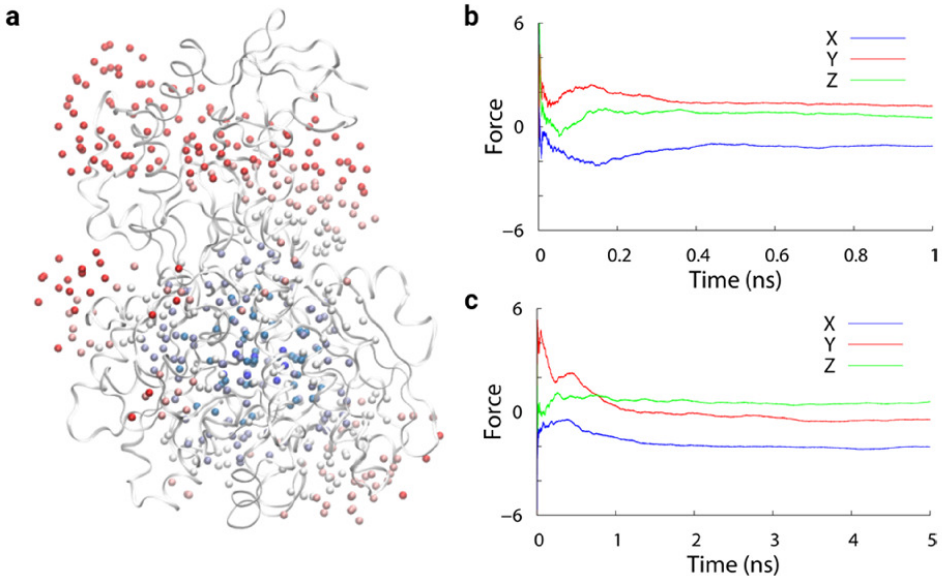



Figure 2.1: (a) The locations of 497 unique centers (the positions of the center-of-mass of $O_2$) chosen after TAMD sampling are highlighted. The protein-backbone is shown as white ribbons and the extracted centers are shown as spheres. The centers are colored from blue to red, based on their distance to the H-cluster. (b) The running-average of mean forces (kcal/mol.Å) for three Cartesian directions for one representative center (blue, X-direction; red, Y-direction; green, Z-direction) are shown from a 1 ns (panel b) and 5 ns (panel c) mean-force calculation.

The mean-force at each of the obtained $k$th center, $f(z_k)$, is calculated using the following time average:

$$f(z_k) = \frac{1}{t}\sum_{j=1}^{t} \kappa(\theta(x(t_j)) - z_k) \tag{2.9}$$

Here, $\kappa$ is the spring constant, $z_k$ is the position vector of the $k$th center, $t$ is the number of time increments in the trajectory and $\theta$ is the center-of-mass of the single molecule as a CV [58, 146]. The mean-force for each center was computed by using an MD simulation with CVs restrained at the center using a coupling constant of 10 kcal/mol$\cdot$ Å$^2$. The $C_\alpha$ atoms of 3 residues were restrained using a spring constant of 2.0 kcal/mol$\cdot$ Å$^2$ to orientally restrained the protein during each force calculation, as has been employed in other applications [57, 58]. The solvent centers around protein were chosen and (Fig. 2.1a) for which mean-forces were nearly negligible. The simulation for all centers are independent of each other, and therefore are independently distributed among different computing nodes [56]. Each mean-force calculation lasts for 1 ns, after which the convergence of mean-force values was observed (Fig. 2.1b). For some selected centers that needed more than 1 ns for force convergence, calculations were continued up to 5 ns (Fig. 2.1c). The histogram of standard deviation computed over the last 10% of each force calculation represent the convergence of mean-forces by delta function at origin (Fig. E.4).

Given the harvested centers and their computed respective mean-forces, the global reconstruction of the free-energy surface (FES), $A(z)$, was carried out as a linear combination of Gaussian radial basis functions [141]. Therefore, the analytical free-energy can be written as:

$$A(z) = \sum_{k=1}^{K} a_k \varphi_\sigma(\mid z - z_k \mid) + C \tag{2.10}$$

Here $\varphi_\sigma$ is a Gaussian with a width of $\sigma$, $K$ is the total number of centers, $a_k$'s are coefficients in expansion, and $C$ is a constant for adjusting the height of $A(z)$. The optimization method based on the least-square fitting was used for obtaining the optimal values of $\sigma$ and $a_k$ [147, 141, 148]. Briefly, optimal values of $\sigma$ and $a_k$ are obtained by minimizing the following error function ($E$):

$$E(a,\sigma) = \sum_{k=1}^{K} \mid \sum_{k'=1}^{K} a_{k'} \nabla_z \varphi_\sigma(\mid z_k - z_{k'} \mid) + f(z_k) \mid^2 \tag{2.11}$$

#### 2.4.1.2 Minimum free-energy pathways (MFEP)

Following the reconstruction of the free-energy surface, the minimum free-energy pathways (MFEP), as pathways of the single molecule diffusion on the FES, can be determined using the zero-temperature string method (ZTS) [149]. These MFEPs are identified as the diffusion pathways of ligands inside the protein [150]. The diffusion pathways are the network of MFEPs that interconnect local minima and the solvent portals. Given an analytical approximation of the free-energy landscape, obtained from the PMF reconstruction step, the ZTS method gives MFEPs as the steepest-descent paths from saddle points on the reconstructed FES [151, 149]. This optimization method starts from an initial guess for a curve on the FES, and in a two-step iterative algorithm finds the optimal pathways; in the first step, the points that were obtained by discretizing the line were allowed to move based on the steepest descent on the FES landscape, while the second reparametrization step kept the distances between the points uniform [151, 152]. MFEPs are the most probable pathways that one can take on the FES between two local minima [151, 152]. Since the three-dimensional FES has been reconstructed and it is available analytically, the ZTS method was employed without requiring further MD simulations [56].

### 2.4.2 Metadynamics Simulations

Metadynamics is an enhanced sampling method for faster exploration of conformational space in a specified set of collective variables (CVs) by augmenting the force-field with a history-dependent biasing potential ($V_{\text{meta}}$) of the following form [153, 154]:

$$V_{\text{meta}}(\xi) = \sum_{t'=\tau_{\text{G}},2\tau_{\text{G}},\ldots}^{t'<t} W \prod_{i=1}^{N_{\text{cv}}} \exp\left(-\frac{[\xi_i - \xi_i(t')]^2}{2\delta_{\xi_i}^2}\right) \tag{2.12}$$

where $\xi_i$ is the current value of the CV and $\xi_i(t')$ is the value of the CV at time $t'$. $V_{\text{meta}}$ is constructed as a sum of $N_{\text{cv}}$-dimensional repulsive Gaussian functions with a chosen height ($W$) and width ($\delta$). The Gaussian functions can be added at a desired frequency $\tau_{\text{G}}$. The three main parameters in metadynamics that control the efficiency and accuracy of the free energy reconstruction from converged metadynamics potential ($V_{\text{meta}}$) [139] are $W$, $\delta$, and $\tau_{\text{G}}$. The metadynamics method has been successfully applied to study many biophysical problems including computational drug design [155, 138, 156, 157, 158, 159, 160].

## 2.5 Procedures for MD Simulations

Three steps for carrying out MD simulations in NAMD are (*i*) solvation and ionization of the initial configuration, (*ii*) simulating the system using the MD software, and (*iii*) analyzing output files. These steps are described below.

All biomolecular systems are found in an aqueous environment with ions, and hence the prepared systems of biomolecules need to be solvated with water and neutralized by adding ions to mimic the physiological environment. The number of added water molecules need to be adjusted, as more water molecules are added the size of the system will increase, thereby increasing the computational cost. The solvate plugin in VMD provides the tool for solvation of the system prior to an MD simulation. The autoionize plugin in VMD provides a tool for adding ions to the solvated system. In this thesis, MD simulations of all systems were carried out using NAMD. The generated trajectories were visualized and analyzed by the

VMD software. In this thesis, I also used MATLAB and python based MD-TASK package for futher analyzing trajectory data.

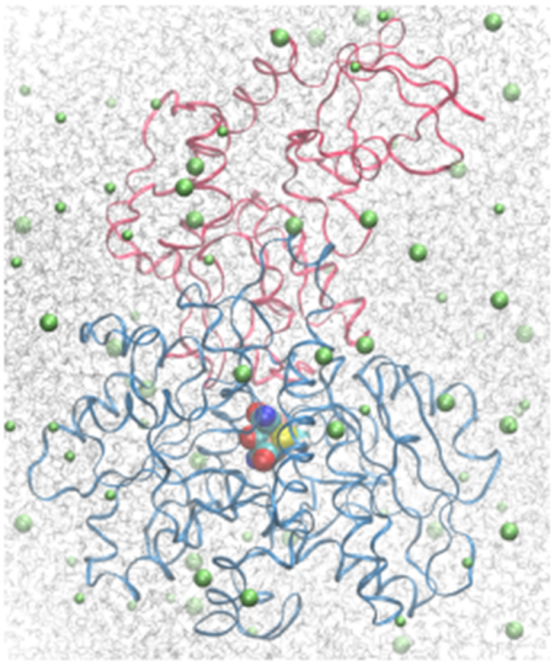


Figure 2.2: A snapshot of an MD simulation domain is shown. The H-cluster of [FeFe]-hydrogenase and ions (green) are shown as space-filling, and water molecules are shown as gray wireframe.

## 2.6 Computational Resources

The computational resources used in this thesis are as the following: (*i*) Lennard is a local linux workstation assigned to me in our laboratory at the University of New Hampshire equipped with 20 processors (model name : Intel(R) Xeon(R) CPU E5-2630 v4 @ 2.20GHz), 2 GB RAM, and a 2 TB internal hard-drive; (*ii*) the Trillian cluster, which was acquired through an NSF Major Research Instrumentation (MRI) grant, with cost sharing from the UNH Space Science Center ,Senior Vice Provost for Research, the College of Engineering and Physical Sciences (CEPS), and the Research Computing Center (RCC). This supercomputer

has 132 compute nodes, each with 2 AMD 16 core 'Abu Dhabi' 2.4 GHz CPUs, for a total of 4,224 cores; (*iii*) the Premise cluster which is made up of 37 compute nodes connected together using 56 Gb FDR Infiniband networking, each node has two 12-core CPUs, all nodes have at least 128GB of main memory. Ten nodes have NVIDIA K80 GPUs which I used mainly in MD simulations because of the higher performance compared to non-GPU compute nodes. The entire cluster shares 225TB of usable Lustre storage; and (*iv*) the Extreme Science and Engineering Discovery Environment (XSEDE) which is a powerful collection of integrated digital resources and services such as supercomputers, visualization and storage systems, collections of data, software, networks, and expert support that scientists, engineers around the world use to advance solving complex problems.

Comet is a dedicated XSEDE cluster, which was designed by Dell using Intel's Xeon Processor E5-2600 v3 family, with two processors per node and 12 cores per processor running at 2.5GHz. Each compute node has 128 GB (gigabytes) of traditional DRAM and 320 GB of local flash memory. Since Comet is designed to optimize capacity for modest-scale jobs, each rack of 72 nodes (1,728 cores) has a full bisection InfiniBand FDR interconnect from Mellanox, with a 4:1 over-subscription across the racks. There are 27 racks of these compute nodes, totaling 1,944 nodes or 46,656 cores. In addition, Comet has four large-memory nodes, each with four 16-core sockets and 1.5 TB of memory, as well as 36 GPU nodes, each with four NVIDIA GPUs (graphic processing units). The GPUs and large-memory nodes are for specific applications such as visualization, and MD simulations.

# CHAPTER 3

# STUDIES ON DIFFUSION PATHWAYS OF INHIBITORY GASES IN [FeFe]-HYDROGENASE

## 3.1 Abstract

The $H_2$ production potential of [FeFe]-hydrogenase, a hydrogen-producing enzyme from green algae, is reported to be promising for economical and large-scale production of $H_2$ as an alternative source of renewable energy. The production of hydrogen takes place at the catalytic center buried in the enzyme core. Unfortunately, binding of $O_2$ and CO to the catalytic center of the enzyme irreversibly inactivates it, essentially blocking hydrogen production. Therefore, a better understanding of the mechanism of $O_2$ and CO entry/exit is necessary to develop strategies for designing oxygen-tolerant enzymes. In this work, I have investigated the pathways and diffusion channels of $O_2$ and CO gases in [FeFe]-hydrogenase. Through exhaustive mapping of $O_2$ and CO diffusion channels, I computed a full thermodynamic map of preferred binding locations of $O_2$ and CO gases within the enzyme interior, which showed that $O_2$ and CO can enter and exit the enzyme through multiple pathways along which are key residues that are known to perturb binding rates of $O_2$ and CO binding. The global minimum for each gas in the free-energy landscape is located near the H-cluster, a key metallic center within the enzyme. Along $O_2$ and CO diffusion channels, I further identified several residues that could be potential candidates for mutations to increase the tolerance of [FeFe]-hydrogenase to both inhibitory gases.

## 3.2 Background

Enzymes use their active sites to exquisitely perform essential processes in living organisms by catalyzing chemical transformations. Given that the active sites are buried within the protein matrices, ligand migration between the solvent and the active site is an indispensable process for chemical reactions to occur. This observation hints at the existence of ligand entry/exit channels through which the activity of enzymes can be modulated by altering ligand diffusion (e.g. limiting the diffusion of inhibitory gases) [161, 162]. Although permanent diffusion channels may not exist in flexible structures of enzymes, transient thermal fluctuations throughout the protein matrix can potentially lead to the formation of a network of metastable pockets that can serve as conduits for site-to-site hopping of ligands.

Described in this chapter are studies on the thermodynamics of ligand diffusion pathways in the FeFe-hydrogenase, an enzyme of the hydrogenase superfamily, that is reported to be a key source of biological hydrogen production due to its ability to catalyze proton reduction and/or $H_2$-oxidation under optimal conditions [163, 164, 165, 166, 167, 168, 169]. The crystal structure of the CpI FeFe-hydrogenase from *Clostridium pasteurianum* revealed a deeply-buried active site (the H-cluster) with three 4Fe4S and one 2Fe2S accessory metal-clusters [170, 171]. The H-cluster is comprised of one $Fe_4S_4$ subcluster and one 2Fe subcluster with a vacant site on $Fe_d$ (distal) which can bind different ligands [167] including inhibitory gases CO and $O_2$ that can severely hamper the enzymatic activity of the FeFe-hydrogenase. Specifically, it has been suggested that $O_2$ attacks and irreversibly deactivates the H-cluster, while CO can compete with $O_2$ for binding to the H-cluster and prevent inhibition by $O_2$ when the enzyme is oxidizing $H_2$ or is at the resting state (no reaction) [172, 173, 174]. Regardless, CO can still irreversibly deactivate the enzyme during $H_2$-production [174]. Importantly, both CO and $O_2$ attack the same site, the $Fe_d$ of the 2Fe subcluster of the H-cluster [175, 176, 177]. Moreover, with CO binding to the active site, $O_2$ cannot access the active site, and previous studies have suggested a faster binding rate for CO than $O_2$ [178, 175].

Since accessing the active site requires channels connecting to the solvent, several possible scenarios can be hypothesized to explain the phenomenon of the protective effect and inhibition by CO: (a) CO diffuses faster than $O_2$ if both gases compete for the same diffusion pathway; (b) CO migrates along shorter diffusion pathways if each gas has an independent pathway; and/or (c) CO has a higher binding-affinity for $Fe_d$ than $O_2$. Therefore, a detailed mapping of diffusion pathways of CO and $O_2$ in the FeFe-hydrogenase is not only needed for resolving these questions, but also for developing approaches to enhance the tolerance of the FeFe-hydrogenase to inhibitory gases.

Furthermore, to improve the tolerance of hydrogenase enzymes to inhibitory gases, mutations in specific amino-acid residues have been carried out. For example, the rates of CO and $H_2$ diffusion in another enzyme of the hydrogenase family, *D. fructosovorans* NiFe-hydrogenase, can be decreased after mutations of two residues (V74 and L122) [175, 178]. However, it is highly challenging to rationally choose residues for mutations among a large number of residues (e.g. 574 residues in the CpI FeFe-hydrogenase), especially without the knowledge of diffusion channels and the location of the free-energy barriers along them. This is evident in an experimental study on mutations of seven residues in the CpI FeFe-hydrogenase (C299, F417, V423, A427, A431, I461, and F493), where none of the residues showed any strong effect on the kinetic performance of this hydrogenase although some of these residues (A427, A431, I461, F493) are located in the pathway A [176]. This suggests the existence of additional pathways for gas diffusion in the FeFe-hydrogenase. Indeed, my previous study [1] disclosed an exhaustive mapping of the diffusion network of $O_2$ in the CpI FeFe-hydrogenase showing several new pathways to access the active site besides those previously suggested. We also suggested several new candidate residues lining diffusion pathways of $O_2$ whose mutations have the potential to significantly disrupt the $O_2$ diffusion network, as these residues are located in the vicinity of the free-energy barriers. However, no studies on the exhaustive mapping of the diffusion pathways of the competing inhibitor CO in the FeFe-hydrogenase have been carried out so far although pathways of CO diffusion in structurally-unrelated

NiFe-hydrogenase have been studied [179].

**Hypothesis:** An interconnected network of migration pathways exists for inhibitory gases throughout CpI [FeFe]-hydrogenase.

Therefore, I have studied the pathways and diffusion channels of $O_2$ and CO in CpI [FeFe]-hydrogenase by employing a judicious combination of enhanced sampling and free-energy reconstruction methods [56]. By finding the three dimensional free-energy of $O_2$ and CO binding within the enzyme, and interconnected network of diffusion pathways of both inhibitory gases, unique insights into $O_2$ and CO diffusion in [FeFe]-hydrogenase have been obtained. The results show that, the global minimum in the free-energy landscape is in the proximity of the H-cluster binding site (location where $O_2$ bind to the H-cluster), and another local minimum is in the proximity of the Xe-binding site (an experimentally found location for xenon presence [180]). The work shows the presence of numerous interconnected pathways of $O_2$ and CO diffusion toward the H-cluster beyond the two main channels, providing enhanced information on other areas of protein affecting $O_2$ diffusion. Moreover, along each pathway, energy barriers and local minima in the free-energy surface can be readily found. The important residues along the obtained pathways are prime candidates for guiding and designing [FeFe]-hydrogenase tolerant to both inhibitory gases.

## 3.3 Methods

### 3.3.1 System Setup

I carried out all MD simulations (section 2.2) using the CHARMM force-field for proteins, ions, water, and metal clusters [122, 181, 182]. I prepared a fully-solvated and ionized simulation domain using the initial coordinates for all protein atoms from the crystal structure of the FeFe-hydrogenase (PDB code 1FEH), solvating with explicit water (TIP3P) molecules,

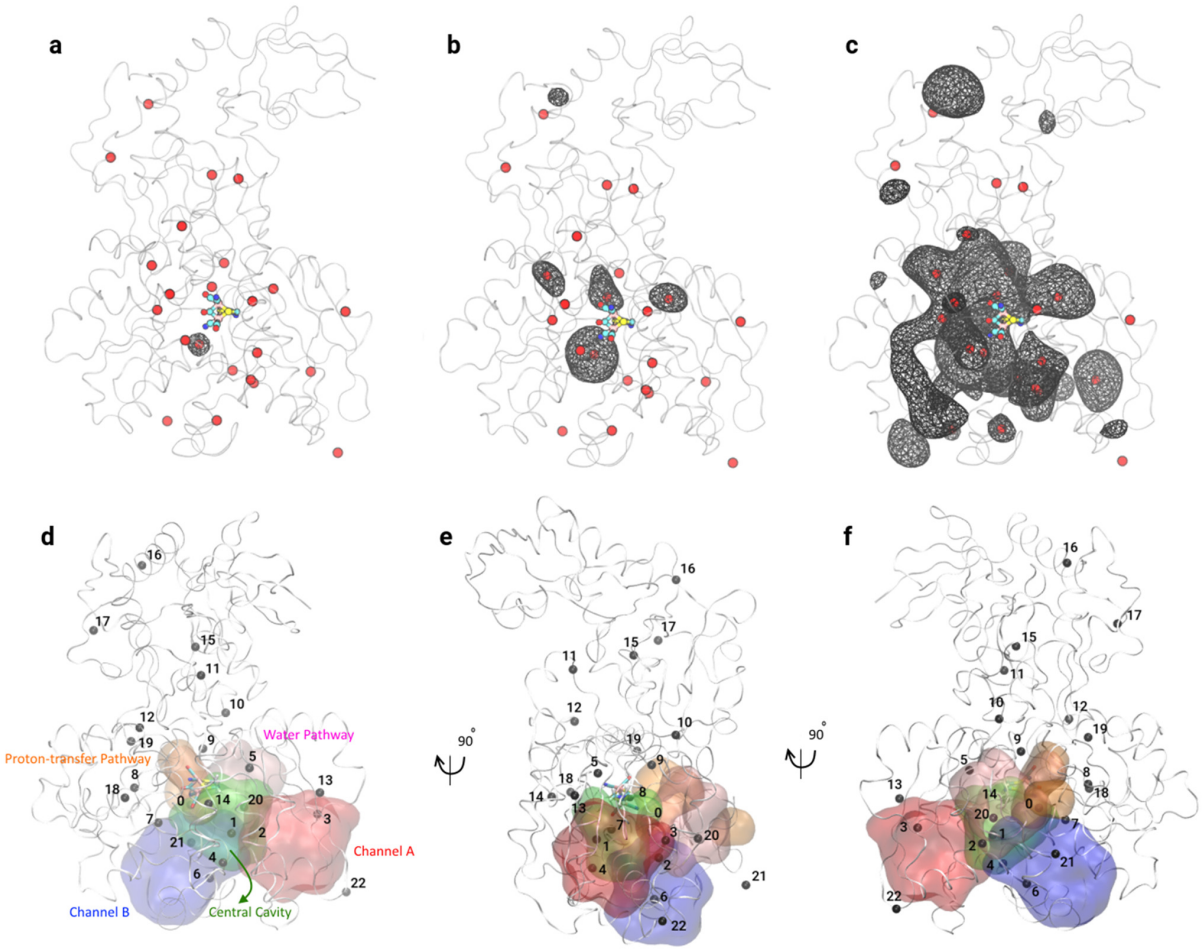



Figure 3.1: Isosurfaces of the 3D-PMF map of $O_2$ inside [FeFe]-hydrogenase are shown in top panels. Energy levels are represented at (a) 3 kcal/mol, (b) 12 kcal/mol, and (c) 17 kcal/mol with respect to the global free-energy minimum in proximity of the H-cluster. The local minima are represented by red spheres. The backbone of protein is shown as ribbons, the isosurfaces are shown as black mesh, and the H-cluster is shown as sticks. In bottom panels, the proximal area of the central cavity (green surface), channel A (red surface) and B (blue surface) [5], water pathway (pink surface) [6], and the proton transport pathway (orange surface) [7] are shown in surface representations as a front-view (d), a side-view (e), and a back-view (f). The local minima are represented by black spheres and labeled. The protein backbone is rendered as ribbons, and the H-cluster is shown as sticks.

and ionizing with NaCl. After an initial energy minimization of this system for 1000 steps, the box-volume was equilibrated by conducting a 1-ns long molecular dynamics (MD) simulation in the NPT ensemble, where the temperature was controlled at 310 K using the

Langevin thermostat and the pressure was controlled using the Nosé-Hoover barostat. An integration time-step of 1 fs, periodic boundary conditions, the particle-mesh Ewald summation (with a grid spacing of 1 Å) for long-range electrostatics, and a cutoff distance of 10 Å for van der Waals interactions was used. For exhaustive exploration of gas diffusion using enhanced sampling methods, I created several initial systems, each differing only in the initial position of gas molecule, which was chosen to span various locations throughout the enzyme matrix. Each of these systems was MD equilibrated for 1 ns in the NPT ensemble using the protocol described above. The coordinates from the end of the MD trajectories of these equilibrated systems were used as initial conditions for TAMD simulations (*vide infra*).

### 3.3.2 TAMD Simulations

I applied TAMD (section 2.4.1) for exhaustively exploring the gas-accessible interior volume of the FeFe-hydrogenase by using the Cartesian coordinates of the center-of-mass of each gas as the three-dimensional CV-space. In TAMD simulations, $\bar{\beta}^{-1}$ values ranging between 2 kcal/mol and 7 kcal/mol, a spring constant $\kappa$ of 200 kcal/mol·Å$^2$, and a fictitious friction coefficient $\bar{\gamma}$ of 50 ps$^{-1}$ were used. Starting from each of the initial locations of each gas, 30 independent 1-ns long TAMD simulations were performed resulting in a total of $\sim 300-600$ ns of simulation time for exhaustive sampling of each gas diffusion.

### 3.3.3 Free-energy Surface Reconstruction and Minimum Free Energy Pathways

To reconstruct the 3D-FES of $O_2$ and CO diffusion in the FeFe-hydrogenase, we used the single-sweep method [141, 56] in which mean-forces are computed at CV positions sampled by TAMD (section 2.4.1). Specifically, it is necessary to choose distinct CV positions, also termed unique centers, throughout the protein matrix. Due to the rotation and translation of the protein, all TAMD trajectories were aligned to a common reference structure for extracting centers. Out of all gas positions sampled by TAMD, we chose several unique centers

(Fig. 2.1a). The centers were chosen by beginning with the first center ($z_1$) and adding new centers such that the distance of each new center from all previous centers exceeds a prescribed cutoff distance.

To compute the mean-force at each of the $k$ centers, $f_k$, we launched a set of new MD simulations in which the center-of-mass of $O_2$ and CO is harmonically restrained about the CV-value at the chosen center. We conducted restrained MD simulations by using a spring constant $\kappa$ of 10 kcal/mol·Å$^2$. Each of these simulations generally lasted for 1 ns within which we observed the convergence of the mean-force (Fig. 2.1b), but some centers required simulations up to 5-ns long for the convergence of the mean-force (Fig. 2.1c). We show the convergence of all mean-force calculations by plotting the histogram of the standard deviation of each mean-force calculation computed over the last 10% of each restrained MD trajectory. This histogram reveals the behavior of a delta function centered at the origin indicating convergence (Fig. E.4).

Using converged mean-forces at all centers, we globally reconstructed the free-energy surface $A(z)$ (section 2.4.1.1) as a linear combination of Gaussian radial basis functions by obtaining the optimized parameters (coefficients, $a_k$; and the Gaussian width, $\sigma$) via a least square fitting procedure per equation [141]. We further studied the 3D-FES using the zero-temperature string (ZTS) method [183, 184, 185, 186] (section 2.4.1.2) and found MFEPs between all pairs of minima as the steepest descent paths from saddle points on the analytically reconstructed free-energy hypersurface, $A(z)$. Therefore, MFEPs are the most probable pathways between pairs of minima on $A(z)$ and are identified as the diffusion pathways for $O_2$ and CO inside the enzyme. We also note that the application of the ZTS method does not require any additional MD simulations since the analytical functional form of the 3D-FES is available from equation.

## 3.4 Results

### 3.4.1 Mean Force Estimation and PMF Map Reconstruction for $O_2$

The centers were selected as described in Chapter 2 (Section 2.4, Fig 2.1a). The converged mean-force values for 497 centers (chapter 2) were used for the reconstruction of the free-energy landscape of $O_2$ diffusion (inside the [FeFe]-hydrogenase), as a linear combination of radial basis functions using the single-sweep method [56]. In Figures 3.1a to 3.1c, I show isosurfaces of the 3D-PMFs, obtained from single-sweep calculations for $O_2$, at three energy levels of 3, 12, and 17 kcal/mol with respect to the global energy minimum in the free-energy landscape. The obtained 3D-PMF of $O_2$ binding sites on scale of the entire protein elucidates the detailed physical insight of $O_2$ transport in this protein. The global minimum in the PMF is located in the vicinity of the active-site pocket of the [FeFe]-hydrogenase (local minimum 0 in Figures 3.1d to 3.1f), and I found a free-energy difference of $\sim$18.5 kcal/mol between the global minimum and the solvent space.

In addition to global minimum, we identified 22 other local minima within the protein matrix (Figures 3.1d to 3.1f). I observed a local minimum in the proximity of the only known Xe-binding site [187] that is near the central cavity [5] and next to the H-cluster (local minimum 1 in Figures 3.1d to 3.1f). The relative free-energy difference for this local minimum is $\sim$13.31 kcal/mol with respect to the global minimum. I further observe local minima 2 and 20 (Figures 3.1d to 3.1f) within a previously identified water pathway [6] (pink surface in Figures 3.1d to 3.1f) that extends from the solvent to the central cavity [6] (green surface in Figures 3.1d to 3.1f). The local minimum 2 is in the proximity of S320 (with a free-energy of 11.72 kcal/mol), and the local minimum 20 is in the proximity of the surface-exposed K571 (with a free-energy of 13.76 kcal/mol). Several other local minima are located throughout the protein matrix with a large number of them in the active-site domain (Figure 3.1).

### 3.4.2 Diffusion Pathways of $O_2$ in [FeFe]-hydrogenase

I characterized diffusion pathways of $O_2$ in [FeFe]-hydrogenase by computing MFEPs between pairs of local minima located in the 3D-PMF (see methods section for details). In Figures 3.1a to 3.1c, I show the locations of all local minima by red spheres. The MFEPs are relatively dense in the vicinity of the H-cluster in the active-site domain of the protein, and in the non-active-site domain, I find MFEPs only near the protein surface (Figure 3.2a). The pathways to the Geminate-site (G-site, a location near the global minimum indicated by 0/0-G in Figure 3.2a,b) are largely originating in the lower side of the active-site domain via the central cavity (green surface in Figures 3.1d to 3.1f), in the water pathway [6] (pink surface in Figures 3.1d to 3.1f), and in the upper-part of the active-site domain.

In Figure 3.2a, I show the diffusion pathways of $O_2$ within the [FeFe]-hydrogenase with color-coded values of the free-energy. The global minimum, located next to the H-cluster, is shown by a blue sphere, on the side of the distal Fe-site of the H-cluster [49]. The local minimum corresponding to the Xe-binding site inside [FeFe]-hydrogenase is designated as minimum 1, and the MFEP that connects the Xe-site to the G-site is shown by a pathway between minima 0 and 1 (Figure 3.2a), and the barrier of this MFEP is in the proximity of A272. The local minimum 2 and 20 are in the proximity of S320 and K571, respectively; both have MFEPs to the G-site (Figure 3.2a). The energy barrier of the MFEP between the local minimum 2 and the G-site is in the proximity of A321.

The local minimum 9 is in the proximity of cluster 1, and the MFEP between the G-site and local minimum 9 shows a pathway for $O_2$ with access both to the H-cluster and to the adjacent [$Fe_4S_4$] cluster. The local minima 7, 8, and 19 with free-energy values of 11.86, 12.18, and 8.88 kcal/mol are located within the active-site domain, and the pathway between them has access to the H-cluster from the local minimum 7, and has access to cluster 1 from the local minimum 19. The access of $O_2$ to the local minima 7, 8, and 19 from pathways originating in the active-site domain is more likely due to lower energy barriers (Figure 3.2a). The Xe-site has access to the local minima 2 and 4 (Figure 3.2a) with the energy barriers

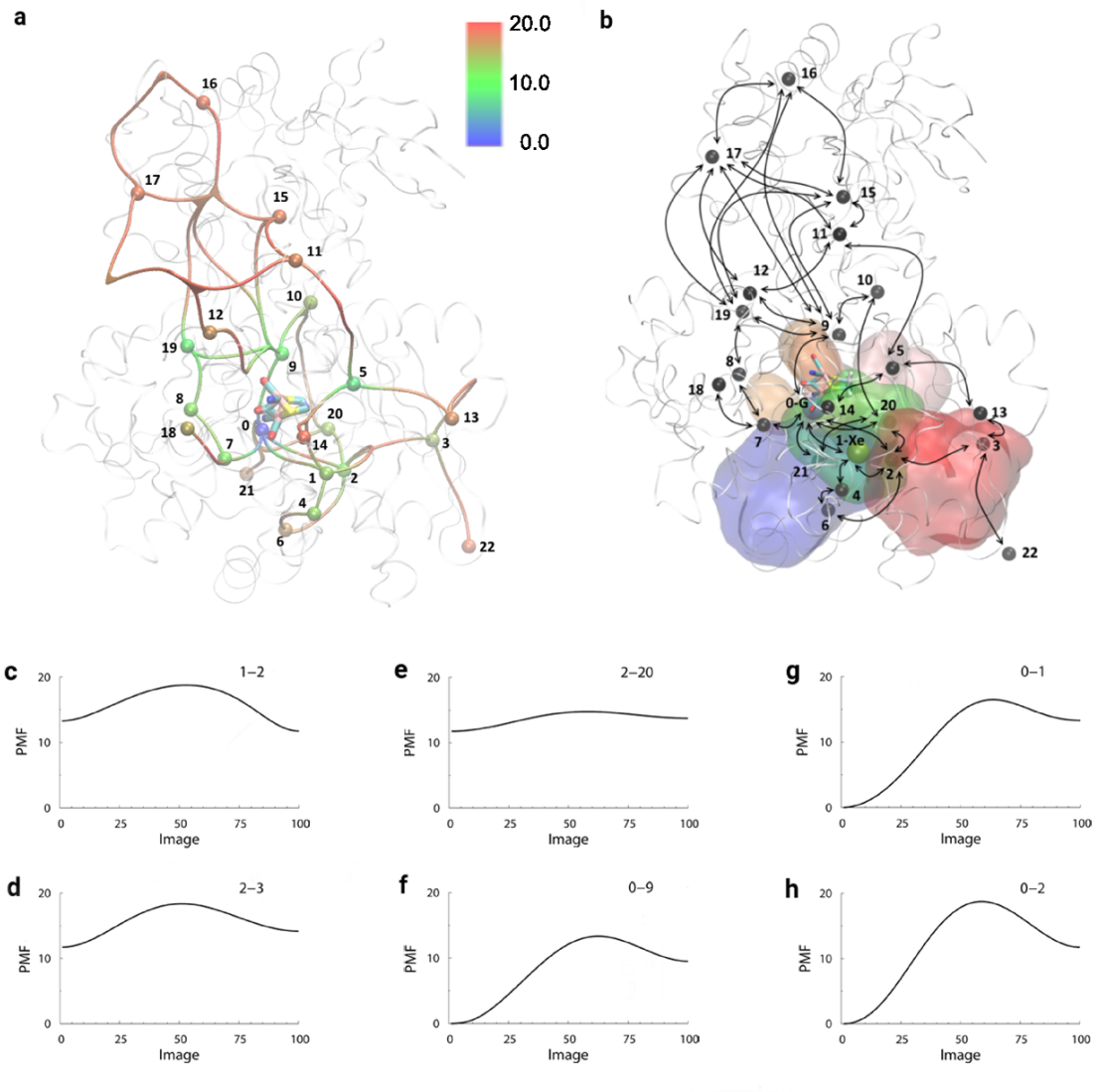



Figure 3.2: (a) Pathways of $O_2$ migration inside [FeFe]-hydrogenase, values of the PMF along each MFEP and local minima, are indicated by color. The pathways are represented as curved lines and the local minima as spheres. (b) The connection between local minima along pathways are shown by black arrows. The local minimum 0 corresponding to the G-site is shown as a blue sphere, and the local minimum 1 corresponding to the Xe-site is shown as a green sphere. The bottom panels show the PMF profiles along MFEPs between local minima 1 and 2 (panel c), local minima 2 and 3 (panel d), local minima 2 and 20 (panel e), local minima 0 and 9 (panel f), local mlocal minima 0 and 1 (panel g), and local minima 0 and 2 (panel h).

of 7.72 and 0.76 kcal/mol, indicating that it is easier for $O_2$ to enter the Xe-site from the local minimum 4 than from the local minimum 2. The local minimum 6 is located within the channel B (blue surface in Figures 3.1d to 3.1f), as described by Cohen et al. [5], and it is connected to local minima 2 and 4 with energy barriers along MFEPs of 1.57 and 1 kcal/mol, suggesting that the energy barriers are lower along the pathway 6-4-1 than the pathway 6-2-1.

The energy barrier for $O_2$ entering the G-site from the Xe-site (pathway 1-0) is 3.6 kcal/mol, whereas the energy barrier from the local minimum 2 to the G-site (pathway 2-0) is 7.29 kcal/mol, suggesting that it is easier for $O_2$ to access the H-cluster from the Xe-site than from the local minimum 2. The sum of energy barriers along three pathways to the H-cluster starting from local minimum 6, pathways 6-4-1-0, 6-2-0, and 6-2-1-0 are 5.36, 8.86, and 12.89 kcal/mol, suggesting that it is unlikely for $O_2$ to take pathway 6-2-1-0 in comparison to pathways 6-2-0 and 6-4-1-0 (Figure 3.2b). The local minimum 3 is located within the channel A (red surface in Figures 3.1d to 3.1f), as identified by Cohen et al.[5], and it is connected to the local minimum 2 with an energy barrier of 4.33 kcal/mol along the MFEP. The two pathways to the H-cluster starting from the local minimum 3, pathways 3-2-0 and 3-2-1-0, have energy barriers of 11.62 and 15.65 kcal/mol, suggesting that $O_2$ diffusion is less likely along the pathway 3-2-1-0 than the pathway 3-2-0 to the H-cluster (Figure 3.2b). A table of energy barrier values between local minima is reported in Table E.5.

The free-energy profiles for MFEPs between local minima 0 and 1, local minima 1 and 2, local minima 2 and 3, local minima 0 and 2, local minima 0 and 9, and local minima 2 and 20 (Figure 3.2c to 3.2h) are showing the change in PMF values as $O_2$ is migrating along the pathway between them, and all PMFs are shown with respect to the global minimum at the G-site (indicated by 0 in Figure 3.2b). The pathways on the upper side of the H-cluster, e.g., pathway 7-8-19-9, are suggesting the existence of a $O_2$ circulation network with relatively low-energy barriers around the H-cluster; however $O_2$ migration is less likely to direct to

pathways in the non-active-site domain due to higher energy barriers, e.g., energy barrier of the pathway 9-12 is 11.96 kcal/mol, suggesting that it is unlikely for $O_2$, after entering the protein from the active-site domain, to access the non-active-site domain versus entering the protein from the active-site domain to access the H-cluster.

The local minima 21 and 22 with free-energy of 16.91 and 18.21 kcal/mol, respectively, are located in the solvent. These PMF values are showing an estimate of the free-energy values in the solvent and suggest an oxygen diffusion gradient toward the protein interior. The local minimum 22 is connected to the local minimum 3 that has access to the H-cluster via channel A (Figures 3.2b) [5]. The local minimum 21 is connected via an MFEP to the local minimum 0 or the G-site with an energy barrier of 1.8 kcal/mol which is in the proximity of F570, suggesting that this particular pathway potentially provides the easiest access route for $O_2$ to the H-cluster. This pathway is located between the channel B [5] and the water pathway [6]. The local minimum 20 with an energy of 15.34 kcal/mol is located within the water pathway, and is connected to the local minimum 2 and the G-site with energy barriers of 1.04 and 3.5 kcal/mol, indicating that it is less likely for $O_2$ to migrate along pathways 20-2-1-0 or 20-2-0 than the pathway 20-0. The pathway 9-0 is in the proximity of residues K358 and E361, defining the water-based proton transfer pathway [188, 7] (orange surface in Figures 3.1d to 3.1f). The free-energy profile of $O_2$ diffusion along this pathway is shown in Figure 3.2f indicating an energy barrier of 4 kcal/mol along this MFEP. The local minimum 10 with a free-energy value of 14 kcal/mol is connected to the local minimum 9. Via this pathway, $O_2$ entry from the local minimum 10 could be directed to the proximity of cluster 1 and to the H-cluster, suggesting that the pathway 10-9-0 could be another access route for $O_2$ to access the H-cluster, and similarly, the pathway 13-5-1-0 shows another possible route to the H-cluster.

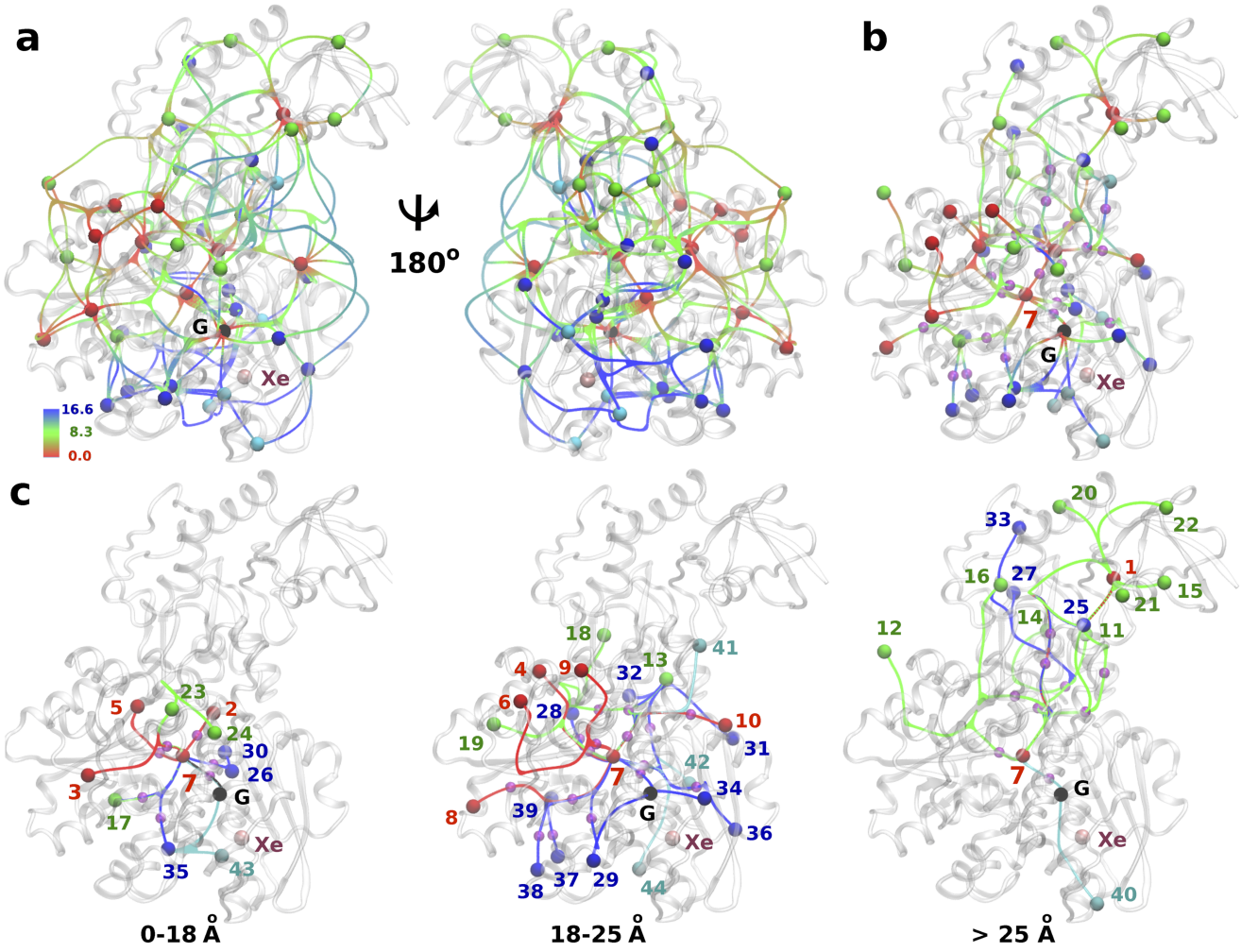



Figure 3.3: (a) Front and back views of MFEPs for CO diffusion are depicted on the structure of the FeFe-hydrogenase along with the positions of all minima (indicated by spheres). (b) The MFEPs for CO diffusion from the other local minima to the local minimum 7 are shown. (c) The MFEPs for CO diffusion within 0-18 Å (leftmost panel), 18-25 Å (middle panel), and over 25 Å (rightmost panel) of the local minimum 7 are shown. The locations of the energy barriers are shown by transparent magenta spheres in panels b and c (This figure was created by my co-author Yong Liu. [3]).

### 3.4.3 Network of CO Diffusion Pathways

We located an MFEP for each pair of local minima using the zero-temperature string (ZTS) method [183, 184, 185] (section 2.4.1.2). The ZTS-algorithm begins with an initial guess for a curve connecting the chosen minima pair on the analytically-reconstructed 3D-FES. This initial guess is then iteratively optimized via a discretized parameterization scheme

by keeping points along the curve equidistant on each iteration while allowing points to move on the basis of the steepest descent on the 3D-FES. Therefore, MFEPs are the pathways of minimal free-energy for CO diffusion in the FeFe-hydrogenase, which are assumed as the most likely diffusion channels for CO.

We show the network of these pathways in Fig. 3.3a, where one can visualize the color-coded free-energy values along with the MFEPs between local minima pairs (depicted as spheres). Consistent with the 3D-FES, we found that the MFEPs of CO diffusion along previously proposed hydrophobic pathways have high free-energy values (over 8 kcal/mol), which are shown as blue curves. We also observed that the MFEPs interconnect multiple local minima between the solvent-space and the H-cluster. We present a center-by-center matrix of energy barriers between all pairs of local minima in Appendix E. We find that major energy barriers along MFEPs originating from the global minimum to neighboring local minima are $\sim$10 kcal/mol, and these barriers are higher than the energy barriers of reverse MFEPs. This suggests a preference for CO diffusion toward the global minimum. Previous studies have suggested that CO binds the $Fe_d$ of the 2Fe subcluster of the H-cluster, thereby competing with the binding of $O_2$ [174, 178, 175]. To explore the CO diffusion pathways connecting the solvent region and the H-cluster, we present the MFEPs of CO diffusion between the local minimum 7 (which is located in the vicinity of the $Fe_d$ of the 2Fe subcluster of the H-cluster, an area known as the geminate site) and other neighboring minima (Fig. 3.3b and 3.3c). The locations of the major energy barriers along MFEPs are indicated by magenta spheres (Fig 3.3c).

Within 18 Å of the H-cluster, we identified 12 local minima, including the global minimum: 5 minima with the free-energy values lower than 4 kcal/mol (red/black spheres), 3 minima with the free-energy values in the range of 4 to 8 kcal/mol (green spheres), 3 minima with the free-energy values in the range of 8 to 12 kcal/mol (blue spheres), and 1 minimum with a free-energy value of 18.20 kcal/mol (cyan sphere).

The energy barriers for pathways 2-7 and 5-7 are 2.45 kcal/mol and 5.23 kcal/mol, respectively. The pathways 3-7, 23-7, and 24-23-7 have energy barriers of 9.29 kcal/mol, 3.61 kcal/mol, and 7.58 kcal/mol, respectively. The pathways 17-7 and 35-7 share a part of their MFEPs, but the major energy barriers are located at distinct positions, in the vicinity of the residue F328 for the pathway 17-7 *vs.* the residue D273 for the pathway 35-7. The MFEP originating at the minimum 30 passes through the minimum 26 before arriving at the minimum 7. Therefore, the pathways 30-7 and 26-7 share the same major energy barrier of 2.49 kcal/mol. The pathway 43-7 passes through the local minimum 35 as well as the global minimum (G), and both MFEPs 43-7 and G-7 have the same major energy barrier of 14.62 kcal/mol. This suggests that a deep local minimum with high energy barriers of CO escape could trap CO.

Within 18 to 25 Å of the H-cluster, we found 20 additional distinct local minima, most of which are located at or near the surface of the enzyme and likely serve as local reservoirs for CO as it enters from the solvent. The local minima 4, 6, 8, 9, and 10 have free-energy values lower than 4 kcal/mol (relative to the global minimum). The pathways 6-7 and 9-7 share part of their MFEPs as well as the location of the major energy barrier in the vicinity of the residue A230, but the values of the free-energy barriers for each pathway are distinct, 9.29 kcal/mol (pathway 6-7) and 8.15 kcal/mol (pathway 9-7). Also, the pathways 6-7 and 9-7, respectively, pass through the minima 3 and 23, two minima located within 18 Å of the H-cluster (left panel; Fig. 3.3c). Similarly, the pathway 4-7 passes through the local minimum 5 and shares a part of its MFEP with the pathway 19-7 as well as the location of the major energy barrier in the vicinity of the residue V352, but the value of the energy barrier for each pathway is different, 5.23 kcal/mol for the pathway 4-7 (same value as for the pathway 5-7) and 8.15 kcal/mol for the pathway 19-7. Even though the pathways from the local minima 8, 37, 38, and 39 to the minimum 7 merge in the pathway 17-7, the locations of the major energy barriers along the MFEPs of 8-7, 37-7, and 38-7 are different. Specifically, the barriers along pathways 8-7, 37-7, and 38-7 are located near residues F348, I567, and

A331, respectively. The energy barrier for the pathway 39-7 is in the vicinity of residues S323 and Q325. Since the pathway 29-7 passes through the local minimum 35, the location of the energy barrier is same as for the pathway 35-7. Several other pathways show shared as well as unique features: the pathways 10-7, 41-7, 32-13-7 and 18-5-7 merge into the pathway 2-7, but the location of the energy barrier for each pathway is different, and the pathways 31-13-30-26-7, 36-26-7, 42-30-26-7, and 44-26-7 merge into the pathway 26-7, but the energy barriers for 31-13-30-26-7 and 36-26-7 are located in the unshared parts of their MFEPs and the energy barriers for 42-30-26-7 and 44-26-7 are located in the shared part 26-7. The pathway 36-7 passes through the global minimum and has the same value for the major energy barrier with the same location as the pathway G-7.

We observed 13 local minima at a distance over 25 Å from the $Fe_d$ of the H-cluster, most of which are located in the non-active-site domain of the enzyme (rightmost panel in Fig. 3.3c). Among these 13 minima, two (12 and 40) are located in the solvent region near the enzyme surface and show MFEPs reaching to the minimum 7: the pathways 12-19-5-7, 19-5-7, 4-5-7, and 5-7 all arrive at the minimum 7 and along which the major energy barrier of 5.23 kcal/mol is shared, and the pathway 40-7 passes through the local minimum 43 as well as the global minimum to arrive at the local minimum 7. The MFEPs of several other minima in the non-active-site domain merge at the local minimum 2, but each MFEP has a different location and magnitude of the energy barrier (see Tables E.5 and E.6).

## 3.5 Discussion

In this chapter, we have presented details on characterizing $O_2$ and CO diffusion channels and their underlying thermodynamics via a combination of several state-of-the-art computational methods [145, 140, 141, 150]. The 3D-PMF for $O_2$ binding shows a preference for $O_2$ diffusion into the protein interior relative to the solvent. The preference for protein environment by $O_2$ has been similarly observed in the earlier work of Cohen et al. [5]. The global minimum

Table 3.1: List of candidate residues for mutagenesis to increase the tolerance of the CpI FeFe-hydrogenase for CO. The symbol ✓ indicates that the residue has also been previously tested/proposed [1, 2] for disrupting the diffusion network of the competing inhibitor $O_2$ in this enzyme (This table was created by my co-author Yong Liu [3]).

| Residue | MFEP(s) | experimental | computational |
|---|---|---|---|
| L191 | 13-2 | | ✓ |
| C193 | 31-26 | | |
| I197 | 25-2 | ✓ | |
| A228 | 6-7 | | ✓ |
| A230 | 9-7, 6-7 | | |
| M295 | 44-26 | | ✓ |
| C300 | 26-7 | | |
| G302 | 42-26, 31-26 | | ✓ |
| S323 | 39-7, 25-7 | | ✓ |
| Q325 | 39-7, 25-7 | | |
| V352 | 19-7, 4-7 | | ✓ |
| K358 | 5-7, 4-7, 19-7, 28-7 | | ✓ |
| T380 | 4-7 | | ✓ |
| I416 | 41-G | | |
| V423 | G-7 | ✓ | |
| A426 | 29-G | | ✓ |
| V496 | 36-26, 44-26 | | |
| H500 | 10-2 | | |
| C503 | 2-7 | | |

in the free-energy is close to the residue P324 and the H-cluster (Figure 3.4a), in the vicinity of an area termed as the **G**-site (Geminate) [10]. The G-site is a location one step away from covalent binding of $O_2$ to the H-cluster that irreversibly inactivates the protein. The noticeable difference in the free-energy between the global minimum and other local minima would make it difficult for $O_2$ to leave the global minimum to other areas of the protein matrix, thus making binding to the H-cluster highly likely, because this is the only path from the G-site that leads to a lower free-energy, as also highlighted in previous studies [10, 5]. All local minima with energies close to the global minimum are in the active-site domain of the protein, whereas those local minima in the non-active-site domain of the protein have energy values closer to those in the solvent.

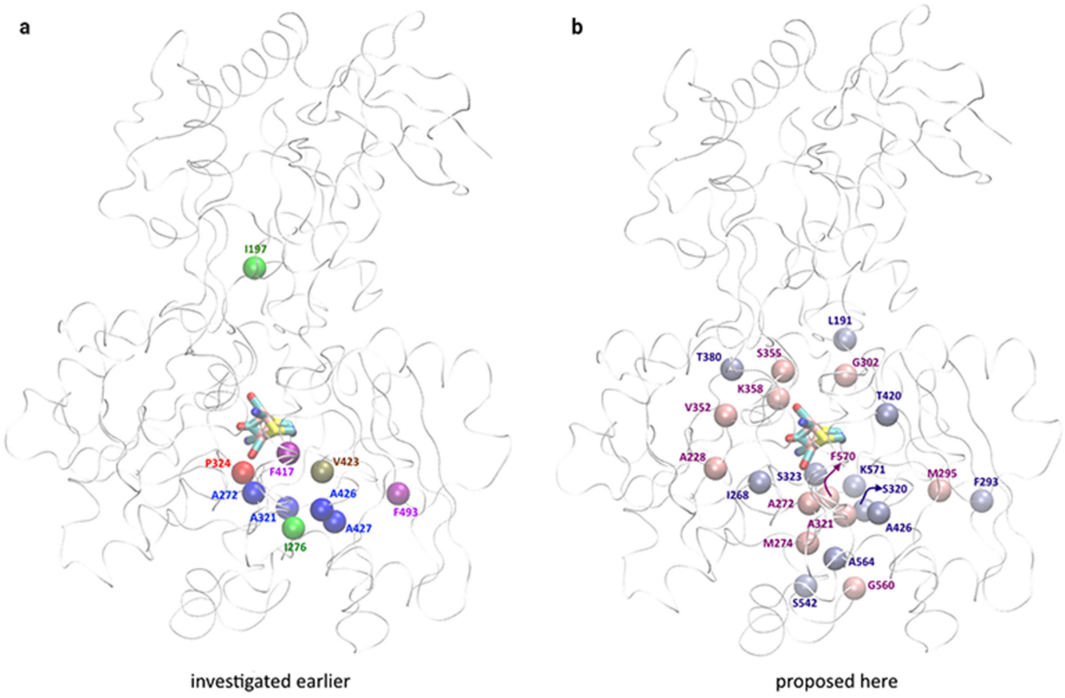



Figure 3.4: The $\alpha$ carbon-atoms of several candidate residues investigated by earlier studies (panel a) and proposed in this work (panel b) are shown. (a) The residue P324 is located near the H-cluster and the G-site. The residues I197 [8, 9], F417 [10], A321 [11], and V423 [11] are experimentally shown to alter $O_2$ diffusion in [FeFe]-hydrogenase. The hydrophobic residues A426, A427 and V423 are located in the proximity of the Xe-site. The residues F493 and A427 are along the pathway to the Xe-site. The residues I276, A272, V423, F493 and A427 are proposed for mutation in the patent application of King et al. [12] (b) Potential candidate residues for mutations to alter the oxygen-diffusion kinetics in [FeFe]-hydrogenase (A321 has been tested in study of Ghirardi et al. [11]) from our work. The residues at the local minima are shown as blue spheres, and near the energy barriers are shown as pink spheres.

### 3.5.1 $O_2$ Pathways in the Vicinity of the Xe Site

In crystallographic studies, Xe as a probe (similar to $O_2$) has been observed to prefer hydrophobic environments in the protein, but it is more electron rich than $O_2$, so it can readily be detected by the crystallographic methods [180]. We observed a local minimum close to the only experimentally known Xe-binding site in this [FeFe]-hydrogenase [187]; as expected, this local minimum is located in a mostly hydrophobic environment of the protein,

e.g., A426, A427, and V423 (Figures 3.4a). The Xe-binding site is reported to be blocked by the side-chains of A427 and F493 in CpI hydrogenase [6]. We found that the computed pathway to this local minimum is well aligned with the side-chains of A427 and F493, along the pathway 1-2. This local minimum is in the proximity of A426 (Figure 3.4a). Importantly, another experimental study on [FeFe]-hydrogenase found that the mutation A426L makes the Xe-binding site smaller [6].

### 3.5.2 $O_2$ Diffusion in Water and Proton-Transport Pathways

The location of two local minima (the local minimum 2 and 20) within the proposed water pathway [6], shows that $O_2$ and water may both use this pathway, as has been suggested by Lautier et al. [6]. The local minimum 2 and 20 are in the proximity of S319 and K571, respectively. In agreement with previous experimental and theoretical work, both residues are found to have contact with stabilized water molecules [170, 33]. Several theoretical and experimental studies show that surprisingly, hydrophobic ligands can occupy and use hydrophilic channels as well. In a previous computational study, MD simulations have revealed that water molecules occupy a Xe-binding site in myoglobin [189], and another study shows that the local minimum of 3D-PMF of water in myoglobin corresponds with the Xe-binding sites [57]. It is therefore interesting to note that Xe or $O_2$ displace water molecules under pressure in protein crystals [180, 190]. Indeed, another theoretical study on catalase using MD simulations found that $O_2$ and water share the same channel [191]. The presence of pathways in the vicinity of water pathway [6] and the proton-transport [188, 7, 192] pathway in [FeFe]-hydrogenase suggest the existence of possible common pathways between water and gas molecules to access the active-site, as also observed in [NiFe]-hydrogenase by Sumner et al. [193].

### 3.5.3 $O_2$ MFEPs Near the G Site and the Xe Site

The $O_2$ migration pathways in [FeFe]-hydrogenase show the areas within the protein through which the access of $O_2$ is most likely. Notably, there is only one pathway found

between the Xe-binding site and the G-site located within the central cavity. The residue A272 is found along the energy barrier of this pathway (Figure 3.4b), and its substitution is proposed in a patent application by King et al. [12] for increasing oxygen tolerance of [FeFe]-hydrogenase. Furthermore, this MFEP is in the proximity of areas of a key transition path to the G-site, as described by Kubas et al. [10], because they found that the mutation F417Y decreases the kinetic rate constant of $O_2$ transition by 100-fold along this path [10]. We observed an increase in the free-energy barrier of oxygen transition in this area by F417Y mutation (see Figure E.1).

### 3.5.4 $O_2$ MFEP to the Accessory Metal Cluster

The MFEPs of $O_2$ diffusion within the protein matrix core could reveal information about the mechanism of [FeFe]-hydrogenase active-site inhibition by $O_2$. The pathway 9-0 (Figure 3.2a) shows a migration path of $O_2$ from the $[Fe_4S_4]$ cluster to the H-cluster. This MFEP shows that the direction of $O_2$ diffusion path from cluster 1 to the G-site is thermodynamically more favorable, which correlates well with the experimental study by Stripp et al. [43]. It is suggested that $O_2$ reacting with the H-cluster produces reactive oxygen species (ROS), that either binds, presumably, at the distal Fe-site of the H-cluster [49] or migrates along a short-path to oxidize cluster 1 [43, 194]. Stripp et al. [43] suggested that ROS very likely is a superoxide that is more potent oxidant than $O_2$ with a higher energy to overcome the energy barriers to leave the H-cluster space for cluster 1 [43], and it is possible that superoxide could potentially migrate along this MFEP for $O_2$ between the G-site and cluster 1.

### 3.5.5 MFEPs and Experimental Understanding of $O_2$ Pathways

Many experimentally tested residues that are known to impact $O_2$ access to the H-cluster are located along our observed pathways (Figure 3.4a) [55, 11, 6]. Using the site-directed mutagenesis, Stapleton et al. [8, 9] found only residue I197 to measurably decrease $O_2$ inhibition of CpI [FeFe]-hydrogenase by a single mutation to V197 [55]. We find that I197

is located along pathways 9-15, 9-16, and 9-17 in the upper-side of the H-cluster that are connected to the local minimum 9 in the proximity of cluster 1, suggesting that the mutation of I197 is impacting the pathways forming a circulation network of $O_2$ on the upper side of the the H-cluster by limiting the access of $O_2$ to the active-site of the protein. Ghirardi et al. [195], using an *E. coli* expression system, made mutations (A321I and V423L) in CpI [FeFe]-hydrogenase that were effective in altering $O_2$-access to the H-cluster [11]. Pathways 3-2-1-0 and 3-2-0 are located in the proximity of the channel A [5] toward the H-cluster, and pathways 6-2-1-0, 6-4-1-0, and 6-2-0 are located in the proximity of the channel B [5] toward the H-cluster (Figures 3.2b). The residue A321 is located along the pathway 2-0 with an energy barrier of 7.29 kcal/mol, consistent with the work of Ghirardi et al. [11], and V423 is located near the pathway 1-0; thus either of these site mutations can impact pathways within the central cavity [5], where pathways from both channels merge. By looking at the full network of pathways, we can find pathways 1-5 and 20-2, that both have accessibility to the solvent other than pathways within the channels A and B. Therefore, the single mutations of V423 and A321 could affect multiple $O_2$ pathways to the H-cluster. In fact, such an unintended effect of these mutations is hypothesized by Ghirardi et al. [11], but is confirmed by our network of $O_2$ pathways within the protein.

The interconnected network of MFEPs shows that it is more likely for $O_2$ to diffuse along channel B than channel A. As there are three possible pathways from the channel B to the H-cluster, and 2 possible pathways from the channel A to the H-cluster, and the energy barriers of these 5 pathways show that the lowest overall energy barriers are along pathways in the channel B. Channel B is a dynamic pathway, first discovered by a theoretical study of $O_2$ diffusion in [FeFe]-hydrogenase [5], and given the presence of many pathways to the H-cluster from the channel B, residues in the channel B along the MFEPs, such as A321 and A564, are prime candidates for mutations (Figure 3.4b). However, a patent application protecting the redesign of [FeFe]-hydrogenase by incorporating bulky amino-acids suggested residues that are concentrated near the central cavity and the channel A [12], neglecting

residues in the channel B. This disparity is also noted by Lautier et al. [6]. Nevertheless, we find many residues along MFEPs, the mutations of which are suggested in the patent application by King et al. [12]: A272 is along the pathway 0-1, I276 is along the pathway 4-1, V423 is near the Xe-site, and A427 and F493 are along the pathway 2-1 toward the Xe-site (Figure 3.4a).

The presence of pathways 20-2-0 and 20-0 (within the water pathway [6]) to the H-cluster corroborates with the proposal [6] that $O_2$ is likely using this hydrophilic channel to access the active-site of [FeFe]-hydrogenase (Figures 3.2b). Although the local minimum 20, close to K571, is located within the water channel, we note that only pathway 20-2 is along the suggested water pathway, and the pathway 2-0 is located in the central cavity. The pathway 20-0 is partially located in an area between the channel B [5] and the water pathway [6]. The pathway 21-0 is perhaps the easiest access route to the H-cluster for $O_2$ located between the channel B [5] and the water pathway [6].

### 3.5.6 Potential Candidate Residues for Future Mutations along $O_2$ MFEPs

The proximity of the pathway 21-0 to these channels might cause the mutations of residues (in the channel B and the water channel) to affect this pathway indirectly, i.e., allosterically. The energy barrier of this pathway is located close to the surface-exposed F570, making this residue a good candidate for mutations to disrupt the $O_2$ diffusion along this pathway. Similarly, $O_2$ diffusion along pathway 13-5-1-0 could be hindered by mutations of residues along energy barriers of pathways 5-1 and 1-0. In a similar fashion, for pathway 10-9-0, the residues close to energy barriers are suitable choices for mutations. All such candidate residues are shown in Figure 3.4b. Our work suggests the existence of multiple pathways for $O_2$ to access the H-cluster beyond previously known channels A and B, such as the presence of pathways in the vicinity of the water pathway [6] and the proton transport [188, 7], and pathways in other areas of the active-site domain. The construction of a systematic network of migration pathways of $O_2$ along MFEPs provides unique physical insight into the complex

nature of ligand diffusion in [FeFe]-hydrogenase, and shows that it is highly informative for the purpose of identifying candidate residues (Figures 3.4b) for future experimental studies by mutagenesis or other methods.

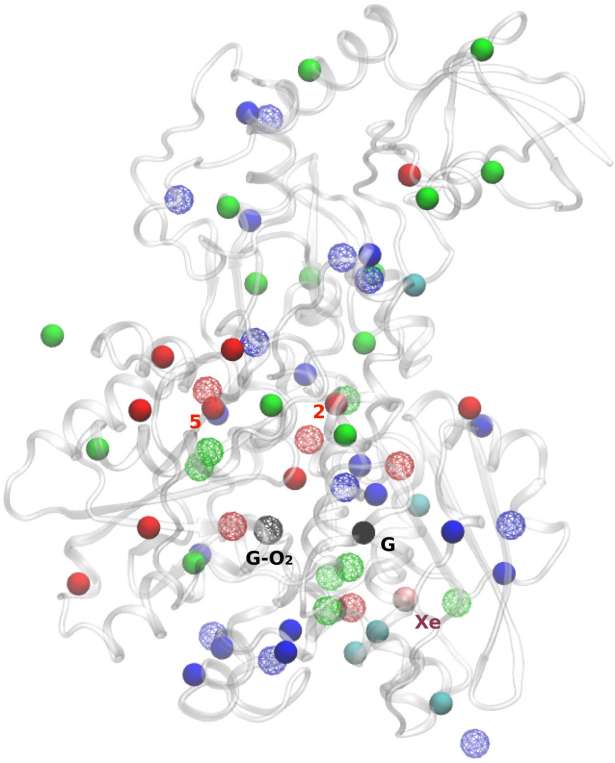



Figure 3.5: Locations of the minima for CO (filled spheres) and $O_2$ (wireframe spheres) diffusion in the CpI FeFe-hydrogenase are shown. The global minimum for both gases are shown in black filled or wireframe spheres and are labeled (G for CO and G-$O_2$ for $O_2$). The Xenon cavity is shown as a pink sphere and labeled as Xe. For CO, colors of other minima are consistent with Fig. 3.3: red (0-4 kcal/mol), green (4-8 kcal/mol), blue (8-12 kcal/mol), and cyan (over 12 kcal/mol). For $O_2$, colors of minima indicate the following energy ranges: red (0-12 kcal/mol), green (12-16 kcal/mol), and blue (over 16 kcal/mol) (This figure was created by my co-author Yong Liu [3]).

### 3.5.7 CO MFEPs to the active site

The MFEPs of CO diffusion from other minima to the local minimum 7 could potentially disclose the mechanism of CO access to the active site. In Fig. 3.3, we show the MFEPs from 44 distinct minima to the local minimum 7 as well as the locations of major barriers along MFEPs. The major energy barriers range between 2.45 kcal/mol and 16.68 kcal/mol,

among which the the pathways 2-7 and 41-7 have the lowest free-energy barriers and the pathway 28-7 has the highest free-energy barrier.

Among thermodynamically favored local minima at the surface of the enzyme (those with free-energy values below 4 kcal/mol relative to the global minimum), the pathway 10-2-7 has the lowest energy barrier of 4.15 kcal/mol, although the pathway 4-5-7 also has a comparable energy barrier of 5.23 kcal/mol. However, the exit route for CO from the H-cluster requires overcoming a barrier of at least 5.04 kcal/mol, which is the energy barrier along the reverse MFEP 4-5-7. The lower energy barriers along the MFEP 4-5-7 for CO entry/exit from the enzyme make this pathway a highly likely route for CO diffusion between the solvent and the H-cluster.

### 3.5.8 Comparisons between CO and $O_2$ diffusion in the enzyme matrix

In our study of $O_2$ diffusion in the CpI FeFe-hydrogenase, we identified 23 minima including the global minimum, which are shown in Fig. 3.5 in a wireframe representation [1]. We observe that two local minima of $O_2$ partially overlapped with the local minima 2 and 5 of CO diffusion, indicating that CO and $O_2$ may reside within the enzyme matrix at shared locations. The global minimum for both inhibitory gases CO and $O_2$ are located near the active site, which suggests a high-affinity for both gases surrounding the active site. The global minimum for CO is located near the residue F417, where F417Y mutation was previously shown to decrease the kinetic rate constant for $O_2$ diffusion [10, 1].

