8-9-2016

Establishing the Relationship Between Function and Dynamics Within the Gated Mechanism of D-arginine Dehydrogenase

Michael Souffrant
msouffrant1@gsu.edu

ESTABLISHING THE RELATIONSHIP BETWEEN FUNCTION AND DYNAMICS WITHIN THE GATED MECHANISM OF D-ARGININE DEHYDROGENASE

by

MICHAEL SOUFFRANT

Under the Direction of Donald Hamelberg, PhD

ABSTRACT

Enzymes are ubiquitous in biological systems. They catalyze chemical reactions and are involved in many biochemical processes. The enzyme of interest is *Pseudomonas aeruginosa* D-arginine dehydrogenase (PaDADH). This flavin-dependent enzyme is composed of approximately 375 amino acid residues and has a broad substrate specificity with D-amino acids. A water recognition motif, observed in roughly 1200 non-redundant protein data bank (PDB) structures, was revealed to be embedded near the active site of PaDADH. This motif coincides with the conformational changes of the enzyme's gated mechanism. Molecular dynamics simulations were carried out to study the gated properties and structural characteristics of PaDADH in solution. Single amino acid mutations were undertaken to further understand the dynamics of the gated mechanism of this enzyme. In addition, $pK_{a,shift}$ analyses were evaluated to probe for the basic catalytic amino acid residue that is suggested to trigger the catalytic mechanism of PaDADH.

INDEX WORDS: Molecular Dynamics, D-arginine dehydrogenase, Water Recognition Motif, Gated Mechanism, Single Amino Acid Mutations, Basic Catalytic Amino Acid.

ESTABLISHING THE RELATIONSHIP BETWEEN FUNCTION AND DYNAMICS WITHIN THE GATED MECHANISM OF D-ARGININE DEHYDROGENASE

by

MICHAEL SOUFFRANT

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science

in the College of Arts and Sciences

Georgia State University

2016

ESTABLISHING THE RELATIONSHIP BETWEEN FUNCTION AND DYNAMICS WITHIN THE GATED MECHANISM OF D-ARGININE DEHYDROGENASE

by

MICHAEL SOUFFRANT

Committee Chair: Donald Hamelberg

Committee: Stuart Allison

Giovanni Gadda

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

July 2016

DEDICATION

To

Michelle & George Souffrant

My dear family members who brought me into this earth and gave me the opportunities necessary to produce such wonderful work.

ACKNOWLEDGEMENTS

Prior to joining Dr. Hamelberg's lab, I was clueless in the research field of computational biophysical chemistry. My original intent was to attend medical school and start a career in health. During my time in his lab, not only I gained much further insight on some of the theoretical approaches of computational chemistry, but I also expanded my understanding in high computing applications. Dr. Hamelberg readily applies this concept of "self-motivation", which enabled me to discover more in regards to my personality and the type of scientist that I wanted to eventually become. He allowed me to freely indulge into my scientific ideas, and use his resources at moment's notice. Thank you for your assistance, patience, and the daily morning lectures which consistently have a significant influence on my life's outcome.

I also wanted to send many thanks to Arghya Barman, a post-doctoral fellow in our lab, for his support and aid on a day-to-day basis throughout my research assignments. Ever since his presence, my comprehension and abilities in the biophysical research field had evolved tenfold. Thank you for your encouragements and your willingness to help with no remorse.

In addition, thank you Dr. Gadda, and Dr. Allison for deciding to be part of my thesis committee on such short notice. Your critiques and philosophical approaches in regards to my work and my knowledge as a scientist is greatly appreciated.

Likewise, I also want to send many thanks to my lab mates for taking the time to criticize my work and becoming my support group in times of need. I truly believe that we are a team in which we compensate for each other's strengths and weaknesses in several ways. Aside from computational chemistry, my dearest memories will be based on our wonderful times together. Thank you once more for having a hand in respect to the type of scientist I have become.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

1.1 Enzymatic Properties of PaDADH

Pseudomonas aeruginosa D-arginine dehydrogenase (PaDADH) is a flavin-depdendent enzyme which converts D-amino acids into their corresponding imino acids.[1-3] Typically, an enzyme has one defined substrate according to its biological function, however, PaDADH has the ability to catalyze every D-amino acid at different efficiencies with the exeption of D-aspartate, D-glutamate, and glycine. An X-ray crystal structure of this enzyme was reported at 1.06 Å resolution, in which case, two conformers of certain residues were established.[1] The ligand-free configuration, with a 70% occupancy in electron density, was denoted to be the open conformation of the enzyme, whereas the product-bound conformer in complex with iminoarginine, with a 30% occupancy in electron density, was designated to be the close conformation based on a tyrosine over the active site. Consequently, the notion of a gate which enhanced substrate binding, enabled product release, and played a role in the broad substrate specificity of PaDADH became plausible. Using the molecular dynamics (MD) approach to study the two conformational states of the gated mechanism of this enzyme and reaching convergence with kinetic assessments became the objective of this investigation. Such study may help into gaining further insight on the catalytic properties of other biological protein structures.

Upon detailed inspection of PaDADH, a water recognition motif, seen in 1200 non-redundant Protein Data Bank (PDB) structures,[4] was determined. Furthermore, this motif was established to be the gate involved in the catalysis of the enzyme. Properties of the motif, in other non-redundant proteins, were suggested to create an electrostatic environment which would enhanced water binding.[4] As proposed, a water molecule was found to be localized through similar secondary characteristics in the motif structure of PaDADH.[1] Such finding not only

generated an interest to further understand the water recognition motif structure, but also, the significant role of the gated mechanism in catalysis. MD methods were applied on the wild type enzyme structure to determine a correlation between the gated mechanism, and the water recognition motif. Single amino acid mutants of PaDADH were simulated to probe the different mechanical properties of the gated mechanism of the enzyme. Thus, motif properties, MD simulations, and kinetic measurements were explored to gain more insight on the strutural components of the two conformers of PaDADH in catalysis.

PaDADH was suggested to oxidize D-amino acids into their corresponding imino acids through the removal of the backbone amino proton which caused electrons to delocolize and lead to a hydride transfer from the alpha carbon to the FAD N5 atom.[1-3] The proton removal of the backbone amino group of the substrate was proposed to be the triggerring step of the oxidative deamination reaction. Thus, a catalytic residue or an ensemble of residues in the active site of the enzyme with acid-base properties was suggested to be the site of deprotonation of the amino proton.[5] $pK_{a,shift}$ calculations were carried out via thermodynamic integration calculations, in which case, electrostatic potentials were evaluated to assess solvent, and protein contributions on the proton affinity of several catalytic residues. By establishing the site(s) of deprotonation in PaDADH another component of its broad substrate specificity may be revealed, where the acid-base properties of the enzyme is influenced by its interaction with different D-amino acids rendering a variety of catalytic efficiencies.

1.2 Molecular Dynamics

Molecular dynamics (MD) is a statistical mechanics method that takes into account the microscopic state ensemble to describe the physical quantities of a system at the macroscopic level. The concept of molecular dynamics evolved from the dynamics of liquids shown by Alder, Wainwright, and Rahman during the years of 1950s and 1960s.[6] As the years passed, both the theoretical aspect of the method and computing technology have greatly improved rendering a viable tool for many areas in research. For example, MD simulations have been used to study different macromolecules such as enzymes, ligand binding receptor proteins, or nucleic acids.

Molecular dynamics methods can be separated into two main categories known as classical and quantum. The classical statistical mechanics approach to MD simulations treats the molecules as the 'ball and stick' model. The ball is a representation of a particular atom whereas the stick is an analogy of a bond between two atoms.[7] The quantum or first-principles statistical mechanics approach to MD simulations takes into account the presence of electrons, however the dynamics of a typical nucleus with its inner electrons is classically integrated. Quantum MD simulations show significant improvements in comparison to the classical approach, however, the method is computationally expensive.[8] Thus, classical MD became the more practical approach for the simulation of large complex systems.

X-ray crystallography or nuclear magnetic resonance (NMR) are amongst the several experimental techniques that provide a detailed three dimensional structure of systems of interest.[9] As a result, computational simulations are a mixture of both experimental (X-ray, NMR) and theoretical (MD) aspect of research.[6] Analyzing the ensemble of a complex system through computer modelling is faster, less expansive, and sometimes provide more atomistic

details as opposed to synthesizing the macromolecule and characterizing its properties, hence, the increase of high computing applications for drug design in pharmaceutical companies.

1.2.1 Molecular dynamics algorithms

The molecular dynamics approach follows the atomic force field model which defines a complex system as a set of atoms held together by interatomic forces.[6] If one was to evaluate the potential energy of N interacting atoms as a function of their positions $U(r_1, \dots, r_N)$, by taking the gradient (∇) of such function with respect to the displacements (xyz coordinates) of the atoms, the force acting upon each atom would be determined as shown in equation 1.[10]

$$F_i = -\nabla_{ri} U(r_1, \cdots, r_N) = -\left(\frac{\partial U}{\partial x_i}, \frac{\partial U}{\partial y_i}, \frac{\partial U}{\partial z_i}\right) \quad (1)$$

Electrons delocalize while being shared by many different nuclei, which in turn displaces into different electric field environments. Thus, it is possible to determine the dynamics of the nuclei without explicitly considering the electrons, which is feasible by considering the potential energy surface. The physics and chemistry of MD simulations are based on the different potentials involving each atom of a particular system.[11]

$$U(r_1, \cdots, r_N) = \sum_{bonds} \frac{a_i}{2}(l_i - l_{i0})^2 + \sum_{angles} \frac{b_i}{2}(\theta_i - \theta_{i0})^2 + \sum_{torsions} \frac{C_i}{2}[1 + \cos(nw_i - \gamma_i] + \sum_{atoms\ pairs} 4_{\varepsilon ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \sum_{atom\ pairs} k\frac{q_i q_j}{r_{ij}} \quad (2)$$

The terms as written in equation 2 correspond to the bond length, bond valence angle, bond dihedral angle, non-bonded van der Waals, and non-bonded electrostatic potentials. Considering Newton's second law of motion, the force acting on a particular atom is equivalent to that atom's mass multiplied by its change in acceleration as define in equation 3.[12] The first

two terms establish the energies generated during the deformation of bond lengths (l_i) and bond angles (θ_i) from their equilibrium values (l_{i0}) and (θ_{i0}). a_i and b_i represent the harmonic force constants which prevent modelling chemical changes. Rotations around the chemical bond is define by the third term, in which case, n and c_i represent the periodicity and heights of rotational barriers respectively. The attractive and repulsive interatomic van der Waals forces in the form of the Lennard-Jones 12-6 potential is defined by the fourth term. Coulomb electrostatic potential is denoted in the fifth term where q_i and q_j represent the partial charges of atom i and j respectively, and r_{ij} expresses the relative distance between both atoms.

$$F_i = m_i \frac{d^2 r_i(t)}{dt^2} \qquad (3)$$

Upon solving for the force vector acting on each atom of a complex system, an algorithm is used in order to generate the MD trajectories. For instance, the Verlet algorithm is usually applied for computing simulations of macromolecules due to its simplicity and stability. Consequently, MD trajectories are described by position and velocity vectors, in which case, time is dependent on both the time step, and the total number of steps undertaken during simulation[6].

$$r_i(t + \Delta t) \cong 2r_i(t) - r_i(t - \Delta t) + \frac{F_i(t)}{m_i}\Delta t^2 \qquad (4)$$

As previously established, the sum in potential is used to generate the net force applied on a particular atom. Newton's second law of motion, and the Verlet algorithm (equation 4) are applied to calculate the change in position of that atom over time. Since the same approach is followed for each and every atom in a system, the change in conformation of a particular macromolecule in time is therefore governed by the fast motions of its atoms. For example, a typical bond between light atoms vibrate at the femtosecond scale, which implies that a sub-

femtosecond time scale should be used for highest accuracy within each integrated step since the sum in potential is based on the localization of electrons surrounding the atoms. Though less accurate, a femtosecond time step is more efficient and detailed enough to thoroughly study a system of interest. However, by continuously increasing the magnitude of time at the femtosecond scale the margin of error increases, hence, a time step of more than 5 femtoseconds would not be a feasible approach for the application of an MD simulation.[12]

1.2.2 Limitations of molecular dynamics

In order to consistently update the positions and velocities of each atom in the system, the sum of the potentials, and the forces must be recomputed at each step. Such aspect of MD trajectories has created a need for better algorithms to compensate for the long-range forces, and new explicit models that would enhance trajectory assessments.[6] For example, sampling methods of the TIP3P water model have decrease computational time of MD simulations more than three folds.[13] High temperature simulations, accelerated molecular dynamics (aMD), and replica exchange molecular dynamics (REMD) were all methods developed to improve sampling.[14] The particle meshed Ewald method enabled efficient computation of long-range interactions, in which case, atoms separated by a distance larger than a certain cut-off aren't neglected.[15]

Physical properties of thermodynamics are largely dependent on the idea of convergence, which is truly achieved when the entire free energy landscape of a system is sampled. Thus, reaching convergence is a challenging task since the integrated time step of a simulation is at the femtosecond scale. Furthermore, the boundaries of the free energy field are not defined, thus, it is currently not possible to conclude whether or not all potential wells are fully sampled within a trajectory.

Classical atomic force field sampling cannot be applied to gain further insight on the acid-base chemistry of electrons, redox reactions, or even proton tunneling because the approach is based on treating the atoms, their valence electrons, and bonds as the "ball and stick"model.[12] In theory, the atomic force field represents the forces that influence the atoms of a system under experimental conditions. Deriving empirical potential energies to be transferable and applicable to many systems of interests at various environmental parameters is challenging, and create a sense of abstraction in the field of computational chemistry. Even through the method is combining experimental data with theoretical calculations, there still exist a need for improvement of the atomic force fields, which would enhance higher timescale simulations, accuracy and stability.

Most biochemical processes occur at the microsecond or millisecond scale. In order to gain further understanding on the characteristics of a system, there has to be enough sampling at the femtosecond time scale to reach higher orders in time (microsecond, millisecond). Such approach has to take into account the atom count of a system since the integrated step is evaluated for each atom. Thus, processing simulations at the microsecond or millisecond time scale requires hundreds of millions or billions of steps, which creates limitations in both computation time and size of biochemical processes. MD methods are currently being explored to design better atomic force fields in order to generate simulations more efficiently.

1.2.3 Thermodynamic cycle

In general, enzyme catalysis involves the binding of a protein with its corresponding ligand(s). The free energy released, and relative affinities of ligands during enzyme-substrate interaction became critical in MD simulations. New method developments such as Docking,[16] can be evaluated through either the 'lock and key' model or the 'induced fit' model. The 'lock

and key' approach is more rigid, in which case, macromolecules are considered to be static analogous to a key opening a door. The 'induced fit' model, however, considers the flexibility of the ligand(s) and its corresponding enzyme.[16] During the formation of enzyme-substrate complex both structures changes in configurations for optimum free energy release. Due to the atomistic movements of the enzyme and substrate structures the 'induce fit' type models can now be investigated using MD simulations. However, the change in Gibbs's free energy related to an enzyme free in solution binding with different ligands is computationally expensive, because it requires the assessment of repeated trajectories, which show the stochastic interactions involved between the enzyme's active site and the various ligands. Nevertheless, calculating the change in Gibb's free energy of the enzyme-ligand complexes and the enzyme with different ligands free in solution is significantly less expensive computationally. Since ΔG is a state function, a thermodynamic cycle between these Gibb's free energy properties can be established, in which case, the relative binding affinities of various ligands to the same protein is evaluated.

1.2.4 Thermodynamic integration method

The thermodynamic cycle can be modified to calculate proton pK_a shifts of a particular model isolated in solution in comparison to the changes of its environment in a complex structure. In general, the interactions between protein residues not only depends on their relative distances, but also their positions, including their proximity to solvent. The evaluation of pK_a shifts can be used as a tool to elucidate protein electrostatics. Once an enzyme complex partially folds or unfolds it causes significant effects on proton binding or release. In addition, the integrity of an enzyme correlates with the pH of its environment.[17]

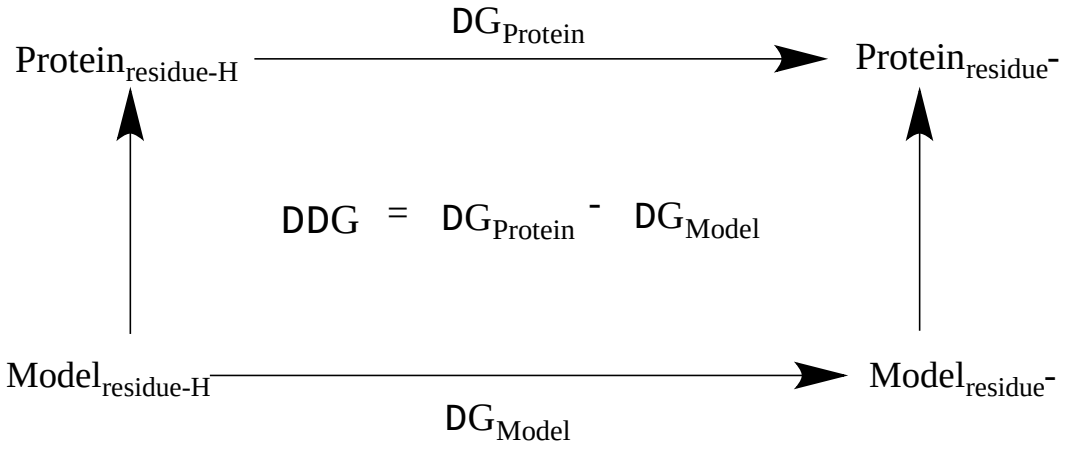


Figure 1.1 Thermodynamic cycle involving proton binding and release.
The change in Gibb's free energy of the deprotonation process of a protein residue is compare to the deprotonated spontaneity of that same residue free in solution. The double free energy difference is denoted as $\Delta\Delta G = \Delta G_{protein} - \Delta G_{model}$.

