UNIVERSITÀ DEGLI STUDI DI MILANO

DIPARTIMENTO DI CHIMICA

Doctoral School of Chemical Sciences

XXVI cycle

Academic year 2012-2013

In silico Study of Protein-Protein Interaction Stabilization and Mechanical Force Application on Biomolecules

Tutor: Dr. Stefano Pieraccini

Co-Tutor: Prof. Maurizio Sironi

Shilpi CHAURASIA

Dedicated to the noble soul

...who perished.

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my tutor Dr. Stefano Pieraccini for his motivation and guidance whilst giving me the freedom to work in my own way. His guidance helped me in all the three years of PhD… not only in the academics, but also in the bureaucracies related to a foreign student.

I would also like to thank my co-tutor Prof. Maurizio Sironi for valuable discussions.

I thank my fellow lab mates in Molecular Modelling Group. Esp. Stefano and Riccardo for the discussions related to projects, courses and seminars, but above all…. for the time we spent together. Thanks to Simone for the valuable discussions. They are great friends. An individual thanks goes to Marina for her company and chatting over many cups of coffee.

I am blessed with a wonderful sister, Rishika, …my best friend, who always stand with me whenever I need her. Special thanks to her, also to my brother, Shivam. You guys are my lifeline. I would never forget to mention my friends, especially Archana, Thanks for being with me. Rajiv, Tanya and Pushkar. We had a wonderful time together…doesn't matter where we are, in Genova… Milan… Trento. Thanks to you all!

I have no words to express my sincere thanks to my mother, Mrs. Ranjana Chaurasia and my father, Mr. M.K. Chaurasia for supporting me throughout my life and having faith in me. Special thanks to my in-laws, they have been really very supportive and friendly.

Last but not the least; I would like to thanks my husband, my best friend, my soul-mate, Amitava for his love, encouragement and support.

Thank you.

TABLE OF CONTENTS

THESIS OVERVIEW

Targeting protein-protein interactions is a challenging task in drug discovery process. Despite the challenges, several studies have provided evidences for the development of small molecules modulating protein-protein interactions. In Part I, it is demonstrated that how a small molecule can induce the formation of an otherwise unstable protein-protein complex. A study of the stabilization of a FKBP12-FRB complex by a small molecule rapamycin is presented. The stability of the complex is analyzed and its interactions are characterized at the atomic level by performing free energy calculations and computational alanine scanning. It is shown that rapamycin stabilizes the complex by acting as a bridge between the two proteins; and the complex is stable only in the presence of rapamycin. The reported results and the good performance of standard molecular modeling techniques in describing the model system can be interesting not only in the design and development of improved molecules acting as FKBP12–FRB protein interaction stabilizers, but also in the somehow neglected study of protein-protein interactions stabilizers in general.

In Part II, studies regarding computational modeling of the application of mechanical force to biomolecules is presented. This part is further divided into two chapters since the investigations have been performed on two biological systems. In the first chapter of Part II (chapter 6), it is described that how the osmolyte molecules affect the mechanical unfolding of a peptide. The mechanical unfolding of peptide has been performed by using Steered Molecular Dynamics. In this study, the effect of four different osmolytes on the free energy difference between the folded and the denatured state have been calculated. The observed trend mirrors the expected behavior of the studied osmolytes and unfolding pathways analysis allows an insight into the mechanism of action of osmolytes.

After the successful application of Steered molecular dynamics technique on the β-hairpin peptide, the same is applied on tubulin heterodimers for the in-depth study of the lateral and longitudinal interactions which are responsible for the stability and dynamics of the microtubules. In the other chapter of Part II (chapter 7), these interactions are studied with the help of mechanical dissociation of the tubulin heterodimers. These studies have allowed the identification of the critical interactions responsible for the binding of tubulin heterodimers laterally as well as longitudinally. The observations obtained could be important for the design of compounds that target these interactions and acts as microtubule inhibitors or stabilizers.

PART I

Molecular Insights into the Stabilization of Protein-Protein Interactions with Small Molecule:

The FKBP12-Rapamycin-FRB Case Study

CHAPTER 1

PROTEIN-PROTEIN INTERACTIONS

1.1 Introduction: Protein-protein Interactions

The basic building block of all living organism is DNA (Deoxyribonucleic acid) which is transcribed into RNA (Ribonucleic acid), which in turn translated to proteins. The proteins are responsible for almost every task of the cellular life. The haploid human genome is composed of three billion base pairs and about 20,000 to 30,000 protein-coding genes which is less than 2% of the whole genome. Protein-coding genes are the most widely studied and best understood components of a genome. These protein coding genes may encode over 500,000 proteins. Proteins are responsible for vast array of biological functions. It has been pointed out that approximately 80% of the proteins in a cell do not stand alone, but rather interact with other proteins to perform a biological function. The complexity of the interaction increases when the cell is in the condition of stress and disease (1). Cellular systems can be considered as complex networks where the molecules are the nodes and the associations between them are the edges of the network. Collectively, all types of interactions can be referred as an interactome. For reference, the size of the binary human interactome has been estimated to comprise about 130,000 Protein-Protein Interactions, of which only 8% have been identified till now. So, there is a great need of enhancing the knowledge in this field (2).

