INVESTIGATION OF DYNAMIC ALLOSTERY IN MOTOR PROTEINS

by

Cihan KAYA

B.S, Chemical Engineering, Boğaziçi University, 2010

B.S, Physics, Boğaziçi University, 2010

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

Graduate Program in Chemical Engineering

Boğaziçi University

2013

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my advisor Prof. Turkan Haliloglu for endless support to my studies and for patience, motivation, enthusiasm and limitless knowledge. Her guidance provided me a lot in all the time of research and writing this thesis. Her continuous and illuminative support in my life as well as my academic studies has been a speechless help. It has been a great pleasure and a privilege to work with such peerless person.

Besides my advisor, I also would like to thank Prof. Pemra Doruker Turgut and Asst. Prof. Hamdi Torun for their encouragement and insightful comments during my studies.

I wish to express my deepest appreciation to my friends and colleagues; Fidan Sumbul and Burcu Aykac Fas for helping on AFM experiments, endless supports throughtout my research, Sina Tureli for his great contribution and effort on the development of force distribution analysis algorithm, Andac Armutlulu and Solen Ekesan for their valuable ideas and advices on the development of MCPath server, Arzu Uyar for being a sister for me, Seren Soner for his intelligent comments, Pemra Ozbek for support in all my studies and all the other people who make office as lively and shaing for me. I would like to hail everyone in Chemical Engineering Department from staff to student and faculty members for their support and help. Also, I would like to thank to friends Irem Sen, Erge Akbas, Begum Alaybeyoglu, Oya Gursoy, Melis Yıldırım and Ferit Ozbakır. Gizem Ozbuyukkaya, thank you for being my side through everything and her enlightening support and incomprehensible help.

Last but not the least, I would like to express my recognition to my family for giving birth to me at first place and supporting me internally throughtout my life.

Financial support provided by TUBITAK through project 110T088 and TUBITAK grad scholarship.

ABSTRACT

INVESTIGATION OF DYNAMIC ALLOSTERY IN MOTOR PROTEINS

Allostery is a key concept in regulation of protein function. The communication pathways which are equivalent of the allosteric signaling or energy propagation pathways are the most significant aspect of allosteric mechanism. A Brownian dynamics simulation approach with force distribution analysis is proposed to determine key residues that have a role in allosteric signal transmission. Along, another stochastic approach -Monte Carlo path generation- is also implemented to define likely allosteric pathways through generating an ensemble of maximum probability paths. The force distribution analysis is tested on methionine repressor and purine repressor. The Monte Carlo path generation method is applied on various proteins such as PDZ domain, bovine rhodopsin and three different structures of myosin. The predicted residues that have a significant effect on the allosteric signaling are in agreement with the previous studies. As a case study, the main focus is Kinesin motor protein, which is a key element of cargo transport in intracellular environment. The ATP binding and the motility are directly related and communicated based on allosteric signaling in this molecule. The analyses based on both methodologies are aimed to define allosteric pathways and key sites in signaling. The predicted regions by both methods are successful in determination of the residues that plays a key role in signaling previously determined by other studies. Based on the graph centrality parameters and force distribution analysis, a novel allosteric site proposed (Cys 176). For the validation of algorithms as well as the specific allosteric sites therein, pulling experiments with atomic force microscopy are considered and performed experiments are simulated with some design plans proposed. The change in the rupture force by application of external force from different sites may provide information about the validity of the computational approaches and validate importance of the predicted sites. As a further study, three predicted sites will be tested via different functionalization methods. Overall, the proposed methodologies have potential to identify residues of importance in signaling event.

ÖZET

MOTOR PROTEİNLERİN DİNAMİK ALLOSTERİSİNİN İNCELENMESİ

Allosteri protein fonksiyonunun düzenlenmesinde anahtar bir kavramdır. İletişim yolları, allosterik sinyal veya enerji yayılma yollarının eşdeğeri olan, allosterik mekanizmasının en önemli yönüdür. Monte Carlo patika üretimi ve kuvvet dağılımı analizi allosterik sinyal iletiminde önemli rolü olan rezidüleri belirlemek için yeni yöntemler olarak önerilmiştir. Metodoloji genellikle hücre içi ortamda kargo taşımacılığı temel unsuru olan motor proteinlere uygulanmıştır. Kinesin proteini, ATP bağlanması ve hareketliliği arasındaki doğrudan ilgi ve allosterik sinyal geçişine dayalı iletişime bağlı olduğu için burada bir vaka çalışması olarak kullanılır. Analizlerin amacı bilinen önemli allosterik bölgeler dışında başka bir anahtar bölge bulmak olup, kullanılan iki metodun sonucunda yeni bir bölge grafik merkezi parametreleri ve kuvvet dağılım analizleri sonucunda önerilmiştir. Monte Carlo yolu oluşturma yöntemi için PDZ proteinleri, sığır rodopsin proteini ve kuvvet dağılımı için metionin represörü ve pürin represörü üzerinde test çalışmaları yapılmıştır. Tahmin edilen allosterik rezidüler önceki çalışmalar ile bir uyum göstermektedir. Algoritmaların doğrulanması için, atomik kuvvet mikroskobu ile çekme deneyleri yapılmaktadır. Farklı bölgelerden kuvvet uygulaması ile kopma kuvveti değişim hesaplanması, yaklaşımların geçerliliği hakkında bilgi verecektir. Önerilen çalışma olarak, yeni bulunan bölgeler farklı işlevlenme yöntemleriyle test edilecektir. Genel olarak, önerilen yöntemler sinyal olaylarının önemi olan rezidüleri tanımlamak için potansiyele sahiptir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

A_i	random force term on ith residue
Å	angstrom
b_i	betweenness value of residue i
dt	time step
E	energy
E_i	total energy residue i
E_{ij}	energy between residue i and residue j
fs	femtosecond
g_{ij}	a variable to hold number of shortest path pairs
g_{ikj}	a variable to hold number of shortest path pairs passing through k
$\boldsymbol{I}$	identity matrix
k_B, k	Boltzmann constant
kcal	kilocalorie
kg	kilogram
l_{ij}	length of shortest path between i and j
m	mass of atomic center
M	molar
ns	nanosecond
O_i	closeness value of residue i
P_{ij}	probability of interaction of i and j
pN	piconewton
r_{cut}	cut-off distance
$\boldsymbol{r}_i$	position vector of ith residue
r_{ij}	distance between residue i and residue j
r_{min}	minimum distance for energy calculation
t	time
T	absolute temperature
u	step function

V	potential energy
W_{ij}	weight of interaction of i and j
α	Alpha
β	normalized friction coefficient
β	temperature factor
δ_{ij}	Dirac delta function
σ	collusion diameter
ε	Lennard-Jones well-depth
μ	micro
ξ	friction coefficient
γ	spring constant

LIST OF ACRONYMS/ABBREVIATIONS

ADP	adenosine di-phosphate
AFM	atomic force microscopy
ATP	adenosine tri-phosphate
BD	Brownian dynamics
BRB80	buffer solution
C	cystine
D	aspartic acid
DNA	de oxi-ribo nucleic acid
E	glutamic acid
F	phenylalanine
G	glycine
GTP	guanosine tri-phosphate
GUN	guanine
H	histidine
I	isoleucine
K	lysine
KIF	kinesin superfamily
kin410-BCCP	biotinylated kinesin from coiled coil region with 410 residue length
L	leucine
M	methionine
MC	Monte Carlo
MCPath	Monte Carlo path generation
MD	molecular dynamics
MT	microtubule
N	asparagine
NMR	nuclear magnetic resonance
PDZ	name of protein family
P	prolin
P	phosphate
PDB	protein data bank

Q	glutamine
R	arginine
RNA	ribo nucleic acid
S	serine
SAM	S-Adenosylmethionine
SEM	scanning electron microscopy
T	threonine
TEM	transmission electron microscopy
W	tryptophan
V	valine
Y	tyrosine

1. INTRODUCTION

1.1. Allostery

Allostery is a vital concept for protein activity especially in regulation. Many significant processes in cellular cycle such as enzymatic activity and binding are tuned by perturbation of the structure from a distal site such as ligand binding or chemical modification sites. In classical view of allostery, the signaling requires conformational changes [1-3]. Yet, the new view of allostery focuses on the presence of different conformational states and dynamics in allostery [4]. Therefore, the new view changes the perspective on allostery directly to multiple pathways instead of single common route [5]. An elusive and very significant concept is the propagation of a signal from binding sites to allosteric sites through non-covalent interactions via multiple pathways.

The previous efforts to elucidate this phenomenon are generally based on experimental evidences. For instance, Nuclear magnetic resonance (NMR) is a commonly used instrumentation for investigation of the allosteric pathways [6-9]. Apart from NMR, the dynamics of allosteric events are key elements for determination of signaling pathways [7]. Another common study for resolving this problem is site-directed mutagenesis which involves in checking the activity of proteins in different conditions involving different mutations on the probable residues that plays a key role in allosteric signaling [10, 11]. As complementary to experimental studies, there are various computational efforts struggled for thoroughly investigation of allosteric signaling [5]. Those approaches involve in energetic connectivity from evolutionary data [12], anisotropic thermal diffusion [13], optimal path generation on residue networks [14], perturbation response scanning [15], Markov processes for information propagation [16], dynamics of proteins from modeled energy landscape [17] , geometrical interpretation of protein structure [18], interaction energy networks [19] and information theory [20].

Along the same line of the latter studies, a new approach based on the Monte Carlo (MC) path generation was recently proposed [21, 22] and implemented as a web server

called MCPath [23]. It aims to characterize all possible allosteric communication pathways using an ensemble of highly probable pathways and also predict functional residues through the graph centrality parameters. As demonstrated with the case studies on PDZ domain representative proteins, bovine rhodopsin and various structures of myosin proteins [23], this provides an ensemble of populated pathways different than generic single allosteric pathway defining algorithms. Here, MCPath has been used to predict plausible allosteric pathways.

1.2. Force Distribution Analysis

The question is how the perturbation that causes allosteric events is transmitted throughout the protein structure and what mechanism is involved in the propagation of signal. The mechanical responses of a protein as a result of force applied by the binding of a ligand may result in unfolding, activation and another binding [3]. It is known that the forces applied by binding events propagate through sites that are not directly related with those events. [24]. Yet, those sites may be related with the different cellular functions such as catalytic activity [25]. However, there is not any strict assumption for the propagation of perturbation that makes specific changes in the protein dynamics as to answer the first question. Therefore, the signaling will not necessarily cause a conformational change [26]. The phenomenon allostery is stated as the transition between states called allosteric states which are pre-existing conformations that are more populated upon allosteric perturbation [27]. It is a known fact that allosteric perturbation at the regulatory site is very similar to the application of natural force such as a ligand force to that site. The perturbation dissipates in the structure as an internal strain and may result in physical events such as conformation changes and mechanical vibrations, yet the effect of perturbation dissipation on the signaling is more effective than the physical events [28]. The signal propagation is effective for cellular function and the efficacy of the protein is enhanced or reduced with this force application [29]. On the other hand, The allosteric effect does not always result in conformational changes as in the case of enzyme-receptor complexes [30]. In this case, the dissipation of the ligand force plays far more fundamental role to resolve allosteric phenomenon. The method called force distribution analysis allows to detect the internal stress propagation [24, 31, 32]. This is a very strong tool to predict the dissipation of forces such as ligand forces that will not lead to conformational changes. Therefore, the force

distribution analysis may be a very significant tool to decipher underlying mechanism beneath allosteric signal propagation. The conformational changes also plays a significant role in allosteric mechanisms however, it is known that the intrinsic dynamics of the structures are robust. Hence, the force distribution analysis may also be useful for systems with conformational changes.

Force distribution analysis needs to be performed to determine the stress propagation computationally. The mechanism of force propagation on a protein structure still remains elusive. Modified molecular dynamics is a significant tool to investigate force dissipation [32]. The signal propagation may result in a change in the conformation, but there are still many cases that avoid conformational changes and triggers another allosteric event at a distal site [33, 34]. Some phenomenon namely mechanism of the robust immunoglobulin domains, stiff allosteric proteins, the von Willebrand factors in blood vessels and silk fibers are resolved with force distribution analysis [32].

1.3. Motor Proteins (Kinesin)

A synthesized molecule in a cell needs to be actively transported through directional cytoskeletal units to their appropriate destinations by molecular motors. Proteins are carried by various membranous organelles and protein complexes [35, 36]. The polarized cells such as neurons are very good case studies to investigate intracellular transport phenomena, since the transportation of proteins from nerve terminal to axons are vital for both survival and function [37].

Three large protein super families have been identified among molecular motor proteins such as kinesins, dyneins and myosins [38]. Kinesins use microtubules as a path to transport cargo along and they use ATP driven chemical energy to perform conformational changes that provides motility [37]. Dyneins use microtubules as kinesins, yet the motion is minus-end directed unlike kinesins which is backward motion of the motor proteins generally to the cell center. Myosins move along actin filaments to perform muscle contraction and short-range transports around plasma membrane. In this study, main focus will be on kinesin super family also known as KIFs.

Kinesin molecule is a homodimer with two identical cores (head) that binds microtubules and ATP. The identical heads are connected with a neck-linker which is a mechanical element that undergoes conformational changes in the presence of different nucleotide that enables movement of kinesin [39]. The neck-linker is connected to coiled coil that leads to binding to cargo protein. To avoid dissociation while making consecutive steps on microtubules, the two head domains needs to operate in coordination [40]. There are two different hypothesis to define motility of kinesin molecules one of which is hand-over-hand walking [41] and the other one is inch-worm model [27]. The inch-worm model states that one head always the trailing head and the other one is leading head. On the other hand, hand-over-hand walking model states the presence of alternating heads as the lead head. In this study, the focus is hand-over-hand walking.

The kinesin molecules move with hydrolyzed ATP. The rear head domain moves around 15 nm, whereas the other head does not translate due to strong binding to microtubule affected by the ATP binding event. This argument was corroborated by fluorescence imaging techniques [42]. In this model, the motion of kinesin molecule is from one monomer to consecutive monomer of microtubule molecule, namely α-tubulin and β-tubulin, similar to the movement of a person on a row of steeping stones [43]. The mechanism that lies under motility of kinesin molecule is elucidated [44] and the binding of ATP affects that motion [45]. The binding of ATP induce a conformational change on the structure of kinesin thus, the motion of heads are performed via energy and the allosteric signaling produced by binding of ATP [45].

In the absence of ATP, bounded ADP to kinesin molecule and the head domains are bound to the two tubulin binding sites. The neck-linker points forward trailing head and rearward on the leading head. The binding of ATP on the trailing head will initiate and complete the docking of neck-linker to the leading head and it will rotate the trailing head to 160 Å forward toward to next tubulin site [46]. After this point, the trailing head moves forward and start to perform a diffusional search, new leading head tightly docks into next tubulin binding site which will complete 80 Å motion of the attached cargo. At this time, new trailing head completes hydrolysis of ATP and ADP-P_i complex remains. Then, another ATP molecule binds to new leading head and neck-linker begins to knot onto the

core [38]. The trailing head detaches its neck-linker after phosphorus release and it will complete its motion around leading head [39].

The allosteric mechanism is a key for motility of kinesin molecule due to signaling event that is created by ATP binding and propagated through functional sites such as tubulin binding sites and neck-linker. In this work, it has mainly investigated by two computational approaches: force propagation analysis and MC path generation. The binding and unbinding events and their effects on the conformational changes are investigated for both monomer and dimer kinesin molecules.

1.4. Atomic Force Microscopy (AFM)

The atomic force microscope is a very well-respected device to study physical and biological structures and provides an exclusive window to the micro and nano scale environment of cells, organelles and biomolecules [47]. In a physiological environment, AFM has strengths compared to SEM and TEM since it does not require any additional requirements such as vacuum or metal coating. [48]. In AFM the imaging is done by the interaction of probe with surface [49].

One of the main advantages of AFM is that it provides imaging of the non-conducting surfaces. Therefore, it could be easily extended to the biological systems as in the case of organic monolayers or amino acid crystals [50]. The analysis involved in studies related to nucleic acid analysis, protein complexes, membranes and peptides. The analysis are held on a coated glass or mica and in a buffer solution [51].

Another useful area that the application of AFM provided significant advantage is the biomolecular interaction studies. The AFM can investigate force measurements for both specific and non-specific interactions such as DNA replication, protein synthesis, protein-protein, antigen-antibody interactions generally governed by intermolecular forces [52]. Also, AFM has an ability to measure the forces that are involved in binding in the order of pico-Newton range. This makes it possible to in biological systems in a large variety such as ligand-receptor interactions [53]. There is one other key aspect of AFM

studies that it can measure micromechanical properties in a living cell and membranes such as elasticity and viscosity [54].

