ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE ENGINEERING AND TECHNOLOGY

MOLECULAR DYNAMICS SIMULATION ANALYSIS OF HIS226 MUTATION ON THE DYNAMICS OF THE ATPASE DOMAIN OF DNAK

M.Sc. THESIS

Elif ÇAKMAK

Department of Molecular Biology-Genetics and Biotechnology

Molecular Biology-Genetics and Biotechnology Programme

JANUARY 2017

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE ENGINEERING AND TECHNOLOGY

MOLECULAR DYNAMICS SIMULATION ANALYSIS OF HIS226 MUTATION ON THE DYNAMICS OF THE ATPASE DOMAIN OF DnaK

M.Sc. THESIS

Elif ÇAKMAK
(521131130)

Department of Molecular Biology-Genetics and Biotechnology

Molecular Biology-Genetics and Biotechnology Programme

Thesis Advisor: Assoc. Prof. Dr. Gizem DİNLER DOĞANAY
Thesis Co-Advisor: Assist. Prof. Dr. Sefer BADAY

JANUARY 2017

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

HİS226 MUTASYONUNUN DNaK ATPAZ DOMENİ ÜZERİNDEKİ DİNAMİKLERİNİN, MOLEKÜLER DİNAMİK SİMULASYONLARI İLE ANALİZİ

YÜKSEK LİSANS TEZİ

Elif ÇAKMAK
(521131130)

Moleküler Biyoloji-Genetik ve Biyoteknoloji Anabilim Dalı

Moleküler Biyoloji-Genetik ve Biyoteknoloji Programı

Tez Danışmanı: Doç. Dr. Gizem DİNLER DOĞANAY
Eş-Danışman : Yrd. Doç. Dr. Sefer BADAY

OCAK 2017

Elif ÇAKMAK, a M.Sc. student of ITU Graduate School of Science Engineering and Technology student ID 521131130, successfully defended the thesis entitled "MOLECULAR DYNAMICS SIMULATION ANALYSIS OF HIS226 MUTATION ON THE DYNAMICS OF THE ATPASE DOMAIN OF DNAK", which she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor :   **Assoc. Prof. Dr. Gizem DINLER DOGANAY** ....
ISTANBUL Technical University

Co-advisor :   **Assist. Prof. Dr. Sefer BADAY**   ....
ISTANBUL Technical University

Jury Members :   **Assoc. Prof. Dr. Ceren CIRACI**   ....
ISTANBUL Technical University

Assoc. Prof. Dr Elif Damla ARISAN   ....
ISTANBUL Kultur University

Assist. Prof. Dr. Ozge SENSOY   ....
ISTANBUL Medipol University

Date of Submission : 05 January 2017
Date of Defense   : 09 January 2017

To my lovely family,

FOREWORD

January 2017   Elif ÇAKMAK
                                                                                                        (Chemist)

TABLE OF CONTENTS

ABREVIATIONS

NBD   : Nucleotide Binding Domain
SBD   : Substrate Binding Domain
3-D   : Three Dimensional
Hsp   : Heat Shock Protein
Hip   : Protonated Histidine
JDP   : J-Domain Protein
ATP   : Adenosine Triphosphate
ADP   : Adenosine Diphosphate
BPTI   : Bovine Pancreatic Trypsin Inhibitor
NMR   : Nuclear Magnetic Resonance
MD   : Molecular Dynamics
IC   : Initial Condition
BC   : Boundary Condition
MM   : Molecular Mechanics
CHARMM   : Chemistry at Harvard Macromolecular Mechanics
DNA   : Deoxiribonucleic Acid
RNA   : Ribonucleic Acid
NAMD   : Nanoscale Molecular Dynamics
RMSD   : Root Mean Square Deviation
PCA   : Principle Component Analysis
PDB   : Protein Data Bank
SD   : Steepest Decent
ABNR   : Adopted Basis Newton-Raphson Model
PME   : Particle Mesh Ewald
VMD   : Visual Molecular Dynamics

SYMBOLS

K_b	: Bond Force Constant
K_{UB}	: Urey-Bradley Force Constant
K_θ	: Angle Force Constant
K_ϕ	: Dihedral Angle Force Constant
K_{imp}	: Improper Dihedral Angle Force Constant
b	: Bond Length
S	: Urey-Bradley 1,3-distance
θ	: Bond Angle
χ	: Dihedral Angle
φ	: Improper Torsion Angle
b_0	: Bond Length at Equilibrium
S_0	: Urey-Bradley 1,3-Distance at Equilibrium
θ_0	: Bond Angle at Equilibrium
φ_0	: Improper Torsion Angle at Equilibrium
ϵ	: Lennard-Jones Well Depth
R_{min}	: Distance at the Lennard-Jones Minimum
q_i	: Partial Atomic Charge
ϵ_1	: Effective Dielectric Constant
R_{ij}	: Distance Between Atoms i and j
d	: Distance
n	: Number Of Pair Equivalent Atom
t	: Time
$r_i(t)$	: Position Vector of Particle i
m_i	: Mass of Particle i
F_i	: Force That Act on Particle i
R_g	: Radius of Gyration
r_i	: Coordinates of Atom i
R_C	: Center of Mass Coordinates
N	: Number of Non-hydrogen Atoms

LIST OF TABLES

LIST OF FIGURES

MOLECULAR DYNAMICS SIMULATION ANALYSIS OF HIS226 MUTATION ON THE DYNAMICS OF THE ATPASE DOMAIN OF DNAK

SUMMARY

Hsp70s are evolutionarily highly conserved ATP-dependent molecular chaperones which are ubiquitously expressed in the cell. They are found in three domains of life and have essential roles in cells, such as aiding proper folding of nascent polypeptides, prevention of polypeptide chains from misfolding and translocation of proteins across membranes. The diverse cellular functions of Hsp70s are based on the recognition of hydrophobic sequences of client protein. Hsp70s corporate with other co-chaperones like nucleotide exchange factors and J-domain proteins.

In our study, we used DnaK which is an *Escherichia coli* homolog of Hsp70. DnaK have an N-terminal nucleotide-binding ATPase domain (NBD) and a C-terminal substrate-binding domain (SBD). These two domains are connected by a highly conserved hydrophobic interdomain linker. There is an allosteric communication between the domains via the hydrophobic linker. Substrate affinity is regulated by ATP binding and hydrolysis, which results in conformational changes in both domains, while ATP hydrolysis is stimulated by substrate binding. In 2007, Swain et al. revealed that the conserved hydrophobic ^{389}VLLL^{392} sequence of the interdomain linker is responsible for the allosteric communication between NBD and SBD. According to this study, DnaK (1-392) behaves like the substrate-stimulated DnaK that is pH-dependent, and shows higher activity than that of the unstimulated full-length protein. In contrast, DnaK (1-388) mimics the activity of the substrate-free form of the full-length DnaK. Which amino acids in the catalytic site are responsible in allosteric communication and pH-dependent ATPase activity in the presence of linker are not enlightened so far, however there are several research trying to find out the key residues and reveal the detailed mechanism of DnaK.

In this study, the effect of linker ^{389}VLLL^{392} on the ATP-bound protein conformation and H226A mutation on ATP-bound DnaK's (1-392) construct were investigated by using molecular dynamics simulations. MD simulation trajectories were analyzed by root mean square deviation (RMSD) and, root mean square fluctuation (RMSF) analysis, also by distance measurement in a time-dependent manner and, principle component analysis. From these analysis it was found that distance between the helices which contain His226 and its neighbour helix is closer to each other in DnaK (1-388) constructs. Moreover, this study reveals that His226 contributes the stabilization of residue Thr199 which is suspicious as a phosphate acceptor after the hydrolysis of ATP.

HİS226 MUTASYONUNUN DnaK ATPaz DOMENİ ÜZERİNDEKİ DİNAMİKLERİNİN, MOLEKÜLER DİNAMİK SİMULASYONLARI İLE ANALİZİ

ÖZET

Proteinler, canlılar için yaşamsal açıdan önemli makromoleküllerdir. Sentezlendikten sonra fonksiyonel özelliklerini kazanmaları içinse doğru şekilde katlanmaları şarttır. Peptid zinciri sentezlenirken, üç boyutlu yapısını kazanması sürecinde görülen ilk etkileşimler hidrofobik etkileşimlerdir. Sonrasında ise iyonik etkileşimler,van der Waals etkileşimleri, dipol-dipol etkileşimleri, hidrojen bağları ile birlikte peptit zincirleri fonksiyonel hale geldiği üç boyutlu yapısına ulaşır. Bazı küçük proteinler bu şekilde tek başına katlanabilirken, pek çok protein katlanabilmek için şaperonlara ihtiyaç duyar. Proteinlerin düzgün katlanamaması ise, Alzheimer, Parkinson gibi çeşitli nörodejeneratif hastalıklara sebebiyet verebilir.

Şaperonlar ilk kez 1978 yılında Ron Laskey tarafından bulundu. 1987 yılında R. John Ellis tarafından yapılan çalışmalarla birlikte, bu konudaki araştırmalar hız kazandı. Eksikliklerinde nörodejeneratif hastalıkların oluştuğunun keşfedilmesi ile beraber, şaperonlar ile ilgili çalışmalar oldukça sık gündeme gelmeye başlamıştır.

Şaperonlar, normal şartlarda hücrede normal seviyelerde sentezlenir. Stres koşulları ise bu durumu değiştirebilir. Şaperonlar hücrenin aşırı ısı ve pH değişiminden olumsuz etkilenmemesini sağlamanın yanında oksidatif stres ve kimyasal stres gibi durumlara karşı da hücreyi korur. Bu gibi streslere karşı hücrenin bir savunma stratejisi olarak, şaperonların ekspresyonları artar ve tehdit altında olan hücre içi yaşamsal öneme sahip enzimlerin ve diğer proteinlerin yapısının bozularak fonksiyonel olarak zarar görmeleri engellenir.

Hsp70, şaperon ailesinin en bilinen ve en yaygın üyesidir. Şaperonlar evrimsel süreçte korunmuş moleküler şaperonlar arasında yer almaktadır. Bakteri, arkea ve ökaryotlarda olmak üzere neredeyse tüm hücrelerde bulunurlar ve hayati açıdan önemli rollere sahiptirler. Örneğin, yeni eksprese edilmiş peptit zincirlerinin katlanması, agregat oluşumunun önlenmesi ya da proteinlerin belirli organellere translokasyonu Hsp70 tarafından gerçekleştirilir. Hsp70, katlanacak proteinin hidrofobik bölgelerine bağlanarak bu bölgelerin birbirleriyle doğru olmayan şekilde etkileşmesini önler, bu sayede proteinin yanlış katlanmasını engeller. Bu sebepten ötürü, Hsp70'in proteinlerin katlanmasında katalizör görevi görmediği, katlanması gereken proteinler için uygun ortam oluşturduğu düşünülmektedir.

Çalışmamızda kullanılan *E.coli* Hsp70 homoloğu olan DnaK iki domenden oluşmaktadır. Bunlardan biri amino (N) ucunda bulunan nükleotit bağlayan ve ATPaz aktivitesi olan 44 kDa'lık Nükleotit Bağlanan Domen (NBD), diğeri ise karboksil (C) ucunda bulunan ve substrat bağlayan 25 kDa'lık Subsrat Bağlanan Domen (SBD)'dir. Bu iki domen ise oldukça korunmuş, domenler arasında bulunan

hidrofobik bir bağlaç ile bağlanır. NBD ve SBD arasında allosterik bir etkileşim mevcuttur. Bu iki domen arasındaki iletişim ise domenler arasındaki bağlaç tarafından sağlanmaktadır. Her iki domende de gerçekleşen konformasyonel değişiklikler sonucunda, ATP bağlanması ve hidrolizi substrat affinitesini düzenlerken, ATP hidrolizi ise substrat bağlanması ile tetiklenmektedir. Yapılan kristalografi ve NMR çalışmaları sonucunda elde edilen yapılarda, ADP bağlı halde, bu iki domen ve bağlaç birbirinden ayrı şekilde gözlemlenmektedir. Bu durumda SBD'nin substrata olan afinitesi yüksektir ve dolayısıyla SBD, katlanacak protein üzerine kapanmış durumdadır. NBD'ye ATP bağlandığında ise, SBD açık bir konformasyonda ve bağlacı NBD'nin içlerine alacak şekilde NBD ile etkileşmektedir. Bu konformasyonda SBD'nin substrata afinitesi düşüktür ve katlanmış proteinin salınması ATP bağlı halde gerçekleştirilir. Sonrasında, SBD'ye bağlanan katlanması gerekli yeni bir protein, NBD'ye bağlı halde bulunulan ATP'nin hidrolizini tetikler. ATP'nin ADP'ye hidrolizi sırasında SBD katlanması gereken proteinin üzerine kapanır, NBD ve SBD birbirinden tekrar ayrılır ve mekanizma bu döngü ile devam eder. Bu döngüye koşaperonlar eşlik eder. Hsp40 ailesi ve nükleotit değişim ailesi bu döngüde görevli koşaperonlar olarak bilinir. Hsp40 ailesi (DnaK için DnaJ), ATP hidrolizini hızlandırmakla görevliyken, nükleotit değişim faktörleri (DnaK için GrpE) ise hidroliz ile oluşan nükleotiti değiştirmek ve Hsp70'yi yeni döngüye hazırlama görevindedir. Mayer ve Bukau'nun 2015 yılında yaptığı son çalışmalar, ATP ile indüklenen substrat salınımının, substratın bağlanmasıyla uyarılan ATP hidrolizine göre şaperonun aktivitesinde daha önemli bir rol oynadığını ortaya koymaktadır.

2007 yılında, Swain grubu tarafından yapılan çalışmalarla, domenler arası bağlacın hidrofobik [389]VLLL[392] sekansının NBD ve SBD arasındaki allosterik ilişkiden sorumlu olduğu bulunmuştur. Yine bu çalışmaya göre, bağlaç varlığında DnaK (1-392), substrat tarafından uyarılmış yabanıl tip DnaK gibi davranmaktadır. Yabanıl tip, substrat ile uyarılmamış DnaK'nin aksine, DnaK (1-392)'nin pH bağımlı ve daha yüksek ATPase aktivitesi vardır. Diğer bir taraftan ise, bağlaç yokluğunda DnaK (1-388), substrat tarafından uyarılmamış yabanıl tip DnaK'yi taklit etmektedir.

Moleküler dinamik simulasyonları, 1950 yıllarının ikinci yarısına doğru ortaya çıkmıştır. 1960'larda daha da gelişmiş ve ilk kez 1977 yılında bovin pankreatik tripsin inibitörü ile birlikte proteinler üzerinde kullanılmaya başlanmıştır. *In vivo* ve *in vitro* olarak gözlemlemekte güç olabilecek bazı konularda fikir vermek için kullanılan önemli yollardan biri haline gelmiştir. Biyokimya ve biyofizikte oldukça geniş bir uygulama alanı bulmuştur.

Bu çalışma DnaK (1-392) ve (1-388) kesik yapılarında, ATP bağlı durumda konformasyonel ve dinamiksel farklılıkları göstermeyi amaçlamanın yanısıra 226. konumdaki histidinin alanine mutasyonu halinde genel konformasyon ise de nasıl bir değişim olduğunu moleküler dinamik analizleriyle göstermeyi amaçlamaktadır.

Kesik yapılar 4JN4 kodlu PDB (protein data bank) dosyasından modifiye edilip, NAMD/CHARMM-GUİ kullanılarak 200 ns boyunca yürütülmüştür. Bu süreçte ortalama kareökten sapma (RMSD) ve ortalama kareök değişimi (RMSF) gibi temel moleküler dinamik analizlerinin yanısıra, Başlıca Komponent Analizi (PCA), hedef amino asitler için zamana bağımlı uzaklık ölçümleri ve proteinin ne kadar kompakt olduğunu ölçmek amacıyla dönüş çapı (Rg) analizi yapılmıştır.

Yapılan analizler sonucu bağlacın [389]VLLL[392] kısmının bulunmadığı DnaK (1-388) kesik yapısının, bağlacın bulunduğu DnaK (1-392) yapısına gore belirli bölgelerde

daha kapalı bir konformasyonda olduğu gözlemlenmiştir. Ayrıca bu iki yapı arasında farklı dinamikler olduğu da görülmektedir. Bir diğer önemli sonuç ise bağlaç varlığında His226'nın bu dinamikte önemli bir rolü olduğu yönündedir. 226 numaralı Histidinin alanine mutasyonu ile birlikte, konformasyonun daha da açıldığı, bazı amino asitlerin de bulunduğu konumdan oldukça saptığı görülmüştür. Bunlardan önemli olarak gördüğümüz katalitik bölgede bulunan ve ATP hidrolizinden sonra fosfatı kabul eden amino asit olduğu düşünülen 199. konumdaki Treonin yakından incelendiğinde simulasyon sürecinin sonlarına doğru ATP'den aniden ve önemli ölçüde uzaklaştığı görülmüştür. Bu durumun doğurabileceği sonuçlar, araştırma grubumuza ait diğer deneysel verileri desteklemekle birlikte Histidinin Hsp70 genel dinamiği için ne kadar önemli olduğunu da göstermektedir. Uzun vadede, yapılan bu moleküler dinamik simulasyon çalışmalarının Hsp70 çalışma mekanizmasını detaylı bir şekilde ortaya koyarak Hsp70 kaynaklı nörodejenaratif hastalıkların tedavisine ışık tutabileceği düşünülmektedir.

1. INTRODUCTION

1.1 Protein Folding Mediated By Molecular Chaperones

Nascent polypeptide chains must fold into their 3D structure to be functionally active. Based on data from *in vitro* experiments, it is known that primary structure of proteins encode their unique three-dimensional structures. However, folding of a protein into its unique 3D structure does not happen spontaneously in the cell because of the crowded nature of the cytosol. Only small and single domain proteins can fold spontaneously in the cell. Evidences show that cells use complex cellular chaperone machineries and metabolic energy to form functionally active proteins from nascent polypeptides. By inhibiting the undesired intermolecular and intramolecular interactions, chaperones prevent the aggregation and misfolding of non-native protein (Hartl, 2012). The models of chaperone-mediated folding of nascent proteins in cytosol eubacteria, archea and eukarya are represented in Figure 1.1 (Hartl, 2002) and related proteostasis network represented in Figure 1.2 (Kim et al., 2013)

1.2 Heat Shock Proteins and Hsp Families

Cellular machinery of chaperones mediate the folding of nascent proteins as well as unfolded proteins due to stress conditions such as high temperature. Expression of most chaperones increases under stress conditions, and these chaperones are called as stress proteins or heat-shock proteins (Hsps) (Hartl, 2002). Hsps are constitutively expressed in cells and also essential for normal conditions (Liang and MacRae, 1997). Chaperones perform their fucntion by recognizing the solvent exposed hydrophobic residues of proteins, which are usually buried in their native states (Hartl, 2002).

Hsps are mainly divided into 5 major groups based on their molecular weights, structure and function: Hsp100, Hsp90, Hsp70, Hsp60 and small Hsps (sHsps) (Liang and MacRae, 1997).

