REPUBLIC OF TURKEY
ACIBADEM MEHMET ALİ AYDINLAR ÜNİVERSİTESİ
INSTITUTE OF NATURAL AND APPLIED SCIENCES

ASSESSING THE PERFORMANCE OF MOLECULAR DYNAMICS SIMULATIONS FOR PREDICTING THE CONFORMATIONS OF FLEXIBLE PROTEIN-PROTEIN COMPLEXES

DİLANUR KAMALI

MASTER THESIS

DEPARTMENT of MOLECULAR
BIOLOGY AND GENETICS

SUPERVISOR
Prof Emel Timuçin

İSTANBUL-2023

REPUBLIC OF TURKEY
ACIBADEM MEHMET ALİ AYDINLAR ÜNİVERSİTESİ
INSTITUTE OF NATURAL AND APPLIED SCIENCES

ASSESSING THE PERFORMANCE OF MOLECULAR DYNAMICS SIMULATIONS FOR PREDICTING THE CONFORMATIONS OF FLEXIBLE PROTEIN-PROTEIN COMPLEXES

DİLANUR KAMALI

MASTER THESIS

DEPARTMENT of MOLECULAR
BIOLOGY AND GENETICS

SUPERVISOR
Prof Emel Timuçin

İSTANBUL-2023

Department:	Institute of Natural and Applied Sciences
Program:	Molecular and Translational Biomedicine
Thesis Title:	Assessing the Performance of Molecular Dynamics Simulations for Predicting the Conformations of Flexible Protein-Protein Complexes
Student's name and Surname:	Dilanur Kamalı
Date of Defence:	04/10/2023

This is to certify that I have examined this copy of master thesis. I have found that she/he prepared after fulfilling the specified requirements in the associated legislations before the final examining committee whose signatures are below.

Jury President	Prof Eda Tahir Turanlı Acıbadem Mehmet Ali Aydınlar Üniversitesi
Supervisor of Thesis	Prof Emel Timuçin Acıbadem Mehmet Ali Aydınlar Üniversitesi
Jury Member	Asst Prof Serkan Belkaya İhsan Doğramacı Bilkent Üniversitesi

DECLARATION

I declare that this thesis work is my own work, I had no unethical behavior at any stages from the planning to the writing of the thesis, I obtained all the information in this thesis in accordance with academic and ethical rules, I cited all the information and comments that were not obtained with this thesis work, and I provided resources in the list of references. I also declare that there was no violation of any patents and copyrights during the study and writing of this thesis.

04.10.2023

Dilanur Kamalı

PREFACE AND ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my advisor, Prof Emel Timuçin, for her unwavering support, invaluable guidance, and endless patience throughout the entire journey of this thesis. Your wisdom and expertise have been the guiding light that illuminated my path in this challenging endeavor.

I am also profoundly thankful to my family, who have been a constant source of encouragement and motivation. Your belief in me, especially during the toughest moments, has been instrumental in my success. To my mother Yücel Öztürk and my siblings Mertcan Kamali, Sercan Kamali, Ahmet Can Kamali, Selin Göker Kamali, and Semra Karaman Kamali, your love and support mean the world to me.

I extend my sincere appreciation to the diligent secretary of the institute, Yağmur Kalafatoğlu, whose assistance and efficiency made administrative tasks manageable, allowing me to focus on my research without unnecessary distractions.

My heartfelt thanks go out to my friends, Msc Başak Kavaklıoğlu, Gülşah Şebnem Özköse, and Msc Merve Gündoğdu, who stood by my side throughout this academic journey. Your friendship, encouragement, and moments of respite provided a much-needed balance to the rigors of thesis writing.

Last but not least, I must acknowledge myself for the dedication, perseverance, and countless hours invested in this journey. It has been an incredibly challenging, yet rewarding experience that has shaped me in countless ways.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND SYMBOLS

Å	Angstrom
AMD	Accelerated Molecular Dynamics
CMD	Conventional (Classical) Molecular Dynamics
DNA	Deoxyribonucleic Acid
FS	Femtosecond
GAMD	Gaussian-Accelerated Molecular Dynamics
K	Kelvin
MD	Molecular Dynamics
NPT	Isothermal–Isobaric Ensemble
NS	Nanosecond
PDB	Protein Data Bank
PME	Particle-Mesh Ewald
RG	Radius of Gyration
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuation
RNA	Ribonucleic Acid
SASA	Solvent Accessible Surface Area
VMD	Visual Molecular Dynamics

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Enhanced comprehension of alterations in protein conformation holds significant potential for unveiling the underlying mechanisms of diverse biological functions and more specifically describing structural transitions. In investigating the binding interactions and conformations of biomolecules, structural dynamics is the key component. Elucidating the mechanism expands our comprehension of the system and impacts the designing of effective drugs and novel treatments. Here, with the help of enhanced molecular dynamic (MD) simulation methods, the backbone of gelsolin with conventional and accelerated MD methods is focused on and its conformational changes that occur during the gelsolin/actin complex formation from the gelsolin plasma free form triggered by calcium concentration is investigated. We found that there is no complete conformational transition in either conventional (classical) MD or accelerated MD. However, when we focused on the gelsolin-like domains, some conformational rearrangements, particularly in G4 and G6 domains adopted similar domain conformations observed in the free form. These changes were especially noted for the accelerated MD simulations with dual boosting. On the other hand, G5 domain was the most rigid part of the structure regardless of the MD acceleration. Based on the findings of this thesis, capturing conformational changes in MD simulations proves to be challenging. One approach could involve focusing on the structures of individual gelsolin domains, rather than the entire structure, and studying their conformational shifts using different accelerated methods. Altogether, this thesis provides insights into the potential use of accelerated MD methods for understanding and capturing important conformational changes in proteins.

Keywords: Conformational change, Conformational selection, Gelsolin, Molecular dynamics simulations, Protein-protein interaction.

ÖZET

Esnek protein-protein komplekslerinin konformasyonlarını tahmin etmek için moleküler dinamik simülasyonlarının performansının değerlendirilmesi

Protein konformasyonundaki değişikliklerin daha iyi anlaşılması, çeşitli biyolojik fonksiyonların altında yatan mekanizmaları ortaya çıkarmak ve daha spesifik olarak yapısal geçişleri tanımlamak için önemli bir potansiyele sahiptir. Biyomoleküllerin bağlanma etkileşimlerini ve konformasyonlarını araştırırken yapısal dinamikler anahtar bileşendir. Mekanizmayı açıklamak, sistem hakkındaki anlayışımızı genişletir ve yeni ilaçların ve yeni tedavilerin tasarlanmasına katkıda bulunur. Burada, geliştirilmiş moleküler dinamik simülasyon yöntemleri yardımıyla, geleneksel ve akselere edilmiş moleküler dinamik yöntemlerle gelsolinin omurgasına odaklandık ve kalsiyum konsantrasyonunun tetiklediği gelsolin plazma serbest formundan gelsolin/aktin kompleksi oluşumu sırasında meydana gelen konformasyonel değişiklikleri incelendi. Ne geleneksel (klasik) MD'de ne de akselere edilmiş MD'de bütüncül bir konformasyonel değişiklik olmadığını tespit edildi. Bununla birlikte, gelsolin benzeri domainlere odaklandığımızda, özellikle G4 ve G6 domainlerindeki bazı konformasyonel değişiklikler, serbest formda gözlemlenen benzer domain konformasyonlarını benimsemiştir. Bu değişiklikler özellikle dual akselerasyon MD simülasyonlarında dikkat çekti. Öte yandan G5 domaini, MD akselerasyonundan bağımsız olarak yapının en hareketsiz kısmıydı. Bu tezin bulgularına dayanarak, MD metotlarının gelsolinin aldığı konformasyonel değişiklikleri yakalamak için performansının yüksek olmadığı görülmüştür. Bu nedenle, gelsolin yapısının tamamına odaklanmak yerine, domainlerinin yapılarına ayrı ayrı odaklanmak ve/veya farklı yöntemlerle, farklı seviyelerde akselere edilmiş MD yaklaşımlarını kullanarak konformasyonel değişimlerin yakalanması mümkün olabilir. Bu tez, proteinlerdeki önemli konformasyonel değişiklikleri anlamak ve yakalamak için akselere edilmiş MD yöntemlerinin potansiyel kullanımına dair katkılar yapmaktadır.

Anahtar Sözcükler: Gelsolin, Konformasyonel değişiklik, Konformasyonel seçim, Moleküler dinamik simülasyonlar, Protein-protein etkileşimleri.

1. INTRODUCTION

Proteins are molecular components of life, and they are large macromolecules that consist of the merging of amino acids. Fundamentally, proteins are an important element of the cellular machinery. They carry out a variety of tasks, within living organisms including promoting growth, facilitating repair, and maintaining proper function. One of the essential roles of proteins is that they are catalysts of reactions. Furthermore, they are also in charge of transport in the cell and they transmit the information to RNA from DNA. They are in charge of both synthesizing new molecules and the degradation process. The broad acknowledgment of their participation in various cellular processes has prompted focused endeavors to anticipate their functions based on their sequences and, when accessible, their structures (1–6).

1.1 Structures of Proteins

Proteins fold and create stable three-dimensional conformations based on their amino acid sequence. According to the final shape of the proteins, after they are folded, they own a specific function. Additionally, with the help of this particular shape, and the bonds inside the molecule help to maintain structural stability (7,8).

The unique three-dimensional shape of a protein is formed by the folding and bonding of its amino acid chain. This process is primarily guided by the protein structure, which is its specific sequence of amino acids (8,9). However, in some cases, particular folding patterns can be induced by the hydrogen bonds formed by carboxyl and amino groups. Alpha helices and beta sheets are known as predictable folding patterns (8,9).

There are 4 different levels of complexity including primary, secondary, tertiary,

and quaternary structure in the complete structure of a protein (9–11).

1.1.1 Primary Structure

Primary structure (also referred as protein sequence), is a linear amino acid sequence in the protein backbone **(Figure 1)**. In 1973, the primary structure of proteins was proved for its role in determining their overall structure. (9,12).