The allostery may be assumed as conformational and dynamics changes due to the ligand binding. The ligand binding event that changes the stress profile of the structure may be mimicked by an external force from the allosteric site. The recent studies investigated the application of force at the ligand binding region by AFM [55]. Therefore, the dissipation of mechanical stress over allosteric sites may result in the change in the inter-residue interactions that leads to the conformational and dynamics changes. Here, we aimed to determine the stress propagation profile on a kinesin structure with the computational method developed in this thesis, based on the force distribution analysis that would help to provide the molecular understanding of the events observed by the AFM experiments as well as validate the computational prediction.

2. METHODS

2.1. Experimental Methods

2.1.1. Atomic Force Microscopy

2.1.1.1. Fundamental elements of the atomic force microscope. In the simplest way, AFM performs scanning with a very small tip stick to the end of a flexible cantilever. The cantilever is controlled over a surface using a piezoelectric actuator. The cantilever deflects when it start to interact with sample molecules on the surface. Therefore, the tip-sample interaction can be monitored by optical means. The beam is reflected to a split photo detector which controls an optical lever which amplifies the deflected beam. This sensor is connected to a feedback circuit and it provides a control on tip-sample distance. This feedback signal is used for reconstruction of 3D sample topography and interactions which generally observed as an image in SEM and TEM. The relevant parts of an AFM device is in Figure 2.1.

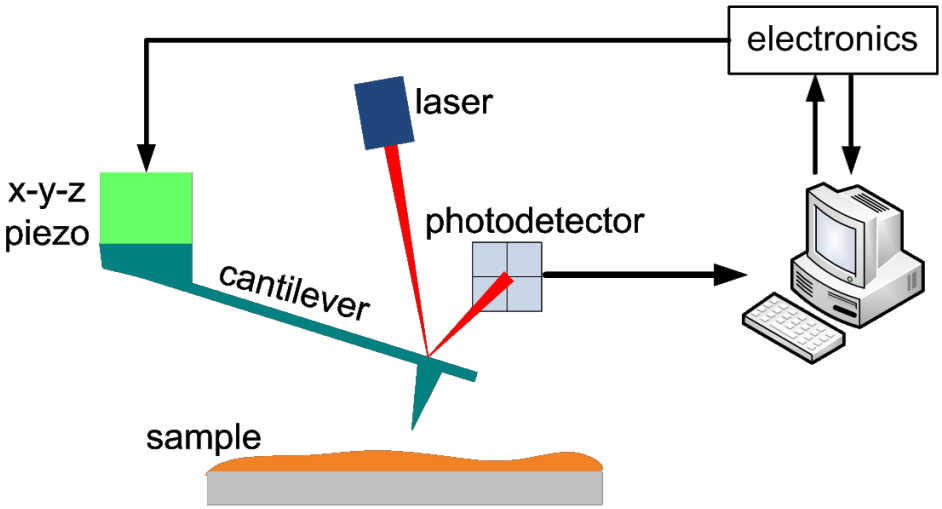


Figure 2.1. Sample afm device.

The common cantilevers are made of silicon materials or silicon nitride. There is a tiny pyramid called as tip at the end of cantilever. Yet, the selection of tips and cantilevers are dependent on the mode of AFM used. In static contact modes, the cantilever needs to have lowest stiffness value possible.

2.1.1.2. Force Spectroscopy. AFM provides measurements of forces in pico-Newton range under physiological conditions. The quantification of forces that are involved in cellular processes is a very interesting topic for many researchers. In single-molecule force spectroscopy, individual molecules are stretched and the force between surface and molecules are measured with respect to change in the loading rate. Generally, the elasticity of the molecules is larger than the binding forces hence the measurement is the binding force. These properties are generally used for identification of mechanical properties of molecules such as elastic modulus, binding landscape or elongation profile.

Various types of experiments such as pulling and rupture force measurements can be performed to determine the forces that involved in molecular recognition and binding. To this end, a functionalized cantilever with a protein and a sample consists of partner molecules are used for such experiments. The functionalized cantilever-protein or cantilever ligand complex by a linker is approached to surface of the sample which has receptors or other proteins that are attached to surface by a linker with a slow rate in micro scale and after surface contact occurred, the sample gently started to pull back surface with a known loading rate until the rupture event between protein and sample occurred. The deflection signal obtained by AFM is a way to determine the amount of force that is exerted for that rupture event. In this experiment, biotin functionalized cantilever is used. Avidin, a strong pair constructing molecule with biotin, is made a complex and approached to surface. The cantilever is slowly pulled off from surface and the deflection signal obtained by diode is transformed to force measurement. A sample force spectroscopy equipment is shown in Figure 2.2.

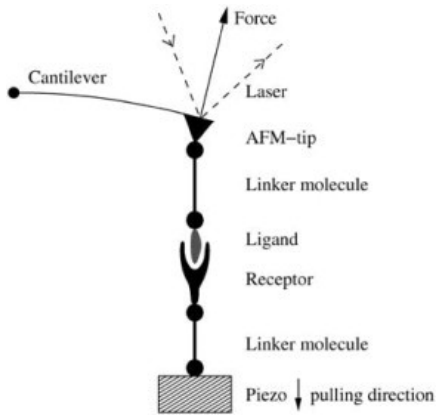


Figure 2.2. Sample force spectroscopy equipment [56].

2.1.2. Materials

2.1.2.1. Kinesin and Microtubules. Kinesin that is used in experiment is expressed for human and extracted from fruit fly (Drosophila Malingaster) and purified in the Center for International Research on Micro Mechatronics (CIRMM) at the Institute of Industrial Science, The University of Tokyo. The kinesin molecules are truncated after 410 amino acids and functionalized with biotin at the tail section.(kin410-BCCP). The motor domain of kinesin that is used in this experiment is in protein family of KIF5. Microtubules are used as sample molecules on surface that kinesin molecules interact with. For this experiment, due to high affinity with silica, microtubules are sticked to surface. Therefore, polymerized microtubules molecules are prepared and used. The microtubule solution has concentration of 20 μM. These molecules are obtained from the Center for International Research on Micro Mechatronics (CIRMM) at the Institute of Industrial Science, The University of Tokyo.

2.1.2.2. Reagent Molecules. Buffer solution that is used in experiment is also known as BRB80. BRB80 means 80mM PIPES buffer at 6.8 pH containing 1mM EGTA and 1mM $MGCl_2$. In the experiment, BRB80 solution is generally used for dilution. Taxol, GTP and $MgSO_4$ are reagents that are used for stabilization of polymerized microtubules. The concentrations of stabilizing agents are 4 mM for taxol, 0.1 M for $MgSO_4$ and 0.1 M for GTP. This solution is obtained from the Center for International Research on Micro Mechatronics (CIRMM) at the Institute of Industrial Science, The University of Tokyo.

Apart from these, streptavidin molecules are purified and obtained from Sanyal Group, Bogazici University.

2.1.3.Experimental Procedures

2.1.3.1. Functionalization of Cantilever Tip. It is possible to purify kinesin molecules with biotin addition on the outer region and for the functionalization of biotin coated cantilever tips streptavidin is used. Biotin coated cantilever and streptavidin mixed in a solution and due to high affinity between biotin and streptavidin, streptavidin molecules are bonded with biotins. The biotinylated kinesins are added to solution and after 30 minutes of incubation biotin coated cantilevers formed a complex of biotin coated cantilever, streptavidin and biotinylated kinesin.

2.1.3.2. Polymerization. For kinesin studies, monomer forms, α/β tubulin molecules, are polymerized. For this purpose, tubulin solution is concentrated to 1mM and GTP-$MgSO_4$ solution is added. The solution is incubated at 37 ℃ for 30 minutes. Finally, taxol solution is added until the taxol concentration level reduce around 20-40 μM after incubation.

2.1.3.3. AFM Experiment. In the experiment, cantilever functionalized with kinesin is approached slowly to surface covered with polymerized microtubule molecules and buffer solution. The cantilever is approached to surface with a velocity of 1, 3 and 10 μm/s. After contacting with surface the cantilever is moved back in same direction with a small velocity. The result of elongation is the rupture of bonds between microtubules and kinesin molecules. To determine the unbinding events, the deflection signal coming from cantilever to diode is used. Deflection data is processed and the force exerted on unbinding are determined. The experimental setup is displayed in Figure 2.3 below.

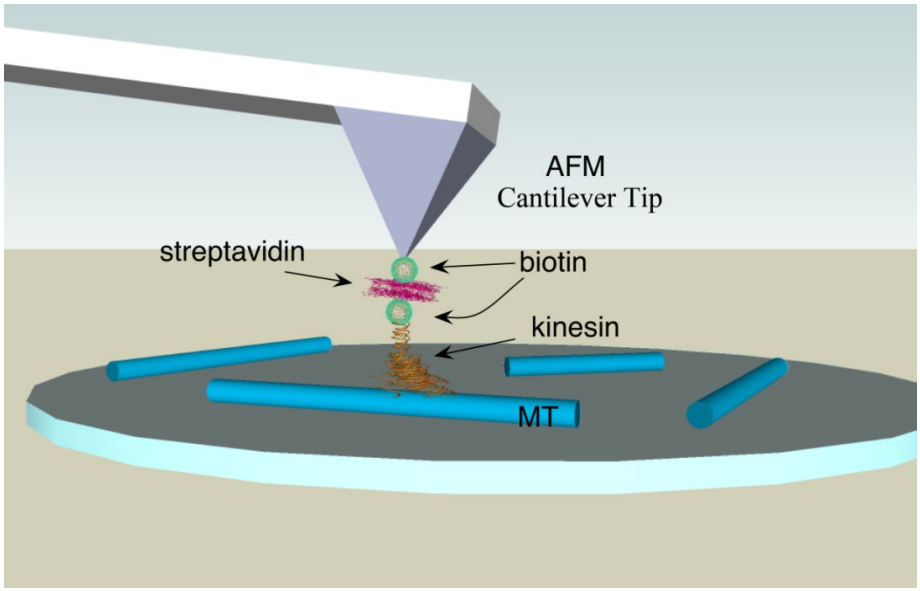


Figure 2.3. Experimental setup.

2.2. Computational Methods

2.2.1. Structures

2.2.1.1. Kinesin. Kinesin is a motor protein that is mainly operated as cargo transporters in cellular environment. In this study, two representatives of kinesin molecules, Drosophila Malingester Dimer Kinesin (PDB ID: 3KIN) and Human monomer kinesin molecule (PDB ID: 1BG2) [57] are investigated.

The X-ray crystal structure of Drosophila Malingester (3KIN) [58] at 3.1 A resolution is shown in Figure 2.4. It is a large dimeric protein with 710 residues. It has two head domains (1-331) and a neck linker (332-337) between two heads (orange). There is also a tail section at C terminus which is consists of two alpha-helices.. In Figure 2.4a, the functional regions for motility are colored: The tubulin binding sites (143-168,234-293)(red and magenta), the ATP binding site (88-93)(blue spheres), the neck (332-337)(orange). The chains A and B are colored in red with green, respectively.

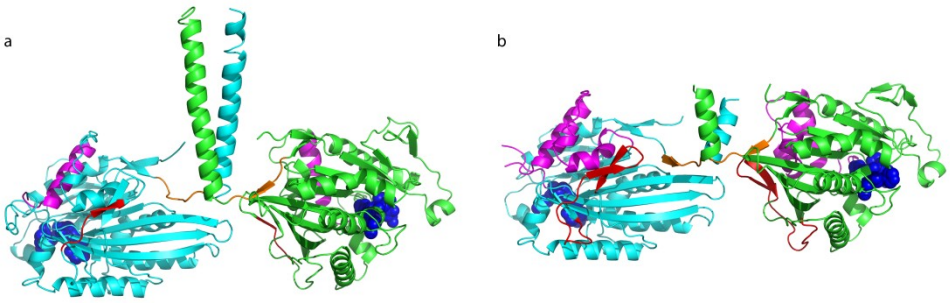


Figure 2.4. Structure of dimer kinesin molecules.

The other kinesin structure investigated in this work is a human kinesin (PDB ID: 1BG2); X-ray crystal structure at 1.9 A resolution. The structure has 340 residues. However, as the dimeric form is not available, the neck-linker and the other domain are modeled using the monomer and another human kinesin structure (PDB ID: 1MKJ). In Figure 2.4b, the functional regions for motion of kinesin structure are colored: The tubulin binding sites (red, magenta), the ATP binding site (blue), the neck linker (orange). Residues in each of these functional sites are as in Drosophila Malingester kinesin structure. Also, the chain A (green) and the chain C(cyan) are marked.

3.2.1.2. Methionine Repressor. Methionine repressor is a very common protein that is investigated for the long-range allosteric communications [59]. Methionine receptor is a transcription factor in E. coli and a controller that regulates genes for methionine biosynthesis. The activation-deactivation scheme of repressor protein is a result of allosteric communication upon binding to operator protein. The protein structures (PDB ID: 1CMA, 1CMB, 1CMC) [60, 61] are investigated here.

The structure of methionine repressor in various states. DNA bound (1CMA, 2.8 Å) [60] , ligand bound (1CMC, 1.8 Å) [61] and unbound (1CMB, 1.8 Å) [61] are shown in Figure 2.5. The structure is a dimer with 208 residues of each monomer. This structure is mainly used as a test case to determine the pathway of the force distributes from the know ligand binding site to the allosteric site namely DNA-binding site. The results of force distribution analysis by Brownian Dynamics simulations in the present work will be compared with those of MD simulations performed for force dissipation [24].

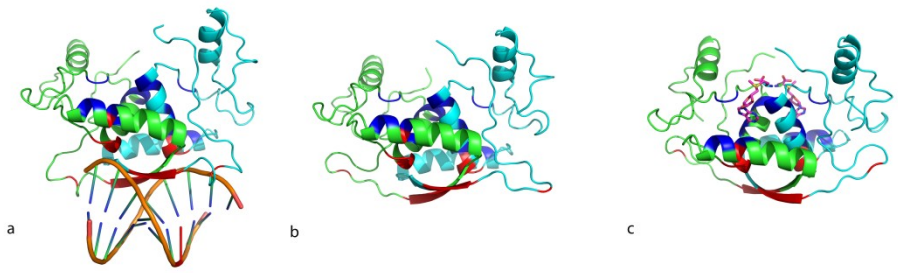


Figure 2.5. Structures of methionine repressor.

In Figure 2.5, three structures of methionine repressor are given. The DNA-bound, unbound and ligand bound structures of methionine repressor are given in a,b,c, respectively. In all figures, the chain A is colored with green and the chain B is colored with cyan. The ligand binding sites are marked as blue and the DNA binding sites are in red. The DNA-binding residues are K17, K22, K23, I24, T25, R40, T52, N53, S54 and the ligand binding residues are E39, R42, R42, L56, E59, H63, A64, F65, L70, P71. The ligand is S-Adenosylmethionine (SAM) molecule.

3.2.1.3. Purine Repressor. The purine repressor is one of the most common protein structures that are used to elucidate long range allosteric signaling phenomena. In this study, the monomer structure of purine repressor is studied. (PDB ID: 1WET) [62]. This transcription factor binds to repressor protein and affects the DNA binding. This structure is used as a test case for long-range allosteric signaling.

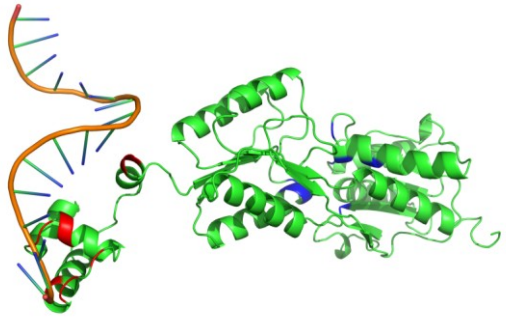


Figure 2.6. Structure of purine repressor.

In Figure 2.6, the monomer structure of purine repressor is visualized as green. The DNA-binding regions are colored as red and the ligand binding site is colored as blue. The

DNA- binding residues are V13, S14, T16, T17, L54 and K55 and, the ligand binding residues are A71, Y73, F74, R190, T192, R196, F221, D275.

2.2.2. Force Distribution Analysis

The motion of the protein is represented by Brownian motion. For each i^{th} interaction center, the Brownian motion is given by Langevin equation.

$$m\frac{d^2\mathbf{r}_i}{dt^2} = -\xi\frac{d\mathbf{r}_i}{dt} - \nabla_i V + mA_i(t) \tag{2.1}$$

where m is the mass of interaction center, **r** is the position vector, ξ is the friction coefficient, V is the potential energy function and A is the Gaussianly distributed stochastic force. This is a second order ordinary differential equation and in high friction limit this equation reduces to first order Langevin equation.

$$\beta\frac{d\mathbf{r}_i}{dt} = -m^{-1}\nabla_i V + A_i(t) \tag{2.2}$$

where β is defined as the ratio ξ/m. For the Gaussianly stochastic force the mean is zero and the covariance matrix is defined as

$$\left(A_i(t)A_j(t')\right) = \left(\frac{2\beta k_B T}{m}\right)\delta_{ij}u(t-t')\mathbf{I}_3 \tag{2.3}$$

where k_B is the Boltzmann constant, T is the absolute temperature and **I** is the identity matrix of order 3 and δ_{ij} is the Dirac delta function and u is unit step function.
The potential energy function used in simulations is GNM potential [63] defined as

$$V_{GNM} = \left[\frac{\gamma}{2}\sum_{i,j}^{N}\left(\Delta\mathbf{r}_j - \Delta\mathbf{r}_i\right)^2\right]u(r_{cut} - r_{ij}) \tag{2.4}$$

where γ is the bond constant, $\Delta \boldsymbol{r}_i$ is the change in the position vector of i^{th} interaction center, u is unit step function, r_{cut} is cutoff distance and r_{ij} is the distance between residues i and j. The force is calculated by taking derivative of this potential with respect to position and the bond constant γ is determined arbitrarily based on the fluctuation profile of the protein. Therefore, the relative values of gamma are more significant for simulations.