Although some common features can be found, CO diffusion and $O_2$ diffusion in the CpI FeFe-hydrogenase differ in many aspects. For example, no local minimum for $O_2$ is found in the immediate neighborhood of the H-cluster, but for CO, the local minimum 7 is situated at a distance of 2.75 Å from the H-cluster. This suggests that CO likely has easier and faster access to the active site than $O_2$, as has been previously suggested [196, 175]. This is further reinforced by the fact that, on comparing the 3D-FES for CO in this work with that of $O_2$ from our previous work [1], we find overall lower free-energy barriers for CO diffusion (with

respect to the global minimum for CO) than for $O_2$ diffusion (with respect to the global minimum for $O_2$), even though the free energy difference ($\Delta$F) between the global minimum and the solvent space for each gas are comparable, 18.81 kcal/mol (CO) and 18.5 kcal/mol ($O_2$).

However, in the regions of the enzyme with two hydrophobic pathways, we find that $O_2$ diffusion is thermodynamically favored over CO diffusion, because five local minima (four shown in green wireframe spheres and one shown as a red wireframe sphere near the Xe cavity in Fig. 3.5) with relatively lower free-energy values of $O_2$ diffusion are located along hydrophobic pathways.

Even though there exists a local minimum in the vicinity of the Xe cavity for both CO and $O_2$, the Xe cavity is located in the lower free-energy region of the 3D-FES for $O_2$ while CO diffusion in the vicinity of the Xe cavity requires overcoming higher free-energy barriers. The Xe cavity is surrounded by three hydrophobic residues (F493, A427, and A431) which may stabilize apolar gases over polar gases [176], indicating that $O_2$ is more likely to be found in the Xe cavity. These differences can also be explained by the observation that CO is a polar molecule with a van der Waals volume of 16.20 $cm^3$/mol, while $O_2$ is an apolar molecule with a van der Waals volume of 13.00 $cm^3$/mol [197]. The effect of the size of gases on diffusion pathways has been demonstrated in previous MD simulations of $H_2$ and $O_2$ diffusion in the CpI FeFe-hydrogenase, where $H_2$ with a smaller van der Waals radius was shown to diffuse in a broader region of the protein matrix and on shorter timescales [198]. Furthermore, CO was shown as a better ligand for binding to electron-rich metals compared with $O_2$ [199]. This is consistent with the observation that CO reacts much faster than $O_2$ with the FeFe-hydrogenase based on electrochemical studies [175, 174, 196, 200].

### 3.5.9 Candidate residues for mutations

To prevent inhibitors (e.g. CO and $O_2$) from binding to the $Fe_d$ of the 2Fe subcluster of the H-cluster, mutagenesis of protein residues is an approach to alter diffusion pathways.

Based on the pathways and thermodynamics of CO diffusion, we propose key candidate residues for mutations to disrupt or block the diffusion of CO (Table 3.1). These 19 residues are located in the vicinity of the energy barriers of the targeted MFEPs for CO diffusion. For example, the residue S323 is located near the energy barrier for the pathway 39-7 and is along the previously proposed water pathway [176]. Similarly, the residue K358 is in the proximity of the energy barrier of the pathway 5-7, and is also one of the residues defining the water-based proton-transfer pathway [1, 188, 7]. We hypothesize that replacing this residue with mutants having larger side-chains could disrupt the CO diffusion along multiple MFEPs ( 5-7, 4-7, 19-7, and 28-7). The presence of MFEPs for CO in the vicinity of the proton-transport pathway [192, 7, 188] as well as the water pathway [176] also suggests the existence of shared pathways between water and gas molecules [193, 57].

Among the proposed 19 residues to disrupt CO diffusion within the enzyme, 11 have been experimentally studied or computationally proposed (marked as ✓in Table 3.1) to decrease the rate of $O_2$ diffusion inside the CpI FeFe-hydrogenase [8, 11, 1]. For example, both I197 [9] and V423 [11] have been experimentally studied to alter the diffusion rate of $O_2$. The other 9 residues have been proposed to block $O_2$ diffusion [1], among which A228, M295, G302, V352, and K358 are located in the proximity of energy barriers for $O_2$ diffusion. This result indicates that mutagenesis of common residues may decrease the diffusion rates of CO and $O_2$ in the CpI FeFe- hydrogenase. We speculate that mutations of I416, V423, and A426 may have a weaker effect on decreasing the rate of CO diffusion because these three residues are located in the enzyme region with higher free-energy values for CO diffusion.

## 3.6 Conclusion

In this work, [FeFe]-hydrogenase is investigated using the single-sweep/string method of Maragliano et al. [56] to map the migration pathways of $O_2$ and CO diffusion in this enzyme. In the reconstructed free-energy map, local minima are notably found near the active-site,

the H-cluster, and a Xe-binding site. By assuming that the diffusion of $O_2$ and CO inside this protein is following the pathways of minimal free-energy, we found that there are multiple interconnected $O_2$ and CO migration pathways from the solvent to the H-cluster. The overall network of pathways shows that $O_2$ and CO transport in [FeFe]-hydrogenase is not limited to previously known pathways [5, 6], and $O_2$ is able to access the H-cluster for inactivation via several alternative routes. Our results provide an enhanced view of gas diffusion in [FeFe]-hydrogenase and suggest new residues (Figure 3.4b), mutations of which could increase oxygen tolerance. Furthermore, 11 of the proposed residues were also found in CO studies. These findings will guide future experimental studies to increase CO and $O_2$ tolerance of hydrogenases, as recently the more tolerant mutant of [FeFe]-hydrogenase found by mutation of residue found in the neighborhood of proposed residues [201].

## 3.7 Publications

The work described in this chapter has resulted in the following journal articles:

- Mohammadi, M., Vashisth, H. (2017). Pathways and Thermodynamics of Oxygen Diffusion in [FeFe]-Hydrogenase. J. Phys. Chem. B, 121(43), 10007-10017.
- Liu, Y., Mohammadi, M., Vashisth, H. (2018). Diffusion network of CO in FeFe-Hydrogenase. J. Chem. Phys., 149(20), 204108.

# CHAPTER 4

# STUDIES ON TDZD (ALIPHATIC VS. AROMATIC) INHIBITORS OF RGS4

## 4.1 Abstract

RGS proteins play a pivotal role in regulation of GPCR signaling and are therefore becoming an increasingly important therapeutic target. Recently discovered thiadiazolidinone (TDZD) compounds that target cysteine residues have shown different levels of specificities and potencies for the RGS4 protein, thereby suggesting intrinsic differences in dynamics of this protein upon binding of these compounds. In this work, we investigated the effect of binding of several small-molecule inhibitors on dynamical motions in RGS4. Specifically, we studied two conformational models of RGS4 in which a buried cysteine residue is solvent-exposed due to side-chain motions or due to flexibility in neighboring helices. We found that TDZD compounds with aromatic functional groups perturb the RGS4 structure more than compounds with aliphatic functional groups. Moreover, small-molecules with aromatic functional groups but lacking sulfur atoms only transiently reside within the protein and spontaneously dissociate to the solvent.

## 4.2 Background

We have previously reported [102] an open-state model of RGS4 (shown as Model 1 in Fig. 4.1d) which is conformationally different from apo-RGS4 (shown as Model 1 in Fig. 4.1c). In this model, the flexibility in the $\alpha$5-$\alpha$6 helical pair facilitates access to the buried C95 residue for covalent-docking of the inhibitor CCG-50014 (compound **1** in Fig. 4.1a). On

inhibitor binding, we found that the the $\alpha$5-$\alpha$6 helical pair remains perturbed and only partially relaxes toward the closed conformation of these helices in the apo-RGS4 structure. However, in the absence of inhibitor we observed that the protein largely reverts to a confor-

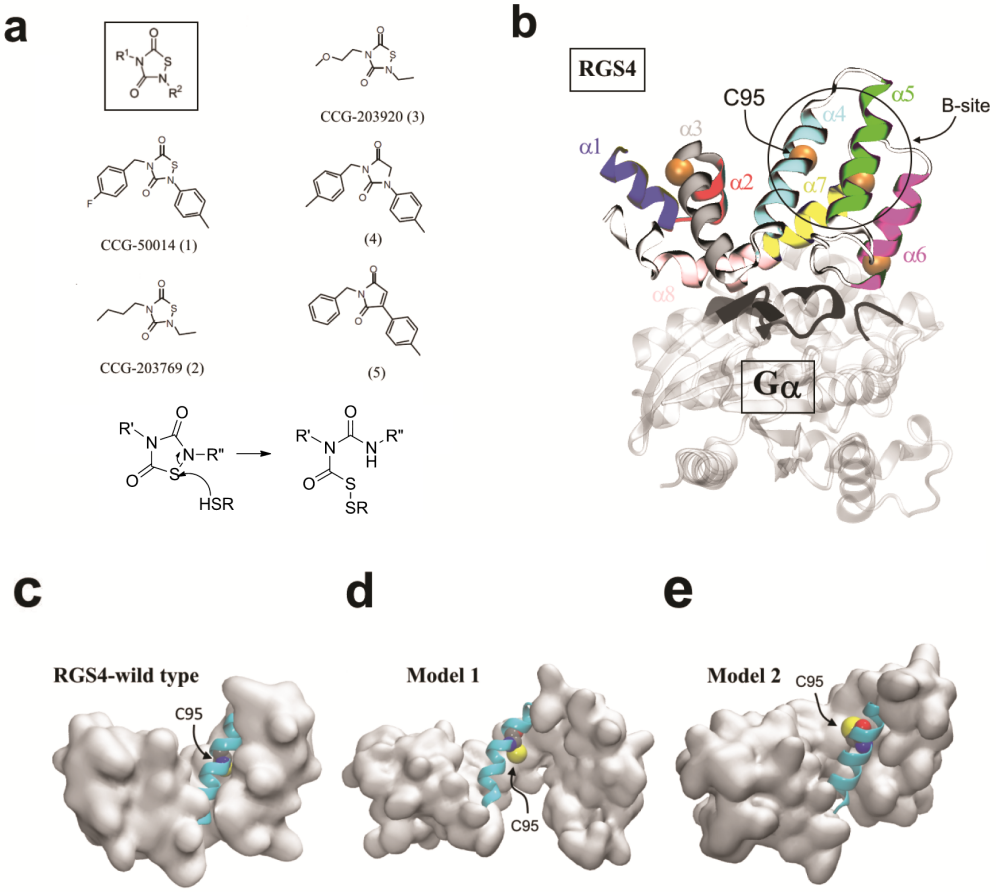



**Figure 4.1.** (a) Small-molecule structure with $R^1$ and $R^2$ functional groups (inset in box) along with the chemical structures of all small-molecules studied and the mechanism of reaction to form an adduct with a cysteine residue on the RGS4 [4]. (b) A cartoon representation of the RGS4-G$\alpha_{i1}$ complex (PDB code 1AGR) is shown. Each of the 9 $\alpha$-helices of RGS4 is colored and labeled, and the location of four cysteine residues are shown by their $C_\alpha$-atoms as orange spheres with the C95 residue labeled. The G$\alpha_{i1}$-subunit is shown in transparent white ribbons, and the loops of G$\alpha_{i1}$ in the proximity of RGS4 are highlighted in black ribbons. (c, d, e) Cartoon representations of RGS4 conformations are shown for the wild-type apo-RGS4, and in its conformationally changed models (Models 1 and 2), respectively. Highlighted in cartoon representations are $\alpha$4-helices as cyan cartoons along with the C95 residues as space-filling. The residue C95 is buried in the wild-type RGS4 structure but it is accessible in Models 1 and 2. For each model, the structure of RGS4 (except the $\alpha$4-helix) is rendered in a white surface representation.

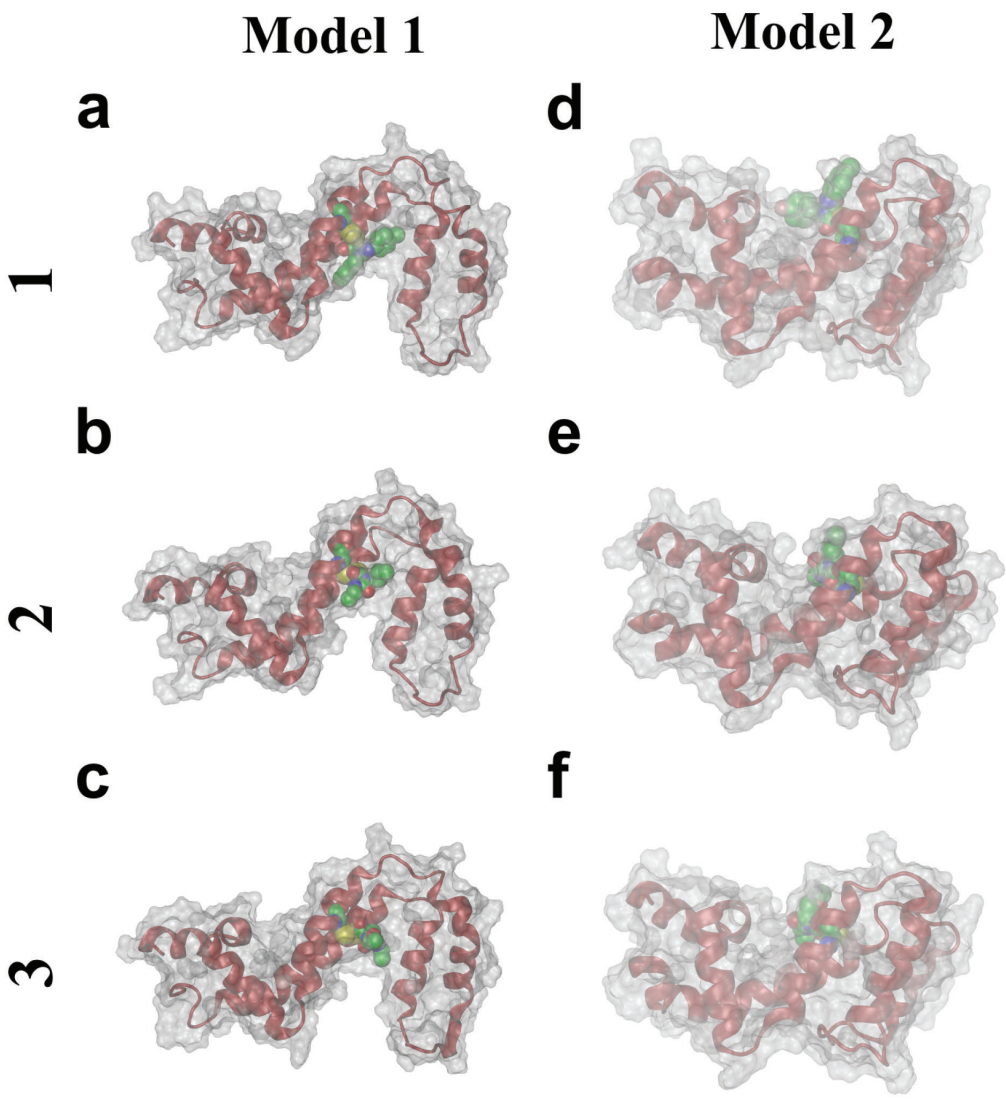



**Figure 4.2.** Docked initial conformations of TDZD compounds **1**, **2**, and **3** in Model 1 (panels a, b, and c) and Model 2 (panels d, e, and f) are shown. In all snapshots, the protein backbone is shown in red ribbons as well as white transparent surfaces, while compounds, along with the cysteine residue C95, are shown in green space-filling representations.

mation similar to apo-RGS4. While we conducted only short time-scale (∼40 ns) molecular dynamics (MD) simulations in that work, parallel NMR HSQC results [15] identified several significant perturbations in residues surrounding the inhibitor binding site. Given the short time-scale of our earlier simulations, it is unclear to what extent the $\alpha$5-$\alpha$6 helical pair will remain perturbed by bound ligand in longer time-scale simulations. Moreover, differences

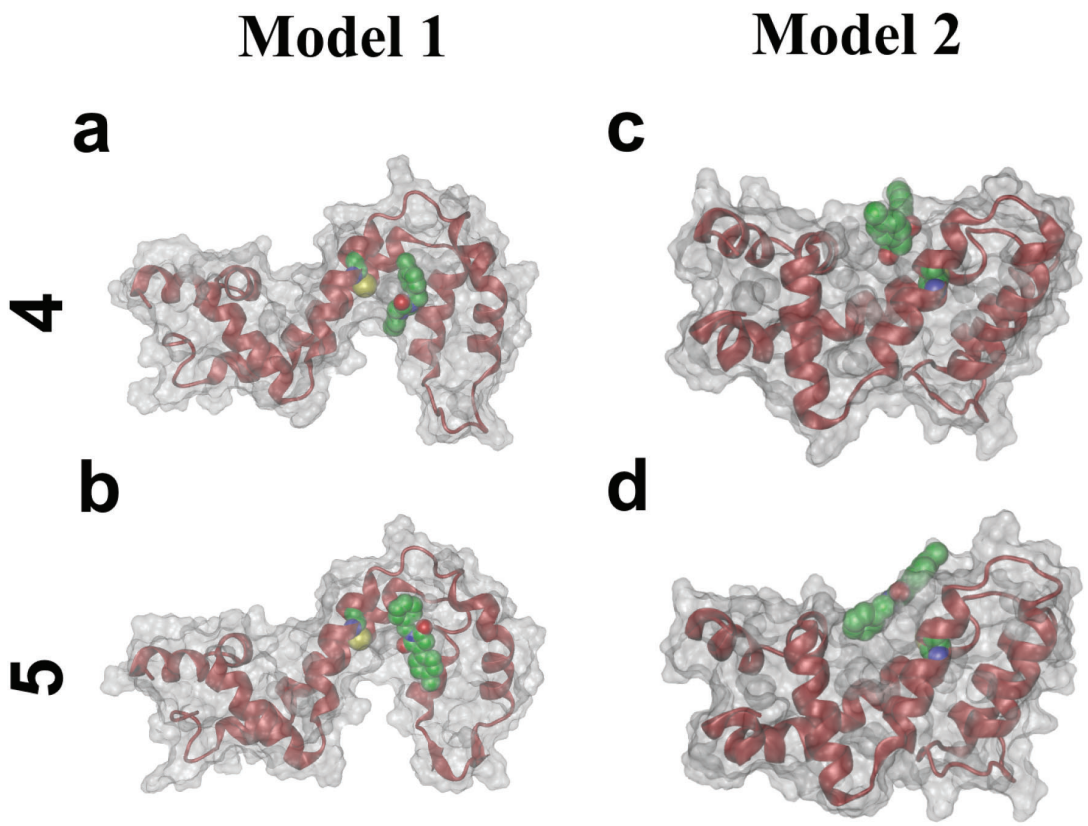



**Figure 4.3.** Docked initial conformations of non-TDZD compounds **4** and **5** in Model 1 (panels a and b) and Model 2 (panels c and d) are shown. Coloring and labeling schemes are identical to Fig. 4.2.

in the degree of helical perturbations by CCG-50014 and its congeners with smaller side-chains [4] (Fig. 4.1a) remain unknown as those small-molecules were not studied previously in the context of our open-state model (Model 1).

Furthermore, using $\mu$s time-scale MD simulations combined with hydrogen-deuterium exchange (HDX) studies of three RGS proteins (RGS4 and its homologues RGS8 and RGS19) [13, 202], we not only observed signatures of flexibility and partial unfolding in helices, but also found that the side-chains of buried cysteine residues (C95 in RGS4) are transiently exposed to solvent while apo proteins largely maintain a closed conformation. In our previous work [13], we did not carry out simulation studies of RGS proteins with inhibitors, but we

hypothesized that inhibitors could potentially access the otherwise buried C95 residue in the cysteine-exposed closed conformation of RGS4 (Model 2 in Fig. 4.1e).

**Hypothesis:** Binding of aromatic and aliphatic TDZD analogues to buried and conserved cysteine residue results in allosteric perturbations.

To determine whether ligand access to the buried cysteine residue (C95) was through large conformational changes (Model 1) or through small local fluctuations in the side-chain of C95 (Model 2), we here performed enhanced MD investigations of the mechanisms of interactions of 5 small-molecules (3 TDZD compounds with aromatic/aliphatic functional groups and 2 non-TDZD analogues lacking sulfur atoms; Fig. 4.1a) with the C95 residue of RGS4. We further performed a flow cytometry-based assay [72] to measure concentrations of three TDZD compounds needed to inhibit the RGS4/G$\alpha$ protein-protein interaction (specifically, the interaction between G$\alpha$ and an RGS4 construct containing only a single-cysteine residue, RGS4 C95). Taken together, these studies highlight the role of cysteine exposure and global protein dynamics in recognition of TDZD compounds by RGS4, and suggest new venues for designing non-covalent small-molecules targeting RGS proteins.

## 4.3 Methods

### MD Simulations

We performed two sets of classical all-atom and explicit-solvent MD simulations for RGS4/small-molecule complexes using the NAMD software [203] and the CHARMM force-field with the CMAP correction [122, 123] (section 2.2). We used VMD for system creation and post-simulation analysis [204]. All small-molecules used in this work are reported in Fig. 4.1a, and their force-fields were parameterized using the Multipurpose Atom-Typer for CHARMM (MATCH) tool [125]. The first set of simulations was for studying small-

molecule analogues in complex with an open-state conformation of RGS4 (Model 1; Fig. 4.1d) [102], and the second set of simulations was for small-molecule analogues in complex with a closed-state conformation of RGS4 (Model 2; Fig. 4.1e) [13].

For simulations of Model 1, the open-state conformation of RGS4 that was reported in our earlier work [102] was used here as an initial state for docking of small-molecules. For specifically understanding the effect of binding of thiadiazolidinone (TDZD) and non-TDZD compounds in a binding pocket in the proximity of the cysteine residue C95 (Fig. 4.1b), a single-cysteine mutant of RGS4 was created, where all cysteine residues (three in total) except C95 were mutated to Ala. Following our previous protocol [102], the open-state con-

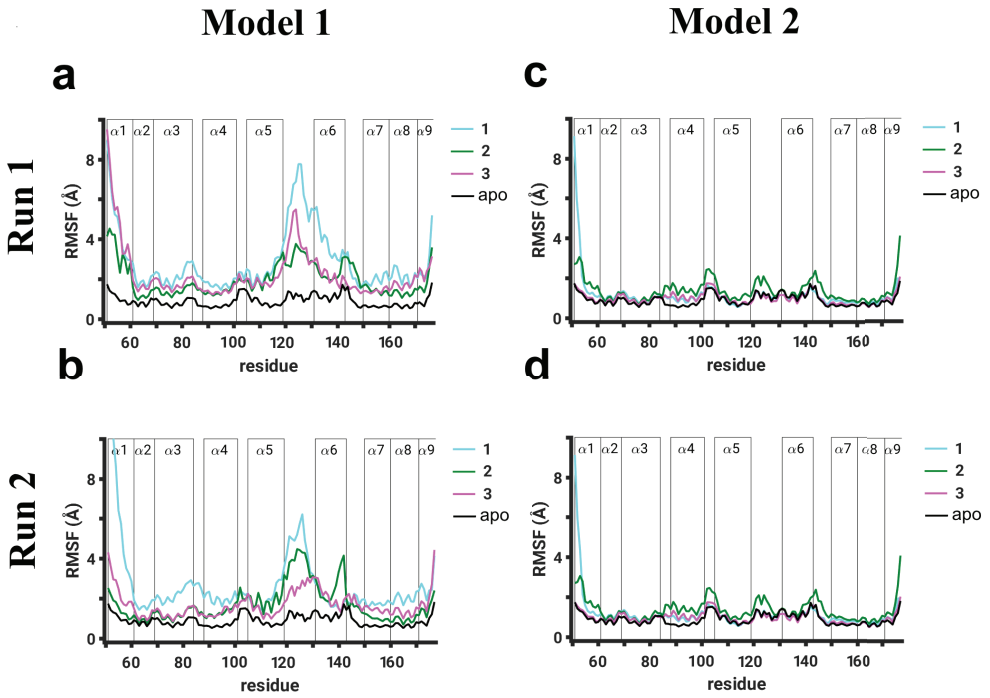



**Figure 4.4.** Root-mean-squared-fluctuation (RMSF) per residue are shown for Model 1 (panels a and b) and Model 2 (panels c and d) of RGS4. The RMSF values are reported from two independent 1 $\mu$s long simulation runs (Run 1 and Run 2) for each model, where simulations were conducted with TDZD analogues (compounds **1**, **2**, and **3** in Fig. 4.1a) covalently-bound to the C95 residue of RGS4. As a baseline reference, the RMSF values of the RGS4 structure without any compound (apo-form; black traces) are also shown from our previous work [13]. The vertical bars labeled $\alpha$1 through $\alpha$9 demarcate the locations of residues in 9 $\alpha$-helices of RGS4.

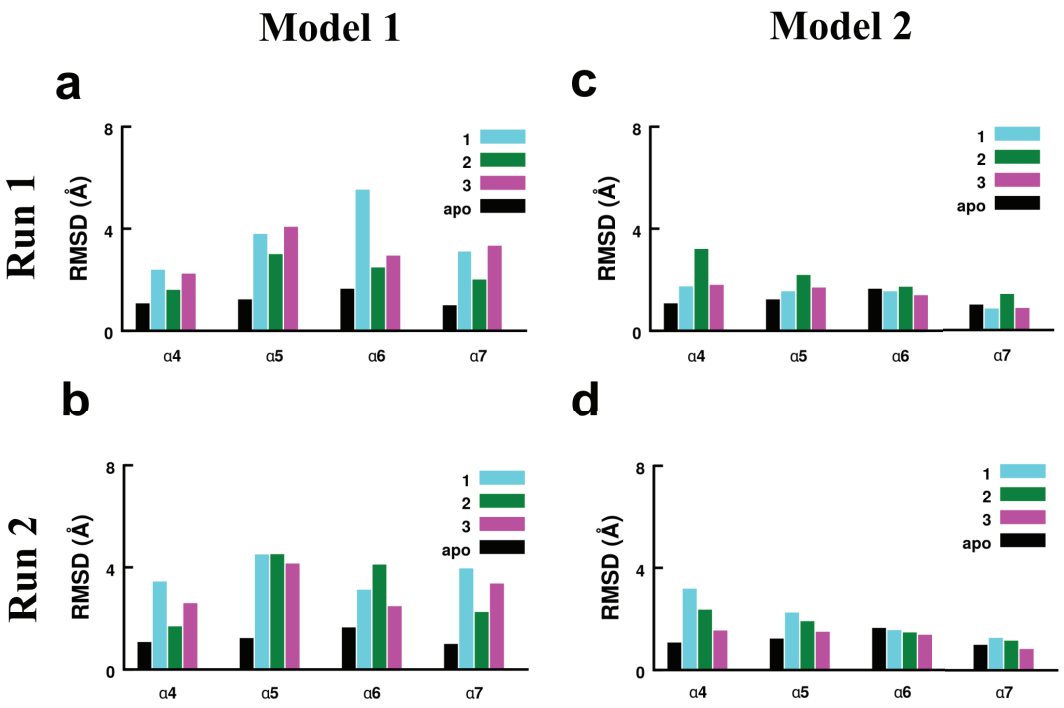



**Figure 4.5.** The histograms of RMSD-averages computed based upon Models 1 and 2 are shown.Panels a and b show data from two independent runs of Model 1, and panels c and d show data from two independent simulations of Model 2.

formation of mutated RGS4 was then used to create 5 docked complexes with 5 compounds (Fig. 4.1), where each TDZD compound is covalently bound to C95 and non-TDZD compounds are non-covalently docked in the same pocket where TDZD compounds are covalently docked (Fig. 4.2a, b, c, and Fig. 4.3a, b). For simulations of Model 2, a protocol similar to Model 1 was followed where the initial state of RGS4 was a closed-conformation in which the key cysteine residue (C95) is surface-exposed (Fig. 4.1e) per our earlier work [13]. The docking of compounds in Model 2 was further facilitated by the Internal Coordinate Mechanics (ICM) software [113] to obtain their energetically favorable conformations (Fig. 4.2d, e, f, and Fig. 4.3c, d).

We solvated all systems for Model 1 and Model 2 using explicit TIP3P water molecules [205], ionized using NaCl, and added all hydrogen atoms. Each solvated and ionized system was

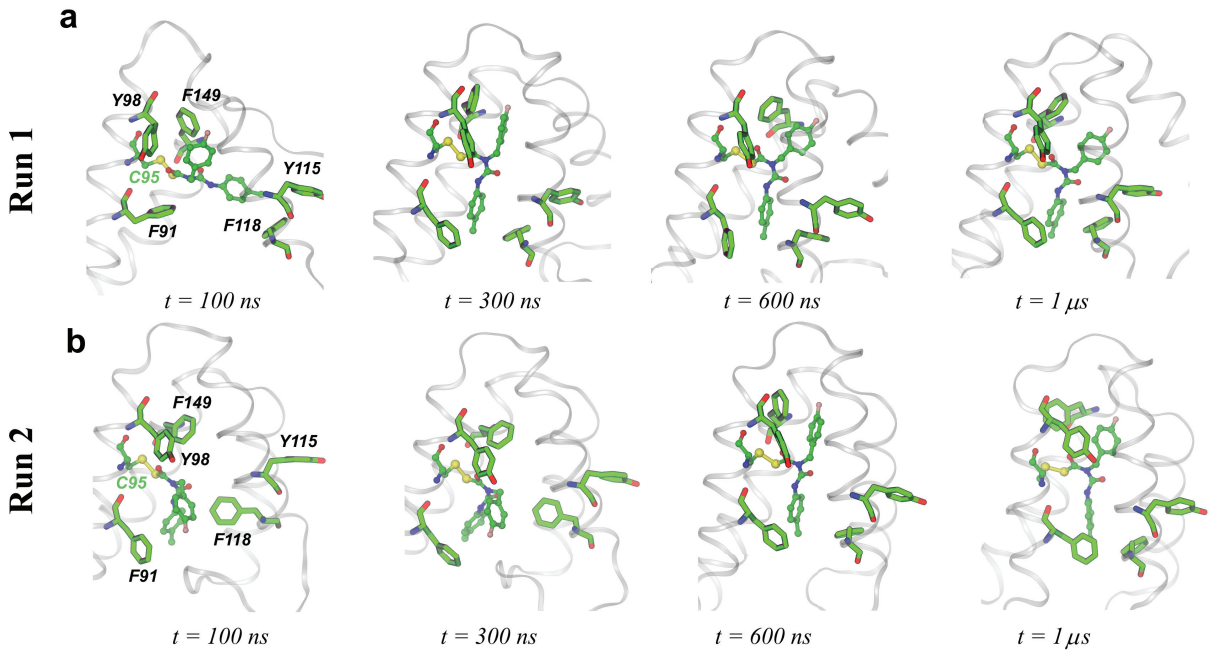



**Figure 4.6.** The side-chains of aromatic residues in the vicinity of covalently-docked compound **1** are shown at various time-points from two independent simulations of Model 1 (panels a and b). The compound **1** is covalently-linked to residue C95, and neighboring residues are labeled and shown in green sticks. The protein backbone in all snapshots is shown in a white transparent cartoon.

then energy minimized for ∼500-1000 cycles via conjugate-gradient optimization, and equilibrated via MD simulations, conducted with a time-step ($\Delta$t) of 2-fs, for 1 $\mu$s in the NPT ensemble where the Langevin thermostat with a damping coefficient of 5 $ps^{-1}$ was used for temperature control and the Nosé-Hoover barostat was used for pressure control. Periodic boundary conditions were used throughout, non-bonded interactions were accounted with a cut-off of 10 Å where smooth switching was initiated at 8 Å, and long-range electrostatic interactions were handled using the Particle Mesh Ewald (PME) method. For each system, two (for TDZD compounds) or three (for non-TDZD compounds) independent MD simulations were carried out. In addition to classical MD simulations, we used metadynamics as an enhanced sampling method (section 2.4.2) to compare thermodynamics of conformational changes between open and closed-states in apo-RGS4 and RGS4/small-molecule complexes.

For metadynamics simulations, we used an eigenvector as a CV [206] that was computed based upon the atomic coordinates of the backbone $C_\alpha$ atoms of all residues in $\alpha_4$ through $\alpha_7$ helices. The eigenvector choice of a CV is a projection of the coordinates of a group of atoms (or more precisely, their deviations from the reference coordinates) onto a linear transformation between two end-states: open and closed conformations of RGS4. The vector is normalized, therefore the CV is 0 when the backbone $C_\alpha$ atoms are at the coordinates of the closed conformation of RGS4 and the CV is 1 when they are at the coordinates of the open conformation of RGS4. All metadynamics simulations were carried out with a time-step of 2-fs, and with $W$, $\delta$, and $\tau_G$ values of 0.05 kcal/mol, 0.025 Å, and 2 ps, respectively. Overall, we carried out 7 metadynamics simulations: one simulation for apo-RGS4 ($\sim$0.25 $\mu$s long), and 3 simulations ($\sim$0.15$-$0.45 $\mu$s long) each for Models 1 and 2 when bound to compounds **1**, **2**, and **3** (Fig. 4.1a), respectively.

## 4.4 Results and Discussion

To understand the effect of binding of small-molecules in each conformational model of RGS4 (Model 1 and Model 2; Fig. 4.1d,e), we docked TDZD as well as non-TDZD compounds (Fig. 4.2 and Fig. 4.3) in binding pockets created near the C95 residue and conducted several independent $\mu$s time-scale MD simulations.The initial conformational state of RGS4 for docking was chosen to either represent flexibility in helices (open-state; Model 1) [102] or cysteine-exposure (closed-state; Model 2) [13]. For each model, we performed two sets of independent simulations (each 1 $\mu$s long) for TDZD congeners having aromatic and aliphatic side-chains (compounds **1**, **2**, and **3** in Fig. 4.1a) covalently-docked at the residue C95 on the $\alpha_4$ helix (highlighted in cyan in Fig. 4.1d,e). We first describe results on various conformational metrics used to characterize the subsequent structural perturbations by three covalently-linking TDZD compounds in each model of RGS4.

Additionally, for two non-covalent TDZD analogues lacking sulfur atoms (compounds **4** and **5** in Fig. 4.1a), we performed three sets of independent simulations (each 1 $\mu$s long) for Model 1 where we hypothesized that non-covalently-docked compounds can transiently reside within the protein due to local interactions. While for Model 2, where compounds **4** and **5** are non-covalently-docked on the protein surface near the exposed cysteine residue C95, our preliminary simulation analyses revealed that compounds **4** and **5** can quickly and spontaneously diffuse into the solvent (*vide infra*). Therefore, we did not perform long time-scale simulations of these two non-TDZD compounds for Model 2. In the following, we describe various conformational metrics used to characterize the subsequent structural perturbations in each model of RGS4.

We performed several $\mu$s-timescale unbiased MD simulations of RGS4 in the open-state conformational model (Model 1) with three covalently-docked (at C95) TDZD congeners (having aliphatic and aromatic side-chains), and two non-TDZD analogues lacking sulfur atoms (Fig. 4.1a). As shown in Fig. 4.1b, the residue C95 is located on the $\alpha_4$ helix (highlighted in cyan) surrounded by $\alpha_5$-$\alpha_7$ helices (highlighted in green, magenta, and yellow, respectively), making it completely buried in the crystal structure of RGS4, but mostly accessible in the open-state model of RGS4 (Model 1 in Fig. 4.1c).

### 4.4.1 Simulations of covalently-bound TDZD inhibitors

**Root-mean-squared-fluctuation/deviation (RMSF/RMSD) analyses:** To resolve residue-level perturbations on binding of TDZD compounds (**1**, **2**, and **3**) in each model of RGS4, we calculated RMSF per residue from two independent sets of simulations for each model (cyan, green, and magenta traces in Fig. 4.4) and compared these values with RMSF per residue values of apo-RGS4 from our previous work (black traces in Fig. 4.4) [13]. In addition to higher fluctuations expected in free terminal helices ($\alpha_1$ and $\alpha_9$), we observed in both simulations of Model 1 (panels a and b in Fig. 4.4) that all compounds induced significant perturbations ($\sim$4-8 Å higher than in apo-RGS4) in helices $\alpha_5$ and $\alpha_6$, and in

the $\alpha_5$-$\alpha_6$ interhelical loop which directly contacts the G$\alpha$-subunit in the RGS4/G$\alpha$ complex (Fig. 4.1b). Moreover, the RMSF values show that compound **1** (with aromatic functional groups) perturbs the $\alpha_5$-$\alpha_6$ interhelical loop more than compounds **2** and **3** (with aliphatic functional groups) (cyan *vs.* green and magenta traces in Fig. 4.4a,b). The perturbations in this interhelical loop also appear to propagate to structural motifs flanking this loop, namely the C-terminus of the $\alpha_5$ helix and/or the N-terminus of the $\alpha_6$ helix. For compound **1** in comparison to compounds **2** and **3**, we also observed marginally higher perturbations in the $\alpha_3$-$\alpha_4$ interhelical loop, another structural region of RGS4 that is known to directly contact the G$\alpha$-subunit in the RGS4/G$\alpha$ complex (Fig. 4.1b). However, in Model 2 (panels c and d in Fig. 4.4), where all three compounds are covalently-docked on the protein surface (Fig. 4.2), we observed no significant perturbations in any structural motif of RGS4 as the RMSF values are comparable to apo-RGS4.

We further report that the $C_\alpha$-RMSD traces for helices $\alpha_4$ through $\alpha_7$ (Figs. F.1 and F.2) and their average RMSD from both simulations (Fig. 4.5), measured relative to the crystallographic conformation of apo-RGS4 (PDB: 1AGR), highlight that all compounds induce greater perturbations when covalently-docked in Model 1 in comparison to Model 2. Also, compounds with aromatic functional groups (compound **1**) induce larger perturbations than compounds with aliphatic functional groups (compound **2** and **3**). The fluctuations of residues located at the interaction site with the G$\alpha$ subunit was shown in our previous work to result in significant weakening of the RGS/G$\alpha$ protein-protein interaction in the presence of compound **1** (CCG-50014) [102]. As also observed here, this was largely attributed to structural rearrangements of the $\alpha_5$-$\alpha_6$ helical pair and the loop connecting them. Specifically, we observed significant allosteric perturbations in two residues (T124 and E126) of the $\alpha_5$-$\alpha_6$ interhelical loop in which we previously reported perturbations on binding of compound **1** based upon NMR HSQC data [102]. We also observed ring-ring interactions (Fig. 4.6) between compound **1** and the side-chains of neighboring aromatic residues some

of which (e.g. F91) were reported as highly perturbed on binding of compound **1** in our previous work [102].

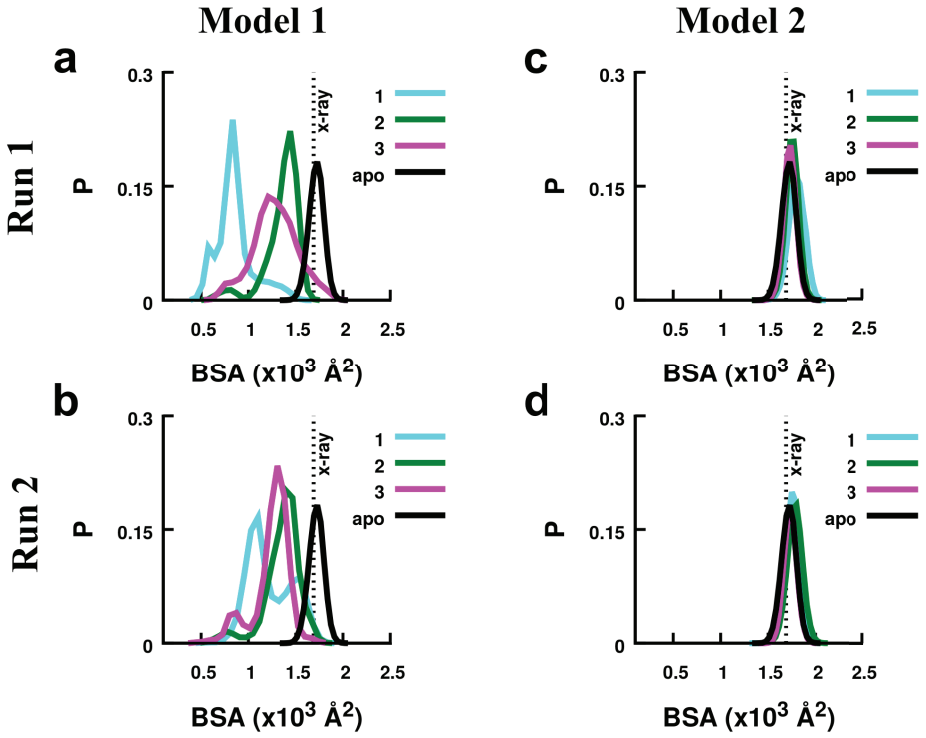



**Figure 4.7.** The histograms of the buried surface area (BSA) between the $\alpha_5$-$\alpha_6$ helical pair and the rest of RGS4 are shown for Model 1 (panels a and b) and Model 2 (panels c and d). Data are shown for simulations of each model conducted with TDZD congeners (compounds **1**, **2**, and **3**). The vertical dotted lines in panels indicate the values of BSA in the RGS4 crystal structure (PDB: 1AGR). The BSA traces for apo-RGS4 computed from a simulation reported in our previous work [13] are also shown (black traces).

**Buried Surface Area (BSA) analysis:** To further investigate conformational changes on binding of compounds in each model of RGS4, we measured the BSA between the $\alpha_5$-$\alpha_6$ helical pair and the rest of the RGS4 structure from two independent simulations (Fig. 4.7). In simulations of Model 1 (Fig. 4.7 a, b), we observed that compound **1** produces a greater shift in the peaks of the BSA-histograms than do compounds **2** and **3** compared to the BSA value in the crystal structure (vertical dotted lines labeled x-ray in Fig. 4.7) or in the apo-RGS4 simulation (black traces in Fig. 4.7 a, b). The BSA traces for all compounds in

Model 1show deviations from the BSA values in the crystallographic or apo-RGS4 simulation indicating that the $\alpha_5$ and $\alpha_6$ helices only partially relax toward their closed conformation in the crystal structure. In both simulations of compound **2** with aliphatic functional groups, the BSA tracesshow a gradual increase in the BSA toward crystallographic values indicating a nearly complete closure of the $\alpha_5$-$\alpha_6$ helical pair. In simulations of Model 2, the peaks of the BSA-histograms (Fig. 4.7 c, d) and the BSA traces show no significant deviation from the crystal structure values, thereby indicating a closed conformation of RGS4 when bound to TDZD compounds. Consistent with RMSD/RMSF trends, these results show that the binding of compounds in Model 1 perturbs the RGS4 structure significantly more than their binding in Model 2. Furthermore, in Model 1, the $\alpha_5$-$\alpha_6$ helical pair, which is critical for RGS/G$\alpha$ binding, only partially relaxes toward the crystallographic conformation in simulations of compound **1**, but significantly recovers in 1 $\mu$s simulations of compound **2**. This suggests that the compounds with aliphatic functional groups (compounds **2** and **3**), when covalently-docked within the $\alpha_4$-$\alpha_7$ helical bundle, are more easily accommodated than those with aromatic functional groups (compound **1**).

**Salt-bridging interactions:** The $\alpha_4$-$\alpha_7$ helical bundle in RGS4 has several charged amino-acids (K, R, D, and E) that likely form stable or intermittent salt-bridges due to electrostatic interactions. To understand the ability of TDZD compounds to perturb interactions between charged residues and thereby between helices, we investigated potential perturbations in several salt-bridge forming residue pairs: D90-K125 ($\alpha_4$-$\alpha_5$), E97-K110 ($\alpha_4$-$\alpha_5$), K99-D150 ($\alpha_4$-$\alpha_7$), E126-R134 ($\alpha_5$-$\alpha_6$), and D130-K155 ($\alpha_6$-$\alpha_7$). In Fig. 4.8, we show histograms of average distances between the center-of-mass of these residue pairs from simulations with compounds (cyan, green, magenta bars) and without compounds (black bars; apo-RGS4). On comparing these data for Model 1, we observed that all compounds perturb salt-bridges between the helical pairs $\alpha_4$-$\alpha_5$ (D90-K125/E97-K110), $\alpha_5$-$\alpha_6$ (E126-R134), and $\alpha_6$-$\alpha_7$ (D130-K155), but only compound **1** perturbs the $\alpha_4$-$\alpha_7$ salt-bridge (K99-D150) marginally higher than compounds **2** and **3**; perturbations by compounds **2** and **3** are compa-

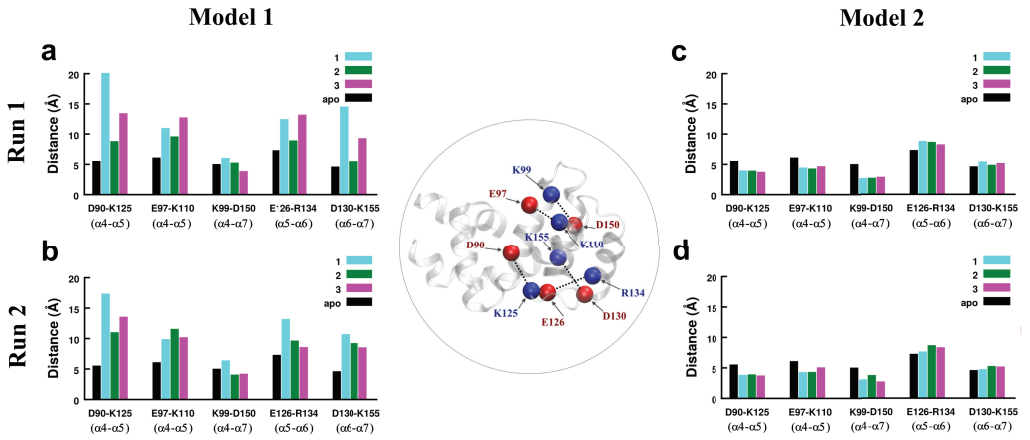



**Figure 4.8.** The histograms of average distances between the centers of mass of residues involved in five salt-bridge-forming residue pairs are shown from two independent simulations of Model 1 and Model 2 for three TDZD compounds. The data for an apo-RGS4 simulation from our previous work [13] are also shown (black histograms). The $C_\alpha$-atoms of all residues involved in salt-bridges are shown and labeled as red/blue spheres on the RGS4 structure (inset in circle).

rable to apo-RGS4. For salt-bridges between the helical pairs $\alpha_4$-$\alpha_5$ (D90-K125) and $\alpha_6$-$\alpha_7$ (D130-K155), compound **1** perturbs salt-bridges more than compounds **2** and **3** in both simulations of Model 1 (Fig. 4.8 a,b). The $\alpha_5$-$\alpha_6$ salt-bridge (E126-R134) is also perturbed more by compound **1** in comparison to compound **2** in both simulations (Fig. 4.8 a,b), and in comparison to compound **3** in the second simulation (Fig. 4.8b). However, compound **2** or compound **3** could perturb one of the $\alpha_4$-$\alpha_5$ salt-bridges (E97-K110) marginally more than compound **1** (Fig. 4.8a,b).

The data from two simulations of Model 2 (Fig. 4.8c,d) reveal no significant perturbations in these salt-bridging interactions although salt-bridges between the helical pairs $\alpha_4$-$\alpha_5$ (D90-K125/E97-K110) and $\alpha_4$-$\alpha_7$ (K99-D150) are marginally stabilized in comparison to apo-RGS4. These data suggest that in Model 1 compounds with aromatic functional groups destabilize interhelical salt-bridging interactions more than the compounds with aliphatic functional groups. Moreover, the perturbations are allosteric since two salt-bridges significantly perturbed by compound **1** (D90-K125 and D130-K155) have residues K125 and D130

that are located in the $\alpha_5$-$\alpha_6$ interhelical loop, away from the docking site residue (C95). As highlighted above, this loop directly participates in the protein-protein interaction between RGS4 and the G$\alpha$-subunit (Fig. 4.1b). Overall, the mode of binding of TDZD compounds in Model 2 resulted in largely insignificant perturbations in the RGS4 structure in comparison to binding of these compounds in Model 1.

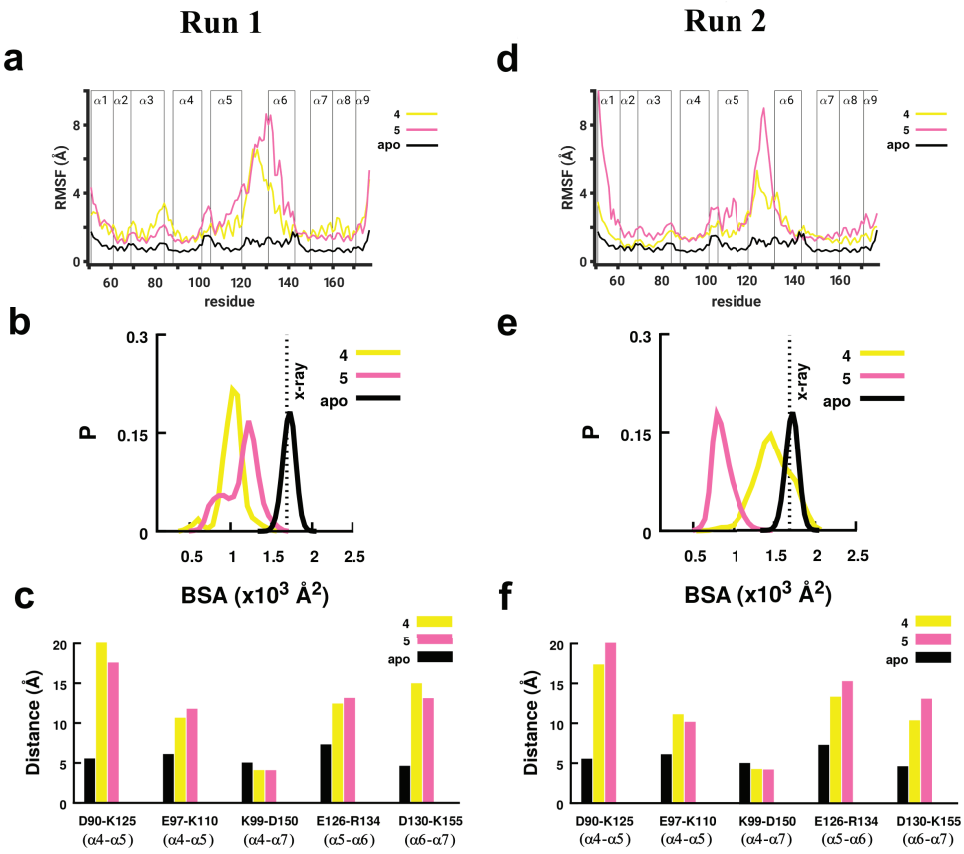



**Figure 4.9.** The data from RMSF (panel a/d), BSA (panel b/e), and salt-bridging interactions (panel c/f) are shown from two simulations of each non-TDZD compound in Model 1 (compound 4, yellow trace; compound 5, magenta trace). Other details in panels a/d, b/e, and c/f are similar to Figs. 4.4, 4.7, and 4.8, respectively.

### 4.4.2 Simulations of non-covalent (non-TDZD) ligands 4 and 5

We further studied compounds **4** and **5** that are analogues of compound **1** with aromatic functional groups but lacking sulfur atoms (Fig. 4.1a). While compound **1** is a known

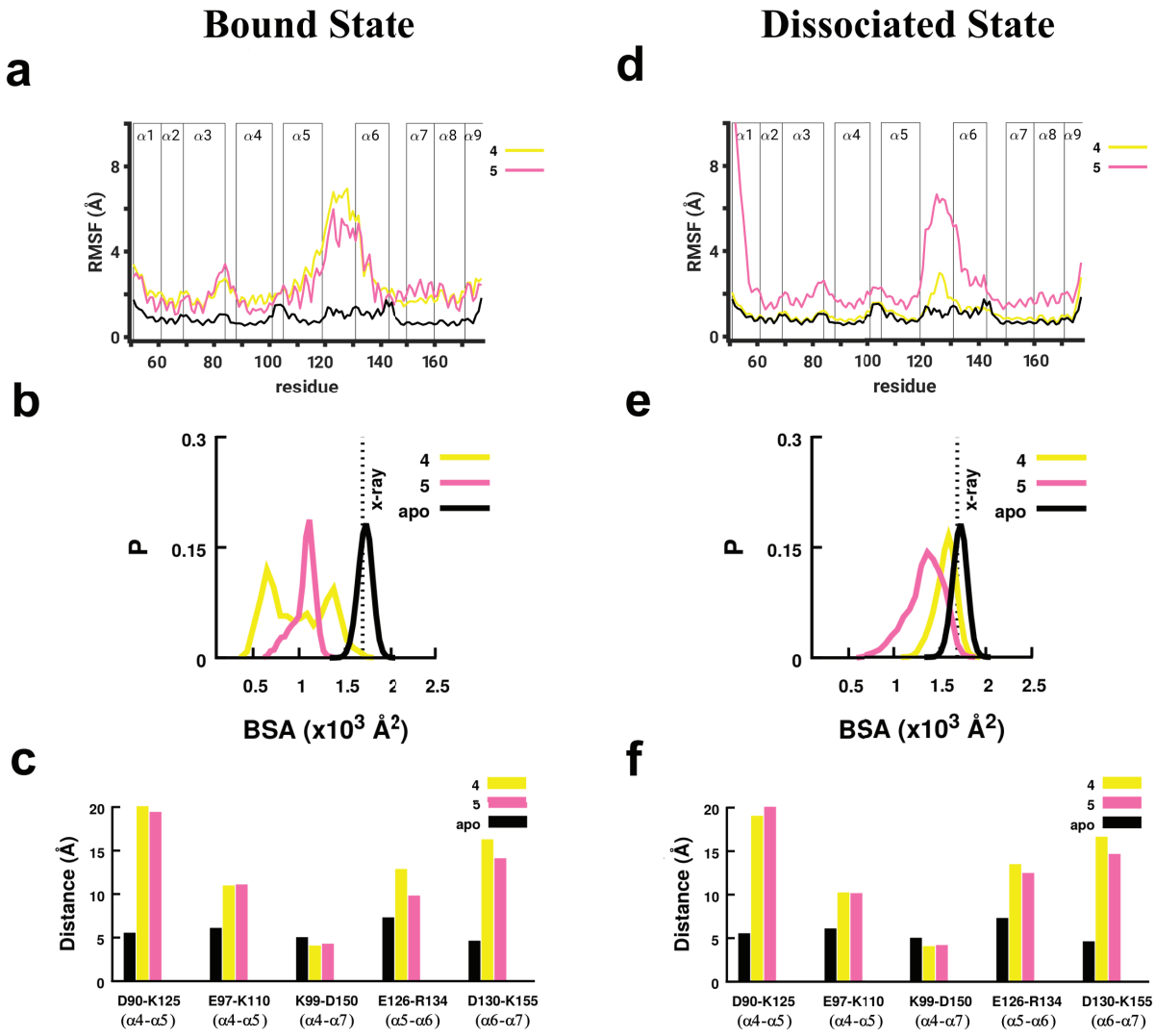



Figure 4.10: Data similar to those presented in Fig. 4.9 are shown from Run3 for compounds **4** and **5** in which diffusion of each compound out of the protein pocket was observed. The left-panels show data for parts of trajectories when compounds still reside within the protein, and the right-panels show data for the remaining parts of trajectories when compounds have diffused out of the pocket.

inhibitor of the wild-type RGS4/G$\alpha$ protein-protein interaction, compounds **4** and **5** do not inhibit this interaction [4]. The mechanistic basis of this observation remains unknown. To test the ability of compounds **4** and **5** to perturb the RSG4 structure, we conducted MD simulations after docking both compounds in each model of RGS4 (Fig. 4.2 and Fig. 4.3). For

Model 1, we conducted three independent 1 $\mu$s long MD simulations of each compoundand observed that in two out of three simulations, these compounds reside within the $\alpha_4$-$\alpha_7$ helical bundle throughout 1 $\mu$s, but in one simulation, each compound diffuses out of the helical bundle into the solvent after transiently residing within the helices. For Model 2 as well, we conducted three independent simulations of each non-TDZD compound. From these simulations, we observed non-specific interactions of compounds (with residues on the protein surface in the vicinity of the docking site) that lead to their rapid dissociation into the solvent. We discontinued these trajectories after the dissociation of each compound. We therefore present analyses from three trajectories of each non-TDZD compound for Model 1 where compounds can transiently reside within the protein.

The data from RMSF, BSA, and salt-bridge measurements from the first two simulations of each compound in Model 1 (where compounds remain bound) are presented in Fig. 4.9 and from the third simulations (where compounds dissociate) are shown in Fig. 4.10. The RMSF data (Fig. 4.9 a,d and Fig. 4.10a) show that both compounds in their transiently bound states perturb the $\alpha_5$ and $\alpha_6$ helices as well as the $\alpha_5$-$\alpha_6$ interhelical loop, and the perturbations by compound 5 are marginally higher than compound 4. The peaks of the BSA histograms (Fig. 4.9 b,e and Fig. 4.10b). for each compound are shifted away from the values in the crystal structure or in the apo-RGS4 simulation, thereby indicating an open conformation of these helices. Accordingly, 4 of 5 interhelical salt-bridges (D90-K125, E97-K110, E126-R134, and D130-K155) are significantly perturbed, while the $\alpha_4$-$\alpha_7$ (K99-D150) is marginally stabilized (Fig. 4.9 c,f and Fig. 4.10c). This is consistent with the observation that in two simulations of Model 1, both non-TDZD compounds reside within the $\alpha_4$-$\alpha_7$ helical bundle. On dissociation of compound 4 in the third simulation, we observed decreased fluctuations in the $\alpha_5$-$\alpha_6$ interhelical loop and a shift of the BSA-histogram peak toward the crystallographic or apo-RGS4 values (yellow traces in Fig. 4.10 d,e *vs.* Fig. 4.10 a,b). After compound 5 moved out of helices, it continued to interact with the protein surface thereby perturbing the $\alpha_5$-$\alpha_6$ interhelical loop and neighboring helices (magenta traces in

Fig. 4.10 d,e). The compounds continue to perturb salt-bridging interactions (Fig. 4.10 c,f) so long as they reside within the helices or near the protein surface on dissociation from the pockets.

In each model, simulations of the non-TDZD compounds (**4** and **5**) show spontaneous dissociation, consistent with the hypothesis that these two compounds do not reside inside RGS4 for longer time-scales, as they cannot covalently bind to the key cysteine residue (C95). However, despite the transient residence of non-TDZD compounds within the protein domain and their subsequent spontaneous dissociation to the solvent, their ability to perturb the RGS4 structure while bound suggests the potential of non-covalent compounds (possibly with higher binding affinities) as promising candidates for developing the next generation of RGS inhibitors.

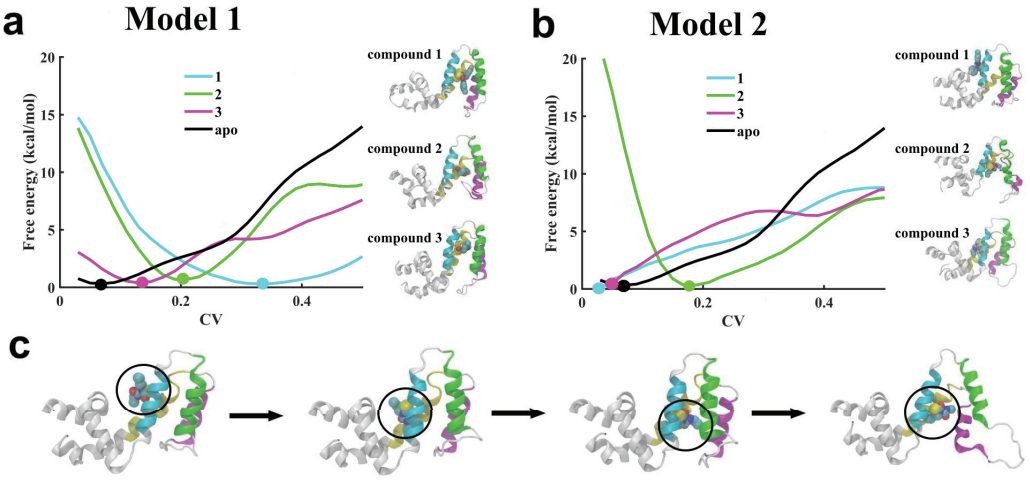



Figure 4.11: (a and b) Free energy profiles are plotted against the collective variable (CV) for structural transitions (between open and closed states) in RGS4 when three TDZD compounds are docked in distinct pockets created in Models 1 and 2. For each model, thermodynamically favorable conformations of RGS4 bound to TDZD compounds are also shown as cartoons in panels a and b. (c) For compound **2** in Model 2, highlighted as cartoons are conformations of RGS4 showing spontaneous diffusion of compound **2** (CCG-203769; Fig. 4.1a) from its initially-docked position on the protein surface to within the $\alpha_4$-$\alpha_7$ helical bundle. The circle in panel c denotes the combined location of covalently-linked residue C95 and compound **2**. For all panels, the protein backbone in snapshots is depicted in white ribbons except helices $\alpha_4$ through $\alpha_7$ that are uniquely colored as in Fig. 4.1b (This figure was created by my co-author Hossein Mohammadiarani [14]).

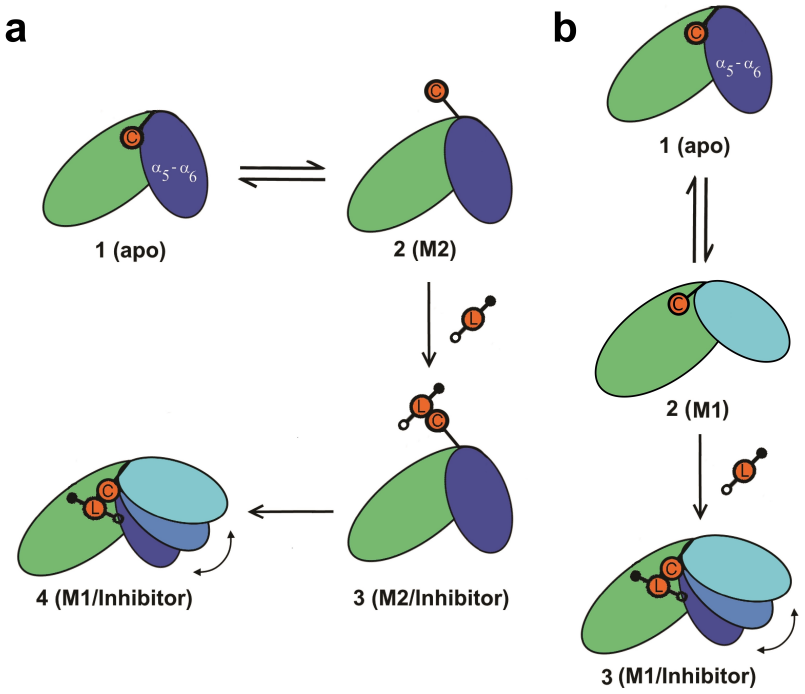



Figure 4.12: (a) A schematic highlighting the proposed mechanism of binding of TDZD small-molecules to RGS4 is shown. In this scheme, the exposure of C95 (orange circle labeled **C**) in the apo-RGS4 conformation (panels 1 and 2) allows initial covalent recognition (panel 3) of small-molecules (orange circle labeled **L** flanked by filled/empty circles indicating $R^1$ and $R^2$ functional groups) and a subsequent migration of compounds to the core of the $\alpha_4$-$\alpha_7$ helical bundle causing allosteric structural perturbations in helices (panel 4), especially in residues in the RGS/G$\alpha$ protein-protein interface. (b) Previously proposed mechanism [15] for the exposure of C95 in the apo-RGS4 conformation (panels 1 and 2) and subsequent binding of compounds causing allosteric structural perturbations (panel 3).

### 4.4.3 Thermodynamic analyses of apo-RGS4 and RGS4/TDZD complexes

To understand thermodynamics of conformational changes in RGS4 in the apo-form as well as when bound to TDZD small-molecules (compounds **1**, **2**, and **3** in Fig. 4.1a), we carried out seven independent enhanced sampling MD simulations using the metadynamics method [153] (see methods). Specifically, we resolved free-energy profiles for structural transitions between the open and closed-states in apo-RGS4 and in Models 1 and 2 of RGS4 when bound to the three TDZD compounds (Fig. 4.11a and 4.11b). The free-energy profiles were resolved in a multidimensional projection of atomic Cartesian coordinates on an eigenvector reaction coordinate (also referred to as a collective variable, CV) such that the CV spans a range between 0 and 1 for closed and open conformations, respectively. The reference conformational states defining the CV were chosen to represent the apo-RGS4 crystal structure (closed conformation of RGS4) and the conformation reported in our previous work (open conformation of RGS4) [102].

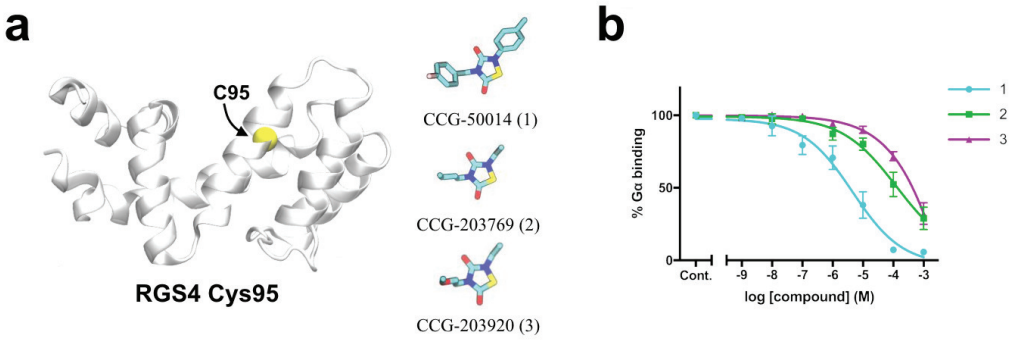



Figure 4.13: (a) The structures of the single-cysteine RGS4 construct (RGS4 C95) and three compounds used in the flow cytometry protein-protein interaction assay are shown. (b) Inhibition of the RGS4 C95/G$\alpha$ protein-protein interaction by compounds **1**, **2** and **3** over a range of concentrations is shown (This figure is provided by my co-author Vincent Shaw [14]).

We observed that the free-energy profile for a structural transition from a closed to open-state in apo-RGS4 (black trace in Fig. 4.11a,b) has a global minimum (marked by a black filled circle in Fig. 4.11a,b) very close to the initial value of CV = 0, which indicates that RGS4 in its apo-state thermodynamically favors a closed-state, as also seen in the crystal structure. On comparing free-energy profiles of protein bound to compound **1** (cyan traces in Fig. 4.11a,b), we observed that the global minimum in Model 1 (marked by a cyan filled circle in Fig. 4.11a) is located at a CV value of ∼0.35, while in Model 2 the global minimum (marked by a cyan filled circle in Fig. 4.11b) is located very close to zero. This suggests that open-like RGS4 states are thermodynamically favored when compound **1** is docked in the binding pocket in Model 1 in comparison to Model 2, where thermodynamically favored conformations are similar to the closed-state, as in apo-RGS4.

However, from the free-energy profiles for compound **2** in both models (green traces in Fig. 4.11a,b), we observed that the global minimum (marked by a green filled circle in Fig. 4.11a,b) is located in the vicinity of a CV value of ∼0.20. This suggests that open-like RGS4 states likely exist for compound **2** not only in Model 1 but also in Model 2. The ability of compound **2** to stabilize open-like conformations in Model 2 is explained by the observation that the compound **2** spontaneously diffuses during the metadynamics simulation from its initially-docked position on the protein surface to its final position within the $\alpha_4$-$\alpha_7$ helical bundle (snapshots in Fig. 4.11c), thereby acquiring conformations similar to Model 1. Importantly, the diffusion of compound **2** from the protein surface to within this helical bundle is driven both by the flexibilities in helices and burial of the side-chain of key cysteine residue C95.

For compound **3**, the location of the global minimum in each free-energy profile (marked by a magenta filled circle in Fig. 4.11a,b) indicates a conformational stability behavior similar to compound **1** in that the open-like RGS4 states are thermodynamically favored in Model 1 and a closed state is favored in Model 2. However, the global minimum for compound **3** in Model 1 is located at a CV value of ∼0.14 smaller than the CV value of ∼0.35 for the global

minimum for compound **1** (magenta *vs.* cyan filled circles in Fig. 4.11a). This suggests that perturbations to the RGS4 structure (relative to the closed state) are smaller for compound **3** (with aliphatic functional groups) than for compound **1** (with aromatic functional groups). Overall, the shift in the free-energy minimum from the higher to lower CV values on binding of TDZD compounds in Model 1 (CV = 0.35, 0.20 and 0.14 for compounds **1**, **2**, and **3**, respectively) is consistent with the perturbation trends observed in classical MD simulations (Fig. 4.4, 4.7, 4.8).