Table 1.1 : Diversity of the major molecular chaperone families (adapted from Ranford, 2000)

Family	Chaperone functions
Chaperonin 10 (cpn10, hsp10, co-chaperonin, early pregnancy factor,GroES)	Co-chaperonin' to chaperonin 60; promotes folding of substrates that are bound to chaperonin 60
Small heat-shock proteins (hsps)	Diverse class of proteins; chaperone function is independent of adenosine triphosphate (ATP); bind non-native proteins
Hsp40 (DnaJ related)	Co-chaperones that regulate the activity of hsp70 proteins; some can bind non-native proteins themselves
Chaperonin 60 (cpn60, hsp60, hsp65, GroEL)	ATP-dependent folding and/or refolding of ~15-30% of total cellular proteins
Hsp70	Prevent the aggregation of unfolded polypeptides; disassemble multimeric protein complexes; involved in protein trafficking; regulate the heat-shock response
Hsp90	Specific functions in regulating signal transduction pathways, through their actions on certain kinases and steroid receptors; might also have 'general' chaperone activity
Hsp100 (Clp)	Disassemble protein oligomers and aggregates
Hsp110	High degree of homology with the hsp70 family; little known about functions

Hsp100s, Clp in bacteria, have a chaperone function. They provide degredation or refolding of aggregated proteins in an ATP-dependent manner (Hartl, 2002). Hsp90 is a highly abundant and ubiqitious chaperone. Its functions vary from folding proteins to different cellular processes such as signal transduction to client protein maturation (Li and Buncher, 2013). Small heat-shock proteins bind to non-native proteins and have a role on inhibition of apoptosis, cytoskeleton organisation and etc. (Haslbeck, 2002).
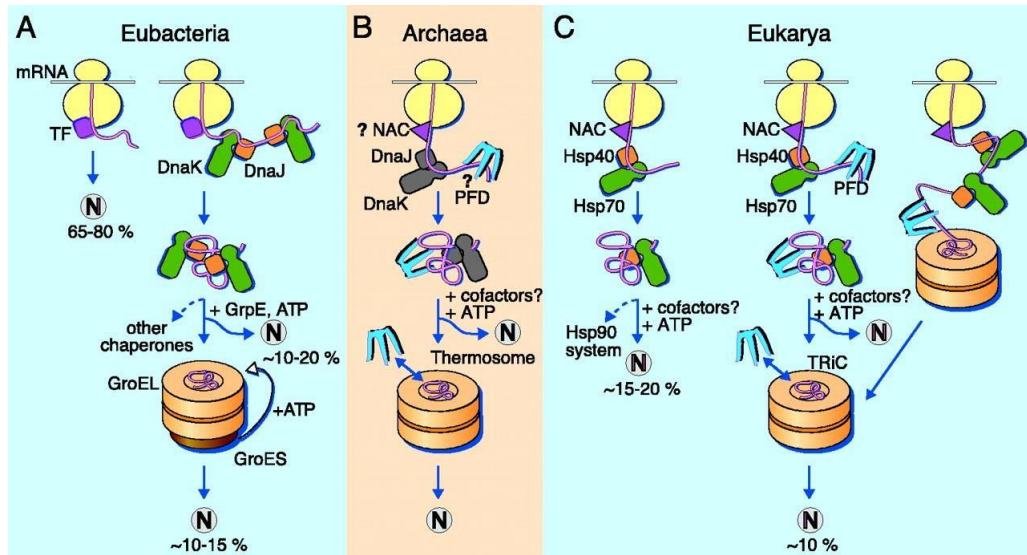
Figure 1.1 : Models for folding of nascent protein in the cytosol of (A) Eubacteria, (B) Archaea and (C) Eukarya (taken from Hartl, 2013).

Protein folding in cells is mediated by cytosolic chaperones with two different mechanisms, and these mechanisms are represented by Hsp70 and Hsp60. Hsp70 binds to 6-9 hydrophobic amino acid residues and use ATP to release their substrates. Hsp60s, chaperonins, is a downstream chaperone which is a nanocage. They have an oligomeric ring shape type of cylindrical structure and provide the folding of larger protein in an ATP-dependent fashion (Figure 1.1) (Liang and MacRae, 1997; Frydman et al., 1994; Hartl, 2002; Kim et al., 2013).
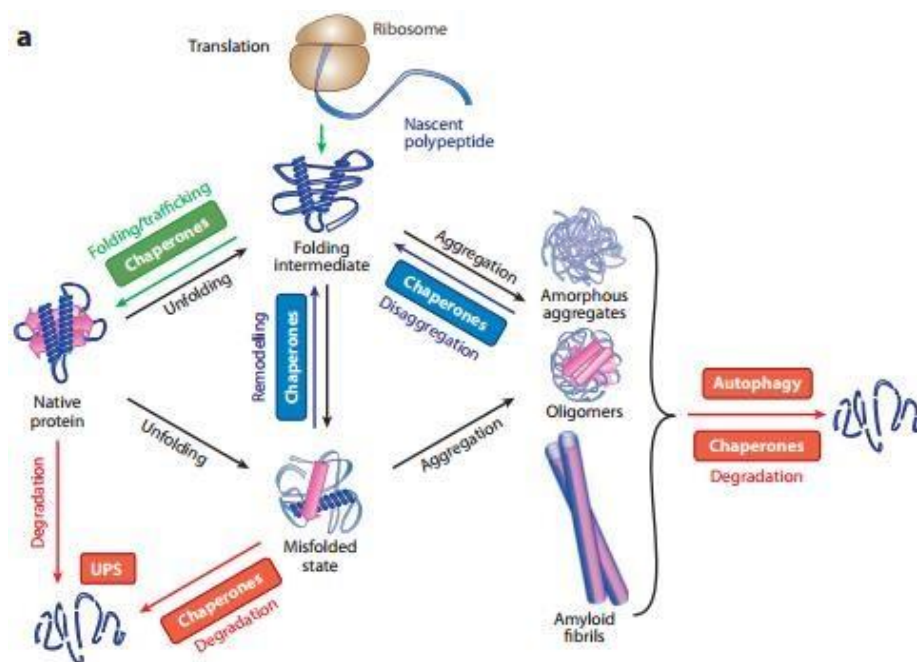
Figure 1.2: The proteostasis network (taken from Kim et al., 2013)

1.3 Hsp 70 Molecular Chaperones

Hsp70s are constitutively expressed (Hsc70) and highly conserved in evolution (Liang and MacRae, 1997). They found in bacteria, eukarya and some archea. Moreover, they may be present in some eukaryotic organelles such as mitochondria and endoplasmic reticulum (Hartl, 2002). Hsp70s can recognize unfolded, misfolded or aggregated states of all proteins without any interaction with folded sides. Besides, they can interact with specific native proteins and regulate their activity, stability and oligomeric conditions (Mayer, 2013). Hsp70s mainly prevent aggregation of non-native proteins, provide appropriate conditions by inhibiting the undesired interactions to reach to their native states, and solubilize and refold aggregated proteins by using ATP (Mayer and Bukau, 2005). Hsp70s owe their versatility to the ability of recognition of short degenerate motives in proteins. These motives, which form the hydrophobic core of native proteins, can be found in almost every 30-40 residues consisting of five hydrophobic core residues and regions surrounded by positively charged amino acids (Rügider et al., 1997; Mayer, 2013). Additionally, Hsp70s work with cochaperones such as J-domain proteins (JDPs) and nucleotide exchange factors (NEFs). While JDPs induce the ATPase activity of Hsp70s (up to >1000 fold when substrate is present), NEFs provide the exchange of nucleotide by stimulating the release of substrates (Mayer, 2013).

1.3.1 Structure of Hsp70

Hsp70s have two domains: ~43 kDa N-terminal nucleotide binding domain (NBD) and ~27 kDa C-terminal substrate binding domain (SBD). These domains are connected by a highly conserved hydrophobic linker (Kityk et al., 2012; Mayer 2013). The structure of NBD is quite similar to the structure of actin. NBD has two lobes (I and II) and four subdomains, which are IA, IIA, IB and IIB. I and II lobes are separated by a hydrophilic nucleotide binding cleft (Mayer, 2013). ‒B‖ subdomains are upper and consisting of α-helices and anti-parallel β-sheets while ‒A‖ subdomains are lower and composed β-sheets surrounded by α-helices (Hurley , 1996). SBD consists of two subdomains, which are SBDβ (a sandwich of two stranded β-sheets and peptide-binding cleft) and SBDα (α-helical lid) (Figure 1.2 and Figure 1.3) (Kityk et al., 2012). The peptide binding cleft is enriched in hydrophobic amino acids like Val, Ile and Leu as well as aromatic and positively

charged amino acids while negatively charged amino acids are not seen occasionally (Rüdiger et al., 1997).

Nucleotide Binding Domain (NBD)

Substrate Binding Domain (SBD)

Figure 1.3 : A model of DnaK (Hsp70 homolog) complex with ADP and substrate (S) (1-605) (adapted from PDB ID: 2KHO).
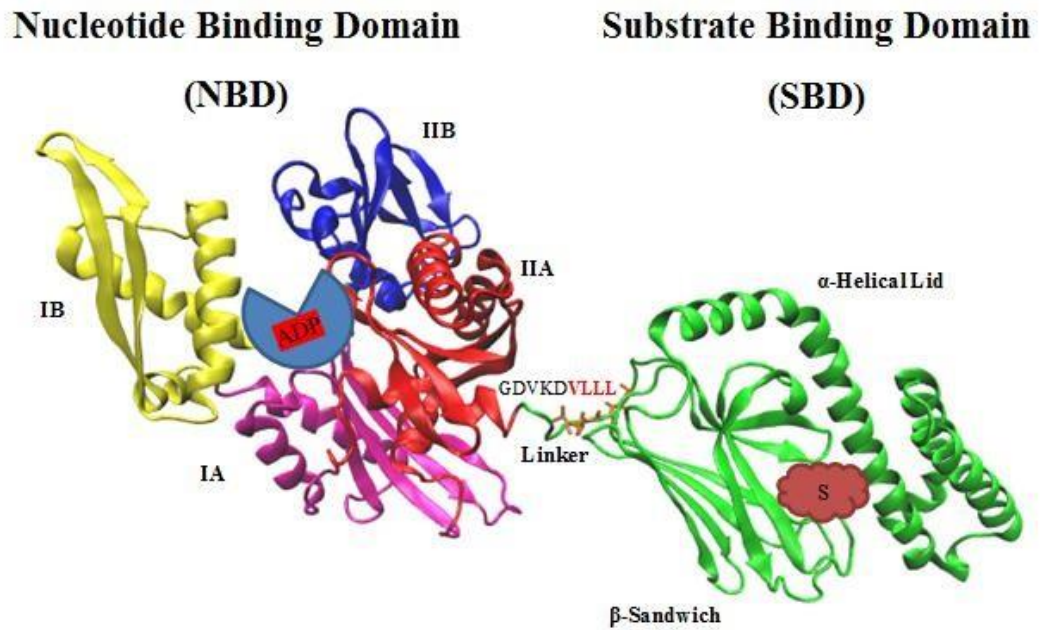
Nucleotide Binding Domain (NBD)

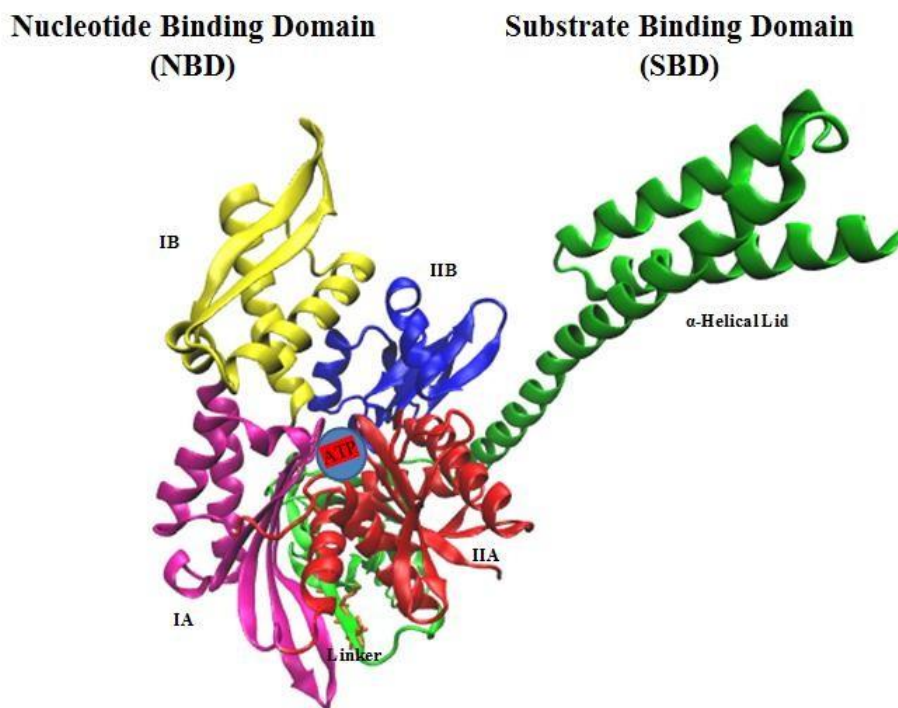**Substrate Binding Domain (SBD)**

Figure 1.4 : Allosteric opening of the polypeptide-binding site when an Hsp70 binds ATP (1-608) (adapted from PDB ID: 4JN4).

PDB structures of DnaK that contains nucleotide binding domain are given in the Table 1.2.

Table 1.2 : DnaK PDB structures that contains nucleotide binding domain.

PDB Code	Name	Method	Resolution	R-Value Free	R-Value Work	Reference
4JN4	Allosteric opening of the polypeptide-binding site when an Hsp70 binds ATP	X-RAY Diffraction	2.3 A°	0.197	0.162	Qi,R. et. al., 2013
4JNE	Allosteric opening of the polypeptide-binding site when an Hsp70 binds ATP	X-RAY Diffraction	1.96 A°	0.201	0.173	Qi,R. et. al., 2013
4B9Q	Open conformation of ATP-bound Hsp70 homolog DnaK	X-RAY Diffraction	2.4 A°	0.239	0.196	Kityk, R. et. al., 2012
2KHO	E. coli Hsp70 (DnaK) chaperone (1-605) complexed with ADP and substrate	Solution NMR	Conformers calculated : 1	Conformers calculated : 1	Conformers calculated : 1	Bertelsen, E. B., et. al., 2009
1DKG	Crystal structure of the nucleotide exchange factor GrpE bound to the ATPase domain of the molecular chaperone DnaK	X-RAY Diffraction	2.8 A°	0.317	0.223	Harrison, C.J. et. al., 1997

The crystal structure 1DKG revealed in 1997, involves nucleotide-free NBD with the nucleotide exchange factor GrpE. According to this study, when GrpE binds to nucleotide-free NBD, structure become resemble to that of nucleotide-bound mammalian Hsp70 homolog. However, only one of the subdomains rotate outwards which prevents the tight nucleotide binding. They identified that two helices extend away from the GrpE dimer. These helices, may suggest a role for GrpE in peptide release from DnaK (Harrison et al., 1997).

In 2009, the structure of *E.coli* DnaK complexed with ADP and substrate (2KHO) was revealed. While loosly linked nucleotide binding domain and substrate binding domain were determined with molecular dynamics simulations, also has been seen that linker region is a dynamic random coil. According to this structure, subdomain IA of NBD is closer to SBD. This finding suggests that SBD interacts with the NBD at this area to constitute an allosteric communication (Bertelsen et.al., 2009).

The crystal structure of ATP-bound open conformation of Hsp70, 4B9Q, was reported in 2012. This study, caught the Hsp70 when NBD and SBD are docked, and both subdomains of the SBD dock onto different sites on NBD. The interaction points belong to the β-sheet subdomain and the NBD reveals the mechanism of allosteric regulation. Kinetic measurements in the study demonstrated the following

communication sequence: i) linker enters into the hydrophobic groove in IA and IIA of hydrophobic binding domain, ii) SBD β docks onto the NBD, iii) SBDα docks onto the NBD (Kityk et al., 2012).

In 2013, two new crystal structures of ATP-bound DnaK were revealed. The first is 4JN4 and the latter is 4JNE (Qi et al., 2013). Their study shows how ATP binding stimulates the polypeptide-binding channel in SBD. In ATP-bound state, NBD and SBD are docked and form extensive contacts which result in radical conformational changes in SBD. As a consequence, NBD pulls away α–helical lid of SBD from SBDβ. Hence, SBD becomes more accessible for peptide binding (Qi et al. 2013).

1.3.2 Conformations and dynamics of Hsp70

Hsp70 perform its function via ATP-dependent mechanism. There is an allostery in this mechanism since binding of nucleotide and substrate results in conformational change in each individual domain (NBD and SBD). The interdomain linker provides the communication between two domains by binding to the hydrophobic pocket between IA and IIA in NBD (Swain et al., 2007). Basically, in the first, when ATP is bound, NBD and SBD domains are docked and act independently. In this state, dissociation and association rates of substrates are higher, showing lower affinity for substrate. Secondly, NBD hydrolysis ATP via its ATPase activity, and dissociation and association of substrates slow down. Thus, the affinity for substrate is higher in this state (Mayer, 2013). Also, recent studies revealed that ATP-induced substrate release is more important than substrate-stimulated ATP hydrolysis in chaperone activity (Kityk et al., 2015)

Furthermore, in 2007 Swain et al. created two different constructs of NBD, which are DnaK (1-388) and DnaK (1-392). 392 construct has a hydrophobic ^{389}VLLL^{392} sequence of the interdomain linker. They measured ATPase rate of substrate-stimulated full length DnaK (WT+p5), unstimulated full-length DnaK (WT), DnaK (1-388) and DnaK (1-392), respectively. While WT and (1-388) were not pH dependent, WT with p5 and (1-392) were pH-sensitive in the ATPase activity, suggesting that DnaK (1-392) behaves like substrate-stimulated form of full-length DnaK (Figure 1.3) (Swain et al., 2007). Their work was crucial for understanding the responsible residues for the pH-dependent ATPase activity.
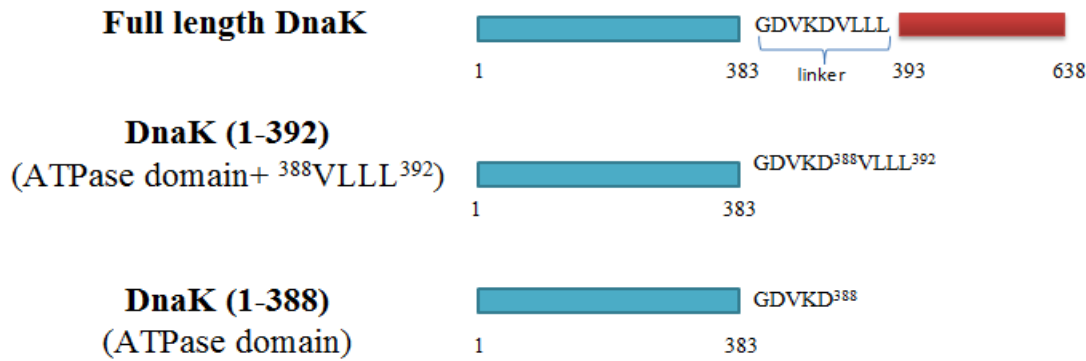
Figure 1.5 : Constructs used in the study (adapted from Swain et al., 2007).

In the closed conformation (ADP-bound), SBDα acts as a lid and covers SBDβ so that peptide dissociation rates are reduced. Additionally, peptide interacts with peptide binding cleft via its hydrophobic residues mostly consisting of Leu, Ill, Val and Phe, and a polar residue Tyr (Mayer, 2013).

In open conformation (ATP-bound), lobe I of NBD rotates and closes the nucleotide-binding cleft while the cleft between IA and IIA is opened. The opening of the cleft causes the interaction of the interdomain linker with the cleft. This interaction is required for the ATP-stimulated peptide dissociation and polypeptide-stimulated ATPase activity. Additionally, these events result in the displacement of the two catalytic residues: Glu171 and Lys70 (*E. coli* DnaK) (Mayer, 2013). To understand the open conformation better, it is necessary to compare this conformation with the closed conformation. Experiments show that three different conformations of SDBβ are present so far: (1) In the open conformation, The upper β-sheet is composed of 5 strands instead of 4 while lower one is consisting of 3 instead of 4. Therefore, the upper one is more twisted than the lower one. (2) In the closed conformation, the angle between strands 1 and 4 is closer than in the open conformation. This provides a scissors like movement of strands 1 and 2. (3) Smaller substrate binding cleft in open conformation is due to the closure of distance between strands 3 and 4 (Mayer, 2013).