In simulations, the mass of interaction centers are taken as 0.14 kg/mol, bond constant is 0.5 kmol/A^2 , β is 1.0 x 10^5 / ns for 400 K. The Brownian dynamics simulations are performed by numerical integration of N differential equation of motions given by (2.2. The second order Runge-Kutta method is adopted for the integration of the differential equation. The details are described in Appendix A. The time step is selected as 10 fs. The process is discrete and Markovian, therefore, the trajectory is generated at the integer multipliers of the time step and the further step is generated from latter step of integration. The interaction centers are selected as the C^{α} atoms and the coordinates are obtained from Protein Data Bank files. The average simulation time for simulations is 10 min/ns.

Two analyses are performed based on the generated trajectories. The first one is the temperature factor analysis based on the mean-square fluctuation of the residues in structure. The beta-factors are directly proportional to mean square fluctuation values.

$$\beta_i = \frac{8\pi^2}{3}\left(\Delta \mathbf{r}_i \cdot \Delta \mathbf{r}_i\right) \tag{2.5}$$

Those obtained temperature factor values are normalized and compared with the experimentally determined ones.

Second analysis is the propagation of energy in the perturbed structure. For this purpose, a force is applied to the system from the ligand binding site and the potential energy contribution of each interaction is calculated as

$$E_i = \sum_j^N E_{ij} \tag{2.6}$$

The residue energy profile is related to the distribution of force on the structure. The results of the ligated and unligated simulations are compared and the distribution of force is suggested.

2.2.3. Monte Carlo Pathway Generation

3.2.3.1. Interresidue Interactions. The generation of pathways and calculation of network parameters such as closeness and betweenness are based on the Interresidue interactions. The interactions between amino acids are determined in atomistic scale and the consolidation of attraction and repulsion are defined by Lennard-Jones 12-6 potential.

$$E(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{2.7}$$

In the formula above, the σ and ε are corresponded to the collusion diameter between two atoms and the minimum energy that is viable between two atoms, respectively. The term r is the distance between two atoms. The σ and ε are specific for residue and atom type and obtained experimentally. To simplify calculations, the energy value for distances less than defined minimum radius are assumed to be on the same energy with the minimum radius. The minimum radius is calculated with following equation:

$$r_{\min} = \sqrt[6]{2}\sigma \tag{2.8}$$

Due to the presence of this simplification in energy function, the repulsion term in Lennard-Jones potential is disregarded. Another simplification is applied to energy function for distances greater than a certain cut-off value. It is assumed that the energies of residues that have a distance greater than the cutoff value are equal to zero. For a protein structure with M atoms and N residues, the interatomic energies are calculated and stored in a matrix with size MxM. The modified Lennard-Jones potential is plotted in Figure 2.7.

After generation of interatomic energy matrix, all energy values are added up to obtain a single value for each residue. Therefore, an energy matrix with NxN dimension is

obtained. Each element on this matrix denotes an interaction energy between i^{th} and j^{th} residues.

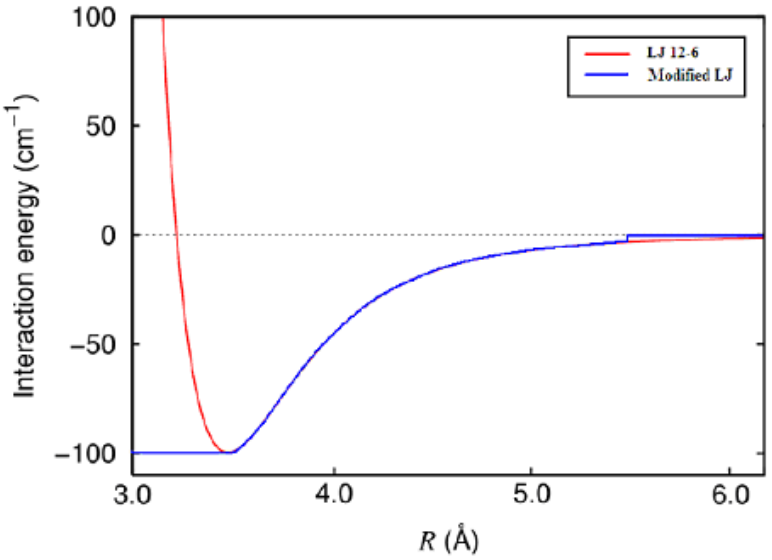


Figure 2.7. Comparison of Lennard-Jones and modified Lennard-Jones potential.

3.2.3.2. Path Generation. From inter-residue energy matrix, by using Boltzmann weight, a statistical weight matrix is obtained with following equation:

$$W_{ij} = \exp\left(-E_{ij} \Big/ kT\right) \tag{2.9}$$

After normalization of the rows of statistical weight matrix, the square probability matrix with dimension NxN is obtained. However, a second normalization is performed to assign zero value to minimum probable ones. Then, the normalization is conducted again and the final version of the probability matrix is obtained with following equation:

$$P_{ij} = \frac{W_{ij}}{\sum_{j=1}^{N} W_{ij}} \tag{2.10}$$

The path generation is performed based on the probability matrix. For generation of pathway from residue i, a uniformly distributed random number between 0 and 1 is generated. By adding each non-zero element, a group of ranges are generated on i^{th} row.

Depending on the range that the random number stands, the path is generated through that residue. After each step, the probability matrix is regenerated by nullifying the columns of the visited residues to prevent pathways to revisit previous residues.

This path generation method is implemented in three different schemes. The first scheme is to generate an ensemble of pathways between initial and final residues. The second scheme is to generate an ensemble of pathways with known initial residue and step length. The final scheme is to generate infinitely long pathways to build up a network of pathways and calculate the centrality parameters for pathway network.

3.2.3.3. Analysis of Pathways. In the first two schemes of implementation of path generation, the frequency of a pathway in the ensemble of pathways is the only key parameter. The pathways are ranked in the order of frequency of a pathway in the ensemble and named as populated pathways. However, it is very hard to find a repeating pathway for pathways with large number of elements. To this end, the probabilities of selected steps are added and the pathways are ranked with their total probability values in case there is not any populated pathway. These pathways are named as probable pathways. Apart from these, Feng-Doolittle algorithm, a multiple alignment algorithm, is consolidated with a single linkage clustering and pathways are clustered and the best representatives of three clusters are selected as most probable or popular pathways. The detail of this algorithm is described in Appendix B.

For the final scheme, an infinitely long path is generated at least in the order of 10^5. These pathways are starting from a random residue and generated until number of steps are reached by passing through residues more than once. A network consists of those pathways are generated and graph centrality calculations are performed. The closeness and betweenness centrality values are used for identification of functional residues.

The closeness of residue i, O_i, is the inverse of the average of the all shortest paths from residue i to all other residues and defined as follows:

$$O_i = \frac{N-1}{\sum_{j=1}^{N} l_{ij}} \tag{2.11}$$

where N is number of residues and l_{ij} is the shortest path between residue i and j.

The betweenness of residue k, b_k, is the ratio of times that the node k stands on any shortest path pair in all network to number of all shortest path pairs and defined as follows:

$$b_k = \sum_{ij} \frac{g_{ikj}}{g_{ij}} \tag{2.12}$$

where g_{ikj} is a binary value and 1 if the node k is on shortest path between i and j, 0 otherwise. g_{ij} is another binary value and 1 if the i and j are not equal, 0 otherwise.

3. RESULTS AND DISCUSSION

3.1. Monte Carlo Path Generation

3.1.1. Analysis of Pathways

The dimer kinesin structure is analyzed with graph centrality parameters and the various pathways generated between different functional sites. The motility of kinesin structure is provided by the ATP binding. Hence, the interaction between tubulin molecules is directly related with the ATP binding. Therefore, there has to be a signaling pathway between ATP binding site and tubulin binding sites. There are two known tubulin binding sites (144-165, 234-293) and one ATP binding site (88-93) and pathways are generated between the ATP binding site and these two tubulin binding site which belongs to same chain for human dimer kinesin structure (PDB ID: 1MKJ) [64]. Also, the analyses are performed for Drosophila Malingester kinesin structure yet not included in the thesis and available at MCPath server.

For human kinesin structure (PDB ID: 1MKJ), three most probable pathways are identified based on sequence clustering of top one hundred most probable pathways and the best members of these three clusters are plotted in Figure 3.1 and Figure 3.2 and tabulated in Table 3.1 and Table 3.2 .

Table 3.1. Best members of the probable pathway clusters between ATP binding site and first tubulin binding site.

Cluster 1	90A	91A	231A	206A	190A	148A	150A
Cluster 2	90A	92A	96A	134A	146A	147A	150A
Cluster 3	90A	92A	231A	206A	145A	189A	148A
	149A	147A	150A				

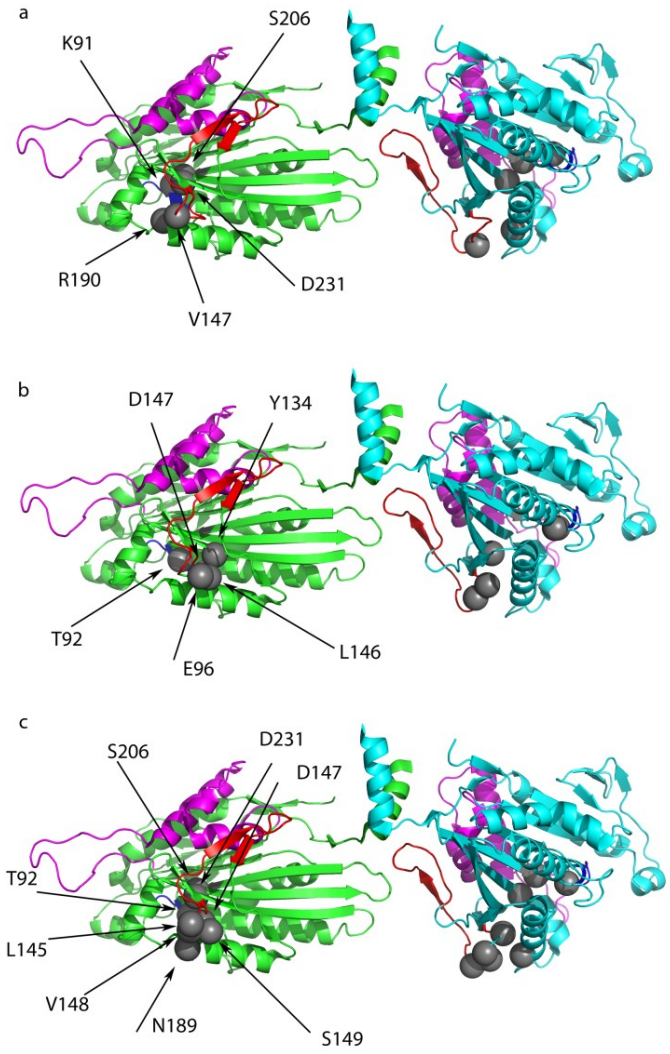


Figure 3.1. Populated pathways between ATP binding site (90) and first tubulin binding site (150).

The tubulin binding sites are marked with red (144-165) and magenta (234-293) and, the ATP binding site (88-93) is colored as blue. The residues belong to the most, second most and third most populated clusters are presented in cartoon in Figure 3.1a,b,c; respectively, in gray spheres. The ensemble of most probable a hundred pathways are clustered into three groups and these residues are generally in the same region of structure (S206) since it is a central position on structure that lies between those two sites. The clustering method works fine because the best members obtained always lies in the top ten pathway means that the other probable pathways are rooted based on the top ten pathways. The ATP binding site lies in the core of protein structure and the tubulin binding site periphery is close to this site in terms of most populated pathways. From these pathways,

the residue S206 is found to be a key in allosteric signaling network which is a residue that does not lie in tubulin binding region unlike all the other residues in populated pathways. Also, the conservation score of this residue is 9 that is a highly conserved residue which may have a functional importance [65].

Table 3.2. Best members of the probable pathway clusters between ATP binding site and second tubulin binding site.

Cluster 1	90A	91A	232A	282A	280A	
Cluster 2	90A	87A	232A	234A	282A	280A
Cluster 3	90A	91A	230A	232A	282A	

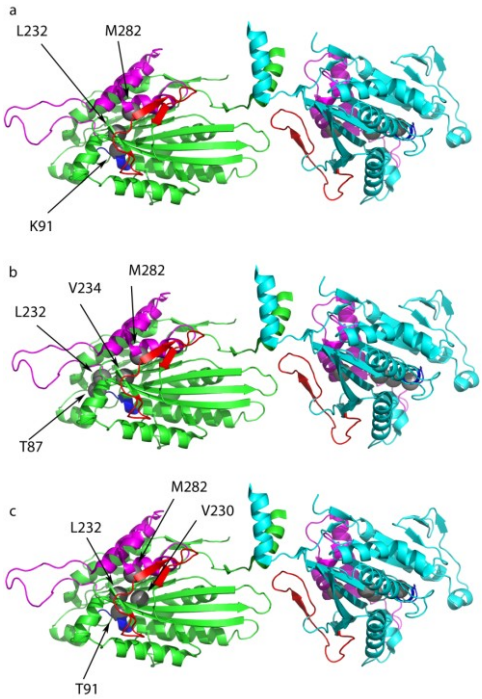


Figure 3.2. Populated pathways between ATP binding site and second tubulin binding site(280).

In Figure 3.2, the best members of the clusters of the most probable hundred pathways between ATP binding site (90) and second tubulin binding site (280), which are selected based on the central position in sites, are visualized in a,b,c with descending probability. The color code is as in Figure 3.3. This tubulin binding site covers a large area.

It starts from G234 and ends at N293. It is seen that in the best members of the most probable clusters, the residue L232 is on the allosteric signaling pathway. The distance between the signaling starting point (90) and the end point (280) is 24 Å however, the signaling pathway is passing through residue L232 and reach the second tubulin binding site. Also, the conservation score of this residue is 9 [65] and a catalytic residue lies at S235 [66]. Hence, the catalytic residue affects the signaling when the ATP binds and the structure display conformational changes.

The motility of the kinesin protein is dependent on the motion of the neck linker between two head parts of dimeric kinesin structure. Therefore, pathways are generated between the two tubulin binding sites (143-168, 234-293) and the neck linker region (332-337) [64]. The best members of the most popular two clusters are shown below. The third cluster is very similar to the first cluster obtained from agglomerative clustering. Therefore, only two clusters are visualized.

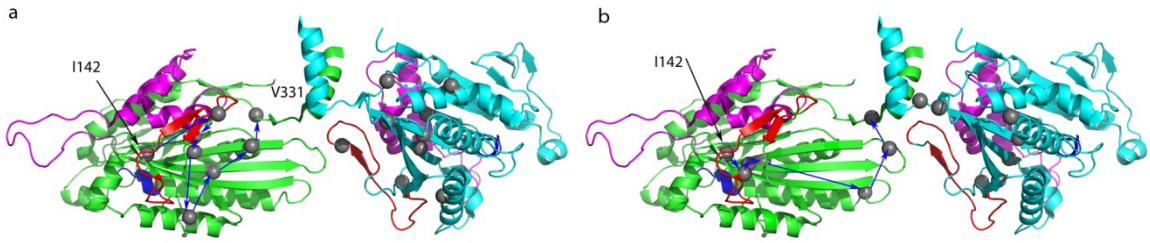


Figure 3.3. Pathways between neck linker region (335) and first tubulin binding site (150).

In Figure 3.3, the human kinesin structure and the populated pathways between neck-linker and first tubulin binding site are visualized. The color code is the same as Figure 3.2. The arrows guide the pathway that allosteric signal propagates. The simulations are performed as taking the tubulin binding site as starting point and the neck linker region as end point. It is seen that the signal leaves from tubulin binding site and propagates through opposite peripheral region of the protein via two parallel pathway. I142 is the common residue in the most popular pathways which has a moderate conservation score. The pre-existence of the parallel pathways is a new notion in explanation of allosteric phenomena [5]. The old view still states that the signaling passes through single path and definitely result in conformational changes.