The molecular dynamics approach gives an ensemble of configurations, which would reflect both protein and solvent structures with their electrostatic and non-electrostatic effects. During a free energy simulation, the system of interest is slowly transformed from an initial state A to a final state B free in solution and within the microenvironment of the enzyme. The energy function U gradually changes with a 'coupling parameter' λ, where λ varies from 0 to 1 in correspondence with U_A to U_B, respectively (equation 5). An intermediate value of λ would be in respect to an hybrid system, that is a mixture of A and B in such way that λ would be greater than 0 but less than 1.[17]

$$U(\lambda) = (1 - \lambda)U_A + \lambda U_B \qquad (5)$$

Removing the proton partial charge by increments causes a change in its electrostatic potential. Subsequently, the average values of the electrostatic potentials are evaluated at their respective λ values, which are then integrated to generate the total change in Gibbs free energy.

$$\partial U/\partial\lambda\,(\lambda) = \left\langle \frac{\partial U}{\partial \lambda} \right\rangle_{\lambda} = \langle U_B - U_A \rangle_{\lambda} = \sum_i \delta q_i \langle V_i \rangle \lambda \qquad (6)$$

The large brackets in regards to equation 6 above defines an ensemble average related to U(λ). The charge increments are written as δqi where *i* is the proton of interest going from state 0 to 1. The proton's electrostatic potential is established as V_i and the sum operator sign is the integrated value of δqi<Vi>λ after being calculated from each lambda. The average quantity of U_B -U_A is defined as the energy required for a virtual ionization reaction while the system is fixed in its reference conformation, also known as the energy gap. By evaluating simulations for a set of λ values between 0 and 1, the Gibbs free energy change of the process can be estimated, hence the name 'thermodynamic integration'.

$$\Delta G = -kT log K_a \qquad (7)$$

$$pK_a = -log_{10} K_a = \frac{1}{2.303kT} \Delta G \qquad (8)$$

$$pK_{a,protein} = pK_{a,model} + \frac{1}{2.303kT} \Delta\Delta G \qquad (9)$$

The standard Gibbs free energy is related to the natural log of the acid dissociation constant through Boltzmann's constant 'k' and temperature 'T' (equation 7). By solving for the negative log of base 10 of the acid dissociation constant, the pK_a value of that reaction can be evaluated (equation 8). The difference between protein pK_a and model pK_a values ($pK_{a,\ shift}$), establishes the relationship with the double free energy difference, ΔΔG (equation 9).[17]

1.3 Motifs

During several structural analyses of enzymes, scientists began to notice repeated patterns in terms of the secondary structures of proteins. These secondary properties were called structural motifs, also known as super-secondary structures. The different conformation of motifs is based on the connectivity between secondary structure elements. For instance, helices and strands have a multitude of ways to interact with one another. Based on such perspective, there is a limited number of ways secondary structure elements of proteins can combine to be biologically viable. In addition, two macromolecules may share the same super-secondary structure but have dissimilar primary sequences.

1.3.1 Beta hairpin motif

One of the most common structural motifs revealed in proteins is known as the beta hairpin motif.[18,19] Due to further understanding of its three-dimensional structure, this super-secondary structure is also referred to as beta-ribbon or beta-beta unit. In its basic form, a beta hairpin is a super-secondary structure with two beta strands similar to the conformation of a hairpin.

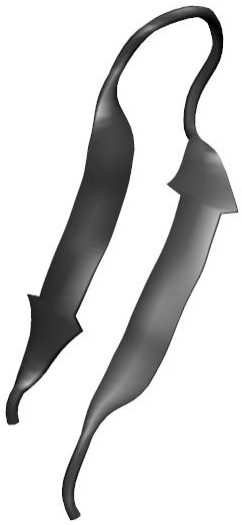


Figure 1.**2 Beta hairpin motif in *Pseudomonas aeruginosa* D-arginine dehydrogenase.**[1]

When present, the two strands are adjacent in primary sequence, however they are antiparallel in the secondary structure of the protein. Each antiparallel strand is then connected by a two to five amino acid loop (Figure 1.2). If a beta hairpin was removed from its protein, the structure will still form. When present, the beta hairpin motif is a necessity for the biological activity of the macromolecule.[18,19]

1.3.2 Greek key motif

Another common super-secondary structure which may be considered an extrapolation of the beta hairpin motif is the Greek key motif as shown in Figure 1.3.[20]

Figure 1.3 Greek key motif in *Pseudomonas aeruginosa* D-arginine dehydrogenase.[1]

This particular motif consists of four strands, in which case, they are all adjacent and antiparallel to one another. Similar to the beta hairpin motif, three antiparallel strands of the Greek key motif are linked amino acid loops. However, the first strand is connected to the fourth by an even longer loop. During protein folding and unfolding, it is common to see Greek key motifs fold and unfold when present.[20]

1.3.3 Zinc finger motif

Aside from these super-secondary protein structures there exist other common motifs, which involve ligand binding. For instance, the zinc finger, shown in Figure 1.4, is a relatively small motif with one or more zinc ions in a coordinated covalent bond with surrounding histidine and cysteine residues to stabilize the fold.

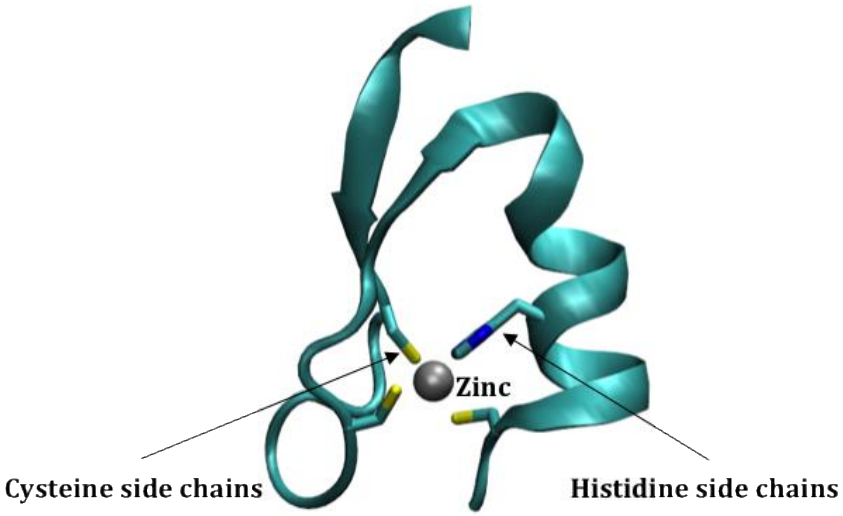


Figure 1.4 Zinc finger motif in Ligand of Numb Protein-X2.[21]

The interesting aspect in regards to this motif is that the presence of the metal is necessary for proper biological activity. For example, in 1983, *Xenopus laevis* transcription factor IIIA was shown to require coordinated ligand binding with zinc metal for proper regulatory activity.[22,23] Additionally, the zinc finger is also known to be a self-contain domain. Zinc fingers can be used to recognize nucleic acid sequences of various lengths; thus this particular motif design can be used to target specific genes.[24]

1.3.4 EF-hand motif

Another super-secondary structure that is most commonly found in calcium binding proteins is the EF-hand motif.[25] This particular motif is a helix-loop-helix structural domain that is sometimes present in the calcium binding protein family. Similar to the zinc finger, the calcium ion is coordinated to both helices included the core of the motif which involves a loop and a strand (Figure 1.5).

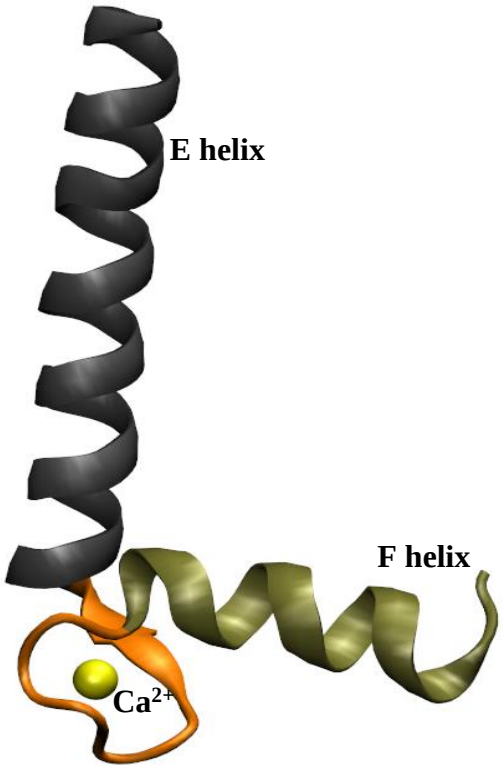


Figure 1.5 EF-hand motif in calmodulin.[26]

This specific motif is found within 66 subfamilies of calcium-binding proteins. The EF hand named was coined about 25 years ago by Kretsinger and Nockolds as a graphical representation of the calcium-binding motif in parvalbumin. In certain cases, the motif has been

shown to not only form coordination with calcium but also magnesium. This particular motif is very significant due to the usual calcium influx within the cytosol of a cell. Meaning, there is a diverse range of enzymes that contain this super-secondary structure which would enable them to bind to calcium in the cell cytosol for proper biological functions. For instance, signal transduction between cellular compartments and muscle contraction are mostly govern by the ability of protein structures to bind to calcium which in turn is related to their EF-hand motif.[25, 27]

1.3.5 Water recognition motif

Peptidyl-propyl cis-trans isomerase NIMA-interacting 1 (Pin1) is an enzyme (PPIase) which catalyzes the cis/trans isomerization of peptide bonds preceding a proline residue.[28,29] Pin1 is a Parvulin protein which has high affinity for the isomerization of the propyl ω-bond in the presence of phosphorylated serine (pSer) or threonine (pThr) as precedents of the propyl bond.[29-32] The enzyme has WW domain and a catalytic domain connected by a flexible loop.[4] NMR experiments have shown the catalytic and WW domains to be allosterically coupled.[33] Due to its catalytic abilities, Pin1 is a subcellular signaling pathway regulator that is involve in DNA repair, cell growth and division, apoptosis and transcription.[34]

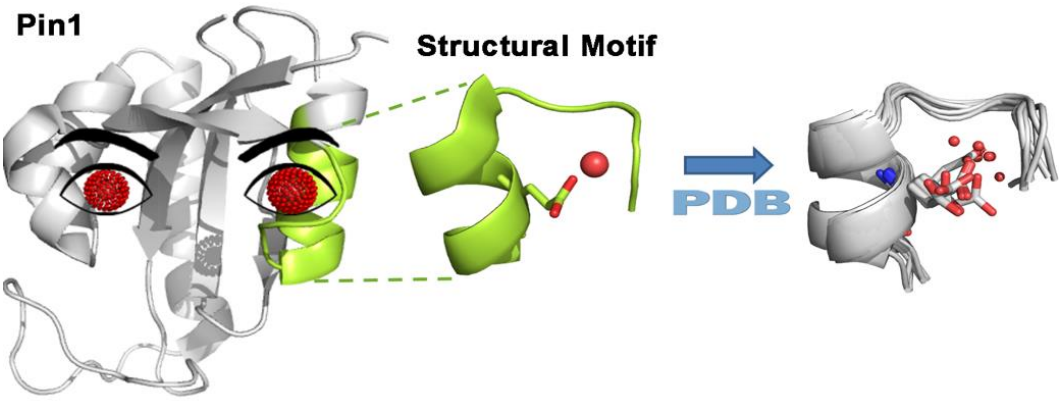


Figure 1.6 Localized conserved water molecules found in Pin1.[4]
Wat1 at site 1 was alluded to be highly localized in comparison to wat2 at site 2 which was suggested to exchange with the water molecules from the solvent. Wat2 at site 2 was found in a variety of non-redundant proteins and was characterized as a distinct water recognition motif.

During the analysis of several protein X-ray crystal structures, the localization of water molecules has been proven to be significant for protein integrity, stability, flexibility, and function.[4] Similar to the metals cofactors involved in coordinated covalent bonds, some of these water molecules must be undergoing some significant favorable interactions to render an enzyme biologically viable. Though the necessity of these localized water molecules is alluded, their role in terms of protein function remains unknown.

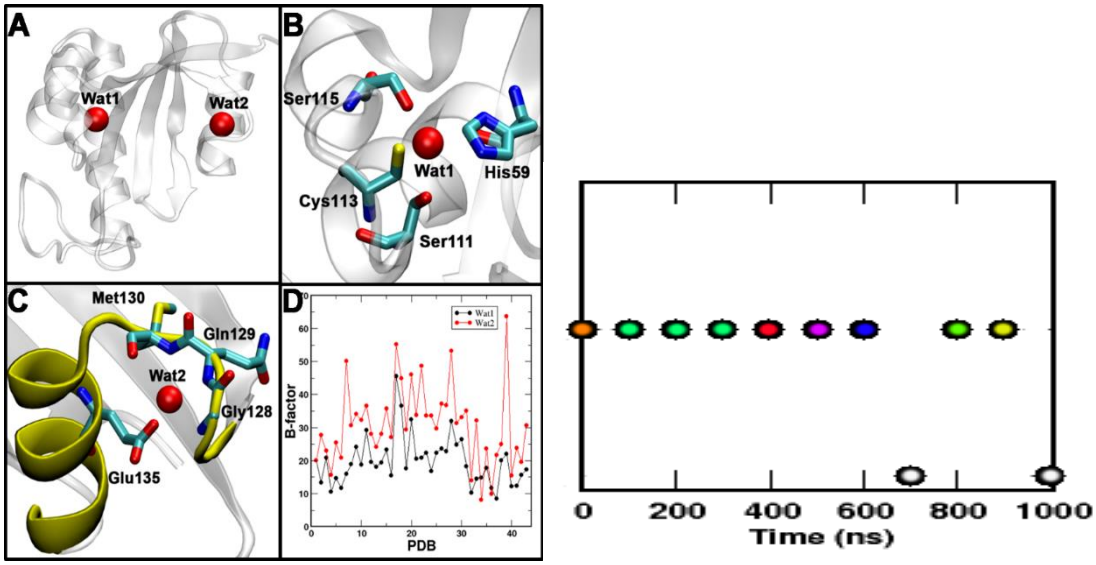


Figure 1.7 Analysis of Conserved Wat1 and Wat2 in Pin1.[4]

Panel A displays the location of site 1 and 2 with their corresponding waters in Pin1. Panel B shows the significant hydrogen bonding involved at site 1 with wat1, which enabled such high localization. Panel C indicates the distinct water recognition motif including the significant interactions with wat2 at site 2. Panel D establishes a comparison of the recorded B factor values of the different PDB structures of Pin1 in regards to wat1 and wat2 at their corresponding sites. In addition, the exchangeable characteristics of wat2 at site2 are displayed over time.

Recently two conserved localized water molecules had been revealed in Pin1. Wat1 was revealed to be located at site 1 and Wat2 at site 2. Wat1 was suggested to be highly localized in site 1 with a favorable binding free energy of -2.8 ± 0.3 kcal/mol. Wat2 showed high exchangeability with the bulk water solvent at site 2 with a lower favorable binding free energy of -0.7 ± 0.2 kcal/mol. In addition, a bioinformatics study was undertaken, in which case, non-redundant X-ray crystallographic protein structures were analyzed in the Protein Data Bank (PDB) using MaDCaT program,[35] and a residue probe from Pin1 consisting of 13 amino acids. The structural properties of the secondary structure of site 2 were found to be present in about 1200 non-redundant protein structures in the PDB. Out of the 1200, more than 880 protein structures had a resolution higher than 2.5 Å. A total of 16 non-redundant protein structures showed a glutamate residue occupying the same position on the helix turn in the presence of water. Out of those 16, 7 structures showed the water molecule to be properly buried. The

remaining 9 structures had water molecules surrounding the super-secondary structure or absent from the site. Wat1 formed hydrogen bonds with neighboring active site residues such as His59, Ser111, and Ser115. In addition, Wat1 is suggested to be involved in a critical interaction with Cys113 which is viable for protein function.[32,36-38] Meaning, in the absence of Wat1 the hydrogen bond interactions would be disrupted and Cys113 would not be properly aligned for proper enzymatic function. Wat2 is involved in interactions with the surrounding residues such as Gln129, Met130, and Glu135.

All of the crystal structures of human Pin1 were downloaded from the PDB, in which case, files with multiple copies of Pin1 were separated into individual copies for a total of 47 crystal structures. The Debye-Waller factors (B factors) of Wat1 and Wat2 were then recorded to compare their localization in their corresponding protein structure. Based on the analysis of Panel 1D (Figure 1.7), Wat1 generally appeared to be less mobile than Wat2. Furthermore, as shown from Figure 1.7 (Panel 2), Wat2 exchanged with the water solvents multiple times during a 1μs simulation.

1 **2**

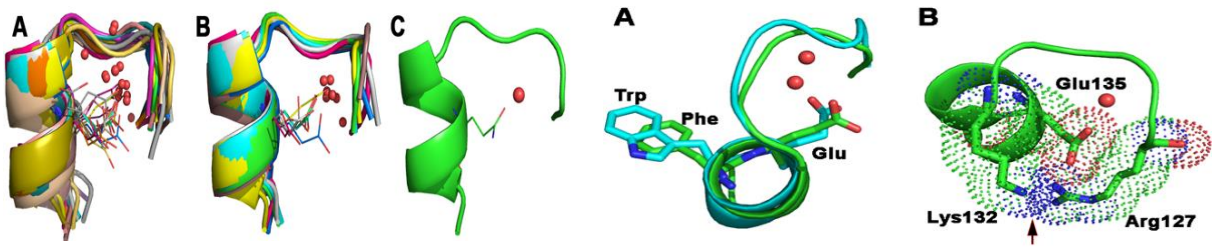


Figure 1.8 Unique properties of water recognition motif.[4]
1A displays 16 water recognition motifs that were found in non-redundant proteins in the PDB database that included a glutamate residue at a similar location. 1B shows 43% of the 16 water recognition motifs, in which case, the water molecule was properly buried. 1C represent the only water recognition motif found where a glutamine residue was present as oppose to a glutamate. 2A represents the superposition of the water recognition motif found in Pin1 (green) and FKBP (cyan) structures. Trp from FKBP and Phe in Pin1, including their relative glutamate residues, occupied the same position on the water recognition motif. 2B shows the highly electropositive environment created by Lys132, and Arg127 residues from Pin1. A strong favorable electrostatic interaction with Glu135 led to a vacant pocket within the motif for water to bind.

