

**INVESTIGATING THE BINDING-RELEASE MECHANISM OF PERIPLASMIC  
FERRIC BINDING PROTEIN BY pH VARIATIONS AND POINT MUTATIONS**

by

**Gökçe GÜVEN**

Submitted to the Graduate School of Engineering and Natural Sciences

in partial fulfillment of

the requirements for the degree of

Master of Science

**Sabancı University**

**August, 2013**

INVESTIGATING THE BINDING-RELEASE MECHANISM OF PERIPLASMIC TERRIC  
BINDING PROTEIN BY pH VARIATIONS AND POINT MUTATIONS

APPROVED BY

Prof. Canan Atılgan (Thesis Supervisor)



Prof. Ali Rana Atılgan (Thesis Co-Supervisor)



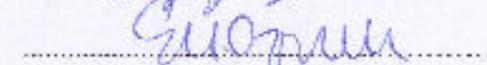
Prof. Dr. Levent Demirel



Assist. Prof. Dr. Alpay Taralp



Assist. Prof. Elif Özkırımılı Ölmez



DATE OF APPROVAL:

01.08.2013

© Gökçe GÜVEN 2013

All Rights Reserved

*To celestials,*

# INVESTIGATING THE BINDING-RELEASE MECHANISM OF PERIPLASMIC FERRIC BINDING PROTEIN BY pH VARIATIONS AND POINT MUTATIONS

Gökçe GÜVEN

Materials Science & Engineering, Master of Science Thesis, 2013

Thesis Supervisor: Prof. Dr. Canan ATILGAN

Thesis Co-Supervisor: Prof. Dr. Ali Rana ATILGAN

Keywords: Perturbation Response Scanning (PRS), Molecular Dynamics (MD), Gram-negative bacteria, Haemophilus Influenza, Periplasmic Ferric Bindin Proteins, pH, point, ionization strength, mutation, protonation

## Abstract

Differences between the conformations of the ligand-bound and unbound forms of proteins provide clues on deciphering residues that have a direct effect on binding mechanisms. Using molecular dynamic (MD) simulations as a basis, conformational changes that take place on time scales much slower than those accessible by MD may be investigated by supplementary methods such as perturbation response scanning (PRS). Since proteins are complex macromolecules, to sample their conformational energy landscapes we applied both global and local perturbations in different combinations. Global perturbations are related with environmental changes such as (i) different values of ionization strength of the solution which mimic the various salt environments experienced by the protein *in vivo*, (ii) or protonation of a group of residues to mimic a different pH environment. Local perturbations are related to specific point perturbations on protein such as, (i) protonation or (ii) mutation of a single residue to locate the points that controls a conformational change in proteins.

In this work, the apo and holo ( $\text{Fe}^{3+}$  bound) forms of the periplasmic ferric binding protein systems (hFBP) of the gram negative bacteria haemophilus influenzae was selected as the model system. PRS studies showed us that D52 and D47 residues are the ones that give the highest value for the fractional contribution of the eigenvalue of the response matrix. Therewithal, using pKa calculations, a particular charged residue (D52) (out of a total of 98) was found to be the most sensitive to subtle pH variations in the physiological range. The effect of single point protonation (D52+) and mutation (D52A) is investigated via a series of MD simulations. The effect of IS (0.15 mM and ~0 mM) and pH (5 and 7.4 ) change is also studied to monitor the conformations sampled. For apo FBP, the kinetics of synergistic anion binding and release was monitored and was consistently manipulated by varying these conditions, while no apparent conformational change was observed in the protein. For holo FBP, protonated and Ala mutants of D52 consistently trigger opening of the iron binding site in an ensemble of simulations, while elevated IS consistently traps the closed forms. We categorized our series of MD trajectories as open, partially open, and closed due to coordination level and mapping of iron ion inside the active site. Our results lend clues as to how the environment versus single residue perturbations may be utilized for regulation of binding modes in hFBP systems.

PERİPLASMIK DEMİR BAĞLAYAN PROTEİNDE BAĞLANMA-BIRAKMA  
MEKANİZMASININ pH VARYASYONLARI VE TEK NOKTA MUTASYONLARI İLE  
İNCELENMESİ

Gökçe GÜVEN

Malzeme Bilimi ve Mühendisliği, Yüksek Lisans Tezi, 2013

Tez Danışmanı: Prof. Dr. Canan ATILGAN

Tez Eşdanışmanı: Prof. Dr. Ali Rana ATILGAN

Anahtar Kelimeler: Etki-Tepki Taraması (ETT), Moleküler Dinamik (MD), Gram-negatif bakteri, Haemophilus İnfluanza, periplazmik demir bağlayan protein, pH, nokta mutasyonu

Özet

Proteinlerin ligand-bağlı ve bağımsız konformasyon modelleri arasındaki farklar, bağlanma mekanizmasında dolaysız kontrolü olan rezidülerin yerlerinin tespitinde önemli ipuçları sağlarlar. Moleküler dinamik (MD) benzetimlerine dayanarak olagelen konformasyonel değişimlerden, MD tarafından elde edilemeyecek kadar yavaş zaman-ölçeklerinde gerçekleşen kiplerin incelenmesi etki-tepki taraması (ETT) gibi destekleyici yöntemlerle yapılabilir. Proteinler girişik iril özdecikler olduğundan, konformasyonel enerji peyzajlarını modellemek için yerel ve evrensel sarsımların farklı fombinasyonlarını uyguladık. Evrensel adıyla belirttiğimiz sarsımlar proteinin bulunduğu ortamın özellikleriyle ilişkilendirilmiştir: (i) proteini canlandıran akışkan ortamın yükünel erkinin farklı ölçekleri, (ii) farklı pH ortamlarının taklidi için belirli bir rezidü öbeğinin protonasyonu gibi. Yerel sarsımlar ise proteindeki özgül noktalarda gerçekleşir: (i) özgül rezidünün protonasyonu veya (ii) mutasyonu ile proteinlerdeki konformasyonel değişimi modellemek mümkündür.

Bu çalışmada, haemophilus influenzae sayırlı bakterisinin periplazmik demir bağlayan protein modelleri hFbpA'nın açık ve kapalı ( $Fe^{3+}$  bağlı) formları çalışılan sistemler olarak seçilmiştir. ETT çalışmalarından D52 ve D47 numaralı rezidülerin tepki matrisinin azami gizdeğerine en yüksek oranda katkı yapan noktalar olarak belirlenmişti. Keza, proteinin pKa hesaplamasında da 98 yüklü rezidü arasından D52 numaralı rezidü ince pH değişimlerinde fizyolojik aralıkta en hassas nokta olarak bulundu. Burada özgül nokta protonasyonlarının (D52+) ve mutasyonlarının (D52A) etkileri MD benzetimi dizileriyle belirlenmiştir. Yükünel erklerin (0.15 mM and ~0 mM) ve sistemlerdeki pH değişimleri de (5 ve 7.4) örneklenen konformasyonları denetlemek için çalışılmıştır. Açık hFbp sistemleri için,  $H_2PO_4^-$  anyonunun bağlanma-kopma hareketinin kinetiğinin değişimi sistemlere uygulanan sarsımların farklı kombinasyonlarıyla kontrol edilebildi ve yapılan analizler sonucunda bağlanma-kopma kinetiği sırasında belirli bir konformasyonel değişikliğe rastlanmadı. Kapalı hFBP sistemlerinde, yüksek yükünel erkli ortama atılan protein MD benzetimlerde kapalı konformasyonda kalmayı tercih ederken, D52 numaralı rezidünün protonlanmış ve Ala mutasyonlarının MD benzetimleri sonucunda gözlemlenen demirin bağlandığı kovuktaki menteşe hareketinin açık konformasyona doğru kayması ETT sonuçlarını desteklemektedir. Demir iyonunun moleküler dinamik benzetimi gezintileri boyunca gözlenen düzenleşim seviyeleri ve komşu rezidülerdeki eşleşmelerine göre elde edilen sistemler açık, yarı açık, ve kapalı olarak adlandırılmıştır. Sonuç olarak, evrensel ve yöresel sarsımların hFbpA proteininin bağlanma hareketlerinin üzerindeki etkilerini incelemiş olduk.

## ACKNOWLEDGEMENTS

As a profound prelude, I would like to express my special thanks and gratitude to my supervisor Prof. Dr. Canan Atilgan who has always been a source of patient guidance, invaluable advices, motivation and inspiration. I know that my concentration level is not always on the top and I know I am able to write this thesis due to Prof. Canan Atilgan's and Prof. Ali Rana Atilgan's steadfastness. That is why I felt that I will concentrate on a problem with passion and come to the lab when the moon rises above the sky and sip my coffee until I saw the morning glories blooms.

I would specially thank to my thesis committee member Assist. Prof. Alpay Taralp and Assist. Prof. Elif Özkırımlı Ölmez for their contributions. I would like to give special thanks to Prof. Levent Demirel and Prof. Mehmet Somer for their support for me and their contribution to my academic career. I would like to thank to the professors that I called them secretly "The Three Musketeers" Prof. Levent Demirel, Prof. Osman Uğur Sezerman and Prof. Mehmet Ali Gülgün also. I wish to extend my thanks to all faculty members of Material Science and Engineering Program for their support and understanding.

I would like to acknowledge Ayşe Özlem Sezerman for teaching me the computational tools I know today and helping me in all points of my study. I convey special acknowledgement to and for their valuable friendship and support. And then there are all the other people who have made Sabanci University a very special place over two years: Kaan Bilge, Meral Yüce, Hasan Kurt, Melike Mercan Yıldızhan, Kinyas Aydın, Yiğit Uysallı, Doğa Ortakoylu, İpek Erdem, "Alice & Alex duality" and Elif Özden Yenigün.

My deepest gratitude goes to my mother, my father, and sisters and brothers in law Hüseyin Çebi, Tanır Özçelebi and my nephew Çınar for their excessive love and superior support throughout my life; this dissertation is simply impossible without them. I am really a lucky to be a "tekne kazıntısı" of my family, and "Minik Kuş" for my sisters. I am such a jammy homosapien to meet these people I mentioned in this acknowledgement.

Finally, I kindly acknowledge the Scientific and Technological Research Council of Turkey Project (grant 110T624) for providing scholarship during my master study.

## Table of Contents

List of Figures.....	vii
List of Tables.....	viii
Abbreviations.....	ix
CHAPTER 1. Introduction .....	1
1.1. Iron Saga .....	1
1.2. Background .....	2
1.2.1. Ferric Binding Phenomena .....	2
1.2.2. Suggested Mechanisms of Transportation of Iron Through Periplasm .....	3
1.2.3. Crystal Structures of Periplasmic FBPs in Gram-Negative Bacterium.....	6
1.3. Problem Statement .....	8
CHAPTER 2. Theoretical and Computational Methods .....	9
2.1. Haemophilus Influenza Periplasmic Ferric Binding Protein Systems (hFbpA) .....	9
2.2. System Construction for Molecular Dynamics Simulations .....	10
2.3. Local and Global Perturbation Tools .....	12
2.3.1. Linear Response Theory .....	12
2.3.2. Perturbation-Response Scanning .....	16
CHAPTER 3. Results and Discussion .....	17
3.1. Analysis of Apo hFBP .....	17
3.1.1. MD Simulation Trajectories for Apo hFBP.....	19
3.1.2. RMSF and RMSD of trajectories .....	23
3.1.3. H <sub>2</sub> PO <sub>4</sub> <sup>-</sup> Bound-Unbound Conformational Search.....	29
3.1.4. Cross Correlations between residues and Correlation between simulations .....	31
3.2. Analysis of Holo hFBP.....	33
3.2.1. MD Simulation Trajectories for Holo hFBP.....	33
3.2.2. Response to the D52 Perturbation by Using Variance-Covariance Matrices.....	37
CHAPTER 4. Conclusion and Future Work.....	42
Appendix.....	45
References.....	53

## LIST OF FIGURES

- Figure 1.1.** Cell Structure differences between gram-negative and gram-positive bacteria
- Figure 1.2.** Ferric ion transport systems in gram-negative bacteria
- Figure 2.1.** Three dimensional structures of the proteins
- Figure 2.2.** hFbpA proteins apo and holo forms
- Figure 2.3.** Free-body diagram of a residue
- Figure 3.1.** Suggested free ion transportation model from transferrin to the cytosol in pathogenic Neisseria and H. influenza
- Figure 3.2.**  $\text{H}_2\text{PO}_4^-$  Binding Sites in Apo-hFbpA.
- Figure 3.3.** Degree of Ionization for apo and holo hFbpA
- Figure 3.4.** Time-Distance ( $\text{H}_2\text{PO}_4^-$  (P) between N175(N)) Graphs for 1d9v Simulations
- Figure 3.5.** RMSF graph for 1D9V (low ionization level) first run
- Figure 3.6.** RMSF graph for pH5, D52A, and original (wild type) simulations of 1D9V
- Figure 3.7.** Cross correlations for original, D52A, and pH5 type 1D9V simulations
- Figure 3.8.** Correlation graph for 1D9V (IS-2) and 1D9V (pH5-2) runs
- Figure 3.9.** 1D9V-1 chunk1 mode 1 and 1D9V-1 chunk 5 mode 1
- Figure 3.10.** Models assigned from MD trajectories of hFbpA
- Figure 3.11.** Radial distribution function graph for the holo hFbpA systems in high IS
- Figure 3.12.** Suggested potential well graph for holo hFbpA simulations

## LIST OF TABLES

**Table 1.1** Crystal Structures of hFbpA

**Table 3.1.** List of MD Simulation Trajectories for Apo hFBP

**Table 3.2.** In-out movementt of  $\text{H}_2\text{PO}_4^-$

**Table 3.2.1** List of MD Simulation Trajectories for Holo hFBP

**Table 3.2.2** List of MD Simulation Trajectories

**Table 3.3.** Comparisions of Bound and Bound (B&B) Time Chunks for Apo Protein by Vector Analysis

**Table 3.4.** Comparisions of Unbound and Unbound (U&U) Time Chunks for Apo Protein by Vector Analysis

**Table 3.5.** Comparisions of Bound and Unbound (B&U) Time Chunks for Apo Protein by Vector Analysis

## ABBREVIATIONS

hFbpA	Haemophilus influenza periplasmic ferric binding protein
RMSD	root mean square deviation
RMSF	root mean square fluctuation
MD	molecular dynamics
PRS	perturbation response scanning
1MRP	Protein Data Bank ID for the crystal structure of wild type hFbpA with Fe <sup>3+</sup>
1D9V	Protein Data Bank ID for the crystal structure of wild type hFbpA without Fe <sup>3+</sup>
hTf	Human transferrin
IS	Ionic Strength
ALA	alanine (A)
ARG	arginine (R)
ASN	asparagine (N)
ASP	aspartic acid (D)
CYS	cysteine (C)
GLN	glutamine (Q)
GLU	glutamic acid (E)
GLY	glycine (G)
HIS	histidine (H)
ILE	isoleucine (I)
LEU	leucine (L)
LYS	lysine (K)
MET	methionine (M)
PHE	phenylalanine (F)
PRO	proline (P)
SER	serine (S)

THR	threonine (T)
TRP	tryptophan (W)
TYR	tyrosine (Y)
VAL	valine (V)

## CHAPTER 1

### Introduction

#### 1.1. Iron Saga

One may liken the functioning of a sequence of amino acids to a necklace with magnetic points that have different electric fields on each of them, the thread is going to fold in a specific way in space to the form of the necklace. In the limit of the perception level of a network of nerve cells, it is not easy to comprehend and solve the dilemma behind the mechanism just by looking from one direction. The question for this work is from which perspective I have to look to a semi light-tight glass so that I will see the inside of the house and see the never-ending parallel mirrors. Parallel mirrors is the moon-glade of the cognizance of science which are the self-sculpturing sand grains of knowledge built by never-ending questioning. Our perspective starts with locating a specific point that will vociferate more among other points when I hit every point one by one with equal random forces so that the other points will drift with the tide and adjust their tone of voice according to the supervisory audio tone. That point is a control point for the overall protein system. Now I will start with playing the dewlap that I am going to place the necklace on. In reality, this is the environment full of water molecules with different concentration of ions that I put my protein in.

Now, in the name of "Joy of Science", Caelum Caeruleo Confido is going to tell you a story from a totally different aspect, just for the motivation. In days of yore, there was a box that had a flying castle (protein) swimming in the sea (molecules making up the solution). In this castle we have the prince, whose name was "ferric ion." We have two twin magicians which are called Y195/Y196 and a princess who loves to escape from the castle secretly from time to time,  $\text{H}_2\text{PO}_4^-$ . Without the  $\text{Fe}^{+3}$ , it was not possible to keep peace in the periplasmic universe. While the prince was away from the castle, the flying castle always kept the gate in an open state. We also have loyal knights which are called  $\text{Cl}^-$  that ride their Sanctus

Chimeras in the opalescent sky of the protein castle. This array of opalescent sky is the reflection of the pH change of our flying castle. When the sky is in high guardianization (IS) level, or when the twin magicians (Y195/196) are activated (deprotonated), the gate is closed so the quicksilver prince cannot leave the castle. The prince cannot return to the castle without the guidance of the no shrinking violet princess. Because our prince gets drunk immediately after leaving the castle from the gate, it is impossible for him to find the way back; he is just too active outside. After all has been said and done they lived in mystery until a scientist with a purple motorbike came up with a solution for this stargazer prince to wake up from his lucid dream and leave the castle for eternity to see the outermembrane universe.

## **1.2. Background**

A problem will have more than one solution. The important thing is choosing the most cost-efficient way. Since cost is time, and life-time is finite for a human-being who is trying to solve a problem, the computationalists and theoreticians are trying to model the biological systems' behaviors with a mathematical perspective through the instrument of specifically developed computer systems.

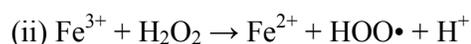
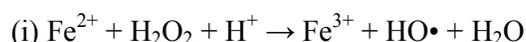
When you try to exemplify the speed of the advancement of the computer technology by some means or other, to get on a Boeing 747 that is large enough to carry twelve thousand people to the moon in about three hours for the round-trip would cost about twelve dollars. Many of the developments in computer hardware and software have produced a new tool for investigating scientific problems. The categories of the scientific research would be taken into account as; (i) theoretical science, (ii) experimental science, and (iii) observational science.

Computational science has emerged over the past 30 years as the application of computational and numerical techniques to solve large and complex problems and systems such as protein folding, or modelling of superconductors. Computational science exploits the advancements in computer hardware, the improvements in computer algorithms, and mathematical techniques. It permits us to compute things that were erenow tough to do due to the intricacy of the mathematics and the large number of calculations involved. It also permits us to create models that predict what might happen in the lab, so that we will interpret what we are seeking. Computational interpretation also will be used as a tool to perform experiments that might be pricey or hazardous to do in the laboratory.

To exemplify the advantages of the computational methodologies, the prediction of the behavior of a new drug in the body by means of computational techniques to reduce the number of animal tests that we might have done heretofore to the improvement of computational pharmacology techniques. While computational models cannot substitute the experimental methodologies, reliable models have become a divaricate part of the overall search for scientific database.