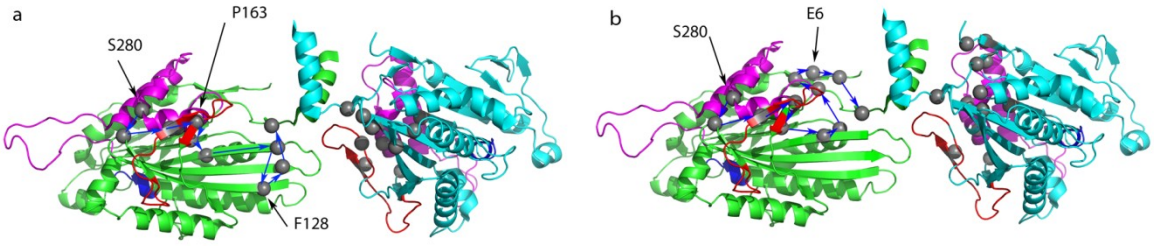


Figure 3.4. Pathways between neck linker region (335) and second tubulin binding site (280).

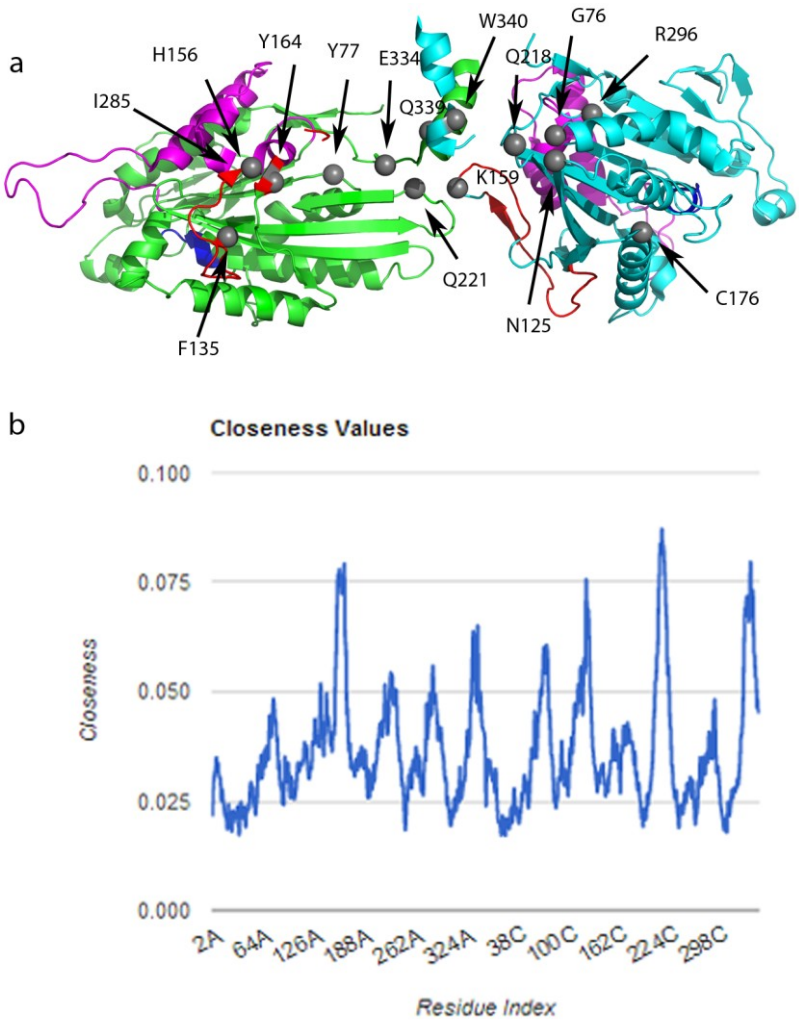
In Figure 3.4, the best members of the most probable pathways between the neck linker region and the second tubulin binding site are visualized on kinesin structure. The color code is the same as in the previous figures. Interestingly, both of the most popular pathways are passing through the first tubulin binding site unlike the pathway generated for first tubulin binding site. It has been known that in the binding and unbinding events of kinesin molecule, both binding sites work cooperatively and the allosteric mechanism induces the events with the signal coming from the motion of neck-linker [67]. Thus, this common route is a sign of cooperativity between two binding sites. Also, the allosteric signal coming from first tubulin binding site may propagate through two distinct paths one of which is on the N-terminal region (1-10) of the protein and the other one passes through the anti-parallel β-sheets in the core of protein structure (210-230).

3.1.2. Graph Centrality Analysis

The centrality parameters of graphs produced from the kinesin structure is also analyzed for prediction of functional sites. The closeness and betweenness centrality values of each residue are calculated and regions that might play a key role in cellular processes or allosteric signaling are determined.

In Figure 3.5, the closeness values (b) and the local maxima of closeness values are shown (a) and tabulated. The functional regions are tubulin binding sites (red and magenta) and ATP binding site (blue) are represented as cartoon. The closeness value of a specific residue is the inverse of the average of the length of shortest paths between a specific residue and all the other residues and calculated from Equation 2.11. In general, the

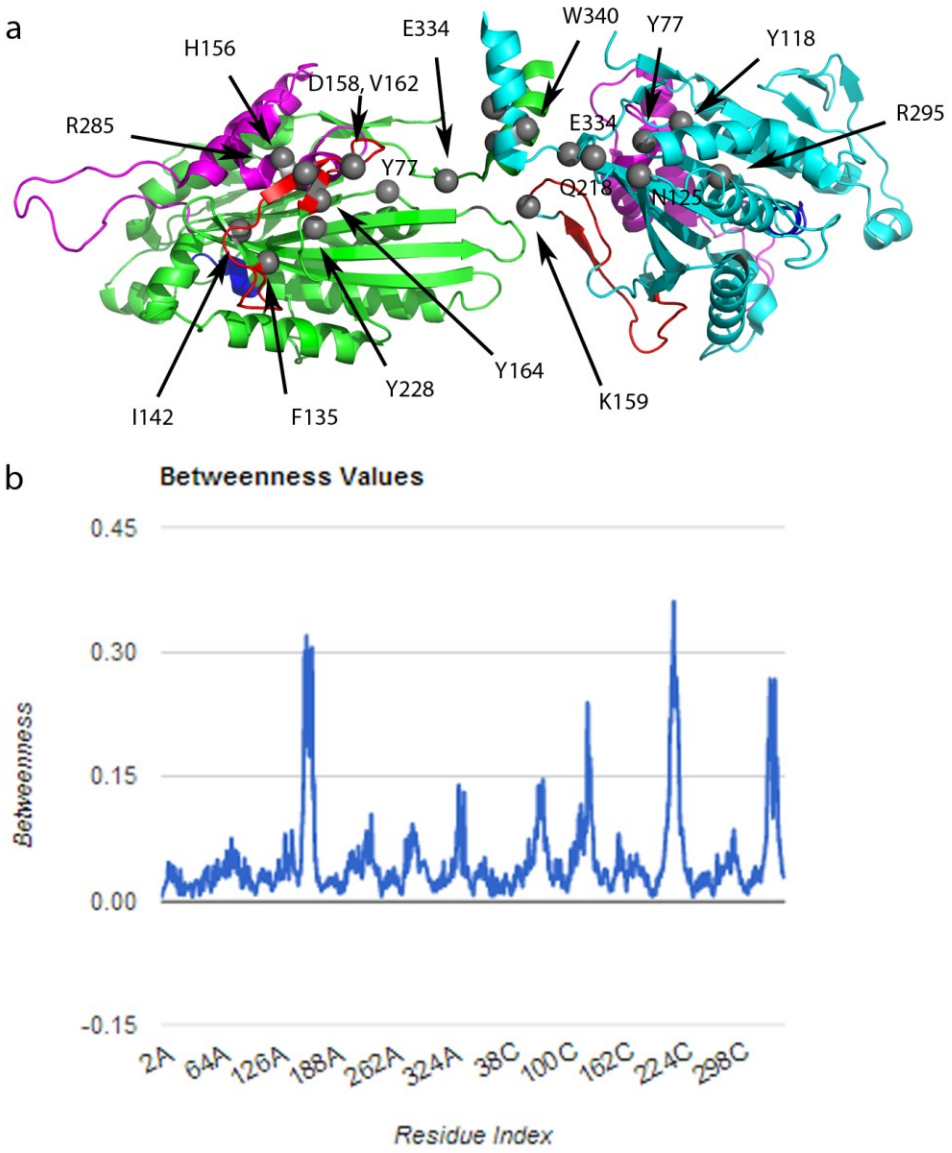
residues with high closeness values are around neck-linker region since there is a strong signal transmission between two heads for the motility of kinesin structure that is provided by the cooperative motion of the heads of kinesin molecule. Apart from those residues in neck linker region, there are also some other functional residues that are scattered around the tubulin binding sites (164, 285) and ATP binding site (77).



77A	135A	164A	221A	285A	334A	340A
76C	125C	159C	176C	218C	295C	339C

Figure 3.5. Closeness values and representation of local maxima values.

In Figure 3.6, the betweenness values (b) and the cartoon representation of local maxima of betweenness values (a) are visualized. The functional elements are marked on the structure such as ATP binding site (blue) and tubulin binding site (red and magenta).



77A	135A	142A	156A	158A	162A	164A	228A
340A	77C	118C	125C	159C	218C	295C	334C

Figure 3.6. Betweenness values and representation of local maxima values.

The betweenness value is a reflection of the likelihood of a residue being in the each shortest path pairs. Due to the characteristics of protein, the local maxima values of betweenness plot are generally around the neck-linker area. All the shortest paths between residues on different heads have to pass through neck linker region, therefore these values

are greater than the other residues. However, there are also some other residues that are involved in strong signal propagation in other words allosteric signaling which are not in the neck-linker region. Those residues are close to tubulin binding sites (142,156,158,162,164) and ATP binding site (77).

To identify a different site that plays a key role in allosteric signaling, the local maxima values of the closeness values are used. It is suggested that C176 is a plausible allosteric signaling site even if it is on outer surface of the protein structure. This residue differs from other residues in terms of conservation score, high closeness value, close distance to hinge residues. This residue could be a plausible site that could be studied by the AFM experiments. It has ten times larger solvent accessible surface area compared to other cystine residues on the structure. The presence of sulfhydryl group in cystine residues allow the functionalization of the kinesin with maleimide group from this region and provide an alternative way to determine some mechanical properties other than the functionalization of coiled coil region with biotin-streptavidin functionalization.

3.2. Force Distribution Analysis

3.2.1. Methionine Repressor Results

Methionine repressor protein is a transcription factor that plays a key role in cellular process transcription. Main function of this protein is the binding of S-Adenosylmethionine (SAM) and following an increase in the affinity of DNA binding. For these two processes inducing each other, there should be allosteric signal propagation between the ligand and DNA binding sites.

In order to investigate the fidelity of the Brownian dynamics (BD) simulations without the application of force, the BD simulations is applied to methionine repressor structure (PDB ID: 1CMA) [60]. The normalized version of the experimental temperature factors and the normalized temperature factors that are obtained from BD simulations for 4 ns are plotted in Figure 3.7.

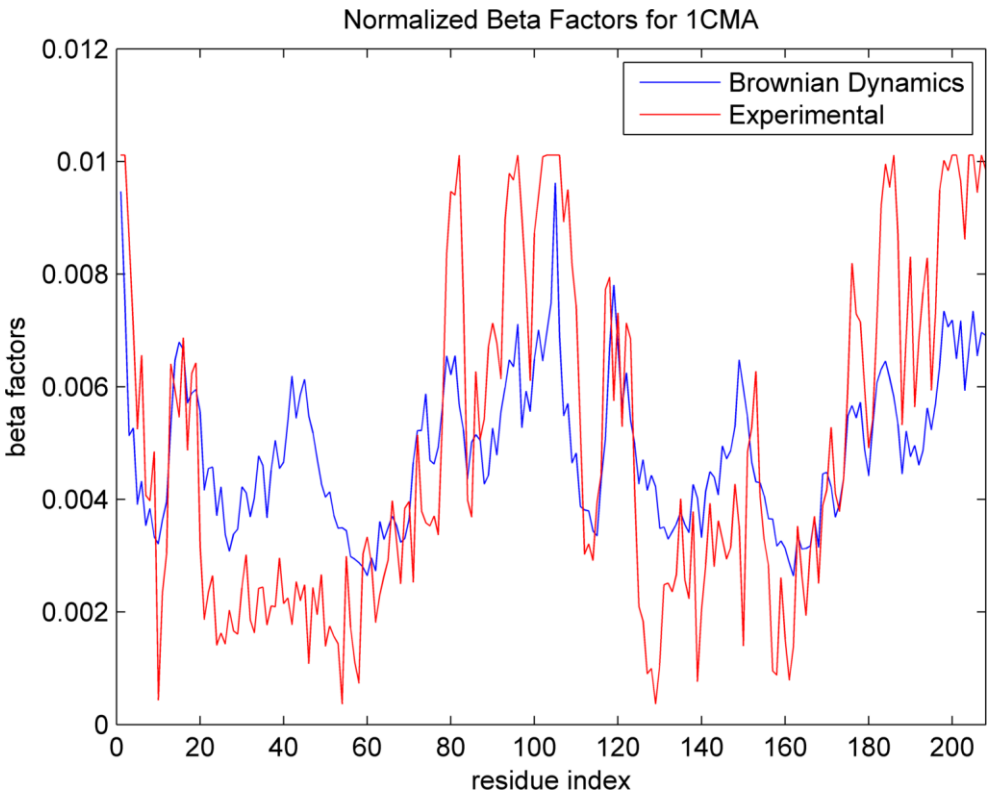


Figure 3.7. Temperature factors for methionine repressor.

The temperature factors that are obtained from the BD simulations without the application of ligand force are similar to the experimental results. For further analysis, the linear correlation coefficient calculated between experimental and simulation results are 0.77. This difference may result in due to the presence of ligand in the real structure. The difference between experimental and simulation results are scattered around the ligand binding sites (39,42-43,56,59,63-65,70-71) and the DNA binding sites (17,22-25, 52-54). This result implies that a fast algorithm with a simple potential provides realistic results.

Following, the ligand force applied to the system and the distribution of the force is investigated with the BD simulations. The residues close to the ligand center are affected by the force. The amount of the force is scaled with the variance of the random force and inversely proportional to the distance to the ligand center. The direction of the force is from center of ligand to the ligand binding residue. The difference between the energy of the unligated and ligated simulations could be considered as an indicator of the force

propagation In Figure 3.8, the energy difference between with and without ligand simulations are plotted.

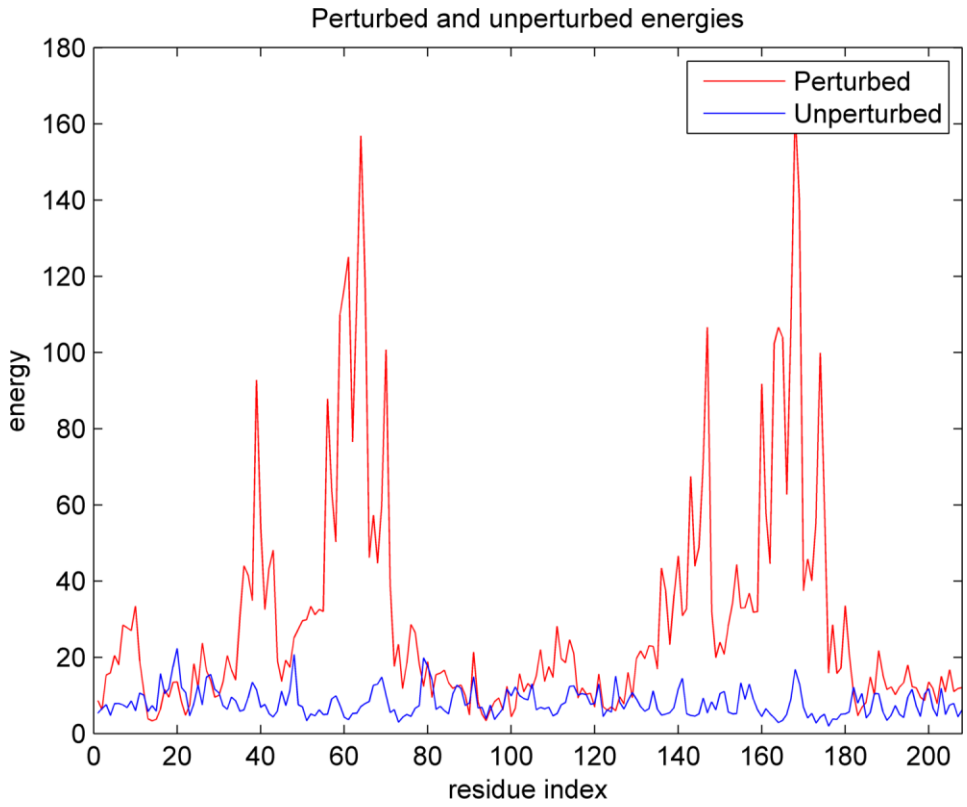


Figure 3.8. The difference in the energy for perturbed and unperturbed methionine repressor.