In general, features of this water recognition motif are suggested to be present in other non-redundant protein structures in the PDB. The establishment of water binding to specific secondary structures may become significant for protein engineering and drug design. It is suggested that the conformation and the electrostatic distribution of the water recognition motif increases the probability of a water molecule to be buried within this super-secondary structure. As a consequence, the water recognition motif is similar in construct to the EF hand, zinc finger, or other ligand binding motifs.

A water recognition motif akin to Pin1's site 2 secondary elements was observed in another PPIase,[39,40] FK506 binding protein (FKBP), with similar residues occupying the same positions on the site. This finding suggests the necessity of a particular distribution of electrostatic environment in order for a water molecule to bind in the pocket of the motif. Thourough analyses suggest that a single point mutation of glutamate to glutamine, would displace water within the motif pocket (Figure 1.10). As seem necessary, such hypothesis was

tested by mutating Glu to Gln in Pin1 followed by molecular dynamics simulations. As a result, Wat2 exchanged with the bulk solvent in less than 50 ns of simulation time. The side chain of Gln partially ocuupied site 2, which left no empty space for the water molecule to bind or re-enter the embedded pocket of the water recognition motif. The presence of Arg127 and Lys132 in Pin1 seems to create an electropositive environment, which interacted with the electronegative glutamate side chain (Glu135) of the PPIase. This created a vacant pocket for a water molecule to bind or exchange with solvent within the water recognition motif. During the mutation of glutamate to glutamine, the interaction with arginine or lysine was either lost or weaken decreasing the space available for water to bind. These results suggest that the secondary elements of the motif are necessary but not sufficient to trap water. In addition, a 100 ns MD simulation of solely the structural motif in solution reveal a random coil conformation. Meaning, this water recognition super-secondary structure requires connection to its enzyme in order to retain its integrity.[4]

2 WATER RECOGNITION MOTIF IN PADADH

2.1 Introduction

The enzyme investigated in this study is *Pseudomonas aeruginosa* D-Arginine dehydrogenase (PaDADH), a monomer with a flask-like cavity and a flavin adenine dinucleotide (FAD) cofactor.[1-3] PaDADH is composed of roughly 375 amino acid residues and has a water recognition motif[4] structure embedded near the site of substrate binding.

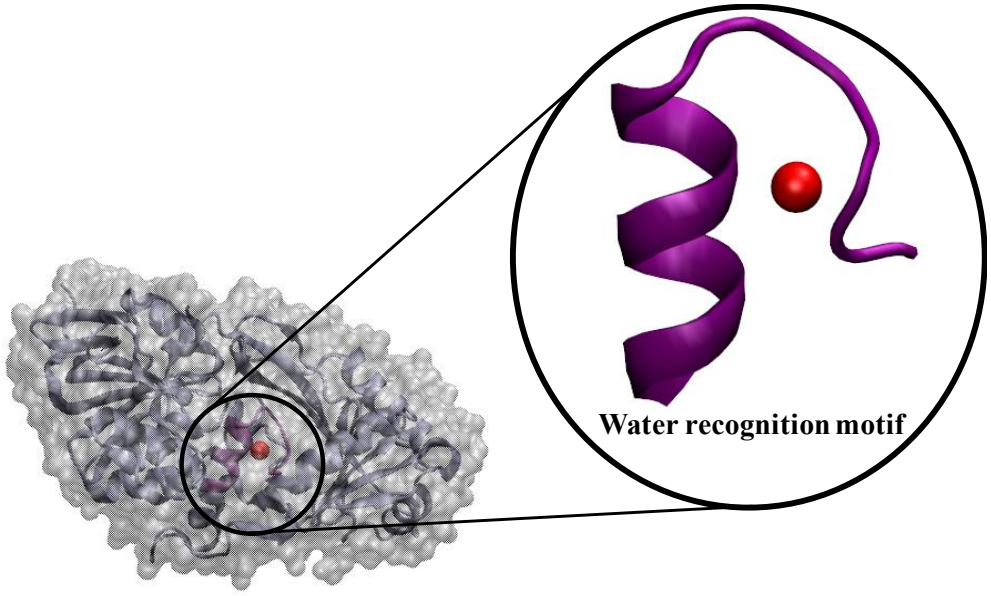


Figure 2.1 Distinct water recognition motif in PaDADH[1].
The water motif features are shown in purple. The water molecule is represented as a red sphere. A 16mer peptide was used to highlight the components of the water recognition motif.

PaDADH is considered to be an oxidoreductase[41,42] that catalyzes the transfer of electrons from its substrate through the removal of hydrogen atoms, during which its coenzyme, FAD, is reduced into $FADH_2$. Thus, FAD is a redox cofactor since it is a non-protein chemical compound required for enzymatic activity.[1] FAD has a flavin component, in which case, its N5 atom is suggested to be the site of electron transfer from the substrate.[1,43,44] In addition, the FAD structure has a straight chain sugar moiety referred to as D-Ribitol, followed by a pyrophosphate in a covalent bond with a cyclic sugar (D-Ribose), and an adenine base pair.

In general, PaDADH is folded into an FAD binding domain and substrate-binding domain with the substrate-binding domain on the *re* face of the FAD 7,8-dimethylisoalloxazine ring (flavin). The adenine portion of the cofactor is well adjusted in a deep pocket of the FAD binding domain while the 7,8-dimethylisoalloxazine ring is at the interface of the 2 domains where the substrate and FAD meet for interactions. In other words, FAD is non-covalently bonded in its domain through an ensemble of interactions with localized PaDADH residues and water molecules. Iminoarginine (Imino-Arg, Figure 2.2) is considered to be the product generated by the enzyme from catalysis with its substrate D-arginine. In addition, there exist four major loops which are suggested to play a role in substrate binding. These four loops (L1, residues 33-56; L2, 244-248; L3, 261-276; L4, 329-336) are suggested to be highly flexible, in which case, loop L1 shows two distinct conformations in the X-ray crystal structure of PaDADH in complex with iminoarginine (1.06 Å resolution; PDB ID 3NYC).[1]

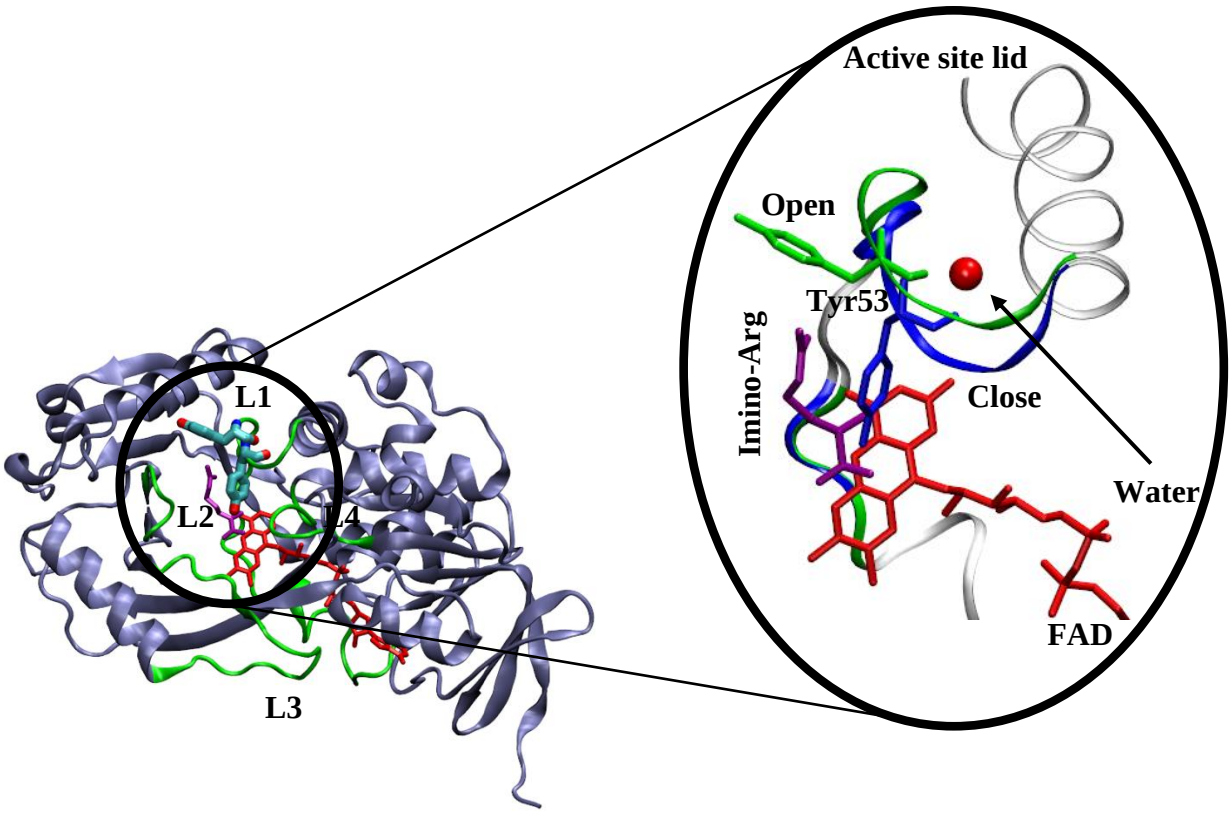


Figure 2.2 Schematic representation of the "active site lid" of PaDADH.[1]
The three dimensional structure of the enzyme is shown in ice blue cartoon display. The four loops (L1-L4) are shown in green. The FAD and Iminoarginine are shown in red and purple sticks respectively. The open and close conformation of the "active site lid" are magnified and colored green and blue respectively. Tyr53 serves as the key amino acid residue of the gate which regulates substrate binding and product release. The water molecule is shown as a red sphere.

Fu and Yuan et al[1-3] appointed loop L1 of PaDADH as an "active site lid." The lid served as a controlling factor for substrate entry and product release of the enzyme. Similar findings concerning other flavin dependent enzymes were also reported in regards to the catalytic involvement of an "active site lid". For example, a loop region established as an "active site lid" was proposed in the structure of porcine D-amino acid oxidase (pDAAO).[1] The 1.06 Å resolution X-ray crystal structure of PaDADH was reported to have two conformations, in which case, the ligand-free and product-bound configurations were denoted in accordance to the open and close state of its amino acid residue Tyrosine 53 (Tyr53). It is important to note that Tyr53 is also part

of loop L1, which is designated as the "active site lid" of the enzyme. This suggested a critical role for Tyr53 as a gated component in regards to enzyme catalysis. Tyr53 undergoes an open and close conformation which regulates the gated mechanism of PaDADH in respect to substrate binding and product release.

Typically, catalysis involves an enzyme binding to its corresponding substrate of high specificity. However, PaDADH, has a catalytic efficiency for every D-amino acid except for D-glutamate, D-aspartate, and glycine (achiral). Furthermore, the enzyme has been shown to not bind to L-arginine due to stereochemistry. As a result, the preference for catalysis is solely based on the configuration of D-amino acids. Experimental analyses were carried out to gain further insight on PaDADH's broad substrate specificity. PaDADH has a proposed catalytic site in decreasing order of efficiency with different D-amino acids from D-arginine to D-threonine. This decrease in efficiency becomes significantly noticeable to the point of unassessed kinetic parameters following D-leucine.[1]

Table 2.1 Steady-state kinetics study of PaDADH with D-amino acids.[1]

[a]These measurements were taken at 25°C, pH 8.7, 1mM PMS in 20mM Tris-HCl. [b]K_m and k_{cat} values were not recorded due to the inability to saturate the enzyme with the substrate. [c]Kinetics with cysteine was not shown due to its reduction with PMS (phenazine methosulfate). D-leucine was the last ligand with three established kinetic parameters (k_{cat}/K_m, k_{cat}, K_m).

Substrates	k_{cat}/K_m, M^{-1} s^{-1}	k_{cat}, s^{-1}	K_m, mM
D-arginine	$(3.4 \pm 0.3) \times 10^6$	204 ± 3	0.06 ± 0.01
D-lysine	$(5.3 \pm 0.2) \times 10^5$	141 ± 3	0.26 ± 0.01
D-tyrosine	27600 ± 3800	23 ± 1	0.8 ± 0.1
D-methionine	14800 ± 600	154 ± 3	10 ± 1
D-phenylalanine	6900 ± 300	75 ± 3	11 ± 1
D-histidine	3140 ± 30	35 ± 1	11 ± 1
D-leucine	515 ± 60	6.4 ± 0.3	12 ± 1
D-proline	420 ± 10	–[b]	–
D-tryptophan	245 ± 3	–	–
D-isoleucine	195 ± 3	–	–
D-valine	47 ± 1	–	–
D-alanine	41 ± 1	–	–
D-glutamine	186 ± 3	–	–
D-asparagine	16 ± 1	–	–
D-serine	3.8 ± 0.1	–	–
D-threonine	0.75 ± 0.01	–	–
D-glutamate	–	–	–
D-asparate	–	–	–
L-arginine	–	–	–
glycine	–	–	–
D-cysteine	nd[c]	nd	nd

X-ray crystal structures of PaDADH[1] suggested that the electrostatic interactions between the guanidinium of the iminoarginine and the side chain carboxyl group of Glu87 played a key role in the broad specificity of the enzyme with different D-amino acids. For example, the 1.30 Å resolution X-ray crystal structure of PaDADH in complex with iminohistidine[1] shows two substrate configurations in relative to one another by approximately 180°. Conformation A of the

iminohistidine is suggested to be catalytically competent due to its interactions with Glu87 and the FAD. However, conformation B of iminohistidine is proposed as being incompetent since the critical interactions with Glu87 and the FAD are lost.[1]

According to Fu et al[1], PaDADH's broad substrate specificity with different D-amino acids may be due to its preference for positively charged side chain ligands. For instance, amino acid residue glutamate 87 of PaDADH (Glu87) was proposed to form strong electrostatic interactions with the substrates D-arginine and D-lysine, hence their high catalytic efficiency with PaDADH during steady-state kinetics analysis.

The pocket entry of the active site of PaDADH is small yet increases in size as it approaches the FAD cofactor.[1] Since the 1.06 Å resolution X-ray crystal structure of PaDADH showed both ligand-free and product-bound conformations, the four designated loops (L1- L4), more specifically the "active site lid" (L1, residues 50 – 56) may be involved in the broad substrate specificity displayed by this enzyme. Thus, an additional suggestion on PaDADH's broad substrate specificity would be its ability to regulate the entrance and release of substrates at different kinetics based on the fluctuation of its "active site lid". The same approach could be taken in terms of binding, in which case, the stability of the "active site lid" in the close conformation would reflect PaDAH's ability to accommodate a variety of substrates.

A similar "loop-and-lid" structure within other glucose-methanol-choline (GMC) family members was revealed in the past. The enantiomer of PaDADH, LADH, have the same conserved interactions in the substrate-binding domain even though both enzymes are mirror images of each other in spatial orientation.[1] PaDADH have been shown to have low sequence identity but high three dimensional structure similarities with other D-amino acid oxidases in the family. For example, PaDADH is less than 17.2% identical in sequence identity to pDAAO with

a rmsd of 2.4 Å for 270 alpha carbons in relative to three-dimensional structure comparisons.[41] Thus, by gaining further insight on the catalytic properties of PaDADH, feasible hypotheses may be developed for other members of the GMC family.

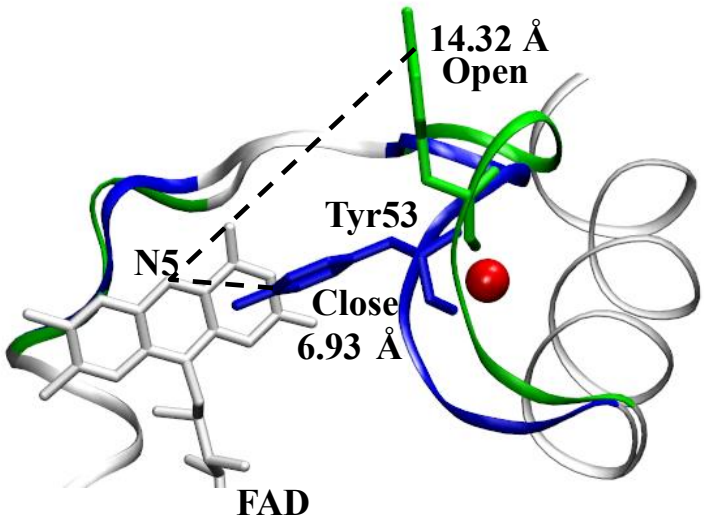


Figure 2.3 Distance betweenTyr53 (CZ) and FAD (N5) with water recognition motif.[1] The water molecule is shown as a red sphere. The open and close conformations are colored green and blue respectively. In addition, the distances between the zeta carbon (CZ) of Tyr53 and N5 atom of FAD in the open and close conformations are labeled accordingly.

Root-mean-square deviation (RMSD) evaluations were calculated to investigate the fluctuation of PaDADH and its cofactor in comparison to their frame of reference at the beginning of the trajectory. It was a necessity to show a localized FAD structure since its N5 atom was chosen as an arbitrary point to measure the change in distance and conformation of Tyr53 over time. Distance analyses between the FAD N5 atom and the zeta carbon (CZ) of Tyr53 were made apparent to investigate the open (A) and close (B) states of PaDADH (PDB ID 3NYC). Following such results, the probability distribution in regards to the presence and

absence of water was plotted in relation to the A and B conformational states of Tyr53 in PaDADH.