### 1.2.1. Ferric Binding Problem

Iron is essential for continuity of an organisms' life and it is abundant in the earth's crust, being the fourth common element; life on this planet will cease without iron [1]. Unfortunately, iron tends to be found in nature in oxygenated form  $\text{Fe}(\text{OH})_3$  and it is almost insoluble in water at physiological pH under aqueous conditions ( $10^{-38}$  M). This means the concentration of free ferric ion ( $\text{Fe}^{3+}$ ) is close to  $10^{-27}$ - $10^{-10}$  M, so that it cannot counterbalance the need for iron for metabolic purposes in the organism ( $10^{-6}$  M) [2]. Also, iron tends to precipitate as ferric hydroxide at low pH, low concentration of free  $\text{Fe}^{3+}$  ( $10^{-18}$  M), and aerobic aqueous conditions to produce hydroxyl radical as a result of Fenton Reaction. Ferrous ion ( $\text{Fe}^{2+}$ ) is oxidized by  $\text{H}_2\text{O}_2$  to the ferric ion ( $\text{Fe}^{3+}$ ), a hydroxyl anion, and a hydroxyl radical.



Mother nature juggles the ferric ion to come up with a solution for insolubility and potential damage by free  $\text{Fe}^{3+}$ . She suggests the usage of various ligands for  $\text{Fe}^{3+}$  chelations so as to render ferric ion bioavailable, more soluble and highly regulated [2]. Since iron is needed for continuity of life, and capturing it in a suitable state is a pretty hard task, the life forms on earth have developed specialized mechanisms to compete for iron within the environment.

Some of the pathogenic bacteria such as, *Neisseria meningitidis*, *Neisseria gonorrhoeae*, *Haemophilus influenza*, *Actinobacillus actinomycetemcomitans*, and *Pasteurella hemolytica* have specific periplasmic iron binding proteins which deliver the iron ion from outer membrane to the cytoplasmic membrane effectively [3]. Gram negative bacteria such as

Haemophilus and Neisseria capture iron from human ferric binding proteins or iron chelators from the outer membrane. Ferric ion is taken by hFbpA at the periplasmic surface of the outer membrane with concomitant  $\text{H}_2\text{PO}_4^-$  [1].

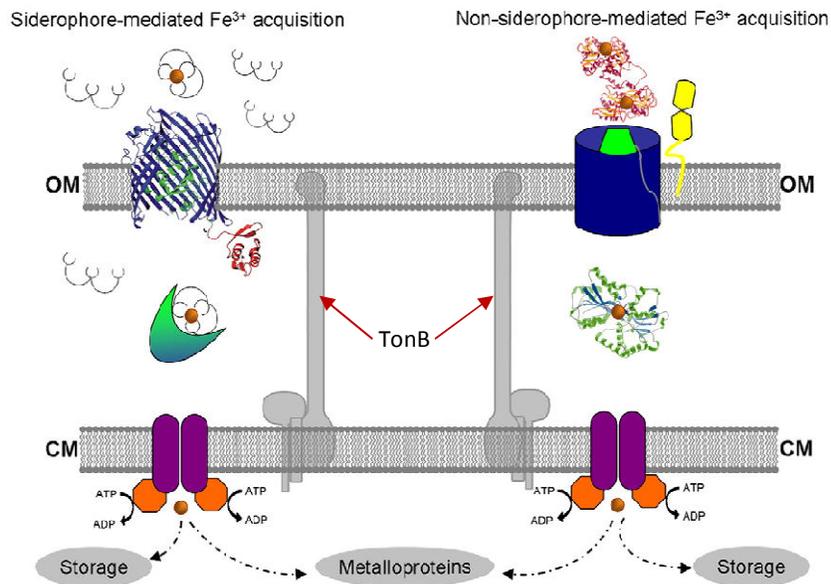
The pivotal role of iron in biological processes is the participation as a single electron acceptor and donor during redox reactions, such as metabolic respiration, oxygen transport, and enzymatic processes [2]. Since the iron is used for biological processes and is required for the continuity of life, understanding the iron binding mechanism is important to prevent the reproduction of these pathogens.

Ferric binding proteins (Fbp) have been called "eukaryotic transferrin" due to their structural similarity to the mammalian transferrin. Based on the structural similarity of hFbpA with eukaryotic transferrins and other prokaryotic periplasmic transport proteins, it can be said that these transport proteins are related by divergent evolution from an anion binding common ancestor. Most of the periplasmic transport proteins including hFbpA are not sequentially related by more than 20% with each other, and their sequential similarity with transferrins is less than 10% [4]. Even though Fbp and Tf are not homologously similar, they both bind directly to  $\text{Fe}^{3+}$  with amino acid side-chain donor groups with a nearly ideal octahedral geometry with ligand- $\text{Fe}^{3+}$ -ligand angles of  $\sim 90^\circ$ , with a deviation of  $6.9^\circ$  [4]. They also capture the iron by Pac-man/Venusfly-trap motion, and they both require a synergistic anion for compact iron binding. [5, 6]. In addition to the sequential dissimilarity of mammalian transferrins and bacterial FbpA systems, the synergistic anions used by proteins and the depth of the active site is different. While hTf uses bidentate carbonate anion ( $\text{CO}_3^{2-}$ ) and five amino acids as ligand, hFbpA uses monodentate phosphate anion ( $\text{H}_2\text{PO}_4^-$ ) as synergistic anion. Also, the hFbpA  $\text{Fe}^{3+}$  active site is more solvent exposed than the transferrin and lactoferrin due to the depth of the cleft, the difference in the relative orientations of the two domains in the proteins, and existence of an additional loop in the transferrins that covers the binding site [4]. That is why it is expected for transferrin to bind the ferric ion with a higher affinity constant ( $K_d = 10^{-21}$  M [7]) than the FbpA proteins ( $K_a = 10^{18}$   $\text{M}^{-1}$ ) in periplasmic space of the gram-negative bacteria. This suggest that there must be special mechanism for iron sequestration from the outer membrane for hFbpA to compete with the power of the hTf iron binding affinity.

### 1.2.2. Suggested Mechanisms of Transport of Iron Through Periplasm

In this thesis, the gram negative pathogenic bacteria *Haemophilus influenzae* periplasmic ferric binding protein apo and holo forms are selected as study systems. There are three different classes of bacteria according to the differences in the composition or thickness of the cell wall structure. These are: (i) gram-positive, (ii) gram-negative, and (iii) acid-fast bacteria. Almost all of the gram negative bacteria are pathogens.

While the gram-positive cell wall consists of a 20-80 nm layer of peptidoglycan, the gram-negative cell wall consists of a much thinner layer of peptidoglycan (1-3 nm). The thick layer of peptidoglycan in gram-positive bacteria make the cell wall behave like an absorbent sponge. Since the gram negative bacteria cell wall is much thinner than the gram-positive bacteria, they need a higher level of protection. For this purpose, gram-negative bacteria have an outer membrane. It is a phospholipid membrane with small openings (porins) that block the entrance of antibiotics or any chemicals that are potentially harmful to the organism [8]. That is the reason why the gram-negative bacteria are much more resistant to the antibiotics than the gram-positive bacteria.



**Figure 1.2** Ferric ion transport systems in gram-negative bacteria. Adapted from [4, 5, 7, 9]

Our sapient protein hFBP is assigned between outer and cytoplasmic membrane of the gram-negative bacteria, called the periplasmic space. There are three basic categories of the bacterial iron uptake mechanisms: (i) ferrous ion transport systems, (ii) siderophore-mediated ferric ion transport systems, (iii) siderophore-independent ferric ion transport systems [2].

Haemophilus influenza periplasmic ferric binding protein's transportation mechanism is a non-siderophore-mediated  $\text{Fe}^{3+}$  acquisition. Siderophores are low molecular weight chelators. The gram-negative pathogenic bacteria such as Neisseria meningitidis, Neisseria gonorrhoeae, and Haemophilus influenza steal iron from host proteins with specific usage of the outer membrane transferrin/lactoferrin receptors located on the cell surface, that exhibit a  $\beta$ -barrel with a plug structure. The receptor is formed by transferrin binding protein to capture the transferrin A (TbpA) and a TbpB to recognize hTf [2]. After being captured by TbpA/B, the iron goes through the pore and is captured by hFbpA which prevails on the inner-membrane surface of the outer membrane (figure 1.2). Questions on how TbpA and TbpB communicate, how crossing of iron through the membrane occurs, and if there exists a conformational change on one lobe of Tf have remained unanswered. It is known that TbpA interacts with apo hFbpA, but the mechanism for iron insertion into hFbpA is not defined [10]. On the surface of the inner cytoplasmic membrane, iron is released from  $\text{Fe}^{3+}$ -hFbpA- $\text{H}_2\text{PO}_4^-$  and passed through hFbpA/B/C which is an (ATP-binding cassette) -ABC transporter-that operates with the energy obtained from ATP to move the  $\text{Fe}^{3+}$  to the cytosol [6].

In this work, we focus solely on FbpA which displays remarkable binding affinity for  $\text{Fe}^{3+}$  ( $K_a = 10^{18} \text{ M}^{-1}$ ) in the presence of phosphate [1]. The question of how  $\text{Fe}^{3+}$  is released is an open problem, the so called "ferric binding dilemma" [11].

### **1.2.3. Crystal Structures of Periplasmic FBPs in Gram-Negative Bacterium**

To prevent the iron binding to the active site for Fbp systems, researchers have tried to control the active site and mutate the coordinator residues around iron to destroy the octahedral symmetry coordination of  $\text{Fe}^{3+}$ . The vicinal residues of ferric ion coordinators Q58, N175, and N193 have been mutated to control and understand the effect of synergistic anion  $\text{H}_2\text{PO}_4^-$  on the binding mechanism [12]. While Q58 is at the N-terminal domain of the active site, N175 and N195 are at its C-terminal domain. The latter is the binding site for

H<sub>2</sub>PO<sub>4</sub><sup>-</sup> in apo form of hFbpA. The results showed that while the C-terminal cleft residue-N175L and N195L- mutation altered the H<sub>2</sub>PO<sub>4</sub><sup>-</sup> binding, the N-terminal cleft residue mutation Q58L abrogated binding of H<sub>2</sub>PO<sub>4</sub><sup>-</sup>. Moreover, all single point mutants of this protein are capable of hijacking ferric ion from transferrin. According to these results, one might say that the transport of ferric ion is not dependent on binding of phosphate in the synergistic anion-binding sites.

**Table 1.1** Crystal Structures of hFbpA

Entry	Method	Resolution(Å)	Chain	Title	Mutant Site	Ligands
1D9V[7]	X-ray	1.75	A	apo	-	PO <sub>4</sub> <sup>3-</sup>
1MRP[4]	X-ray	1.60	A	holo	-	Fe <sup>3+</sup> PO <sub>4</sub> <sup>3-</sup>
1NNF[13]	X-ray	1.10	A	holo	H9Q	EDT Fe
1QVS[14]	X-ray	2.10	A	holo	H9A	Fe <sup>3+</sup> PO <sub>4</sub> <sup>3-</sup>
1QW0[14]	X-ray	1.90	A	holo	N175L	Fe <sup>3+</sup> PO <sub>4</sub> <sup>3-</sup>
2O68[15]	X-ray	1.70	A	holo	Q58L	Fe <sup>3+</sup> PO <sub>4</sub> <sup>3-</sup>
2O69[15]	X-ray	2.00	A	holo	N193L	Fe <sup>3+</sup>
2O6A[16]	X-ray	1.80	A	holo	E57A	Fe <sup>3+</sup> PO <sub>4</sub> <sup>3-</sup>
3KN7[12]	X-ray	1.71	A	holo	Y195A	Fe <sup>3+</sup> PO <sub>4</sub> <sup>3-</sup>
3KN8[12]	X-ray	1.89	A	holo	Y196A	Fe <sup>3+</sup> PO <sub>4</sub> <sup>3-</sup>

The x-ray structures of the mutation studies of active site residues for hFbpA have H9A, H9Q, E57A, Q58L, Y195A, Y196A, N175L, and N193L. These point mutants are available in the literature as shown in table 1.1. The same method to mutate the active site residues have also been applied to the hTf and Lf. For Tf and Lf, it is observed that the Fe<sup>3+</sup> bound forms of single point mutations are crystallized in the closed conformation while

hFbpA metal-bound forms of the site-directed mutants result in an  $\text{Fe}^{3+}$  bound open conformation due to the depth difference of the cleft for the binding site of the ferric ion. For Tf and Lf, it is thought that the additional interactions coming from a deeper cleft are also one of the causes to stabilize the closed conformation obtained after site-directed mutations [14].

The tyrosine residues inside the cleft of the Fbp systems in gram negative bacteria and hTf has a role for iron transportation mechanisms. To state collectively, it is believed that the Y195 and Y196 in hFbpA are not only responsible for binding the ferric ion but also hijacking iron from the hTf. When the Y195 and Y196 residues are mutated, it is observed that the synergistic  $\text{H}_2\text{PO}_4^-$  anion was not present in the crystal structure so suggesting that initial iron capturing is not always controlled by the  $\text{H}_2\text{PO}_4^-$  but also there are other pathways for iron capturing, e.g. by Y195 and Y196 residues [12]. Y195 and Y196 are two of the binding site residues for  $\text{H}_2\text{PO}_4^-$ , located in the C-terminal domain of the active site. It is also known that the wild type hFbpA will be found in both of the open and closed state. Thus after mutation of Y195 and Y196 residues it is expected for  $\text{H}_2\text{PO}_4^-$  to leave its own side (C-terminal side of the cleft) and ferric is controlled only by H9, E57, and  $\text{H}_2\text{PO}_4^-$ . It is possible that after mutation,  $\text{H}_2\text{PO}_4^-$  leaves the cleft since it loses two of the coordinators.

### 1.3. Problem Statement

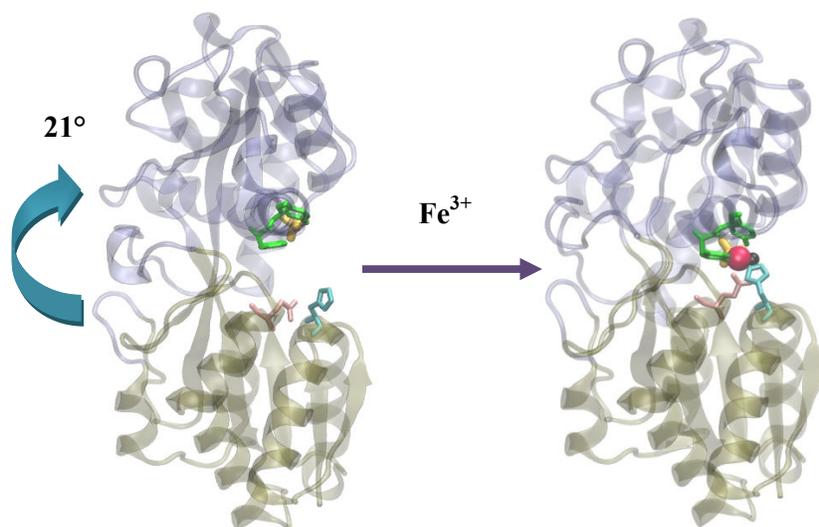
In this work, we focus on controlling the conformations of apo and holo hFbpA. To this end, we use single residue perturbations and charge states of pH sensitive residues in the pH range of biological relevance (5.0-6.5). For the holo form, we focus on the  $\text{H}_2\text{PO}_4^-$  coordinated locations of the  $\text{Fe}^{3+}$  ion while for the apo form we monitor directly the  $\text{H}_2\text{PO}_4^-$  binding/release kinetics. The former problem addresses the questions of ferric ion binding/release phenomena in FbpA [17]. The latter, on the other hand, lends clues on the role of synergistic anions in initiating iron binding [1].

## CHAPTER 2

### Theoretical and Computational Methods

#### 2.1. Haemophilus influenzae periplasmic ferric binding protein systems (hFbpA)

*H. influenzae* was the first free-living organism to have its entire genome sequenced [18]. hFbpA protein systems consists of 309 amino acids made of N-lobe (residues 1-82, 88-101, 226-276) and C-lobe (83-87, 102-225, 277-309). While apo-hFbpA has  $\text{H}_2\text{PO}_4^-$  anion only, holo-hFbpA has  $\text{Fe}^{3+}$  cation also. The ferric ion binds to the active site residues of Y195, Y196, E57, H9 and a water molecule with a synergistic anion as  $\text{H}_2\text{PO}_4^-$  in the structure of an octahedral co-ordination complex as determined in the x-ray crystal structure [PDB-ID:1MRP (holo form) [4], 1D9V (apo form) [7]]. This coordination is similar to the active sites of a single lobe of



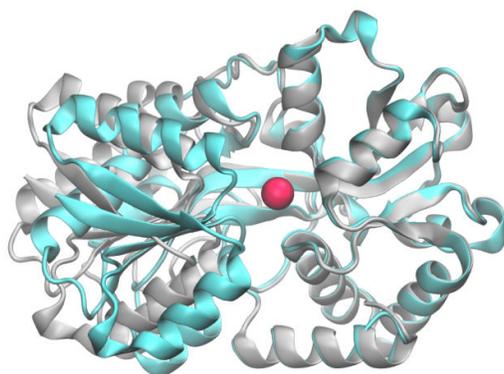
**Figure 2.1.** Three dimensional structures of the proteins. The left structure is the apo (PDB ID: 1D9V) form, the right one is the holo (PDB ID: 1MRP). N-terminal domain (tan), C-terminal domain (iceblue),  $\text{Fe}^{3+}$  (red), water (black),  $\text{H}_2\text{PO}_4^-$  (yellow), Y195/Y196 (green), H9 (light blue), E57 (pink) transferrin or lactoferrin [19].

Apo and holo forms of hFbpA differ by 21° rotation of two domains which makes a hinge movement around central  $\beta$ -strands (figure 2.1). This is called "Venus fly trap movement" [19].

Overall, the three dimensional structure of hFBPA is similar to one of the lobes of human transferrin. It has two-domain structure consisting of an alternating  $\alpha$ -helix/ $\beta$ -sheet structure connected by two antiparallel  $\beta$ -strands even though they share less than 20% sequence identity [9].

## 2.2. System Construction for Molecular Dynamics Simulations

We analyze FBP in detail, using both PRS and MD. The apo and holo forms of FBP have PDB codes 1D9V and 1MRP, respectively. In the latter case, the Fe ion is treated as an additional node of the network. The protein has two domains, and upon binding one moves relative to the other as shown in figure 2.1.



**Figure 2.2.** hFbpA proteins apo and holo forms. The two structures are superimposed on the fixed domain (residues 83-87, 102-225, 277-307). The Fe<sup>3+</sup> ion is shown as a red sphere.