The force applied from residue scattered around two ligand binding sites (38,41,42,55,58,59,66,69,70) on both chains. The energy is generally transferred to regions with functional importance. The Arg40 is known as the residue that plays a key role in the propagation of the energy. The region lies between residues 52 and 54 is a DNA-binding region and it has an energy gain in the application of the ligand force even though it is lower than the ligand binding residues need to be decreased for ligated state. However, the energy increase in a residue is an implication of the force dissipation through that point. Also, there is not any force application on the N-terminal region of the protein (5-15) yet, there is an energy increase in those parts of protein which are at a moderate distance from the ligand binding site. Hence, the force is propagated through these regions via the changes in the dynamics of the ligand binding regions. The distribution of the force on the

structure is displayed on the Figure 3.9. The peaks on the energy values are marked with wheat spheres. Those residues are on the path from ligand binding site to the DNA-binding site.

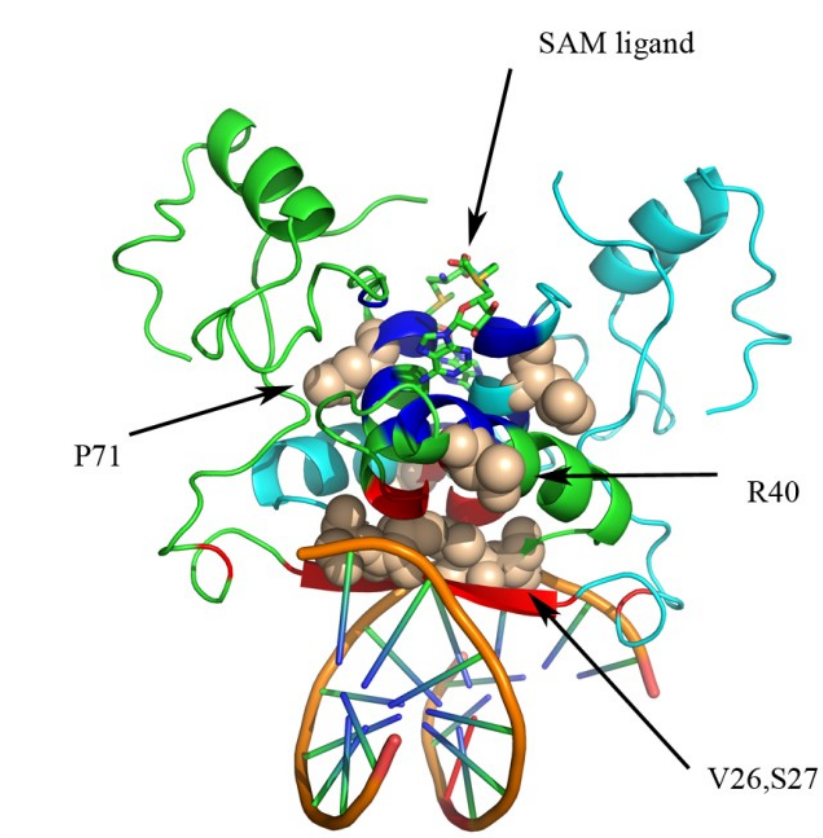


Figure 3.9. The force propagation on the structure of methionine repressor.

There is another force distribution software is available as a patch in molecular dynamics program Gromacs [68]. The results obtained from force distribution algorithm for methionine repressor are very similar with the results from Gromacs.

The change in force with respect to time is displayed in A. It is seen that there is a strict increase in the energy of beta sheet near the DNA-binding region (27-29) in the first cycles of simulations yet, the energy drop downs in the equilibrated structure. This is an indication of fast force propagation through this region. The energy values are higher compared to unligated simulations. The reason is simple that the perturbation applied to system generally increases the motility of the protein. Therefore, the energy values are

higher unless there is an energy barrier or force exertion due to the functional events such as DNA-binding.

3.2.2. Purine Repressor Results

Purine repressor is a very well-known protein in the lac repressor family. It is a transcription factor that involves in coding of genes in the synthesis of proteins that are significantly useful in lactose metabolism. The main function of the purine repressor is the binding of guanine (GUN) and the induction of DNA-binding via allosteric signaling.

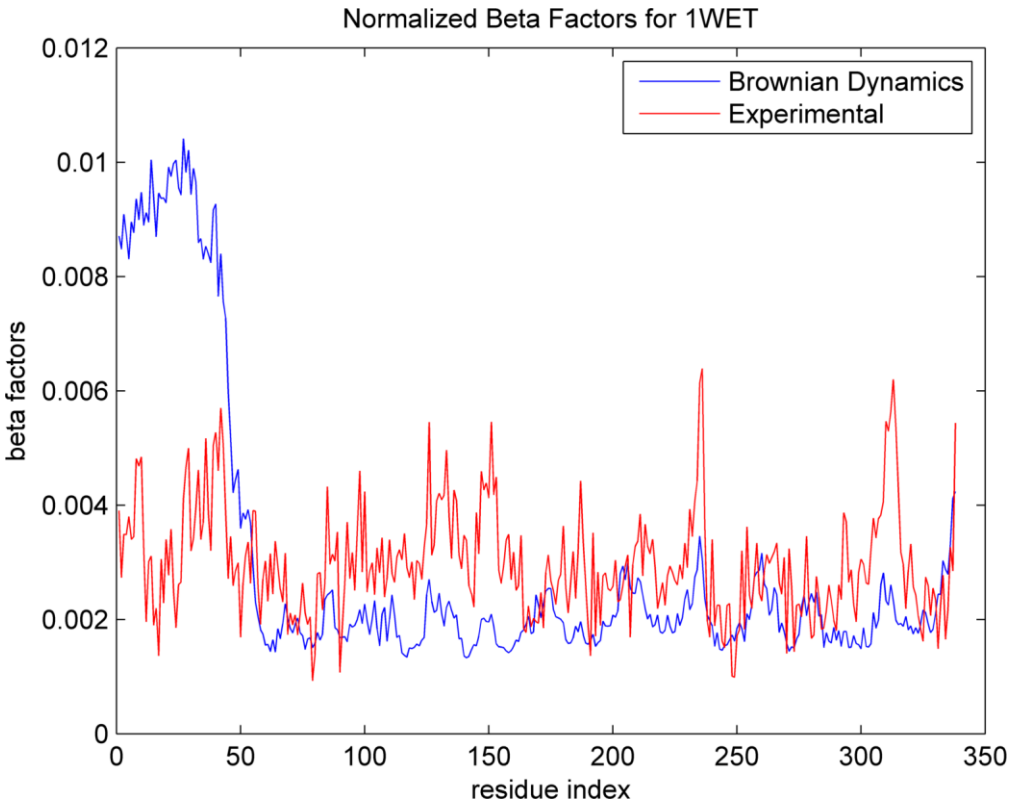


Figure 3.10. Temperature factors for purine repressor.

Figure 3.10 shows the normalized residue mean-square fluctuations by the BD simulations and by the x-ray crystallographic temperature factors. The temperature factors are similar for the regulatory domain (56-338), however due to the absence of DNA in simulations, the DNA-binding region has high fluctuation values compared to the

experimental ones. The addition of the DNA-molecule into the simulations may result in better temperature factors.

The distribution of forces applied from the binding site passes through the linker region to the DNA-binding region. The ligand force is applied as in the case of methionine repressor to ligand binding sites (65-66, 69-72, 122, 184-185, 188-191, 218-219, 229, 244-246, 248, 272-273). In Figure 3.11, the change in the energy of each residue with the application of ligand force after 30 ns.

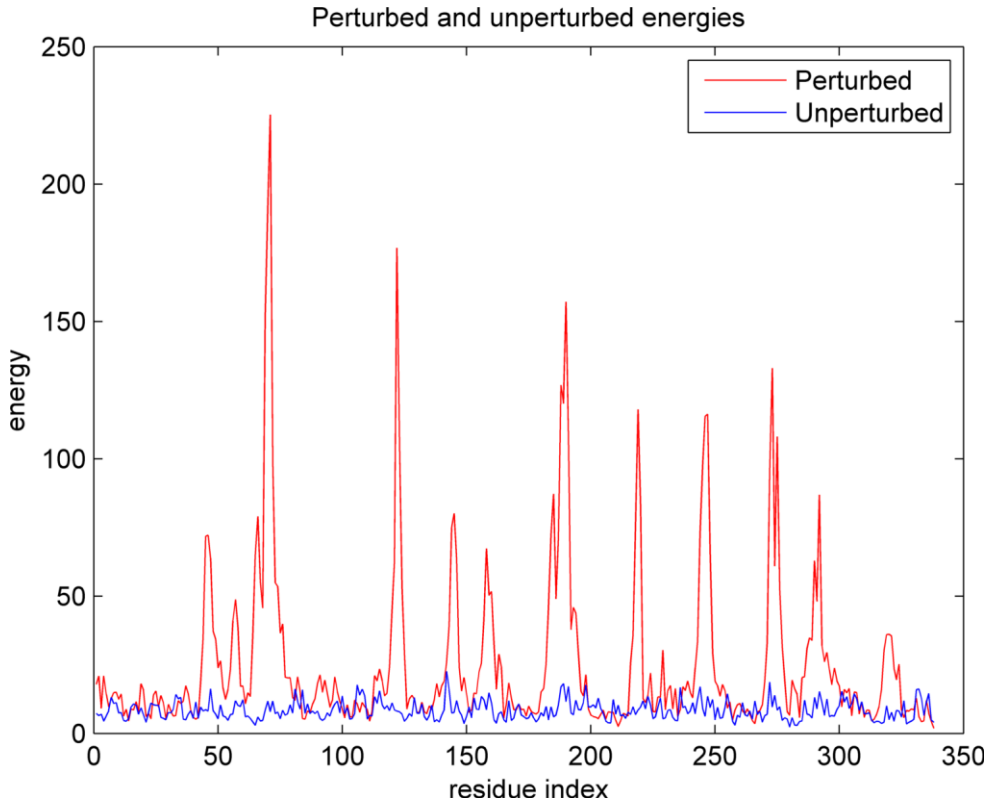


Figure 3.11. The difference in the energy for perturbed and unperturbed methionine repressor.

The linker region between DNA binding domain (1-56) and regulatory domain (60-338) is mobile/flexible and has a high energy (57-59). T59 is on the linker region and has a high energy. Also, S48 is close to DNA-binding regions and may play a key role on the allosteric signaling. The residue A71 is the ligand binding site. Due to the external force on this region, it has a high energy value. The residues between residues with following id numbers: 118-121, 142-144 and 156-159 are the internal β-sheets that are involved in

signal propagation between linker region and the ligand binding sites. All the other high energy values are in the ligand binding domain and mostly affected by the close distance to the ligand binding region unlike the linker region. The regions that have high energy gain are shown in Figure 3.12.

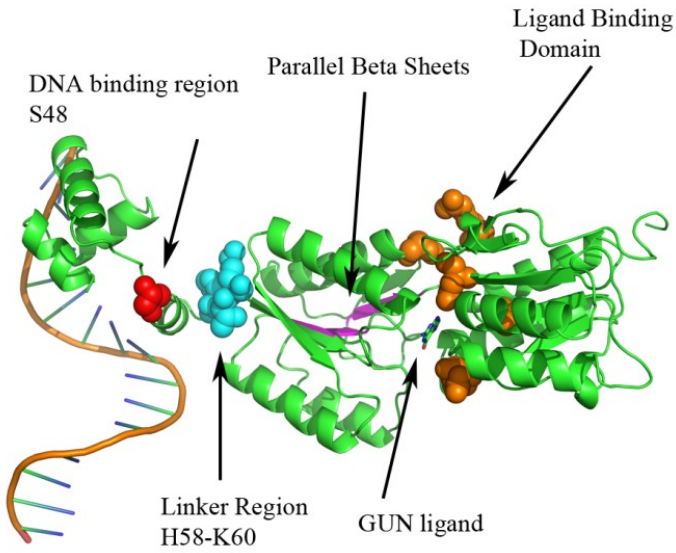


Figure 3.12. Energetically affected regions on purine repressor.

In Figure 3.12, the DNA binding site (red), linker region (cyan), parallel β-sheets (magenta) and ligand binding region (orange) are colored. It is seen that applied force from the ligand binding site only propagates in the direction to DNA binding domain. The consolidation of these ligand binding regions, internal parallel β-sheets and linker region construct the pathway between ligand and the DNA. Therefore, the induction of transcription event is triggered with the binding of ligand to the ligand binding region.

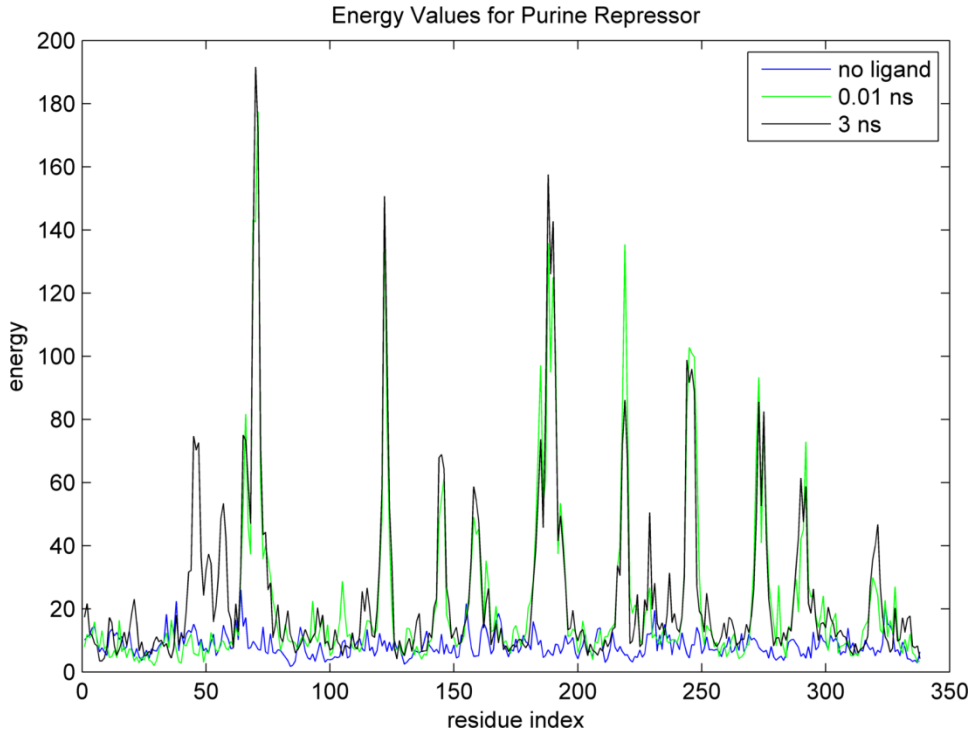


Figure 3.13. Purine repressor force change with respect to time.

In Figure 3.13, the energy values for each residue at different times of simulation are plotted. It is seen that the ligand binding regions directly start with high energy compared to other residues. However, after a certain time, there are some residues, especially in the linker region and DNA-binding region, start to gain some energy after internal strain propagation to these sites.

3.2.2. Kinesin Results

Kinesin protein is one of the motor proteins which has a lot of functionalities in cellular transport. This protein has a walking motion on the microtubules via different conformation changes. This motion is induced by the binding of ATP and allosteric signaling to the neck linker and tubulin binding sites. Also, the AFM experiment is mimicked with the force distribution analysis method.

Firstly, the unperturbed kinesin simulations are performed and the normalized temperature factors are compared with the experimental results in Figure 3.14. The values

are generally correlated and the calculated linear correlation coefficient is 0.78. Only some regions that are affected by the ATP binding or tubulin binding are mostly differing from the experimental results.

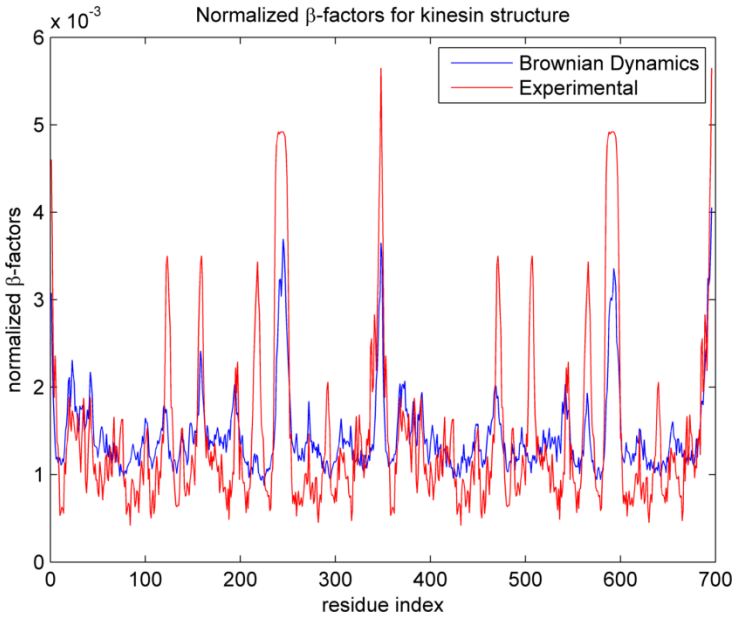


Figure 3.14. Temperature factors for kinesin molecule.

