INVESTIGATION OF PTP1B ACTIVATION MECHANISM BY COMPUTATIONAL METHODS

by

Burcu Özkaral

B.S., Chemical Engineering, İzmir Institute of Technology, 2007

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Chemical Engineering
Boğaziçi University

2010

ACKNOWLEDGEMENTS

I would like to express my gratitude to my thesis supervisor and co-supervisor, Assist. Prof. Elif Özkırımlı Ölmez and Assist. Prof. Burak Alakent, for their invaluable guidance, suggestions and support they have given me during my research.

I would like to thank my thesis committee members Prof. Pemra Doruker Turgut, Prof. Türkan Haliloğlu and Assist. Prof. Bülent Balta for spending their time to read my thesis and for their constructive comments.

I am grateful to all my friends, especially Aysun Eren, Celal Ceylan and Kübra Büyükyanbolu Karagöz for their encouragements, for sharing all the good and bad times and for being such great friends. I am grateful to Pınar Kanlıkılıçer and Ahmet Özcan for their helps, suggestions and sincere friendship. My sincere thanks go to my friends Mine Gün, İpek Paksoy, Yasemen Güngörmez, Aslıgül Doğan, Pınar Derin, Ömür Akdağ, Ezgi Akkaya and Nilay Büdeyri for their support and for giving me their friendship. My special thanks go to Murat Tergip for his never-ending support, encouragement and love. Finally, the biggest thanks go to my mother Nermin Özkaral, my father Mahmut Özkaral and my sister Duygu Alıcıoğlugil for their ever-lasting encouragement and support. This thesis is dedicated to them.

TUBITAK Project No. 107T863 is gratefully acknowledged for the funding.

ABSTRACT

INVESTIGATION OF PTP1B ACTIVATION MECHANISM BY COMPUTATIONAL METHODS

Protein tyrosine phosphatase 1B (PTP1B) is a negative regulator of insulin and leptin signaling, and is a major molecular target for the treatment of type II diabetes and obesity. WPD loop is a key element in the mechanism of PTP1B catalysis. In the apo form of the enzyme, WPD loop is usually in an "open" conformation, whereas it closes over the active site upon substrate binding. Here, targeted molecular dynamics (TMD) simulations are reported to examine the transition of the WPD loop from the open to closed states as well as the effect of this motion on the PTP1B conformational activation mechanism. Targeting potential was applied to the WPD loop only and to the whole protein in different sets of simulations. Residue-residue interactions and dihedral angles that contribute to the WPD loop conformational transition were identified. Two major conformational transitions between the open and closed states of WPD loop were observed using PCA and K-means clustering analysis of the C_α atoms in all TMD simulations, except TMD4 simulation. The first transition was the backbone dihedral angle rotation of Ser187 or Pro188 at the WPD loop C-terminus. The second transition was the simultaneous rotations of Trp179 and Arg221 sidechain dihedral angles, and the formation of a polar interaction between the WPD loop and the Arg221 sidechain. A third subtle conformational change, which was not observed in the clustering analysis, was the rotation of the backbone dihedral angle between Asp181 and Phe182 backbone dihedral angles, resulting in the formation of hydrogen bonds between Asp181-Gly183 and between Phe182-Gln262 at. It was also was found that WPD loop closing motion was hindered due to the alternative conformations of the Arg221 and the closed conformation of the R loop. Hydrophobic interactions between the WPD loop and the regions around it, such as the active site, helices α3, α6, α7, loop between β9-β10 (loop11) and S loop were also investigated and it was found that these regions may act as steric hinderance that could influence the WPD loop closure during the simulations. Elucidating the detailed mechanism of PTP1B conformational activation will guide future drug design efforts toward type II diabetes and obesity.

ÖZET

PTP1B AKTİVASYON MEKANİZMASININ HESAPSAL YOLLARLA İNCELENMESİ

Tirosin Fosfataz 1B (PTP1B) enzimi diyabet ve obezite hastalıklarındaki rolü nedeniyle önemli bir akıllı ilaç tasarımı hedefidir. WPD döngüsü PTP1B'nin kataliz mekanizmasında önemli bir elementtir. Ligand bağlı olmadığı durumlarda, WPD döngüsü genellikle açık halde iken ligand bağlandığı durumda WPD döngüsü aktif bölgeye doğru kapanır. Bu çalışmada, hedefli moleküler dinamik (TMD) simülasyonları, WPD döngüsünün açık ve kapalı durumları arasındaki geçişini ve aynı zamanda bu hareketin PTP1B yapısal aktivasyon mekanizması üzerindeki etkilerini incelemek için kullanılmıştır. Simülasyonların değişik setlerinde, kuvvet sadece WPD döngüsüne veya bütün protein atomlarına uygulanmıştır. WPD döngüsünün yapısal geçişinde katkısı olan kalıntı-kalıntı etkileşimleri ve dihedral açı geçişleri tanımlanmıştır. C_α atomlarının Temel Bileşenler Analizi ve K-ortalamaları kümelenmesi analizleri kullanılarak WPD döngüsünün açık ve kapalı durumları arasında iki ana yapısal geçiş bulunmuştur. Birinci geçiş, WPD döngüsünün N-ucunun (184-189 kalıntıları) hedef yapıya yaklaştığını gösteren Ser187 yada Pro188 kalıntılarının omurga dihedral açı dönüşleridir. İkinci geçiş ise Trp179 ve Arg221 yan zincirlerinin dihedral açılarının eşzamanlı dönüşleri ve WPD döngüsü ile Arg221 yan zinciri arasında kutupsal etkileşimleridir. Kümelenme analizinde gözlenmeyen üçüncü bir yapısal değişiklik de Asp181-Gly183 ve Phe182-Gln262 kalıntıları aralarında hidrojen bağları oluşumuyla sonuçlanan Asp181 ve Phe182 kalıntılarının omurga dihedral açı dönüşleriyle tanımlanmıştır. İki ayrı TMD simülasyonunda, Arg221 kalıntısının alternative yapıları ve R döngüsünün "kapalı" konformasyonu WPD döngüsü kapanışına engel olmuştur. WPD döngüsü ile aktif bölge, α3, α6, α7 sarmalları, β9-β10 yaprakları arasında döngü, S döngüsü gibi bölgeler arasındaki hidrofobik etkileşimlerin, sterik engellemeler ile WPD döngüsünün kapanmasında etkili olduğu görülmüştür. PTP1B detaylı yapısal aktivasyon mekanizmasının açığa kavuşturulması, bu konuda yapılacak benzer çalışmalara ışık tutacak, obesite ve diyabet gibi hastalıklara karşı gelecek ilaç tasarımı girişimlerine yol gösterecektir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS / ABBREVIATIONS

a_i	Acceleration of particle *i*
b_0	minimum energy bond length
F_i	Force acting on particle *i*
Da	Dalton
K_b	Bond force constant
K_0	Angle force constant
K_i	Inhibition constant
K_φ	Dihedral force constant
k_w	Force constant
k-Dal	kilo Dalton
m_i	Mass of particle *i*
ns	Nanosecond
$\langle R_i \rangle$	Vector of time average of the cartesian coordinates of the C_α atom of the same residue
$R_i(t)$	Vector of the cartesian coordinates of C_α atom of residue i at time t.
V	Potential energy of the system

α	Alpha
Å	Angstrom
β	Beta
χ	Chi
°	Degree
δ	Phase shift
Φ	Phi
Ψ	Psi
θ	Angle

Ala	Alanine
Arg	Arginine
Asp	Aspartic Acid
BPPM	Bis-(p-phosphophenyl) methane
C	Carbon
C-Kit	Cell factor receptor
Cys	Cysteine
Gln	Glutamine
H	Hydrogen
IRK	Insulin Receptor Kinase
IZD	Isothiazolidinone
JMR	Juxtamembrane region
KIM	Kinase Interaction Motif
L	Loop
L11	Loop 11
LacI	Lactase receptor protein
MD	Molecular Dynamics
MM PBSA	Molecular Mechanics Poisson Boltzmann Solvent Accessible Surface Area
MSF	Mean Squared Fluctuation
N	Nitrogen
NtrC	Nitrogen regulatory protein C
O	Oxygen
PCA	Principal Component Analysis
PDB	Protein Data Bank
Phe	Phenylalanine
Pro	Proline
PME	Particle Mesh Ewald
PTK	Protein Tyrosine Kinase
PTP1B	Protein Tyrosine Phosphatase 1B
PTyr	Phosphotyrosine
RMSD	Root Mean Square Deviation
Ser	Serine

TC-PTP	T-cell Protein Tyrosine Phosphatase
TMD	Targeted Molecular Dynamics
Trp	Tryptophan
Tyr	Tyrosine
vdW	Van der Waals
VMD	Visual Molecular Dynamics

1. INTRODUCTION

Reversible protein tyrosine phosphorylation reactions are of great importance in the regulation of many biological processes in eukaryotic cells [1]. Tyrosine phosphorylation is regulated by protein tyrosine kinases (PTKs), which catalyze the addition of phosphate to tyrosine residues, and by protein tyrosine phosphatases (PTPs), which remove the phosphate group from phosphotyrosine (pTyr) [2]. The sequence of the human genome completed in 2001 identified 107 human PTP genes that code for 81 catalytically active PTPs [3]. PTPs are a diverse class of enzymes that have been further subdivided into 17 subfamilies based on their degree of homology between functional and architectural domains. Protein Tyrosine Phosphatase 1B (PTP1B) is a member of the class I cysteine-based PTPs in the classical non-receptor PTP (NR-PTP) subfamily. PTP1B acts as a negative regulator of insulin and leptin signaling, so it is considered to be an important molecular target for the treatment of type 2 diabetes and obesity [4]. The native protein consists of 435 amino acid residues. Residues 30-278 define the catalytic domain and the 35 C-terminal residues target the protein to the cytosolic face of the endoplasmic reticulum [5]. In 1994, the first crystal structure of PTP1B was solved [6]. Lately, Barr et al. recently published a compendium of 22 new phosphatase X-ray crystal structures including at least one member of every PTP subgroup [7]. The WPD loop (residues 176- 187) assumes two different conformations in the available crystal structures. In the apo form, WPD-loop is usually in an “open” conformation, whereas it closes over the active site upon substrate binding, and forms a tight binding pocket for the substrate [6,8]. However, structures in which the WPD loop in the closed conformation in the absence of a ligand exist [9].

In this study, considering the importance of the WPD loop conformation and dynamics on the full catalytic activity of PTP1B [10], the conformational transition mechanism of the WPD loop was investigated. Equilibrium molecular dynamics (MD) and targeted molecular dynamics (TMD) simulations have been performed to elucidate the atomistic details of the loop motion, and to identify the residue interaction networks participating in the closing/opening motion of WPD loop. Elucidating the conformational activation mechanism will guide future drug design efforts toward type II diabetes and

obesity. TMD simulations were performed using different initial and target structures, different force constant values, and applying force on different regions of the protein. Two major conformational transition of WPD loop between the open and closed states were found using Principal Component Analysis (PCA) and k-means clustering analysis of the C_α atoms. Important polar and hydrophobic interactions in the vicinity of the WPD loop were investigated in detail.

The outline of this thesis is as follows: the biological significance, tertiary structure, catalytic mechanism of PTP1B, the interactions in the PTP1B-ligand complexes, the significance of the WPD loop for the function of PTP1B, and the simulation studies on PTP1B are discussed in Section 2. In Section 3, the protein structures, forcefield parameters and protocols of the MD and TMD simulations are discussed. Simulation trajectories are analyzed and discussed to identify the residue-residue interactions between the WPD loop and the regions around it, the dihedral angle transitions of WPD loop residues and some hydrophobic interactions during the WPD loop closure/opening in the TMD simulations in Section 4. A summary of the results and recommendations for future work are included in Section 5.

2. PROTEIN TYROSINE PHOSPHATASE 1B

2.1. Biological Significance of Protein Tyrosine Phosphatase 1B

Protein tyrosine phosphorylation is the predominant post-translational modification utilized by the cells to regulate signal transduction and to control numerous cellular functions, such as growth, survival, differentiation, apoptosis, metabolism and gene transcription [11]. Tyrosine phosphorylation is tightly controlled by the coordinated actions of protein tyrosine kinases (PTKs), which catalyze the addition of phosphate to tyrosine residues and phosphatases (PTPs), which remove the phosphate group from phosphotyrosine (pTyr) [9]. Loss of this control leads to aberrant tyrosine phosporylation and has been implicated in the development of many human diseases, such as diabetes, obesity, cancer, inflammation and neurodegenerative diseases [11]. Hence, the role of PTPs in diverse pathophysiological conditions results in great interest in PTPs as molecular targets for the development of novel therapeutic agents [12].

There are about 100 PTP genes encoded within the human genome, including transmembrane, receptor-like and intracellular enzymes. The prototypical member of the PTP family is the enzyme PTP1B. PTP1B is located on the cytoplasmic side of the endoplasmic reticulum (ER) and is expressed in the insulin-targeted tissues such as liver, muscle and fat. There is evidence from biochemical, genetic and pharmacological studies that PTP1B plays an important role as a negative regulator in both insulin and leptin signalling (Figure 2.1). There are numerous protein tyrosine phosphatases carrying out this role, such as PTPα LAR, CD45, PTPε, SHP2, T-cell protein tyrosine phosphatase (TCPTP) and PTP1B. These PTPs control insulin signaling in vitro when overexpressed in cell lines, but only PTP1B has significant effect on insulin signaling in knockout mice [13]. Inhibition of PTP1B increases insulin sensitivity and resistance to obesity, and does not cause abnormalities in growth of fertility or other pathogenetic effects in mice [4]. While overexpression of PTP1B in cell cultures decreases insulin-stimulated phosphorylation of IR/IRS-1, decrease in PTP1B level increases the insulin intiated signaling [14]. PTP1B also participates in leptin signaling by binding and dephosphorylating JAK2, which is downstream of the leptin receptor [12]. The hormone leptin plays a key role in body mass

regulation [14]. Consequently, these genetic, biochemical and pharmacological studies have shown that inhibition of PTP1B may address both diabetes and obesity and make PTP1B an attractive target for drug design improvement [12].

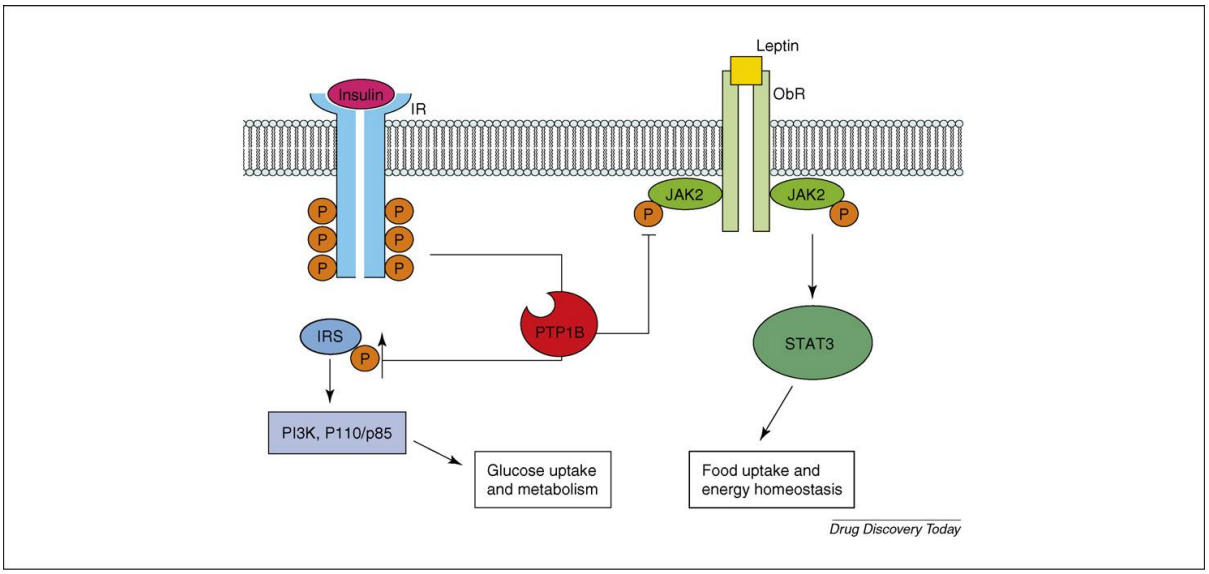

Figure 2.1. The role of PTP1B in insulin and leptin signalling. In the insulin pathway, PTP1B can associate with and dephosphorylate activated Insulin recptor (IR), or insulin receptor substrates (IRS). In the leptin pathway, PTP1B binds and dephosphorylates JAK2, which is downstream of the leptin receptor, ObR. STAT3 and P110/p85 are downstream targets of JAK2 and IRS1, respectively. P110/p85 is a specific form of PI3K responsive to insulin signalling [12].

There is also evidence that PTP1B has an oncogenic role. It is overexpressed in some human cancers and acts as a positive regulator of tumor onset and progression [11]. PTP1B is also a positive regulator of Erb2-PTK, which is overexpressed in many human breast cancers. PTP1B deficiency or inhibition destroys Erb2-induced tumorogenesis and protects from lung metastasis [11].

2.2. Tertiary Structure of PTP1B

PTP1B is the prototypical member of a family of about 130 PTPs [3] and it is the first PTP structure to be resolved crystallographically [6]. Its structure is composed of a single domain with the polypeptide chain organized into eight α helices and 12 β strands

with a 10-stranded mixed β-sheet that adopts a highly twisted conformation, spanning the entire length of the molecule (Figure2.2). The final β strand, β12, in the primary sequence, is located close to the center of the sheet within the parallel β strand region with strand order β3, β12, β4, and β11. Antiparallel β strands flank the four central parallel β strands. The center of the sheet is buried by α helices with α2 on one side and α3 and α4 on the other. Following α4, the chain folds into α5 and α6, which, with α3 and α4, form a four-helix bundle. The COOH-terminus of the conserved catalytic domain corresponds to the COOH terminus of α6. A short α helix (α1) found at the top of the β sheet close to β1, indicating the beginning of the conserved PTP domain. An antiparallel β sheet (β5 and β6) is located above the central β sheet. The nonconserved NH_2-terminal of 30 residues folds into two α helices (α1' and α2') [6].

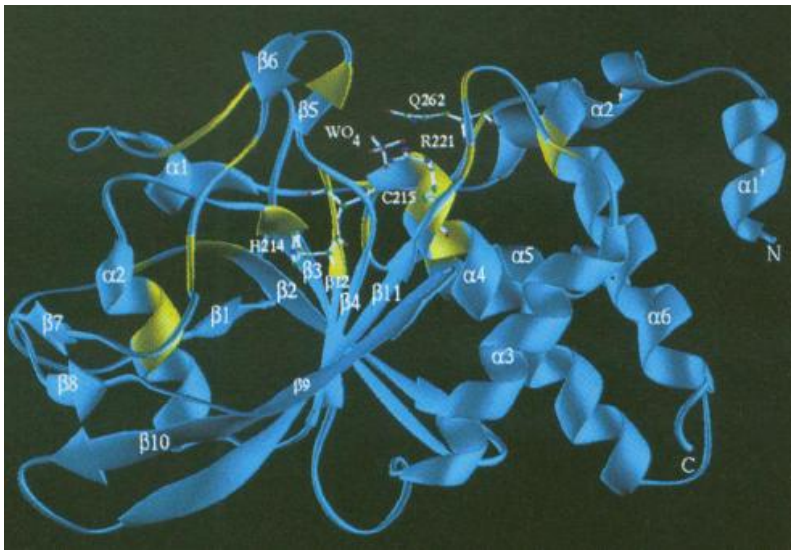

Figure 2.2. Cartoon representation of PTP1B crystal structure (PDB ID: 2HNP) in which WPD loop is in open conformation indicating the secondary structure elements, the catalytic site, and invariant residues (yellow). The α helices and 13 β strands are labeled. The tungstate ion and the side chains of His214, Cys215, Arg221, and Gln262 are shown. The catalytic site is located at the COOH terminus of the central parallel region of the sheet. [6].

The catalytically and structurally important regions and the residues on PTP1B tertiary structure are listed in Table 2.1.

Table 2.1. The important regions and residues on PTP1B

Regions on PTP1B	**Residues**
P loop (Active Site) [6, 9, 13, 15]	214-221
WPD loop [8,9,10,13,23]	176-187
pTyr recognition loop [8, 9, 19, 20, 21]	46-49
Secondary binding site [13, 16, 17, 18]	258, 256, 260, 262, 266, 24, 254
Allosteric binding site [22]	188, 191, 192, 193, 196, 276, 280
R loop [6, 9, 23]	114-119
Q loop [8,9]	261-265
S loop [10, 23]	198-209
α3 helix [10, 13]	190-198
α6 helix [10]	264-280
α7 helix [10, 13]	282-298

Figure 2.3 shows the structurally and biochemically important regions of PTP1B. The catalytic site is located within a crevice on the molecular surface, defined by the position of the important cysteine residue (Cys215). The base of the catalytic site is made up of the residues from His214 to Arg221, corresponding to the conserved motif characteristic of PTPs [6]. The active site residues that interact with substrates/inhibitors are located in the active site (P loop), and four other loop regions that surround the active site: pTyr-recognition loop (residues 46-49), the R loop (residues 113-118), the Q loop (residues 261-265), and WPD loop (residues 176-187). The flexible WPD loop is found to be either in open (WPD_{open}) or closed (WPD_{closed}) conformation in the resolved structures

of PTP1B. The S loop (201-209) has been suggested to play an important role in WPD loop dynamics [23]. Analysis of all available PTP1B crystal structures shows the existence of an additional structural element, named α7 helix (residues 282-299) which is resolved only in the holo form of the enzyme.

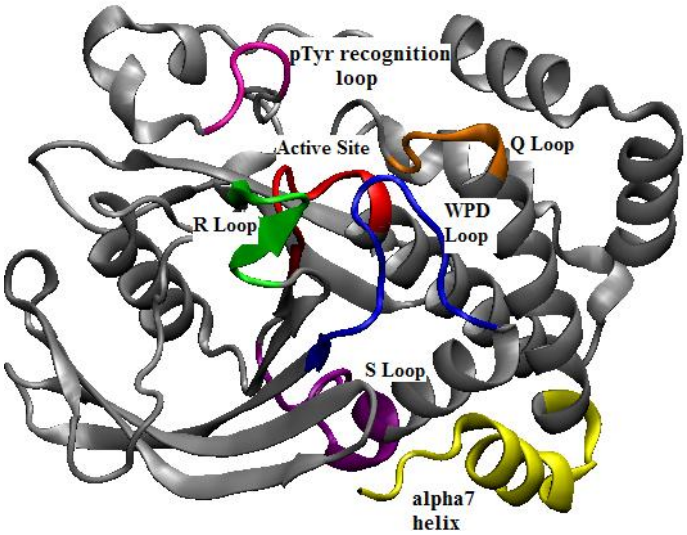

Figure 2.3. A schematic diagram of PTP1B in the closed conformation (PDB ID: 1SUG). The WPD loop (residues 176–187) is shown in blue, the active site (residues 214–221) in red, the R-loop (residues 113–118) in green, the S-loop (residues 198– 209) in purple, the pTyr recognition loop (residues 46-49) in pink, the Q loop (residues 261-265) in orange and the α7-helix (residues 280–298) is in yellow.

The conformations of WPD loop and R loop and the presence of the ligand and α7 helix structures are important structural features in PTPs, listed in Table 2.2. The PTP1B structures the crystal structures with PDB codes 1SUG, 2HNP, 2F6F, 1PTU, 1T48 and 1Q6J while the other PTP structures which are not PTP1B are the crystal structures with PDB codes 1L8K, 1YTS, 2A8B, 2A3K and 1FPR. In all PTPB structures except 1Q6J, R loop is in the closed conformation, which is close to the active site. On the other side, it is found to be in an alternative conformation or disordered in some of the other PTPs. Alternative “open” or “closed” conformation means that R loop is in such a conformation that it has a different conformation in open or closed state. 1SUG, 2HNP and 2F6F PTPB structures are not liganded while 1PTU crystal structure of PTP1B is the PTP1B-pTyr peptide substrate. The other PTP structures are enzyme-inhibitor complexes.

Table 2.2. The important structural features of PTP1B and some other PTPs

PDB code	**WPD loop**	**ligand**	**α7 helix**	**R loop**
1SUG [9]	closed	-	present	closed
2F6F [24]	open	-	present	closed
2HNP [6]	open	-	not present	closed
1PTU [8]	closed	bound	present	closed
1T48 [22]	open	bound	not present	closed
1L8K [25]	open	bound	not present	alternative "open"
1YTS [15]	closed	bound	not present	disordered
2A8B [26]	intermediate	bound	not present	alternative "closed"
2A3K [27]	intermediate	bound	not present	alternative "open"
1FPR [28]	intermediate	bound	not present	alternative "open"
1Q6J [29]	closed	bound	not present	alternative "open"

2.3. PTP1B Catalytic Mechanism

Structural and kinetic studies have revealed substrate recognition and catalytic mechanisms of PTPs. PTP-mediated catalysis proceeds by a two-step mechanism shown in Figure 2.4 [1]. The catalytic domain of PTPs contains the characteristic active-site sequence known as the PTP signature motif, HCXXGXGR(S/T), where X represents variant amino acids. In the first step, Cys215 makes a nucleophilic attack by the sulfur atom of its thiolate side chain on the substrate phosphate. Conserved Arg221 is essential for substrate binding and transition-state stabilization. Asp181 is located on the WPD-loop, which closes on the pTyr side chain of the peptide, acts as a general acid in the first step of the reaction, which is the formation of the phosphoenzyme complex (cysteinyl-phosphate catalytic intermediate) by protonating the leaving group. In the second step, Asp181 acts as a general base by taking a proton from a catalytic water molecule, and the activated water molecule releases the inorganic phosphate. A conserved glutamine (Gln262), located on the Q loop, coordinates this water molecule and Asp181. The pTyr loop contains a tyrosine residue (conserved Tyr46 in PTP1B), which defines the depth of the cleft and confers to the absolute specificity that PTP1B displays for phosphotyrosine-containing substrates. The smaller phosphoserine and phosphothreonine residues would not reach down to the phosphate binding site. Phe182 on WPD loop stacks against the phenyl side chain of the substrate pTyr residue. This leads to stabilization of the closed conformation of the loop, and helps Asp181 to function. The Gln262Ala mutant form of PTP1B allows trapping and visualization of the catalytic intermediate due to impairment of hydrolysis. The WPD loop closes over the active site, sequestering the phosphocysteine intermediate with water molecule at the catalytic center and preventing the transfer of phosphate to other phosphoryl acceptors. This explains the requirement for hydrolysis in the second catalytic step [1].

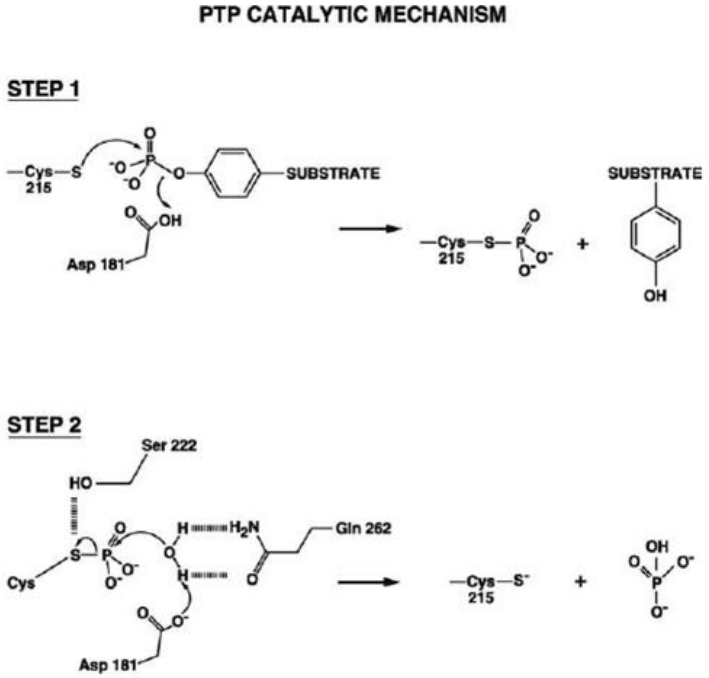

Figure 2.4. Schematic representation of the catalytic mechanism of PTP1B [1]

2.4. Interactions Stabilizing PTP1B-Ligand Complexes

In this section, PTP1B-ligand interactions will be examined by focusing on four different regions: catalytic loop, the secondary aryl-phosphate-binding site, the pTyr Recognition Loop, and the allosteric inhibition site.

2.4.1. Interactions in the Active Site

The catalytic site (residues 215- 221) corresponding to the conserved PTP signature motif, is located at the base of the pTyr binding pocket. Jia and co-workers determined the structures of a catalytically inactive Cys215Ser mutant form of the enzyme complexed with pTyr and the high affinity peptide DADEpYL-NH_2 to examine the structural basis of substrate recognition by PTP1B (PDB ID: 1PTU) [8]. Main chain amide groups of Ser216-Arg221 loop and the guanidium group of Arg221 form six hydrogen bonds and two salt bridges, respectively, with the three terminal phosphate oxygens of pTyr. The phenolic

oxygen atom of pTyr forms a network of hydrogen bonds with the side chain of Asp181 [8].

Fang and co-workers [17] investigated the interactions between five bidentate inhibitors and PTP1B, TCPTP, SHP-2 and compared these interactions by using MD simulations. The five bidentate inhibitors were extracted from five PTP1B complexes (PDB entries: 1NNY, 2B07, 2F6Z, 1Q1M, and 1PYN). They found that hydrogen bonds were formed between the residues in the active site and the carboxyl and carbonyl groups of the ligands. Hydrophobic interactions formed between the residues in the active site and the aromatic rings of the ligands. Some residues near the PTP1B active site such as Tyr46, Lys120, and Phe182 are also involved in hydrogen bonds and hydrophobic interactions [17].

Ala and his co-workers [21] determined the crystal structure of PTP1B in complex with an unsaturated IZD-phenol potent inhibitor to determine the structural features responsible for the high potency of IZD-compounds. In this structure, extensive interactions form between the inhibitor and the active site (Figure 2.5). The heterocycle binds at the center of the phosphate-binding loop (Cys215-Arg221), and the sulfone oxygens form hydrogen bonds with N-η of Arg221 and the backbone nitrogens of Ser216, Ala217, Gly218, Ile219, and Gly220. The nitrogen anion interacts with the helix dipole of α4 (Arg221–Arg238) and hydrogen bonded to N-ε and backbone NH of Arg221, and the carbonyl oxygen hydrogen bonded to the backbone nitrogen of Phe182 and side chain of Gln266. In addition, they observed a hydrogen bond between between the Asp181 side chain and the π electrons of the unsaturated IZD. They also found that the hydrophobic side of the heterocycle and phenyl ring of the inhibitor attached to PTP1B interact with Tyr46, Val49, Ala217, Ile219, Phe182. They concluded that the van der Waals interactions and the extensive network of hydrogen bonds explain the potency of IZD-phenol [21].

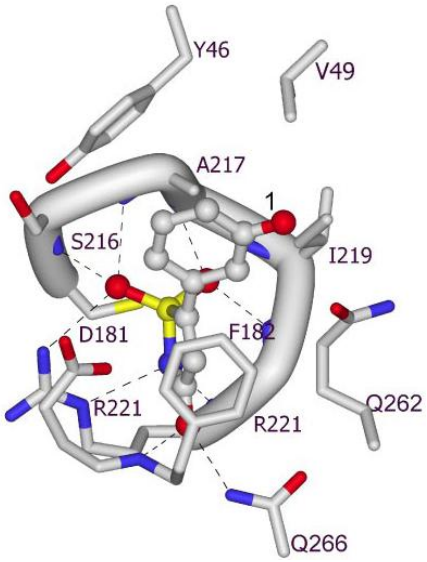

Figure 2.5. Unsaturated IZD-phenol 1 inhibitor (ball-and-stick) bound in the active site. Dashed lines indicate 10 potential hydrogen bonds between the inhibitor and the active site residues [21].

2.4.2. Interactions in the Second Aryl-Phosphate Binding Site

Crystallographic studies of the complex between *bis*-(*p*-phosphophenyl) methane (BPPM) and the Cys215Ala mutant resulted in the discovery of a second aryl-phosphate-binding site adjacent to the catalytic site (Figure 2.6) [19]. This second binding site is catalytically inactive and even though it forms weaker binding interactions compared with the primary catalytic site due to higher exposure to solvent, more potent inhibitors that bind to the active site as well as the secondary binding site can be designed [19]. The second binding site of PTP1B comprises Met258, Gly259, Leu260, Gln262 and Gln266 and Arg24, Arg254, which are the key arginine residues [17].

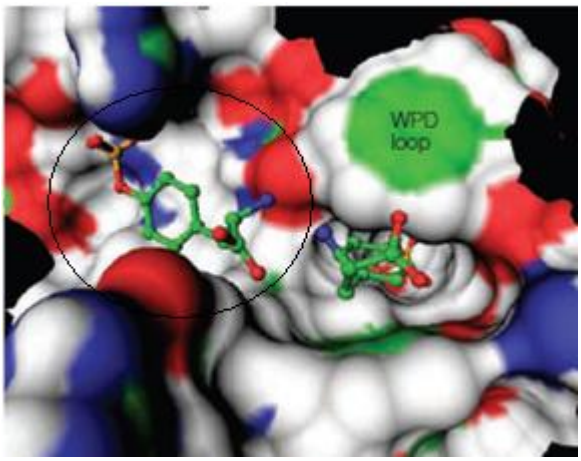

Figure 2.6. PTP1B binds phosphotyrosine in the active site and also in an adjacent site. A secondary aryl-phosphate-binding site is on the left (shown with circle), adjacent to the primary phosphate site (active site) on the right. The WPD loop is closed [19].

A comparison of PTP1B, TCPTP, SHP2 with respect to binding of five bidentate inhibitors to the secondary binding site showed that PTP1B and TCPTP are similar but PTP1B and SHP-2 are different [17]. The two essential arginine residues in site B (Arg24 and Arg254 in PTP1B, Arg22 and Arg248 in TCPTP) have similar interactions with NNY inhibitor in PTP1B and TCPTP. Gln262 in PTP1B forms a hydrogen bond with NNY inhibitor while the corresponding Gln256 in TCPTP does not form any hydrogen bond with the ligands. According to the MD simulations, MM/PBSA free energy calculations, and MM/GBSA free energy decomposition analyses, the obtained results are the following: (i) the NNY inhibitor which is not so selective, has strong interactions with the arginine residues in site B of the three PTP proteins whereas Q1M which is selective forms effective interactions with the arginine residues in site B of PTP1B. Hence, the second binding site is very important for designing inhibitors with good selectivity over low homology protein like SHP-2; (ii) most of the inhibitors except B07, the polarity of the part that binds to site B of the protein is positive to the potential of the inhibitor and negative to the selectivity of the inhibitor [17].

The inhibitor INTA can fully occupy both of the two aryl phosphate binding sites, with the thiophene moiety at the active site and the benzylsulfonamide moiety at the second binding site [30]. Wang and co-workers analyzed the interactions between the inhibitors INTA and INTB with PTP1B. The high potency and selectivity of INTA is due

to the hydrogen bond between one of the sulfonyl oxygens and the backbone amides of Leu260, and the water mediated hydrogen bonds between the other oxygens and Arg24 and Arg254 in the second binding site. On the other hand, inhibitor INTB does not occupy the second phosphate-binding site so it has fewer hydrogen bond interactions with PTP1B. Nevertheless, a strong van der Waals interaction between the methyl groups of INTB and Met258 of PTP1B may compensate for the loss of the hydrogen bond interactions of INTB in the second aryl phosphate-binding site [18].

2.4.3. Interactions in the pTyr Recognition Loop

Selectivity considerations in the design of phosphatase inhibitors may provide important clues for PTP protein sequence and structural similarity. PTP1B selectivity has been achieved by taking advantage of dissimilarities within the PTP family. Targeting interactions with Arg47 and Asp48 close to the active site has allowed selectivity over the other PTPs. Arg47 and Asp48 form a charged region at the top of the active site of PTP1B. This region is referred as the YRD (tyrosine, arginine, aspartic acid) motif (Figure 2.7). Targeting interactions between Arg47 and Asp48, which are close to the active site, allow selectivity to be achieved over most other PTPs [19]. However, in the KIM-containing (kinase interaction motif) PTPs like PTPN5, PTPN7, PTPRR, this sequence is YKT (Tyr-Lys-Thr) [26].

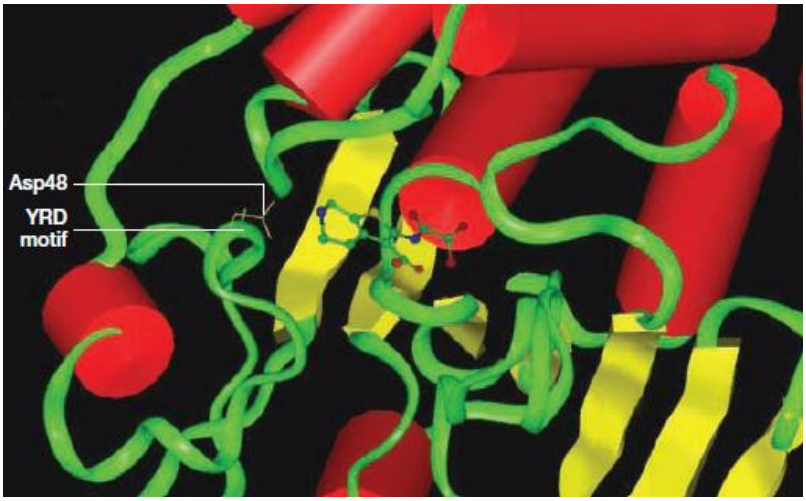

Figure 2.7. The YRD motif provides an opportunity for selective PTP inhibition.Formation of a salt bridge between compound 12 and aspartic acid Asp48 in PTP1B [19].

The interactions of the SNA inhibitor with PTP1B (PDB ID: 1PXH) beyond the active site were examined using MD simulations since the interactions between the inhibitor and the active site of PTP1B was extensively investigated previously and the interactions between the remainder of SNA and PTP1B beyond the active site was crucial in increasing inhibitory potency and selectivity [20]. The main chain nitrogen of Arg47 makes hydrogen bond with the carbonyl of the distal 4-phosphonodifluoromethylphenylacetyl group of the inhibitor during the simulation. A hydrogen bond forms between the nitrogen atoms of the guanidinium group of Arg47 and the Asp side chain of SNA and a polar interaction forms between a fluorine atom of the distal difluorophosphonate group and the guanidinium group of Arg47 during the simulation. In addition, the side chain of Asp48 forms a hydrogen bond with the main chain nitrogen of F2Pmp of SNA. Two or three hydrogen bonds between the amino group of Lys41 and the phosphonate oxygen atoms were observed. Based on the the number and distribution of hydrogen bonds for Lys41, Arg47, and Asp48, Arg47 was proposed to be the most important residue for inhibitor potency and selectivity [20].

The crystal structures of PTP1B in complex with isothiazolidinone (IZD)-containing peptides identified the key binding interactions outsite of the site A and gave important clues for the design of nonpeptidic inhibitors [21]. The inhibitor is bound in the active site and extends into the pTyr recognition site. The distal heterocycle is disordered but interacts with Ly41 the side chain and the aryl ring attached to it is in van der Waals contact with the hydrophobic portion of the Arg47 side chain of the inhibitor. The carboxylate of Asp48 forms hydrogen bonds with the carboxamide and α-amino groups were. In addition, the backbone nitrogen of Arg47 forms hydrogen bond with the oxygen of the most distal amide of the inhibitor [21].

2.4.4. Interactions in the Allosteric Site

The crystal structure of PTP1B-inhibitor complex (PDB: 1T48) revealed the location of an allosteric inhibitor binding site, about 20 Å distant from the active site [22] (Figure 2.8). The benzofuran core of the inhibitor is located in a hydrophobic pocket formed by the

side chains of Leu192, Phe196 and Phe280. The ketone has hydrogen bond with the side chain of Asn193 and the sulfonamide nitrogen atom has hydrogen bond with the carboxylate of Glu276. In addition, a water-mediated hydrogen bond exists between the phenol group and the main chain carbonyl of Phe196 (Figure 2.8b) [22]. Allosteric inhibitor stabilizes the WPD loop in the open conformation.

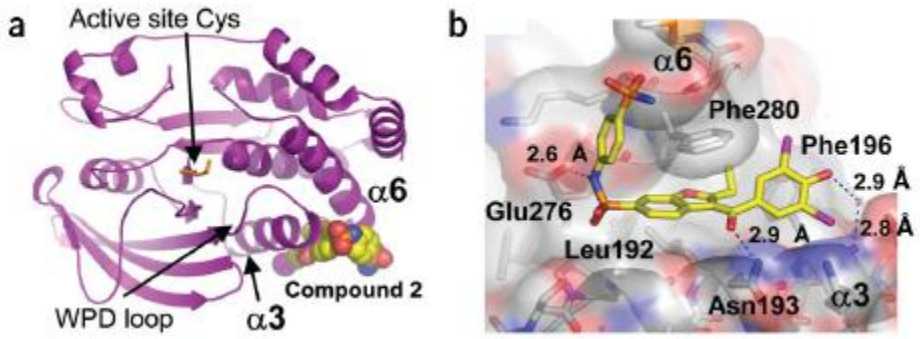

Figure 2.8. The allosteric site in PTP1B. (a) Inhibitor bound to the site formed by helices α3 and α6, distance from the active site Cys is 20Å. (b) .Detailed view of the allosteric inhibitor binding site. The hydrophobic pocket formed by Leu192, Phe196 and Phe280 and the hydrogen bonds with side chains of Glu276, Asn193 and the main chain carbonyl of Phe196 (water-mediated) are shown [22].

2.5. Importance of the WPD Loop and the Surrounding Regions on The Function of PTP1B

Residues 177-185 of PTP1B include a flexible loop, known as the WPD loop. In the unliganded form, WPD loop is generally found to be in the open conformation, whereas it closes over the active site upon substrate binding [8]. However, an open WPD-loop has been also observed in the structures with bound ligands in some cases [9]. In addition, the WPD loop is in the open state in complex structures with a group of non-pTyr-like allosteric inhibitors, which bind to a site located about 20 Å away from the catalytic site [22].

2.5.1. Effect of Ligand Binding on WPD loop Conformation

Examination of the crystallographic structures of the catalytically inactive PTP1B mutants has revealed the conformational activation mechanism of PTP1B. It has been observed that the WPD loop of PTP1B is flexible and may change between the open and closed conformation upon ligand binding (Figure 2.9). Generally, the WPD loop is observed in the open conformation when no ligand is bound in the active site. Upon ligand binding, the WPD loop (residues 176-187) moves up to 12 Å to close down on the phenyl ring of the substrate, maximizing hydrophobic interactions [19]. In substrate and inhibitor complexes, the WPD loop is usually in the closed conformation. On the other hand, Pedersen and co-workers demonstrated that the WPD loop of PTP1B can also assume the closed conformation in the apo form (PDB ID : 1SUG, [9]). In addition, Romsicki and his co-workers found that the structure of an apo PTP1B C215D mutant has WPD loop in a closed conformation (PDB ID: 1PA1, [31]). Comparison of WPD loop conformation of the PTPN5, PTPRR and PTPN7 structures with that of PTP1B (PDB codes 1SUG for the closed conformation and 2HNP for the open conformation) suggests that PTPN7 (PDB code 2A3K) and PTPRR (PDB code 2A8B) are in an alternative intermediate state, while PTPN5 (PDB code 2BIJ) is in an open conformation (Figure 2.10). Moreover, comparison of PTPRR structure with that of its mouse homologue, PTPSL (PDB ID: 1JLN), indicates that both WPD loop conformations are in an intermediate state, with PTPSL more closely resembling the conformation in PTP1B [26]. These findings support that the WPD loops of PTPs are relatively flexible and can fluctuate between the open and closed conformation in the absence of a ligand.

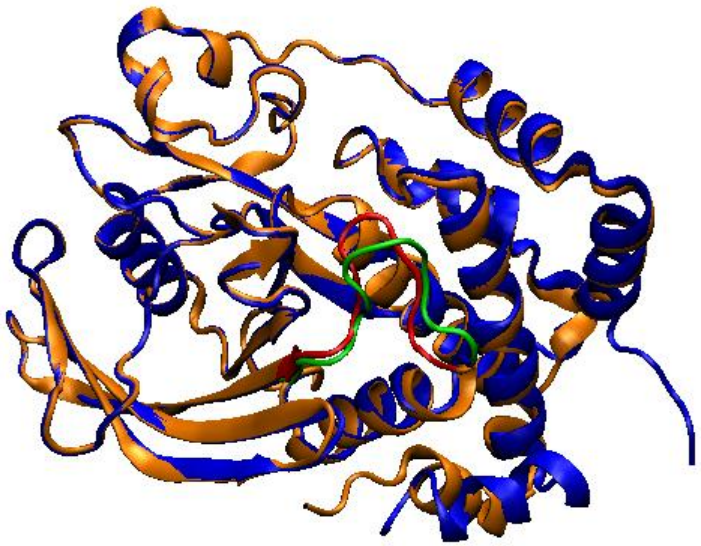

Figure 2.1. Cartoon representation of the WPD_{open} (2F6F in blue, WPD loop in green) and WPD_{closed} (1SUG in orange, WPD loop in red) structures of PTP1B.

Figure 2.2. Structural overlay of WPD loops in the closed conformation (PDB code 1SUG; purple); in the open conformation PTPN7 (PDB ID: 2A3K; black); PTPRR (green); PTPSL (PDB ID: 1JLN; red); (PDB ID: 2HNP; blue) and PTPN5 (PDB ID: 2BIJ; yellow) [26].

Structural comparison of the available PTP structures show that the WPD loop assumes four main conformations: a closed state, an intermediate state, an open state, and an atypically open state present in STEP, LYP, and GLEPP1 (Figure 2.11). The atypically open state is associated in STEP (PDB ID: 2BIJ) with a stabilizing α3 helix C-terminal to

the WPD loop and in LYP (PDB ID: 2P6X) and GLEPP1 (PDB ID: 2GJT) an extra turn of helix α3 following the WPD loop [7].

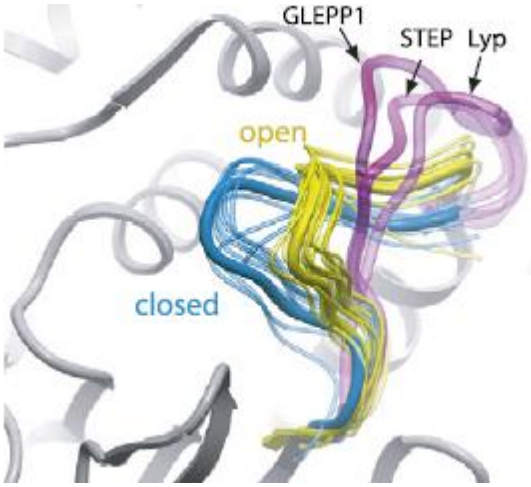

Figure 2.3. WPD loop conformations are shown by a PTP representative of each state:closed (blue, PTP1B, PDB: 1SUG); open (yellow, PTP1B, PDB: 2HNP); and atypical (magenta, GLEPP1, PDB: 2GJT; STEP, PDB: 2BIJ; Lyp, PDB: 2P6X). The intermediate WPD loop conformation of PCPTP1 (PDB: 2A8B) is not shown for clarity. Other PTP structures are shown with a thin transparent line tracing the backbone and are colored according to conformation [7].