Evidences from recent methyl nuclear magnetic resonance (NMR) studies show that there is an intermediate (allosterically active) state between open and closed conformation. In this conformation, Hsp70 (DnaK) is found as a complex with ATP and substrate. Based on data of methyl-TROSY, Zhuravleva et al. concluded that in ATP-substrate DnaK complex, the linker binds to NBD, and SBD is undocked from

NBD. However, it was understood that ATP and substrate-bound DnaK population is consisted with not only linker-bound conformations but also linker-unbound conformations as shown in Figure 1.4 (Zhuravleva et al., 2012).
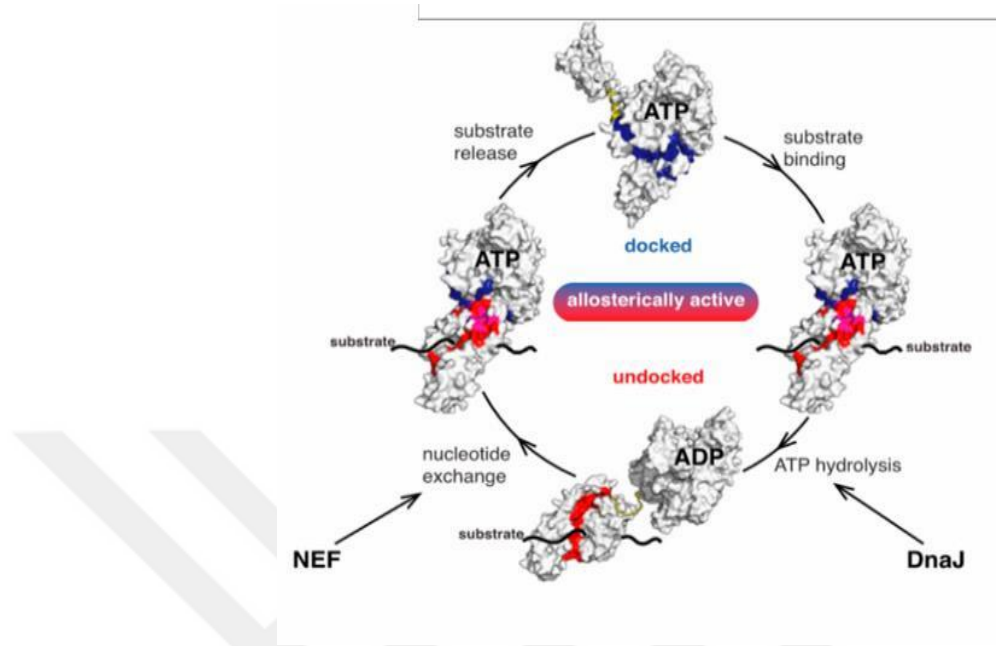
Figure 1.6 : A model for DnaK allosteric cycle (Taken from Zhuravleva et al., 2012)

1.4 Molecular Modeling and Simulations

It is important to obtain dynamic alterations of proteins and one of the methods to use for that purpose is to do molecular dynamics simulations in combination with X–ray crystallography and NMR to reveal in detail the working mechanism of proteins (Karplus et al., 2005).

Atomic trajectories in MD is basicaly generated by numerical integration of Newton's equation of motion. Specific interatomic potential that system's have, can be calculated for certain initial condition (IC) and boundary condition (BC) (Li, J.,2005). Newton's equation of motion has shown in eqn (1.1).

$$F_i = m_i \frac{d^2 r_i(t)}{dt^2} \tag{1.1}$$

While m_i is the mass of particle ‒i‖ , $r_i(t)$ is the position vector of it and F_i is the force that acts on particle i, at time of t. For this reason, some algorithms like

Verlet,leap-frog,velocity verlet have been used. Purpose of the numerical integration is to find $r_i (t + \Delta t)$ at time of $(t + \Delta t)$ (Meller, 2001).

Molecular modeling and molecular dynamics are used for molecular interaction investigations, biomolecular motor design, protein folding and distributed computing, computational drug design and biocomputing, quantum dynamics of enzyme reactions and etc. (Ramachandran et al., 2008).

Details of basic concepts of programs and analysis methods present in this thesis are mentioned under methods section.

1.5 Previous Molecular Dynamics Simulation Studies on Hsp70

In order to understand the dynamics of Hsp70, several studies have been done. In 2002, Golas et al. performed coarsed-grained molecular dynamics studies on DnaK for three states: i) with the two halves of the NBD unrestrained relative to each other, (ii) with the two halves of the NBD restrained in an ‒open‖ geometry as in the SBD-closed form of DnaK (2KHO), and (iii) with the two halves of NBD restrained in a ‒closed‖ geometry as in known experimental structures of ATP-bound NBD forms of Hsp70. Since there was not any experimental structure of ATP-bound state avaliable on those days, they used complexed structures of ATP-bound Hsp70 with Hsp110. Their study was the first that simulated complete spontaneous transition from the SBD-closed to SBD-open conformation of the Hsp70 chaperone from *E. coli*. They saw that SBD-NBD docked open conformation formed spontaneously during all simulations; number of transitions are larger in simulations run with ‒closed‖ NBD and smaller for ‒open‖ NBD (Golas et al.,2012).

Another study which is done in 2012 by Chiappori and her collugues, proposed a molecular mechanism for the allosteric signal propagation of the ATP-encoded conformational signal. The hydrogen bond-based connections between the NBD and the linker include the residues K67, I137, V139, R164, I166 and E168 which are conserved for all DnaK systems and are also responsible for linking the nucleotide to NBD. They revealed that while this pathway was more stable in the presence of ATP because of the coordination between the phosphate group and K67 and E168, it was distrupted in ADP-bound DnaK. Another nucleotide-modulated interaction according to this study, is the conserved loop 195. In ATP-bound DnaK; G194, G195 and T196 contribute the binding of ATP. However in the ADP-bound form, the coordination to

the loop is reduced, only G194 is connected to the nucleotide. Also R342 has a role on hydrogen bond network of ATP binding. R342 is localized at the interface between subdomain IIA and IIB, and they revealed that this interaction may involved in the rotation of IIB, which causes opening of NBD in ATP-bound state. Other important residues for allosteric signal propagation are D385 and D390. These two residues both involved in the ADP-bound DnaK while for ATP-bound state only D390 is included. D385 and D390 flank the [389]VLLL[392] sequence of the interdomain linker which contributes the allosteric coupling of NBD and SBD (Chiappori et al., 2012).

In 2012 Kityk et al., and in 2013 Qi et al. revealed ATP-bound crystal structures 4B9Q and 4JN4, 4JNE, respectively. These crystal structures explain the substrate release by ATP binding better than triggering of ATP hydrolysis by substrate binding (Qi et. al,2013; Kityk et al., 2012).
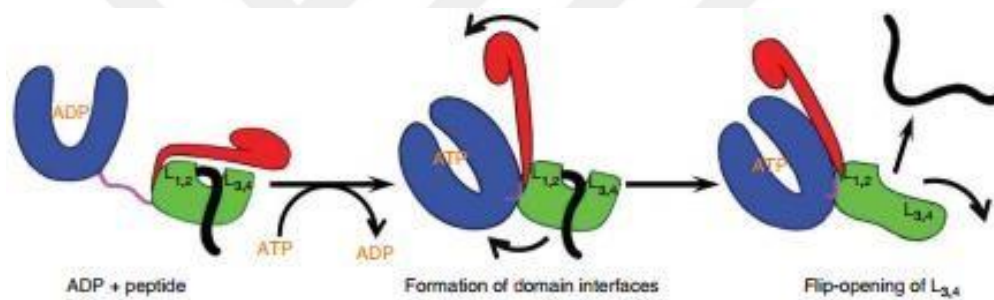
Figure 1.7 : Model for allosteric opening of the polypeptide-binding site when an Hsp70 binds ATP (Taken from Qi et al., 2013)

Another molecular dynamics study which is done by Nicolai et al. revealed that, when ATP binds to the NBD, which leads docking of SBD onto NBD lobe I, protein reached an intermediate plausible ATP-bound state (ATP*) with free-energy landscape analysis (Nicolai et al., 2013).

In 2015, Stetz and Verkhivker performed a study that try to reveal atomistic picture of signal propagation and energetics of dynamics-based communication with molecular dynamics simulation and network-centric modeling. According to their study ADP-bound DnaK has larger movements of subdomain IIA and α–helical domain of SBD. In ATP-bound form, there are smaller fluctuations compared to ADP-bound DnaK which may be a sign of conformational dynamics providing structural tightening of NBD core and ATP-binding site (Stetz and Verkhivker, 2015).

In the noval study of Chiappori et.al. in 2016, they aimed to explain the molecular determinants of the allosteric communication in both directions, from NBD to SBD and vice versa. They proposed a model for the allosteric communication in DnaK which emphasized the importance of linker connecting NBD and SBD. Effects of ATP and the model substrate NR-peptide were combined in order to have a stable-linker docked complex. As a result they showed that convergence of the dynamics, the persistence of the beta strand content of the docked linker and increased coordination in the ATP binding site in agreement with the hypothesis that this structure is an intermediate in ATPase cycle (Chiappori et al., 2016).

1.6 Purpose of This Study

Based on the previous studies, it is known that ^{389}VLLL^{392} sequence, which is an evolutionary highly conserved hydrophobic linker, provides the allosteric communication between substrate binding domain (SBD) and nucleotide binding domain (NBD). Furthermore, it is revealed that the ATPase activity of DnaK (1-392), which is NBD with hydrophobic ^{389}VLLL^{392} sequence, is higher than DnaK (1-388) in the absence of linker, and DnaK (1-392) has a pH-dependent ATPase activity while DnaK (1-388) has not (Swain et al., 2007).

In this study, we aimed to investigate the differences in the dynamics of DnaK (1-388) and (1-392) constructs when ATP is bound. Additionally, we focused on the effect of the mutation of histidine residue which is positioned at 226 to alanine on dynamics of protein.

2. METHODS

In this study, molecular dynamics simulations were performed by the nanoscale molecular dynamics (NAMD) program (Phillips et al., 2008), using the CHARMM22 force field parameters (Brooks et al., 2009). To prepare the simulation input, CHARMM-GUI web server was used (Jo and Kim et al., 2008).

Firstly, the protein models are uploaded to CHARMM-GUI to construct the simulation input which is prepared from the crystal (PDB ID: 4JN4) (Qi, Sarbeng and Liu et.al., 2013). As a first step, for input generation in server, the solvation of the protein performed with a buffer of TIP3 water and the total electric charge of system was adjusted to neutral state by placement of potassium ions. Next, periodic boundary conditions (PBC) was chosen for all simulations, and then canonical ensemble (NVT) dynamics in which amount of substance (N), volume (V) and temperature (T) are conserved was chosen to perform equilibration. As a final setup of parameters, temperature was adjusted to 303,15 K with Nose-Hoover method (Nose, 1984).

To start the simulation, step size was set to 1 fs in conjunction with the SHAKE algorithm (Ryckaert et al., 1977) and VVER (velocity verlet) was used for integration. Then, Steepest Decent (SD) and adopted basis Newton Raphson method (ABNR) was used for minimization. For electrostatic interactions, Particle Mesh Ewald (PME) method was chosen (Essmann et al., 1995).

Temperature was adjusted to 303,15 with Hoover and pressure was adjusted to 1 atm. Simulation time step was set to 2 fs to, perform SHAKE. PME was chosen for electrostatic interactions same as in the equilibration step. CPT (constant pressure-temperature) leap used for integration. Constructs simulated for 200 ns. Each step is 5 ps and coordinates were saved for each 2500 step.

RMSD, RMSF and radius of gyration analysis results were performed with the program visual molecular dynamics (VMD) program (Humphrey et al., 1996), which

was also used to draw the molecular representations. PCA was performed with R programs' Bio3D package (Grant et al., 2006).

2.1 MM Force Fields

A force field can simply be stated as a mathematical expression which gives information about system's energy dependency on its particles' coordinates. It includes the interatomic potential enegy in analytical form and some parameters which can also participate into this form. The parameters are procured from *ab initio* or semi-empirical quantum mechanical calculations. Also it can be obtained from NMR, X-RAY, infra-red (IR), neutron spectroscopy etc. Different force fields, having different complexity, can be found in the literature in order to be used in various systems (Gonzalez, 2011).

2.2 Chemistry at Harvard Macromolecular Mechanics (CHARMM)

CHARMM is one of the force fields that can be used in molecular dynamics simulations. While CHARMM27 is mostly use for MD simulastions of DNA, RNA and lipids, for pure proteins, there are united-atom CHARMM19 and all-atom CHARMM22 subtypes which are released in 90s (MacKerell et al., 1998).

Energy function is given with the equation below.

$$
\begin{aligned}
U(\vec{R}) = &\sum_{\text{bonds}} K_{\text{b}}(b - b_0)^2 + \sum_{\text{UB}} K_{\text{UB}}(S - S_0)^2 + \\
&\sum_{\text{angle}} K_{\theta}(\theta - \theta_0)^2 + \sum_{\text{dihedrals}} K_{\chi}(1 + \cos(n\chi - \delta)) + \\
&\sum_{\text{impropers}} K_{\text{imp}}(\varphi - \varphi_0)^2 + \\
&\sum_{\text{nonbond}} \epsilon \left[\left(\frac{R_{\text{min}_{ij}}}{r_{ij}} \right)^{12} - \left(\frac{R_{\text{min}_{ij}}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{\epsilon_1 r_{ij}}
\end{aligned}
\tag{2.1}
$$

2.3 Nanoscale Molecular Dynamics (NAMD)

NAMD is a parallel molecular dynamics code which can be used with AMBER (Assisted Model Building with Energy Refinement) and CHARMM for large biomolecular systems. NAMD, allows producing classical simulation with high performance in factual environments of 100,000 atoms or more. It is based on

Charm^{++} parallel programming. NAMD sees the simulation cell as 3-D crazy quilt, and in NAMD, each patch interacts with the closest patches. Bonded, short-range electrostatic and van der Waals interactions are calculated for each patch and their neighbours. NAMD and molecular graphics/sequence analysis software Visual Molecular Dynamics (VMD) complement each other.

NAMD has a Tcl (Tool command language) scripting interface and ability of working with parallel performance. So, user can benefit from calculations which are used in making simulations with direct proportion. Tcl, advantageously saves time for user (Phillips et al., 2005).

2.4 Root Mean Square Deviation (RMSD)

RMSD measures the similarity of the quantitative degree between two, 3-D protein structures. It calculates this similarity with distance between equivalent atom pairs. For identical structures, RMSD is zero (Carugo and Pongor, 2001).

$$\text{RMSD}(t_1, t_2) = \left[\frac{1}{N} \sum_{i=1}^{N} \| \mathbf{x}_i(t_2) - \mathbf{x}_i(t_1) \|^2 \right]^{1/2} \quad (2.2)$$

While N is the number of atoms in the molecule, $\mathbf{x}_i(t)$ is the position of atom i at time t. In molecular dynamic simulations, RMSD can tell us if the simulation reached equilibrium. As long as RMSD does not attain a stationary shape, one can conclude that simulation has not converged yet (Schreiner et al., 2012).

2.5 Root Mean Square Fluctuation (RMSF)

RMSF, either can be calculated from molecular dynamics simulation with formula (Frenkel and Smit, 2002);

$$\text{MSD} \equiv \langle (x - x_0)^2 \rangle = \frac{1}{N} \sum_{n=1}^{N} (x_n(t) - x_n(0))^2 \quad (2.3)$$

and B-factors or Debye-Waller factors in X-RAY experiments with using the formula below (Kuzmanic and Zagrovic, 2010).

$$RMSF_i^2 = \frac{3B_i}{8\pi^2}.$$

(2.4)

RMSF gives information about thermal stability and flexibility of the protein. It is also used in the prediction of binding pockets and active sites. If RMSF is high, that means relevant atom fluctuation by the time (Kuzmanic and Zagrovic, 2010).

2.6 Principle Component Analysis (PCA)

Principle component analysis is the oldest and the most popular technique all over multivariate statistical techniques. PCA analyzes a data table representing observations described by several dependent variables, which are generally inter-correlated. Aim of PCA is to take the substantial information from that data table and reflect this griff as the principal components. Principle components can define new orthogonal variables. It is also possible to see the pattern of similarity of the variables and observations by visualizing them with points in a map (Abdi and Williams, 2010). Mathematical background is mentioned below.

$$\mathrm{cov}(X, Y) = E[E[X] - X] \cdot E[E[Y] - Y]$$

(2.5)

This equation expresses the covariance, showing two random variables' leaning to vary together. E[X] is the expected value of X. For sampled data this equation can be arranged as;

$$\mathrm{cov}(X, Y) = \sum_{i=1}^{N} \frac{(x_i - \bar{x})(y_i - \bar{y})}{N}$$

(2.6)

While $\bar{\ }$ is mean of (X) , $\bar{\ }$ is mean of (Y). It is important to remember that, cov(X,X) = var(X) and if X and Y independent from each other cov(X,Y) = 0. Covariance matrix can be shown for matrix A, $A_{i,j}$ = cov(i,j). Note that, covariance matrix is symmetric and square. If the variables are independent, the covariance matrix is going to be a diagonal matrix which has variances along the diagonal. For calculating the covariance matrix, first step that must be done is centering the data by

substracting the mean of each sample vector. If we take in consideration that the data matrix A's columns as the sample vectors, elements of covariance matrix C can be written as follows:

$$c_{ij} = \frac{1}{N}\sum_{i=1}^{N} a_{ij}a_{ji} \qquad (2.7)$$

We can also express it in matrix in another form,

$$C = \frac{1}{N}AA^{T} \qquad (2.8)$$

− shows the scale factor which is distributed along the matrix. Covariance matrix expressed simply as AA^{T} (Coombe, 2006).

PCA can be applied to several systems. For instance, Saal et al. used microarrays for measuring the expression levels of 27,648 genes in 105 breast tumor samples and do PCA to detect gene expressions with dominant patterns (Ringnér, M., 2008). Another system that PCA can be used is proteins. Maisuradze et al. used PCA for observing folding dynamics of Formin binding protein 28 (FBP) (Maisuradze et al., 2009).

2.7 Radius of gyration

Radius of gyration (R_g) is a parameter that defines the compactness of a protein structure. Rg can be calculated in two steps. First center of mass coordinates Rc specified with not taking into a account hydrogen atoms, from following equation.

$$\sum m_i(r_i - R_C) = 0 \qquad (2.9)$$

Mass of the i-th atom is m_i and coordinates showed as r_i.

If atoms assumed as points in a 3D space, the Rg can be obtained as

$$R_g^2 = \sum m_i (r_i - R_C)^2 / M \tag{2.10}$$

where mass of atoms in a protein indicated as M.

For proteins, it is feasible to assume masses for all nonhydrogen atoms. So that where N is the number of atoms which are not hydrogens in a protein.

$$R_g^2 \cong \sum_{i=1}^{N} (r_i - R_C)^2 / N \tag{2.11}$$

Difference between Radius of gyration values calculated from equations (2.11) and (2.12) is within various hundredths of angstrom. It is significant to consider atoms as balls with the radius R, instead of material points in terms of accuracy. If we assume all atoms have the same radius (1.5A°), the equation becomes more exact:

$$R_g^2 = \sum_{i=1}^{N} (r_i - R_C)^2 / N + \frac{3}{5} R^2 \tag{2.12}$$

To conclude of those equations, radius of gyration is an indicator of protein compactness (Lobanov et.al., 2008).

3. RESULTS

To initiate the study, truncated models DnaK (1-388) and DnaK (1-392) were constructed based on the crystal structure of ATP-bound DnaK (PDB: 4JN4). In the crystal structure, there was only T199A mutation on the nucleotide binding domain, so we mutated back T199A to T199 in the crystal structure in order to observe real dynamics of the protein. Conditions of this arrangement are explained in methods section and observations are stated below.

3.1. Different Overall Protein Dynamics Observed for DnaK (1-388) and (1-392).

We studied two ATPase domain constructs, one having the entire linker sequence DnaK (1-392), and one missing the hydrophobic part of the linker sequence ^{389}VLLL^{392}, which is DnaK (1-388). RMSD, RMSF and principal component analysis were performed for observing the differences between DnaK (1-392) and DnaK (1-388).

The most important difference was observed for two helices, one containing residues between 71 to 90 and the other 224 to 250 during the simulation (Figure 3.1). These helices correspond to subdomain IB and, end of IIA beginning of IIB, respectively. It was observed that, two α-helices are closer in DnaK (1-388) while they become distant in DnaK (1-392). This subject mentioned under results section 3.2 and discussion section 4.

From two α-helices, first one which belongs to subdomain IA, contains the residues 148 to 161 and second helix which belongs to subdomain IA and IB, contains the residues 114 to 132 is observed closer and stable in DnaK (1-392) either.

Oppositely located β-sheet containing the residues 279 to 285 (corresponds to IIB) and α-helix containing the residues 256 to 274 (corresponds to IIB) are observed closer for DnaK (1-392) compared to DnaK (1-388) in the last nanoseconds of the trajectory.

19

Comparison of two α-helices that are located opposite of each other, one is located at an α-helix containing residues 52 to 59 (corresponds to IB) and the other one containing the residues 256 to 274 (corresponds to IIB) showed that the first α-helix moves away from the second in the y axis in DnaK (1-388) (Figure 3.3).