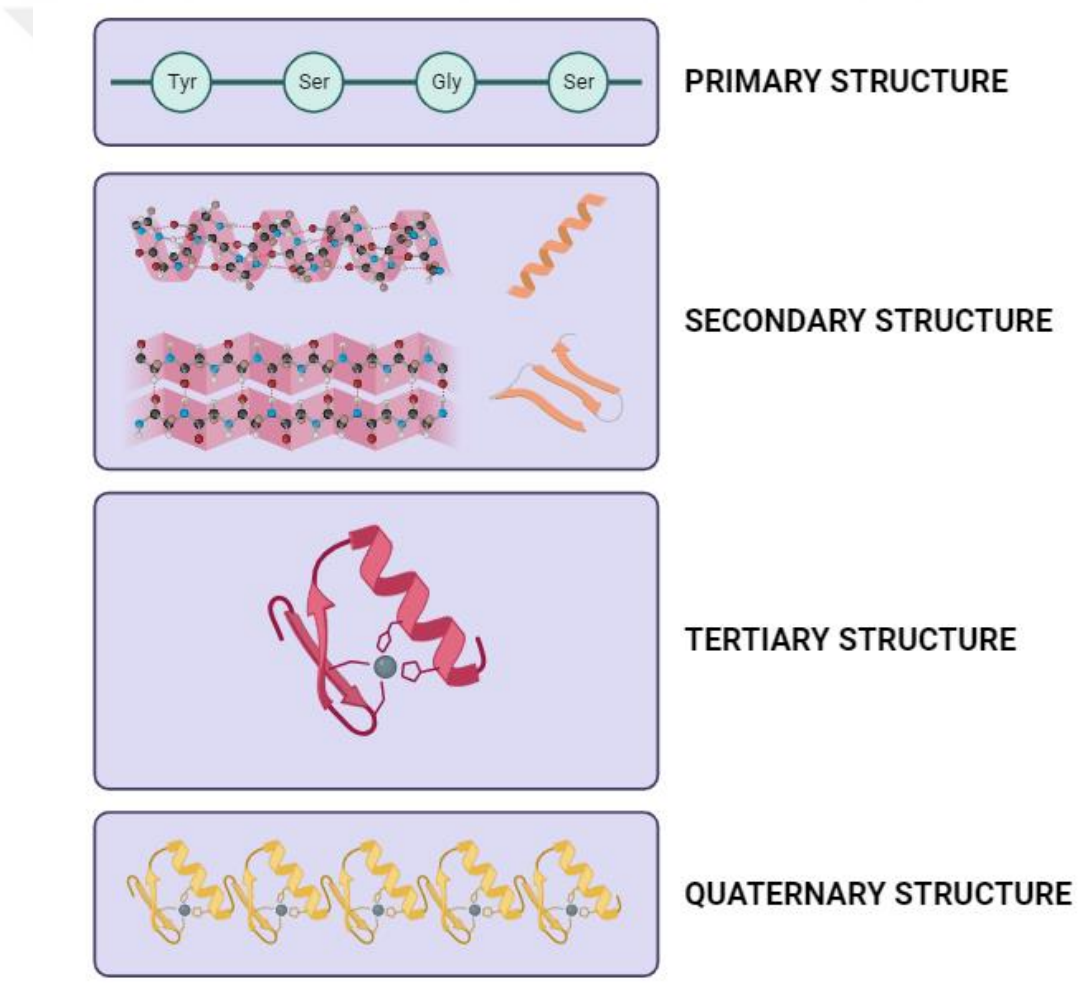

Figure 1. Four types of protein structure (created with Biorender).

However, it's worth noting that changes in the surrounding environment can sometimes disrupt the three-dimensional arrangement of a protein. As an example, ligand interactions or substrates can trigger controlled modifications, in their shape potentially leading to consequences (12). Furthermore, scientists have discovered that certain proteins possess regions that lack a defined structure. Hence while the idea of a proteins sequence dictating its structure remains powerful it's important to remember that this principle may be influenced by the local environment and may not always apply universally (13,14).

1.1.2 Secondary Structure

Protein secondary structure refers to the specific arrangement of the polypeptide backbone's atoms, excluding the side chains. The main secondary structural components are alpha helices and beta sheets, with occasional occurrences of beta turns and omega loops (9). These secondary structure elements usually take shape as an interim step before the tertiary structure is formed by the protein **(Figure 1)**.

The secondary structure is established when the amino hydrogen and carboxyl oxygen atoms along the backbone interact, through specific hydrogen bonds. Alternately, it can be characterized by the consistent arrangement of dihedral angles within a specific area of the Ramachandran plot, without considering the hydrogen bonds (9–11).

1.1.3 Tertiary Structure

The tertiary structure consists of secondary structure components. In other words, the polypeptide folds and arranges itself into an overall three-dimensional organization

(9,11). The exact spatial positioning of secondary structure components and the positions of all functional groups within a single polypeptide chain are included in the tertiary structure (Figure 1) (10,11).

The tertiary structure of a protein can often be broken down into domains, which are distinct and compact folding units typically consisting of 100 to 200 amino acid residues. While large proteins consist of multiple domains, small proteins may have only one domain in their structure (2,3). The specific arrangement on the secondary structure is called a "fold" and it creates a domain-like structure (9). Some folds are associated with particular functions, like those involved in binding to nucleotides (15).

A "structural motif" is smaller, and it is an essential unit for folds. It is basically similar to a fold, however, some structural motifs, such as β barrels can be found in a range of unrelated proteins with different variations. On the other hand, certain motifs like zinc fingers, which are associated with specific biochemical functions are highly preserved and frequently occur in protein domains that have similar functions (9).

It's essential to understand that the concept of folds, motifs, and domains are alternatively with blurred distinctions, particularly for broader and certain structural patterns is extremely conserved. (9).

1.1.4 Quaternary Structure

Proteins can exist as complexes those with a size of over 100 kDa, where multiple polypeptide chains come together (Figure 1). The specific arrangement of these units, within a protein is called as a structure. (14).

Structural Motifs	Main Features
α-Helix Bundles	These are arrangements of helices in parallel or antiparallel orientations, found in proteins like globins and cytochrome c oxidase.
Coiled-Coil	Special helix bundles where long helices intertwine to form supercoiled structures, often found in proteins with specific repeats.
β-Hairpin	A twisted, two-strand antiparallel β-sheet, less regular compared to extended β-sheets.
Greek Key and Jelly-Roll Motifs	Arrangements of four adjacent strands in β-sheets resembling Greek key patterns or forming barrel-like jelly rolls.
β-Sandwich	Two β-sheets packed face-to-face with parallel, antiparallel, or mixed strand orientations.
Mixed α/β Sandwich	Layered structures with combinations of α-helices and β-sheets, common in nucleotide-binding proteins.
β-Barrel	Barrel-like structures formed by β-strands, with varying sizes and orientations.
α/β-Barrel	Alternating α and β structures, seen in enzymes like triosephosphate isomerase.
β-Propeller	Disk-shaped arrangements of four to eight small β-sheets around a central tunnel, found in proteins with diverse functions, including enzymatic catalysis and ligand binding.

Table 1. Structural motifs and their main features. (9–11).

1.2 Functions of Proteins

The function of a protein can be predicted by identifying its binding partners. Given that a significant portion of a protein's responsibilities within living cells relies on interactions with other proteins, determining the function of even one of these interacting components can help in understanding the overall function, and importance of a protein within cellular pathways. Furthermore, their intricate cross-connectivity within cellular pathways can be mapped by exploring the network of protein-protein interactions (16–20).

The study of how proteins interact, and the identification of these interactions are essential to understanding their dynamic regulation. This knowledge is central to functional genomics and vital for drug discovery. However, predicting protein associations is not only important but also challenging, as well.

Protein-protein interactions are primarily influenced by the hydrophobic effect (21–23). Additionally, while hydrogen bonding and electrostatic interactions are essential for this process, covalent bonds also play an important role in this context (24–27).

1.3 Protein-Protein Complexes and Their Flexibility

Protein-protein interactions have an essential role in governing and executing almost all cellular processes. These interactions occur when two or more protein molecules come together and engage in a coordinated manner. The nature of these interactions can vary, ranging from temporary to lasting and they are vital for various functions like transmitting signals, catalyzing enzymes expressing genes, and providing structural support within cells (15).

Protein-protein complexes often involve binding sites and complementary shapes on the interacting proteins allowing them to fit together like interlocking puzzle pieces. However, this binding model does not like ‘lock and key’ model (28). More specifically, macromolecules undergo recognition through binding interactions, which can lead to conformational changes either before or, after the binding process. These two primary mechanisms are known as " selection" and "induced fit," respectively (28,29). Determining how these systems contribute to any binding interaction is still a difficult task of general scientific interest.

In the context of ligand binding, there are two mechanisms: conformational selection and induced fit (28). While conformational selection is flexible and can explain the kinetics of many systems, induced fit has limited applicability and can be ruled out in some cases where conformational selection is the only plausible explanation (29). This shift in perspective highlights the importance of conformational

selection as the primary mechanism for ligand binding, aligning with the idea that macromolecules exist as conformational ensembles, and the ligand selects the most suitable initial fit to trigger a biological response (28,29).

Overall, by forming these complexes, proteins can accomplish tasks that would be difficult or even impossible as units (15). Therefore, study of protein-protein complexes by more powerful methods and understand the landscape of the biomolecule is essential, in unraveling the molecular mechanisms behind cellular functioning and/or diseases (15,30).

1.4 Computational Methods for Protein-Protein Complexes

To understand the biomolecular function of the protein at a molecular level, structural details of the protein-protein interactions need to be observed in three-dimension. Since the experimental structural determination of these structures is quite challenging, the computational methods are turning into powerful techniques, in this regard (31).

Molecular modeling is an extremely invaluable tool for investigating the nature of proteins, studying biological mechanisms, and designing an effective drug (32,33). Various methods, such as structure generation and molecular docking are employed. Molecular dynamics (MD) simulations help to observe how macromolecule structures change over time providing information, about dynamic processes (34).

Another important approach in simulating systems is the use of MD simulations. These simulations explore the changes in the coordinates and structural state of a macromolecule over time (35). This trajectory or evolution is crucial in studying time-

dependent phenomena, including changes in molecular surface interactions between small molecules and glycoproteins antigen-antibody interactions, at epitope paratope interfaces the appearance and disappearance of specific channels or cavities, and the fusion of glycoprotein membranes (35,36). MD simulation was initially introduced in the 1980s to take advantage of increasing computational capabilities and has been widely used to address different aspects of protein folding problems (37). It helps us understand how protein dynamics influence catalysis and ligand binding (38).