The possibility of binding of TDZD analogues (**1**, **2**, and **3**) to C95 in a closed-state conformation of RGS4 due to a transient exposure of C95 (Model 2) [13] shows that multiple binding mechanisms may exist by which these small-molecules can access the otherwise buried cysteine residue C95. In Model 2, all compounds are docked on the protein surface located near the "B-site" of RGS4 [65], entirely outside of the $\alpha_4$-$\alpha_7$ helical bundle. Therefore, the initial binding of TDZD analogues to C95 is not dependent on a significant opening of the $\alpha_5$-$\alpha_6$ helical pair. This mechanism of small-molecule recognition by C95 is distinct from our previously suggested mechanism [102] that highlighted the flexibility in the $\alpha_5$-$\alpha_6$ helical pair as a potential route for compound binding. However, our enhanced sampling metadynamics simulations showed that compounds covalently-bound to the exposed C95 residue in Model 2 (e.g. compound **2**) could translocate from the protein surface to the core of the $\alpha_4$-$\alpha_7$ helical bundle (Fig.4.11c) and stabilize open-like states similar to Model 1 while remaining covalently-bound, thereby suggesting that Model 2 can evolve toward Model 1 (Fig. 4.12). Therefore, we propose that these mechanisms of binding of compounds may not be mutually exclusive. This substantiates the importance of coupling between local conformational flexibilities in protein side-chains (e.g. exposure of C95) with global protein dynamics (e.g. flexibilities in RGS4 helices) to facilitate small-molecule recognition and allow allosteric inhibition of the protein-protein interface.

### 4.4.4 Functional analysis of the inhibition of the RGS4-C95/G$\alpha$ protein-protein interaction by TDZD compounds

To further investigate the effect of binding of TDZD compounds **1**, **2** and **3** on the interaction of single-cysteine RGS4 (RGS4 C95) and G$\alpha$ *in vitro*, we utilized a flow cytometry protein interaction assay (FCPIA) (see methods) [72]. A single-cysteine (RGS4 C95) mutant was used to limit compound binding to C95, such that the results reflect compound action at the same cysteine to which compound was covalently docked in simulations. The concentration-response curves (Fig. 4.13) show that the single-cysteine protein is inhibited by compound **1** with an $IC_{50}$ value of $\sim$4.5 $\mu$M, and less inhibited by compounds **2** and **3** (Fig. 4.13b). The magnitude of structural perturbations and deviations from the native RGS4 conformation (Fig. 4.4, 4.7, 4.8, and 4.11) induced by the compounds in modeling and simulation studies correlate well with the results of inhibition experiments. The use of single-cysteine mutants resulted in lower potencies of inhibition by each compound compared to wild-type proteins [4]. Notably, the difference in potencies between aromatic compound (1) and aliphatic compounds (2 and 3) is more pronounced in single-cysteine RGS4 C95 than in WT RGS4. This may be because in WT proteins, differences in dynamics between compounds bound to C95 are masked by the action of compounds at other cysteines. Compound **1** causes the greatest perturbations in the RGS4 structure and also showed the greatest potency to inhibit RGS4/G$\alpha$ binding. Compounds **2** and **3** caused smaller perturbations in the RGS4 structure and accordingly showed a reduced ability to inhibit RGS4/G$\alpha$ binding.

## 4.5 Conclusion

In this chapter, I have presented modeling and simulation studies predicting structural perturbations in the RGS4 protein on binding to various small-molecules. We found that compounds with aromatic functional groups significantly perturb the protein structure in com-

parison to those with the aliphatic functional groups. Non-covalent compounds only transiently perturb the protein as these compounds spontaneously dissociate to solvent. Thermodynamic analyses of RGS4/small-molecule complexes suggest that two distinct modes of binding can lead to a two-step mechanism in which compounds are initially recognized by the exposed side-chains of conserved and buried cysteine residues (C95 in RGS4) followed by their migration to the helical core of RGS4 that leads to significant allosteric perturbations in those RGS4 residues that are located in the RGS4/G$\alpha$ protein-protein interface. These findings will inform future drug development efforts focused on the discovery of non-covalent compounds capable of inducing similar allosteric perturbations.

## 4.6 Publications

The work described in this chapter has resulted in the following journal article:

- Mohammadi, M., Mohammadiarani, H., Shaw, V. S., Neubig, R. R., Vashisth, H. (2019). Interplay of cysteine exposure and global protein dynamics in small-molecule recognition by a regulator of G-protein signaling protein. Proteins: Structure, Function, and Bioinformatics, 87(2), 146-156.

The work has also appeared on journal cover [17]. The cover image is included in Appendix I.

# CHAPTER 5

# STUDIES ON COUPLING OF PROTEIN DYNAMICS AND SALT-BRIDGING INTERACTIONS IN RGS PROTEINS

## 5.1 Abstract

RGS proteins modulate receptor signaling by binding to activated G-protein $\alpha$-subunits, accelerating GTP hydrolysis. Selective inhibition of RGS proteins increases G-protein activity and may provide unique tissue specificity. Thiadiazolidinones (TDZDs) are covalent inhibitors that act on cysteine residues to inhibit RGS4, RGS8 and RGS19. There is a correlation between protein flexibility and potency of inhibition by the TDZD CCG-50014. In the context of a single conserved cysteine residue on the $\alpha$4 helix, RGS19 is the most flexible and most potently inhibited by CCG-50014, followed by RGS4 and RGS8. We hypothesized that interhelical salt-bridge forming residues are responsible for differences in both flexibility and potency of inhibition among RGS isoforms. RGS19 lacks a charged residue on the $\alpha$4 helix that is present in RGS4 and RGS8. Introducing a negative charge at this position (L118D) increased the thermal stability of RGS19 and decreased the potency of inhibition by CCG-50014. Mutations which eliminated salt bridge formation in RGS8 and RGS4 led to decreased thermal stability in RGS8 and increased potency of inhibition of both RGS4 and RGS8. Molecular dynamics (MD) simulations with an added salt bridge in RGS19 (L118D) showed reduced RGS19 flexibility. Hydrogen-deuterium exchange (HDX)* studies showed striking differences in flexibility in the $\alpha$4 helix of RGS4, 8, and 19 with salt bridge modifying mutations. These results show that an $\alpha$4 salt bridge-forming residue controls flexibility in several RGS isoforms and supports a causal relationship between RGS flexibility and the potency of TDZD inhibitors.

* The experimental works (mutegenesis, HDX, and compound potency studies) performed by my collaborator Vincent Shaw at Michigan State University.

## 5.2 Background

Drug specificity is often considered to be like a key fitting into a complementary shaped lock. It has become clear recently that protein dynamics can play an important role in drug discovery. RGS proteins bind to activated $G_\alpha$ subunits of G-proteins, thereby accelerating GTP hydrolysis and attenuating G-protein signaling. In regulating G-Protein Coupled Receptor (GPCR) signaling, RGS proteins play a role in the physiology of numerous systems. By inhibiting RGS proteins, signaling via a GPCR may be enhanced. There are twenty RGS isoforms, each with a different tissue distribution. Combination of GPCR agonists with inhibitors specific for a single RGS isoform should limit effects on GPCR signaling to a subset of target tissues that intersects with the distribution of the GPCR. This has the potential to reduce agonist off-target effects, and makes RGS proteins an attractive target for modulation of GPCR signaling.

RGS inhibitors discovered to date are covalent modifiers of cysteine residues and are selective for RGS4 and RGS1 [84, 207]. These have four and three cysteines, respectively, in the RGS homology domain, which is more than in most other RGS proteins. RGS4 has been linked to nervous system related disease states in which RGS4 inhibition may be desirable, including seizures [208] and Parkinson's disease [90, 91, 93]. Continued efforts to seek non-covalent inhibitors are attractive, because the lower risk associated with non-covalent inhibitors is considered safer and may ease further development [209]. In addition, it would be valuable to discover RGS inhibitors with other specificities since other RGS proteins less potently inhibited by covalent modifiers have been implicated as potential targets, including RGS17 in cancer [210, 211] and RGS19 in depression [212]. To identify noncovalent

inhibitors with novel specificities, it will be necessary to understand what factors apart from the number of cysteines in the RGS domain drive selectivity of RGS inhibitors.

The RGS homology domain contains nine helices. A cysteine residue on $\alpha 4$, which faces the interior of the $\alpha 4$-$\alpha 7$ helical bundle, is conserved among all 20 RGS isoforms with the exception of RGS6 and RGS7 [100]. Interestingly, when RGS proteins are mutated to contain only this single, shared cysteine, there are still dramatic differences in potencies by which different isoforms are inhibited [13]. RGS19, which contains only the shared $\alpha 4$ cysteine, is more potently inhibited than single-cysteine versions of RGS4 and RGS8 [13, 202].

**Hypothesis:** Inhibitor potency can be modulated by salt-bridging interactions.

Previously, we found using MD simulations that RGS19 is more flexible than RGS4 and RGS8 [13]. In these modeling studies we also found that the extent of perturbations of salt bridge interactions by inhibitor compounds correlated with structural flexibility in RGS4 [14, 17]. In this work, we sought to identify the cause of flexibility differences among these isoforms and hypothesize that mutations aimed at salt bridge interactions that enhance RGS protein flexibility can increase the potency of inhibitors such as CCG-50014.

## 5.3 Methods

We performed two sets of classical all-atom and explicit-solvent MD simulations (section 2.2) for single-cysteine RGS4/ RGS4 D90L, single-cysteine RGS8/ RGS8 E84L, and RGS19/ RGS19 L118D using the NAMD software [111] and the CHARMM force-field with the CMAP correction [122, 123]. We used VMD for system creation and post-simulation analysis [110]. The initial coordinates were obtained from the protein data bank files with codes 1AGR

(RGS4), 2ODE (RGS8), and 1CMZ (RGS19). Except Cys95 in RGS4 and Cys107[1] RGS8, all cysteines were mutated to alanines. Each protein was then solvated in a simulation box of TIP3P water molecules [205] and charge-neutralized with NaCl. Each solvated and ionized system was energy minimized for ∼500-1000 cycles via conjugate-gradient optimization, then equilibrated via 1$\mu$s MD simulations conducted with a time-step ($\Delta$t) of 2 fs. The NPT ensemble with a Langevin thermostat and a damping coefficient of 5 $ps^{-1}$ was used for temperature control and the Nosé-Hoover barostat was used for pressure control. Periodic boundary conditions were used throughout; non-bonded interactions were accounted for with a cut-off of 10 Å where smooth switching was initiated at 8 Å. Long-range electrostatic interactions were handled using the Particle Mesh Ewald (PME) method. The dynamic cross-correlation (DCC) maps were created, and salt-bridge interaction analysis were performed as described in section 2.3.2.

## 5.4 Results and Discussion

Comparison of the structures for RGS19 (PDB 1CMZ[213]), RGS4 (PDB 1AGR[101]), and RGS8 (PDB 2ODE [214]) shows that there are differing numbers of interhelical salt bridges among their $\alpha$4-$\alpha$7 helix bundles. Some of these may contribute to differences in stability and dynamics among the RGS isoforms.

RGS19 has only one interhelical salt bridge in this bundle, between E125 ($\alpha$4) and K138 ($\alpha$5) (Fig. 5.1E). This salt bridge is well conserved among all three proteins, however, so it is unlikely to contribute to observed differences in flexibility [13].RGS4 has this salt bridge (E97-K110, Fig. 5.1B) as well as two additional interhelical salt bridge locations, between D130 ($\alpha$6) and K155 ($\alpha$7) (Fig. 5.1D); and a salt bridge network that connects D90 ($\alpha$4), K125 on the $\alpha$5-$\alpha$6 interhelical loop, and E117 on $\alpha$5 (Fig. 5.1B). RGS8 has four interhelical

[1]Note, amino acid numbering follows that for RGS8, Isoform 1, NCBIRefSeq$NP_2$03131.1

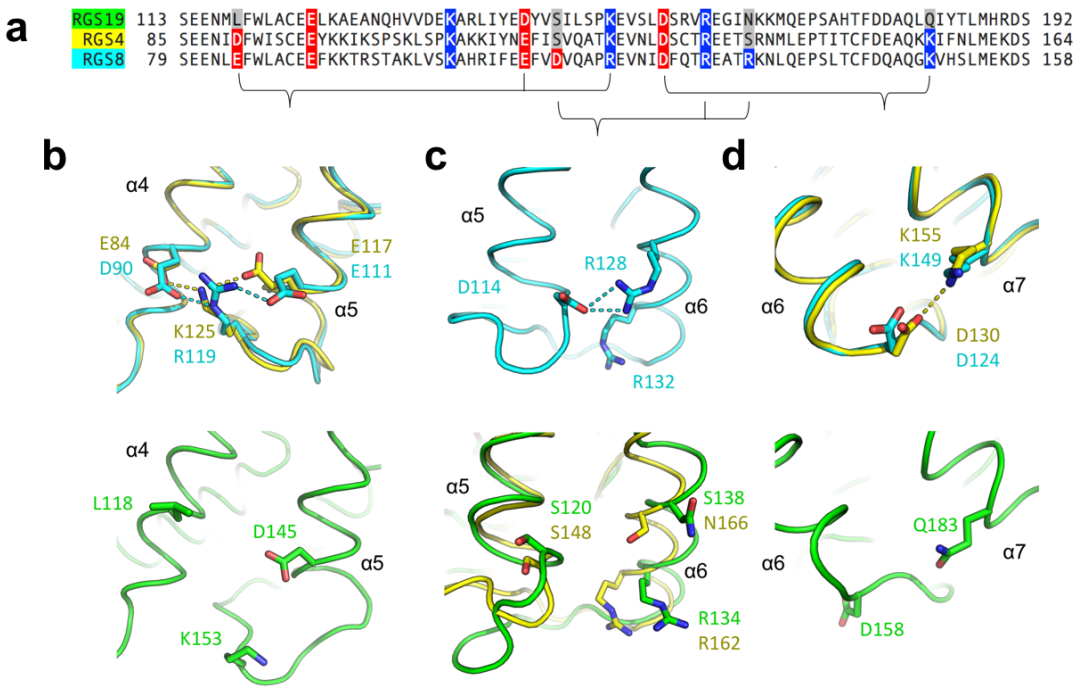



**Figure 5.1.** (a) Alignment of RGS19, RGS4, and RGS8 sequences in $\alpha$4-$\alpha$7 helix bundle. Charged residues that make interhelical contacts are indicated in red and blue. Structural alignments of $\alpha$4-$\alpha$5 (b and e), $\alpha$5-$\alpha$6 (c and f), and $\alpha$6-$\alpha$7 (d and g) helix pairs are shown, with highlighted residues in panel a rendered as sticks. RGS19 (PDB 1CMZ) is in green, RGS4 (PDB 1AGR) is in yellow, and RGS8 (PDB 2ODE) is in cyan (This figure was created by my co-author Vincent Shaw [16]).

salt bridges in the $\alpha$4-$\alpha$7 bundle. One salt bridge (E91-K104, Fig. 5.1B) is shared by both RGS19 and RGS4. Two salt bridges (D124-K149, Fig. 5.1D, and E84-R119-E111, Fig. 5.1B) are shared with RGS4. Finally, a salt bridge between D114 ($\alpha$5) and R132 ($\alpha$6) is unique to RGS8 (Fig. 5.1C).

To estimate the relevance of each of these salt bridges in maintenance of helix bundle rigidity, the time each amino acid in a charged pair spent within a Å of one another over the course of a long timescale MD simulation was measured [13]. The $\alpha$6-$\alpha$7 salt bridge, which is present in RGS4 and RGS8 but absent in RGS19, was stably maintained. It occupied a salt bridge-forming distance for 31.5% of the simulation in RGS4 and 36.1% in RGS8. The

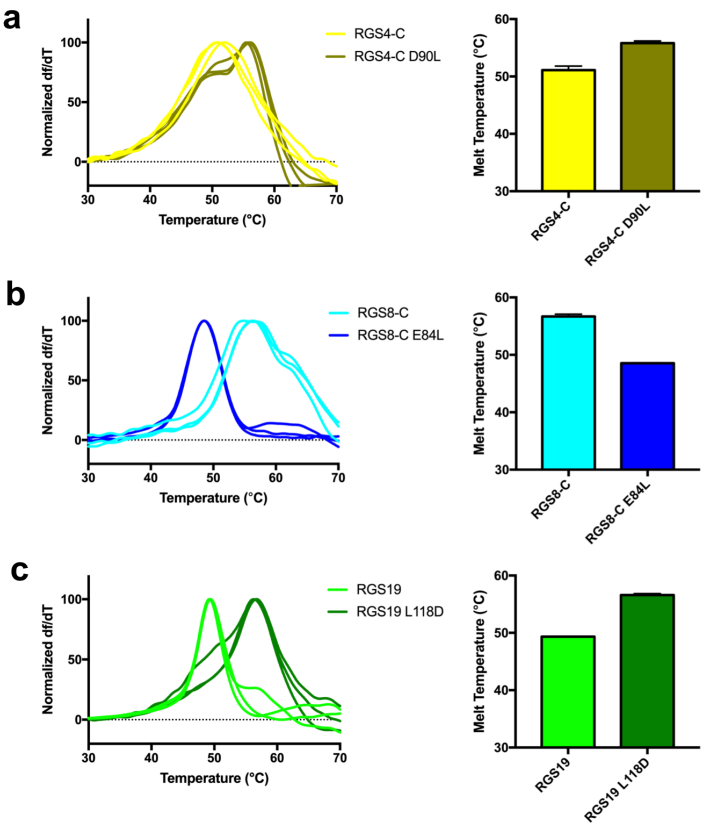



**Figure 5.2.** Thermal stability was determined by differential scanning fluorimetry. A) The L118D mutation in RGS19 increased melting temperature by 7 °C compared to WT. B) The E84L mutation in RGS8 decreased melting temperature by 8 °C. C) The RGS4 D90L mutation introduced a biphasic melt curve and increased melting temperature by 5 °C. For each pair, derivative melt curves are shown on the left and melt temperatures are shown on the right. Error bars represent SD. n=3. Analyzed by 1-way ANOVA with Sidak's Multiple Comparisons test (This figure is provided by my co-author Vincent Shaw [16]). ****p $<$ 0.0001

salt bridge interaction between residues of $\alpha$4 and $\alpha$5-$\alpha$6 interhelical loop, also not present in RGS19, was maintained for 58.7% of time in RGS4 and 44.2% in RGS8. The charged pair that is unique to RGS8 between $\alpha$5 and $\alpha$6 helices remained in contact for 47.5% of the simulation.

Based on these MD results, we elected to make mutations that altered interhelical ($\alpha$4-$\alpha$5 and $\alpha$6-$\alpha$7) salt bridges to test their functional roles. In helix $\alpha$4, L118 in RGS19 (Fig. 5.1E)

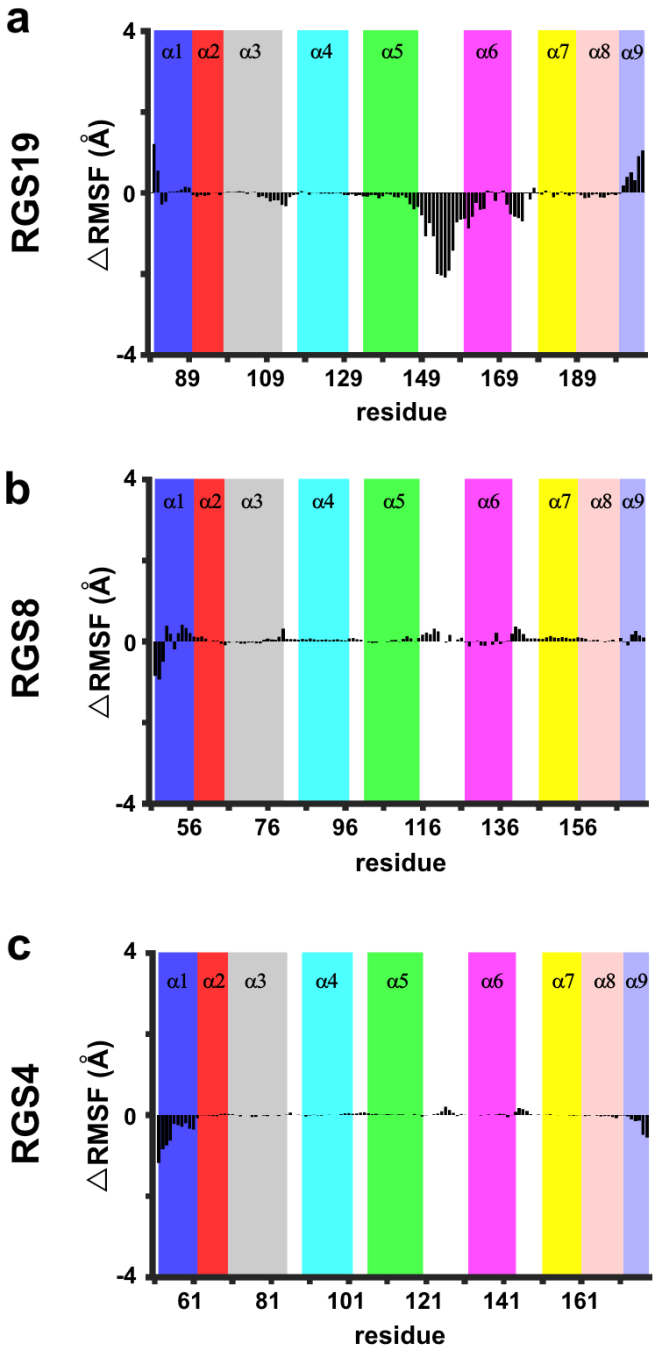



**Figure 5.3.** Change in RMSF per residue (ΔRMSF) between wild-type RGS proteins and RGS proteins with mutation in the $\alpha$4-$\alpha$5 salt bridge forming residue. (a) L118D in RGS19 (b) E84L in RGS8 and (c) D90L in RGS4. Data represent differences in RMSF from two independent MD simulations of the mutated and unmutated forms of RGS proteins.

was mutated to an aspartate to introduce the salt bridge found in RGS4 and RGS8 (Fig. 5.1B). In helix $\alpha$7, Q183 in RGS19 (Fig. 5.1G) was mutated to a lysine, to introduce the $\alpha$6-$\alpha$7 salt bridge found in RGS4 and RGS8 (Fig. 5.1D). In order to eliminate confounding effects due to multiple cysteines in inhibitor potency experiments, all proteins, with and without salt-bridge mutations, used a single-cysteine protein background. Each construct has only the conserved cysteine in helix $\alpha$4 of the RGS domain.

To determine how disruption of the salt bridges D90-K125 in RGS4 and E84-R131 in RGS8 and addition of a salt bridge by L118D mutation in RGS19 may alter protein structure or dynamics, thermal stability was measured by differential scanning fluorimetry. As expected, addition of a salt bridge in RGS19 by the L118D mutation caused a 7 °C increase in thermal stability compared to WT (Fig 5.2A). Removal of a salt bridge by the E84L mutation in RGS8 caused an 8 °C decrease in thermal stability (Fig 5.2B). Unexpectedly, RGS4 showed a more complex pattern, in which D90L mutation resulted in a biphasic melt curve and a 5 °C increase in melting temperature rather than a decrease (Fig 5.2C).

To probe the molecular details of changes in structural flexibility in the mutated RGS4, RGS8, and RGS19, we conducted microsecond timescale classical MD simulations in explicit-solvent for RGS4 D90L, RGS8 E84L, and RGS19 L118D. To understand the effect of the mutations on the protein structures, particularly in helices in the vicinity of the mutated site, we computed the root-mean-squared-fluctuation (RMSF) per residue from two independent MD simulations of mutated and WT RGS4, RGS8, and RGS19. There were minimal changes in RMSF in RGS4 in comparison to RGS8 and RGS19 (Fig. 5.3). We find a modest increase in fluctuation of residues in mutant (E84L) RGS8 vs. the wild-type structure. These changes are in the loop region connecting helices $\alpha$5 and $\alpha$6, the $\alpha$6 helix, and the loop connecting helices $\alpha$6 and $\alpha$7. Similar changes in lesser extend were found in the mutant (D90L) RGS4. The calculated change in RMSF per residue of the mutant (L118D) RGS19

from wild-type RGS19 reveals a strong stabilization and decrease in fluctuations of residues located in helices $\alpha$4-$\alpha$7 and the interhelical loops between these helices. There is a particularly pronounced decrease in motion in the $\alpha$5-$\alpha$6 interhelical loop (Fig. 5.3). Additionally, the RMSF values of residues in helices $\alpha$3 and $\alpha$8 of the mutated RGS19 are more stable in compared to wild-type than in mutated RGS4 and RGS8 (Fig. 5.3).

To further investigate whether the mutations D90L in RGS4, E84L in RGS8, and D118L in RGS19 affect residue-residue interactions, we calculated the dynamic cross-correlation matrix for the C$\alpha$ atoms in all MD trajectories. For wild-type RGS8, we find that the motions of residues in the $\alpha$4 helix (highlighted by the dashed lines in Fig. 5.4) and the $\alpha$5 helix (highlighted by the solid-lines in Fig. 5.4) are marginally positively correlated. This positive correlation between the $\alpha$4 and $\alpha$5 helices remains in the RGS8 E84L mutant, but shows a modest shift in areas of correlation away from the loop connecting $\alpha$4-$\alpha$5 to mid-regions of the $\alpha$4 and $\alpha$5 helices (see arrows). For WT RGS4 and RGS19 the correlation matrices show slight positive correlation between the residues of the $\alpha$4 helix and the residues of the $\alpha$5 helix. For the RGS19 L118D mutant, we find higher residue-residue correlations between residues in helices $\alpha$4 and $\alpha$5 in comparison to unmutated RGS19 (see arrows).

In order to experimentally determine which regions in WT and mutant proteins were affected by the salt bridge mutations, hydrogen-deuterium exchange was performed. After exposure to solvent containing 90% D2O, proteins were digested with pepsin and deuterium incorporation (DI) was measured by mass spectrometry as previously reported [13]. In RGS4, the fragment surrounding the salt-bridge mutation site (aa 88-91) took up deuterium very slowly in both the WT and D90L mutant constructs, reaching 8.1% and 6.7% DI, respectively. However, the D90L mutation led to a substantial increase in deuterium exchange in the 92-97 fragment surrounding Cys95, from 17.5% to 37.0% DI. The RGS4 D90L mutant also trended toward increased DI across all protein fragments compared to WT RGS4, espe-

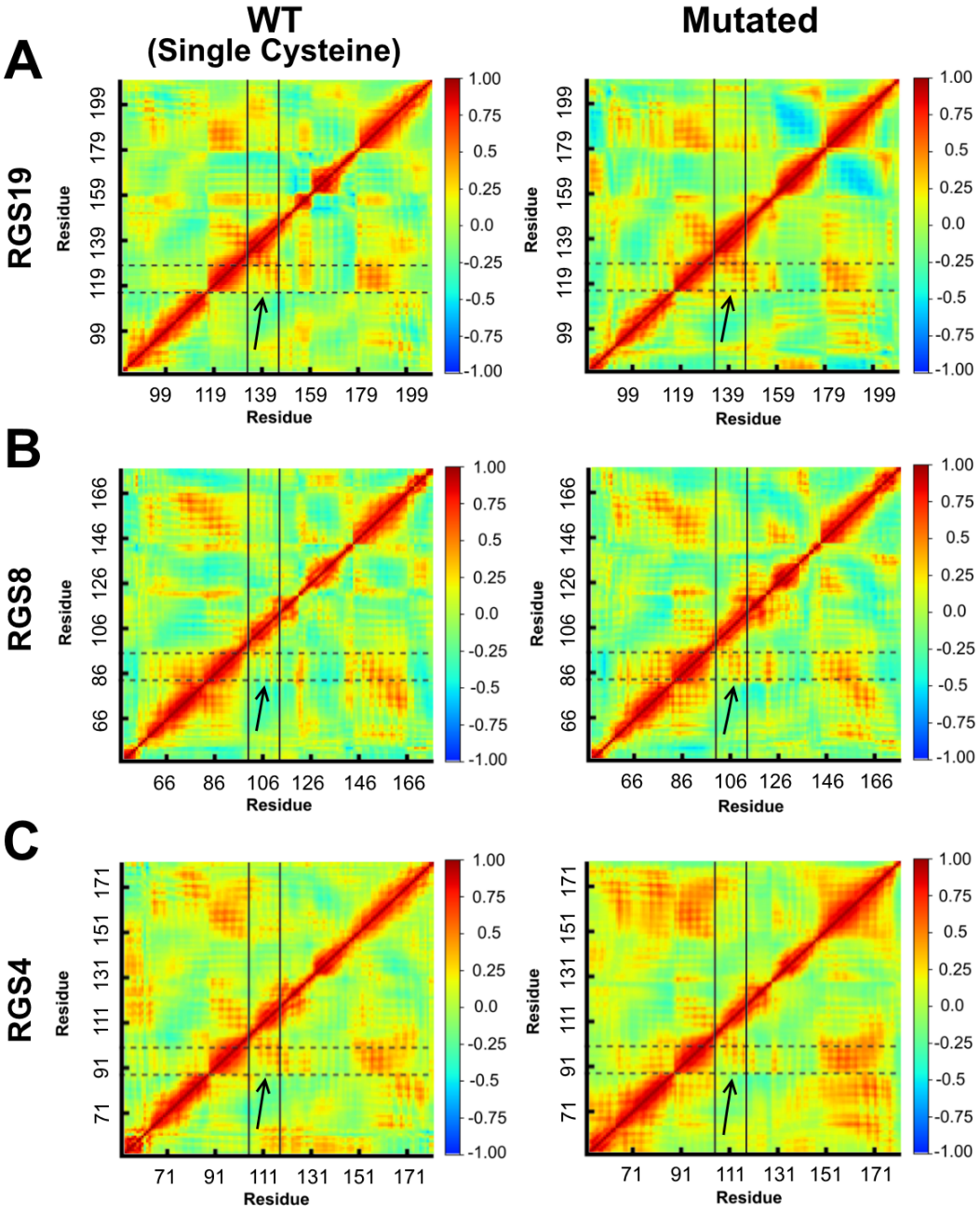



**Figure 5.4.** (a) Dynamic cross-correlation matrix calculated for the C$\alpha$ atoms of (a) RGS19/RGS19 L118D, (b) RGS8/RGS8 E84L, (c) RGS4/RGS4 D90L. Horizontal dotted lines indicate the regions of the $\alpha$4 helix, while vertical solid lines indicate the regions of the $\alpha$5 helix for each protein. The color scheme ranges from anticorrelation (-1.0, blue), no correlation (0, green), and positive correlation (+1.0, red). Values are the average for the two independent simulation runs of each protein.

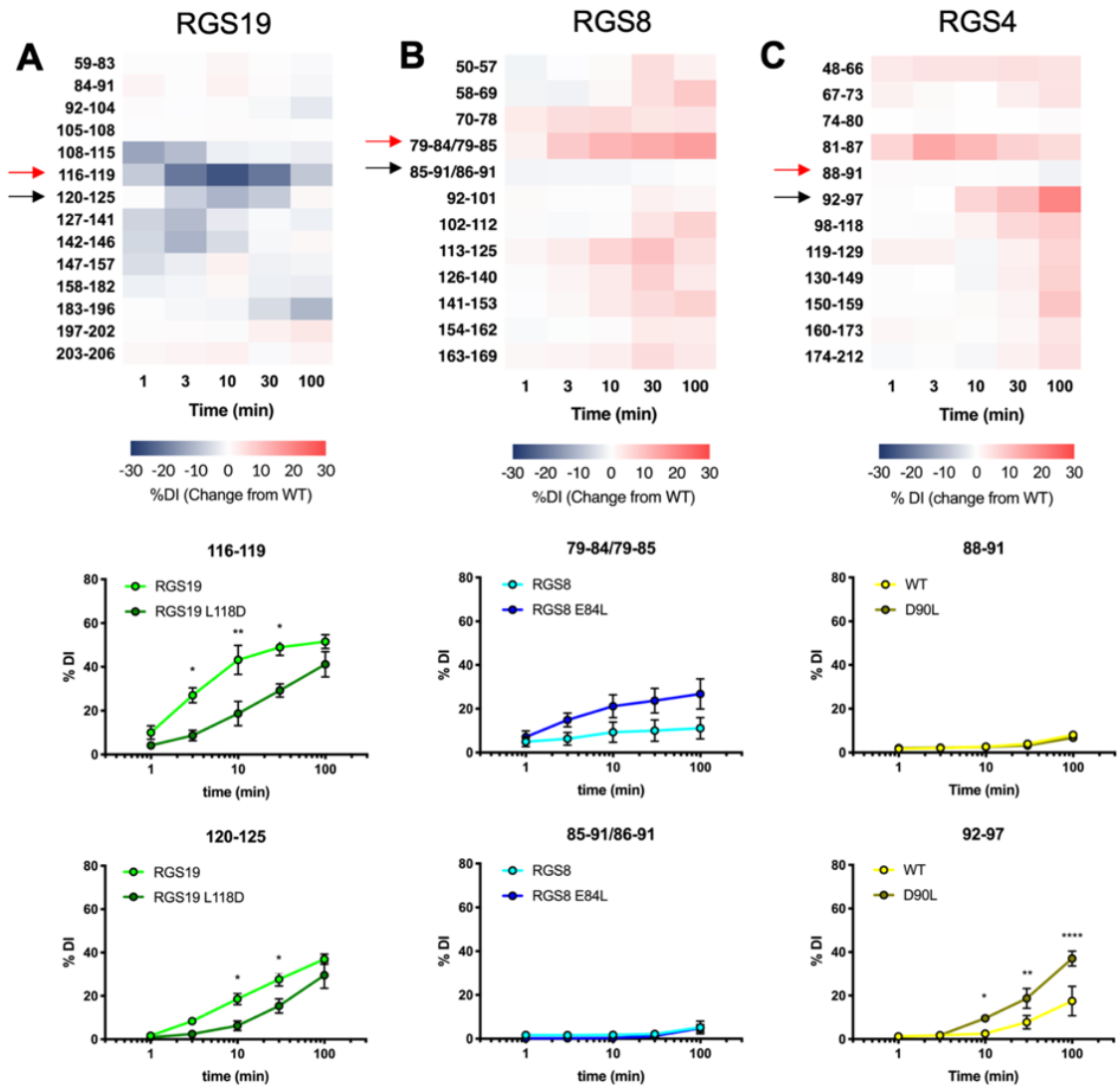



**Figure 5.5.** Difference in % deuterium incorporation (Δ%DI) between mutated and unmutated proteins in RGS19 L118D (A), RGS8 E84L (B), and RGS4 D90L (C) fragments, as measured by HDX. Red arrows indicate fragments containing mutated residue, and black arrows indicate fragments containing conserved $\alpha 4$ cysteine. Kinetics of deuterium incorporation in these fragments for individual constructs are shown below. n=3. Error bars represent SD. Analyzed by 2-way ANOVA with Sidak's multiple comparisons test (This figure is provided by my co-author Vincent Shaw [16]). *$p < 0.05$, **$p < 0.01$, ****$p < 0.0001$

cially at higher timepoints (Fig. 5.5A). In RGS8, removal of the salt-bridge forming residue by the E84L mutation did not cause a significant change in DI in either of the fragments of the $\alpha 4$ helix but trended toward a global increase in DI throughout the protein (Fig.

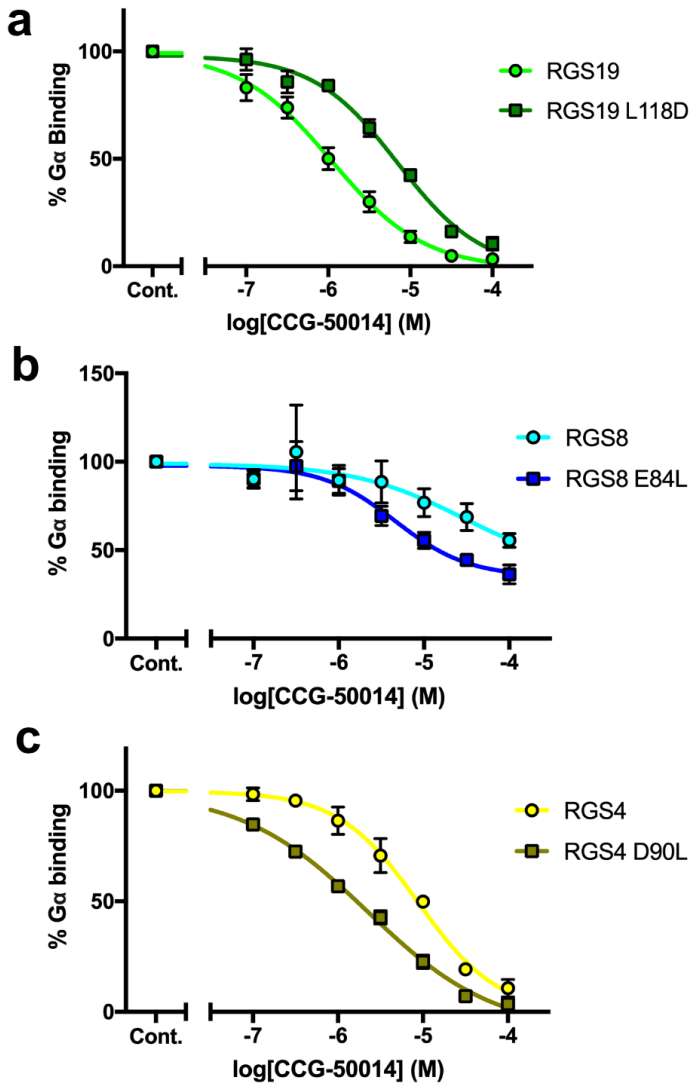



**Figure 5.6.** Potency of inhibition of CCG-50014 against $\alpha 4$ is altered in salt bridge mutants of RGS proteins. (A) RGS4 $IC_50$: 8.8 $\mu$M, RGS4 D90L $IC_50$: 2.2 $\mu$M. (B) RGS8 $IC_50$: 29 $\mu$M, RGS8 E84L $IC_50$: 4.6 $\mu$M (C) RGS19 $IC_50$: 7.0 $\mu$M, RGS19 L118D $IC_50$: 1.1 $\mu$M. n=3 (This figure is provided by my co-author Vincent Shaw [16]).

5.5B). In RGS19, mutation of L118 to a salt bridge-forming residue, aspartic acid, caused significant decreases in DI in both $\alpha 4$ helical fragments, aa 116-119 and aa 120-125. In the 116-119 fragment, WT RGS19 had reached 43.1% DI by 10 minutes, while the RGS19 L118D mutant showed less than half as much DI (18.7%). In fragment 120-125, WT RGS19 reached 18.5% DI at 10 minutes, while the RGS19 L118D mutant reached only 6.2%. Unlike RGS4

and RGS8, the RGS19 L118D mutant's changes in DI were more restricted to fragments from helices neighboring the mutation site, and were most pronounced in the early (1 to 10 minute) timescale (Fig. 5.5C).

Finally, to see whether differences in the $\alpha$4 salt-bridge forming residue altered compound potency, we used a flow-cytometry based protein-protein interaction assay (FCPIA) [215, 216] to measure the binding of RGS proteins to G$\alpha$o and inhibition by CCG-50014. The L118D mutation in RGS19 induced an increase in $IC_50$ from 1.1 $\mu$M (WT) to 7.0 $\mu$M (L118D). Conversely, removal of this charged $\alpha$4 residue in RGS4 and RGS8 induced a decrease in $IC_50$. CCG-50014 inhibited the RGS-G$\alpha$ interaction with an $IC_50$ of 8.8 $\mu$M for WT RGS4 and 2.2 $\mu$M for the RGS4 D90L mutant. It showed a potency of 29 $\mu$M for WT RGS8 and 4.6 $\mu$M for the RGS8 E84L mutant. None of the mutations to salt bridge-forming residues on the $\alpha$4 helix caused notable changes in affinity between G$\alpha$o and RGS proteins. The L118D mutation in RGS19 shifted the Kd of the G$\alpha$o interaction from 16.6 nM to 20.2 nM, the E84L mutation in RGS8 shifted the Kd from 5.9 nM to 4.8 nM, and the D90L mutation in RGS4 shifted the Kd from 5.2 nM to 3.9 nM (Fig. 5.6).

### 5.4.1 The analysis of mutated RGS proteins flexibility

A comparison of the crystal structures of the three RGS proteins studied here revealed several differences in charged residue contacts among the proteins. We first observed that RGS19 has fewer interhelical salt bridges between helices in its $\alpha$4-$\alpha$7 helical bundle than RGS4 or RGS8, which may be responsible for the high flexibility and low melting temperature previously observed in WT RGS19 [13]. RGS8 has four distinct interhelical salt bridge locations within the helical bundle, while RGS4 has three and RGS19 has one (Fig 5.1A), correlating with previously observed flexibility differences. RGS19 is most flexible, followed

by RGS4 and RGS8 [13]. This further suggests a role of salt bridges in RGS protein flexibility.

To determine which of these salt bridge locations was most relevant to the overall stability of the RGS proteins, we analyzed the stability of salt bridges via $\mu$s-timescale MD simulations. A salt bridge between a glutamate on $\alpha$4 and lysine on $\alpha$5 is well conserved among RGS proteins (Fig. 5.1A, 5.1B, and 5.1E). Therefore, it is unlikely to contribute to differences in flexibility among RGS isoforms. This leaves two positions at which interhelical salt bridges are shared by RGS4 and RGS8 but lack necessary residues in RGS19: L118 ($\alpha$4) and Q183 ($\alpha$7). Because these positions were most likely to contribute to differences in RGS19 protein flexibility, we chose to mutate each of these positions in RGS19 (L118D and Q183K) in order to replicate the residues in RGS4 and RGS8. The RGS19 L118D mutation significantly increased the melting temperature compared to WT (Fig 5.2A). This increased stability suggests that addition of a D118-K125 salt bridge interaction in RGS19 was successful. Attempts to add a salt bridge at the $\alpha$6-$\alpha$7 position in RGS19 by Q183K mutation, however, produced no significant change in either thermal stability or potency of inhibition. This correlates with the observation that the $\alpha$6-$\alpha$7 salt bridge (D130-K155 in RGS4 and D124-K149 in RGS8) was less stably maintained in simulations than was the $\alpha$4-$\alpha$5 salt bridge (D90-K124 in RGS4 and E84-R119 in RGS8). It is possible that hydrogen bonding contacts between Gln183 on $\alpha$7 and Asp158 on $\alpha$6 in RGS19 are sufficient to compensate for the lack of a salt bridge at this position. In light of these results, we found it unlikely that the difference between Gln183 in $\alpha$6 of RGS19 and the lysines found in RGS4 and RGS8 (K155 and K149 respectively) play a major role in the flexibility differences between these proteins. Rather, the salt bridge-forming residue on $\alpha$4 is a stronger driver of differences in protein flexibility.

To verify the importance of differences in the $\alpha$4-$\alpha$5 salt bridge in protein structure and function, we also tested whether the elimination of this salt bridge in RGS4 and RGS8

could reduce the thermal stability of the proteins. The acidic residues D90 (RGS4) and E84 (RGS8) were each mutated to a leucine, which is the corresponding residue in RGS19. As expected, this significantly reduced thermal stability in RGS8, suggesting a successful removal of this salt bridge (Fig 5.2B). However, while the thermal stability of the RGS4 D90L mutant was significantly altered compared to WT RGS4, the mutation resulted in a biphasic melt curve shape with increased overall thermal stability (Fig 5.2C). This may indicate that the mutation had a more complex effect on the protein structure and its dynamics than a simple disruption of the salt bridge.

To determine the effects of mutations in salt bridge-forming residues on protein dynamics, both an in silico approach (all-atom MD simulations) and an experimental approach (hydrogen-deuterium exchange) were employed. MD simulations demonstrated that the addition of a salt bridge by the L118D mutation in RGS19 drastically reduces fluctuations throughout the $\alpha$3-$\alpha$8 helical bundle. This stabilization was particularly dramatic in the $\alpha$6 helix and the $\alpha$5-$\alpha$6 interhelical loop (Fig 5.3C). Likewise, in HDX analysis, decreases in DI were observed in the RGS19 L118D mutant compared to WT. As expected, the highest magnitude of decrease was observed in the fragment containing the mutation site, suggesting a strong reduction in solvent exposure at the $\alpha$4 helix. Other decreases in DI were most prominent at 1 to 3 minutes in fragments from the $\alpha$4 and $\alpha$5 helices and the $\alpha$5-$\alpha$6 interhelical loop (Fig. 5.5C). The $\alpha$4 helix changes are of particular interest given the location of the Cys123 as a target of the TDZD compounds. Conversely, mutations that eliminated salt bridges in RGS4 and RGS8 increased deuterium incorporation. Both mutants had increased DI in some fragments from the $\alpha$4 helix (Fig. 5.5A and 5.5B), but the RGS4 D90L mutant did not have increased DI in the fragment spanning the mutation site (Fig. 5.5A). This fits with the thermal stability data that suggest that the effect of the D90L mutation in RGS4 is more complex than disruption of a single interacting pair.

### 5.4.2 MD simulations to probe mutated RGS proteins conformations

In MD simulations, the RGS4 D90L and RGS8 E84L mutations did not have as large an effect on the magnitude of residue fluctuations as did mutation of L118D in RGS19 (Fig 5.3A and B). This may be because differences become apparent on shorter timescales in RGS19 than in RGS4 and RGS8, so simulations on $\mu$s timescales may not have captured all of the differences in dynamics caused by mutations in RGS4 (D90L) and RGS8 (E84L). Indeed, in HDX studies, stronger differences in DI were observed between RGS19 and RGS19 L118D at shorter timepoints (1 and 3 minutes) than in RGS4 D90L and RGS8 E84L (Fig 5.5A-C).

The $\mu$s-timescale MD simulations of mutated (D90L) RGS4 and WT RGS4 captured positive residue-residue ($C_\alpha$-$C_\alpha$) correlations between the $\alpha$4 and $\alpha$5 helices. This indicates a complex effect of the D90L mutation on the protein structure and dynamics, especially alongside data indicating increased thermal stability in mutant (D90L) RGS4. We also observed a change in the pattern of positive correlations between the $\alpha$4 and $\alpha$5 helices in the RGS8 E84L mutant from those of WT RGS8, possibly resulting from the removal of the interhelical salt-bridge. The marked increase in positive correlation between residues in the $\alpha$4 and $\alpha$5 helices in the RGS19 L118D mutant likely results from the introduced interhelical salt-bridge.

Finally, to determine how changes in protein flexibility affected the potency of inhibition by an RGS inhibitor, we used a flow-cytometry based protein-protein interaction assay (FCPIA) to evaluate the inhibition of G$\alpha$ binding by CCG-50014. Importantly, manipulation of RGS protein flexibility induced the expected changes in the potency of inhibition by TDZD covalent modifiers. The inhibition curve of the RGS19 L118D mutant was shifted rightward in comparison to RGS19, indicating that the more flexible mutant is less potently inhibited. Conversely, the inhibition curves for the RGS4 D90L mutant and the RGS8 E84L mutant were shifted leftward in comparison to those for RGS4 and RGS8, respectively. Thus,

enhancing flexibility by removal of salt bridge-forming residues increased the potency of inhibition by CCG-50014. These results support a causal relationship between the protein flexibility and the potency of inhibition.

## 5.5 Conclusion

In conclusion, differences in flexibility among RGS isoforms are capable of driving differences in potency of inhibition of a covalent inhibitor, CCG-50014. The differences in isoform flexibility in turn are strongly influenced by the presence or absence of an $\alpha$4-$\alpha$5 salt bridge and manipulation of this salt bridge is sufficient to induce changes in inhibitor potency among single-cysteine RGS proteins. Developing a deeper understanding of these differences in flexibility may enable the development of a new generation of RGS inhibitors with novel specificities.

## 5.6 Publications

The work described in this chapter has submitted in the following journal article:

- Shaw, V.S., Mohammadi, M., Quinn, J.A., Vashisth, H., and Neubig, R.R. (20XX). An Interhelical Salt Bridge Controls Flexibility and Inhibitor Potency For Regulators of G-protein Signaling (RGS) Proteins 4, 8, and 19. Molecular Pharmacology, (Submitted).

# CHAPTER 6

# OTHER SYSTEMS STUDIED: PHOSPHODIESTERASE 4 (PDE4) ENZYME

## 6.1 Abstract

Novel chemical controls are needed that selectively target human, animal, and plant parasitic nematodes with reduced adverse effects on the host or the environment. We hypothesize that the phosphodiesterase (PDE) enzyme family represents a potential target for development of novel nematicides and anthelmintics. MD simulations were used to evaluate differences in binding interactions of these inhibitors within the PDE4 catalytic domain. Of 32 residues within 5 Å of the ligand binding site, five revealed significant differences in non-bonded interaction energies that could account for the differential binding affinities of roflumilast and zardaverine. One site (Phe506 in human PDE4, Tyr253 in the *C. elegans* PDE4 catalytic domain) is predicted to alter the binding conformation of roflumilast and zardaverine (but not IBMX) into a less energetically favorable state. These results support the feasibility of designing the next generation of anthelmintics/nematicides that could selectively bind to nematode PDEs.

## 6.2 Background

Cyclic nucleotide metabolism is of central importance for a wide range of physiological processes in nematodes, as attested by the presence in *C. elegans* of 38 genes that synthesize cAMP or cGMP [217], as well as six PDE genes [218, 219, 220]. The cAMP second messenger has been linked to various behaviors including feeding and locomotion [221, 222]. Neural

pathways responsible for sensory signaling (chemoreceptors, thermotaxis, phototaxis) appear to be regulated through cGMP signaling pathways [218, 223, 224, 225]. Furthermore, RNAi, gene deletion, and pharmacological studies have shown that altered PDE activity can cause lethality, sterility, aberrant locomotion, lethargy, and altered development in *C. elegans* [226, 227, 228, 229, 230]. Some of the observed phenotypes (e.g., lethality, sterility) are clearly relevant to developing effective anthelmintics nematicides that specifically target parasitic nematodes-but not other animal phyla. Indeed, PDEs and their inhibitors are being investigated for their therapeutic potential in combatting various protozoal diseases [231]. Another advantage to targeting phytoparasitic nematode PDEs for the development of nematicides is the greatly reduced likelihood that a PDE inhibitor-based nematicide would have adverse effects on plants, since Class I PDEs have not been identified to date in plants [232].

The Class I PDE superfamily in vertebrates consists of eleven PDE families that have been identified throughout the animal kingdom [233]. The eleven families are distinguished by differences in their substrate specificity, modes of regulation, pharmacological properties, and tissue distribution [234]. However, all Class I PDE enzyme families share a conserved Prosite domain signature (Prosite PS00126[1]) in the catalytic domain consisting of the amino acid sequence pattern HD[LIVMFY]xHx[AG]xx[NQ]x[LIVMFY]. The crystal structures of the catalytic domains of almost all the PDE families have been solved, providing atomic-level details on the enzymatic and pharmacological properties of this enzyme superfamily [108]. The catalytic domains of the Class I PDE superfamily are made up of 330 amino acids whose secondary structure consists of 16 $\alpha$-helices. These $\alpha$-helices create three subdomains [109], which form a deep catalytic pocket at their center. The active site is composed of two sub-pockets, which bind two divalent metal ions and the substrate, respectively [108]. Zinc

[1]https://prosite.expasy.org/PDOC00116

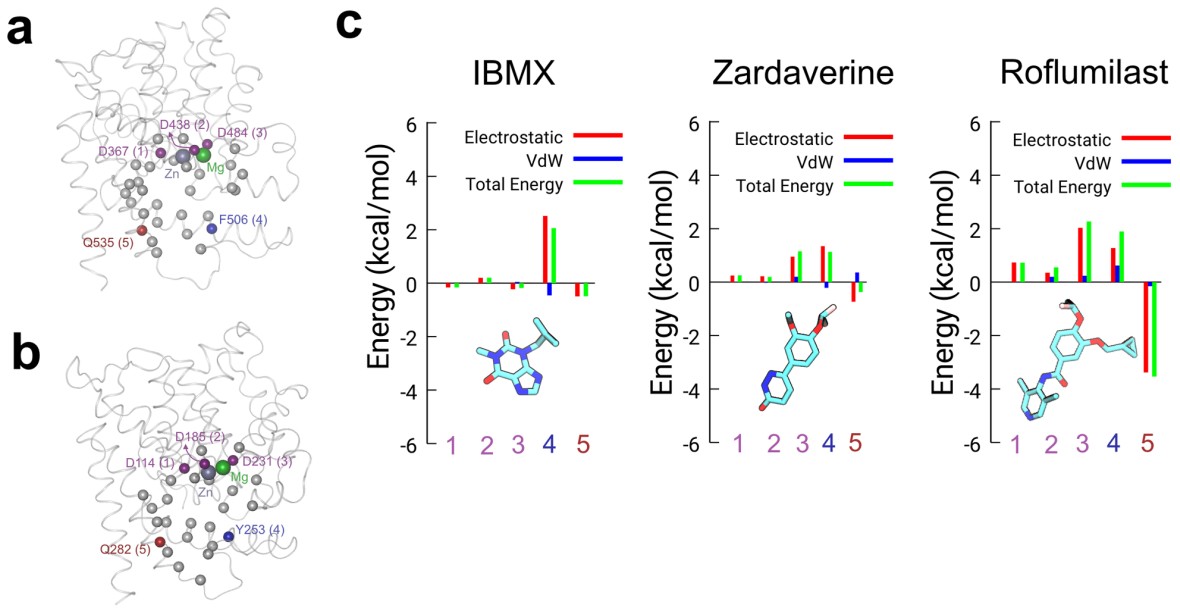



**Figure 6.1.** Interactions of five key $C_\alpha$-atoms of the 32 residues that interact with ligands in (a) human PDE4D and (b) *C. elegans* PDE4 binding pocket. The Zn and Mg ions are shown as gray and green spheres, respectively. The three Asp residues coordinating with Zn and Mg ions are highlighted by purple spheres. The Gln and Phe/Tyr residues are shown as red and blue spheres, respectively. The protein backbone is represented as ribbons, the $C_\alpha$-atoms of residues of the binding pocket and ions are shown as space-filling. (c) Changes in total non-bonded interaction energy and its components for *C. elegans* PDE4 relative to human PDE4D are shown for selected residues in the binding pocket (labeled 1-5 in panel a and b).

and magnesium ions are stabilized by conserved His and Asp residues in the metal binding pocket [108]. The crystal structure of human PDE4 catalytic domain in a complex with 5′-AMP [235] has revealed that cyclic nucleotides are stabilized by ionic interactions with the bound divalent cations and with Asp and His residues in the metal binding pocket, as well by hydrophobic interactions with conserved Gln and Phe residues in the hydrophobic pocket. The invariant Gln residue of PDEs have been shown to be critical for substrate and inhibitor binding [108].

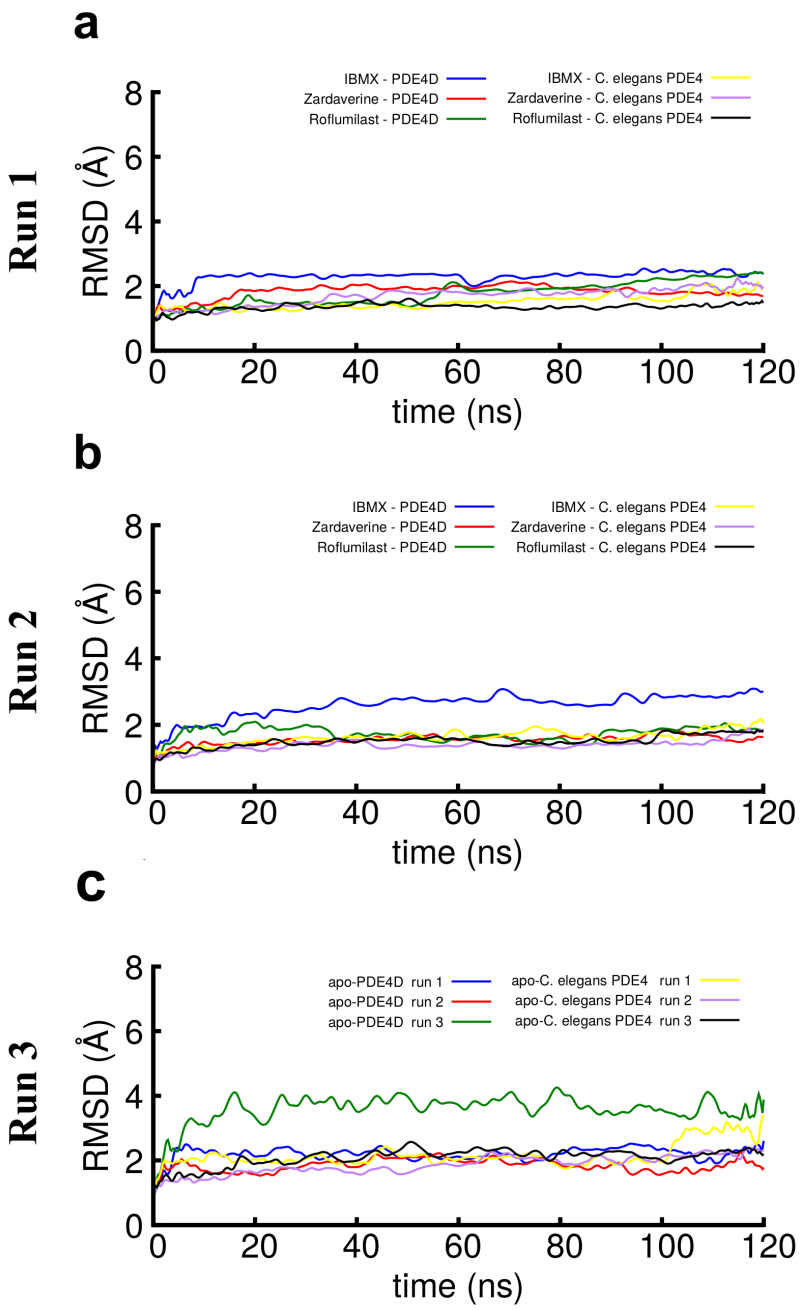



**Figure 6.2.** The traces of root-mean-squared-deviation (RMSD) vs. simulation time (ns) for PDE4D and *C. elegans* PDE4. (a and b) Two independent simulation runs for complexes of human PDE4D and *C. elegans* PDE4 with IBMX, zardaverine, or roflumilast. (c) RMSD traces of three independent simulation runs of apo-PDE4D and apo-*C. elegans* PDE4.

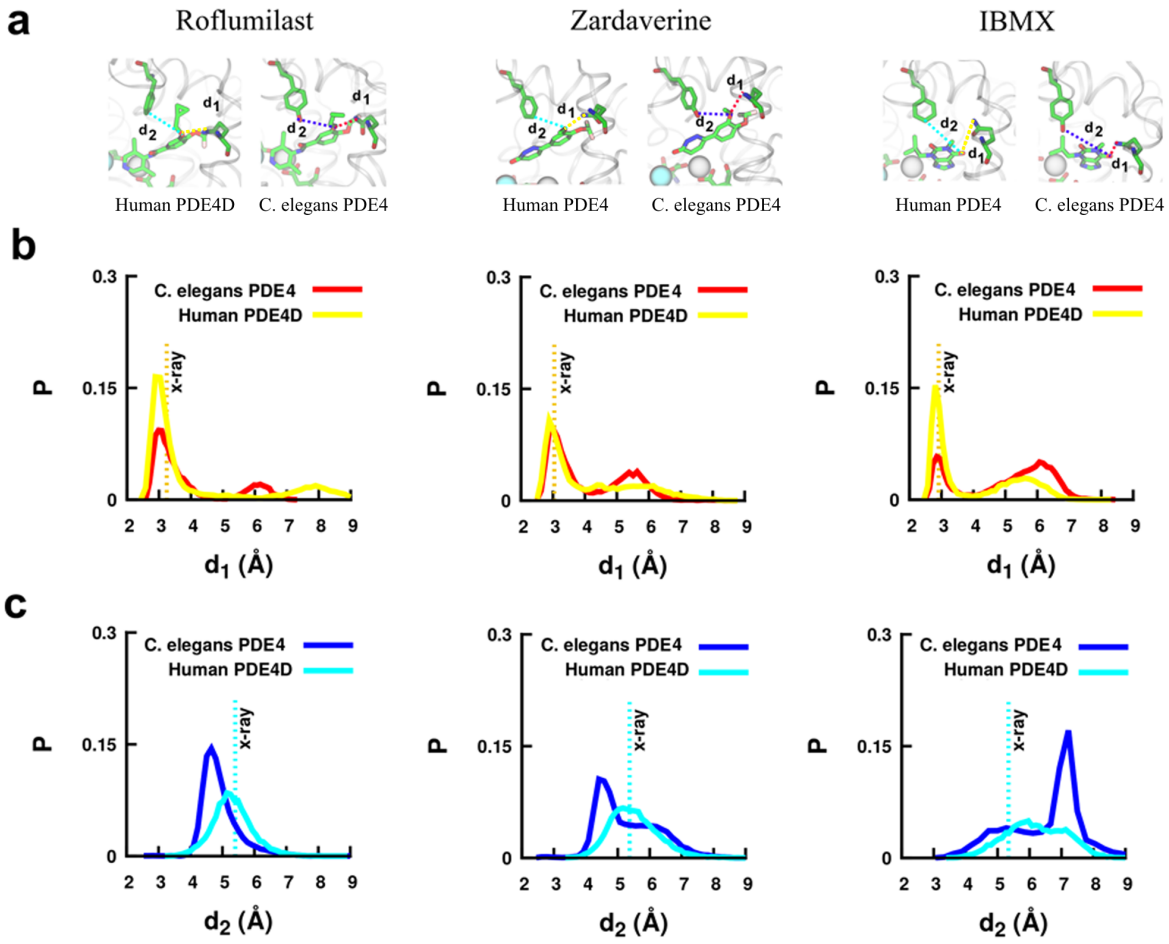



**Figure 6.3.** Probability (P) distributions of interatomic distances between ligand (O4 atom in roflumilast or zardaverine, or O6 atom in IBMX) and binding pocket residues. a) shows measurements of d1 [between the oxygen atom on the ligand and the N$\delta$ atom on Gln535(human)/Gln282(*C. elegans*)] and of d2 [between the oxygen atom of the ligand and the C4 atom of Phe506 (human)/O atom on the side chain of Tyr253(*C. elegans*)]. b) illustrates the distributions for distance d1 for *C. elegans* PDE4 (red) and human PDE4D (yellow). c) shows the distributions for the distance d2 for *C. elegans* PDE4 (blue) and human PDE4D (cyan). Vertical dotted lines in panels b and c indicate the distances in the crystal structures of inhibitors bound to human PDE4D.

**Hypothesis:** PDEs in parasitic nematodes represent a viable target for anthelmintic or nematicidal compounds.

The extensive literature on human PDE inhibitor pharmacology [236] and the commercial availability of many types of family-specific PDE inhibitor compounds enabled us to experimentally evaluate the potential of PDE inhibitors to serve as chemical nematicides targeting parasitic nematodes. In this chapter, I have used atomistic MD simulations (an approach used previously to compare the binding of inhibitors to different molecular human PDE4 isoforms [237, 238]) to investigate the role of structural differences in inhibitor interactions in human and nematode PDE4 that underlie the different pharmacological properties of nematode and human PDE4. These results support the idea that differences in the inhibitor binding site of nematode PDEs can be exploited to rationally design nematode-selective PDE inhibitors that act as an anthelmintic or nematicide without adverse effects on vertebrate animals or crops.

## 6.3 Methods

The initial coordinates for protein structures were obtained from the crystallographic structures of PDE4D bound to ligands with PDB codes: 1ZKN [239], 3G4L [240] [241], and 1MKD; in each instance, only one polypeptide chain of the overall structure was modeled, since the crystal structures of PDE4D consisted of identical PDE catalytic domains that co-crystallized into an oligomeric crystal. Numbering of amino acid residues for human PDE4D sequences used for MD simulations reflect the PDE4D 3 amino acid sequence. For simulation studies of *C. elegans* PDE4 with bound inhibitors, we created structural models using SwissModel[2] [242] with each of the previously mentioned PDE4D crystal structures as templates.

Upon comparing the *C. elegans* homology models with their templates, we observed that each model had initial mean-squared deviations (relative to their templates) of 0.12 Å,

[2]https://swissmodel.expasy.org/

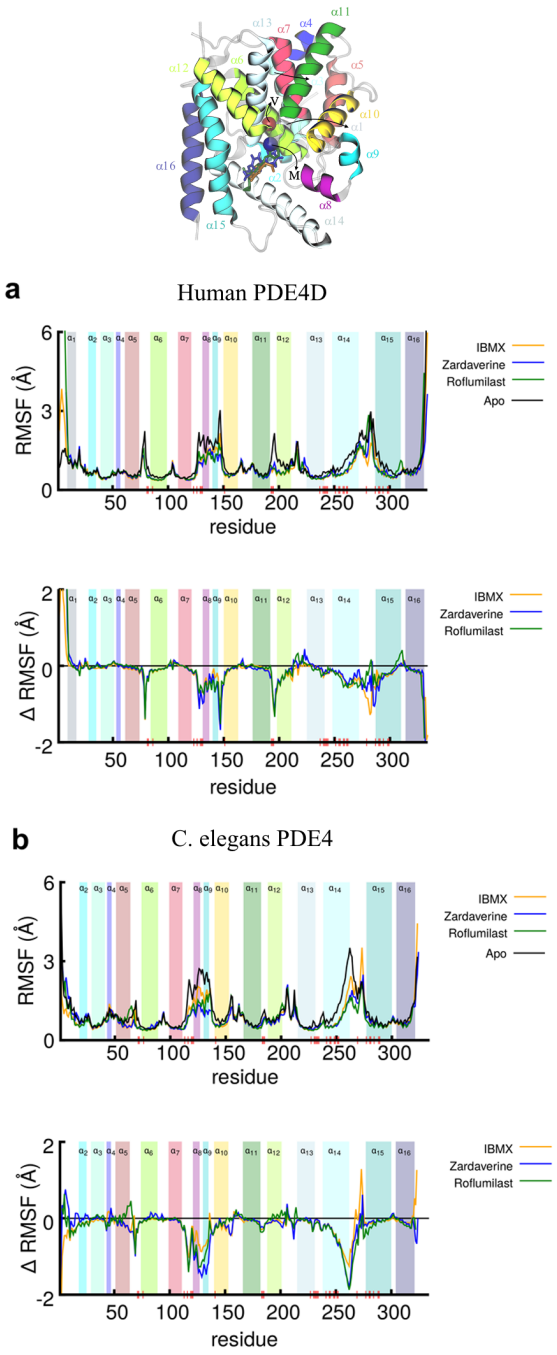



**Figure 6.4.** Root-mean-squared-fluctuation (RMSF) per residue (top panel) and the change in RMSF (ΔRMSF) per residue (bottom panel) are shown for (a) human PDE4D and (b) *C. elegans* PDE4 complexes with IBMX, zardaverine, and roflumilast. The superimposed structures for human PDE4D/*C. elegans* PDE4 along with superposition of IBMX, zardaverine, and roflumilast ligands (sticks) are shown at the top. The colored helices and vertical bars labeled $\alpha 1$ through $\alpha 16$ highlight the location of residues in the 16 helices in the catalytic domain. The Val334/81 and Met439/186 residues for human/*C. elegans* PDE4 are shown by red and blue spheres, respectively.

indicative of the homology models having a high structural similarity to the human PDE4 crystal structures. The stability of homology models was further tested using all-atom and explicit-solvent MD simulations. The stability of the *C. elegans* structural homology models in our MD simulations [in conjunction with the highly conserved nature of the PDE catalytic domain structure (PDEase_I; Pfam PF00233)] support their usefulness in the absence of experimentally determined structures.

We prepared six systems for MD simulations: three each for human PDE4D and *C. elegans* PDE4 such that each of the starting structures had a PDE inhibitor bound to it. Each system was then solvated with explicit TIP3P [205] water molecules, and charge-neutralized with counter-ions resulting in various system sizes. We used the software NAMD [111] for all MD simulations, and VMD [110] for system setup and post-processing analysis. CHARMM36 [238] force field was used including the CMAP correction [122, 123] for protein structures, and developed force-fields for all inhibitors using MATCH [125].

We used periodic boundary conditions [243] and computed long range-electrostatics using the particle-mesh Ewald summation [244] with a grid spacing of 1 Å, an integration time-step of 2 fs, and a cutoff-distance of 10 Å for van der Waals interactions; these settings are typically used for conducting MD simulations of solvated systems of proteins using NAMD [111]. We first energy minimized each system, and continued production runs of each system in the NPT-ensemble for 120 ns using a Langevin thermostat and Nosé-Hoover barostat [133, 134, 135]. We also carried out an independent run with the same length of simulation for each system, giving two production runs for each prepared system. Additionally, we carried out simulations of the same length for the apo states of human PDE4D and *C. elegans* PDE4.

### 6.3.1 Nonbonding interaction energy calculations

To investigate the role of individual amino acids in the binding pocket of each protein/ligand complex, we computed non-bonded interaction energies between all atoms of ligands and those of residues forming the binding pocket (i.e., within 5 Å from each ligand). Interaction energy values were estimated by splitting them into electrostatic and van der Waals interactions, as follows:

$$\Delta E_{non-bonded} = \Delta E_{elec} + \Delta E_{vdW} \tag{6.1}$$

We carried out these calculations by including all frames in each MD trajectory.