The α-helix containing the residues 79 to 88 (corresponds to IB) and the opposite neighbour β-sheet containing the residues 92 to 95 (corresponds to IB), observed closer to eachother in DnaK (1-392) construct. Also it has been seen that α-helix is more mobile in DnaK (1-392) construct.

The investigation between the β-sheet sequence 335 to 341 and the α-helix sequence 300 to 327 (both corresponds to subdomain IIA and IIB) showed that in DnaK (1-388) construct, end of the β-sheet is closer to α-helix. On the other hand in general overview β-sheet stays more stable in DnaK (1-388) compared to DnaK (1-392).

When two α-helices that are located opposite of each other, were compared, one is located at an α-helix containing residues 310 to 328 (corresponds to IIA) and the other one containing the residues 347 to 387 (corresponds to IIA) showed that helices are closer in DnaK (1-392).

From oppositely located α-helices the former containing the residues 170 to 181 (corresponds to IA) and the latter 369 to 381 (corresponds to IIA) showed more stable motions in DnaK (1-392).
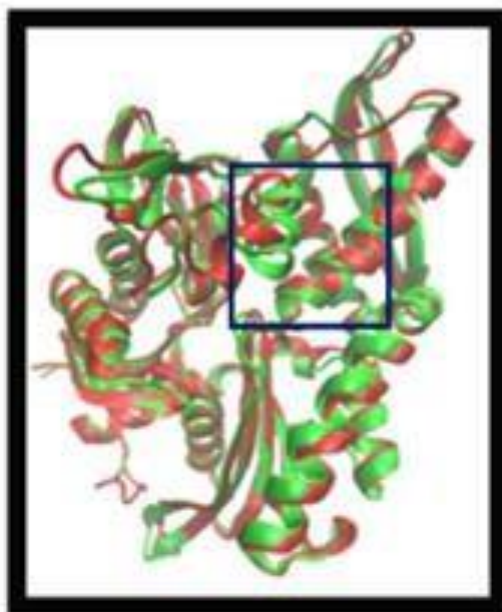
The α-helices that composed of the residues 255 to 275 (corresponds to IIB) and 51 to 62 (corresponds to IB), have closer bearing in DnaK (1-392).
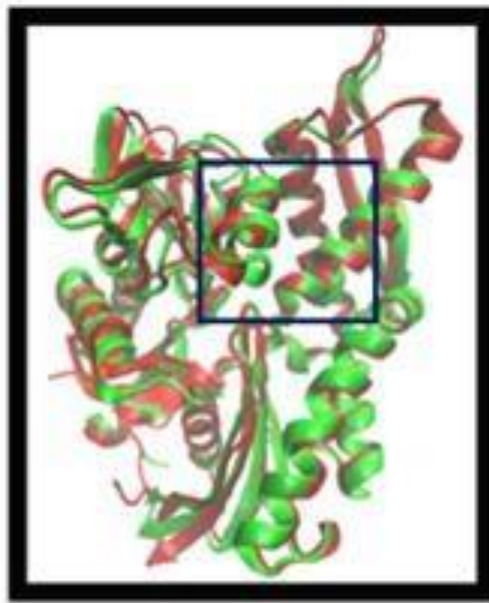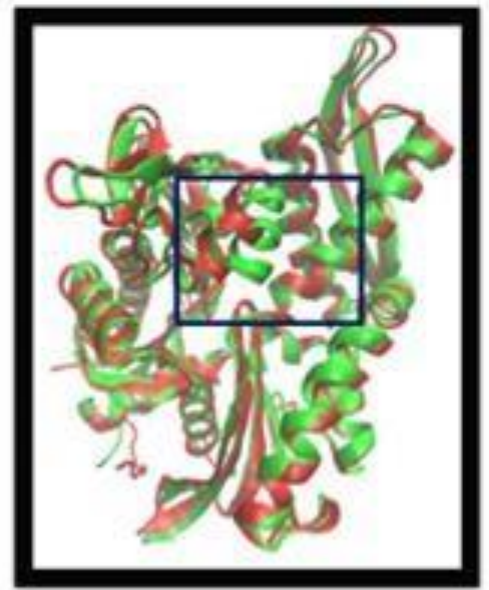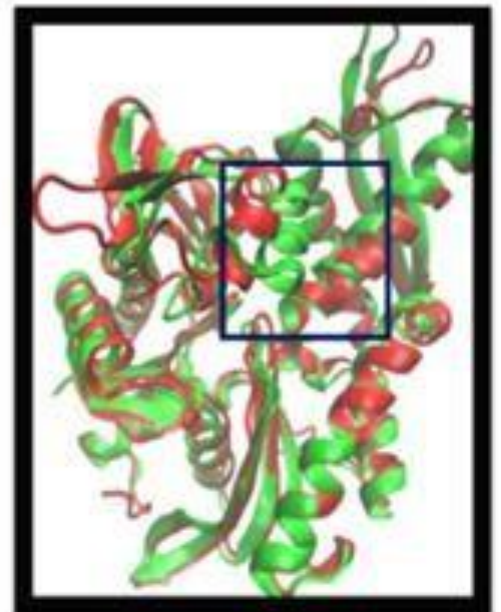
Figure 3.1 : DnaK (1-388) and (1-392) constructs screenshots from particular point of time (green and red,respectively). Blue square shows the significant change in the helix contains residues between 71 to 90.

We performed RMSD analysis for the DnaK (1-388) and DnaK (1-392) constructs to see the main difference for these two constructs (Figure 3.2).
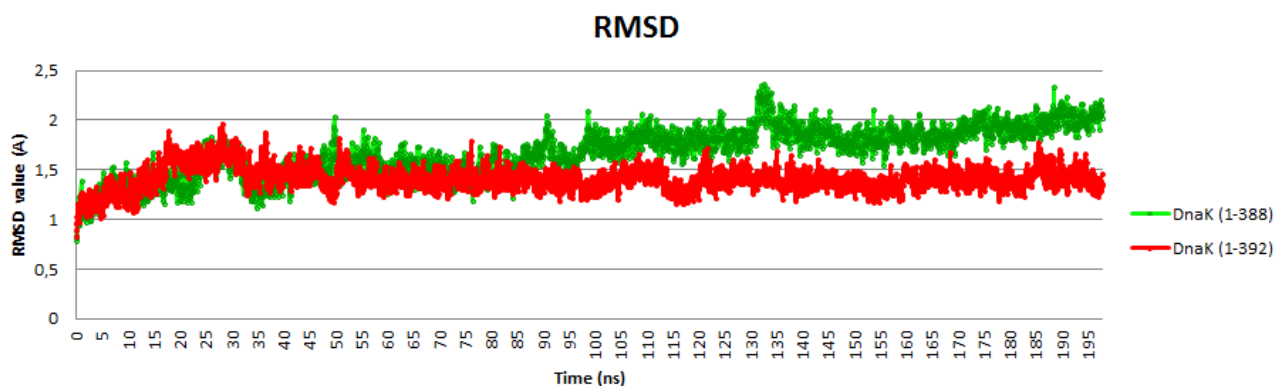

Figure 3.2 : RMSD plot of the model DnaK (1-388) and DnaK (1-392).

RMSD results showed that there are differences between the DnaK (1-388) and DnaK (1-392) constructs. While DnaK (1-392) construct reached the equilibrium faster, for DnaK (1-388) it took more time. It seems that the structure of DnaK (1-388) deviates from its initial conformation.

RMSF plots also showed us significant differences in protein flexibility for particular residues (Figure 3.3).
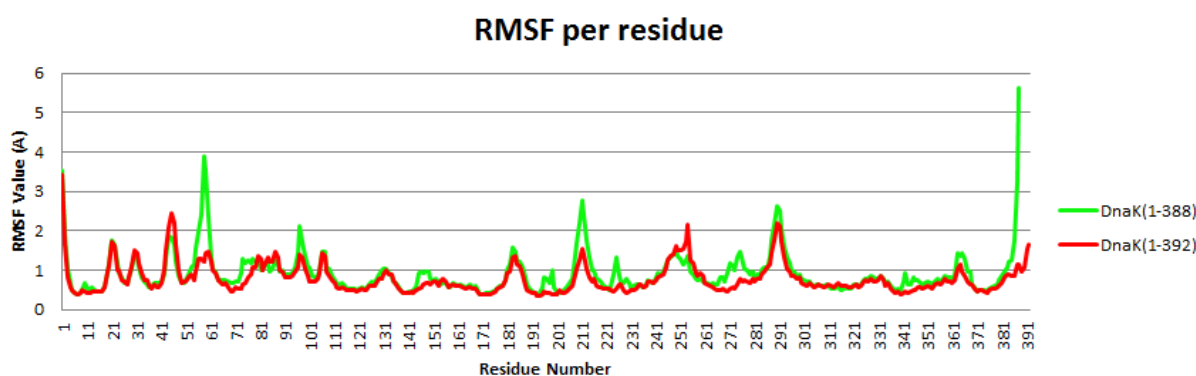

Figure 3.3 : RMSF plot of the model DnaK (1-388) and DnaK (1-392).

In DnaK (1-392) constructs, lower RMSF value for C-terminus has observed compared to DnaK (1-388). Also it has been seen that there are substantial differences for the residues; 58 and 97 in IB, 145 in IA, 196, 211, 225, 226 and 342 in IIA, 254 and 274 in IIB.

Radius of gyration analysis was done for examining and comparing compactness of DnaK (1-388) and DnaK (1-392) (Figure 3.4).
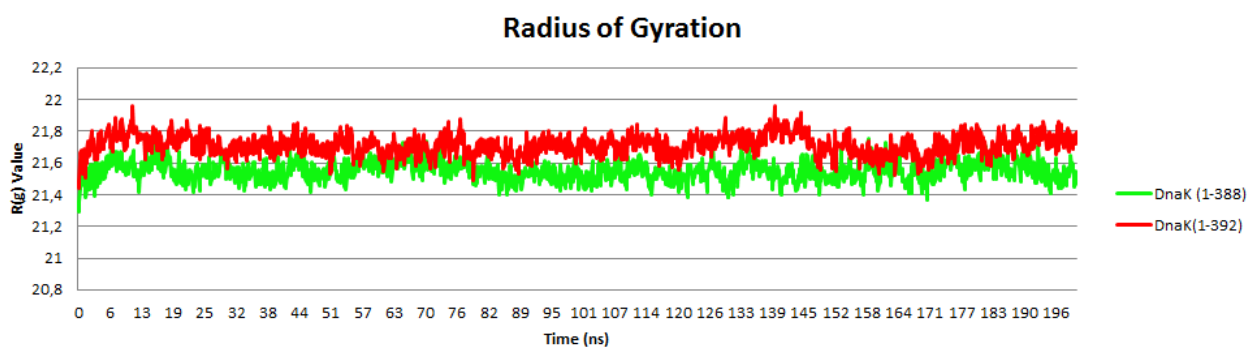
Radius of Gyration

Figure 3.4 : Radius of gyration plot of DnaK (1-388) and DnaK (1-392).

Radius of gyration results revealed that DnaK (1-388) have a more compact conformation compared to DnaK (1-392) with a little difference.

Most of functional processes contains slow and large conformational states. PCA selects collective degrees of freedom which has a role on atomic displacements seen in a trajectory. In terms of determining dominant dynamics differences and correlation of the relevant structures, principle component analysis was performed (Figure 3.5).
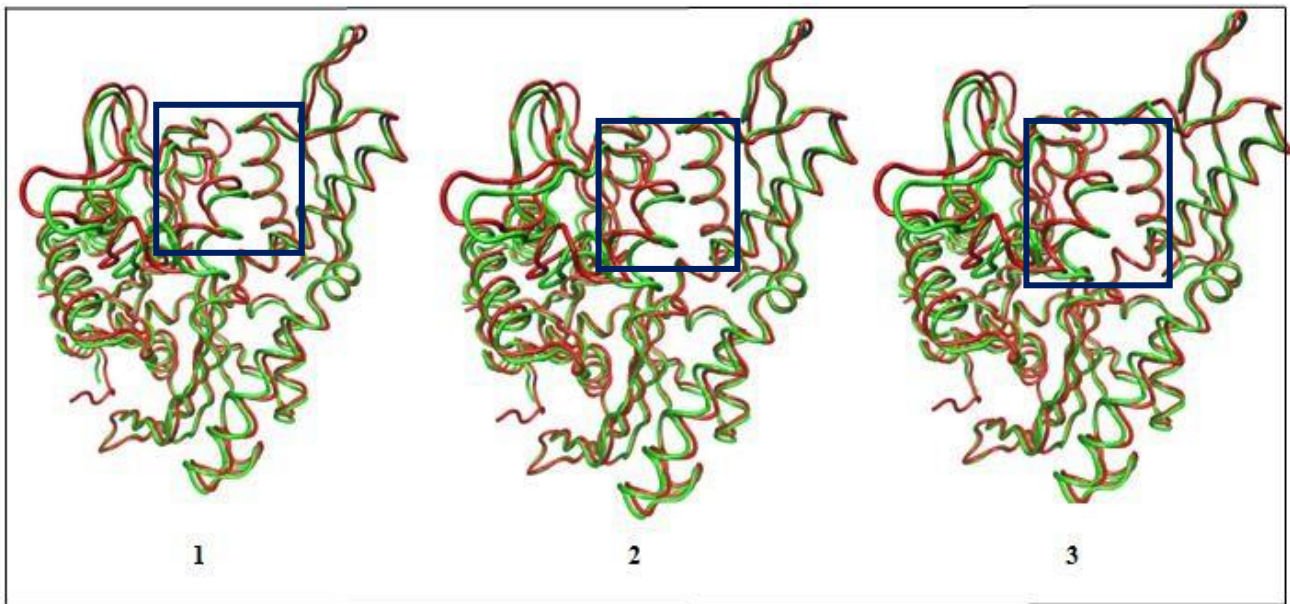
Figure 3.5 : First principal component analysis of DnaK (1-388) and
DnaK (1-392), green and red, respectively. Blue square shows the
significant change in the helix contains residues between 71-90. (1,2,3
indicates the time dependent screenshots obtained from reconstructing
protein structure using first principal component)

Analysis of first principal component showed that dominant groups in protein and
their correlations are different for DnaK (1-388) and DnaK (1-392) constructs. DnaK
(1-388) have more flexible motions, especially for the helix includes the resdiues 71-
90 which is indicated in square. While first principle component describes the
13.83% of the all motions for DnaK (1-392), it describes 29.04% for DnaK (1-388).

3.2. Different Dynamics Observed for Regions Including Sites His226, Asp224, Asp231 and Arg71, Glu80, Glu81, Arg84, Asp85.

The first region that we realized different dynamics is the area between two
neighbour helices which contains His226, Asp224, Asp231 and Arg71, Glu80,
Glu81, Arg84, Asp85. For better understanding where these two helices are located
within protein, relevant structures can be seen in Figure 3.6 and Figure 3.7. Closer
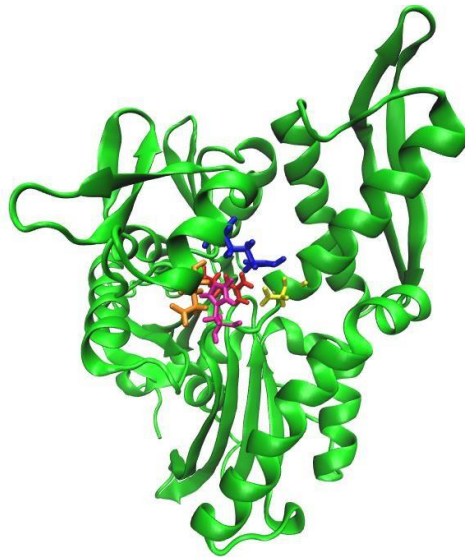look to helices can be seen in Figure 3.8.

Figure 3.6 : V-shaped NBD of DnaK (1-388) with residues H226, E81, D231, R84 and R71 (magenta, orange, yellow, blue and red, respectively).
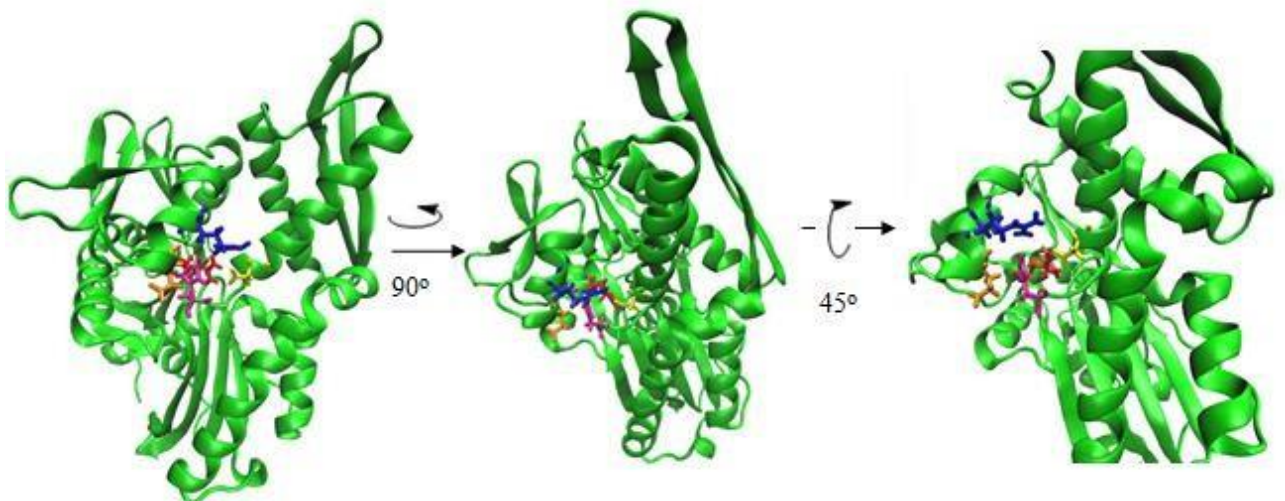

Figure 3.7 : Positions of DnaK (1-388) with residues H226, E81, D231, R84 and R71 (magenta, orange, yellow, blue and red, respectively).
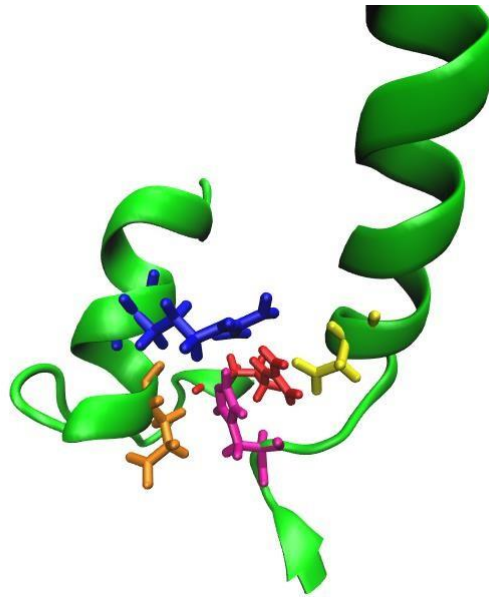
Figure 3.8 : Closer look to relevant helices and residues H226, E81, D231, R84 and R71 (magenta, orange, yellow, blue and red, respectively).