Schubert and co-workers examined the structural and functional details of the conformational change in the WPD loop based on a comparison of the structure of inactive Yersinia PTP Y403S mutant in the apo form and in complex with sulfate. They found that the WPD loop (residues 351-359 in Yersinia PTP, corresponding to residues 177-184 in PTP1B) main-chain atoms in the sulfate-bound complex moved as much as 7 Å away from the active site, in contrast to the unliganded structure. The average B factor of WPD loop in the complex structure is 15 $Å^2$ whereas it is 18 $Å^2$ in the apo form, showing that the WPD loop in the unliganded structure is more flexible than in the complex structure. Upon sulfate binding, Arg 409 (corresponding to Arg221 in PTP1B) guanidinium group rotates 90° around its χ3 torsion angle to form a salt bridge with two of the sulfate oxygens, and it forms a hydrogen bond with Glu290 (Glu115 in PTP1B). The conformational change of the WPD loop results in a new hydrogen bond between Arg 409 Nηl and the carbonyl of Trp 354 (corresponding to Trp179 in PTP1B) (Figure 2.12A). The conformational change of WPD loop also moves the Asp356 (Asp181 in PTP1B) carboxylate Oδ1 atom to within

3.6 Å from the sulfate oxygen O4 (Figure 2.12B). Residues 355-358 (corresponding to residues 180-183 in PTP1B) forms a type I β-turn at the center of the WPD loop in the unliganded structure (Figure 2.13A) while the peptide bond between Asp 356 and Gln 357 (β-turn residues *i* + 1 and *i* + 2) flips into a type II β -turn configuration in the sulfate bound structure (Figure 2.13B). They concluded that this β-turn conformation could bring Asp356 side chain close to the active site and orient the Gln 357 (Phe182 in PTP1B) backbone amide to hydrogen bond to a water molecule found in the sulfate-bond structures [15].

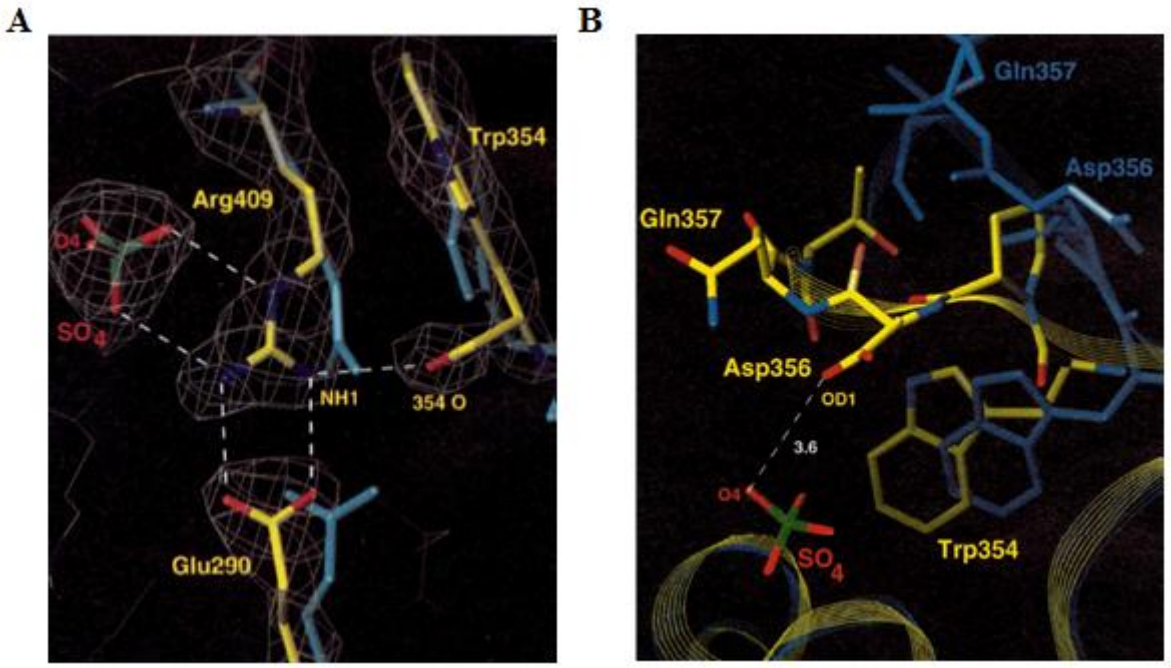

Figure 2.4. The conformational changes in Yersinia PTP upon sulfate binding (A) The active site Arg409 reorients in response to sulfate binding. (B) Conformational change in the WPD loop moves Asp 356 into the active site. The unbound structure (residues 354-358) is shown in blue and the sulfate-bound structure is shown in yellow with red oxygens and blue nitrogens [15].

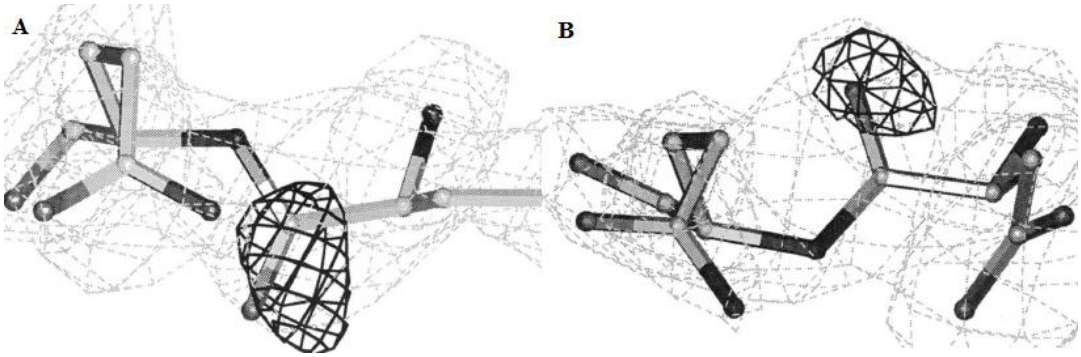

Figure 2.5. The peptide bond between residues 356 and 357 changes orientations between unliganded and sulfate-bound structures. Residues Asp 356 and Gln 357 in (A) the unliganded and (B) the sulfate-bound crystal form [15].

Li and co-workers cocrystallized a trapping mutant of PTP1B with phosphorylated insulin receptor kinase (IRK) domain (PDB ID: 2B4S) in order to examine the molecular determinants of PTP1B specificity for the insulin receptor. In this structure, the WPD loop containing the PTP-conserved catalytic residue Asp181 forms interactions with the residues in and preceding αJ of IRK (Figure 2.14). Hydrophobic interactions between Pro180 (WPD loop) with Val1274 (αJ of IRK) and Phe182 (WPD loop) with Pro1269 (αI-αJ loop of IRK) were observed. In addition, the backbone of Phe182 in the WPD loop forms a hydrogen bond with H1268 in the αI-αJ loop of IRK. A closed conformation of WPD loop was maintained by these interactions between αJ of IRK and the WPD loop [32].

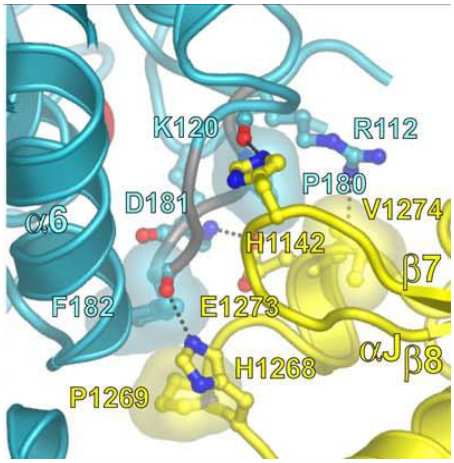

Figure 2.6. The interactions between αJ of IRK and the WPD loop of PTP1B (PDB ID: 2B4S). Carbon atoms are colored cyan (PTP1B) or yellow (IRK), oxygen atoms are colored red, and nitrogen atoms are colored blue. Hydrogen bonds are represented by dashed lines [32].

2.5.2. Effect of Allosteric Inhibition on the WPD loop

Crystal structure of the allosteric inhibitor bound PTP1B complex (PDB : 1T48) showed the location of the allosteric site, about 20 Å from the active site [22]. The PTP1B-allosteric inhibitor complex has the WPD loop in the open conformation (Figure 2.15A). The backbone atoms of the WPD loop and helices α3 and α6 were shifted relative to their positions in the closed conformation (Figure 2.15B,C), and the side chains of Phe192, Phe280 and Glu276 moves to accommodate the allosteric inhibitors (Figure 2.15D). It was also observed that the allosteric-bound conformation and the closed conformation were not only dissimilar but also mutually exclusive. They concluded that allosteric inhibitors blocked the atomic interactions and side chain movements required for WPD loop to close [22].

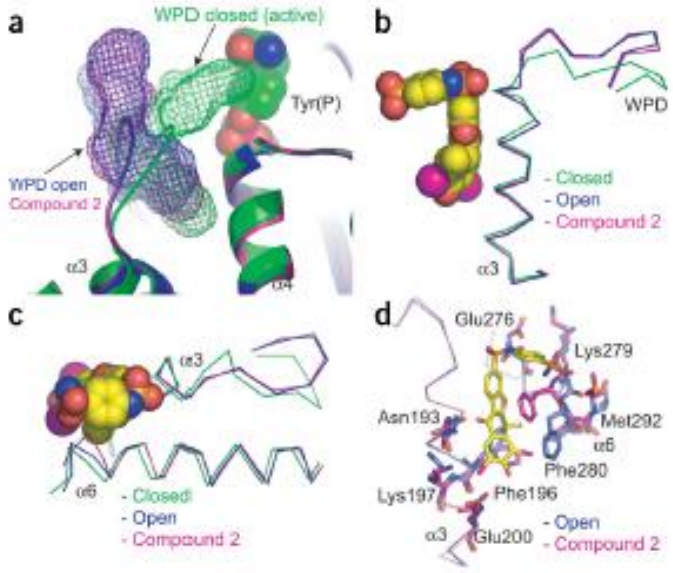

Figure 2.7. Conformations of some important regions in closed conformation of PTP1B bound to pTyr in green (1PTY), open conformation in blue (2HNP) and PTP1B bound to the allosteric inhibitor in magenta (1T48) (a) Overview of the active site. (b) Same overlay as in a with view of α3 main chain seen from interior of PTP1B. The allosteric inhibitor is shown as spheres. (c) Same overlay as in a with view of α3 and α6 main chains seen from the protein surface. (d) Overlay of the allosteric inhibitor-bound (inhibitor, magenta) and open conformations (PDB entry 2HNP, blue). Several side chains (sticks) must move to accommodate the allosteric inhibitor [22].

The allosteric inhibitor occupies the site of Trp291 and prevents the interaction between α7 and α3-α6, which might be a key feature in controlling the mobility of the WPD loop (Figure 2.16a-d). The direct interaction of the inhibitor with Pro188 and Leu192, along with side chain movements in helices α3 and α6, might stabilize the enzyme in an open, inactive conformation (Figure 2.16d). Phe191 occupies the site that Trp179 must enter in order for the WPD loop to close over the Tyr(P) pocket (Figure 2.15d). The α7 helix is disordered in the presence of an allosteric inhibitor (Figure 2.16c), suggesting that allosteric inhibitors might act by blocking the interaction between α7 and α3-α6 [22].

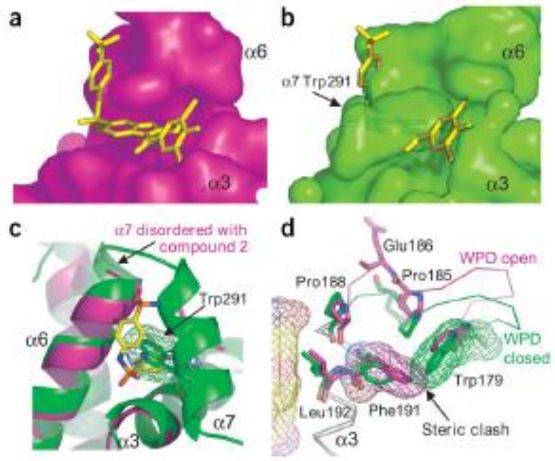

Figure 2.8. View of the important regions in presence of compound 2 and in the closed conformation (a) View of allosteric site in presence of compound 2. The surfaces of helices α3 and α6 are shown. (b) View of the corresponding region in the closed conformation. Trp291 from helix α7 is shown. (c) Overlay of PTP1B in closed conformation (green, 1PTY) and bound to compound 2 (magenta and yellow). (d) Same superposition as in c with view of α3 and the WPD loop from the interior of PTP1B [22].

Kamerlin and co-workers [10] used MD simulations to understand the effect of the allosteric inhibitors on WPD loop mobility. They developed force field parameters for one of Wiesmann's inhibitors and used them to carry out MD and targeted molecular dynamics (TMD) simulations. They investigated three structures, the apo-enzyme (PDB ID: 2HNP), PTP1B complexed with the allosteric inhibitor BB3 (PDB ID:1T48) and PTP1B complexed with pTyr (PDB ID:1PTU) by using MD simulations. The WPD loop in the apo PTP1B is more flexible than in liganded states [10]. This finding is in agreement with previous comparisons of crystallographic B factors [15]. Although the allosteric site is over 10 Å away from the WPD loop, mobility of the WPD loop is reduced, suggesting that the reduction in mobility is related to the allosteric inhibitor hindering WPD loop closure. The S loop in the allosteric inhibited form is also more rigid than in the apo-form and pTyr-bound states, highlighting the close coupling between S loop and WPD loop motions. The α3 helix mobility is also reduced in the presence of the allosteric inhibitor. The reduced mobility of this helix might reduce the mobilities of WPD loop and S loop, although the allosteric inhibitor does not have direct interaction with these two loops [10].

2.5.3. The Effect of α7 Helix on WPD loop Stability

Crystal structures of PTP1B with WPD loop in the closed conformation were obtained in the presence of the pTyr containing hexapeptide, and these structures showed interactions between α7, α3 and α6 [8]. Allosteric inhibitors prevent this interaction and closing motion of the WPD loop. This reveals a relationship between α7 and the WPD loop. The α7 helix is proposed to be a regulatory helix stabilizing the closure of the WPD loop and is needed for full enzyme activity. Indeed, truncation of α7 leads to a decrease in catalytic activity, suggesting the important role in full enzyme activity [22]. Moreover, conformational flexibility in α7 and along helices α3 and α6 might be important for accommodating and stabilizing the various motions relative to the mobility of the WPD loop [22].

Montalibet and co-workers [24] identified several PTP1B residues that confer inhibitor resistance to PTP1B when they are mutated (PDB ID: 2F6F). The inhibitor resistant mutations are clustered around helices α3, α6 and α7. The S295F mutation on α7 helix was examined in detail and the crystal structure was resolved. On the other hand, in the wild-type (PDB ID: 1SUG), the Ser295 side chain occupies a pocket formed by helices α3, α7 and loop β9-β10 (L11) and forms a hydrogen bond with Asn193 side chain on α3 helix. Asn193 also hydrogen bonds to the side chains of Phe152 on L11 and Glu297 on α7 helix. Phe152 interacts with the backbone carbonyl of Ser190 on α3 helix. The side chain of Glu297 on α7 helix has a hydrogen bond with the side chain of Phe153 on L11, while the main chain carbonyl interacts with the main chain nitrogen of Ser151 on L11. The α7 helix interacts with α3 helix via another hydrogen bond between Gln288 and Glu200. Moreover, Trp291 is buried in the hydrophobic pocket between α3 and α6. The crystal structure of the S295F mutant PTP1B (PDB ID: 2F6F, [24]) showed that replacement of Ser (Ser295) with Phe resulted in the repositioning and partial disordering of helix α7 and blocking this extensive network of interactions spanning the α7, α3 and L11 (Figure 2.17). The simplified scheme of this extensive network is shown in Figure 2.18. Only the hydrogen bond between the side chains of Tyr152 and Asn193 was maintained. It was suggested that the local variations resulting from the disruption in α7 helix may give rise to a transduction of changes to helix α6 through helix α3 directly connected to the WPD loop [24].

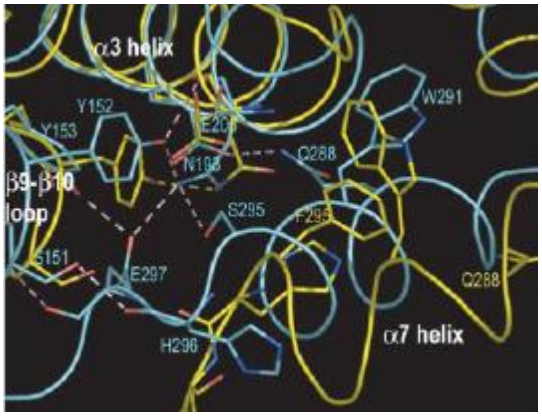

Figure 2.9. Interactions of α7 helix with the PTP1B core. 1SUG (wild-type structure) is shown in blue and the S295F mutant is shown in yellow [24].

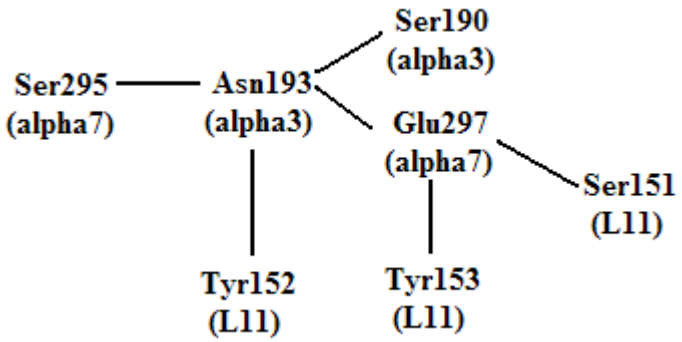

Figure 2.10. The extensive network of interactions formed by α3, α6 and α7 helices in the 2F6F crystal structure of PTP1B.

2.5.4. Conformation of R loop in the PTP Family

The PTP1B region comprising residues Val113- Ser118 is known as the "R loop" because of a conserved arginine residue (Arg112 in PTP1B). In PTP1B structures, the R loop points toward the active site, irrespective of the WPD loop conformation [9]. The R loop has been identified as an important structural feature of PTP1B by MD simulations performed on the PTP1B complexed with pTyr (PDB ID: 1PTV) [23]. The displacement of the R loop is smaller compared to that of WPD loop and S loop in the MD simulations. The R loop moved slightly toward the active site upon WPD loop closure in TMD simulations, and it was suggested that this displacement resulted in a tighter binding pocket for ligand binding [23].

MD simulations on PTP1B in the apo form showed the concurrent movement of the WPD loop and the R loop and suggesting that both WPD and R loops should change their conformations upon substrate binding. Either the Asp181 side chain has to move between Glu115 and Lys116 (located on R loop), or residues 116-120 have to move to allow WPD loop to close. Glu115 forms a salt bridge with Arg221 and the conformation of these residues changes depending on the conformation of the WPD loop. The movement of Arg221 is also proposed to be coupled to the movement of the WPD loop. Hence, the interaction between Glu115 and Arg221 may change upon WPD loop conformational transition, allowing the WPD loop to close without interfering Lys116. Following WPD loop closure, the R loop could move to its original location and the interaction between Glu115 and Arg221 could form again [9]. The R loop of Yernisia PTP (PDB ID: 1YTS) structure [15] was found to be in a "open" conformation similar to that in the TC-PTP structure [25]. Based on this point, Pedersen and co-workers speculated that the efficiency of the Yernisia PTP might be a result of the open R loop conformation which does not prevent free WPD loop motion [9].

The TC-PTP structure (PDB ID : 1L8K) has Lys122 (Lys120 in PTP1B) in an open conformation (R loop), pointing away from the active site, contrary to other PTPs that have the corresponding lysine residues pointing toward the active site [25]. In addition, the salt bridge between Glu115 and Arg221 (in PTP1B), seen in all other PTPs, is not present in the TC-PTP structure. Lys122 in R loop (Lys120 in PTP1B) is forced (by crystal packing

interactions) into a conformation pointing away the active site pocket. The salt bridge, which is normally formed between Glu117 (Glu115 in PTP1B) and Arg222 (Arg221 in PTP1B) in all other PTPs, anchors the R loop and defines the architecture and function of the P loop. This salt bridge is not seen in TCPTP structure (PDB ID: 1LK8), resulting in Glu117 pointing into the solvent. However, a small shift in the active site and a limited rotation of the guanidinium group of Arg222 are observed in TC-PTP structure when the active site of TC-PTP and PTP1B structures are compared. Hence, although the movement of R loop and Glu117 is a result of crystal packing and does not reflect the normal structure of TCPTP, it may be speculated that (i) the importance of the salt bridge between Glu117 and Arg222 in stabilizing the P loop is overestimated (can be compensated for by ligand binding) and (ii) such dramatic movement of R loop can be induced by high affinity, low molecular weight inhibitors [25].

The crystal structures of PTP1B in complex with different inhibitors (PDB ID: 1Q6J) showed that the residues Met114-Leu119 assume different conformations [29]. In general, the R loop is in a closed conformation shielding the active site from bulk solvent. The R loop is in a closed conformation in one of these complexes and formed hydrophobic interactions with the difluorobenzene. In the other complexes reported in this study, R loop is in the closed state but it might oscillate between different positions due to the extremely weak electron density, and the very high temperature factors associated with its atoms. Based on this information, it was suggested that R loop conformation may vary upon the exact chemical nature of the substituent present in this pocket [29].

In addition to these structures, R loop is disordered in some PTP structures, such as the SHP-1 complexed with peptide substrate (PTP name: PTP1C, PDB ID : 1FPR, [28]), the crystal structure of the catalytic D1 domain of human receptor-type protein tyrosine phosphatase (RPTPκ) (PTPN7, PDB ID : 2A3K, [26]) and the structure of catalytic domain of human KIM (kinase-interaction motif)-containing PTPRρ (PTPRR, PDB ID : 2A8B, [26]). The WPD loop conformations are different on these PTP structures and listed in Table 2.2.

2.6. Computational Studies on PTP1B

2.6.1. Molecular Dynamics Simulations on PTP1B

Molecular dynamics programs simulate the behavior of biomolecular systems, leading to understanding of their functions [33]. MD simulations are usually performed on PTP1B to investigate the changes in the structure and dynamics in PTP1B upon ligand binding, the importance of protein and substrate flexibility for binding in the PTP1B-ligand complexes, the physical forces involved in enzyme-ligand interactions and to reveal the possible mechanisms of ligand recognition and inhibition [23,10,34,35,20,36,18]. In these studies, the topology files and force field parameters for the inhibitor were generated by the program PRODRG.

Peters and co-workers [34] investigated the concerted motions in the apo PTP1B (PDB ID: 2HNQ) and studied the effect of substrate binding in the PTP1B-peptide complex (PDB ID: 1PTU) by using MD simulations. They applied the essential dynamics analysis technique to analyze MD simulation trajectories and to identify collective motions in PTP1B. Conformational changes in several regions of the protein are found to be associated with the displacement of WPD loop. Fluctuations in several regions and the flexibility of the enzyme are reduced upon substrate binding [34]. In a following study, the same group performed MD simulation of both PTP1B in complex with the peptide-based substrate DADEpYL (PDB ID: 1PTU) and the free substrate to investigate (i) the influence of substrate flexibility on the mobility of binding pocket, (ii) the interactions between the substrate and residues in the binding pocket, and (iii) the electrostatic properties of PTP1B [35]. They also performed essential dynamics analysis of the PTP1B-substrate complex to examine the effect of a peptide on PTP1B dynamics. The N- and C-terminus of the substrate show high flexibility and the P loop (His214-Gly219) is rigid while weak fluctuations of the WPD loop are observed along several eigenvectors. On the comparison of the interaction energies of the complexed structure and free substrate, van der Waals interaction and electrostatic energies between the substrate and its surrounding for the substrate-enzyme complex are lower than for the free substrate [35].

Kamerlin and co-workers [23] also performed MD simulations of both unbound PTP1B (PDB ID: 2HNP) and PTP1B complexed with a pTyr containing hexapeptide (PDB ID: 1PTV) to examine the changes in the structure and dynamics in PTP1B upon ligand binding. The largest differences between the unbound PTP1B an PTP1B complex are in the residues 113-142, 198-209, 280-298 and, as expected, the WPD loop from the mean RMSD fluctuation calculations [23]. Subsequently, they performed MD simulations of the unbound PTP1B (PDB ID: 2HNP), PTP1B complexed with allosteric inhibitor BB3 (PDB ID: 1T48) and PTP1B complexed with pTyr (PDB ID: 1PTV) to examine the effects of the allosteric inhibitor on WPD loop mobility [10]. The WPD loop in the apo form (2HNP) is much more flexible than in two other structures. Not surprisingly, the loop is more mobile in the unbound PTP1B compared with the bound form (1PTV) but the loop mobility reduces in the allosteric inhibitor bound form (1T48) although the ligand is not interacting with the loop [10]. Bharatham and co-workers developed this study by performing MD simulations to investigate the molecular level interactions between the allosteric inhibitor and the PTP1B allosteric site, especially α7 [36]. The simulation results show that the benzofuran core of the inhibitor is in a hydrophobic pocket by Trp291, Phe280, Phe196 and Leu192, suggesting that the inhibitor is stabilized in the α7-α3-α6 groove. The Ser295 (on α7 helix) initially forms extensive H-bonds with α3 helix but it gradually loses its interactions with α3 and forms with the allosteric inhibitor. This shows that the allosteric inhibitor prevents the α3-α6-α7 interaction pattern for the WPD loop to close [36].

The MD simulations of the uncomplexed PTP1B (PDB ID: 2HNP) and the PTP1B complexed with a bidentate inhibitor SNA (PDB ID: 1PXH) were performed to analyze the protein-ligand interactions and display the possible mechanisms of ligand recognition and inhibition [20]. The RMSD from the starting structure is calculated. The RMSD for the uncomplexed PTP1B is larger than that for the complexed structure, suggesting that th flexibility of PTP1B decreases upon ligand binding. From the principal component analysis (PCA), the significant fluctuations are observed on the residues 113-120, 128-132, 162-168, 178-185, 234-242, the N- and C-terminal regions for uncomplexed PTP1B. Similar regions have high fluctuations for complexed PTP1B but the WPD loop decrease to allow WPD loop to move toward the inhibitor and interact with it [20].

Wang and co-workers [18] performed MD simulations of the PTP1B complexed with INTA and INTB (PDB IDs: 2QBP and 2QBQ, respectively) to analyze the enzyme-inhibitor interactions for revealing the powerful inhibition and high selective recognition mechanism. The RMSD from the initial structure is an important criterion used to measure the convergence of the protein system concerned. The RMSD for the uncomplexed PTP1B is a little higher than the complexed PTP1B-INTA and PTP1B-INTB showing the flexibility of PTP1B reduces upon inihibitor binding. By monitoring the distances of key atoms during MD simulations, an extensive hydrogen bond network is found between the oxygen atoms of the thiophene moiety of inhibitors and the residues in the active site. From the data of the average structures in MD simulations, it is found that INTA inhibitor makes interactions with Leu260, Arg24 and Arg254, resulting in high potency and selectivity of INTA [18].

2.6.2. Background of Targeted Molecular Dynamics (TMD) Simulations and TMD on PTP1B

MD simulation is a valuable tool to investigate the dynamic behavior of stable macromolecules at finite temperatures. Nevertheless, in standard MD simulations, considerable conformational transitions occur only accidentally or at exceptionally high temperatures. TMD is a method to induce a conformational change to a known target structure at ordinary temperature by applying a time-dependent, purely geometrical constraint. The transition is enforced independently of the height of energy barriers, while the dynamics of the molecule is only minimally influenced by the constraint [37]. The essential point of TMD method is to use simulations on the nanosecond time scale to obtain information about much slower events [38].

The TMD algoritm was developed [37] and then incorporated into CHARMM [39]. In the initial algoritm, a holonomic constraint was used to linearly decrease the RMSD with respect to a target structure. Schlitter and co-workers found that the transitions of insulin comprised unfolding of an α-helical portion and refolding from an extended conformation in the reverse direction [37]. Ma and Karplus performed TMD simulations of GTP-bound Ras (PDB ID: 5P21) and GDP-bound Ras (PDB ID: 1Q21) to determine possible paths for the conformational transitions in the switch I (residues 30-38) and switch

II region (residues 60-76) in going from GTP-bound to GDP-bound form [39]. In the current study, a targeting potential was applied to a specific set of atoms to decrease the RMSD.

Kamerlin and co-workers [23] performed TMD simulations of both unbound PTP1B (PDB ID: 2HNP) and PTP1B complexed with a pTyr containing hexapeptide (PDB ID: 1PTU) to obtain a more detailed examination of WPD loop movement. The TMD method makes it easy to examine full loop opening and closure. TMD simulations shows that (i) the distance between the WPD loop and the active site is dramatically reduced; (ii) the R loop moves slightly towards the active site upon WPD loop closure to create a tighter binding pocket once ligand is bound; (iii) the S loop closest to the active site (residues 204-209) changes its conformation simultaneously with the WPD loop movement and (iv) the extremities of the α7 helix move away from the S loop and the WPD loop while its central part moves closer to both loops upon WPD loop closure [23]. In their following study, they performed a set of TMD simulations of the conformational transition from the open structure complexed with allosteric inhibitor BB3 (PDB ID: 1T48) to the fully closed structure PTP1B complexed with pTyr (PDB ID: 1PTV) [10]. From the RMSD evolution during the TMD simulations, as expected, the largest RMSD changes are seen in WPD loop and S loop and also a significant change is observed in α3 helix. This helix is is slightly displaced to allow WPD loop closure. TMD simulations show that the WPD loop closure is assisted by a hydrogen bond formation between Trp179 and Arg221 [10].

Recently, TMD simulations have been performed to investigate the conformational transitions of some important regions in proteins from one state to another [40],[41],[42],[38]. Cheng and co-workers performed TMD simulations of human α7 nicotinic acetylcholine receptor (nAChR) and TMD forces were applied to the backbone atoms of the C-loop residues (Arg186 to Glu193) [40]. Xu and co-workers performed TMD of Lactose repressor protein (LacI) to investigate the transition between inducer- and operator-bound states [42]. Zou and co-workers investigated the detailed conformational dynamics of juxtamembrane region (JMR) and activation loop (A loop) regulating the stem cell factor receptor (c-Kit) kinase activity by peforming TMD simulations [41]. TMD was also performed on the receiver domain of nitrogen regulatory protein C (NtrC) to

determine a possible pathway of interconversion between the active and the inactive state [38].

3. METHODS

3.1. Molecular Dynamics Simulations

Computer simulations act as a bridge between microscopic length and time scales and the macroscopic world of the laboratory. Simulation acts as bridge in another sense: between theory and experiment [43]. One of the principal tools in the theoretical study of biological molecules is the method of molecular dynamics (MD) simulations. This computational method calculates the time dependent behavior of a molecular system. Details on the fluctuations and conformational changes of proteins and nucleic acids have been examined by MD simulations. These methods are used to investigate the structure, dynamics and thermodynamics of biological molecules and their complexes. Researchers believe that MD simulations are very important in translating structure information into mechanisms highlighting biomolecular functions. Successful simulations of biomolecular systems can give rise to a good understanding of basic biological processes and they can provide new insight to drug design [33].

In MD Simulations, motions of atoms are simulated as a function of time according to Newton's equation of motion given by,

$$F_i = m_i a_i \tag{3.1}$$

The force can be also expressed as the gradient of the potential energy,

$$F_i = -\nabla_i V \tag{3.2}$$

Hence, the following expression is obtained by combining two equations,

$$-\frac{dV}{dr_i} = m_i \frac{d^2 r_i}{dt^2} \tag{3.3}$$

where *V* is the potential energy of the system. Newton's equation of motion can then relate the derivative of the potential energy to the changes in position as a function of time.

MD simulation of a protein includes three stages: minimization, equilibration and data collection stages. Energy minimization plays an essential role in starting dynamics simulations. The purpose is to bring the molecule to a conformation where its potential is at the minimum. The aim of equilibration is to bring the system to a favorable conformation at the target temperature and pressure. In the data collection period, MD trajectory is recorded for further analysis.

3.1.1. CHARMM Force Field

Theoretical chemistry is associated with the application of mathematical equations describing the relationship of chemical structure to energy. The combination of this information with statistical mechanics allows for the properties of the system to be calculated. However, energies of all possible conformations of a chemical system cannot be calculated. Therefore, simple mathematical functions are used to treat the structure-energy relationship. These methods are valid for the study of biological molecules, in which the molecular weight of the molecules is large (> 10kDal) and the aqueous environment must be included. This approach is named as molecular mechanics or empirical force field calculations. Different parameters should also be included in the mathematical equation since the equation alone is not enough for computation of structure-energy relationships. The empirical force field used in this study is CHARMM27. The intramolecular potential energy function and the intermolecular potential energy function in CHARMM are shown in Equations 3.4 and 3.5.

$$\sum_{bonds} K_b(b-b_o)^2 + \sum_{angles} K_\theta(\theta-\theta_o)^2 + \sum_{torsions} K\phi\ 1+\cos\ n\phi-\delta\ + \sum_{impropers} K_\varphi(\varphi-\varphi_o)^2 + \sum_{Urey-Bradley} K_{UB}(r_{1,3}-r_{1,3,o})^2 \quad (3.4)$$

$$\sum_{electrostatics} \frac{q_i q_j}{r_{ij}} + \sum_{VDW}\left[\left(\frac{R_{\min,ij}}{r_{ij}}\right)^{12} - 2\left(\frac{R_{\min,ij}}{r_{ij}}\right)^{6}\right] \quad (3.5)$$

In equation 3.4, the first term represents the covalent bond strecthing interaction between two atoms. Here, b_o is the minimum energy bond length, so $b-b_o$ indicates the distance that the atom moves from the equilibrium, and K_b is the bond force constant. The second term represents the bond angles: K_o is the angle force constant and $\theta-\theta_o$ is the deviation of the angle between three bonded atoms from their equilibrium value. The third term represents the dihedral angle interactions: K_φ is the dihedral force constant, n is the multiplicity of the function, φ is the dihedral angle and φ is the phase shift. The fourth term represents the impropers, where k_w is the force constant and $w-w_o$ means the out of plane angle. The fifth term is the Urey-Bradley component, in which k_u is the respective force constant and U is the distance between 1,3 atoms in the harmonic potential. The nonbonded interactions between pairs of atoms (i,j) are represented by Equation 3.5. A standard 12-6 Lennard-Jones potential and electrostatic energy with Coulombic potential are used to calculate the van der Waals (VDW) energy. In the Lennard-Jones potential in Equation 3.5, $R_{\min,ij}$ represents the value where the Lennard-Jones potential crosses the x-axis [33].

3.1.2. NAMD

NAMD2 program is a parallel program developed for the purpose of utilizing large parallel machines in a scalable manner [33]. The non-bonded force computations require calculation of pairwise interactions between atoms. A cutoff distance, RC, is used in most methods. NAMD2 uses a unique decomposition strategy that combines the advantages of spatial decomposition and force decomposition, which permits the program to utilize a large number of processors.

NAMD2 uses an adaptive measurement-based periodic load balancing strategy for achieving good performance even for difficult MD simulations (e.g., those involving non-periodic configurations with high variations in atom density). NAMD2 allows force field to be divided into three parts upon their variation frequency. All bonded forces such as bonds, angles, dihedrals, impropers are assumed to be quickly varying; while non-bonded forces including electrostatics, Lennard Jones within a cutoff are considered to change slower and long-range electrostatics are considered to be the slowest.

The three overall goals of NAMD2 are sequential performance, scalability with respect to both simulation size and number of processors and modifiability. NAMD2 uses CHARMM force fields, X-PLOR coordinate and molecular structure files. NAMD2 can use periodic boundary conditions over any combination of the three coordinate axes as well as the non-periodic simulations. It performs cutoff simulations or full-electrostatic simulations employing multiple time stepping. NAMD2 can be connected to VMD to monitor the interactions with ongoing simulations [33].

3.1.3. MD Simulation Structures and Parameters

The atomic coordinates for the WPD_{open} conformation (2F6F) [24] and the WPD_{closed} conformation (1SUG) [9] were obtained from the protein data bank [44]. All simulations were carried out using the NAMD program and the CHARMM27 force field. In the 2F6F crystal structure, Ser295 is mutated to Phe, which exists in wild-PTP1B, using VMD program. Missing residues and hydrogens were built using the psfgen module of VMD.

The structures were surrounded by a periodic box of 86.3×68.9×60.8 Å and approximately 11500 TIP3P water molecules. The SHAKE algorithm was applied to keep all bonds involving hydrogen bonds fixed. Long range electrostatic interactions were computed by the particle-mesh-Ewald (PME) method and the non-bonded cutoff was taken to be 14 Å. The energy of the whole system was minimized for 2000 steps. The simulations were carried out for 40 ns at 300 K and 1 atm with an integration time step of 1 fs.

3.2. Targeted Molecular Dynamics Simulations

An initial and desired final configuration of a system is used in the targeted molecular dynamics (TMD) method to establish a possible pathway between the given configurations. The purpose for this simulation is to drive the evolution of the simulation toward the given target conformation. TMD has been used in the prediction of pathways between particular protein conformations and in protein folding [45]. TMD constrains an initial molecular structure to move progressively toward a targeted final structure. The constraint is controlled by the root-mean-square deviation (RMSD) of the current structure relative to the final structure for TMD [46].

In TMD, subset of atoms in the simulation is guided towards a final 'target' structure by means of steering forces. At each time step, the RMSD between the current coordinates and the target structure is computed. The force on each atom is given by the gradient of the potential;

$$U_{TMD} = \frac{1}{2}\frac{k}{N}[RMS(t) - RMS^{*}(t)]^{2} \tag{3.6}$$

where RMS(t) is the instantaneous best-fit RMSD of the current coordinates from the target coordinates, and RMS*(t) evolves linearly from the initial RMSD at the first TMD step to the final RMSD at the last TMD step. The spring constant is scaled down by the number N of targeted atoms.

3.2.1. TMD Simulation Structures

The initial and target structures for TMD were representative equilibrium structures from the MD simulations. The initial structure was the representative equilibrium structure obtained from the MD simulation of the WPD_{open} conformation (2F6F) and the target structure was that of the WPD_{closed} conformation (1SUG). In TMD simulations, the occupancy (O) of the target atoms should be nonzero in TMD PDB file. The TMD file must contain the same number of atoms as the structure file. The coordinates for the target structure are also taken from the targeted atoms in this file and the non-targeted atoms are ignored. Therefore, the atoms of the target structure are either removed or added in the TMD PDB file in order to have the same number of atoms in both target coordinates (the WPD_{closed} conformation) and the initial coordinates (the WPD_{open} conformation). Then, the missing atoms in the target coordinates were built and coordinates were guessed by the psfgen module of VMD. Finally, the occupancy values of the target atoms in TMD file (i.e., all atoms of the residues 176-187 for the applied force on only WPD loop or residues (-3)-298 for the applied force on the whole protein) were set to 1 (nonzero). After the target TMD file was prepared for TMD simulation, the simulation was performed using the configuration file given in Appendix D.

3.2.2. TMD Simulation Parameters

All TMD simulations were also carried out using the NAMD 2.7 program and the CHARMM 27 force field. The SHAKE algorithm was applied to keep all bonds involving hydrogen bonds fixed and Particle Mesh Ewald (PME) method was used for long range electrostatic interactions. Electrostatic interactions were computed by the particle-mesh-Ewald method in TMD simulations and the non-bonded cutoff was taken to be 14 Å. The length of the TMD simulations was 5 ns and the time step for updating the potential was 0.5 ps. In one set of simulations, the force (k=300 kcal/mol/$Å^2$) was applied only to all atoms of WPD loop. In a second set of simulations, the force (k=5000 kcal/mol/$Å^2$) was applied to the all atoms of the whole protein. The simulations were carried out at 300 K and 1 atm. The energy of the system was not minimized in TMD simulations since the

structures used in TMD simulations were already energy minimized and equilibrated by MD simulations.

3.3. Trajectory Analysis

3.3.1. Root Mean Square Deviations (RMSD)

Root mean square deviation (RMSD) between equivalent atoms of two superimposed structures was used to determine structural similarity. The root mean square deviation of the Cα atoms was calculated between the simulation structures and the target structure throughout the TMD simulations after superpositioning all Cα atoms based on the target structure (Equation 3.7).

$$rmsd = \frac{\sqrt{\sum_{i=1}^{N} d_i^2}}{N} \tag{3.7}$$

3.3.2. Mean Square Fluctuations (MSF)

The mean square fluctuation was calculated as the deviation from the average structure (Equation 3). The initial structure and the trajectory of the simulation were aligned before the calculation of the fluctuations.

$$MSF = \left\langle \left(R_i(t) - \langle R_i \rangle \right) \left(R_i(t) - \langle R_i \rangle \right) \right\rangle \tag{3.8}$$

3.3.3. Principal Component Analysis (PCA)

PCA, which is an eigenvalue decomposition of the covariance matrix, gives a good description of collective motions in protein trajectory. This method is used to reduce large dimensional data sets to data sets with a few dimensions which still have most of the information in the original data matrix.

From a given data matrix X with n cases and p variables (X_1, X_2,…, X_p variables), a linear transformation to a new set of variables Y_1, Y_2,…, Y_p is calculated by using Equation 3.9, Equation 3.10 and Equation 3.11. The first principal component explains the major variation in the original p-dimensional data set. The second principal component explains the second largest percentage of the variation and so on. Principal components can simplify several analyses by reducing the dimensionality of the original data. Because, generally the first few principal components explain the most of the variation and the rest of the principal components do not make a considerable contribution to explain the variation.

$$Y_1 = a_{11}X_1 + a_{21}X_2 + ... + a_{p1}X_p \tag{3.9}$$

$$Y_2 = a_{12}X_1 + a_{22}X_2 + ... + a_{p2}X_p \tag{3.10}$$

...

$$Y_p = a_{1p}X_1 + a_{2p}X_2 + ... + a_{pp}X_p \tag{3.11}$$

4. RESULTS AND DISCUSSION

Seven different targeted molecular dynamics (TMD) simulations between the WPD_{open} and WPD_{closed} conformations of PTP1B were performed to elucidate the conformational activation mechanism of PTP1B. The simulation conditions and the initial structures are listed in Table 4.1. Representative snapshots from equilibrium MD simulations on the WPD_{open} structure (2F6F) were used as the initial structures in the TMD simulations and a representative snapshot from an equilibrium MD simulation on the WPD_{closed} conformation (1SUG) was taken as the target structure except in TMD4, in which case the initial and final structures were reversed. Subscripts show the different equilibrium structures obtained by MD simulations on the crystal structures; e.g. $2F6F_1$ denotes one equilibrated structure sampled from the MD simulation starting from the crystal structure 2F6F, $2F6F_2$ denotes another and so forth.

To demonstrate the difference between the WPD_{open} and WPD_{closed} conformations, the distances between C_α atoms of the initial and target structures in the TMD1 simulation were calculated after aligning the C_α atoms (Figure 4.1). The WPD loop (average RMSD between the two structures: around 4 Å), R loop (6 Å), the loop between β1-β2 sheets (L2) (3 Å) and α7 helix (6 Å) differ the most between the two structures. We performed five TMD simulations in which the force was applied to the WPD loop and two TMD simulations in which the force was applied to the whole protein. To examine the high dimensional energy surface spanned by PTP1B during the TMD simulation trajectories more thoroughly, TMD simulations were repeated using different initial and target structures, different force constant values, and by applying force on different regions of the protein.

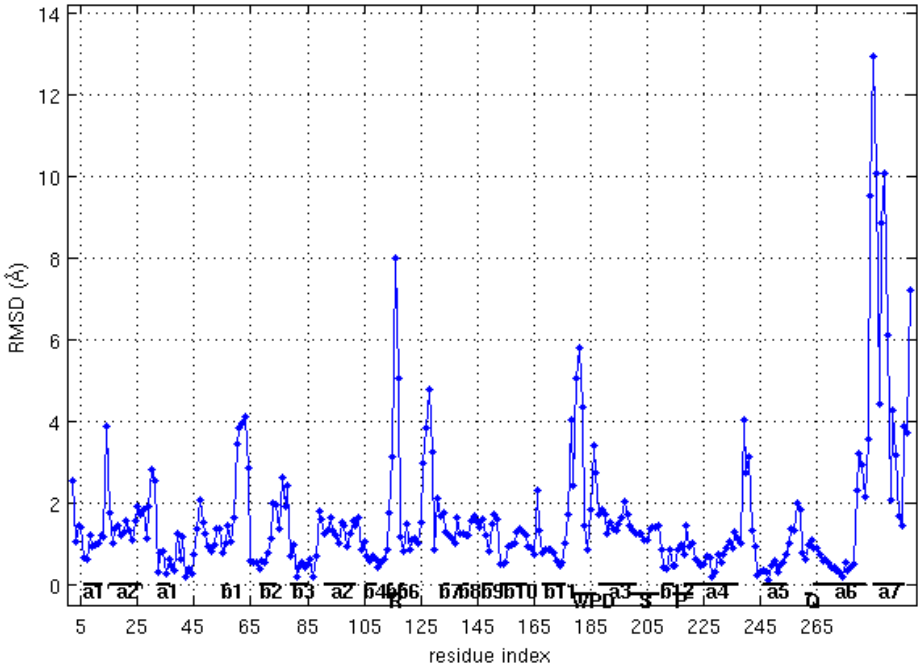

Figure 4.1. Distance between C_α atoms of the initial structure of TMD1 simulation ($2F6F_1$) and the target structure ($1SUG_1$).

Table 4.1. List of the TMD simulations and simulation conditions

RUN NAME	Initial Structure	Target Structure	Force Constant (kcal/mol/Å^2)	Targeted Region	Simulation Length (ns)
TMD1	$2F6F_1$	$1SUG_1$	300	WPD loop	5
TMD2	$2F6F_2$	$1SUG_1$	300	WPD loop	5
TMD3	$2F6F_3$	$1SUG_1$	300	WPD loop	5
TMD4	$1SUG_1$	$2F6F_1$	300	WPD loop	5
TMD5	$2F6F_4$	$1SUG_1$	300	WPD loop	5
TMD6	$2F6F_1$	$1SUG_1$	5000	Protein	5
TMD7	$2F6F_4$	$1SUG_1$	5000	Protein	5

4.1. Overall Structure Analysis During WPD Closing and Opening Motions in TMD1 and TMD4 Simulations

TMD1 (WPD_{open} to WPD_{closed}) and TMD4 (WPD_{closed} to WPD_{open}) simulations were compared to determine whether similar intermediate states were visited in the forward and reverse pathways. The initial structure in the TMD1 simulation was a representative snapshot from the equilibrium MD simulation on the WPD_{open} conformation ($2F6F_1$) and the target structure was a representative snapshot from the equilibrium MD simulation on the WPD_{closed} conformation ($1SUG_1$). Another TMD simulation (TMD2, WPD_{open} to WPD_{closed}) was performed using different initial coordinates to see whether a similar conformational activation pathway was visited. The initial structure in TMD2 was another representative snapshot from the equilibrium MD simulation on the WPD_{open} conformation ($2F6F_2$). In TMD4, the initial structure was the WPD_{closed} conformation ($1SUG_1$) and the target was the WPD_{open} conformation ($2F6F_1$). The force (force constant is 300 kcal/mol/Å^2) was applied to all atoms of the WPD loop and the simulation length was 5 ns in these three simulations. Overall structural changes in the protein were examined by calculating the RMSD of C_α atoms of the TMD simulation structures to the target conformation after aligning all C_α atoms of PTP1B with respect to the target conformation.

Figure 4.2A shows that the RMSD increases by ~0.4 Å within 4 ns of the simulation, then decreases approximately to its initial value, showing that the overall protein structure does not move toward the target structure in TMD1. A similar result was obtained in TMD2 (Figure A.1). A slight decrease of about 0.3 Å in the RMSD of all atoms in TMD4 (Figure 4.2B) suggests that the overall protein structure moves toward the target conformation as WPD loop opens in TMD4.