### 6.3.2 Analysis of salt-bridging interactions

The salt-bridging interaction analysis was carried out using VMD based on a distance criterion uniformly applied to determine the existence of salt-bridges for each frame in all trajectories [14]. Specifically, the formation of a salt-bridging interaction was considered if the distance between any of the oxygen atoms of acidic residues and the nitrogen atoms of basic residues were within a cut-off distance of 3.2 Å.

## 6.4 Results and Discussion

### 6.4.1 MD simulations to predict inhibitor binding conformations

To investigate the mechanistic details of differences in binding of each inhibitor, we performed two independent sets of MD simulations (120 ns each) of human and *C. elegans* PDE4 with each inhibitor (IBMX, zardaverine and roflumilast) (Fig. 6.1), and also carried out simulations of each enzyme without inhibitors. The root-mean-squared-deviation (RMSD) measured relative to initial structures revealed deviations below 2 Å indicating stable structures for both enzymes (Fig 6.2). Through visual analyses of these simulations,

we identified 32 residues in the immediate vicinity of bound ligands (defined as within 5 Å of any of the inhibitors; Figs 6.1a-b) as forming a binding pocket and then computed interaction energies of inhibitors with each of these 32 residues. In Figs H.1 and H.2, we present non-bonded interaction energies (van der Waals and electrostatic) for each of the 32 residues where energies were computed based upon all atoms of each residue and of the inhibitor molecule. These analyses for human PDE4D and *C. elegans* PDE4 resulted from two independent sets of simulations. From our interaction energy analyses, we identified five key residues showing differences between the human and *C. elegans* PDE4 (Fig 6.1a-b): (a) three conserved Asp residues (residues 367, 438, and 484 in human PDE4D corresponding to residues 114, 185, and 231 in *C. elegans* PDE4; purple spheres) that are critical for the coordination of the zinc and magnesium ions; (b) a conserved Gln residue (Gln535 in human PDE4D and Gln282 in *C. elegans* PDE4; red spheres) that stabilizes ligand binding via non-covalent interactions; and (c) a Phe residue in human PDE4D (Phe506; blue sphere) and a Tyr residue at the same position in *C. elegans* (Tyr253; blue sphere) that show differences in non-bonded interactions. Fig 6.1c presents the differences in the total non-bonded interaction energy and its components ($\Delta E$) for *C. elegans* PDE4 relative to the human PDE4D at these five sites. A positive value of $\Delta E$ indicates a higher non-bonded interaction energy of a given residue with the inhibitor in *C. elegans* in comparison to human PDE4D, and a negative value of $\Delta E$ indicates a lower, non-bonded interaction energy. We observed positive $\Delta E$ values for the three conserved Asp residues for roflumilast and, to a lesser extent, zardaverine, indicating stronger interactions in *C. elegans* relative to human PDE4D. In contrast, for IBMX the $\Delta E$ values between *C. elegans* PDE4 and human PDE4D are comparable. For all three inhibitor complexes with *C. elegans* PDE4, Tyr253 showed higher nonbonded interaction energy with the inhibitors in comparison to the corresponding Phe506 residue in human PDE4D. Based on the interaction energy analysis, we observed a correlation between the change in the interaction energy at the non-conserved and conserved residue sites. Primarily for roflumilast and to a lesser extent for zardaverine and IBMX, we

observed that an increase in the total non-bonded interaction energy at the non-conserved site (F506 in human vs. Y253 in *C. elegans*; labeled as the residue 4 in Fig 6.1) is correlated with a decrease in the total non-bonded interaction energy at the conserved site (Q535 in human vs. Q282 in *C. elegans*; labeled as the residue 5 in Fig 6.1). Similar correlation was observed between an increase in the total non-bonded interaction energy at the conserved site (D484 in human vs. D231 in *C. elegans*; labeled at the residue 3 in Fig 6.1) and a decrease at the conserved site (Q535 in human vs. Q282 in *C. elegans*; labeled as the residue 5 in Fig 6.1).

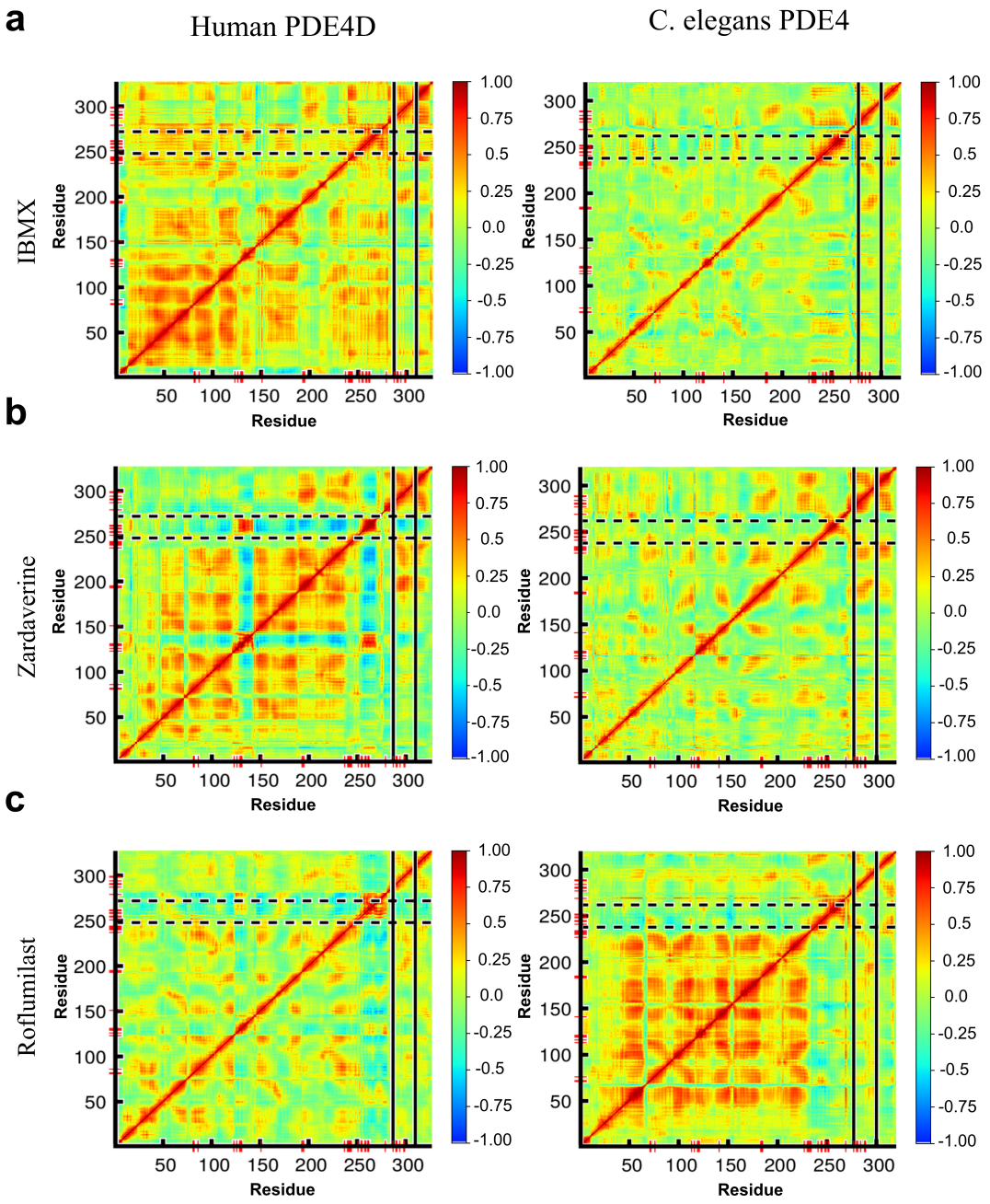



**Figure 6.5.** Dynamic cross correlation matrices calculated for the C$\alpha$ atoms of human PDE4D and *C. elegans* PDE4 complexed with IBMX (a), zardaverine (b), and roflumilast (c). Residues in the $\alpha$14 and $\alpha$15 helices are shown by areas between dashed-lines and solid-lines, respectively. Red tick-marks on the axes represent the 32 residues in the binding site (as depicted in Fig. 6.1a-b). The color scheme ranges from anticorrelation (-1.0, blue), no correlation (0, green), and positive correlation (+1.0, red). Values are the average for the two independent simulation runs.

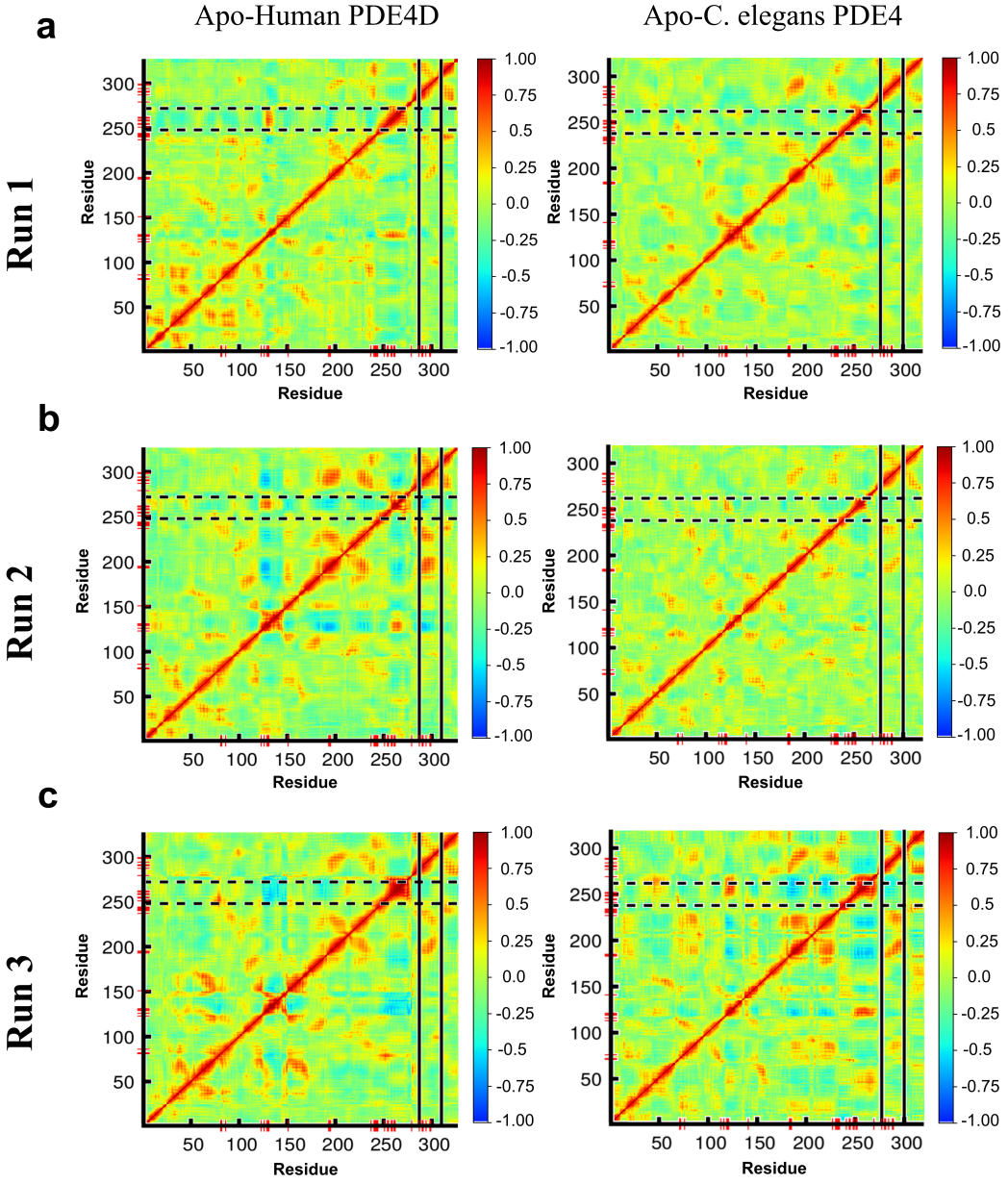



**Figure 6.6.** Dynamic cross correlation matrices calculated for the C$\alpha$ atoms of human PDE4D and *C. elegans* PDE4 in their apo state. Color scheme is the same as for Fig. 6.5. Panels a-c represent three independent simulations.

To investigate the variation in the docked positions of ligands in the binding pockets of human and *C. elegans* PDE4, I computed the interatomic distances between specific atoms

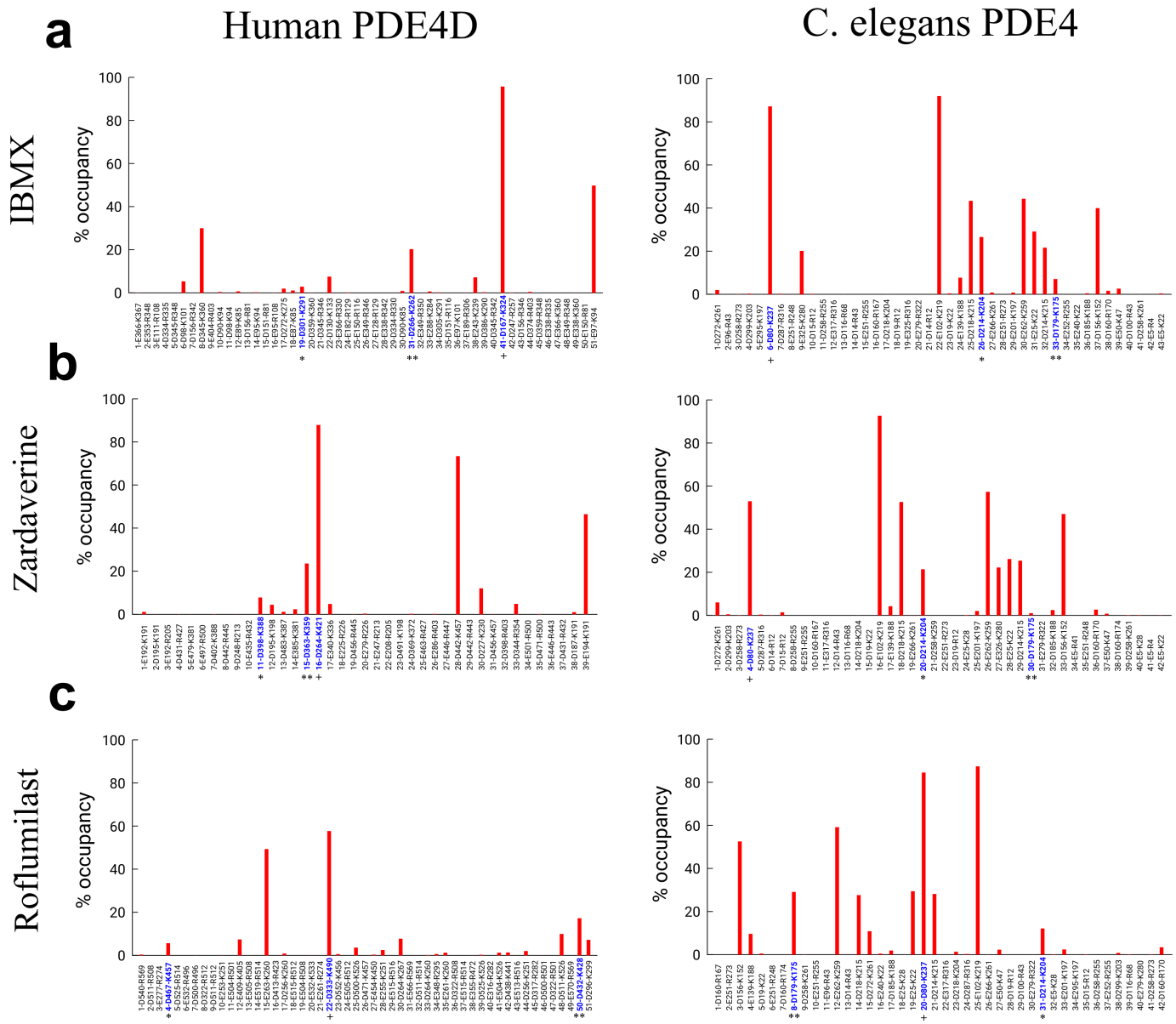



**Figure 6.7.** Key salt-bridging interactions are shown based upon the first set of MD simulations of human PDE4D and *C. elegans* PDE4 with IBMX (a), zardaverine (b), and roflumilast (c). Three conserved salt-bridges are labeled in blue.

in the ligands and the nearby Gln535(human)/Gln282(*C. elegans*) residues (d1 in Fig 6.3). For *C. elegans* PDE4, the distributions of d1 are bimodal (red traces in Fig 6.3b) for all three inhibitors, with roflumilast and zardaverine having a higher probability of being in

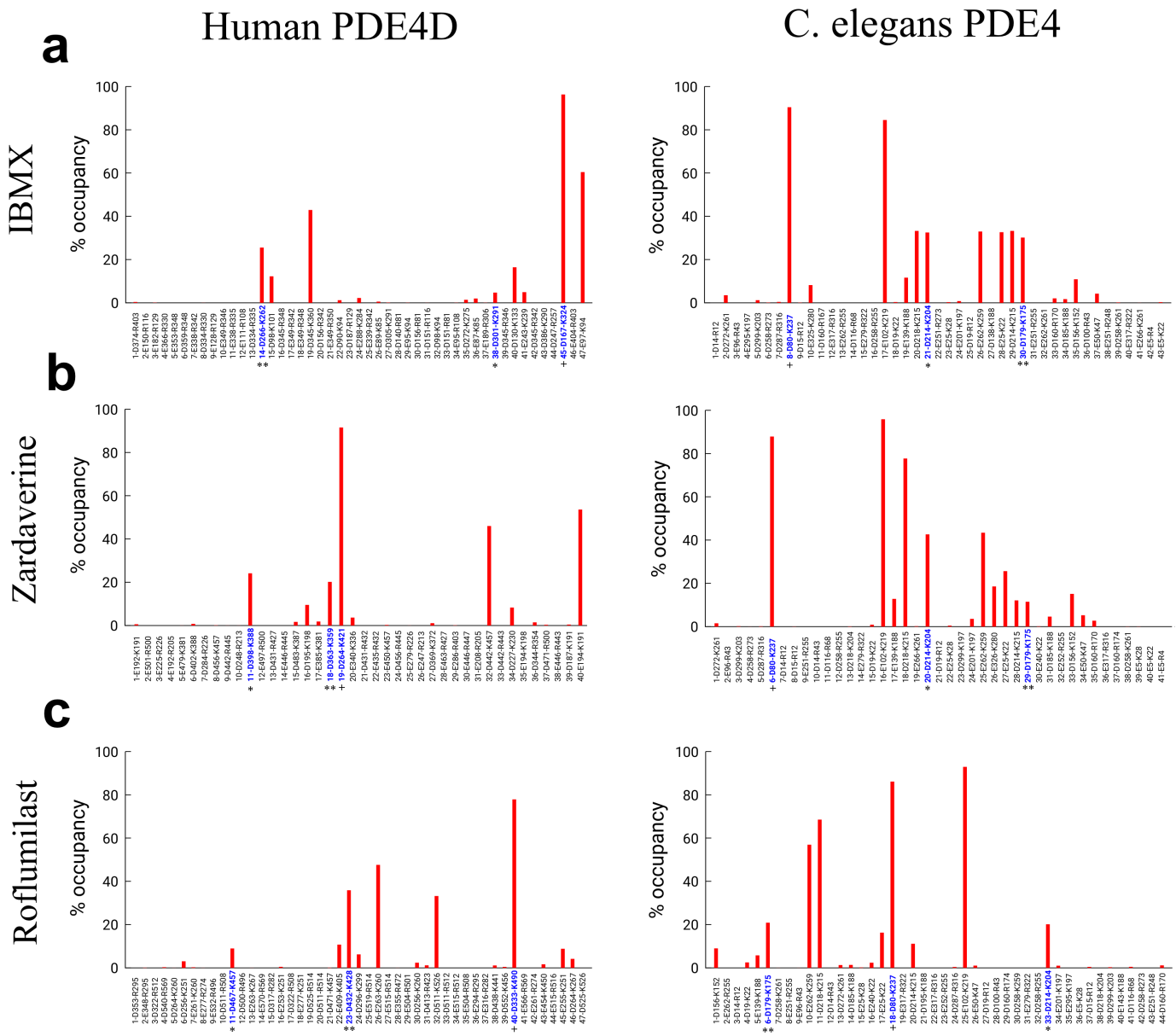



**Figure 6.8.** Data similar to Fig. 6.7 are shown for a second set of MD simulations with the three inhibitors.

states with d1∼ 3Å, while for IBMX both states at d1∼ 3Å and ∼ 6Å are equally probable. For human PDE4D, all inhibitors show an increased probability of being in states at shorter distances (∼ 3Å) but bimodal distributions with lower probabilities of states at larger dis-

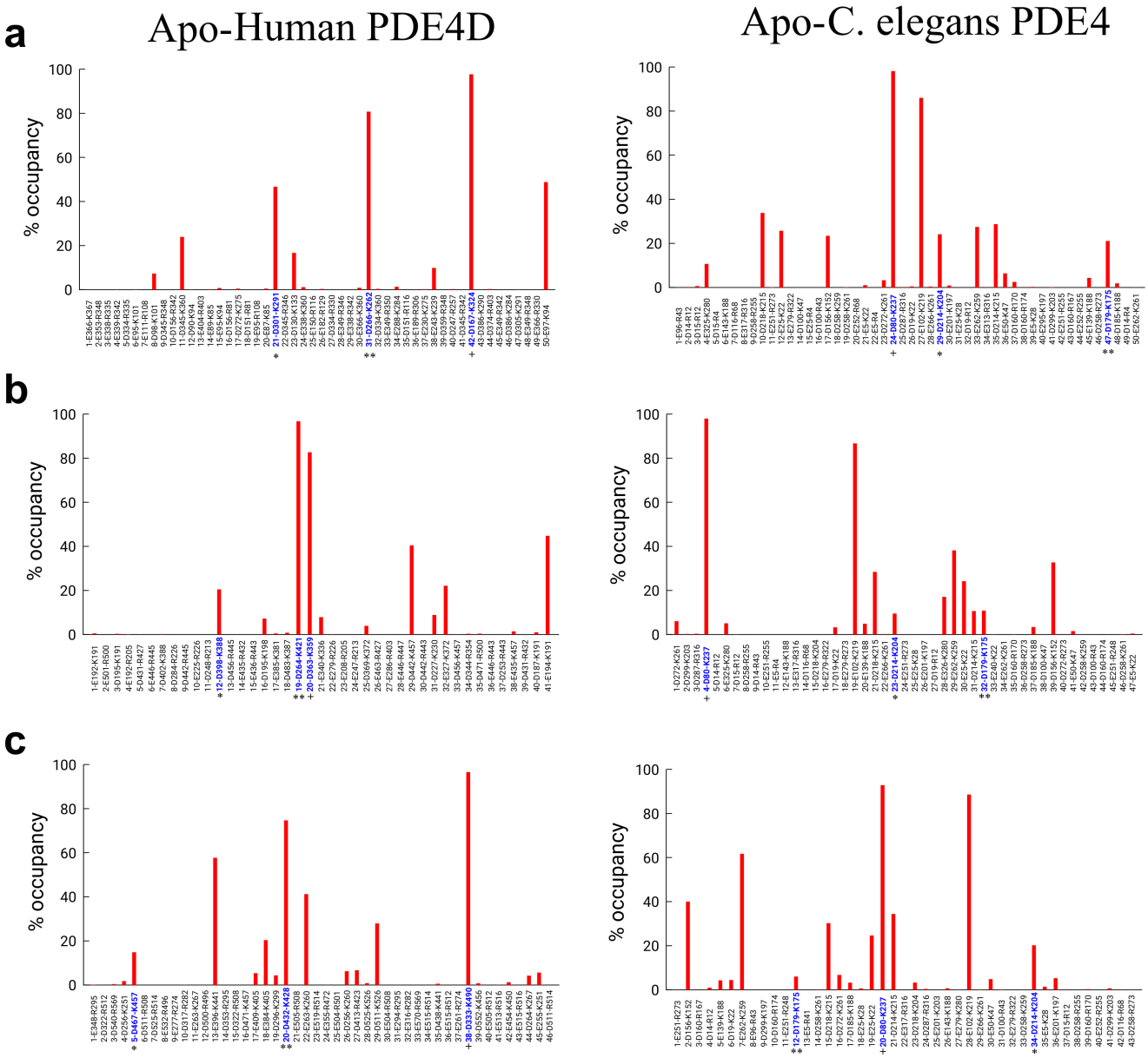



**Figure 6.9.** Data similar to Fig. 6.7 are shown for three independent sets of MD simulations of apo human PDE4D and apo *C. elegans* PDE4.

tances are observed for zardaverine and IBMX. These observations suggest that zardaverine and IBMX are more likely to transition between two distinct states within the binding pocket

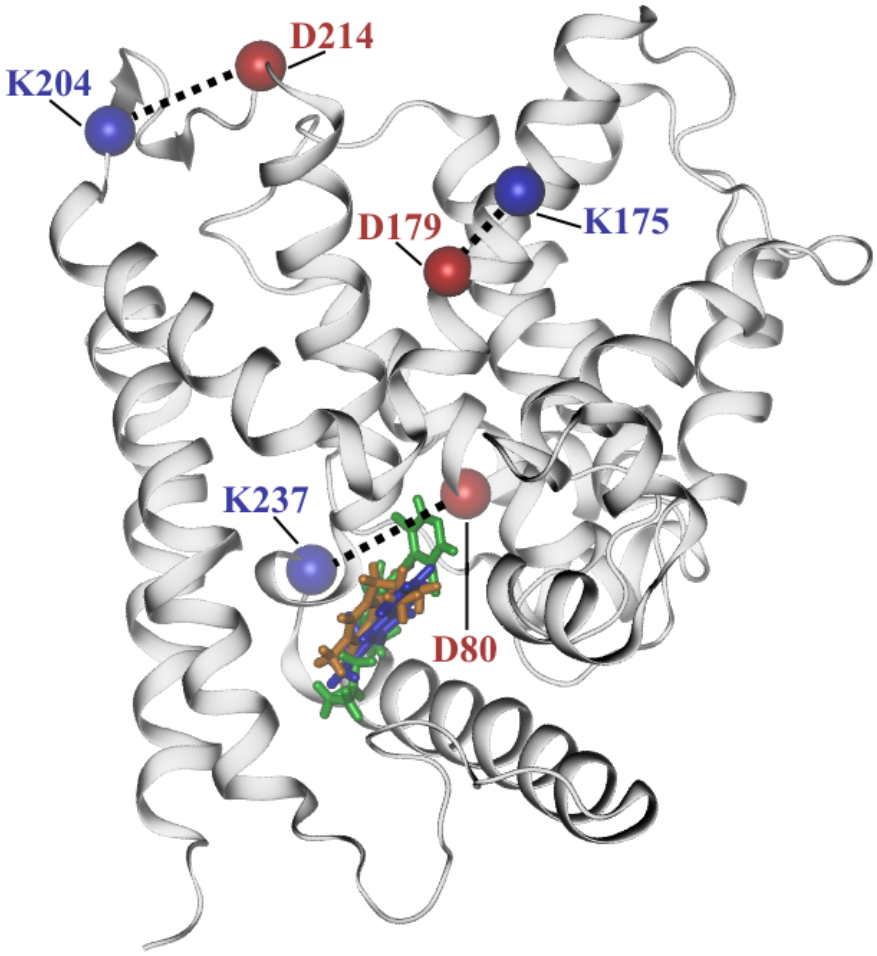



**Figure 6.10.** *C. elegans* PDE4 catalytic domain illustrating three conserved salt-bridges. Residues participating in each salt-bridge are colored and labeled. The three inhibitors are shown as sticks.

in comparison to roflumilast, which appears to be stably bound, largely in a single state.

In addition, we measured the interatomic distance (d2 in Fig 6.3) between the oxygen atom of the inhibitors to the side-chains of Tyr253 (*C. elegans*) or Phe506 (human PDE4D). In *C. elegans* PDE4, the distance distributions are bimodal and span a larger distance range

418 (∼3-9 Å) for zardaverine and IBMX in comparison to a unimodal and narrower (∼4-6 Å) distribution for roflumilast. In human PDE4D, the distributions are unimodal for roflumilast and zardaverine and bimodal for IBMX, with mean values distinct from those in *C. elegans* PDE4. Overall, the measurements on these distances suggest distinct positioning of inhibitors in the proximity of Phe506 and Gln535 residues in human PDE4D and the corresponding residues Tyr253 and Gln282 in *C. elegans*. Specifically, roflumilast is significantly more stable than other inhibitors in the binding pocket of the human PDE4D and showed a higher non-bonded interaction energy with the Gln535 residue in human PDE4D (Fig 6.1c).

To further probe per-residue perturbations on binding of each inhibitor in both enzymes, we have computed the per-residue root-mean-squared fluctuation (RMSF) of the liganded enzyme structures (top panels in Figs 6.4a-b) and the change in per-residue RMSF relative to their unliganded apo-forms (ΔRMSF) (bottom panels in Figs 6.4a-b). Among binding pocket residues, we observed that ligand binding increased fluctuations in Val334 and Met439 in human PDE4D (corresponding to Val81 and Met186 in *C. elegans* PDE4). However, the residues located in loops connecting $\alpha$5-$\alpha$6, and $\alpha$11-$\alpha$12 helices are more stabilized by the ligands in human PDE4D in comparison to *C. elegans* PDE4. The residues located in the M loop between $\alpha$8 and $\alpha$9 helices are more stabilized in *C. elegans* PDE4 by zardaverine and roflumilast and to a lesser extent by IBMX in comparison to human PDE4D. Residue Phe506 (human)/Tyr253 (*C. elegans*) is located in $\alpha$14-helix which appear more stabilized by ligands in *C. elegans* PDE4 in comparison to human PDE4D. Residue Gln535 (human)/Gln282 (*C. elegans*) is located in the $\alpha$15-helix which is perturbed to a greater extent in human PDE4 than *C. elegans* PDE4 (Fig 6.4). The fluctuations in residues of the binding pocket as observed in the RMSF analyses are correlated with the analyses of non-bonded interaction energies.

To further investigate whether the higher flexibility of the $\alpha$14-helix in human PDE4D complexes with bound ligands affects the motion of the residues belonging to the $\alpha$15-helix, we calculated the dynamic cross-correlation matrix for the $C_\alpha$ atoms in all MD trajectories. For human PDE4D, the correlation matrices showed neither significant positive correlation nor significant anti-correlation between the residues of the $\alpha$14-helix (highlighted by the dashed-lines in Figs 6.5-6.6) and the residues of the $\alpha$15-helix (highlighted by the solid-lines in Figs 6.5-6.6). However, we find that the motion of residues in the $\alpha$14 and $\alpha$15 helices are marginally more correlated in *C. elegans* PDE4 in comparison to human PDE4D. The correlation between the $\alpha$14 and $\alpha$15 helices is mostly found between neighboring residues of Tyr253 (*C. elegans*) and Gln282 (*C. elegans*). We also observed (Fig 6.5) significantly higher positive residue-residue ($C_\alpha$-$C_\alpha$) correlation within human PDE4D complexes with IBMX and zardaverine in comparison to *C. elegans* PDE4, whereas the complexes with roflumilast showed significantly lower positive correlation in human PDE4D in comparison to *C. elegans* PDE4. This indicates that roflumilast induces a different pattern of correlated motions in the protein backbone in comparison to IBMX and zardaverine, and comparable to those of the apo states (Fig 6.6).

To better understand the effect of ligands on the protein structure outside the binding site, we identified all possible salt-bridging interactions within human PDE4D and *C. elegans* PDE4 (Figs 6.7-6.10). Qualitatively, the salt-bridging interactions are observed to occur with a lower frequency in human PDE4D in comparison to *C. elegans* PDE4. We also found a smaller number of salt-bridging interactions in human PDE4D complexed with roflumilast, but these salt-bridging interactions were comparatively stable for longer times during simulations. Furthermore, we identified three salt-bridge pairs conserved between human PDE4D and *C. elegans* PDE4 with higher occupancy number: D179-K175 (*C. elegans*), D204-K214 (*C. elegans*), and D80-K237 (*C. elegans*) (Fig 6.7, and labeled in Figs 6.8 and 6.9). D179-K175 (*C. elegans*) is an intra-helical salt-bridge in the $\alpha$11-helix, and

D204-K214 (*C. elegans*) is in the loop connecting the $\alpha$12 and $\alpha$13 helices. The D179-K175 (*C. elegans*) interaction pair in the human PDE4D complexes with ligands showed a lower occupancy in comparison to apo human PDE4D. The occupancy of the D204-K214 (*C. elegans*) is higher in the apo-human PDE4D in comparison to *C. elegans* PDE4, whereas it has a higher occupancy in *C. elegans* PDE4 complexes with ligands in comparison to human PDE4D complexes with ligands (Figs 6.7-6.10). The D80-K237 (*C. elegans*) salt-bridge is located near the binding pocket, the D80 residue is in the $\alpha$6-helix and the K237 residue is in the loop connecting $\alpha$13 and $\alpha$14. I observed that the occupancy of the D80-K237 (*C. elegans*) salt-bridge was significantly suppressed by roflumilast in human PDE4D in comparison to zardaverine and IBMX, while the occupancy of this salt-bridge is not affected by the presence of IBMX and zardaverine (Figs 6.7-6.10).

Both dynamic cross-correlation analysis and salt-bridging interactions revealed allosteric effects of each ligand on the protein structure. Unlike IBMX and zardaverine, roflumilast induced distinct patterns of structural perturbations outside of the binding pocket for human PDE4D compared with *C. elegans* PDE4. Specifically, we observed lower residue-residue correlations for roflumilast in comparison to zardaverine and IBMX in human PDE4D in comparison to *C. elegans* PDE4. While we observed overall a smaller number of salt-bridging interactions in human PDE4D in comparison to *C. elegans* PDE4, the salt-bridging interactions in human PDE4D were significantly more stable for roflumilast. In contrast, roflumilast significantly perturbed some salt-bridging interactions (D88-K237) more than zardaverine and IBMX in *C. elegans* PDE4. Therefore, we suggest that modulation of salt-bridging interactions could be one of the factors that contribute to an altered binding affinity of an inhibitor among different protein isoforms. Taken together, these structural analyses provide a molecular basis for better understanding differential binding of inhibitory compounds in the human PDE4 versus *C. elegans* PDE4 catalytic domain.

### 6.4.2 Atomistic simulations provide insight into altered pharmacological properties

The substantial differences in the primary sequence of nematode and vertebrate PDEs and pharmacological results revealed significant changes in binding affinities of compounds that were designed for human PDE4D. To gain further insights into differences in the ligand binding sites of *C. elegans* PDE4 and human PDE4D that could explain the reduced affinity of *C. elegans* PDE4 for compounds optimized as human PDE inhibitors, I used homology models and all-atom explicit-solvent MD simulations. The use of homology models has been successfully used in previous studies to identify amino acid residues that are responsible for differences in binding of inhibitors to PDE5 and PDE6 [238]. From the 32 amino acid residues that we defined as constituting the inhibitor-binding site, only five sites differed between the two enzymes and four of those were conservative substitutions that preserved the polar or hydrophobic nature at its position and thus are unlikely to drastically change the binding conformation or energy. However, I observed differences in nonbonded interaction energies due to the movement of inhibitors in the vicinity of residue Phe506 in *H. sapiens* (corresponding to Tyr253 in *C. elegans*). Overall, this Tyr residue contributes significantly more total non-bonded interaction energy than the Phe in the same position for all three inhibitors (Figs H.1 and H.2), likely due to the hydrogen bonding that results from the addition of a hydroxyl group at the 4-C of the aromatic ring.

For IBMX, the Phe to Tyr substitution appears to have little impact on the overall binding of IBMX to either human PDE4D or *C. elegans* PDE4. This is likely a result of IBMX interacting solely with the hydrophobic pocket of PDE4, as suggested by IBIS [245, 246]. Despite the polar Tyr residue coordinating to the ketone at position 6 in the purine ring of IBMX in *C. elegans*, our MD results indicate that the interactions and conformation of IBMX are very similar in the two PDE4 catalytic domains. This is consistent with the ob-

| Inhibitor | Human $IC_{50}$ (μM) | *C. elegans* $IC_{50}$ (μM) | Fold difference |
|---|---|---|---|
| IBMX | 15.8 ± 1.7 (n=4) | 34.1 ± 8.7 (n=5) | 2 |
| Zardaverine | 1.9 ± 0.6 (n=5) | 146 ± 34 (n=8) | 77 |
| Roflumilast | 0.0046 ± 0.0006 (n=3) | 0.73 ± 0.13 (n=5) | 159 |

**Figure 6.11.** Enzyme activity was tested over a range of inhibitor concentrations with 1 $\mu$M cAMP substrate 323 concentration. The dose-response relationship was fit to a 3-parameter logistic equation to obtain the 324 $IC_{50}$ and the standard error of the mean for the indicated number of experiments (This figure was created by my co-author Kevin Schuster).

served similarity in $IC_{50}$ values for IBMX with the two enzymes (Fig 6.11).

In contrast, the MD simulations suggest that the binding conformation of zardaverine or roflumilast are altered as a result of the substitution of Tyr253 for Phe506 at this site in the binding pocket of the two enzymes, consequently inducing a different pattern of correlated motions in the protein backbone. In *C. elegans*, the hydroxyl group of the Tyr residue co-ordinates strongly with the methoxyphenyl and cyclopropylmethoxyl group of zardaverine and roflumilast, respectively. This increase in energy contribution appears to result in a displacement of both ligands away from the hydrophobic sub-pocket that further disrupts stabilization by the conserved glutamine residues (Gln535 in human PDE4; Fig 6.1c).

It has been previously reported that Tyr495, Phe506, Gln535 are critical for stabilizing PDE4 inhibitors [237]. This disruption of the desired binding conformation in *C. elegans* could partially explain the reduced $IC_{50}$ values for these two compounds (Fig 6.11). While per-residue fluctuations are also found to be correlated with non-bonded interaction energy

analyses, other analyses (e.g. interatomic distances and residue-residue correlations) suggest that, among the three inhibitors, roflumilast is more stable in the binding pockets and induces a distinct pattern of correlated motions in comparison to zardaverine and IBMX. In summary, these MD analyses highlight the importance of considering not only differences in residue substitutions (e.g. Tyr253 in *C. elegans* vs. Phe506 in *H. sapiens*) but also allosteric perturbations and overall inhibitor stabilization of catalytic domain conformation in future efforts to design and optimize nematode-specific PDE inhibitor compounds using in silico approaches such as virtual screening and fragment-based drug design [247, 248].

## 6.5 Conclusion

MD simulations were used to support the hypothesis that the nematode PDE enzyme family differs sufficiently from the vertebrate PDE orthologs to validate the feasibility of developing PDE inhibitor compounds as potent and selective anthelmintics/nematicides. While analysis of the differences in the amino acid sequence or structure-activity relationships for selected PDE inhibitor compounds did not immediately identify which sites of interaction may have been disrupted in *C. elegans* PDE4 for inhibitors designed for human PDE4D, MD simulations revealed the importance of Phe506 (human)/Tyr253 (*C. elegans*) substitution and demonstrated that changes in the conformation of the catalytic domain may collectively lead to inhibitor discrimination in the binding pocket, based on the following analyses: (1) non-bonded interaction energy analysis; (2) changes in the ligand orientation in the binding pocket; (3) RMSF analysis; (4) cross correlation analysis; and (5) salt-bridge interaction analysis.

Collectively, the results indicate that future efforts to discover inhibitor compounds specifically targeting nematode PDE4 must take into consideration not only the molecular architecture of the inhibitor binding site, but also the conformational dynamics of the entire

catalytic domain of the enzyme. Insights gained from this study will advance efforts to rationally design inhibitor compounds that selectively and potently inhibit plant and animal parasitic nematode, PDEs to disrupt their lifecycle, thereby enhancing public health and agricultural productivity.

## 6.6 Publications

The work described in this chapter has resulted in the following journal article:

- Schuster, K.D., Mohammadi, M., Cahill, K.B., Matte, S.L., Maillet, A.D., Vashisth, H., and Cote, R.H. (2019). Pharmacological and molecular dynamics analyses of differences in inhibitor binding to human and nematode PDE4: implications for management of parasitic nematodes. PLoS ONE 14(3): e0214554.

# CHAPTER 7

# CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

In this thesis, I used MD simulations described in section 2.2 and enhanced sampling techniques such as TAMD described in section 2.4.1, and single-sweep method [141] described in sections 2.4.1, 2.4.1.1, and 2.4.1.2. Additionally, I used various conformational metrics calculations based on MD simulations (e.g. RMSD, RMSF, DCC, and non-bonded interaction energy). The overall understanding gained from work presented in this thesis could further guide mutagenesis studies in [FeFe]-hydrogenase or drug discovery efforts aimed at RGS proteins and nematode PDEs.

The primary outcome of this thesis work has been an enhanced understanding of the protein-inhibitor interactions in the studied cases. Specifically, the studies of protein conformational changes show transient states (beyond static protein states) that play a key role in inhibitor recognition and migration within protein-cavities. By characterizing the interconnected pathways of inhibitory gases within [FeFe]-hydrogenase, more effective mutations in residues could be performed for increasing the enzyme tolerance to inhibitory gases. The studies of inhibitory compounds of RGS4 protein provide enhanced view of the mechanism of action of TDZD inhibitors among isoforms of RGS proteins, and further elucidate the alternative binding mechanisms for these allosteric inhibitors and correlations between their potencies and the protein dynamics. The studies of human PDE4D inhibitors and *C. elegans* PDE4 provide enhanced insight into effects of inhibitors on the catalytic domain of the enzyme.

In chapter 3, I studied the interconnected network of pathways of inhibitory gases ($O_2$ and CO) within [FeFe]-hydrogenase. The differences in the free-energy levels of pathways between $O_2$ and CO show that any residue mutation directed toward one gas might affect the enzyme structure that may affect the interconnected network of the other gases (e.g. hydrogen). In addition to the challenge of increasing tolerance of the enzyme to inhibitory gases, one should consider that the successful mutagenesis studies would require characterization of pathways of other molecules that compete with $O_2$ and CO such as $H_2$ and $H_2O$ within the enzyme. Therefore, one future work may be the studies of interconnected network of pathways of $H_2$ and $H_2O$ within [FeFe]-hydrogenase. The final step in understanding the network of pathways of all of these molecules in the enzyme would be the computation of kinetic rates across each migration pathway for each gas/solvent molecule. One can conduct the calculation for finding the kinetic rates, and further enhance the experimental mutagenesis studies aimed at $H_2$ production from [FeFe]-hydrogenase.

In chapters 4 and 5, I studied the covalent inhibitors of RGS proteins that inhibit the RGS/$G_\alpha$ protein-protein interaction. Here, the computational studies were solely directed toward the RGS proteins and not when they are bound to $G_\alpha$ proteins. The differences in potency of the inhibitor molecules across isoforms of RGS proteins suggest similar differences may exist in the interaction of the RGS-inhibitor complex with $G_\alpha$ subunits. The final goal of these inhibitors is to prevent the formation of the RGS-$G_\alpha$ complex. Therefore, one future work may be the MD simulation of the RGS-$G_\alpha$ complex with and without inhibitor molecules. Another future work may be the studies to enhance our view of binding and unbinding rate [249] of the inhibitory compounds to RGS proteins.

In chapter 6, I studied the PDE inhibitor compounds designed for human PDE4D on *C. elegans* PDE4. In this study, I found key sites in the catalytic domain of the enzyme for interaction with the selected PDE inhibitory compounds. The knowledge of differences

in the conformation of the catalytic domain may be supplemented with the calculation of free-energy of binding of inhibitory molecules. The free-energy of binding of compounds can be computed using MM/PBSA and MM/GBSA methods [250]. Another future work may be the computation of network of pathways of inhibitory compounds within the catalytic domain. This study can be performed using the single-sweep method [141] described in sections 2.4.1, 2.4.1.1, and 2.4.1.2.

# BIBLIOGRAPHY


[1] Mohammadjavad Mohammadi and Harish Vashisth. Pathways and thermodynamics of oxygen diffusion in [FeFe]-hydrogenase. *J. Phys. Chem. B*, 121(43):10007–10017, 2017.

[2] James A Stapleton and James R Swartz. Development of an in vitro compartmentalization screen for high-throughput directed evolution of [FeFe] hydrogenases. *PLoS ONE*, 5(12):e15275, 2010.

[3] Yong Liu, Mohammadjavad Mohammadi, and Harish Vashisth. Diffusion network of CO in FeFe-hydrogenase. *J. Chem. Phys.*, 149(20):204108, 2018.

[4] Emma M Turner, Levi L Blazer, Richard R Neubig, and Stephen M Husbands. Small molecule inhibitors of regulators of G-protein signaling (RGS) proteins. *ACS Med. Chem. Lett.*, 3(2):146–150, 2011.

[5] Jordi Cohen, Kwiseon Kim, Paul King, Michael Seibert, and Klaus Schulten. Finding gas diffusion pathways in proteins: Application to $O_2$ and $H_2$ transport in CpI [FeFe]-hydrogenase and the role of packing defects. *Structure*, 13(9):1321–1329, 2005.

[6] Thomas Lautier, Pierre Ezanno, Carole Baffert, Vincent Fourmond, Laurent Cournac, Juan C Fontecilla-Camps, Philippe Soucaille, Patrick Bertrand, Isabelle Meynial-Salles, and Christophe Léger. The quest for a functional substrate access tunnel in [FeFe]-hydrogenase. *Faraday Discuss.*, 148:385–407, 2011.

[7] Bojana Ginovska-Pangovska, Ming-Hsun Ho, John C Linehan, Yuhui Cheng, Michel Dupuis, Simone Raugei, and Wendy J Shaw. Molecular dynamics study of the proposed

proton transport pathways in [FeFe]-hydrogenase. *Biochim. Biophys. Acta. Bioenerg.*, 1837(1):131–138, 2014.

[8] James A Stapleton and James R Swartz. A cell-free microtiter plate screen for improved [FeFe]-hydrogenases. *PloS ONE*, 5(5):e10554, 2010.

[9] James A Stapleton and James R Swartz. Development of an in vitro compartmentalization screen for high-throughput directed evolution of [FeFe]-hydrogenases. *PloS ONE*, 5(12):e15275, 2010.

[10] Adam Kubas, Christophe Orain, David De Sancho, Laure Saujet, Matteo Sensi, Charles Gauquelin, Isabelle Meynial-Salles, Philippe Soucaille, Hervé Bottin, Carole Baffert, Vincent Fourmond, Robert Best, Jochen Blumberger, and Christophe Léger. Mechanism of O2 diffusion and reduction in [FeFe]-hydrogenases. *Nat. Chem.*, 9:88–95, 2016.

[11] Maria L Ghirardi, Jordi Cohen, Paul King, Klaus Schulten, Kwiseon Kim, and Michael Seibert. [FeFe]-hydrogenases and photobiological hydrogen production. In *Proceedings of SPIE-The International Society for Optical Engineering, San Diego, CA, Aug 14-17, 2006*, number 6340. International Society for Optics and Photonics.

[12] P. King, M.L. Ghirardi, and M. Seibert. Oxygen resistant hydrogenases and methods for designing and making same, April 16. U.S. Patent Appl. 10/553,097, Apr 16, 2004.

[13] Vincent S. Shaw, Hossein Mohammadiarani, Harish Vashisth, and Richard R. Neubig. Differential protein dynamics of regulators of G-protein signaling: Role in specificity of small-molecule inhibitors. *J. Am. Chem. Soc.*, 140(9):3454–3460, 2018.

[14] Mohammadjavad Mohammadi, Hossein Mohammadiarani, Vincent S Shaw, Richard R Neubig, and Harish Vashisth. Interplay of cysteine exposure and global protein dynamics in small-molecule recognition by a regulator of G-protein signaling protein. *Proteins*, 87(2):146–156, 2019.

[15] Harish Vashisth and Cameron F. Abrams. All-atom structural models of insulin binding to the insulin receptor in the presence of a tandem hormone-binding element. *Proteins*, 81:1017–1030, 2013.

[16] Vincent S. Shaw, Mohammadjavad Mohammadi, Richard R. Neubig, and Harish Vashisth. An interhelical salt bridge controls flexibility and inhibitor potency for regulators of G-protein signaling (RGS) proteins 4, 8, and 19. *unpublished*.

[17] Mohammadjavad Mohammadi, Hossein Mohammadiarani, Vincent S. Shaw, Richard R. Neubig, and Harish Vashisth. Cover image, volume 87, issue 2. *Proteins*, 87(2):C1–C1, 2019.

[18] Jürgen Drews and Stefan Ryser. Drug development: The role of innovation in drug development. *Nat. Biotechnol.*, 15(13):1318, 1997.

[19] Andrew L Hopkins and Colin R Groom. The druggable genome. *Nat. Rev. Drug Discov.*, 1(9):727, 2002.

[20] Alan C Cheng, Ryan G Coleman, Kathleen T Smyth, Qing Cao, Patricia Soulard, Daniel R Caffrey, Anna C Salzberg, and Enoch S Huang. Structure-based maximal affinity model predicts small-molecule druggability. *Nat. Biotechnol.*, 25(1):71, 2007.

[21] James A Wells and Christopher L McClendon. Reaching for high-hanging fruit in drug discovery at protein–protein interfaces. *Nature*, 450(7172):1001, 2007.

[22] Anastasios Melis, Liping Zhang, Marc Forestier, Maria L Ghirardi, and Michael Seibert. Sustained photobiological hydrogen gas production upon reversible inactivation of oxygen evolution in the green *alga chlamydomonas reinhardtii*. *Plant Physiol.*, 122(1):127–136, 2000.

[23] Anastasios Melis. Green alga hydrogen production: Progress, challenges and prospects. *Int. J. Hydrogen Energy*, 27(11):1217–1228, 2002.

[24] David B Levin, Lawrence Pitt, and Murray Love. Biohydrogen production: Prospects and limitations to practical application. *Int. J. Hydrogen Energy*, 29(2):173–185, 2004.

[25] Maria L Ghirardi, Paul W King, Matthew C Posewitz, P Ching Maness, Alexander Fedorov, Kwiseon Kim, Jordi Cohen, Klaus Schulten, and Michael Seibert. Approaches to developing biological $H_2$-photoproducing organisms and processes. *Biochem. Soc. Trans.*, 33:70–72, 2005.

[26] Maria Lucia Ghirardi, Alexandra Dubini, Jianping Yu, and Pin-Ching Maness. Photobiological hydrogen producing systems. *Chem. Soc. Rev.*, 38(1):52–61, 2009.

[27] Michel Frey. Hydrogenases: Hydrogen-activating enzymes. *ChemBioChem*, 3(2-3):153–160, 2002.

[28] Wolfgang Lubitz, Hideaki Ogata, Olaf Rudiger, and Edward Reijerse. Hydrogenases. *Chem. Rev.*, 114(8):4081–4148, 2014.

[29] Vincent Fourmond, Carole Baffert, Kateryna Sybirna, Sébastien Dementin, Abbas Abou-Hamdan, Isabelle Meynial-Salles, Philippe Soucaille, Hervé Bottin, and Christophe Léger. The mechanism of inhibition by $H_2$ of $H_2$-evolution by hydrogenases. *Chem. Commun.*, 49(61):6840–6842, 2013.

[30] Patrick Kwan, Chelsea L McIntosh, David P Jennings, R Chris Hopkins, Sanjeev K Chandrayan, Chang-Hao Wu, Michael WW Adams, and Anne K Jones. The [NiFe]-hydrogenase of pyrococcus furiosus exhibits a new type of oxygen tolerance. *J. Am. Chem. Soc.*, 137(42):13556–13565, 2015.

[31] David W Mulder, Eric M Shepard, Jonathan E Meuser, Neelambari Joshi, Paul W King, Matthew C Posewitz, Joan B Broderick, and John W Peters. Insights into [FeFe]-hydrogenase structure, mechanism, and maturation. *Structure*, 19(8):1038–1052, 2011.

[32] Brian J Lemon and John W Peters. Binding of exogenously added carbon monoxide at the active site of the iron-only hydrogenase (CpI) from *Clostridium pasteurianum*. *Biochemistry*, 38(40):12969–12973, 1999.

[33] Arti S Pandey, Travis V Harris, Logan J Giles, John W Peters, and Robert K Szilagyi. Dithiomethylether as a ligand in the hydrogenase H-cluster. *J. Am. Chem. Soc.*, 130(13):4533–4540, 2008.

[34] Gustav Berggren, A Adamska, C Lambertz, TR Simmons, J Esselborn, M Atta, S Gambarelli, J-M Mouesca, E Reijerse, W Lubitz, et al. Biomimetic assembly and activation of [FeFe]-hydrogenases. *Nature*, 499(7456):66–69, 2013.

[35] Alexey Silakov, Brian Wenk, Eduard Reijerse, and Wolfgang Lubitz. $^{14}$N HYSCORE investigation of the H-cluster of [FeFe] hydrogenase: Evidence for a nitrogen in the dithiol bridge. *Phys. Chem. Chem. Phys.*, 11(31):6592–6599, 2009.

[36] Anne Volbeda, Claudine Darnault, Alison Parkin, Frank Sargent, Fraser A Armstrong, and Juan C Fontecilla-Camps. Crystal structure of the $O_2$-tolerant membrane-bound hydrogenase 1 from *escherichia coli* in complex with its cognate *cytochrome b*. *Structure*, 21(1):184–190, 2013.

[37] Johannes Fritsch, Patrick Scheerer, Stefan Frielingsdorf, Sebastian Kroschinsky, Bärbel Friedrich, Oliver Lenz, and Christian MT Spahn. The crystal structure of an oxygen tolerant hydrogenase uncovers a novel iron sulphur centre. *Nature*, 479(7372):249–252, 2011.

[38] Abbas Abou Hamdan, Pierre-Pol Liebgott, Vincent Fourmond, Oscar Gutiérrez-Sanz, Antonio L De Lacey, Pascale Infossi, Marc Rousset, Sébastien Dementin, and Christophe Léger. Relation between anaerobic inactivation and oxygen tolerance in a large series of [NiFe]-hydrogenase mutants. *Proc. Natl. Acad. Sci. U. S. A.*, 109(49):19916–19921, 2012.

[39] Nicolas Plumeré, Olaf Rüdiger, Alaa Alsheikh Oughli, Rhodri Williams, Jeevanthi Vivekananthan, Sascha Pöller, Wolfgang Schuhmann, and Wolfgang Lubitz. A redox hydrogel protects hydrogenase from high potential deactivation and oxygen damage. *Nat. Chem.*, 6(9):822–827, 2014.

[40] Vincent Fourmond, Stefanie Stapf, Huaiguang Li, Darren Buesen, James Birrell, Olaf RuÌĹdiger, Wolfgang Lubitz, Wolfgang Schuhmann, Nicolas Plumeré, and Christophe Léger. Mechanism of protection of catalysts supported in redox hydrogel films. *J. Am. Chem. Soc.*, 137(16):5494–5505, 2015.

[41] Timothy Flynn, Maria Lucia Ghirardi, and Michael Seibert. Accumulation of $O_2$-tolerant phenotypes in $H_2$-producing strains of *chlamydomonas reinhardtii* by sequential applications of chemical mutagenesis and selection. *Int. J. Hydrogen Energy*, 27(11):1421–1430, 2002.

[42] Gongyi Hong and Ruth Pachter. Inhibition of biocatalysis in [FeFe]-hydrogenase by oxygen: Molecular dynamics and density functional theory calculations. *ACS Chem. Biol.*, 7(7):1268–1275, 2012.

[43] Sven T Stripp, Gabrielle Goldet, Caterina Brandmayr, Oliver Sanganas, Kylie A Vincent, Michael Haumann, Fraser A Armstrong, and Thomas Happe. How oxygen attacks [FeFe]-hydrogenases from photosynthetic organisms. *Proc. Natl. Acad. Sci. U. S. A.*, 106(41):17331–17336, 2009.

[44] Martin T Stiebritz and Markus Reiher. Theoretical study of dioxygen induced inhibition of [FeFe]-hydrogenase. *Inorg. Chem.*, 48(15):7127–7140, 2009.

[45] Adam Kubas, David De Sancho, Robert B Best, and Jochen Blumberger. Aerobic damage to [FeFe]-hydrogenases: Activation barriers for the chemical attachment of $O_2$. *Angew. Chem. Int. Ed.*, 126(16):4165–4168, 2014.

[46] Christophe Orain, Laure Saujet, Charles Gauquelin, Philippe Soucaille, Isabelle Meynial-Salles, Carole Baffert, Vincent Fourmond, Hervé Bottin, and Christophe Léger. Electrochemical measurements of the kinetics of inhibition of two [FeFe]-hydrogenases by $O_2$ demonstrate that the reaction is partly reversible. *J. Am. Chem. Soc.*, 137(39):12580–12587, 2015.

[47] Carole Baffert, Marie Demuez, Laurent Cournac, Benedicte Burlat, Bruno Guigliarelli, Patrick Bertrand, Laurence Girbal, and Christophe Leger. Hydrogen activating enzymes: Activity does not correlate with oxygen sensitivity. *Angew. Chem. Int. Ed.*, 47(11):2052–2054, 2008.

[48] Camilla Lambertz, Nils Leidel, Kajsa GV Havelius, Jens Noth, Petko Chernev, Martin Winkler, Thomas Happe, and Michael Haumann. $O_2$ reactions at the six iron active site (H-cluster) in [FeFe]-hydrogenase. *J. Biol. Chem.*, 286(47):40614–40623, 2011.

[49] Kevin D Swanson, Michael W Ratzloff, David W Mulder, Jacob H Artz, Shourjo Ghose, Andrew Hoffman, Spencer White, Oleg A Zadvornyy, Joan B Broderick, Brian Bothner, Paul W. King, and John W. Peters. [FeFe]-hydrogenase oxygen inactivation is initiated at the h-cluster 2fe subcluster. *J. Am. Chem. Soc.*, 137(5):1809–1816, 2015.

[50] Pierre-Pol Liebgott, Fanny Leroux, Bénédicte Burlat, Sébastien Dementin, Carole Baffert, Thomas Lautier, Vincent Fourmond, Pierre Ceccaldi, Christine Cavazza, Isabelle Meynial-Salles, Philippe Soucaille, Juan Carlos Fontecilla-Camps, Bruno Guigliarelli, Patrick Bertrand, Marc Rousset, and Christophe Léger. Relating diffusion along the substrate tunnel and oxygen sensitivity in hydrogenase. *Nat. Chem. Biol.*, 6(1):63–70, 2010.

[51] Karin Nienhaus, Pengchi Deng, John S Olson, Joshua J Warren, and G Ulrich Nienhaus. Structural dynamics of myoglobin: Ligand migration and binding in valine 68 mutants. *J. Biol. Chem.*, 278(43):42532–42544, 2003.

[52] Karin Nienhaus, Pengchi Deng, Jan M. Kriegl, and G. Ulrich Nienhaus. Structural dynamics of myoglobin: Effect of internal cavities on ligand migration and binding. *Biochemistry*, 42(32):9647–9658, 2003.

[53] Jordi Cohen, Kwiseon Kim, Matthew Posewitz, Maria L , Klaus Schulten, Michael Seibert, and Paul King. Molecular dynamics and experimental investigation of H2 and O2 diffusion in [Fe]-hydrogenase. *Biochem. Soc. Trans.*, 33:80–82, 2005.

[54] B Ginovska, S Raugei, and WJ Shaw. Molecular dynamics studies of proton transport in hydrogenase and hydrogenase mimics. *Methods Enzymol.*, 578:73–101, 2016.

[55] Alyssa S Bingham, Phillip R Smith, and James R Swartz. Evolution of an [FeFe]-hydrogenase with decreased oxygen sensitivity. *Int. J. Hydrogen Energy*, 37(3):2965–2976, 2012.

[56] Luca Maragliano, Grazia Cottone, Giovanni Ciccotti, and Eric Vanden-Eijnden. Mapping the network of pathways of CO diffusion in myoglobin. *J. Am. Chem. Soc.*, 132(3):1010–1017, 2009.

[57] Mauro Lapelosa and Cameron F Abrams. A computational study of water and CO migration sites and channels inside myoglobin. *J. Chem. Theory Comput.*, 9(2):1265–1271, 2013.

[58] Anthony Bucci and Cameron F Abrams. Oxygen pathways and allostery in monomeric sarcosine oxidase via single-sweep free energy reconstruction. *J. Chem. Theory Comput.*, 10(7):2668–2676, 2014.

[59] Andrew L Hopkins and Colin R Groom. The druggable genome. *Nat. Rev. Drug Discov.*, 1(9):727–730, 2002.

[60] Antonio Carrieri, Violeta I Perez-Nueno, Giovanni Lentini, and David W Ritchie. Recent trends and future prospects in computationaldrug discovery: from virtual screening to polypharmacology. *Curr. Top. Med. Chem.*, 13(9):1069–1097, 2013.

[61] Hilary Highfield Nickols and P Jeffrey Conn. Development of allosteric modulators of GPCRs for treatment of cns disorders. *Neurobiol. Dis.*, 61:55–71, 2014.

[62] Miles Congreve, João M Dias, and Fiona H Marshall. Structure-based drug design for G-protein-coupled receptors. *Prog. Med. Chem.*, 53:1–63, 2014.

[63] Arthur Christopoulos. Advances in G-protein-coupled receptor allostery: from function to structure. *Mol. Pharmacol.*, 86(5):463–478, 2014.

[64] AJ Venkatakrishnan, Xavier Deupi, Guillaume Lebon, Christopher G Tate, Gebhard F Schertler, and M Madan Babu. Molecular signatures of G-protein-coupled receptors. *Nature*, 494(7436):185–194, 2013.

[65] Huailing Zhong and Richard R Neubig. Regulator of G-protein signaling proteins: novel multifunctional drug targets. *J. Pharmacol. Exp. Ther.*, 297(3):837–845, 2001.

[66] Richard R Neubig and David P Siderovski. Regulators of G-protein signalling as new central nervous system drug targets. *Nat. Rev. Drug Discov.*, 1(3):187–197, 2002.

[67] Susanne Hollinger and John R Hepler. Cellular regulation of RGS proteins: modulators and integrators of G-protein signaling. *Pharmacol. Rev.*, 54(3):527–559, 2002.

[68] Adam J Kimple, Dustin E Bosch, Patrick M Giguère, and David P Siderovski. Regulators of G-protein signaling and their $G_\alpha$ substrates: promises and challenges in their use as drug discovery targets. *Pharmacol. Rev.*, 63(3):728–749, 2011.

[69] David M Berman, Tohru Kozasa, and Alfred G Gilman. The GTPase-activating protein RGS4 stabilizes the transition state for nucleotide hydrolysis. *J. Biol. Chem.*, 271(44):27209–27212, 1996.

[70] Levi L Blazer and Richard R Neubig. Small molecule protein–protein interaction inhibitors as cns therapeutic agents: current progress and future hurdles. *Neuropsychopharmacology*, 34(1):126–141, 2009.

[71] Michelle R Arkin and James A Wells. Small-molecule inhibitors of protein–protein interactions: progressing towards the dream. *Nat. Rev. Drug Discov.*, 3(4):301–317, 2004.

[72] David L Roman, Jeffery N Talbot, Rebecca A Roof, Roger K Sunahara, John R Traynor, and Richard R Neubig. Identification of small-molecule inhibitors of RGS4 using a high-throughput flow cytometry protein interaction assay. *Mol. Pharmacol.*, 71(1):169–175, 2007.

[73] Nir London, Barak Raveh, and Ora Schueler-Furman. Druggable protein-protein interactions–from hot spots to hot segments. *Curr. Opin. Chem. Biol.*, 17(6):952–959, 2013.

[74] Amit S Kalgutkar and Deepak K Dalvie. Drug discovery for a new generation of covalent drugs. *Expert Opin. Drug Discov.*, 7(7):561–581, 2012.

[75] Asher Mullard. Protein–protein interaction inhibitors get into the groove. *Nat. Rev. Drug Discov.*, 11(3):173–175, 2012.

[76] Xavier Morelli, Raphaël Bourgeas, and Philippe Roche. Chemical and structural lessons from recent successes in protein–protein interaction inhibition (2P2I). *Curr. Opin. Chem. Biol.*, 15(4):475–481, 2011.

[77] B O Villoutreix, C M Labbe, David Lagorce, Guillaume Laconde, and Olivier Sperandio. A leap into the chemical space of protein-protein interaction inhibitors. *Curr. Pharm. Des.*, 18(30):4648–4667, 2012.

[78] Michelle R Arkin, Yinyan Tang, and James A Wells. Small-molecule inhibitors of protein-protein interactions: progressing toward the reality. *Chem. Biol.*, 21(9):1102–1114, 2014.

[79] Benita Sjögren and Richard R Neubig. Thinking outside of the "RGS box": new approaches to therapeutic targeting of regulators of G-protein signaling. *Mol. Pharmacol.*, 78(4):550–557, 2010.

[80] David L Roman and John R Traynor. Regulators of G-protein signaling (RGS) proteins as drug targets: Modulating G-protein-coupled receptor (GPCR ) signal transduction: Miniperspective. *J. Med. Chem.*, 54(21):7433–7440, 2011.

[81] Isaac J Dripps, Qin Wang, Richard R Neubig, Kenner C Rice, John R Traynor, and Emily M Jutkiewicz. The role of regulator of G-protein signaling 4 in delta-opioid receptor-mediated behaviors. *Psychopharmacology*, 234(1):29–39, 2017.

[82] Benita Sjögren. Regulator of G-protein signaling proteins as drug targets: current state and future possibilities. *Adv. Pharmacol. Sci.*, 62:315, 2011.

[83] Levi L Blazer, David L Roman, Alfred Chung, Martha J Larsen, Benjamin M Greedy, Stephen M Husbands, and Richard R Neubig. Reversible, allosteric small-molecule inhibitors of regulator of G-protein signaling proteins. *Mol. Pharmacol.*, 78(3):524–533, 2010.

[84] David L Roman, Levi L Blazer, C Aaron Monroy, and Richard R Neubig. Allosteric inhibition of the regulator of G-protein signaling–g$\alpha$ protein–protein interaction by CCG-4986. *Mol. Pharmacol.*, 78(3):360–365, 2010.

[85] Levi L Blazer, Haoming Zhang, Emma M Casey, Stephen M Husbands, and Richard R Neubig. A nanomolar-potency small molecule inhibitor of regulator of G-protein signaling proteins. *Biochemistry*, 50(15):3181–3192, 2011.

[86] Andrew J. Storaska, Jian P. Mei, Meng Wu, Min Li, Susan M. Wade, Levi L. Blazer, Benita Sjogren, Corey R. Hopkins, Craig W. Lindsley, Zhihong Lin, Joseph J. Babcock, Owen B. McManus, and Richard R. Neubig. Reversible inhibitors of regulators of G-protein signaling identified in a high-throughput cell-based calcium signaling assay. *Cell. Signal.*, 25(12):2848 – 2855, 2013.

[87] David L Roman, Shodai Ota, and Richard R Neubig. Polyplexed flow cytometry protein interaction assay: a novel high-throughput screening paradigm for RGS protein inhibitors. *J. Biomol. Screen*, 14(6):610–619, 2009.

[88] Stephen J Gold, Yan G Ni, Henrik G Dohlman, and Eric J Nestler. Regulators of G-protein signaling (RGS) proteins: region-specific expression of nine subtypes in rat brain. *J. Neurosci.*, 17(20):8024–8037, 1997.

[89] Christopher Larminie, Paul Murdock, Jean-Philippe Walhin, Malcolm Duckworth, Kendall J Blumer, Mark A Scheideler, and Martine Garnier. Selective expression of regulators of G-protein signaling (RGS) in the human central nervous system. *Mol. Brain Res.*, 122(1):24–34, 2004.

[90] Talia N Lerner and Anatol C Kreitzer. RGS4 is required for dopaminergic control of striatal LTD and susceptibility to parkinsonian motor deficits. *Neuron*, 73(2):347–359, 2012.

[91] Weixing Shen, JoshuaL. Plotkin, Veronica Francardo, WaiKinD. Ko, Zhong Xie, Qin Li, Tim Fieblinger, Jürgen Wess, RichardR. Neubig, CraigW. Lindsley, P.Jeffrey Conn, Paul Greengard, Erwan Bezard, M.Angela Cenci, and D.James Surmeier. M4 muscarinic receptor signaling ameliorates striatal plasticity deficits in models of L-DOPA-induced dyskinesia. *Neuron*, 88(4):762 – 773, 2015.

[92] Wai Kin D Ko, Marie-Laure Martin-Negrier, Erwan Bezard, Alan R Crossman, and Paula Ravenscroft. RGS4 is involved in the generation of abnormal involuntary move-

ments in the unilateral 6-OHDA-lesioned rat model of parkinson's disease. *Neurobiol. Dis.*, 70:138–148, 2014.