These two helices are closer in DnaK (1-388) rather than DnaK (1-392). Distance measurements for some residues of interests are given in Figure 3.9.
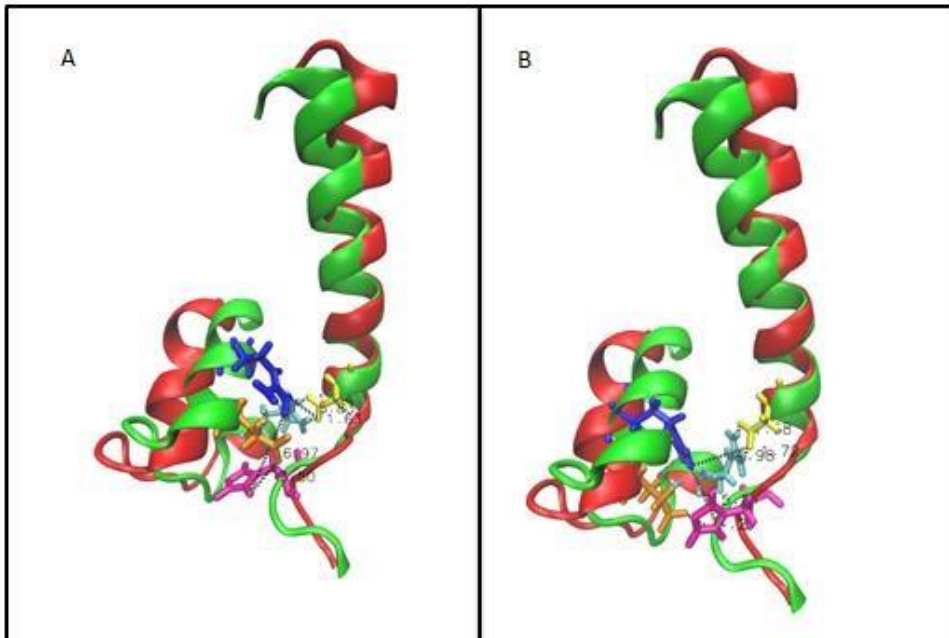

Figure 3.9 : Closer look to relevant helices at (A) DnaK (1-388) (green) and (B) DnaK (1-392) (red) with residues H226, E81, D231, R84 and R71 with a instantaneous value of distances at 200 ns (magenta, orange, yellow, blue and cyan, respectively).

26

Residues picked from the construct DnaK (1-388) (A) for measuring distances instantaneously. H226-E81, R71-D231 with two interaction points and D231-R84 have the distances 3.51, 1.63, 1.64 and 5,83 A°, respectively.

Residues picked from the construct DnaK (1-392) (B) for measuring distances instantaneously. H226-E81, R71-D231 with two interaction points, R84-D231 have the distances 4,21, 1,72, 1,68 and 4,98 A°, respectively.

Detailed distance measurement analysis along simulations for DnaK (1-388) and DnaK (1-392) was performed. Distances between alpha carbons for the relevant residue pairs can be seen in Figure 3.10 to Figure 3.18.
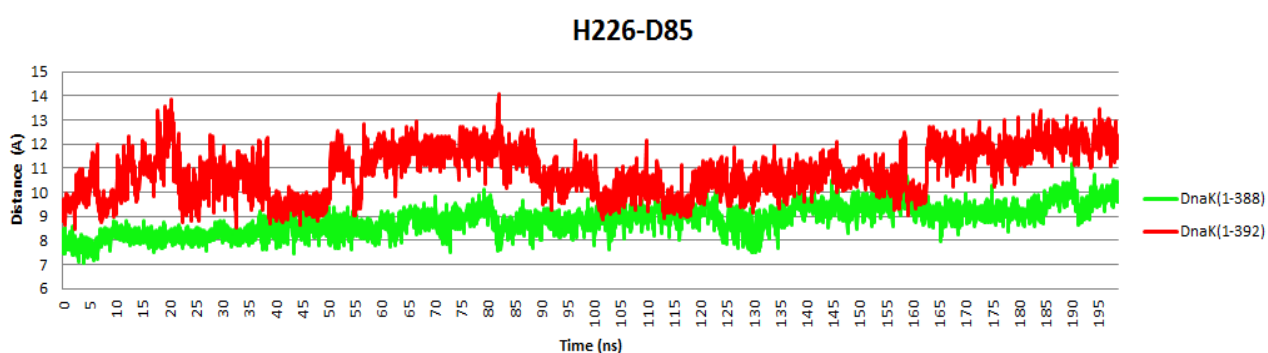

Figure 3.10 : αC distance between H226-D85 residues for model DnaK (1-388) and DnaK (1-392).

It is seen that distance between His226 and Asp85 is closer in DnaK (1-388) . Also, DnaK (1-392) does not showed a stable pattern compared to DnaK (1-388).
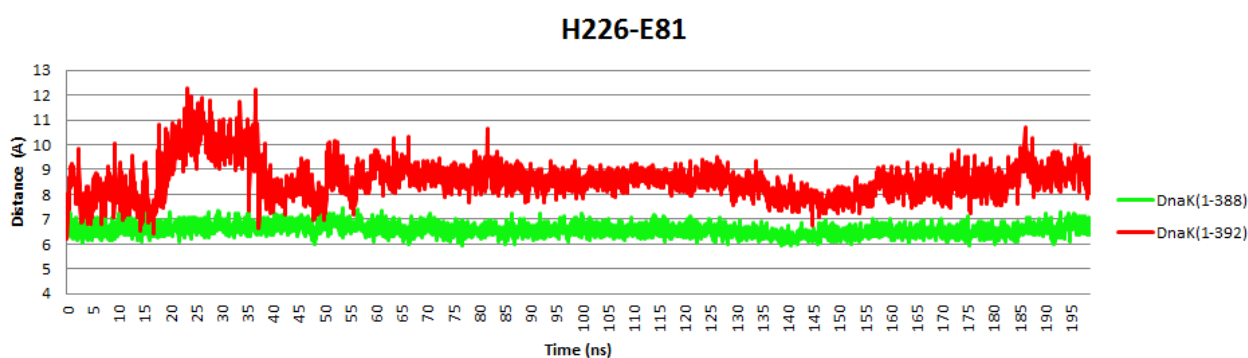

Figure 3.11 : αC distance between H226-E81 residues for model DnaK (1-388) and DnaK (1-392).

Distance between the residues His226 and Glu81 is higher for DnaK (1-392) construct during the whole simulation.
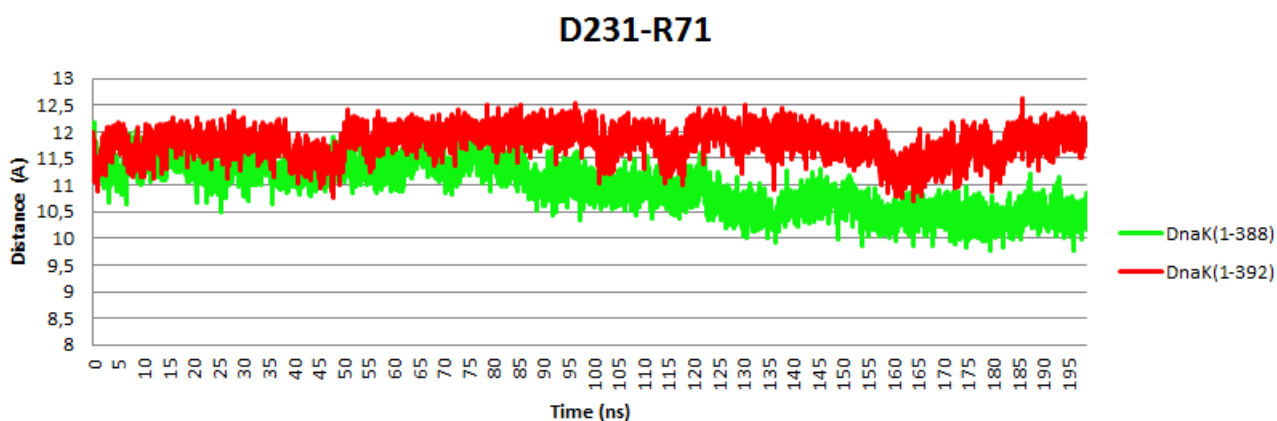
.

Figure 3.12 : αC distance between D231-R71 residues for model DnaK (1-388) and DnaK (1-392).

Distance between the residues Asp231 and Arg71 is decreased by the time for DnaK (1-388).

Figure 3.13 : αC distance between D231-R84 residues for model DnaK (1-388) and DnaK (1-392).

At a first glance it is seen that distance between Asp231 and Arg84 is higher for DnaK (1-392). In addition, it is also seen that there is an increase between those residues in DnaK (1-392) construct after 145 ns. Also, distance measuremenrs according to time showed that DnaK (1-388) has more prominent profile compared to DnaK (1-392).

Figure 3.14 : αC distance between H226-R71 residues for model DnaK (1-388) and DnaK (1-392).

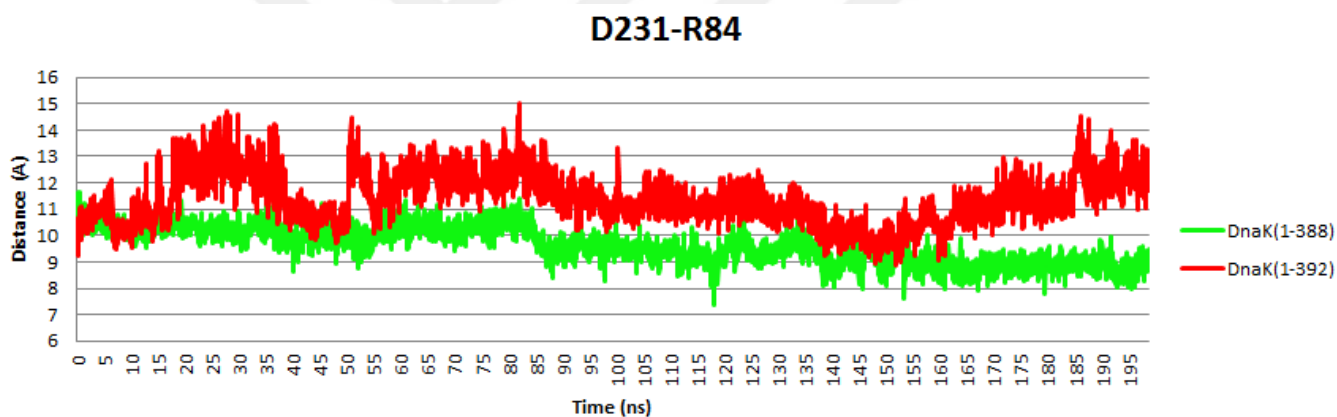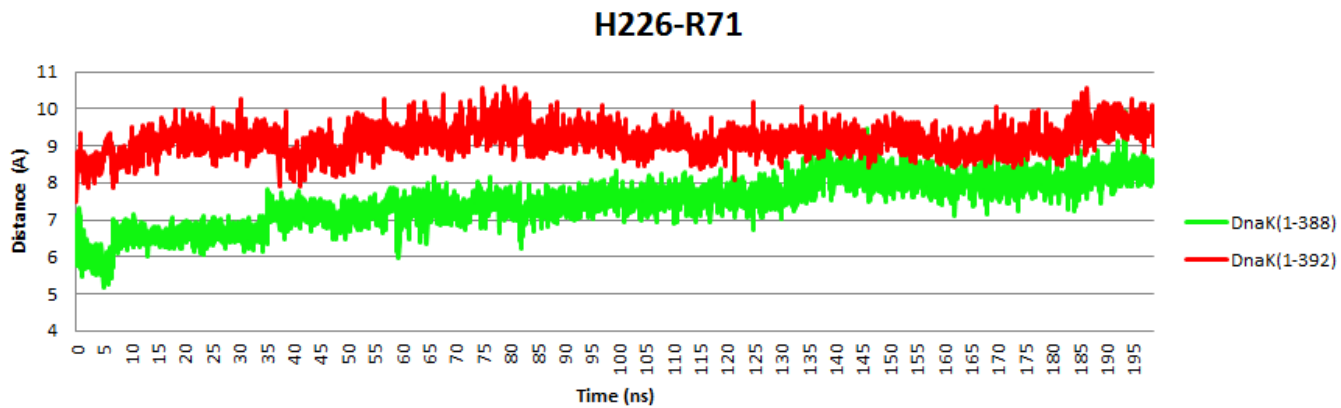Distance between the residues His226 and Arg71 is fewer for DnaK (1-388).
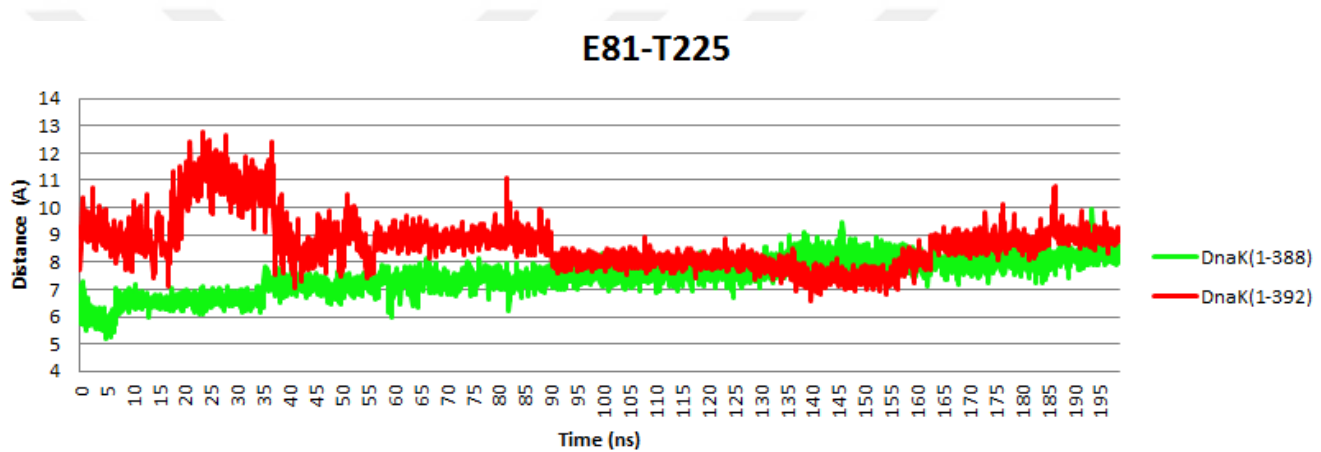
Figure 3.15 : αC distance between E81-T225 residues for model DnaK (1-388) and DnaK (1-392).

Distance between Glu81 and Thr225 is not show us a stable profile for DnaK (1-392). But it is seen that after 160 ns it is increasing.

Figure 3.16 : αC distance between R75-D224 residues for model DnaK (1-388) and DnaK (1-392).

Distance between the Arg75 and Asp224 is higher for construct DnaK (1-392).
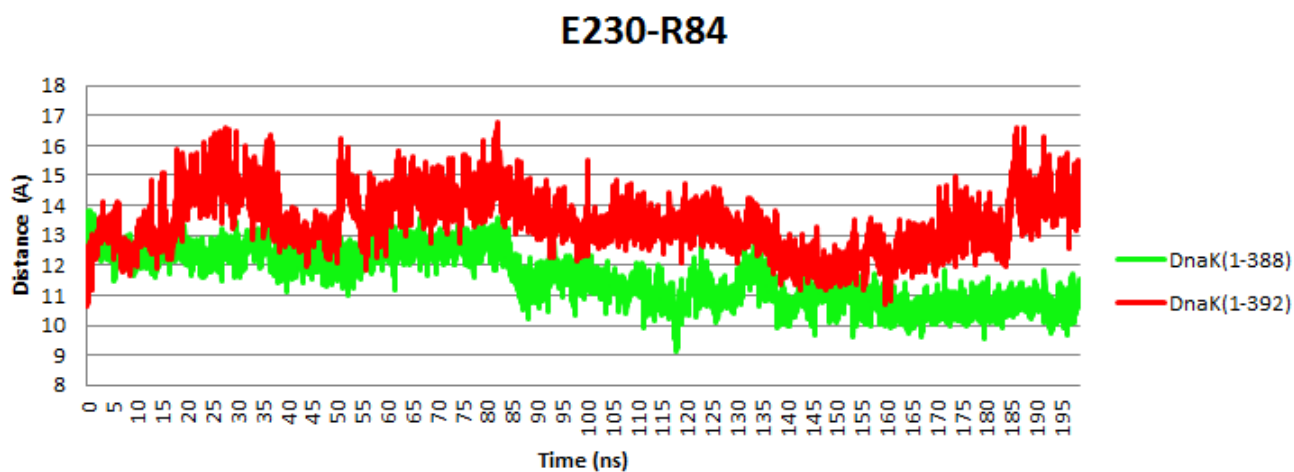
Figure 3.17 : αC distance between E230-R84 residues for model DnaK (1-388) and DnaK (1-392).

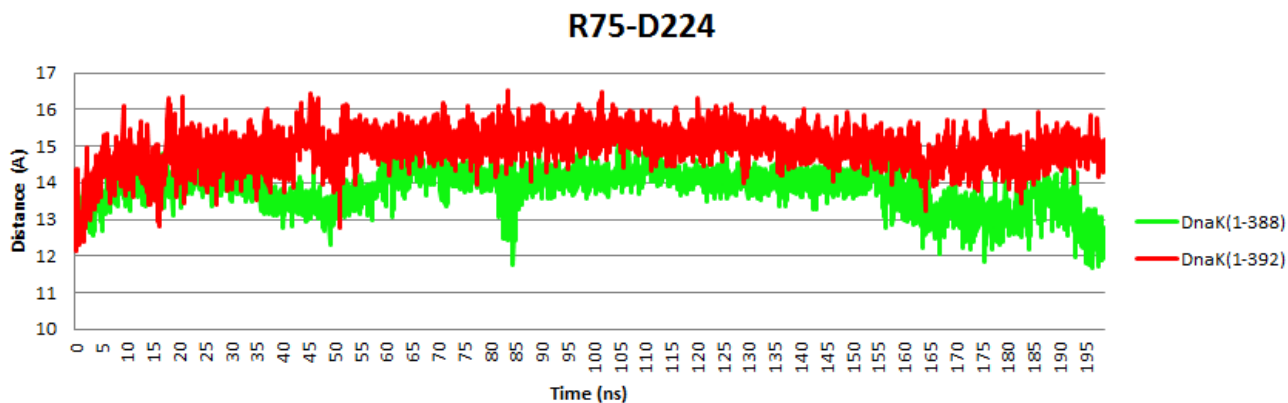Distance between Glu230 and Arg84 is higher for DnaK (1-392). Also, significant increase in 160 ns is observed for DnaK (1-392).
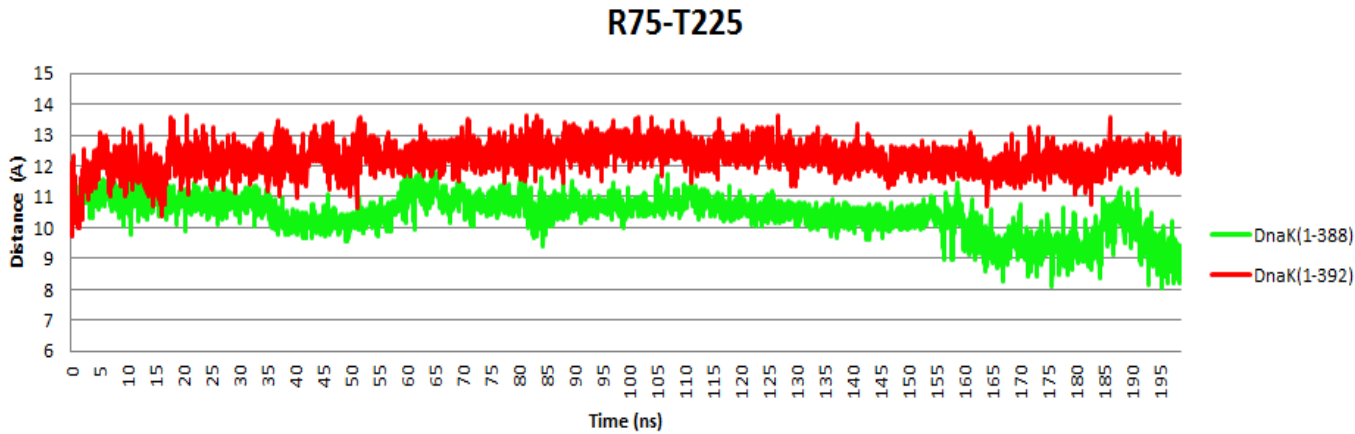
Figure 3.18 : αC distance between R75-T225 residues for model DnaK (1-388) and DnaK (1-392).

The residues Arg75 and Thr225 are closer to each other in DnaK (1-388).

Distance measurement analyses testify that member of helices 71 to 90 and 224 to 250 are closer in DnaK (1-388) construct.

3.3. Ala Mutation at His226 Revealed More Open Conformation of the ATPase Domain When Linker Sequence ^{389}VLLL^{392} is Present.