MD simulations often fall short of capturing the full range of relevant conformational substates, particularly those associated with biological function. Enhanced sampling algorithms have emerged as a valuable tool to bridge this gap by extending the timescales and improving the exploration of complex systems' configuration space (31,39). The limitations of MD simulations in terms of timescales compared to experiments and introduces enhanced sampling methods as a solution. Enhance sampling algorithms introduces as a remedy to the limitations of MD simulations regarding the experiments comparing to the timescales (31,39).

In drug discovery and structural biology, using computational methods for studying protein-protein interactions has an exceptional role (40–42). These methods involve a wide range of techniques and algorithms aimed at understanding the interactions, dynamics, and structural conformations of protein-protein complexes at the molecular level (42).

One key approach offers molecular docking, which predicts the binding modes of proteins within a complex. On the other hand, MD simulations, allow researchers to track the dynamic behavior of protein-protein complexes over time, providing insights into their conformational changes and stability (42).

1.4.1 Molecular Dynamics (MD) Simulations

The structural analysis of the 3D structure of a protein is essential to understanding the biomolecular function of the protein at the molecular level. Since experimental structural determination of the complexity of the protein structure, is quite challenging, the computational methods rise as powerful techniques, in this regard. Especially, with the help of enhanced MD simulations, discovering the landscapes of a protein is possible (31). These simulations rely on solving Newton's equations of motion using methods that allow to simulation of the movement of atoms within a complex system (43). Methods based on thermodynamics, including MD simulations are invaluable for understanding the conformational aspects of proteins at the atomic level. By calculating the forces between atoms and continuously updating their positions and velocities MD simulations provide insights into the dynamic behavior of biological macromolecules such as proteins (43–46).

Proteins have the ability to be flexible. MD simulations provide information, about the changes happening within a protein or a protein-ligand complex (43). Knowing the functional mechanisms of proteins can be helpful in understanding the structural concept of diseases. Additionally, it can contribute to the improvement and development of proteins and peptides, as well (47,48). To understand how atoms and molecules move and behave over a period of time and investigate the dynamic nature of protein-protein complexes, MD simulations are incredibly valuable computational techniques (42,44).

MD simulations have become a tool in structural biology offering a detailed perspective, on protein-protein interactions, and mutational studies (43). Furthermore, it is extremely valuable for assessing the conformational changes, and stability of protein structure (43,49). The foundation of MD simulations revolves around the idea of integrating equations of motion. This involves calculating the forces that act on each atom determining the accelerations updating the positions and velocities of the atoms and repeating these steps in small time intervals called timesteps (34). In order to

accurately represent the system being studied researchers must take into account factors such as force fields (which describe interactions, between atoms) conditions (including positions and velocities), and the temperature at which the simulation is conducted (34,43,49). MD simulations typically run for durations ranging from nanoseconds to microseconds enabling scientists to explore atomic-level conformations and dynamics that may occur (34).

Overall, simulating the behavior of proteins and their complexes through dynamics (MD) is a valuable method that can address a wide range of inquiries. To effectively utilize this approach, it is important to consider the choice of software algorithms and force fields that govern the behavior of protein molecules during the simulation (50). MD simulations provide insights into the flexibility and mobility of different regions, within protein structures improve the accuracy of modeled structures refine three-dimensional configurations and shed light on various biological processes involving proteins.

2 BACKGROUND AND AIM OF THE STUDY

2.1 Background of Gelsolin

Gelsolin is a protein that have a role in regulating the shape and movement of cells through its ability to sever and cap actin (51). It is made up of six domains called G1 to G6 (52), which appear to have developed one after another from a globular protein with a single domain. This evolution happened because of a process where the gene was triplicated and then duplicated again (53,54). The history of gelsolins domain (G0) and these duplication events can be seen in the existence of proteins with one or three domains that have a common origin with gelsolin and carry out similar functions (55–57). The robust structural similarities are found when comparing the domains G1 and G4, G2 and G5 as well as G3 and G6 within the two halves of gelsolin. Furthermore, while the domains within each triplet from G1 to G3 and from G4, to G6 share some resemblances they also have distinct features (52).

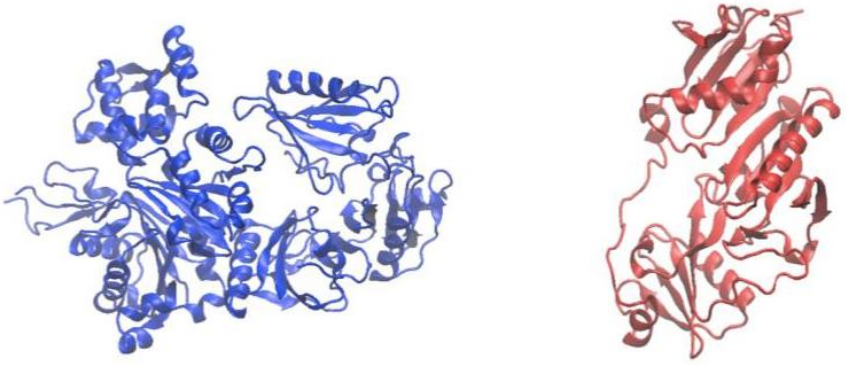

Figure 2. The crystal structure of Gelsolin G4-G6/actin complex (1H1V) (shown in blue) and Calcium-free equine plasma gelsolin (1D0N) (shown in red) (created with VMD).

2.2 Activation of Gelsolin

Activation of gelsolin is regulated by an increase in calcium ion concentration and

involves multiple levels of intervention (58–60). Its function involves the modulation of actin by binding to the growing ends of actin filaments, thereby preventing the exchange of subunits. Gelsolin has the ability to initiate the formation of new filaments (nucleation) and sever existing ones (52,58,60). When the free form of the gelsolin in the plasma is regulated by calcium, it tends to bind the actin and creates the gelsolin/actin complex (58,61). During the formation of this complex, conformational alterations occur which is why this protein aligns with the conformational selection model (61).

Level 1	Aktin bağlama yüzeylerinin aktine maruz kalmaktan sterik olarak maskelendiği altı alanın kompakt küresel düzenlemesi açılmalıdır. Bu, G2'deki aktin bağlama yerini ortaya çıkarmak için G6'dan ve C-terminal sarmalından G2 ile temasların serbest bırakılmasını gerektirir.
Level 2	Within each half of gelsolin, the third domain needs to disengage from the extended central β-sheet, shifting aside to uncover the actin-binding areas on G1 and G4 domains.
Level 3	Calcium directly engages in the interaction between gelsolin and actin by occupying sites that are mutually shared by both proteins.

Table 2. 3 level of actin binding (61).

The half of the gelsolin that is closer to the c-terminal is bound to monomeric actin. It is revealed that there are two calcium-binding sites, located in G5 and G6 domains (61). Moreover, in G1 and G4 domains, there is anothe calcium binding site known as type 1 calcium binding sites (58,61). The conservation of certain amino acid residues across the gelsolin domains suggests that type 2 sites exist in each of the domains. It was proposed that the type 2 site in domain G6 is the critical site that initiates the activation of gelsolin by releasing the tail latch that locks calcium-free gelsolin in a conformation unable to bind actin (57,61).

Since the gelsolin/actin complex is one of the difficult models in the benchmark (62), understanding the mechanisms by which gelsolin is activated in response to calcium ions is extremely invaluable (61). The structural features and interactions involved in gelsolin activation are elucidated by using computational methods. Specifically, the binding of gelsolin to actin filaments and the role of calcium ions in this process is investigated (58,61). These studies provide insights into the complex

regulation of gelsolin and its interaction with actin, which is crucial for understanding the control of cell shape changes, motility, and other cellular processes. The structure of actin is bound to the C-terminal half of gelsolin, revealing new calcium-binding sites and providing insights into the calcium control of gelsolin (57,61).

3. MATERIALS AND METHODS

3.1. Structures

Crystal structures of the gelsolin-actin complex (gelsolin G4-G6/actin complex) (PDB ID: 1H1V, resolution=2.99 Å) (61) and gelsolin plasma free-form (calcium-free equine plasma gelsolin) (PDB ID:1D0N, resolution=2.50 Å) (52) conformations of gelsolin were obtained from Protein Data Bank.

Before the simulations, the ligands and water molecules were removed from the structures, while the calcium ions were kept. No atoms with multiple occupancies were reported in either of the structures.

3.2 Input Generation

To generate input for CHARMM, CHARMM-GUI (a web-based graphical user interface for CHARMM software) was used in this study. This graphical user interface is used to simulate and analyze biomolecular systems. CHARMM Graphical User Interface (CHARMM-GUI) is a user interface designed for the CHARMM software that is commonly used in both chemistry and MD simulations. The primary objective is to investigate the behavior and structure of molecules and biomolecular systems (63).

The main advantage of CHARMMGUI lies in its approach to setting up running and analyzing molecular simulations using CHARMM. Moreover, it offers support for tasks associated with molecular modeling and simulation. These tasks include protein-

ligand docking, protein-protein docking, and protein structure refinement among others. Additionally, it seamlessly integrates with tools and databases to streamline the preparation and analysis of molecular systems (63,64).

Overall CHARMMGUI serves as a resource for researchers interested in utilizing the capabilities of the CHARMM software suite for their molecular modeling and simulation endeavors. It significantly enhances accessibility by providing a user approach, to these processes.

3.3 Molecular Dynamics (MD) Simulations

MD systems were generated by placing the gelsolin structure in a cubic water box of 10 Å in each dimension. These systems were mixed with Na and Cl ions allowing the system to be neutral. Final systems were simulated by the NAMD engine (v2.13) with the CHARMM36 force field including the correction map (65–68). Water was modeled using the TIP3P model (69).

Initially, gelsolin molecules were energy minimized for 10,000 steps and equilibrated for 0.25 ns in NPT ensemble. Then, the same relaxation steps were carried out on the whole system without fixing the protein. Next, the systems were equilibrated for 10ns in NPT ensemble at 1 bar and 298 K using Langevin dynamics temperature and Langevin piston pressure methods (70). All production simulations were performed as NPT ensembles with the periodic boundary conditions applied in all directions. A 2fs integration time was applied in all the simulations. The particle-mesh Ewald (PME) algorithm was used to compute the long-range Coulomb interactions with a grid spacing set to 1.2 Å and a cutoff radius set to 12 Å (71).

To observe the molecular structure and dynamics of gelsolin G4-G6/actin complex (1H1V) and calcium-free plasma gelsolin (1D0N), MD simulations performed for 100ns. Starting MD structures were illustrated in Figure 2. These structures were analyzed by MD simulations by using with/without boosting. For boosting, we have applied accelerated systems using with/without gaussian method.