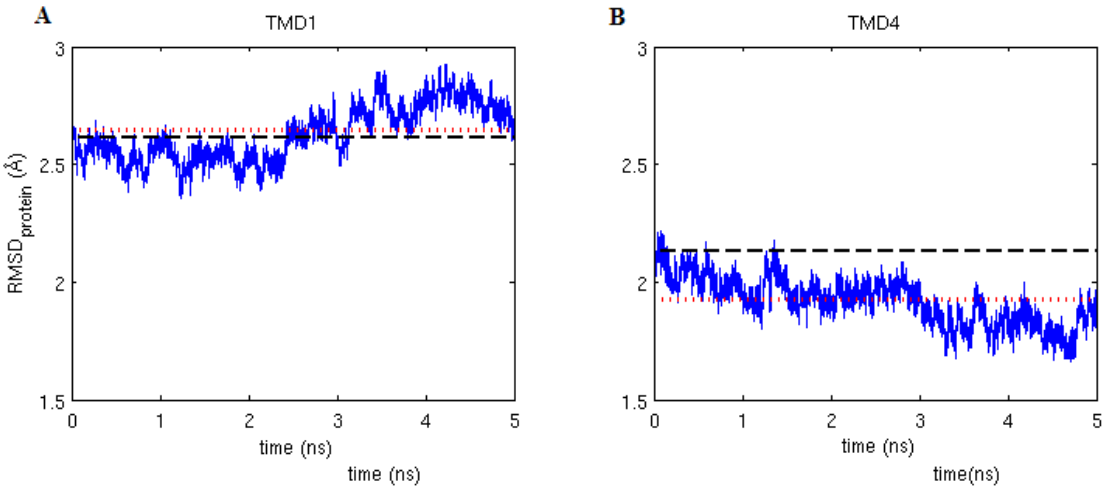

Figure 4.2. RMSD of the whole protein relative to the target structure during (A) TMD1 and (B) TMD4 simulations. Dashed lines show the RMSD value (average value of the first 100 snapshots) and dotted lines show the final value (average value of the last 100 snapshots).

In order to examine the contribution of the individual structural elements to the change of RMSD to target values in TMD1 or TMD4, distance between C_α atoms of the final and target structures, and distance between C_α atoms of the initial and target structures were calculated after aligning the structures with respect to the C_α atoms (Figure 4.3). As expected, the most significant difference is in the WPD loop, which moves toward its target conformation in all simulations due to the targeting potential. L2 moves away from its conformation in the target. In TMD2, the loop between α2'-α1 moves toward its conformation in the target, and the R loop moves away from its conformation in the target (Figure A.2). In TMD4, the only region that moves away from the target (WPD_{open}) structure is the loop between α2'-α1 helices, while L3, R loop and L8 move toward their respective target states (Figure 4.4). These results show that the WPD loop motion may be

coupled with the R-loop, which is in the vicinity of the WPD loop, and the α2'-α1 loop, which is relatively distant from the WPD loop.

In order to examine whether other parts of the protein move away from their initial conformations during WPD closing and opening motions, distances between the C_α atoms of the initial and final structures in the TMD simulations were calculated after aligning the C_α atoms with respect to their initial structures (Figure 4.5). The loop between α2'-α1 helices, the loop between β1-β2 strands (L2 Loop), R loop, WPD loop and α7 helix move away from their initial conformations in TMD1. These regions are shown in Figure 4.6. α7 helix moves by about 2 Å between the initial and final states in TMD1, but its conformation is different from the initial or the target structure. In TMD2, L2 loop does not move away as much as it does in TMD1, but WPD loop, R loop and α7 helix move farther away from their initial conformations, compared to TMD1 (Figure A.3).

After the distances between the C_α atoms of the initial and final structures in the TMD simulations were calculated (Figure 4.5, Figure 4.7 and Figure A.3), the correlation between these values in TMD1, TMD2 and TMD4 simulations were determined to investigate whether the displacements in the protein are located in similar regions in these simulations or not. Table 4.2 shows that the correlation of the residue based displacements in TMD1 and TMD2 is high (0.75 including the WPD loop displacements, 0.65 excluding WPD loop displacements), showing that displacements in the protein are located in similar regions, irrespective of the initial conditions used in TMD simulations. On the other hand, the correlation value between TMD1 and TMD4 is relatively lower (0.56 including the WPD loop displacements, 0.32 excluding WPD loop displacements). This correlation values are important measurements to show the similarities or differences between the residue based displacements in TMD1, TMD2 and TMD4. Therefore, it can be showed that the residues that have highest displacements during the TMD simulations are consistent or not in these three simulations.

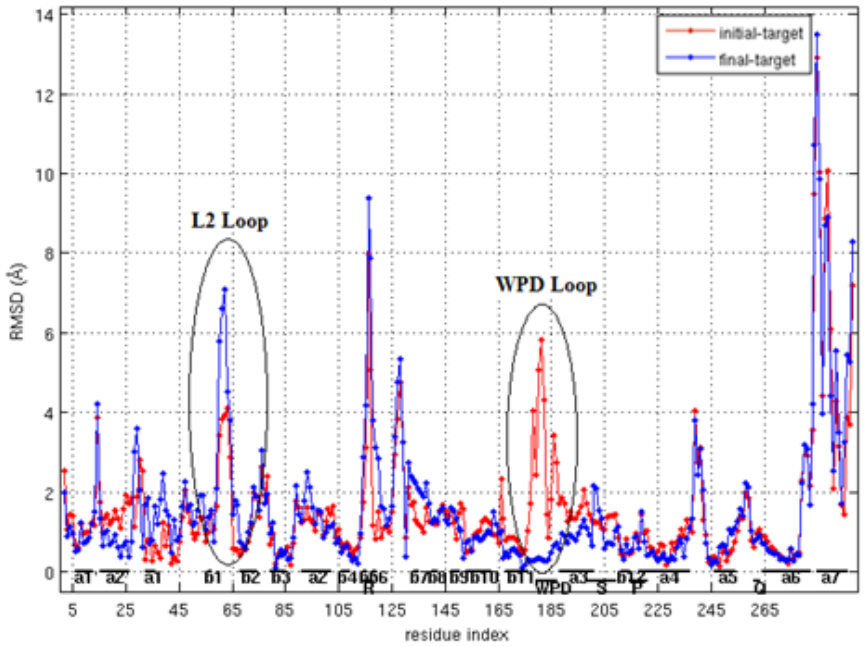

Figure 4.3. Residue based distances for all C_α atoms between the target structure and the initial structure (red), and distances between the target structure and the final structure (blue) in TMD1 simulation. (The initial structure is the average value of the first 100 snapshots and the final value is the average value of the last 100 snapshots).

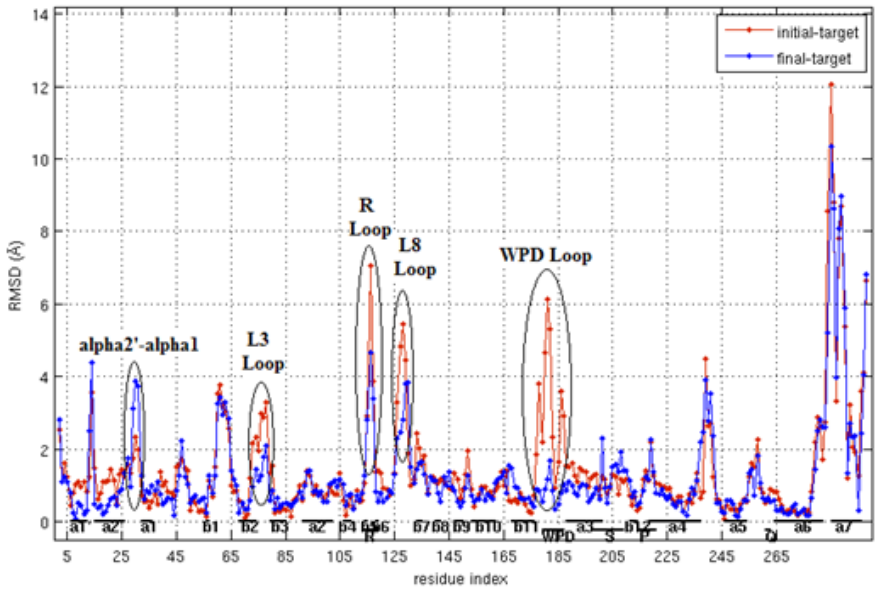

Figure 4.4. Distances between C_α atoms of the target structure and the initial structure (red), and distances between C_α atoms of the target structure and the final structure (blue) in TMD4. (The initial structure is the average value of the first 100 snapshots and the final value is the average value of the last 100 snapshots).

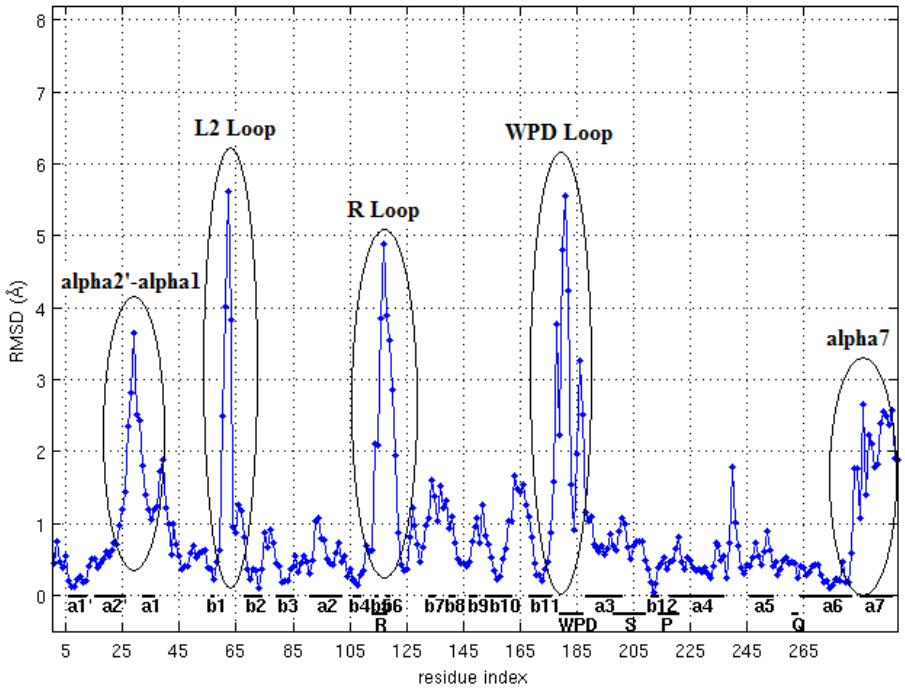

Figure 4.5. Distances between the C_α atoms of the initial and the final structures in the TMD1 simulation. (The initial structure is the average value of the first 100 snapshots and the final value is the average value of the last 100 snapshots).

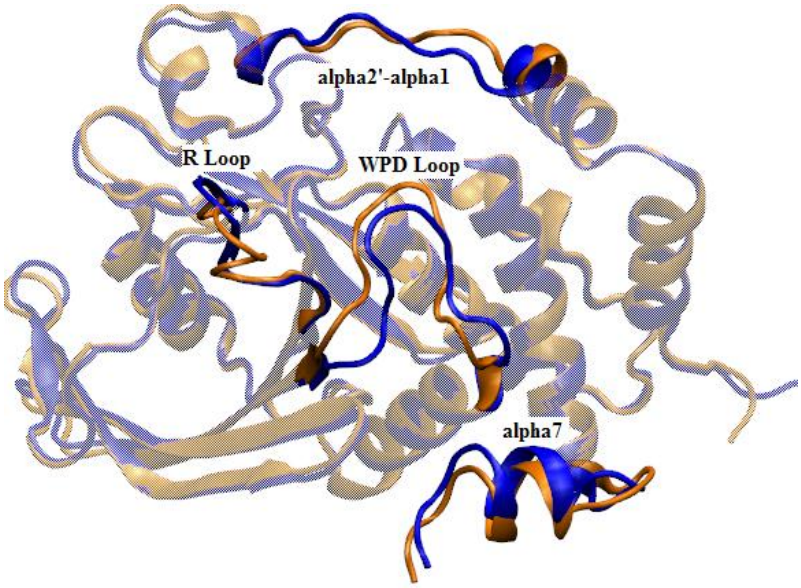

Figure 4.6. The cartoon representation of the initial structure (in WPD_{open} form) shown in blue and the final structure from TMD1 simulation (in WPD_{closed} form) shown in orange. Regions that move away by more than 2 Å from their initial conformations are labeled on the structure.

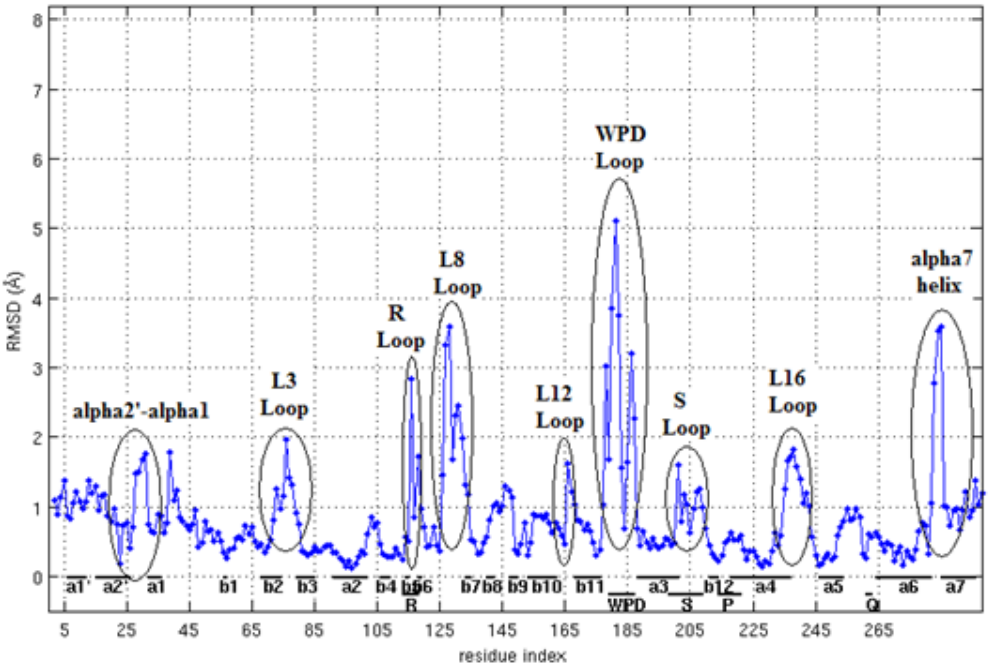

Figure 4.7. RMSDs of the C_α atoms between the initial and the final structures in the TMD4 simulation. (The initial structure is the average value of the first 100 snapshots and the final value is the average value of the last 100 snapshots).

Table 4.2. The correlation values between the C_α atom displacements in TMD1, TMD2 and TMD4 simulations

Correlation Value	TMD1-TMD4	TMD1-TMD2
WPD loop displacements excluded	0.32	0.65
WPD loop displacements included	0.56	0.75

4.2. Comparison of Residue Mobility During Closing/Opening Motion of WPD Loop

Mobility of different structural elements was determined by calculating residue based MSF of the simulation structures around the average structure of the TMD simulations during TMD1 and TMD4 (Figure 4.8). The average MSF value of TMD1, 0.85 $Å^2$ (TMD2: 0.71 $Å^2$) and that of TMD4, 0.49 $Å^2$ suggests that the protein fluctuates more in the

closing, TMD1, simulation. The regions of highest mobility in the TMD1 simulation are the loop between α2'-α1 helices, the loop between β1-β2 sheets (L2 loop), R loop, WPD loop, S loop and α7 helix (Figure 4.8A). These regions are highlighted on the protein (Figure 4.9A). Despite an overall drop in mobility, similar regions have high mobility in TMD4, with additional regions that are also highly mobile; the loop between β6-β7 sheets (L8 loop), the loop between β10-β11 sheets (L12 loop), the loop between α4- α5 helices (Figure 4.8B). Regions of high mobility in TMD4 are highlighted on the protein (Figure 4.9B).

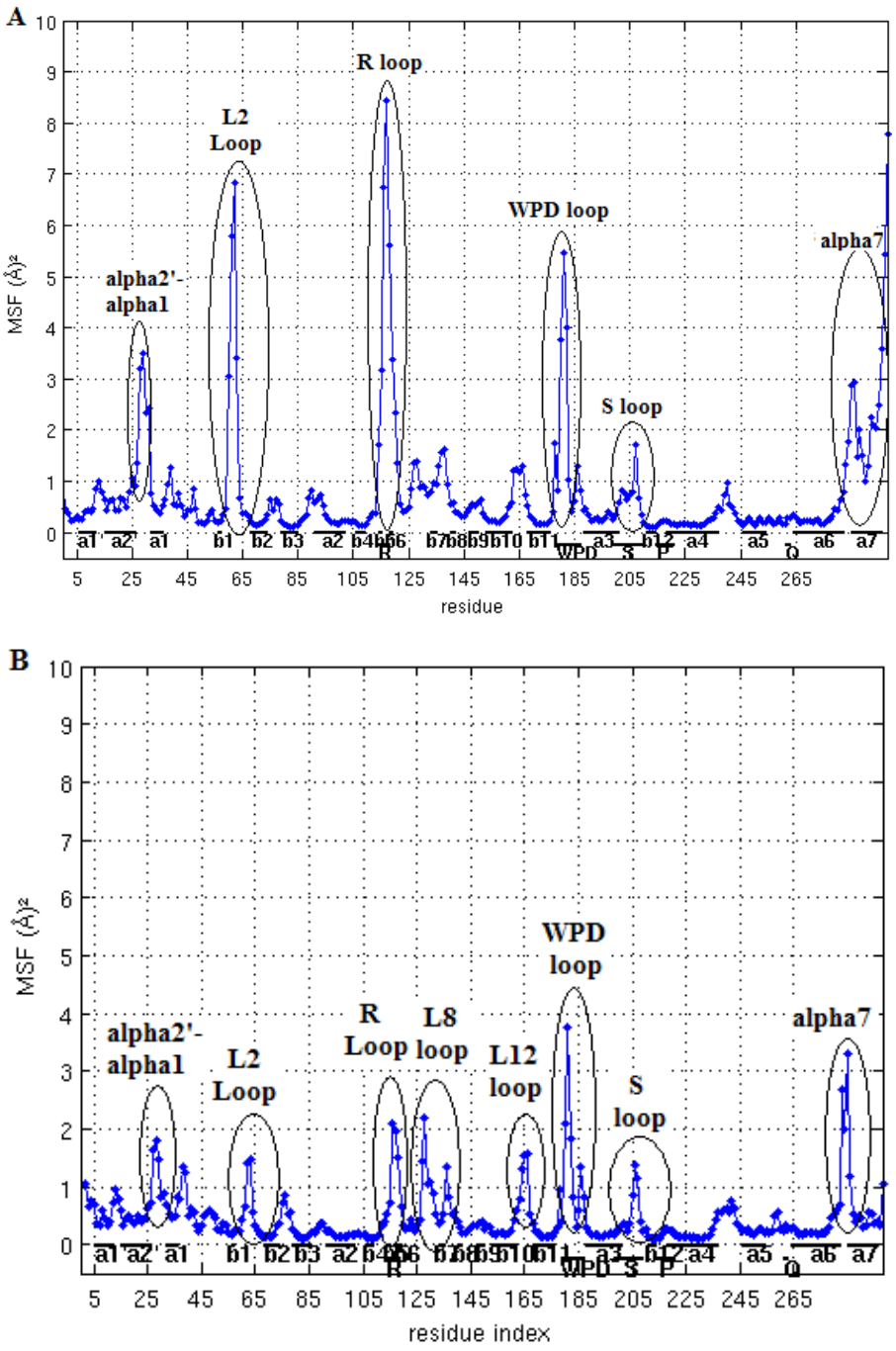

Figure 4.8. MSF of the residues during the (A) TMD1 simulation, and (B) TMD4 simulation.

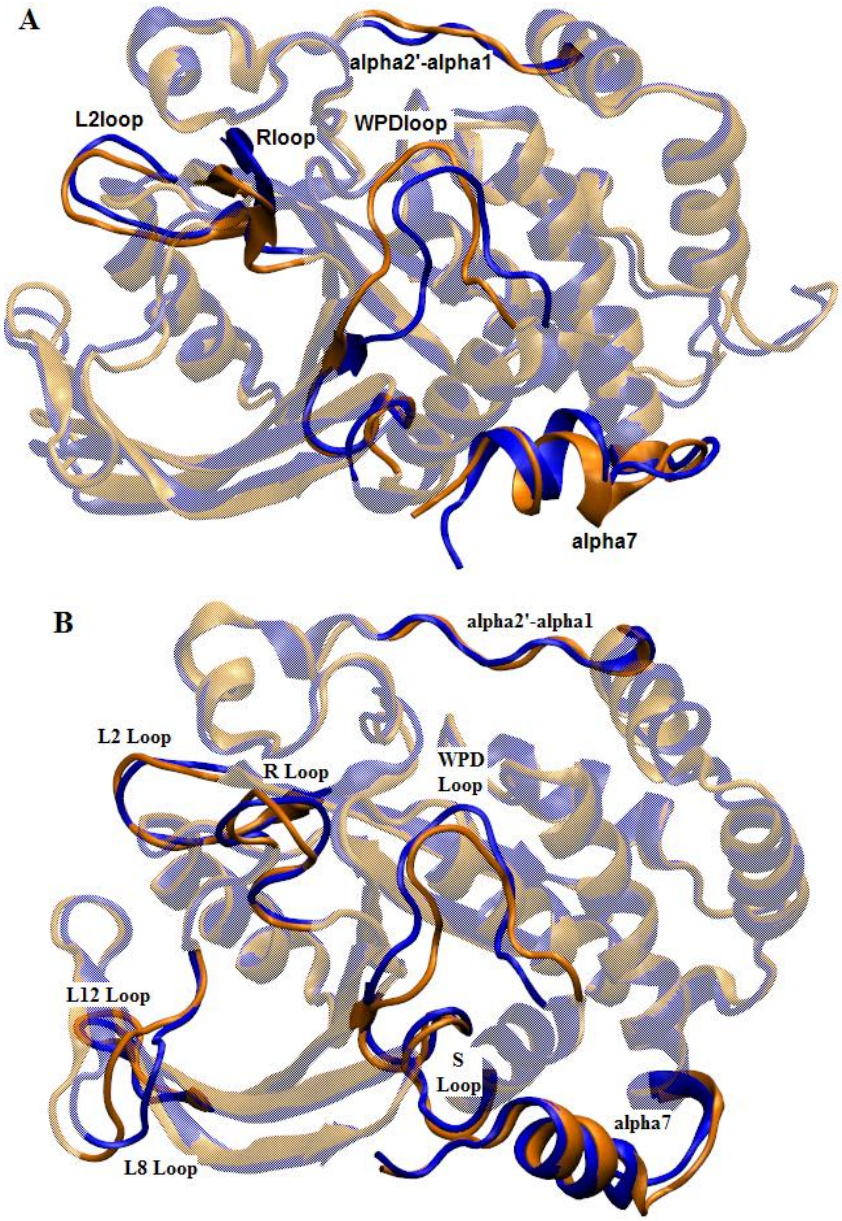

Figure 4.9. The cartoon representation of (A) initial structure (in WPD_{open} form) shown in blue and that of final structure (in WPD_{closed} form) shown in orange in TMD1 and (B) initial structure (in WPD_{closed} form) shown in blue and that of final structure (in WPD_{open} form) shown in orange in TMD4. Regions with high MSF values are labeled on the structure.

Residue based fluctuations in TMD1 were compared with those in an equilibrium MD simulation on 2F6F (Figure 4.10). The residue based fluctations of the equilibrium MD simulations were obtained by taking an average value for each residue of the 5 ns segments from the simulation trajectory (Figure 4.10). Mobility of the regions of the loop

between α2'-α1 helices, L2 loop (the loop between β1-β2), R loop, WPD loop, S loop and α7 helix are higher in TMD1 compared to the MD simulation. As expected, WPD loop fluctuates much more in TMD compared to MD due to the force applied. In addition, R loop also fluctuates more in TMD simulation than in equilibrium MD simulation (Figure 4.10). In TMD2, R loop, WPD loop and α7 helix are also more mobile regions than in equilibrium MD but the loop between α2'-α1 helices and the loop between β1-β2 (L2 loop) are not more mobile in TMD2 (Figure A.4). Comparison of the residue mobility values in TMD4 simulation with those in the MD simulation on 1SUG structure shows that the R loop, the loop between β6-β7 sheets (L8 loop), the loop between β10-β11 sheets (L12 loop), WPD loop, S loop and α7 helix are more mobile in TMD4. The loop between α4- α5 helices (L16 loop) is the only region that is more mobile in the MD simulation compared to TMD4 simulation (Figure 4.11). Note that, although S loop does not fluctuate in equilibrium MD simulations and does not move away from its initial conformation in TMD1 (Figure 4.5), it is mobile in TMD simulations from its MSF value, showing that S loop should be flexible for WPD loop closing and opening.

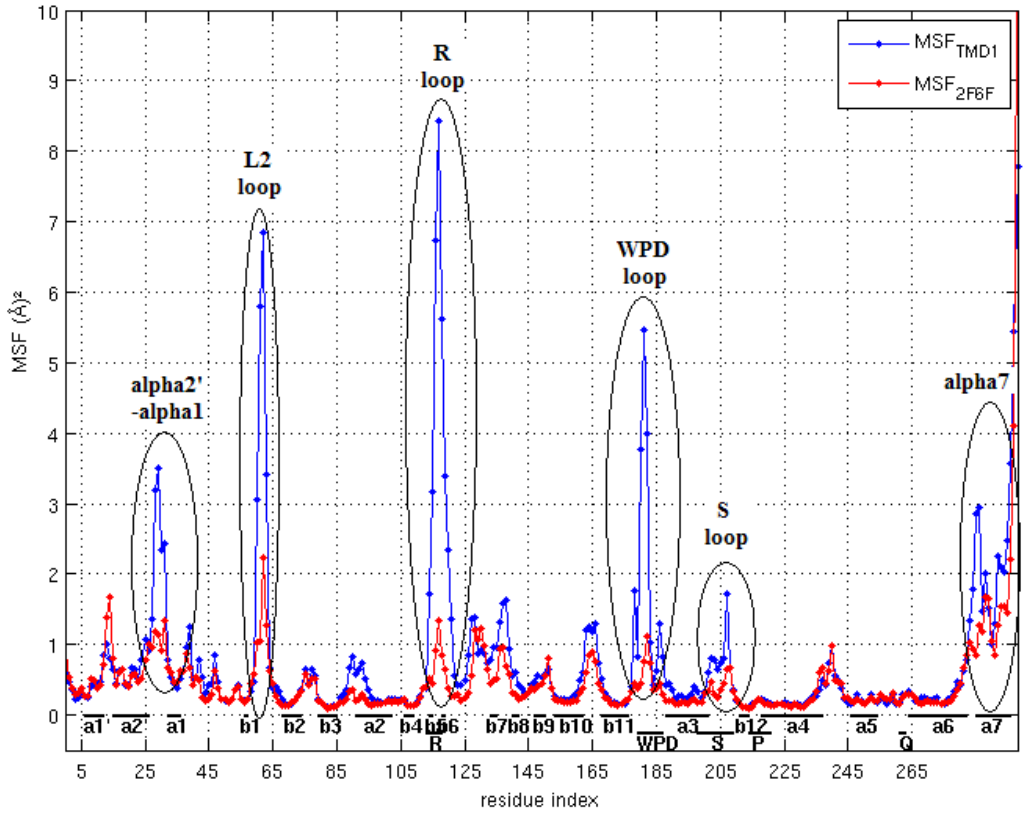

Figure 4.10. Residue based MSF values in TMD1 simulation (blue) and in the MD simulation of the 2F6F structure (red).

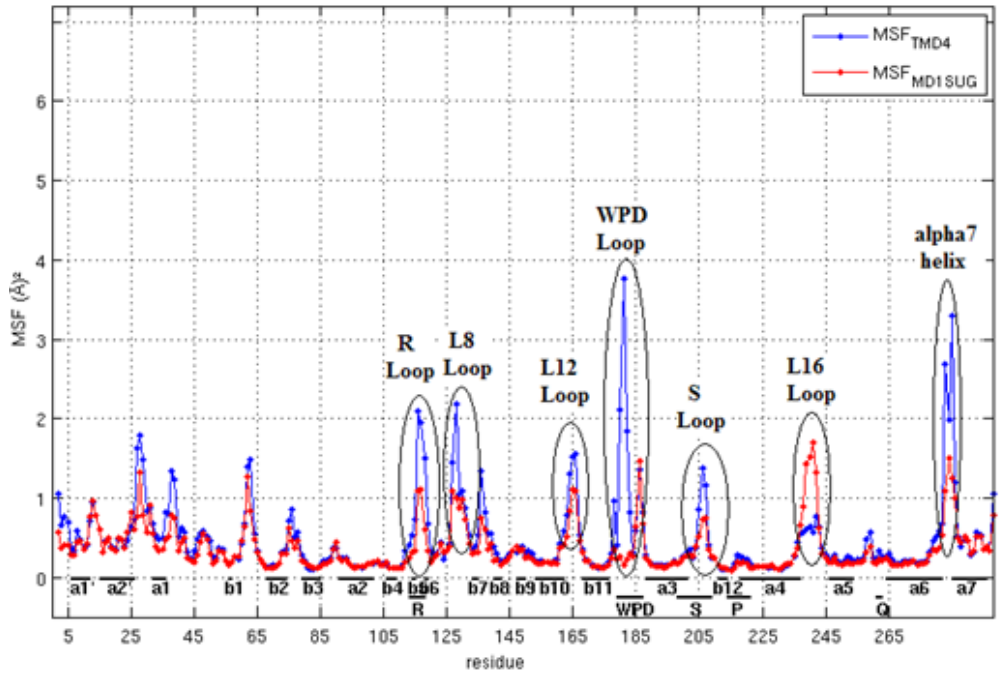

Figure 4.11. Residue based MSF values in TMD4 (blue) and in the MD simulation of the 1SUG structure (red).

4.3. Clustering of WPD Loop Conformations in the TMD Simulation Trajectories

RMSD of the WPD loop between the conformation in the simulation and that in the target structure should decrease and the WPD loop should move toward the target conformation during the TMD simulation due to the force applied on WPD loop. The progress of the WPD loop motion was monitored by calculating the RMSD of C_α atoms of the WPD loop to its target conformation, after aligning all the C_α atoms of the protein with respect to the target conformation. Figure 4.12A shows that RMSD of the WPD loop decreases from 4.75 Å to about 0.5 Å in TMD1, showing that the WPD loop closes. There is a decrease in the RMSD values at ~0.75ns, a sharp decrease is also clearly seen at ~2.75 ns and another decrease is observed at ~4.5 ns after a plateau in the RMSD values. In order to observe the conformational change of WPD loop at these time steps, the conformations of the WPD loop before and after these time steps in TMD1 are shown in Figure 4.13. The C terminus of the WPD loop (residues 184-189) move toward the target conformation at ~0.75 ns (Figure 4.13A) while the N terminus and tip of the loop move toward the target at

~2.75 (Figure 4.13C) and residues 182-183 move toward the target at ~4.5 ns (Figure 4.13E). TMD2 simulation has a similar RMSD profile (Figure A.5B); there is a sharp decrease at ~2.75 ns, and there are two other decreases at ~3.75 ns and 4.75 ns. In TMD4, RMSD of the WPD loop decreases from 3.5 Å to ~1 Å, showing that the WPD loop moves in the desired direction, but it does not open completely (Figure 4.12B). There is a decrease in the RMSD of WPD loop at ~1ns, a sharp decrease at ~3 ns and another decrease at 4.25 ns in TMD4. The conformations of WPD loop before and after these time steps are shown in Figure 4.9. The opening motion of C terminus of WPD loop occurs at ~0.1 ns (Figure 4.13B) while the N terminus and tip of the loop move toward the target at ~3 ns (Figure 4.13D) and residues 181-183 move toward the target at ~4.25 ns (Figure 4.13F). Therefore, the WPD loop closing and opening motions occur with a similar order of events.

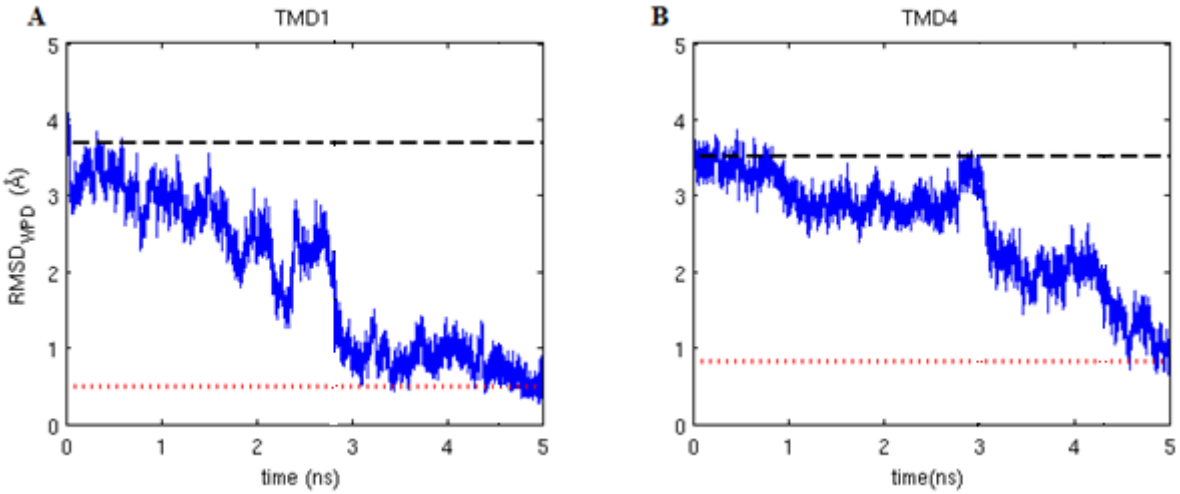

Figure 4.12. RMSD of C_α atoms of the WPD loop relative to the target structure during (A) TMD1 and (B) TMD4 simulations (Dashed lines represent the initial value, dotted lines represent the final value).

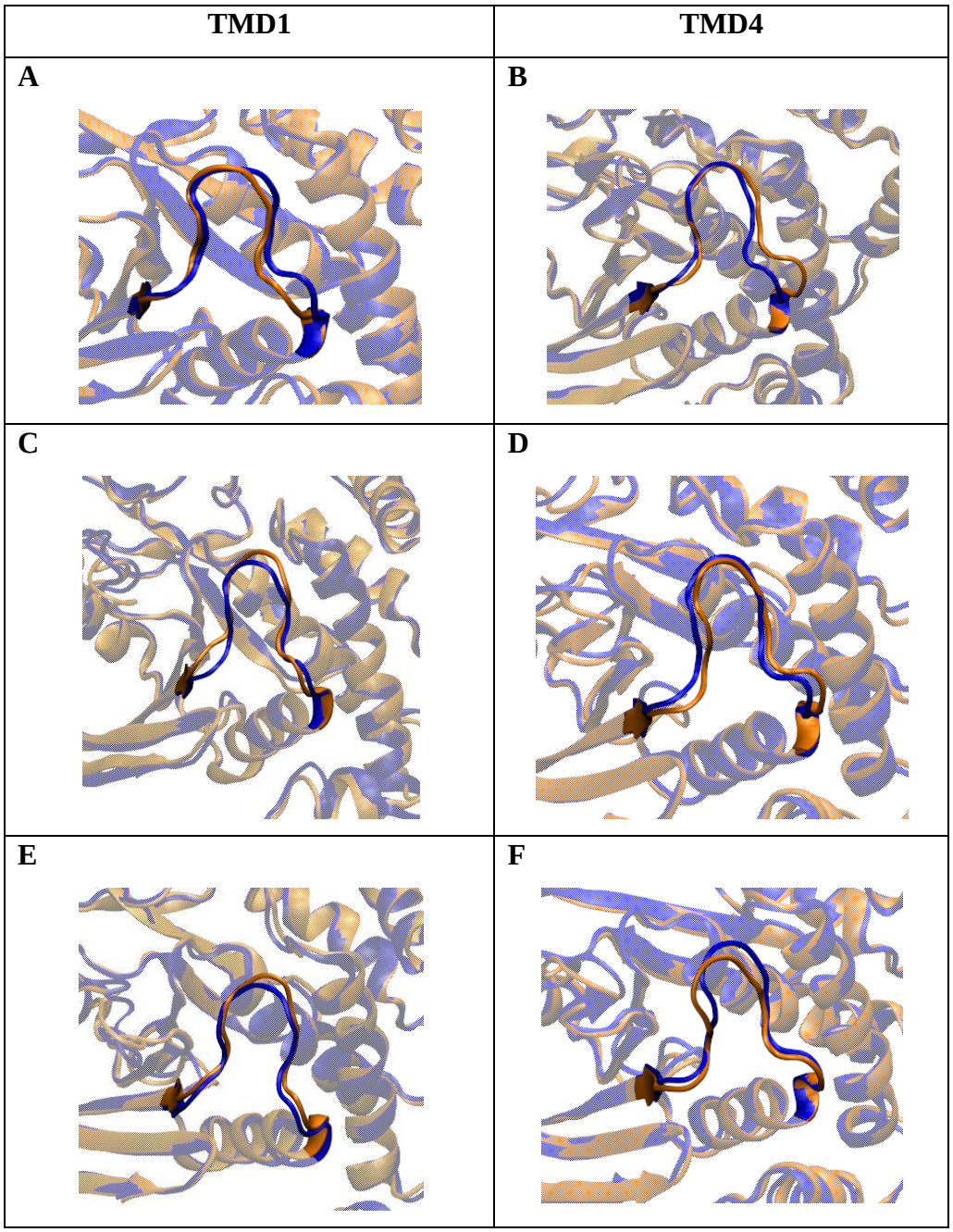

Figure 4.13. Conformations of the WPD loop in TMD1 (left panel) and TMD4 (right panel) snapshots representing the conformations before and after the sharp decrease in the RMSD values of C_α atoms of WPD loop in these simulations. In TMD1, snapshots at (A) 0.5ns (blue) and 1ns (orange), (C) 2.5ns (blue) and 3 ns (orange) and (E) 4 ns (blue) and 4.9 ns (orange) are shown. In TMD4, snapshots at (B) 0.5ns (blue) and 1.5 ns (orange), (D) 2.5ns (blue) and 3ns (orange) and (F) 4 ns (blue) and 4.75 ns (orange) are shown.

In order to cluster the conformations visited by the WPD loop during the closing/opening motions, PCA was applied to the C_α atoms of the WPD loop in TMD1, TMD2 and TMD4. The percentage of the WPD loop closing or opening motion explained by the first five principal components (PCs) in TMD simulations is shown in Table 4.3. The first PC (PC1) is the direction of the motion with the largest fluctuations; the second PC (PC2) is the direction of the motion with the largest variance in the remaining unexplained covariance matrix.

Table 4.3. Percentage of the WPD loop motion explained by the first five principal components

TMD Simulation	Principal Component Number	% Explanation of WPD loop Motion
TMD1	PC1	77
	PC2	5.7
	PC3	4.2
	PC4	2.3
	PC5	1.7
TMD2	PC1	77
	PC2	10
	PC3	2.3
	PC4	2.2
	PC5	1.2
TMD4	PC1	59
	PC2	13
	PC3	8.0
	PC4	3.6
	PC5	3.1

In TMD1, PC1 explains ~77 % of the overall WPD loop closing motion (Table 4.3). Examination of the explanatory power of PC1 on residue basis shows that ~80 % of the motions of the residues 178-182 (the tip of the loop) are explained, whereas ~50% of the motions of the residues 183-189 (C-terminus of the loop) are explained by PC1. The

general closing motion of the WPD loop is described by the first PC (Figure 4.14A-B). On the other hand, PC2 explains ~30% of the localized motion of the residues 183-189 in TMD1 (Figure 4.14C-D). Therefore, PCA shows that a significant percentage of the WPD loop C-terminal motion is independent of the rest the loop.

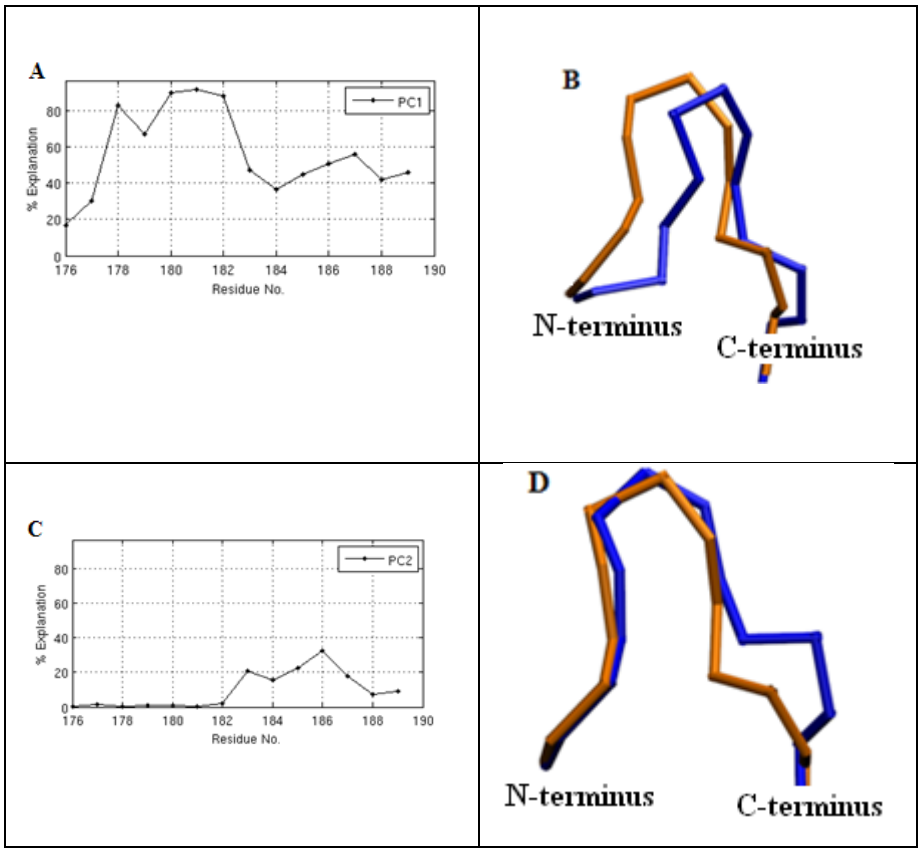

Figure 4.14. In TMD1, (A) the percentage of WPD loop motion explained by PC1, and (B) representative projection snapshots along PC1. (C) The percentage of WPD loop motion explained by PC2, and (D) representative projections along PC2. In (B) and (D), the first and the final projection snapshots are shown in blue and orange, respectively.

WPD loop trajectory TMD1 is projected onto the PC1-PC2 subspace and the conformations adopted by the WPD loop are clustered on this two-dimensional reduced subspace using the k-means algorithm (Figure 4.15). Visual examination of the two-dimensional plot shows three separate clusters.

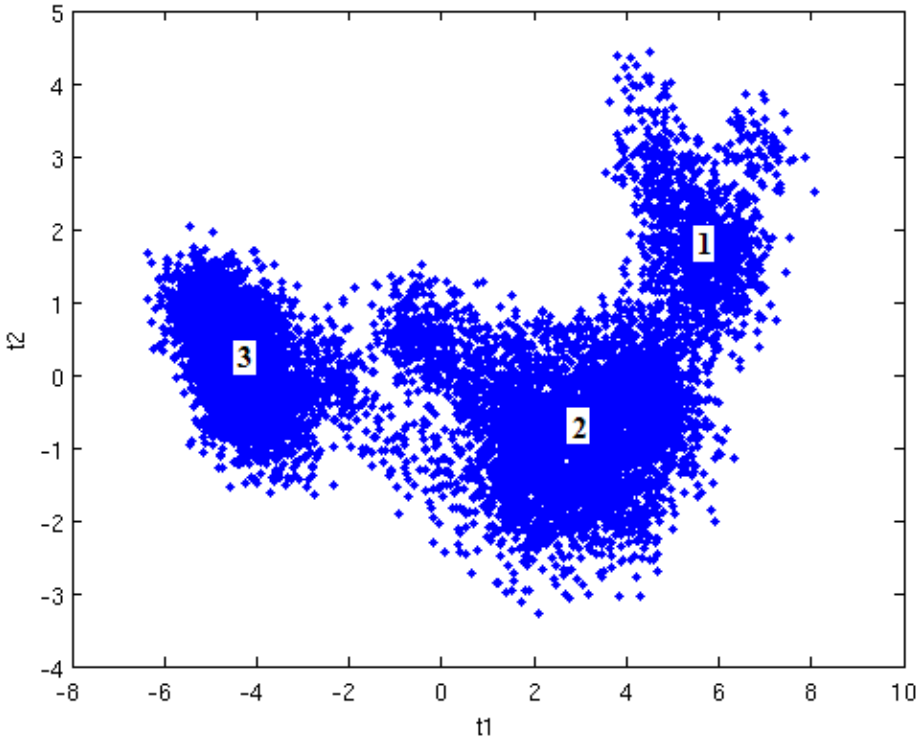

Figure 4.15. Projection of the WPD loop trajectory in TMD1 on the reduced subspace of PC1 and PC2.

WPD loop trajectory is projected onto PC1 and PC2, separately, and each cluster is shown with a different color in Figure 4.16. The transition of the WPD loop conformation to the next cluster occurs at 0.5-1.5 ns and 2.5-3 ns. The difference between the first and second clusters is more evident in the projection on PC2, while the second and third clusters can be easily distinguished on PC1. One can visualize the TMD trajectory by using Figure 4.16 for TMD1 and 4.19 for TMD4. The C-terminus of the WPD loop moves toward the target conformation at the beginning of the simulation (motion along PC2), and the WPD loop makes a conformational jump as a whole in the middle of the simulation (motion along PC1).

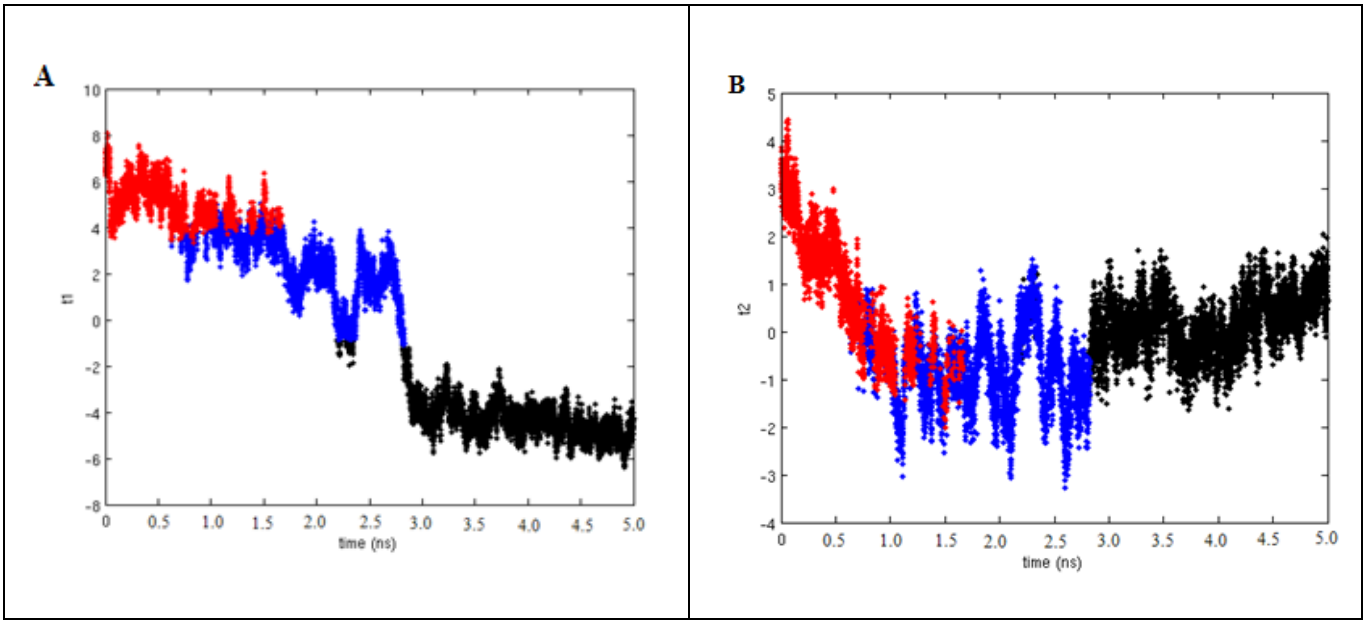

Figure 4.16. The projection of the WPD loop motions on (A) PC1 (B) and PC2 in TMD1. Three different clusters obtained based on Figure 4.15 are shown with different colors.