MD simulations on both the apo and the holo forms of FBP in water with different combinations of environmental - NaCl concentrations were performed. We utilize two different ionic strengths, one at 150mM which we called high ionization strength (IS or H) and the other is neutralized systems at low ionic strength (denoted L) with only three chloride anions. We also introduce local perturbations by a single point D52A mutation or by protonation of the same residue. Finally, we study two pH values of 5.5 and 6.5 obtained by

protonation of all histidine residues in the system as well as D52 having an upshifted pKa value. The total list of all MD simulations carried out is provided in Table 3.1 and Table 3.2.1 under the results and discussion part of the chapter 3 along with their average RMSD values for the protein, C-terminal domain RMSD, and N-terminal domain RMSD values.

The protein-water systems are prepared using the VMD 1.8.6 program, autoionize and solvate plugins [20] and the NAMD software package is used to produce the molecular dynamics simulations of the systems prepared [21]. For the holo systems, the length of the simulations were decided on a case-by-case basis. For the apo systems we run for 200 ns to calculate the rate of binding for  $\text{H}_2\text{PO}_4^-$  anion.

After soaking the protein in a water box such that there is at least a 10 Å layer of water molecules in each direction from any atom of the protein to the edges of the box, the systems are neutralized with different concentrations of  $\text{Na}^+$  and  $\text{Cl}^-$  ions to observe the change in the dynamics of hFBPs' behavior in different ionization states. The simulated protein-water complexes have approximately 9370-9640 water molecules. The CharmM22 force field parameters are used for protein and water molecules [22]. The binding parameters for the synergistic anion  $\text{H}_2\text{PO}_4^-$ ,  $\text{Fe}^{+3}$  ions were taken from the literature [20, 23] For the  $\text{Fe}^{+3}$  ion, an effective van der Waals interaction term in addition to electrostatics is included in the spirit done for other ions in the literature [20]. Since the parameters for  $\text{Fe}^{3+}$  do not appear in literature, in our group's previous study the values have been self-consistently parameterized so that the six liganded coordination within  $2.0 \pm 0.2$  Å average distance of the ion [4] is maintained after energy minimization and 200 ps long MD simulations [17]. The optimal values of the Lennard-Jones parameters were found to be  $-0.1$  kcal/mol for well-depth and 2.6 Å for the separation at the minimum. Long range electrostatic interactions were calculated using particle mesh Ewald (PME) method [24]. The cutoff distance for non-bonded van der Waals interactions was set to 12 Å with a switching function cutoff of 10 Å. Rattle algorithm was used to fix the bond lengths to their average values. During the simulations, periodic boundary conditions were used and the equations of motion were integrated using the Verlet algorithm with a step size of 2 fs [25]. Temperature control was carried out by Langevin dynamics with a dampening coefficient of 5/ps and pressure control was attained by a Langevin piston. Volumetric fluctuations were preset to be isotropic in the runs that were carried out at constant pressure and temperature (NPT ensemble).

Both systems were first subjected to energy minimization with the conjugate gradients algorithm until the gradient tolerance was less than  $10^{-2}$  kcal/mol/Å. The final structures were then run in the NPT ensemble at 1 atm and 310 K until volumetric fluctuations were stable to maintain the desired average pressure. Finally, the runs in the NPT ensemble were extended to a total of 20-500 ns. The coordinate sets were saved at 2 ps intervals for subsequent analysis.

### 2.3. Local and Global Perturbation Tools

PRS is a tool used in this work for the analysis of remote control strategies utilized by proteins is based on applying forces at a given residue as a perturbation, and recording the displacements of all the residues as the response. Since the procedure is repeated sequentially for all the residues in the protein, we term the technique, perturbation-response scanning (PRS) [17].

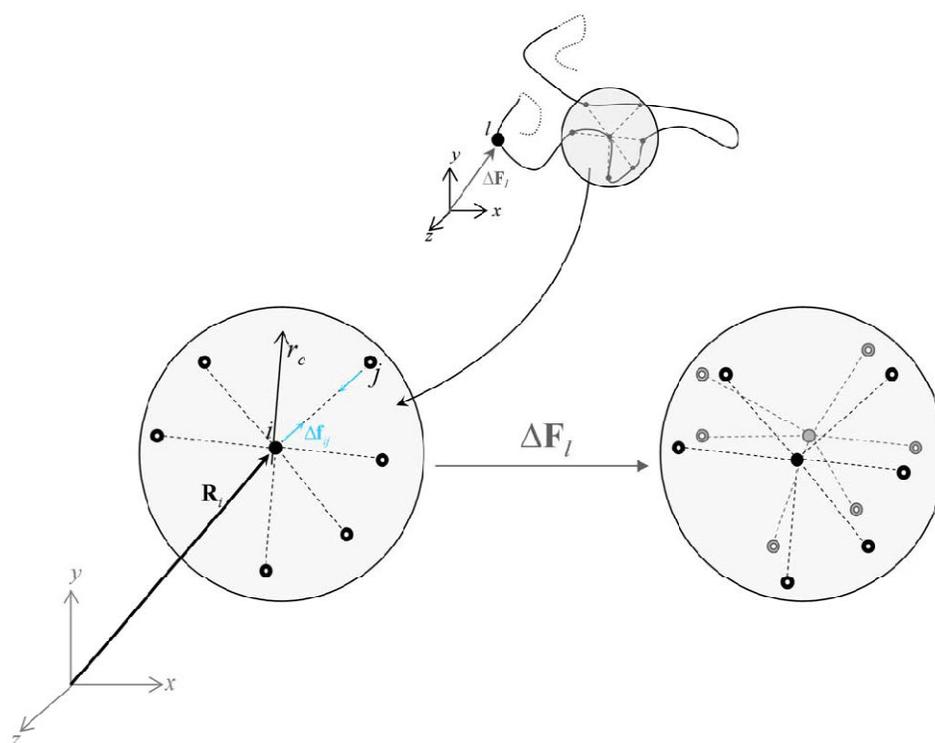
#### 2.3.1. Linear Response Theory

LRT is a derivation of how a structure may be manipulated by external forces [26, 27]. We construct the protein as a residue network of  $N$  nodes that are centered on the  $C_\alpha$  atoms. Any given pair of nodes are assumed to interact via a harmonic potential, if they are within a cut-off distance  $r_c$  of each other (Figure 2.3). In the notation used,  $r$  and  $f$  refer to the bond and internal force vectors along the edge connecting any two nodes, respectively. On the other hand,  $\mathbf{R}$  and  $\mathbf{F}$  are vectors on the nodes and are referred to as the position and external force vectors, respectively. There are  $m$  interactions pertaining to each residue (Figure 2.3, as an example, schematically illustrates the interactions for a residue that has six interactions, i.e.,  $m = 6$ ), and a total of  $M$  interactions for the system of  $N$  residues. In the absence of an external force acting on the system, the equilibrium condition for each residue,  $i$ , necessitates that the summation of the internal, residue-residue interaction forces must be zero for each residue. Therefore,

$$\mathbf{b}\Delta\mathbf{f}_i = 0 \tag{1}$$

where the  $3 \times m$  coefficient matrix  $\mathbf{b}$  consists of the direction cosines of each force representing the residue-residue interaction. The row indices of  $\mathbf{b}$  are  $x$ ,  $y$ , or  $z$ . Here  $\Delta\mathbf{f}_i$  is an  $m \times 1$  column

matrix of forces aligned in the direction of the bond between the two interacting residues. For instance, in Figure 2.3, residue  $i$  has six contacts; and, thus,  $\Delta\mathbf{f}_i$  is a  $6 \times 1$  column matrix. Following the example outlined in Figure 2.3, equation 1 sums up the projection of these six forces on the  $x$ ,  $y$ , and  $z$ -axes. This algebra gives rise to three independent equations involving six unknown interaction forces, which are the residual interaction forces of residue  $i$  with its contacting neighbors. One can write the equilibrium condition (equation 1) for each residue.



**Figure 2.3.** Free-body diagram of a residue

Excerpted from the protein chain (upper panel), scheme depicting the free body diagram of a  $C_{\alpha i}$  atom coordinated by  $C_{\alpha j}$ 's within a cut-off radius  $r_c$  (lower left).  $\Delta f_{ij}$  denotes the interaction force between  $i$  and  $j$ . Under an external force applied on residue  $l$ ,  $\Delta F_l$ , the residues are displaced in space (from the black to the gray nodes in the lower right). The contacting pairs are assumed not to change under this force.

This results in a total of  $N$  sets of equations, each of which involves the summation of forces in three respective directions. Consequently, generalizing equation 1 to the whole system of  $N$  nodes and a total of  $M$  interactions, one can write the following algebraic system of a total of  $3N$  number of equations consisting of  $M$  number of unknown residue-residue interaction forces

$$\mathbf{B}\Delta\mathbf{f}_i = 0 \quad (2)$$

with the  $3N \times M$  direction cosine matrix  $\mathbf{B}$  and the  $M \times 1$  column matrix of residue-residue interaction forces,  $\Delta\mathbf{f}$ . It is straightforward to generate the matrix  $\mathbf{B}$  from the topology of the native structure (i.e., the protein data bank (PDB) file [28]) with a specified  $r_c$ . As an example, apo FBP has 309 residues and a total number of 1542 interactions when the cut-off distance of 8 Å is selected.

In the presence of an external force,  $\Delta\mathbf{F}$  (Figure 2.3), the equilibrium consideration for each residue dictates that the summation of the residue-residue interaction forces for each residue must be equal to the external, applied force on the same residue. Then, equation 2 may be cast into the following form

$$\mathbf{B}_{3N \times M} \Delta\mathbf{f}_{M \times 1} = \Delta\mathbf{F}_{3N \times 1} \quad (3)$$

Under the action of external forces, each residue experiences a displacement,  $\Delta\mathbf{R}$ , which is termed the positional displacement vector. Moreover, the bond distance between any two residues changes in the amount of  $\Delta\mathbf{r}$  in accord with the positional displacements of the two residues which participate in the bonding. Therefore, there must be compatibility between the total of  $3N$  number of positional displacements and the changes that take place in the intra-residual distances, a total of  $M$  number of distortions. This compatibility is very similar to the form given in equation 3 [27]:

$$\mathbf{B}_{M \times 3N}^T \Delta\mathbf{R}_{3N \times 1} = \Delta\mathbf{r}_{M \times 1} \quad (4)$$

Within the scope of an elastic network of residues that are connected to their neighbors with linear-elastic springs, the residual interaction forces,  $\Delta\mathbf{f}$ , are related to the changes in the contact distances,  $\Delta\mathbf{r}$ , through Hooke's law by

$$\mathbf{K}_{M \times M} \Delta\mathbf{r}_{M \times 1} = \Delta\mathbf{f}_{M \times 1} \quad (5)$$

where the coefficient matrix  $\mathbf{K}$  is diagonal. Inasmuch as the native structures are stabilized predominantly by homogeneous tertiary contacts rather than specific interactions [29] we take the entries of  $\mathbf{K}$  to be equivalent in PRS.

Thus, rearranging equations 3–5, one gets the forces necessary to induce a given point-by-point displacement of residues:

$$(\mathbf{BKB}^T)\Delta\mathbf{R} = \Delta\mathbf{F} \quad (6)$$

On the other hand, one may choose to perturb a single or a set of residues, and follow the response of the residue network through,

$$(\mathbf{BKB}^T)^{-1}\Delta\mathbf{F} = \Delta\mathbf{R} \quad (7)$$

where the  $\Delta\mathbf{F}$  vector will contain the components of the externally applied force vectors on the selected residues. The  $(\mathbf{BKB}^T)$  matrix is equivalent to the Hessian [26] and its inverse has six zero eigenvalues, corresponding to the global translational and rotational degrees of freedom of the system. The elements of the inverse of the Hessian,  $\mathbf{G}=\mathbf{H}^{-1}$ , may be used to predict the auto- and crosscorrelations of residues.  $\mathbf{G}$  may be viewed as an  $N\times N$  matrix whose  $ij$ th element is the 3x3 matrix of correlations between the  $x$ -,  $y$ -, and  $z$ -components of the fluctuations  $\Delta R_i$  and  $\Delta R_j$  of residues  $i$  and  $j$ ; i.e.,  $z$ -components of the fluctuations  $\Delta R_i$  and  $\Delta R_j$  of residues  $i$  and  $j$ ; i.e.,

$$\mathbf{G}^{ij} = \begin{bmatrix} \langle \Delta X_i \Delta X_j \rangle & \langle \Delta X_i \Delta Y_j \rangle & \langle \Delta X_i \Delta Z_j \rangle \\ \langle \Delta Y_i \Delta X_j \rangle & \langle \Delta Y_i \Delta Y_j \rangle & \langle \Delta Y_i \Delta Z_j \rangle \\ \langle \Delta Z_i \Delta X_j \rangle & \langle \Delta Z_i \Delta Y_j \rangle & \langle \Delta Z_i \Delta Z_j \rangle \end{bmatrix} \quad (8)$$

The cross-correlations between residue pairs are obtained from the trace of its components:

$$\langle \Delta\mathbf{R}_i \cdot \Delta\mathbf{R}_j \rangle = tr(\mathbf{G}^{ij}) \quad (9)$$

Equation 9 has been shown to reproduce the cross-correlations obtained from MD simulations and molecular mechanics [30, 31]. In this work, we shall not be directly interested in the correlations, but rather shall use  $\mathbf{G}$  as a kernel to predict the response of other residues to applied perturbations on selected ones as we discuss next.

### 2.3.3. Perturbation-Response Scanning

PRS analysis is based on a systematic application of equation 7. We apply a force on the  $C_\alpha$  atom of each residue by forming the  $\Delta\mathbf{F}$  vector in such a way that all the entries, except those corresponding to the residue being perturbed, are equal to zero. For a selected residue  $i$ , the force  $\Delta\mathbf{F}^i$  is  $(\Delta F_x^i \ \Delta F_y^i \ \Delta F_z^i)$  so that the external force vector is constructed as

$$(\Delta\mathbf{F})^T = \{000 \dots \Delta F_x^i \ \Delta F_y^i \ \Delta F_z^i \dots 000\}_{1 \times 3N} \quad (10)$$

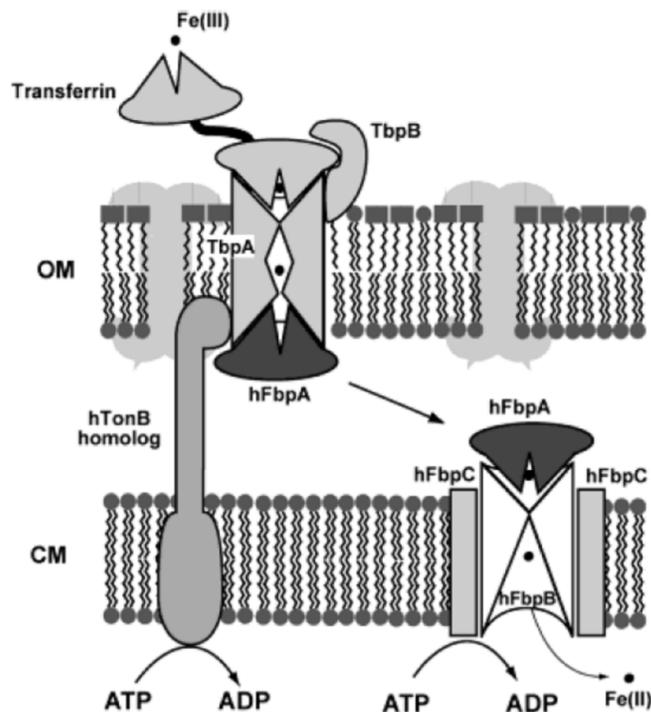
The direction of the applied force vector deserves special attention. Here the forcing direction is chosen randomly, attributing no bias due to the specific contact topology or the solvent exposed nature of the residue being perturbed. The forcing directions are uniformly distributed within a sphere enveloping the residue; therefore, the forcing may well be termed isotropic. It is definitely possible to favor specific directions leading to anisotropy in forcing, since there are no intrinsic constraints in the methodology dictating the opposite. A plausible forcing scenario for contact with a ligand, similar to one in [32] may also be conceived to determine the associated conformational changes. Then the resulting  $(\Delta\mathbf{R})$  vector of the protein is computed through equation 7.

## CHAPTER 3

### Results and Discussion

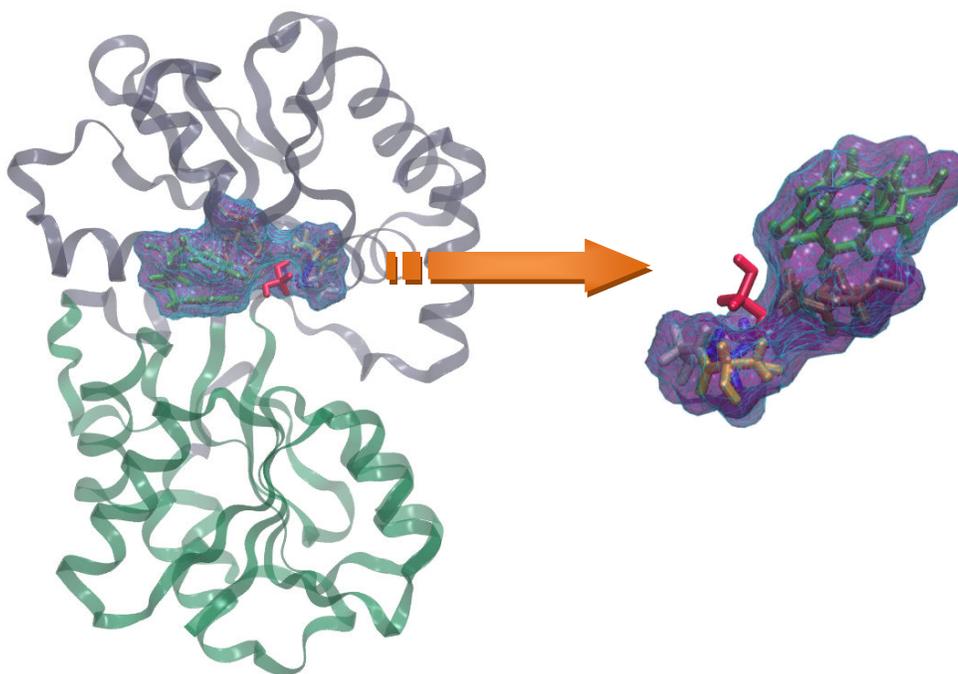
#### 3.1. Analysis of Apo hFBP

The crystal structure of the apo form (PDB ID:1D9V) of the pathogenic bacteria *Haemophilus influenzae* periplasmic ferric binding protein hFbpA has been determined to 1.75 Å resolution and it is observed that the synergistic anion  $\text{H}_2\text{PO}_4^-$  remains bound to the C-terminal domain [7]. Experimentally, it is found that iron-free form of hFbpA binds phosphate anion with an affinity  $K_d$  of  $2.3 \times 10^{-3}$  M [7]. The use of phosphate as a synergistic anion for the binding process of ferric ion to the hFbpA ensues an extreme affinity constant for  $\text{Fe}^{3+}$  binding process (nFbpA  $K'_{\text{eff}} = 2.4 \times 10^{18}$   $\text{M}^{-1}$ ) [7]. According to a previous experimental study,  $\text{Fe}^{3+}$  and  $\text{H}_2\text{PO}_4^-$  can be removed from *Neisseria gonorrhoeae* (nFbpA) at pH 4.5 [15].