Accelerated Molecular Dynamics (aMD) is an enhanced sampling technique that speeds up transitions between different low-energy states in biomolecules by adding a non-negative boost potential to the potential energy surface (72). Unlike conventional MD simulations, aMD does not require predefined reaction coordinates, making it suitable for exploring biomolecular conformations without prior knowledge or restraints (73,74). However, aMD suffers from energetic noise during reweighting, especially for large proteins (75).

To address this issue, Gaussian Accelerated Molecular Dynamics (GaMD) is introduced, which uses harmonic functions to create an adaptively adjusted boost potential with a Gaussian distribution. This approach reduces energetic noise and allows for accurate reweighting of simulations (76).

In this study, accelerated MD simulations were carried out both using gaussian and non-gaussian acceleration (72). Accelarated simulations were carried out using the formulations below,

$$V^*(\boldsymbol{r}) = V(\boldsymbol{r}) + \Delta V(\boldsymbol{r})$$

wherein, $V^*(\boldsymbol{r})$ is the modified potential, $V(\boldsymbol{r})$ is the potential energy of the system and $\Delta V(\boldsymbol{r})$ is the boost potential. Boost potential, $\Delta V(\boldsymbol{r})$, is calculated as follows,

$$\Delta V(\boldsymbol{r}) = \begin{cases} 0, & V(\boldsymbol{r}) \geq E \\ \dfrac{(E - V(\boldsymbol{r}))^2}{\alpha + E - V(\boldsymbol{r})}, & V(\boldsymbol{r}) < E \end{cases}$$

wherein, α is calculated by multiplying the number of residues with 0.8 and and E is the threshold energy that is calculated by summing up the average energy and 4 times of number of residues (72).

3.4 Data Analysis and Visulization

The visualization and analysis of the trajectories were carried out by different plugins of VMD (77) and UCSF Chimera v1.17 (67). The analysis of trajectories encompasses the root mean square deviation (RMSD) of the protein backbone, fluctuations of protein backbone Cα atoms, solvent accessible surface area (SASA), and radius of gyration (Rg).

4. RESULTS

4.1 Selection of Target Protein

In this thesis, the target protein was chosen from ‘Protein-Protein Docking Benchmark 5.5’ (62) to assess the performance of the MD simulations for predicting the conformations of the flexible protein-protein complexes. These benchmarks consist of sets of protein-protein complexes and the performance of computational docking algorithms and approaches is evaluated using these sets of protein-protein complexes known as protein-protein docking benchmarks. In these benchmarks, there are a number of protein complexes with three-dimensional structures, they are used to assess the efficiency and accuracy of computational methods such as MD systems and protein-protein docking software (78).

In the benchmark that, there are three levels of complexity: rigid body, medium difficulty, and difficult. From the protein set of benchmark, gelsolin-actin complex (1H1V A:G) was decided to be used as a target protein for this study. This protein is identified as 'difficult' set due to its high structural complexity. Gelsolin-actin complex consists of the combination of free-form gelsolin and actin molecules. During the formation Gelsolin-actin complex, the structures undergo an extensive conformation re-arrangement.

4.2. Displacement of Gelsolin Backbone

To assess the stability of the simulated systems, the root mean square deviation (RMSD) of gelsolin backbone atoms was computed. MD simulations indicated stabilized RMSD profiles of gelsolin conformation in the G4-G6/actin complex

(1H1V) in all of the stimulated systems. Pairwise RMSD plots showed a highly rigid gelsolin backbone in classical MD simulation (Figure 3) and Gaussian-accelerated MD simulation with dihedral boosting (Figure 4). However, the RMSD profiles of the accelerated MD simulations particularly with dual boosting setting showed highly flexible protein backbone suggesting a drastic change in the initial structure (Figure 4). Gaussian accelerated dynamics enhances the exploration of conformations by incorporating a harmonic boost potential that helps to refine the overall energy landscape of the system (76).

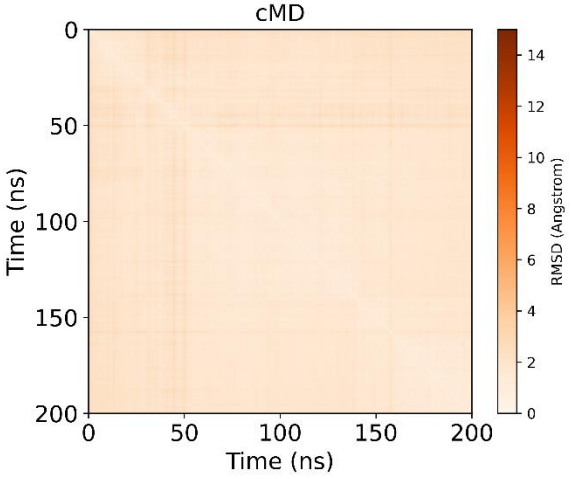

Figure 3. Pair-wise RMSD profiles of the gelsolin complex structures generated by classical MD system.

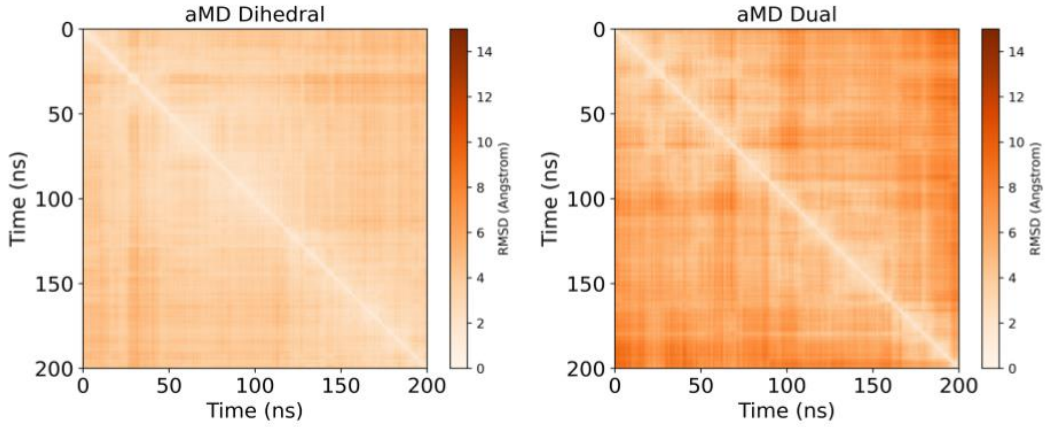

Figure 4. Pair-wise RMSD profiles of the gelsolin complex structures generated by accelerated MD with dihedral boosting and dual boosting systems.

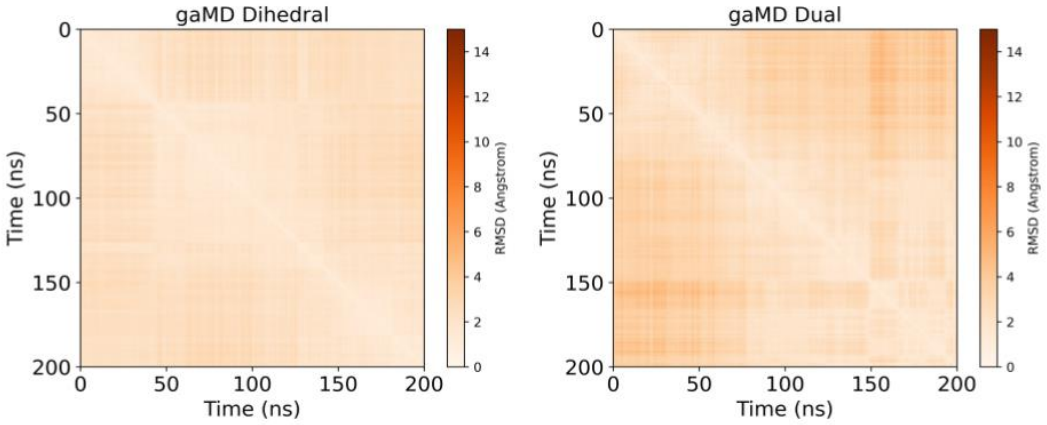

Figure 5. Pair-wise RMSD profiles of the gelsolin complex structures generated by Gaussian-accelerated MD with dihedral boosting and dual boosting systems.

Out of all MD simulations, Gaussian-accelerated MD with dihedral boosting was the closest one to the classical MD structure (Figure 5). For the rest of the simulations, Gaussian-accelerated MD with dual boosting, accelerated MD with dihedral boosting, and accelerated MD with dual boosting showed gradual increase in the backbone flexiblity.

4.3 Radius of Gyration

The radius of gyration (Rg) analysis of a protein backbone corresponds to the average distance of its constituent atoms, typically the alpha carbons (Cα), from the center of mass. A smaller Rg represents a more compact, globular structure, while a larger Rg suggests a more extended or elongated conformation (69)

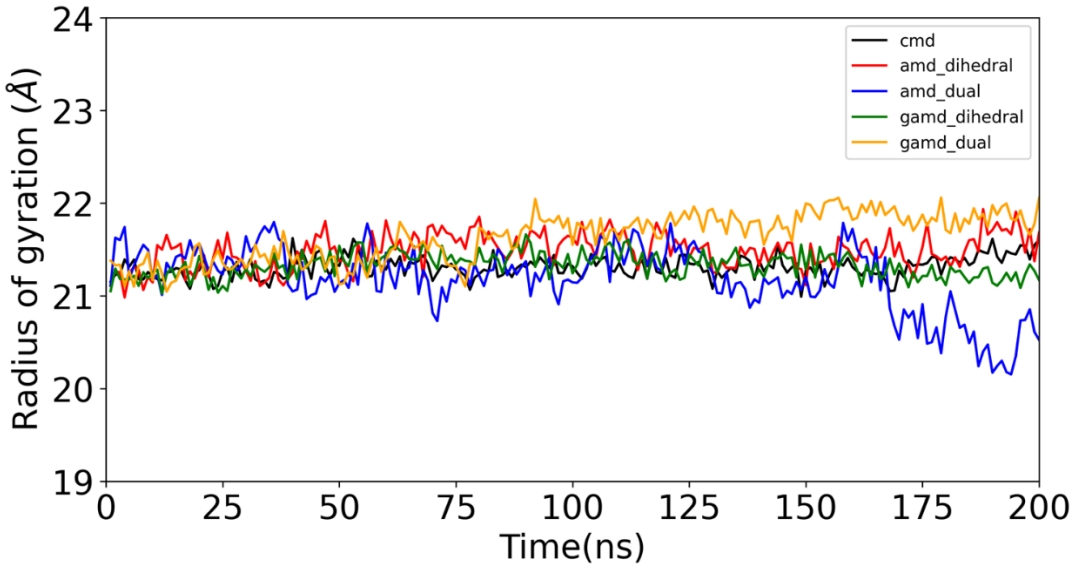

Figure 6. Radius of gyration analysis of the gelsolin complex backbone.