PC results obtained for TMD2 are similar to those for TMD1. In TMD2, PC1 also explains ~72 % of the loop motion (Table 4.3) and it describes the general closing motion (Figure A.5B). It explains ~80 % of the motions of the residues 178-182, and ~40% of the motions of residues 183-189 (Figure A.6A). PC2 explains ~35-40 % of WPD loop C-terminus motion (Figure A.6C- A.6D). In the clustering of WPD loop conformations by PC1 and PC2 in TMD2, an additional cluster is visible, therefore K-means routine is used with four clusters. (Figure A.7). The four clusters are separated at 1-1.5 ns, 2.5-3 ns and 3.5-4 ns, (Figure A.8A). Similar to TMD1, the first cluster is due to PC2, while the motion on PC1 is more significant in distinguishing the other two clusters.

Collective motions of the WPD loop in TMD4 also resemble those in TMD1. In TMD4, PC1 explains ~60 % of the motion (Table 4.3). More than 80% of the motion of residues 178-182 is explained by PC1, while ~30 % of the motions of residues 183-189 are explained by PC1 (Figure 4.17A-B). The PC2 explains ~30-40 % of the C-terminus of the WPD loop (residues 185-188) (Figure 4.17C-D). Projection of the WPD loop trajectory on the reduced two-dimensional subspace of PC1 and PC2 shows the existence of four clusters (Figure 4.18). Similar to TMD1, the difference of the first cluster from the other clusters is well-pronounced on PC2, while the second and third clusters can easily be

distinguished on PC1 (Figure 4.19). The transition to the fourth conformation occurs at ~4.5 ns, manisfesting itself on both PC1 and PC2.

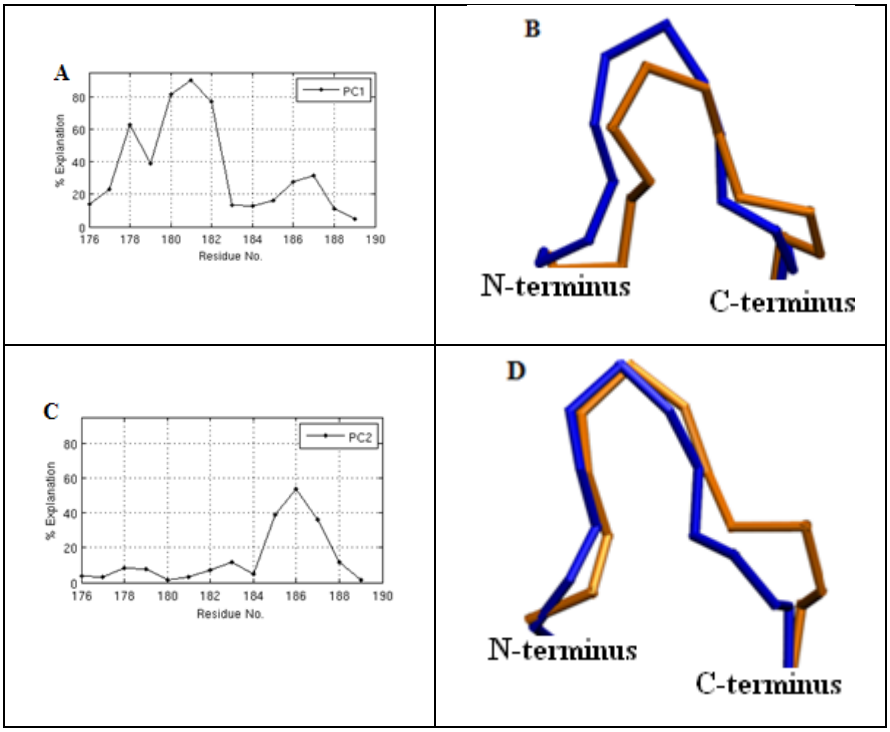

Figure 4.17. In TMD4, (A) the percentage of WPD loop motion explained by PC1, and (B) representative snapshots along PC1. (C) The percentage of WPD loop motion explained by PC2, and (D) representative projections along PC2. In (B) and (D), the first and the final conformations are shown in blue and orange, respectively.

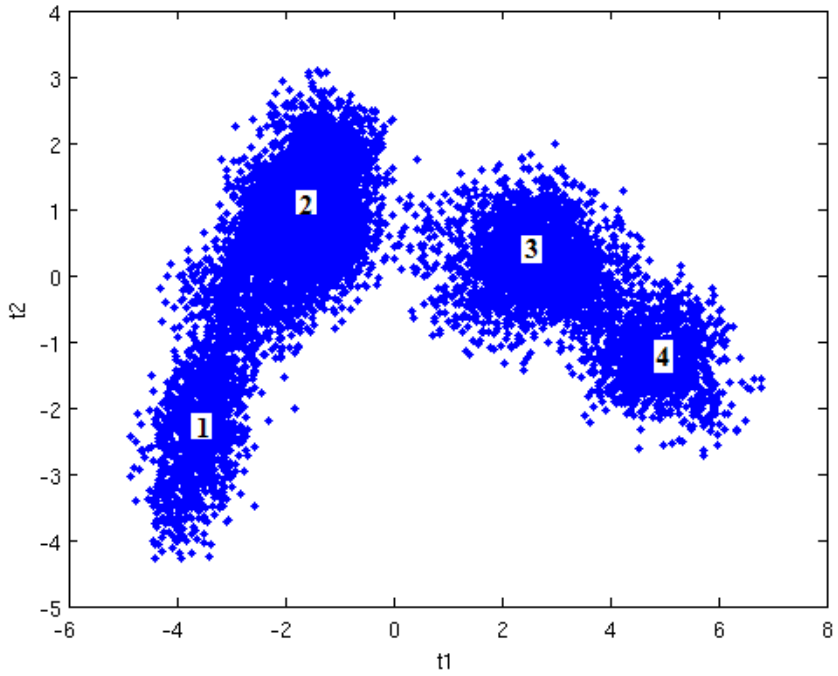

Figure 4.18. Projection of the WPD loop trajectory in TMD4 on the reduced subspace of PC1 and PC2.

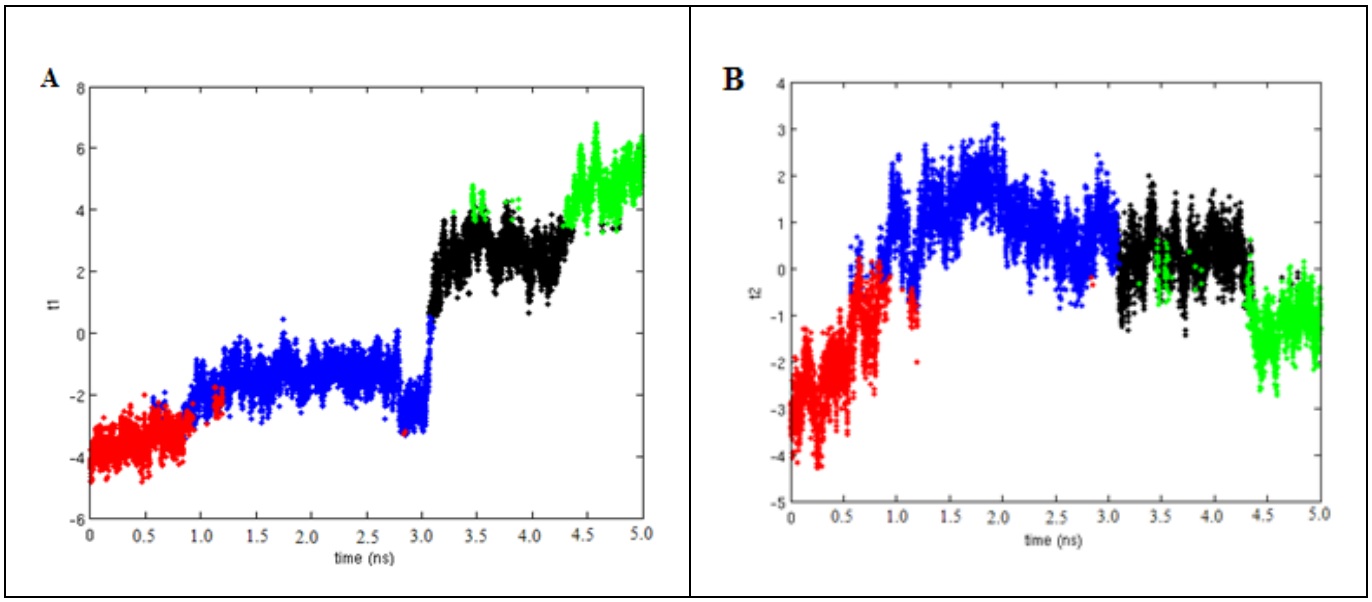

Figure 4.19. The projection of the WPD loop motions on (A) PC1 (B) and PC2 in TMD4. Three different clusters are shown with different colors.

4.4. Conformational Transitions During the WPD Loop Closing/Opening

In Section 4.3 the two major conformational transitions of the WPD loop between the closed and open states were discussed. The closing motion of the WPD-loop C-terminus (residues 183-189) respresents the first conformational transiton of the WPD loop. The closing motions of the WPD-loop N-terminus (Thr177-Trp179) and the tip of the WPD loop (Pro180-Phe182) are mainly responsible for the second conformational transition. The additional conformational jump at 3.75 ns in TMD2 simulation was examined in detail and it was seen that this transition corresponds to the formation of hydrogen bonds between the P-loop Ser216 sidechain with Asp181 and Arg221. This hydrogen bond was not observed in TMD1, suggesting that these interactions are not essential for WPD loop closing, but they may be part of a path visited in some cases. Three conformations of the WPD loop, averaged during 0-1 ns (red), 2-3 ns (blue), and 4-5 ns (black) are shown in Figure 4.20.

It should be recalled that these stages are determined by PCA and k-means clustering analysis of the C_α atoms. However, it is quite likely that there may be subtle changes, which do not manifest themselves as significant displacements in the backbone coordinates. This type of subtle transition has been observed in the detailed visual examination of the TMD1 and TMD2 trajectories, as discussed in Section 4.3 shown in Figures 4.12A and A5B.

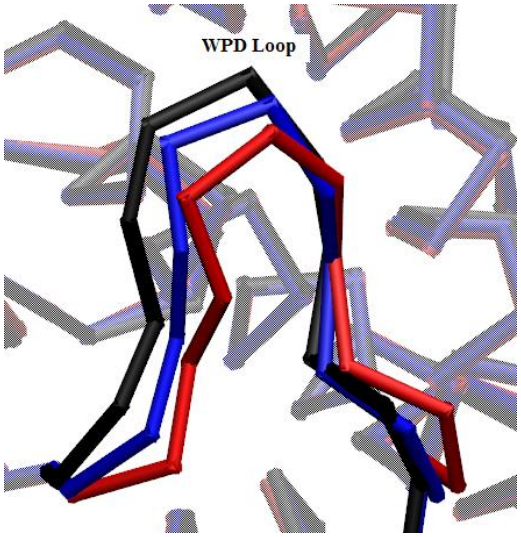

Figure 4.20. The WPD loop conformations in the first (red), second (blue) and third (black) clusters are the average conformations of the snapshots between 0-1 ns, 2-3 ns, and 4-5 ns, during the closing simulation.

4.4.1. First Conformational Transition of the WPD Loop

Clustering analysis and visual examination of the trajectories show that the C-terminus of the WPD loop (residues 183-189) moves toward the target conformation at ~0.5-1.5 ns in TMD1, TMD2 and TMD4 simulations. Examination of the backbone dihedral angles of this region shows that the backbone dihedral angles of Ser187 or Pro188 rotate in this time time interval in these simulations. The Φ angle of Ser187 rotates by ~50° between 0.5-1 ns in TMD1 (Figure 4.21A), and Φ angle of Pro188 rotates by ~20° at ~1 ns during the WPD loop opening in TMD4 (Figure 4.21B). In addition, Ψ angle of Pro188 dihedral angle rotates by 20-25° at ~1.25 ns in TMD2 simulation, resulting in the closure of the C-terminus part of the loop (Figure A.9). In both directions, the WPD loop C terminus moves first, suggesting that this is a relatively facile motion not hindered by energy barrier.

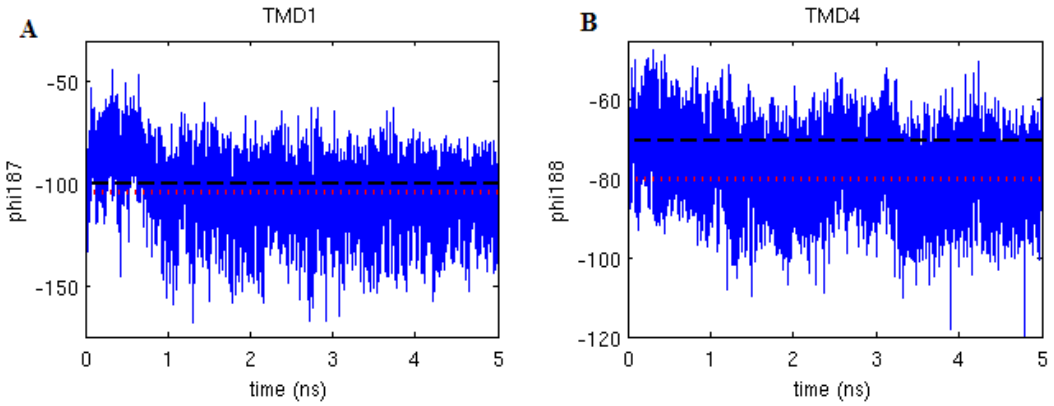

Figure 4.21. Time evolution in (A) Ser187 Φ angle in TMD1 and, (B**)** Pro188 Φ angle in TMD4. (Dashed lines represent the initial value, dotted lines represent the final value).

4.4.2. The Second Conformational Transition of the WPD Loop

The crystallographic studies on Yersinia PTP have revealed that the formation of a potential hydrogen bond between Trp354 and Arg409 (corresponding to Trp179 on the WPD loop and Arg221 of the active site in PTP1B) residues may mediate the closure of the WPD loop [15]. In an effort to determine whether formation of the Trp179-Arg221 interaction contributes to the conformational activation mechanism in PTP1B, interactions between Trp179-Arg221 and Asp181-Arg221 during the TMD simulations were monitored.

In TMD1, Trp179 χ2 angle rotates gradually by 80°-90° difference until 4 ns.(Figure 4.22A). A similar behavior is seen in TMD2 (Figure A.10B), in which χ2 increases steeply between 1.5-4 ns. The reverse of this motion is seen in TMD4 (Figure 4.22B). As Trp179 χ2 rotates, and WPD loop moves closer to the target conformation, the backbone carbonyl oxygen of Trp179 rotates toward the active site, and the distance between Trp179 and Arg221 decreases from 5 Å to 2.5 Å, and a hydrogen bond between Trp179 backbone O-Arg221 side chain NH1 is formed at ~2.5 ns (Figure 4.22C). This interaction is lost in a single step in TMD4 (Figure 4.22D). Visual examination shows that the Arg221 side chain χ2 rotates at the same time in TMD1. As the WPD loop closes, distance between Asp181 side chain and the Arg221 side chain decreases from 7 Å to 2.5 Å, and the salt bridge formation between Asp181 and Arg221 coincides with the Trp179O-

Arg221 hydrogen bond formation in TMD1 (Figure 4.22E) and TMD2 (Figure A.10E). Visual examination of the trajectory of the TMD1 simulation has shown that neither the side chain nor the backbone of Asp181 rotates at ~2.5 ns of the TMD simulations. The rotation of the side chain of Trp179, and Arg221, formation of Trp179-Arg221 and Asp181-Arg221 take place approximately at the same time as the second conformational transition seen in TMD simulations in both closing (TMD1, TMD2) and opening motions (TMD4) of WPD loop.

The Trp179-Asp181-Arg221 interactions are shown in the snapshots before and after the conformational transition in Figure 4.23A and B. In TMD1, the snapshot sampled at 2 ns shows that the WPD loop is away from the active site and Trp179-Arg221 and Asp181-Arg221 do not interact, while closing of the WPD loop at 3 ns enables the formation of these bonds (Figure 4.23A). In TMD4, the snapshot at 2.5 ns shows that WPD is close to the active site enabling the formation of polar interactions between Trp179-Arg221 and Asp181-Arg221, while these interactions are lost at 3.5 ns in TMD4 (Figure 4.23B).

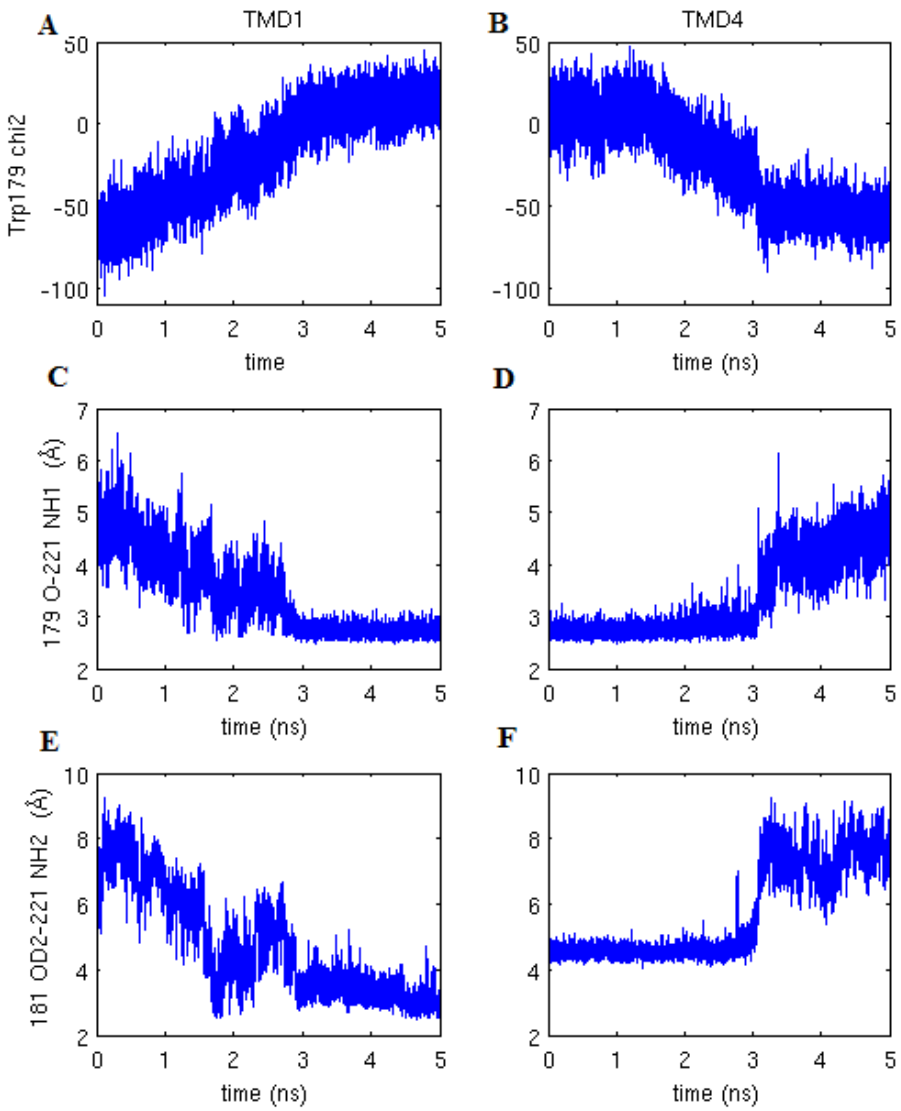

Figure 4.22. The time evolution of (A,B) Trp179 χ2 angle , (C,D) Trp179 O - Arg221 NH1 distance , and (E,F) Asp181 OD2 - Arg221 NH2 distance.

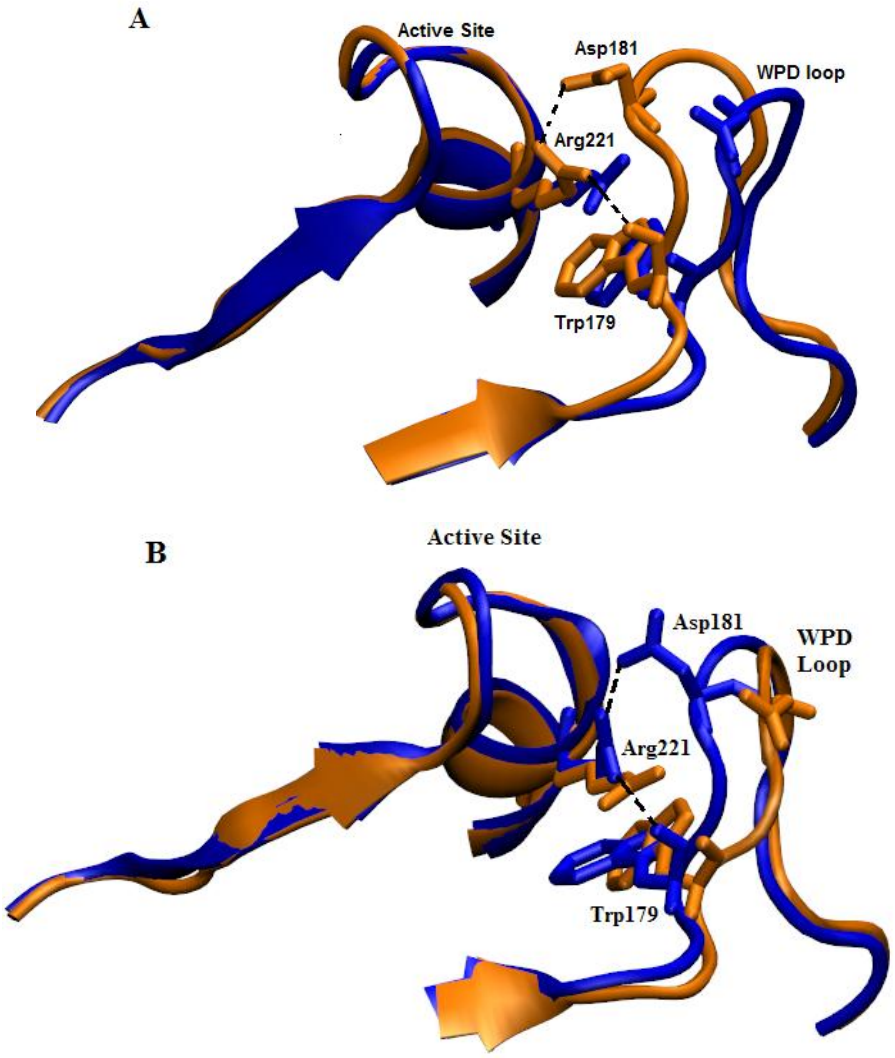

Figure 4.23. Conformation of the WPD loop and the active site (A) at 2 ns (before the conformational transition in ~2.75 ns) shown in blue, and at 3 ns (after the conformational transition) shown in orange in TMD1. (B) Conformations before (blue) and after (orange) the transition in TMD4, sampled at 2.5 ns and 3.5 ns. Dashed lines represent the polar interactions between Trp179 and Asp181 with Arg221.

The interaction between Phe182 and Thr263 forms in TMD1 and breaks in TMD4 during this step. In TMD1, Phe182 O- Thr263 OG1 distance decreases from ~6 Å to ~2.75 Å at ~2.75 ns, showing that as Trp179 and Asp181 make polar contacts with Arg221, a hydrogen bond between Phe182 O -Thr263 OG1 is formed at the same time (Figure 4.24A). The formation of this hydrogen bond between these atoms is also seen in TMD2 (Figure A.11). This hydrogen bond breaks at ~3 ns during opening of the WPD loop in TMD4 (Figure 4.24B).

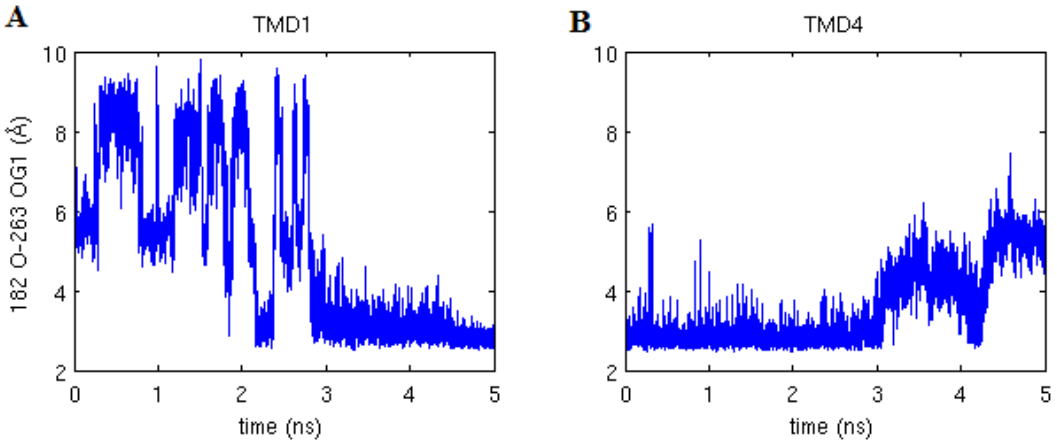

Figure 4.24. Change in the Phe182 O – Thr263 OG1 distance in (A) TMD1, and in (B) TMD4.

As stated in Section 4.3, an additional conformational transition was observed at ~3.75 ns in the TMD2 simulation. Examination of the WPD loop residues and their interactions showed that the distance between Gly183 O and Trp179 OG atoms decreased sharply from 3 Å to 2 Å at ~3.75 ns (Figure 4.25A), resulting in a hydrogen bond between these WPD loop residues. Simultaneously, the Asp181 OD1 - Ser216 OG distance decreased from 5 Å to 2.5 Å (Figure 4.25B), and the Arg221 NH1 - Ser216 OG distance decreased from 4 Å to 3 Å (Figure 4.25C). The absence of this intermediate stage in TMD1 is related to an absence of change in Ser216, which maintains its interaction with solvent, pointing away from the active site (Figure 4.26).

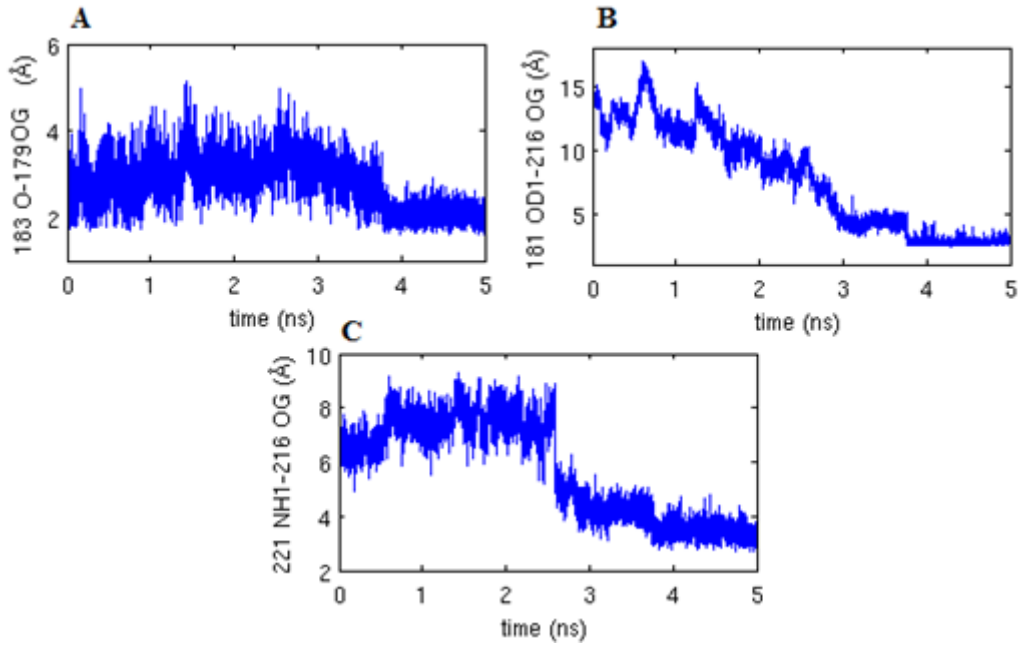

Figure 4.25. Time evolution of (A) Gly183 O – Trp179 OG distance**,** (B) Asp181 OD1 - Ser216 OG distance, and (C) Arg221 NH1-Ser216 OG in TMD2**.**

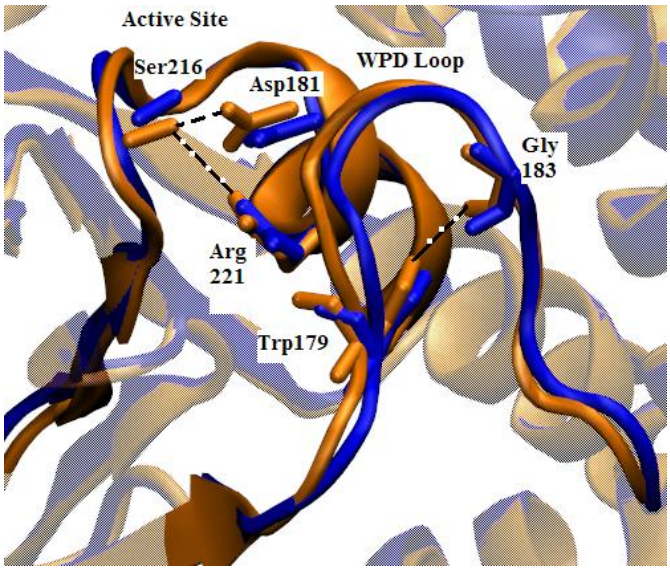

Figure 4.26. Conformation of the WPD loop and P-loop in TMD2, shown as snapshots sampled at 3 ns shown in blue, and at 4 ns shown in orange. Dashed lines represent the hydrogen bonds between Arg221 and Asp181 with Ser216, and between Trp179 and Gly183.

4.4.3. Third Conformational Transition of the WPD Loop

This conformational transition was detected as a separate cluster using k-means clustering in MATLAB only in TMD4 (Section 4.3). Inspection of the backbone dihedral angles of the tip of the WPD loop, and the interactions between WPD loop and the Q loop shows the existence of this conformational transition, which manifests itself as a small change of WPD loop structures. The third stage occurs at ~4.25-4.75 ns in TMD1, TMD2 and TMD4 simulations.

In TMD1, the Asp181 backbone Ψ angle rotates by about 100° at 4.5 ns (Figure 4.27A). This sharp change in the backbone dihedral angle of Asp181 occurs simultaneously with the decrease in the RMSD of WPD loop at 4.5 ns (Figure 4.12) and brings the WPD loop closer to the active site such a hydrogen bond is formed between Phe182 O and Gln262 OE1 at ~4.5ns (Figure 4.27C) resulting in a decrease in electrostatic energy between them (Figure 4.27E). The formation of this hydrogen bond occurring simultaneously with the rotation of Ψ of Asp181 is also seen in TMD2 (Figure A.12A-B). Gln262 is a catalytically important residue for coordinating a water molecule during the hydrolysis step, thus the Asp181 backbone dihedral angle transition, which stabilizes Gln262 in a conformation pointing to the active site pocket, may be important for catalysis (Pedersen et al., 2004). In the reverse motion of the WPD loop, the backbone Asp181 Ψ angle, does not rotate in TMD4 (Figure 4.27B). It is possibly due to lack of this conformational transition that the WPD loop does not completely open in TMD4, suggesting that the Asp181 rotation might be an energy barrier in the conformational activation mechanism (Figure 4.12). The magnitude of the applied force may not be enough to overcome the energy barrier for the rotation of Asp181-Phe182 backbone dihedral angles in TMD4. The distance between Phe182 N atom-Gln262 OE1 atom decreases from ~3.5Å to ~8Å (Figure 4.27D) and the interaction between Phe182 and Gln262 is totally lost (Figure 4.27F).

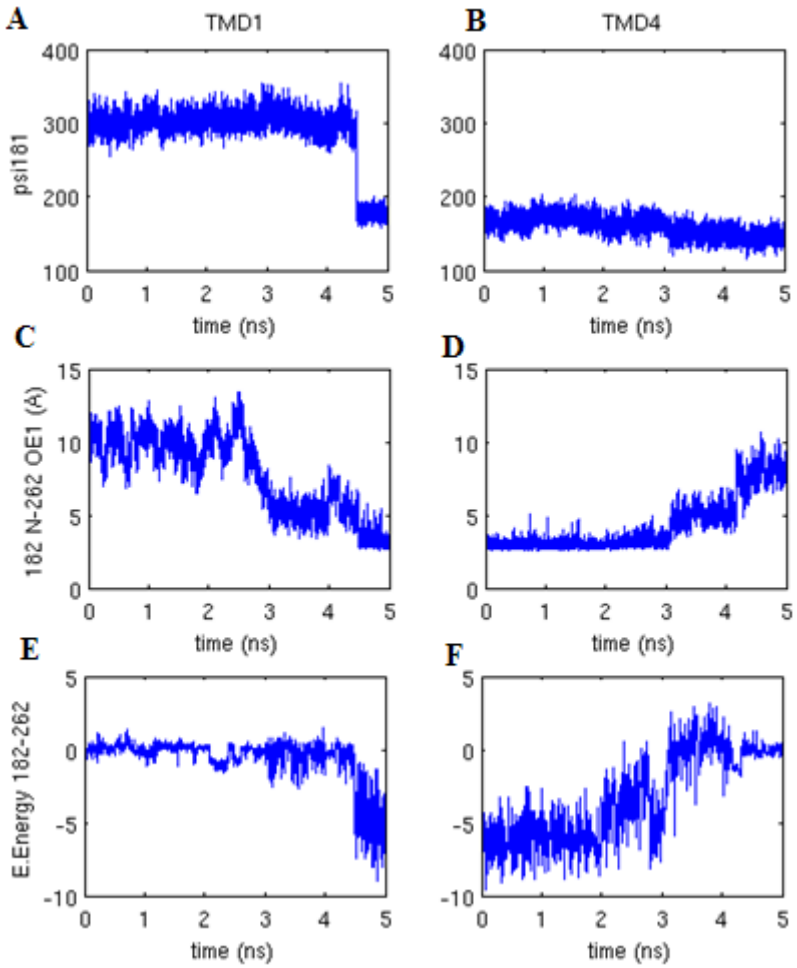

Figure 4.27. Time evolution of (A,B) Asp181 Ψ backbone dihedral angle, (C,D) the distance between Phe182 N atom-Gln262 OE1 atom and (E,F) the electrostatic energy between Phe182-Gln262 in TMD1 (left panel) and TMD4 (right panel).

When the Asp181 Ψ backbone dihedral angle changes at 4.5 ns in TMD1, Asp181 moves closer to Gly183 (Figure 4.28A) and the electrostatic energy between Gly183 and Asp181 residues decreases (Figure 4.28C). The formation of this interaction between the two residues is also observed in TMD2 (Figure A.12C). A hydrogen bond between Gly183 O atom and Trp179 HE1 atom is formed at ~4.5 ns (Figure 4.28E) simultaneously with the Asp181 Ψ transition. These interactions of Gly183 with Trp179 and Asp181 might stabilize the WPD loop in the closed state. Because Asp181 does not rotate, the interaction between Gly183 and Asp181 is not broken during the reverse motion of the WPD loop in TMD4 (Figures 4.28B-D), resulting in an incomplete opening of the WPD in TMD4 simulation. However, the Gly183 and Trp179 interaction in the WPD_{closed} state is

broken at ~4.25 ns in TMD4 (Figures 4.28F) Therefore, disruption of Gly183 interactions with both Trp179 and Asp181 may be required for the complete opening of the WPD loop.

Figure 4.29 shows the interactions formed during the third conformational transition. When the backbone Φ angle of Phe182 rotates as a result of the rotation of the Asp181 Ψ dihedral angle, a hydrogen bond between Phe182 and Gln262 forms simultaneously with the movement of the backbone of Gln262 toward the WPD loop at 4.5 ns in TMD1.

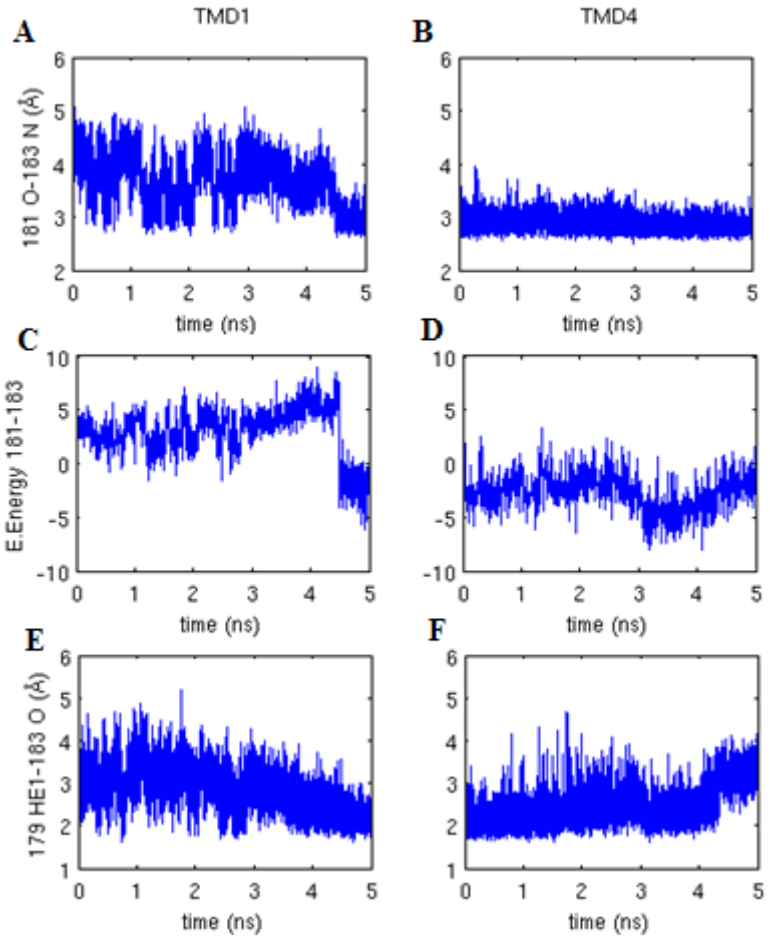

Figure 4.28. Time evolution of (A,B) the electrostatic energy between Asp181-Gly183, (C,D) Trp179 HE1-Gly183 O distance and (E,F) Asp181 O -Gly183 N distance in TMD1 (left panel) and in TMD4 (right panel).

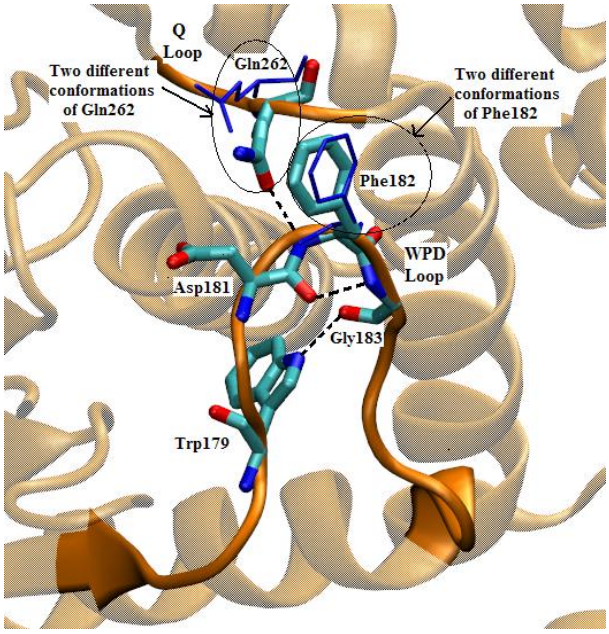

Figure 4.29. Conformation of the WPD loop and Q-loop in TMD1 sampled at 4.75 ns (after the conformational transition at 4.5 ns). The conformations of Phe182 and Gly262 in at 3 ns are shown in thin blue licorice and those at 4.75 ns are shown in thick licorice. Dashed lines represent the hydrogen bonds between Phe182-Gln262, Trp179-Gly183 and Asp181-Gly183.

4.5. Changes in Hydrophobic Interactions During WPD Loop Motion

Numerous hydrophobic interactions in the vicinity of the WPD loop stabilize the WPD loop conformation. The roles of the hydrophobic interactions between the WPD loop and the surrounding residues were examined to elucidate their contribution to the WPD loop motions. Figure 4.30 shows a simplified scheme of the hydrophobic interactions between the WPD loop and the regions around it. Some of these residues must change their conformations to avoid the steric clashes shown by arrows. For instance, the arrow between Ser201 and Phe174 shows that Phe174 changes its conformation during the TMD simulation, such that Phe174 and Ser201 would have clashed, if Phe174 had not moved out of the way for Ser201.

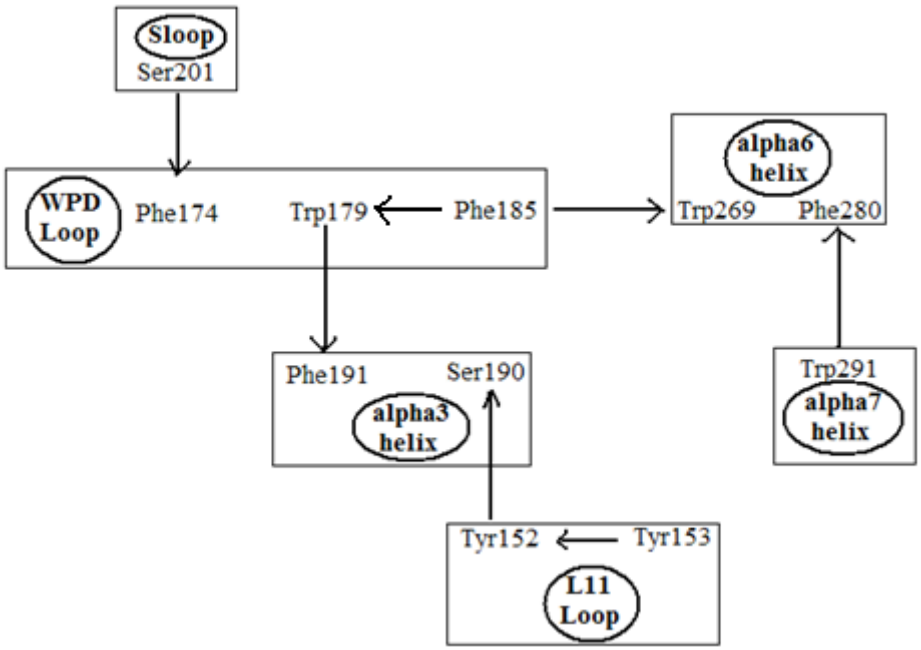

Figure 4.30. The simplified map of hydrophobic interactions around WPD loop during WPD loop closure.

4.5.1. Hydrophobic interactions between WPD Loop, α3 Helix and Loop 11

In an effort to examine the response of the surrounding structural elements to WPD loop closure, hydrophobic interactions with the α3 helix and L11 loop (residues 150-154) were monitored. All atoms of the simulation snapshots were aligned with respect to the initial structure and the distance between Phe185 of the simulation structures and Trp179 of the initial structure was measured. This gives an indication of whether the trajectory followed by Phe185 would result in a steric clash with Trp179, if Trp179 had not moved out of the way. The sum of the van der Waals radii of two carbon atoms is 3 Å, therefore two carbon atoms cannot move closer than this distance. During WPD loop closure, Phe185 moves into the pocket occupied by Trp179 in WPD_{open} conformation in the first nanosecond (before the first conformational transition) (Figure 4.31A), pushing Trp179 (Figure 4.31B). In addition, the distance between Trp179 in the TMD simulation and Phe191 in the initial structure reaches 2 Å at ~2.75 ns (the second conformational transition). It should be noted that the initial and final values of the distance are around 3 Å (Figure 4.32A), and the distance is 3 Å throughout the simulation showing that as Trp179

moves, Phe191 is pushed out of the way. In other words, Phe191 and Trp179 would have clashed, if Phe191 had not moved out of the way of Trp179 (Figure 4.32B).

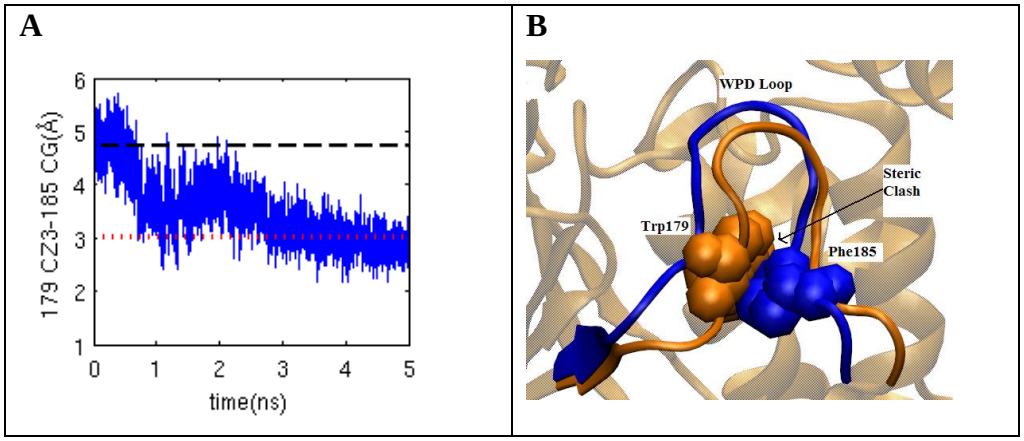

Figure 4.31. (A) Time evolution of the Trp179 CE2 - Phe185 CD distance in TMD1. Dashed line represents the initial value (average value of the first 100 snapshots) and dotted line represents the final value (average value of the last 100 snapshots). (B) The initial structure (orange) and the simulation snapshot at 2.5 ns (blue). Trp179 in the initial structure (orange spheres) and Phe185 in the snapshot at 2.5 ns (blue spheres) are in steric conflict.

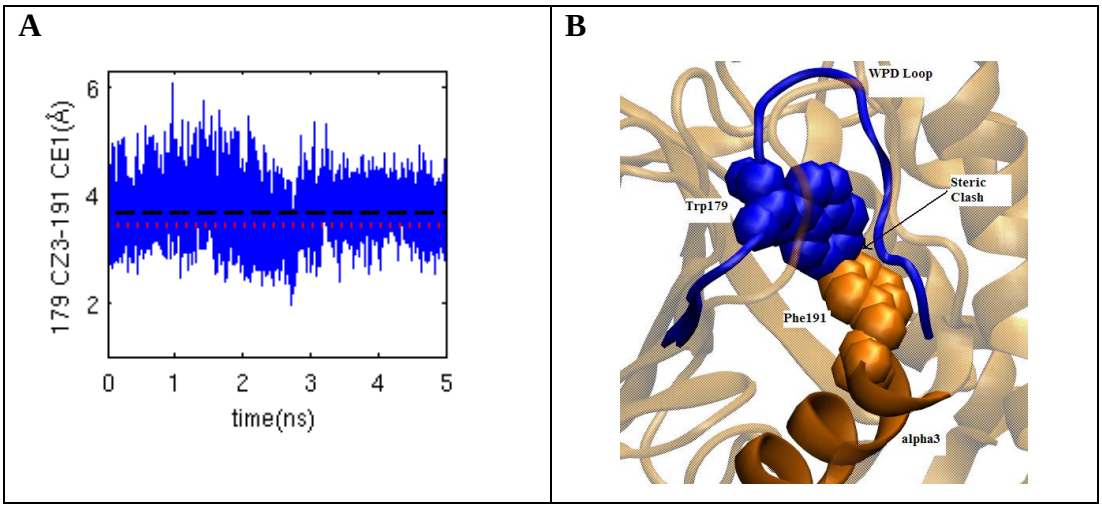

Figure 4.32. (A) Time evolution of the Trp179 CZ3 - Phe191 CE1 distance in TMD1. Dashed line represents the initial value (average value of the first 100 snapshots) and dotted line represents the final value (average value of the last 100 snapshots). (B) The initial structure (orange) and the simulation snapshot at 2.75 ns (blue). Phe191 in the initial structure (orange spheres) and Trp179 in the snapshot at 2.75 ns (blue spheres) are in steric conflict.