2.2 Experimental Procedures

Both PaDADH wild type free in solution and in complex with its D-arginine were constructed from an X-ray crystal structure of the enzyme with a resolution of 1.06 Å (Protein Data Bank (PDB) identification no. 3NYC).[1] D-arginine's partial charges were obtained with the application of the two-step RESP method[45] and Gaussian03[46] electrostatic potentials. AmberTools xleap software was used to construct the appropriate system required for each MD simulation of PaDADH. The FAD parameters originated from the Gaff amber force field.

Amber14 simulation package[47] was used to carry out all simulations with modified force field parameters of Cornell et al. (1995)[48] by Perez et al. (2007),[49] The solvation of each system was defined by a 10 Å TIP3P explicit water model in octahedron box. Once solvated, each system was neutralized accordingly with sodium ions. Both systems were minimized for a 1000 steps with a 100 kcal/mol/Å in harmonic constraints. Then, the systems were equilibrated in three steps at a constant pressure and temperature of 1 bar and 300K respectively. During the three equilibrated steps, the harmonic constraints were reduced in decreasing order of 50 kcal/mol/Å and 25 kcal/mol/Å, and 0 kcal/mol/Å. Both minimization and equilibration constraints were only applied on the enzyme, its cofactor (FAD), and the substrate when in complex with PaDADH. The first equilibrated step was carried out for 0.1 nsec with a time step of 2 fsec. The last two equilibrated steps were separately simulated for 0.2 nsec with also a 2 fsec time step.

Temperature regulation (300K) was carried out by using the Langevin thermostat with a collision frequency of 1 $psec^{-1}$. A 2 fsec time step was established to calculate Newton's equation of motion. The particle mesh Ewald (PME) method[50] was applied to evaluate the effects of long-range non-bonded interactions with a 9 Å cut off. The constraints on bonds involving hydrogen

atoms was taking into account by the SHAKE algorithm.[51] As a result, each system was simulated for a total of 1.1 μsec because the equilibration was extended for the first 100 nsec of the run.

Root mean square deviation of DADH was solely evaluated in respect to the backbone heavy atoms of each residue within PaDADH. RMSD of FAD was targeted towards the flavin component, in which case, the N1, N5, and C9 atoms were taken into consideration. This approach was plausible due to the planarity of the flavin component in regards to the N1, N5, and C9 atom positions. Amber14 package[47] was used for all distance analyses and to record the residue number of the water molecule that was closest to the water recognition motif in each frame. Then, water molecules that were located 3.5 Å away or less to either two of the carbonyl backbone of Thr50, Val51, and Tyr53 of the motif were taken into account. This particular strategy was plausible due to the position of the localized water molecule within the water recognition motif from the X-ray crystal structure of PaDADH and throughout the restart files of the two simulations. Whenever a water molecule was considered buried within the water recognition motif, it was in close interaction with either two of the backbone carbonyls of Val51, Tyr53, Thr50. Graphical representations of all data analyses were made possible through the application of the xmgrace software.

2.3 Results & Discussions

2.3.1 PaDADH free in solution

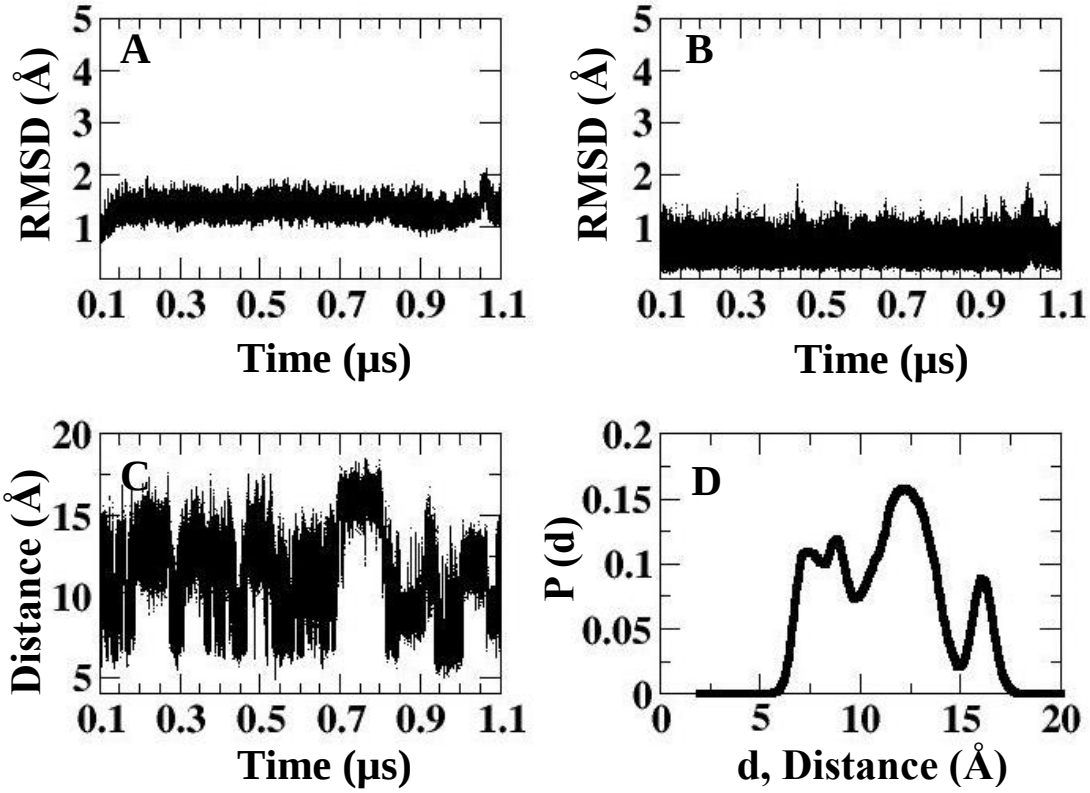


Figure 2.4 Dynamics analyses of PaDADH free in solution.
Panel A and B represent the RMSD assessment of PaDADH and FAD respectively. Panel C displays the distance between the Zeta carbon of Tyr53 and N5 of FAD over time. Panel D illustrates the probability distribution of the distance between the zeta carbon of Tyr53 and the N5 atom of FAD.

The root-mean-square deviation (RMSD) of PaDADH and its FAD cofactor in relative to their frame of reference at 0.1 μs indicate that both structures were relatively localized throughout the trajectory. The distances between the PaDADH Tyr53 zeta carbon and the N5 atom of FAD were established at the open and close conformations to be 14.32 Å, and 6.93 Å respectively. The fluctuation in distance between these two atoms over time (Figure 2.4, panel C) indicate that Tyr53 was indeed undergoing two conformational states while the rest of the

enzyme and FAD were low in mobility. Thus, the gated mechanism of PaDADH was supported by the MD simulation of the enzyme free in solution, in which case, ligand-free and product-bound conformations of the "active site lid" were occurring stochastically. The establishment of additional close and open populations in the probability distributions of PaDADH free in solution was due to the translational motion of the "active site lid" throughout the trajectory. According to the probability distribution analysis of PaDADH free in solution, an interesting characteristic of the enzyme was its ability to easily exist between open and close populations of its gated mechanism throughout the simulation.

It is plausible to suggest this particular enzyme to be one of the machineries of *Pseudomonas aeruginosa*, that is constantly experiencing two states for catalysis of various structures of D-amino acids. The probability distribution of the open and close conformations of Tyr53 while in PaDADH free in solution became evident to study different properties of the gated mechanism of the enzyme. According to panel D, there exist four dominant populations in regards to this gated catalysis. Based on the open and close distances established from the 1.06 Å resolution X-ray crystal structure of PaDADH,[1] and monitored observations of the simulation over time, the first and last two populations were defined as product-bound and ligand-free configurations respectively. An interesting characteristic of unbound PaDADH in solution was the ability to easily exist between open and close populations of its gated mechanism. The probability distribution of the different populations of the gated mechanism of PaDADH became significant to examine any shift in distribution of its "active site lid" at different enzymatic conditions or parameters.

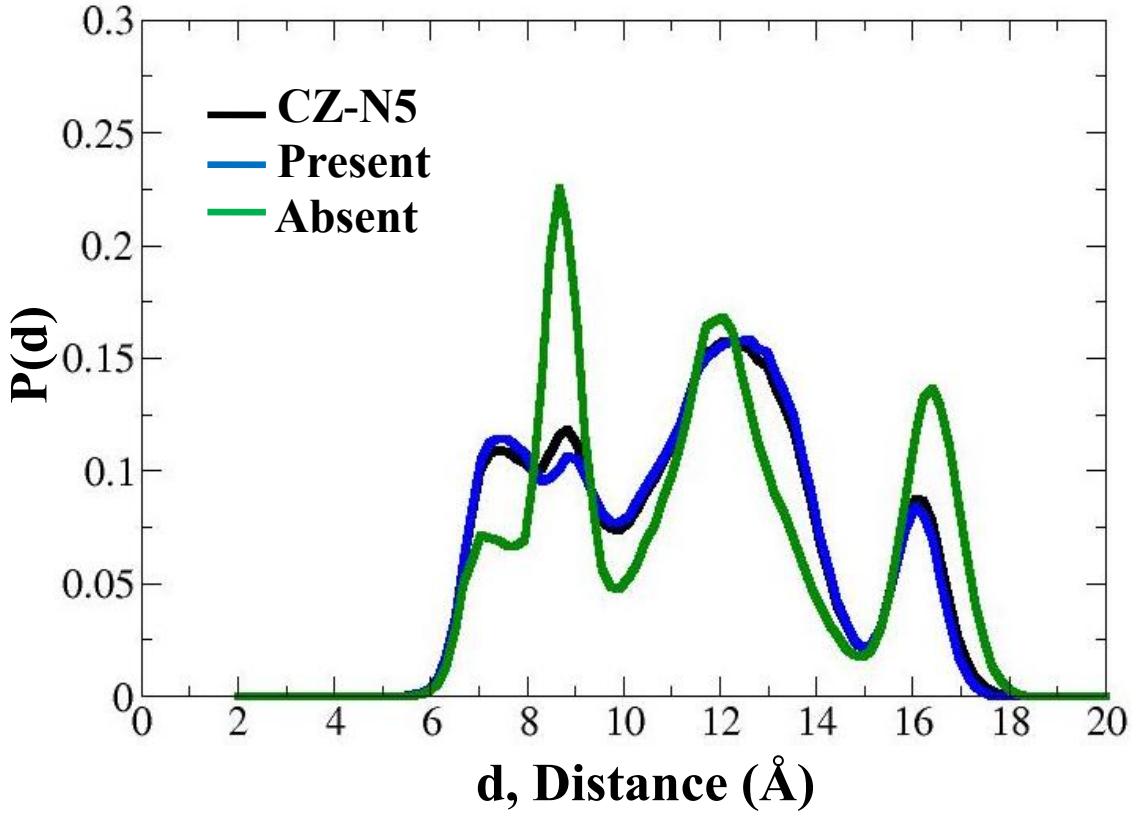


Figure 2.5 Distribution of buried water molecules in the water recognition motif. The probability distribution of the distance between the zeta carbon (CZ) of Tyr53 in PaDADH and the N5 atom of FAD is illustrated in black. The probability distribution of the presence and absence of water in respect to the distance between PaDADH Tyr53 (CZ) and FAD (N5) is shown in blue and green respectively.

While PaDADH was free in solution, water was established to be buried within the water recognition motif at a percentage of 88%. Under the same conditions, water was defined absent from the enzyme's water recognition motif pocket 12% of the time. In other words, during the simulation, water molecules were consistently exchanging within the vacant pocket of the water recognition motif of PaDADH. Even though the percentage of absent water molecules was minimal in comparison to the percentage of water presence, distinct enzymatic characteristics of PaDADH became evident whenever water was not embedded in the motif's pocket during the simulation. According to Figure 2.5, when water was concluded buried within the water recognition motif, a probability distribution very similar in magnitude and shape to the

probability distribution of the gated mechanism of PaDADH free in solution was established. However, during the period of water absence from the motif's pocket, the probability distribution of the populations of the gated mechanism increase in intensity in relative to each other. Since Gibb's free energy is equivalent to the gas constant multiplied by temperature and the natural log of probability, increases in the probability distribution peaks correlate with higher energy barriers between the observed gated populations of PaDADH.

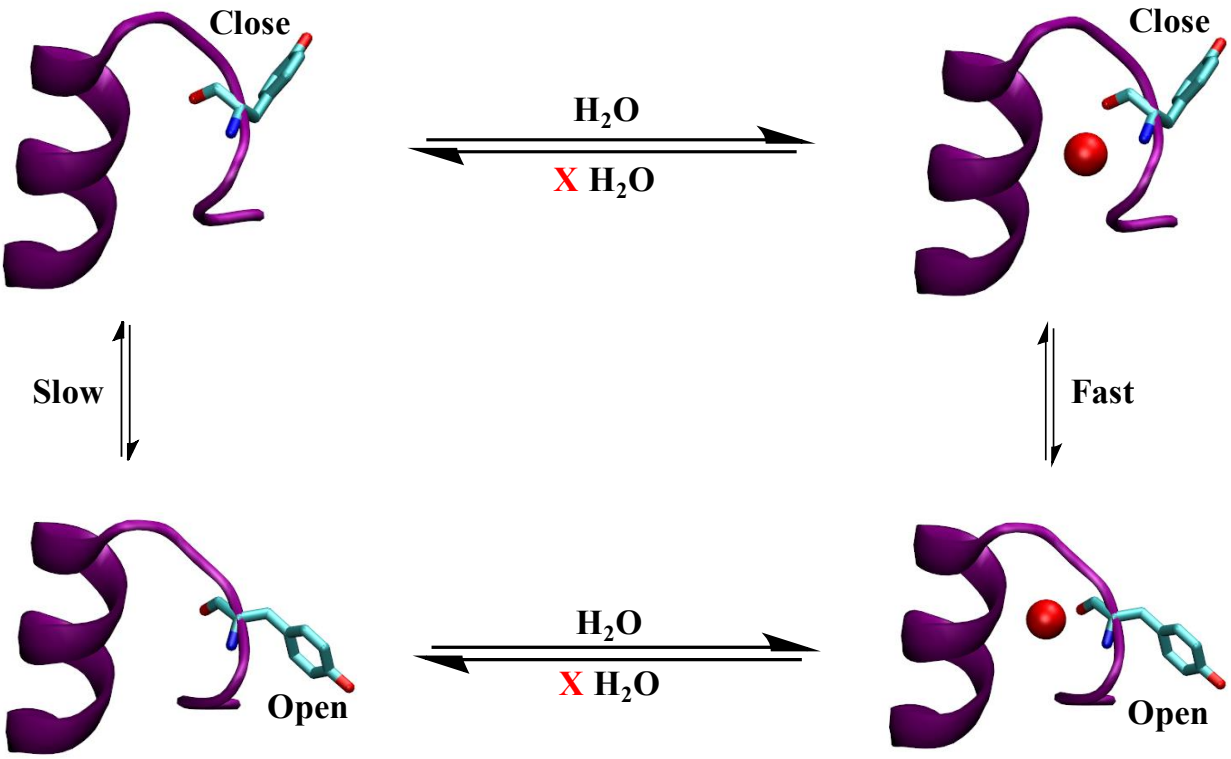


Figure 2.6 Schematic representation of the absence and presence of water.
A 15mer peptide was selected to illustrate the water recognition motif in purple color. Water (not labeled) is represented as red spheres, and PaDADH Tyr53 is shown as sticks, in which case, its open and close conformations were established.

The water recognition motif is the "active site lid" that is involved in the gated mechanism of PaDADH. As previously established, this gate can fluctuate stochastically between the open and close conformations of the enzyme over time. In addition, water within this "active site lid" can be buried or considered absent in both conformations. The key was to

establish any given trend(s) between buried water molecules and "active site lid" conformations. Consequently, the analyses of embedded water molecules in the motif pocket during the simulation were evaluated to decrease the energy barriers between the gated populations of PaDADH. Thus, buried water molecules in the water recognition motif over time facilitated the change in conformations of the "active site lid". However, whenever water molecules were not included in the motif pocket PaDADH underwent a much slower fluctuation in its gated mechanism.