According to Rg analysis of MD simulations with various boosting methods, gelsolin structures did not undergo notable changes affecting their Rg. All structures displayed a Rg range of 21-22 Å with the exception of the gelsolin structure from the accelerated system with dual boosting. This system led the gelsolin structure to alter its radius, particularly reducing its Rg by 1 Å. This change is visible in Figure 6, showing a decrease in the accelerated dual boost system towards the end of 200 ns. Classical MD simulation, accelerated MD with dihedral boosting, and Gaussian-accelerated MD with dihedral boosting structure showed similar Rg values throughout the simulation of 200 ns. While the accelerated MD with dual boosting was almost stable until 170 ns, it shows a change after 170 ns. Additionally, Gaussian-accelerated MD with dual boosting led to a slightly higher Rg of gelsolin.

4.4 Surface Accessible Solvent Area of Gelsolin Structure

The Surface Accessible Solvent Area (SASA) for a protein backbone represents the total surface area of the protein that is accessible to solvent molecules, particularly

water. It serves as a crucial metric to evaluate solvent accessibility, identifying regions on the protein's surface that interact with other molecules. High SASA values often point to functional sites and protein-protein interaction regions, while lower values indicate the hydrophobic core and structural stability. Changes in SASA can signify conformational adjustments and are essential for understanding the structural and functional aspects of a protein.

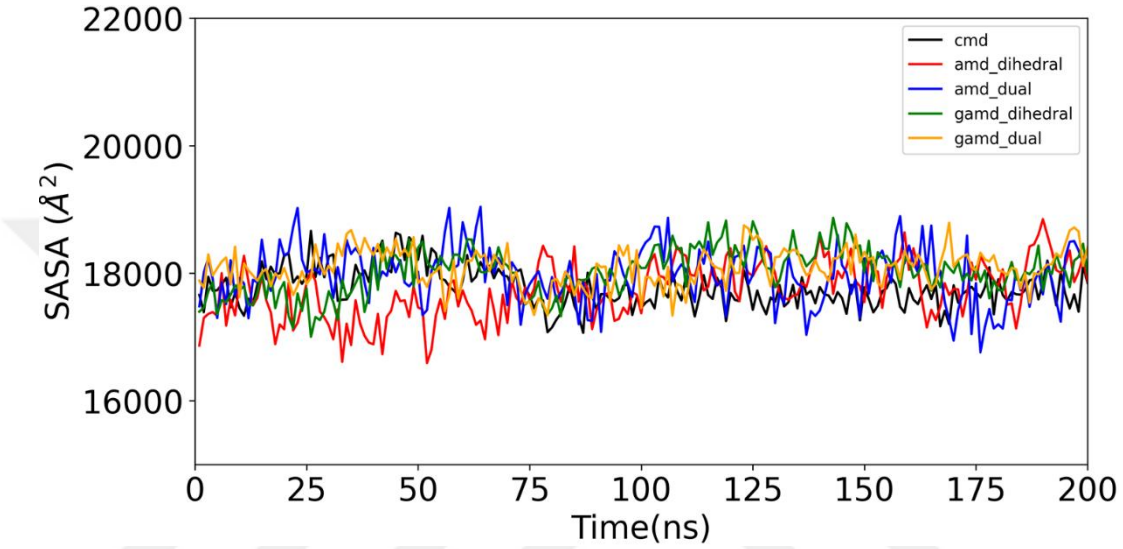

Figure 7. The change of gelsolin complex SASA analysis in the structures created by enhanced MD simulations.

According to surface accessible solvent area (SASA) analysis, notable changes are not observed. Gelsolin structures showed a flat SASA profile for all simulations including the accelarated system with dual boosting that consistently cuased changes in the RMSD and Rg measurements (Figures 4, 5, 6). Overall, three gelsolin-like domains maintained a constant SASA value around 180 nm^2. Although SASA and Rg measurements are often highly similar to each other, in line with them being measures of protein shape, gelsolin structures analyzed here showed a distinction between Rg and SASA measurements particularly for the accelerated system with dual boosting (Figures 6, 7). These kind of changes between Rg and SASA measurements could be due to the fact that the initial shape of the protein is not ideally globular.

4.5 Backbone Fluctuations of Gelsolin

To identify the mobile regions of the protein, backbone fluctuations were analyzed in terms of root-mean-square fluctuation (RMSF) of Cα atoms (Figure 8). In line with the pairwise-RMSD profiles (Figure 3, 4, 5), accelerated MD simulations with dual boosting led to the highest mobility in the backbone throughout the entire sequence.

Given the relative locations of the Gesolin-like domains; G4, G5 and G6 (Figure 8b), Cα fluctuations showed that the G5 that was found in the central region moved the least compared to the other two gelsolin-like domains. Following G5, G6 was the second domain that showed limited flexibility. On the other hand G4, was the most mobile domain that especially moved extensively in the aMD simulations regardless of the boosting type, dual or dihedral. These results that were obtained from dynamical analysis were in fact in line with the findings obtained for the static structures of gelsolin, free and actin-bound forms (52,61). As such, the central gelsolin-like domain, G5 was rigid upon actin binding. On the other hand, G4 and G6 domains went a drastic conformation change due to gelsolin binding to actin (61).

Overall, gelsolin fluctuations showed limited mobility for the simulations that were not accelarated or accelatated by using Gaussian dihedral or dual boosting. For all systems, gelsolin structure showed peaked fluctations in case of accelarated simulations for the regions corresponding to 455-460, 530-535 and 715-725. In addition to these regions, the structure showed a pronounced increase in the fluctuations particularly for the conventioanl acceralted methods and for the regions encompassing the residues 430-440 and 470-500. These differences between conventional MD simulations and accelareted simulations might suggest that gelsolin structure reponds differently to accelaration methods.

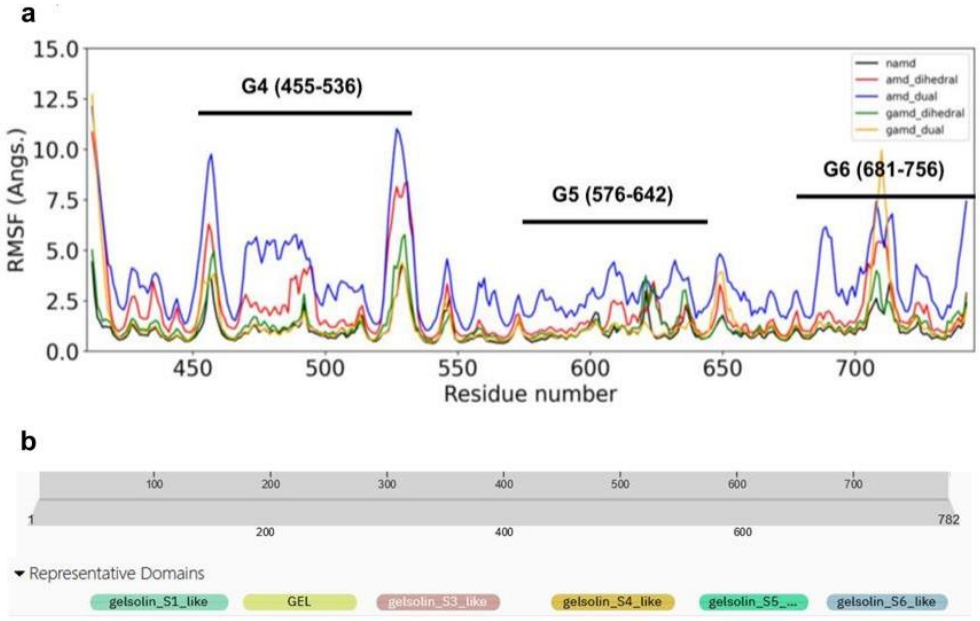

Figure 8. Domain-based mobility of gelsolin complex structures generated by various boosting methods: **(a)** RMSF analysis of the gelsolin complex structures, **(b)** Interpro (79)

4.6 Conformational Change of Gelsolin

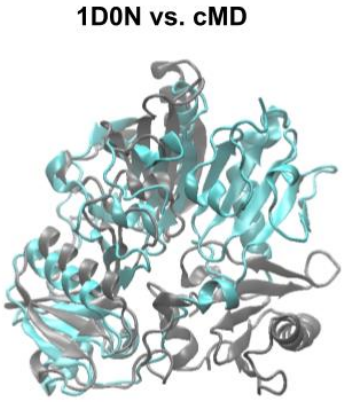

Figure 9. Aligned crystal structures of gelsolin free form (shown in cyan) and gelsolin complex structures generated by classical MD (shown in gray).

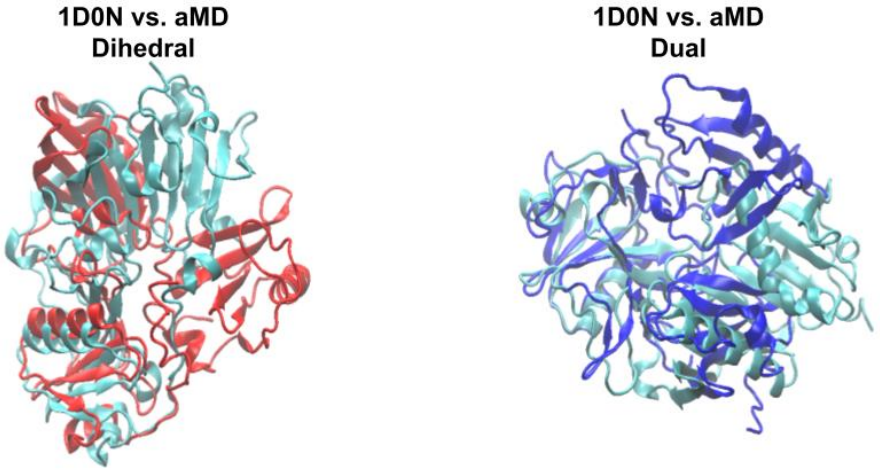

Figure 10. Aligned crystal structures of gelsolin free form (shown in cyan) and gelsolin complex structures generated by accelerated MD dihedral (shown in red) and Dual shown in blue).