**Figure 3.1** Suggested free ion transportation model from transferrin to the cytosol in pathogenic *Neisseria* and *H. Influenza* adapted from [7]

Transportation of iron through transferrin-TbpA-hFbpA into the periplasm remains ill-defined. With the release of  $\text{Fe}^{3+}$ , the number of favorable hydrogen bonding interactions between C and N-terminal domains decreases from 19 to 8 in hFbpA holo to apo transformation [7].

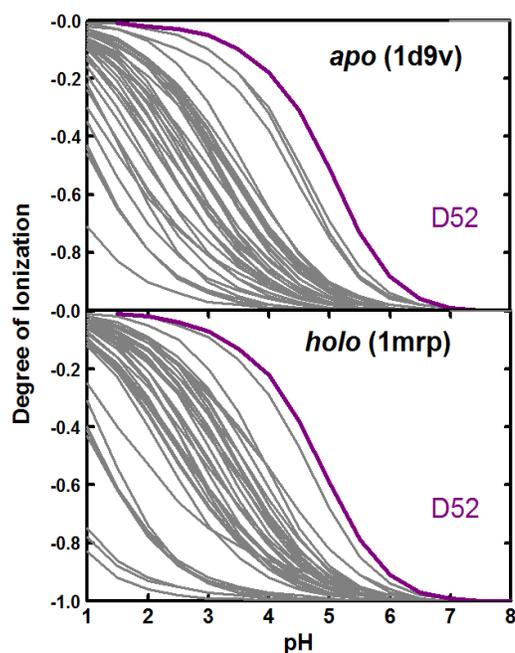


**Figure 3.2**  $\text{H}_2\text{PO}_4^-$  Binding Sites in Apo-hFbpA.  $\text{H}_2\text{PO}_4^-$  (red) is bound to the cleft from the C-lobe. The surrounding residues for  $\text{H}_2\text{PO}_4^-$  are S139, G140, A141, N175, N193, Y195, Y196. C lobe is gray, N lobe is in green.

Apo form of hFbpA takes  $\text{Fe}^{3+}$  from the opening of the channel in the outer membrane (TbpA/B heterodimeric receptor) and captures the iron with an affinity constant of  $K_a = 10^{18} \text{ M}^{-1}$ ; the protein then goes through the periplasm. When hFbpA- $\text{Fe}^{3+}$ - $\text{H}_2\text{PO}_4^-$  complex sticks to the ATP binding cassette (ABC) for free ion transportation process encoded by the cytoplasmic membrane permease (FbpB) and an ATP binding protein (FbpC the hFbpA releases the iron for transportation to the cytoplasm (figure 3.1) [7].

### 3.1.1. MD Simulation Trajectories for Apo hFBP

For the 1MRP and 1D9V, the protonation states of the residues are determined so as to select the most probable state of each ionizable residue at pH 5 and pH 6.5. This latter is the pH of the periplasm for gram-negative bacteria. Four different servers H++ [33], pKd [34], propKa 3.1 [35-38], and PHEMTO [39] were used to calculate the pKa of all the titratable groups. The intrinsic pKa value is defined as the modification of the pKa in the model molecules by the Born energy and the contributions from the partial charges of interacting atoms. Starting from a set of initial values, the electrostatic free energy is calculated iteratively until the pKa values converge. The pKa values calculated for all the titratable residues are listed in Appendix-C. Using pKa calculations (PHEMTO [39]) we find a particular charged residue (D52) (out of a total of 98) to be the most sensitive to subtle pH variations in the physiological range (Figure 3.3). PRS studies showed us that D52 and D47 residues are the ones that give the highest value for the fractional contribution of the highest eigenvalue of the response matrix. According to the residue conservation calculated by ConSurf-DataBase D52 has the highest possible degree of conservation [40].



**Figure 3.3.** Degree of ionization curves for all acidic residues of apo and holo hFbpA. D52 which has the largest upshift in its pKa value is colored purple.

We performed five different types of runs for the apo form of hFbpA and repeated each one starting from a different set of initially assigned random velocities. After locating the perturbation sites we applied different combinations for local and global perturbations. The runs are labeled and listed in Table 3.1. Abbreviations are as follows:

L: neutralized systems with addition of a few  $\text{Cl}^-$  anion. (Low ionization state(IS))

H: Addition of 0.15mM  $\text{Na}^+\text{Cl}^-$  salt to create an approximate model for in vivo environment of the periplasm. (High ionization state)

pH5: protonation of histidines throughout the hFbpA, and protonation of D52 residue.

D52A: Alanine mutation of D52 residue

**Table 3.1.** List of MD Simulation Trajectories for Apo hFBP

simulations	IS	Time (ns)	RMSD (Å)	RMSD (Å)	RMSD (Å)	State	Sim. #
	(H/L)		(equil. state)	C lobe	N lobe		
Wild/apo	L	200	2.4±0.4	1.8±0.4	1.3±0.1	Open	25
Wild/apo	L	200	2.1±0.3	1.8±0.3	1.7±0.3	Open	26
pH5-1	L	200	1.9±0.4	1.6±0.3	1.2±0.1	Open	27
pH5-2	L	200	2.0±0.3	1.6±0.3	1.3±0.2	Open	28
D52A-1	L	200	2.0±0.4	1.5±0.4	1.3±0.2	Open	29
D52A-2	L	200	2.1±0.3	1.6±0.3	1.4±0.2	Open	30
Wild/apo	H	200	2.6±0.7	2.6±0.7	1.7±0.4	Open	31
Wild/apo	H	200	2.0±0.3	1.3±0.1	1.6±0.3	Open	32
$\text{H}_2\text{PO}_4^-$ away	L	100	1.9±0.4	1.5±0.3	1.2±0.2	Open	33
$\text{H}_2\text{PO}_4^-$ away	H	60	1.7±0.3	1.6±0.3	1.0±0.2	Open	34

For wild/apo (H&L IS), pH5, D52A simulations the kinetics for  $\text{H}_2\text{PO}_4^-$ -hFbpA is interesting. It is observed that the monodentate synergistic anion  $\text{H}_2\text{PO}_4^-$  binds to the cleft at the C-terminal domain as in the X-ray crystal structure.  $\text{H}_2\text{PO}_4^-$  is shuttled between the

solution and the cleft several times throughout the simulation. For different types of simulations such as pH5 and D52A the time spent inside the cleft differs. These observations are reproducible, as corroborated by two independent runs of 200 ns each. This showed us that the binding mechanism of  $\text{H}_2\text{PO}_4^-$  to the protein is dependent on the salt concentration, single point mutation, and the pH of the protein. It is surprising that single point mutation on a point that is  $\sim 30$  Å far away from the cleft affects the binding affinity. The  $\text{H}_2\text{PO}_4^-$  stayed bound 19% of the time for wild/apo simulation without any salt and when it is mutated,  $\text{H}_2\text{PO}_4^-$  was bound to the cleft of 52% of the time. Lowering the pH almost completely eliminated shuttling of phosphate and locked it into the binding cleft. Results are summarized in Table 3.2 whereby A and B refer to bound and unbound fractions of the trajectories, respectively.

**Table 3.2.** In-out movement of  $\text{H}_2\text{PO}_4^-$

simulations	IS (H/L)	Sim. #	B	A	Keq= [22]/[22]
Wild/apo	L	1	0.200	0.800	0.236
		2	0.182	0.818	
Wild/apo	H	1	0.088	0.912	0.324
		2	0.402	0.598	
pH5	L	1	1	0	14.385
		2	0.870	0.130	
D52A	L	1	0.562	0.438	1.10
		2	0.488	0.512	

(free)  $\text{A} \leftrightarrow \text{B}$  (bounded)

The fractions A/B are calculated from the  $\text{H}_2\text{PO}_4^-$ -N175 distance plots, whereby a cutoff distance of  $10\text{Å}$  is used to differentiate if the synergistic anion exists in the binding cleft or in the solvent (figure 3.4).

After observing the in-out movement of  $\text{H}_2\text{PO}_4^-$ , we performed control simulations:

- i.  $\text{H}_2\text{PO}_4^-$  away from the cleft simulations in high and low ionization states
- ii.  $\text{H}_2\text{PO}_4^-$  and  $\text{Fe}^{3+}$  away and separated from the cleft in high and low ionization states

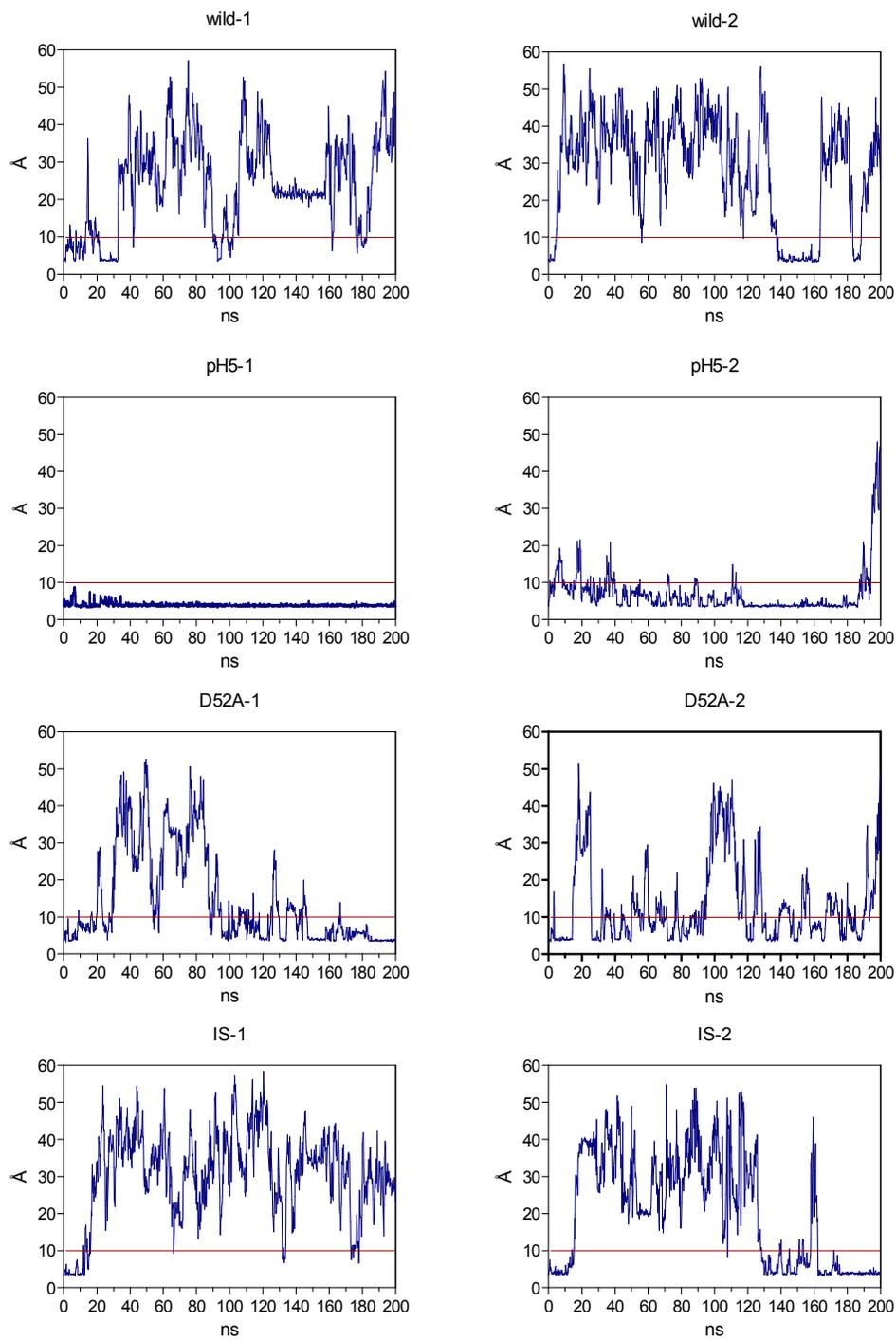
According to the MD simulations, in case (i), the  $\text{H}_2\text{PO}_4^-$  goes and sticks to the cleft and performs in-out movement again, displaying the same dynamics as in the simulations which starts with the  $\text{H}_2\text{PO}_4^-$  inside the cleft. For case (ii), it is observed that  $\text{Fe}^{+3}$  cannot stick to the cleft without its  $\text{H}_2\text{PO}_4^-$  anion. When these two are alone and cannot find each other throughout the periplasmic world, generally  $\text{Fe}^{+3}$  has tendency to stick to the glutamic acids along the protein surface without visiting the cleft by itself. For  $\text{H}_2\text{PO}_4^-$  behavior when it unbound, it goes and finds an lysine or asparagine amino acid while searching for the cleft. The interactions of  $\text{H}_2\text{PO}_4^-$  and side chains of the protein on the surface is not as powerful as the  $\text{Fe}^{+3}$  and glutamic acid side chain interaction which has a +3 charge so that  $\text{H}_2\text{PO}_4^-$  leaves the ASN or LYS and goes and sticks to the C-terminal domain part of the cleft.

There will be more than one mechanism working in vivo. This protein is able to bind other ions such as niobium and gallium [41, 42]. If we assume that the mechanism for iron binding that we mentioned earlier in section 3.1 is the real and only mechanism for binding, mutation of D52 to Ala will increase the fraction of hFbpA- $\text{H}_2\text{PO}_4^-$  complex available in the periplasmic space, and as a result there will be a larger amount of the high iron affinity protein (hFbpA- $\text{H}_2\text{PO}_4^-$  complex) for transportation of iron from the outer membrane to the cytoplasmic membrane. High ionization level for the periplasmic solution will also increase the percentage of these complexes. As a next step we focus on the reason behind the change of  $\text{H}_2\text{PO}_4^-$ .

### 3.1.2. RMSF and RMSD of Trajectories of APO hFBP

The mean square fluctuation MSF is a measure of the deviation between the position of  $\text{C}_\alpha^i$  atom along the peptide chain and reference position as the first frame of MD trajectories.

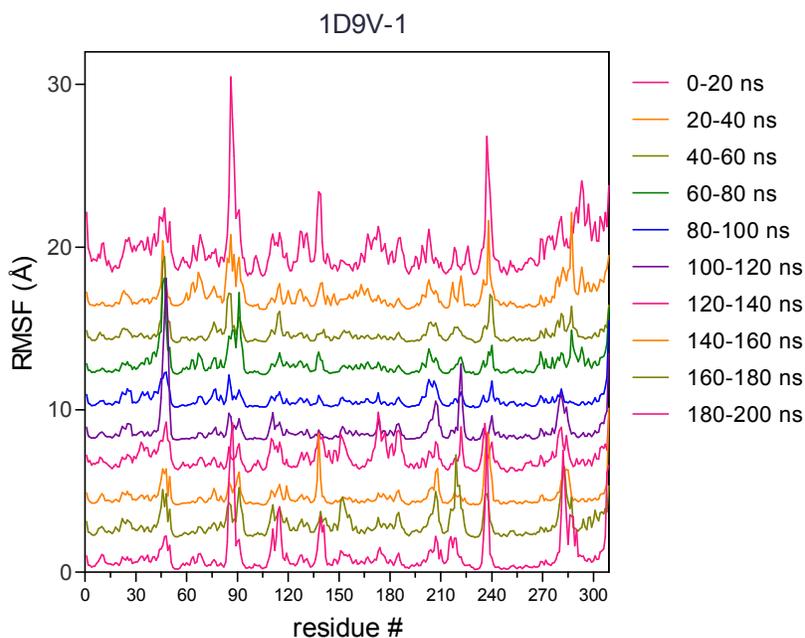
$$MSF = \frac{1}{T} \sum_{t_j=1}^T (x_i(t_j) - \tilde{x}_i)^2 \quad (11)$$



**Figure 3.4.** Time-Distance ( $\text{H}_2\text{PO}_4^-(\text{P})$  between N175(N-side chain)) Graphs for 1D9V Simulations

where  $T$  is the time over which one wants to average, and  $\tilde{x}_i$  is the reference position of particle  $i$ . The reference position will be the time-averaged position of the same  $C_\alpha^i$  atom, i.e.  $\tilde{x}_i$ . The difference between RMSD and RMSF is that with the latter the average is taken over time, giving a value for each  $C_\alpha^i$  atom. With RMSD the average is taken over the all  $C_\alpha^i$  atoms along the peptide chain that constitute the protein giving time specific values. The RMSD trajectories are presented in Figure B1.

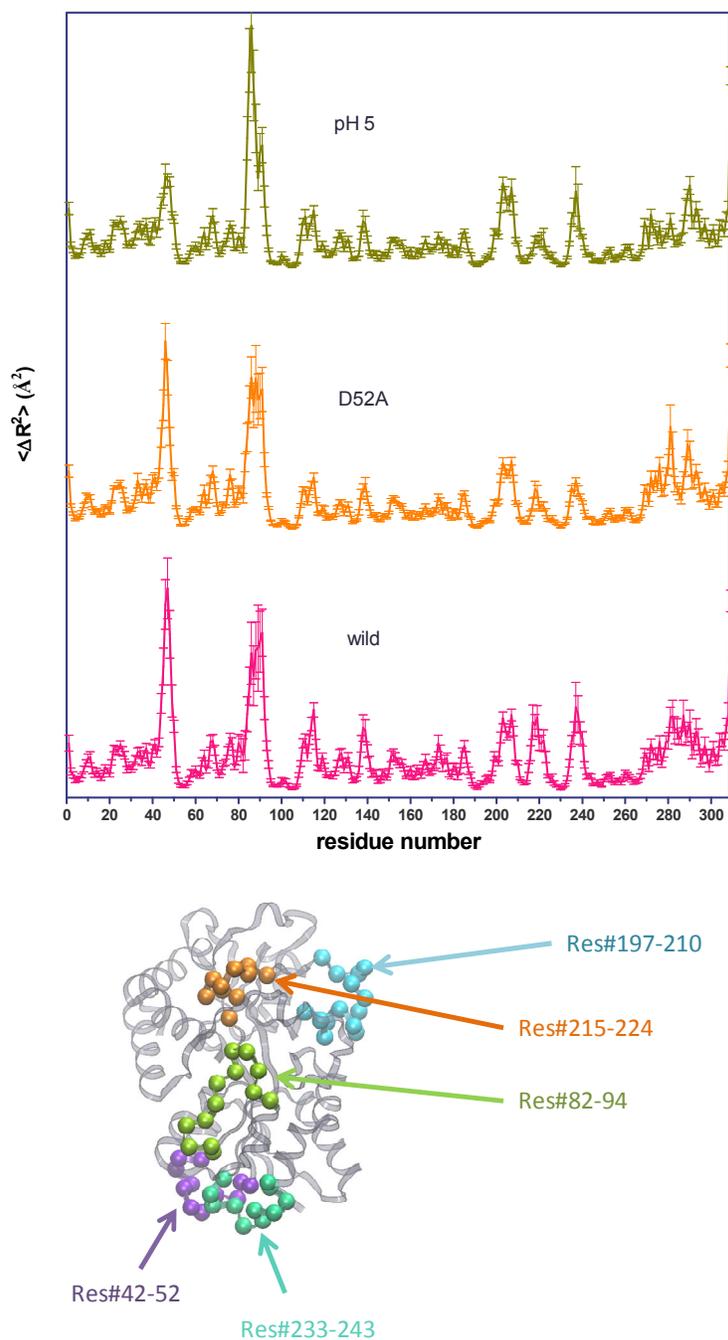
The RMSF graphs are obtained from variance-covariance matrices obtained from MD trajectories. From all the RMSF graphs of apo systems, similar fluctuation pattern is observed. There are three main peak areas of fluctuations. These constitute the neighbourhood of D52, residues 234 to 240, and 282 to 290.