Suspicion about a network between those relevant helices took us a step forward. H226A mutation was performed in DnaK (1-392) to see effect of this residue on general dynamics of protein. Time dependent screenshots can be seen in Figure 3.19.
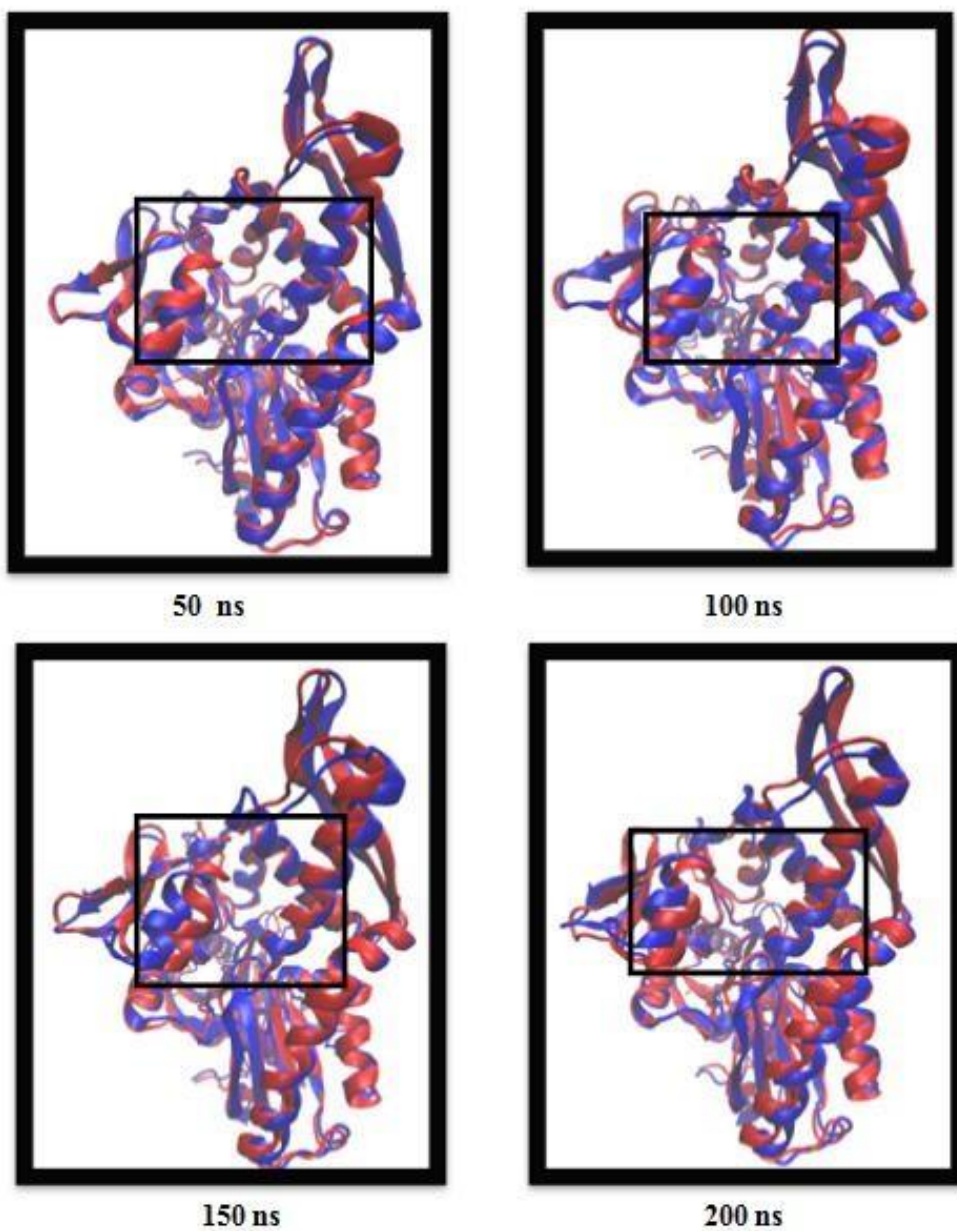
Figure 3.19 : DnaK (1-392) and H226A DnaK (1-392) constructs screenshots from particular point of time (red and blue, respectively). Black square shows relevant helices.

RMSD and RMSF analysis were performed for constructs DnaK (1-392) and mutant H226A DnaK (1-392) (Figure 3.20 and Figure 3.21).

32
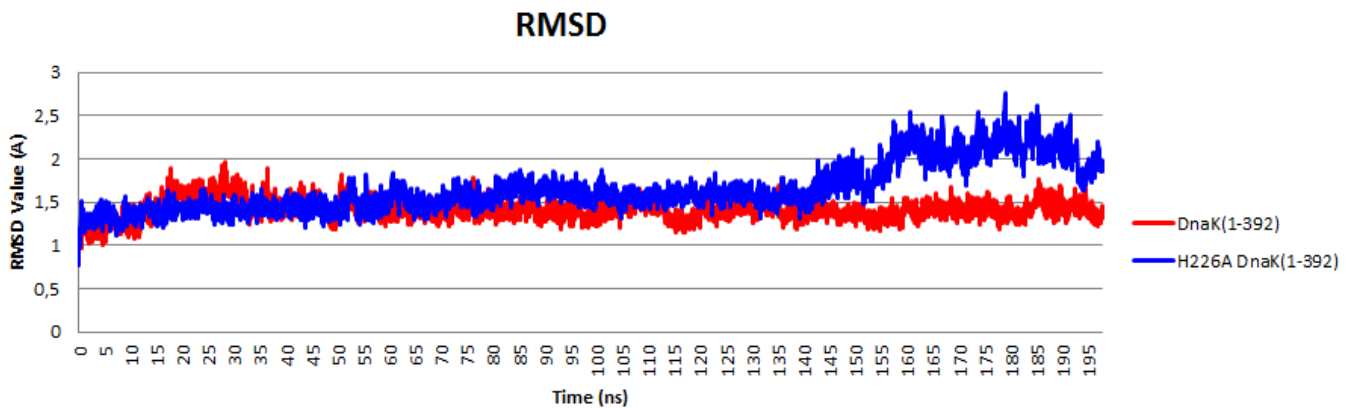
Figure 3.20 : RMSD plot of the model DnaK (1-388) and H226A DnaK (1-392).

RMSD plot showed similar pattern until 145 ns. However RMSD of mutant H226A DnaK (1-392) increased in the last quadrant while DnaK (1-392) showed a stable pattern after relaxation.
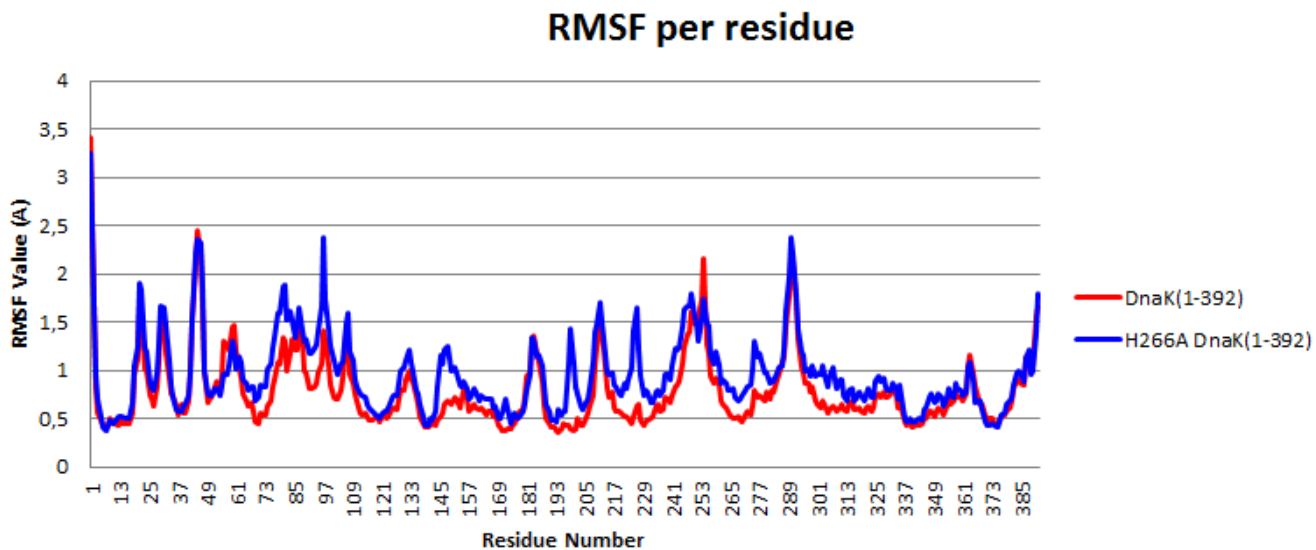
Figure 3.21 : RMSF plot of the model DnaK (1-392) and H226A DnaK (1-392).

RMSF plot showed that generally higher fluctuations belongs to H226A DnaK (1-392) mutant especially residues; 60 to 110, 131, 145 to 170, 199, 226 and 275. Residues between 60 to 110, belongs to subdomain IB and residues between 145 to 170 belongs to subdomain IA. Those subdomains contains critical residues for ATP hydrolysis.

33

Radius of gyration analysis was done for comparing the compactness of DnaK (1-392) and mutant H226A DnaK (1-392) (Figure 3.22).
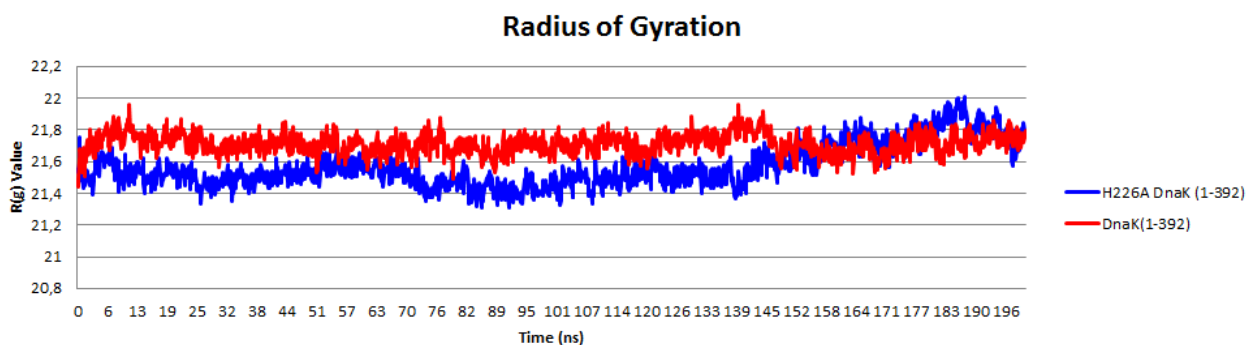
Figure 3.22 : Radius of gyration plot for DnaK (1-392) and H226A DnaK (1-392).

Radius of gyration results revealed that H226A DnaK (1-392) started to lose its compactness after 140 ns.
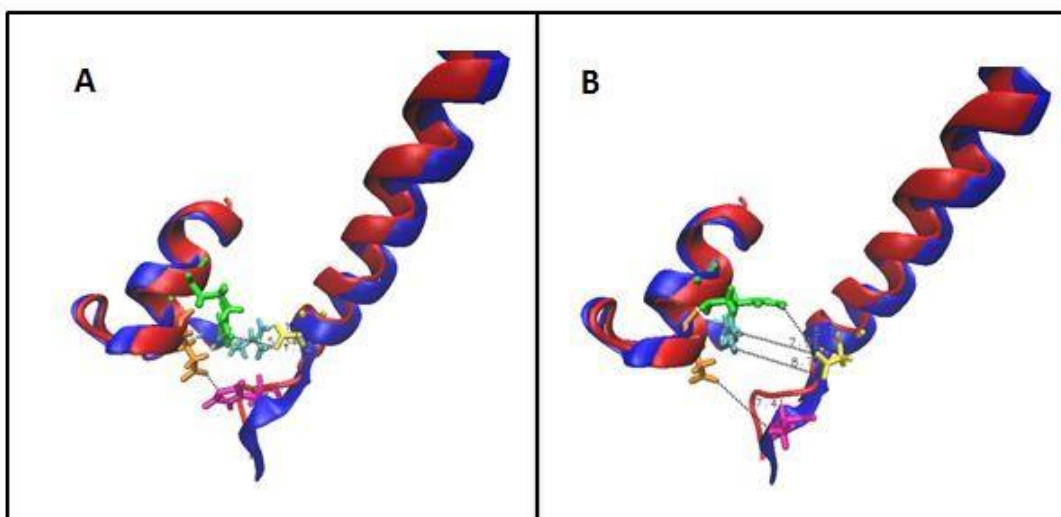
Figure 3.23 : Closer look to relevant helices at (A) DnaK (1-392) (red) and (B) H226A DnaK (1-392) (blue) with residues H226, E81, D231 and R71 with an instantaneous value of distances at 200 ns (magenta, orange, yellow and cyan, respectively).

Distances between the residues belonging to DnaK (1-392) construct (A) measured instantaneously at 200 ns. H226-E81, R71-D231 with two interaction points and R84-D231 have the distances 4.21, 1.68, 1.72 and 5.34 A^o, respectively.

Distances between the residues belong the mutant H226A DnaK (1-392) construct (B) measured instantaneously at 200 ns. H226A-E81, R71-D231 for two interaction points, R84-D231 have the distances 7.41 ,7.92, 8.77 and 5.75 A^o, respectively.

Detailed distance measurement analysis along simulations for DnaK (1-392) and H226A DnaK (1-392) were performed. Distance between relevant residue pairs for alpha carbons can be seen in Figure 3.24 to Figure 3.31.
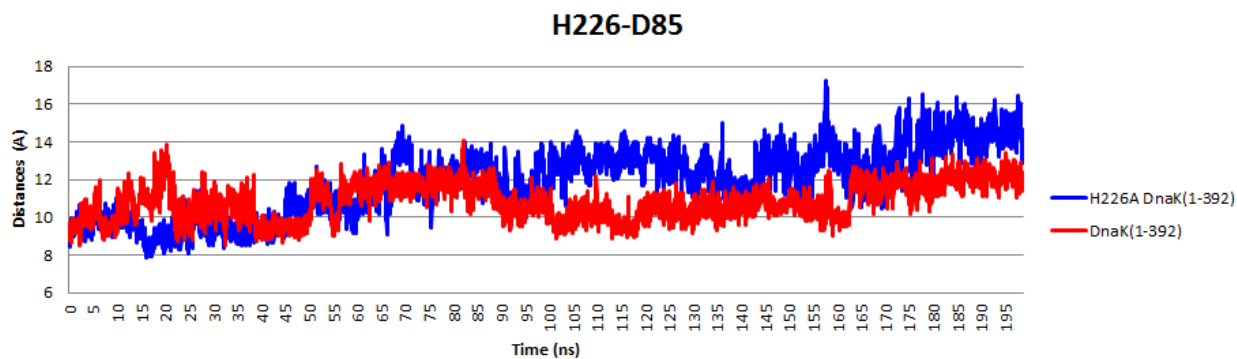
Figure 3.24 : αC distance between H226-D85 residues for model H226A DnaK (1-392) and DnaK (1-392).

Distance between His226 and Asp85 was observed to be increased for both constructs. However this increase is higher for H226A mutant.
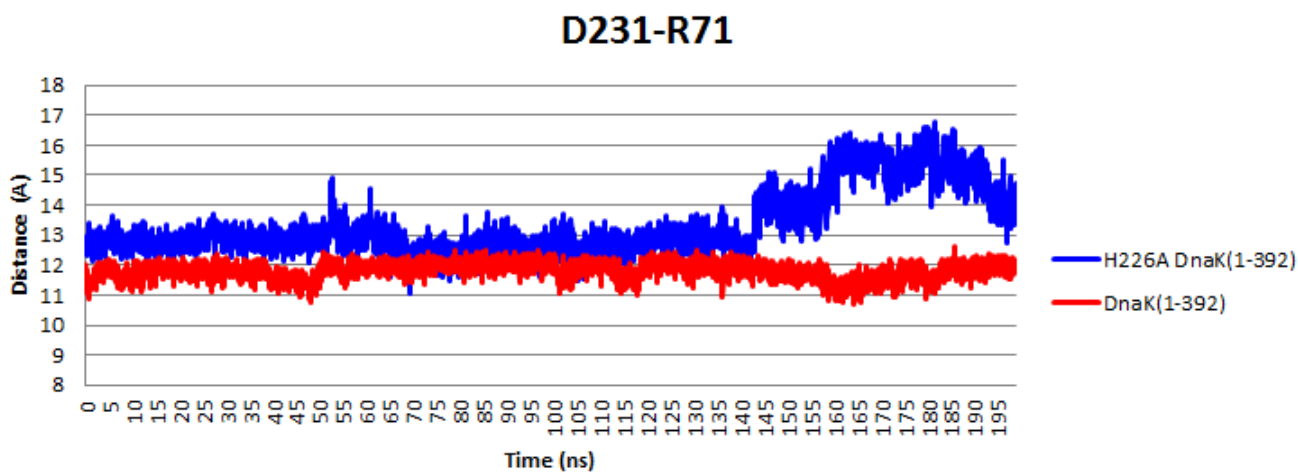
Figure 3.25 : αC distance between D231-R71 residues for model H226A DnaK (1-392) and DnaK (1-392).

Distance between the residues Asp231 and Arg71 were higher for mutant H226A during the whole simulation, however a significant change is observed after 140 ns.
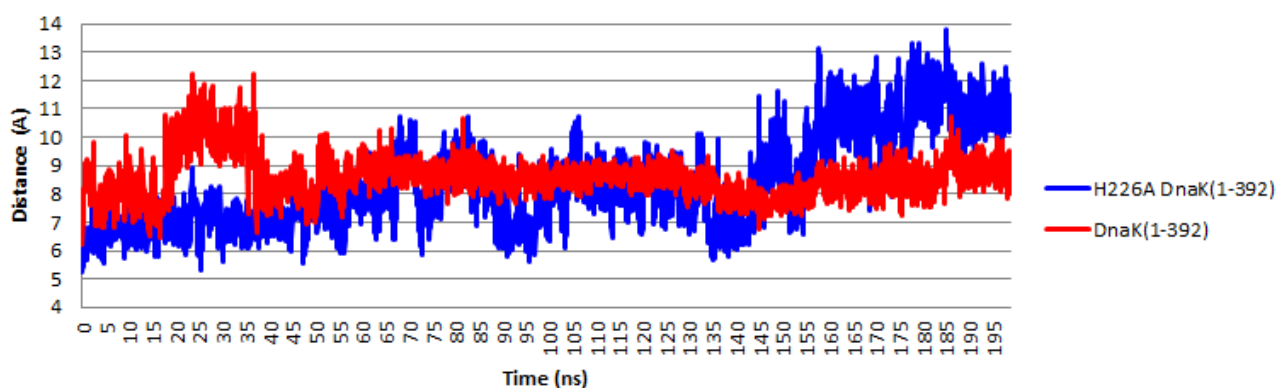
Figure 3.26 : αC distance between H226-E81 residues for model H226A DnaK (1-392) and DnaK (1-392).

Distance between His226 and Glu81 were higher for mutant H226A DnaK (1-392). This difference is especially seen after 145 ns.

Figure 3.27 : αC distance between D231-R84 residues for model H226A DnaK (1-
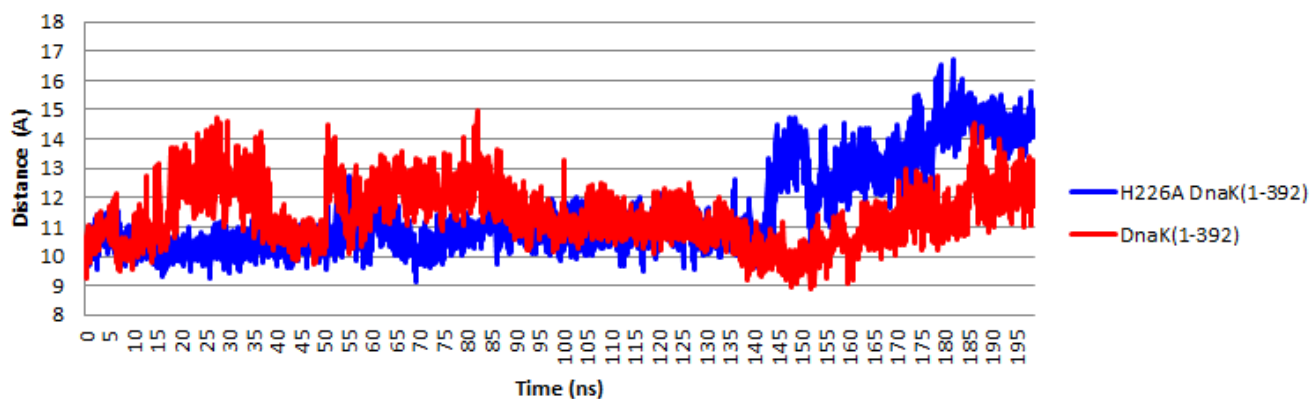
392) and DnaK (1-392).