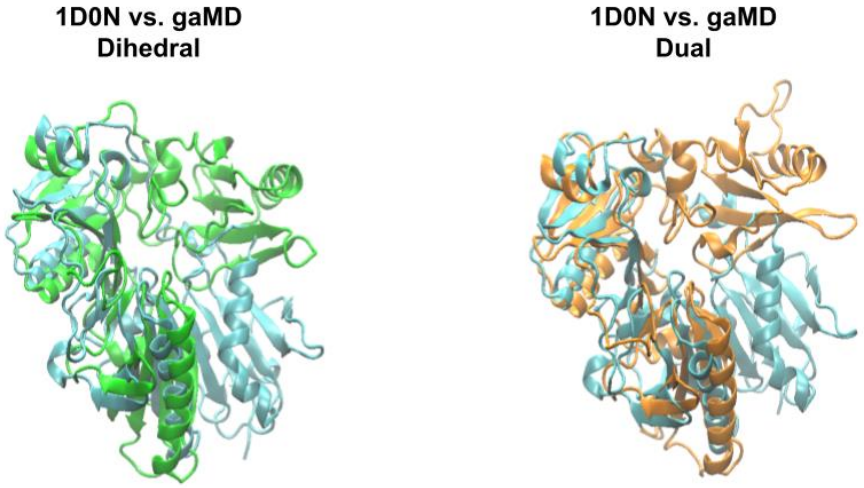

Figure 11. Aligned crystal structures of gelsolin free form (shown in cyan) and gelsolin complex structures generated by gaussian-accelerated MD dihedral (shown in green) and Dual shown in orange).

Overall, various MD simulations, both accelarated and classical simulations, showed that gelsolin in the actin-bound conformation cannot readily change its bound conformation to the unbound one wherein it is characterized by a more compact G4-G5-G6 domain organization and a close interaction of G4-G6 domains.

5 DISCUSSION AND CONCLUSION

5.1. Activation of Gelsolin

Gelsolin, an important protein in cell biology, plays a vital role in coordinating essential cellular functions like cell movement, cytokinesis and apoptosis. This versatile protein consists of six domains that form the foundation for its interaction with actin filaments. What makes gelsolin particularly fascinating is how it is tightly regulated by calcium ions. At levels of calcium, inside cells its structure hides the surfaces that bind to actin making it inactive. However when calcium concentrations rise, gelsolin undergoes structural changes that reveal these crucial actin binding sites allowing it to connect with and break apart actin filaments.

Evolutionarily, gelsolin's journey is characterized by gene duplications that transformed it from a single-domain globular protein into the multi-domain structure. This evolutionary history remains apparent in the presence of one and three-domain proteins that share a common ancestry with gelsolin and perform similar functions.

The activation of gelsolin is a process that involves multiple stages of structural modifications. These modifications include unfolding its arranged structure moving domains to expose areas that bind to actin and the direct interaction, between calcium ions, gelsolin and actin. Although the exact mechanisms governing these changes are still being researched they highlight the nature of gelsolin activation.

One fascinating aspect of gelsolin's regulation is the existence of two distinct modes of activation. Mode 1 operates at submicromolar calcium levels within the cytoplasm, where partially activated gelsolin slowly severs actin filaments through a

unique calcium-binding mechanism. This mechanism depends on a calcium binding site with affinity that is not in equilibrium and it ensures some level of severing even when the cell calcium levels are, at rest. On the other hand Mode 2 comes into play when the blood plasma calcium levels are higher and it doesn't require the unique calcium binding mechanism. Instead it achieves the severing rates directly.

In conclusion, gelsolin's intricate interplay with calcium ions and its multifaceted modes of activation, including conformational selection, make it a central figure in the dynamic world of cellular processes. Gaining an understanding of gelsolin structure, evolution and ways it is activated offers valuable insights into how it influences cell shape changes, movement, cytokinesis (cell division), and apoptosis (programmed cell death). The impressive capability of gelsolin to adjust its functions based on the levels of calcium along, with its conformational selection process highlights its crucial role in regulating the dynamics within cells.

5.2. Assessment of Enhanced MD Methods Based on the Gelsolin Dynamics

The evaluation of improved MD approaches that examine the dynamics of gelsolin signifies an advancement in our comprehension of the complex realm of cellular dynamics. Employing computational methods in the investigation of gelsolin enables us to explore its protein landscape specifically focusing on activation and the selection process, for different conformations. By harnessing the potential of dynamics (MD) techniques we can further our understanding of how gelsolin responds to varying calcium levels and undergoes structural modifications. This knowledge is not crucial for comprehending the role of gelsolin in cellular processes like shape changes, movement, cell division and programmed cell death but also holds potential for developing effective treatments, for various diseases where gelsolin plays a crucial part. These enhanced MD methods offer a pathway to uncover the hidden intricacies of cellular dynamics.

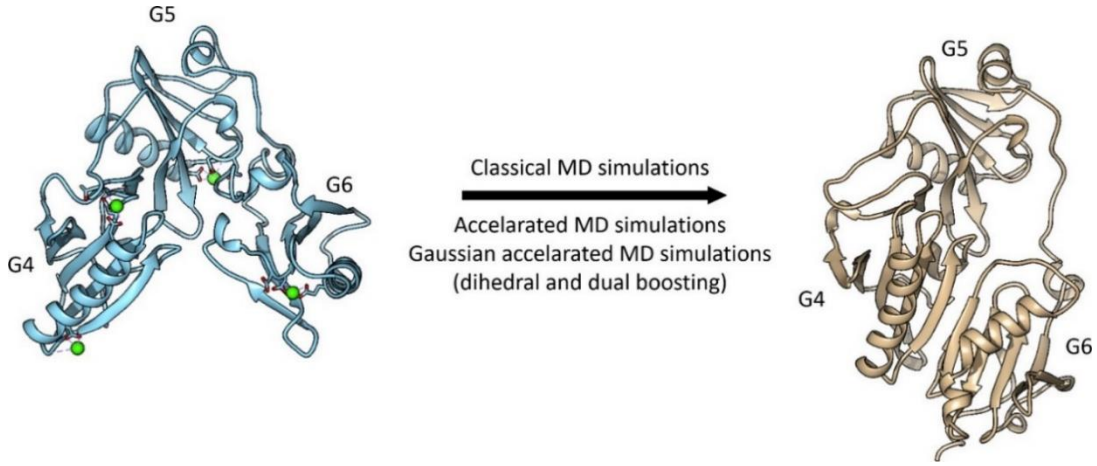

Figure 12. Schematic representations of the study: Gelsolin complex (1h1v) is shown in blue and gelsolin free form (1D0N) is shown in yellow.

As shown in the Figure 6, we assessed and compare the conformational changes on the enhanced MD structures for both complex form and the free form of the Gelsolin.

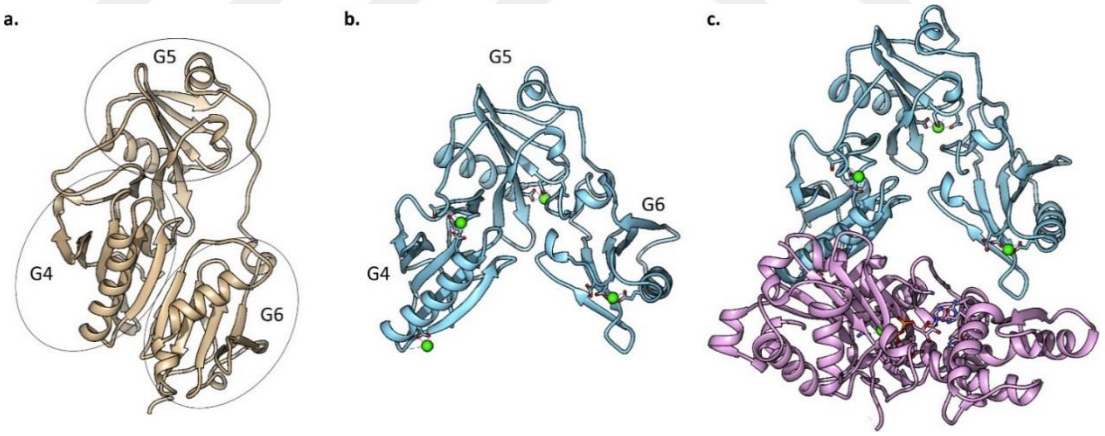

Figure 13. Schematic representations of Gelsolin free form (1D0N) is shown in yellow, Gelsolin complex (1H1V) is shown in blue , and complete structure of the Gelsolin with its 6 domains.

According to this assessment, we observed that instead of complete change in the mobility of the structures, there are domain-based changes in the specific regions.

5.3 Domain Based Mobility of Gelsolin Structure

According to enhanced MD simulation performed by accelerated and gaussian accelerated dihedral/dual boosting methods, it is obvious that the different domain has different mobility traits (Figure 8).