2.3.2 PaDADH in complex with D-arginine

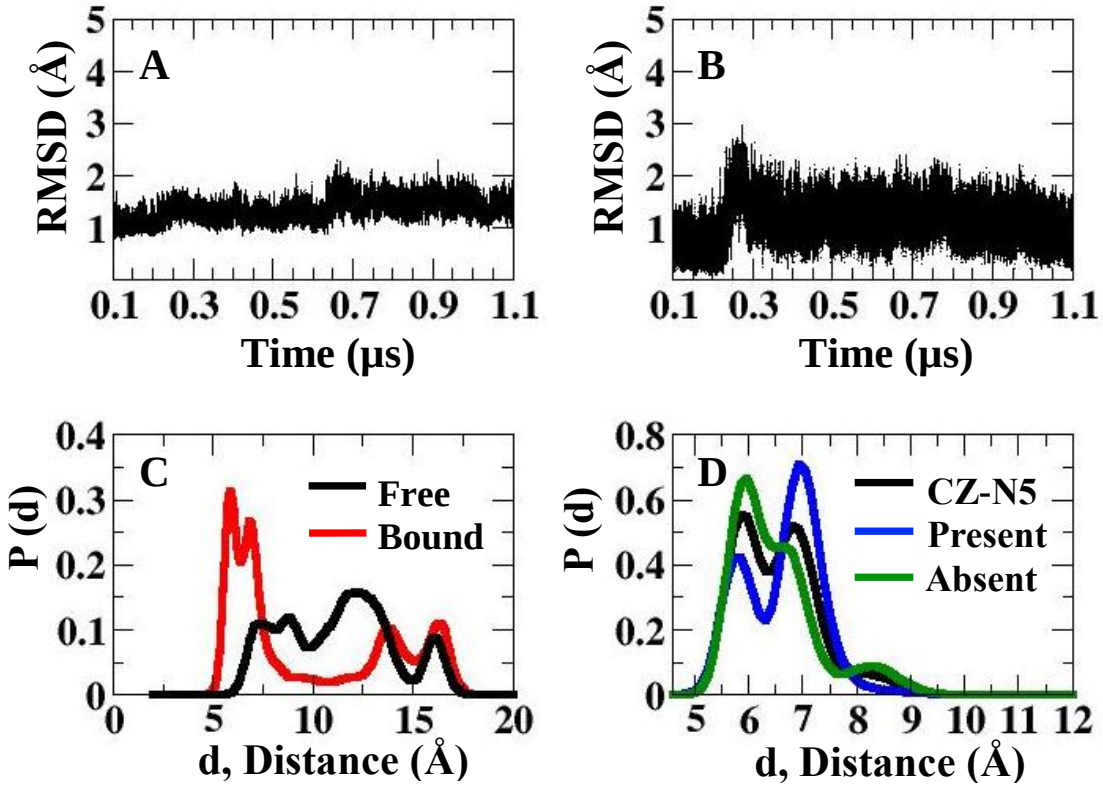


Figure 2.7 Dynamics analyses of PaDADH in complex with D-arginine.
RMSD calculation of PaDADH and FAD are shown in panel A and B respectively. Panel C indicates the probability distribution of the distance between Tyr53 (CZ) and FAD (N5) of PaDADH free in solution and in complex with D-arginine. Panel D shows the distribution of the absence and presence of buried water molecules within the water recognition motif during the conformational changes of Tyr53 while the enzyme is in its close state.

RMSD calculations of PaDADH in complex with D-arginine showed localization of the enzyme and its FAD cofactor. Using the same analytical approach of PaDADH free in solution, the probability distribution of the gated mechanism of the enzyme in complex with D-arginine indicated a shift towards its close configurated populations.

In regards to the presence or absence of buried water molecules within the water recognition motif of PaDADH, there existed a clear preference for the two close gated conformations of Tyr53 in relative to FAD. When in complex with D-arginine, the probability of water being buried in the water recognition motif of PaDADH was lower once the distance between the zeta carbon of Tyr53 and the FAD N5 atom decreased to ~ 5 Å. As the FAD (N5) – Tyr53 (CZ) distance increased to ~ 6 Å, the likelihood of embedded water molecules in the motif was significantly higher. Such findings corresponded well with the results seen with PaDADH free in solution. During enzyme – substrate binding, the closing of the gate at ~ 5 Å was suggested to be catalytically favorable, thus, a lower probability of water being buried would increase the energy barrier necessary for the "active site lid" to change in conformation. As the distance increased to ~ 6 Å, a higher probability of water was now required to facilitate the change in configuration of the gated mechanism.

2.4 Conclusions

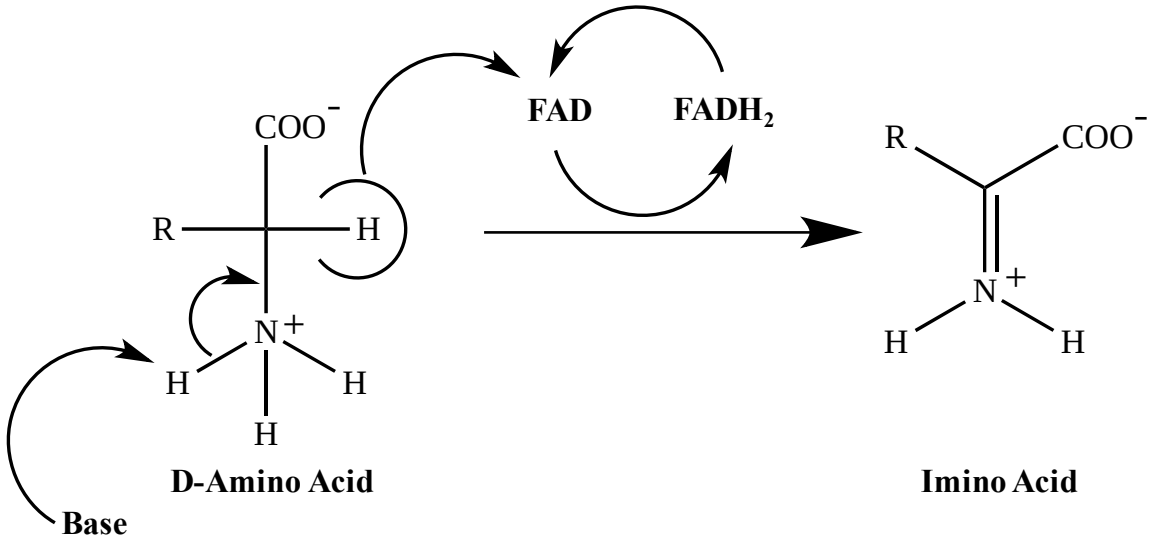
As established with PaDADH free in solution, ligand-free and product-bound conformations of the "active site lid" were occurring via a stochastic distribution over time. The presence of water within the vacant pocket of the water recognition motif facilitated the open and close configurations of the gated mechanism of PaDADH, while the absence of water increased the energy barriers between the gated populations of the enzyme. PaDADH in complex with D-arginine showed a preference of the gated mechanism of the enzyme towards its close conformations. The close gated configuration of the enzyme that was suggested to be catalytically competent showed a preference for water absence in the motif's pocket, while water presence was required to facilitate configurational changes of the "active site lid".

3 PADADH Y53F, S45A, AND A46G MUTANT DYNAMICS

3.1 Introduction

PaDADH is a flavin-dependent enzyme that catalyzes the oxidative deamination of different D-amino acids into their corresponding imino acids. The zwitterionic forms of its substrates are suggested to be the preferred protonation state for ligand-enzyme binding. PaDADH is suggested to deprotonate the backbone amino groups of its various D-amino acid substrates, which in turn causes electrons to delocalize and create a hydride transfer from their alpha carbons to the N5 atom of the FAD cofactor. This reaction is hypothesized to be triggered by a basic catalytic amino acid residue in the enzyme's active site. A similar reaction is seen with choline oxidase where a hydrogen and proton transfer are proposed to coincide in catalysis.[1-3,5] Furthermore, a suggestion concerning the concurrence of a proton and hydride transfer using FAD as a cofactor was proposed from the X-ray crystal structure of D-amino acid oxidase in *Rhodotorula gracilis*.[52]

Scheme 3.1 Enzymatic mechanism of PaDADH.



pH effects on the kinetic parameters of PaDADH suggested that an unprotonated group with a pK_a of 9.6 was the basic catalytic amino acid residue which triggered electron delocalization within the enzyme.[5] Such pK_a value became apparent to either the backbone amino groups of PaDADH D-amino acid substrates (pKa ~ 9) or a tyrosine side chain residue of the enzyme (pKa ~ 10). The hydroxyl group of Tyr53 from the "active site lid" of PaDADH was observed to be 3.74 Å away from the backbone nitrogen of iminoarginine in the product-bound conformation.[5] Thus, PaDADH Tyr53, in close proximity for acid-base chemistry, was proposed to be the catalytic base, which triggered the oxidative deamination reaction of the enzyme.

Table 3.1 Kinetic study of PaDADH Y53F mutant variant.[53]

k_{cat} and k_{cat}/K_m with D-Arginine at pH 8.7		
Enzymes	k_{cat} (s^{-1})	k_{cat}/K_m (M^{-1} s^{-1})
Wild Type	204 ± 6	3,400,000 ± 0.3
Y53F	420 ± 4	1,200,000 ± 6

As a consequence, a single amino acid mutation of PaDADH was carried out, in which case, Tyr53 was mutated into a phenylalanine (Y53F).[53] The goal of the experiment was to render the enzyme nonfunctional since the oxidative deamination reaction would no longer be triggered due to the absence of the hydroxyl group of tyr53. However, PaDADH remained functional with a significant decrease in efficiency.

Another region of interest in PaDADH, with two conformational states, is the Serine 45 (Ser45) to Alanine 47 (Ala47) sequence located near the *si* face of the flavin component of its FAD cofactor.[1] For instance, in the ligand-free conformation, there existed a hydrogen bond between the backbone carbonyl group of Ser45 and the N5 atom of the FAD. However, in the product-bound conformation of PaDADH, the hydrogen bond was replaced with a polar interaction, which involved the backbone amino group of Ala46 and FAD (N5). Other interesting findings showed that interactions between the FAD N5 atom and analogs of Ala46 within that

region (Ser45 – Ala47) remained conserved in other FAD-dependent enzymes.[1] Furthermore, members of the glucose-methanol-choline (GMC) family, like PaDADH, shared a similar Ser45 – Ala47 sequence near the *si* face of the FAD flavin component, in which case, bulkier and smaller groups within this sequence were suggested to correspond to an enzyme with low and high reactivity with oxygen respectively. For example, the reduced FAD of L-galactono-γ-lactone dehydrogenase increased in reactivity with oxygen by approximately 400-fold during a single amino acid mutation from Alanine 113 to Glycine.[1]

Table 3.2 Kinetic study of PaDADH S45A mutant variant.[58]

High pH k_{cat} and k_{cat}/K_m with D-Arginine		
Enzymes	**k_{cat} (s^{-1})**	**k_{cat}/K_m (M^{-1} s^{-1})**
Wild Type	**160 ± 20**	**2,600,000 ± 78,000**
S45A	**200 ± 30**	**290,000 ± 50,000**
High pH k_{cat} and k_{cat}/K_m with D-Leucine		
Enzymes	**k_{cat} (s^{-1})**	**k_{cat}/K_m (M^{-1} s^{-1})**
Wild Type	**100**	**11,000**
S45A	**nd**	**65**

Table 3.3 Kinetic study of PaDADH A46G mutant variant.[58]

High pH k_{cat} and k_{cat}/K_m with D-Arginine		
Enzymes	**k_{cat} (s^{-1})**	**k_{cat}/K_m (M^{-1} s^{-1})**
Wild Type	**160 ± 20**	**2,600,000 ± 78,000**
A46G	**240 ± 20**	**270,000 ± 70,000**

High pH k_{cat} and k_{cat}/K_m with D-Leucine		
Enzymes	**k_{cat} (s^{-1})**	**k_{cat}/K_m (M^{-1} s^{-1})**
Wild Type	**100**	**11,000**
A46G	**nd**	**40**

Consequently, a similar approach was taken into consideration in regards to the structural stability and enzymatic properties of PaDADH.[58] For example, a single amino acid mutation of Ser45 to alanine (S45A) was carried out in hopes of increasing reactivity of the reduced FAD cofactor of the enzyme with oxygen. As a result, PaDADH reactivity with oxygen remained minimal while the enzyme became significantly less efficient catalytically with different substrates. In addition, assessable kinetic parameters of D-leucine from wild type PaDADH could no longer be measured in the S45A mutant variant of the enzyme. Following such approach, PaDADH Alanine 46 was mutated to Glycine (A46G) to increase the reactivity of the enzyme's reduced FAD with oxygen. Subsequently, PaDADH A46G mutant variant had a greater decreased in catalytic efficiency in relative to the S45A mutation of the enzyme while maintaining similar level of reactivity with oxygen.

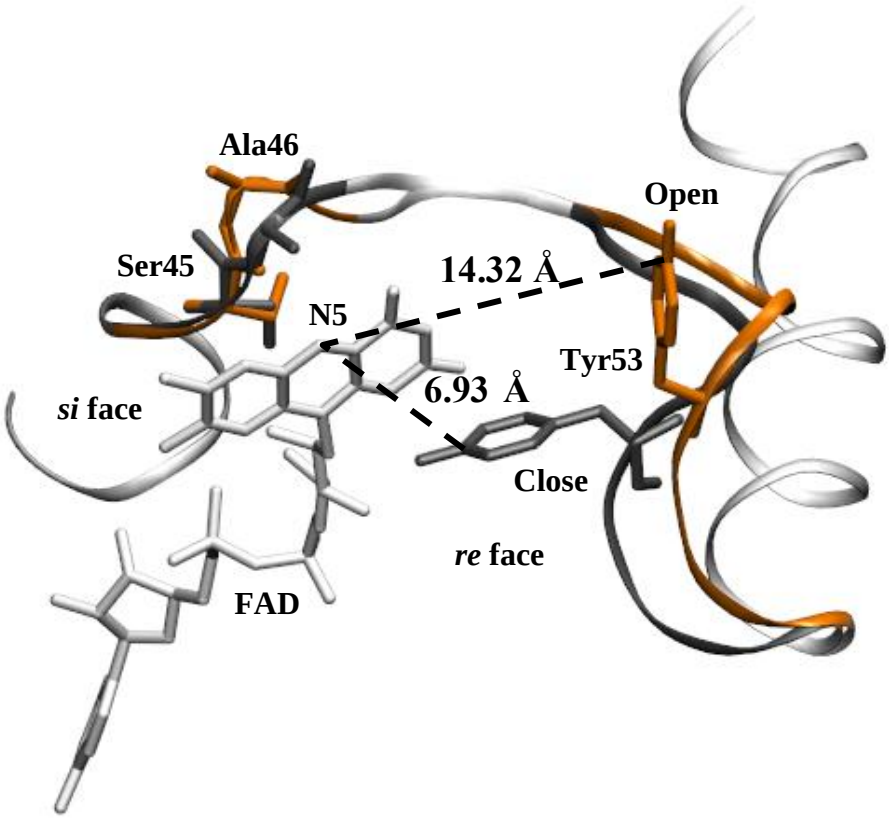


Figure 3.1 Distance betweenTyr53 (CZ) and FAD (N5) with site of mutations.[1]
Ala46 and Ser45 are located on the *si* face of the FAD flavin component with Tyr53 positioned on the *re* face of the structure. The open and close conformations are colored orange and black respectively with their corresponding distances labeled.

In all three cases, as indicated from the kinetic analyses, PaDADH mutant variants remained functional but decreased in catalytic efficiency. All three mutation sites were part of the flexible loop (L1), which included the "active site lid" of the enzyme. MD simulation results of wild type PaDADH free in solution supported a stochastic distribution of both ligand-free and product-bound configurations of the enzyme over time. Thus, carrying MD simulations of single amino acid mutations (Y53F, S45A, and A46F) of PaDADH free in solution would indicate the effect of dynamics on the mechanical properties of the enzyme.

Following each MD simulation, both FAD and apoenzyme RMSDs were evaluated to investigate the fluctuation of PaDADH and its cofactor over time. In addition, a more localized

FAD fluctuation was required for distance analyses between the N5 atom of the FAD and the zeta carbon (CZ) of Tyr53 (Phenylalanine 53 in case of Y53F PaDADH mutant variant). The probability distribution of this distance for each mutation of PaDADH (Y53F, S45A, and A46G) was then overlapped with the result obtained from wild type enzyme free in solution. Through the comparison of open and close probability distributions of mutant variants and the wild type form of PaDADH, the dynamics of the enzyme throughout each trajectory became more evident.

3.2 Experimental Procedures

The Y53F, S45A, and A46G mutant variants of PaDADH were derived from a 1.06 Å resolution X-ray crystal structure of the enzyme (PDB ID 3NYC).[1] All MD simulated variant system of PaDADH were constructed via AmberTools xleap software. Gaff amber force field parameters were used to derive the parameters for the FAD structure.[47]

All simulations were carried out using amber14 softwares[47] under the same force field parameters.[48,49] Each system was properly solvated with a 10 Å TIP3P water model octahedron box, in which case, sodium ions were subsequently added for neutralization. All simulations were minimized for a 1000 steps with harmonic constraints of a 100 kcal/mol/Å on PaDADH and FAD. Following minimization, all systems underwent three rounds of equilibration with harmonic constraints in decreasing order of 50 kcal/mol/Å, 25 kcal/mol/Å, and 0 kcal/mol/Å on the enzyme and its cofactor. All equilibrations were processed at a constant temperature, and pressure of 300K, and 1bar respectively. Each equilibration was parameterized with a time step of 2 fsec, however, the first equilibrated round was carried out for 0.1 nsec, while the last two rounds were individually processed for 0.2 nsec.

Langevin thermostat with a collision frequency of 1 $psec^{-1}$ was the applied method to regulate temperature. Newton's equation of motion was consistently evaluated through a time step of 2fsec. Long-range nonbonded interactions, with a cut off of 9 Å, were establishing by using the particle mesh Ewald (PME) method.[50] The SHAKE algorithm[51] was applied to constrained bonds involving hydrogen atoms. The trajectory of each mutant variant was carried out for 1.1 µsec, in which case, the first 100 nsec was considered to be further equilibration with no constraints. The xmgrace software was used for the graphical representations of all data analyses of the investigation.