**Figure 3.5.** RMSF graph for 1D9V (low ionization level) first run

When we averaged the RMSF values per chunk for each simulation as shown in figure 3.6, we observed that the fluctuation except the sections of residues with higher fluctuations those are residue 44-54, 84-94, and triple middle peaks of residue 194-242 is the lowest in value for pH5 simulation. After pH5 simulations of apo hFbpA, D52A simulations comes as the second place in terms of fluctuations except for the areas that we mentioned above. It is

easier to comprehend the damping of the cross correlations from 3-D cross correlations graphs in figure 3.7.



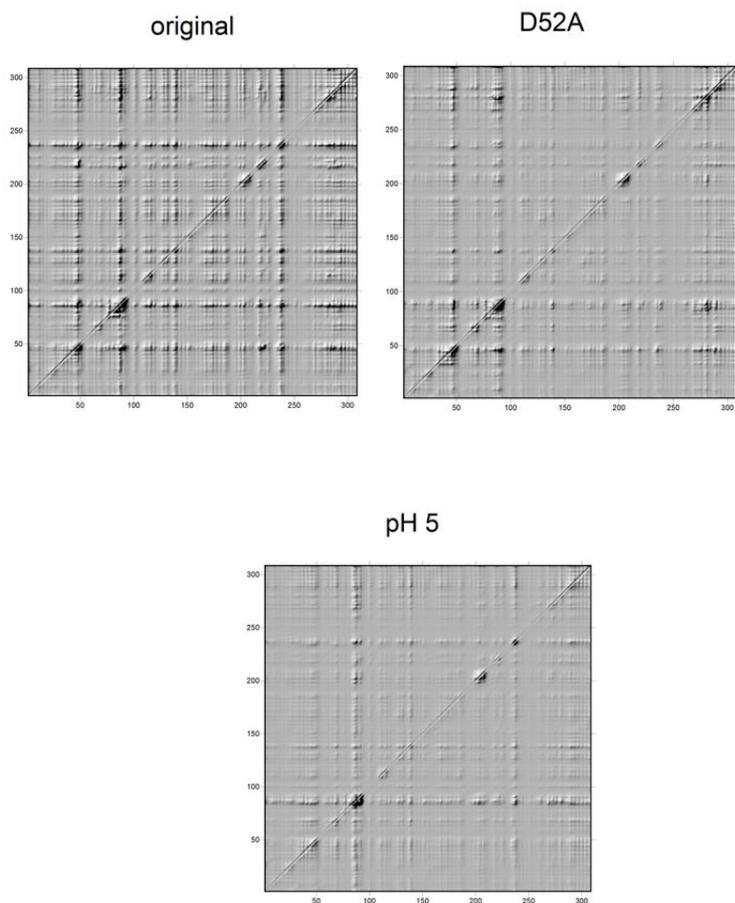
**Figure 3.6** RMSF graph for pH5, D52A, and original (wild type) simulations of 1D9V

From the original system to the pH5 system, cross correlations values gradually fade. If you think about the Pac-man motion of the protein to capture the  $\text{Fe}^{3+}$  or  $\text{H}_2\text{PO}_4^-$ , for two side of the mouth must come closer and get away from each other at the same time. This harmonic-like motion necessitates cross correlation between the two sides of the mouth. Otherwise it is only an independent probability that the 309 point in the protein system have to locate in the space in such a way that the system will be in the state of capturing or releasing the particles.

As displayed in Table 3.1, the RMSD values for three case of the 1D9V systems, Wild/apo, pH5, and D52A, are replicated in bothsimulations. In particular, the C-terminal domain, which is called the fixed domain of the protein for hFbpA systems (residue 83-87, 102-225, 277-309) remains the same. The  $\text{H}_2\text{PO}_4^-$  away simulations were interrupted once the synergistic anion sticks to the cleft for  $\text{H}_2\text{PO}_4^-$  that is located inside of the C terminal domain. For high IS case of  $\text{H}_2\text{PO}_4^-$  away simulations, synergistic anion found the cleft at 48 ns, and for the low IS case anion went to the cleft around 72 ns. For these control simulations, our purpose was to understand whether the  $\text{H}_2\text{PO}_4^-$  will go to the cleft when we have free  $\text{H}_2\text{PO}_4^-$  inside the solution or not. Another observation for these two simulations is that  $\text{H}_2\text{PO}_4^-$  anion tends to stick to lysine and  $\text{Fe}^{3+}$  tends to stick to glutamic acid amino acids.

### 3.1.3 $\text{H}_2\text{PO}_4^-$ Bound-Unbound Conformational Search in holo hFBP

Since we have the MD trajectories of the simulations, we will calculate the correlation between bound and unbound states from covariance matrix that we obtained from the coordinates of  $\text{C}_\alpha$  atoms on the peptide chain. First, we determined the bound and unbound states from the distance measurement to the coordinators of the  $\text{H}_2\text{PO}_4^-$  inside the cleft at C-terminal lobe via VMD [43] package. We measured the distance between phosphate atom of  $\text{H}_2\text{PO}_4^-$  and nitrogen atom of the side chain of N175 (175 numbered asparagine residue) at 0.2 ns time intervals (see figure 3.4 and Appendix B). The same procedure was repeated with two of the other coordinators of  $\text{H}_2\text{PO}_4^-$  which are Y196 and Y195 (oxygen atom on the side chain) (see Appendix B). We observed that  $\text{H}_2\text{PO}_4^-$  is captured by the (cleft at C-terminal domain), and is in close proximity to residues S139, G140, A141, N175, N193, Y195, Y196.

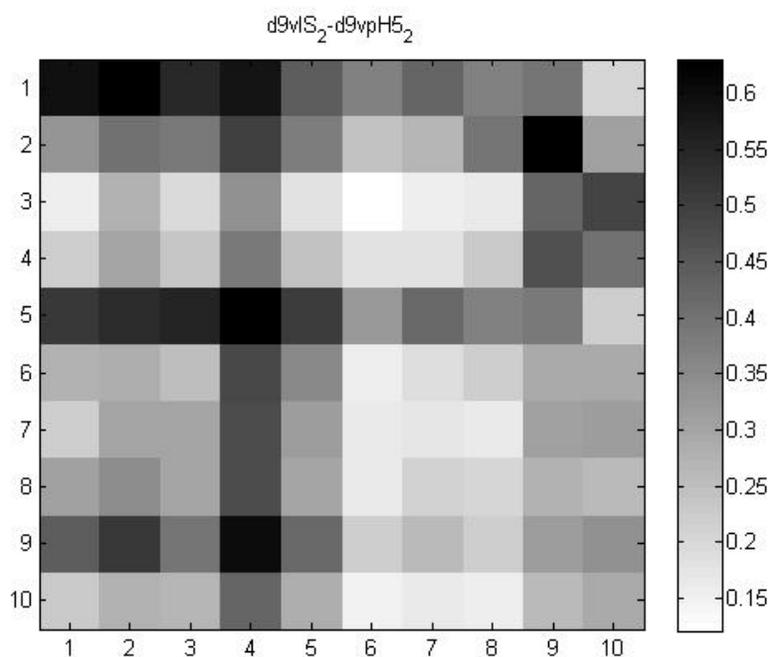


**Figure 3.7** Cross correlations for original, D52A, and pH5 type 1D9V simulations

After observing the coordination of  $\text{H}_2\text{PO}_4^-$  inside the cleft (figure 3.2), we calculated the correlation between bound and unbound states by taking 20 ns bound and unbound time intervals from the trajectories and sampled them (Tables 3.3, 3.4, and 3.5). The results showed that there is a correlation between consecutive bound states, but as the time interval between chunks increases these correlations disappear. The same shrinkage is observed between consecutive unbound states, but the correlation between consecutive unbound states were lower than the consecutive bound states. Thus, the binding of  $\text{H}_2\text{PO}_4^-$  provides a more stable conformation and a slower disappearance of the correlation between chunks. Finally, we investigate the correlation between unbound and bound states and show that there is no correlation between them as anticipated.

To make this analysis visually tracktable, we prepared 20 ns of variance-covariance matrices for each simulation leading to eighty sets which are further analyzed and compared. By writing an elementary level matlab code (Appendix-A), the variance-covariance matrices of each chunk of each simulation was examined for any type of correlations between bound and unbound states of hFbpA-  $\text{H}_2\text{PO}_4^-$  - and hFbpA systems. One example comparison is shown in figure 3.8.

We also calculated the modes of conformational change for trajectory pieces of 20 ns length, each of which we call a “chunk”. From variance-covariance matrix we obtained from MD trajectories, we calculated the eigenvalues and the eigenvectors. We chose the eigenvectors that have the highest eigenvalue (mode 1) and draw them using the VMD package (figure 3.9).



**Figure 3.8** Correlation graph for 1D9V (IS-2) and 1D9V (pH5-2) runs



**Table 3.4: Comparisons of Unbound and Unbound (U&U) by vector analysis**

U-U	U	hIS_1										hIS_2										M_1		IIS_2					
		C	C2	C3	C5	C6	C8	C10	C2	C3	C4	C5	C8	C10	C4	C5	C8	C10	C4	C5	C8	C10	C4	C5	C8	C10			
hIS_1	C2	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
	C3	0.2649	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
	C5	0.2958	0.2689	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	C6	0.2025	0.1996	0.3358	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	C8	0.2779	0.3061	0.3228	0.3109	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	C10	0.2137	0.1935	0.2120	0.2138	0.3240	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	hIS_2	C2	0.2545	0.1463	0.1184	0.1749	0.1939	0.1903	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
		C3	0.1209	0.1157	0.1513	0.1553	0.1640	0.1100	0.4162	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
		C4	0.1685	0.1120	0.1105	0.1918	0.1594	0.1207	0.5041	0.5490	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		C5	0.4084	0.2173	0.2207	0.2512	0.2783	0.2348	0.5155	0.2690	0.3603	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
C8		0.2128	0.1244	0.1125	0.0991	0.1607	0.1317	0.3386	0.2994	0.3451	0.3965	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
C10		0.1820	0.1203	0.1361	0.1268	0.1706	0.1206	0.3307	0.3070	0.3022	0.3691	0.3281	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
M_1		C4	0.2953	0.2248	0.2580	0.2710	0.2621	0.2356	0.2535	0.1286	0.1822	0.3701	0.1696	0.1642	1	-	-	-	-	-	-	-	-	-	-	-	-		
IIS_2		C2	0.2817	0.2078	0.3180	0.1976	0.2528	0.1903	0.2441	0.1229	0.1332	0.3753	0.1532	0.1703	0.3260	1	-	-	-	-	-	-	-	-	-	-	-		
		C5	0.1405	0.1829	0.3088	0.2092	0.1839	0.1090	0.1436	0.1775	0.2187	0.1382	0.1409	0.1539	0.1752	0.2484	1	-	-	-	-	-	-	-	-	-	-		

\*C1:0-20ns / C2:20-40ns / C3:40-60ns / C4:60-80ns / C5:80-100ns / C6:100-120ns / C7:120-140ns / C8:140-160ns / C9:160-180ns / C10:180-200ns

\*C:Chunk

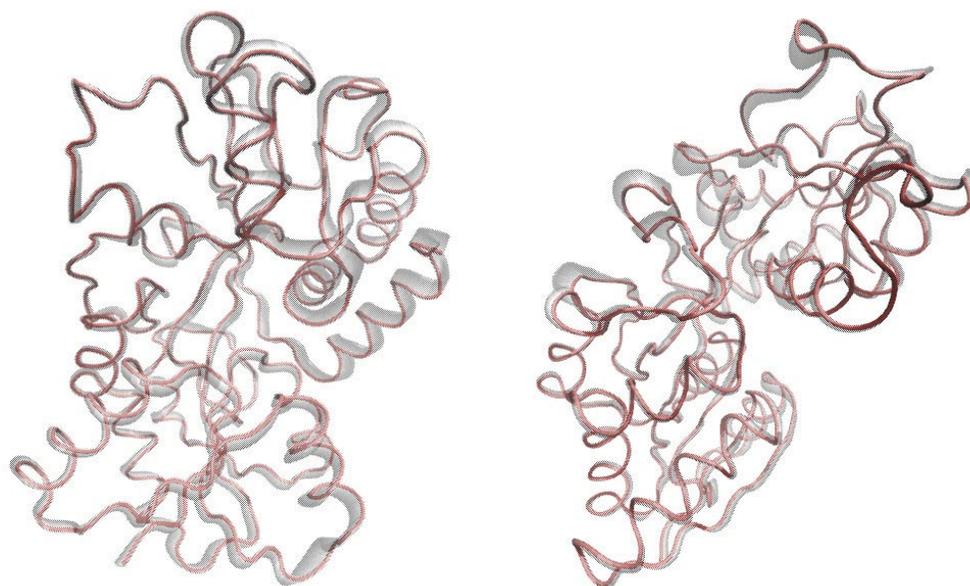
**Table 3.5:** Comparisons of Unbound and Unbound (B&U) by vector analysis

B-U	U	hIS_1										hIS_2					M_1		IIS_2	
		C2	C3	C5	C6	C8	C10	C2	C5	C4	C5	C4	C2	C5						
pH5_1	C	0.2425	0.1508	0.1736	0.1831	0.1982	0.1955	0.2848	0.4532	0.2916	0.2853	0.4533								
	C1	0.3117	0.17579	0.1943	0.2162	0.2321	0.2146	0.3901	0.4993	0.3395	0.2954	0.4994								
	C2	0.2591	0.1392	0.1111	0.1315	0.1994	0.1933	0.3402	0.4917	0.2460	0.2675	0.4917								
	C3	0.2628	0.1455	0.1641	0.1690	0.2050	0.2079	0.4350	0.5135	0.3308	0.4123	0.5135								
	C4	0.3800	0.1675	0.1521	0.1324	0.2115	0.1947	0.3279	0.4421	0.2084	0.2671	0.4422								
	C5	0.2101	0.1195	0.1215	0.1296	0.1772	0.1630	0.2453	0.3611	0.1656	0.1939	0.3611								
	C6	0.1925	0.1180	0.1090	0.1115	0.1480	0.1223	0.2250	0.3398	0.1384	0.1977	0.3399								
	C7	0.1218	0.1228	0.1998	0.1683	0.1521	0.1077	0.1866	0.2044	0.1319	0.1592	0.2044								
	C8	0.1408	0.1290	0.1973	0.1817	0.1569	0.1129	0.2109	0.1919	0.1456	0.1719	0.1919								
	C10	0.1273	0.1294	0.1672	0.1630	0.1819	0.1515	0.1839	0.1591	0.1720	0.1888	0.1591								
pH5_2	C5	0.3870	0.3503	0.2313	0.1895	0.2939	0.2399	0.4891	0.3869	0.2987	0.3266	0.2284								
	C7	0.3260	0.3131	0.1761	0.1188	0.2064	0.1828	0.1771	0.1430	0.3163	0.1772	0.1431								
	C9	0.2605	0.1603	0.1932	0.2193	0.2203	0.1926	0.1732	0.1208	0.2151	0.1732	0.1210								
hIS_2	C10	0.1903	0.1100	0.2348	0.1229	0.1317	0.1207	0.3307	0.3691	0.1642	0.1703	0.1540								
	C8	0.1184	0.1417	0.1702	0.1336	0.1503	0.1012	0.1314	0.1011	0.1036	0.1532	0.2076								
IIS_2	M_1	0.2667	0.2029	0.3236	0.2255	0.2543	0.2039	0.2651	0.3405	0.3077	0.4137	0.2787								
	C10																			

\*C1:0-20ns / C2:20-40ns / C3:40-60ns / C4:60-80ns / C5:80-100ns / C6:100-120ns / C7:120-140ns / C8:140-160ns / C9:160-180ns / C10:180-20

\*C:Chunk

To perceive the magnitude of the contributions per section on protein we visualized the slowest nontrivial mode (mode 1) for the first chunk of the 1D9V-1<sup>st</sup> simulation. These are displayed from front and side view in figure 3.9. While during 0-20 ns the C lobe residues (S139, G140, A141, N175, N193, Y195, Y196) gave the highest contribution to the first mode for chunk 1, the back side residues (83-94, 236-264) control the second mode in chunk 5. The first and the second modes of each 20 ns chunk for each simulation were observed and no significant result is obtained to explain the binding-release mechanisms of  $\text{H}_2\text{PO}_4^-$  to hFbpA and  $\text{Fe}^{3+}$  to  $\text{H}_2\text{PO}_4^-$ -hFbpA.