Hydrophobic interactions between L11 and α3 are mediated through Tyr152, Tyr153 of L11 and Ser190 of α3. Tyr152 would have clashed with Tyr153, resulting in an interatomic distance less than 3 Å at the beginning of the simulation, if Tyr152 did not move out of the way for Tyr153 (Figure 4.33A). During WPD loop closure, Tyr153 moves into the pocket occupied by Tyr152 in WPD_{open} conformation, pushing Tyr152 (Figure 4.33B). Ser190 of α3 helix and Tyr152 of L11 would have clashed if Ser190 had not changed its conformation from that in the initial structure (Figure 4.34A-B). Although the initial and final distance values are indistinguishable, there would have been steric clashes mostly in the first 1.5 ns, suggesting that L11-α3 hydrohobic interactions may be related to WPD loop C-terminus conformational transition, which occurs simultaneously. Similar results are obtained for TMD2 (Figure A.13).

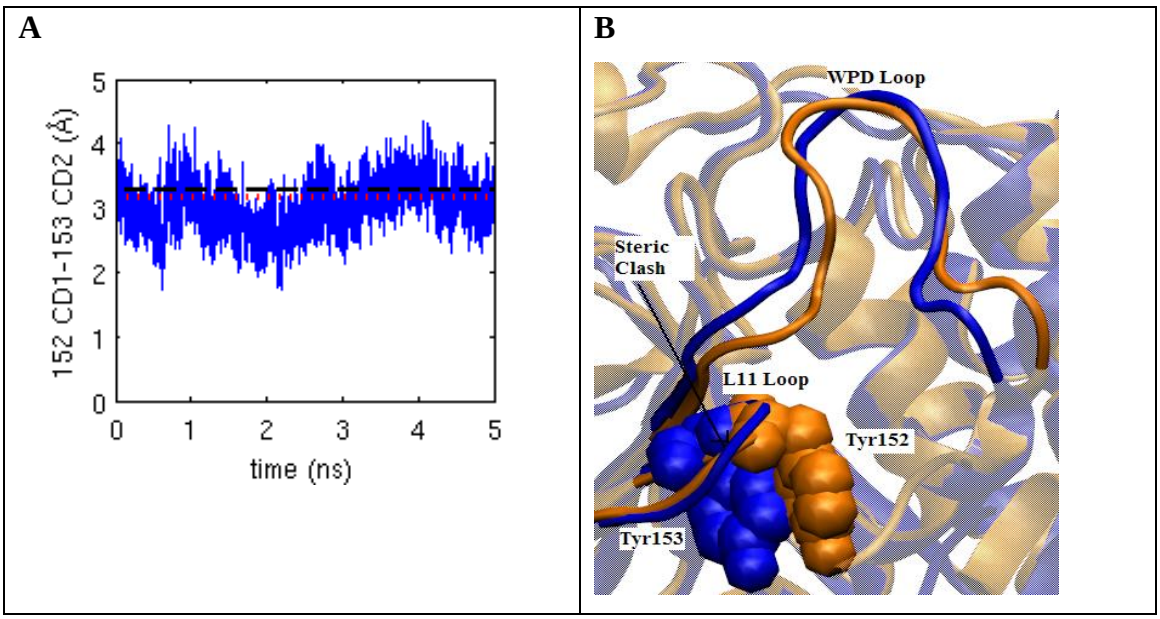

Figure 4.33. (A) Time evolution of the Tyr152 CD1 - Tyr153 CD2 distance in TMD1. Dashed line represents the initial value (average value of the first 100 snapshots) and dotted line represents the final value (average value of the last 100 snapshots). (B) The initial structure (orange) and the simulation snapshot at 2.15 ns (blue). Tyr152 in the initial structure (orange spheres) and Tyr153 in the snapshot at 2.15 ns (blue spheres) are in steric conflict.

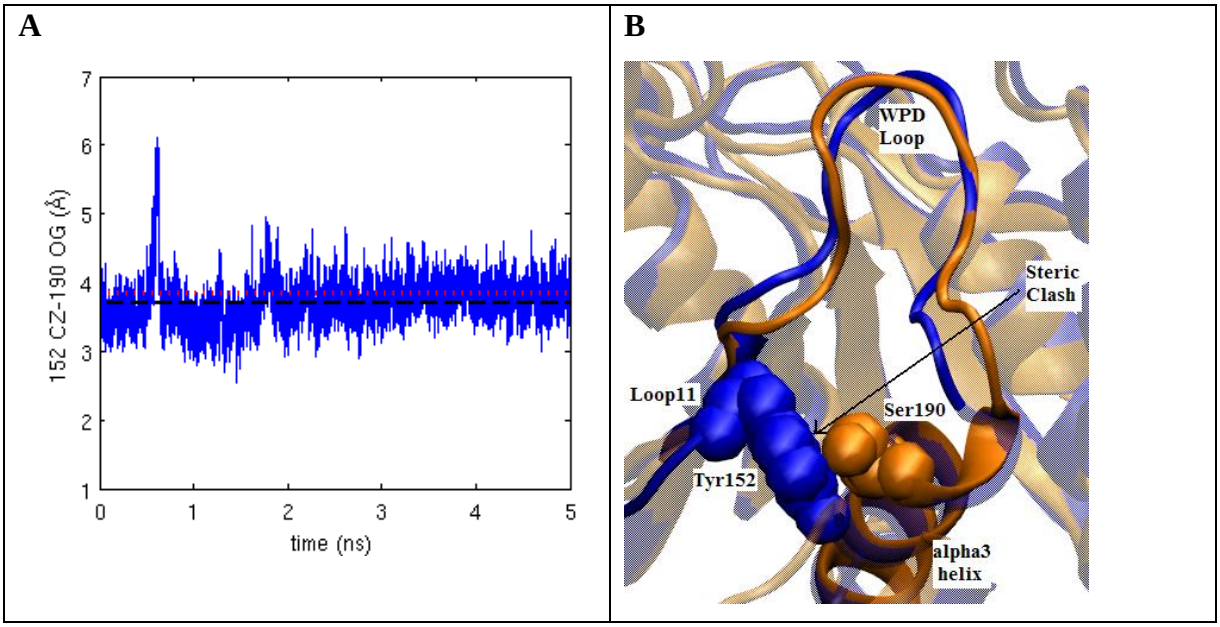

Figure 4.34. (A) Time evolution of the Ser190 O - Tyr152 CD1 distance in TMD1. Dashed line represents the initial value (average value of the first 100 snapshots) and dotted line represents the final value (average value of the last 100 snapshots). (B) The initial structure (orange) and the simulation snapshot at 1.75 ns (blue). Ser190 in the initial structure (orange spheres) and Tyr152 in the snapshot at 1.75 ns (blue spheres) are in steric conflict.

4.5.2. Hydrophobic interactions between WPD Loop, α6 and α7

How α6 and α7 helices move in response to WPD loop closure was determined by investigation of the hydrophobic interactions between WPD loop and these regions. All atoms of the simulation snapshots were aligned with respect to the initial structure and the distance between Phe269 of the first structure and Pro185 of the simulation structures was monitored. Phe269 of α6 helix and Pro185 of WPD loop would clash as Pro185 moves toward the closed conformation, resulting in an interatomic distance less than 3 Å, if Phe269 did not move out of the way for Pro185 (Figure 4.34A). During WPD loop closure, Pro185 moves into the pocket occupied by Phe269 in WPD_{open} conformation (Figure 4.34B). Pro185-Phe269 distance shows a similar behavior in TMD2 (Figure A.14B). Furthermore, it is interesting that both distance trajectories experience a steep decrease between 2-3 ns, followed by a stable distance value. This shows that displacement of Phe269 may be related with the second conformational transition, which takes place at 2-3 ns (corresponding to the second decrease in the RMSD of WPD loop).

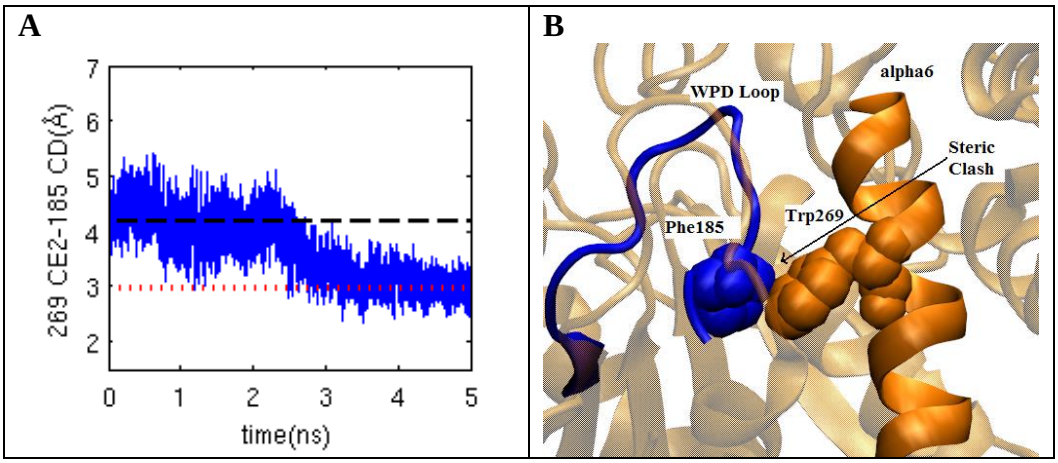

Figure 4.35. (A) Time evolution of the Phe269 CE2 - Pro185 CD distance in TMD1. Dashed line represents the initial value (average value of the first 100 snapshots) and dotted line represents the final value (average value of the last 100 snapshots). (B) The initial structure (orange) and the simulation snapshot at 3.5 ns (blue). Phe269 in the initial structure (orange spheres) and Pro185 in the snapshot at 3.5 ns (blue spheres) are in steric conflict.

Trp291 and Phe280 would also clash, resulting in an interatomic distance less than 3 Å, if Phe280 did not move out of the way for Trp291 (Figure 4.36A). During WPD loop closure, Trp291 moves into the pocket occupied by Phe280 in WPD_{open} conformation (Figure 4.36B). A similar behavior is also seen in TMD2 (Figure A.14D).

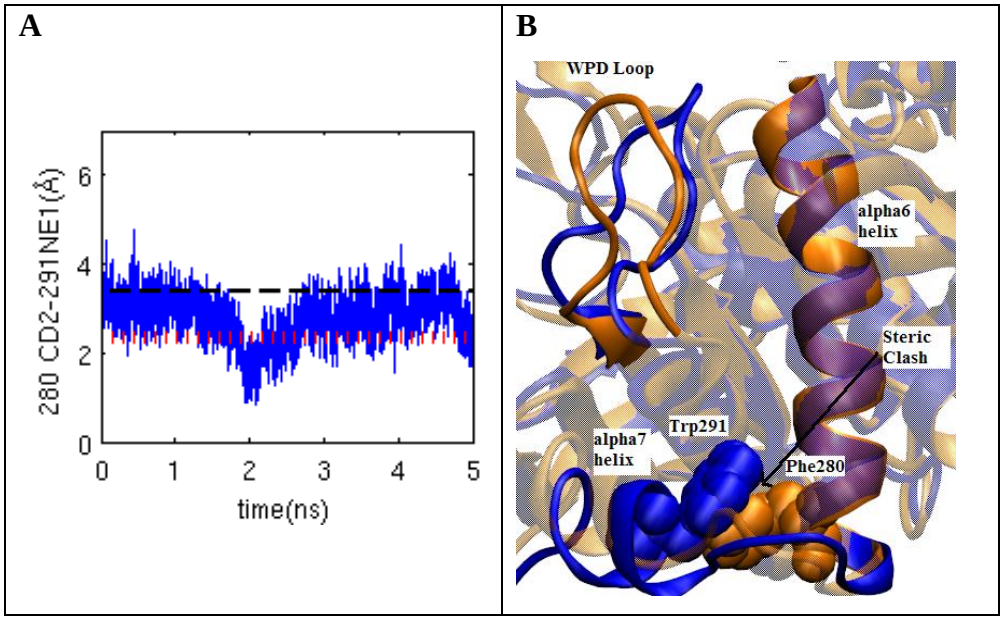

Figure 4.36. (A) Time evolution of the Trp291 NE1 - Phe280 CD2 distance in TMD1. Dashed line represents the initial value (average value of the first 100 snapshots) and dotted line represents the final value (average value of the last 100 snapshots). (B) The initial structure (orange) and the simulation snapshot at 2 ns (blue). Phe280 in the initial structure (orange spheres) and Trp291 in the snapshot at 2 ns (blue spheres) are in steric conflict.

4.5.3. Hydrophobic Interactions between WPD Loop and S Loop

All atoms of the simulation snapshots were aligned with respect to the initial structure and the distance between Phe174 of the first structure and Ser201 of the simulation structure was measured. Phe174 of WPD loop and Ser201 of S loop would clash, resulting in an interatomic distance less than 3 Å, if Phe174 did not move out of the way for Ser201 (Figure 4.37A). As WPD loop closes due to the targeting potential in TMD1, the S loop and especially Ser201 undergoes a conformational change, as a result Ser201 moves into the position occupied by Phe174 in WPD_{open} conformation (Figure 4.37B). The steep decrease seen at ~2-3 ns shows a possible relation to the second conformational transition of the WPD loop. However, a similar decrease is seen during the first 1 ns of TMD2 (Figure A.15B), therefore it can be said that the displacement of S loop is not coincident with any of the previously identified conformational transition steps of the WPD loop.

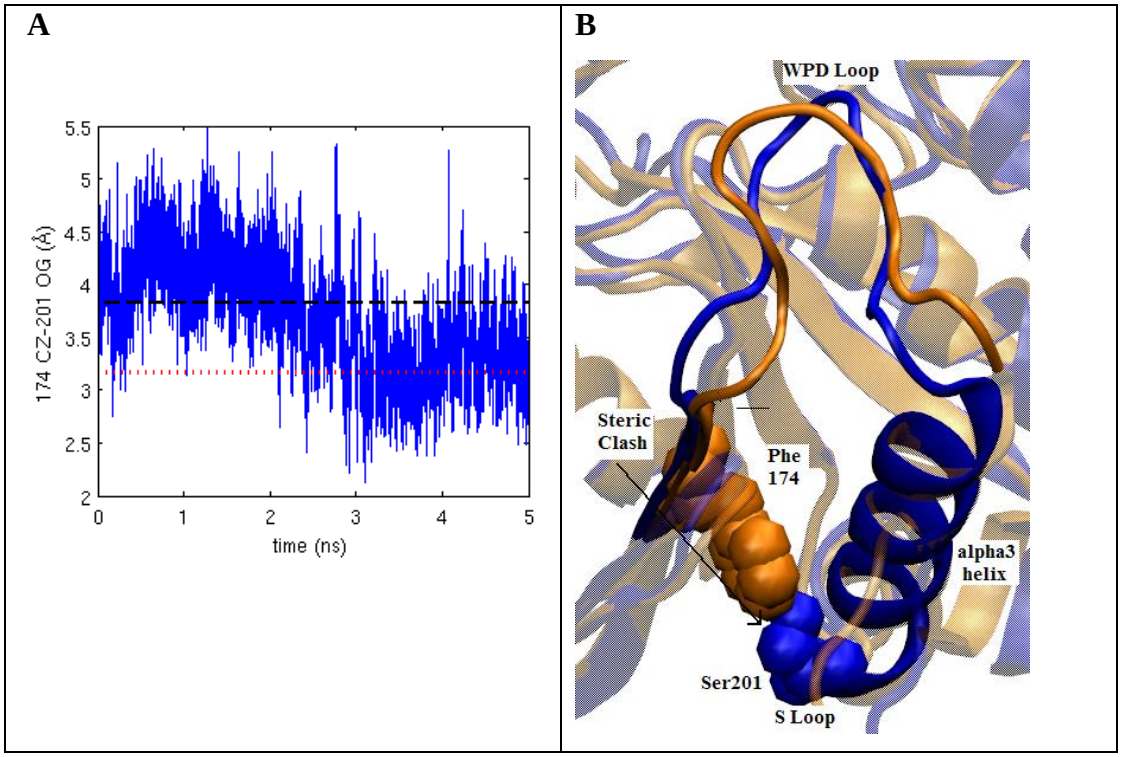

Figure 4.37. (A) Time evolution of the Phe174 CZ - Ser201 OG distance in TMD1. Dashed line represents the initial value (average value of the first 100 snapshots) and dotted line represents the final value (average value of the last 100 snapshots). (B) The initial structure (orange) and the simulation snapshot at 3.5 ns (blue). Phe174 in the initial structure (orange spheres) and Ser201 in the snapshot at 3.5 ns (blue spheres) are in steric conflict.

4.6. Movement of α6 and α7 Helices During the WPD Loop Closure

The α7 helix is proposed to be a regulatory helix stabilizing the closure of the WPD loop and is needed for full enzyme activity. The conformational flexibility in the helices α7, α3 and α6 might be important for accommodating and stabilizing the various conformers of the WPD loop [22]. Crystal structures of PTP1B with WPD loop in the closed conformation were obtained in the presence of the pTyr containing hexapeptide, and these structures showed interactions between α7, α3 and α6 [8]. Allosteric inhibitors prevent this interaction and the closing motion of the WPD loop. This reveals a relationship between α7 and the WPD loop [22]. This observation is consistent with the available structures in the Protein Data Bank [44], we have found that 25 of 26 structures

with the WPD loop in the open conformation do not have α7 helix coordinates and 45 of 47 structures with the WPD loop in the closed conformation have an ordered α7 helix. The structure with the WPD loop in open state as well as α7 coordinates is a complex of PTP1B with the FOMT (4'-*O*-[2-(2-fluoromalonyl)]-Ltyrosine) peptide inhibitor (PDB ID: 1BZH, [47]), in which α7 helix is in an alternative conformation. The two structures with the WPD loop in the closed state but with not α7 coordinates belong to PTP1B in complex with an unsaturated IZD-phenol potent inhibitor (PDB ID: 2CNE; [21]) and PTP1B in complex with monoacid-based inhibitors (PDB ID: 1QXK; [48]).

The motions of the regions around WPD loop such as α6 and α7 helices might provide important clues about the mechanism of WPD loop closure. The direction of motion for these helices was investigated during closing of the WPD loop in TMD1.

PCA was applied separately to trajectories of the α6 and α7 helix in TMD1 after aligning the Cα atoms of the whole protein with respect to its initial structure in order to determine the major directions of motions of these helices during WPD loop closure. First principal component (PC1) of the α6 helix trajectory explains 34% and the second PC explains 16% of α6 helix movement. Along PC1, α6 helix fluctuates toward the WPD loop (Figure 4.38A), in which it firstly moves away from the WPD loop in the first ~0.5 ns, and then comes back to its initial position at ~3 ns (Figure 4.39A) and along PC2, α6 helix moves along a direction between its N- and C-terminus laterally (Figure 4.38B). These two dominant α6 helix motion are also observed in TMD2 (Figures A.16 and A.17). Projection on PC2, shows that α6 helix fluctuates throughout the whole trajectory (Figure 4.39B). PC1 of α6 helix seems to describe a motion correlated with the conformational transitions in WPD loop because the first movement of α6 helix occuring at ~0.5 ns is coincident with the first transition of conformational change of WPD loop and the second movement occurring at ~3 ns is synchronized with the second transition.

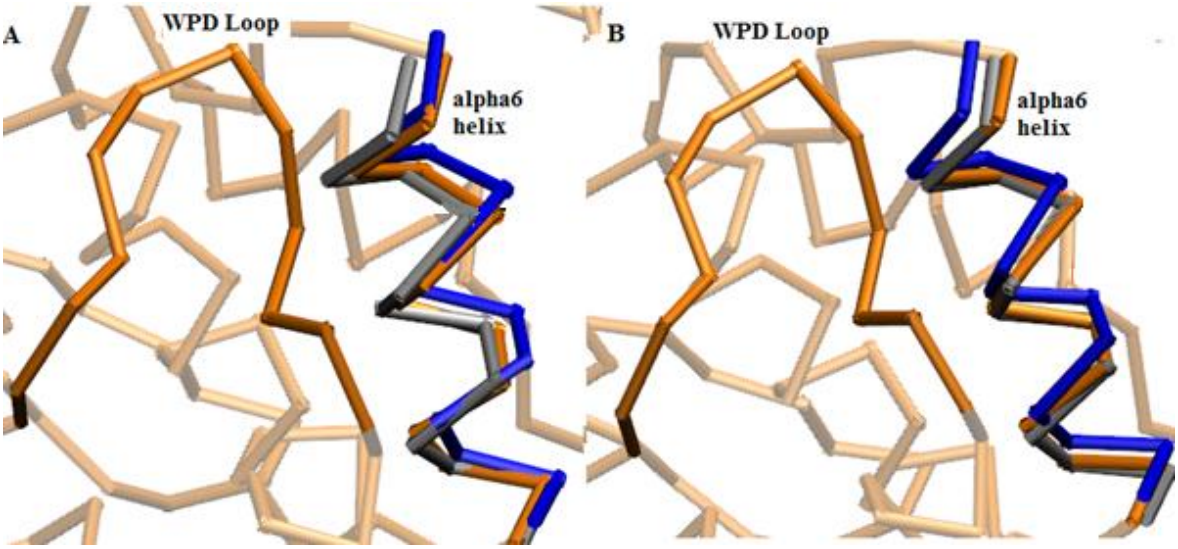

Figure 4.38. The representation of the movement of α6 helix explained by (A) PC1 and (B) PC2 on the protein structure in TMD1. The target structure (orange), the first snapshot along the projection on the PCs (blue) and the final snapshot on the PCs (grey) are shown.

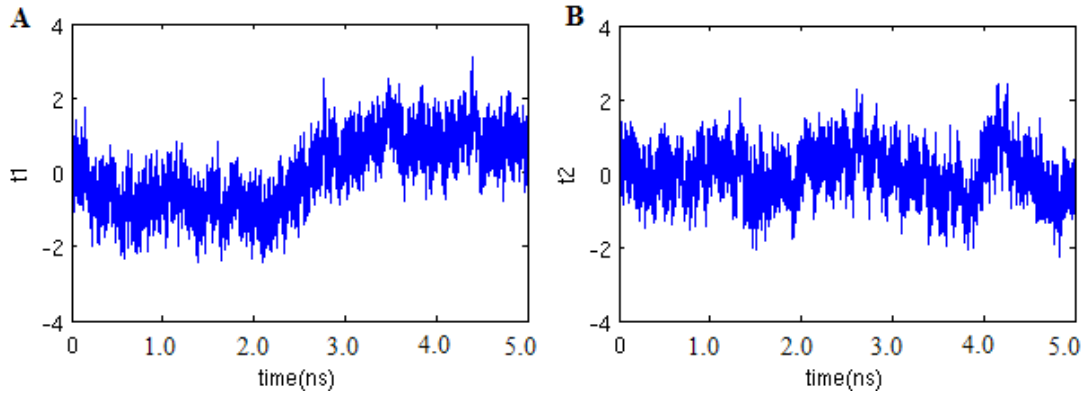

Figure 4.39. The projection of the α6 helix motions on (A) PC1 (B) and PC2 in TMD1.

PCA on α7 helix has shown that the first principal component (PC1) α7 of explains 47% and the second PC explains 20% of α7 helix movement. On PC1, α7 helix fluctuates toward the WPD loop (Figure 4.40A), and PC2 shows a longitudinal type of motion, both of which are similar to α6 motions (Figure 4.40B). These fluctuation directions are identical in TMD2 simulation (Figure A.18). The projection of α7 helix on PC1 has shown that α7 helix does not move until ~1.5 ns, then it moves toward the WPD loop, with large oscillations (Figure 4.41A). However, Figure A.19A shows that projection of α7 on PC1 of TMD2 does not show a similar pattern. In TMD1, projection of α7 on PC2 shows several

large amplitude fluctuation(Figure 4.41B), while there is one sudden increase in those of TMD2 (Figure A.19B) These results show that although motions of α7 do not seem to be synchronized with major conformational transitions of the WPD loop (decreases in the RMSD of WPD loop), fluctuations of α7 occur in the same conformational subspace, and high flexibility of α7 may be required for WPD loop motions.

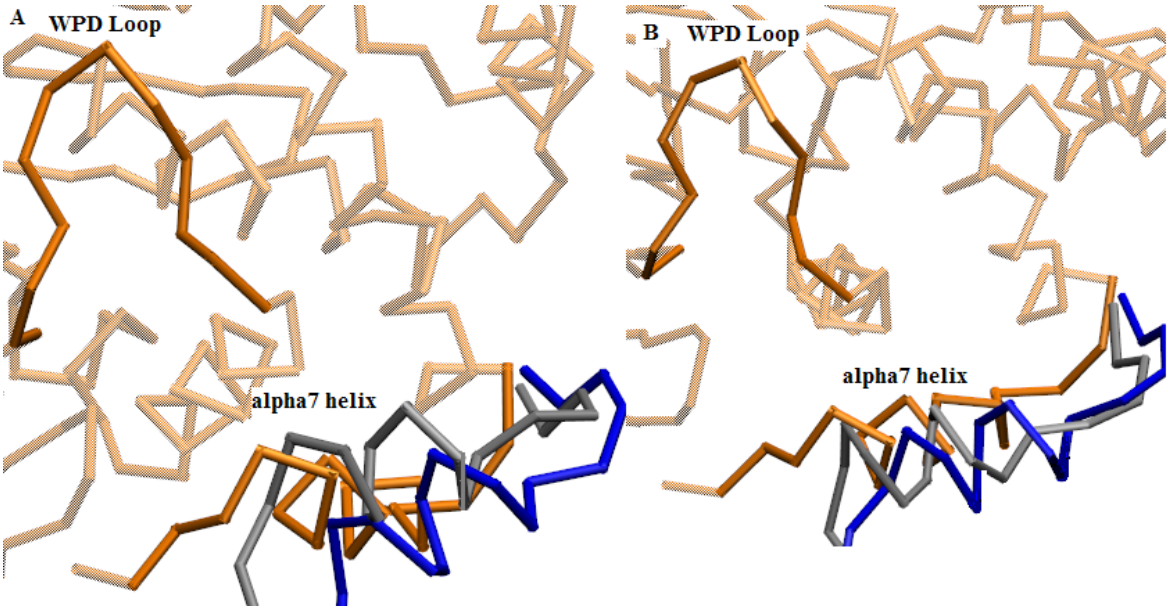

Figure 4.40. Projection along (A) PC1 and (B) PC2 of α7 helix motion in TMD1. The target structure (orange), the first snapshot along the PCs (blue) and the final snapshot along the PCs (orange) are shown. The arrows point from negative to positive directions along the PCs.

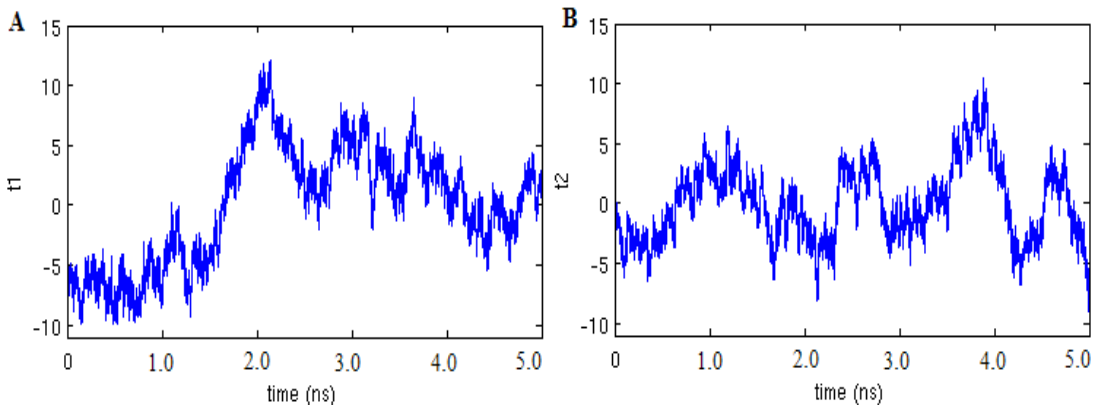

Figure 4.41. The projection of the α7 helix motions on (A) PC1 (B) and PC2 in TMD1.

4.7. Effect of the Conformational Change of Arg221 on WPD Loop Closure

In TMD3, initial structure was another representative snapshot from the equilibrium MD simulation on the WPD_{open} conformation ($2F6F_3$) and the target structure was the same as that used in TMD1 ($1SUG_1$). The force constant and simulation length were identical to those of TMD1 and TMD2. TMD1 and TMD3 were compared to determine whether similar intermediate states were visited in these simulations. There is no considerable change in the RMSD of all atoms to the target conformation in TMD3 (Figure 4.42B), similar to that seen in TMD1 (Figure 4.42A), however the overall initial and final structures are much closer to each other.

In TMD3, the RMSD of the WPD loop to the target conformation decreases from 4 Å to ~2.8 Å within 2 ns, then it remains at ~2.5 Å until the end of the simulation, indicating that WPD loop does not reach its target closed conformation (Figure 4.42D). This is an unexpected result, since WPD loop coordinates are targeted toward the closed state by a targeting potential and the WPD loop reached the target conformation in both TMD1 and TMD2, which differ only by their initial coordinates.

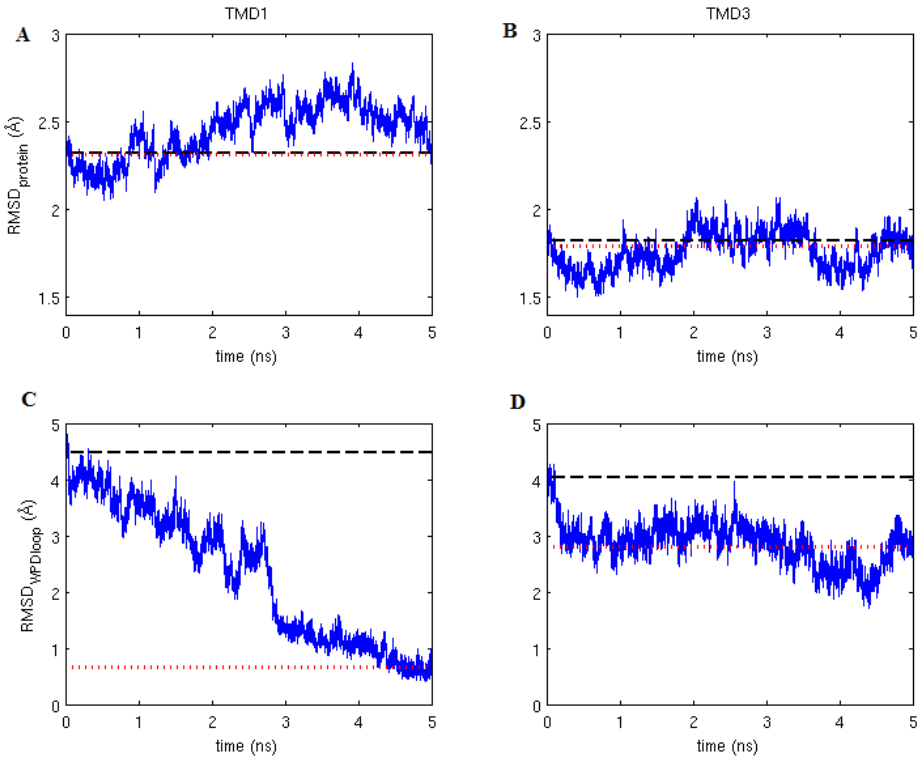

Figure 4.42. RMSD of (A,B) all protein and (C, D) WPD loop atoms only relative to the target structure during TMD1 (first column) and TMD3 (second column) simulations (Dashed line represents the initial value, dotted line represents the final value).

In order to determine why WPD loop does not close during TMD3, the conformational transition of the WPD loop and its interactions with surrounding structural elements were investigated. The interactions formed by Trp179-Arg221 and Asp181-Arg221 were monitored to determine whether similar intermediate states were visited during the second stage of WPD loop closure in TMD1 and TMD3 simulations. In TMD1, when χ2 of Trp179 side chain has rotated by 80°-90° (Figure 4.43A) and χ2 of Arg221 side chain rotates, the Trp179 O- Arg221 NH1 distance decreases from 5 Å to 2.5 Å at ~ 2.5 ns (Figure 4.43C). Similarly, In TMD3, when χ2 of Trp179 dihedral angle has rotated by 80°-90° (Figure 4.43B), the distance between Trp179 O - Arg221 NH1 distance decreases from 5.5 Å to 2.5 Å to form an interaction (Figure 4.43D). Note that, the hydrogen bond between Trp179 and Arg221 is formed at ~ 2.5 ns of TMD1, while it is formed earlier at ~ 1.75 ns of TMD3. On the other hand, while the distance between Asp181 OD1- Arg221 NH1 decreases from 7 Å to ~2.5Å at 2 ns in TMD1 (Figure 4.43E), the hydrogen bond between Asp181 OD1 and Arg221 NH1 does not form until 3.75 ns (Figure 4.43F).

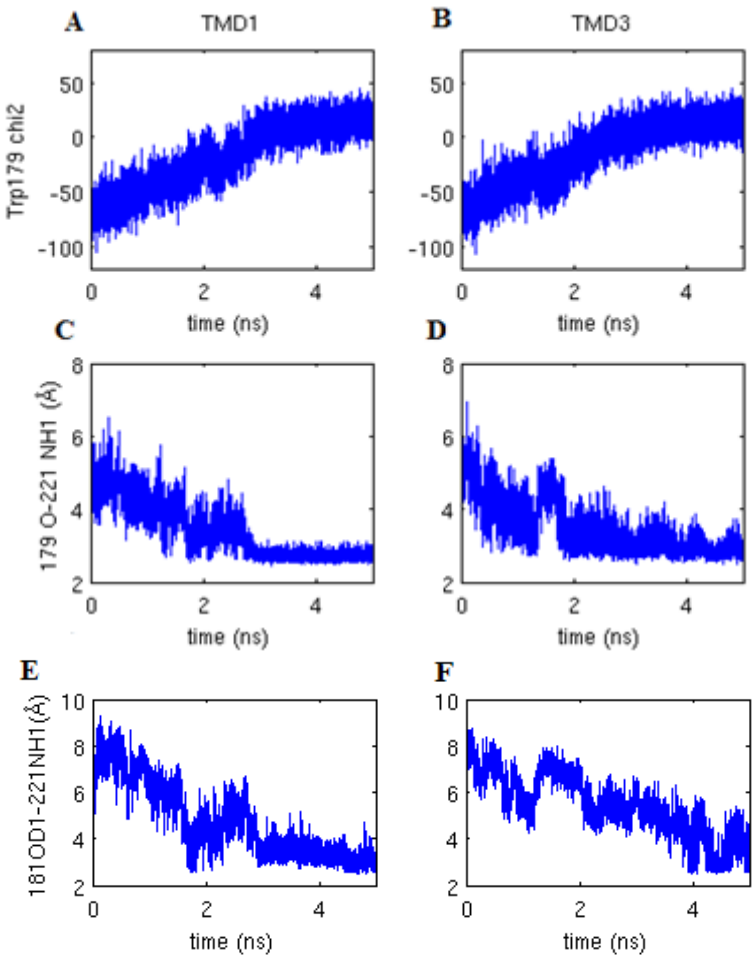

Figure 4.43. The changes in (A), (B) χ2 of Trp179, (C), (D) Trp179 O - Arg221 NH1 distance and (E), (F) Asp181 OD2 - Arg221 NH2 distance during TMD1 (first column) and TMD3 (second column) simulations.

A detailed visual examination of the trajectory has shown us that the backbone oxygen atom of Asp181 and the side chain nitrogen atom of Arg221 make a hydrogen bond from the beginning of the simulation to ~3.75 ns in TMD3 (Figure 4.44B), while this bond does not exist in TMD1 (Figure 4.44A). It can be seen that the χ3 of Arg221 side chain does not rotate in TMD3 (Figure 4.44D) whereas it rotates by 100° at ~2.5 ns in TMD1 (Figure 4.44C). Moreover, in TMD2, although Arg221 does not rotate to the desired position (Figure A.10H), it can form hydrogen bonds with Asp181 and Trp179 allowing the WPD loop to close (Figure A.10D-F). In TMD3, Arg221 residue obtains a new alternative position at ~3.75 ns, resulting in breakage of its interaction with the backbone oxygen of Asp181 (Figure 4.44B). However, a hydrogen bond between

Arg221NH1- Leu110 O is formed (Figure 4.44F). After this conformational change of Arg221 at 3.75 ns, it is kept in this new location as a result of the hydrogen bonds of Leu110 O and Trp179 O with Arg221 NH1, preventing WPD loop closure (Figure 4.45). In addition, the early formations of the hydrogen bonds between Asp181O-Arg221NH1 and Trp179O-Arg221NH1 hinder the movement of Asp181 for WPD loop to close.

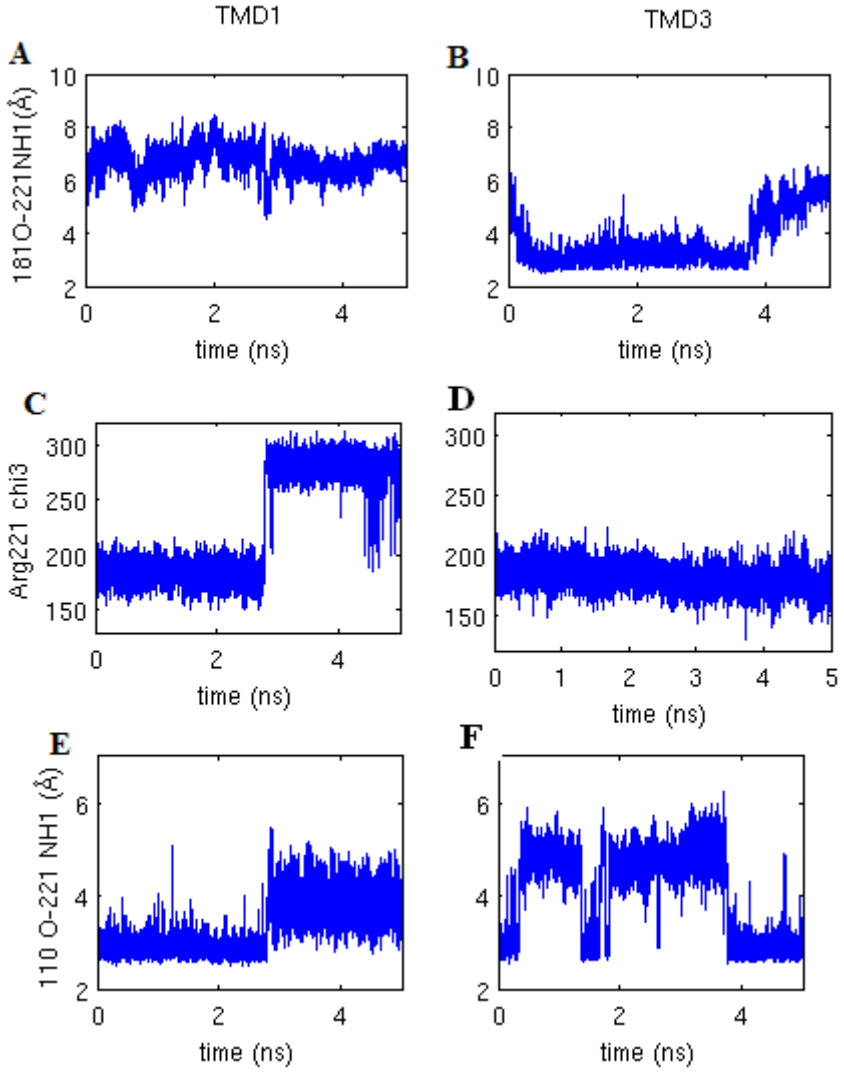

Figure 4.44. The changes in (A), (B) Asp181 O - Arg221 NH1, (C), (D) Arg221 $\chi 3$ dihedral angle and (E), (F) Leu110 O - Arg221 NH1 distance in TMD1 and TMD3.

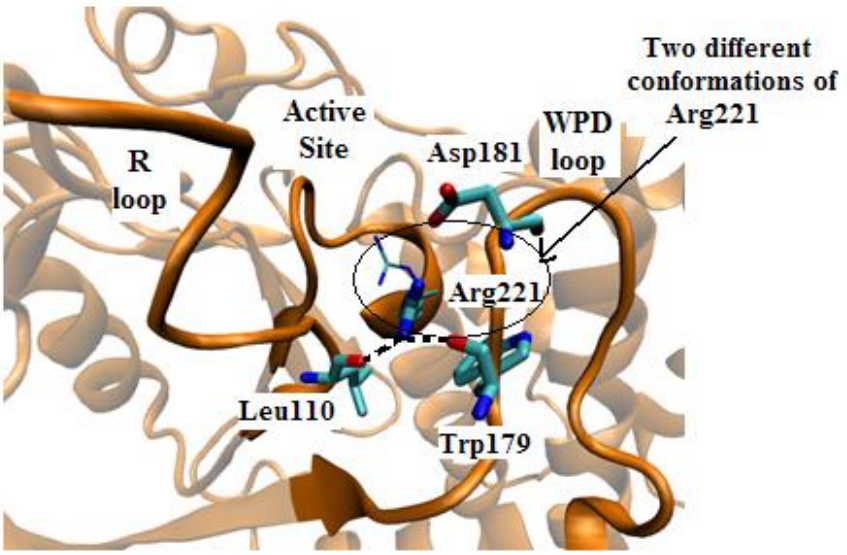

Figure 4.45. Conformation of WPD loop, active site and Leu110 residue in a TMD3 snapshot at 4 ns. Dashed lines represent the hydrogen bonds between Trp179 and Leu110 with Arg221. The conformation of Arg221 in the target structure is shown in thin licorice and that in TMD3 is shown in bold licorice.

Ψ dihedral angle transition of Asp181 and the interactions formed by Phe182-Gln262 and Asp181-Gly183 were investigated to determine whether similar intermediate states were visited during the third stage of WPD loop closure in TMD1 and TMD3 simulations. Ψ backbone dihedral angle of Asp181 rotates about 100° at 4.75 ns in TMD3 (Figure 4.46B). The hydrogen bond between Asp181-Gly183 bond is formed (Figure 4.46F) but the hydrogen bond does not form between Phe182 O atom and Gln262 OE1 atom at ~4.75ns (Figure 4.46D). This shows that the rotation of Asp181 backbone leads to the interaction between Asp181 and Gly183, but this rotation may not be enough to allow formation of the hydrogen bond between Phe182 and Gln262, since WPD loop is away from the target conformation in TMD3.

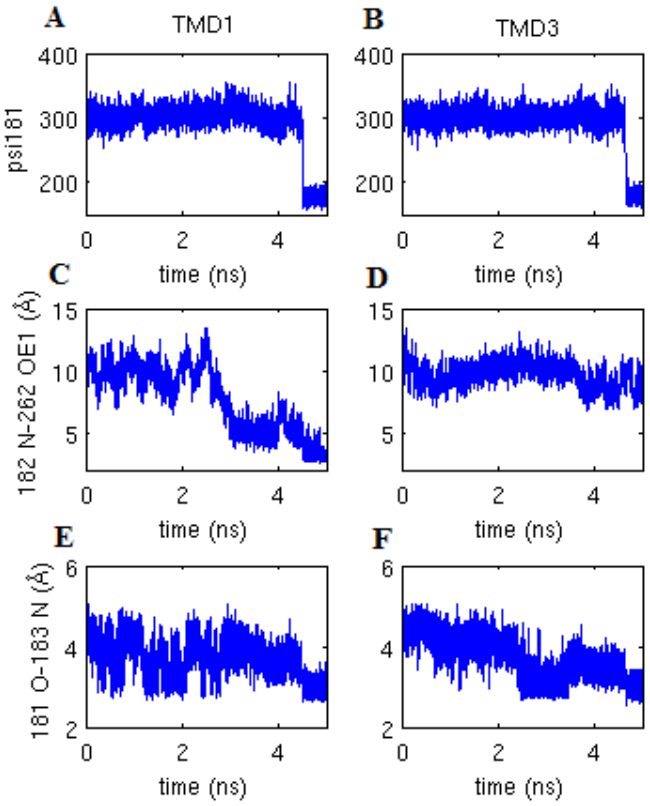

Figure 4.46. Time evolution of (A), (B) Ψ backbone dihedral angle of Asp181, (C), (D) Phe182 N -Gln262 OE1 distance (E), (F) and Asp181 O -Gly183 N distance in TMD1 (left panel) and TMD3 (right panel).

4.8. Effect of R Loop Conformation on WPD Loop Closure

An additional TMD simulation (TMD5) with another initial structure ($2F6F_4$) ,in which the R loop is closer to the active site compared to that in the TMD1, was performed. This conformation of R loop in $2F6F_4$ structure is generally seen in all PTP1B crystal structures, except PTP1B complexed with first/second generation compounds (PDB ID: 1Q6J , 1Q6M) [29]. In this simulation, WPD loop closure han been hindered by the closed conformation of R loop. Arg221 residue has its conformation in WPD_{closed} structure. The WPD loop did not close completely and the RMSD of the WPD loop to the target conformation decreased from 5.5 Å to 3.5 Å within 2 ns, then it remained at ~3.5 Å until the end of the simulation. RMSD of the whole protein to the target conformation shows a slight decrease of ~0.1 Å in TMD5 (Figure B.1B). In our study, we also applied force with

force constant 500 kcal/mol/Å^2 and 800 kcal/mol/Å^2 to all atoms of the WPD loop in two other TMD simulations. We observed that the WPD loop could not be closed in these simulations and finally it can only be closed via the force (5000 kcal/mol/Å^2) applied to the whole protein. This suggests that the closed conformation of the R loop is an important energy barrier for WPD loop to close.

In order to determine why the WPD loop does not close during TMD5, the WPD loop conformational transitions stages, which were previously identified, were investigated. χ2 of Trp179 side chain rotates by 80°-90° (Figure 4.47B), similar to that seen in TMD1, but unlike TMD1, Arg221 side chain cannot form hydrogen bonds with Trp179 and Asp181 during TMD5 simulation (Figure 4.47D and Figure 4.47F). The reason for the absence of these interactions is that Trp179 and Asp181 cannot move toward their conformations in the target structure due to the steric hinderance of the R loop, which does not allow Asp181 to move closer to the active site.