[93] Levi L. Blazer, Andrew J. Storaska, Emily M. Jutkiewicz, Emma M. Turner, Mariangela Calcagno, Susan M. Wade, Qin Wang, Xi-Ping Huang, John R. Traynor, Stephen M. Husbands, Michele Morari, and Richard R. Neubig. Selectivity and anti-parkinson's potential of thiadiazolidinone RGS4 inhibitors. *ACS Chem. Neurosci.*, 6(6):911–919, 2015.

[94] Lars Tatenhorst, Volker Senner, Sylvia Püttmann, and Werner Paulus. Regulators of G-protein signaling 3 and 4 (RGS3, RGS4) are associated with glioma cell motility. *J. Neuropathol. Exp. Neurol.*, 63(3):210–222, 2004.

[95] Dragomira Nikolaeva Nikolova, Hitoshi Zembutsu, Tanio Sechanov, Kalin Vidinov, Low Siew Kee, Radina Ivanova, Elitza Becheva, Miriana Kocova, Draga Toncheva, and Yusuke Nakamura. Genome-wide gene expression profiles of thyroid carcinoma: identification of molecular targets for treatment of thyroid carcinoma. *Oncol. Rep.*, 20(1):105–121, 2008.

[96] Jillian H Hurst, Nisha Mendpara, and Shelley B Hooks. Regulator of G-protein signalling expression and function in ovarian cancer cell lines. *Cell. Mol. Biol. Lett.*, 14(1):153, 2008.

[97] Carlo Cifelli, Robert A Rose, Hangjun Zhang, Julia Voigtlaender-Bolz, Steffen-Sebastian Bolz, Peter H Backx, and Scott P Heximer. RGS4 regulates parasympathetic signaling and heart rate control in the sinoatrial node. *Circ. Res.*, 103(5):527–535, 2008.

[98] Yunjia Chen, Yin Liu, Christopher Cottingham, Lori McMahon, Kai Jiao, Paul Greengard, and Qin Wang. Neurabin scaffolding of adenosine receptor and RGS4 regulates anti-seizure effect of endogenous adenosine. *J. Neurosci.*, 32(8):2683–2695, 2012.

[99] Qin Wang, Lee-Yuan Liu-Chen, and John R Traynor. Differential modulation of $\mu$- and $\delta$-opioid receptor agonists by endogenous RGS4 protein in SH-SY5Y cells. *J. Biol. Chem.*, 284(27):18357–18367, 2009.

[100] John JG Tesmer. Structure and function of regulator of G-protein signaling homology domains. *Prog. Mol. Biol. Transl. Sci.*, 86:75–113, 2009.

[101] John JG Tesmer, David M Berman, Alfred G Gilman, and Stephen R Sprang. Structure of RGS4 bound to ALF 4–activated $G_{i\alpha 1}$: stabilization of the transition state for gtp hydrolysis. *Cell*, 89(2):251–261, 1997.

[102] Harish Vashisth, Andrew J Storaska, Richard R Neubig, and Charles L Brooks III. Conformational dynamics of a regulator of G-protein signaling protein reveals a mechanism of allosteric inhibition by a small molecule. *ACS Chem. Biol.*, 8(12):2778–2784, 2013.

[103] Serguei G Popov, U Murali Krishna, JR Falck, and Thomas M Wilkie. Ca2+/calmodulin reverses phosphatidylinositol 3, 4, 5-trisphosphate-dependent inhibition of regulators of G-protein-signaling GTPase-activating protein activity. *J. Biol. Chem.*, 275(25):18962–18968, 2000.

[104] Masaru Ishii, Satoru Fujita, Mitsuhiko Yamada, Yukio Hosaka, and Yoshihisa Kurachi. Phosphatidylinositol 3, 4, 5-trisphosphate and ca2+/calmodulin competitively bind to the regulators of G-protein-signalling (RGS) domain of RGS4 and reciprocally regulate its action. *Biochem. J.*, 385(1):65–73, 2005.

[105] Boakye A. Boatin, María-Gloria Basáñez, Roger K. Prichard, Kwablah Awadzi, Rashida M. Barakat, Héctor H. García, Andrea Gazzinelli, Warwick N. Grant, James S. McCarthy, Eliézer K. N'Goran, Mike Y. Osei-Atweneboana, Banchob Sripa, Guo-Jing Yang, and Sara Lustigman. A research agenda for helminth diseases of humans: Towards control and elimination. *PLOS Negl. Trop. Dis.*, 6(4):1–10, 04 2012.

[106] Ray M Kaplan. Drug resistance in nematodes of veterinary importance: a status report. *Trends parasitol.*, 20(10):477–481, 2004.

[107] Joseph A Beavo, Sharron H Francis, and Miles D Houslay. *Cyclic nucleotide phosphodiesterases in health and disease*. CRC Press, 2006.

[108] Hengming Ke and Huanchen Wang. Crystal structures of phosphodiesterases and implications on substrate specificity and inhibitor selectivity. *Curr. Top. Med. Chem.*, 7(4):391–403, 2007.

[109] Robert X Xu, Anne M Hassell, Dana Vanderwall, Millard H Lambert, William D Holmes, Michael A Luther, Warren J Rocque, Michael V Milburn, Yingdong Zhao, Hengming Ke, et al. Atomic structure of pde4: insights into phosphodiesterase mechanism and specificity. *Science*, 288(5472):1822–1825, 2000.

[110] William Humphrey, Andrew Dalke, and Klaus Schulten. VMD: Visual molecular dynamics. *J. Mol. Graphics*, 14(1):33–38, 1996.

[111] James C Phillips, Rosemary Braun, Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D Skeel, Laxmikant Kale, and Klaus Schulten. Scalable molecular dynamics with NAMD. *J. Comput. Chem.*, 26(16):1781–1802, 2005.

[112] David K Brown, David L Penkler, Olivier Sheik Amamuddy, Caroline Ross, Ali Rana Atilgan, Canan Atilgan, and Özlem Tastan Bishop. MD-TASK: a software suite for analyzing molecular dynamics trajectories. *Bioinformatics*, 33(17):2768–2771, 2017.

[113] Ruben Abagyan, Maxim Totrov, and Dmitry Kuznetsov. ICM–a new method for protein modeling and design: applications to docking and structure prediction from the distorted native conformation. *J. Comp. Chem.*, 15(5):488–506, 1994.

[114] Sunhwan Jo, Taehoon Kim, Vidyashankara G Iyer, and Wonpil Im. CHARMM-GUI: a web-based graphical user interface for charmm. *J. Comput. Chem.*, 29(11):1859–1865, 2008.

[115] Martin Karplus and John Kuriyan. Molecular dynamics and protein function. *Proc. Natl. Acad. Sci. U. S. A.*, 102(19):6679–6685, 2005.

[116] Tamar Schlick. *Molecular modeling and simulation: an interdisciplinary guide: an interdisciplinary guide*, volume 21. Springer Science & Business Media, 2010.

[117] Rodney J Baxter. *Exactly solved models in statistical mechanics*. Elsevier, 2016.

[118] J Willard Gibbs. *Elementary principles in statistical mechanics*. Courier Corporation, 2014.

[119] R Leach Andrew. Molecular modeling principles and applications. *2nd, editor.: Pearson Education Limited*, 2001.

[120] Martin Karplus and J Andrew McCammon. Molecular dynamics simulations of biomolecules. *Nat. Struct. Mol. Biol*, 9(9):646, 2002.

[121] Daan Frenkel and Berend Smit. Understanding molecular simulation: From algorithms to applications. *Comput. Sci.*, 1:1–638, 2002.

[122] Alex D MacKerell Jr, Donald Bashford, MLDR Bellott, Roland Leslie Dunbrack Jr, Jeffrey D Evanseck, Martin J Field, Stefan Fischer, Jiali Gao, H Guo, Sookhee Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiòrkiewicz-Kuczera, D. Yin, and M. Karplus. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B*, 102(18):3586–3616, 1998.

[123] Alexander D MacKerell, Michael Feig, and Charles L Brooks III. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.*, 25(11):1400–1415, 2004.

[124] Bernard R Brooks, Charles L Brooks III, Alexander D Mackerell Jr, Lennart Nilsson, Robert J Petrella, Benoît Roux, Youngdo Won, Georgios Archontis, Christian Bartels, Stefan Boresch, et al. CHARMM: the biomolecular simulation program. *J. Comput. Chem.*, 30(10):1545–1614, 2009.

[125] Joseph D Yesselman, Daniel J Price, Jennifer L Knight, and Charles L Brooks III. MATCH: An atom-typing toolset for molecular mechanics force fields. *J. Comput. Chem.*, 33(2):189–202, 2012.

[126] Axel Brünger, Charles L Brooks III, and Martin Karplus. Stochastic boundary conditions for molecular dynamics simulations of ST2 water. *Chem. Phys. Lett.*, 105(5):495–500, 1984.

[127] Bimal Mishra and Tamar Schlick. The notion of error in langevin dynamics. i. linear analysis. *J. Chem. Phys.*, 105(1):299–318, 1996.

[128] Tamar Schlick, Eric Barth, and Margaret Mandziuk. Biomolecular dynamics at long timesteps: Bridging the timescale gap between simulation and experimentation. *Annu. Rev. Biophys.*, 26(1):181–222, 1997.

[129] Wilfred F van Gunsteren, Phillippe H Huenenberger, Alan E Mark, Paul E Smith, and Ilario G Tironi. Computer simulation of protein motion. *Comput. Phys. Commun.*, 91(1-3):305–319, 1995.

[130] Ahmad Jabbarzadeh and Roger I Tanner. Molecular dynamics simulation and its application to nano-rheology. *J. Rheol.*, 2006:165, 2006.

[131] Nevena Todorova, Fabrizio Marinelli, Stefano Piana, and Irene Yarovsky. Exploring the folding free energy landscape of insulin using bias exchange metadynamics. *J. Phys. Chem. B*, 113(11):3556–3564, 2009.

[132] Hans C Andersen. Molecular dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys.*, 72(4):2384–2393, 1980.

[133] Shūichi Nosé. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.*, 52(2):255–268, 1984.

[134] William G Hoover, Anthony JC Ladd, and Bill Moran. High-strain-rate plastic flow studied via nonequilibrium molecular dynamics. *Phys. Rev. Lett.*, 48(26):1818, 1982.

[135] William G Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A*, 31(3):1695, 1985.

[136] Juan R Perilla and Klaus Schulten. Physical properties of the HIV-1 capsid from all-atom molecular dynamics simulations. *Nat. Commun.*, 8:15959, 2017.

[137] Cameron F. Abrams and Giovanni Bussi. Enhanced sampling in molecular dynamics using metadynamics, replica-exchange, and temperature-acceleration. *Entropy*, 16(1):163–199, 2014.

[138] Daniele Granata, Carlo Camilloni, Michele Vendruscolo, and Alessandro Laio. Characterization of the free-energy landscapes of proteins by NMR-guided metadynamics. *Proc. Natl. Acad. Sci. U.S.A.*, 110(17):6817–6822, 2013.

[139] Yanier Crespo, Fabrizio Marinelli, Fabio Pietrucci, and Alessandro Laio. Metadynamics convergence law in a multidimensional system. *Phys. Rev. E.*, 81(5), 2010.

[140] Luca Maragliano and Eric Vanden-Eijnden. A temperature accelerated method for sampling free energy and determining reaction pathways in rare events simulations. *Chem. Phys. Lett.*, 426(1):168–175, 2006.

[141] Luca Maragliano and Eric Vanden-Eijnden. Single-sweep methods for free energy calculations. *J. Chem. Phys.*, 128(18):184110–184120, 2008.

[142] Harish Vashisth, Luca Maragliano, and Cameron F. Abrams. "DFG-flip" in the insulin receptor kinase is facilitated by a helical intermediate state of the activation loop. *Biophys. J.*, 102:1979–1987, 2012.

[143] Harish Vashisth, Georgios Skiniotis, and Charles L. Brooks III. Using enhanced sampling and structural restraints to refine atomic structures into low-resolution electron microscopy maps. *Structure*, 20:1453–1462, 2012.

[144] Harish Vashisth and Charles L. Brooks III. Conformational sampling of maltose-transporter components in cartesian collective variables is governed by the low-frequency normal modes. *J. Phys. Chem. Lett.*, 3 :3379–3384, 2012.

[145] Cameron F Abrams and Eric Vanden-Eijnden. Large scale conformational sampling of proteins using temperature accelerated molecular dynamics. *Proc. Natl. Acad. Sci. U. S. A.*, 107(11):4961–4966, 2010.

[146] Giovanni Ciccotti, Raymond Kapral, and Eric Vanden-Eijnden. Blue moon sampling, vectorial reaction coordinates, and unbiased constrained dynamics. *ChemPhysChem*, 6(9):1809–1814, 2005.

[147] Michele Monteferrante, Sara Bonella, Simone Meloni, Eric Vanden-Eijnden, and Giovanni Ciccotti. Calculations of free energy barriers for local mechanisms of hydrogen diffusion in alanates. *Sci. Model Simul.*, pages 187–206, 2008.

[148] Michele Monteferrante, Sara Bonella, Simone Meloni, and Giovanni Ciccotti. Modified single-sweep method for reconstructing free energy landscapes. *Mol. Simul.*, 35(12-13):1116–1129, 2009.

[149] Weinan E, Weiqing Ren, and Eric Vanden-Eijnden. String method for the study of rare events. *Phys. Rev. B*, 66(5):052301, 2002.

[150] Luca Maragliano, Alexander Fischer, Eric Vanden-Eijnden, and Giovanni Ciccotti. String method in collective variables: Minimum free energy paths and isocommittor surfaces. *J. Chem. Phys.*, 125(2):024106, 2006.

[151] E Weinan, Weiqing Ren, and Eric Vanden-Eijnden. Simplified and improved string method for computing the minimum energy paths in barrier-crossing events. *J. Chem. Phys.*, 126(16):164103, 2007.

[152] Luca Maragliano and Eric Vanden-Eijnden. On the fly string method for minimum free energy paths calculation. *Chem. Phys. Lett.*, 446(1):182–190, 2007.

[153] Alessandro Laio and Michele Parrinello. Escaping free-energy minima. *Proc. Natl. Acad. Sci. U.S.A.*, 99(20):12562–12566, 2002.

[154] Alessandro Laio and Francesco L Gervasio. Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Rep. Prog. Phys.*, 71(12):126601, 2008.

[155] Vittorio Limongelli, Luciana Marinelli, Sandro Cosconati, Concettina La Motta, Stefania Sartini, Laura Mugnaini, Federico Da Settimo, Ettore Novellino, and Michele Parrinello. Sampling protein motion and solvent effect during ligand binding. *Proc. Natl. Acad. Sci. U.S.A.*, 109(5):1467–1472, 2012.

[156] Alessandro Barducci, Massimiliano Bonomi, Meher K Prakash, and Michele Parrinello. Free-energy landscape of protein oligomerization from atomistic simulations. *Proc. Natl. Acad. Sci. U.S.A.*, 110(49):E4708–E4713, 2013.

[157] Hossein Mohammadiarani and Harish Vashisth. All-atom structural models of the transmembrane domains of insulin and type 1 insulin-like growth factor receptors. *Front. Endocrinol.*, 7, 2016.

[158] Hossein Mohammadiarani and Harish Vashisth. Insulin mimetic peptide S371 folds into a helical structure. *J. Comput. Chem.*, 38(15):1158–1166, 2017.

[159] Xevi Biarnés, Salvatore Bongarzone, Attilio Vittorio Vargiu, Paolo Carloni, and Paolo Ruggerone. Molecular motions in drug design: the coming age of the metadynamics method. *J. Comput. Aided Mol. Des.*, 25(5):395–402, 2011.

[160] Matteo Incerti, Simonetta Russo, Donatella Callegari, Daniele Pala, Carmine Giorgio, Ilaria Zanotti, Elisabetta Barocelli, Paola Vicini, Federica Vacondio, Silvia Rivara, Riccardo Castelli, Massimiliano Tognolini, and Alessio Lodola. Metadynamics for perspective drug design: computationally driven synthesis of new protein–protein interaction inhibitors targeting the epha2 receptor. *J. Med. Chem.*, 60(2):787–796, 2017.

[161] Riccardo Baron, Conor Riley, Pirom Chenprakhon, Kittisak Thotsaporn, Remko T Winter, Andrea Alfieri, Federico Forneris, Willem JH van Berkel, Pimchai Chaiyen, Marco W Fraaije, et al. Multiple pathways guide oxygen diffusion into flavoenzyme active sites. *Proc. Natl. Acad. Sci. U. S. A.*, 106(26):10603–10608, 2009.

[162] Robert Y Igarashi and Lance C Seefeldt. Nitrogen fixation: the mechanism of the mo-dependent nitrogenase. *Crit. Rev. Biochem. Mol. Biol.*, 38(4):351–384, 2003.

[163] W Lubitz and W Tumas. Hydrogen: an overview. *Chem. Rev.*, 107(10):3900–3903, 2007.

[164] Gregory J Kubas. Fundamentals of $H_2$ binding and reactivity on transition metals underlying hydrogenase function and $H_2$ production and storage. *Chem. Rev.*, 107(10):4152–4205, 2007.

[165] Cédric Tard, Xiaoming Liu, Saad K Ibrahim, Maurizio Bruschi, Luca De Gioia, Siân C Davies, Xin Yang, Lai-Sheng Wang, Gary Sawers, and Christopher J Pickett. Synthesis of the H-cluster framework of iron-only hydrogenase. *Nature*, 433(7026):610, 2005.

[166] Gustav Berggren, A Adamska, C Lambertz, TR Simmons, J Esselborn, M Atta, S Gambarelli, J-M Mouesca, E Reijerse, W Lubitz, T. Happe, V. Artero, and M. Fontecave. Biomimetic assembly and activation of [FeFe]-hydrogenases. *Nature*, 499(7456):66, 2013.

[167] Juan C Fontecilla-Camps, Anne Volbeda, Christine Cavazza, and Yvain Nicolet. Structure/function relationships of [NiFe]-and [FeFe]-hydrogenases. *Chem. Rev.*, 107(10):4273–4303, 2007.

[168] Paulette M Vignais and Bernard Billoud. Occurrence, classification, and biological function of hydrogenases: an overview. *Chem. Rev.*, 107(10):4206–4272, 2007.

[169] Zhi-Pan Liu and P Hu. Mechanism of H2 metabolism on Fe-only hydrogenases. *J. Chem. Phys.*, 117(18):8177, 2002.

[170] John W Peters, William N Lanzilotta, Brian J Lemon, and Lance C Seefeldt. X-ray crystal structure of the fe-only hydrogenase (CpI) from clostridium pasteurianum to 1.8 angstrom resolution. *Science*, 282(5395):1853–1858, 1998.

[171] Arti S Pandey, Travis V Harris, Logan J Giles, John W Peters, and Robert K Szilagyi. Dithiomethylether as a ligand in the hydrogenase h-cluster. *J. Am. Chem. Soc.*, 130(13):4533–4540, 2008.

[172] MW Adams. The mechanisms of $H_2$ activation and CO binding by hydrogenase i and hydrogenase ii of clostridium pasteurianum. *J. Biol. Chem.*, 262(31):15054–15061, 1987.

[173] Gabrielle Goldet, Caterina Brandmayr, Sven T Stripp, Thomas Happe, Christine Cavazza, Juan C Fontecilla-Camps, and Fraser A Armstrong. Electrochemical kinetic investigations of the reactions of [FeFe]-hydrogenases with carbon monoxide and oxygen: comparing the importance of gas tunnels and active-site electronic/redox effects. *J. Am. Chem. Soc.*, 131(41):14979–14989, 2009.

[174] Carole Baffert, Luca Bertini, Thomas Lautier, Claudio Greco, Kateryna Sybirna, Pierre Ezanno, Emilien Etienne, Philippe Soucaille, Patrick Bertrand, Hervé Bottin, Isabelle Meynial-Salles, Luca De Gios, and Christophe Léger. CO disrupts the reduced H-cluster of FeFe hydrogenase. a combined DFT and protein film voltammetry study. *J. Am. Chem. Soc.*, 133(7):2096–2099, 2011.

[175] Pierre-Pol Liebgott, Fanny Leroux, Bénédicte Burlat, Sébastien Dementin, Carole Baffert, Thomas Lautier, Vincent Fourmond, Pierre Ceccaldi, Christine Cavazza, Isabelle Meynial-Salles, Philippe Soucaille, Juan Carlos Fontecilla-Camps, Bruno Guigliarelli, Patrick Bertrand, Marc Rousset, and Léger. Relating diffusion along the substrate tunnel and oxygen sensitivity in hydrogenase. *Nat. Chem. Biol.*, 6(1):63, 2010.

[176] Thomas Lautier, Pierre Ezanno, Carole Baffert, Vincent Fourmond, Laurent Cournac, Juan C Fontecilla-Camps, Philippe Soucaille, Patrick Bertrand, Isabelle Meynial-Salles, and Christophe Léger. The quest for a functional substrate access tunnel in fefe hydrogenase. *Faraday Discuss.*, 148:385–407, 2011.

[177] Mohammad Mirmohades, Agnieszka Adamska-Venkatesh, Constanze Sommer, Edward Reijerse, Reiner Lomoth, Wolfgang Lubitz, and Leif Hammarstrom. Following [FeFe] hydrogenase active site intermediates by time-resolved mid-ir spectroscopy. *J. Phys. Chem. Lett.*, 7(16):3290–3293, 2016.

[178] Fanny Leroux, Sébastien Dementin, Bénédicte Burlat, Laurent Cournac, Anne Volbeda, Stéphanie Champ, Lydie Martin, Bruno Guigliarelli, Patrick Bertrand, Juan

Fontecilla-Camps, Marc Rousset, and Christophe Léger. Experimental approaches to kinetics of gas diffusion in hydrogenase. *Proc. Natl. Acad. Sci. U.S.A*, 105(32):11188–11193, 2008.

[179] Po-hung Wang and Jochen Blumberger. Mechanistic insight into the blocking of CO diffusion in [NiFe]-hydrogenase mutants through multiscale simulation. *Proc. Natl. Acad. Sci. USA*, 109:6399, 2012.

[180] Thierry Prange, Marc Schiltz, Lucile Pernot, Nathalie Colloc'h, Sonia Longhi, William Bourguet, and Roger Fourme. Exploring hydrophobic sites in proteins with xenon or krypton. *Proteins*, 30(1):61–73, 1998.

[181] Christopher H Chang and Kwiseon Kim. Density functional theory calculation of bonding and charge parameters for molecular dynamics studies on [FeFe]-hydrogenases. *J. Chem. Theory Comput.*, 5(4):1137–1145, 2009.

[182] Martin McCullagh and Gregory A Voth. Unraveling the role of the protein environment for [FeFe]-hydrogenase: A new application of coarse-graining. *J. Phys. Chem. B*, 117(15):4062–4071, 2013.

[183] E Weinan, Weiqing Ren, and Eric Vanden-Eijnden. String method for the study of rare events. *Phys. Rev. B*, 66(5):052301, 2002.

[184] E Weinan, Weiqing Ren, and Eric Vanden-Eijnden. Simplified and improved string method for computing the minimum energy paths in barrier-crossing events. *J. Chem. Phys.*, 126(16):164103, 2007.

[185] Luca Maragliano, Alexander Fischer, Eric Vanden-Eijnden, and Giovanni Ciccotti. String method in collective variables: Minimum free energy paths and isocommittor surfaces. *J. Chem. Phys.*, 125(2):024106, 2006.

[186] Luca Maragliano and Eric Vanden-Eijnden. On-the-fly string method for minimum free energy paths calculation. *Chem. Phys. Lett.*, 446(1-3):182–190, 2007.

[187] Juan C Fontecilla-Camps, Anne Volbeda, Christine Cavazza, and Yvain Nicolet. Structure/function relationships of [NiFe]-and [FeFe]-hydrogenases. *Chem. Rev.*, 107(10):4273–4303, 2007.

[188] Hai Long, Paul W King, and Christopher H Chang. Proton transport in clostridium pasteurianum [FeFe]-hydrogenase I: A computational study. *J. Phys. Chem. B*, 118(4):890–900, 2014.

[189] Mariano Andrea Scorciapino, Arturo Robertazzi, Mariano Casu, Paolo Ruggerone, and Matteo Ceccarelli. Heme proteins: the role of solvent in the dynamics of gates and portals. *J. Am. Chem. Soc.*, 132(14):5156–5163, 2010.

[190] Bryan J Johnson, Jordi Cohen, Richard W Welford, Arwen R Pearson, Klaus Schulten, Judith P Klinman, and Carrie M Wilmot. Exploring molecular oxygen pathways in hansenula polymorpha copper-containing amine oxidase. *J. Biol. Chem.*, 282(24):17767–17776, 2007.

[191] Patricia Amara, Pierre Andreoletti, HéLène Marie Jouve, and Martin J Field. Ligand diffusion in the catalase from proteus mirabilis: A molecular dynamics study. *Protein Sci.*, 10(10):1927–1935, 2001.

[192] Olaseni Sode and Gregory A Voth. Electron transfer activation of a second water channel for proton transport in [FeFe]-hydrogenase. *J. Chem. Phys.*, 141(22):12B630_1, 2014.

[193] Isaiah Sumner and Gregory A Voth. Proton transport pathways in [NiFe]-hydrogenase. *J. Phys. Chem. B*, 116(9):2917–2926, 2012.

[194] Jason C Crack, Jeffrey Green, Myles R Cheesman, Nick E Le Brun, and Andrew J Thomson. Superoxide mediated amplification of the oxygen induced switch from [4Fe-4S] to [2Fe-2S] clusters in the transcriptional regulator FNR. *Proc. Natl. Acad. Sci. U. S. A.*, 104(7):2092–2097, 2007.

[195] Matthew C Posewitz, Paul W King, Sharon L Smolinski, Liping Zhang, Michael Seibert, and Maria L Ghirardi. Discovery of two novel radical s-adenosylmethionine proteins required for the assembly of an active [Fe]-hydrogenase. *J. Biol. Chem.*, 279(24):25711–25720, 2004.

[196] Carole Baffert, Marie Demuez, Laurent Cournac, Benedicte Burlat, Bruno Guigliarelli, Patrick Bertrand, Laurence Girbal, and Christophe Léger. Hydrogen-activating enzymes: Activity does not correlate with oxygen sensitivity. *Angew. Chem.*, 47(11):2052–2054, 2008.

[197] A. Bondi. Van der waals volumes and radii. *J. Phys. Chem.*, 68(3):441–451, 1964.

[198] Jordi Cohen, Kwiseon Kim, Paul King, Michael Seibert, and Klaus Schulten. Finding gas diffusion pathways in proteins: application to $O_2$ and $H_2$ transport in CpI [FeFe]-hydrogenase and the role of packing defects. *Structure*, 13(9):1321–1329, 2005.

[199] Udo Radius, F Matthias Bickelhaupt, Andreas W Ehlers, Norman Goldberg, and Roald Hoffmann. Is CO a special ligand in organometallic chemistry? theoretical investigation of AB, Fe (CO) 4AB, and Fe (AB) 5 (AB= N2, CO, BF, SiO). *Inorg. Chem.*, 37(5):1080–1090, 1998.

[200] Sven T Stripp, Gabrielle Goldet, Caterina Brandmayr, Oliver Sanganas, Kylie A Vincent, Michael Haumann, Fraser A Armstrong, and Thomas Happe. How oxygen attacks [FeFe] hydrogenases from photosynthetic organisms. *Proc. Natl. Acad. Sci. U. S. A.*, 106(41):17331–17336, 2009.

[201] Jamin Koo and James R Swartz. System analysis and improved [FeFe] hydrogenase $O_2$ tolerance suggest feasibility for photosynthetic $H_2$ production. *Metab. Eng.*, 49:21–27, 2018.

[202] Hossein Mohammadiarani, Vincent S Shaw, Richard R Neubig, and Harish Vashisth. Interpreting hydrogen–deuterium exchange events in proteins using atomistic simulations: Case studies on regulators of G-protein signaling proteins. *J. Phys. Chem. B*, 122(40):9314–9323, 2018.

[203] James C Phillips, Rosemary Braun, Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D Skeel, Laxmikant Kale, and Klaus Schulten. Scalable molecular dynamics with NAMD. *J. Comput. Chem.*, 26(16):1781–1802, 2005.

[204] William Humphrey, Andrew Dalke, and Klaus Schulten. VMD: Visual molecular dynamics. *J. Mol. Graph.*, 14(1):33–38, 1996.

[205] William L Jorgensen, Jayaraman Chandrasekhar, Jeffry D Madura, Roger W Impey, and Michael L Klein. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, 79(2):926–935, 1983.

[206] Giacomo Fiorin, Michael L Klein, and Jérôme Hénin. Using collective variables to drive molecular dynamics simulations. *Mol. Phys.*, 111(22-23):3345–3362, 2013.

[207] Michael P Hayes, Christopher R Bodle, and David L Roman. Evaluation of the selectivity and cysteine dependence of inhibitors across the regulator of G-protein–signaling family. *Mol. Pharmacol.*, 93(1):25–35, 2018.

[208] Derek W Morris, Alana Rodgers, Kevin A McGhee, Siobhan Schwaiger, Paul Scully, John Quinn, David Meagher, John L Waddington, Michael Gill, and Aiden P Corvin. Confirming RGS4 as a susceptibility gene for schizophrenia. *Am. J. Med. Genet.*, 125(1):50–53, 2004.

[209] Michele H Potashman and Mark E Duggan. Covalent modifiers: an orthogonal approach to drug design. *J. Med. Chem.*, 52(5):1231–1246, 2009.

[210] Michael A James, Yan Lu, Yan Liu, Haris G Vikis, and Ming You. RGS17, an overexpressed gene in human lung and prostate cancer, induces tumor cell proliferation through the cyclic AMP-PKA-CREB pathway. *Cancer Res.*, 69(5):2108–2116, 2009.

[211] Christopher R Bodle, Duncan I Mackie, and David L Roman. RGS17: an emerging therapeutic target for lung and prostate cancers. *Future Med. Chem.*, 5(9):995–1007, 2013.

[212] Qin Wang, Akiko Terauchi, Christopher H Yee, Hisashi Umemori, and John R Traynor. 5-HT1A receptor-mediated phosphorylation of extracellular signal-regulated kinases (ERK1/2) is modulated by regulator of G-protein signaling protein 19. *Cell. Signal.*, 26(9):1846–1852, 2014.

[213] Eva de Alba, Luc De Vries, Marilyn Gist Farquhar, and Nico Tjandra. Solution structure of human GAIP ($G_\alpha$ interacting protein): a regulator of G-protein signaling. *J. Mol. Biol.*, 291(4):927–939, 1999.

[214] Veronica G Taylor, Paige A Bommarito, and John JG Tesmer. Structure of the regulator of G-protein signaling 8 (RGS8)-$G_{\alpha q}$ complex molecular basis for $G_\alpha$ selectivity. *J. Biol. Chem.*, 291(10):5138–5145, 2016.

[215] Levi L Blazer, David L Roman, Molly R Muxlow, and Richard R Neubig. Use of flow cytometric methods to quantify protein-protein interactions. *Curr. Protoc. Cytom.*, pages 13–11, 2010.

[216] David L Roman, Jeffery N Talbot, Rebecca A Roof, Roger K Sunahara, John R Traynor, and Richard R Neubig. Identification of small-molecule inhibitors of RGS4 using a high-throughput flow cytometry protein interaction assay. *Mol. Pharmacol.*, 71(1):169–175, 2007.

[217] Christopher O. Ortiz, John F. Etchberger, Shoshana L. Posy, Christian Frøkjær-Jensen, Shawn Lockery, Barry Honig, and Oliver Hobert. Searching for neuronal left-/right asymmetry: Genomewide analysis of nematode receptor-type guanylyl cyclases. *Genetics*, 173(1):131–149, 2006.

[218] Jie Liu, Alex Ward, Jingwei Gao, Yongming Dong, Nana Nishio, Hitoshi Inada, Lijun Kang, Yong Yu, Di Ma, Tao Xu, et al. C. elegans phototransduction requires a G-protein–dependent cgmp pathway and a taste receptor homolog. *Nat. Neurosci.*, 13(6):715, 2010.

[219] Mamoru Usuyama, Chisato Ushida, and Ryuzo Shingai. A model of the intracellular response of an olfactory neuron in *Caenorhabditis elegans* to odor stimulation. *PloS ONE*, 7(8):e42907, 2012.

[220] Lindy Holden-Dye and R Walker. Anthelmintic drugs and nematocides: Studies in *caenorhabditis elegans*. *WormBook: the online review of C. elegans biology*, pages 1–29, 2014.

[221] Sylvana Papaioannou, Lindy Holden-Dye, and Robert J Walker. Evidence for a role for cyclic AMP in modulating the action of 5-HT and an excitatory neuropeptide, FLP17A, in the pharyngeal muscle of caenorhabditis elegans. *Invertebr. Neurosci.*, 8(2):91, 2008.

[222] Michael A Schade, Nicole K Reynolds, Claudia M Dollins, and Kenneth G Miller. Mutations that rescue the paralysis of *Caenorhabditis elegans ric-8* (Synembryn) mutants activate the $G_{\alpha_s}$ pathway and define a third major branch of the synaptic signaling network. *Genetics*, 169(2):631–649, 2005.

[223] DL Noelle and Cornelia I Bargmann. Olfaction and odor discrimination are mediated by the *C. elegans* guanylyl cyclase ODR-1. *Neuron*, 25(3):575–586, 2000.

[224] Suk-Woo Cho, Kyu Yeong Choi, and Chul-Seung Park. A new putative cyclic nucleotide-gated channel gene, cng-3, is critical for thermotolerance in *Caenorhabditis elegans*. *Biochem. Biophys. Res. Commun.*, 325(2):525–531, 2004.

[225] Renate K Hukema, Suzanne Rademakers, Martijn PJ Dekkers, Jan Burghoorn, and Gert Jansen. Antagonistic sensory cues generate gustatory plasticity in *Caenorhabditis elegans*. *EMBO J.*, 25(2):312–322, 2006.

[226] Nicole K Charlie, Angela M Thomure, Michael A Schade, and Kenneth G Miller. The dunce CAMP phosphodiesterase pde-4 negatively regulates $G_{\alpha_s}$-dependent and $G_{\alpha_s}$-independent camp pools in the *Caenorhabditis elegans* synaptic signaling network. *Genetics*, 173(1):111–130, 2006.

[227] David M Raizen, John E Zimmerman, Matthew H Maycock, Uyen D Ta, Young-jai You, Meera V Sundaram, and Allan I Pack. Lethargus is a *Caenorhabditis elegans* sleep-like state. *Nature*, 451(7178):569, 2008.

[228] Julian Ceron, Jean-François Rual, Abha Chandra, Denis Dupuy, Marc Vidal, and Sander van den Heuvel. Large-scale RNAi screens identify novel genes that interact with the *C. elegans* retinoblastoma pathway as well as splicing-related components with synMuv B activity. *BMC Dev. Biol.*, 7(1):30, 2007.

[229] Seongseop Kim, J Amaranath Govindan, Zheng Jin Tu, and David Greenstein. Sacy-1 dead-box helicase links the somatic control of oocyte meiotic maturation to the sperm-to-oocyte switch and gamete maintenance in *Caenorhabditis elegans*. *Genetics*, 192(3):905–928, 2012.

[230] Caroline Schmitz, Parag Kinge, and Harald Hutter. Axon guidance genes identified in a large-scale RNAi screen using the RNAi-hypersensitive *Caenorhabditis elegans* strain nre-1 (hd20) lin-15b (hd126). *Proc. Natl. Acad. Sci. U. S. A.*, 104(3):834–839, 2007.

[231] Thomas Seebeck, Geert Jan Sterk, and Hengming Ke. Phosphodiesterase inhibitors as a new generation of antiprotozoan drugs: exploiting the benefit of enzymes that are highly conserved between host and parasite. *Future Med. Chem.*, 3(10):1289–1306, 2011.

[232] Inonge Gross and Jörg Durner. In search of enzymes with a role in 3′, 5′-cyclic guanosine monophosphate metabolism in plants. *Front. Plant. Sci.*, 7:576, 2016.

[233] Marco Conti and Joseph Beavo. Biochemistry and physiology of cyclic nucleotide phosphodiesterases: essential components in cyclic nucleotide signaling. *Annu. Rev. Biochem.*, 76:481–511, 2007.

[234] Sharron H Francis, Mitsi A Blount, and Jackie D Corbin. Mammalian cyclic nucleotide phosphodiesterases: molecular mechanisms and physiological functions. *Physiol. Rev.*, 91(2):651–690, 2011.

[235] Qing Huai, John Colicelli, and Hengming Ke. The crystal structure of AMP-bound PDE4 suggests a mechanism for phosphodiesterase catalysis. *Biochemistry*, 42(45):13220–13226, 2003.

[236] Donald H Maurice, Hengming Ke, Faiyaz Ahmad, Yousheng Wang, Jay Chung, and Vincent C Manganiello. Advances in targeting cyclic nucleotide phosphodiesterases. *Nat. Rev. Drug Discov.*, 13(4):290, 2014.

[237] Pasqualina D'Ursi, Sara Guariento, Gabriele Trombetti, Alessandro Orro, Elena Cichero, Luciano Milanesi, Paola Fossa, and Olga Bruno. Further insights in the binding mode of selective inhibitors to human pde4d enzyme combining docking and molecular dynamics. *Mol. Inform.*, 35(8-9):369–381, 2016.

[238] Jing Huang and Alexander D MacKerell Jr. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J. Comput. Chem.*, 34(25):2135–2145, 2013.

[239] Alex B Burgin, Olafur T Magnusson, Jasbir Singh, Pam Witte, Bart L Staker, Jon M Bjornsson, Margret Thorsteinsdottir, Sigrun Hrafnsdottir, Timothy Hagen, Alex S Kiselyov, et al. Design of phosphodiesterase 4D (PDE4D) allosteric modulators for enhancing cognition with improved safety. *Nat. Biotechnol.*, 28(1):63, 2010.

[240] Qing Huai, Yudong Liu, Sharron H Francis, Jackie D Corbin, and Hengming Ke. Crystal structures of phosphodiesterases 4 and 5 in complex with inhibitor 3-isobutyl-1-methylxanthine suggest a conformation determinant of inhibitor selectivity. *J. Biol. Chem.*, 279(13):13095–13101, 2004.

[241] Mi Eun Lee, Joseph Markowitz, Jie-Oh Lee, and Hayyoung Lee. Crystal structure of phosphodiesterase 4D and inhibitor complex. *FEBS Lett.*, 530(1-3):53–58, 2002.

[242] Andrew Waterhouse, Martino Bertoni, Stefan Bienert, Gabriel Studer, Gerardo Tauriello, Rafal Gumienny, Florian T Heer, Tjaart A P de Beer, Christine Rempfer, Lorenza Bordoli, et al. Swiss-model: homology modelling of protein structures and complexes. *Nucleic Acids Res.*, 46(W1):W296–W303, 2018.

[243] Mike P Allen and Dominic J Tildesley. *Computer simulation of liquids*. Oxford university press, 1989.

[244] Tom Darden, Darrin York, and Lee Pedersen. Particle mesh ewald: An N.log(N) method for ewald sums in large systems. *J. Chem. Phys.*, 98(12):10089–10092, 1993.

[245] Benjamin A Shoemaker, Dachuan Zhang, Ratna R Thangudu, Manoj Tyagi, Jessica H Fong, Aron Marchler-Bauer, Stephen H Bryant, Thomas Madej, and Anna R Panchenko. Inferred biomolecular interaction serverâĂŤa web server to analyze and predict protein interacting partners and binding sites. *Nucleic Acids Res.*, 38(suppl_1):D518–D524, 2009.

[246] Benjamin A Shoemaker, Dachuan Zhang, Manoj Tyagi, Ratna R Thangudu, Jessica H Fong, Aron Marchler-Bauer, Stephen H Bryant, Thomas Madej, and Anna R

Panchenko. Ibis (inferred biomolecular interaction server) reports, predicts and integrates multiple types of conserved interactions for proteins. *Nucleic Acids Res.*, 40(D1):D834–D840, 2011.

[247] Yu Chen and Brian K Shoichet. Molecular docking and ligand specificity in fragment-based inhibitor discovery. *Nat. Chem. Biol.*, 5(5):358, 2009.

[248] Chen Wang, Pan Xu, Luyu Zhang, Jing Huang, Kongkai Zhu, and Cheng Luo. Current strategies and applications for precision drug design. *Front Pharmacol.*, 9:787, 2018.

[249] Pratyush Tiwary, Vittorio Limongelli, Matteo Salvalaglio, and Michele Parrinello. Kinetics of protein–ligand unbinding: Predicting pathways, rates, and rate-limiting steps. *Proc. Natl. Acad. Sci. U. S. A.*, 112(5):E386–E391, 2015.

[250] Samuel Genheden and Ulf Ryde. The mm/pbsa and mm/gbsa methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.*, 10(5):449–461, 2015.

# APPENDIX A

# MATLAB SCRIPTS

## A.1 Sample Code for Free-energy reconstruction using Radial Basis Functions

The following script can be used for reconstruction of the free-energy surface (FES) as a linear combination of Gaussian radial basis functions (see section 2.4.1.1) [141].

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% The sample script is tested for PMF reconstruction                          %
% (c) Luca Maragliano, Istituto Italiano di Tecnologia, Genoa, Italy          %
% Modified (Lines 97-127)and compiled by M. Mohammadi                         %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
clear all
load meanforces_new.dat
zkcsa=meanforces_new;
n1 = size(zkcsa,1);
zc=zkcsa(1:n1,1:3);
xc=zc';
dz=zkcsa(1:n1,4:6);
fc=-dz';
%% plot centers & forces
plot(.5*(xc(1,:)+xc(2,:)),xc(3,:),'.')
figure
plot3(xc(1,:),xc(2,:),xc(3,:),'.b')
hold on
quiver3(xc(1,:),xc(2,:),xc(3,:),fc(1,:),fc(2,:),fc(3,:),'r')
grid on
axis equal
axis vis3d
xlabel('z1')
ylabel('z2')
zlabel('z3')
%%
d1 = norm(xc(:,1)-xc(:,2))
[ac,sigma] = reconstruct2(xc,fc);
%% build potential
n1 = size(xc,2);
bord=.5;
sigma0 = sigma(1,1);
sigma=sigma0*ones(1,n1);
```

```
 xmin=min(xc(1,:))-bord;
 xmax=max(xc(1,:))+bord;
 ymin=min(xc(2,:))-bord;
 ymax=max(xc(2,:))+bord;
 zmin=min(xc(3,:))-bord;
 zmax=max(xc(3,:))+bord;
[xx,yy,zz] = meshgrid(linspace(xmin,xmax,100),...
    linspace(ymin,ymax,100),linspace(zmin,zmax,100));
clear r2;
V3 = xx.*0;
        for j=1:n1
            r2 = (xx-xc(1,j)).^2+(yy-xc(2,j)).^2 + (zz-xc(3,j)).^2;
            V3 = V3 + ac(j)*exp(-r2/sigma(j)^2*0.5);
        end;
V3m=V3-min(V3(:));
%% plot potential
figure
t=0.
p = patch(isosurface(xx,yy,zz,V3m,t));
isonormals(xx,yy,zz,V3m,p)
set(p,'FaceColor','red','EdgeColor','none');
camlight
%axis([-22 18 -12 10 -10 18])
view([-156 16])
t=1
p = patch(isosurface(xx,yy,zz,V3m,t));
isonormals(xx,yy,zz,V3m,p)
set(p,'FaceColor','yello','EdgeColor','none');
t=6
p = patch(isosurface(xx,yy,zz,V3m,t));
isonormals(xx,yy,zz,V3m,p)
set(p,'FaceColor','blue','EdgeColor','none');
alpha(.4)
t=7.7
p = patch(isosurface(xx,yy,zz,V3m,t));
isonormals(xx,yy,zz,V3m,p)
set(p,'FaceColor','green','EdgeColor','none');
alpha(.4)
grid on
%axis equal
axis vis3d
xlabel('z1')
ylabel('z2')
zlabel('z3')
function [ac,sigma] = reconstruct2(xc,fc)
% xc is an nd x n1 matrix with the centers coordinate:
% nd is the dimensionality of the free energy space,
% n1 is the number of centers
% fc is a nd x n1 matrix with the forces
% the outputs are: ac, the coefficients of the radial-basis fcts, and
% sigma, the width of the radial-basis fcts.
% the function also plots the relative residual and prints the condition
% number of the matrix A at the optimal sigma
nd = size(xc,1);
```

```
n1 = size(xc,2);
% distance between centers
d1 = norm(xc(:,1)-xc(:,2));
bb = reshape(fc',nd*n1,1);
AA = zeros(n1*nd,n1);
% minimization loop
% sigmav = linspace(0.5*d1,2*d1,1e2);
% sigmav = linspace(5*pi,6*pi,1e1);
sigmav = linspace(d1,2*d1,2e2);
ss1 = size(sigmav,2);
resvrel = Inf(1,ss1);
        B1 = xc'*xc;
        r2 = diag(B1)*ones(1,n1)+ones(n1,1)*diag(B1)'-2*B1;
        for i=1:ss1
            sigma = sigmav(i);
            sigma2 = 0.5./sigma.^2;
            rr = exp(-r2.*sigma2);
            for j = 1:nd
                AA((j-1)*n1+1:j*n1,:) = (xc(j,:)'*ones(1,n1)-ones(n1,1)*xc(j
                    ,:)).*rr;
            end
            AA = AA./sigma.^2;
            if rank(AA)<n1; break; end;
            gg = AA\bb;
            resrelv(i) = norm(AA*gg-bb)/norm(bb);
            resv(i) = norm(AA*gg-bb);
            if i>1; if resrelv(i)>resrelv(i-1); break; end; end;
            fprintf('iteration number = %3d; sigma = %f; residual = %f; cond.
                number = %e\n',...
                 i,sigmav(i),resv(i)/n1,cond(AA)')
        end
figure(1);clf;
plot(sigmav(1:i-1),resv(1:i-1)/n1,'o-')
set(gca,'FontSize',16);
title('residual per center |AA*gg-bb|/n1','FontSize',16)
xlabel('\sigma','FontSize',16)
ylabel('residual per center','FontSize',16)
[i1 i2] = min(resrelv);
sigma = sigmav(i2);
%% reconstruction step
        sigma2 = 0.5/sigma^2;
        rr = exp(-r2*sigma2);
        for j = 1:nd
            AA((j-1)*n1+1:j*n1,:) = (xc(j,:)'*ones(1,n1)-ones(n1,1)*xc(j,:)).*
                rr;
        end
        AA = AA/sigma^2;
% ac are the coefficients in the radial-basis representation of the free
% energy
ac = AA\bb;
sigma = sigma*ones(1,n1);
fprintf('\n')
fprintf('# of centers = %3d; optimal sigma = %f; rel. residual = %f; cond.
    number = %e\n',...
```

```
    n1,sigma(1,1),norm(AA*ac-bb)/norm(bb),cond(AA))
fprintf('# of centers = %3d; optimal sigma = %f; residual / cents = %f; cond. number = %e\n',...
    n1,sigma(1,1),norm(AA*ac-bb)/n1,cond(AA))
fprintf('\n')
end
```

The following is the input file ("meanforces_new.dat") containing the coordinates of the centers and the corresponding mean-force values.

```
1.470000 -2.120000 -7.710000 -9.785590 4.990049 -0.927257
1.470000 -2.120000 -8.280000 -9.812816 4.286304 -0.473017
1.560000 -2.410000 -9.890000 -6.371151 0.564879 0.349201
1.670000 -2.270000 -8.780000 -4.508113 1.930594 -0.806485
1.780000 -2.170000 -7.310000 -4.063549 1.879347 3.359805
1.800000 -2.210000 -9.440000 -3.680580 3.214869 -2.117216
1.840000 -1.670000 -5.830000 -6.623003 5.578274 3.747816
1.880000 -2.770000 -9.020000 -0.032609 -4.296424 0.694036
1.890000 -2.390000 -10.450000 -1.470600 0.430322 0.394696
1.890000 -2.870000 -7.610000 0.880890 -4.438182 3.351721
10.160000 0.380000 -5.280000 0.329397 -12.581186 5.083637
10.180000 2.250000 -5.310000 -0.890966 7.292744 -0.146103
10.210000 1.030000 -5.470000 -1.288463 -0.375553 0.266809
10.210000 2.750000 -5.890000 -0.472463 5.437764 -6.729431
10.250000 1.750000 -4.570000 -0.497051 2.693496 10.859090
10.320000 2.090000 -6.060000 0.531865 7.153510 -7.271102
10.330000 1.710000 -5.160000 -1.305893 5.445777 2.823317
10.360000 1.310000 -4.320000 -1.538975 -0.944368 11.377527
10.380000 0.900000 -6.190000 -0.067118 -1.471345 -5.396713
10.400000 0.270000 -6.030000 -1.254845 -9.305528 -0.620820
10.400000 3.840000 -5.520000 -1.538517 1.110011 -1.950121
10.470000 2.190000 -6.600000 -0.763081 8.472491 -8.433132
10.520000 1.420000 -6.180000 -0.929578 5.664027 -6.990566
10.530000 3.020000 -6.600000 -0.656510 2.564594 -5.712497
10.610000 2.920000 -5.320000 -0.558751 1.545318 0.086698
10.700000 1.360000 -5.350000 -0.630059 3.129134 0.481480
10.730000 2.560000 -4.840000 -0.295766 5.813913 5.884236
10.780000 0.960000 -6.870000 -0.274745 2.583686 -6.032612
10.790000 0.510000 -4.670000 -0.604078 -10.895482 12.771118
10.790000 1.000000 -5.860000 -1.229582 0.004796 -3.664434
10.800000 0.510000 -5.400000 -0.466494 -10.689803 1.947389
10.910000 -0.160000 -6.220000 -0.513956 -7.874482 -1.876247
10.950000 2.820000 -5.840000 -0.296766 2.573710 -5.011045
11.010000 2.040000 -6.080000 -0.665404 7.711624 -7.861041
11.020000 1.010000 -4.930000 -0.517248 -1.643624 8.777645
11.080000 1.460000 -6.130000 -0.271051 6.087383 -6.318860
11.130000 1.840000 -4.860000 -0.373715 5.104568 7.103900
11.180000 1.800000 -4.330000 -0.208607 2.583141 8.694338
11.190000 1.040000 -4.400000 -0.188699 -3.668468 12.136767
11.200000 3.000000 -4.150000 -0.625936 3.359043 4.409662
11.220000 0.650000 -6.370000 -0.324947 -3.706987 -5.135758
```

```
11.230000 0.670000 -5.740000 -0.090179 -4.240865 -0.574641
11.410000 0.460000 -4.650000 -0.181423 -10.398926 12.095535
11.420000 2.350000 -5.390000 -1.269670 7.852689 -1.725379
11.430000 1.370000 -4.980000 0.033519 2.122918 6.755236
11.500000 1.580000 -7.050000 -0.227878 7.616850 -6.348085
11.570000 1.310000 -6.220000 -0.112256 5.913774 -6.734546
11.600000 1.580000 -5.550000 -0.216202 5.784586 -2.326657
11.600000 1.960000 -4.030000 -0.249035 2.671545 6.284807
11.690000 1.880000 -4.560000 -0.120765 3.634639 9.053479
11.720000 0.630000 -5.560000 0.308886 -7.141562 0.668203
11.810000 0.560000 -6.690000 -0.013099 -4.162477 -6.105688
11.860000 2.000000 -6.290000 0.119091 7.982332 -7.276114
11.930000 0.930000 -4.470000 -0.032969 -3.756202 12.239510
11.930000 1.350000 -6.620000 -0.201110 4.670160 -7.090052
11.970000 1.100000 -4.970000 -0.309290 -1.145643 7.833637
11.980000 0.460000 -5.080000 -0.350485 -10.708945 6.540088
12.020000 2.730000 -5.770000 -0.098955 4.416381 -5.927511
12.080000 1.620000 -5.820000 -0.434442 6.149895 -5.374572
12.180000 2.170000 -5.330000 -0.993701 7.545303 -0.509524
12.190000 0.190000 -4.510000 -0.504289 -10.683225 11.935753
12.230000 0.630000 -6.110000 -0.674993 -3.808016 -3.392482
12.230000 3.020000 -5.070000 -0.723733 0.707082 3.381953
12.250000 0.330000 -5.650000 -0.225927 -12.308530 -0.174033
12.260000 3.780000 -5.510000 -0.832179 -0.529532 -1.538337
12.310000 1.130000 -6.200000 -0.592559 2.412709 -6.282849
12.320000 1.060000 -5.570000 -0.609169 1.395056 -1.438595
12.400000 0.200000 -6.610000 -0.288645 -6.167720 -5.226179
12.400000 1.530000 -5.140000 -1.124196 3.479462 3.749120
12.410000 0.720000 -4.790000 -0.562788 -7.009042 11.317314
12.480000 3.250000 -3.740000 -0.503941 0.605541 0.389208
2.020000 -1.880000 -8.320000 -2.020616 5.168058 -0.834765
2.020000 -2.030000 -9.960000 -1.620296 4.503261 1.162102
2.040000 -2.450000 -7.980000 0.073659 -0.876309 -1.257002
2.040000 -2.660000 -8.540000 1.127652 -4.122007 0.563482
2.050000 -1.690000 -6.490000 -4.047220 7.870610 0.628047
2.060000 -2.550000 -9.740000 0.461196 -2.254338 -1.683790
2.080000 -2.210000 -6.390000 -0.676476 -5.404862 2.885443
2.100000 -1.840000 -7.000000 -2.777419 3.180030 5.175967
2.100000 -1.930000 -7.660000 -1.785870 5.148049 -1.229410
2.150000 -1.220000 -8.520000 -4.318319 15.486063 -2.222311
2.170000 -1.900000 -9.020000 -1.299470 6.323809 -0.809971
2.320000 -3.040000 -9.810000 4.911817 -6.365213 1.211002
2.350000 -2.420000 -8.860000 2.637871 -3.227060 -0.664330
2.420000 -1.790000 -5.860000 -3.144319 1.032389 3.857516
2.470000 -2.990000 -7.880000 6.539072 -6.672430 0.550712
2.500000 -1.620000 -9.330000 -3.054592 9.481162 -2.412334
2.500000 -2.640000 -10.150000 3.527522 -6.034656 1.114888
2.540000 -3.000000 -8.740000 6.265618 -9.942799 0.130870
2.550000 -1.970000 -6.810000 0.439530 -2.067726 6.908680
2.550000 -1.970000 -9.800000 -0.940055 4.023278 0.039894
2.560000 -1.910000 -8.680000 -0.381733 3.782025 -0.168624
2.560000 -2.240000 -7.460000 0.969988 -3.825991 1.928867
2.600000 -2.340000 -8.400000 1.891884 -3.560201 0.438040
2.610000 -0.760000 -8.700000 -5.530802 14.634261 -1.594754
```

```
2.660000 -1.900000 -7.930000 -1.109495 2.146963 -1.607371
2.670000 -3.250000 -8.280000 3.157939 -3.486741 -1.781799
2.710000 -2.110000 -10.280000 0.270010 0.184418 1.445464
2.740000 -1.930000 -6.300000 -2.207743 -1.718878 1.570019
2.780000 -2.420000 -6.710000 3.678854 -4.129190 2.967938
2.850000 -1.730000 -7.360000 -2.375362 3.999146 0.111007
2.900000 -1.990000 -9.300000 -0.972644 0.563813 -1.047719
2.940000 -1.140000 -6.380000 -5.559317 11.024156 -1.825307
2.940000 -1.540000 -8.770000 -5.549079 8.218361 -0.677302
2.990000 -2.530000 -9.380000 4.276616 -8.264305 -0.789744
3.000000 -1.030000 -5.170000 -7.008680 6.814184 5.735588
3.010000 -2.820000 -8.790000 0.951670 -5.546646 1.371082
3.030000 -1.870000 -9.860000 -2.863721 1.601365 -0.521689
3.070000 -2.750000 -7.780000 1.635657 -4.038255 0.695380
3.080000 -2.260000 -7.680000 1.876295 -7.781997 0.635802
3.120000 -1.260000 -7.090000 -4.728191 9.068522 0.430987
3.130000 -1.680000 -6.790000 -0.885649 -0.628200 5.988530
3.130000 -2.610000 -7.260000 0.925797 -10.283233 0.693989
3.210000 -1.440000 -7.960000 -4.838812 6.758490 -2.426404
3.220000 -1.800000 -5.430000 -3.362104 -7.704335 7.602412
3.230000 -1.300000 -9.980000 -6.704521 9.601720 1.268847
3.230000 -1.550000 -10.460000 -4.596837 6.005295 0.493376
3.230000 -2.210000 -5.820000 -5.736266 -10.563909 7.563982
3.230000 -2.540000 -8.260000 3.307465 -10.315886 -0.258160
3.270000 -4.240000 -7.850000 -1.445826 -0.675387 4.983644
3.290000 -2.000000 -8.920000 0.503675 -3.282220 -0.698471
3.300000 -1.480000 -9.140000 -6.800216 6.508543 -0.800016
3.310000 -1.840000 -7.530000 -1.662242 -3.074127 2.088493
3.330000 -2.060000 -10.490000 0.113140 -4.688570 0.718155
3.380000 -2.040000 -8.160000 -0.152428 -6.648301 -0.529211
3.400000 -3.430000 -8.030000 -1.692701 -2.792933 -1.202766
3.430000 -2.630000 -10.140000 -0.738600 -6.175346 1.232766
3.450000 -1.510000 -5.990000 0.115159 -5.966683 7.742243
3.450000 -4.290000 -8.640000 0.279142 -1.377745 -1.265891
3.500000 -2.380000 -9.630000 3.154513 -9.911830 -1.664458
3.510000 -2.370000 -6.460000 2.927198 -15.818146 2.007969
3.510000 -3.070000 -6.260000 -3.161518 -18.433693 11.068323
3.520000 -1.020000 -6.770000 -3.424899 8.114120 1.784130
3.540000 -0.800000 -7.560000 -7.025334 9.459388 -0.289985
3.540000 -2.730000 -8.810000 -3.272266 -5.279181 -1.238291
3.550000 -1.580000 -9.640000 -2.130468 2.722263 -2.072569
3.550000 -2.970000 -8.140000 -3.163119 -3.457108 -0.964448
3.590000 -1.390000 -7.590000 -3.249656 3.947318 1.177240
3.610000 -2.290000 -8.580000 2.735607 -9.292341 -0.888280
3.630000 -0.150000 -7.410000 -7.766808 9.189951 -0.475645
3.640000 -0.810000 -8.380000 -7.267304 11.606289 -2.581550
3.650000 -1.570000 -6.470000 1.162952 -4.141665 7.503477
3.660000 -0.980000 -8.920000 -4.490235 9.457324 -0.109234
3.680000 -1.280000 -10.350000 -2.787161 6.820652 0.874688
3.690000 0.120000 -5.780000 -6.372637 2.239877 -0.305453
3.700000 -2.130000 -6.940000 1.409731 -7.795042 6.556067
3.760000 -2.080000 -10.040000 2.099290 -4.971352 0.903574
3.780000 -2.570000 -7.580000 -3.065761 -4.391506 2.295742
3.790000 -1.520000 -7.120000 2.604494 -0.294103 2.959327
```

```
3.790000 -2.820000 -9.520000 -3.012678 -5.863907 -1.780807
3.790000 -3.140000 -10.210000 -2.723597 -5.628188 0.808057
3.820000 -1.370000 -8.510000 0.117832 3.867569 -2.052579
3.820000 -4.460000 -8.210000 0.772448 -5.439717 3.618620
3.820000 -4.820000 -10.020000 1.015474 1.153112 0.444472
3.830000 -1.730000 -8.100000 2.382660 -3.871209 -1.386598
3.870000 -1.950000 -9.050000 -0.778882 -0.102451 -0.631741
3.880000 -0.440000 -6.240000 0.349090 -1.640122 -1.758444
3.890000 -1.010000 -5.130000 -2.813109 4.148113 5.958722
3.970000 -1.710000 -10.310000 4.767743 -3.070539 1.235396
4.000000 -1.740000 -7.560000 5.206293 -5.599441 1.658151
4.020000 -1.460000 -9.310000 4.442137 0.699896 -1.343906
4.040000 -2.440000 -8.220000 -2.390232 -4.325837 0.072488
4.060000 -0.830000 -7.240000 -0.403484 5.645580 3.967544
4.060000 -1.120000 -6.810000 3.328613 2.868491 4.076771
4.080000 -2.310000 -6.500000 -1.381826 -9.075440 1.049330
4.100000 -1.250000 -9.850000 3.599599 3.537593 0.336824
4.130000 -0.520000 -7.920000 -3.225067 7.669725 -3.988029
4.140000 -1.880000 -5.730000 -1.389061 -3.635261 5.614716
4.140000 -2.350000 -10.420000 -0.085269 -2.765264 0.968779
4.140000 -4.950000 -9.090000 1.041130 2.274613 -0.229843
4.150000 -1.090000 -7.750000 3.057731 4.131807 -0.640695
4.160000 -1.170000 -6.270000 4.723494 -1.666095 4.575410
4.170000 -2.600000 -8.710000 0.294758 -6.096677 -0.550167
4.180000 -2.590000 -9.890000 0.494912 -5.032680 0.418716
4.180000 -3.370000 -8.830000 1.951023 -5.788995 -1.321839
4.190000 -1.180000 -8.880000 4.772582 3.340467 0.007011
4.190000 -2.940000 -6.440000 0.285906 -8.209997 7.013270
4.190000 -3.600000 -6.990000 0.495233 -8.116226 4.869160
4.190000 0.310000 -5.920000 -5.875549 -5.888099 0.831154
4.190000 0.730000 -7.170000 -5.480088 8.180393 -3.603162
4.200000 -1.830000 -8.590000 3.647515 -1.879768 -1.340319
4.210000 -0.080000 -7.580000 -5.382754 7.006465 -2.415755
4.210000 -2.030000 -6.990000 2.688249 -3.912597 5.742791
4.220000 -0.530000 -8.490000 -2.488143 7.118298 -2.298807
4.220000 -2.420000 -9.200000 0.319968 -4.373995 -0.413001
4.230000 -5.100000 -10.240000 1.872499 2.316112 1.220353
4.240000 -1.040000 -5.610000 4.805478 -6.281026 5.954221
4.270000 -2.680000 -7.600000 -0.027437 -11.381655 -2.144056
4.270000 -3.020000 -8.060000 2.405375 -5.822452 -0.615311
4.290000 -1.440000 -8.250000 6.813849 -0.321762 -1.935248
4.290000 -1.780000 -9.630000 3.458713 0.795462 -1.249165
4.310000 -0.460000 -6.810000 -0.954586 2.871875 3.820897
4.320000 -1.510000 -7.110000 7.691676 -3.497254 4.865637
4.320000 0.180000 -6.550000 -1.403350 -5.783934 4.431160
4.330000 -2.050000 -7.820000 0.688616 -1.286833 1.513804
4.340000 -0.330000 -5.310000 0.268397 -3.452452 1.169911
4.340000 -0.960000 -10.370000 4.821383 4.327572 0.824766
4.380000 -1.900000 -5.110000 -0.042070 -10.142058 15.852719
4.380000 0.730000 -5.330000 -6.830413 0.573644 3.417525
4.390000 -0.580000 -5.820000 -0.806802 2.181452 -0.866439
4.400000 -2.100000 -6.140000 1.038072 -6.348476 2.449267
4.400000 -4.880000 -9.820000 2.368387 2.986602 -0.159066
4.420000 0.720000 -6.090000 -6.008680 3.253870 -2.514157
```

```
4.440000 0.380000 -4.930000 -4.666213 -8.625240 7.486384
4.460000 -0.720000 -7.590000 3.086580 3.153877 -0.737242
4.480000 -1.750000 -6.590000 1.308740 4.159887 -1.034718
4.500000 -1.510000 -10.160000 6.703095 -0.391844 1.595561
4.500000 -2.900000 -10.280000 4.377151 -7.758788 0.832771
4.510000 -2.480000 -6.970000 3.726110 -1.892715 2.051724
4.520000 1.200000 -7.160000 -7.566908 16.874228 -4.073073
4.540000 -0.900000 -6.950000 6.342940 0.016932 4.588568
4.570000 -2.890000 -8.860000 4.609167 -8.046249 0.561148
4.600000 -2.180000 -5.630000 1.735749 -12.507701 8.796243
4.630000 -5.300000 -9.440000 4.628955 0.755702 -0.262614
4.640000 -2.290000 -8.530000 3.397215 -3.900287 -0.563893
4.670000 -1.630000 -5.690000 2.511834 1.990933 3.175886
4.670000 -2.460000 -7.800000 4.120627 -6.442432 1.278630
4.670000 0.200000 -7.330000 -2.734028 3.678339 1.063417
4.680000 -0.420000 -6.270000 -0.037796 1.173538 -2.268326
4.680000 -1.140000 -8.700000 7.454345 2.141599 -2.258181
4.680000 -1.300000 -6.360000 5.297482 4.457757 -3.508386
4.690000 -0.260000 -7.660000 1.360590 2.313850 -2.382844
4.700000 -1.520000 -7.760000 1.766637 6.600609 0.940029
4.710000 -5.460000 -10.200000 5.508190 -1.893559 1.051645
4.720000 -0.910000 -5.050000 0.169857 3.918847 8.182927
4.720000 -2.660000 -9.470000 5.272360 -6.445159 -1.442031
4.720000 -2.720000 -8.270000 3.769986 -7.816313 -0.133827
4.750000 -1.020000 -7.930000 7.276638 1.175270 -2.425706
4.760000 -2.060000 -8.970000 3.356487 0.279907 -0.837866
4.760000 1.090000 -5.500000 -6.101597 9.359843 0.203263
4.800000 0.430000 -6.420000 0.467138 -4.873386 4.792695
4.810000 -1.190000 -7.370000 7.241381 -0.728716 3.841719
4.820000 0.290000 -5.750000 -1.780848 -8.969543 0.837479
4.830000 -1.330000 -9.520000 2.955362 11.655929 -1.018340
4.830000 -2.100000 -6.690000 4.741772 -2.847821 5.423159
4.850000 -0.320000 -7.010000 3.437222 -1.181741 4.264129
4.860000 0.970000 -6.130000 -3.615946 6.265728 -3.571929
4.880000 -2.590000 -10.120000 4.951982 -6.331773 1.560828
4.920000 -2.550000 -6.420000 4.253723 -14.110398 2.705759
4.940000 0.930000 -7.110000 -5.500820 10.785299 3.028568
4.950000 -1.180000 -5.420000 1.194126 6.554039 3.443597
4.960000 -1.930000 -7.230000 5.394983 -1.395456 3.952844
4.970000 -2.920000 -7.750000 6.492949 -7.132185 0.864996
4.980000 -2.040000 -9.800000 5.358382 0.239460 -0.313196
5.000000 -1.440000 -4.970000 1.933756 0.447480 15.704404
5.000000 -1.620000 -8.240000 6.362592 -0.586621 -1.766079
5.000000 0.600000 -5.180000 -1.855560 -5.553953 4.787424
5.020000 -2.020000 -5.890000 3.516523 -7.076170 2.975502
5.030000 1.460000 -7.160000 -5.047748 17.038799 -3.534401
5.050000 -1.390000 -8.940000 6.381571 1.575937 0.620047
5.070000 -0.890000 -5.920000 0.886655 9.224613 -2.262109
5.070000 -2.070000 -10.410000 5.444685 -1.265097 0.907437
5.140000 -0.550000 -5.220000 -0.858790 3.126429 2.643051
5.150000 0.850000 -4.730000 -1.016898 -2.063311 9.431469
5.160000 -0.250000 -6.150000 0.242950 2.200820 -3.348519
5.170000 -1.470000 -7.120000 3.543595 5.844263 0.879714
5.180000 -2.450000 -7.170000 7.247987 -5.202087 3.189444
```