It is seen that distance between Asp231 and Arg84 is increasing for both H226A DnaK (1-392) mutant and DnaK (1-392) after 140 ns. However, this increase is higher for mutant H226A DnaK (1-392).
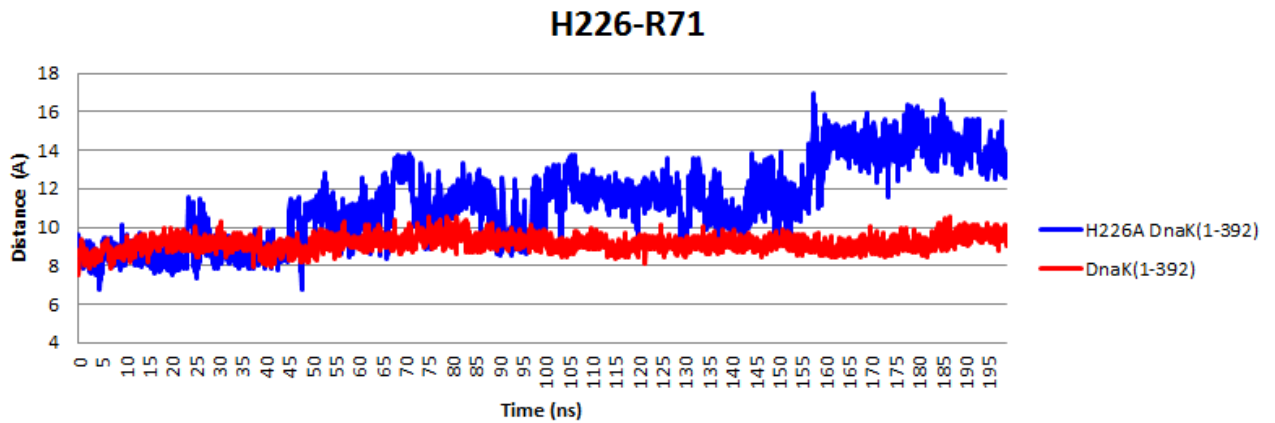
Figure 3.28 : αC distance between H226-R71 residues for model H226A DnaK (1-392) and DnaK (1-392).

While distance between His226 and Arg71 has a stable profile in DnaK (1-392), when His226 is mutated a considerable change has observed after 160 ns.
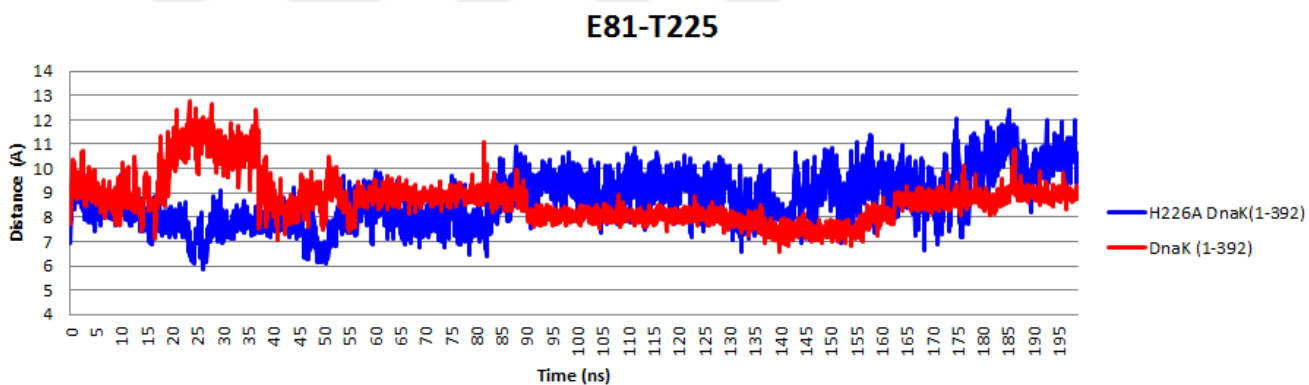
Figure 3.29 : αC distance between E81-T225 residues for model H226A DnaK (1-392) and DnaK (1-392).

Distance between the residues Glu81 and Thr225 is higher for mutant H226A after 90 ns.
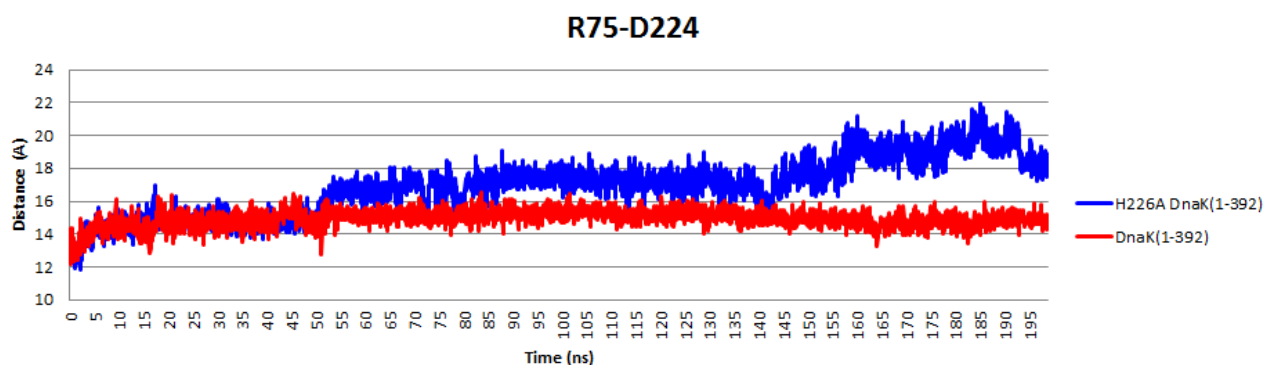
Figure 3.30 : αC distance between R75-D224 residues for model H226A DnaK (1-392) and DnaK (1-392).

Distance between Arg75 and Asp224 began to increase after 50 ns for H226A mutant.
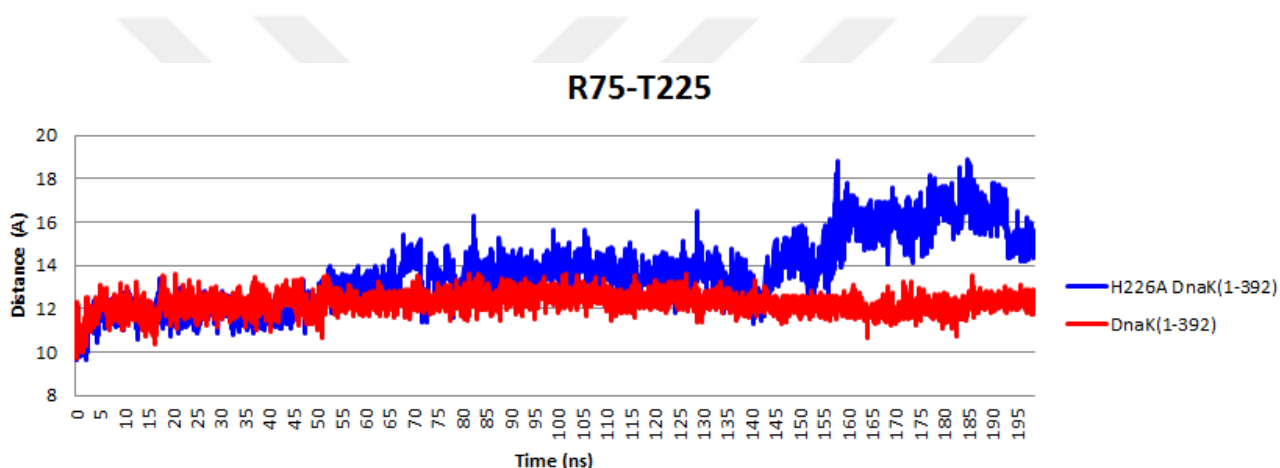
Figure 3.31 : αC distance between R75-T225 residues for model H226A DnaK (1-392) and DnaK (1-392).

Distance between Arg75 and Thr225 is higher for H226A DnaK (1-392) mutant especially after 140 ns. Distance measurement analysis testify that member of relevant helices' are close to each other in DnaK (1-392) construct and so H226A DnaK (1-392) shows more open conformation. In addition, we see bounces at 140 and 160 ns in some couples of distance measurement which is going to be discussed in the section 4.

Principle Component Analysis was performed to see dominant dynamics differences of the relevant structures (Figure 3.32).
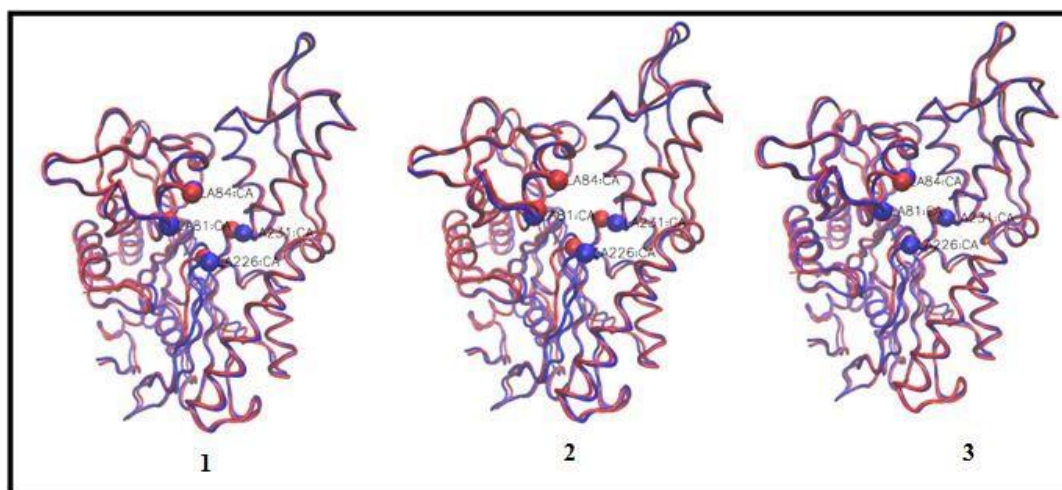
Figure 3.32 : First principal component analysis of DnaK (1-392) and H226A DnaK (1-392), red and blue, respectively. (1,2,3 indicates the time dependent screenshots)

PCA results showed that H226A DnaK (1-392) has flexible motions rather than DnaK (1-392), especially for the helix which includes H226 and D231. While 13.83% of all motions described with first principle component for DnaK (1-392), 30.81% was described for mutant H226A DnaK (1-392).

3.4. His226 Plays a Role on Stabilization of the Active Site by Positioning Thr199 at the Right Position.

Since we realize a big difference in RMSF plot for T199, we decided to focus on His226 and Thr199 relationship. Both in simulations and principal component analysis, it is seen that His226 is significant for catalytically important residue Thr199 positioning.

Molecular dynamics simulations showed that while DnaK (1-392) stable during the simulation time, in H226A DnaK (1-392), Thr199 suddenly moved away at 160 ns for DnaK (1-392) construct and completed the simulation away from ATP (Figure 3.33).
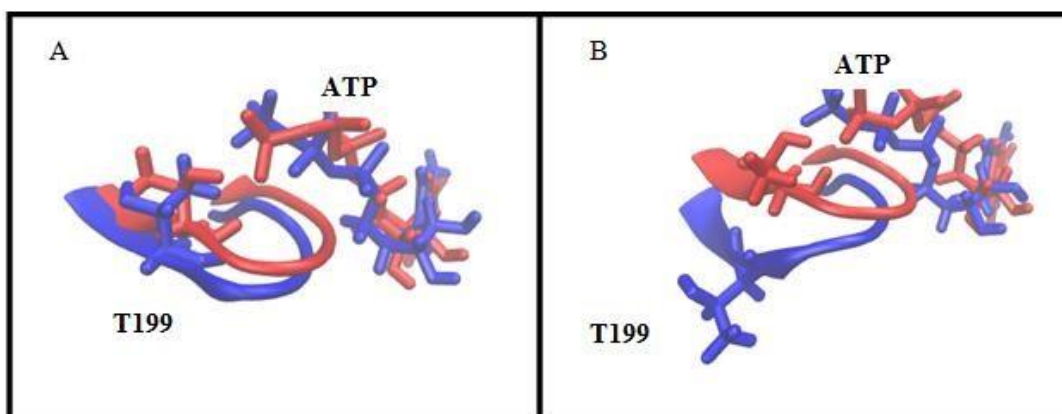
Figure 3.33 : DnaK (1-392) and H226A DnaK (1-392) seen with residue T199 and ATP, red and blue, respectively.(A: Screenshot from 50 ns time point, B: Screenshot from 170 ns time point).

PCA was performed for the DnaK (1-392) and H226A DnaK (1-392). As a result, principal component analysis showed that T199 in H226A DnaK (1-392), has more flexible characteristic (Figure 3.34).
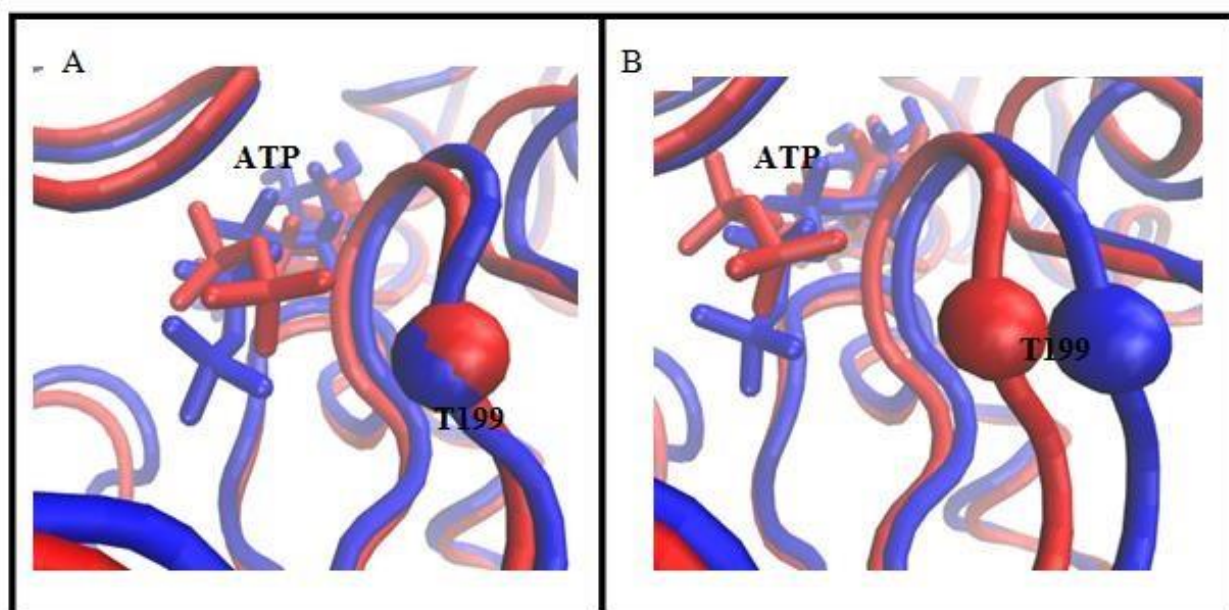
Figure 3.34 : PCA screenshots of different time points DnaK (1-392) and H226A DnaK (1-392), red and blue, respectively, seen with residue T199 and ATP.

Distance measurement analysis between the Thr199 and ATP for the constructs DnaK (1-392) and mutant H226A DnaK (1-392) was performed (Figure 3.34).

T199-ATP
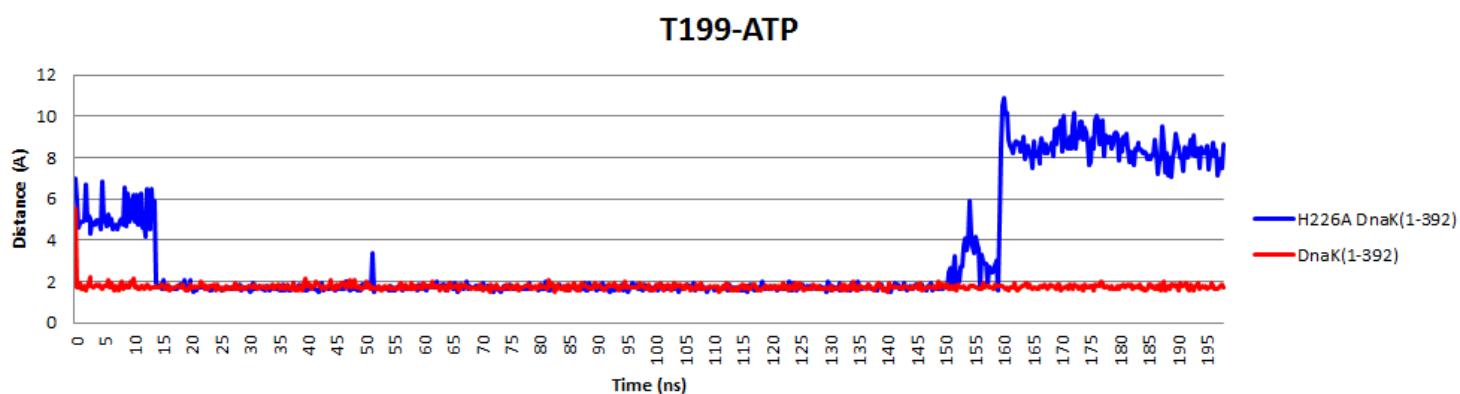
Figure 3.35 : Distance between ATP and Thr199 during simulations.

During the simulation time, distance between Thr199 and ATP is protected for DnaK (1-392). On the other hand, for mutant H226A DnaK (1-392), it is observed that distance between T199 and ATP considerably increase at 160 ns.

4. DISCUSSION AND CONCLUSION

Our findings contribute in the understanding of linker induced dynamic alterations of the ATPase domain conformation in the ATP-bound state. Previous studies showed that the ATPase activity of DnaK (1-392) mimics the ATPase activity of the substrate-stimulated form of the full-length DnaK, whereas DnaK (1-388) has a basal level of ATPase activity mimicking the full-length DnaK in the absence of the substrate. Here, when we analyzed DnaK (1-388) and DnaK (1-392) conformational and dynamical differences in the ATP-bound state, we observed the major effect on the two α-helices that are located on the subdomains IB (residues 71 to 90) and IIB (residues 224 to 250), showing rearrangements at the subdomain interface, agreeing with the previous NMR study that revealed large perturbations with nucleotide binding (Zhuravleva and Gierasch, 2011). This observation compares well with the preferential domain adaptation in the ATP-bound state conformation for linker binding.

Our simulations revealed major differences on the dynamics, compactness and flexibility of subdomains of DnaK (1-388) and DnaK (1-392). Differences observed in the helices belongs to IB (residues 71 to 90) and IIB (residues 224 to 250) in DnaK(1-388) construct agrees well with the experimental findings of their studies that these two helices are more compact than that of DnaK(1-392) for the regions of upper arms of V-shaped NBD (subdomains IB and IIB) when ATP is bound. The inter-subdomain distance of IB-IIB domains is increased with the hydrophobic linker ^{389}VLLL^{392} in the ATP-bound state. It is revealed that presence of ATP in the nucleotide binding domain increases the chance of linker binding by allowing conformational change of the NBD to a more linker interacting position as observed with the consequence of chemical-shift changes on this domain (Zhureavleva and Gierasch,2011). Additionaly, these findings suggested that interface between the NBD lobes is perturbed because of linker binding on the hyrophobic cleft (between IA and IIA) showing an allosterically-induced lobe reorientation. In their NMR results. Also, they observed C-terminus domain of DnaK (1-388) to be more mobile

as observed in our simulations. When linker interacts with the ATPase domain it binds to an edge strand of the β-sheet (residues between 215 and 220) in subdomain IIA in the ATP-bound state. So, the presence of the linker sequence ^{389}VLLL^{392} results in a reduced flexibility of subdomain IIA and increased compactness of subdomain IA observed in our simulations which are in agreement with the findings of Zhueravleva, 2011 and Chiappori, 2016, indicating possible linker interactions changing the flexibility and compactness of these domains which are important for allosteric communication. Thus, according to our study, the two below arms (IA-IIA) of V-shaped NBD structure becomes more stable with ^{389}VLLL^{392}, whereas the two upper arms (IB-IIB) of V-shaped NBD structure gets wider. As it is known that ATPase reaction occurs in the center of the V-shaped cleft (Zuiderweg et.al., 2013), the wider the cleft in DnaK (1-392) can results in a higher rate of ATPase activity.