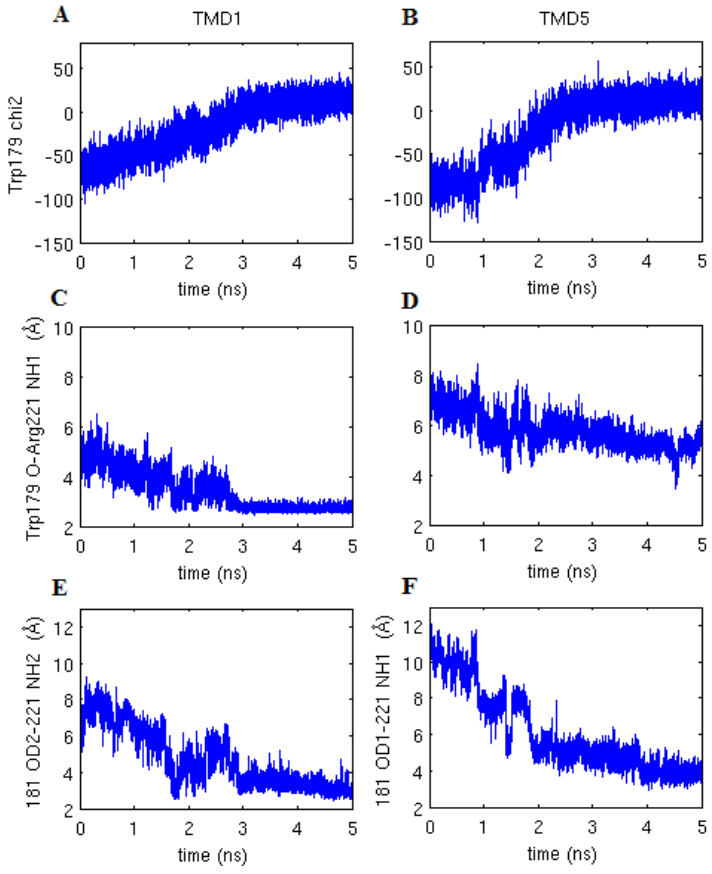

Figure 4.47. The changes in (A), (B) χ2 of Trp179, (C), (D) Trp179 O - Arg221 NH1 distance, (E), (F) and Asp181 (OD1) OD2 - Arg221 (NH1) NH2 distance.

Ψ dihedral angle transition of Asp181 and the interactions formed by Phe182-Gln262 and Asp181-Gly183 were investigated to determine whether similar intermediate states were visited during the third stage of WPD loop closure in TMD5 simulations. The Ψ backbone dihedral angle of Asp181 rotates about 100° (Figure 4.48B) and the hydrogen bond between Asp181 and Gly183 bond forms at ~4.75ns (Figure 4.48F), but the hydrogen bond between Phe182 O and Gln262 OE1 does not form (Figure 4.48D). The incomplete WPD loop closure in the absence of this interaction may suggest that this interaction is an important energy barrier in WPD loop motion. Eventhough Asp181 Ψ dihedral angle rotates in TMD5, visual examination of Asp181 motion shows that it does not move toward its target conformation. This may be the reason for the absence of 182-262 interaction in TMD5 in which WPD loop does not close. The formation of this interaction may be a necessary step toward WPD loop closure.

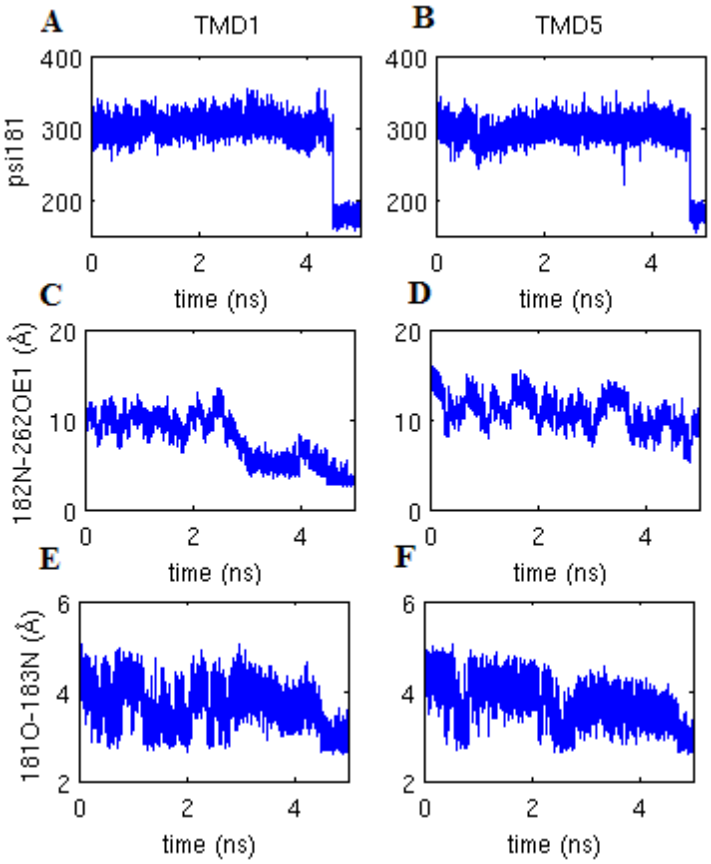

Figure 4.48. Time evolution of (A), (B) Ψ backbone dihedral angle of Asp181, (C), (D) Phe182 N -Gln262 OE1 distance and (E), (F) Asp181 O -Gly183 N distance in TMD1 (left panel) and TMD5 (right panel).

4.8.1. The Interactions of the R loop with the WPD Loop and the Active Site

In the initial structure of the TMD5 simulation, R loop is closer to the active site in the initial structure of the TMD5 simulation compared to that in the TMD1 simulation. To elucidate the relationship between the WPD loop motions and the R loop conformation, the interactions of the WPD loop and the active site with the R loop were compared in TMD1 and TMD5.

The interaction between the R loop residue Glu115 OE2 and Arg221 NH1 is almost never formed in TMD1 (Figure 4.49A), whereas it is never lost in TMD5 simulation (Figure 4.49B). This shows that the interaction between the R loop and the active site is maintained in TMD5, and R loop does not move away from its initial conformation.

As the WPD loop moves toward the closed conformation in TMD1, the interactions of Glu115 and Arg112 with Asp 181 of the WPD loop are lost at ~2.75 ns (Figure 4.49C, E), showing that the interaction between R-loop and WPD is lost. However in TMD5, the interaction between Glu115 and Asp181 forms at 2 ns (Figure 4.49D), and the hydrogen bond between Arg112 and Asp181 is formed at 1 ns, and they are maintained for the rest of the simulation. It is possible that the R loop, which maintains its orientation in TMD5, prevents WPD loop closure. Indeed, in TMD1, in which the WPD loop closes, change in RMSD for all atoms of the R loop with respect to the initial structure shows that the R loop moves away from its initial conformation, (Figure 4.49G) while it maintains its initial conformation in TMD5 (Figure 4.49H).

The conformations of Trp179-Asp181-Glu115-Arg221 are shown in a TMD5 snapshot at 3 ns in Figure 4.50. The interactions of Trp179 and Asp181 with Arg221 can not form, since the interaction between Glu115 and Arg221 is never lost and Glu115 hinders the conformational change of Asp181. The conformation of Glu115 in this TMD5 snapshot is different from that in a TMD1 snapshot, where Glu115 does not interact with the active site (Figure 4.50).

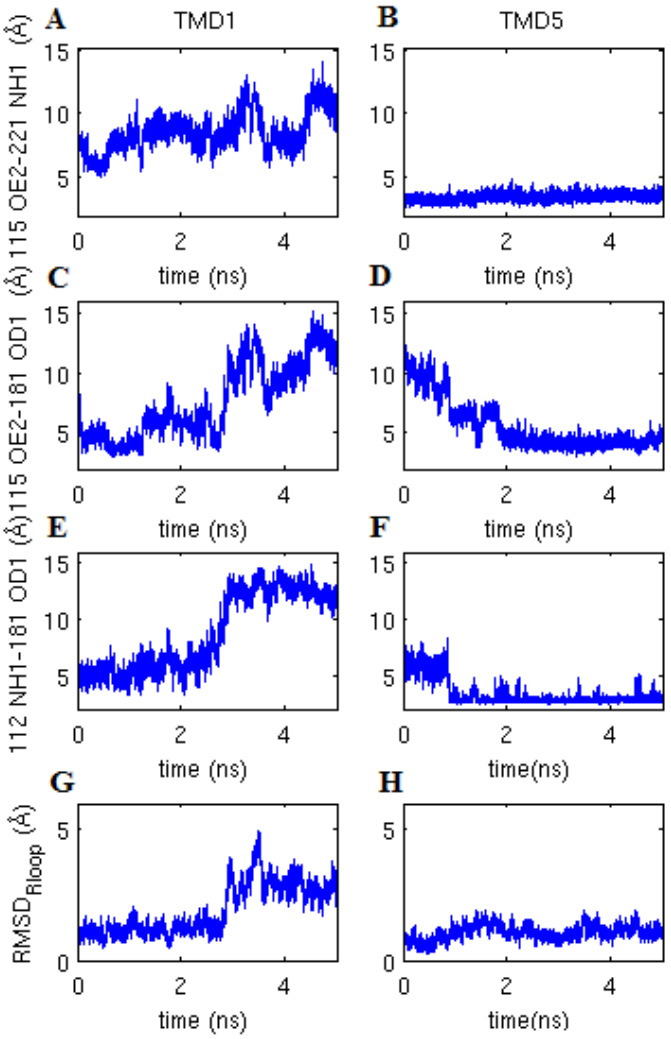

Figure 4.49. Time evolution of (A), (B) Glu115 OE2 - Arg221 NH1distance, (C), (D) Glu115 OE2 - Asp181 OD1 distance, (E), (F) Asp181 OD1 - Arg112 NH1distance, (G), (H) and the time profile of the RMSD for all atoms of the R loop relative to its initial conformation in TMD1 (left panel) and TMD5 (right panel).

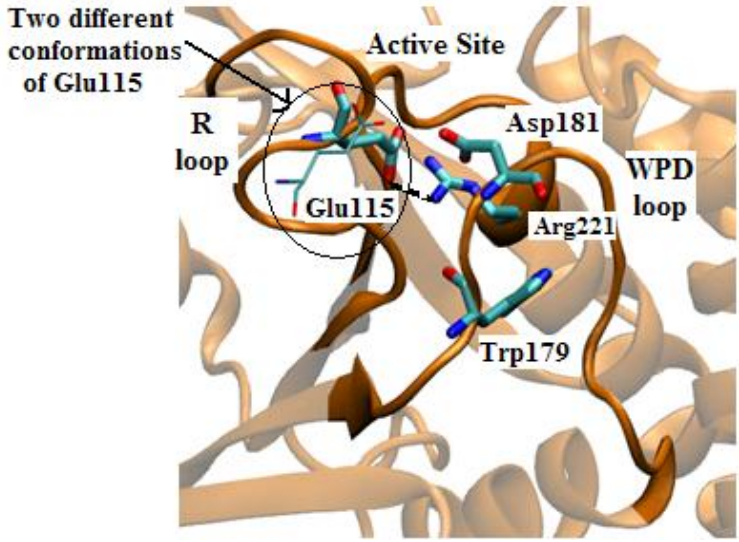

Figure 4.50. Conformation of the WPD loop, active site and R loop in a TMD5 snapshot at 3 ns. The conformation of Glu115 in a TMD1 snapshot at 3 ns is shown in thin licorice and that in a TMD5 is shown in bold licorice. Dashed lines represent the hydrogen bonds between Glu115 and Arg221.

4.8.2. R Loop Conformation in PTP Family

The results in Section 4.8.1 suggest that the conformation of the R loop and its interaction with the active site may be important for the WPD loop closure. To examine the conformations of the R loop samples in different PTP1B structures, the R-loop in 81 different PTP1B structures was compared by calculating the RMSD of the R loop Cα atoms to the 1SUG crystal structure, after aligning the C_α atoms.(Figure 4.51). In the 1SUG crystal structure, the R loop is close to the active site. In 78 crystal structures of PTP1B, R loop is in this conformation, with an average RMSD of ~0.5Å from the 1SUG conformation, while in three crystal structures, namely TCPTP (PDB ID: 1L8K) [25], PTP1B complexed with first/second generation compounds (PDB ID: 1Q6J , 1Q6M) [29], the R loop is in an alternative “open” conformation (Figure 4.52). 1L8K crystal structure of TCPTP is included in this comparison due to its high sequence identity with PTP1B. As it is discussed in Section 2.5.4, 1L8K crystal structure has Lys122 (Lys120 in PTP1B) in an open conformation that might be induced by high affinity or low molecular weight

inhibitors [25]. In the 1Q6J crystal structure of PTP1B, the R loop conformation has been variable upon the exact chemical nature of the substituent present in this pocket [29].

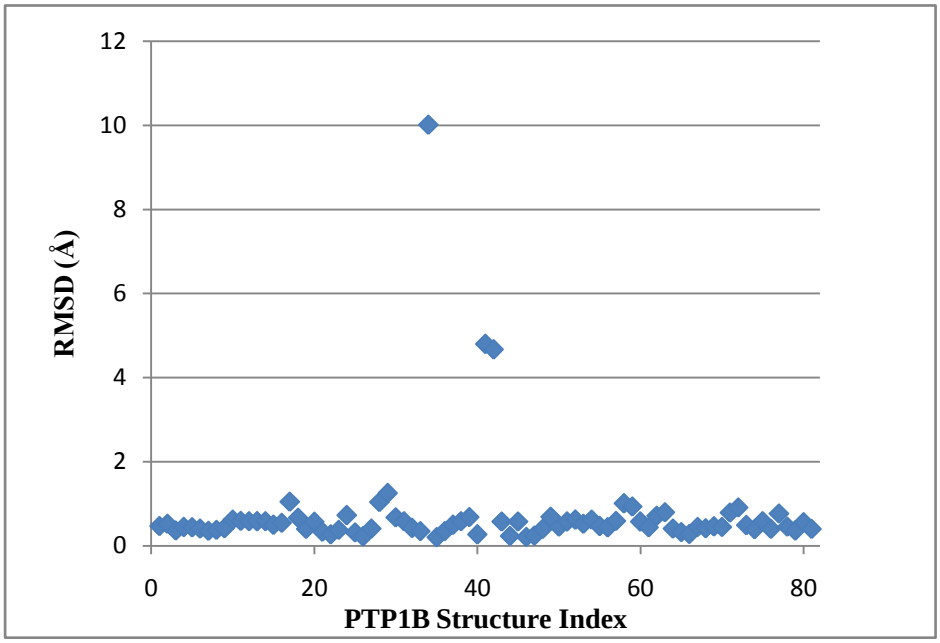

Figure 4.51. The RMSD for Cα atoms of R loop in 81 different PTP1B crystal structures relative to the 1SUG crystal structure (The crystal structure codes are represented in Table B.1 in Appendix).

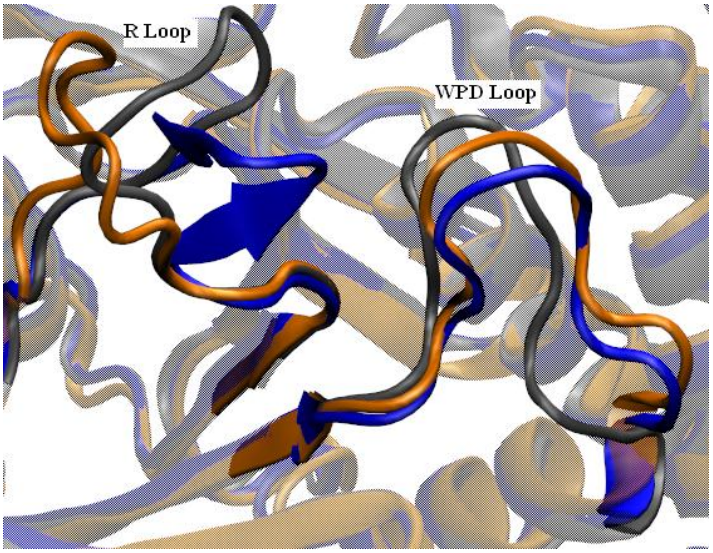

Figure 4.52. The conformations of WPD loop and R loop in the 1L8K (in orange) and 1Q6J (in grey) crystal structures superimposed with 2HNP crystal structure (in blue).

Besides these PTP1B crystal structures that have alternative R loop conformations, a number of other members in PTP family have their R-loop in a different conformation (or disordered), as explained in Section 2.5.4. The R-loop conformations in the crystal structures of PTPRR (PDB ID: 2A8B) [26], PTPN7 (PDB ID: 2A3K) [26], PTP1C (PDB ID: 1FPR) [28]and Yersinia PTP (PDB ID: 1YTS) [15] overimposed with PTP1B structure (PDB ID: 2HNP) are shown in Figure 4.53. It should be noted that these open conformations of R-loop are seen for the ligand-bound cases only.

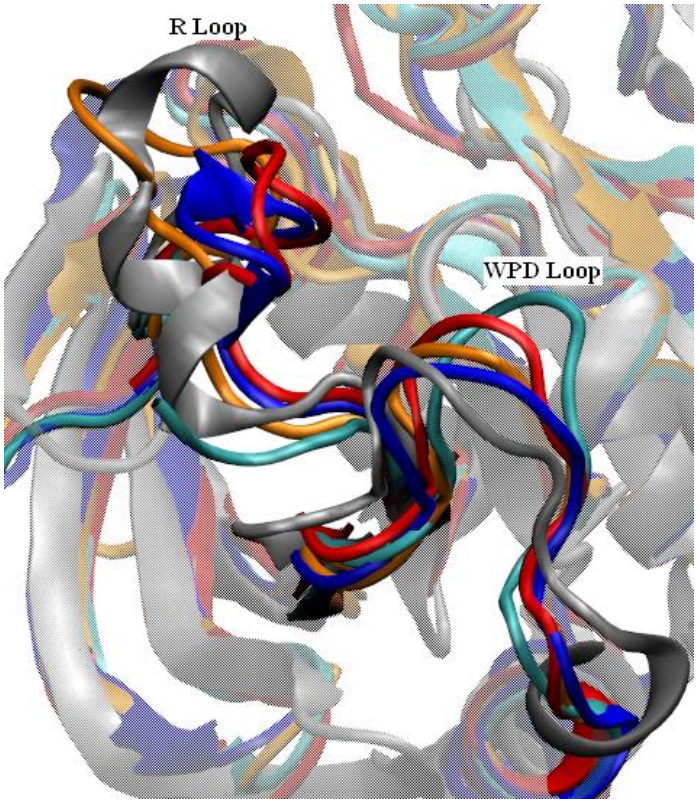

Figure 4.53. The conformations of WPD loop and R loop in the PTPRR (PDB ID: 2A8B in red), PTPN7 (PDB ID: 2A3K in cyan), PTP1C (PDB ID: 1FPR in orange) and Yersinia PTP (PDB ID:1YTS in grey) crystal structures superimposed with PTP1B crystal structure (PDB ID: 2HNP in blue).

4.8.3. Effect of Targeted Region on the WPD Loop Closure in Open/ Closed Conformations of the R loop

In the last two TMD simulations, the targeting potential (force constant is 5000 kcal/mol/Å^2) was applied to all atoms of the protein and the simulation length was 5 ns. It is important to monitor the dynamic response of the protein, when the force is applied to all atoms. The initial structure used in TMD6 simulation was the same initial structure used in TMD1 ($2F6F_1$), in which WPD loop was in the closed conformation and The R loop in the open conformation, while that used in TMD7 was the same as the initial structure in TMD5 ($2F6F_4$), in which WPD loop did not close due to the hindrance of R-loop. In the $2F6F_4$ structure, Arg221 residue has its conformation in WPD_{closed} structure. However, when we investigate the conformation of Arg221 in TMD1, it has its conformation in WPD_{open} structure like 2F6F crystal structure. The same target structure ($1SUG_1$) with the other simulations was used in both simulations. The RMSD of all atoms decreases from ~2.7 Å to 0.5 Å in TMD6 (Figure C.1A) and it decreases from ~1.75 Å to 0.25 Å in TMD7 (Figure C.1B), showing that the overall structure moves toward the target due to the targeting potential on the whole protein in both simulations. The RMSD for the WPD loop Cα atoms to the target conformation decreases from ~4.75 Å to 0.5 Å in TMD6 (Figure C.1C) and it decreases from ~4.25 Å to 0.25 Å in TMD7 (Figure C.1D) showing that WPD loop closes in both simulations. Note that, WPD loop closes in TMD7 although it did not close in TMD5, even though the same initial structure, in which the R loop is close to the active site, and the same target structure was used. We observed that the WPD loop could not be closed when the forces with force constant 500 kcal/mol/Å^2 and 800 kcal/mol/Å^2 to all atoms of the WPD loop were applied in two other TMD simulations. Finally, it can only be closed via the force (5000 kcal/mol/Å^2) applied to the whole protein. This suggests that the closed conformation of the R loop is an important energy barrier for WPD loop to close.

Figure 4.54 shows the conformational changes of the WPD loop during its closure in TMD6 (left panel) and TMD7 (right panel). The stages of WPD loop conformational transition were determined based on the sharp decreases in the RMSD for the WPD loop Cα atoms to the target conformation (Figure C.1). In TMD6, the WPD loop N-terminus moves toward the target slightly; but in TMD1, only C-terminus of the loop has moved

toward the target in the first 0.5-1 ns (Figure 4.54A). This shows that the first conformational transition of WPD loop is not seen in TMD6. The WPD loop closes via one major transitional conformational change, occuring at ~2.25 ns in TMD6 (Figure 4.54C). In TMD7, neither the N- or C-terminus but the tip of the WPD loop moves toward the target at the first conformational transition (Figure 4.54B). The complete closure of WPD loop occurs with two different conformational transitions at ~2 ns (Figure 4.54D) and ~3 ns (Figure 4.54F) in TMD7.

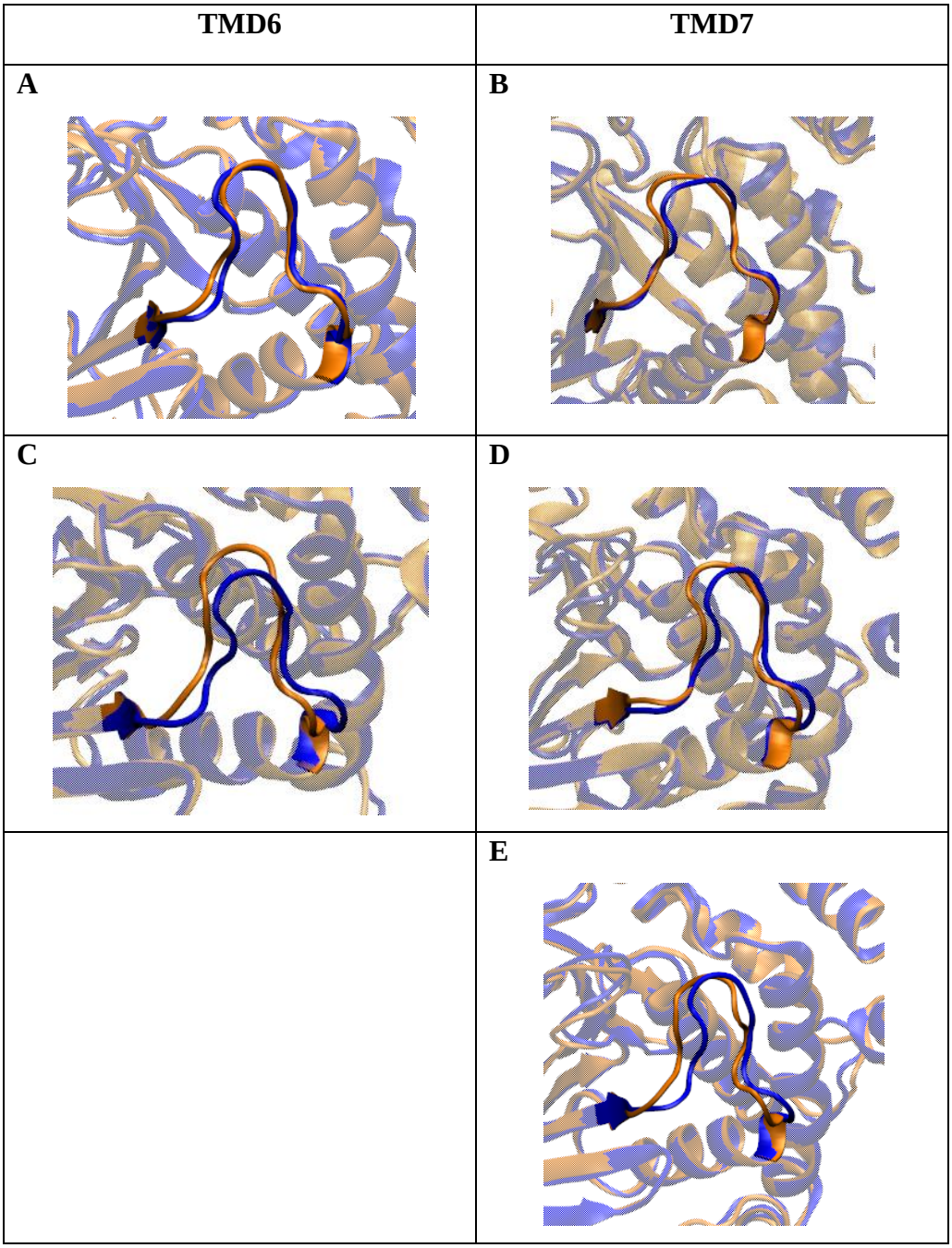

Figure 4.54. Conformations of WPD loop in TMD6 and TMD7 snapshots representing the conformations before and after the sharp decrease in the RMSD of C_α atoms of WPD loop in these simulations. TMD6 snapshots at (A) 1ns (blue) and 1.75 ns (orange); (C) 2 ns (blue) and 3 ns (orange) are shown. TMD7 snapshots at (B) 0.5ns (blue) and 1.5 ns (orange); (D) 1.75ns (blue) and 2.5 ns (orange) and (E) 2.75 ns (blue) and 3.5 ns (orange) are shown.

The residue-residue interactions and dihedral angle transitions related to the conformational transition of WPD loop were monitored in TMD6 and TMD7. The change in the $\chi 2$ of Trp179 dihedral angle, the interactions of Trp179 and Asp181 with Arg221 and the interaction between Phe182 and Thr263 were seen to be similar to those previously seen (Figure C.2). The Ψ dihedral angle transition of Asp181 and the interactions formed by Phe182-Gln262 and Asp181-Gly183 were also found to be similar in TMD6 and TMD7 to the previous simulations (Figure C.3 and Figure C.4).

The hydrophobic interactions during WPD loop closure were also investigated in TMD6 and TMD7 simulations. The potential steric clashes between the hydrophobic residues were obtained as the same in TMD6 and TMD7 with the previous simulations such as TMD1 and TMD2 (Figures C.5-C.6-C.7).

4.8.3.1. Interactions of the R loop with the WPD loop and the active site : Although there is no interaction between Glu115 OE2 and Arg221 NH1 at the beginning of the simulation, this salt bridge is formed at ~2 ns, just before the conformational jump of the WPD loop (Figure 4.55A). In TMD7, this salt bridge is maintained throughout the whole simulation (Figure 4.55B). It should be noted that, WPD loop was able to close in the presence of this interaction. Visual examination of the simulations shows that the R loop (especially residues 116-118) moves slightly away from its initial closed conformation to allow Asp181 to move toward its conformation in the target structure for the WPD loop closure and then returns to its initial conformation in TMD7 (Figure 4.55H). In TMD6, R loop moves away from its initial conformation and remains at ~3.5 Å away from its initial conformation, showing that R loop moves from the open to closed conformations. (Figure 4.55G). The loss of the Arg112- Asp181 interaction and formation of Glu115- Asp181 interaction take place simultaneously with the WPD loop closure (Figures 4.55C-F).

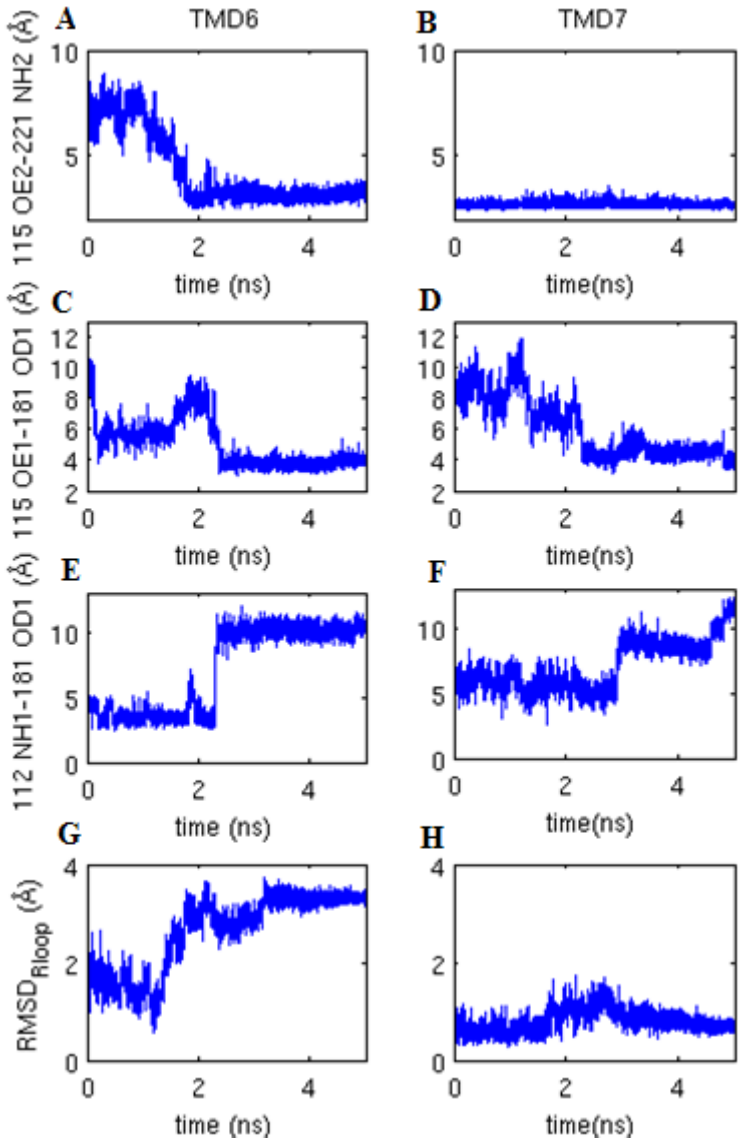

Figure 4.55.Time evolution of (A), (B) Glu115 OE2 - Arg221 NH1 distance, (C), (D) Glu115 OE2 - Asp181 OD1 distance, (E), (F) Asp181 OD1 - Arg112 NH1 distance, (G), (H) and the time evolution of the RMSD for all atoms of R loop relative to its initial conformation in TMD6 and TMD7.

In order to determine the correlation between the R and WPD loops, the RMSDs of Cα atoms of R and WPD loops calculated relative to the target are shown for TMD6 and TMD7 (Figure 4.56). It is clearly seen that R-loop moves toward the target conformation as WPD loop closes in TMD6. As it is discussed in Section 4.8.2, R loop is close to the active site in all 78 PTP1B crystal structures. Therefore, R loop moves toward the target conformation in TMD6 since it is far away from the active site in the initial structure of this simulation. In TMD7, since R loop (especially residues 116-118) moves away from its

target conformation to allow WPD loop to close and then turns its target conformation back (Figure 4.57). The centroid point represent the center point of each cluster visited by WPD loop during its closure in TMD6 and TMD7 simulations.

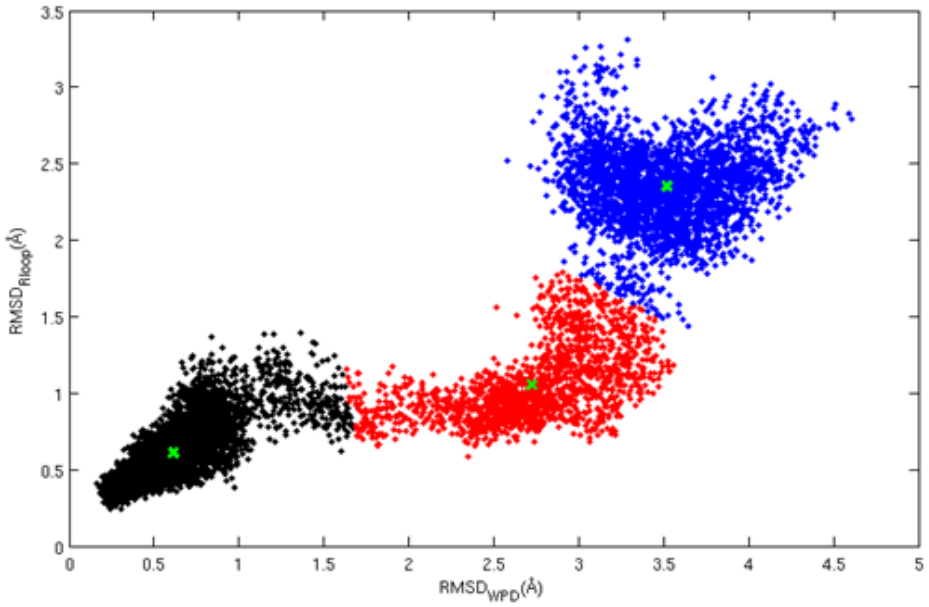

Figure 4.56. RMSD trajectory of Cα atoms of R and WPD loops relative to the target structure in TMD6. (x represents the centroid points of each cluster).

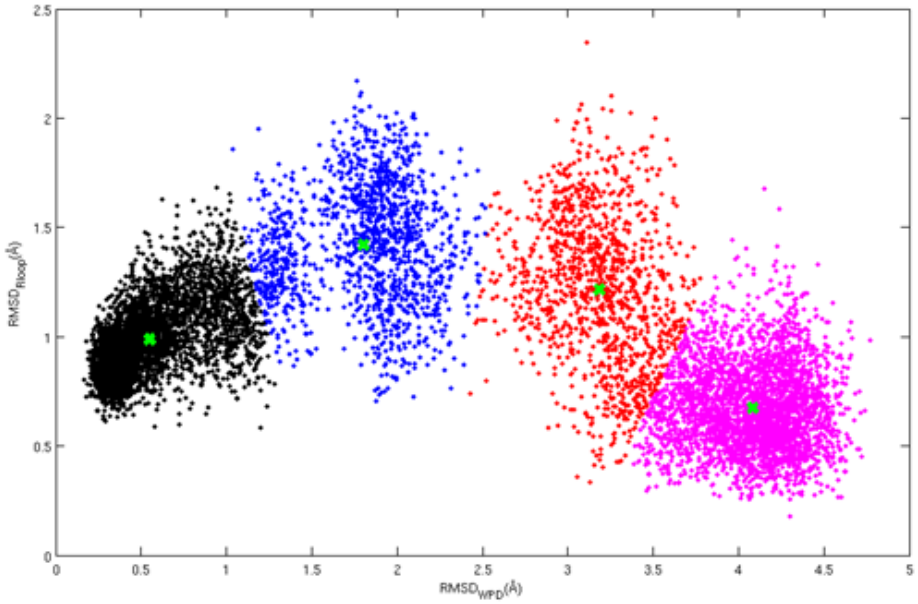

Figure 4.57. RMSD trajectory of Cα atoms of R and WPD loops relative to the target structure in TMD7. (x represents the centroid points of each cluster).

4.8.3.2. Mechanism of the WPD loop closing in closed conformation of the R loop : In TMD7, Asp181 was able to move toward the target, and WPD loop was able to close, even though the R loop was in closed conformation, and even though the Glu115- Arg221 interaction was maintained. The complete closure of the WPD loop can be reached by the steps shown in Figure 4.58: (i) the hydrogen bond between Trp179 O and Arg112 NH is maintained until 2.8 ns (Figure 4.58A). However, as WPD loop starts to move toward the target conformation due to the targeting potential, Trp179 starts to move. Hence, the $\chi 4$ of Arg112 side chain rotates slightly at ~2 ns to maintain the interaction with Trp179 (Figure 4.58B). (ii) Arg112 and Glu115 would have made a steric clash, if Glu115 did not move out of the way for Arg112 (Figure 4.58C). Hence, Glu115 should change its position in order to avoid the steric clash. Arg112 $\chi 4$ rotates slightly at ~2ns (Figure 4.58B), simultaneously with a small change in $\chi 1$ of Glu115 (Figure 4.58F) and a significant change in $\chi 3$ of Glu115 (Figure 4.58E). This shows that WPD loop has approached the active site, and Arg112 sidechain should change its position to maintain the hydrogen bond with Trp179 (Figure 4.58A), and Glu115 $\chi 1$ sidechain angle approaches its value in the target conformation, while maintaining its interaction with Arg112 backbone. The increase in the R loop RMSD (Figure 4.55G) coincides with this event. (iii) At ~2.5 ns, the Glu115 side chain $\chi 1$ angle rotates by about -60^{o} (shown as 300° in Figure 4.58 F), a new conformation not seen in the WPD_{open} and WPD_{closed} structures, and stays there between 2.5-3 ns. During this period, Glu115 sidechain oxygens make a sudden rotation at 2.8 ns (Figure 4.58E) breaking the hydrogen bond between Glu115 OE1 and Arg112 NH (Figure 4.58D). This conformational change in Glu115 allows Asp181, thus WPD loop, to close over the active site (Figure 4.58A). After WPD loop moves closer to the active site, Glu115 retains its hydrogen bond with Arg112 backbone (Figure 4.58D) and Glu115 sidechain returns to its conformation before the transition after visiting a novel sidechain conformation between 3 and 4.2 ns (Figure 4.58F). Asp181 CB and Glu115 O would have made a steric clash and Asp181 could not have moved for the WPD loop to close, if Glu115 did not move out of the way for Asp181 (Figure 4.59A). Therefore, this allowance of Glu115 results from the movement of CB atom of Glu115 in TMD6 (in green) and TMD7 (in red) simulations for Asp181 to move toward the target and WPD loop to close in these simulations (Figure 4.59B).

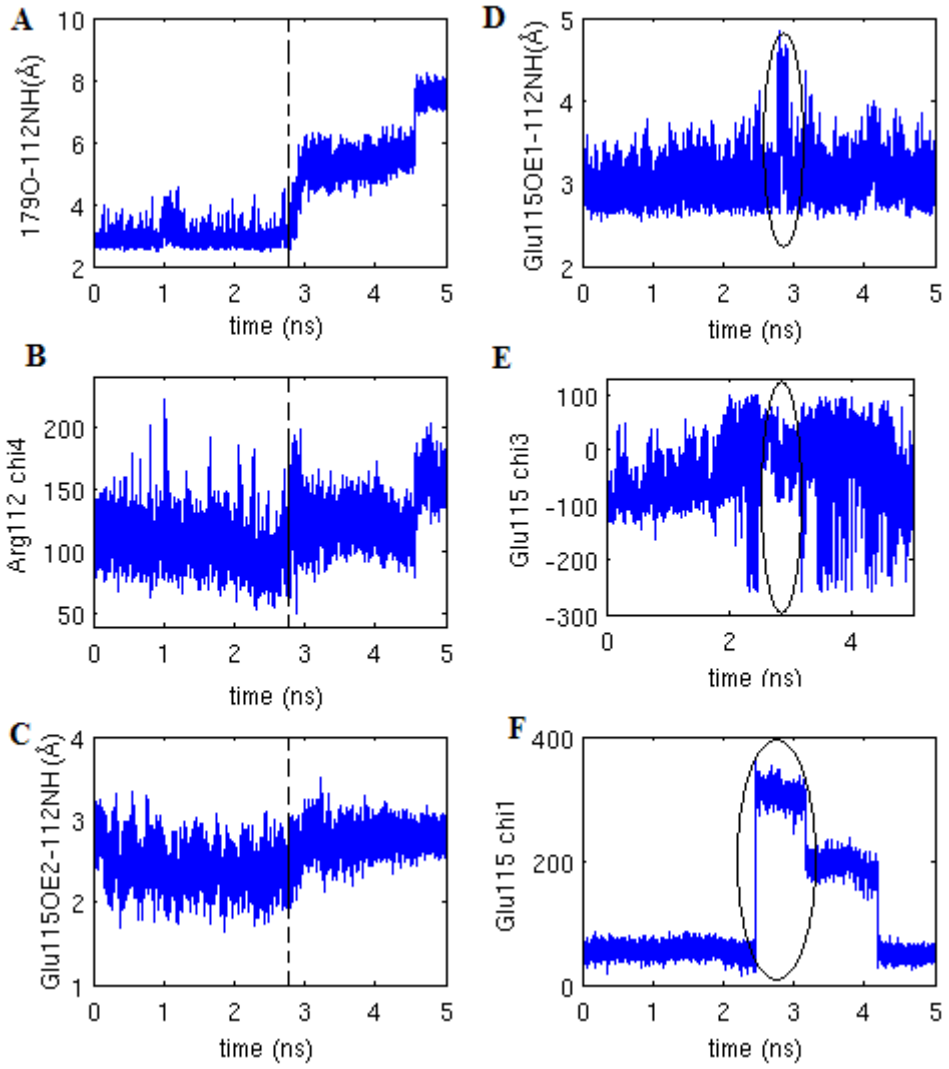

Figure 4.58. Time evolution in (A) Trp179 O-Arg112NH distance, (B) χ4 of Arg112, (C) Glu115 OE2(from the initial structure)- Arg112 NH (from the simulation structure) distance, (D) Glu115 OE1-Arg112 NH distance, (E) χ3 of Glu115, and (F) χ1 of Glu115 in TMD7.

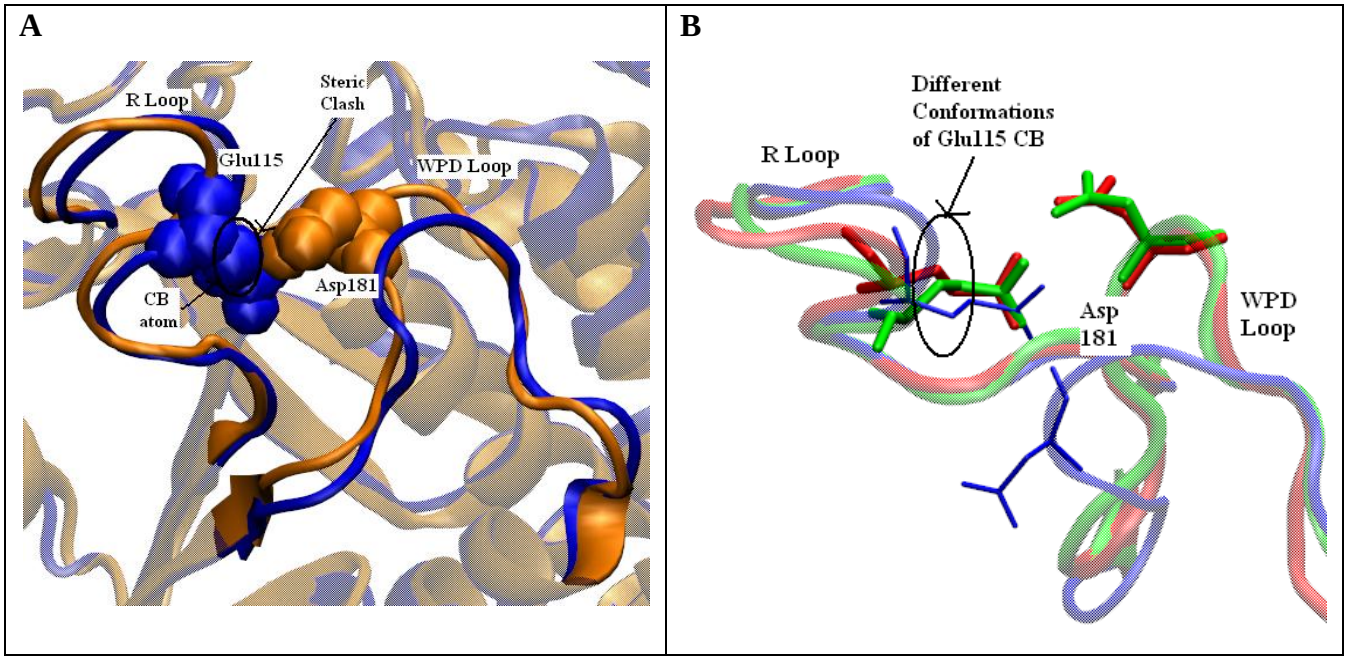

Figure 4.59. (A) Conformations of the WPD loop and the R loop in a TMD7 snapshot at 2.6 ns (in orange) and in the initial structure (in blue). Glu115 (in blue spheres) and Asp181 (in orange spheres) are shown. (B) Different conformations of Glu115 and Asp181 residues in initial structure of TMD7 (in blue), in a TMD7 snapshot at 2.9 ns (in orange) and in a TMD6 snapshot at 2.45 ns (in green). The conformational change of Glu115 CB atom in these structures is shown in circle.

5. CONCLUSIONS AND RECOMMENDATION FOR FUTURE STUDIES

5.1. Conclusions

In this thesis, targeted molecular dynamic (TMD) simulations between the WPD_{open} and WPD_{closed} conformations of PTP1B were performed to elucidate the conformational activation mechanism of PTP1B. Initial structures in the TMD simulations were representative snapshots from equilibrium MD simulations on the WPD_{open} structures (2F6F), and the target structure was a representative snapshot from an equilibrium MD simulation on the WPD_{closed} structure (1SUG). TMD simulations were repeated using different initial conformations and force constants, as well as by employing TMD potential on different regions of the protein. The most flexible regions in the TMD simulations, in which TMD potential was only applied to the WPD loop, were found to be the regions between α2'-α1 helices, the loop between β1-β2 sheets (L2 loop), R loop, WPD loop, S loop and α7 helix. We observed that the loop between α2'-α1 helices, R loop, WPD loop, S loop and α7 helix moved away from their initial conformations, but only the loop between α2'-α1 helices and WPD loop move toward the target conformation. It is important to note that mobility of the S loop, which does not move toward its conformation in the target structure, is higher in TMD simulations compared to MD simulations, suggesting that flexibility of the S loop may be important for WPD loop conformational activation.

Visual examination of the simulation trajectories as well as the change in RMSD of Cα atoms of the TMD simulation structures with respect to the target conformations showed that there were two major conformational transitions between the open and closed states of WPD loop. The conformational states (clusters) between these transitions were determined by PCA and K-means clustering analysis of the C_α atoms. During the first transition, the WPD loop N-terminus (residues 184-189) moved toward the target structure as a result of the rotation of the backbone dihedral angles of Ser187 or Pro188 at ~1- 2 ns. The second transition took place as a consequence of the simultaneous rotation of Trp179 and Arg221 sidechain dihedral angles, and the formation of a polar interaction between the

WPD loop and the Arg221 sidechain at ~2.75- 3 ns. A third subtle conformational change, which was not revealed in the clustering analysis, was observed by a visual examination of the TMD trajectories. During this third transition, the backbone dihedral angles between Asp181 and Phe182 underwent conformational change, resulting in the formation of the hydrogen bonds between the C and N terminus sides of the WPD loop (Asp181-Gly183), making the tip of the loop narrower, and between Phe182-Gln262, bringing the catalytically important Gln262 closer to the active site, at ~4.25-4.75 ns.

In one of the simulations, WPD loop did not close over the active site. Arg221 residue formed a hydrogen bond, which was not seen in other TMD simulations, with the backbone oxygen of Asp181 from the early stages of the simulation up to 3.75 ns, and then the Arg221 sidechain rotated to an alternative conformation, which was not seen in the crystal structures. Hydrogen bonds with Leu110, Trp179 and Arg221 help stabilize this novel conformation of Arg221, and prevent the closing of Asp181 (thus WPD loop) on the active site.

In the other simulation in which WPD loop did not close, the closed conformation of R loop was found to be responsible from preventing the closure of the WPD loop. Salt bridges between Glu115-Arg221 and Arg112-Asp181 were not disrupted during this simulation, and R loop maintained its starting conformation, which hindered the WPD loop motion toward the target conformation. Investigation of the TMD simulations, in which WPD loop was closed, showed that open conformation of the R loop was stabilized by rotation of Arg112 side chain $\chi 4$ angle and simultaneous rotation of Leu110 backbone, leading to the disruption of the hydrogen bond between Leu110-Arg221. Our results suggest that the R loop conformation and its interaction with the active site might be important for WPD loop closure. The R loop is close to the active site in most of the PTP1B structures, while in three PTP1B crystal structures and in four structures of the other PTPs, it was found in an alternative “open” conformation (or disordered). This clearly showed that the R loop could be stabilized and even disordered in open conformation in the PTP-ligand complexes. In our TMD simulations, we observed that one of the most important barriers to WPD loop closure was the closed conformation of the R loop. Therefore, we might suggest that in a possible intermediate structure, the R loop

moves away from the active site and remains in an open conformation for the WPD loop to close.