```
5.190000 0.380000 -6.860000 -0.025898 1.250865 3.502156
5.200000 -0.090000 -5.570000 0.087813 -8.404875 0.473690
5.200000 0.520000 -6.110000 0.467435 -3.455571 -2.608300
5.210000 -2.150000 -9.350000 5.476206 -2.749068 -1.138429
5.260000 -1.230000 -6.280000 2.189622 9.663296 -2.963120
5.310000 0.540000 -7.630000 -1.874913 2.550372 -1.471901
5.330000 -0.110000 -7.510000 3.376774 -2.620245 -0.574225
5.340000 -2.220000 -8.420000 5.083711 -4.508308 -1.000604
5.370000 0.170000 -5.090000 0.232935 -9.805058 5.698012
5.380000 1.570000 -4.840000 -4.419427 12.692011 6.333938
5.400000 -1.930000 -7.850000 5.018345 -2.792842 1.909229
5.450000 1.030000 -5.270000 -0.577851 3.179862 2.316042
5.460000 0.110000 -6.470000 1.581735 -8.000470 -3.449772
5.480000 -1.860000 -6.220000 4.560303 -0.890183 -0.294788
5.520000 -0.370000 -7.090000 0.531156 -1.252364 -2.644332
5.520000 -2.330000 -10.350000 5.853322 -5.025993 0.854952
5.520000 1.510000 -5.440000 -2.230416 12.202447 -0.385406
5.530000 -0.960000 -7.550000 3.987823 11.770851 -3.538373
5.550000 -1.510000 -6.790000 4.475621 6.599496 -1.931426
5.550000 0.960000 -6.110000 0.789043 4.809904 -4.019577
5.560000 -1.630000 -8.310000 5.365530 4.559041 -1.996384
5.560000 -1.650000 -5.430000 3.212084 -2.608166 6.615633
5.570000 -1.950000 -8.850000 4.798322 -0.391977 -2.439367
5.640000 1.070000 -6.640000 1.207548 5.886533 -5.675743
5.650000 1.330000 -4.310000 -1.457491 4.520629 10.358261
5.660000 -2.510000 -8.050000 4.028168 -9.029806 -3.423799
5.710000 0.790000 -8.110000 -3.554358 2.904050 -2.298852
5.740000 0.240000 -8.080000 -0.721254 -4.413234 -1.459978
5.750000 -0.380000 -7.790000 -1.373982 -1.437502 -2.613933
5.750000 -1.420000 -6.080000 4.563731 5.754505 -2.253122
5.750000 -1.820000 -7.130000 5.698422 -1.387835 3.877744
5.750000 1.420000 -7.150000 -0.401374 11.007213 -3.888427
5.780000 -0.680000 -6.460000 -0.391152 4.965101 -3.784924
5.820000 0.560000 -5.630000 2.890539 -6.396539 0.541520
5.830000 -2.210000 -6.060000 4.781170 -11.652579 2.739130
5.850000 0.810000 -7.110000 2.555555 1.336511 -4.327075
5.860000 2.000000 -5.530000 -3.603476 14.668583 -2.166133
5.870000 0.140000 -4.960000 -1.419550 -9.399266 5.992914
5.930000 -0.380000 -5.590000 0.528337 0.490034 -1.294186
5.940000 1.650000 -6.150000 -1.009981 12.416338 -5.603166
5.950000 -1.260000 -5.540000 4.359454 4.469948 1.812648
5.990000 0.060000 -6.700000 3.291725 -7.901625 -4.060096
6.010000 -1.840000 -6.390000 5.871569 -1.725065 -1.739007
6.020000 0.570000 -6.260000 3.493596 -3.525321 -4.344670
6.020000 1.440000 -4.940000 -1.268812 7.358745 5.312328
6.030000 1.400000 -6.640000 -0.995089 8.623967 1.814714
6.060000 -1.730000 -7.900000 5.574838 0.245764 0.463545
6.130000 -1.300000 -6.810000 4.986794 8.583773 -3.595565
6.130000 -2.670000 -6.520000 4.532879 -16.382965 2.617609
6.130000 1.480000 -4.350000 -2.631996 4.853879 8.800497
6.140000 -0.570000 -7.490000 -0.952977 3.764521 -0.457009
6.140000 0.770000 -5.300000 0.380271 -6.085596 3.282303
6.190000 1.100000 -5.900000 -0.359627 3.592734 -4.143387
6.190000 1.250000 -8.110000 -7.029285 10.120116 -0.966697
```

```
6.230000 -1.340000 -5.050000 2.408447 -0.495421 12.519599
6.240000 -0.470000 -6.200000 -0.502339 2.151677 -3.715754
6.330000 -0.430000 -5.080000 0.280964 -0.235228 3.233716
6.330000 0.850000 -4.820000 -1.218289 -5.465954 7.861419
6.360000 -1.410000 -6.090000 3.669792 4.471261 -1.351240
6.360000 -1.980000 -5.870000 4.831958 -7.806591 2.895534
6.380000 0.900000 -7.150000 -1.442137 -0.576524 -5.142725
6.420000 -0.560000 -7.030000 3.069367 -2.227903 -3.250582
6.420000 0.630000 -5.900000 -1.046445 -5.798025 -2.383686
6.420000 1.570000 -5.680000 -2.960701 9.172733 -4.174515
6.450000 1.180000 -3.970000 -3.549352 -3.004582 10.652515
6.480000 -1.990000 -6.500000 4.043524 -2.847972 -1.121834
6.530000 0.710000 -6.570000 -0.588745 -4.027726 -5.542502
6.570000 -0.200000 -5.800000 1.504181 -6.073162 -1.061145
6.610000 -1.600000 -5.650000 5.287634 -1.337155 2.173533
6.620000 1.540000 -4.570000 -2.859815 3.541240 5.403079
6.630000 -1.090000 -6.780000 3.664742 8.018772 -3.706543
6.630000 1.710000 -5.240000 -3.770365 9.505173 0.621020
6.670000 1.100000 -5.110000 -1.049995 0.135023 3.169182
6.690000 0.520000 -5.060000 -0.075904 -10.513305 5.914184
6.810000 1.180000 -5.990000 -0.662750 2.993023 -5.205589
6.850000 0.840000 -4.420000 -0.099106 -7.133138 11.288634
6.850000 1.360000 -6.510000 -0.596096 6.408150 -6.173361
6.910000 1.340000 -5.520000 -0.534212 4.565864 -1.636322
6.910000 2.070000 -5.800000 -1.432949 9.057551 -5.034982
6.920000 2.570000 -5.580000 -4.829117 8.977786 -4.277987
7.010000 1.260000 -4.150000 0.055078 -1.672152 8.924435
7.050000 1.850000 -4.010000 0.046123 4.027859 7.862283
7.080000 -1.850000 -5.730000 3.822482 -7.600072 3.217529
7.100000 -0.730000 -7.070000 6.218815 8.665473 -2.926175
7.110000 -0.730000 -6.270000 3.795763 5.222723 -4.287915
7.170000 0.680000 -5.860000 1.859171 -6.265047 -2.351190
7.190000 0.350000 -5.350000 1.971966 -11.064725 2.725801
7.210000 1.670000 -7.190000 2.547221 9.167685 2.324850
7.320000 2.070000 -4.630000 2.390077 5.542002 4.430370
7.370000 1.450000 -5.200000 2.721028 5.053285 2.840806
7.390000 1.190000 -4.620000 3.878594 -0.887461 8.373271
7.440000 1.540000 -5.720000 3.241314 7.036184 -3.927956
7.450000 0.150000 -6.820000 5.234042 -11.940313 -5.242416
7.510000 1.390000 -6.380000 3.648288 5.388274 -6.766823
7.510000 2.470000 -5.500000 2.035588 1.841560 -2.741476
7.580000 1.020000 -6.000000 4.524123 -0.519741 -4.711188
7.630000 -1.830000 -5.690000 5.568798 -3.779837 4.768283
7.710000 0.890000 -4.910000 2.770515 -4.067473 7.296069
7.770000 2.660000 -6.110000 0.466204 -0.316426 -4.973325
7.810000 0.970000 -5.550000 1.861250 0.435538 -0.650927
7.820000 3.010000 -4.740000 -0.869983 2.620221 4.012755
7.850000 1.140000 -4.430000 2.293750 -1.102156 9.739676
7.870000 3.100000 -5.770000 2.315459 0.595940 -5.309249
7.930000 1.550000 -5.060000 0.771882 5.946854 5.398545
8.020000 2.270000 -4.920000 1.866762 3.743695 2.417959
8.230000 3.660000 -4.830000 -0.685348 7.291451 3.529174
8.320000 0.960000 -5.610000 0.484645 -2.857555 -1.676067
8.370000 0.620000 -6.460000 -0.238463 -5.273331 -5.725803
```

```
8.430000 1.950000 -5.410000 -1.539673 8.510531 -1.092240
8.510000 1.030000 -4.710000 -0.949979 -3.224230 11.938912
8.690000 1.470000 -6.100000 -0.672348 6.303684 -6.792371
8.850000 2.640000 -6.260000 -2.609535 6.346974 -8.074174
8.940000 0.350000 -4.760000 -1.009582 -12.140244 11.065896
8.970000 0.760000 -5.760000 -1.365027 -4.264435 -2.156946
8.980000 1.380000 -4.980000 -0.956414 2.911630 6.951530
8.990000 0.540000 -5.260000 -2.014932 -9.671615 4.194186
9.070000 0.010000 -5.560000 -2.203293 -13.727834 2.322717
9.100000 1.170000 -6.610000 -1.846779 3.245957 -6.790586
9.100000 1.270000 -5.700000 -1.061001 2.991233 -2.734217
9.120000 2.160000 -5.350000 -1.493177 7.751459 -1.027296
9.160000 1.780000 -6.050000 -1.327887 9.236950 -5.988685
9.180000 0.960000 -7.770000 -1.268384 1.385118 0.208353
9.220000 0.430000 -6.300000 -2.510071 -4.167695 -4.276783
9.430000 0.820000 -4.780000 -1.322281 -7.393560 11.321846
9.440000 0.390000 -7.060000 -1.679670 -4.249693 -6.689306
9.490000 1.170000 -6.290000 -2.293362 3.027558 -6.488503
9.510000 1.730000 -5.640000 -1.915270 6.374895 -3.513869
9.620000 1.010000 -5.320000 -1.155763 -1.444269 2.390249
9.730000 0.470000 -5.870000 -0.330852 -8.231715 -2.128090
9.840000 1.480000 -5.050000 -0.266062 3.405301 5.359532
9.880000 1.720000 -6.640000 0.204631 7.355597 -7.061658
9.920000 2.300000 -6.770000 -1.450918 7.423427 -7.131328
9.940000 3.490000 -5.280000 -2.341432 2.985128 0.767734
9.970000 4.160000 -5.750000 -2.319189 0.357524 -2.823965
```

The following output figures are generated from the script.

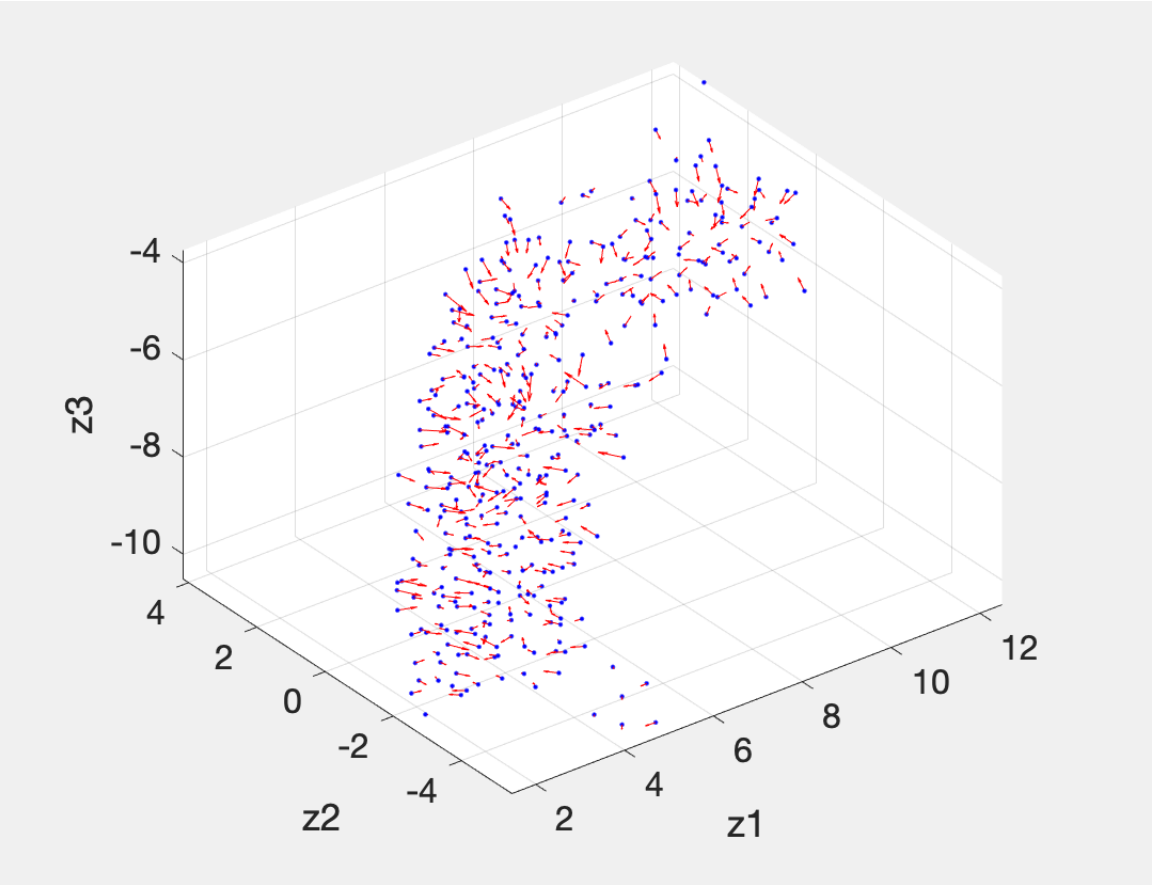



**Figure A.1.** The locations of centers are shown as blue points, and the mean-forces at each point are shown as red arrow.

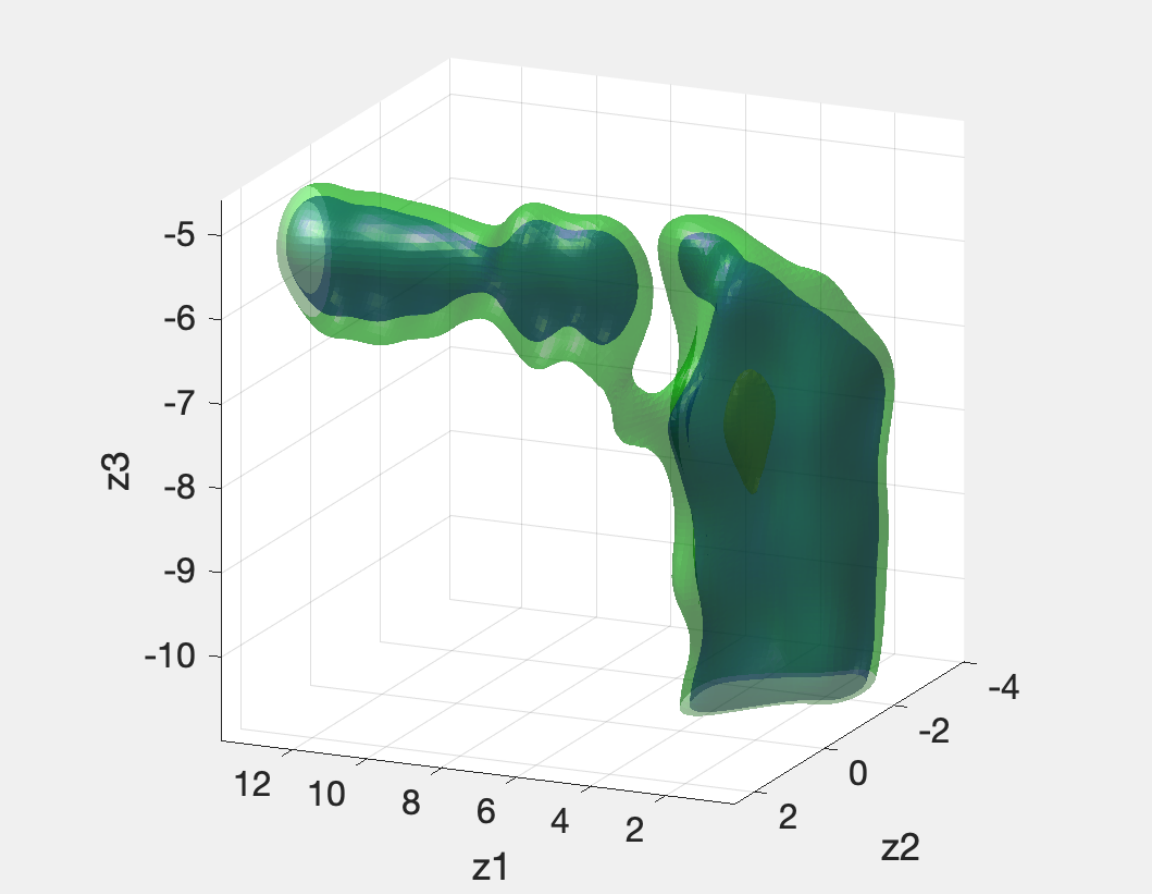



**Figure A.2.** The reconstructed 3D free energy surfaces are shown. The lower isosurfaces with lower energy values are shown with darker colors.

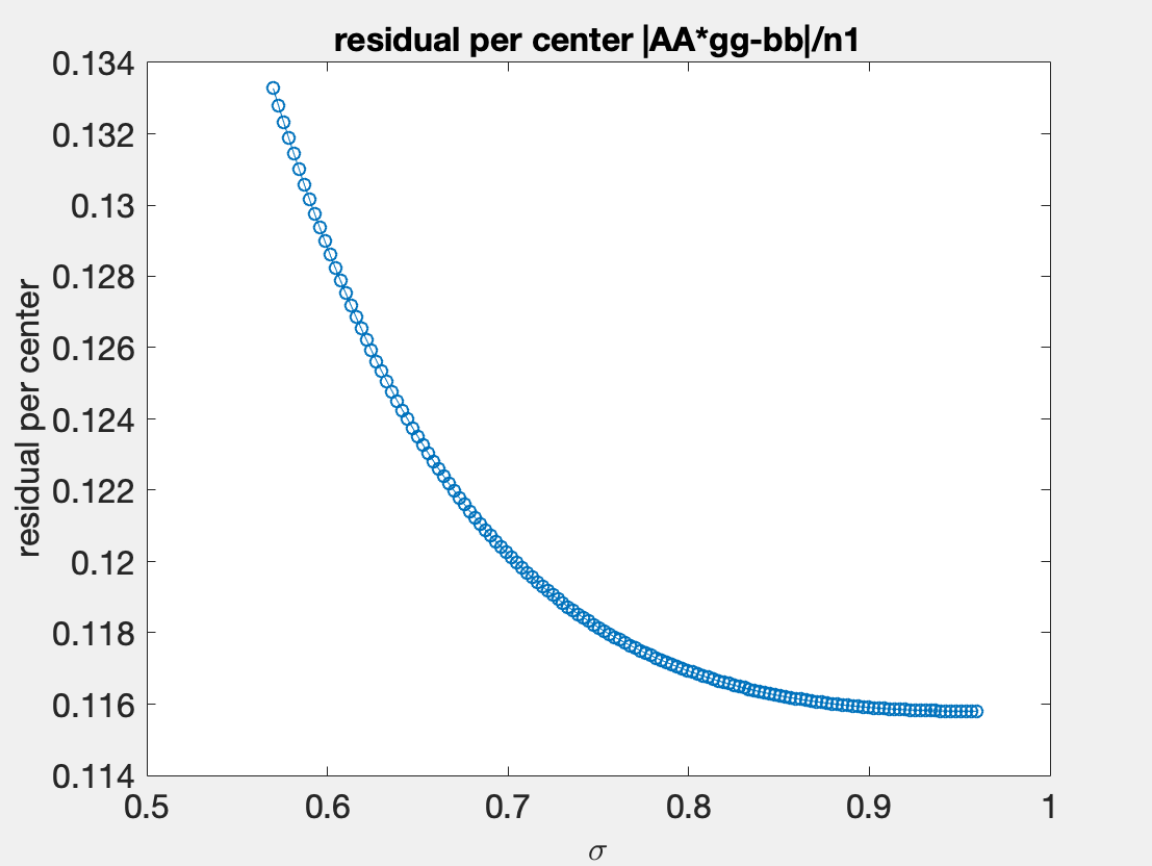



**Figure A.3.** The optimization profile of $\sigma$ is shown.

## A.2 Sample code for zero temperature string method (ZTS)

The minimum free-energy pathways (MFEP) can be determined using the ZTS method [149] (see section 2.4.1.2).

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Zero-temperature string method (ZTS) code                           %
% The script is only tested for string                                 %
% optimization, and nothing else.                                      %
% (c) Eric Vanden-Eijnden, NYU, New York, NY                           %
% Modified (Lines 10-30) and compiled by M. Mohammadi                  %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
clear all
set(0,'DefaultTextFontName','TimesRoman')
set(0,'DefaultAxesFontSize',16)
% max number of iterations
nstepmax = 2e3;
% frequency of plotting
nstepplot = 1e1;
% plot string every nstepplot if flag1 = 1
flag1 = 1;
% parameter used as stopping criterion
tol1 = 1e-7;
% number of images along the string (try from  n1 = 3 up to n1 = 1e4)
n1 = 25;
% time-step (limited by the ODE step on line 83 & 84 but independent of n1)
h = 1e-4;
% end points of the initial string
% notice that they do NOT have to be at minima of V -- the method finds
% those automatically
xa = -1;
ya = 0.5;
xb = 0.7;
yb = 0.5;
% initialization
g1 = linspace(0,1,n1);
x = (xb-xa)*g1+xa;
y = (x-xa)*(yb-ya)/(xb-xa)+ya;
dx = x-circshift(x,[0 1]);
dy = y-circshift(y,[0 1]);
dx(1) = 0;
dy(1) = 0;
lxy = cumsum(sqrt(dx.^2+dy.^2));
lxy = lxy/lxy(n1);
x = interp1(lxy,x,g1);
y = interp1(lxy,y,g1);
xi = x;
yi = y;
% parameters in Mueller potential
aa = [-1 -1 -6.5 0.7];
bb = [0 0 11 0.6];
```

```
cc = [-10 -10 -6.5 0.7];
AA = [-200 -100 -170 15];
XX = [1 0 -0.5 -1];
YY = [0 0.5 1.5 1];
[xx,yy] = meshgrid(-1.5:0.01:1.2,-0.2:0.01:2);
V1 = AA(1)*exp(aa(1)*(xx-XX(1)).^2+bb(1)*(xx-XX(1)).*(yy-YY(1))+cc(1)*(yy-YY
    (1)).^2);
for j=2:4
    V1 =  V1 + AA(j)*exp(aa(j)*(xx-XX(j)).^2+bb(j)*(xx-XX(j)).*(yy-YY(j))+cc(j
        )*(yy-YY(j)).^2);
end;
figure(1);clf;
contourf(xx,yy,min(V1,200),40);
hold on
plot(xi,yi,'.-w','MarkerSize',14)
set(gca,'XTick',-1.5:.5:1,'YTick',0:.5:2);
xlabel('x','FontAngle','italic');
ylabel('y','FontAngle','italic');
title('Initial string');
drawnow
figure(2);clf;
contourf(xx,yy,min(V1,200),40);
whitedots = line(x,y,'linestyle','none','marker','.','color','w','MarkerSize'
    ,14);
set(gca,'XTick',-1.5:.5:1,'YTick',0:.5:2);
xlabel('x','FontAngle','italic');
ylabel('y','FontAngle','italic');
title('String evolution')
drawnow
%% Main loop
for nstep=1:nstepmax
% calculation of the x and y-components of the force, dVx and dVy respectively
    ee = AA(1)*exp(aa(1)*(x-XX(1)).^2+bb(1)*(x-XX(1)).*(y-YY(1))+cc(1)*(y-YY
        (1)).^2);
    dVx = (2*aa(1)*(x-XX(1))+bb(1)*(y-YY(1))).*ee;
    dVy = (bb(1)*(x-XX(1))+2*cc(1)*(y-YY(1))).*ee;
    for j=2:4
        ee = AA(j)*exp(aa(j)*(x-XX(j)).^2+bb(j)*(x-XX(j)).*(y-YY(j))+cc(j)*(y-
            YY(j)).^2);
        dVx = dVx + (2*aa(j)*(x-XX(j))+bb(j)*(y-YY(j))).*ee;
        dVy = dVy + (bb(j)*(x-XX(j))+2*cc(j)*(y-YY(j))).*ee;
    end;
    x0 = x;
    y0 = y;
% string steps:
% 1. evolve
    x = x - h*dVx;
    y = y - h*dVy;
% 2. reparametrize
    dx = x-circshift(x,[0 1]);
    dy = y-circshift(y,[0 1]);
    dx(1) = 0;
    dy(1) = 0;
    lxy = cumsum(sqrt(dx.^2+dy.^2));
```

```
    lxy = lxy/lxy(n1);
    x = interp1(lxy,x,g1);
    y = interp1(lxy,y,g1);
    if and(flag1 == 1,mod(nstep,nstepplot) == 0)
        set(whitedots,'xdata',x,'ydata',y)
        drawnow
        pause(0.025)
    end
    tol = (norm(x-x0)+norm(y-y0))/n1;
    if tol <= tol1; break; end;
end
%% Output
fprintf('\n')
    fprintf('\n')
    fprintf('ZTS calculation with %d images\n',n1)
if tol > tol1
    fprintf('The calculation failed to converge after %d iterations\n',nstep)
else
    fprintf('The calculation terminated after %d iterations\n',nstep)
end;
figure(2);clf;
contourf(xx,yy,min(V1,200),40);
hold on
plot(x,y,'.-w','MarkerSize',14)
set(gca,'XTick',-1.5:.5:1,'YTick',0:.5:2);
xlabel('x','FontAngle','italic');
ylabel('y','FontAngle','italic');
title('Final string');
drawnow
% Energy along MEP
tx = circshift(x,[0 -1])-circshift(x,[0 1]);
ty = circshift(y,[0 -1])-circshift(y,[0 1]);
% potential computed as integral of projection of gradV on string tangent
Vz=cumtrapz(tx.*dVx + ty.*dVy);
Vz=0.5*Vz;
Vz=Vz-min(Vz);
ntxy = sqrt(tx.*tx+ty.*ty);
tx = tx./ntxy;
ty = ty./ntxy;
% err is an estimate of the error between disrectized MEP and actual one
% err scale as 1/n1
err = trapz(1-(tx.*dVx+ty.*dVy).^2./(dVx.*dVx+dVy.*dVy))/(n1-1);
fprintf('Estimate of difference between discretized MEP and actual one: %f\n',
    err)
%
% Reinterpolate string with lots of points to get accurate energy along it
g0 = linspace(0,1,1e3);
x0 = interp1(lxy,x,g0);
y0 = interp1(lxy,y,g0);
ee = AA(1)*exp(aa(1)*(x0-XX(1)).^2+bb(1)*(x0-XX(1)).*(y0-YY(1))+cc(1)*(y0-YY
    (1)).^2);
V0=ee;
dVx = (2*aa(1)*(x0-XX(1))+bb(1)*(y0-YY(1))).*ee;
dVy = (bb(1)*(x0-XX(1))+2*cc(1)*(y0-YY(1))).*ee;
```

```
for j=2:4
    ee = AA(j)*exp(aa(j)*(x0-XX(j)).^2+bb(j)*(x0-XX(j)).*(y0-YY(j))+cc(j)*(y0-
        YY(j)).^2);
    V0 = V0 + ee;
    dVx = dVx + (2*aa(j)*(x0-XX(j))+bb(j)*(y0-YY(j))).*ee;
    dVy = dVy + (bb(j)*(x0-XX(j))+2*cc(j)*(y0-YY(j))).*ee;
end;
V0=V0-min(V0);
figure(3);clf;
hold on
title('Energy along MEP');
plot(g1,Vz)
plot(g0,V0,'r')
box on
legend('Thermodynamic integration along string','Exact','Location','South')
legend('boxoff')
```

The following uotpot figures are generated from the zts script.

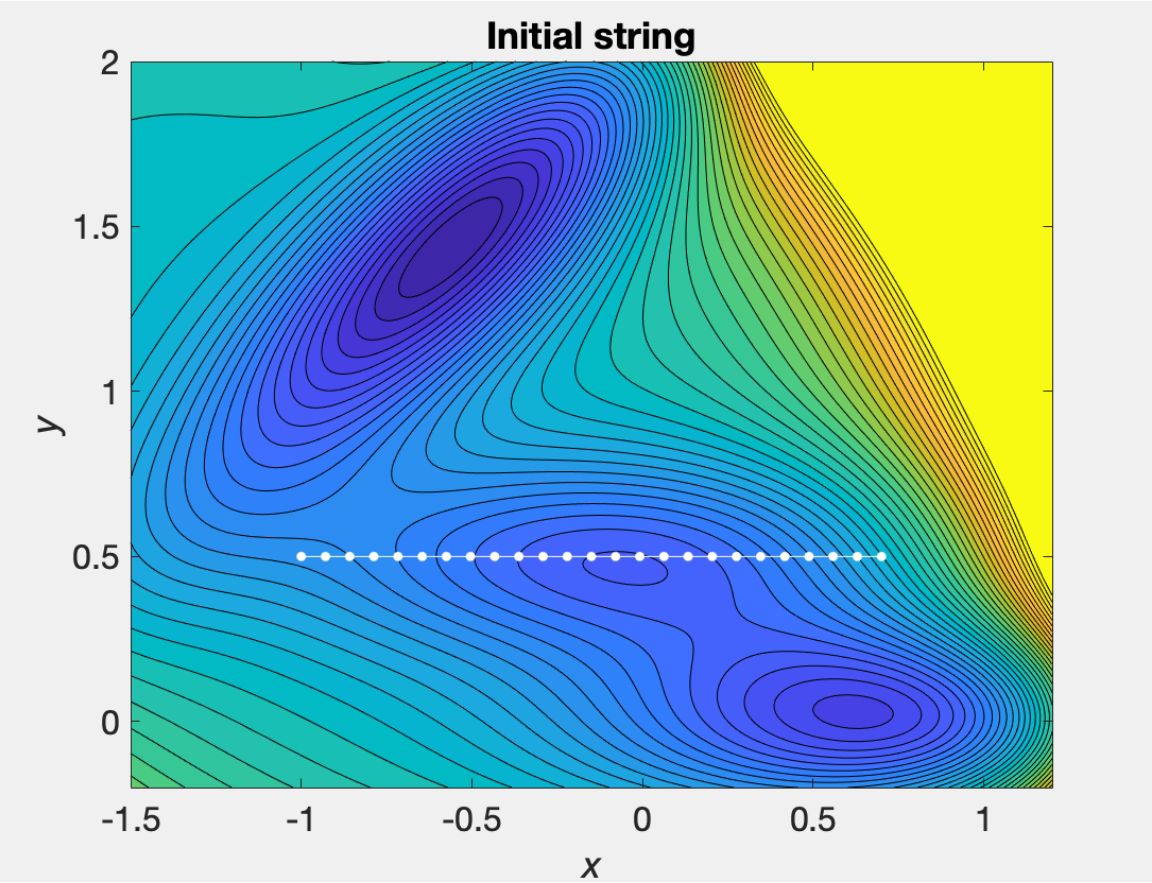



**Figure A.4.** The initial string before initiation of the script.

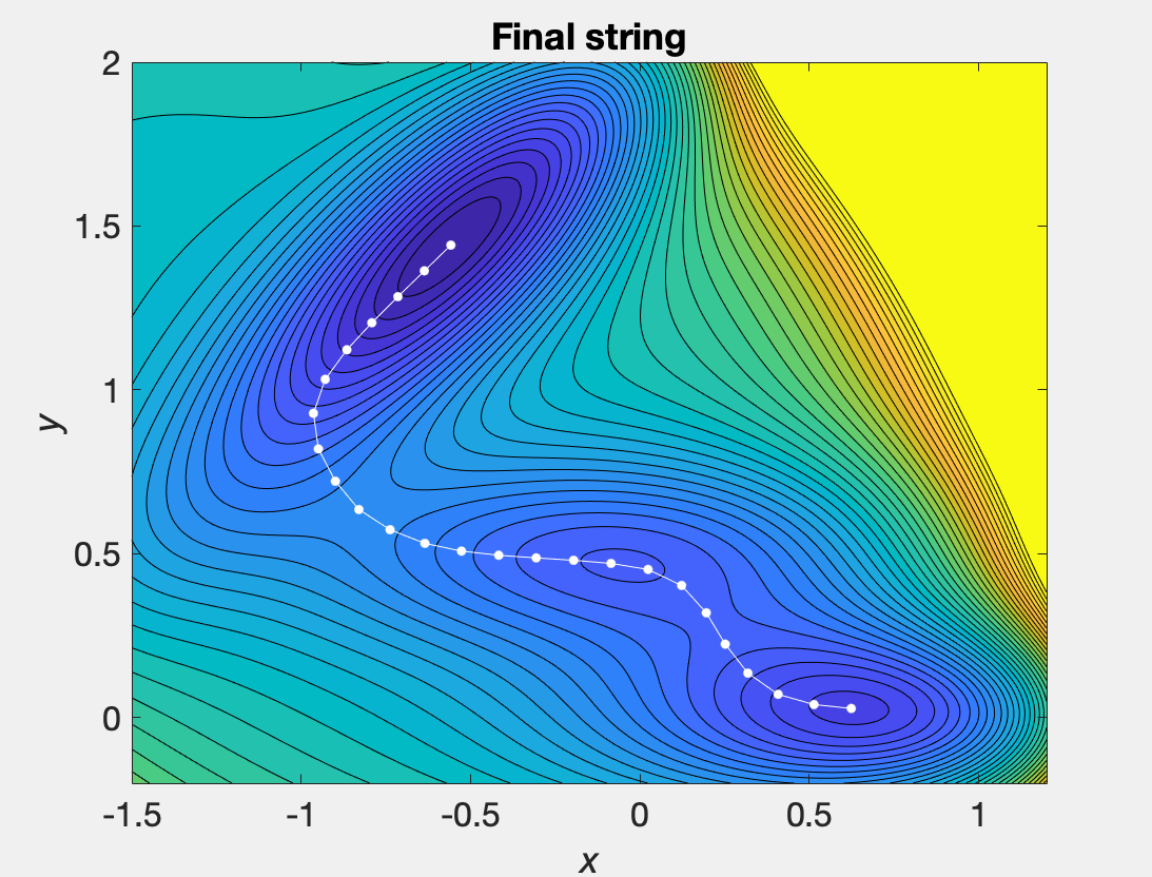



**Figure A.5.** The final string after completion of the script.

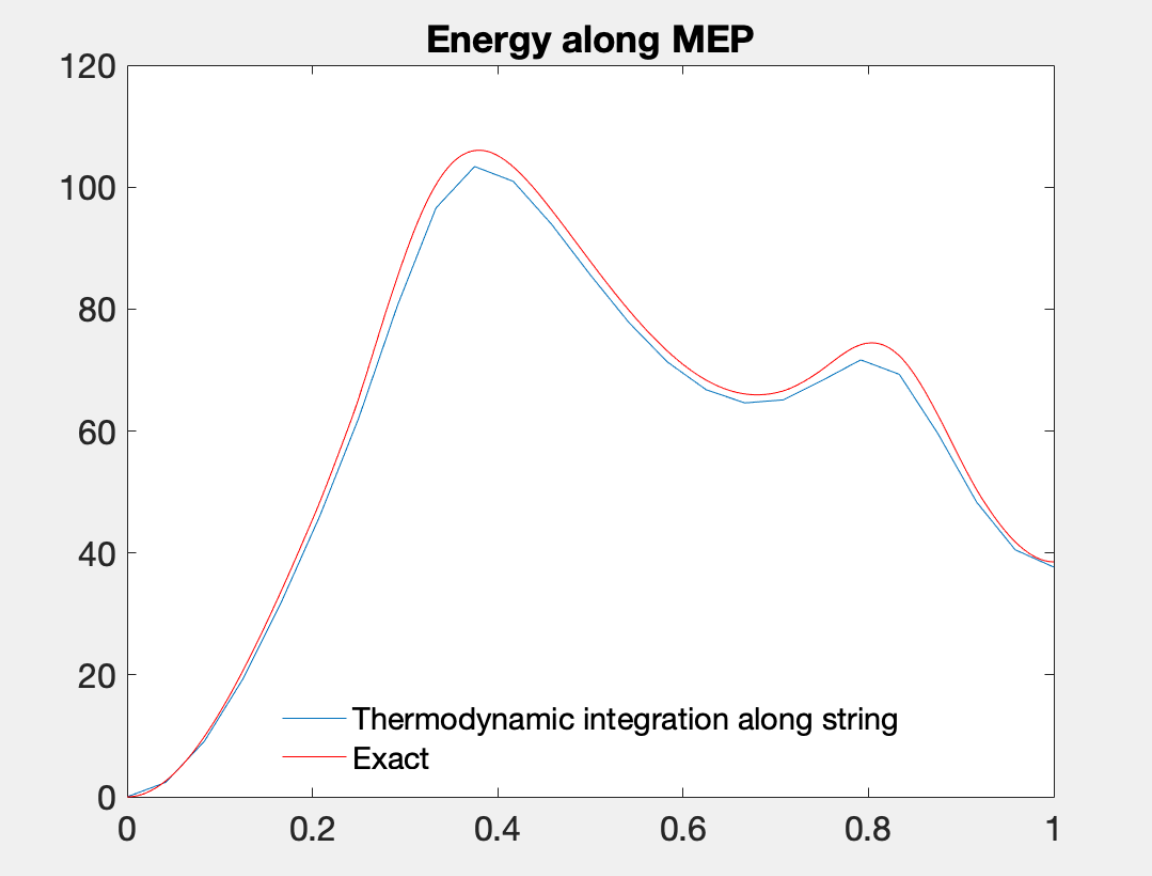



**Figure A.6.** The 2D PMF, showing the free-enrgy change acrosss the final string.

# APPENDIX B

# TCL SCRIPTS

## B.1 Sample Code for selecting centers

The exploration of the protein interior space via TAMD results in a dense cloud of ligand coordinates (section 2.4.1.1). I used the following script for selecting the centers within the protein matrix for mean-force calculation.

```
1  ##################################################################
2  ## Point Clustering (Version 3.0)                                ##
3  ## This script has been only tested for selecting positions of ##
4  ## ligand covering various areas of the protein by a distance  ##
5  ## cut-off value. These centers could be used for mean-force   ##
6  ## calculation described in chapter 3.                           ##
7  ##################################################################
8
9  set outfile [open ox.dat w];
10 set nf [molinfo top get numframes]
11 set sel [atomselect top "segname OX"]
12 set com_list {}
13 set refcom {}
14 set newcom {}
15 set com_list2 {}
16 set chosen_frames {}
17 proc lremove {list match} {
18     set idx_list [lsearch -all $list $match]
19     foreach idx [lreverse $idx_list] {
20         set list [lreplace $list $idx $idx]
21     }
22     return $list
23 }
24 for {set i 0 } { $i < $nf } { incr i } {
25     set temp_ox {}
26  $sel frame $i
27     set com_ox [measure center $sel weight mass]
28  lappend temp_ox $com_ox
29  lappend com_list $temp_ox
30 }
31 set com_list2 $com_list
32 set size [llength $com_list]
33 set uni {}
34 set k 1
35 set g 1
```

```
puts "Initial total COM: $size "
set refcom0 {}
set newcom5 {}
set new2 {}
 set refcom0 [lindex $com_list 0]
 lappend uni $refcom0
 set ref [lindex $refcom0 0]
 for {set i 1} { $i < [llength $com_list] } { incr i} {
    set newcom5 [lindex $com_list $i]
    set new5 [lindex $newcom5 0]
       set dd [vecdist $new5 $ref]
     set g 1
     if { $dd >= 2.5 } {
    for {set j 0 } { $j < [llength $uni] } { incr j } {
    set newcom [lindex $uni $j]
    set s1 [lsearch -all $com_list $newcom]
    set new2 [lindex $com_list $s1]
       set new3 [lindex $new2 0]
   set dddd [vecdist $new5 $new3]
    if {$dddd > 0 && $dddd < 3.0 } {
    set g 0 }
      }
   if {$g == 1 } {
      lappend uni $newcom5
      set refcom0 $newcom5
      set ref [lindex $refcom0 0]
   }
 }
 }
set refcom1 {}
set newcom1 {}
#  for {set r 0 } { $r < [llength $uni] } { incr r } {
#  set refcom [lindex $uni $r]
#  set ref [lindex $refcom 0]
#  for {set k 0} { $k < [llength $uni] } { incr k } {
#  set newcom [lindex $uni $k]
#  set new [lindex $newcom 0]
#  set ddd [vecdist $new $ref]
#  if {$ddd < 2.51 && $ddd > 0} {
#  puts "$r $k $ddd \ " }
#  }
# }
 #set chosen_frames {}
 draw delete all
 for {set i 0 } { $i < [llength $uni] } { incr i } {
 set refcom1 [lindex $uni $i]
 set s [lsearch -exact $com_list2 $refcom1]
 lappend chosen_frames $s
 puts $outfile "$s $refcom1 \ "
 set ref [lindex $refcom1 0]
# puts "$ref"
 draw color red
 draw sphere $ref radius 0.5
# animate write pdb ${s}.pdb beg $s end $s
```

```
 }
 set size [llength $uni]
 puts "Final total COM: $size "
close $outfile
puts "frames chosen are: $chosen_frames"
```

The following typical output data file shows how coordinates of the centers are shown in "ox.dat".

```
     0     7.1351    85.4152    76.9845
   156     7.0427    82.6447    75.2778
   191    4.10029    81.7408     76.387
   206    1.98574    82.4221    74.2702
   211  -0.356568    84.6737    74.5849
   247    2.60047    84.1852    71.0736
   290    5.08997    82.5053    72.0253
   338    3.94729    86.4296    74.0891
   733     7.2174    85.6596    73.1914
  2580    5.02837    87.7864     76.841
  2728  -0.823304    81.2919    75.0923
  2762   -1.12958    82.0141    71.8188
  5280   -2.86603    82.4115    69.3282
  5372    3.35803    79.5987     72.587
  5414    2.97107    79.2217    77.5833
  6315    7.91289    82.0163    71.0953
  6363    10.2266    84.9011    76.4793
  7211   -3.16977    85.1162    73.4462
  7358    11.4153    85.3845    71.0905
  7455    9.24379     87.798    71.7588
  7903    8.47505    88.5709    77.1181
  7909    6.31513    88.2072    79.6226
  9779  -0.598473    83.0012    67.4441
 10410     9.0344    85.1746    79.3655
 12211   -5.08412    84.4729    70.8897
 12220   -3.59797    87.3056     70.978
 12290  -0.088821    86.7477    72.3108
 12298   -3.06006    89.8068    72.6884
 12341   -3.01366    89.6928    75.7531
 12480   -5.77485    88.7539    68.2326
 12898    12.6081    85.8561    78.2051
 13693    1.18522    78.7637     75.204
 14215  -0.810877    87.7092    75.1475
 14429   -6.34716    88.7802    73.7231
 14447   -8.36029    89.9232    70.7358
 14480    -10.123    92.6985    69.7389
 14486   -7.82316    91.0486    67.4253
 14912    7.04897    82.8316    79.1575
 15000   -2.01535    78.1745    82.9957
 15002    0.43548    77.3959    81.3736
 15071   -1.03908    76.8404    78.4965
 15157    2.13227    74.1964    82.2809
 15195   -0.71501    73.5197    83.2207
 15225   -4.70946    77.7479    84.7837
 15326   -1.78592    78.8838    86.0158
```

| | | | | |
|---|---|---|---|---|
| 46 | 15343 | 0.56893 | 77.6384 | 88.2306 |
| 47 | 15362 | 3.55881 | 77.3868 | 88.7032 |
| 48 | 15709 | -3.18309 | 78.6372 | 79.9735 |
| 49 | 15839 | 1.5584 | 78.7178 | 84.0627 |
| 50 | 15860 | 2.89458 | 76.0987 | 86.0084 |
| 51 | 15993 | -0.0992651 | 75.958 | 85.5896 |
| 52 | 16422 | 3.42125 | 78.6341 | 81.5722 |
| 53 | 16782 | -1.61375 | 76.4536 | 74.5981 |
| 54 | 16789 | 1.50726 | 75.1149 | 73.5906 |
| 55 | 17207 | -5.92617 | 79.5784 | 79.0649 |
| 56 | 17214 | -7.95139 | 80.7052 | 80.9874 |
| 57 | 17268 | -8.08941 | 77.5371 | 77.6604 |
| 58 | 17340 | -5.80022 | 80.3998 | 83.165 |
| 59 | 17725 | -6.94952 | 77.566 | 82.4301 |
| 60 | 19413 | -4.72763 | 78.4518 | 88.2021 |
| 61 | 19475 | -4.35415 | 80.6275 | 85.8048 |
| 62 | 19979 | -6.52378 | 83.3175 | 83.6854 |
| 63 | 20307 | -5.43351 | 82.4185 | 80.2318 |
| 64 | 20409 | -7.71119 | 80.8883 | 85.9511 |
| 65 | 20695 | -3.38147 | 73.8707 | 84.8278 |
| 66 | 20705 | -4.76022 | 75.2638 | 82.0215 |
| 67 | 20914 | 0.252021 | 77.2081 | 91.2985 |
| 68 | 21256 | -5.5409 | 76.0863 | 78.6958 |
| 69 | 21746 | -0.272053 | 78.0284 | 71.8256 |
| 70 | 21756 | -1.79839 | 75.0986 | 71.1601 |
| 71 | 21776 | -0.0915667 | 77.1357 | 68.8906 |
| 72 | 21790 | 2.99189 | 75.6838 | 70.3448 |
| 73 | 24403 | 0.986457 | 80.5307 | 87.4241 |
| 74 | 25209 | -9.63618 | 76.6114 | 80.1423 |
| 75 | 25221 | -9.37912 | 79.2163 | 83.6128 |
| 76 | 25227 | -9.65463 | 75.566 | 84.4429 |
| 77 | 25708 | -9.59972 | 80.313 | 77.5849 |
| 78 | 25711 | -12.2069 | 81.7913 | 78.0756 |
| 79 | 25718 | -10.143 | 84.0751 | 78.0835 |
| 80 | 25762 | -13.0779 | 82.0587 | 74.9921 |
| 81 | 25779 | -12.054 | 84.5249 | 73.4851 |
| 82 | 25788 | -9.05334 | 83.682 | 73.6551 |
| 83 | 25807 | -7.94901 | 85.7751 | 76.29 |
| 84 | 25828 | -7.8868 | 85.6192 | 79.4377 |
| 85 | 25841 | -7.32407 | 87.5269 | 82.0868 |
| 86 | 25850 | -4.27308 | 86.5287 | 81.0078 |
| 87 | 25868 | -3.79369 | 84.1495 | 78.3871 |
| 88 | 25898 | -1.34692 | 84.6306 | 80.4013 |
| 89 | 26697 | -5.72582 | 72.2962 | 83.6206 |
| 90 | 27184 | -1.77601 | 70.1802 | 82.7375 |
| 91 | 27280 | -10.2919 | 78.37 | 75.3051 |
| 92 | 27284 | -8.37523 | 75.7658 | 75.0296 |
| 93 | 27292 | -7.24866 | 73.1177 | 76.4577 |
| 94 | 27673 | -2.45819 | 74.3541 | 77.3623 |
| 95 | 27725 | -11.6236 | 75.728 | 76.2037 |
| 96 | 27746 | -8.19481 | 76.1339 | 71.4791 |
| 97 | 27756 | -9.86193 | 73.1767 | 70.6255 |
| 98 | 27775 | -9.97434 | 74.6542 | 67.6926 |
| 99 | 27783 | -8.19527 | 73.6674 | 64.7872 |

| | | | | |
|---|---|---|---|---|
| 100 | 27790 | -5.55023 | 72.5659 | 66.2084 |
| 101 | 27807 | -6.28999 | 75.873 | 66.0208 |
| 102 | 27842 | -5.88627 | 78.481 | 67.5104 |
| 103 | 27860 | -3.21747 | 75.8665 | 68.5378 |
| 104 | 28443 | 0.903283 | 74.8342 | 89.3632 |
| 105 | 29289 | -4.8102 | 74.0598 | 74.455 |
| 106 | 29368 | 0.854802 | 73.6352 | 70.882 |
| 107 | 29711 | -12.2702 | 77.3339 | 82.1299 |
| 108 | 29915 | -1.24518 | 79.9476 | 90.7084 |
| 109 | 30000 | 14.9703 | 93.7883 | 94.715 |
| 110 | 30001 | 13.7146 | 90.0929 | 93.2049 |
| 111 | 30094 | 12.8017 | 89.2761 | 90.1279 |
| 112 | 30173 | 11.2603 | 87.7012 | 88.0044 |
| 113 | 30210 | 8.00348 | 88.852 | 88.0841 |
| 114 | 30484 | 15.8416 | 90.9543 | 91.1764 |
| 115 | 30727 | 8.20999 | 85.9679 | 89.5407 |
| 116 | 31210 | 10.3387 | 90.5778 | 91.4926 |
| 117 | 33805 | 18.1487 | 93.1588 | 90.0519 |
| 118 | 33925 | 16.4352 | 91.4383 | 87.9478 |
| 119 | 34363 | 15.0601 | 87.0594 | 89.7112 |
| 120 | 34412 | 10.7545 | 86.6302 | 91.5772 |
| 121 | 35210 | 5.08066 | 90.0703 | 88.5942 |
| 122 | 35228 | 5.18456 | 87.7336 | 90.5426 |
| 123 | 35341 | 5.45502 | 90.6981 | 91.5825 |
| 124 | 35396 | 11.006 | 86.8117 | 94.9752 |
| 125 | 35430 | 10.7992 | 83.8352 | 94.1879 |
| 126 | 35746 | 13.8828 | 92.6993 | 89.1085 |
| 127 | 36407 | 13.4493 | 87.3994 | 96.8475 |
| 128 | 36478 | 10.896 | 89.8585 | 94.5575 |
| 129 | 36846 | 7.85948 | 87.7463 | 92.0737 |
| 130 | 37360 | 16.2352 | 90.8196 | 94.6945 |
| 131 | 38409 | 5.73247 | 88.3188 | 95.3002 |
| 132 | 38764 | 9.00363 | 88.541 | 85.083 |
| 133 | 39764 | 5.5124 | 92.4056 | 86.5484 |
| 134 | 39831 | 4.29214 | 93.023 | 89.9503 |
| 135 | 40256 | 4.35224 | 89.4817 | 85.5505 |
| 136 | 40264 | 3.62183 | 92.2582 | 83.6031 |
| 137 | 40335 | 3.44229 | 94.5511 | 87.4491 |
| 138 | 41144 | 17.4094 | 88.5241 | 88.3844 |
| 139 | 41281 | 5.85185 | 94.4184 | 84.1123 |
| 140 | 41478 | 3.05991 | 95.719 | 83.5125 |
| 141 | 41784 | 6.77811 | 92.354 | 82.1059 |
| 142 | 41841 | 5.2955 | 97.198 | 86.818 |
| 143 | 41861 | 8.15189 | 96.1233 | 87.371 |
| 144 | 41900 | 9.87323 | 98.124 | 89.372 |
| 145 | 41909 | 7.75953 | 96.6024 | 91.6622 |
| 146 | 41957 | 6.56776 | 98.4494 | 89.3902 |
| 147 | 41979 | 2.83732 | 98.934 | 88.2184 |
| 148 | 42363 | 11.3277 | 96.8024 | 86.7179 |
| 149 | 42822 | 3.7294 | 96.8739 | 90.24 |
| 150 | 42841 | 5.65124 | 98.6003 | 93.25 |
| 151 | 42863 | 9.12351 | 99.3475 | 92.9977 |
| 152 | 42909 | 5.87241 | 98.6047 | 96.8971 |
| 153 | 42930 | 3.9719 | 96.0604 | 96.4887 |

| | | | | |
|---|---|---|---|---|
| 154 | 42978 | 2.7325 | 99.0494 | 96.0453 |
| 155 | 43114 | 14.4279 | 88.3379 | 87.025 |
| 156 | 43484 | 4.96169 | 99.4478 | 84.0545 |
| 157 | 43498 | 4.03853 | 100.659 | 81.1839 |
| 158 | 44184 | 10.2113 | 84.003 | 87.8207 |
| 159 | 45000 | 14.74 | 84.0507 | 71.6804 |
| 160 | 45043 | 15.5001 | 80.96 | 73.3332 |
| 161 | 45192 | 12.5429 | 82.7204 | 73.2925 |
| 162 | 45715 | 9.96288 | 91.3777 | 71.2315 |
| 163 | 45746 | 11.4495 | 91.1158 | 67.7748 |
| 164 | 45763 | 11.745 | 89.9624 | 64.6253 |
| 165 | 45786 | 14.1821 | 91.7683 | 64.7346 |
| 166 | 45842 | 16.8772 | 93.6042 | 66.9184 |
| 167 | 45850 | 19.6844 | 92.9532 | 65.7277 |
| 168 | 45863 | 23.0417 | 92.0531 | 66.1996 |
| 169 | 45909 | 20.4208 | 92.3354 | 68.8185 |
| 170 | 45925 | 19.267 | 90.9382 | 62.6974 |
| 171 | 45986 | 19.4953 | 89.1783 | 59.7148 |
| 172 | 46144 | 17.5247 | 83.8643 | 72.8341 |
| 173 | 46219 | 11.9341 | 86.5122 | 74.0527 |
| 174 | 46340 | 11.7277 | 88.4249 | 76.5302 |
| 175 | 46746 | 12.1384 | 88.1819 | 69.2111 |
| 176 | 46903 | 16.7171 | 86.4872 | 71.1689 |
| 177 | 46909 | 15.2958 | 83.7865 | 75.7012 |
| 178 | 47838 | 14.3498 | 92.3488 | 68.7978 |
| 179 | 47854 | 17.0665 | 95.7922 | 69.8577 |
| 180 | 47902 | 23.1192 | 94.49 | 68.4883 |
| 181 | 48275 | 8.21782 | 89.1538 | 68.355 |
| 182 | 48328 | 10.3978 | 93.2381 | 73.7507 |
| 183 | 48342 | 12.5865 | 94.9603 | 75.1012 |
| 184 | 48379 | 10.2766 | 97.3114 | 75.1652 |
| 185 | 48398 | 13.2073 | 98.7562 | 73.1421 |
| 186 | 48426 | 10.3022 | 98.1995 | 71.8297 |
| 187 | 48479 | 9.54532 | 100.94 | 70.8073 |
| 188 | 49447 | 6.80723 | 92.3623 | 71.7377 |
| 189 | 49477 | 6.47318 | 95.6562 | 69.599 |
| 190 | 49490 | 6.53173 | 93.1849 | 67.7629 |
| 191 | 50379 | 8.48999 | 94.4875 | 75.961 |
| 192 | 50413 | 9.67346 | 96.4316 | 78.0274 |
| 193 | 51789 | 14.6155 | 88.7566 | 64.1894 |
| 194 | 51933 | 22.5554 | 89.5591 | 64.312 |
| 195 | 51976 | 20.7023 | 93.6561 | 62.5712 |
| 196 | 51982 | 20.8669 | 92.2675 | 59.6872 |
| 197 | 51991 | 24.056 | 91.9896 | 58.2814 |
| 198 | 52937 | 25.4646 | 89.6201 | 63.4884 |
| 199 | 52946 | 23.3845 | 89.8324 | 61.1029 |
| 200 | 53280 | 8.56636 | 90.1671 | 65.3873 |
| 201 | 53404 | 26.2739 | 94.7062 | 67.5974 |
| 202 | 53409 | 25.4519 | 93.1284 | 70.672 |
| 203 | 53418 | 27.5991 | 91.6226 | 68.9346 |
| 204 | 53425 | 27.3047 | 91.3423 | 65.4401 |
| 205 | 53436 | 30.092 | 90.6348 | 66.684 |
| 206 | 53479 | 24.8038 | 92.7822 | 63.0657 |
| 207 | 53486 | 26.4353 | 90.0741 | 60.2316 |

| | | | | |
|---|---|---|---|---|
| 208 | 53488 | 28.9856 | 91.4444 | 61.1333 |
| 209 | 53768 | 4.02202 | 93.1006 | 65.746 |
| 210 | 53780 | 5.22809 | 94.5092 | 62.8748 |
| 211 | 53786 | 8.54911 | 92.727 | 62.9972 |
| 212 | 53981 | 12.162 | 93.0452 | 62.5108 |
| 213 | 53990 | 15.1121 | 93.2763 | 61.6849 |
| 214 | 54898 | 15.766 | 95.4148 | 73.7298 |
| 215 | 55225 | 4.85793 | 90.9854 | 74.1641 |
| 216 | 55282 | 5.23304 | 91.5737 | 61.9299 |
| 217 | 55315 | 9.49886 | 91.8281 | 59.9016 |
| 218 | 55324 | 9.10584 | 94.7849 | 60.7624 |
| 219 | 55364 | 14.395 | 95.5462 | 59.7567 |
| 220 | 55368 | 13.4163 | 92.4716 | 58.5636 |
| 221 | 55416 | 14.0593 | 96.0322 | 62.7901 |
| 222 | 55436 | 17.6941 | 92.8711 | 59.3607 |
| 223 | 55455 | 14.9974 | 95.8191 | 56.6137 |
| 224 | 55479 | 13.0262 | 98.1597 | 55.0293 |
| 225 | 55484 | 15.4991 | 97.7228 | 52.1518 |
| 226 | 55901 | 30.4263 | 93.1498 | 68.4496 |
| 227 | 55909 | 29.3813 | 92.8719 | 72.0127 |
| 228 | 55928 | 29.0693 | 87.7593 | 67.7686 |
| 229 | 55936 | 32.1188 | 88.4597 | 67.9656 |
| 230 | 55961 | 28.8597 | 88.6311 | 63.7859 |
| 231 | 56284 | 13.5151 | 90.3631 | 61.86 |
| 232 | 56409 | 21.7956 | 92.8813 | 72.0858 |
| 233 | 56414 | 22.1254 | 89.876 | 72.2056 |
| 234 | 56436 | 26.3882 | 88.2905 | 73.2902 |
| 235 | 56814 | 14.383 | 89.6632 | 67.1647 |
| 236 | 57307 | 17.5176 | 95.0115 | 63.7838 |
| 237 | 57399 | 26.146 | 88.6418 | 66.6097 |
| 238 | 57417 | 28.2311 | 89.2433 | 70.6998 |
| 239 | 57428 | 29.1795 | 85.471 | 65.3813 |
| 240 | 57486 | 28.141 | 89.0276 | 57.6636 |
| 241 | 57491 | 31.7401 | 89.1265 | 57.1754 |
| 242 | 57998 | 28.0457 | 94.3463 | 59.5314 |
| 243 | 58268 | 6.65527 | 87.1762 | 65.82 |
| 244 | 58282 | 7.1354 | 87.3422 | 62.0094 |
| 245 | 58288 | 10.2792 | 87.2154 | 61.7247 |
| 246 | 58658 | 16.6952 | 78.4282 | 72.2514 |
| 247 | 58845 | 12.1338 | 91.8763 | 76.0824 |
| 248 | 58909 | 11.175 | 92.7272 | 79.4656 |
| 249 | 58978 | 7.36898 | 95.4803 | 73.0651 |
| 250 | 59398 | 17.319 | 96.574 | 60.7571 |
| 251 | 59403 | 15.088 | 98.6497 | 59.2278 |
| 252 | 59498 | 13.8043 | 99.6462 | 50.3247 |
| 253 | 59681 | 15.3074 | 75.6769 | 73.0649 |
| 254 | 59705 | 15.0263 | 77.2339 | 69.612 |
| 255 | 59709 | 10.3717 | 76.0132 | 69.1277 |
| 256 | 59719 | 11.2703 | 79.2119 | 70.0091 |
| 257 | 59746 | 12.6423 | 77.4116 | 65.856 |
| 258 | 59756 | 9.98813 | 75.282 | 63.834 |
| 259 | 59780 | 9.38158 | 77.8198 | 62.2754 |
| 260 | 59792 | 13.017 | 74.2403 | 63.6795 |
| 261 | 59860 | 14.4586 | 74.6581 | 66.5257 |

| | | | | |
|---|---|---|---|---|
| 262 | 59863 | 17.9848 | 75.7387 | 67.3083 |
| 263 | 59901 | 20.9957 | 76.6113 | 67.212 |
| 264 | 59909 | 19.9066 | 75.829 | 71.1131 |
| 265 | 59923 | 23.0008 | 74.4934 | 68.4037 |
| 266 | 59929 | 20.7535 | 72.0209 | 67.7305 |
| 267 | 59933 | 23.478 | 71.2819 | 69.3446 |
| 268 | 59982 | 21.2707 | 74.164 | 65.1748 |
| 269 | 59988 | 24.5697 | 74.6582 | 65.7611 |
| 270 | 60000 | 38.6844 | 93.357 | 58.0567 |
| 271 | 60051 | 40.5804 | 90.8549 | 59.4269 |
| 272 | 60206 | 36.4636 | 94.1357 | 55.9041 |
| 273 | 60224 | 35.7699 | 91.6259 | 57.8926 |
| 274 | 60335 | 33.3502 | 93.5191 | 59.5877 |
| 275 | 60842 | 41.1888 | 93.6283 | 60.8847 |
| 276 | 61361 | 42.8916 | 89.6518 | 61.5272 |
| 277 | 61898 | 43.9644 | 93.3944 | 62.9925 |
| 278 | 61916 | 42.7333 | 92.9752 | 66.0196 |
| 279 | 61998 | 44.1118 | 94.4425 | 60.1263 |
| 280 | 62185 | 39.4632 | 90.513 | 56.5748 |
| 281 | 62210 | 33.0591 | 93.834 | 56.5127 |
| 282 | 62275 | 32.1715 | 94.9969 | 53.2237 |
| 283 | 62331 | 31.9768 | 97.9989 | 52.5177 |
| 284 | 62342 | 34.2792 | 97.4988 | 54.697 |
| 285 | 63942 | 45.1992 | 88.1281 | 59.5719 |
| 286 | 65184 | 40.6073 | 87.745 | 58.7014 |
| 287 | 65380 | 41.4359 | 91.7454 | 63.3436 |
| 288 | 65411 | 43.791 | 89.9484 | 66.0217 |
| 289 | 65434 | 46.3555 | 88.0146 | 65.0996 |
| 290 | 65496 | 45.9073 | 89.4194 | 62.4605 |
| 291 | 65769 | 31.1146 | 95.9421 | 50.3899 |
| 292 | 65786 | 33.1677 | 96.7005 | 48.1053 |
| 293 | 65807 | 33.0543 | 99.5594 | 49.7603 |
| 294 | 65835 | 30.7121 | 101.205 | 52.7759 |
| 295 | 65862 | 34.3535 | 102.983 | 53.0802 |
| 296 | 65879 | 32.1505 | 105.11 | 54.4152 |
| 297 | 65904 | 35.169 | 106.307 | 52.191 |
| 298 | 65962 | 33.3459 | 103.347 | 50.108 |
| 299 | 65989 | 33.6168 | 105.797 | 47.9588 |
| 300 | 66296 | 29.2612 | 98.2719 | 48.6756 |
| 301 | 66479 | 31.196 | 106.126 | 51.2357 |
| 302 | 66863 | 43.7688 | 91.497 | 59.3224 |
| 303 | 67435 | 46.751 | 91.9163 | 60.9192 |
| 304 | 68226 | 30.2018 | 92.3172 | 56.7756 |
| 305 | 68289 | 33.6758 | 94.0451 | 46.724 |
| 306 | 69682 | 38.1888 | 85.831 | 58.3762 |
| 307 | 69695 | 36.6053 | 88.3853 | 58.8253 |
| 308 | 69726 | 33.9817 | 86.5053 | 58.563 |
| 309 | 69747 | 36.4678 | 89.267 | 55.9564 |
| 310 | 69756 | 34.3408 | 87.9183 | 53.8812 |
| 311 | 69780 | 34.8667 | 90.6135 | 50.9737 |
| 312 | 69868 | 37.7656 | 87.2181 | 53.8966 |
| 313 | 70173 | 43.2625 | 86.0259 | 57.7865 |
| 314 | 70180 | 41.6479 | 84.2867 | 59.6296 |
| 315 | 70185 | 41.1105 | 81.6516 | 58.0926 |

| | | | | |
|---|---|---|---|---|
| 316 | 70192 | 38.539 | 82.9687 | 59.9469 |
| 317 | 70403 | 47.4857 | 91.7645 | 63.903 |
| 318 | 70448 | 44.6488 | 86.6435 | 62.995 |
| 319 | 70795 | 31.3777 | 97.633 | 45.8635 |
| 320 | 70814 | 34.0369 | 99.5849 | 46.3521 |
| 321 | 70863 | 39.2573 | 104.826 | 52.8019 |
| 322 | 70868 | 37.1413 | 103.235 | 50.4358 |
| 323 | 70898 | 37.7274 | 108.135 | 53.0657 |
| 324 | 70909 | 37.8061 | 108.634 | 56.3451 |
| 325 | 70957 | 41.1063 | 108.549 | 52.984 |
| 326 | 70980 | 37.8221 | 110.59 | 50.352 |
| 327 | 70984 | 40.8773 | 109.738 | 49.3581 |
| 328 | 70998 | 41.1812 | 111.358 | 46.4168 |
| 329 | 71350 | 37.7664 | 98.379 | 55.0397 |
| 330 | 71363 | 41.0268 | 100.018 | 56.3091 |
| 331 | 71402 | 41.8071 | 102.834 | 57.3825 |
| 332 | 71409 | 39.5281 | 100.932 | 58.8567 |
| 333 | 71447 | 38.178 | 101.423 | 54.2224 |
| 334 | 71490 | 41.9839 | 103.074 | 52.1858 |
| 335 | 71761 | 32.4447 | 85.849 | 55.9095 |
| 336 | 71780 | 32.9563 | 86.7426 | 51.1933 |
| 337 | 71788 | 35.7298 | 85.4724 | 51.8931 |
| 338 | 72727 | 33.2821 | 90.1495 | 60.4076 |
| 339 | 73279 | 39.9296 | 93.0622 | 54.8181 |
| 340 | 73341 | 41.8858 | 96.5049 | 61.4788 |
| 341 | 73783 | 34.8629 | 88.6419 | 48.3954 |
| 342 | 73805 | 36.9979 | 91.4334 | 48.5528 |
| 343 | 73842 | 37.7461 | 92.3367 | 51.4583 |
| 344 | 73863 | 40.9231 | 89.998 | 51.9206 |
| 345 | 73928 | 38.3432 | 83.8801 | 53.7603 |
| 346 | 73982 | 39.5339 | 86.1558 | 50.199 |
| 347 | 74211 | 30.6099 | 95.7225 | 56.5923 |
| 348 | 74481 | 29.5939 | 104.194 | 49.3172 |
| 349 | 74497 | 31.176 | 104.972 | 46.1078 |
| 350 | 75140 | 31.6289 | 91.2614 | 62.8747 |
| 351 | 75163 | 31.8913 | 88.234 | 62.3119 |
| 352 | 75184 | 30.2776 | 86.5345 | 60.1126 |
| 353 | 75209 | 24.9201 | 89.0394 | 57.7541 |
| 354 | 75263 | 25.1058 | 88.8183 | 54.7432 |
| 355 | 75283 | 26.023 | 88.6782 | 51.6782 |
| 356 | 75306 | 28.7641 | 89.7452 | 53.1146 |
| 357 | 75344 | 27.118 | 91.2198 | 55.5535 |
| 358 | 75482 | 28.017 | 93.5992 | 53.2126 |
| 359 | 75680 | 27.1967 | 86.7408 | 61.4738 |
| 360 | 76291 | 27.791 | 86.9577 | 55.4679 |
| 361 | 77268 | 22.1058 | 90.08 | 56.9053 |
| 362 | 78775 | 23.7425 | 91.7425 | 53.4331 |
| 363 | 78878 | 27.1233 | 94.1834 | 56.4755 |
| 364 | 80256 | 22.2289 | 88.7297 | 50.1237 |
| 365 | 80275 | 23.4285 | 90.5491 | 47.7484 |
| 366 | 80287 | 26.6141 | 90.1341 | 45.7716 |
| 367 | 80299 | 25.8931 | 92.7881 | 47.8424 |
| 368 | 80335 | 26.3547 | 94.8723 | 50.3131 |
| 369 | 80364 | 29.5411 | 92.3513 | 50.1328 |