3.3 Results & Discussions

3.3.1 Y53F

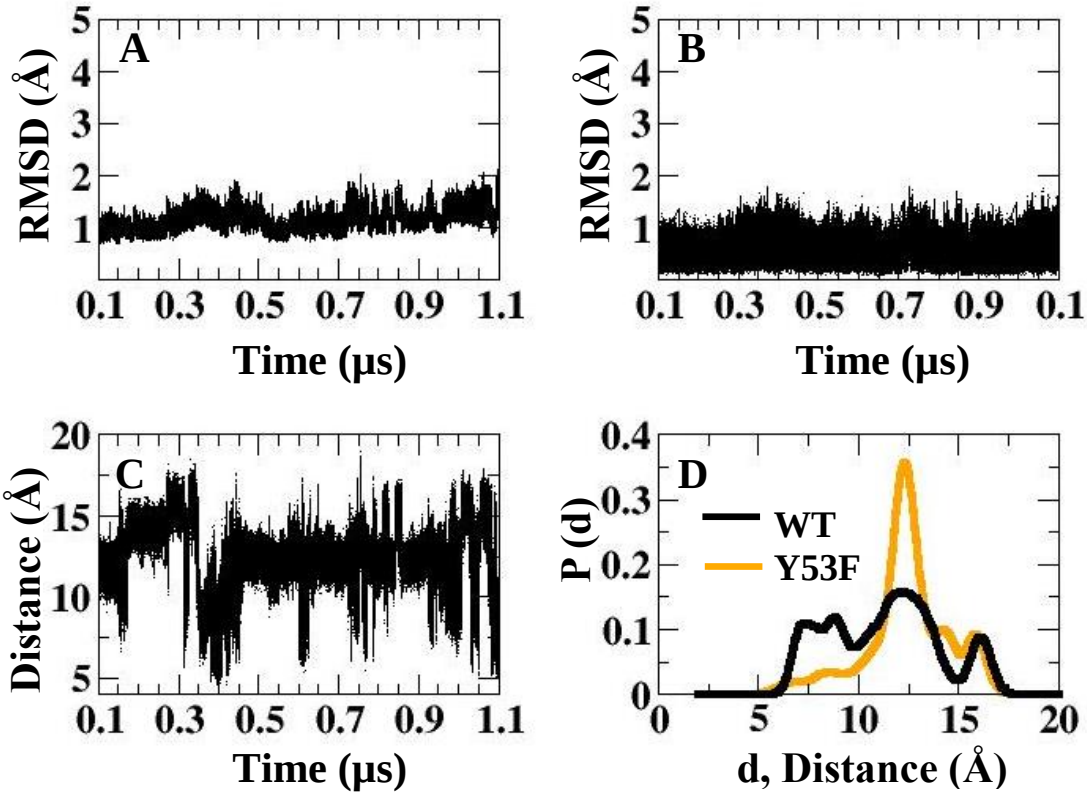


Figure 3.2 Dynamics analyses of PaDADH Y53F mutant variant.
Panel A and B indicate the RMSD results for Y53F PaDADH and FAD respectively. The distance between PaDADH Tyr53 (CZ) and FAD (N5) over time is shown in panel C. The probability distributions of the distance concerning Tyr53 (CZ) – FAD (N5) in respect to both the wild type (WT) PaDADH and the Y53F mutant variant of the enzyme are shown in panel D.

As seen from the wild type enzyme free in solution and in complex with D-arginine, the PaDADH Y53F mutant variant and its FAD cofactor were relatively low in mobility. As established from previous distance approaches, the initial distance between the CZ atom of Tyr53 and the N5 atom of FAD were measured from the PDB structure of PaDADH in complex with iminoarginine (PDB ID 3NYC).[1] Based on the fluctuation in distance between these two atoms, the Y53F mutant variant of the enzyme indicated a tendency of the "active site lid" to

remained open for a longer period in time (panel C). Upon overlapping, the Y53F mutant variant, and wild type distance probability distributions of PaDADH, a noticeable shift towards the open conformation of the gated mechanism of the enzyme was established. This simulation result converged with the catalytic kinetic measurements (k_{cat}, K_m, k_{cat}/K_m) of the Y53F mutant variant of the enzyme with different D-amino acid substrates.[53] As the open conformation probability of the "active site lid" increased, PaDADH became less efficient catalytically. Mutating Tyr53 into a phenylalanine, did not remove the catalytic basic residue of the enzyme which triggered the oxidative deamination process of various D-amino acids. However, the Y53F mutation of PaDADH affected the mechanical aspect of its "active site lid", which in turn correlated with the kinetics of the enzyme with different D-amino acid substrates.

3.3.2 S45A

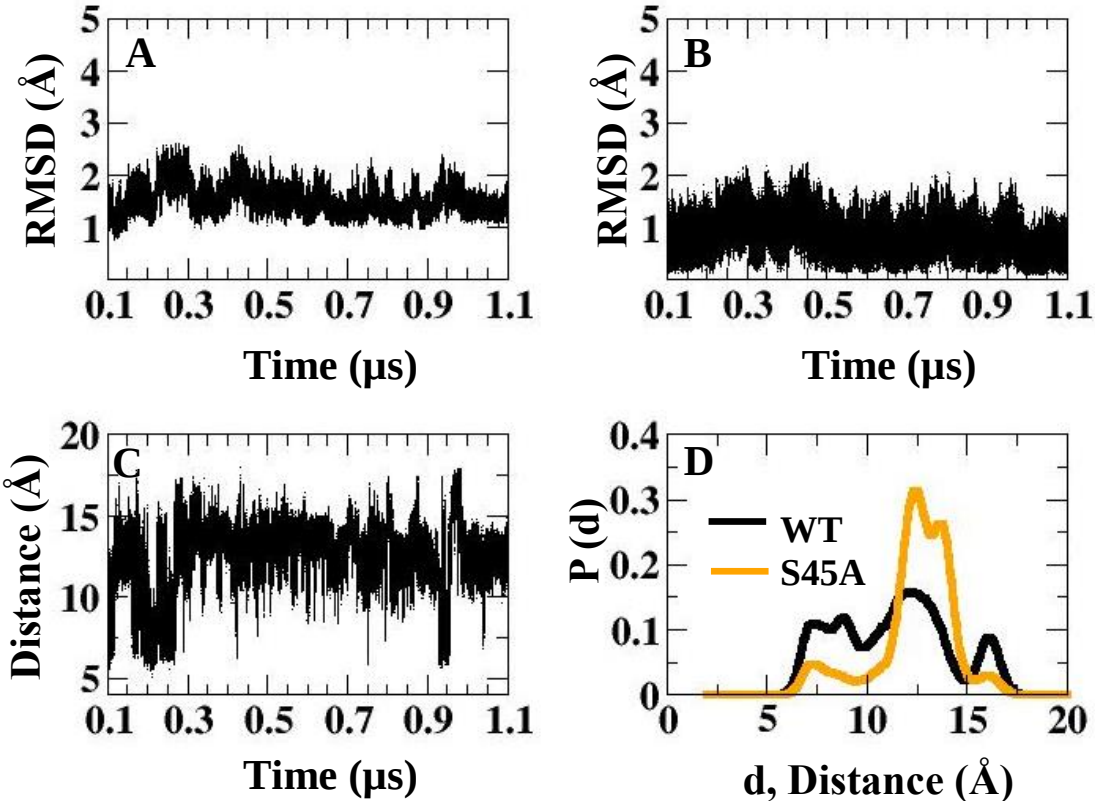


Figure 3.3 Dynamics analyses of PaDADH S45A mutant variant.
RMSD evaluations of S45A PaDADH and FAD are established in panel A and B respectively. Panel C indicates the change in distance between the zeta carbon (CZ) of Tyr53 and N5 atom of FAD over time. Panel D shows the probability distribution of the Tyr53 (CZ) – FAD (N5) distance in respect to the wild type (WT) and the S45A mutant variant of the enzyme.

This was another case of high localization of PaDADH S45A mutant variant and its FAD cofactor. Tyr53 (CZ) – FAD (N5) distance analysis of S45A PaDADH indicated a preference of the "active site lid" to its open conformation over time (panel C). In addition, comparison of probability distributions involving wild type and S45A PaDADH free in solution (panel D), established a shift towards the ligand-free configuration in the S45A mutation of the enzyme. Similar to the dynamics of the Y53F mutant variant of PaDADH, the gated mechanism of the enzyme was mechanically affected when Serine 45 was mutated into an alanine. The S45A mutant variant of PaDADH did not significantly increase in reactivity with oxygen, however, the

catalytic efficiency of the enzyme was noticeably lowered. Furthermore, MD analyses of S45A PaDADH established a correlation between the gated mechanism and the kinetics (k_{cat}, K_m, k_{cat}/K_m) of the enzyme with different D-amino acid substrates. For instance, after the S45A mutation, certain kinetic parameters for D-leucine, established with wild type PaDADH kinetic measurements, could no longer be recorded. Nevertheless, though low in efficiency, kinetic parameters for D-arginine with both wild type and S45A PaDADH remained assessable. D-arginine has the ability to form a favorable electrostatic interaction with PaDADH Glutamate 87 (Glu87), but D-leucine cannot form such contacts due to its short nonpolar side chain.[1] This suggest that D-arginine has an affinity to PaDADH even when the "active site lid" is primarily open. On the other hand, D-leucine requires the distribution of the lid to be shifted towards the close conformation for favorable enzyme-substrate binding to occur.

3.3.3 A46G

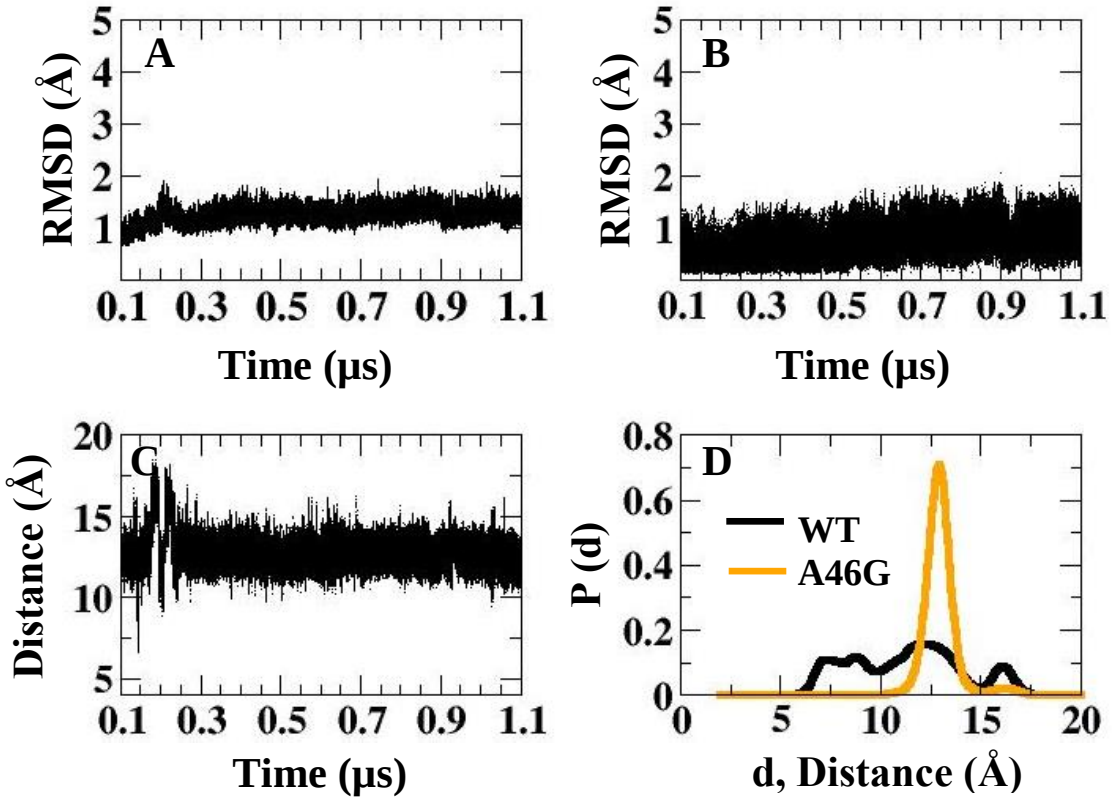


Figure 3.4 Dynamics analyses of PaDADH A46G mutant variant.
Panel A and B refers to the RMSD calculations of A46G PaDADH and FAD respectively. The change in distance between Tyr53 (CZ) and FAD (N5) over time is established in panel C. Panel D illustrates the probability distribution of the distance between the zeta carbon (CZ) of Tyr53 and the N5 atom of FAD in regards to the wild type and the A46G mutant variant of the enzyme.

The A46G mutant variant of PaDADH, including its FAD cofactor displayed high localization properties from the RMSD assessments. Mutating PaDADH Alanine 46 to glycine resulted into an "active site lid" that was predominately open over time (panel C). Similar to previous single amino acid mutations of the enzyme, certain mechanical aspects of the "active site lid" were altered in A46G PaDADH. In addition, the kinetic measurements (k_{cat}, K_m, k_{cat}/K_m) of the A46G mutant variant of PaDADH indicated a decrease in efficiency. Comparison of S45A and A46G results showed noticeable convergences between kinetic and dynamics analyses of the enzyme. For instance, the "active site lid" of A46G had a probability distribution that was more

prominent than S45A towards the open conformation when both enzymes were simulated free in solution. Evaluation of kinetic properties with different D-amino acid substrates established lower enzyme efficiency in both cases, however, the results were more evident with the A46G mutation of the enzyme. Both cases also indicated a connection between substrate binding and the gated mechanism. As seen from the S45A mutant variant of PaDADH, certain kinetic parameters for D-leucine became inaccessible once Alanine 46 was mutated to glycine. However, while low in efficiency, D-arginine kinetic measurements were evaluated for both wild type and A46G mutant variant of PaDADH due to its favorable interaction with residue Glu87 of the active site and overall enhanced binding with the enzyme.[1] Thus, as the mechanics of the gate is increasingly altered towards the open conformation, the kinetic deficiencies of the enzyme become more apparent.

3.4 Conclusions

The gated mechanism of PaDADH was mechanically affected as seen from the investigated MD results of the single amino acid mutations of the enzyme. The dynamics analysis of each mutant variant of PaDADH (Y53F, S45A, A46G), while free in solution, displayed convergence with previous corresponding kinetic measurements of the enzyme with different D-amino acid substrates. Even though, both S45A and A46G mutant variants of PaDADH showed similar substrate-binding and "active site lid" dynamics properties, the affects were more pronounced in A46G relative to S45A.

4 PADADH TYR53, AND GLU246 PKA SHIFT CALCULATIONS

4.1 Introduction

Bacteria such as *Pseudomonas aeruginosa* have the ability to grow solely with D-arginine as a carbon and nitrogen source.[42,54] For example, *Pseudomonas aeruginosa* uses DADH in a two-enzyme-coupled system to catalyze the D- to L- conversion of amino acids. PaDADH converts D-arginine into 2-ketoarginine and ammonia, and PaLADH catalyzes the process of 2-ketoarginine to L-arginine.[1-3,42]

After its oxidative deamination reaction, PaDADH can either form an enzyme-imino acid complex or release the product into its aqueous environment. Thus, hydrolysis of the imino acid can either occur by water molecules within the enzyme's active site or in the surrounding environment. Previous crystallography reports have shown the X-ray crystal structure of PaDADH in complex with iminoarginine[1,3], however, an unexpected 1.07 Å resolution crystal structure of a covalent N(5) flavin adduct was discovered in the active site of the enzyme after crystallization with D-leucine.[2] This PaDADH acyl adduct complex was replicated by photoreduction of the enzyme in the presence of ketoleucine. The acyl adduct structure within the active site of PaDADH was suggested to be the product of a nucleophilic substitution reaction, in which case, the N(5) atom of the flavin component of the FAD cofactor was reduced by the X-ray beam from the crystallization process, and formed a covalent bond with the α-carbon of ketoleucine to generate 4-methly-2-pentanone-FAD and CO_2.

Further studies on another flavin-dependent enzymatic structure of the GMC family, *Rhodotorula gracilis* D-amino acid oxidase (RgDAAO),[44] have suggested RgDAAO Serine 335 (Ser335) to be the catalytic residue with acid-base properties, which removed the backbone amino proton of D-amino acids to initiate the delocalization of electrons to the FAD. Similar to

PaDADH, the zwitterion form of D-amino acids was suggested to be preferential for favorable enzyme-substrate binding with RgDAAO. These hypotheses were in agreement with suggestions made in respect to flavin-dependent *Rhodococcus opacus* (roLAAO), in which case, roLAAO Aspartate 227 (Asp227) was suspected to abstract the proton on the backbone amino group of D-amino acids through a relay of water molecules.[52]

The oxidative deamination reaction of RgDAAO was suggested to be reversible due to the co-planarity of the FAD in relative to the backbone of D-amino acid substrates. Based on the 1.06 Å resolution X-ray crystal structure of PaDADH,[1] the FAD N5 atom was established to be at a distance of 3.3 Å from the α-carbon of the iminoarginine product, which was attuned with the proposed hydride transfer mechanism to the FAD. Closing of the "active site lid" of PaDADH was suggested to increase the hydrophobicity of the enzyme's active site, to facilitate the hydride transfer between the alpha carbon of D-amino acid substrates and the N5 atom of FAD.

Based on X-ray crystal structure assessments, the structural equivalent residue(s) of PaDADH in relative to RgDAAO Ser335, could either be Tyr53 or Glutamate 246 (Glu246). According to the 1.06 Å resolution X-ray crystal structure of PaDADH, Tyr53 is less than 4 Å away from the nitrogen on the imino group of iminoarginine in the product-bound configuration of the enzyme. In addition, Glu246 is less than 3 Å away from Tyr53, through a favorable side chain hydrogen bonding interaction, with PaDADH in its close conformation. In addition, Glu246 is less than 7 Å away from the nitrogen backbone imino group of iminoarginine, which does not make it a logical candidate for the basic catalytic residue that triggers the oxidative deamination of D-amino acids. Nevertheless, similar to the suggestion concerning the deprotonation of the backbone amino group of D-amino acids in roLAAO, proton removal may

occur through a relay of hydrogen bonding with water molecules. Hence, Glu246 may be the enzymatic amino acid residue of PaDADH with acid-base properties. The Y53F[53] mutant variant of PaDADH invalidated the hypothesis of Tyr53 having the ability to remove a proton since the enzyme remained functional after the mutation. Yet, $pK_{a,shift}$ of Tyr53 in relative to Glu246 may support or refute the hypothesis of proton relay with water molecule(s) and surrounding enzymatic residue(s) of PaDADH. As a result, the thermodynamic integration method was carried out to evaluate the shift in pK_a of the backbone hydroxyl and carboxyl groups of Tyr53 and Glu246 respectively, with PaDADH in its open conformation free in solution (A), and in its close conformation in complex with D-arginine (B).