An AFM experiment is performed by applying force from the coiled coil region (A338-A345). Therefore, the same effect is simulated and the distribution of the force is investigated.

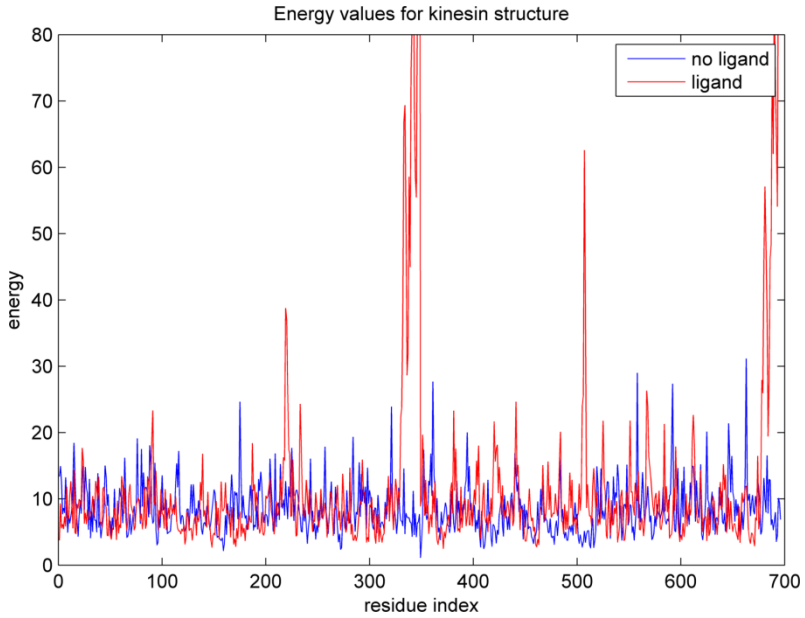


Figure 3.15. The difference in the energy for perturbed and unperturbed kinesin molecules.

The Figure 3.15 shows the distribution of the force applied from coiled-coil region to both monomers on the kinesin structure. It is seen that the force applied from the coiled-coil region (A338-A345) affects the functional sites in general. Two of the peaks show the ATP binding site (A91, A436). The ATP binding site is one of the significant sites for the motility of the kinesin structure. The peaks at the core of protein (A224,A232) are in the agreement with the predictions by the Monte Carlo path generation. Also, the suggested residue for experiment, Cys 176, has an interesting response to the ligand binding. This residue decreased its energy upon ligand binding which may be another ligand binding region or force application point for the structure. Also, the sites are displayed in cartoon representation in Figure 3.16 . The color code is same as in previous kinesin figures. Also, the orange residues are the allosteric residues that are evaluated with both methodologies and the gray residues are the proposed allosteric site and also a tubulin binding site (C162) is evaluated on the other monomer.

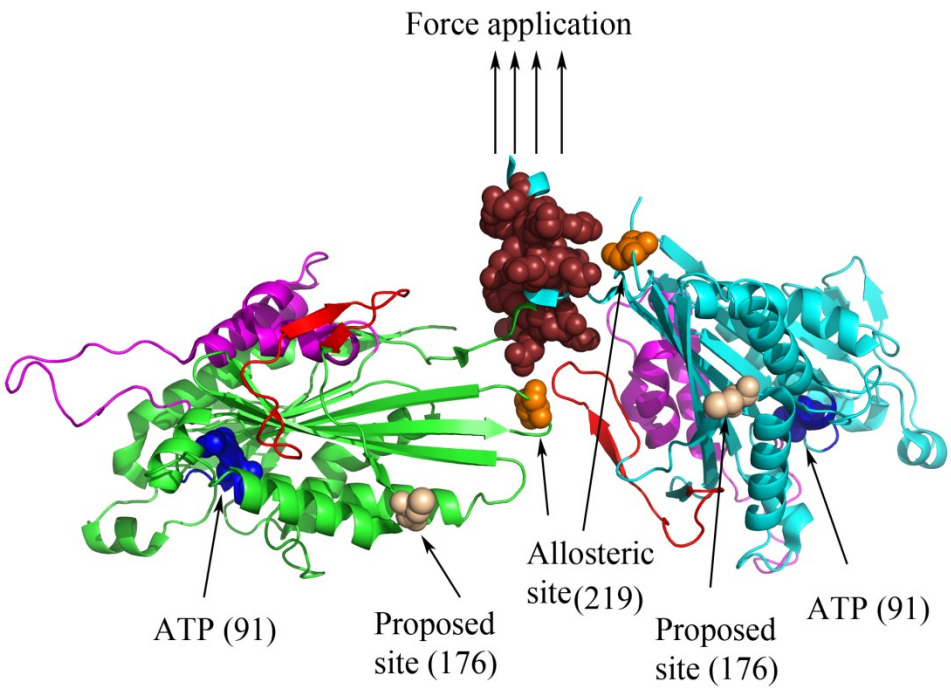


Figure 3.16. Structural representation of force distribution for kinesin.

The time dependency of the force is also investigated. The energy distribution at different points of simulations is plotted in Figure 3.17. The initial impact of the perturbation affects mostly ligand binding sites and interestingly tubulin binding sites (A144,C162). However, the perturbation stabilizes in the end of a certain period. The proposed region lost energy rapidly and after a certain while the energy lost decays. Yet, the ligand binding sites and affected regions such as neck linker have still energy gain.

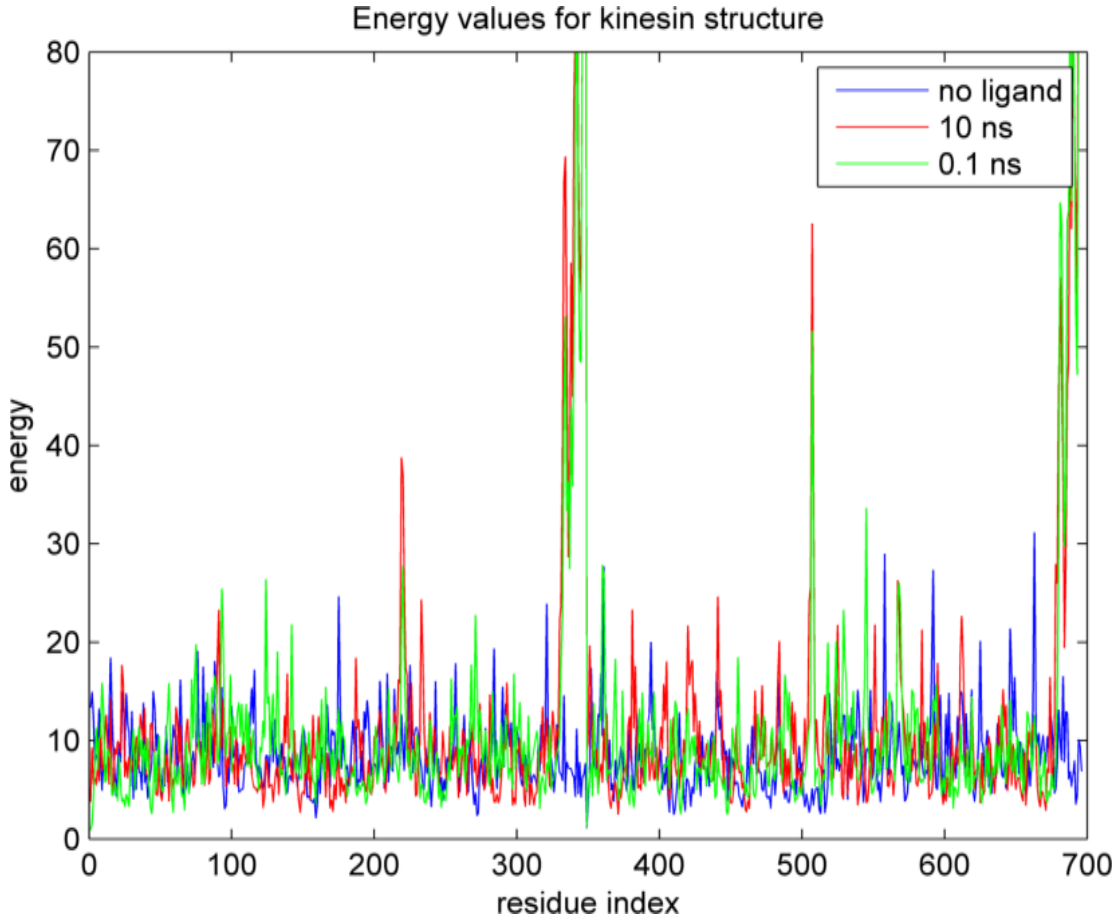


Figure 3.17. Kinesin molecule force change with respect to time.

To understand the force propagation through ATP binding, an external force is applied from the ATP binding sites on one of the monomers since in reality the two ATP molecules bind to kinesin at distinct times (A91). The results are shown in Figure 3.18. The ATP binding effects firstly the ATP binding site and the sites around the tubulin binding sites (A230,A232). Also, the energy of tubulin binding site residues are generally close to the energy of the unperturbed structure except the tubulin binding site on the second monomer. Therefore, the signaling is propagated through distribution of forces until reaching the allosteric site. The distribution of the forces on the structure is visualized in Figure 3.19. The figure is not symmetric since the force is applied from only one of the ATP binding sites.

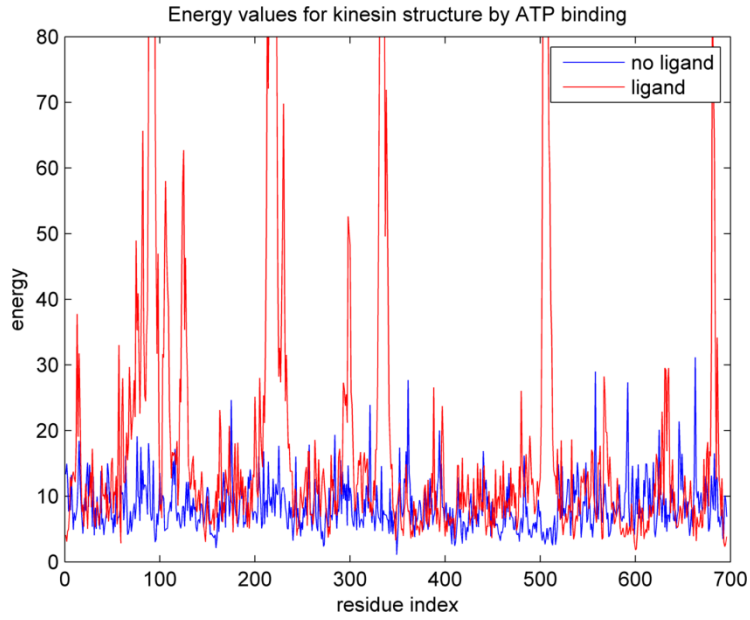


Figure 3.18. The difference in the energy for perturbed and unperturbed kinesin molecules via ATP binding.

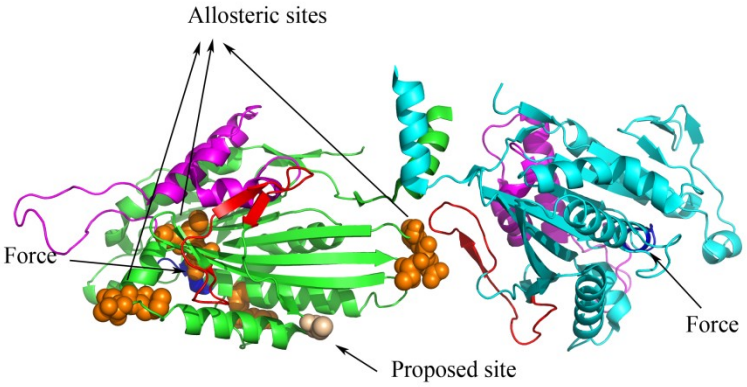


Figure 3.19. Distribution of forces by ATP binding.

The time dependence of the energy is also plotted in Figure 3.20. It is seen that generally tubulin binding regions (A145,A 238,A264,A283) are affected by the application of force at the start of simulation in the monomer that force is applied. However, those residues will come to equilibration and the other regions especially neighboring residues of the starting and the end points of the tubulin binding sites (A230,A232, A302,A306) are mostly affected. Also, it is seen that the ATP binding site is affected most at all instances.

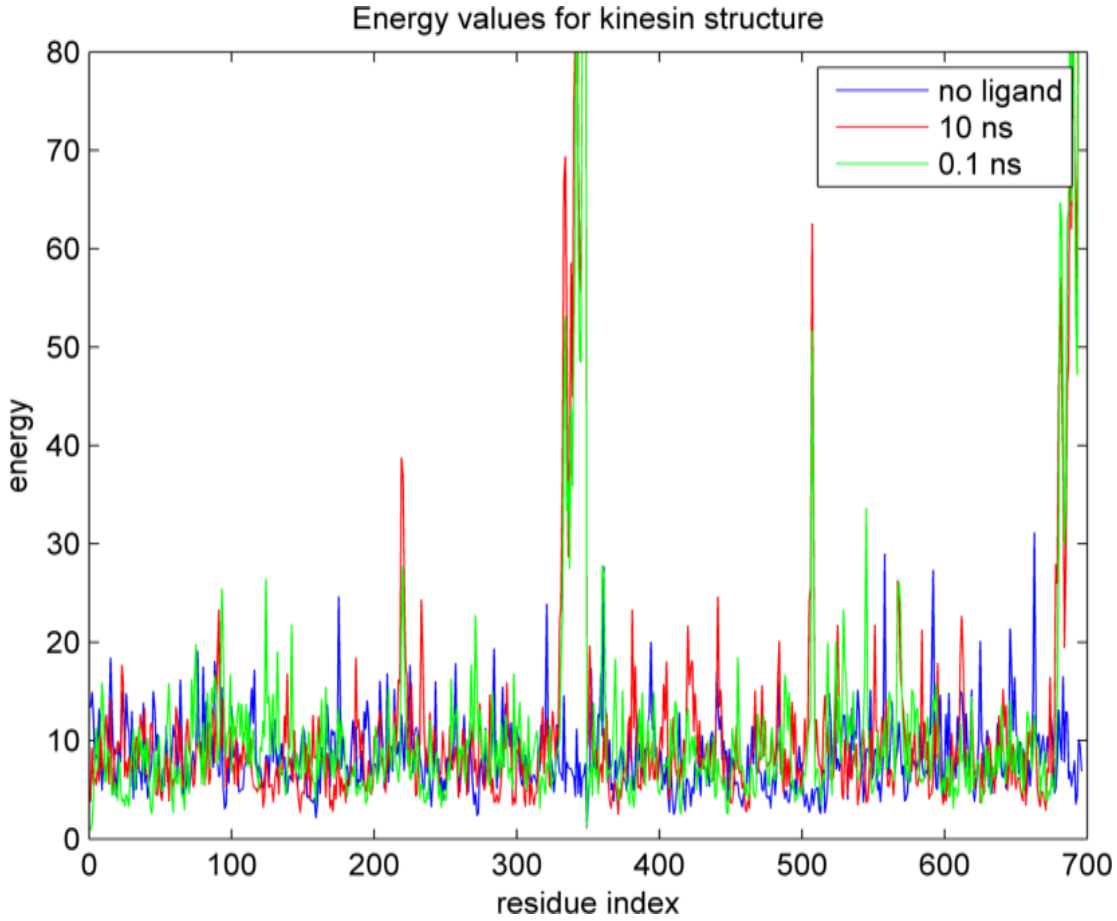


Figure 3.20. Kinesin molecule force change with respect to time upon ATP binding.

Monte Carlo simulations and the results obtained from various other tools e.g GNM, evolutionary conservation [65], suggested an allosteric residue to test in AFM experiments. Therefore, the force distribution algorithm is applied from this residue. The energy values are shown in Figure 3.21.

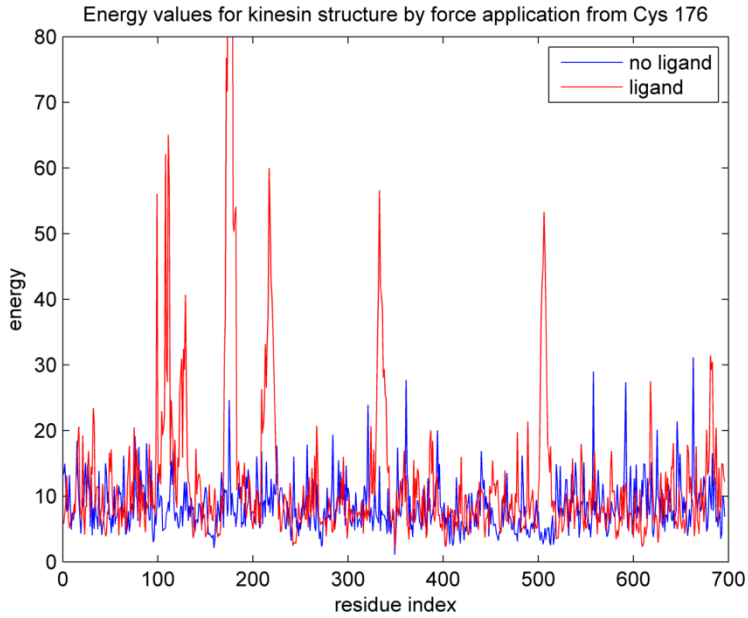


Figure 3.21. The difference in the energy for perturbed and unperturbed kinesin molecules via the proposed site.