**Figure 3.9** 1D9V-1 chunk1 (0-20 ns) mode 1 and 1D9V-1 chunk 5 (80-100 ns) mode 1

### 3.2. Analysis of Holo hFBP

Holo hFbpA is the system with the ferric ion inside the cleft. Our aim is to control and understand the hinge/Venusfly-Trap/Pacman movement of the cleft for the binding/release of iron in hFbpA. For this purpose we applied the following perturbations by to the systems in the MD simulations:

**Table 3.2.1** List of MD Simulation Trajectories for Holo hFBP

simulations	IS	Time (ns)	RMSD (equil. state)	RMSD C lobe	RMSD N lobe rmsd	State	Sim. #
Wild_1	L	144	2.2±0.2	1.8± 0.2	1.9±0.2	SC1	1
Wild_2	L	200	2.3±0.3	1.6±0.3	1.6±0.2	SC2	2
Wild_3	L	92	1.7±0.3	1.3±0.2	1.3±0.2	SC2	3
Wild_1	L	40	2.3±0.5	1.6±0.2	1.6±0.4	SC2	4
Wild_2	L	120	1.5±0.2	1.1±0.1	1.0±0.2	SC2	5
Wild_3	L	40	1.3±1.1	0.9±0.1	1.1±0.2	SC1	6
Wild_4	L	60	1.5±0.1	1.3±0.2	1.2±0.1	C1	7
Wild_1	H	20	1.5±0.2	1.1±0.1	1.0±0.1	C1	8
Wild_2	H	20	1.6±0.3	1.1±0.1	1.1±0.1	C1	9
Wild_3	H	20	1.7±0.2	1.2±0.2	1.1±0.1	C1	10
Wild_4	H	20	1.6±0.2	1.2±0.1	1.2±0.1	C1	11
Wild_195/196Y_neg	L	20	1.4±0.2	1.2±0.1	1.0±0.1	C2	12
Wild_195/196Y_neg	H	20	1.3±0.1	1.1±0.1	1.1±0.1	C2	13
D52A_1 (120-300)	L	300	3.3±0.6	1.9±0.4	2.2±0.4	SC2	14
D52A_2	L	200	2.1±0.2	1.3±0.2	1.6±0.2	SC2	15
D52A	H	60	1.6±0.2	1.1±0.2	1.2±0.2	SC2	16
D52A_Y195/196_neg	H	60	1.5±0.1	1.1±0.1	1.1±0.1	C2	17
D52A_Y195/196_neg	L	60	1.2±0.1	1.1±0.1	1.1±0.1	C2	18
pH5_1 (200-500)	L	500	2.8±0.2	2.7±0.2	2.2±0.4	SC1	19
pH5_2 (120-300)	L	300	2.5±0.5	2.2±0.5	2.3±0.4	SC1	20
D52+_1	L	194	2.5±0.3	1.7±0.3	1.5±0.4	SC2	21
D52+_2	L	250	2.1±0.2	1.7±0.2	1.5±0.4	SC2	22
Fe <sup>+3</sup> H <sub>2</sub> PO <sub>4</sub> <sup>-</sup> away	L	100	1.9±0.3	1.3±0.2	1.4±0.2	-	23
Fe <sup>+3</sup> H <sub>2</sub> PO <sub>4</sub> <sup>-</sup> away	H	100	1.5±0.2	1.0±0.1	1.2±0.1	-	24

\*SC1 (semiclosed1), SC2(semiclosed2), C1(Closed1), C2(Closed2)

- (i) Wild: Wild type holo hFbpA with low and high IS
- (ii) Wild\_195/196Y\_neg: Deprotonation of tyrosine 195-196 (Y195/196) residues in the cleft at C-terminal domain coordinators of  $\text{H}_2\text{PO}_4^-$  as a control run with high and low IS.
- (iii) D52A: Mutation of D52 to alanine (D52A) at both high and low IS
- (iv) D52A\_Y195/196\_neg: Mutation of D52A and deprotonation of Y195/Y196 residues.
- (v) pH5: Protonation of all four histidine residues in the polypeptide chain and protonation of D52
- (vi) D52+ : Protonation of D52 only
- (vii)  $\text{Fe}^{+3}$   $\text{H}_2\text{PO}_4^-$  away: Control runs to see the binding behaviour of hFbpA/  $\text{H}_2\text{PO}_4^-/\text{Fe}^{3+}$  starting from the initial condition that they are not connected in the solution environment. They are placed  $\sim 30$  Å from each other to prevent any direct interaction between them.

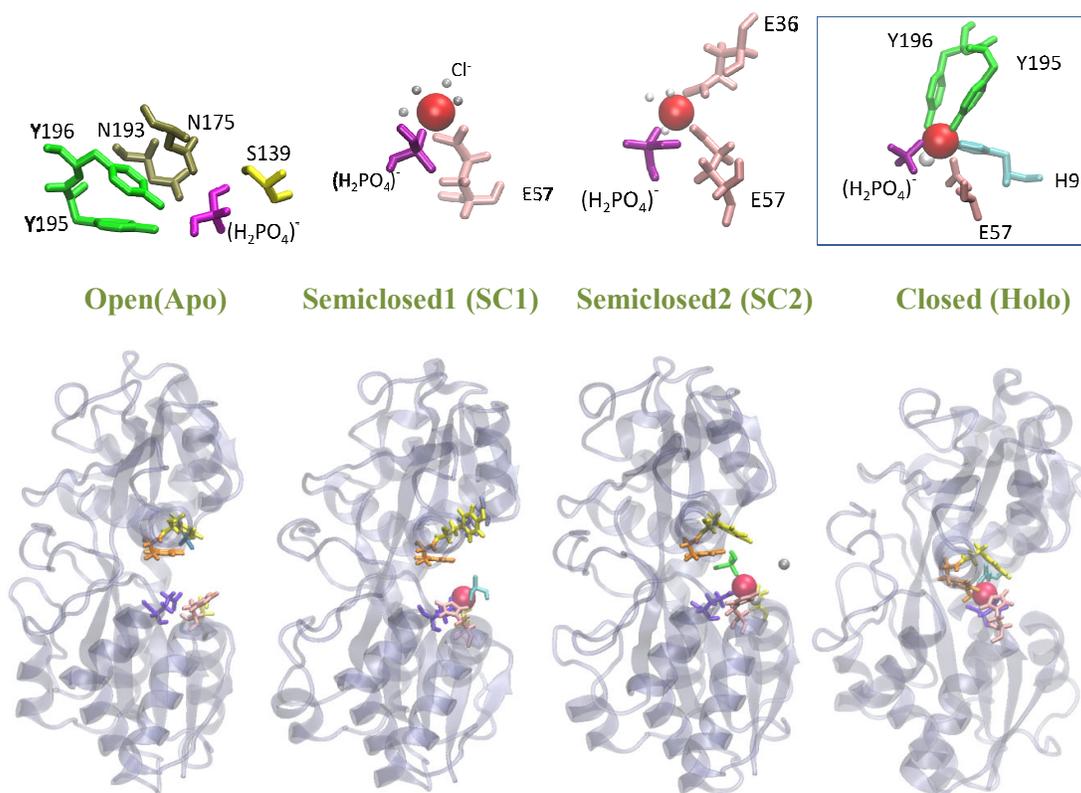
The simulation details are summarized in Table 3.2.1 and different conformations observed are displayed in figure 3.10. The final conformations obtained are classified as Closed 1 (C1), Closed 2 (C2), semiclosed1 (SC1), and semiclosed 2 (SC2). Their details will be provided in the next subsection.

### 3.2.1. MD Simulation Trajectories for Holo hFBP

Wild-type hFbpA could be found in an open or a closed state due to the crystallization conditions [12]. We note that semiclosed experimental structures have not been reported for hFbpA, but a similar situation was recently observed for hTf [44]. Holo hFBP has an octahedral coordination of the  $\text{Fe}^{+3}$  cation in the X-ray structure of the closed state [4]. After equilibration time, the octahedral coordination of ferric ion changes with the different combinations of perturbations applied to the system. For all of the simulations studied, it is observed that free  $\text{Fe}^{3+}$  has tendency to stick to a glutamic acid residue.

For the simulations of wild type holo hFbpA at low IS, four out of seven lead to SC2. For SC2 conformations, the cleft is semiopen and  $\text{Fe}^{3+}$  cation stays between E36 and Y195/196 residues. In these simulations, one of the chlorine anions comes to the active site (without  $\text{Cl}^-$

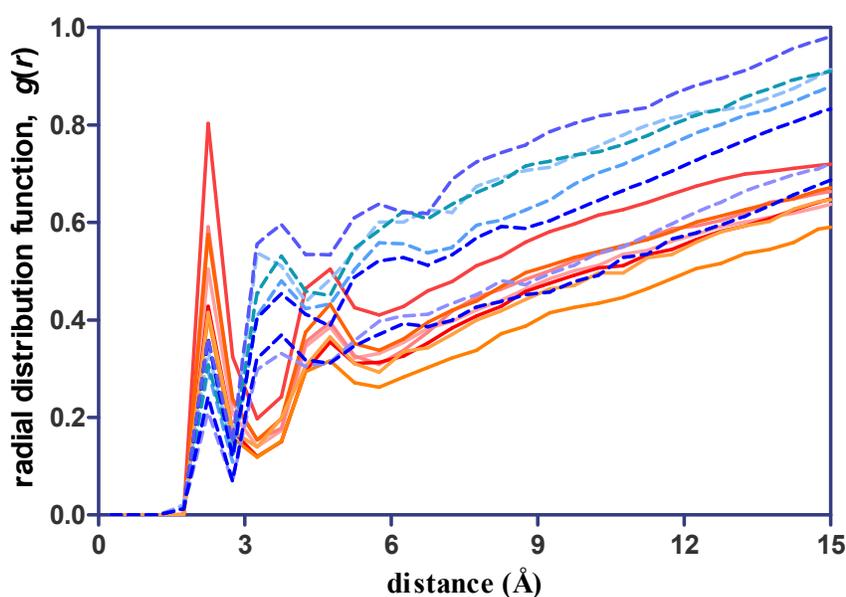
forming any type of permanent ion pairing) while E57/Fe<sup>+3</sup>/H<sub>2</sub>PO<sub>4</sub><sup>-</sup> triplet stays in the middle of Y196/195 and E36. In the case of semiclosed1 (SC1) conformation, as soon as a Cl<sup>-</sup> anion comes to the vicinity of the active site (cleft), monodentate phosphate anion/Fe<sup>+3</sup> cation/E57 triplet starts to move away from the active site residues and distort the octahedral symmetry. E36 and E37 start to keep Fe<sup>+3</sup> cation bound to the active site while phosphate anion faces the active site residues H9, Y195, Y196. Once H<sub>2</sub>PO<sub>4</sub><sup>-</sup>/Fe<sup>3+</sup> doublet goes and binds to the E36. This doublet does not form an octahedral coordination of the ferric ion



**Figure 3.10** Models assigned from MD trajectories of hFbpA

All of the wild type hFbpA systems at high IS and one out of seven low IS wild type hFbpA shows behavior of C1. In this conformation, the Fe<sup>+3</sup>/H<sub>2</sub>PO<sub>4</sub><sup>-</sup>/E57 triplet moves away from Y196/195 and H9, but apart from that no significant change occurs in the system and the system stays closed. The phosphate anion interacts closely with the Y195/196 residues while H9 residue switches its conformation in and out of Fe<sup>+3</sup>. Two of the chloride anions stay close to the Fe<sup>+3</sup> and the monodentate phosphate anion. When there is Na<sup>+</sup>/Cl<sup>-</sup> ions around the

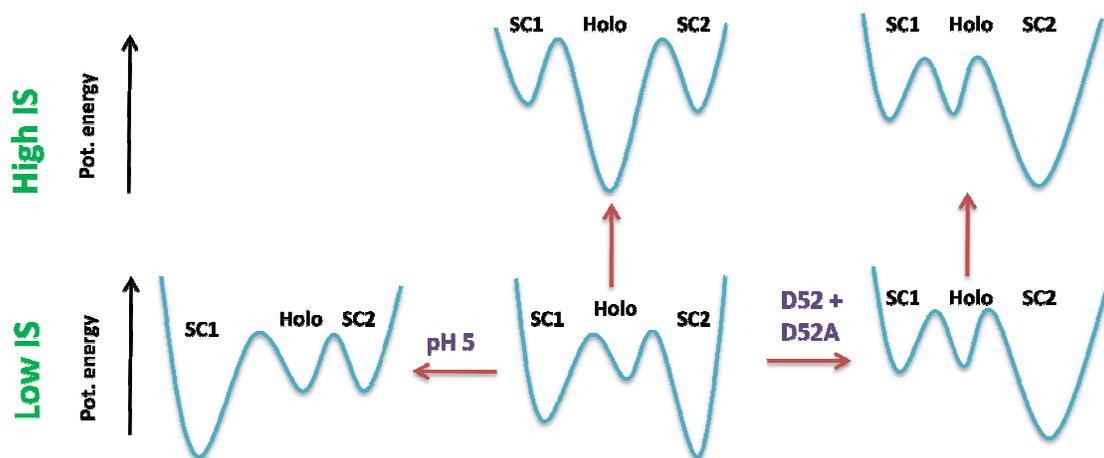
protein at high IS, they have different levels of electronic attraction to the protein so that their placement near the binding cleft is different. According to the radial distribution function of the ions with the protein heavy atoms (figure 3.11), the highest density of  $\text{Na}^+$  and  $\text{Cl}^-$  ions occur at the same distance from the protein, but the density of the  $\text{Na}^+$  cations are higher than the  $\text{Cl}^-$  anions. The outermost shell of the protein is richer in  $\text{Cl}^-$  anion. hFBP is neutral in total; when the tyrosine doublet Y195/Y196 in the active site is deprotonated, the entrance of the cleft is in the strongest closed state C2 which is the closest strict structure to the crystal structure of the 1MRP.



**Figure 3.11** Radial distribution function graph for the holo hFbpA systems in high IS. The red curves are for the protein heavy atoms and  $\text{Na}^+$  ions in the various runs, while the blue curves are their counterparts for protein heavy atoms and  $\text{Cl}^-$  ions.

When we change the pH of the holo hFbpA, the protein has tendency to be in the SC1 conformation described above. Finally, the system chooses to be in SC2 conformation when single point perturbation is applied onto the system (D52A/D52+). With these observations in mind, we propose that shifts in the energy landscape may be promoted using different perturbations on the protein, as depicted in figure 3.12. These refer to what happens to the system once the Y195/Y196 pair is protonated. Starting from the neutral protein at low IS which has the propensity to sample both the SC1 (2 runs) and SC2 (4 runs) conformations or

the closed state (1 run) ,increasing the salt concentration shifts the equilibrium towards the closed (C1) state (all four runs). Conversely, lowering the pH or perturbing the D52 residue stabilizes the SC1 and SC2 conformations, respectively, the latter being irrespective of the salt concentration.



**Figure 3.12** Suggested potential well graph for holo hFbpA simulations

## CHAPTER 4

### 1. Conclusion and Future Work

For the hTf it is known that with an intervention such as an active site residue mutation, it is possible to have a semiclosed state [44]. For periplasmic ferric binding proteins in gram-negative bacteria, even without a mutation, open conformations with ferric ion inside the cleft was obtained earlier by using experimental methods [2]. We obtained these structures by using various computational tools and observed four different states as: closed, semiclosed-1, semiclosed-2, and open. These four states are obtained with different perturbations to the systems, including changing the pH of the protein, mutation or protonation of a single residue away from the active site, and changing the ionization strength of the solution. Perhaps our most important finding is that, our MD simulations of holo hFbpA support the importance of D52 in manipulating the conformational change of the protein system. Even though when there is no  $\text{Fe}^{+3}$  ion in the system for the apo hFbpA it is observed that the D52 perturbations affects the  $\text{HPO}_4^-$  binding mechanism. It is observed that  $\text{Fe}^{3+}$  has a tendency to stick a glutamic acid residue whenever it is near a glutamic acid residue. In the active site there are two glutamic acid residues; one of them is inside the cleft (E57) and the other one is around the mouth of the cleft (E36). These always capture the  $\text{Fe}^{3+}$  ion before it has a chance to leave the cleft. We also performed two control simulations with  $\text{Fe}^{3+}$  and  $\text{H}_2\text{PO}_4^-$  kept  $\sim 30\text{\AA}$  away from the protein and more so from each other in the initial condition. From the control simulations it is observed that when the ferric ion is inside the solution it cannot find the active site, at least for 100ns either with or without  $\text{H}_2\text{PO}_4^-$ . For the low IS case, the ferric ion is captured by the E305 which is on the surface of the protein, and does not leave the sling shot like structure of the carboxylic group of the glutamic acid. For the high IS case, ferric ion and  $\text{H}_2\text{PO}_4^-$  found each other inside the solution but together they could not reach the cleft within our window of observation. For the apo-hFbpA systems, it is observed that the  $\text{H}_2\text{PO}_4^-$  has the ability to perform binding and release motions find the cleft. We find that these motions can be consistently controlled under different perturbation scenarios and they display the same kinetics in repeated runs.

In future work, we propose that the energy differences between the ferric ion binding partners under different scenarios should be studied by quantum mechanical calculations. For a deeper understanding of the binding/release kinetics of phosphate in the absence of iron, free energy calculations should be carried out. Furthermore, the dynamics of other synergistic anions should be studied to better understand the role of these anions in iron transport by FBP and related proteins.

## APPENDIX A

### Code A1 Calculation of correlation between two vectors

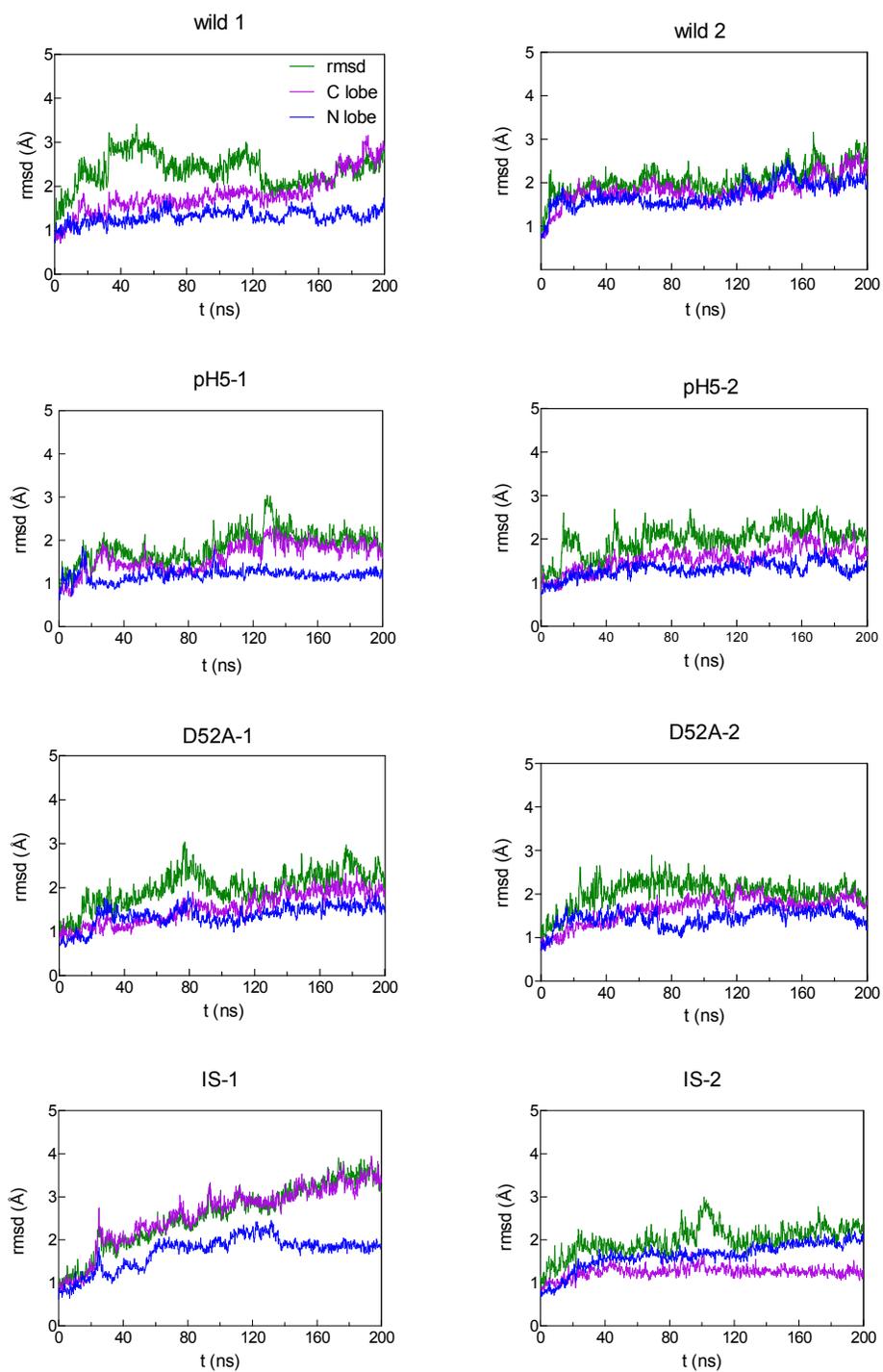
*%this code calculates the correlation between two vectors*

```
clear all
clc
M=dlmread('varcovar_d9vpH5_1.txt');
N=dlmread('varcovar_d9vpH5_2.txt');
%number of residues
noRES=309;
dimen=noRES*3;
line=dimen^2;
Vone=M(:,3);
Vtwo=N(:,3);
vector1=zeros(line,1);
vector2=zeros(line,1);
counter=0;
matrix=zeros(10,10);
v1=cell(10,1);
v2=cell(10,1);
for i=1:10
    for k=1:line
        counter=counter+1;
        vector1(k,1)=Vone(counter,1);
        vector2(k,1)=Vtwo(counter,1);
    end
    v1{i,1}=vector1;
    v2{i,1}=vector2;
end
corr_matrix=zeros(10,10);
for a=1:10
    for b=1:10
        bir=v1{b,1};
```