The biggest difference in the upper arms of V-shaped NBD, especially for the helices, which containing the residues 71 to 90 (IB) and 224 to 250 (IIB), for DnaK (1-388) and DnaK (1-392) constructs call into doubt if these helices may have a role in determining the rate limiting step as a gate since it is on the way of P_i release. Because previous studies support that the helix contains the hinge residues which are located in IIA and IIB, G223, L227, G228 and G229, may play a role in nucleotide-dependent structural allostery and transitions (Ung et.al., 2013, Chiappori et.al., 2012). On the other hand, residues interacting with the nucleotide exchange factor GrpE, are located in the helices which contains the residues 255 to 275 (IIB) and 51 to 62 (IB) (Bukau et.al., 2001) give us thoughts about the ADP-release pathway. In our study, it is observed that way for ADP release is nearly closed because of the approximate positioning of the suspected helices in the DnaK (1-392). This fact, may lead to a slower ADP release in DnaK (1-392) which is consistent with the experiments that are previously done (Imamoglu, 2014, Tubitak project: 110T434). The difference in distance of these gates, is clearly the consequence of the linker interactions. Previously, it is known from Swain et al. study (2007) that the ADP release is the rate limiting step of DnaK (1-392) constructs. Findings from this study and the data from our study together, reveal a model for pathways for ATP hydolysis products (Figure 4.1).
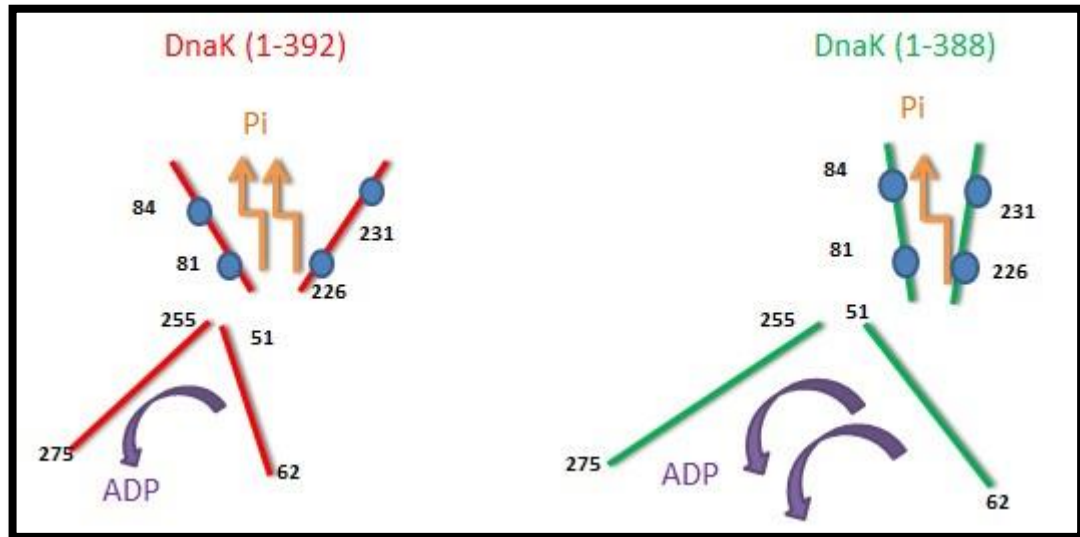
Figure 4.1: Possible releasing pathways for ATP hydrolysis products in DnaK (1-388) and DnaK (1-392). (Two arrows indicates higher release rate)

While the openness of the oppositely located helices which contain the residues between 71 to 90 and 214 to 250, may facilitate easier and faster release of P_i through this gate, closer bearing of the oppositely located helices which contain the residues between 255 to 275 and 51 to 62, may result in slower the ADP release. This allows us to detect the rate limiting step as the ADP release rate for DnaK (1-392) construct. On the other hand, from our simulations for DnaK (1-388), it is seen that way for ADP release is more open and suggested gate for P_i release is closed on the contrary to DnaK (1-392). Since it is clarified before, ADP release is not the rate-limiting step for DnaK (1-388) (Swain et.al., 2007, Tubitak project: 110T434), we suggest that the closed route may be the sign of the rate-limiting step for P_i release. These suggestions are also consistent with experimentally measured overall ATPase and single turnover ADP release rates which are previously done by our research group (Imamoglu, 2014, Tubitak project 110T434).

Finally, we realized that His226 has an important role on the stabilization of Thr199 in DnaK (1-392) construct. Thr199, is a highly conserved residue in Hsp70 family, from DnaK to human Hsc70 (Hunt and Moromoto, 1985). Thr199 in all members of Hsp70 family, is positioned preciding aspartic acid and three glycines and followed by a phenylalanine and an aspartic acid. This conservation of Thr199, makes the residue important for the active site. Such that, when Thr199 mutation to alanine, showed an nearly abolished, ATPase activity of DnaK. Thr199 is located very close to a γ-phosphate of ATP which also makes Thr199 as a suspected leading actor of the

autophosphorilation site for DnaK, which means the first residue that interacting with the detached P_i (McCarty and Walker, 1991). In our study the sudden conformational change in H226A DnaK (1-392) construct between 140 and 160 ns, causes the relevant gate (helices contains the residues 71 to 90 and 224 to 250) to open, leading to a sharp orientational rearrangement of the sidechain of Thr199 that increases the distance between Thr199 and ATP. This observation could affect the proper coordination of P_i and may results in higher P_i release rate. When T199 loses its interaction with P_i, the release of ATP hydrolysis products might be faster. Moreover, when effect of this situation combined with a more open conformation for mutant H226A DnaK (1-392), it becomes consistent with an experimental three fold change in the ATP hydrolysis rates and ADP release rates which is previously done by our research group (Tubitak project 110T434).

To conclude, with this study, we show the effect of the linker sequence ^{389}VLLL^{392}, and the importance of residue His226 for the general conformation and dynamics of the ATPase domain in a linker-induced conformation as observed for DnaK (1-392) construct. Further investigation must be done for critical residues which belong to the ‒gate‖ either *in silico* or *in vitro* experiments especially for enlightening the P_i release pathway.

REFERENCES

Abdi, H., Williams, L. J. (2010). Principal Component Analysis. *WIREs Comp Stat*, (**2**): 433-459. doi: 10.1002/wics.101

Arakawa,A., Handa,N., Shirouzu,M., Yokoyama,S. (2011). Biochemical and structural studies on the high affinity of Hsp70 for ADP. *Protein Science* (**20**):1367—1379

Barry J. Grant, B. J., Rodrigues, A. P. C., ElSawy, K. M., McCammon, J. A., Caves, L. S. D. (2006) Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics* 22 (**21**): 2695-2696. doi: 10.1093/bioinformatics/btl461.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland,G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E. (2000). The Protein Data Bank, *Nucl. Acids Res.* 28 (**1**):235-242.doi: 10.1093/nar/28.1.235.

Bertelsen, E.B., Chang, L., Gestwicki, J.E., Zuiderweg, E.R.P. (2009). Solution conformation of wild-type E. coli Hsp70 (DnaK) chaperone complexed with ADP and substrate. *PNAS* vol. 106 (**21**)**:** 8471–8476

Brehmer, D., Rüdiger, S., Gässler, C. S., Klostermeier, D., Packschies, L., Reinstein, J. et al. (2001) Tuning of chaperone activity of Hsp70 proteins by modulation of nucleotide exchange. *Nat. Struct. Biol.* (**8**)**:** 427–432.

Brooks, B. R., Brooks III, C. L., MacKerell Jr., A. D., Nilsson, L. , Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A.R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M., Karplus, M. (2009). CHARMM: The Biomolecular Simulation Program. *J Comput Chem.* July 30; 30(**10**): 1545–1614. doi:10.1002/jcc.21287.

Calloni, G., Chen ,T., Schermann , S. M., Chang , H., Genevaux , P., Agostini F., Tartaglia , T. T., Hayer-Hartl , M., Hartl , F. U. (2012). DnaK Functions as a Central Hub in the *E.coli* Chaperone Network. *Cell Reports* 1, 251–264.

Carugo , O., Pongor , S. (2001) A Normalized Root-Mean-Spuare Distance For Comparing Protein Three-Dimensional Structures. *Protein Sci. 2001* Jul; 10(**7**): 1470–1473.

Chiappori, F., **Merelli, I., Colombo, G., Milanesi, M., Morra, G.** (2012). Molecular Mechanism Of Allosteric Communication İn Hsp70 Revealed By Molecular Dynamics Simulations. *PLOS Comp. Bio.* Volume 8, Issue 12, e1002844.

Chiappori, F., Merelli, I., Milanesi, L., Colombo, I., Morra, G. (2016). An Atomistic View Of Hsp70 Allosteric Crosstalk: From The Nucleotide To The Substrate Binding Domain And Back. *Scientific Reports,* (**6**)**,** Article number: 23474.

Coombe, G. An Introduction to Principal Component Analysis and Online Singular Value Decomposition. (2006). https://www.cs.unc.edu/~coombe/research/phd/svd.pdf

Essmann, U., Perera1, L., Berkowitz1, M. L., Darden, T., Lee, H., Pedersen, L. G. (1995). A smooth particle mesh Ewald method. *J. Chem. Phys*. 103, 8577.

Frenkel, D., Smit, B., Understanding molecular simulation: From algorithms to applications. Academic Press, 196 (2nd Ed.), p. 97.

Frydman, J., Nimmesgern, E., Ohtsuka, K., and Hartl, F. U. (1994). Folding of nascent polypeptide chains in a high molecular mass assembly with molecular chaperones. *Nature* **370**, 111 - 117; doi:10.1038/370111a0

Golas,E, Maisuradze, G.G., Senet,P., Oldziej, S., Czaplewski, C., Scheraga, H.A., Liwo,A. (2012). Simulation of the Opening and Closing of Hsp70 Chaperones by Coarse-Grained Molecular Dynamics. *J. Chem. Theory Comput.* **8**, 1750−1764.

Gonzalez, M. A. (2011) Force fields and molecular dynamics simulations. *Collection SFN* 12 169-200.

Harrison, C.H., Hayer-Hartl, M., Liberto, M., Hartl, F.U., Kuriyan, J. (1997). Crystal Structure of the Nucleotide Exchange Factor GrpE Bound to the ATPase Domain of the Molecular Chaperone DnaK. *Science* **276**: 431-435.

Hartl, F. U.,and Hayer-Hartl, M. (2002). Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science*, *295*(5561), 1852-1858.

Haslbeck, M. (2002). sHsps and their role in chaperone network. *Cell Mol Life Sci.* Oct;59(10):1649-57.

Hunt, C., Moromoto, R.I. (1985) Conserved features of eukaryotic hsp70 genes revealed by comparison with the nucleotide sequence of human hsp70. *Proc. Natl. Acad. Sci. USA,* **82**, 6455-6459.

Hurley, J. H. (1996). The Sugar Kinase/Heat Shock Protein 70/Actin Superfamily: Implications of Conserved Structure for Mechanism. *Annu. Rev. Biophys. Biomol. Struct.* *25*, pp. 137-162.

Karplus, M., Kuriyan , J. (2005) Molecular dynamics and protein function. *PNAS* vol. 102 , no.19 , 6679-6685.

Kim, Y. E., Hipp , M.S., Bracher , A., Hayer-Hartl , M., Hartl , F. U. (2013) Molecular Chaperone Functions in Protein Folding and Proteostasis. *Annu. Rev. Biochem.* 2013. 82:323–55 ; doi : 10.1146/annurev-biochem-060208-092442

Kityk, R., Vogel , M., Schlecht , R., Bukau ,B., Mayer , M. P. (2015). Pathways of allosteric in Hsp70 chaperones. *Nature Communications* | 6:8308 | DOI: 10.1038/ncomms9308.

Kityk, R., Kopp, J., Sinning, I., and Mayer, M. P. (2012). Structure and Dynamics of the ATP-Bound Open Conformation of Hsp70 Chaperones. *Molecular Cell*. *48*, pp. 1-12.

Kuzmanic,A, Zagrovic,B. (2010). Determination of Ensemble-Average Pairwise Root Mean-Square Deviation from Experimental B-Factors. *Biophys J*. 2010 Mar 3; 98(5): 861–871. doi: 10.1016/j.bpj.2009.11.011.

Li, J. & Buncher, J. (2013) Structure, Function and Regulation of the Hsp90. Machinery *Biomed J* 36,106-117.

Li, J. (2005) Handbook of Materials Modeling, Springer, pp 565-588.

Liang, P. & MacRae, T. H. (1997) Molecular chaperones and the cytoskeleton. *Journal of*

Cell Science 110, 1431-1440.

Lobanov, M.Y.,Bogatyreva, N.S., Galzitskaya, O.V. (2008).Radius of Gyration as an Indıcator of Protein Structure Compactness. ISSN 0026-8933, *Molecular Biology*, 2008, Vol. 42, No. 4, pp. 623–628.

MacKerell Jr. , A. D. §, Bashford , D., Bellott , M., Dunbrack Jr. , R. L., Evanseck , J. D., Field , M. J., Fischer , S., Gao , J., Guo , H., Ha , S., Joseph-McCarthy , D., Kuchnir , L., Kuczera , K., Lau F. T. K., Mattos , C., Michnick , S., Ngo T., Nguyen , D. T., Prodhom , B., Reiher , W. E., Roux , B., Schlenkrich , M., Smith , J. C., Stote ; R., Straub , J., Watanabe , M., Wiórkiewicz-Kuczera , J., Yin , D. and Karplus , M. (1998). All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B*, 102 (18), pp 3586–3616 ; doi : 10.1021/jp973084f

Maisuradze, G. G., Liwo, A., Scheraga, H. A. (2009) Principal component analysis for protein folding dynamics. *J Mol Biol*. January 9; 385(1): 312–329. doi:10.1016/j.jmb.2008.10.018

Mayer, M. P. (2013). Hsp70 chaperone dynamics and molecular mechanism. *Trends in biochemical sciences*, *38*(10), 507-514.

Mayer, M. P., Bukau, B. (2005). Hsp70 chaperones: cellular functions and molecular mechanism. *Cellular and molecular life sciences*, *62*(6), 670-684.

McCarty, J.S., Walker, G.C., (1991). DnaK as a thermometer: Threonine-199 is site of autophosphorilation and is critical for ATPase activity. *Proc. Natl. Acad. Sci..USA*. Nov 1;88(21):9513-7.

Meller, J., Molecular Dynamics, (2001) , Encyclopedia Of Life Sciences / 2001 *Nature* Publishing Group / www.els.net

Nicolai, A., Delarue, P., Senet, P. Decipher the Mechanisms of Protein Conformational Changes Induced by Nucleotide Binding through Free Energy Landscape Analysis: ATP Binding to Hsp70. (2013). *PLoS Comput Biol.* 9(12): e1003379. doi:10.1371/journal.pcbi.1003379.

Nosé, S. (1984). A unified formulation of the constant temperature molecular-dynamics methods. *Journal of Chemical Physics.* 81 (1): 511–519.

Phillips , J.C., Braun , R., Wang , W., Gumbart , J., Tajkhorshid , E., Villa , E., Chipot , C., Skeel , R. D., Kalé , L., Schulten , K. (2005). Scalable molecular dynamics with NAMD. *J Comput Chem.* (16):1781-802.

Qi, R., Sarbeng, E.B., Liu, Q., Le, K.Q., Xu, X., Xu, H., Yang, J., Wong, J.L., Vorvis, C., Hendrickson, W.A., Zhou, L., Liu, Q. (2013). Allosteric opening of the polypeptide-binding site when an Hsp70 binds ATP. *Nat.Struct.Mol.Biol.* 20**: 900-907.**

Ramachandran, K. I., Deepa, G., Namboori, K. (2008). Computational Chemistry and Molecular Modeling, Principles and Applications. *Heidelberg : Springer*

Ringner, M., (2008) What is principal component analysis?. *Nature Biotechnology*. 26, 303-304.

Rüdiger, S., Germeroth, L., Schneider-Mergener, J., and Bukau, B. (1997). Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *The EMBO journal*, *16*(7), 1501-1507.

Ryckaert, J.P., Ciccotti, G., Berendsen, H. J. C. (1977). Numerical integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. Journal Of Computational. *Physics* 23, 321-341.

Schreiner, W., Karch , R., Knapp , B., Ilieva , N. (2012). Relaxation Estimation of RMSD in Molecular Dynamics Immunosimulations. *Computational and Mathematical Methods in Medicine* , Vol 2012, Article ID 173521 , 9 pages ; doi: http://dx.doi.org/10.1155/2012/173521

Stetz, G., Verkhivker, G.M. (2015). Dancing through Life: Molecular Dynamics Simulations and Network-Centric Modeling of Allosteric Mechanisms in Hsp70 and Hsp110 Chaperone Proteins. *PLoS ONE* 10(11): e0143752.

doi:10.1371/journal. pone.0143752

Swain, J., Dinler, G., Sivendran, R., Montgomery, DL., Stotz, M., Gierasch, L.M. (2007). Hsp70 chaperone ligands control domain association via an allosteric mechanism mediated by the interdomain linker. *Molecular Cell 26,* 27-39.

Ung, PM-U., Thompson, A.D., Chang, L., Gestwicki, J.E., Carlson, H.A. (2013) Identification of key hinge residues important for nucleotide-dependent allostery in *E. coli* Hsp70/DnaK. *PLoS Comput Biol* **9**(11): e1003279. doi:10.1371/journal.pcbi.1003279

Yip, S. (ed.). (2005). Handbook of Materials Modeling, *Springer* 565–588.

Zhuravleva, A., Clerico, E. M., and Gierasch, L. M. (2012). An interdomain energetic tug-of-war creates the   allosterically active state in Hsp70 molecular chaperones. *Cell*, *151*(6), 1296-1307.

Zhuravleva, A., Gierasch, L.M. (2011). Allosteric signal transmission in the nucleotide-binding domain of 70-kDa heat shock protein (Hsp70) molecular chaperones. *Proc Natl Acad Sci USA* 108: 6987–699

CURRICULUM VITAE

Elif CAKMAK

Personal Information

Birthday: 13/08/1991
Ulus mah. Oztopuz Cad. No:36/7
Besiktas/ ISTANBUL/TURKEY
M: +905453683211
e-mail: cakmakelif@itu.edu.tr

Education

- **February 2014 – December 2016**

ITU Graduate School of Science, Engineering & Technology – Molecular Biology, Genetics & Biotechnology Master Programme ISTANBUL/TURKEY

- **September 2009 – February 2014**

Yildiz Technical University- Science & Letter Faculty – Chemistry ISTANBUL/TURKEY

- **February 2013 – July 2013**

Universidade de Lisboa Faculdade de Ciencias – Quimica & Bioquimica - Erasmus Student Exchange Program LISBON/PORTUGAL

- **September 2012- present**

Anadolu University – Open Education Faculty - Business Administration

Work Experiences

- Gebze Technical University – Project Assistant (July 2016 – present)
- Mene Research – Clinical Research Site Coordinator (June 2015 - September 2015)

- Istanbul Technical University (in Molecular Biology & Genetics Student Laboratories) – Student Assistant (September 2014- June 2015)

- Cerrahpaşa Medical Faculty – Project based Researcher (June 2012- August 2013)

- IN-KA Young Human Resources– Organization Staff (June 2010- September 2012)

- Pfizer –Intern (August – October 2011)

Certificates

ISO 9001:2008
ISO/IEC 17025
Istanbul Language Center- Spanish Course (A1)
Universidade de Lisboa – Portuguese Course (A2)
YTU Quality & Productivity Club- Kariyerimin Kimyası Organization Team
ITU International Molecular Biology & Genetics Student Congress Organization Team
GC –ETC Training and Consulting
HPLC – ETC Training and Consulting
EuPA 2016 -Biosimilar Drug Analysis Workshop

Interests

Tango,Puzzle,Jogging