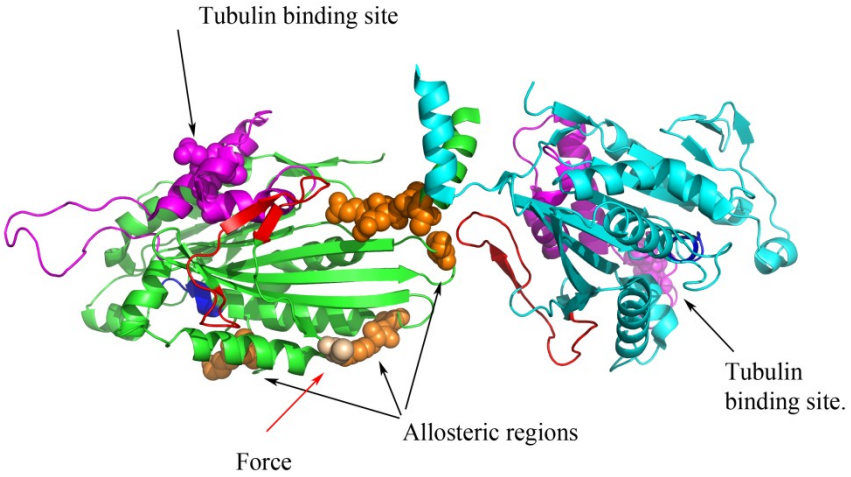


Figure 3.22. Distribution of forces by proposed allosteric site.

The force applied from proposed site on the first monomer propagates to an allosteric site which is the starting point of the tubulin binding site (A232), the ATP binding site (A91) and to the neck-linker region (A334). It also affects the tubulin binding site on the other monomer even though the force is applied from the other monomers proposed allosteric region. The distance between proposed site and the tubulin binding site of the other monomer is nearly 45 Å which is a very long range. Therefore, it can be stated that the application of the force will propagate through the other monomer's tubulin binding

site. As a result, pulling experiments may result in the induce the motility of the protein due to the lack of the signal. The representation of effected regions on the structure is shown in Figure 3.22.

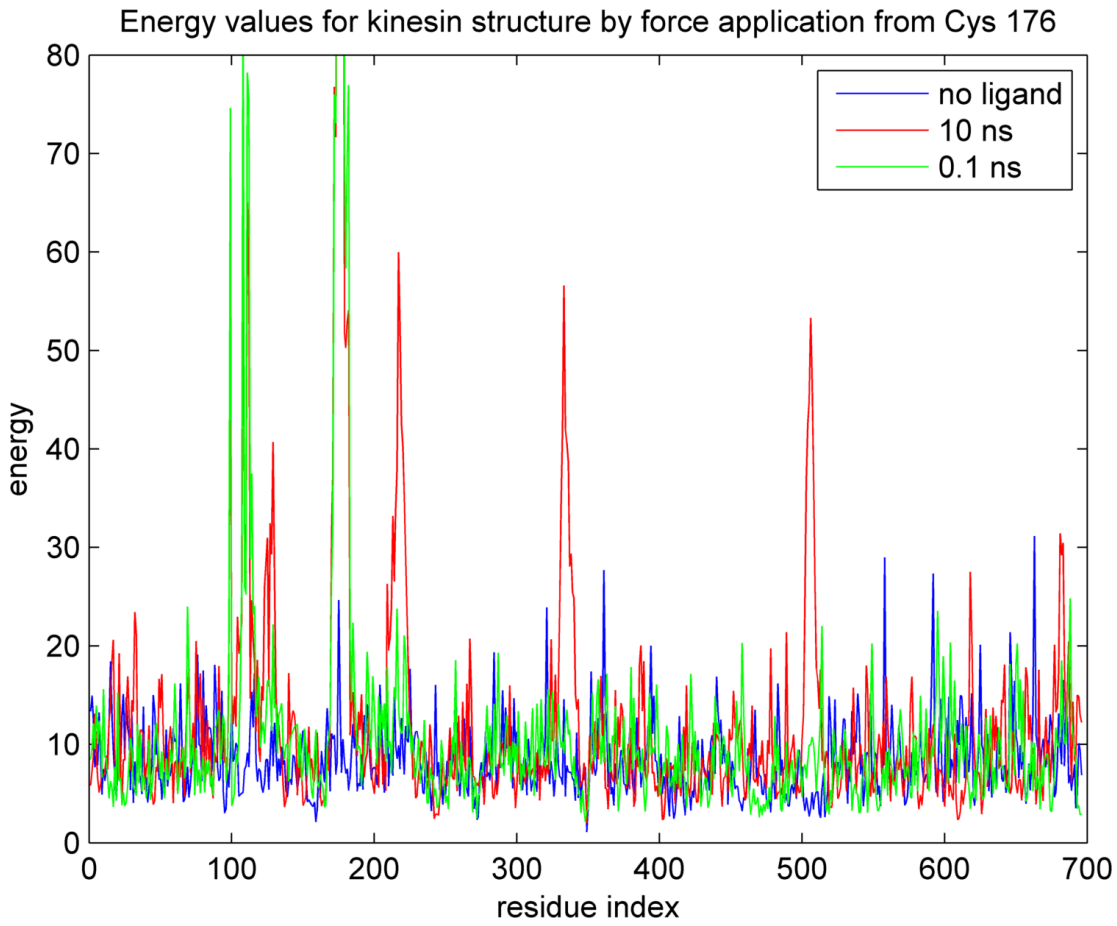


Figure 3.23. Kinesin molecule force change with respect to time upon force application on proposed site.

The change in the forces with respect to time is plotted in Figure 3.23. The initial impact of the application of the force on this region affects the ATP binding site (A88-A93) on the monomer that the force is applied and some unstable regions that are on the tubulin binding site (A255,A292). However, the propagation of force will complete at least in one ns. Hence, the initial energy profile changes during simulations and tubulin binding regions on this monomer stabilizes. The neck linker region (A335), other monomers tubulin binding site (C162) and the allosteric region (A176) on this monomer are gained energy at 10 ns unlike at 0.1 ns. The energy of the ATP binding site is stabilized after a certain while. For the rest of structure, the residues did not get affected.

3.3. Atomic Force Microscopy Results

The force exerted due to the interaction between kinesin and microtubule molecules are investigated. Firstly, the functionalized tip is approached to a hard surface which cannot deform at contact and calibrated. The tip of cantilever retracted with piezo and the voltage that is measured by photo detector. These voltage values are synchronized with displacement-force values. These forces are resulting from the defflection of cantilever. The spring constant of cantilever is measured with thermal noise method.

Pulling experiments are performed with calibrated and functionalized AFM cantilevers by approaching polymerized microtubules molecules on the Petri dish. In the experiment, the kinesin molecule on functionalized cantilever is approached to microtubule coated surface until the binding between kinesin and microtubule molecules occur. After contact, cantilever is retracted from surface and the bond between kinesin and microtubule ruptured. The rupture forces are investigated with different retraction velocities such as 1μm/s , 5 μm/s and 10 μm/s. In Figure 3.24, an example displacement-force curves is plotted. The y-axis shows the force between kinesin and microtubule molecule and x-axis shows the displacement that the piezo retracted. The zero point on x-axis indicates the contact point of the cantilever tip to the surface. The blue region is the approach part and the red region is the retract part of cantilever. The Figure 3.24a, the cantilever approaches to surface from 1 μm and after contact point, cantilever retracted from surface without an unbinding or rupture event. On the other hand, Figure 3.24b shows a rupture event that happens at 0.5 μm. The difference between on top and the bottom of rupture part is the indication of the rupture event and in this experiment, a force around 250 pN is measured as rupture force.

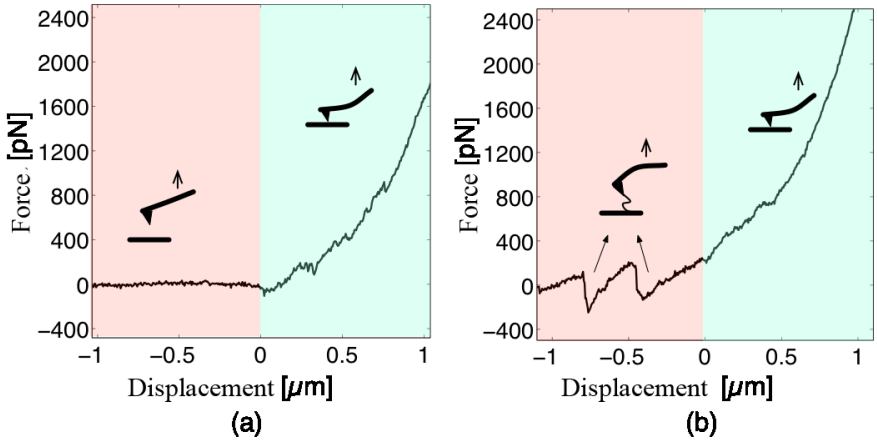


Figure 3.24. Sample approach-retract curve in AFM experiment.

Due to high viscosity in the media, the viscous forces affecting the cantilever may mislead most of the results. Also, the lateral Young's modulus of the polymerized microtubule molecules result in indentation and the many of the retract curves are obtained without a rupture event. Apart from these two problems, the polymerized microtubule molecules are scattered in media as bundles. Therefore, more than one unbinding events with very different magnitudes are observed which is an implication of the rupture between microtubule-microtubule molecules or microtubule-surface.

To overcome these issues, another experiment is designed. To this end, the cantilever tip is again functionalized with biotinylated kinesin molecules with streptavidin yet, the surface is coated with monomer form of microtubules namely tubulin with a dilute concentration and the experiment is repeated.

Based on the results of computer simulations, there is a new site suggested on the outer surface of the molecule. Cys 176 is a residue with high solvent accessible surface area, high closeness value in Monte Carlo pathway generations and close to a hinge region. Due to the presence of sulfhydryl group, as explained in previous section, enables the functionalization with maleimide. Therefore, the mechanical properties, binding events, energy landscape of the kinesin microtubule interactions could be further investigated, through this site.

4. CONCLUSIONS AND RECOMMENDATIONS

4.1. Conclusions

To mimic atomic force microscopy experiments and the ligand binding events, force distribution analysis method with Brownian dynamics simulations is proposed here. As demonstrated with methionine and purine repressor proteins, the force applied from the ligand binding sites is able to distribute through the structure to the DNA binding site. For kinesin structure, the proposed and performed experiments are simulated with application of pulling force and the results of the analysis imply that the internal strain propagates to the functional regions in the structure. The propagation path is determined to be directly passing through tubulin binding sites (A159,A163,A280) and ATP binding sites which are the main key residues for the motility and the main cargo transport function of the kinesin molecule.

The Monte Carlo Path generation method has shown to be a useful tool to investigate the allosteric signaling mechanism on protein structure. The generated pathways for kinesin structure suggest novel sites that play a key role in the allosteric signaling. There are parallel pathways in signaling between neck linker and tubulin binding sites as moving through core of protein or the N-terminal region. The centrality parameters of residues indicate that there are still some residues that are far from neck linker region and has still high closeness and betweenness values even if those values are not as high as linker region, implying that the function of this protein may rely on some other regions (CYS176) as well as the neck-linker region.

Allostery is an interesting phenomenon yet remains evanescent. The induction of different function by a catalytic or a binding event is intriguing, yet, there has to be a connection via a form of signal between the functional sites. The simulations performed for the thesis to investigate the regions that are highly correlated with the signal propagation mechanism display that the force distribution analysis and the Monte Carlo path generation

are promising, fast and reliable. Further, the present analysis provides preliminary efforts to combine simulations and the atomic force microscopy experiments.

4.2. Recommendations for Future Studies

The force distribution analysis method with Brownian dynamics is a new methodology. In dynamics simulations, the potential used is the GNM potential, yet a comprehensive potential which takes other interactions into account such as non-bonded interactions, bond angles, torsion angles; more accurate simulations with extended CPU times may be obtained. The algorithm is strict and software for force distribution analysis will be produced in future. The FDA algorithm shows the first order Langevin equation, however, it is also integrate the second order Langevin equation which may result in better energy distribution. Due to stochastic nature of the Langevin equation, parallel runs needs to be performed to reduce the effect of probability. Apart from these issues, the numerical integration is held with second order Runge-Kutta method and a higher order integration method such as fourth order Runge-Kutta is better in terms of accuracy and precision yet this method is slower compared to second order. Also, the simulations may be performed at different temperatures.

The number of AFM experiments is very few and since the experiment is stochastic, the experiments need to be repeated. Also, new sites on the periphery which may play a key role in the allosteric signaling needs to be determined and experiments rerun with different functionalization. Also, if the predicted key residue for signaling is in the core region of protein, it may be better to mutate those residues and check the difference in the rupture force.

APPENDIX A. RUNGE-KUTTA METHOD

Langevin equation which is the dynamics model of the low-resolution protein structure is a stochastic differential equation. A stochastic differential equation does not have an exact solution. Therefore, it is required to integrate the differential equation numerically. The statistically correct results are obtained from Runge-Kutta method with order k and time step s.

General form of a stochastic differential equation in Brownian dynamics simulations is as follows:

$$\frac{dx}{dt} = f(x) + A(t) \tag{A.1}$$

In this equation, A(t) are the normally distributed random variables with zero mean and covariance

$$\langle A(t)A(t')\rangle = \xi\delta(t - t') \tag{A.2}$$

The Eq. (A.1 can be integrated in a time step as following:

$$x = x_0 + \int_0^s f(x) + w_0(s) \tag{A.3}$$

Here, f(x) can be extended as

$$f(x) = f_0 + f_0^{'}(x - x_0) + ... \frac{1}{p!} f_0^{(p)}(x - x_0)^p \tag{A.4}$$

for all the points of x_0 accessible to trajectory. To calculate the trajectories the second order Runge-Kutta method which is an approximation of Eq. (A.4 is used in following form.

$$
\begin{aligned}
g_1 &= f(x_0 + s^{1/2}\xi^{1/2}\lambda_1 Z) \\
g_2 &= f(x_0 + s\beta_{21}g_1 + s^{1/2}\xi^{1/2}\lambda_2 Z \\
x_s &= x_0 + s(A_1 g_1 + A_2 g_2) + s^{1/2}\xi^{1/2}\lambda_0 Z
\end{aligned}
\tag{A.5}
$$

The solution set of parameters are the following:

$$
\begin{aligned}
&A_1 = A_2 = 1/2 \\
&\beta_{21} = 1 \\
&\lambda_0 = 1, \lambda_1 = 0, \lambda_2 = 1
\end{aligned}
\tag{A. 6}
$$

The first and second order perturbations of f which is the derivative of the potential energy function calculated in the program. Z is a single random variable from uniformly distributed variables between 0 and 1 and; ξ is equal to the covariance of the Gaussianly distributed force in Eq. (2.1.

APPENDIX B. FENG-DOOLITTLE ALGORITHM

The most probable pathways are differing very little from each other. Therefore, it is required to cluster that sequences. However, the length of the sequences are not exactly equal so the pathways needs to be aligned and clustering needs to be performed after alignment. For this purpose, a very well-established multiple alignment algorithm Feng-Doolittle algorithm is used.

The algorithm divided into three different steps such as pair-wise alignment, clustering and construction of multiple alignment. For the first step, all the sequences are aligned with a global pair-wise alignment method such as Needleman-Wunsch algorithm. The calculated alignment score is normalized and a pair-wise distance matrix has been built.

In the second step, based on the distance matrix evaluated an agglomerative clustering such as single linkage clustering is applied. A guide tree is produced to determine the order of the sequences in the construction of multiple alignment.

The third step involves in the construction of the multiple alignment. From the deep leaves of the guide tree, sequences with strong similarities are aligned and appended in the multiple alignment sequence in traverse order. The third step is repeated for each child node on the guide tree.

REFERENCES

1. Cui, Q. and M. Karplus, "Allostery and Cooperativity Revisited", *Protein Science*, Vol. 17, No. 8, pp. 1295-307, 2008.

2. Changeux, J. P., "Allostery and the Monod-Wyman-Changeux Model after 50 Years", *Annual Reviews of Biophysics*, Vol. 41, No. pp. 103-33, 2012.

3. Goodey, N. M. and S. J. Benkovic, "Allosteric Regulation and Catalysis Emerge Via a Common Route", *Nature Chemical Biology*, Vol. 4, No. 8, pp. 474-82, 2008.