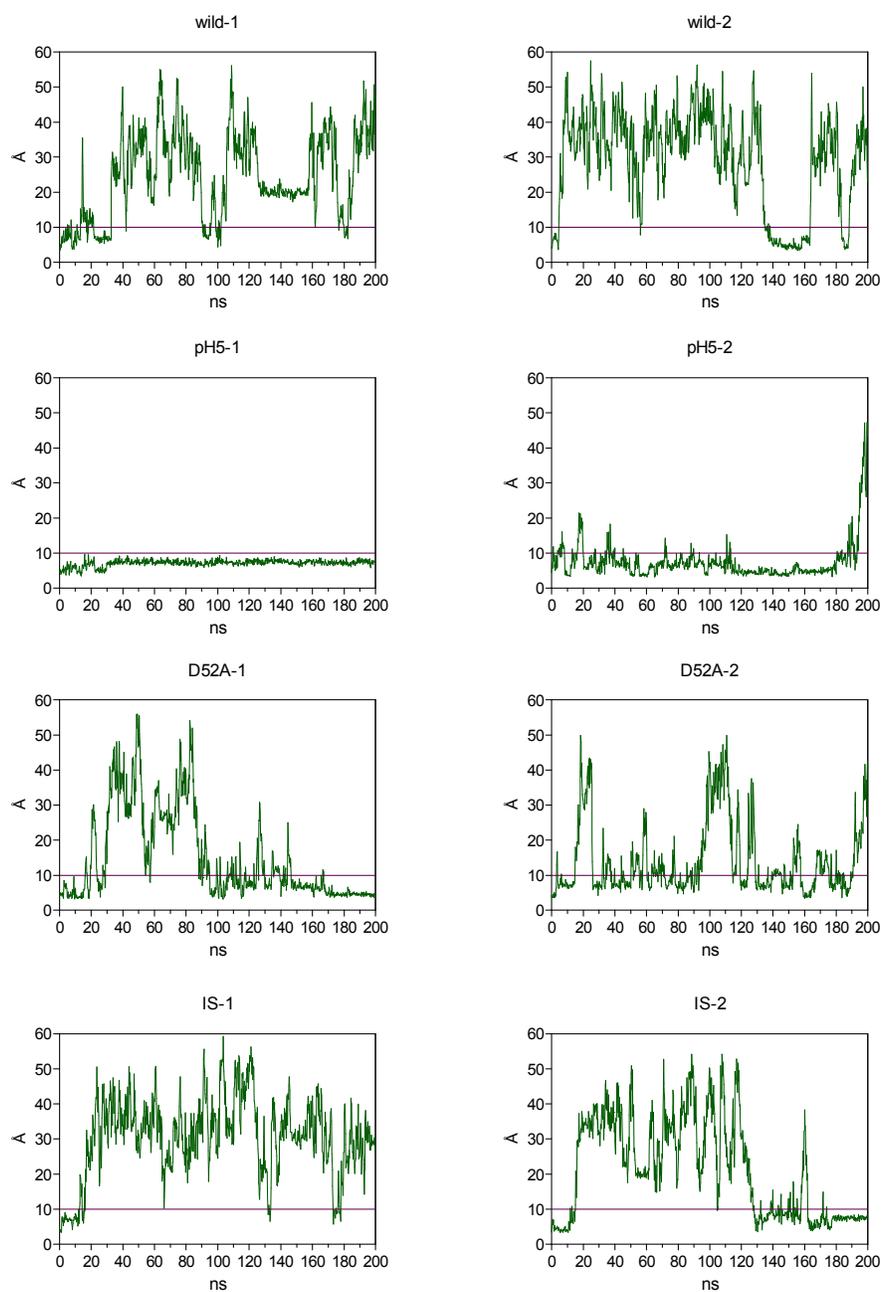
```
    iki=v2{a,1};  
    corr_matrix(b,a)=corr(bir,iki);  
end  
end  
imagesc(corr_matrix);  
colormap(flipud(gray));  
colorbar;  
set(gca,'YTick',[1 2 3 4 5 6 7 8 9 10]);  
set(gca,'XTick',[1 2 3 4 5 6 7 8 9 10]);  
title('d9vpH5_1-d9vpH5_2');  
xlswrite('d9vpH5_1-d9vpH5_2_chunk_corr_all.xls', corr_matrix);
```

## APPENDIX B

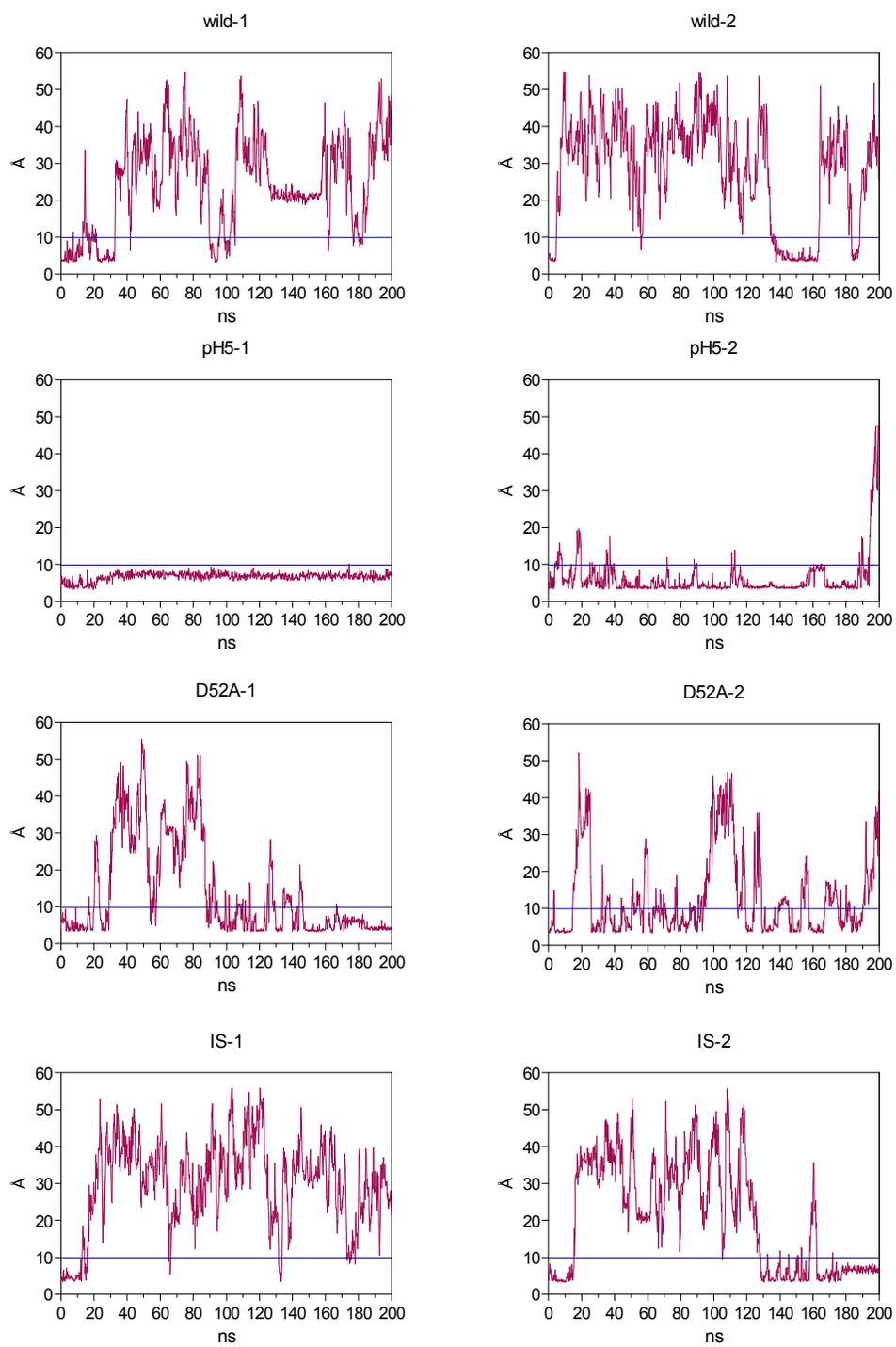
Figure B1 RMSD graphs of apo hFbpA systems



**Figure B2** Time-Distance ( $\text{H}_2\text{PO}_4^-$  (P) between Y195(O)) Graphs for 1D9V Simulations



**Figure B3** Time-Distance ( $\text{H}_2\text{PO}_4^-(\text{P})$  between Y196(O)) Graphs for 1D9V Simulations



## APPENDIX C

**Table C1** Holo hFbpA pKa calculations

pKa values calculated for 1MRP by various methodologies. Ionic strength is set to 150 mM, external and internal dielectric constants are 80 and 10, respectively.

residue type	residue index	model pKa values (H++)	H++	propka 3.111	PKd	Phemto	average	std dev
<b>TYR</b>	5	9.60	10.52	11.50	16.06	-	12.69	2.96
<b>HIS</b>	9	6.60	0.75	6.96	0.41	4	3.03	3.08
<b>LYS</b>	10	10.40	9.16	10.30	9.42	9.5	9.59	0.49
<b>GLU</b>	11	4.40	5.65	4.05	3.61	4	4.33	0.90
<b>LYS</b>	18	10.40	9.99	10.48	10.76	10	10.31	0.38
<b>GLU</b>	21	4.40	4.77	3.56	3.69	3	3.75	0.74
<b>GLU</b>	23	4.40	5.00	4.47	4.07	2.5	4.01	1.08
<b>LYS</b>	27	10.40	9.69	10.67	11.08	9.5	10.24	0.76
<b>LYS</b>	34	10.40	9.09	10.32	10.30	3	8.18	3.50
<b>GLU</b>	36	4.40	3.52	6.34	1.68	13	6.14	4.96
<b>LYS</b>	43	10.40	9.06	11.51	13.41	-	11.33	2.18
<b>GLU</b>	44	4.40	4.39	4.65	4.66	3.5	4.30	0.55
<b>GLU</b>	45	4.40	1.05	3.24	2.10	4.5	2.72	1.48
<b>ASP</b>	47	4.00	5.08	3.07	3.51	1.5	3.29	1.47
<b>LYS</b>	48	10.40	10.21	11.25	10.99	12	11.11	0.74
<b>ASP</b>	52	4.00	0.99	3.49	2.95	5	3.11	1.66
<b>TYR</b>	55	9.60	13.96	13.45	14.64	11.5	13.39	1.35
<b>GLU</b>	57	4.40	5.02	1.63	0.00	2	2.16	2.09
<b>ASP</b>	64	4.00	4.82	2.24	0.00	-	2.35	2.41
<b>GLU</b>	67	4.40	4.36	4.70	4.23	3.5	4.20	0.51
<b>GLU</b>	76	4.40	4.60	3.50	4.05	4	4.04	0.45
<b>LYS</b>	85	10.40	9.54	10.31	10.83	10	10.17	0.54
<b>LYS</b>	92	10.40	10.04	10.42	10.85	10.5	10.45	0.33
<b>LYS</b>	93	10.40	9.68	10.82	10.84	10.5	10.46	0.54
<b>ASP</b>	94	4.00	3.66	4.07	2.93	2.5	3.29	0.71
<b>ARG</b>	101	12.00	4.23	15.43	0.00	-	6.55	7.97
<b>ARG</b>	103	12.00	7.46	9.79	19.01	13	12.31	5.01
<b>TYR</b>	107	9.60	12.59	11.23	13.61	11.5	12.23	1.09
<b>ASP</b>	108	4.00	3.38	3.05	2.36	2.5	2.82	0.47
<b>HID</b>	109	7.00	6.47	6.34	6.84	4	5.91	1.29
<b>LYS</b>	111	10.40	10.38	10.44	11.23	10	10.51	0.52
<b>GLU</b>	114	4.40	4.50	4.69	3.33	2.5	3.76	1.03

<b>LYS</b>	115	10.40	9.42	10.42	10.27	10.5	10.15	0.50
<b>ASP</b>	116	4.00	3.41	2.47	3.40	3	3.07	0.44
<b>GLU</b>	118	4.40	5.31	3.51	2.06	2.5	3.35	1.45
<b>LYS</b>	119	10.40	10.35	10.47	11.03	11	10.71	0.35
<b>ASP</b>	123	4.00	3.92	4.63	3.82	2.5	3.72	0.89
<b>TYR</b>	124	9.60	15.53	13.33	0.00	-	9.62	8.40
<b>LYS</b>	128	10.40	9.73	11.49	13.69	13	11.98	1.76
<b>LYS</b>	130	10.40	10.26	10.29	10.57	9.5	10.16	0.46
<b>LYS</b>	132	10.40	10.99	10.26	11.08	10.5	10.71	0.39
<b>TYR</b>	135	9.60	13.19	11.73	12.53	11	12.11	0.95
<b>GLU</b>	144	4.40	3.81	6.37	0.00	3.5	3.42	2.62
<b>LYS</b>	151	10.40	9.36	10.56	15.33	13	12.06	2.66
<b>LYS</b>	153	10.40	9.08	10.43	10.42	10.5	10.11	0.68
<b>ASP</b>	155	4.00	5.27	2.49	2.11	-	3.29	1.73
<b>LYS</b>	156	10.40	10.26	10.46	10.57	10.5	10.45	0.13
<b>LYS</b>	163	10.40	10.18	10.39	10.93	10.5	10.50	0.31
<b>LYS</b>	166	10.40	10.59	10.42	11.75	11	10.94	0.59
<b>GLU</b>	167	4.40	4.97	4.49	4.11	2.5	4.02	1.07
<b>LYS</b>	170	10.40	10.23	10.52	11.29	10.5	10.64	0.46
<b>TYR</b>	172	9.60	15.28	12.51	15.82	12	13.90	1.93
<b>LYS</b>	174	10.40	8.96	10.06	9.91	9.5	9.61	0.49
<b>GLU</b>	183	4.40	2.22	3.50	0.00	3.5	2.30	1.65
<b>GLU</b>	186	4.40	4.27	3.84	3.99	3	3.77	0.55
<b>TYR</b>	195	9.60	14.52	-	19.88	10.5	14.97	4.71
<b>TYR</b>	196	9.60	8.69	13.52	6.89	-	9.70	3.43
<b>TYR</b>	198	9.60	11.13	11.70	11.81	12.5	11.78	0.56
<b>LYS</b>	202	10.40	9.87	10.46	11.24	11	10.64	0.61
<b>GLU</b>	203	4.40	4.47	4.68	3.46	4	4.15	0.54
<b>LYS</b>	204	10.40	10.03	10.44	10.90	10.5	10.47	0.35
<b>GLU</b>	207	4.40	4.52	4.25	4.23	3	4.00	0.68
<b>LYS</b>	210	10.40	10.41	10.58	11.52	11	10.88	0.50
<b>ARG</b>	212	12.00	12.61	11.76	17.05	-	13.81	2.84
<b>TYR</b>	214	9.60	13.98	11.45	15.60	13	13.51	1.74
<b>ARG</b>	217	12.00	11.14	13.15	18.29	-	14.19	3.69
<b>HIS</b>	218	6.60	6.49	6.41	7.05	6.5	6.61	0.29
<b>ASP</b>	220	4.00	3.96	3.71	0.34	3	2.75	1.66
<b>TYR</b>	227	9.60	14.93	14.18	14.92	12.5	14.13	1.14
<b>LYS</b>	234	10.40	10.21	9.98	11.37	11	10.64	0.65
<b>LYS</b>	237	10.40	10.59	10.31	11.58	11	10.87	0.55
<b>GLU</b>	241	4.40	6.51	3.73	3.04	3.5	4.20	1.57
<b>LYS</b>	244	10.40	9.94	10.39	10.71	10.5	10.39	0.32

<b>ASP</b>	247	4.00	4.16	4.51	3.84	3	3.88	0.65
<b>LYS</b>	252	10.40	10.24	11.41	14.28	-	11.98	2.08
<b>LYS</b>	253	10.40	10.18	10.73	10.93	11	10.71	0.37
<b>GLU</b>	256	4.40	6.09	3.71	3.47	1	3.57	2.08
<b>ARG</b>	262	12.00	7.80	12.39	18.12	-	12.77	5.17
<b>GLU</b>	264	4.40	5.88	2.88	0.00	-	2.92	2.94
<b>TYR</b>	265	9.60	19.09	14.03	19.77	12	16.22	3.81
<b>ARG</b>	268	12.00	11.07	13.80	16.81	-	13.90	2.87
<b>ASP</b>	270	4.00	5.29	1.84	0.61	1	2.19	2.13
<b>GLU</b>	278	4.40	5.90	4.13	2.61	1	3.41	2.09
<b>TYR</b>	280	9.60	13.97	11.53	13.04	12	12.64	1.09
<b>GLU</b>	281	4.40	4.48	4.11	3.98	3.5	4.02	0.41
<b>LYS</b>	282	10.40	9.89	10.53	11.10	11	10.63	0.55
<b>GLU</b>	284	4.40	4.50	3.98	3.67	3.5	3.91	0.44
<b>ASP</b>	295	4.00	3.80	2.39	1.50	1	2.17	1.23
<b>LYS</b>	296	10.40	6.63	10.74	11.31	12	10.17	2.42
<b>GLU</b>	297	4.40	4.60	4.52	3.85	2.5	3.87	0.97
<b>HIS</b>	298	6.60	4.74	7.33	7.92	-	6.67	1.69
<b>LYS</b>	301	10.40	10.20	10.61	11.69	11	10.87	0.63
<b>GLU</b>	304	4.40	4.61	4.91	4.86	3.5	4.47	0.66
<b>GLU</b>	305	4.40	5.22	4.55	4.65	3.5	4.48	0.72
<b>LYS</b>	309	10.40	7.83	3.31	9.87	9.5	7.63	3.01