| | | | | |
|---|---|---|---|---|
| 370 | 80368 | 28.3792 | 89.5369 | 49.2271 |
| 371 | 80488 | 29.854 | 90.7421 | 45.0148 |
| 372 | 80498 | 29.2553 | 92.4854 | 42.3315 |
| 373 | 80833 | 26.2475 | 91.4852 | 50.5396 |
| 374 | 81256 | 21.7263 | 89.1577 | 53.3164 |
| 375 | 81684 | 27.6699 | 83.8427 | 62.7429 |
| 376 | 81698 | 24.4008 | 86.3436 | 62.5116 |
| 377 | 81727 | 22.2985 | 85.8972 | 59.6308 |
| 378 | 81737 | 23.4906 | 86.0971 | 56.6518 |
| 379 | 81756 | 22.3084 | 86.1951 | 53.8775 |
| 380 | 81787 | 24.4573 | 87.2692 | 48.6541 |
| 381 | 82184 | 27.9787 | 84.4605 | 58.4956 |
| 382 | 82536 | 28.7468 | 93.8865 | 63.0177 |
| 383 | 82860 | 30.8261 | 92.1178 | 53.6098 |
| 384 | 82991 | 28.1043 | 95.5177 | 47.2876 |
| 385 | 83830 | 27.5138 | 96.6518 | 53.3359 |
| 386 | 85656 | 30.5137 | 82.9544 | 59.3211 |
| 387 | 85681 | 27.7306 | 81.0812 | 59.3852 |
| 388 | 85684 | 28.3489 | 78.6527 | 57.3693 |
| 389 | 85692 | 24.4962 | 81.004 | 58.2003 |
| 390 | 85703 | 22.5304 | 83.2718 | 55.9334 |
| 391 | 85710 | 19.2534 | 86.5151 | 54.5126 |
| 392 | 85728 | 19.5134 | 84.8532 | 57.4599 |
| 393 | 85756 | 22.5055 | 85.1964 | 49.706 |
| 394 | 85776 | 21.2651 | 86.4879 | 46.9104 |
| 395 | 85781 | 23.2734 | 87.9498 | 45.036 |
| 396 | 85787 | 26.1244 | 86.9024 | 45.2619 |
| 397 | 86211 | 18.7137 | 90.3395 | 55.4939 |
| 398 | 86248 | 21.6722 | 91.7633 | 50.1848 |
| 399 | 86261 | 18.7958 | 89.0472 | 49.5443 |
| 400 | 86275 | 19.4712 | 91.1445 | 46.6483 |
| 401 | 86789 | 27.9465 | 86.0356 | 47.6391 |
| 402 | 87198 | 25.0303 | 84.5905 | 59.6473 |
| 403 | 87484 | 28.0525 | 89.3096 | 43.0765 |
| 404 | 89216 | 20.5036 | 92.6082 | 53.955 |
| 405 | 89278 | 23.5655 | 92.4472 | 44.774 |
| 406 | 89444 | 32.0623 | 93.0243 | 44.3804 |
| 407 | 89472 | 30.3518 | 95.8711 | 43.0343 |
| 408 | 89486 | 30.9222 | 94.878 | 40.2397 |
| 409 | 89491 | 33.5377 | 96.9462 | 40.2731 |
| 410 | 95673 | 26.0745 | 84.3853 | 55.9713 |
| 411 | 95710 | 16.4406 | 87.8914 | 53.3631 |
| 412 | 95728 | 17.0207 | 87.4968 | 56.4494 |
| 413 | 96755 | 19.8108 | 86.2553 | 50.7753 |
| 414 | 96867 | 28.255 | 86.5206 | 52.0834 |
| 415 | 96934 | 28.6661 | 82.782 | 54.0303 |
| 416 | 96946 | 26.1825 | 83.055 | 52.1309 |
| 417 | 99312 | 23.0765 | 93.9322 | 47.9197 |
| 418 | 100863 | 33.8043 | 91.3473 | 54.4324 |
| 419 | 102715 | 15.546 | 90.6606 | 55.0112 |
| 420 | 102999 | 26.6686 | 93.332 | 44.0585 |
| 421 | 103270 | 17.4372 | 87.3008 | 47.2752 |
| 422 | 103279 | 18.4729 | 88.8681 | 44.6482 |
| 423 | 103990 | 31.9126 | 89.9914 | 50.3358 |

| | | | | |
|---|---|---|---|---|
| 424 | 104498 | 29.9696 | 101.218 | 47.708 |
| 425 | 105000 | 44.2066 | 102.155 | 65.9959 |
| 426 | 105139 | 46.9223 | 101 | 67.1848 |
| 427 | 105280 | 44.9848 | 104.24 | 63.2626 |
| 428 | 105305 | 46.4989 | 104.231 | 66.1887 |
| 429 | 105338 | 48.4888 | 104.2 | 68.8257 |
| 430 | 105360 | 50.259 | 101.593 | 70.9061 |
| 431 | 105401 | 53.5728 | 101.516 | 69.9864 |
| 432 | 105484 | 51.8713 | 102.77 | 67.6231 |
| 433 | 105715 | 43.7239 | 105.624 | 66.0976 |
| 434 | 106409 | 46.0492 | 102.293 | 70.0835 |
| 435 | 107410 | 52.7777 | 101.977 | 72.9777 |
| 436 | 107844 | 45.8206 | 106.034 | 68.8815 |
| 437 | 107916 | 48.8054 | 103.458 | 72.9513 |
| 438 | 108710 | 41.6658 | 104.541 | 63.6906 |
| 439 | 109182 | 42.9229 | 99.6262 | 64.6816 |
| 440 | 110403 | 53.56 | 104.567 | 69.5585 |
| 441 | 110498 | 51.4674 | 106.214 | 66.2446 |
| 442 | 110786 | 47.7289 | 103.485 | 62.0441 |
| 443 | 110882 | 51.3412 | 105.055 | 71.9384 |
| 444 | 110907 | 55.3443 | 104.288 | 72.542 |
| 445 | 111435 | 51.591 | 99.148 | 69.6598 |
| 446 | 112595 | 46.1796 | 100.609 | 64.0389 |
| 447 | 113291 | 49.2123 | 102.651 | 65.0806 |
| 448 | 113340 | 51.7085 | 107.21 | 69.7235 |
| 449 | 113363 | 56.0926 | 106.964 | 71.016 |
| 450 | 113413 | 56.4933 | 107.746 | 74.4604 |
| 451 | 113436 | 58.4697 | 105.03 | 72.1628 |
| 452 | 113486 | 57.6742 | 105.415 | 68.7294 |
| 453 | 113497 | 59.9748 | 106.569 | 66.8707 |
| 454 | 115156 | 44.6561 | 98.4349 | 67.1681 |
| 455 | 115215 | 40.2773 | 107.724 | 62.7579 |
| 456 | 115247 | 40.8529 | 109.301 | 59.7641 |
| 457 | 115278 | 40.1412 | 111.03 | 55.7638 |
| 458 | 115286 | 44.2274 | 110.141 | 55.9517 |
| 459 | 115307 | 45.2747 | 113.085 | 56.3029 |
| 460 | 115321 | 46.9275 | 110.435 | 57.4536 |
| 461 | 115335 | 46.296 | 114.24 | 58.902 |
| 462 | 115341 | 48.9619 | 114.711 | 60.617 |
| 463 | 115359 | 52.1751 | 114.422 | 60.0674 |
| 464 | 115363 | 55.9925 | 113.145 | 61.1176 |
| 465 | 115368 | 54.4061 | 112.42 | 58.598 |
| 466 | 115409 | 56.4225 | 113.227 | 64.1421 |
| 467 | 115424 | 58.6723 | 112.289 | 62.1732 |
| 468 | 115489 | 58.8803 | 113.656 | 59.4195 |
| 469 | 115498 | 58.2396 | 114.77 | 56.6617 |
| 470 | 115863 | 56.5037 | 102.417 | 69.5209 |
| 471 | 115936 | 56.2888 | 100.337 | 72.3219 |
| 472 | 115987 | 54.6659 | 104.268 | 65.9115 |
| 473 | 116256 | 40.458 | 106.984 | 56.697 |
| 474 | 116286 | 43.6556 | 108.904 | 51.4399 |
| 475 | 116306 | 44.8257 | 111.86 | 53.4335 |
| 476 | 116331 | 42.1358 | 113.592 | 55.4481 |
| 477 | 116404 | 49.1075 | 116.275 | 57.7059 |

| | | | | |
|---|---|---|---|---|
| 478 | 116431 | 48.8825 | 112.614 | 56.3765 |
| 479 | 116480 | 44.1612 | 115.262 | 53.4978 |
| 480 | 116488 | 47.5714 | 115.898 | 53.7411 |
| 481 | 116498 | 46.105 | 116.354 | 50.7057 |
| 482 | 116907 | 56.4768 | 102.224 | 75.5882 |
| 483 | 116917 | 53.3798 | 101.624 | 77.1346 |
| 484 | 116933 | 55.6788 | 98.0079 | 74.7437 |
| 485 | 117291 | 45.3159 | 107.445 | 54.0051 |
| 486 | 117402 | 55.4575 | 116.421 | 61.1546 |
| 487 | 117409 | 53.7218 | 116.586 | 64.2046 |
| 488 | 117475 | 53.6401 | 116.47 | 58.2615 |
| 489 | 117486 | 55.2147 | 115.765 | 55.3992 |
| 490 | 117692 | 40.9907 | 102.819 | 66.2049 |
| 491 | 117781 | 40.5898 | 112.339 | 52.5721 |
| 492 | 117835 | 44.0065 | 115.856 | 56.4589 |
| 493 | 117854 | 47.5962 | 118.041 | 59.7813 |
| 494 | 117974 | 51.5467 | 117.787 | 56.4117 |
| 495 | 117984 | 52.5634 | 117.759 | 53.5008 |
| 496 | 118498 | 55.8243 | 117.147 | 52.3829 |
| 497 | 118913 | 50.6577 | 114.306 | 63.1101 |
| 498 | 118951 | 52.1204 | 113.555 | 56.9154 |
| 499 | 119238 | 41.9957 | 106.32 | 60.373 |
| 500 | 119315 | 47.7193 | 110.454 | 54.411 |
| 501 | 119418 | 50.7962 | 111.673 | 59.2568 |
| 502 | 119424 | 51.6589 | 110.281 | 56.6445 |
| 503 | 119480 | 50.1562 | 112.779 | 53.4196 |
| 504 | 119486 | 52.7088 | 111.05 | 51.8936 |
| 505 | 119912 | 53.4162 | 110.944 | 62.1228 |
| 506 | 119933 | 55.3974 | 108.84 | 58.2424 |
| 507 | 119959 | 54.362 | 111.674 | 55.6412 |
| 508 | 120000 | 18.953 | 79.5523 | 82.011 |
| 509 | 120002 | 22.0036 | 78.7428 | 81.2804 |
| 510 | 120069 | 20.7699 | 81.1297 | 79.2364 |
| 511 | 120211 | 16.4911 | 81.7329 | 80.2903 |
| 512 | 120262 | 19.3688 | 79.6948 | 76.9109 |
| 513 | 120279 | 20.4945 | 79.3009 | 73.6495 |
| 514 | 120291 | 22.3662 | 77.039 | 74.3727 |
| 515 | 120376 | 18.1417 | 76.7232 | 75.4285 |
| 516 | 120413 | 20.1691 | 77.3978 | 78.8793 |
| 517 | 120745 | 24.1418 | 80.8811 | 78.6827 |
| 518 | 120764 | 24.1967 | 82.5052 | 74.7373 |
| 519 | 120787 | 26.9023 | 83.7686 | 73.6383 |
| 520 | 120835 | 24.9283 | 86.3467 | 75.947 |
| 521 | 120863 | 30.1305 | 83.5625 | 75.3522 |
| 522 | 120909 | 26.9759 | 83.226 | 76.7673 |
| 523 | 121247 | 18.6207 | 83.9159 | 77.0771 |
| 524 | 121296 | 16.3345 | 85.7534 | 79.4698 |
| 525 | 121307 | 18.2223 | 87.7678 | 80.7654 |
| 526 | 121363 | 18.6982 | 84.5049 | 81.91 |
| 527 | 121417 | 13.9186 | 85.9697 | 81.5082 |
| 528 | 121478 | 15.4563 | 88.902 | 80.3228 |
| 529 | 121658 | 24.7281 | 76.6211 | 80.3658 |
| 530 | 122141 | 26.3983 | 79.405 | 80.9773 |
| 531 | 122158 | 27.5543 | 74.3727 | 80.4158 |

| | | | | |
|---|---|---|---|---|
| 532 | 122166 | 25.2587 | 73.7715 | 82.4485 |
| 533 | 122182 | 23.7861 | 70.7096 | 81.3516 |
| 534 | 122208 | 22.0043 | 74.8598 | 79.6791 |
| 535 | 122264 | 24.6848 | 78.0655 | 77.7256 |
| 536 | 122288 | 28.9554 | 80.6778 | 75.2217 |
| 537 | 122292 | 28.6257 | 78.7502 | 78.3209 |
| 538 | 122338 | 28.9001 | 81.2923 | 81.5429 |
| 539 | 122362 | 31.9645 | 81.8943 | 82.187 |
| 540 | 122368 | 31.4314 | 81.0189 | 79.1994 |
| 541 | 122415 | 30.2304 | 82.7532 | 84.5578 |
| 542 | 122728 | 21.3261 | 77.1544 | 83.7643 |
| 543 | 122979 | 23.3961 | 85.3197 | 73.353 |
| 544 | 123673 | 25.2833 | 72.7584 | 78.9825 |
| 545 | 123762 | 27.3629 | 73.5829 | 76.7484 |
| 546 | 123780 | 28.9096 | 74.0266 | 74.0836 |
| 547 | 123786 | 32.271 | 73.2042 | 73.8527 |
| 548 | 123802 | 31.5367 | 74.1317 | 76.6624 |
| 549 | 123807 | 34.3934 | 76.9684 | 76.02 |
| 550 | 123821 | 34.7419 | 73.5767 | 75.8454 |
| 551 | 123830 | 35.1875 | 76.3217 | 79.1187 |
| 552 | 123841 | 36.7071 | 77.5257 | 82.0445 |
| 553 | 123860 | 40.0247 | 75.6943 | 81.6668 |
| 554 | 123863 | 42.9344 | 77.7425 | 82.8337 |
| 555 | 123882 | 38.4553 | 76.3057 | 84.7259 |
| 556 | 123909 | 42.9115 | 76.3875 | 87.5331 |
| 557 | 123924 | 44.2058 | 74.7622 | 85.13 |
| 558 | 123986 | 43.7521 | 76.3962 | 80.1336 |
| 559 | 124330 | 24.1933 | 81.1404 | 82.9413 |
| 560 | 124680 | 19.6266 | 75.5897 | 81.6119 |
| 561 | 124696 | 16.6151 | 74.8446 | 80.9911 |
| 562 | 124711 | 13.579 | 75.8572 | 79.2193 |
| 563 | 124727 | 16.9325 | 72.2923 | 82.6776 |
| 564 | 124752 | 19.101 | 74.1719 | 78.2375 |
| 565 | 124782 | 18.4213 | 72.7819 | 75.5622 |
| 566 | 124860 | 25.1572 | 76.5968 | 83.72 |
| 567 | 124937 | 28.2361 | 75.8027 | 83.1322 |
| 568 | 125195 | 22.5192 | 73.8388 | 83.887 |
| 569 | 125214 | 18.5924 | 79.3364 | 85.1297 |
| 570 | 125227 | 20.1638 | 75.3789 | 86.5236 |
| 571 | 125290 | 29.8896 | 76.735 | 80.6465 |
| 572 | 125363 | 35.3671 | 83.436 | 83.2209 |
| 573 | 125409 | 34.4127 | 84.1936 | 86.5397 |
| 574 | 125431 | 36.9498 | 80.4721 | 83.5972 |
| 575 | 125436 | 39.9858 | 80.5769 | 84.7838 |
| 576 | 125482 | 36.3664 | 80.8938 | 80.0141 |
| 577 | 125488 | 39.2943 | 80.0475 | 80.7649 |
| 578 | 125863 | 31.574 | 76.7378 | 84.1932 |
| 579 | 125899 | 29.6274 | 75.4902 | 86.2952 |
| 580 | 125910 | 31.4705 | 75.4849 | 89.0238 |
| 581 | 125925 | 31.3491 | 72.6293 | 85.5569 |
| 582 | 125936 | 34.4816 | 72.0094 | 86.588 |
| 583 | 126247 | 20.8972 | 82.2915 | 75.8689 |
| 584 | 126405 | 31.6317 | 83.0449 | 72.5863 |
| 585 | 126411 | 32.6013 | 82.2587 | 76.5508 |

| | | | | |
|---|---|---|---|---|
| 586 | 126490 | 34.9999 | 82.7893 | 70.7637 |
| 587 | 126842 | 19.466 | 89.532 | 82.8651 |
| 588 | 126881 | 17.3827 | 91.0188 | 85.0443 |
| 589 | 126980 | 15.8769 | 91.2212 | 82.2961 |
| 590 | 127989 | 35.2254 | 83.8848 | 76.1612 |
| 591 | 128215 | 16.1023 | 78.0534 | 82.8708 |
| 592 | 128228 | 18.0093 | 73.6495 | 85.3294 |
| 593 | 128290 | 27.0191 | 73.262 | 84.9973 |
| 594 | 128306 | 27.2804 | 77.4332 | 86.154 |
| 595 | 128327 | 27.2788 | 77.0475 | 89.2301 |
| 596 | 128335 | 25.3414 | 78.6977 | 90.902 |
| 597 | 128338 | 28.509 | 79.2378 | 90.9366 |
| 598 | 128345 | 28.0691 | 76.9534 | 93.5888 |
| 599 | 128360 | 31.0288 | 74.3423 | 93.508 |
| 600 | 128363 | 34.8596 | 76.1697 | 92.9075 |
| 601 | 128399 | 33.7218 | 75.6123 | 96.1379 |
| 602 | 128410 | 34.9001 | 76.4598 | 99.0427 |
| 603 | 128433 | 34.5249 | 72.2692 | 97.4436 |
| 604 | 128470 | 33.0866 | 72.1102 | 94.0482 |
| 605 | 130184 | 23.9439 | 67.4599 | 80.3345 |
| 606 | 130196 | 20.5191 | 70.1337 | 82.3105 |
| 607 | 130225 | 18.3102 | 73.7099 | 88.8235 |
| 608 | 130236 | 21.5675 | 72.7104 | 87.3349 |
| 609 | 130341 | 23.4984 | 79.7731 | 86.0651 |
| 610 | 130370 | 31.2862 | 78.1645 | 86.9037 |
| 611 | 130410 | 31.2183 | 78.2889 | 92.1407 |
| 612 | 130427 | 28.3307 | 73.3692 | 89.2716 |
| 613 | 130436 | 31.7829 | 72.4825 | 90.599 |
| 614 | 130470 | 29.0479 | 70.5954 | 87.9195 |
| 615 | 130789 | 30.8362 | 76.9803 | 75.9241 |
| 616 | 130892 | 32.827 | 84.1408 | 79.9812 |
| 617 | 130902 | 35.9437 | 83.8916 | 79.9616 |
| 618 | 130999 | 34.5735 | 85.0187 | 73.0541 |
| 619 | 131409 | 30.2056 | 75.3662 | 97.1366 |
| 620 | 131430 | 29.8995 | 71.53 | 94.1349 |
| 621 | 131491 | 33.2826 | 74.8288 | 86.5458 |
| 622 | 131842 | 28.7526 | 88.2128 | 77.2808 |
| 623 | 131863 | 32.8883 | 89.4595 | 77.9321 |
| 624 | 131878 | 29.828 | 91.062 | 78.0509 |
| 625 | 131903 | 33.1767 | 93.5581 | 78.075 |
| 626 | 131909 | 31.6412 | 90.2136 | 81.2517 |
| 627 | 131982 | 31.6441 | 92.1283 | 75.5176 |
| 628 | 131989 | 33.9699 | 94.0072 | 74.7202 |
| 629 | 132403 | 37.0119 | 85.9903 | 85.4499 |
| 630 | 132413 | 35.8241 | 82.8704 | 88.8553 |
| 631 | 132478 | 33.2338 | 85.566 | 83.3277 |
| 632 | 132711 | 15.5395 | 78.4889 | 75.8464 |
| 633 | 132780 | 19.6392 | 79.171 | 70.6692 |
| 634 | 132786 | 23.2614 | 77.2184 | 70.4608 |
| 635 | 132825 | 24.7563 | 74.6535 | 71.2111 |
| 636 | 132842 | 26.8295 | 77.4298 | 71.8677 |
| 637 | 132858 | 28.6612 | 73.9001 | 70.5185 |
| 638 | 132864 | 32.6014 | 72.8171 | 69.7769 |
| 639 | 132876 | 29.6448 | 71.5821 | 72.3498 |

| | | | | |
|---|---|---|---|---|
| 132901 | 35.0738 | 71.6998 | 73.0027 |
| 132912 | 33.8484 | 70.8564 | 78.6322 |
| 132928 | 33.846 | 69.2363 | 74.2766 |
| 132935 | 36.2367 | 67.8181 | 73.0238 |
| 132963 | 32.5181 | 69.6903 | 71.4375 |
| 132986 | 33.3601 | 71.0312 | 66.2788 |
| 132991 | 36.2441 | 71.9441 | 66.5981 |
| 133490 | 32.3849 | 75.0382 | 81.3756 |
| 133860 | 29.6136 | 85.0134 | 78.0958 |
| 133935 | 31.7834 | 86.4579 | 75.9626 |
| 133997 | 31.7782 | 88.4073 | 72.9392 |
| 134203 | 21.1163 | 69.4115 | 79.3805 |
| 134209 | 17.7349 | 69.9137 | 77.0093 |
| 134233 | 21.1766 | 70.9923 | 74.8239 |
| 134236 | 20.119 | 67.5346 | 75.4493 |
| 134246 | 20.4907 | 71.618 | 70.9707 |
| 134256 | 18.6703 | 69.1948 | 69.2343 |
| 134276 | 19.3189 | 70.4019 | 65.1386 |
| 134280 | 19.5323 | 72.7602 | 62.9868 |
| 134286 | 23.258 | 71.6708 | 62.8961 |
| 134290 | 25.2224 | 69.6034 | 63.8868 |
| 134316 | 26.5619 | 73.0335 | 64.0945 |
| 134329 | 26.2331 | 74.3662 | 68.2771 |
| 134368 | 29.7235 | 72.2161 | 67.4427 |
| 134403 | 33.6404 | 76.2365 | 70.2513 |
| 134481 | 31.6346 | 74.6404 | 66.2478 |
| 134497 | 34.7738 | 74.1346 | 64.0088 |
| 134711 | 15.1197 | 72.1864 | 76.9903 |
| 135000 | 29.0459 | 89.6405 | 88.3355 |
| 135017 | 29.7089 | 89.4147 | 85.4052 |
| 135139 | 32.7202 | 88.5042 | 84.8918 |
| 135205 | 28.043 | 90.6592 | 83.1813 |
| 135211 | 24.9868 | 89.9837 | 83.7963 |
| 135228 | 25.4922 | 88.3592 | 86.6472 |
| 135256 | 26.4729 | 87.4144 | 82.9137 |
| 135330 | 31.7874 | 91.4274 | 84.2028 |
| 135407 | 34.1534 | 88.9165 | 87.6539 |
| 136354 | 25.9137 | 91.697 | 86.5181 |
| 137848 | 31.1939 | 91.7012 | 88.9097 |
| 137982 | 30.8762 | 93.9244 | 86.8683 |
| 138307 | 35.0116 | 90.6104 | 81.8309 |
| 138338 | 37.7309 | 92.0036 | 82.0862 |
| 138350 | 39.5242 | 91.7616 | 84.7472 |
| 138363 | 43.184 | 91.5402 | 85.3436 |
| 138407 | 43.5204 | 92.8336 | 88.3856 |
| 138433 | 44.2816 | 89.5986 | 87.8424 |
| 138835 | 34.7847 | 91.8395 | 84.6473 |
| 138989 | 45.5002 | 93.1503 | 84.029 |
| 139335 | 23.6202 | 90.1418 | 88.4485 |
| 139863 | 37.5877 | 89.3288 | 87.1202 |
| 139899 | 36.5508 | 88.6826 | 90.3456 |
| 141355 | 40.2376 | 93.3599 | 87.6434 |
| 141377 | 41.1722 | 90.3503 | 87.6322 |
| 141410 | 44.0243 | 93.2672 | 91.3458 |

| | | | | |
|---|---|---|---|---|
| 694 | 141498 | 42.7751 | 93.6531 | 82.4144 |
| 695 | 141757 | 28.7347 | 88.4752 | 81.1797 |
| 696 | 141790 | 33.7175 | 88.0714 | 80.6796 |
| 697 | 141862 | 35.7372 | 88.9865 | 84.4535 |
| 698 | 141925 | 39.7277 | 88.9495 | 83.0981 |
| 699 | 142663 | 30.5177 | 86.0296 | 85.0819 |
| 700 | 142928 | 23.219 | 86.4031 | 88.1165 |
| 701 | 143182 | 24.0513 | 85.4065 | 82.8477 |
| 702 | 143209 | 22.4031 | 88.9085 | 81.5483 |
| 703 | 143486 | 42.0706 | 90.5368 | 81.0036 |
| 704 | 143898 | 46.8342 | 92.0132 | 87.1536 |
| 705 | 143913 | 47.1151 | 91.2019 | 90.4712 |
| 706 | 143930 | 47.3199 | 89.0836 | 87.9926 |
| 707 | 143983 | 47.3622 | 90.4546 | 84.2808 |
| 708 | 145215 | 25.2176 | 92.9834 | 81.6979 |
| 709 | 145256 | 25.7349 | 90.0135 | 80.5335 |
| 710 | 145718 | 26.9247 | 95.6636 | 82.1338 |
| 711 | 145732 | 27.4853 | 94.4994 | 85.8176 |
| 712 | 145796 | 25.755 | 97.642 | 86.5393 |
| 713 | 145803 | 27.8665 | 98.5886 | 88.6297 |
| 714 | 145838 | 30.2117 | 101.078 | 88.9476 |
| 715 | 145842 | 31.9797 | 101.073 | 91.4798 |
| 716 | 145860 | 33.921 | 97.318 | 92.6079 |
| 717 | 145863 | 37.0564 | 99.8205 | 92.4365 |
| 718 | 145868 | 34.9227 | 99.1155 | 90.3953 |
| 719 | 145879 | 34.1388 | 99.6864 | 95.04 |
| 720 | 145912 | 36.8513 | 99.2215 | 97.1264 |
| 721 | 145933 | 37.2424 | 95.258 | 95.125 |
| 722 | 146981 | 37.8177 | 90.7886 | 79.0639 |
| 723 | 147211 | 22.577 | 92.0831 | 85.8736 |
| 724 | 147220 | 25.0907 | 94.5315 | 88.0473 |
| 725 | 147260 | 28.3461 | 97.4529 | 85.0337 |
| 726 | 147278 | 29.0675 | 99.9565 | 82.6742 |
| 727 | 147288 | 32.2775 | 99.0437 | 81.693 |
| 728 | 147303 | 32.9286 | 100.165 | 84.7108 |
| 729 | 147328 | 35.0559 | 100.55 | 87.11 |
| 730 | 147341 | 34.8769 | 103.375 | 88.5201 |
| 731 | 147358 | 38.4842 | 103.808 | 88.4414 |
| 732 | 147363 | 42.315 | 102.255 | 89.2497 |
| 733 | 147368 | 40.8679 | 102.31 | 86.4285 |
| 734 | 147403 | 43.6443 | 105.568 | 88.9804 |
| 735 | 147409 | 42.581 | 103.882 | 92.0858 |
| 736 | 147436 | 45.7118 | 102.463 | 90.1809 |
| 737 | 147480 | 40.4706 | 105.29 | 86.2957 |
| 738 | 147487 | 42.8443 | 104.313 | 84.6272 |
| 739 | 147768 | 24.0576 | 95.2447 | 84.9896 |
| 740 | 147845 | 29.6293 | 98.7402 | 91.9923 |
| 741 | 148737 | 24.1935 | 89.0342 | 78.1453 |
| 742 | 149403 | 37.4235 | 91.9942 | 89.381 |
| 743 | 149412 | 37.3978 | 89.5144 | 93.4226 |
| 744 | 149722 | 21.1081 | 84.8307 | 84.5977 |
| 745 | 149791 | 26.9624 | 84.7207 | 84.4126 |
| 746 | 149841 | 27.9464 | 86.7248 | 87.2188 |
| 747 | 150063 | 0.773091 | 91.053 | 82.0184 |

| | | | | |
|---|---|---|---|---|
| 748 | 150093 | 2.78027 | 88.8576 | 81.3657 |
| 749 | 150290 | 5.62595 | 87.6153 | 82.5326 |
| 750 | 150694 | -1.59979 | 89.6136 | 83.316 |
| 751 | 150706 | -1.58749 | 92.386 | 80.4727 |
| 752 | 150722 | -3.79459 | 94.4117 | 80.688 |
| 753 | 150746 | -3.06832 | 93.9142 | 77.109 |
| 754 | 151210 | -4.94726 | 91.2632 | 81.5575 |
| 755 | 151255 | -5.007 | 91.3878 | 78.4521 |
| 756 | 151279 | -5.51102 | 93.1721 | 74.7199 |
| 757 | 151711 | -7.35783 | 93.4426 | 80.0594 |
| 758 | 151788 | -3.56995 | 92.8891 | 72.2841 |
| 759 | 151853 | -4.84338 | 96.8489 | 76.031 |
| 760 | 151898 | -2.82231 | 96.5366 | 78.6331 |
| 761 | 151915 | -3.53869 | 97.7547 | 81.6842 |
| 762 | 153756 | -2.86001 | 89.5364 | 79.9348 |
| 763 | 153873 | 1.45289 | 88.9699 | 84.4974 |
| 764 | 153965 | 2.90268 | 86.5006 | 85.5496 |
| 765 | 155780 | -5.32016 | 96.5822 | 72.4123 |
| 766 | 155858 | -0.278967 | 98.1172 | 79.0824 |
| 767 | 155907 | 0.632117 | 99.978 | 81.2631 |
| 768 | 155912 | 0.124685 | 99.5026 | 84.2533 |
| 769 | 155979 | -4.19098 | 100.385 | 80.3043 |
| 770 | 155998 | -2.18842 | 100.639 | 78.08 |
| 771 | 157182 | -0.324432 | 87.8975 | 80.9699 |
| 772 | 157261 | -7.98004 | 92.2618 | 76.571 |
| 773 | 157278 | -7.37256 | 93.7743 | 72.3892 |
| 774 | 157382 | -5.95177 | 95.9453 | 79.1129 |
| 775 | 157480 | -7.26707 | 99.9589 | 79.0237 |
| 776 | 157980 | -8.6012 | 97.1406 | 73.1437 |
| 777 | 160455 | -2.64491 | 100.738 | 83.1272 |
| 778 | 160658 | 3.53484 | 85.3781 | 82.1067 |
| 779 | 160904 | 1.71756 | 100.823 | 78.4633 |
| 780 | 160998 | -0.214703 | 101.841 | 76.1668 |
| 781 | 161778 | -6.92091 | 96.2931 | 69.8813 |
| 782 | 161784 | -5.34655 | 93.6329 | 69.0403 |
| 783 | 161814 | -2.60049 | 95.9501 | 70.4629 |
| 784 | 161841 | -2.7855 | 98.0858 | 73.5275 |
| 785 | 161861 | 0.000899255 | 96.329 | 72.6689 |
| 786 | 161903 | 2.01595 | 98.0474 | 71.1514 |
| 787 | 161978 | -1.8836 | 98.9224 | 70.3448 |
| 788 | 161991 | 0.512871 | 100.356 | 69.2235 |
| 789 | 161998 | -2.46085 | 100.421 | 67.7863 |
| 790 | 163508 | 1.26684 | 93.6631 | 85.2003 |
| 791 | 163717 | -9.09902 | 95.9062 | 79.3308 |
| 792 | 163727 | -10.1451 | 93.1266 | 82.2272 |
| 793 | 163766 | -9.73341 | 95.3182 | 75.8675 |
| 794 | 164397 | 4.59494 | 100.043 | 72.5609 |
| 795 | 164409 | 4.42454 | 99.5809 | 75.6508 |
| 796 | 164484 | 3.50253 | 101.479 | 67.3089 |
| 797 | 164991 | 2.71349 | 102.817 | 74.909 |

# APPENDIX C

# NAMD SAMPLE INPUT FILES

## C.1 Sample NAMD configuration file for TAMD simulation

The following script is for a TAMD simulation using the NAMD software.

```
#NAMD CONFIG FILE FOR TAMD runs
# input
# set path                /export/users/harishv/abc/tamd_dimer_3kt/
set input               ionized_1feh_o2
coordinates             $input.pdb
structure               $input.psf
bincoordinates         npt.restart.coor
extendedsystem          npt.restart.xsc
parameters          par_all22_prot_cmap.inp
parameters          toppar_all22_prot_heme.str
#parameters          ct800342w_si_003.txt
parameters          par2.inp
paratypecharmm          on
# output
set output              tamd77
outputname              $output
dcdfile                 ${output}.dcd
xstFile                 ${output}.xst
dcdfreq                 2000
xstFreq                 10000
binaryoutput            yes
binaryrestart           yes
outputEnergies          10000
restartfreq             10000
fixedAtoms              off
# Basic dynamics
exclude                 scaled1-4
1-4scaling              1
COMmotion               no
#dielectric             1.0
# Simulation space partitioning
switching               on
switchdist              8.0
cutoff                  10
pairlistdist            12
# Multiple timestepping
```

```
set dt 1
firsttimestep           0
timestep                $dt
stepspercycle           10
nonbondedFreq           1
fullElectFrequency      2
rigidbonds              none
# Periodic Boundary Conditions
if {0} {
cellBasisVector1     94.549    0.    0.
cellBasisVector2     0.     80.798   0.
cellBasisVector3     0.       0.    87.958
cellOrigin           24.078893661499023 91.66455078125 70.47941589355469
}
wrapAll             on
#PME (for full-system periodic electrostatics)
PME                 yes
PMEGridSizeX        95
PMEGridSizeY        82
PMEGridSizeZ        90
# Temperature control
set temperature         310
temperature             $temperature
# Langevin Dynamics
langevin                on
langevinDamping         5
langevinTemp            $temperature
langevinHydrogen        off
seed                    1125784
# external forces
if 0 {
constraints                     on
consKCol                        B
ConsExp                         2
consRef                         fixed.cnst
consKFile                       fixed.cnst
}
#set hvcv                        /export/users/harishv/hvcv/
tclforces           on
tclforcesscript   cfacv_tclforces.tcl
set labelPDB   label.pdb
set cvINP   cv.inp
set restrINP   restr.inp
set TAMDof                      100
# Scripting
run  1000000
```

## C.2 Sample NAMD configuration files

In this section, I provide general input scripts for running MD simulations in NAMD using the NPT ensemble for RGS proteins and human PDE 4.

### C.2.1 RGS4 equilibration

```
# equilibration in NPT ensemble run
#############################################################
## ADJUSTABLE PARAMETERS                                   ##
#############################################################
set input      ionized_conjugated_1agr95C
structure          $input.psf
coordinates        $input.pdb
set outputname     npt
set temperature    310
#continuing a run
#set inputname      npt.restart      ;# only need to edit this in one place!
#binCoordinates     $inputname.coor  ;# coordinates from last run (binary)
#extendedSystem    $inputname.xsc   ;# cell dimensions from last run
firsttimestep      0
#############################################################
## SIMULATION PARAMETERS                                   ##
#############################################################
# Input
paraTypeCharmm     on
#parameters          1a.prm
parameters          par_all36m_prot.prm
parameters          toppar_water_ions_namd.str
temperature         $temperature
COMmotion           no
# Force-Field Parameters
exclude             scaled1-4
1-4scaling          1.0
cutoff              10.
switching           on
switchdist          8.
pairlistdist        12
# Integrator Parameters
timestep            2.0
rigidBonds          all
nonbondedFreq       1
fullElectFrequency  2
stepspercycle       20
# Periodic Boundary Conditions
if {1} {
cellBasisVector1     89.55800247192383     0.   0.
cellBasisVector2     0.     66.03299713134766   0.
cellBasisVector3     0.       0.    57.72800064086914
cellOrigin          -0.7978525757789612 0.15729907155036926 0.8326889276504517
```

```
}
wrapAll             on
#PME (for full-system periodic electrostatics)
PME                 yes
PMEGridSizeX        91
PMEGridSizeY        68
PMEGridSizeZ        59
# Constant Temperature Control
langevin            on    ;# do langevin dynamics
langevinDamping     5     ;# damping coefficient (gamma) of 5/ps
langevinTemp        $temperature
langevinHydrogen    off    ;# don't couple langevin bath to hydrogens
# Constant Pressure Control (variable volume)
useGroupPressure      no
useFlexibleCell       no
useConstantArea       no
langevinPiston        on
langevinPistonTarget  1.01325 ;#  in bar -> 1 atm
langevinPistonPeriod  100.
langevinPistonDecay   50.
langevinPistonTemp    $temperature
# Output
outputName          $outputname
restartfreq         10000
dcdfreq             10000
outputEnergies      10000
outputPressure      10000
xstFreq             10000
minimize             1000
reinitvels           $temperature
run   500000000
```

### C.2.2 Human PDE 4 equilibration

```
# equilibration in NPT ensemble run
#############################################################
## ADJUSTABLE PARAMETERS                                   ##
#############################################################
set input          ionized
structure          ${input}_pde4ibm.psf
coordinates        ${input}_pde4ibm.pdb
set outputname     npt
set temperature    300
#continuing a run
#set inputname      npt.restart      ;# only need to edit this in one place!
#binCoordinates     $inputname.coor  ;# coordinates from last run (binary)
#extendedSystem    $inputname.xsc   ;# cell dimensions from last run
firsttimestep      0
#############################################################
## SIMULATION PARAMETERS                                   ##
#############################################################
# Input
paraTypeCharmm     on
parameters          par_all36m_prot.prm
parameters          toppar_water_ions_namd.str
parameters          ibm.prm
temperature         $temperature
COMmotion           no
# Force-Field Parameters
exclude             scaled1-4
1-4scaling          1.0
cutoff              10.
switching           on
switchdist          8.
pairlistdist        12
# Integrator Parameters
timestep            2.0
rigidBonds          all
nonbondedFreq       1
fullElectFrequency  2
stepspercycle       20
# Periodic Boundary Conditions
if {1} {
cellBasisVector1   80.03099822998047 0 0
cellBasisVector2   0 88.6500015258789 0
cellBasisVector3   0 0 80.71900177001953
cellOrigin         20.97156524658203 6.474047660827637 24.07125473022461
}
wrapAll             on
#PME (for full-system periodic electrostatics)
PME                 yes
#PMEGridSpacing        1
PMEGridSizeX        83
PMEGridSizeY        91
PMEGridSizeZ        83
```

```
# Constant Temperature Control
langevin            on    ;# do langevin dynamics
langevinDamping     5     ;# damping coefficient (gamma) of 5/ps
langevinTemp        $temperature
langevinHydrogen    off    ;# don't couple langevin bath to hydrogens
# Constant Pressure Control (variable volume)
if {1} {
useGroupPressure      no
useFlexibleCell       no
useConstantArea       no
langevinPiston        on
langevinPistonTarget  1.01325 ;#  in bar -> 1 atm
langevinPistonPeriod  100.
langevinPistonDecay   50.
langevinPistonTemp    $temperature
}
# Output
outputName          $outputname
restartfreq         10000
dcdfreq             10000
outputEnergies      10000
outputPressure      10000
xstFreq             10000
reinitvels           $temperature
run   30000000
```

# APPENDIX D

# LINUX AND BASH SCRIPTS

## D.1 Sample code catdcd

Catdcd functions much like the Unix "cat" command: it concatenates DCD files into a single DCD file. You can also use catdcd to write only selected atoms to the final DCD file. Starting with version 4.0, CatDCD is now built as part of the VMD Plugin tree, and shares the same reader/writer with VMD. CatDCD 4.0 can read/write any of the structure/trajectory formats that are supported by VMD by virtue of the plugin interface[1].

```
1 CatDCD 4.0
2    catdcd -o outputfile [-otype <filetype>] [-i indexfile]
3       [-stype <filetype>] [-s structurefile]
4       [-first firstframe] [-last lastframe] [-stride stride]
5       [-<filetype>] inputfile1 [-<filetype>] inputfile2 ...
6 * EXAMPLES *
7 catdcd -num eq01.dcd eq02.dcd
8  Prints the number of frames in the two DCD files, then the total,
9  then exits.
```

## D.2 Parallel calculations for per residue non-bonded interaction energy

VMD has an ability for handling analysis with high computational cost using parallel calculations. The following script distributes analysis task on N number of CPUs on the Cluster. In my thesis, I used this feature for calculating per residue non-bonded interaction energy described in chapter 6.

```
1 ##############################################################################
2 # (c) Hossein Mohammadiarani                                                 #
3 # Modified (Lines 7-37) and compiled by Mohammadjavad Mohammadi              #
4 ##############################################################################
5
6 #!/bin/bash
7 #SBATCH --job-name=zarn
```

[1] https://www.ks.uiuc.edu/Development/MDTools/catdcd/

```
#SBATCH --partition=shared
#SBATCH --nodelist=node105
#SBATCH -N 1
#SBATCH --cpus-per-task=24
#SBATCH --time 240:00:00
#SBATCH --output test.out
#SBATCH --get-user-env
#SBATCH --mail-type=ALL
#SBATCH --mail-user=mm14@wildcats.unh.edu
#SBATCH --exclusive
#SBATCH --gres=gpu:0
/mnt/lustre/chem-eng/mm14/
module load mpi/openmpi-x86_64
module load vmd/vmd-1.9.2-text
psf=/mnt/lustre/chem-eng/mm14/120/zar_nematode_resenergy/ionized_pde4zarc.psf
dcd=/mnt/lustre/chem-eng/mm14/120/zar_nematode_resenergy/npt_zc.dcd
NumCpus=24; # the number of cpus that deployed for calculations
totalnf=6000; # the total length of dcd file to be analyzed
FramesCpu=$(($totalnf/$NumCpus+1)); # the number of frames for each cpu
j=0;

iter=$(($NumCpus-1))
for j in $(seq 0 $iter )
do
                start=$(( $FramesCpu * $j))
                end=$(($FramesCpu * ($j+1)-1))
                srun --exclusive  -n 1 -c 1 -s -v vmd -e intermedcalc.vmd -
                   args $psf $dcd $start $end _out_$j.dat >& analysis_$j.log
                   &
                echo "last job number is : $j";
                j=$(expr $j + 1);
done

wait
```

# APPENDIX E

# SUPPORTING INFORMATION FOR STUDIES REPORTED IN CHAPTER 3

This appendix provides additional analyses related to Chapter 3.

## E.1 Supplemental Results: Testing the effect of F417Y mutation

Kubas et al. [10] have previously shown that mutating F417Y results in a decrease in the rate of oxygen diffusion toward the G-site. To demonstrate the effect of this specific mutation (F417Y) on oxygen diffusion along the MFEP between the Xe-site and the G-site (pathway 1-0 in Fig 3.2), I prepared three systems each for the wildtype (WT) and the mutant protein that differed in the location of oxygen along the MFEP.

Specifically, three locations of oxygen were chosen (see Fig E.1): one corresponding to the Xe-site (system 1), another in the proximity of residues F417/Y417 (system 2), and the third away from the F417/Y419, but toward the G-site (system 3). The WT systems contained 62,306 atoms and the mutant systems contained 62,307 atoms. All systems were energy minimized and briefly equilibrated in the NPT ensemble to equilibrate the volumes of simulation domains. I carried out mean-force calculations using 5-ns long restrained MD simulations (following the same protocols as for all centers reported in section 2.4.1.1) for each of the 3 oxygen locations along the same MFEP in the WT and the mutant proteins, and integrated these mean-forces to obtain the PMFs (Fig E.1c) for oxygen diffusion from the Xe-site toward the G-site. These results show an increase in the free-energy barrier of 2.4 kcal/mol for oxygen diffusion along the MFEP from the Xe-site toward the G-site. Importantly, the location of the barrier is near the side-chain of Y417, and the free-energy values after crossing this barrier are significantly lower toward the G-site (located near point 3 along the MFEP; see Fig E.1a,b) with respect to the Xe-site (point 1 along the MFEP; see Fig E.1a,b). The existence of a free-energy barrier for oxygen diffusion along this MFEP in the mutant protein implies a decreased diffusion rate, as suggested by Kubas et al. [10].

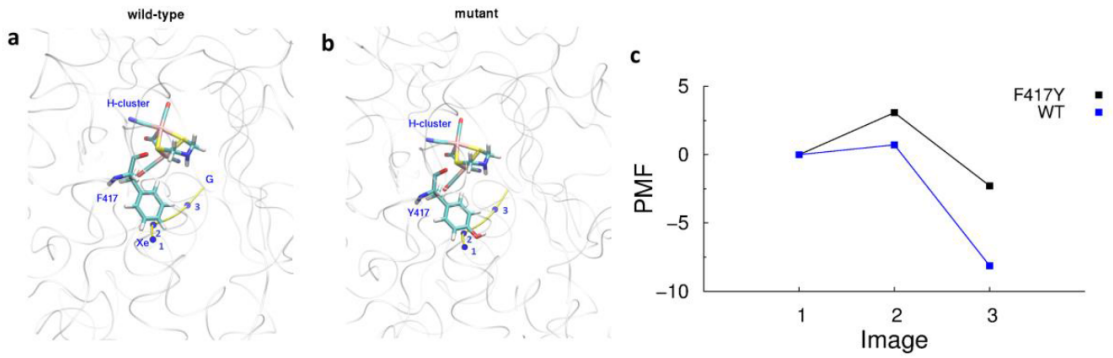



**Figure E.1.** The effect of F417Y mutation on the PMFs along the pathway 1-0 (see Fig 3.2) in [FeFe]-hydrogenase. (a and b) Snapshots highlighting the location of the wildtype F417 residue and the mutant Y417 residue along the pathway 1-0 (yellow curve). Three points (in the collective variable space) chosen to map the PMF are labeled and depicted as blue spheres in panel a. The approximate locations of the Xe-site and the G-site are marked by Xe and G, respectively. The H-cluster in each panel is labeled and highlighted in a stick representation. (c) The PMF values (kcal/mol) calculated with respect to the first point are shown for the WT protein by blue filled squares, and for the mutated protein by black filled squares. The increase in the free-energy value of the second point closest to the mutated residue is showing a higher energy-barrier of oxygen transition around Y417 toward the H-cluster in the mutated protein.

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | | 16.68 | 18.98 | | | | | 13.87 | | 13.51 | | | | | | | | | | | 17.26 | 18.7 | |
| 1 | 3.6 | | 5.55 | | 0.31 | 3.91 | | | | | | | | | 5.38 | | | | | | | | |
| 2 | 7.29 | 7.72 | | 6.76 | | | 6.28 | | | | | | | | | | | | | | 3.07 | | |
| 3 | | | 4.33 | | | | | | | | | | | 4.18 | | | | | | | | | 4.09 |
| 4 | | 0.76 | | | | | 4.58 | | | | | | | | | | | | | | | | |
| 5 | | 9.51 | | | | | | | | | | 12.6 | | 10.09 | 10.97 | | | | | | | | |
| 6 | | | 1.57 | | 1.0 | | | | | | | | | | | | | | | | | | |
| 7 | 2.01 | | | | | | | | 2.22 | | | | | | | | | | 9.83 | | | | |
| 8 | | | | | | | | 1.9 | | | | | | | | | | | | 0.54 | | | |
| 9 | 4.0 | | | | | | | | | | 4.57 | | 11.96 | | | 8.77 | 10.95 | 8.84 | | 8.04 | | | |
| 10 | | | | | | | | | | | | | | | | | | | | | 3.55 | | |
| 11 | | | | | | 2.9 | | | | | | | 2.22 | | | 1.43 | | 1.42 | | | | | |
| 12 | | | | | | | | | | 4.71 | | 2.89 | | | | | | 2.01 | | | | | |
| 13 | | | | 1.14 | | 0.59 | | | | | | | | | | | | | | | | | |
| 14 | | | | | | | | | | | | | | | | | | | | | | | |
| 15 | | | | | | | | | | 0.4 | | 0.96 | | | | | 2.57 | 1.14 | | 3.63 | | | |
| 16 | | | | | | | | | | 2.1 | | | | | | 2.11 | | 0.63 | | 3.17 | | | |
| 17 | | | | | | | | | | 1.04 | | 1.57 | 2.17 | | | 0.87 | 0.81 | | | 0.79 | | | |
| 18 | | | | | | | | 6.34 | | | | | | | | | | | | | | | |
| 19 | | | | | | | | | 3.83 | 3.78 | | | | | | 12.63 | 12.64 | 10.06 | | | | | |
| 20 | 3.5 | | 1.04 | | | | | | | | 3.79 | | | | | | | | | | | 22.2 | |
| 21 | 1.8 | | | | | | | | | | | | | | | | | | | | 19.05 | | |
| 22 | | | | | | | | | | | | | | | | | | | | | | | |

**Figure E.2.** Energy barriers (kcal/mol) between local minima in the 3D PMF map (Figure 3.2 in the main article). The value on the entry i j is the free energy barrier when going from local minimum i to local minimum j. The empty cells indicate negligible energy barriers or minima pairs that are not connected. The G-site and the Xe-site are shown by local minimum 0, and local minimum 1, respectively (Figure 3.2b).

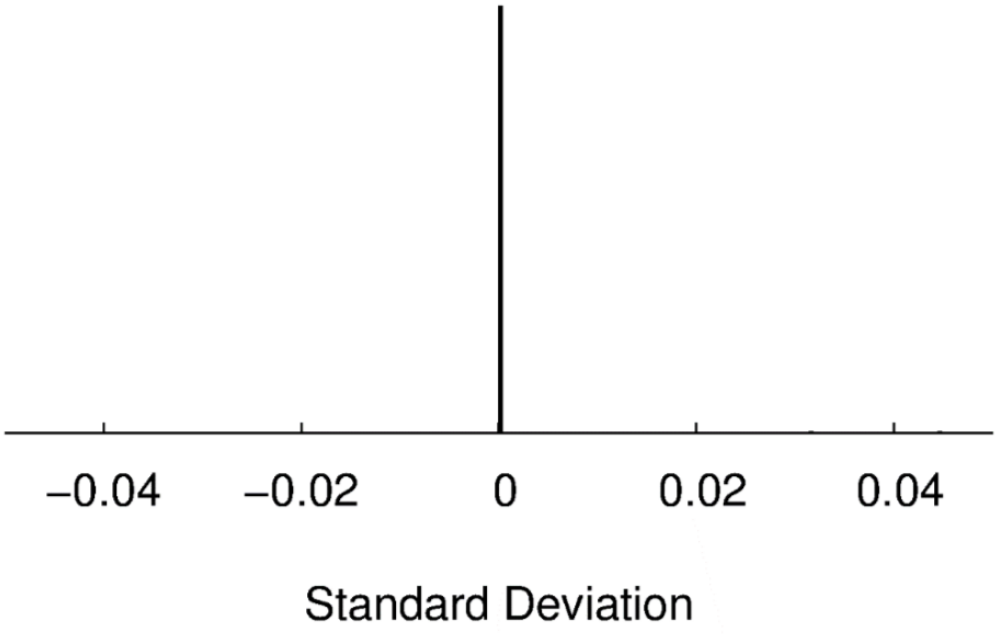



**Figure E.3.** Histogram of the standard deviation of each mean-force over the last 10% for all centers of $O_2$ in [FeFe]-hydrogenase. The convergence of mean-force values for all centers are represented as a delta function at the origin.

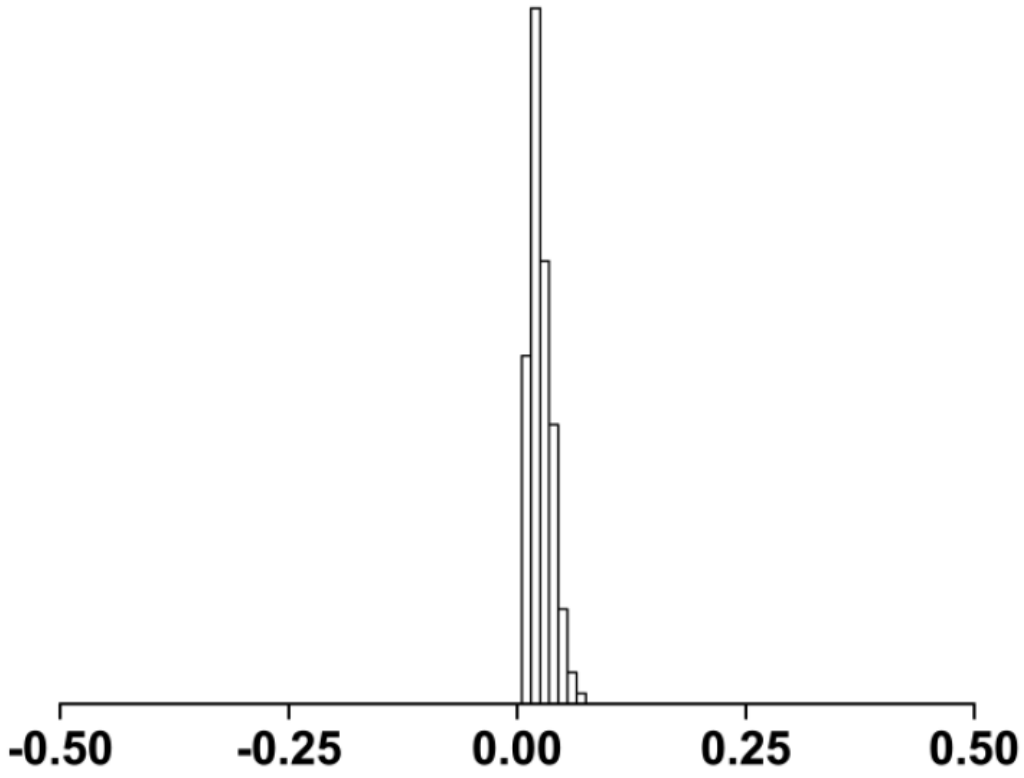



**Figure E.4.** Histogram of the standard deviation of each mean-force over the last 10% of each mean-force simulation. Data are from mean-force simulations for all 635 unique centers of CO in [FeFe]-hydrogenase (This figure was created by my co-author Yong Liu [3]).

| | G | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| G | - | 15.47 | 14.68 | 14.69 | 18.63 | 14.62 | 14.67 | 14.62 | 16.68 | 14.4 | 10.38 | 14.63 | 14.56 | 14.65 | 14.61 | 14.58 | 15.66 | 14.74 | 14.61 | 14.67 | 14.61 | 14.62 | 14.58 |
| 1 | 13.99 | - | 11.95 | 11.95 | 11.57 | 11.95 | 11.59 | 11.58 | 11.95 | 11.61 | 12.17 | 8.88 | 8.88 | 8.88 | 8.88 | 4.54 | 8.88 | 11.95 | 8.87 | 11.58 | 7.83 | 6.99 | 7.85 |
| 2 | 12.6 | 11.35 | - | 9.48 | 8.26 | 8.26 | 8.26 | 2.45 | 9.18 | 8.26 | 5.94 | 8.19 | 8.25 | 8.19 | 8.19 | 13.37 | 11.35 | 7 | 9.18 | 8.26 | 11.34 | 12.9 | 13.36 |
| 3 | 12.42 | 11.32 | 9.29 | - | 6.3 | 6.72 | 6.3 | 9.29 | 5.24 | 6.3 | 9.29 | 8.07 | 6.05 | 8 | 9.61 | 11.25 | 9.34 | 6.22 | 9.2 | 5.73 | 8.78 | 11.22 | 12.37 |
| 4 | 16.05 | 10.47 | 7.77 | 6 | - | 2.87 | 2.46 | 5.23 | 6.45 | 3.57 | 7.74 | 9.06 | 3.47 | 8.21 | 9.29 | 11.46 | 5.73 | 7.18 | 8.92 | 6.72 | 8.47 | 11.51 | 10.08 |
| 5 | 11.93 | 11.16 | 7.66 | 6.3 | 2.76 | - | 2.76 | 5.23 | 8.7 | 8.19 | 7.65 | 8.15 | 8.14 | 8.19 | 9.19 | 11.13 | 8.95 | 7.07 | 8.94 | 5.33 | 8.95 | 11.21 | 11.34 |
| 6 | 11.76 | 10.16 | 7.64 | 5.67 | 2.13 | 2.87 | - | 9.29 | 5.67 | 3.24 | 7.62 | 11.16 | 3.14 | 8.14 | 9.29 | 11.8 | 5.71 | 7.49 | 9.05 | 6.39 | 9.75 | 11.44 | 10.77 |
| 7 | 11.67 | 11.34 | 2.24 | 8.62 | 5.04 | 5.03 | 8.62 | - | 8.32 | 8.68 | 5.94 | 8.18 | 5.57 | 8.19 | 10.38 | 13.36 | 9.53 | 6.14 | 9.81 | 5.12 | 11.27 | 12.5 | 11.1 |
| 8 | 13.7 | 10.86 | 8.28 | 4.53 | 6.05 | 8.41 | 6.29 | 8.29 | - | 6.25 | 8.27 | 8.28 | 6.05 | 8.28 | 12.49 | 11.57 | 8.95 | 8.29 | 9.19 | 6.02 | 8.94 | 10.8 | 11.08 |
| 9 | 10.97 | 9.68 | 7.64 | 5.62 | 2.73 | 7.35 | 2.73 | 8.15 | 5.67 | - | 11.17 | 11.15 | 3.36 | 11.17 | 10.79 | 11.34 | 9.25 | 7.06 | 9.04 | 6.7 | 9.24 | 10.65 | 10.39 |
| 10 | 6.5 | 9.78 | 4.15 | 9.47 | 8.26 | 8.26 | 8.2 | 4.15 | 9.16 | 12.35 | - | 7.93 | 8.6 | 7.93 | 7.93 | 9.92 | 13.16 | 8.63 | 7.93 | 5.88 | 12.78 | 10.1 | 10.09 |
| 11 | 12.53 | 6.15 | 6.06 | 8.16 | 7.42 | 8.23 | 9.05 | 6.06 | 8.56 | 9.03 | 7.6 | - | 7.62 | 1.49 | 5.69 | 6.15 | 5.46 | 6.76 | 7.06 | 7.68 | 6.44 | 6.91 | 7.83 |
| 12 | 11.86 | 6.06 | 7.76 | 4.01 | 1.74 | 6.52 | 1.74 | 5.23 | 4.72 | 3.54 | 7.99 | 8.55 | - | 9.16 | 7.58 | 8.77 | 4 | 5.18 | 8.84 | 4.01 | 6.75 | 10.77 | 8.26 |
| 13 | 12.51 | 6.15 | 5.13 | 7.48 | 8.26 | 8.25 | 8.22 | 5.13 | 7.55 | 8.22 | 6.66 | 0.56 | 7.96 | - | 5.69 | 9.61 | 6.12 | 6.37 | 6.13 | 5.78 | 7.8 | 8.29 | 7.83 |
| 14 | 12.44 | 6.12 | 6.06 | 6.49 | 6.47 | 6.49 | 6.5 | 7.07 | 10.08 | 9.1 | 7.54 | 4.51 | 6.49 | 4.51 | - | 6.92 | 5.33 | 6.46 | 4.86 | 7.44 | 6.38 | 6.91 | 8.17 |
| 15 | 13.09 | 0.22 | 13.91 | 12.01 | 12.43 | 11.92 | 13.06 | 13.96 | 12.09 | 13.15 | 11.98 | 8.88 | 11.51 | 12.12 | 9.74 | - | 8.88 | 13.92 | 8.88 | 12.2 | 7.87 | 6.88 | 5.81 |
| 16 | 12.51 | 4.55 | 7.63 | 5.8 | 2.49 | 5.83 | 2.49 | 7.63 | 6.32 | 6.86 | 9.65 | 3.87 | 2.5 | 5.46 | 5.55 | 4.56 | - | 7.07 | 3.87 | 5.32 | 5.25 | 8.22 | 7.85 |
| 17 | 11.47 | 11.08 | 2.71 | 4.28 | 3.39 | 3.39 | 6.3 | 2.71 | 4.89 | 6.12 | 9.55 | 8.21 | 3.73 | 8.22 | 11.13 | 13.34 | 8.94 | - | 9.2 | 4.39 | 8.93 | 11.18 | 11.31 |
| 18 | 12.49 | 5.76 | 5.13 | 5.31 | 5.55 | 5.58 | 5.55 | 8.19 | 8.55 | 5.59 | 6.65 | 4.66 | 7.91 | 4.66 | 5.68 | 5.1 | 3.07 | 7.06 | - | 5.29 | 4.85 | 7.42 | 7.84 |
| 19 | 12.38 | 10.37 | 7.65 | 1.03 | 2.31 | 1.04 | 2.31 | 5.23 | 2.56 | 3.56 | 5.92 | 9.2 | 1.34 | 8.23 | 8.66 | 11.16 | 5.3 | 6.16 | 9.19 | - | 7.47 | 11.86 | 10.04 |
| 20 | 13.12 | 2.21 | 11.93 | 7.24 | 3.96 | 5.73 | 6.84 | 11.88 | 6.3 | 6.07 | 13.97 | 4.13 | 3.96 | 8.86 | 5.64 | 4.54 | 3.96 | 7.06 | 3.87 | 4.78 | - | 6.99 | 2.45 |
| 21 | 12.53 | 1.17 | 13.27 | 11.42 | 12.25 | 11.98 | 7.06 | 13.02 | 11.74 | 6.77 | 11.83 | 9.13 | 11.84 | 11.94 | 8.88 | 4.5 | 12.3 | 11.89 | 11.38 | 12.65 | 7.78 | - | 7.85 |
| 22 | 13.09 | 1.48 | 13.96 | 13 | 10.91 | 11.72 | 12.11 | 11.9 | 12.17 | 12.25 | 11.92 | 8.87 | 10.27 | 8.87 | 9.25 | 3.77 | 11.09 | 11.5 | 8.87 | 10.98 | 1.7 | 6.92 | - |
| 23 | 10.95 | 9.36 | 7.65 | 1.88 | 2.74 | 0.87 | 2.73 | 3.61 | 7.14 | 0.03 | 7.64 | 11.17 | 3.47 | 8.16 | 9.74 | 9.26 | 5.64 | 7.08 | 8.92 | 5.3 | 9.24 | 5.96 | 9.19 |
| 24 | 6.5 | 5.01 | 6.53 | 5.81 | 4.83 | 4.79 | 2.73 | 7.58 | 6.1 | 2.04 | 6.52 | 11.16 | 4.85 | 8.17 | 8.19 | 5.73 | 9.24 | 7.57 | 11.14 | 5.3 | 9.1 | 5.93 | 7.81 |
| 25 | 12.59 | 0.3 | 6.6 | 5.69 | 6.06 | 8.07 | 5.24 | 6.59 | 8.19 | 5.24 | 7.2 | 8.87 | 4.75 | 6.69 | 5.56 | 4.5 | 4.8 | 6.98 | 6.76 | 8.26 | 7.82 | 6.65 | 7.85 |
| 26 | 11.66 | 13.36 | 2.49 | 8.61 | 5.02 | 5.02 | 8.58 | 2.49 | 8.32 | 8.14 | 4.16 | 2.97 | 7.99 | 2.97 | 7.32 | 11.5 | 11.12 | 5.89 | 6.1 | 5.03 | 11.32 | 12.51 | 13.28 |
| 27 | 12.43 | 5.49 | 5.12 | 5.48 | 5.5 | 5.53 | 5.49 | 5.13 | 7.62 | 3.36 | 6.65 | 0.53 | 2.79 | 4.65 | 5.59 | 5.83 | 2.95 | 7.06 | 0.08 | 6.97 | 4.84 | 6.92 | 7.84 |
| 28 | 11.63 | 8.16 | 8.22 | 5.56 | 3.23 | 3.21 | 3.31 | 5.56 | 6.78 | 8.26 | 9.05 | 3.77 | 3.23 | 3.77 | 3.21 | 7.97 | 3.07 | 6.05 | 2.83 | 3.72 | 4.78 | 10.63 | 7.83 |
| 29 | 5.73 | 13.83 | 14.6 | 6.87 | 13.26 | 13.5 | 6.25 | 16.68 | 6.86 | 13.12 | 8.23 | 14.59 | 6.85 | 14.55 | 9.61 | 14.33 | 13.9 | 7.98 | 14.43 | 6.87 | 14.13 | 13.68 | 18.72 |
| 30 | 11.67 | 12.16 | 2.48 | 8.62 | 8.12 | 6.09 | 5.25 | 2.49 | 8.31 | 8.23 | 4.19 | 2.53 | 7.96 | 2.52 | 7.29 | 11.57 | 6.05 | 6.13 | 6.04 | 5.06 | 7.8 | 13.37 | 13.2 |
| 31 | 6.54 | 8.5 | 4.14 | 8.58 | 8.17 | 8.19 | 8.21 | 6.27 | 8.22 | 8.2 | 1.92 | 3.69 | 7.95 | 3.69 | 5.01 | 4.89 | 8.29 | 6.16 | 6.12 | 6.11 | 7.79 | 9.96 | 11.58 |
| 32 | 11.62 | 6.15 | 5.12 | 5.31 | 3.05 | 3.21 | 3.01 | 5.13 | 7.81 | 8.04 | 6.66 | 2.24 | 7.42 | 2.8 | 2.2 | 6.14 | 3.06 | 6.08 | 2.75 | 5.3 | 6.28 | 8.28 | 7.83 |
| 33 | 12.57 | 2.21 | 7.63 | 5.82 | 2.49 | 5.83 | 2.47 | 7.51 | 7.8 | 6.84 | 13.62 | 3.86 | 2.5 | 3.88 | 5.62 | 4.53 | 0.54 | 7.07 | 3.87 | 4.01 | 0.05 | 6.98 | 7.66 |
| 34 | 0.61 | 3.07 | 14.59 | 14.43 | 14.35 | 14.59 | 13.97 | 14.62 | 16.23 | 14.26 | 10.13 | 11.4 | 10.55 | 11.69 | 9.89 | 3.74 | 14.38 | 14.62 | 14.56 | 14.63 | 9.85 | 3.3 | 9.79 |
| 35 | 5.48 | 14.58 | 6.2 | 6.42 | 13.27 | 5.88 | 13.82 | 5.88 | 7.69 | 13.82 | 10.33 | 8.16 | 11.51 | 8 | 8.15 | 14.59 | 8.93 | 5.93 | 9.18 | 13.7 | 11.23 | 14.29 | 14.5 |
| 36 | 1.46 | 9.78 | 4.12 | 14.66 | 14.49 | 5.19 | 14.64 | 3.18 | 8.24 | 14.29 | 1.46 | 9.08 | 8 | 9.02 | 8.96 | 11.12 | 11.33 | 6.14 | 6.1 | 5.01 | 11.54 | 10.44 | 10.39 |
| 37 | 11.66 | 11.81 | 2.97 | 4.87 | 3.42 | 4.93 | 6.26 | 2.92 | 5.03 | 8.61 | 9.5 | 8.42 | 3.39 | 8.36 | 9.09 | 13.33 | 8.94 | 2.92 | 9.19 | 3.39 | 11.08 | 12.47 | 11.26 |
| 38 | 6.63 | 11.16 | 3.67 | 4.9 | 5.72 | 3.67 | 6.7 | 3.65 | 5.05 | 7.55 | 12.22 | 8.05 | 6.19 | 8.04 | 10.89 | 13.36 | 8.94 | 3.72 | 9.19 | 5.7 | 8.94 | 14.36 | 11.31 |
| 39 | 11.62 | 10.68 | 4.85 | 3.37 | 3.36 | 4.62 | 6.26 | 2.71 | 4.91 | 6.12 | 9.5 | 10.94 | 4.06 | 8.08 | 5.33 | 10.84 | 8.9 | 1.26 | 11.2 | 3.37 | 9.18 | 11.56 | 11.05 |
| 40 | 6.07 | 15.42 | 14.59 | 6.87 | 14.63 | 14.61 | 14.64 | 14.62 | 7.53 | 14.37 | 5.28 | 12.48 | 6.03 | 12.51 | 9.69 | 10.12 | 14.58 | 6.04 | 14.51 | 12.86 | 13.34 | 13.75 | 13.68 |
| 41 | 12.59 | 0.18 | 2.07 | 8.09 | 8.22 | 8.26 | 5.23 | 2.45 | 7.86 | 5.24 | 5.92 | 8.83 | 5.82 | 7.83 | 8.88 | 4.54 | 4.78 | 7 | 6.55 | 8.02 | 7.81 | 6.8 | 10.13 |
| 42 | 8.06 | 8.56 | 2.44 | 8.61 | 8.22 | 5.81 | 7.12 | 2.49 | 10.45 | 8.18 | 7.83 | 2.47 | 8 | 2.51 | 7.31 | 8.46 | 6.11 | 10.42 | 6.12 | 5.54 | 7.76 | 11.56 | 10.98 |
| 43 | 0.45 | 15.37 | 14.6 | 14.57 | 14.56 | 14.62 | 14.04 | 14.62 | 8.28 | 14.66 | 10.32 | 14.52 | 5.87 | 14.58 | 14.55 | 14.21 | 14.6 | 6.06 | 14.57 | 8.6 | 14.53 | 14.65 | 14.47 |
| 44 | 11.53 | 13.35 | 2.49 | 11.25 | 4.96 | 4.89 | 13.83 | 2.49 | 4.96 | 8.58 | 3.8 | 2.98 | 5 | 2.97 | 9.74 | 10.8 | 9.51 | 11.85 | 9.08 | 3.5 | 11.33 | 14.31 | 11.41 |