4.2 Experimental Procedures

Amber14 simulation package[47] was used to generate the appropriate systems for pKa,shift analysis. An X-ray crystal structure with a resolution of 1.06 Å (Protein Data Bank (PDB) identification no. 3NYC),[1] was the investigated model used for thermodynamic integrations of PaDADH. The proton of the hydroxyl group of Tyr53 and the carboxyl group on the side chain of Glu246 were probed for $pK_{a,shift}$ analysis. By changing the partial charges of the residues of interest from the protonated to the deprotonated state, the changes in Gibbs free energy of the protein and the model were generated. λ(0) and λ(1) were in correspondence to the protonated state and deprotonated state respectively. The potential energy function, $\langle\frac{\partial V(\lambda)}{\partial\lambda}\rangle_{\lambda(0\text{-}1)}$, was applied to calculate the change in Gibbs free energies.[17] The evaluation of ΔG's were generated using a nine-point Gaussian quadrature of λ values (0.01592, 0.08198, 0.19331, 0.33787, 0.5, 0.98408, 0.91802, 0.80669, 0.66213), in which case, each Gaussian point simulation was independently processed.[47]

The partial charges of D-arginine and the deprotonated form of Tyr53 were established via two-step RESP method[45] with the electrostatic potentials from Gaussian03.[46] The FAD parameters were acquired from GAFF AMBER force field. Coordinate parameters for both A (open) and B (close) conformations of PaDADH, including the complexed imioarginine structure were incorporated within the same PDB file. The A conformation free in solution was constructed by removing all B configured residues and the iminoarginine. The B conformation in complex with D-arginine was established by deleting all A configured residues and modifying atom names of iminoarginine into D-arginine.

All simulations were carried out under the same force field parameters.[48,49] The TIP3P[55] water model 10 Å octahedron box was used to explicitly solvate all systems. Prior to the

simulations of each enzyme complex, sodium ions were added to establish neutralization. Each model systems were capped with N-terminal (N-acetyl, ACE) and C-terminal (N-methylamide, NME) protecting groups.[47]

Each lambda simulated trajectory was carried out within two rounds of minimizations (1000 steps each) with harmonic constraints in decreasing order of 200 kcal/mol/Å and 100 kcal/mol/Å. The harmonics constraints were parameterized on the enzyme and its corresponding ligands (FAD & substrate) excluding ions and water molecules. Once minimized all systems were then equilibrated at a constant temperature, and pressure of 300K, and 1 bar respectively. A set of harmonic constraints in decreasing order of 50 kcal/mol/Å, 25 kcal/mol/Å and 0 kcal/mol/Å was carried out during the equilibration process. Each constraint was applied for 0.2 nsec with a time step of 2 fsec on PaDADH, its cofactor (FAD), and D-arginine when present. The model systems were simulated using the same trajectory parameters including identical constraints on the residue of interest and its protecting groups.

Temperature regulation was processed via Langevin thermostat with a collision frequency of 1 $psec^{-1}$. A time step of 2 fsec was designated to generate Newton's equation of motion. A cut off of 9 Å was established for all long-range nonbonded interactions, in which case, electrostatic interactions were evaluated via the particle mesh Ewald (PME) method.[50] Bonds involving hydrogen atoms were constrained as a consequence of the SHAKE algorithm.[51]Each simulation was carried out for 1 ns, in which case, the MD trajectories were individually repeated for a total of three trials. Replication of trials was necessary to generate the average Gibbs free energy values, and standard deviations of thermodynamic integration calculations.

4.3 Results & Discussions

4.3.1 Open conformation free in solution (A)

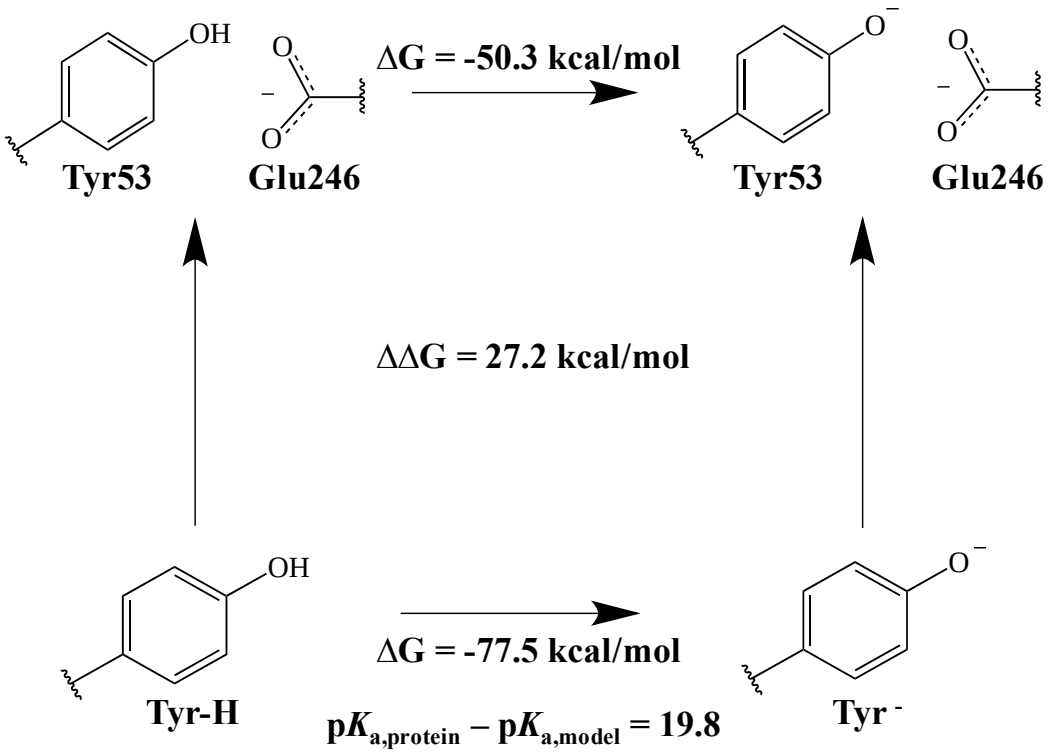


Figure 4.1 Thermodynamic cycle involving PaDADH Tyr53 (A).
Deprotonation within the protein included the enzyme free in solution, in the open conformation, with its cofactor FAD. Tyr53 and Glu246 are the only components shown. Deprotonation within the model system only involved the amino acid tyrosine.

Removing a proton from the hydroxyl group on the side chain of Tyr53 yielded an average Gibbs free energy of -50.3 kcal/mol. Since such quantities were lower in magnitude than the average Gibbs free energy of the model, -77.5 kcal/mol, the $pK_{a,shift}$ was assessed to be positive. A shift in pK_a of 19.8 indicated a high favorability towards protonation for PaDADH Tyr53.

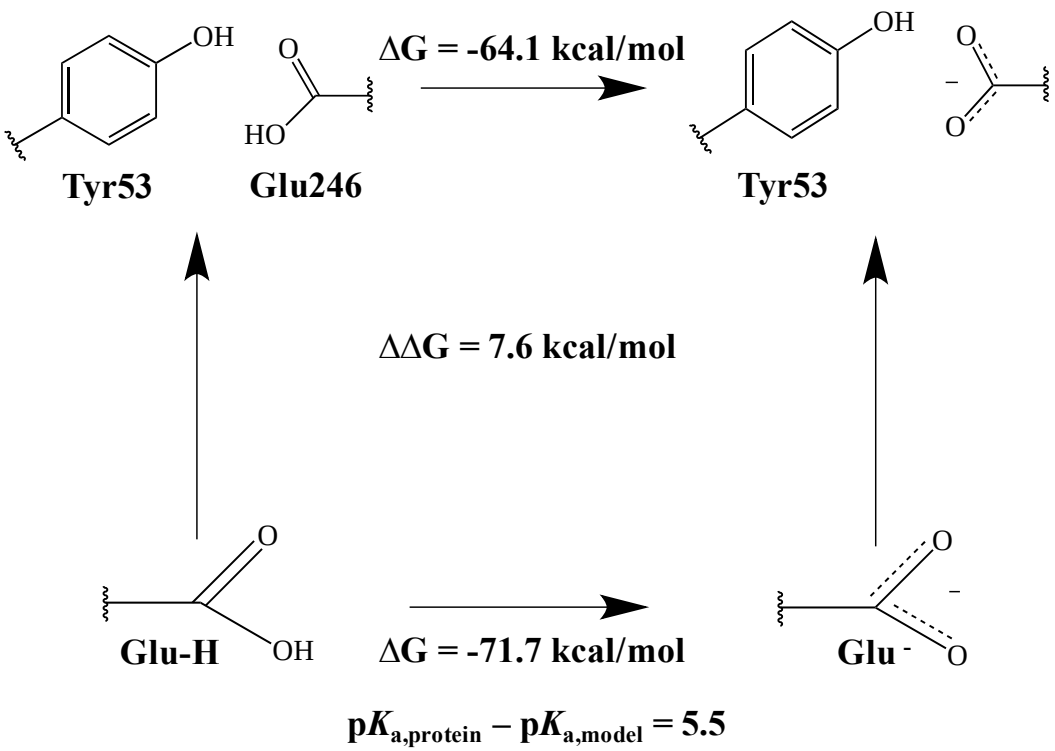


Figure 4.2 Thermodynamic cycle involving PaDADH Glu246 (A).
Deprotonating Glu246 in the protein included the unbound conformation of the haloenzyme without its substrate. Within the model system the carboxyl group on the R chain of glutamic acid was deprotonated while in solution.

Deprotonation of the side chain carboxyl group of Glu246 generated an average value of -64.1 kcal/mol in Gibbs free energy. This resulted into a lower magnitude than the average Gibbs free energy of the model system (-71.7 kcal/mol). Thus, the $pK_{a,shift}$ of the proton on the carboxyl group of Glu246 was a positive value of 5.5, indicating a favorability towards protonation. Such finding was unusual, since the side chain of glutamic acid has a standard pKa value of ~ 4[56] and is more likely to be deprotonated in solution.

4.3.2 Close conformation in complex with D-arginine (B)

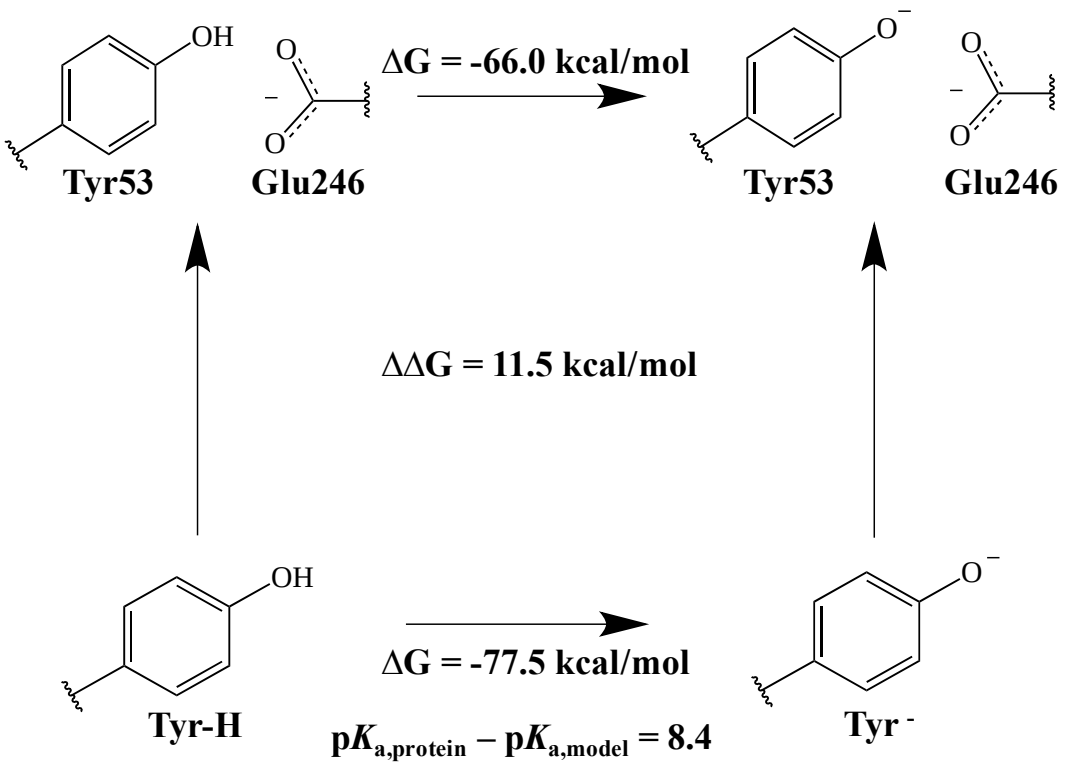


Figure 4.3 Thermodynamic cycle involving PaDADH Tyr53 (B).
Deprotonation of Tyr53 in the protein included the halo form of PaDADH in complex with D-arginine, in the close configuration of the enzyme. Tyrosine was the only amino acid considered in the model system, in which case, the hydroxyl group on its R chain was the only component evaluated for deprotonation.

Proton removal of the hydroxyl group of Tyr53, produced an average Gibbs free energy of -66.0 kcal/mol. Since the model system had a negative average Gibbs free energy of -77.5 kcal/mol, the $pK_{a,shift}$ was generated to be 8.4, showing favorability towards protonation. Nevertheless, this protonated favorability seems to be lower in magnitude in relative to Tyr53 state A $pK_{a,shift}$ results (~ 19.8).

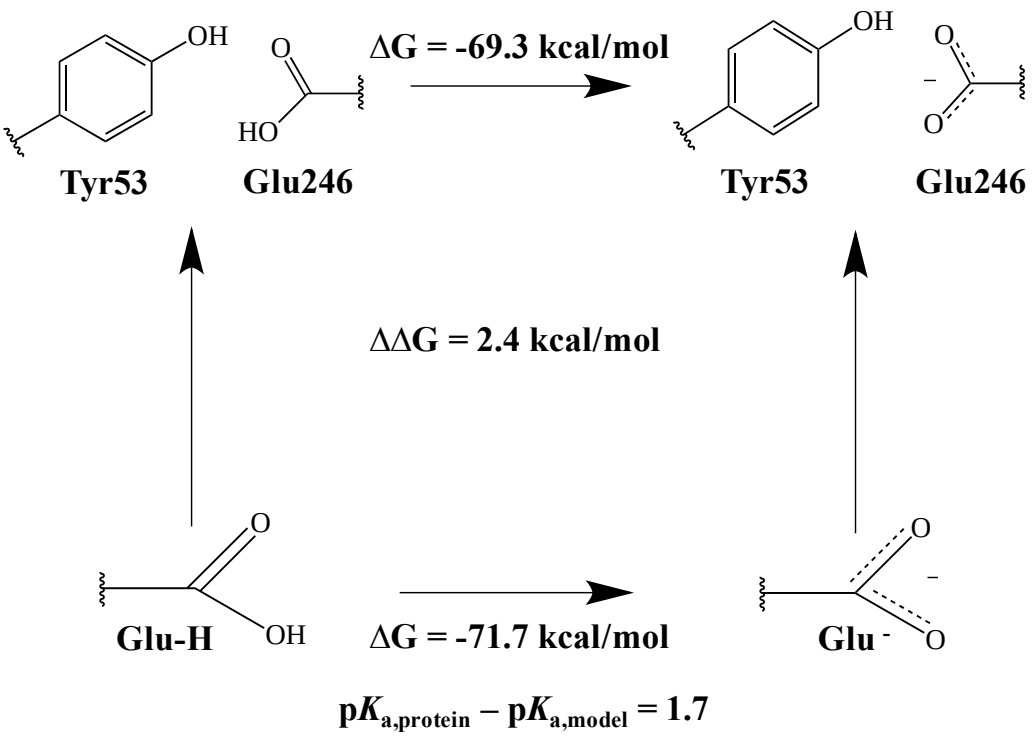


Figure 4.4 Thermodynamic cycle involving PaDADH Glu246 (B).
Glu246 deprotonation was evaluated as an enzyme residue in the close conformation of PaDADH in complex with D-arginine, and its FAD cofactor. Deprotonation within the model system only included glutamic acid.

Removing the side chain carboxyl proton of Glu246 generated an average value of -69.3 kcal/mol in Gibbs free energy. A $pK_{a,shift}$ of 1.7 was evaluated, since the average Gibbs free energy of the model was assessed to -71.7 kcal/mol. Thus, Glu246 showed favorability towards protonation as an enzymatic residue in the PaDADH-FAD-substrate complex, which is in contrast with its standard behavior in solution ($pK_a \sim 4$).[56] However, its protonation favorability is less pronounced in comparison to the generated value (~ 5.5) of wild type PaDADH Glu246 with the enzyme free in solution in the open conformation.

4.3.3 Individual trial thermodynamic integration values

Table 4.1 Quantitative analysis of thermodynamic integration values.
For each probing proton of interest, three trials were carried out, in which case, the averages and standard deviations were assessed. The difference between the average Gibbs free energies were calculated to solve for the $pK_{a,shift}$ of deprotonation within the protein complex in relative to the model system.

Structure	Trial 1 (kcal/mol)	Trial 2 (kcal/mol)	Trial 3 (kcal/mol)	Average (kcal/mol)	Stdev	$Pk_{a,shift}$
Tyr53_protein (A)	-50.67	-50.44	-49.77	-50.29	0.47	19.84
Glu246_protein (A)	-63.62	-64.24	-64.46	-64.11	0.44	5.51
Tyr53_protein (B)	-66.01	-66.03	-66.03	-66.03	0.01	8.38
Glu246_protein (B)	-69.17	-69.56	-69.16	-69.30	0.23	1.73
Tyr_model	-77.72	-77.97	-76.91	-77.53	0.56	-
Glu_model	-71.66	-71.68	-71.67	-71.67	0.01	-

This table represent the individual processed trials of Gibb's free energy evaluations for $pK_{a,shift}$ calculations. The thermodynamic integration is a stochastic approach, hence the calculated Gibbs free energy values for a multitude of trials. Therefore, an average estimation with a low value in standard deviation was critical for proper assessment of $pK_{a,shifts}$.