**Table C2** Apo hFbpA pKa calculations

pKa values calculated for 1D9V by various methodologies. Ionic strength is set to 150 mM, external and internal dielectric constants are 80 and 10, respectively.

residue type	residue index	model pKa values (H++)	H++	propka 3.1	PKd	PHEMTO	Average	std dev
<b>TYR</b>	5	9.60	10.21	11.50	15.76	4.50	10.49	4.65
<b>HIS</b>	9	6.60	1.17	6.96	0.76	-	2.96	3.47
<b>LYS</b>	10	10.40	9.83	10.30	10.41	10.50	10.26	0.30
<b>GLU</b>	11	4.40	5.15	4.05	4.03	4.00	4.31	0.56
<b>LYS</b>	18	10.40	9.87	10.48	10.69	10.00	10.26	0.39
<b>GLU</b>	21	4.40	5.64	3.56	3.62	2.50	3.83	1.31
<b>GLU</b>	23	4.40	4.57	4.47	4.29	3.00	4.08	0.73
<b>LYS</b>	27	10.40	9.85	10.67	11.47	11.00	10.75	0.68
<b>LYS</b>	34	10.40	9.39	10.32	10.74	10.50	10.24	0.59
<b>GLU</b>	36	4.40	3.44	6.34	2.81	3.00	3.90	1.65

<b>LYS</b>	43	10.40	7.57	11.51	9.61	10.00	9.67	1.62
<b>GLU</b>	44	4.40	4.51	4.65	5.02	3.50	4.42	0.65
<b>GLU</b>	45	4.40	0.86	3.24	3.12	4.50	2.93	1.52
<b>ASP</b>	47	4.00	5.21	3.07	2.74	1.50	3.13	1.54
<b>LYS</b>	48	10.40	10.20	11.25	11.61	11.50	11.14	0.64
<b>ASP</b>	52	4.00	2.20	3.49	1.76	5.00	3.11	1.46
<b>TYR</b>	55	9.60	12.82	13.45	13.36	11.50	12.78	0.90
<b>GLU</b>	57	4.40	2.94	1.63	1.39	1.50	1.87	0.72
<b>ASP</b>	64	4.00	3.81	2.24	1.88	1.50	2.36	1.01
<b>GLU</b>	67	4.40	4.57	4.70	4.27	3.50	4.26	0.54
<b>GLU</b>	76	4.40	4.81	3.50	3.23	2.50	3.51	0.96
<b>LYS</b>	85	10.40	10.49	10.31	10.83	10.50	10.53	0.22
<b>LYS</b>	92	10.40	9.94	10.42	12.46	10.00	10.70	1.19
<b>LYS</b>	93	10.40	8.98	10.82	10.76	11.00	10.39	0.94
<b>ASP</b>	94	4.00	4.14	4.07	3.71	2.50	3.60	0.76
<b>ARG</b>	101	12.00	9.87	15.43	0.00	-	8.43	7.81
<b>ARG</b>	103	12.00	8.24	9.79	18.35	13.00	12.35	4.47
<b>TYR</b>	107	9.60	13.17	11.23	13.76	12.00	12.54	1.14
<b>ASP</b>	108	4.00	3.85	3.05	3.14	3.00	3.26	0.40
<b>HID</b>	109	7.00	6.31	6.34	6.38	4.00	5.76	1.17
<b>LYS</b>	111	10.40	10.10	10.44	11.73	10.50	10.69	0.71
<b>GLU</b>	114	4.40	4.63	4.69	3.20	2.50	3.75	1.08
<b>LYS</b>	115	10.40	10.42	10.42	10.71	10.50	10.51	0.14
<b>ASP</b>	116	4.00	4.39	2.47	3.71	3.00	3.39	0.84
<b>GLU</b>	118	4.40	4.32	3.51	3.94	4.50	4.07	0.44
<b>LYS</b>	119	10.40	9.98	10.47	11.27	11.00	10.68	0.57
<b>ASP</b>	123	4.00	3.91	4.63	3.16	2.50	3.55	0.92
<b>TYR</b>	124	9.60	15.71	13.33	0.00	13.00	10.51	7.11
<b>LYS</b>	128	10.40	9.00	11.49	10.37	10.00	10.21	1.03
<b>LYS</b>	130	10.40	10.69	10.29	10.93	10.00	10.48	0.41
<b>LYS</b>	132	10.40	10.40	10.26	11.26	10.50	10.61	0.45
<b>TYR</b>	135	9.60	12.98	11.73	12.63	11.00	12.08	0.89
<b>GLU</b>	144	4.40	4.76	6.37	0.98	2.50	3.65	2.39
<b>LYS</b>	151	10.40	9.30	10.56	15.89	-	11.91	3.50
<b>LYS</b>	153	10.40	9.70	10.43	10.50	10.00	10.16	0.38
<b>ASP</b>	155	4.00	2.35	2.49	1.12	1.00	1.74	0.79
<b>LYS</b>	156	10.40	10.27	10.46	10.52	10.50	10.44	0.12
<b>LYS</b>	163	10.40	8.63	10.39	9.02	9.00	9.26	0.77
<b>LYS</b>	166	10.40	10.54	10.42	11.18	10.50	10.66	0.35
<b>GLU</b>	167	4.40	4.86	4.49	4.24	3.00	4.15	0.81
<b>LYS</b>	170	10.40	9.73	10.52	10.74	10.50	10.37	0.44

<b>TYR</b>	172	9.60	13.84	12.51	12.31	11.50	12.54	0.97
<b>LYS</b>	174	10.40	9.34	10.06	10.14	10.00	9.89	0.37
<b>GLU</b>	183	4.40	2.30	3.50	0.00	2.00	1.95	1.45
<b>GLU</b>	186	4.40	4.88	3.84	4.07	3.00	3.95	0.77
<b>TYR</b>	195	9.60	9.90	-	10.33	12.00	10.75	1.11
<b>TYR</b>	196	9.60	9.08	13.52	7.18	10.50	10.07	2.67
<b>TYR</b>	198	9.60	12.38	11.70	11.21	12.00	11.82	0.49
<b>LYS</b>	202	10.40	9.95	10.46	10.44	9.50	10.09	0.46
<b>GLU</b>	203	4.40	4.49	4.68	3.56	2.50	3.81	1.00
<b>LYS</b>	204	10.40	9.76	10.44	11.07	11.00	10.57	0.61
<b>GLU</b>	207	4.40	5.07	4.25	3.97	3.00	4.07	0.86
<b>LYS</b>	210	10.40	10.02	10.58	13.24	11.00	11.21	1.41
<b>ARG</b>	212	12.00	12.20	11.76	16.85	-	13.60	2.82
<b>TYR</b>	214	9.60	14.61	11.45	15.36	13.00	13.60	1.74
<b>ARG</b>	217	12.00	10.87	13.15	18.31	-	14.11	3.81
<b>HIS</b>	218	6.60	6.44	6.41	6.77	5.00	6.15	0.79
<b>ASP</b>	220	4.00	1.29	3.71	1.16	3.50	2.41	1.38
<b>TYR</b>	227	9.60	13.50	14.18	18.23	-	15.30	2.56
<b>LYS</b>	234	10.40	10.35	9.98	10.87	10.00	10.30	0.42
<b>LYS</b>	237	10.40	10.89	10.31	11.31	11.00	10.88	0.42
<b>GLU</b>	241	4.40	6.86	3.73	2.89	3.50	4.24	1.78
<b>LYS</b>	244	10.40	10.21	10.39	10.96	11.00	10.64	0.40
<b>ASP</b>	247	4.00	4.95	4.51	4.80	4.50	4.69	0.22
<b>LYS</b>	252	10.40	9.95	11.41	12.77	12.50	11.66	1.28
<b>LYS</b>	253	10.40	9.77	10.73	10.70	10.50	10.43	0.45
<b>GLU</b>	256	4.40	6.05	3.71	3.53	2.00	3.82	1.67
<b>ARG</b>	262	12.00	9.01	12.39	16.64	-	12.68	3.82
<b>GLU</b>	264	4.40	4.63	2.88	0.00	3.50	2.75	1.97
<b>TYR</b>	265	9.60	15.01	14.03	19.27	12.00	15.08	3.06
<b>ARG</b>	268	12.00	11.38	13.80	16.25	-	13.81	2.43
<b>ASP</b>	270	4.00	5.15	1.84	1.21	1.00	2.30	1.93
<b>GLU</b>	278	4.40	7.21	4.13	2.26	2.00	3.90	2.40
<b>TYR</b>	280	9.60	16.32	11.53	14.80	12.00	13.66	2.29
<b>GLU</b>	281	4.40	4.78	4.11	4.04	3.00	3.98	0.74
<b>LYS</b>	282	10.40	10.21	10.53	11.20	11.00	10.74	0.45
<b>GLU</b>	284	4.40	5.63	3.98	3.22	3.50	4.08	1.08
<b>ASP</b>	295	4.00	5.09	2.39	1.00	-	2.83	2.08
<b>LYS</b>	296	10.40	8.39	10.74	12.16	11.50	10.70	1.64
<b>GLU</b>	297	4.40	4.52	4.52	4.07	3.50	4.15	0.48
<b>HIS</b>	298	6.60	4.98	7.33	5.28	5.50	5.77	1.06
<b>LYS</b>	301	10.40	9.76	10.61	11.10	10.50	10.49	0.55

<b>GLU</b>	304	4.40	4.93	4.91	3.50	2.00	3.84	1.40
<b>GLU</b>	305	4.40	4.88	4.55	4.24	3.00	4.17	0.82
<b>LYS</b>	309	10.40	8.95	9.53	11.03	11.50	10.25	1.21

## References

1. Weaver, K.D., et al., *Role of Citrate and Phosphate Anions in the Mechanism of Iron(III) Sequestration by Ferric Binding Protein: Kinetic Studies of the Formation of the Holoprotein of Wild-Type FbpA and Its Engineered Mutants*. *Biochemistry*, 2010. **49**(29): p. 6021-6032.
2. Siburt, C.J.P., T.A. Mietzner, and A.L. Crumbliss, *FbpA - A bacterial transferrin with more to offer*. *Biochimica Et Biophysica Acta-General Subjects*, 2012. **1820**(3): p. 379-392.
3. Guo, M.L., et al., *Synergistic anion and metal binding to the ferric ion-binding protein from Neisseria gonorrhoeae*. *Journal of Biological Chemistry*, 2003. **278**(4): p. 2490-2502.
4. Bruns, C.M., et al., *Structure of Haemophilus influenzae Fe<sup>3+</sup>-binding protein reveals convergent evolution within a superfamily*. *Nature Structural Biology*, 1997. **4**(11): p. 919-924.
5. Ferreira, C., M.T. Criado, and J.A. Gomez, *The neisserial 37 kDa ferric binding protein (FbpA)*. *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology*, 1999. **123**(1): p. 1-7.
6. Mietzner, T.A., et al., *Fe(III) periplasm-to-cytosol transporters of gram-negative pathogens*. *Bacterial Infection: Close Encounters at the Host Pathogen Interface*, 1998. **225**: p. 113-135.
7. Bruns, C.M., et al., *Crystallographic and biochemical analyses of the metal-free Haemophilus influenzae Fe<sup>3+</sup>-binding protein*. *Biochemistry*, 2001. **40**(51): p. 15631-15637.
8. Beveridge, T.J., *Structures of gram-negative cell walls and their derived membrane vesicles*. *Journal of Bacteriology*, 1999. **181**(16): p. 4725-4733.
9. Krewulak, K.D. and H.J. Vogel, *Structural biology of bacterial iron uptake*. *Biochimica Et Biophysica Acta-Biomembranes*, 2008. **1778**(9): p. 1781-1804.
10. Siburt, C.J.P., et al., *Hijacking transferrin bound iron: protein-receptor interactions involved in iron transport in N-gonorrhoeae*. *Metallomics*, 2009. **1**(3): p. 249-255.
11. Chen, C.Y., et al., *The Ferric Iron-Binding Protein of Pathogenic Neisseria SPP Functions as a Periplasmic Transport Protein in Iron Acquisition from Human Transferrin*. *Molecular Microbiology*, 1993. **10**(2): p. 311-318.

12. Khambati, H.K., et al., *The role of vicinal tyrosine residues in the function of Haemophilus influenzae ferric-binding protein A*. Biochemical Journal, 2010. **432**: p. 57-64.
13. Shouldice, S.R., et al., *High resolution structure of an alternate form of the ferric ion binding protein from Haemophilus influenzae*. Journal of Biological Chemistry, 2003. **278**(13): p. 11513-11519.
14. Shouldice, S.R., et al., *Presence of ferric hydroxide clusters in mutants of Haemophilus influenzae ferric ion-binding protein A*. Biochemistry, 2003. **42**(41): p. 11908-11914.
15. Khan, A.G., et al., *The role of the synergistic phosphate anion in iron transport by the periplasmic iron-binding protein from Haemophilus influenzae*. Biochemical Journal, 2007. **403**: p. 43-48.
16. Khan, A.G., et al., *High-affinity binding by the periplasmic iron-binding protein from Haemophilus influenzae is required for acquiring iron from transferrin*. Biochemical Journal, 2007. **404**: p. 217-225.
17. Atilgan, C. and A.R. Atilgan, *Perturbation-Response Scanning Reveals Ligand Entry-Exit Mechanisms of Ferric Binding Protein*. Plos Computational Biology, 2009. **5**(10): p. 1-14.
18. Fleischmann, R.D., et al., *Whole-Genome Random Sequencing and Assembly of Haemophilus-Influenza RD*. Science, 1995. **269**(5223): p. 496-512.
19. Mukherjee, A., et al., *Ferric ion (hydr)oxo clusters in the "Venus flytrap" cleft of FbpA: Mossbauer, calorimetric and mass spectrometric studies*. Journal of Biological Inorganic Chemistry, 2012. **17**(4): p. 573-588.
20. Beglov, D. and B. Roux, *Finite Representation of an Infinite Bulk System - Solvent Boundary Potential for Computer-Simulations*. Journal of Chemical Physics, 1994. **100**(12): p. 9050-9063.
21. Gunasekaran, K., B.Y. Ma, and R. Nussinov, *Is allostery an intrinsic property of all dynamic proteins?* Proteins-Structure Function and Bioinformatics, 2004. **57**(3): p. 433-443.
22. MacKerell, A.D., et al., *All-atom empirical potential for molecular modeling and dynamics studies of proteins*. Journal of Physical Chemistry B, 1998. **102**(18): p. 3586-3616.

23. Yang, W., et al., *The missing link between thermodynamics and structure in F1-ATPase*. Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(3): p. 874-879.
24. Tsai, C.J., A. del Sol, and R. Nussinov, *Allostery: Absence of a change in shape does not imply that allostery is not at play*. Journal of Molecular Biology, 2008. **378**(1): p. 1-11.
25. Swope, W.C., et al., *A Computer-Simulation Method for the Calculation of Equilibrium-Constants for the Formation of Physical Clusters of Molecules - Application to Small Water Clusters*. Journal of Chemical Physics, 1982. **76**(1): p. 637-649.
26. Atilgan, A.R., et al., *Anisotropy of fluctuation dynamics of proteins with an elastic network model*. Biophysical Journal, 2001. **80**(1): p. 505-515.
27. Yilmaz, L.S. and A.R. Atilgan, *Identifying the adaptive mechanism in globular proteins: Fluctuations in densely packed regions manipulate flexible parts*. Journal of Chemical Physics, 2000. **113**(10): p. 4454-4464.
28. Berman, H.M., et al., *The Protein Data Bank*. Acta Crystallographica Section D-Biological Crystallography, 2002. **58**: p. 899-907.
29. Bahar, I. and R.L. Jernigan, *Inter-residue potentials in globular proteins and the dominance of highly specific hydrophilic interactions at close separation*. Journal of Molecular Biology, 1997. **266**(1): p. 195-214.
30. Baysal, C. and A.R. Atilgan, *Coordination topology and stability for the native and binding conformers of chymotrypsin inhibitor 2*. Proteins-Structure Function and Genetics, 2001. **45**(1): p. 62-70.
31. Baysal, C. and A.R. Atilgan, *Elucidating the structural mechanisms for biological activity of the chemokine family*. Proteins-Structure Function and Genetics, 2001. **43**(2): p. 150-160.
32. Ikeguchi, M., et al., *Protein structural change upon ligand binding: Linear response theory*. Physical Review Letters, 2005. **94**(7): p. 1-4.
33. Anandakrishnan, R., B. Aguilar, and A.V. Onufriev, *H<sup>++3.0</sup>: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations*. Nucleic Acids Research, 2012. **40**(W1): p. W537-W541.

34. Tynan-Connolly, B.M. and J.E. Nielsen, *pKD: re-designing protein pK(a) values*. Nucleic Acids Research, 2006. **34**: p. W48-W51.
35. Bas, D.C., D.M. Rogers, and J.H. Jensen, *Very fast prediction and rationalization of pK(a) values for protein-ligand complexes*. Proteins-Structure Function and Bioinformatics, 2008. **73**(3): p. 765-783.
36. Li, H., A.D. Robertson, and J.H. Jensen, *Very fast empirical prediction and rationalization of protein pK(a) values*. Proteins-Structure Function and Bioinformatics, 2005. **61**(4): p. 704-721.
37. Olsson, M.H.M., et al., *PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pK(a) Predictions*. Journal of Chemical Theory and Computation, 2011. **7**(2): p. 525-537.
38. Sondergaard, C.R., et al., *Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pK(a) Values*. Journal of Chemical Theory and Computation, 2011. **7**(7): p. 2284-2295.
39. Kantardjiev, A.A. and B.P. Atanasov, *PHEMTO: protein pH-dependent electric moment tools*. Nucleic Acids Research, 2009. **37**: p. W422-W427.
40. Goldenberg, O., et al., *The ConSurf-DB: pre-calculated evolutionary conservation profiles of protein structures*. Nucleic Acids Research, 2009. **37**: p. D323-D327.
41. Shi, Y.B., et al., *Niobium Uptake and Release by Bacterial Ferric Ion Binding Protein*. Bioinorganic Chemistry and Applications, 2010: p. 1-11.
42. Weaver, K.D., et al., *Ga(3+) as a mechanistic probe in Fe(3+) transport: characterization of Ga(3+) interaction with FbpA*. Journal of Biological Inorganic Chemistry, 2008. **13**(6): p. 887-898.
43. Humphrey, W., A. Dalke, and K. Schulten, *VMD: Visual molecular dynamics*. Journal of Molecular Graphics & Modelling, 1996. **14**(1): p. 33-38.
44. Yang, N., et al., *Iron and bismuth bound human serum transferrin reveals a partially-opened conformation in the N-lobe*. Scientific Reports, 2012. **2**: p. 1-7.