4. Tsai, C. J., A. del Sol, and R. Nussinov, "Allostery: Absence of a Change in Shape Does Not Imply That Allostery Is Not at Play", *Journal of Molecular Biology*, Vol. 378, No. 1, pp. 1-11, 2008.

5. del Sol, A., C. J. Tsai, B. Ma, and R. Nussinov, "The Origin of Allosteric Functional Modulation: Multiple Pre-Existing Pathways", *Structure*, Vol. 17, No. 8, pp. 1042-50, 2009.

6. Boehr, D. D., H. J. Dyson, and P. E. Wright, "An Nmr Perspective on Enzyme Dynamics", *Chemical Reviews*, Vol. 106, No. 8, pp. 3055-79, 2006.

7. Kern, D. and E. R. Zuiderweg, "The Role of Dynamics in Allosteric Regulation", *Current Opinions in Structural Biology*, Vol. 13, No. 6, pp. 748-57, 2003.

8. Swain, J. F., E. G. Schulz, and L. M. Gierasch, "Direct Comparison of a Stable Isolated Hsp70 Substrate-Binding Domain in the Empty and Substrate-Bound States", *Journal of Biological Chemistry*, Vol. 281, No. 3, pp. 1605-11, 2006.

9. Volkman, B. F., D. Lipson, D. E. Wemmer, and D. Kern, "Two-State Allosteric Behavior in a Single-Domain Signaling Protein", *Science*, Vol. 291, No. 5512, pp. 2429-33, 2001.

10. Matsui, H., S. Lazareno, and N. J. Birdsall, "Probing of the Location of the Allosteric Site on M1 Muscarinic Receptors by Site-Directed Mutagenesis", *Molecular Pharmacology*, Vol. 47, No. 1, pp. 88-98, 1995.

11. Valentini, G., L. Chiarelli, R. Fortin, M. L. Speranza, A. Galizzi, and A. Mattevi, "The Allosteric Regulation of Pyruvate Kinase", *Journal of Biological Chemistry*, Vol. 275, No. 24, pp. 18145-52, 2000.

12. Lockless, S. W. and R. Ranganathan, "Evolutionarily Conserved Pathways of Energetic Connectivity in Protein Families", *Science*, Vol. 286, No. 5438, pp. 295-9, 1999.

13. Ota, N. and D. A. Agard, "Intramolecular Signaling Pathways Revealed by Modeling Anisotropic Thermal Diffusion", *Journal of Molecular Biology*, Vol. 351, No. 2, pp. 345-54, 2005.

14. Atilgan, A. R., D. Turgut, and C. Atilgan, "Screened Nonbonded Interactions in Native Proteins Manipulate Optimal Paths for Robust Residue Communication", *Biophysical Journal*, Vol. 92, No. 9, pp. 3052-62, 2007.

15. Gerek, Z. N. and S. B. Ozkan, "Change in Allosteric Network Affects Binding Affinities of Pdz Domains: Analysis through Perturbation Response Scanning", *Plos Computational Biology*, Vol. 7, No. 10, pp. e1002154, 2011.

16. Chennubhotla, C. and I. Bahar, "Markov Propagation of Allosteric Effects in Biomolecular Systems: Application to Groel-Groes", *Moleculat Systems Biology*, Vol. 2, No. pp. 36, 2006.

17. Weinkam, P., J. Pons, and A. Sali, "Structure-Based Model of Allostery Predicts Coupling between Distant Sites", *Proceedings of National Academy of Science U S A*, Vol. 109, No. 13, pp. 4875-80, 2012.

18. Mitternacht, S. and I. N. Berezovsky, "A Geometry-Based Generic Predictor for Catalytic and Allosteric Sites", *Protein Engineering Design & Selection*, Vol. 24, No. 4, pp. 405-9, 2011.

19. Vijayabaskar, M. S. and S. Vishveshwara, "Interaction Energy Based Protein Structure Networks", *Biophysical Journal*, Vol. 99, No. 11, pp. 3704-15, 2010.

20. Cilia, E., G. W. Vuister, and T. Lenaerts, "Accurate Prediction of the Dynamical Changes within the Second Pdz Domain of Ptp1e", *PLoS Computational Biology*, Vol. 8, No. 11, pp. e1002794, 2012.

21. Armutlulu, A., *Monte Carlo (Mc) Path Generation and Small-World Network Approach to Identify Functional Residues in Proteins,* M.S. Thesis, Bogazici University, 2009.

22. Ekesan, S., *Prediction of Allosteric Key Residues and Their Role in Protein Folding,* M.S. Thesis, Bogazici University, 2009.

23. Kaya, C., A. Armutlulu, S. Ekesan, and T. Haliloglu, "Mcpath: Monte Carlo Path Generation Approach to Predict Likely Allosteric Pathways and Functional Residues", *Nucleic Acids Res*, 2013.

24. Stacklies, W., F. Xia, and F. Grater, "Dynamic Allostery in the Methionine Repressor Revealed by Force Distribution Analysis", *Plos Computational Biology*, Vol. 5, No. 11, pp. e1000574, 2009.

25. Helmstaedt, K., S. Krappmann, and G. H. Braus, "Allosteric Regulation of Catalytic Activity: Escherichia Coli Aspartate Transcarbamoylase Versus Yeast Chorismate Mutase", *Microbiology and Molecular Biology Reviews*, Vol. 65, No. 3, pp. 404-21, table of contents, 2001.

26. Tsai, C. J., A. del Sol, and R. Nussinov, "Allostery: Absence of a Change in Shape Does Not Imply That Allostery Is Not at Play", *Journal of Molecular Biology*, Vol. 378, No. 1, pp. 1-11, 2008.

27. Hua, W., J. Chung, and J. Gelles, "Distinguishing Inchworm and Hand-over-Hand Processive Kinesin Movement by Neck Rotation Measurements", *Science*, Vol. 295, No. 5556, pp. 844-8, 2002.

28. Helmes, M., K. Trombitas, T. Centner, M. Kellermayer, S. Labeit, W. A. Linke, and H. Granzier, "Mechanically Driven Contour-Length Adjustment in Rat Cardiac Titin's Unique N2b Sequence - Titin Is an Adjustable Spring", *Circulation Research*, Vol. 84, No. 11, pp. 1339-1352, 1999.

29. Thomas, W., M. Forero, O. Yakovenko, L. Nilsson, P. Vicini, E. Sokurenko, and V. Vogel, "Catch-Bond Model Derived from Allostery Explains Force-Activated Bacterial Adhesion", *Biophysical Journal*, Vol. 90, No. 3, pp. 753-764, 2006.

30. Berg, J. M., J. L. Tymoczko, and L. Stryer, *Biochemistry*. 6th ed New York: W. H. Freeman. 1 v. (various pagings), 2007.

31. Stacklies, W., M. C. Vega, M. Wilmanns, and F. Grater, "Mechanical Network in Titin Immunoglobulin from Force Distribution Analysis", *Plos Computational Biology*, Vol. 5, No. 3, pp. e1000306, 2009.

32. Stacklies, W., C. Seifert, and F. Graeter, "Implementation of Force Distribution Analysis for Molecular Dynamics Simulations", *Bmc Bioinformatics*, Vol. 12, No. pp. 2011.

33. Miyashita, O., J. N. Onuchic, and P. G. Wolynes, "Nonlinear Elasticity, Proteinquakes, and the Energy Landscapes of Functional Transitions in Proteins", *Proceedings of National Academy of Science USA*, Vol. 100, No. 22, pp. 12570-12575, 2003.

34. Miyashita, O., P. G. Wolynes, and J. N. Onuchic, "Simple Energy Landscape Model for the Kinetics of Functional Transitions in Proteins", *Journal of Physical Chemistry B*, Vol. 109, No. 5, pp. 1959-1969, 2005.

35. Hirokawa, N., "Kinesin and Dynein Superfamily Proteins and the Mechanism of Organelle Transport", *Science*, Vol. 279, No. 5350, pp. 519-26, 1998.

36. Hirokawa, N. and Y. Noda, "Intracellular Transport and Kinesin Superfamily Proteins, Kifs: Structure, Function, and Dynamics", *Physiological Reviews*, Vol. 88, No. 3, pp. 1089-118, 2008.

37. Hirokawa, N. and R. Takemura, "Molecular Motors and Mechanisms of Directional Transport in Neurons", *Nature Reviews Neuroscience*, Vol. 6, No. 3, pp. 201-14, 2005.

38. Vale, R. D., "The Molecular Motor Toolbox for Intracellular Transport", *Cell*, Vol. 112, No. 4, pp. 467-80, 2003.

39. Rice, S., A. W. Lin, D. Safer, C. L. Hart, N. Naber, B. O. Carragher, S. M. Cain, E. Pechatnikova, E. M. Wilson-Kubalek, M. Whittaker, E. Pate, R. Cooke, E. W. Taylor, R. A. Milligan, and R. D. Vale, "A Structural Change in the Kinesin Motor Protein That Drives Motility", *Nature*, Vol. 402, No. 6763, pp. 778-84, 1999.

40. Vale, R. D. and R. J. Fletterick, "The Design Plan of Kinesin Motors", *Annual Reviews of Cell and Developmental Biology*, Vol. 13, No. pp. 745-77, 1997.

41. Yildiz, A., M. Tomishige, R. D. Vale, and P. R. Selvin, "Kinesin Walks Hand-over-Hand", *Science*, Vol. 303, No. 5658, pp. 676-8, 2004.

42. Korten, T., B. Nitzsche, C. Gell, F. Ruhnow, C. Leduc, and S. Diez, "Fluorescence Imaging of Single Kinesin Motors on Immobilized Microtubules", *Methods in Molecular Biology*, Vol. 783, No. pp. 121-37, 2011.

43. Svoboda, K., C. F. Schmidt, B. J. Schnapp, and S. M. Block, "Direct Observation of Kinesin Stepping by Optical Trapping Interferometry", *Nature*, Vol. 365, No. 6448, pp. 721-7, 1993.

44. Friedman, D. S. and R. D. Vale, "Single-Molecule Analysis of Kinesin Motility Reveals Regulation by the Cargo-Binding Tail Domain", *Nature Cell Biology*, Vol. 1, No. 5, pp. 293-7, 1999.

45. von Massow, A., E. M. Mandelkow, and E. Mandelkow, "Interaction between Kinesin, Microtubules, and Microtubule-Associated Protein 2", *Cell Motility and the Cytoskeleton*, Vol. 14, No. 4, pp. 562-71, 1989.

46. Vale, R. D. and R. A. Milligan, "The Way Things Move: Looking under the Hood of Molecular Motor Proteins", *Science*, Vol. 288, No. 5463, pp. 88-95, 2000.

47. Lal, R. and S. A. John, "Biological Applications of Atomic-Force Microscopy", *American Journal of Physiology*, Vol. 266, No. 1, pp. C1-&, 1994.

48. Durfrene, Y. F., "Atomic Force Microscopy Provides a New Means for Looking at Microbial Cells", *Asm News*, Vol. 69, No. 9, pp. 438-442, 2003.

49. Shegaonkar, A. C. and S. M. Salapaka, "Feedback Based Simultaneous Correction of Imaging Artifacts Due to Geometrical and Mechanical Cross-Talk and Tip-Sample Stick in Atomic Force Microscopy", *Reviews of Scientific Instrumentation*, Vol. 78, No. 10, pp. 2007.

50. Hansma, P. K., V. B. Elings, O. Marti, and C. E. Bracker, "Scanning Tunneling Microscopy and Atomic Force Microscopy - Application to Biology and Technology", *Science*, Vol. 242, No. 4876, pp. 209-216, 1988.

51. Richter, R. P. and A. R. Brisson, "Following the Formation of Supported Lipid Bilayers on Mica: A Study Combining Afm, Qcm-D, and Ellipsometry", *Biophysical Journal*, Vol. 88, No. 5, pp. 3422-3433, 2005.

52. van Noort, J., F. Orsini, A. Eker, C. Wyman, B. de Grooth, and J. Greve, "DNA Bending by Photolyase in Specific and Non-Specific Complexes Studied by Atomic Force Microscopy", *Nucleic Acids Research*, Vol. 27, No. 19, pp. 3875-3880, 1999.

53. Li, M., X. B. Xiao, L. Q. Liu, N. Xi, Y. C. Wang, Z. L. Dong, and W. J. Zhang, "Imaging and Measuring the Molecular Force of Lymphoma Pathological Cells Using Atomic Force Microscopy", *Scanning*, Vol. 35, No. 1, pp. 40-46, 2013.

54. Shroff, S. G., D. R. Saner, and R. Lal, "Dynamic Micromechanical Properties of Cultured Rat Atrial Myocytes Measured by Atomic-Force Microscopy", *American Journal of Physiology-Cell Physiology*, Vol. 269, No. 1, pp. C286-C292, 1995.

55. Alsteens, D., M. C. Garcia, P. N. Lipke, and Y. F. Dufrene, "Force-Induced Formation and Propagation of Adhesion Nanodomains in Living Fungal Cells", *Proceedings of National Academy of Science USA*, Vol. 107, No. 48, pp. 20744-20749, 2010.

56. Raible, M., M. Evstigneev, F. W. Bartels, R. Eckel, M. Nguyen-Duong, R. Merkel, R. Ros, D. Anselmetti, and P. Reimann, "Theoretical Analysis of Single-Molecule Force Spectroscopy Experiments: Heterogeneity of Chemical Bonds", *Biophysical Journal*, Vol. 90, No. 11, pp. 3851-64, 2006.

57. Kull, F. J., E. P. Sablin, R. Lau, R. J. Fletterick, and R. D. Vale, "Crystal Structure of the Kinesin Motor Domain Reveals a Structural Similarity to Myosin", *Nature*, Vol. 380, No. 6574, pp. 550-5, 1996.

58. Kozielski, F., S. Sack, A. Marx, M. Thormahlen, E. Schonbrunn, V. Biou, A. Thompson, E. M. Mandelkow, and E. Mandelkow, "The Crystal Structure of Dimeric Kinesin and Implications for Microtubule-Dependent Motility", *Cell*, Vol. 91, No. 7, pp. 985-94, 1997.

59. Perutz, M. F., "Mechanisms of Cooperativity and Allosteric Regulation in Proteins", *Quarterly Reviews of Biophysics*, Vol. 22, No. 2, pp. 139-237, 1989.

60. Somers, W. S. and S. E. Phillips, "Crystal Structure of the Met Repressor-Operator Complex at 2.8 a Resolution Reveals DNA Recognition by Beta-Strands", *Nature*, Vol. 359, No. 6394, pp. 387-93, 1992.

61. Rafferty, J. B., W. S. Somers, I. Saint-Girons, and S. E. Phillips, "Three-Dimensional Crystal Structures of Escherichia Coli Met Repressor with and without Corepressor", *Nature*, Vol. 341, No. 6244, pp. 705-10, 1989.

62. Schumacher, M. A., A. Glasfeld, H. Zalkin, and R. G. Brennan, "The X-Ray Structure of the Purr-Guanine-Purf Operator Complex Reveals the Contributions of Complementary Electrostatic Surfaces and a Water-Mediated Hydrogen Bond to Corepressor Specificity and Binding Affinity", *Journal of Biological Chemistry*, Vol. 272, No. 36, pp. 22648-53, 1997.

63. Bahar, I., A. R. Atilgan, and B. Erman, "Direct Evaluation of Thermal Fluctuations in Proteins Using a Single-Parameter Harmonic Potential", *Folding and Design*, Vol. 2, No. 3, pp. 173-81, 1997.

64. Sindelar, C. V., M. J. Budny, S. Rice, N. Naber, R. Fletterick, and R. Cooke, "Two Conformations in the Human Kinesin Power Stroke Defined by X-Ray Crystallography and Epr Spectroscopy", *Nature Structural Biology*, Vol. 9, No. 11, pp. 844-8, 2002.

65. Ashkenazy, H., E. Erez, E. Martz, T. Pupko, and N. Ben-Tal, "Consurf 2010: Calculating Evolutionary Conservation in Sequence and Structure of Proteins and Nucleic Acids", *Nucleic Acids Res*, Vol. 38, No. Web Server issue, pp. W529-33, 2010.

66. Laskowski, R. A., E. G. Hutchinson, A. D. Michie, A. C. Wallace, M. L. Jones, and J. M. Thornton, "Pdbsum: A Web-Based Database of Summaries and Analyses of All Pdb Structures", *Trends in Biochemical Science*, Vol. 22, No. 12, pp. 488-90, 1997.

67. Duselder, A., C. Thiede, C. F. Schmidt, and S. Lakamper, "Neck-Linker Length Dependence of Processive Kinesin-5 Motility", *Journal of Molecular Biology*, Vol. 423, No. 2, pp. 159-68, 2012.

68. Stacklies, W., C. Seifert, and F. Graeter, "Implementation of Force Distribution Analysis for Molecular Dynamics Simulations", *BMC Bioinformatics*, Vol. 12, No. pp. 101, 2011.