4.4 Conclusions

PaDADH Tyr53 and Glu246 showed favorability towards protonation in both the close conformation of the enzyme while in complex with D-arginine and the open conformation of the enzymatic structure free in solution. However, the spontaneity to be protonated for both residues lowered noticeably from the ligand-free configuration of PaDADH free in solution to its product-bound state in complex with D-arginine.

5 CONCLUSIONS

5.1 Concluding Remarks

In order to gain further insight on the gated mechanism of PaDADH, a comparison between the wild type and mutated enzyme could be undertaken using an analogy. Imagine a room with an open door, in which case, the room and the door is a representation of PaDADH and its "active site lid" respectively. Over time, the door of this room is revolving at a hypothetical probability distribution, which was the case for wild type PaDADH free in solution. The opening and closing of this door is analogous to the ligand-free and product-bound conformations of the enzyme correspondingly. Upon ligand entry inside this room, during catalysis, the door is more likely to close to enhance enzymatic activity, as seen with PaDADH in complex with D-arginine. Buried water molecules within the "active site lid" over time, would be equivalent to the lubricant on the hinges of the door. Thus, in the presence of water, the door would be able to open or close with less resistance. However, once a ligand is inside this room, during catalysis, there would be a preference for closing the door. This would result into a higher probability of water absence to add resistance against the opening of the door.

By considering the same representation for the single amino acid mutations of PaDADH, a greater understanding on the mechanical aspect of the enzyme may be developed. The door knob would be the equivalence of PaDADH Tyr53 of the enzyme's "active site lid". The top or bottom hinge of the door would be analogous to PaDADH Ser45, whereas the middle hinge would correspond to PaDADH Ala46. Mutating Tyr53 into phenylalanine could be envisioned as modifying the knob of the door by removing its latch. Thus, a Y53F mutation of PaDADH will likely remained open, since the latch and strike plate interaction of the door is prevented. The S45A mutant variant of this enzyme would be similar to detaching either the top or bottom hinge

of the door to this room, which would cause the door to tilt and shift its probability distribution towards the open conformation. Consequently, detaching the middle hinge (A46G) of the door would result into an even greater tilt and a higher likelihood to its open configuration.

The results of the shifts in pK_a analyses could further be evaluated through the appliance of the Henderson Hasselbalch equation.[57] This approach is based on the probability of finding either the salt or acid form of a molecule in solution. For instance, an increase or decrease in pH corresponds with a higher or lower possibility of locating the deprotonated form of a compound in its environment. This base/acid ratio is established on the log of base ten and the pK_a value (equation 1). Thus, a small difference between pH and pK_a is significant to shift the protonation state of a molecule.

$$pH = pK_a + \log\frac{[Base]}{[Acid]} \qquad (1)$$

The $pK_{a,shift}$ of Tyr53 while the enzyme is in its open conformation free in solution and in complex with D-arginine in its close conformation, showed favorability towards protonation. In addition, kinetic results from the Y53F mutant variant of PaDADH[53] disproved the hypothesis of Tyr53 being the catalytic enzymatic residue which triggered the oxidative deamination reaction of D-amino acids.

The shift in pK_a of Glu246 in both the close state of the enzyme in complex with D-arginine and the open state free in solution was evaluated to be ~ 1.7 and ~ 5.5 respectively. According to pH profile analyses on the kinetics of PaDADH, the amino acid residue responsible for triggering the electron delocalization of D-amino acids was suggested to have a pK_a of 9.6.[5] Addition of Glu246 $pK_{a,shift}$ estimation (~ 5.5) with the standard pK_a value of a glutamic acid side chain (~ 4) generates a pK_a of ~ 9.5, which is comparable with kinetic pH profile results. Following a similar approach with PaDADH in complex with D-arginine in its close

conformation, Glu246 has a different pK_a value of ~ 5.7. All MD simulations were carried out at an environmental pH of 7, thus, prior to enzyme-substrate binding PaDADH Glu246 is more likely to be protonated, which could be a necessity to maintain the enzyme's integrity and the probability distribution of its gated mechanism. However, during the formation of the enzyme-substrate complex, the deprotonated form of Glu246 is more favorable in order to remove a proton from the backbone amino group of the substrate through a relay of hydrogen bonds with water and/or surrounding enzymatic residues. In addition, a deprotonated Glu246 would form a hydrogen bond with Tyr53 that is less than ~ 3 Å away in the close configuration of the enzyme.[1] Therefore, Glu246 may indeed be the catalytic residue, which triggers the catalysis of PaDADH with D-amino acids.

6 REFERENCES

1. Fu, G.; Yuan, H.; Li, C.; Lu, C.-D.; Gadda, G.; Weber, I. T., Conformational Changes and Substrate Recognition in Pseudomonas aeruginosa d-Arginine Dehydrogenase. *Biochemistry* **2010,** *49* (39), 8535-8545.

2. Fu, G.; Yuan, H.; Wang, S.; Gadda, G.; Weber, I. T., Atomic-Resolution Structure of an N(5) Flavin Adduct in d-Arginine Dehydrogenase. *Biochemistry* **2011,** *50* (29), 6292-6294.

3. Yuan, H.; Fu, G.; Brooks, P. T.; Weber, I.; Gadda, G., Steady-State Kinetic Mechanism and Reductive Half-Reaction of d-Arginine Dehydrogenase from Pseudomonas aeruginosa. *Biochemistry* **2010,** *49* (44), 9542-9550.

4. Barman, A.; Smitherman, C.; Souffrant, M.; Gadda, G.; Hamelberg, D., Conserved Hydration Sites in Pin1 Reveal a Distinctive Water Recognition Motif in Proteins. *Journal of Chemical Information and Modeling* **2016,** *56* (1), 139-147.

5. Yuan, H.; Xin, Y.; Hamelberg, D.; Gadda, G., Insights on the Mechanism of Amine Oxidation Catalyzed by d-Arginine Dehydrogenase Through pH and Kinetic Isotope Effects. *Journal of the American Chemical Society* **2011,** *133* (46), 18957-18965.

6. Nosé, S., An extension of the canonical ensemble molecular dynamics method. *Molecular Physics* **1986,** *57* (1), 187-191.

7. Mackerell, A. D.; Feig, M.; Brooks, C. L., 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. *Journal of Computational Chemistry* **2004,** *25* (11), 1400-1415.

8. Piana, S.; Bucher, D.; Carloni, P.; Rothlisberger, U., Reaction Mechanism of HIV-1 Protease by Hybrid Car-Parrinello/Classical MD Simulations. *The Journal of Physical Chemistry B* **2004,** *108* (30), 11139-11149.

9. Kallen, J.; Spitzfaden, C.; Zurini, M. G. M.; Wider, G.; Widmer, H.; Wuthrich, K.; Walkinshaw, M. D., Structure of human cyclophilin and its binding site for cyclosporin A determined by X-ray crystallography and NMR spectroscopy. *Nature* **1991,** *353* (6341), 276-279.

10. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M., CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry* **1983,** *4* (2), 187-217.

11. Brooks, C. L., Simulations of protein folding and unfolding. *Current Opinion in Structural Biology* **1998,** *8* (2), 222-226.

12. Meller, J. a., Molecular Dynamics. In *eLS*, John Wiley & Sons, Ltd: 2001.

13. Mark, P.; Nilsson, L., Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *The Journal of Physical Chemistry A* **2001,** *105* (43), 9954-9960.

14. Jiang, W.; Roux, B., Free Energy Perturbation Hamiltonian Replica-Exchange Molecular Dynamics (FEP/H-REMD) for Absolute Ligand Binding Free Energy Calculations. *Journal of Chemical Theory and Computation* **2010,** *6* (9), 2559-2565.

15. Cerutti, D. S.; Duke, R. E.; Darden, T. A.; Lybrand, T. P., Staggered Mesh Ewald: An Extension of the Smooth Particle-Mesh Ewald Method Adding Great Versatility. *Journal of Chemical Theory and Computation* **2009,** *5* (9), 2322-2338.

16. Carlsson, J.; Boukharta, L.; Åqvist, J., Combining Docking, Molecular Dynamics and the Linear Interaction Energy Method to Predict Binding Modes and Affinities for Non-nucleoside

Inhibitors to HIV-1 Reverse Transcriptase. *Journal of Medicinal Chemistry* **2008,** *51* (9), 2648-2656.

17. Simonson, T.; Carlsson, J.; Case, D. A., Proton Binding to Proteins: pKa Calculations with Explicit and Implicit Solvent Models. *Journal of the American Chemical Society* **2004,** *126* (13), 4167-4180.

18. Sibanda, B. L.; Blundell, T. L.; Thornton, J. M., Conformation of β-hairpins in protein structures. *Journal of Molecular Biology* **1989,** *206* (4), 759-777.

19. Milner-White, E. J.; Poet, R., Four classes of beta-hairpins in proteins. *Biochemical Journal* **1986,** *240* (1), 289-292.

20. Hutchinson, E. G.; Thornton, J. M., The Greek key motif: extraction, classification and analysis. *Protein Engineering* **1993,** *6* (3), 233-245.

21. Nayak, D.; Sivaraman, J., Structural basis for the indispensable role of a unique zinc finger motif in LNX2 ubiquitination. *Oncotarget* **2015,** *6* (33), 34342-34357.

22. Hanas, J. S.; Hazuda, D. J.; Bogenhagen, D. F.; Wu, F. Y.; Wu, C. W., Xenopus transcription factor A requires zinc for binding to the 5 S RNA gene. *Journal of Biological Chemistry* **1983,** *258* (23), 14120-14125.

23. Berg, J. M., Zinc fingers and other metal-binding domains. Elements for interactions between macromolecules. *Journal of Biological Chemistry* **1990,** *265* (12), 6513-6516.

24. Klug, A., The discovery of zinc fingers and their development for practical applications in gene regulation and genome manipulation. *Quarterly Reviews of Biophysics* **2010,** *43* (01), 1-21.

25. Lewit-Bentley, A.; Réty, S., EF-hand calcium-binding proteins. *Current Opinion in Structural Biology* **2000,** *10* (6), 637-643.

26. O'Donnell, S. E.; Newman, R. A.; Witt, T. J.; Hultman, R.; Froehlig, J. R.; Christensen, A. P.; Shea, M. A., Chapter 21 Thermodynamics and Conformational Change Governing Domain–Domain Interactions of Calmodulin. In *Methods in Enzymology*, Academic Press: 2009; Vol. Volume 466, pp 503-526.

27. Nakayama, S.; Kretsinger, R. H., Evolution of the EF-Hand Family of Proteins. *Annual Review of Biophysics and Biomolecular Structure* **1994,** *23* (1), 473-507.

28. Lang, K.; Schmid, F. X.; Fischer, G., Catalysis of protein folding by prolyl isomerase. *Nature* **1987,** *329* (6136), 268-270.

29. Yaffe, M. B.; Schutkowski, M.; Shen, M.; Zhou, X. Z.; Stukenberg, P. T.; Rahfeld, J.-U.; Xu, J.; Kuang, J.; Kirschner, M. W.; Fischer, G.; Cantley, L. C.; Lu, K. P., Sequence-Specific and Phosphorylation-Dependent Proline Isomerization: A Potential Mitotic Regulatory Mechanism. *Science* **1997,** *278* (5345), 1957-1960.

30. Bayer, E.; Goettsch, S.; Mueller, J. W.; Griewel, B.; Guiberman, E.; Mayr, L. M.; Bayer, P., Structural Analysis of the Mitotic Regulator hPin1 in Solution: INSIGHTS INTO DOMAIN ARCHITECTURE AND SUBSTRATE BINDING. *Journal of Biological Chemistry* **2003,** *278* (28), 26183-26193.

31. Schutkowski, M.; Bernhardt, A.; Zhou, X. Z.; Shen, M.; Reimer, U.; Rahfeld, J.-U.; Lu, K. P.; Fischer, G., Role of Phosphorylation in Determining the Backbone Dynamics of the Serine/Threonine-Proline Motif and Pin1 Substrate Recognition. *Biochemistry* **1998,** *37* (16), 5566-5575.

32. Ranganathan, R.; Lu, K. P.; Hunter, T.; Noel, J. P., Structural and Functional Analysis of the Mitotic Rotamase Pin1 Suggests Substrate Recognition Is Phosphorylation Dependent. *Cell* **1997,** *89* (6), 875-886.

33. Wilson, K. A.; Bouchard, J. J.; Peng, J. W., Interdomain Interactions Support Interdomain Communication in Human Pin1. *Biochemistry* **2013,** *52* (40), 6968-6981.

34. Lu, K. P.; Zhou, X. Z., The prolyl isomerase PIN1: a pivotal new twist in phosphorylation signalling and disease. *Nat Rev Mol Cell Biol* **2007,** *8* (11), 904-916.

35. Zhang, J.; Grigoryan, G., Mining Tertiary Structural Motifs for Assessment of Designability. *Methods in enzymology* **2013,** *523*, 21-40.

36. Barman, A.; Hamelberg, D., Cysteine-Mediated Dynamic Hydrogen-Bonding Network in the Active Site of Pin1. *Biochemistry* **2014,** *53* (23), 3839-3850.

37. Xu, N.; Tochio, N.; Wang, J.; Tamari, Y.; Uewaki, J.-i.; Utsunomiya-Tate, N.; Igarashi, K.; Shiraki, T.; Kobayashi, N.; Tate, S.-i., The C113D Mutation in Human Pin1 Causes Allosteric Structural Changes in the Phosphate Binding Pocket of the PPIase Domain through the Tug of War in the Dual-Histidine Motif. *Biochemistry* **2014,** *53* (34), 5568-5578.

38. Wang, J.; Tochio, N.; Kawasaki, R.; Tamari, Y.; Xu, N.; Uewaki, J.-i.; Utsunomiya-Tate, N.; Tate, S.-i., Allosteric Breakage of the Hydrogen Bond within the Dual-Histidine Motif in the Active Site of Human Pin1 PPIase. *Biochemistry* **2015,** *54* (33), 5242-5253.

39. Szep, S.; Park, S.; Boder, E. T.; Van Duyne, G. D.; Saven, J. G., Structural coupling between FKBP12 and buried water. *Proteins: Structure, Function, and Bioinformatics* **2009,** *74* (3), 603-611.

40. Park, S.; Saven, J. G., Statistical and molecular dynamics studies of buried waters in globular proteins. *Proteins: Structure, Function, and Bioinformatics* **2005,** *60* (3), 450-463.

41. Lu, C.-D., Pathways and regulation of bacterial arginine metabolism and perspectives for obtaining arginine overproducing strains. *Applied Microbiology and Biotechnology* **2006,** *70* (3), 261-272.

42. Li, C.; Lu, C.-D., Arginine racemization by coupled catabolic and anabolic dehydrogenases. *Proceedings of the National Academy of Sciences* **2009,** *106* (3), 906-911.

43. Gannavaram, S.; Gadda, G., Relative Timing of Hydrogen and Proton Transfers in the Reaction of Flavin Oxidation Catalyzed by Choline Oxidase. *Biochemistry* **2013,** *52* (7), 1221-1226.

44. Umhau, S.; Pollegioni, L.; Molla, G.; Diederichs, K.; Welte, W.; Pilone, M. S.; Ghisla, S., The x-ray structure of d-amino acid oxidase at very high resolution identifies the chemical mechanism of flavin-dependent substrate dehydrogenation. *Proceedings of the National Academy of Sciences* **2000,** *97* (23), 12463-12468.

45. Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A., A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *The Journal of Physical Chemistry* **1993,** *97* (40), 10269-10280.

46. Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara,

A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A., Gaussian 03. **2004**.

47. D.A. Case, V. B., J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T.S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K.M. Merz, F. Paesani, D.R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu and P.A. Kollman AMBER 14. **2014**.

48. Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M., Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of α/γ Conformers. *Biophysical Journal* **2007,** *92* (11), 3817-3829.

49. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A., A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules J. Am. Chem. Soc. 1995, 117, 5179−5197. *Journal of the American Chemical Society* **1996,** *118* (9), 2309-2309.

50. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G., A smooth particle mesh Ewald method. *The Journal of Chemical Physics* **1995,** *103* (19), 8577-8593.

51. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C., Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics* **1977,** *23* (3), 327-341.

52. Faust, A.; Niefind, K.; Hummel, W.; Schomburg, D., The Structure of a Bacterial l-Amino Acid Oxidase from Rhodococcus opacus Gives New Evidence for the Hydride Mechanism for Dehydrogenation. *Journal of Molecular Biology* **2007,** *367* (1), 234-248.

53. Gannavaram, S.; Sirin, S.; Sherman, W.; Gadda, G., Mechanistic and Computational Studies of the Reductive Half-Reaction of Tyrosine to Phenylalanine Active Site Variants of d-Arginine Dehydrogenase. *Biochemistry* **2014,** *53* (41), 6574-6583.

54. Haas, D.; Matsumoto, H.; Moretti, P.; Stalon, V.; Mercenier, A., Arginine degradation in Pseudomonas aeruginosa mutants blocked in two arginine catabolic pathways. *Molecular and General Genetics MGG* **1984,** *193* (3), 437-444.

55. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **1983,** *79* (2), 926-935.

56. Bashford, D.; Case, D. A.; Dalvit, C.; Tennant, L.; Wright, P. E., Electrostatic calculations of side-chain pKa values in myoglobin and comparison with NMR data for histidines. *Biochemistry* **1993,** *32* (31), 8045-8056.

57. Po, H. N.; Senozan, N. M., The Henderson-Hasselbalch Equation: Its History and Limitations. *Journal of Chemical Education* **2001,** *78* (11), 1499.

58. Ouedraogo, D., Mechanistic Investigation of the Flavin-Neighboring Residues S45A, A46G and I335 in Pseudomonas Aeruginosa D-arginine Dehyfrogenase. *Thesis Master* **2015**, 1-74.