To close the WPD loop in the closed conformation of the R loop, force constant was increased to 500 kcal/mol/Å^2, and then to 800 kcal/mol/Å^2, but WPD loop still did not close. Therefore, TMD potential was applied to all atoms of the whole protein with a higher force constant of 5000 kcal/mol/Å^2. A conformational transition pathway followed by the WPD loop during its closing motion was obtained, despite the presence of the Glu115- Arg221 salt bridge, which was throughout the TMD simulation. Closing motion of the WPD loop took place in a number steps: the sidechain of Arg112 moved to maintain the polar interaction with the backbone of Trp179. To avoid a steric clash with Arg112, the Glu115 changed its position, which allowed Asp181 to move toward the target structure. In another TMD simulation, in which the same TMD potential but an initial structure with open R-loop conformation was used, the R loop moved from the open to the closed conformation, followed by the simultaneous loss of interactions between Arg112 with Asp 181, and formation of Arg221- Asp181 interaction.

5.2. Recommendations for Future Studies

The conformation of the R loop was found to be an important element in the WPD loop conformational activation. As a future study, TMD potential may be applied only on the atoms of the R loop to examine whether the R loop motion will affect WPD loop dynamics. In addition, the Arg221 conformation in the initial structure is obtained as one of the most important parameters for WPD loop closure. In order to understand the effect of Arg221 conformation on the WPD loop closure, one can perform a TMD simulation by using an initial structure in which Arg221 is in an alternative conformation as a future study. TMD potential may be applied on Glu115 and WPD loop only to investigate whether Arg221 and the R loop will maintain their conformations during WPD loop closure. If so, this may show that the function of Glu115 is to keep Arg221 in a convenient conformation during WPD loop closure. In this thesis, the TMD simulations were performed on apo structures of PTP1B.

In future studies, TMD simulations may be performed to examine the conformational activation of liganded structures, such as PTP1B structures complexed with peptides, with pTyr or with pTyr mimetic inhibitors. This will help elucidate the contribution of the ligand to the conformational activation of the WPD loop.

APPENDIX A: RESULTS OF THE TMD2 SIMULATION ANALYSIS

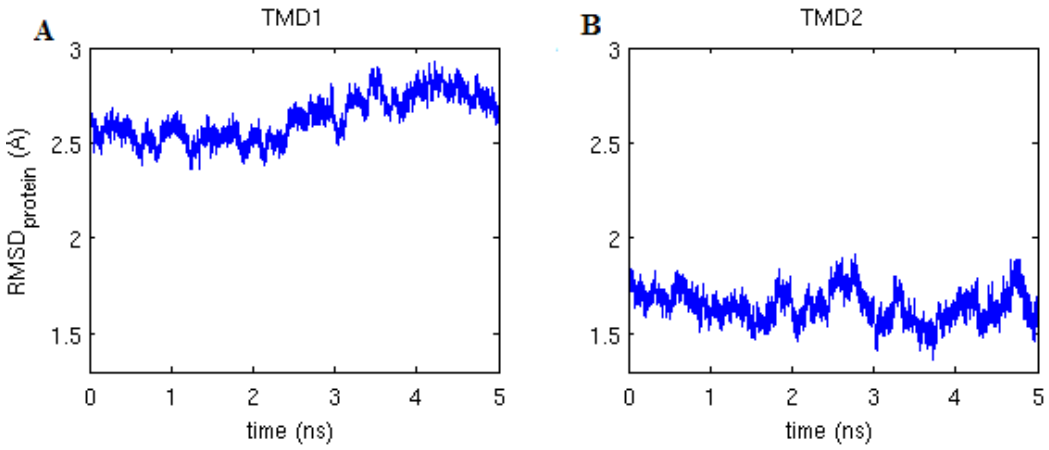

Figure A.1. RMSD of all protein atoms relative to the target structure during TMD1 and TMD2 simulations.

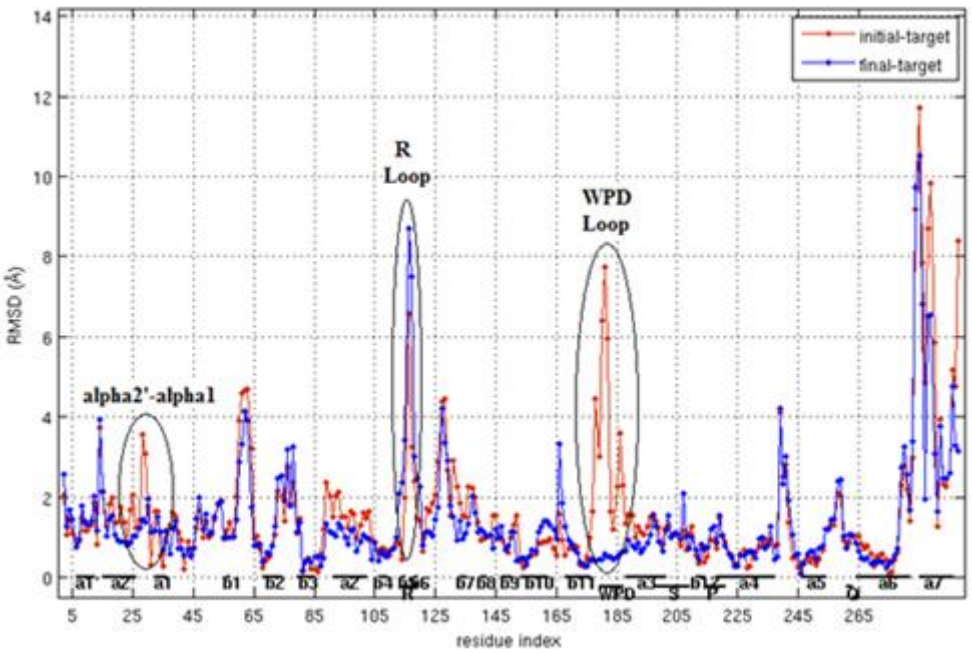

Figure A.2. Residue based RMSDs for all C_α atoms between the target structure and the initial structure (red), and RMSDs between the target structure and the final structure (blue) in TMD2 simulation.

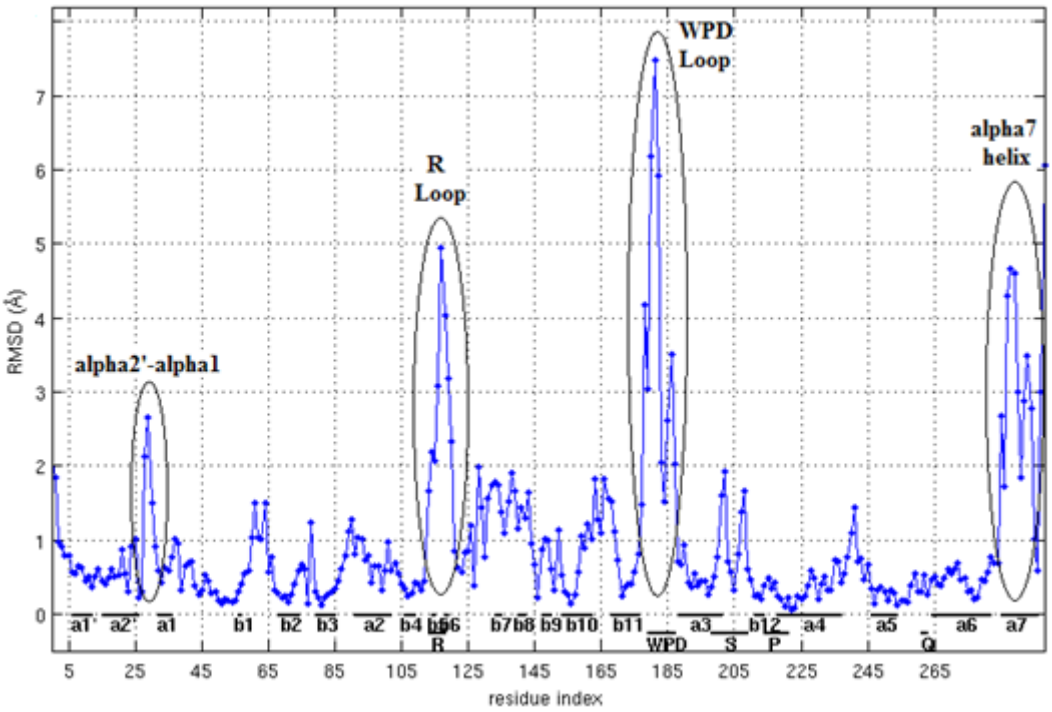

Figure A.3. RMSDs of the C_α atoms between the initial and the final structures of of the TMD2 simulation.

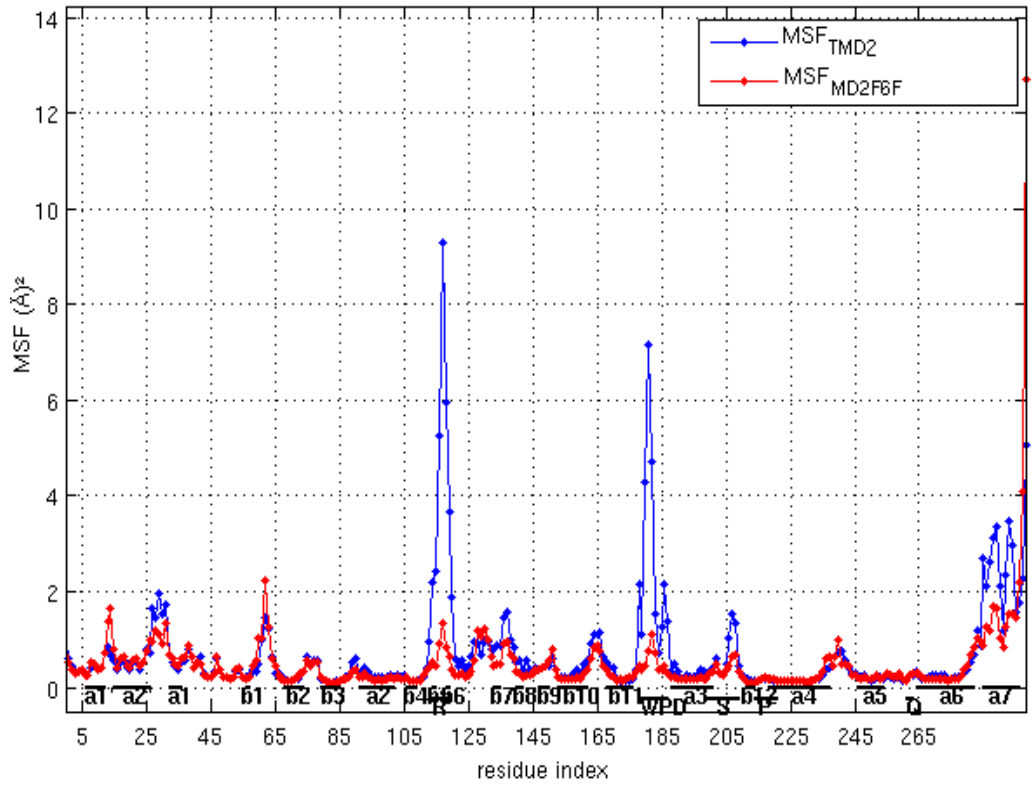

Figure A.4. MSF of residues in TMD2 simulation (blue) and in the MD simulation of the 2F6F structure (red).

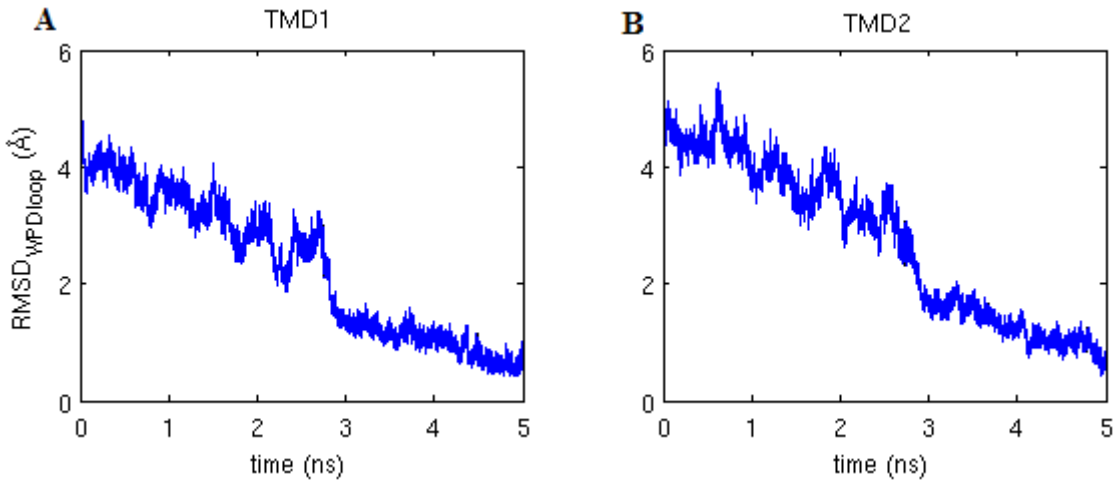

Figure A.5. RMSD of WPD loop atoms relative to the target structure during (A) TMD1 and (B) TMD2 simulations

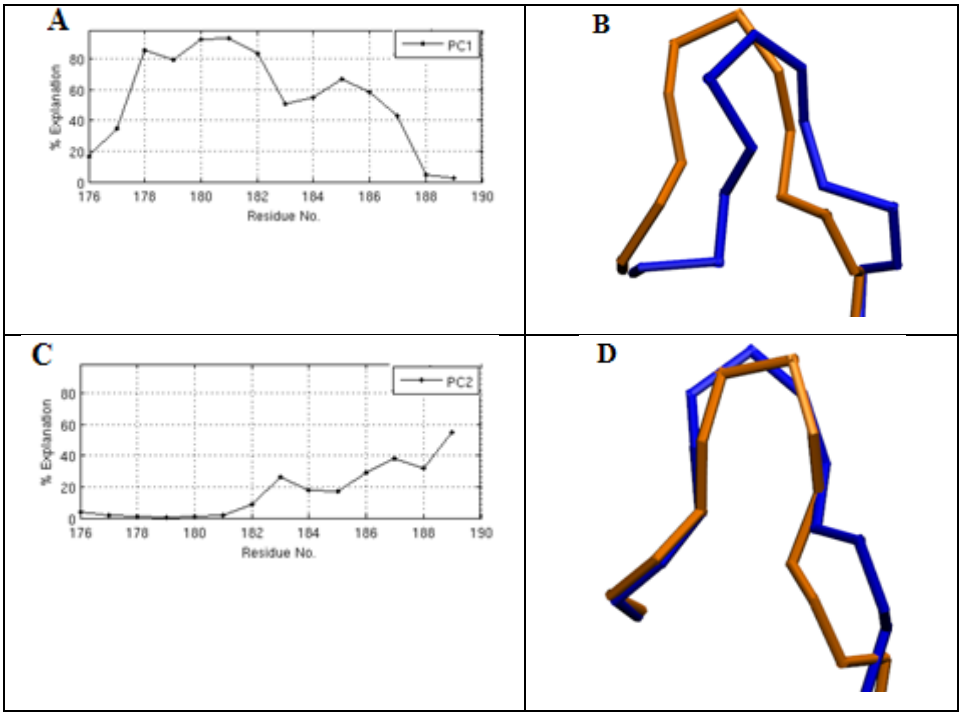

Figure A.6. In TMD 2, (A) the percentage of WPD loop motion explained by PC1, and (B) representative snapshots along PC1. (C) The percentage of WPD loop motion explained by PC2, and (D) representative snapshots along PC2. In (B) and (D), the first and the final conformations are shown in blue orange, respectively.

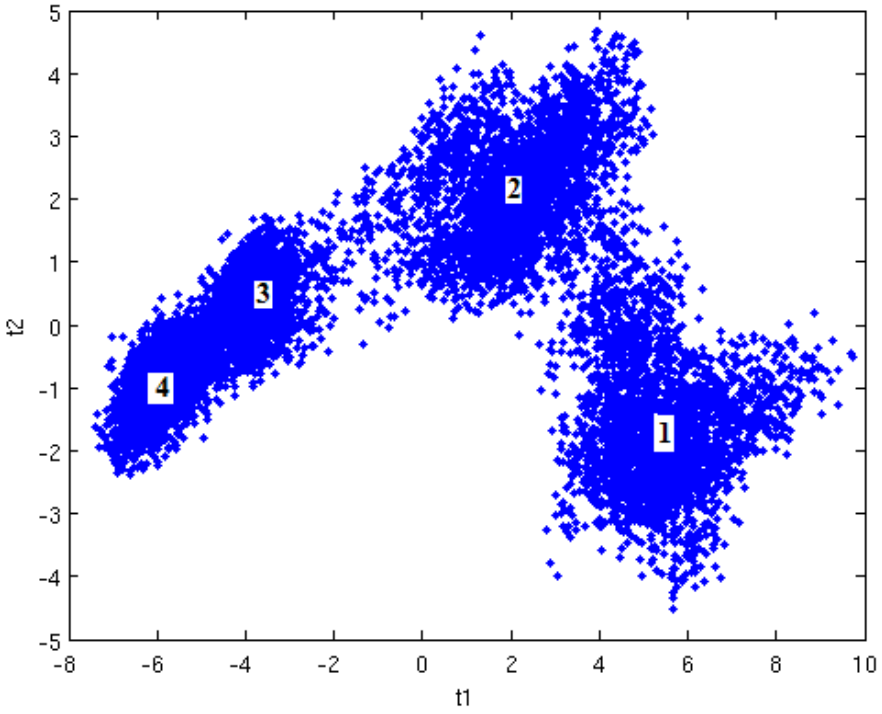

Figure A.7. Projection of the WPD loop trajectory in TMD2 on the reduced subspace of PC1 and PC2.

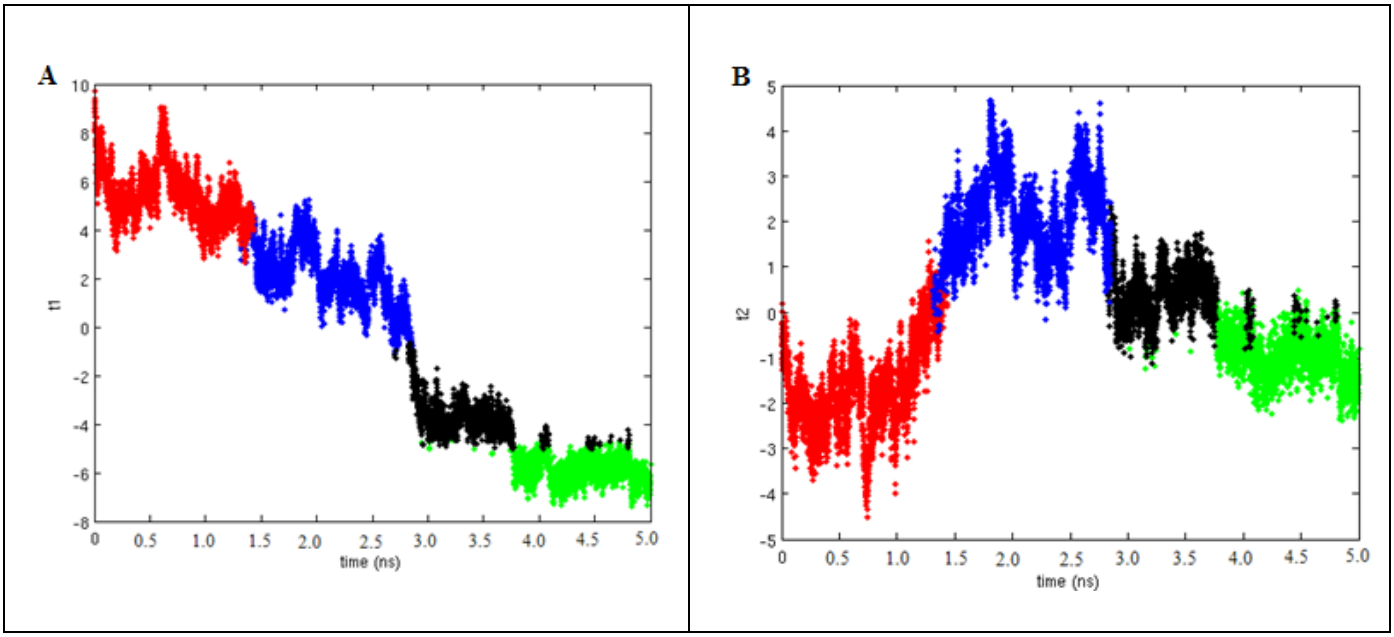

Figure A.8. The projection of the WPD loop motions on (A) PC1 (B) and PC2 in TMD2. Three different clusters are shown with different colors.

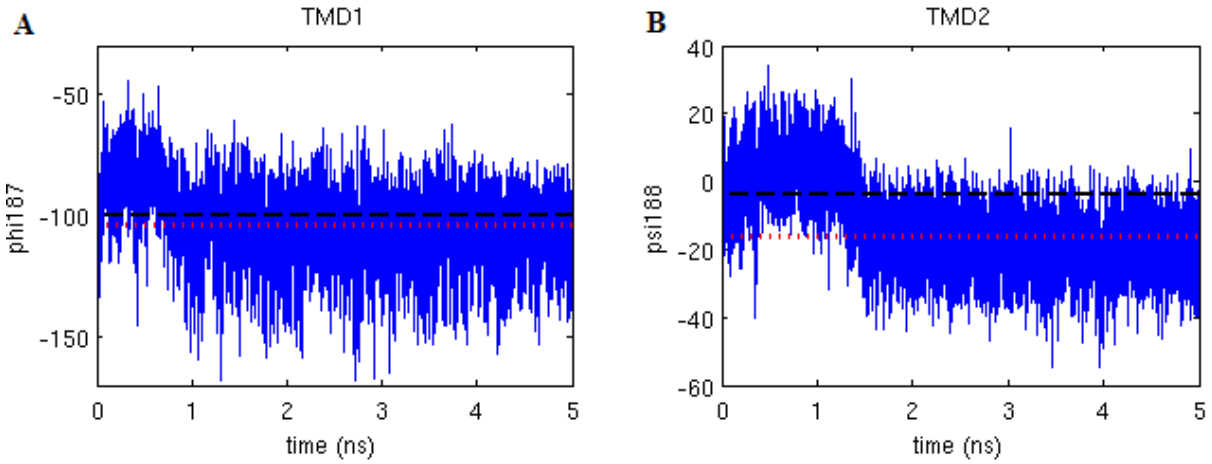

Figure A.9. The changes (A) in Φ of Ser187 in TMD1 and (B) Ψ of Pro188 in TMD2.

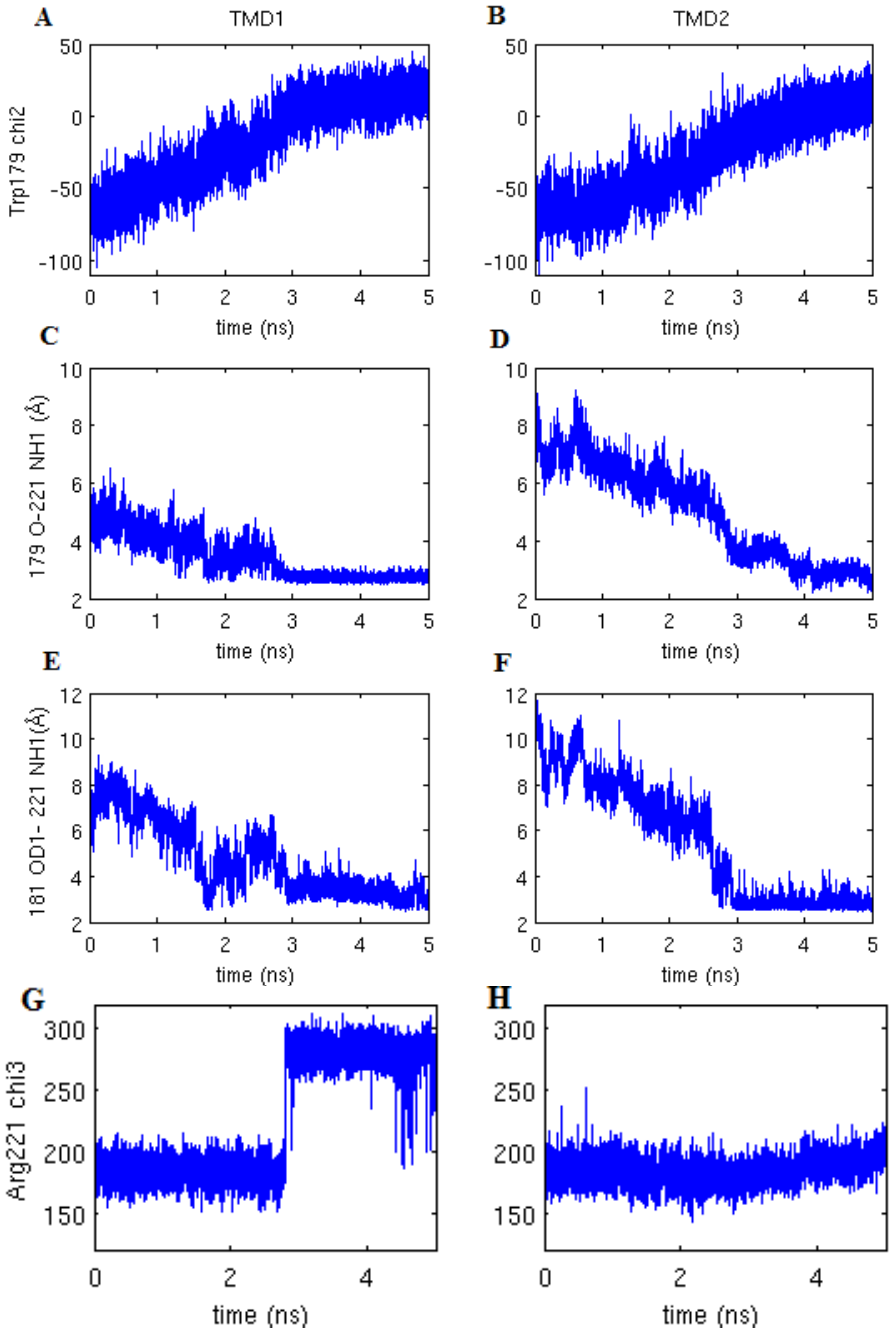

Figure A.10. The changes in (A), (B) χ2 of Trp179 , (C), (D) Trp179 O - Arg221 NH1 distance, (E), (F) Asp181 (OD1) OD2 - Arg221 (NH1) NH2 distance and (G), (H) the χ3 of Arg221 in TMD1 and TMD2.

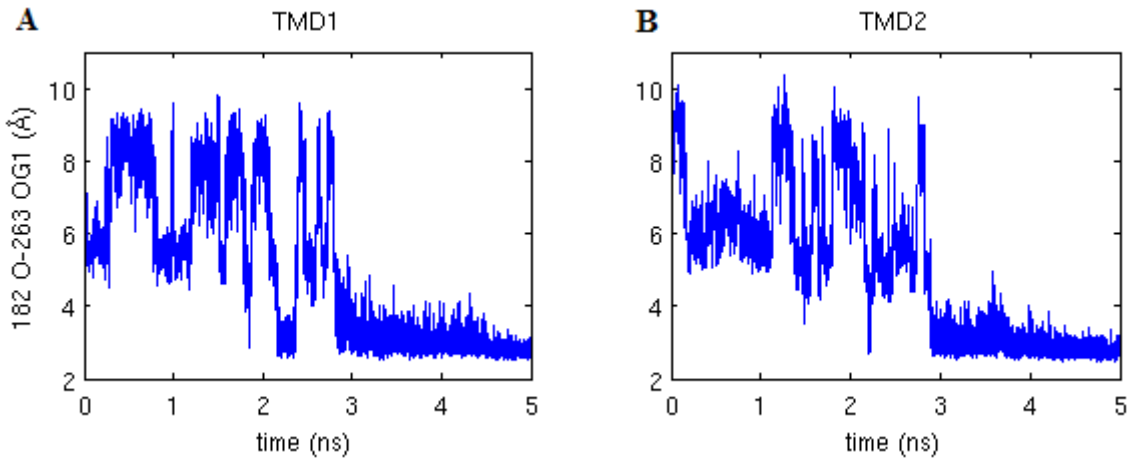

Figure A.11. The changes in Phe182 O – Thr263 OG1 distance in (A) TMD1 and (B) in TMD2.

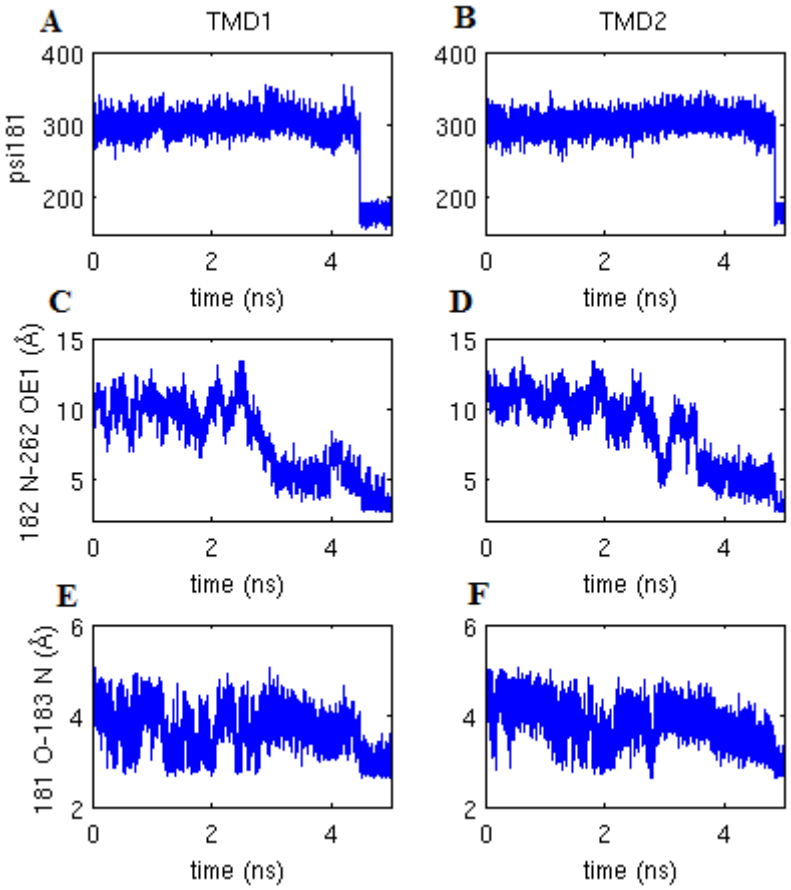

Figure A.12. Time profiles of (A), (B) Ψ backbone dihedral angle of Asp181, (C), (D) the distance between Phe182 N atom-Gln262 OE1 atom, (E), (F) and the distance between Asp181 O atom-Gly183 N atom in TMD1 and TMD2.

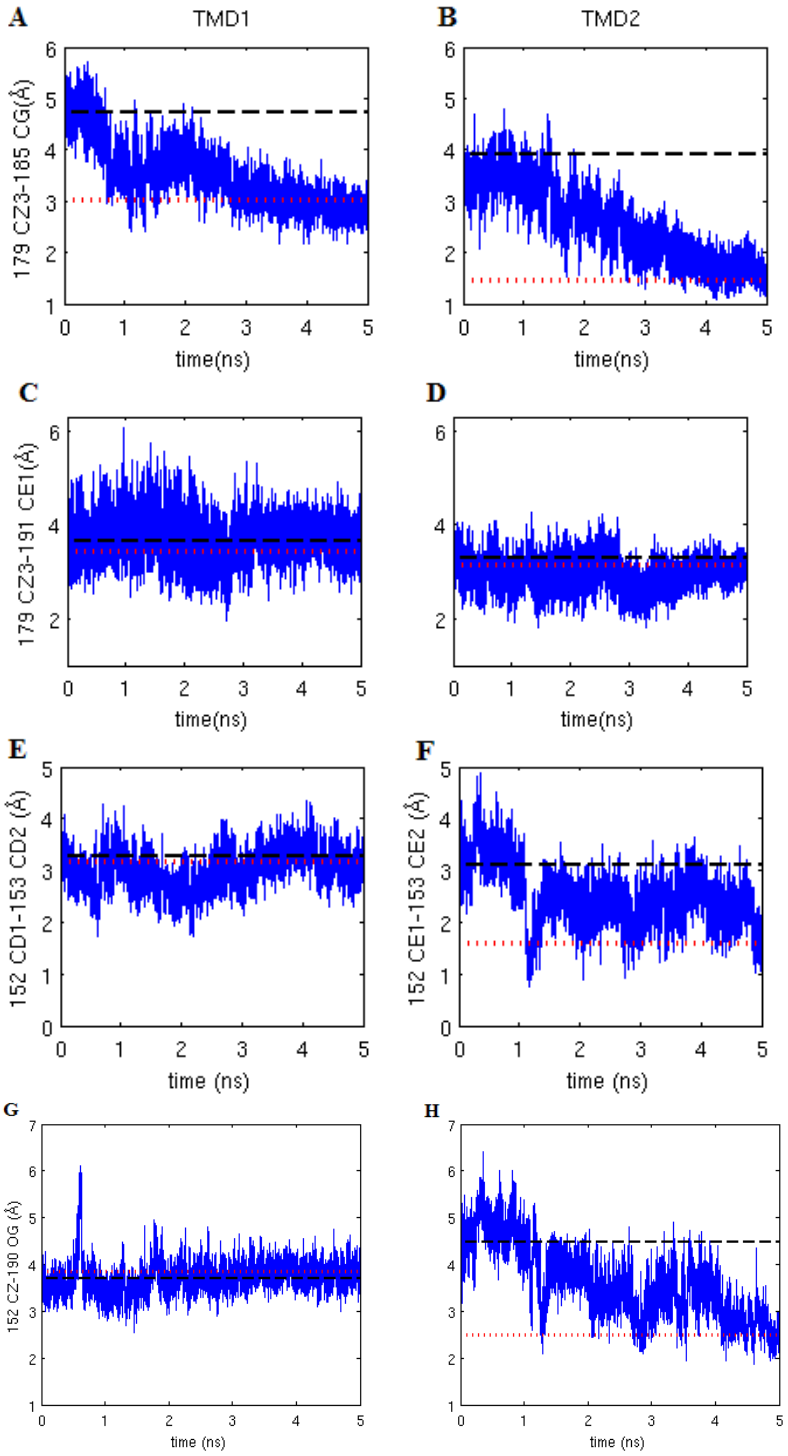

Figure A.13. Time profile of (A), (B) the distance between Trp179 CE2 atom- Phe185 CD atom, (C), (D) the distance between Trp179 CZ3 atom- Phe191 CE1, (E), (F) the distance between Ser190 O atom- Tyr152 CD1 and (G), (H) the distance Tyr152 CZ-Ser190 OG in TMD1 and TMD2.

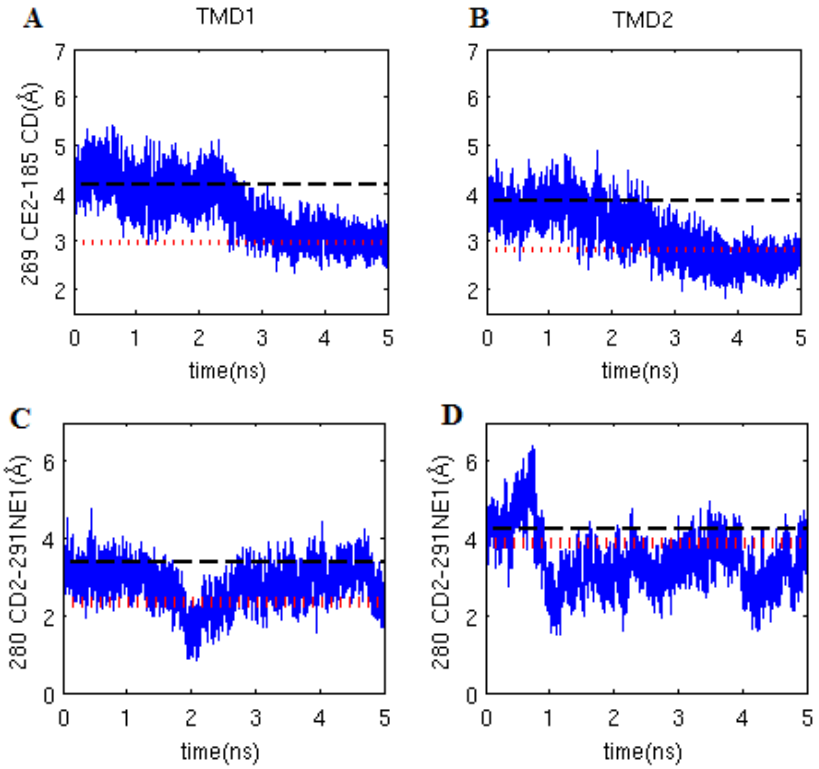

Figure A.14. Time profile of (A), (B) the distance between Trp269 CE2 atom- Phe185 CD atom, (C), (D) and the distance between Trp291 NE1 atom- Phe280 CD2 atom in TMD1 and TMD2.

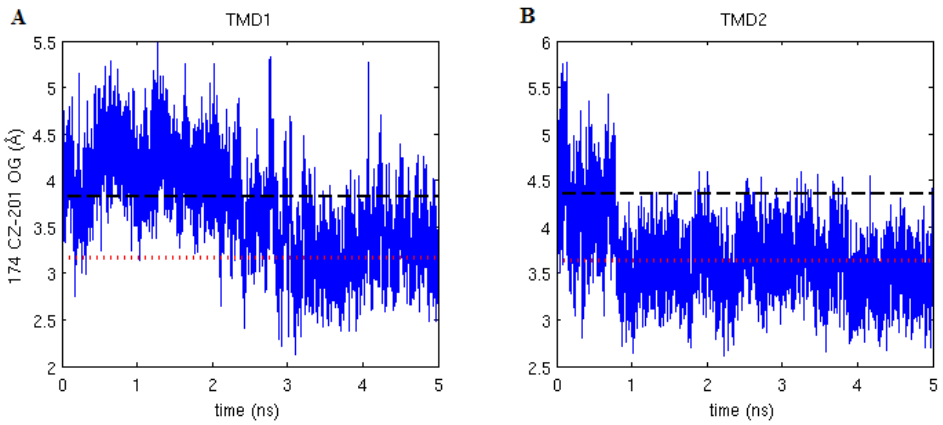

Figure A.15. Time profile of the distance between Phe174 CZ atom- Ser201 OG atom in (A) TMD1 and (B) TMD2.

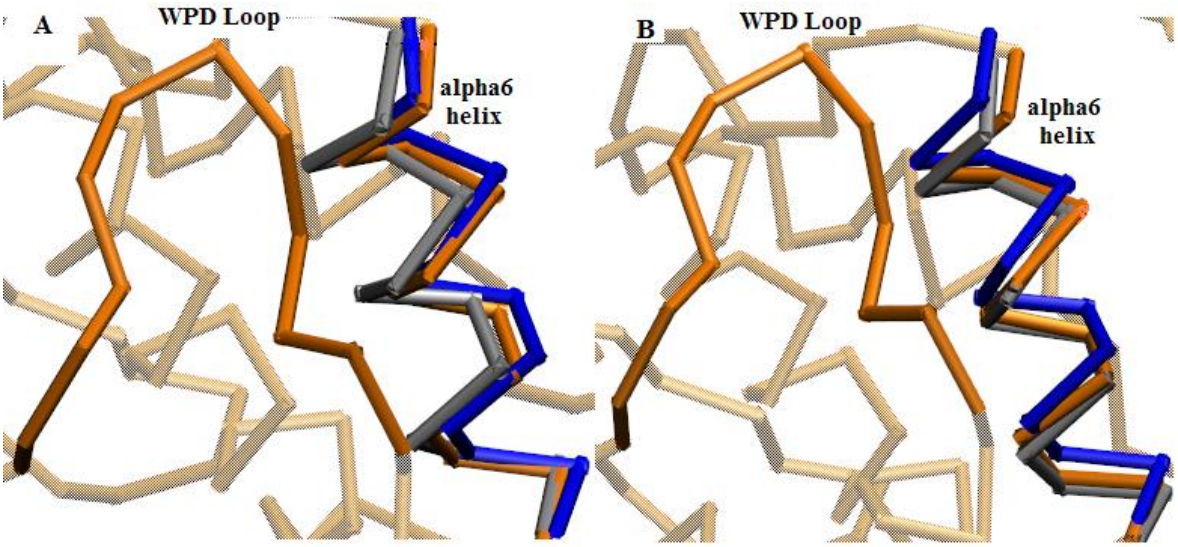

Figure A.16. The representation of the movement of α6 helix explained by (A) PC1 and (B) PC2 on the protein structure in TMD2. The target structure (orange), the first snapshot of PCs (blue) and the final snapshot of PCs (silver) are shown.

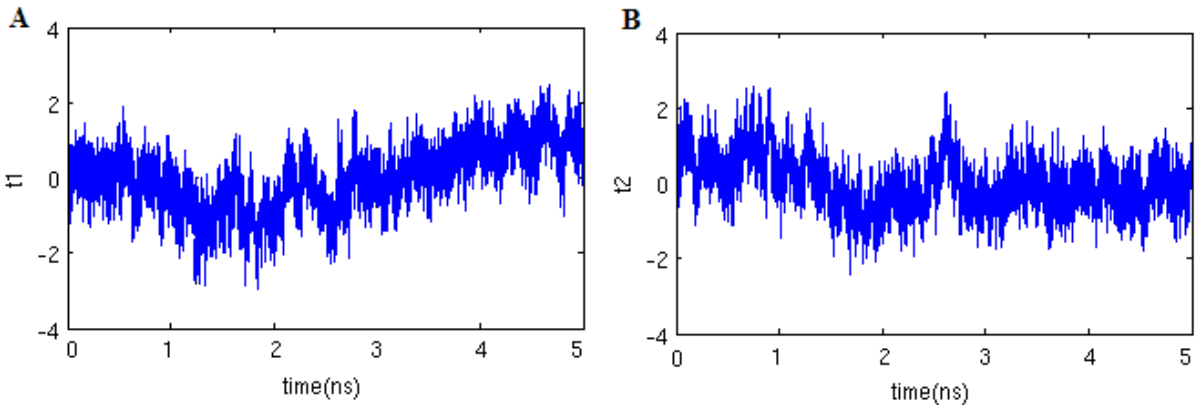

Figure A.17. The projection of the α6 helix motions on (A) PC1 (B) and PC2 in TMD2.

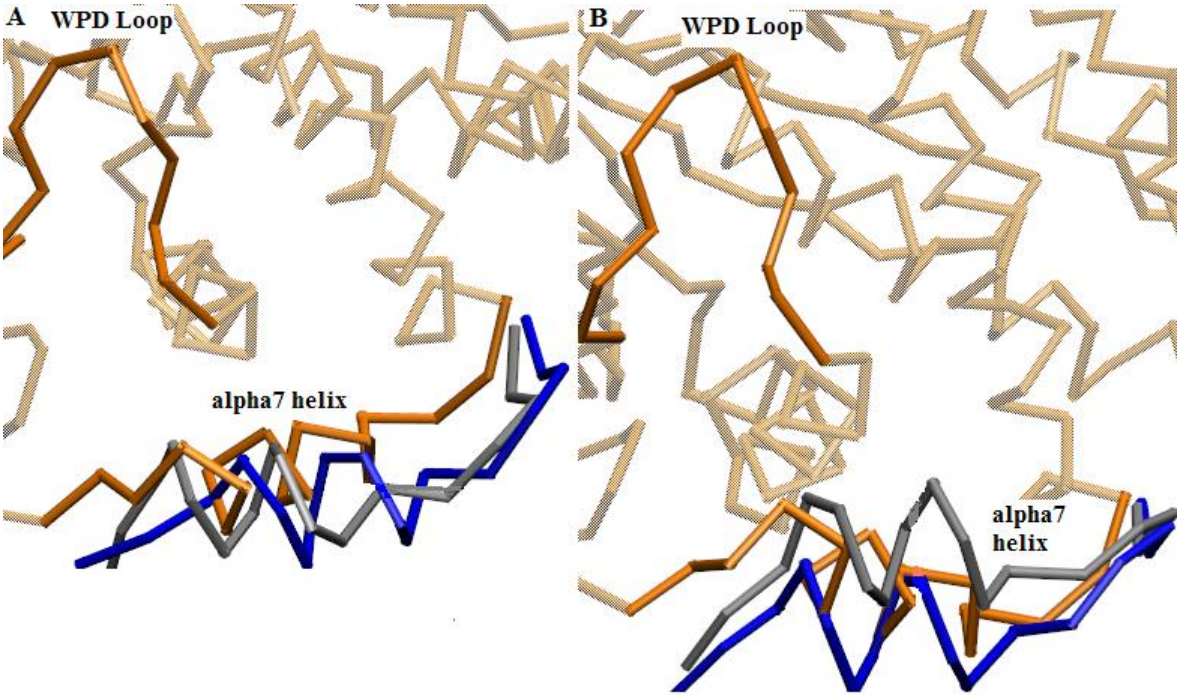

Figure A.18. The representation of the movement of α7 helix explained by (A) PC1 and (B) PC2 on the protein structure in TMD2. The target structure (orange), the first snapshot of PCs (blue) and the final snapshot of PCs (silver) are shown.

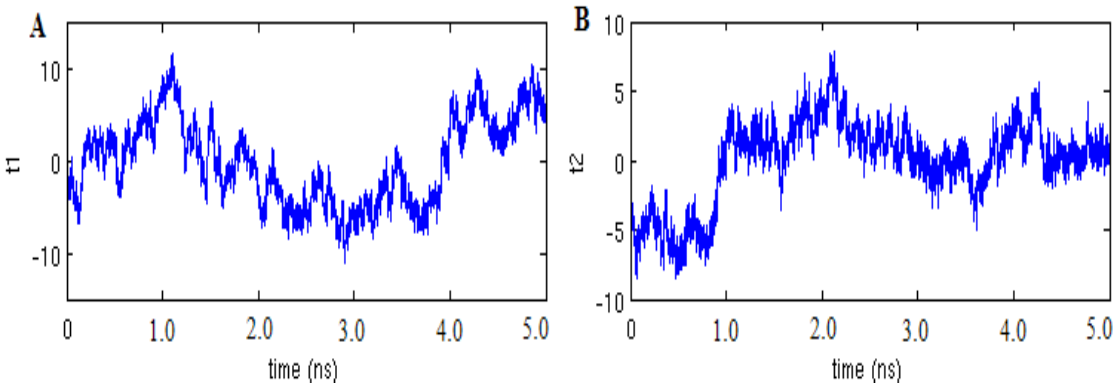

Figure A.19. The projection of the α7 helix motions on (A) PC1 (B) and PC2 in TMD2.