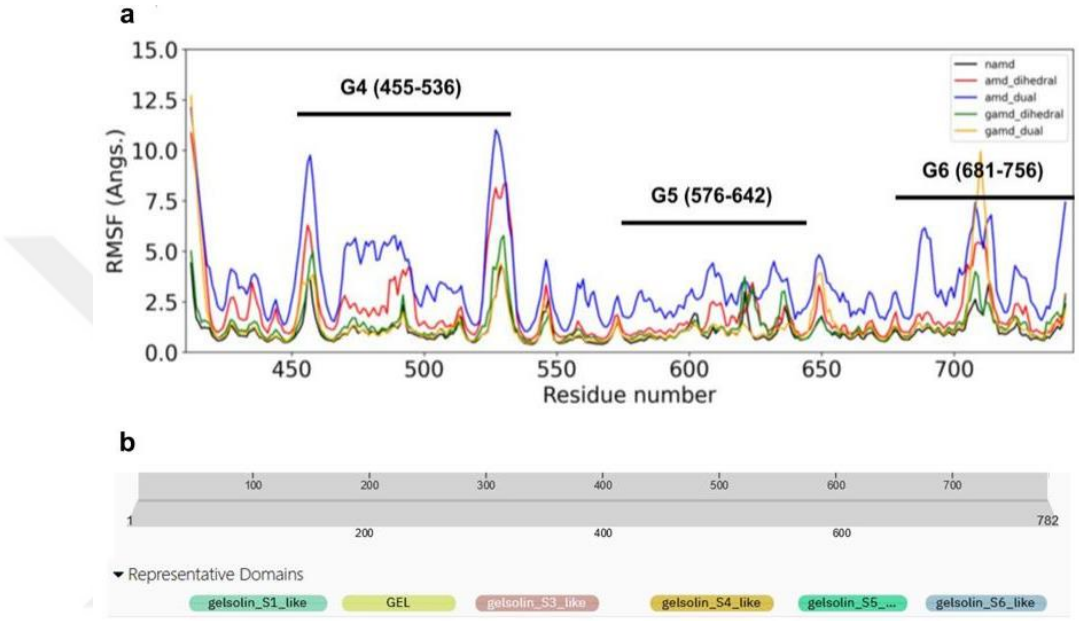

Figure 8. Domain-based mobility of gelsolin complex structures generated by various boosting methods: **(a)** RMSF analysis of the gelsolin complex structures, **(b)** Interpro (79)

Especially in the G4 domain of the Gelsolin complex, the protein structure has large movement compared to the G5 and G6 domains for all of the structures (Figure 5). The lower mobility of G5 domain for all structures, indicates that the domain is more stable compared to the other domains. However, in the G6 domain, there are highly mobile structures either such as gaussian-accelerated MD simulation with dual boosting (shown in orange), accelerated MD simulation with dual boosting (shown in blue), and accelerated MD simulation with dihedral boosting (shown in red).

Collectively, enhanced MD simulations cannot address the conformational selection of the complete structure. However, it can address domain by domain.

Therefore, it is crucial to choose appropriate parameters for enhanced MD simulation considering the domains.

5.4. Challenges and Future Directions in Gelsolin Dynamics Research

One of the challenges of this study is that the enhanced MD simulations should be performed repeatedly to be certain about their nature. Here, we have faced multiple computational power problems during the processes therefore, both complex form and free form have only one series of simulations. Hence, it is a limitation for statistical analysis of the possibility of the conformations.

The study of gelsolin dynamics poses challenges and presents future research directions. Despite the advancements brought by enhanced MD methods there are still challenges to overcome in order to achieve an understanding. One major challenge is reconciling the differing data and interpretations from studies. The intricate interplay between gelsolins domains and the complexities of calcium regulation make it a difficult system to simulate accurately. Additionally there is room for improvement in force fields simulation timescales and sampling techniques. Looking ahead, an integrated approach that combines enhanced MD methods with data holds promise for validating and refining computational models thereby increasing the reliability of our findings. Exploring gelsolin dynamics through enhanced MD methods continues to offer potential for advancing our knowledge of cellular processes and as computational techniques progress further we can expect even deeper insights, into the mysteries surrounding this crucial protein.

According to recent pan-cancer research, it is found that gelsolin tends to be downregulated in most tumors (80). The study also discovered that it shows varying levels of expression within molecular and immunological subtypes across various

cancer types (80–82). As a result, gelsolin has impacts on the prognosis of specific tumor types acts as a unique prognostic factor for certain categories of cancer, and plays an important role in cancer-related processes such, as proteoglycans chemokine signaling pathways, immune-related pathways, DNA methylation, and the cell cycle (80). Considering all these important features and functions of this protein, knowing its landscape makes it easier to understand these processes and more importantly make a contribution to the field.

Overall, knowing the nature of a protein and its landscape is promising for contributing to the field in many terms. For his purpose MD simulations are promising and extremly invaluable methods to understand the nature of a protein.

6 REFERENCES

1. Aytuna AS, Gursoy A, Keskin O. Prediction of protein-protein interactions by combining structure and sequence conservation in protein interfaces. Bioinforma Oxf Engl. 2005 Jun 15;21(12):2850–5.

2. Malmström L, Riffle M, Strauss CEM, Chivian D, Davis TN, Bonneau R, et al. Superfamily Assignments for the Yeast Proteome through Integration of Structure Prediction with the Gene Ontology. PLOS Biol. 2007 Mar 20;5(4):e76.

3. Nariai N, Kolaczyk ED, Kasif S. Probabilistic Protein Function Prediction from Heterogeneous Genome-Wide Data. PLOS ONE. 2007 Mar 28;2(3):e337.

4. Punta M, Forrest LR, Bigelow H, Kernytsky A, Liu J, Rost B. Membrane Protein Prediction Methods. Methods San Diego Calif. 2007 Apr;41(4):460–74.

5. Sharan R, Ulitsky I, Shamir R. Network-based prediction of protein function. Mol Syst Biol. 2007 Mar 13;3:88.

6. Watson JD, Sanderson S, Ezersky A, Savchenko A, Edwards A, Orengo C, et al. Towards Fully Automated Structure-Based Function Prediction In Structural Genomics: A Case Study. J Mol Biol. 2007 Apr 13;367(5):1511–22.

7. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. Protein Function. In: Molecular Biology of the Cell 4th edition [Internet]. Garland Science; 2002 [cited 2023 Oct 16]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK26911/

8. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. The Shape and Structure of Proteins. In: Molecular Biology of the Cell 4th edition [Internet]. Garland Science; 2002 [cited 2023 Oct 16]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK26830/

9. Sun PD, Foster CE, Boyington JC. Overview of Protein Structural and Functional Folds. Curr Protoc Protein Sci. 2004 Feb;35(1):1711–171189.

10. Shirley BA. Protein Stability and Folding [Internet]. Vol. 40. New Jersey: Humana Press; 1995 [cited 2023 Oct 16]. Available from: http://link.springer.com/10.1385/0896033015

11. Murphy KP. Protein Structure, Stability, and Folding [Internet]. Vol. 168. New Jersey: Humana Press; 2001 [cited 2023 Oct 16]. Available from: http://link.springer.com/10.1385/1592591930

12. Christian B. Anfinsen - Profiles in Science [Internet]. 2019 [cited 2023 Oct 16]. Protein Folding and the Thermodynamic Hypothesis, 1950-1962. Available from: https://profiles.nlm.nih.gov/spotlight/kk/feature/protein

13. Turoverov KK, Kuznetsova IM, Uversky VN. The protein kingdom extended: Ordered and intrinsically disordered proteins, their folding, supramolecular complex formation, and aggregation. Prog Biophys Mol Biol. 2010 Jun 1;102(2):73–84.

14. Uversky VN. A decade and a half of protein intrinsic disorder: Biology still waits for

physics. Protein Sci. 2013;22(6):693–724.

15. Parca L, Gherardini PF, Truglio M, Mangone I, Ferrè F, Helmer-Citterich M, et al. Identification of Nucleotide-Binding Sites in Protein Structures: A Novel Approach Based on Nucleotide Modularity. PLoS ONE [Internet]. 2012 [cited 2023 Oct 16];7(11). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3507729/

16. Ewing RM, Chu P, Elisma F, Li H, Taylor P, Climie S, et al. Large-scale mapping of human protein–protein interactions by mass spectrometry. Mol Syst Biol. 2007 Mar 13;3:89.

17. Hart GT, Ramani AK, Marcotte EM. How complete are current yeast and human protein-interaction networks? Genome Biol. 2006 Dec 1;7(11):120.

18. Kim PM, Lu LJ, Xia Y, Gerstein MB. Relating Three-Dimensional Structures to Protein Networks Provides Evolutionary Insights. Science. 2006 Dec 22;314(5807):1938–41.

19. Baudot A, Angelelli JB, Guénoche A, Jacq B, Brun C. Defining a Modular Signalling Network from the Fly Interactome. BMC Syst Biol. 2008 May 19;2:45.

20. Yip AM, Horvath S. Gene network interconnectedness and the generalized topological overlap measure. BMC Bioinformatics. 2007 Jan 24;8(1):22.

21. Tsai CJ, Lin SL, Wolfson HJ, Nussinov R. Studies of protein-protein interfaces: a statistical analysis of the hydrophobic effect. Protein Sci Publ Protein Soc. 1997 Jan;6(1):53–64.

22. Tsai CJ, Xu D, Nussinov R. Structural motifs at protein-protein interfaces: protein cores versus two-state and three-state model complexes. Protein Sci Publ Protein Soc. 1997 Sep;6(9):1793–805.

23. Young L, Jernigan RL, Covell DG. A role for surface hydrophobicity in protein-protein recognition. Protein Sci Publ Protein Soc. 1994 May;3(5):717.

24. Norel R, Sheinerman F, Petrey D, Honig B. Electrostatic contributions to protein–protein interactions: Fast energetic filters for docking and their physical basis. Protein Sci Publ Protein Soc. 2001 Nov;10(11):2147–61.

25. Sheinerman FB, Norel R, Honig B. Electrostatic aspects of protein–protein interactions. Curr Opin Struct Biol. 2000 Apr 1;10(2):153–9.

26. Sheinerman FB, Honig B. On the Role of Electrostatic Interactions in the Design of Protein–Protein Interfaces. J Mol Biol. 2002 Apr 19;318(1):161–77.

27. Xu D, Lin SL, Nussinov R. Protein binding versus protein folding: the role of hydrophilic bridges in protein associations11Edited by B. Honig. J Mol Biol. 1997 Jan 10;265(1):68–84.

28. Vogt AD, Di Cera E. Conformational selection or induced-fit? A critical appraisal of the kinetic mechanism. Biochemistry. 2012 Jul 31;51(30):5894–902.

29. Vogt AD, Di Cera E. Conformational selection is a dominant mechanism of ligand

binding. Biochemistry. 2013 Aug 27;52(34):5723–9.

30. Rao VS, Srinivas K, Sujini GN, Kumar GNS. Protein-Protein Interaction Detection: Methods and Analysis. Zhou Y, editor. Int J Proteomics. 2014 Feb 17;2014:147648.

31. Yang YI, Shao Q, Zhang J, Yang L, Gao YQ. Enhanced sampling in molecular dynamics. J Chem Phys. 2019 Aug 21;151(7):070902.

32. Pensak DA. Molecular modelling: scientific and technological boundaries. Pure Appl Chem. 1989 Jan 1;61(3):601–3.

33. Forster MJ. Molecular modelling in structural biology. Micron Oxf Engl 1993. 2002;33(4):365–84.

34. Uberuaga BP, Montalenti F, Germann TC, Voter AF. Accelerated Molecular Dynamics Methods. In: Yip S, editor. Handbook of Materials Modeling: Methods [Internet]. Dordrecht: Springer Netherlands; 2005 [cited 2023 Nov 6]. p. 629–48. Available from: https://doi.org/10.1007/978-1-4020-3286-8_32

35. Paquet E, Viktor HL. Molecular dynamics, monte carlo simulations, and langevin dynamics: a computational review. BioMed Res Int. 2015;2015:183918.