**Figure E.5.** Major energy barriers (kcal/mol) along MFEPs between minima pairs (MFEPs are shown in Fig. 3.3) . The value on the entry i j is the free energy barrier of the rate-limiting step from the local minima i to the local minima j (This figure was created by my co-author Yong Liu [3]).

| | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 18.9 | 14.57 | 14.69 | 14.61 | 14.59 | 14.62 | 14.43 | 14.61 | 11.22 | 14.61 | 14.71 | 11.35 | 14.3 | 12.67 | 14.63 | 18.38 | 14.61 | 18.91 | 15.45 | 16.42 | 18.8 | 22.55 |
| 1 | 12.02 | 11.59 | 7.08 | 13.97 | 8.88 | 11.95 | 13.54 | 15.83 | 12.22 | 10.97 | 8.69 | 12.33 | 14.39 | 12.18 | 13.17 | 13.92 | 11.95 | 18.83 | 12.09 | 14.09 | 18.7 | 19.65 |
| 2 | 8.26 | 12.51 | 11.35 | 8.76 | 9.18 | 9.2 | 14.29 | 8.79 | 9.12 | 10.38 | 11.35 | 12.53 | 14.91 | 10.58 | 12.61 | 13.34 | 11.29 | 18.85 | 13.37 | 13.52 | 18.79 | 19.05 |
| 3 | 7.56 | 7.82 | 11.23 | 9.29 | 9.33 | 11.98 | 13.3 | 9.3 | 10.69 | 9.62 | 9.35 | 12.4 | 13.3 | 12.66 | 12.37 | 13.13 | 11.1 | 13.3 | 11.27 | 12.33 | 18.34 | 19.41 |
| 4 | 6.25 | 7.52 | 10.61 | 8.26 | 9.04 | 9.35 | 12.89 | 9.17 | 10.81 | 9.32 | 8.08 | 11.8 | 12.96 | 12.53 | 12.06 | 13.12 | 10.8 | 18.83 | 11.43 | 13.5 | 18.4 | 18.55 |
| 5 | 6.13 | 7.41 | 11.16 | 7.88 | 8.95 | 9.21 | 13.13 | 9.17 | 10.94 | 9.21 | 8.95 | 11.63 | 13.99 | 9.97 | 11.94 | 12.74 | 10.69 | 18.79 | 11.34 | 13.39 | 18.43 | 18.34 |
| 6 | 5.08 | 7.19 | 10.59 | 8.65 | 9.03 | 9.09 | 11.08 | 8.26 | 9.07 | 9.3 | 8.06 | 11.48 | 13.47 | 12.62 | 12.33 | 16.16 | 11.07 | 18.85 | 10.63 | 13.28 | 17.8 | 20.34 |
| 7 | 8.62 | 8.62 | 11.34 | 7.9 | 9.14 | 11.3 | 14.15 | 7.9 | 10.13 | 10.27 | 11.16 | 11.72 | 13.73 | 9.72 | 11.7 | 12.46 | 10.43 | 18.87 | 13.37 | 12.67 | 18.6 | 18.18 |
| 8 | 7.55 | 7.82 | 11.29 | 8.29 | 8.91 | 12.49 | 13.28 | 8.42 | 10.03 | 11.25 | 8.71 | 13.68 | 15.51 | 9.68 | 11.66 | 12.43 | 10.41 | 15.9 | 11.09 | 16.29 | 15.45 | 19.56 |
| 9 | 4.57 | 6.68 | 10.08 | 8.71 | 6.88 | 9.4 | 12.72 | 8.73 | 9.08 | 10.21 | 9.24 | 10.97 | 13.48 | 12.53 | 11.63 | 16.25 | 10.68 | 18.87 | 10.12 | 13.46 | 18.78 | 18.17 |
| 10 | 8.26 | 12.43 | 12.92 | 8.64 | 7.93 | 9.11 | 14.3 | 8.67 | 7.33 | 8.58 | 11.53 | 11 | 14.44 | 8.79 | 10.76 | 18.02 | 10.34 | 14.24 | 13.35 | 11.73 | 18.79 | 17.25 |
| 11 | 9.05 | 12.48 | 7.07 | 7.11 | 4.86 | 8.24 | 14.27 | 7.13 | 8.76 | 7.69 | 6.44 | 10.82 | 14.78 | 8.77 | 11.69 | 12.55 | 11.47 | 18.85 | 12.04 | 11.39 | 18.75 | 16.92 |
| 12 | 6.25 | 7.5 | 10.43 | 8.13 | 5.37 | 7.61 | 13.27 | 8.58 | 10.22 | 8.62 | 6.35 | 11.23 | 16.74 | 9.98 | 10.32 | 12.42 | 9.12 | 15.62 | 10.95 | 12.91 | 15.44 | 18.52 |
| 13 | 8.23 | 12.5 | 8.29 | 6.18 | 6.12 | 7.32 | 14.29 | 6.2 | 7.83 | 7.32 | 6.44 | 11.09 | 14.69 | 8.72 | 11.7 | 13.18 | 10.51 | 18.87 | 9.76 | 10.46 | 18.75 | 16 |
| 14 | 8.23 | 12.84 | 8.72 | 7.07 | 4.76 | 6.51 | 13.86 | 7.09 | 8.7 | 6.47 | 6.38 | 10.76 | 13.9 | 8.64 | 10.07 | 12.04 | 12.08 | 17.33 | 12.09 | 10.45 | 18.42 | 15.81 |
| 15 | 11.99 | 12.28 | 7.03 | 13.94 | 8.88 | 11.93 | 14.23 | 13.88 | 11.43 | 10.95 | 8.67 | 11.57 | 14.23 | 12.44 | 13.93 | 16.56 | 11.92 | 15.54 | 12.09 | 14.07 | 18.73 | 18.94 |
| 16 | 5.32 | 6.85 | 7.34 | 8.7 | 3.87 | 5.71 | 13.54 | 6.1 | 7.83 | 5.71 | 4.85 | 12.51 | 13.98 | 10.56 | 11.9 | 12.73 | 10.66 | 18.86 | 7.73 | 10.45 | 18.79 | 19.04 |
| 17 | 6.13 | 8.52 | 11.35 | 7.77 | 9.2 | 11.21 | 13.54 | 7.88 | 10.01 | 11.25 | 8.94 | 11.66 | 13.53 | 9.71 | 8.27 | 9.11 | 7.02 | 12.51 | 13.36 | 13.9 | 15.42 | 14.75 |
| 18 | 6.11 | 8.18 | 8.24 | 6.16 | 2.02 | 4.9 | 14.03 | 6.08 | 7.82 | 4.9 | 4.85 | 12.43 | 13.95 | 7.52 | 11.88 | 12.71 | 10.33 | 18.86 | 8.26 | 10.45 | 18.75 | 18.31 |
| 19 | 6.13 | 7.41 | 11.39 | 7.88 | 8.18 | 7.63 | 13.28 | 8.39 | 10.48 | 9.21 | 6.35 | 12.38 | 15.73 | 9.68 | 9.34 | 13.1 | 7.77 | 17.09 | 11.25 | 12.9 | 16.15 | 17.46 |
| 20 | 5.81 | 12.01 | 7.08 | 11.94 | 3.86 | 5.62 | 13.79 | 9.8 | 11.46 | 7.48 | 3.07 | 12.28 | 14.46 | 13.88 | 12.39 | 12.72 | 10.65 | 18.82 | 12.08 | 14.07 | 18.75 | 18.99 |
| 21 | 11.79 | 6.7 | 6.81 | 12.94 | 8.94 | 11.76 | 13.86 | 13.25 | 12.14 | 11.94 | 8.65 | 6.75 | 14.21 | 8.76 | 12.96 | 17.98 | 11.78 | 18.79 | 11.96 | 13.81 | 18.8 | 20.9 |
| 22 | 12 | 12.3 | 7.08 | 13.84 | 8.88 | 11.91 | 13.95 | 13.97 | 12.23 | 10.41 | 8.5 | 11.58 | 14.22 | 11.72 | 13.1 | 16.73 | 11.8 | 18.67 | 12.09 | 14.03 | 18.68 | 19.58 |
| 23 | - | 6.56 | 9.7 | 8.61 | 8.94 | 9.45 | 14.18 | 8.73 | 9.09 | 10.28 | 9.23 | 11.32 | 13.65 | 12.58 | 11.69 | 16.14 | 10.68 | 18.86 | 11.4 | 13.47 | 18.77 | 18.17 |
| 24 | 4.45 | - | 5.01 | 14.6 | 10.73 | 9.19 | 14.29 | 8.76 | 8.92 | 10.34 | 9.44 | 11.32 | 14.23 | 12.66 | 15.68 | 18.34 | 10.41 | 18.86 | 5.47 | 14.87 | 18.8 | 21.12 |
| 25 | 5.24 | 4.81 | - | 8.76 | 4.8 | 9.13 | 14.09 | 9.23 | 9.1 | 7.32 | 4.85 | 5.53 | 14.32 | 9.65 | 12.49 | 13.14 | 11.26 | 18.79 | 5.28 | 13.12 | 18.77 | 19.03 |
| 26 | 9.34 | 14.61 | 13.37 | - | 6.1 | 11.54 | 14.28 | 0.67 | 7.83 | 7.29 | 6.09 | 11.02 | 13.73 | 4.31 | 11.67 | 15.43 | 10.4 | 18.8 | 13.37 | 7.25 | 18.79 | 12.77 |
| 27 | 6.13 | 8.07 | 7.77 | 6.16 | - | 4.9 | 13.26 | 6.13 | 8.75 | 7.57 | 4.85 | 10.89 | 13.96 | 7.48 | 11.8 | 12.68 | 10.74 | 18.63 | 12.08 | 11.38 | 18.75 | 18.2 |
| 28 | 6.13 | 8.81 | 10.67 | 8.22 | 2.83 | - | 13.56 | 8.98 | 7.85 | 1.38 | 4.84 | 11.65 | 13.57 | 9.35 | 11.56 | 12.04 | 8.92 | 15.65 | 10.92 | 9.31 | 15.25 | 17.57 |
| 29 | 14.56 | 14.4 | 14.2 | 14.6 | 13.56 | 11.12 | - | 14.54 | 12.49 | 8.57 | 13.36 | 11.21 | 3.31 | 12.64 | 9.72 | 9.72 | 7.99 | 16.44 | 15.42 | 15.57 | 9.74 | 14.12 |
| 30 | 8.23 | 12.56 | 14.78 | 0.21 | 6.13 | 8.85 | 14.27 | - | 7.88 | 7.17 | 6.08 | 10.89 | 13.72 | 4.31 | 11.7 | 12.43 | 10.6 | 18.8 | 11.57 | 6.81 | 18.44 | 12.77 |
| 31 | 8.24 | 12.32 | 13.35 | 6.18 | 4.86 | 7.35 | 14.25 | 6.11 | - | 7.32 | 6.44 | 11.05 | 13.76 | 10.86 | 15.84 | 15.42 | 11.43 | 14.22 | 8.44 | 10.46 | 18.73 | 12.77 |
| 32 | 8.13 | 12.47 | 8.29 | 6.15 | 4.78 | 0.4 | 13.93 | 5.98 | 7.83 | - | 6.29 | 11.18 | 13.47 | 7.46 | 11.57 | 12.09 | 8.65 | 17.34 | 8.58 | 10.44 | 18.43 | 13.95 |
| 33 | 6.85 | 7.68 | 7.25 | 6.15 | 3.87 | 5.69 | 12.99 | 6.13 | 8.73 | 7.5 | - | 11.04 | 14.45 | 8.77 | 11.87 | 12.73 | 9.44 | 18.78 | 12.07 | 11.38 | 18.78 | 15.94 |
| 34 | 14.38 | 14.38 | 3.05 | 12.33 | 10.964 | 14.5 | 14.16 | 12.26 | 11.04 | 12.49 | 13.24 | - | 14.41 | 12.55 | 17.91 | 18.26 | 14.55 | 18.88 | 13.34 | 15.45 | 18.79 | 22.58 |
| 35 | 13.96 | 14.37 | 14.58 | 7.89 | 9.19 | 11.14 | 1.2 | 7.9 | 9.46 | 8.43 | 9.46 | 11.37 | - | 12.65 | 8.27 | 7.59 | 10.39 | 8.08 | 14.53 | 12.5 | 7.65 | 11.82 |
| 36 | 14.47 | 14.39 | 12.44 | 3.08 | 6.07 | 8.72 | 14.27 | 2.76 | 7.21 | 7.26 | 9.52 | 11.09 | 18.83 | - | 13.17 | 18.84 | 12.47 | 6.92 | 12.8 | 7.25 | 18.79 | 11.66 |
| 37 | 8.63 | 13.45 | 11.78 | 7.89 | 9.14 | 11.21 | 6.69 | 7.89 | 9.41 | 11.22 | 9.19 | 11.62 | 12.01 | 10.2 | - | 9.02 | 7 | 10.48 | 13.39 | 7.58 | 15.82 | 11.64 |
| 38 | 6.87 | 14.17 | 11.24 | 7.88 | 9.17 | 11.05 | 6.66 | 7.87 | 10.06 | 14.43 | 8.94 | 11.29 | 6.63 | 12.66 | 8.25 | - | 7 | 8.08 | 13.39 | 12.27 | 7.61 | 16.13 |
| 39 | 6.12 | 8.61 | 11.49 | 7.87 | 5.22 | 5.47 | 10.31 | 8.56 | 11.05 | 5.2 | 6.65 | 11.31 | 13.71 | 9.02 | 8.27 | 9.04 | - | 16.17 | 13.35 | 7.15 | 15.38 | 12.04 |
| 40 | 14.66 | 14.39 | 14.51 | 12.51 | 14.43 | 11.01 | 12.3 | 12.49 | 7.3 | 9.69 | 14.55 | 11.2 | 6.04 | 5.29 | 11.27 | 7.61 | 11.73 | - | 15.43 | 7.69 | 6.04 | 9.75 |
| 41 | 8.21 | 4.8 | 0.15 | 8.76 | 8.87 | 9.1 | 14.28 | 8.12 | 7.73 | 7.32 | 9.153 | 12.47 | 14.42 | 10.02 | 12.61 | 13.34 | 11.23 | 18.86 | - | 11.72 | 18.79 | 18.71 |
| 42 | 8.19 | 13.83 | 12.97 | 0.74 | 4.86 | 9.09 | 14.29 | 0.75 | 7.83 | 7.3 | 6.42 | 11.04 | 13.63 | 4.31 | 4.43 | 12.09 | 10.36 | 14.19 | 11.58 | - | 14.2 | 12.19 |
| 43 | 14.65 | 14.41 | 14.6 | 14.59 | 14.52 | 11.05 | 1.18 | 14.67 | 12.46 | 14.6 | 14.48 | 11.34 | 0.09 | 12.37 | 16.6 | 7.64 | 10.31 | 0.52 | 15.45 | 7.18 | - | 4.3 |
| 44 | 8.61 | 14.4 | 13.47 | 0.87 | 8.98 | 10.68 | 2.57 | 0.88 | 10.41 | 10.54 | 6.04 | 11.3 | 2.36 | 2.61 | 11.64 | 11.4 | 12.04 | 2.33 | 13.04 | 12.79 | 2.4 | - |

**Figure E.6.** Continuation of major energy barriers (kcal/mol) along MFEPs between minima pairs from Fig E.5 (This figure was created by my co-author Yong Liu [3]).

# APPENDIX F

# SUPPORTING INFORMATION FOR STUDIES REPORTED IN CHAPTER 4

This appendix provides additional figures related to Chapter 4.

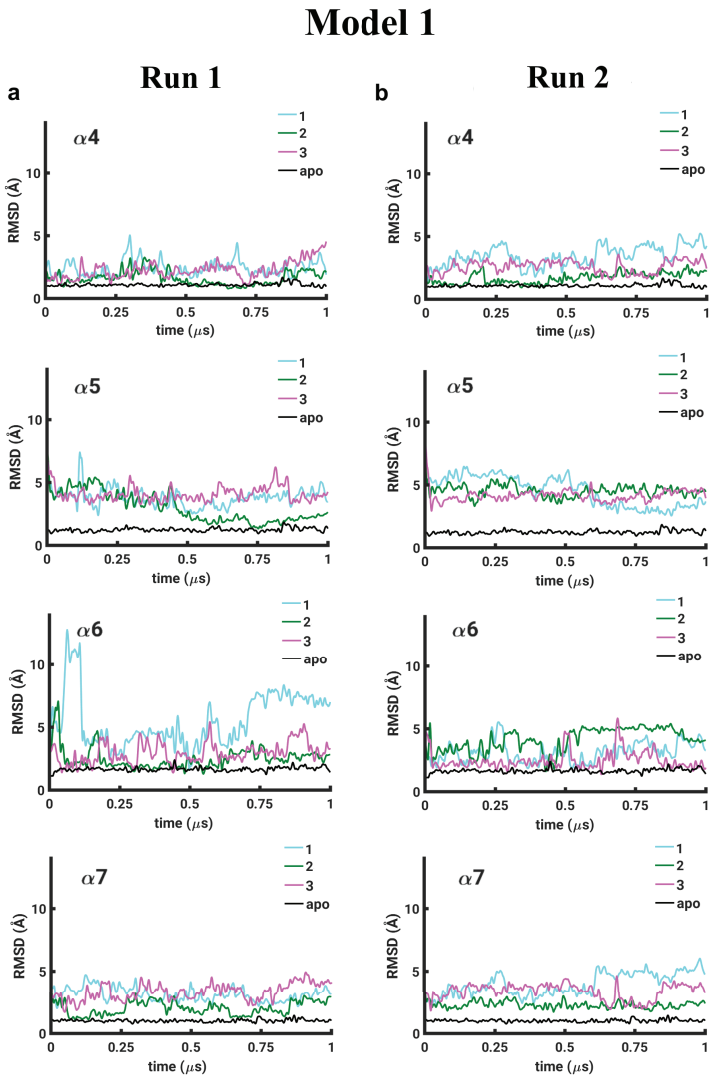



Figure F.1: The traces of root-mean-squared-deviation (RMSD) *vs.* simulation time ($\mu$s) for 4 helices in the $\alpha_4$-$\alpha_7$ helical bundle of RGS4 are shown from two independent simulation runs (Run1, panel a; Run2, panel b) for complexes of RGS4 with TDZD compounds **1** (cyan trace), **2** (green trace), and **3** (magenta trace). The black traces show data for an apo-RGS4 simulation from our previous work [13].

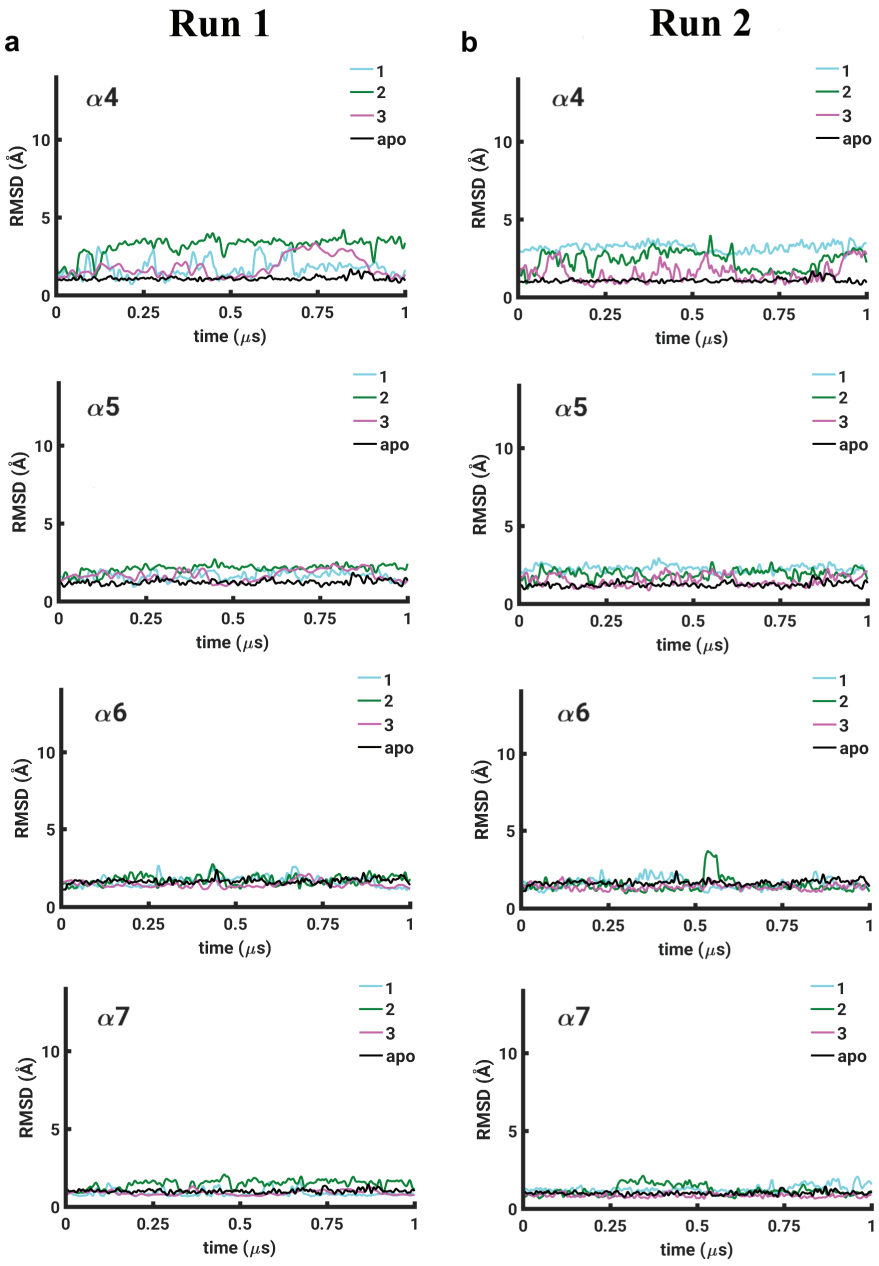



Figure F.2: Same data as in Fig. F.1 are shown for Model 2.

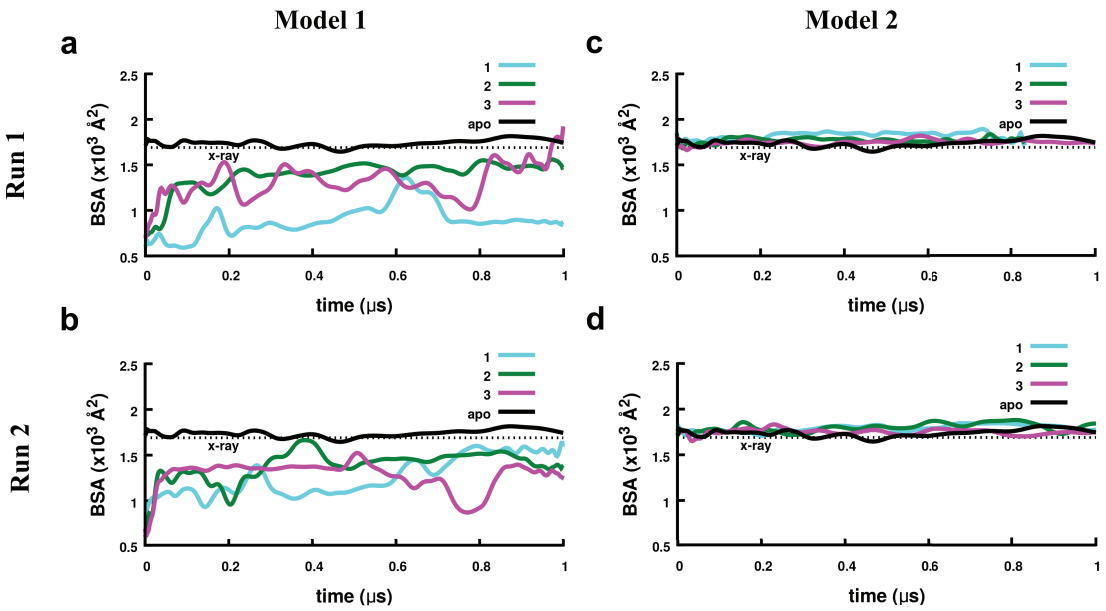



Figure F.3: The traces of buried surface area (BSA) between the $\alpha_5$-$\alpha_6$ helices and the rest of RGS4 *vs.* simulation time ($\mu$s) are shown from two independent simulation runs for each Model (Models 1 and 2). The BSA traces are shown for three TDZD compounds (cyan, green, and magenta traces) and from a simulation of apo-RGS4 (black traces). The dotted horizontal line in each panel highlights the BSA-value in the crystal structure of RGS4 (PDB: 1AGR).

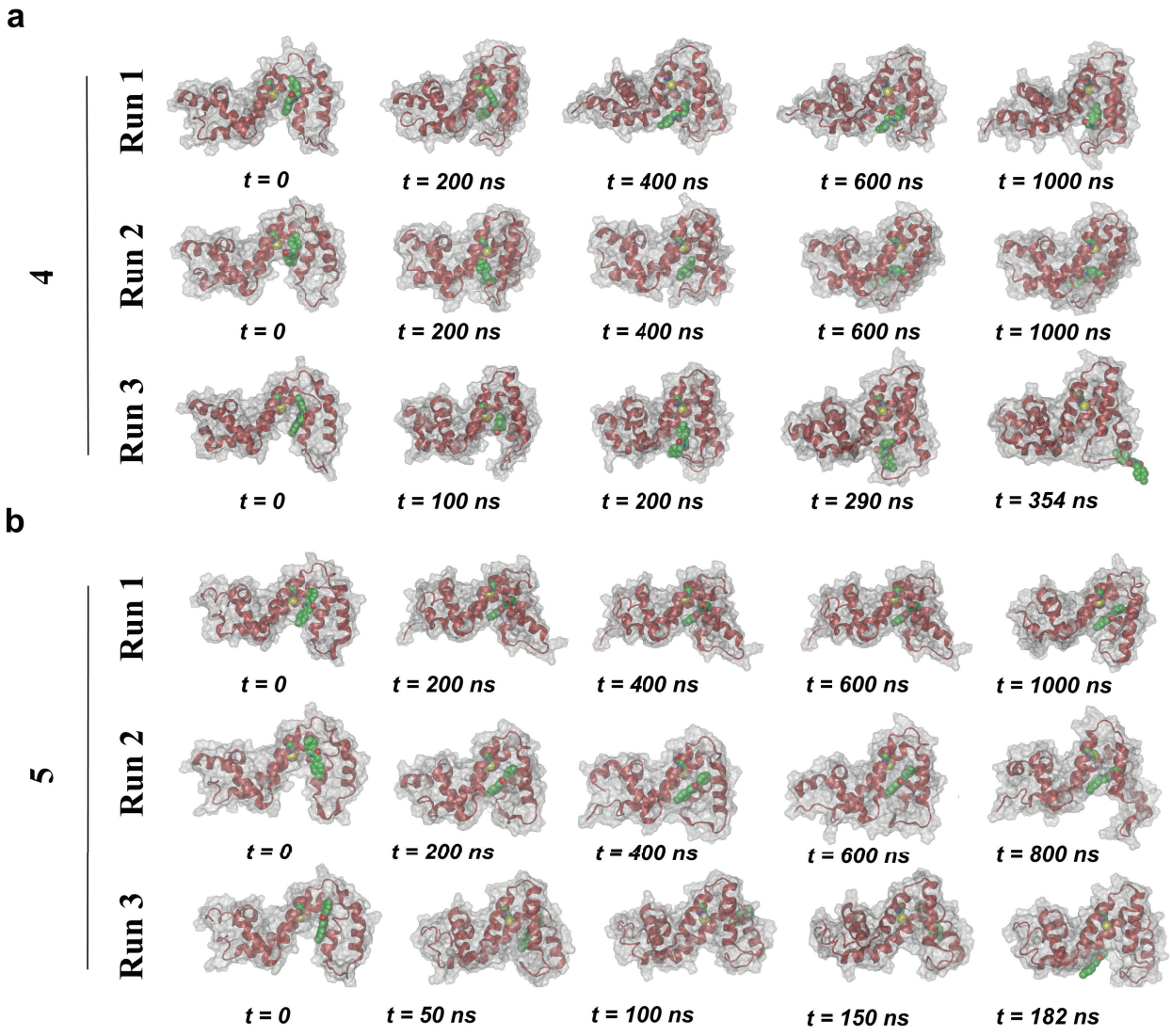



Figure F.4: Snapshots at various time-points for conformational evolution of complexes of non-TDZD compounds **4** (panel a) and **5** (panel b) with RGS4 (Model 1). In each panel, snapshots from three independent simulation runs are shown for each compound. Coloring and labeling schemes are identical to initial states shown in Fig. 4.3.

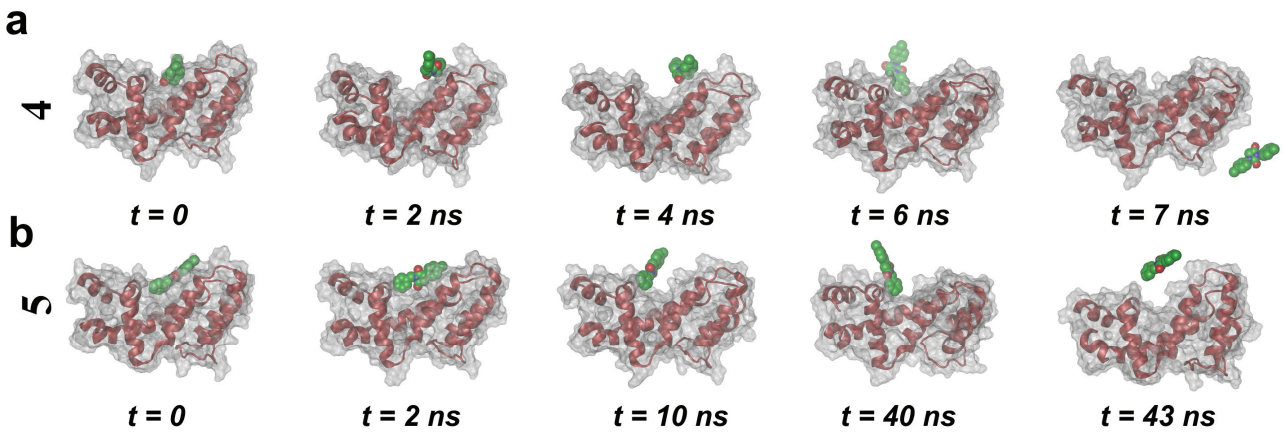



Figure F.5: Snapshots at various time-points for conformational evolution of complexes of non-TDZD compounds **4** (panel a) and **5** (panel b) with RGS4 (Model 2). Coloring and labeling schemes are similar to Fig. F.4.

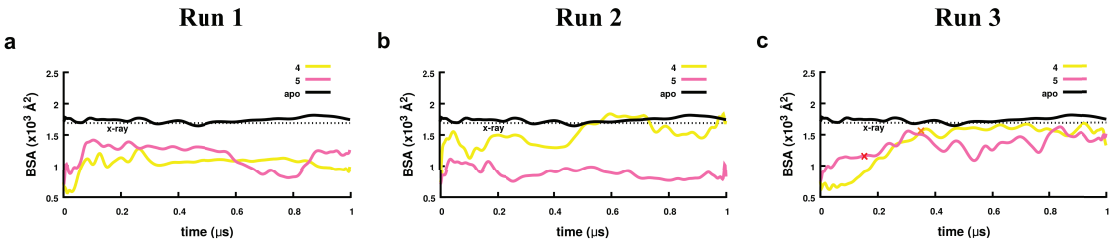



Figure F.6: The traces of buried surface area (BSA) between the $\alpha_5$-$\alpha_6$ helices and the rest of RGS4 *vs.* simulation time ($\mu$s) are shown from three independent simulation runs for Model 1. The BSA traces are shown for two non-TDZD compounds (magenta and yellow traces) and from a simulation of apo-RGS4 (black traces). The dotted horizontal line in each panel highlights the BSA-value in the crystal structure of RGS4 (PDB: 1AGR). The symbols ($\times$) on the BSA traces mark the locations of time-points in Run3 of each compound (panel c) after which compounds diffuse out of the binding pocket.

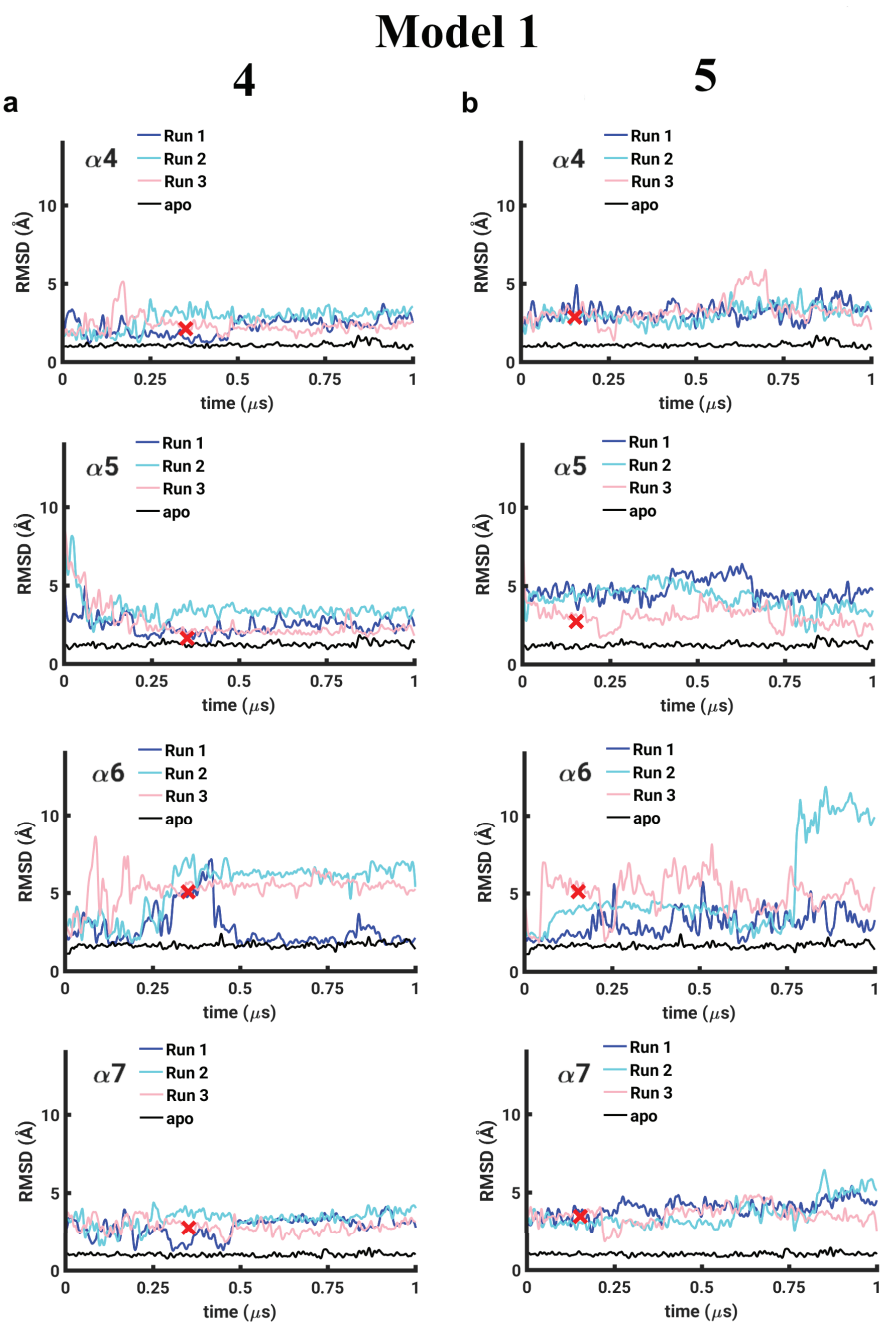



Figure F.7: The RMSD data similar to Fig. F.1 are shown for non-TDZD compounds **4** and **5** from three independent simulations. The red symbol (×) marks the locations of time-points in Run3 of each compound after which compounds diffuse out of the binding pocket (see snapshots in Fig. F.4).

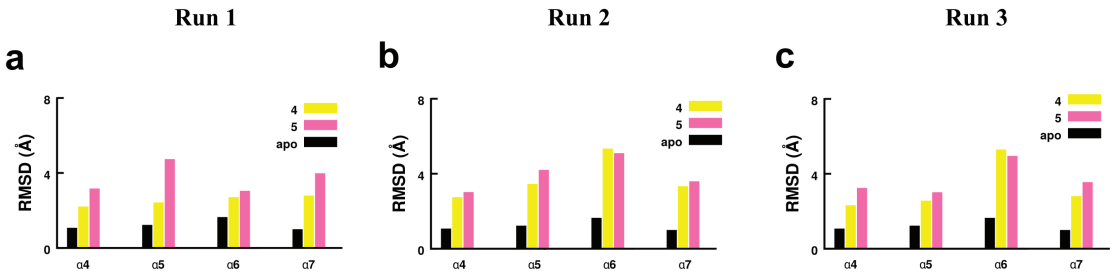



Figure F.8: The histograms of RMSD-averages computed based upon data from each run in Fig. F.7 are shown.

# APPENDIX G

# SUPPORTING INFORMATION FOR STUDIES REPORTED IN CHAPTER 5

This appendix provides additional figures related to Chapter 5.

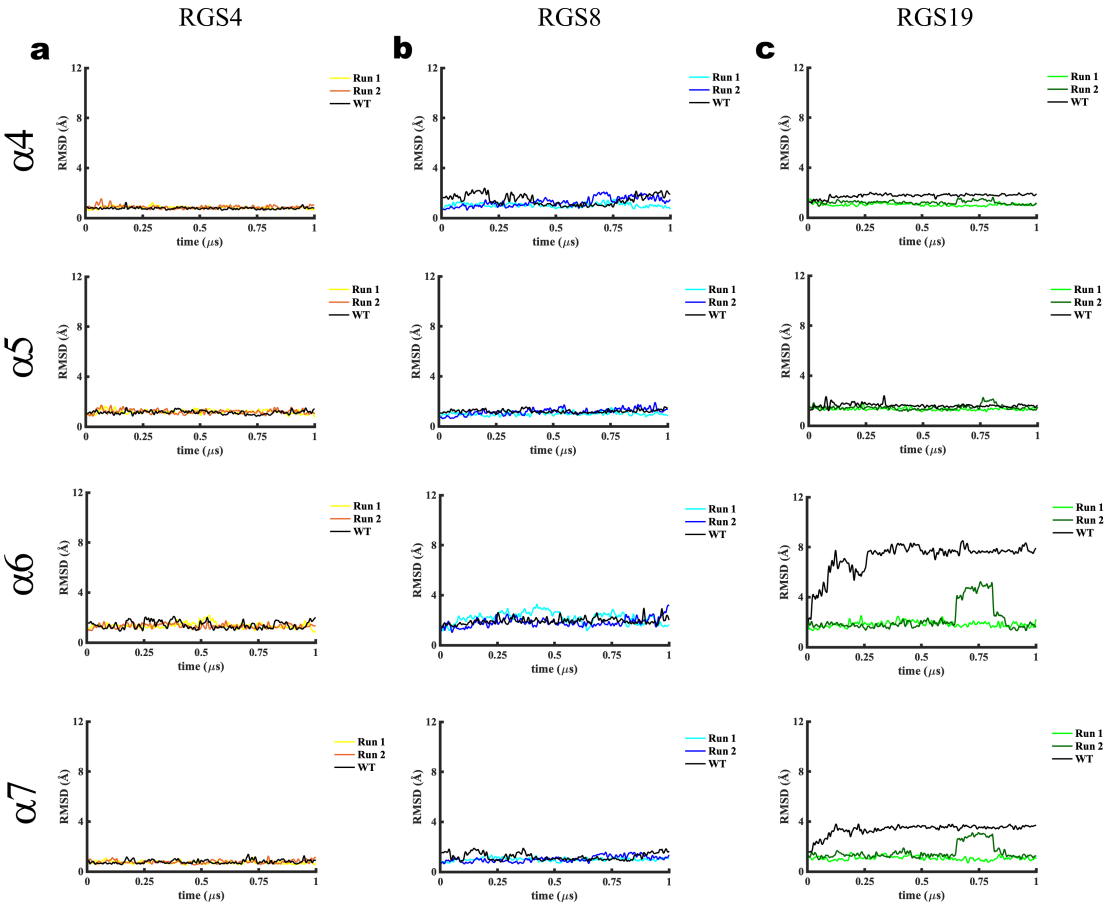



**Figure G.1.** The traces of root-mean-squared-deviation (RMSD) vs. simulation time ($\mu$s) for (a) RGS4 D90L, (b) RGS8 E84L, and (c) RGS19 L118D.Two independent simulation runs for each structure are presented, and the wild-type runs are presented from our previous work [13].

| | α4-α5 | | % of sim within 4 Å | α5-α6 | | % of sim within 4 Å | α6-α7 | | % of sim within 4 Å | CCG-50014 $IC_{50}$ (μM) |
|---|---|---|---|---|---|---|---|---|---|---|
| RGS4 | D90 | K125 | 58.7 | S120 | S138 | - | D130 | K155 | 31.5 | 8.5 |
| RGS8 | E84 | R119 | 44.2 | D114 | R128 | 47.5 | D124 | K149 | 36.1 | >1000 |
| RGS19 | L118 | K153 | - | S148 | N166 | - | D158 | Q183 | - | 1.1 |

**Figure G.2.** The salt-bridge interaction within the $\alpha$4-$\alpha$7 bundle of helices in single-cysteine structures of RGS4, RGS8, and RGS19 from MD simulations and potency of CCG-50014 inhibition of single-cysteine RGS proteins from our previous work [13] (This figure was created by my co-author Vincent Shaw).

# APPENDIX H

# SUPPORTING INFORMATION FOR STUDIES REPORTED IN CHAPTER 6

This appendix provides additional figures related to Chapter 6.

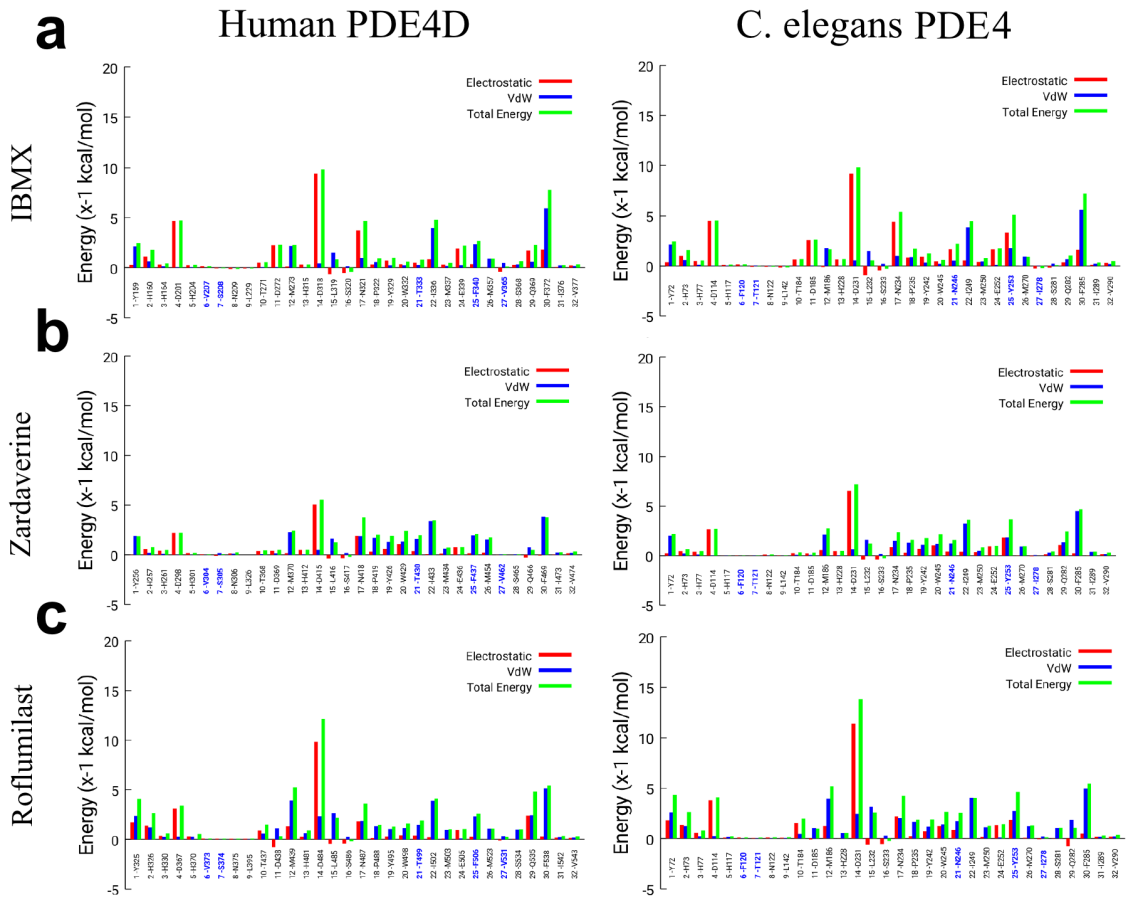



**Figure H.1.** The nonbonded interaction energy analysis between residues in the inhibitor binding pocket of PDE4D and *C. elegans* PDE4 for the first simulation run. See Fig. 6.1a and 6.1b for depictions of the binding pocket and the 32 residues analyzed with bound (a) IBMX, (b) zardaverine, and (c) roflumilast. Amino acid residues in blue text denote residues that differ between human and *C. elegans* PDE4 sequences.

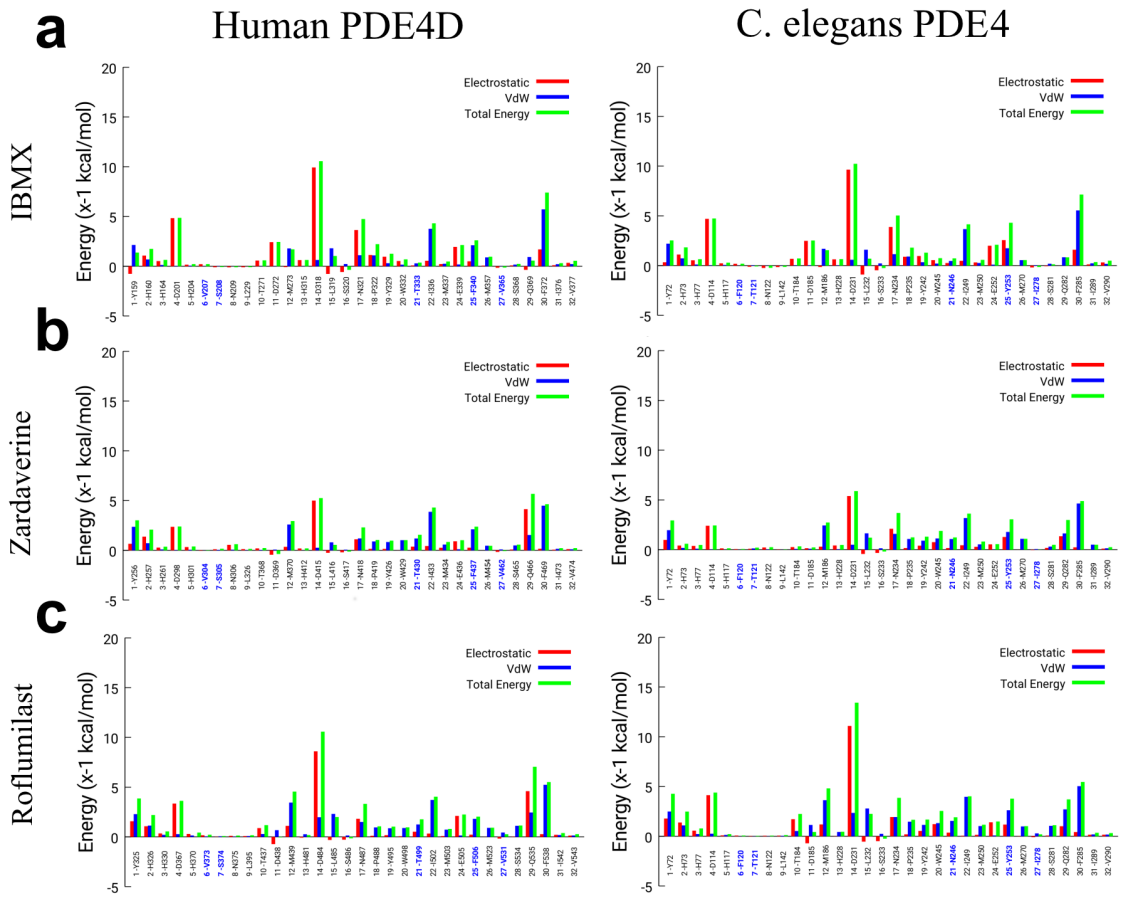



**Figure H.2.** The nonbonded interaction energy analysis between residues in the inhibitor binding pocket of PDE4D and *C. elegans* PDE4 for the second simulation run. (a) IBMX, (b) zardaverine, and (c) roflumilast. Amino acid residues in blue text denote residues that differ between human and *C. elegans* PDE4.

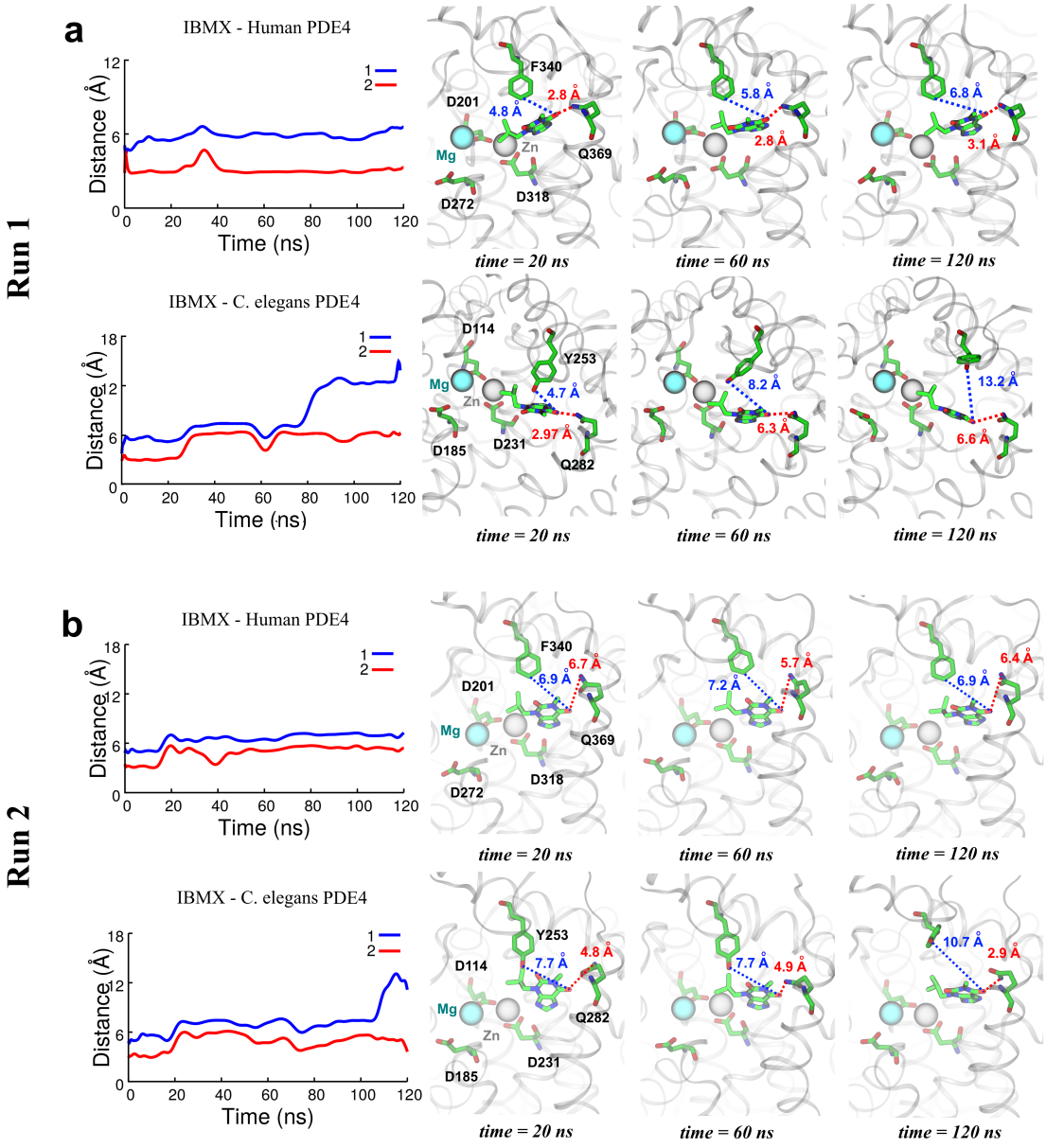



**Figure H.3.** Interatomic distances between C4 atom of F506(human)/O atom on the side chain of Y253(*C. elegans*) (blue dashed line, labeled 1) or the N$\delta$ atom of Q369/Q282 (red dashed line, labeled 2) and the O6 oxygen of IBMX bound to human PDE4D or *C. elegans* PDE4 obtained from two independent MD simulation runs, (a) run 1 and (b) run 2.

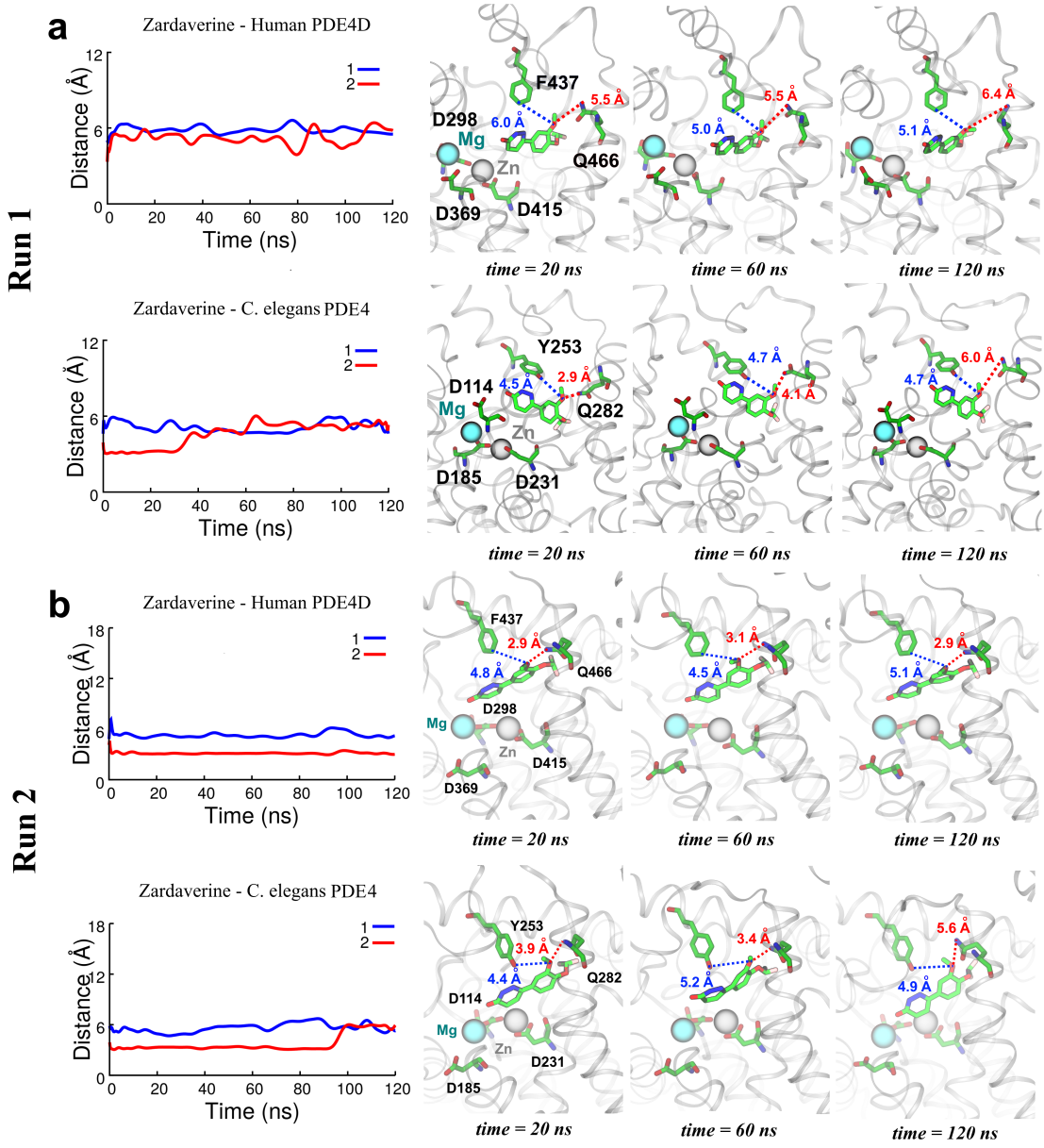



**Figure H.4.** Interatomic distances between C4 atom of F506(human)/O atom on the side chain of Y253(*C. elegans*) (blue dashed line, labeled 1) or the N$\delta$ atom of Q369/Q282 (red dashed line, labeled 2) and the O4 oxygen of zardaverine bound to human PDE4D or *C. elegans* PDE4 obtained from two independent MD simulation runs, (a) run 1 and (b) run 2.

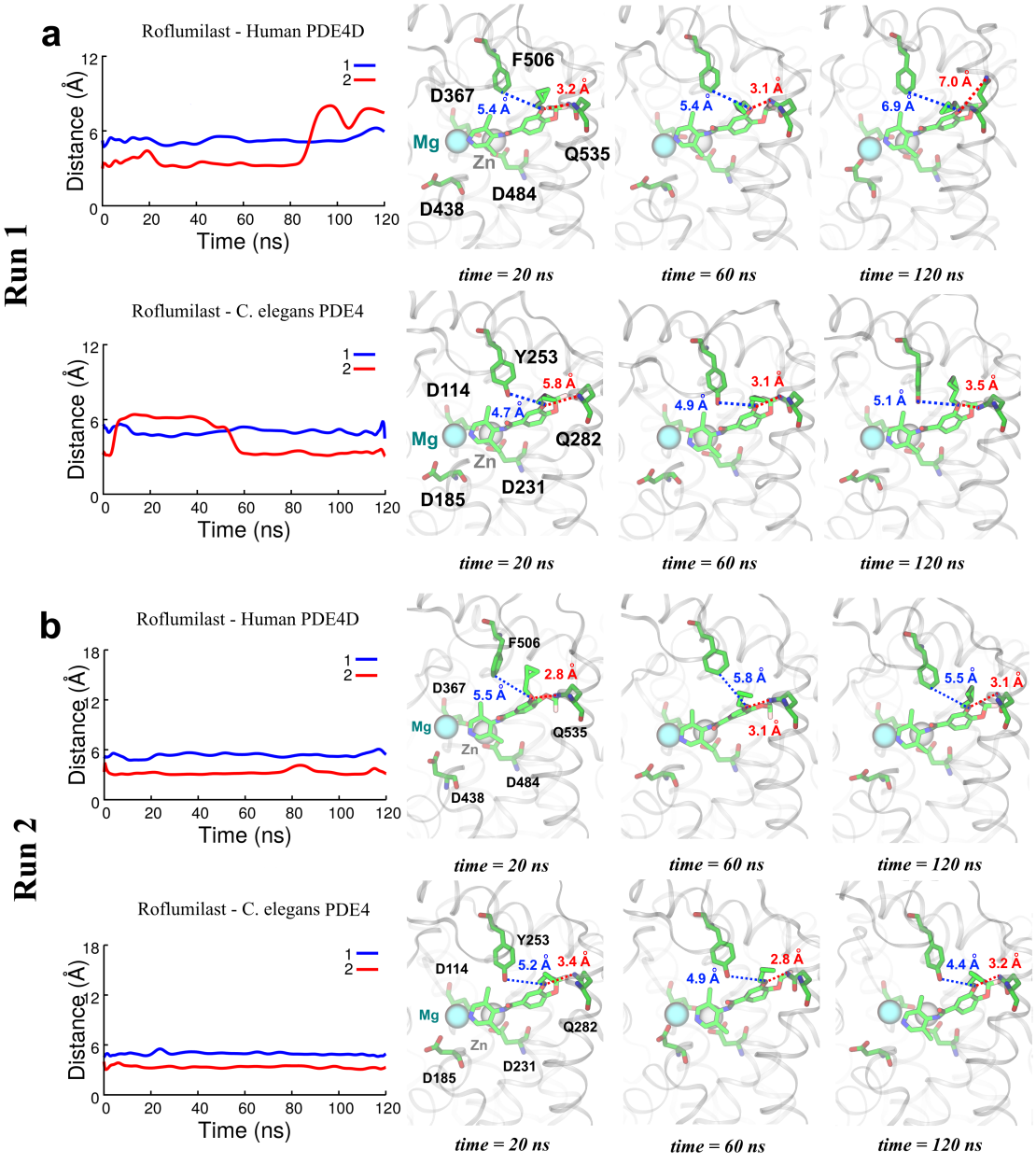



**Figure H.5.** Interatomic distances between C4 atom of F506(human)/O atom on the side chain of Y253(*C. elegans*) (blue dashed line, labeled 1) or the N$\delta$ atom of Q369/Q282 (red dashed line, labeled 2) and the O4 oxygen of roflumilast bound to human PDE4D or *C. elegans* PDE4 obtained from two independent MD simulation runs, (a) run 1 and (b) run 2.

# APPENDIX I

# JOURNAL COVER IMAGE

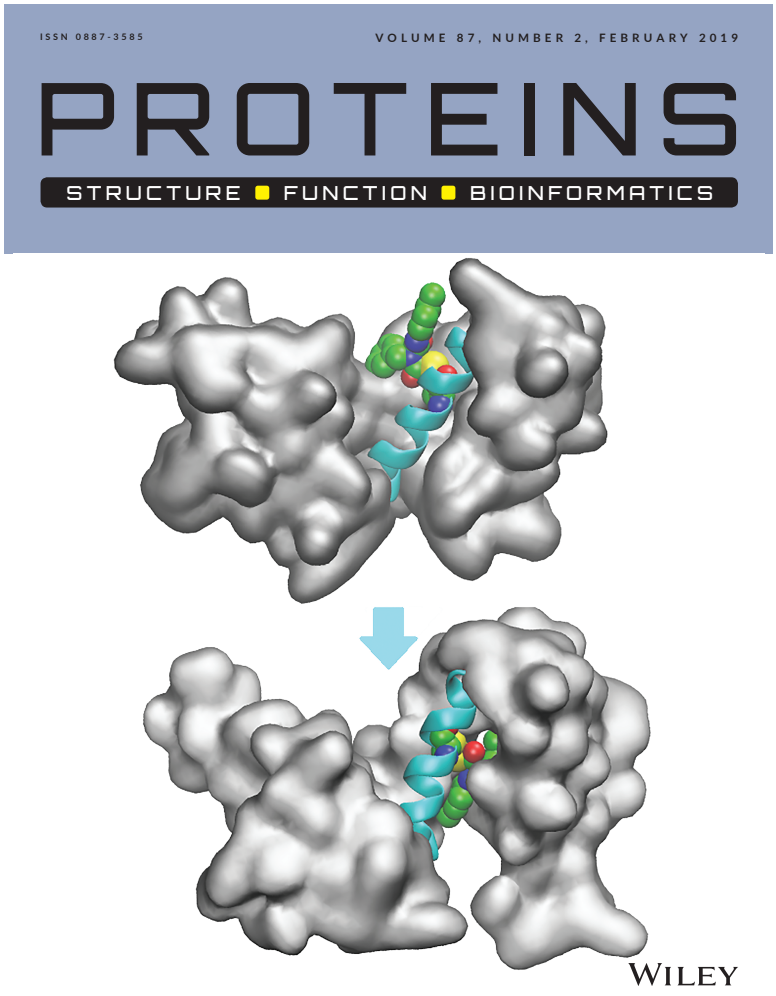



**Figure I.1.** Shown is the journal cover image [17] related to work described in chapter 4 and in Ref. [14].

# APPENDIX J

# CURRICULUM VITAE

## J.1 Education

**University of New Hampshire**, Ph.D., Chemical Engineering, May 2019

**University of Southern California**, M.Sc. in Chemical Engineering, May 2013

**Amirkabir University of Technology**, B.Sc. in Chemical Engineering, August 2010

## J.2 Journal Publications

Shaw, V.S., **Mohammadi, M.**, Quinn, J.A., Vashisth, H., and Neubig, R.R. (20XX). "An interhelical salt bridge controls flexibility and inhibitor potency for regulators of G-protein signaling (RGS) proteins 4, 8, and 19." *Molecular Pharmacology*, (Submitted).

Schuster, K.D., **Mohammadi, M.**, Cahill, K.B., Matte, S.L., Maillet, A.D., Vashisth, H., and Cote, R.H. (2019). "Pharmacological and molecular dynamics analyses of differen- ces in inhibitor binding to human and nematode PDE4: implications for management of parasitic nematodes." *PloS ONE* 14(3): e0214554. (PDF)

**Mohammadi, M.**, Mohammadiaran, H., Shaw, V.S., Neubig, R.R., and Vashisth, H. (2019). "Interplay of cysteine exposure and global protein dynamics in small-molecule recognition by a regulator of G-protein signaling protein." *Proteins: Structure, Function, and Bioinformatics*, 87(2), 146-156 (Featured on the Journal Cover). (PDF)

Liu, Y., **Mohammadi, M.**, and Vashisth, H. (2018). "Diffusion networks of CO in FeFe-Hydrogenase." *The Journal of Chemical Physics* 149(20), 204108. (PDF)

**Mohammadi, M.**, and Vashisth, H. (2017). "Pathways and thermodynamics of oxygen diffusion in [FeFe]-Hydrogenase." *The Journal of Physical Chemistry B*, 121(43), 10007-10017. (PDF)

## J.3 Conference Publications

Shaw, V.S., Quinn, J.A., **Mohammadi, M.**, Vashisth, H., and Neubig, R.R. (2019). "A salt bridge between $\alpha 4$ and $\alpha 5$ helices drives differences in flexibility and potency of inhibition

among regulator of G-protein signaling (RGS) proteins" *The FASEB Journal 33:1_supplement 784.16*, Experimental Biology 2019 Meeting, Orlando, FL, USA.

Shaw, V.S., Quinn, J.A., **Mohammadi, M.**, Vashisth, H., and Neubig, R.R. (2018). "A Salt Bridge Between $\alpha 4$ and $\alpha 5$ Helices Drives Flexibility and Potency Differences Among Regulator of G-protein Signaling (RGS) Proteins" *The* $19^{th}$ *Annual G-Protein Coupled Receptors Retreat*, Detroit, MI, USA.

**Mohammadi M.** and Vashisth H. (2018). "Computational studies of regulators of G-proteins signaling small-molecule inhibitors" *Abstracts of papers of the American Chemical Society*, $256^{th}$ American Chemical Society National Meeting, Boston, MA, USA.

Liu Y., **Mohammadi M.**, and Vashisth H. (2018). "Network of inhibitor diffusion pathways in hydrogenase enzymes" *Abstracts of papers of the American Chemical Society*, $256^{th}$ American Chemical Society National Meeting, Boston, MA, USA.

Liu Y., **Mohammadi M.** and Vashisth H. (2018). "Diffusion Network of CO in [FeFe]-Hydrogenase" *Proc.* $78^{th}$ *Annual Physical Electronics Conference*, Durham, NH, USA.

**Mohammadi, M.** and Vashisth, H. (2018). "Molecular dynamics studies of regulators of signaling proteins inhibitors" *University of New Hampshire Bioengineering Symposium*, Durham, NH, USA.

**Mohammadi, M.** and Vashisth, H. (2018). "Molecular dynamics studies of allosteric inhibitors for protein-protein interfaces" $20^{th}$ *Annual Northeast Student Chemistry Resea- rch Conference*, Boston, MA, USA.

**Mohammadi M.** and Vashisth H. (2017). "Mapping of gas diffusion pathways in [FeFe]-Hydrogenase" *Proc. 2017 American Institute of Chemical Engineers Annual Meeting*, Minneapolis, MN, USA.

**Mohammadi, M.** and Vashisth, H. (2017). "Mapping of gas diffusion pathways in hydrogenase enzymes" $19^{th}$ *Annual Northeast Student Chemistry Research Conference*, Boston, MA, USA.

## J.4 Honors and Awards

| | |
|---|---|
| Best Poster Award, University of New Hampshire Bioengineering Symposium | Spring 2018 |
| Anton Supercomputer Workshop Invitee, Pittsburgh Supercomputing Center | Fall 2017 |
| Member, Golden Key International Honor Society | Fall 2015 |