APPENDIX B : RESULTS OF THE TMD5 SIMULATION ANALYSIS

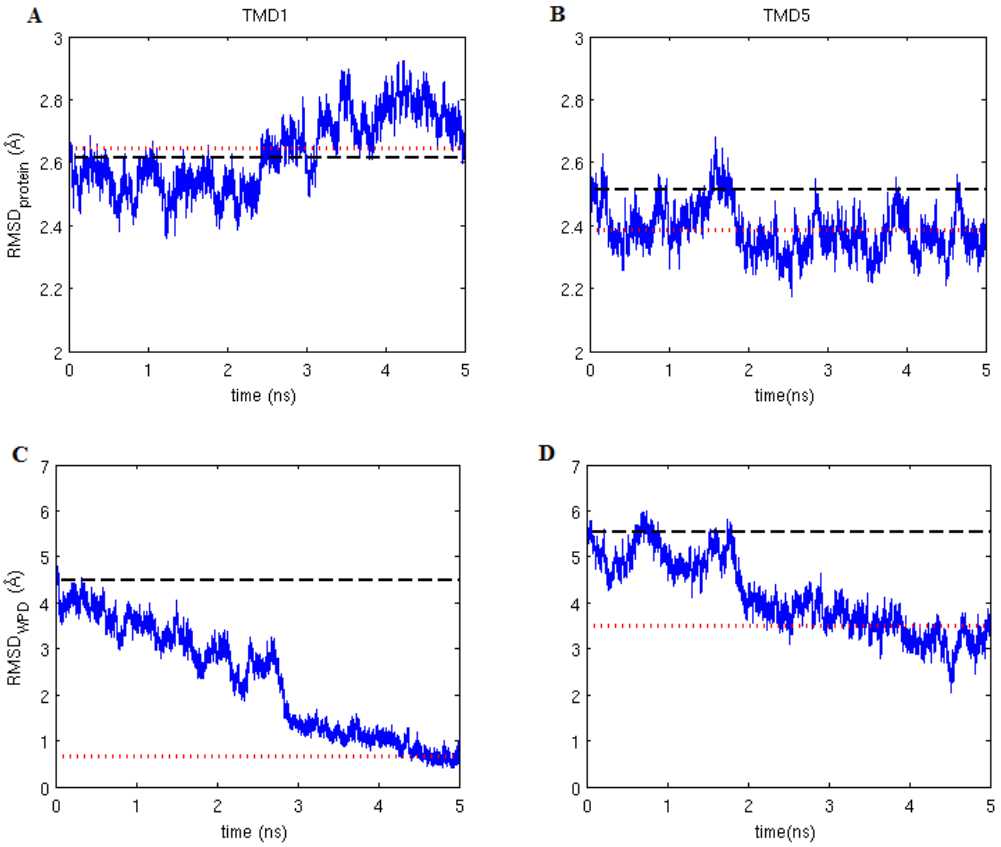

Figure B.1. RMSD of all protein and only WPD loop atoms relative to the target structure during (A), (C) TMD1 and (B), (D) TMD5 simulations (dashed line represents the initial value, dotted line represents the final value).

Table B.1. The Structure Indices Corresponding to the PTP1B Crystal Structure PDB ID

Structure Index	PTP1B Structures (PDB ID)	Structure Index	PTP1B Structures (PDB ID)
1	3eax	41	1q6j
2	3eb1	42	1q6m
3	2zmm	43	1pyn
4	2zn7	44	1q1m
5	3d9c	45	1pxh
6	2qbp	46	1oem
7	2qbq	47	1oes
8	2qbr	48	1oet
9	2qbs	49	1oeu
10	2veu	50	1oev
11	2vev	51	1nz7
12	2vew	52	10ny
13	2vex	53	1onz
14	2vey	54	1nl9
15	2nt7	55	1nny
16	2nta	56	1no6
17	2cne	57	1nwl
18	2cnf	58	1kak
19	2cng	59	1kav
20	2cnh	60	1l8g
21	2h4g	61	1g7g
22	2h4k	62	1g7f
23	2hb1	63	1een
24	2cm2	64	1eeo
25	2cm7	65	1c83
26	2cm8	66	1c84
27	2cma	67	1c85
28	2cmb	68	1c86
29	2cmc	69	1c87
30	2f6z	70	1c88
31	2b07	71	1eav
32	2bgd	72	1bzc
33	2bge	73	1bzh
34	1l8k	74	1bzj
35	1wax	75	1q5y
36	1xbo	76	1pty
37	1t48	77	1ptt
38	1t49	78	1ptu
39	1t4j	79	1ptv
40	1qxk	80	2hnp
		81	2hnq

APPENDIX C : RESULTS OF THE ANALYSIS OF TMD6 AND TMD7 SIMULATIONS

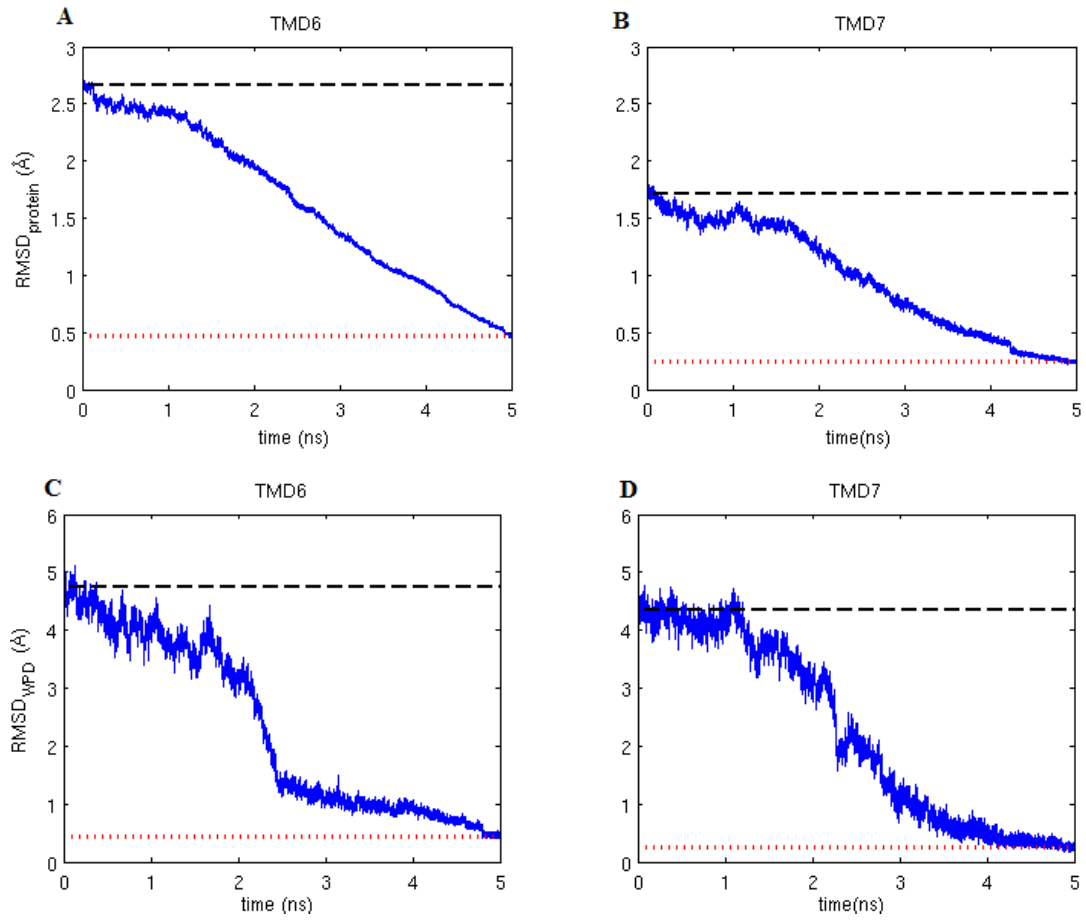

Figure C.1. RMSD of all protein and only WPD loop atoms relative to the target structure during (A), (C) TMD6 and (B), (D) TMD7 simulations (Dashed line represents the initial value, dotted line represents the final value).

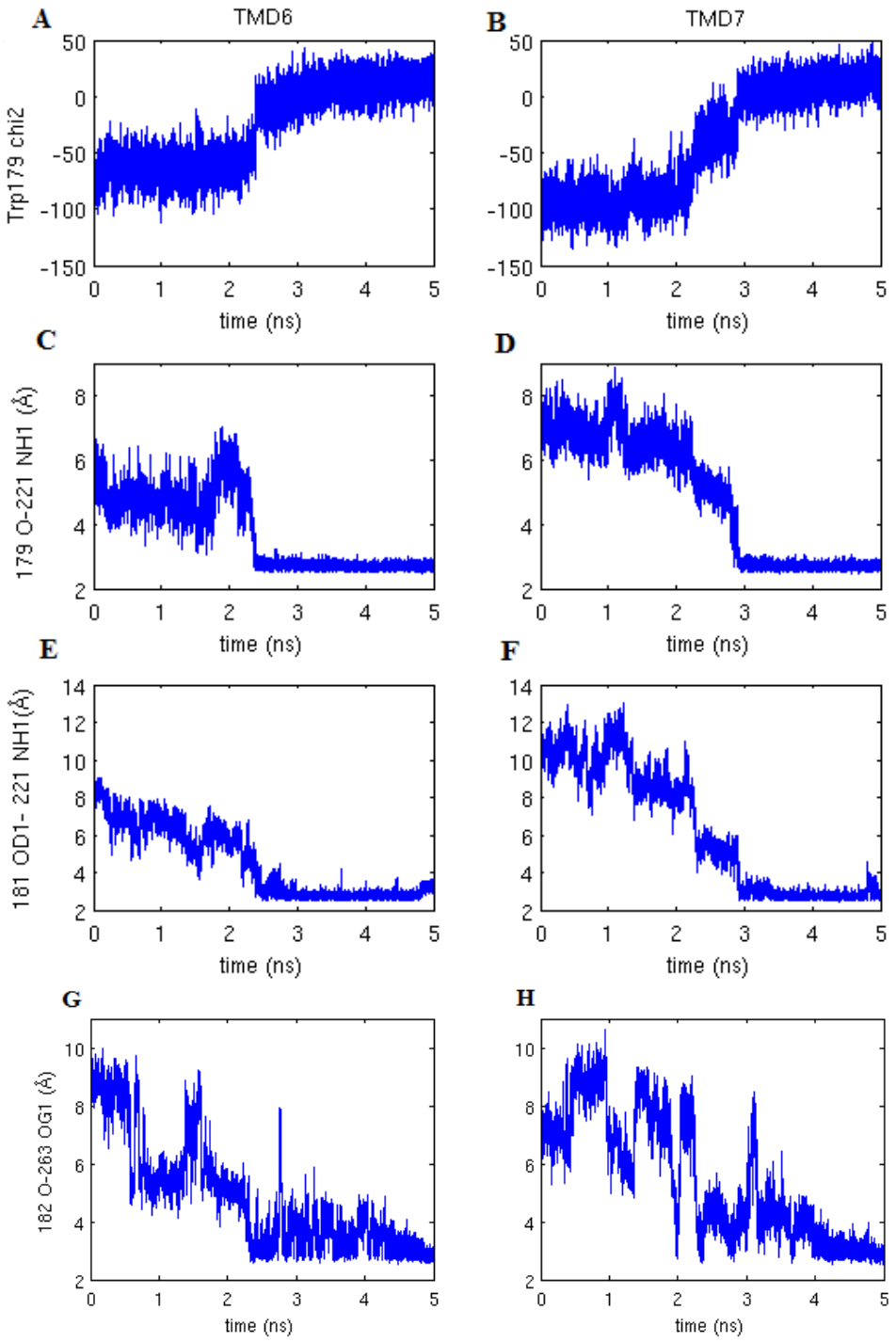

Figure C.2. The changes in (A), (B) χ2 of Trp179, (C), (D) Trp179 O - Arg221 NH1 distance, (E), (F) Asp181 (OD1) OD2 - Arg221 (NH1) NH2 distance, (G), (H) and Phe182 O – Thr263 OG1 distance in TMD6 and TMD7.

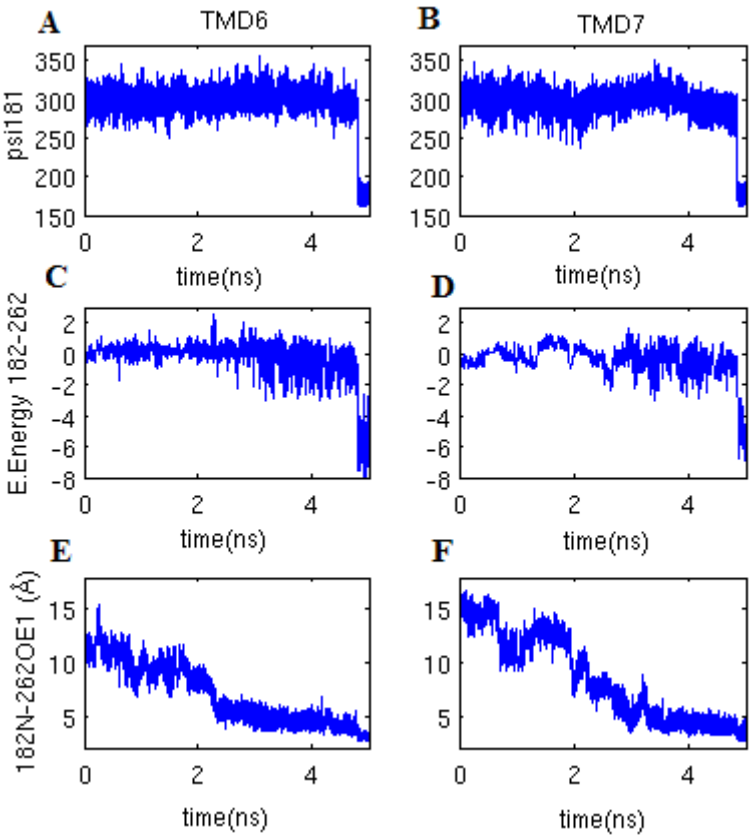

Figure C.3. Time evolution of (A), (B) Ψ backbone dihedral angle of Asp181, (C), (D) the electrostatic energy between Phe182-Gln262, (E), (F) and the distance between Phe182 N atom-Gln262 OE1 atom in TMD6 and TMD7.

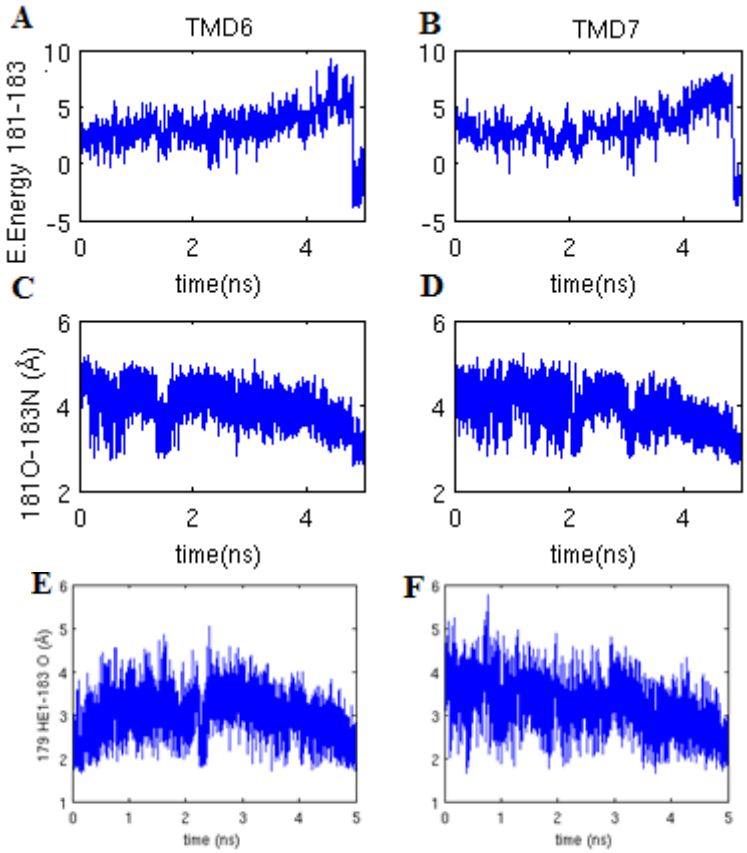

Figure C.4. Time profiles of (A), (B) the electrostatic energy formed between Asp181-Gly183, (C), (D) the distance between Asp181 O atom-Gly183 N atom, (E), (F) and the distance between Trp179 HE1-Gly183 O atom in TMD6 and TMD7.

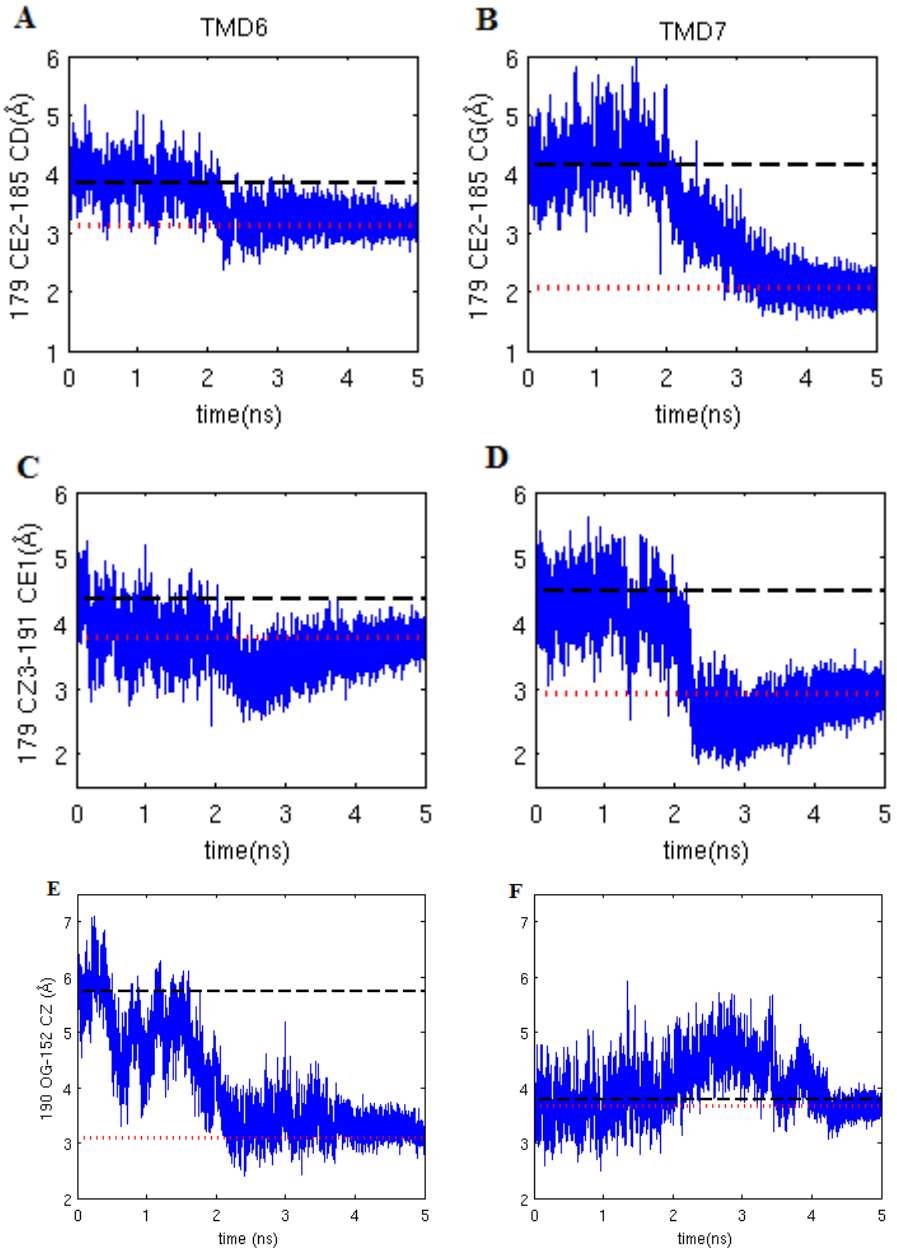

Figure C.5. Time profile of (A), (B) the distance between Trp179 CE2 atom- Phe185 CD atom, (C), (D) the distance between Trp179 CZ3 atom- Phe191 CE1, (E), (F) and the distance between Ser190 O atom- Tyr152 CD1 in TMD 6 and TMD7.

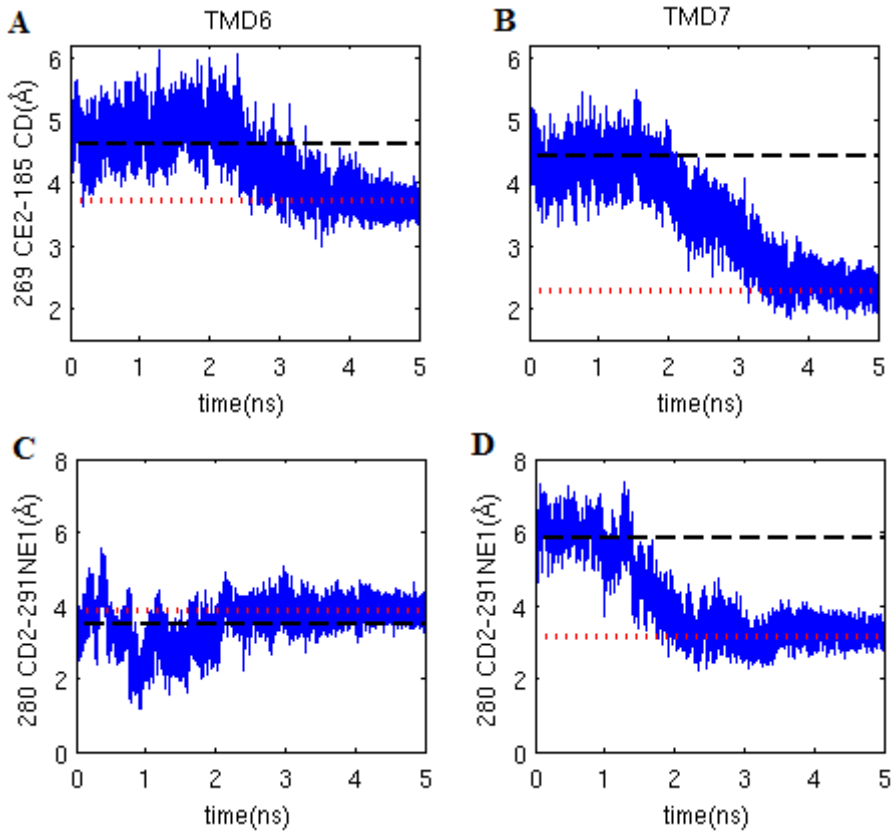

Figure C.6. Time evolution of (A), (B) the distance between Trp269 CE2 atom- Phe185 CD atom and (C), (D) the distance between Trp291 NE1 atom- Phe280 CD2 atom in TMD6 and TMD7.

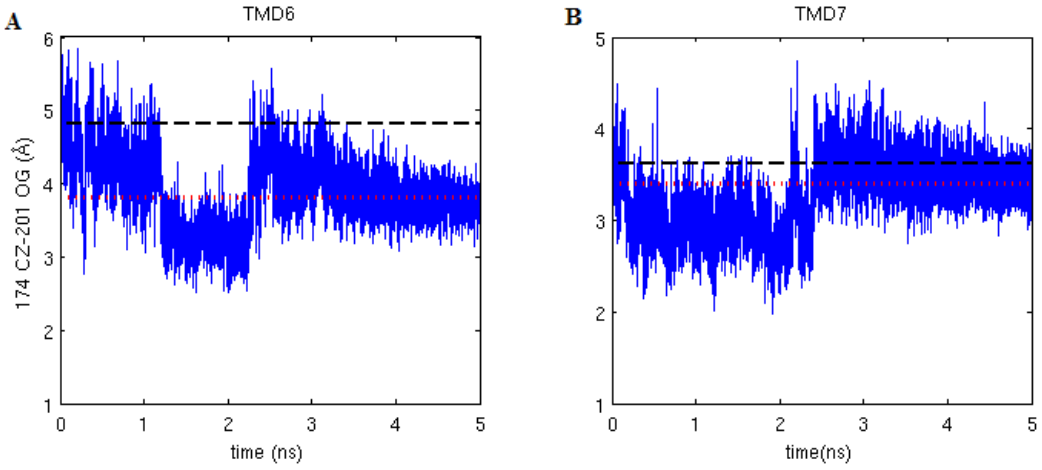

Figure C.7. Time evolution of the distance between Phe174 CZ atom- Ser201 OG atom in (A) TMD6 and (B) TMD7.

APPENDIX D: INPUT FILES USED IN COMPUTATIONAL STUDIES

NAMD Configuration File (TMD Simulation)

```
#############################################################
#############################################################
### TMD simulation of 2f6f in a water box               ###
#############################################################
#############################################################

set pdb                2f6f
set outputfile         2f6f.2f6fto1sug
set temperature        300

#############################################################
## INPUT FILES (and TEMPERATURE)                          ##
#############################################################

structure          /home/kb407/burcu/2f6fto1sugk300/ionized.psf
coordinates        /home/kb407/burcu/2f6fto1sugk300/ionized.pdb
bincoordinates     2f6f.md14.coor
binvelocities      2f6f.md14.vel
extendedSystem     2f6f.md14.xsc

############################################################
## SIMULATION PARAMETERS                                  ##
############################################################

paraTypeCharmm         on
parameters         /home/kb407/burcu/TOPPAR/par_all27_prot_lipid.prm

parameters         /home/kb407/burcu/TOPPAR/thiolate.prm

#############################################################
## STANDARD OUTPUT                                        ##
#############################################################

outputname         $outputfile
binaryoutput       yes
outputEnergies     500
outputPressure     500
```

```
#####################################################################
## FORCEFIELD PARAMETERS                                           ##
#####################################################################

# Related with cut-offs
exclude           scaled1-4
1-4scaling        1.0
cutoff            12.

switching         on
switchdist        10.

pairlistdist      14.
stepspercycle     10
nonbondedFreq     1
fullElectFrequency  2

# No ShakeH and SETTLE for any molecule
rigidBonds        none
useSettle         off


#####################################################################
## PERIODIC BOUNDARY CONDITIONS                                    ##
#####################################################################

# Wrap the water molecules
margin          2.5
wrapWater       on
wrapAll                 on

wrapNearest     off


# PME (for full-system periodic electrostatics)
# Grid sizes: 48  54 60 64 72 75 80 81 90
PME             yes
PMEGridSizeX      90
PMEGridSizeY      80
PMEGridSizeZ      72
```

```
#############################################################
## INTEGRATION PARAMETERS - related to MD sim.           ##
#############################################################
firsttimestep     0
timestep          1

#############################################################
## OUTPUT FILES - related to MD sim.                    ##
#############################################################
restartname       restart14
restartfreq       1000
DCDfile           2f6fto1sug_k300.dcd
DCDfreq           500
XSTfile           celldim.md14.xst
XSTfreq           500


#############################################################
## TEMPERATURE CONTROL (Langevin) - related to MD sim.     ##
#############################################################

langevin          on
langevinDamping     5.0
langevinTemp        $temperature
langevinHydrogen    off


#############################################################
## PRESSURE CONTROL (Langevin) - related to MD sim.        ##
#############################################################

useGroupPressure      yes ;# needed for rigidBonds
useFlexibleCell       no
useConstantArea       no

langevinPiston        on
langevinPistonTarget  1.01325 ;#  in bar -> 1 atm
langevinPistonPeriod  100.
langevinPistonDecay   50.
LangevinPistonTemp    $temperature


#############################################################
## EXECUTION SCRIPT                                    ##
#############################################################
# TMD
   TMD on

   TMDk 300
```

```
   TMDOutputFreq 1

   TMDFile /home/kb407/burcu/2f6fto1sugk300/target_tmdready.pdb

   TMDFirstStep 0

   TMDLastStep 5000000

   TMDInitialRMSD 0

   TMDFinalRMSD 0


#############################################################
## Molecular Dynamics                                      ##
#############################################################
run               5000000
```

PGN File used to build the PSF file

1SUG crystal structure

```
package require psfgen
topology top_all27_prot_lipid_na.inp
topology thiolate.inp
pdbalias residue HIS HSE
pdbalias atom ILE CD1 CD
segment PTP {pdb 1sugp.pdb

  mutate  25  HSD
  mutate  60  HSD
  mutate  94  HSD
  mutate  175 HSP
  mutate  214 HSP

  mutate  215 CYM
}

coordpdb 1sugp.pdb PTP
guesscoord

writepdb 1sug.pdb
writepsf 1sug.psf
```

2F6F crystal structure psf generation script

```
package require psfgen
topology /home/kb407/burcu/TOPPAR/top_all27_prot_lipid_na.inp
topology /home/kb407/burcu/TOPPAR/thiolate.inp
pdbalias residue HIS HSE
pdbalias atom ILE CD1 CD
segment PTP {pdb 2f6fp.pdb

  mutate  25  HSD
  mutate  60  HSD
  mutate  94  HSD
  mutate  175 HSP
  mutate  214 HSP

  mutate  215 CYM
  mutate  295 SER
}

coordpdb 2f6fp.pdb PTP
guesscoord

writepdb 2f6f.pdb
writepsf 2f6f.psf
```

Topology and parameters for thiolate

Thiolate.prm

```
read param card append
* additional parameters for CYM
*

BONDS
CS   CT1   222.500     1.5380 ! From CT2  CT1

ANGLES
SS   CS   CT1    58.000   112.5000 ! From S    CT2  CT1 (NF)
CS   CT1  C      52.000   108.0000 !From CT2  CT1  C (NF)
HA   CS   CT1    33.430   110.10   22.53   2.17900 !From HA   CT2  CT1 (NF)
HB   CT1  CS     35.000   111.0000 ! From HB   CT1  CT2 (NF)
NH1  CT1  CS     70.000   113.5000 ! From NH1  CT1  CT2 (NF)


DIHEDRALS
CS CT1 NH1 C 1.8000 1 0.00
H NH1 CT1 CS 0.0000 1 0.00 ! For CYM/thiolate
N C CT1 CS 0.0000 1 0.00 ! For CYM/thiolate
```

```
NH1 C CT1 CS 0.0000 1 0.00 ! For CYM/thiolate
NH1 C CS HA 0.0000 3 0.00 ! For CYM/thiolate
O C CT1 CS 1.4000 1 0.00 ! For CYM/thiolate
X CT1 CS X 0.2000 3 0.00 ! For CYM/thiolate
```

Tcl Scripts used in VMD

Distance.tcl

```
proc distance {molid residue1 type1 residue2 type2} {
set a1 [atomselect $molid "protein and resid $residue1 and name $type1"]
set a2 [atomselect $molid "protein and resid $residue2 and name $type2"]
set ind1 [$a1 get index]
set ind2 [$a2 get index]
label add Bonds $molid/$ind1 $molid/$ind2 }
```

Distance_modified.tcl (used to find the distance between the atoms of the residues in different structures)

```
proc FindDist3 {molid1 residue1 type1 molid2 residue2 type2} {
  #
  # Find the distance between two selected atoms and save them in a .txt
  # file
  #
  # Format: FindDist molid residue1 name1 residue2 name2
  #
  # ex: FindDist 0 178 CA 2 214 SG
  #

  set noframes [molinfo $molid2 get numframes]

  if { $molid1 == "top" } {
     set molid1 [molinfo top]
  }

  if { $molid2 == "top" } {
     set molid2 [molinfo top]
  }

  set a1 [atomselect $molid1 "resid $residue1 and name $type1"]
  set ind1 [$a1 get index]

  set fid [open "dist${residue1}${type1}_${residue2}${type2}.txt" w]

  set dist {}

  for { set i 0 } { $i <= $noframes } { incr i } {
    set a2 [atomselect $molid2 "resid $residue2 and name $type2" frame $i]
    set ind2 [$a2 get index]
```

```
    lappend dist [measure bond "{$ind2 $molid2} {$ind1 $molid1}" frame $i]
  }
  puts $fid $dist
  close $fid
}
```

Dihedral.tcl (used to find Φ and Ψ angles)

```
proc FindDihed { molid residue } {
  #
  # Find the phi-psi of a residue
  #
  # If the molid contains multiple frames, dihedral angles
  # trajectory will be written to a file named dihed<resid>.txt.
  #
  # Ex: FindDihed 0 178
  #     FindDihed top 182
  #

  set noframes [molinfo $molid get numframes]

  if { $molid == "top" } {
     set molid [molinfo top]
  }

  set prevres [expr $residue-1]
  set nextres [expr $residue+1]
  set a1 [atomselect $molid "resid $prevres and name C"]
  set a2 [atomselect $molid "resid $residue and name N"]
  set a3 [atomselect $molid "resid $residue and name CA"]
  set a4 [atomselect $molid "resid $residue and name C"]
  set a5 [atomselect $molid "resid $nextres and name N"]
  set ind1 [$a1 get index]
  set ind2 [$a2 get index]
  set ind3 [$a3 get index]
  set ind4 [$a4 get index]
  set ind5 [$a5 get index]

  if {$noframes == 1} {

     set dihed1 [measure dihed " {$ind1 $molid} {$ind2 $molid} {$ind3 $molid} {$ind4
$molid} "]
     set dihed2 [measure dihed " {$ind2 $molid} {$ind3 $molid} {$ind4 $molid} {$ind5
$molid} "]
     puts "phi = ${dihed1}   psi = ${dihed2}"
     label add Dihedrals $molid/$ind1 $molid/$ind2 $molid/$ind3 $molid/$ind4
     label add Dihedrals $molid/$ind2 $molid/$ind3 $molid/$ind4 $molid/$ind5
  } else {
     set fid [open "dihed${residue}.txt" "w"]
```

```
    for { set i 0 } { $i < $noframes } { incr i } {
      molinfo $molid set frame $i
      set dihed1 [measure dihed " {$ind1 $molid} {$ind2 $molid} {$ind3 $molid} {$ind4
$molid} "]
      set dihed2 [measure dihed " {$ind2 $molid} {$ind3 $molid} {$ind4 $molid} {$ind5
$molid} "]
      puts $fid "$dihed1  $dihed2"
    }
    close $fid
  }
}
```

REFERENCES

1. Tonks, N., 2003, "PTP1B: From the Sidelines to the Front Lines", *Febs Letters*, Vol. 546, pp. 140-148.

2. Neel, B., and N. Tonks, 1997, "Protein Tyrosine Phosphatases in Signal Transduction", *Current Opinion in Cell Biology*, Vol. 9, pp. 193-204.

3. Andersan, J., O. Mortensen, G. Peters, P. Jansen, H. Andersen, T. NP, and H. Moller, 2001, "Structural and Evalutionary Relationships Among Protein Tyrosine Phosphatase Domains", *Molecular And Cellular Biology*, Vol. 21, pp. 7117-7136.

4. Cheng, A., N. Uetani, P. Simoncic, V. Chaubey, A. Lee-Loy, C. McGlade, B. Kennedy, and M. Tremblay, 2002, "Attenuation of Leptin Action and Regulation of Obesity by Protein Tyrosine Phosphatase 1B", *Developmental Cell*, Vol. 2, pp. 497-503.

5. Combs, A. P., "Recent Advances in the Discovery of Competitive Protein Tyrosine Phosphatase 1B Inhibitors for the Treatment of Diabetes, Obesity, and Cancer", *Journal of Medicinal Chemistry*, Vol. 53, pp. 2333-2344.

6. Barford, D., A. Flint, and N. Tonks, 1994, "Crystal-Structure of Human Protein-Tyrosine-Phosphatase 1B", *Science*, Vol. 263, pp. 1397.

7. Barr, A. J., E. Ugochukwu, W. H. Lee, O. N. F. King, P. Filippakopoulos, I. Alfano, P. Savitsky, N. A. Burgess-Brown, S. Muller, and S. Knapp, 2009, "Large-Scale Structural Analysis of the Classical Human Protein Tyrosine Phosphatome", *Cell*, Vol. 136, pp. 352-363.

8. Jia, Z., D. Barford, A. Flint, and N. Tonks, 1995, "Structural Basis for Phosphotyrosine Peptide Recognition by Protein-Tyrosine-Phosphatase 1B", *Science*, Vol. 268, pp. 1754-1758.

9. Pedersen, A., G. Peters, K. Moller, L. Iversen, and J. Kastrup, 2004, "Water-Molecule Network and Active-Site Flexibility of Apo Protein Tyrosine Phosphatase 1B", *Acta Crystallographica Section D-Biological Crystallography*, Vol. 60, pp. 1527-1534.

10. Kamerlin, S., R. Rucker, and S. Boresch, 2007, "A Molecular Dynamics Study of WPD-loop Flexibility in PTP1B", *Biochemical and Biophysical Research Communications*, Vol. 356, pp. 1011-1016.

11. Ottana, R., R. Maccari, R. Ciurleo, P. Paoli, M. Jacomelli, G. Manao, G. Camici, C. Laggner, and T. Langer, 2009, "5-Arylidene-2-phenylimino-4-thiazolidinones as PTP1B and LMW-PTP Inhibitors", *Bioorganic & Medicinal Chemistry*, Vol. 17, pp. 1928-1937.

12. Zhang, S., and Z. Zhang, 2007, "PTP1B as a Drug Target: Recent Developments in PTP1B Inhibitor Discovery", *Drug Discovery Today Theurapeutic Strategies*, Vol. 12, pp. 373-381.

13. Montalibet, J., and B. Kennedy, 2005, "Therapeutic Strategies for Targeting PTP1B in Diabetes", *Drug Discovery Today Theurapeutic Strategies*, Vol. 2, pp. 129-135.

14. Zabolotny, J., K. Bence-Hanulec, A. Stricker-Krongrad, F. Haj, Y. Wang, Y. Minokoshi, Y. Kim, J. Elmquist, L. Tartaglia, B. Kahn, and B. Neel, 2002, "PTP1B Regulates Leptin Signal Transduction in Vivo", *Developmental Cell*, Vol. 2, pp. 489-495.

15. Schubert, H., E. Fauman, J. Stuckey, J. Dixon, and M. Saper, 1995, "A Ligand-Induced Conformational Change in the Yersinia Protein-Tyrosine-Phosphatase", *Protein Science*, Vol. 4, pp. 1904-1913.

16. Zhang, Z., and S. Lee, 2003, "PTP1B Inhibitors as Potential Therapeutics in the Treatment of Type 2 Diabetes and Obesity", *Expert Opinion on Investigational Drugs*, Vol. 12, pp. 223-233.

17. Fang, L., H. Zhang, W. Cui, and M. Ji, 2008, "Studies of the Mechanism of Selectivity of Protein Tyrosine Phosphatase 113 (PTP1B) Bidentate Inhibitors Using Molecular Dynamics Simulations and Free Energy Calculations", *Journal of Chemical Information and Modeling*, Vol. 48, pp. 2030-2041.

18. Wang, J., K. Gong, D. Wei, Y. Li, and K. Chou, 2009, "Molecular Dynamics Studies on the Interactions of PTP1B with Inhibitors: From the First Phosphate-Binding Site to the Second One", *Protein Engineering Design & Selection*, Vol. 22, pp. 349.

19. Johnson, T., J. Ermolieff, and M. Jirousek, 2002, "Protein Tyrosine Phosphatase 1B Inhibitors for Diabetes", *Nature Reviews Drug Discovery*, Vol. 1, pp. 696-709.

20. Liu, G., J. Tan, C. Niu, J. Shen, X. Luo, X. Shen, K. Chen, and H. Jiang, 2006, "Molecular Dynamics Simulations of Interaction Between Protein-Tyrosine Phosphatase 1B and a Bidentate Inhibitor", *Acta Pharmacologica Sinica*, Vol. 27, pp. 100-110.

21. Ala, P., L. Gonneville, M. Hillman, M. Becker-Pasha, M. Wei, B. Reid, R. Klabe, E. Yue, B. Wayland, B. Douty, P. Polam, Z. Wasserman, M. Bower, A. Combs, T. Burn, G. Hollis, and R. Wynn, 2006, "Structural Basis for Inhibition of Protein-Tyrosine Phosphatase 1B by Isothiazolidinone Heterocyclic Phosphonate Mimetics", *Journal of Biological Chemistry*, Vol. 281, pp. 32784-32795.

22. Wiesmann, C., K. Barr, J. Kung, J. Zhu, D. Erlanson, W. Shen, B. Fahr, M. Zhong, L. Taylor, M. Randal, R. McDowell, and S. Hansen, 2004, "Allosteric Inhibition of Protein Tyrosine Phosphatase 1B", *Nature Structural & Molecular Biology*, Vol. 11, pp. 730-734.

23. Kamerlin, S., R. Rucker, and S. Boresch, 2006, "A Targeted Molecular Dynamics Study of WPD loop Movement in PTP1B", *Biochemical and Biophysical Research Communications*, Vol. 345, pp. 1161-1166.

24. Montalibet, J., K. Skorey, D. McKay, G. Scapin, E. Asante-Appiah, and B. Kennedy, 2006, "Residues Distant from the Active Site Influence Protein-Tyrosine Phosphatase 1B Inhibitor Binding", *Journal of Biological Chemistry*, Vol. 281, pp. 5258-5266.

25. Iversen, L. F., K. B. Moller, A. K. Pedersen, G. H. Peters, A. S. Petersen, H. S. Andersen, S. Branner, S. B. Mortensen, and N. P. H. Moller, 2002, "Structure Determination of T Cell Protein-Tyrosine Phosphatase", *Journal of Biological Chemistry*, Vol. 277, pp. 19982-19990.

26. Eswaran, J., J. Von Kries, B. Marsden, E. Longman, J. Debreczeni, E. Ugochukwu, A. Turnbull, W. Lee, S. Knapp, and A. Barr, 2006, "Crystal Structures and Inhibitor Identification for PTPN5, PTPRR and PTPN7: A Family of Human Mapk-Specific Protein Tyrosine Phosphatases", *Biochemical Journal*, Vol. 395, pp. 483-491.

27. Debreczeni, J., J. Eswaran, E. Longman, A. Barr, and S. Knapp, 2006, "The Crystal Structure of Human Receptor Protein Tyrosine Phosphatase Kappa Phosphatase Domain 1", *Protein Science*, Vol. 15, pp. 1500-1505.

28. Yang, J., Z. Cheng, T. Niu, X. Liang, Z. Zhao, and G. Zhou, 2000, "Structural Basis for Substrate Specificity of Protein-Tyrosine Phosphatase SHP-1", *Journal Of Biological Chemistry*, Vol. 275, pp. 4066-4071.

29. Scapin, G., S. Patel, J. Becker, Q. Wang, C. Desponts, D. Waddleton, K. Skorey, W. Cromlish, C. Bayly, M. Therien, J. Gauthier, C. Li, C. Lau, C. Ramachandran, B. Kennedy, and E. Asante-Appiah, 2003, "The Structural Basis for the Selectivity of Benzotriazole Inhibitors of PTP1B", *Biochemistry*, Vol. 42, pp. 11451-11459.

30. Wilson, D., Z. Wan, W. Xu, S. Kirincich, B. Follows, D. Joseph-McCarthy, K. Foreman, A. Moretto, J. Wu, M. Zhu, E. Binnun, Y. Zhang, M. Tam, D. Erbe, J. Tobin, X. Xu, L. Leung, A. Shilling, S. Tam, T. Mansour, and J. Lee, 2007, "Structure-based Optimization of Protein Tyrosine Phosphatase 1B Inhibitors: From the Active Site to the Second Phosphotyrosine Binding Site", *Journal of Medicinal Chemistry*, Vol. 50, pp. 4681-4698.

31. Romsicki, Y., G. Scapin, V. Beaulieu-Audy, S. Patel, J. Becker, B. Kennedy, and E. Asante-Appiah, 2003, "Functional Characterization and Crystal Structure of the C215D Mutant of Protein-Tyrosine Phosphatase-1B", *Journal of Biological Chemistry*, Vol. 278, pp. 29009-29015.

32. Li, S., R. Depetris, D. Barford, J. Chernoff, and S. Hubbard, 2005, "Crystal Structure of a Complex Between Protein Tyrosine Phosphatase 1B and the Insulin Receptor Tyrosine Kinase", *Structure*, Vol. 13, pp. 1643-1651.

33. Kale, L., R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan, and K. Schulten, 1999, "NAMD2: Greater Scalability for Parallel Molecular Dynamics", *Journal of Computational Physics*, Vol. 151, pp. 283-312.

34. Peters, G. H., T. M. Frimurer, J. N. Andersen, and O. H. Olsen, 1999, "Molecular Dynamics Simulations of Protein-Tyrosine Phosphatase 1B. I. Ligand-induced Changes in the Protein Motions", *Biophysical Journal*, Vol. 77, pp. 505-515.

35. Peters, G. H., T. M. Frimurer, J. N. Andersen, and O. H. Olsen, 2000, "Molecular Dynamics Simulations of Protein-Tyrosine Phosphatase 1B. II. Substrate-enzyme Interactions and Dynamics", *Biophysical Journal*, Vol. 78, pp. 2191-2200.

36. Bharatham, K., N. Bharatham, Y. J. Kwon, and K. W. Lee, 2008, "Molecular Dynamics Simulation Study of PTP1B with Allosteric Inhibitor and its Application in Receptor Based Pharmacophore Modeling", *Journal of Computer-Aided Molecular Design*, Vol. 22, pp. 925-933.

37. Schlitter, J., M. Engels, P. Kruger, E. Jacoby, and A. Wollmer, 1993, "Targeted Molecular-Dynamics Simulation of Conformational Change-Application to the T[--]R Transition in Insulin", *Molecular Simulation*, Vol. 10, pp. 291-&.

38. Lei, M., J. Velos, A. Gardino, A. Kivenson, M. Karplus, and D. Kern, 2009, "Segmented Transition Pathway of the Signaling Protein Nitrogen Regulatory Protein C", *Journal of Molecular Biology*, Vol. 392, pp. 823-836.

39. Ma, J. P., and M. Karplus, 1997, "Molecular Switch in Signal Transduction: Reaction Paths of the Conformational Changes in Ras p21", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 94, pp. 11905-11910.

40. Cheng, X. L., H. L. Wang, B. Grant, S. M. Sine, and J. A. McCammon, 2006, "Targeted Molecular Dynamics Study of C-loop Closure and Channel Gating in Nicotinic Receptors", *Plos Computational Biology*, Vol. 2, pp. 1173-1184.

41. Zou, J., Y. D. Wang, F. X. Ma, M. L. Xiang, B. Shi, Y. Q. Wei, and S. Y. Yang, 2008, "Detailed Conformational Dynamics of Juxtamembrane Region and Activation Loop in C-Kit Kinase Activation Process", *Proteins-Structure Function and Bioinformatics*, Vol. 72, pp. 323-332.

42. Xu, J., and K. S. Matthews, 2009, "Flexibility in the Inducer Binding Region is Crucial for Allostery in the Escherichia Coli Lactose Repressor", *Biochemistry*, Vol. 48, pp. 4988-4998.

43. Allen, M., 2004, "Introduction to Molecular Dynamics Simulation" ", *Nic Series*, Vol. 23, pp. 1-28.

44. Berman, H., J. Westbrook, Z. Feng, G. Gilliland, T. Bhat, H. Weissig, I. Shindyalov, and P. Bourne, 2000, "The Protein Data Bank", *Nucleic Acids Research*, Vol. 28, pp. 235-242.

45. Adcock, S., and J. McCammon, 2006, "Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins", *Chemical Reviews*, Vol. 106, pp. 1589-1615.

46. Aci, S., S. Mazier, and D. Genest, 2005, "Conformational Pathway for the Kissing Complex - Extended Dimer Transition of the SL1 Stem-loop from Genomic HIV-1

RNA as Monitored by Targeted Molecular Dynamics Techniques", *Journal of Molecular Biology*, Vol. 351, pp. 520-530.

47. Groves, M. R., Z. J. Yao, P. P. Roller, T. R. Burke, and D. Barford, 1998, "Structural Basis for Inhibition of the Protein Tyrosine Phosphatase 1B by Phosphotyrosine Peptide Mimetics", *Biochemistry*, Vol. 37, pp. 17773-17783.

48. Xin, Z. L., G. Liu, C. Abad-Zapatero, Z. H. Pei, P. J. Hajduk, S. J. Ballaron, M. A. Stashko, T. H. Lubben, J. M. Trevillyan, and M. R. Jirousek, 2003, "Identification of a Monoacid-Based, Cell Permeable, Selective Inhibitor of Protein Tyrosine Phosphatase 1B", *Bioorganic & Medicinal Chemistry Letters*, Vol. 13, pp. 3947-3950.