36. Salmaso V, Moro S. Bridging Molecular Docking to

mics in Exploring Ligand-Protein Recognition Process: An Overview. Front Pharmacol. 2018;9:923.

37. Adcock SA, McCammon JA. Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins. Chem Rev. 2006 May 1;106(5):1589–615.

38. Salsbury FR. Molecular dynamics simulations of protein dynamics and their relevance to drug discovery. Curr Opin Pharmacol. 2010 Dec;10(6):738–44.

39. Bernardi RC, Melo MCR, Schulten K. Enhanced sampling techniques in molecular dynamics simulations of biological systems. Biochim Biophys Acta BBA - Gen Subj. 2015 May 1;1850(5):872–7.

40. Sliwoski G, Kothiwale S, Meiler J, Lowe EW. Computational Methods in Drug Discovery. Pharmacol Rev. 2014 Jan;66(1):334–95.

41. Śledź P, Caflisch A. Protein structure-based drug design: from docking to molecular dynamics. Curr Opin Struct Biol. 2018 Feb 1;48:93–102.

42. Meng XY, Zhang HX, Mezei M, Cui M. Molecular Docking: A powerful approach for structure-based drug discovery. Curr Comput Aided Drug Des. 2011 Jun 1;7(2):146–57.

43. Frenkel D, Smit B. Chapter 4 - Molecular Dynamics Simulations. In: Frenkel D, Smit B, editors. Understanding Molecular Simulation (Second Edition) [Internet]. San Diego: Academic Press; 2002 [cited 2023 Nov 6]. p. 63–107. Available from:

https://www.sciencedirect.com/science/article/pii/B9780122673511500067

44. Salo-Ahen OMH, Alanko I, Bhadane R, Bonvin AMJJ, Honorato RV, Hossain S, et al. Molecular Dynamics Simulations in Drug Discovery and Pharmaceutical Development. Processes. 2021 Jan;9(1):71.

45. Thompson AP, Aktulga HM, Berger R, Bolintineanu DS, Brown WM, Crozier PS, et al. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. Comput Phys Commun. 2022 Feb 1;271:108171.

46. Hospital A, Goñi JR, Orozco M, Gelpí JL. Molecular dynamics simulations: advances and applications. Adv Appl Bioinforma Chem AABC. 2015 Nov 19;8:37–47.

47. Zaidi ZH, Smith DL, editors. Protein Structure — Function Relationship [Internet]. Boston, MA: Springer US; 1996 [cited 2023 Nov 6]. Available from: http://link.springer.com/10.1007/978-1-4613-0359-6

48. Rauer C, Sen N, Waman VP, Abbasian M, Orengo CA. Computational approaches to predict protein functional families and functional sites. Curr Opin Struct Biol. 2021 Oct 1;70:108–22.

49. Singh S, Singh VK. Molecular Dynamics Simulation: Methods and Application. In: Singh DB, Tripathi T, editors. Frontiers in Protein Structure, Function, and Dynamics [Internet]. Singapore: Springer; 2020 [cited 2023 Nov 6]. p. 213–38. Available from: https://doi.org/10.1007/978-981-15-5530-5_9

50. Collier TA, Piggot TJ, Allison JR. Molecular Dynamics Simulation of Proteins. Methods Mol Biol Clifton NJ. 2020;2073:311–27.

51. Sun HQ, Yamamoto M, Mejillano M, Yin HL. Gelsolin, a Multifunctional Actin Regulatory Protein*. J Biol Chem. 1999 Nov 19;274(47):33179–82.

52. Burtnick LD, Koepf EK, Grimes J, Jones EY, Stuart DI, McLaughlin PJ, et al. The Crystal Structure of Plasma Gelsolin: Implications for Actin Severing, Capping, and Nucleation. Cell. 1997 Aug 22;90(4):661–70.

53. Way M, Weeds A. Nucleotide sequence of pig plasma gelsolin: Comparison of protein sequence with human gelsolin and other actin-severing proteins shows strong homologies and evidence for large internal repeats. J Mol Biol. 1988 Oct 20;203(4):1127–33.

54. Van Troys M, Dewitte D, Verschelde JL, Goethals M, Vandekerckhove J, Ampe C. Analogous F-actin Binding by Cofilin and Gelsolin Segment 2 Substantiates Their Structural Relationship*. J Biol Chem. 1997 Dec 26;272(52):32750–8.

55. Van Troys M, Vandekerckhove J, Ampe C. Structural modules in actin-binding proteins: towards a new classification. Biochim Biophys Acta BBA - Mol Cell Res. 1999 Jan 11;1448(3):323–48.

56. Puius YA, Mahoney NM, Almo SC. The modular structure of actin-regulatory proteins. Curr Opin Cell Biol. 1998 Feb 1;10(1):23–34.

57. Lin KM, Mejillano M, Yin HL. Ca2+ Regulation of Gelsolin by Its C-terminal Tail*. J Biol Chem. 2000 Sep 8;275(36):27746–52.

58. Kinosian HJ, Newman J, Lincoln B, Selden LA, Gershman LC, Estes JE. Ca2+ Regulation of Gelsolin Activity: Binding and Severing of F-actin. Biophys J. 1998 Dec 1;75(6):3101–9.

59. Allen PG, Janmey PA. Gelsolin displaces phalloidin from actin filaments. A new fluorescence method shows that both Ca2+ and Mg2+ affect the rate at which gelsolin severs F-actin. J Biol Chem. 1994 Dec 30;269(52):32916–23.

60. Weeds AG, Gooch J, McLaughlin P, Pope B, Bengtsdotter M, Karlsson R. Identification of the trapped calcium in the gelsolin segment 1—actincomplex: Implications for the role of calcium in the control of gelsolin activity. FEBS Lett. 1995 Mar 6;360(3):227–30.

61. Choe H, Burtnick LD, Mejillano M, Yin HL, Robinson RC, Choe S. The calcium activation of gelsolin: insights from the 3A structure of the G4-G6/actin complex. J Mol Biol. 2002 Dec 6;324(4):691–702.

62. Vreven T, Moal IH, Vangone A, Pierce BG, Kastritis PL, Torchala M, et al. Updates to the Integrated Protein-Protein Interaction Benchmarks: Docking Benchmark Version 5 and Affinity Benchmark Version 2. J Mol Biol. 2015 Sep 25;427(19):3031–41.

63. Jo S, Kim T, Iyer VG, Im W. CHARMM-GUI: A web-based graphical user interface for CHARMM. J Comput Chem. 2008;29(11):1859–65.

64. Suh D, Feng S, Lee H, Zhang H, Park SJ, Kim S, et al. CHARMM-GUI Enhanced Sampler for various collective variables and enhanced sampling methods. Protein Sci. 2022;31(11):e4446.

65. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. J Comput Chem. 2005 Dec;26(16):1781–802.

66. MacKerell ADJr, Bashford D, Bellott M, Dunbrack RLJr, Evanseck JD, Field MJ, et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. J Phys Chem B. 1998 Apr 1;102(18):3586–616.

67. Huang J, MacKerell AD. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. J Comput Chem. 2013 Sep 30;34(25):2135–45.

68. Brooks BR, Brooks CL, MacKerell AD, Nilsson L, Petrella RJ, Roux B, et al. CHARMM: The Biomolecular Simulation Program. J Comput Chem. 2009 Jul 30;30(10):1545–614.

69. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys. 1983 Jul 15;79(2):926–35.

70. Pastor RW, Brooks BR, Szabo A. An analysis of the accuracy of Langevin and molecular dynamics algorithms. Mol Phys. 1988 Jan 1;65:1409–19.

71. Darden T, York D, Pedersen L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. J Chem Phys. 1993 Jun 15;98(12):10089–92.

72. Hamelberg D, Mongan J, McCammon JA. Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. J Chem Phys. 2004 Jun 22;120(24):11919–29.

73. Wereszczynski J, McCammon JA. Nucleotide-dependent mechanism of Get3 as elucidated from free energy calculations. Proc Natl Acad Sci U S A. 2012 May 15;109(20):7759–64.

74. Pierce LCT, Salomon-Ferrer R, Augusto F. de Oliveira C, McCammon JA, Walker RC. Routine Access to Millisecond Time Scale Events with Accelerated Molecular Dynamics. J Chem Theory Comput. 2012 Sep 11;8(9):2997–3002.

75. Shen T, Hamelberg D. A statistical analysis of the precision of reweighting-based simulations. J Chem Phys. 2008 Jul 21;129(3):034103.

76. Miao Y, Feher VA, McCammon JA. Gaussian Accelerated Molecular Dynamics: Unconstrained Enhanced Sampling and Free Energy Calculation. J Chem Theory Comput. 2015 Aug 11;11(8):3584–95.

77. Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. J Mol Graph. 1996 Feb 1;14(1):33–8.

78. Vakser IA. Protein-Protein Docking: From Interaction to Interactome. Biophys J. 2014 Oct 21;107(8):1785–93.

79. Gelsolin (P06396) - protein - InterPro [Internet]. [cited 2023 Nov 6]. Available from: https://www.ebi.ac.uk/interpro/protein/UniProt/P06396/

80. Wang Y, Bi X, Luo Z, Wang H, Ismtula D, Guo C. Gelsolin: A comprehensive pan-cancer analysis of potential prognosis, diagnostic, and immune biomarkers. Front Genet [Internet]. 2023 [cited 2023 Oct 26];14. Available from: https://www.frontiersin.org/articles/10.3389/fgene.2023.1093163

81. Chen MH, Chang SC, Lin PC, Yang SH, Lin CC, Lan YT, et al. Combined Microsatellite Instability and Elevated Microsatellite Alterations at Selected Tetranucleotide Repeats (EMAST) Might Be a More Promising Immune Biomarker in Colorectal Cancer. The Oncologist. 2019 Dec;24(12):1534–42.

82. Kim JC, Ha YJ, Tak KH, Roh SA, Kwon YH, Kim CW, et al. Opposite functions of GSN and OAS2 on colorectal cancer metastasis, mediating perineural and lymphovascular invasion, respectively. PloS One. 2018;13(8):e0202856.

7 CURRICULUM VITAE