Proteins mutually interact with each other for specific biological functions. Protein–protein Interactions **(PPIs)** are defined as physical interaction between two or more proteins that occur within a cell as well as between cells. PPIs are parts of PPI networks that control the flow of information both within and between biological processes (3). **Figure 1** illustrates a case of PPI, where Interleukin-2 and alpha subunit of Interleukin's receptor interact to form a complex which in turn is responsible for performing a specific biological function.

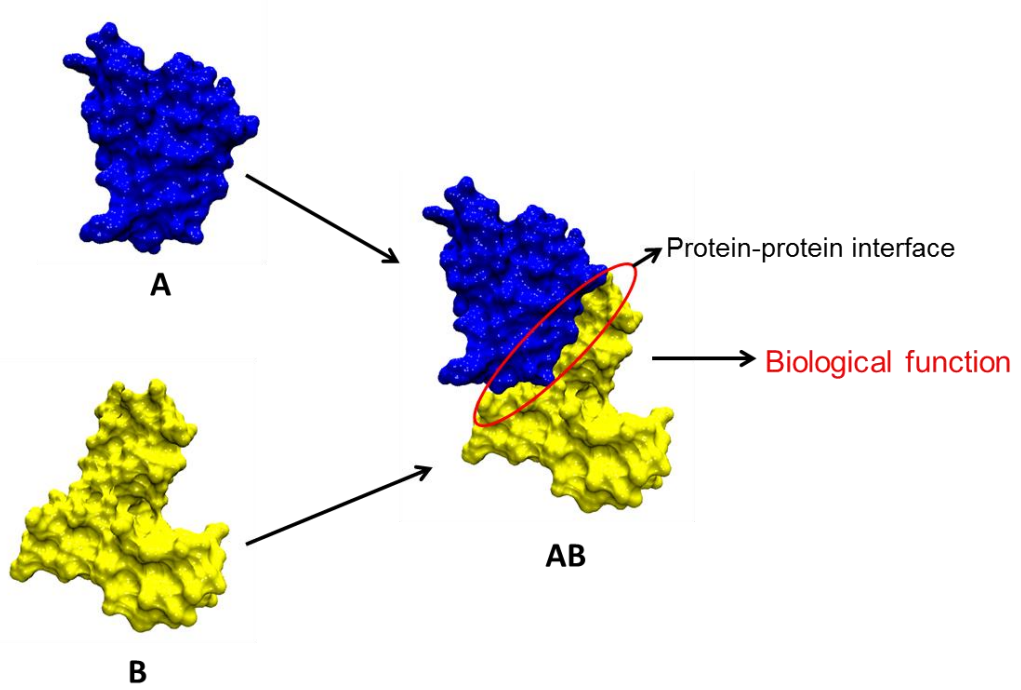

Figure 1: The illustration of PPI which represents the two participant proteins **A** and **B**. They interact to form a complex **AB** which is responsible for performing specific biological function. Here, **AB** represents the structure of the complex of interleukin-2 with its alpha receptor. **A** represents Interleukin-2 and **B** represents alpha subunit of its alpha receptor. The area marked with red represents the interface of the complex (protein-protein interface).

PPIs are essential components of almost all cellular processes, which leads to a great deal of biological functions like cell proliferation, growth differentiation,

modulating or initiating signal transduction, gene transcription, providing cytoskeletal stability, promoting cellular replication or death (2). To perform a specific biological function, PPIs need to be at the right location and at the right time. Any disruption in the PPI network has been shown to result in various diseases including cancer (4). Thus, understanding how the proteins interact with each other has allowed us to elucidate the molecular basis of different diseases. This suggests the possibility of targeting PPIs for drug discovery, which, as we shall explain later, is more challenging than the traditional drug discovery, and that is why; there is a great necessity for the detailed understanding of the network of these interactions at the atomic level.

1.2 Structural characteristics of Protein–protein Interfaces

Proteins interact with each other through their surfaces, thus a great focus has been made to the analysis of the geometrical, structural and dynamical aspects of protein surfaces mediating PPIs.

1.2.1 Types of Protein-protein Interactions

PPIs can be classified as permanent or transient according to their binding lifetime.

Permanent PPIs: In permanent/obligate type of PPIs, proteins form multi-subunit complexes and the participant proteins are permanently bound with one another. For example, hemoglobin and RNA polymerase form stable complex. Participant proteins of this group of PPIs bind more strongly and their interfaces are larger and exhibit high shape complementarity. They have tighter packing and very few water molecules are trapped between the participant proteins. These interfaces contain an over-abundance of non-polar amino acids, especially near the center and contact surfaces (5).

Transient PPIs: In transient/non-obligate type of PPIs, the interactions are temporary in nature, i.e. they associate and dissociate under particular physiological conditions like phosphorylation and conformational changes. The strength of these interactions is strongly dependent on the modulation by other molecules like cofactors. Transient interactions can be fast or slow, strong or weak. Transient PPIs does not have much higher percentage of non-polar amino acids (3).

The main interactions that are responsible for the binding of proteins are hydrophobic in nature. Polar and electrostatic interactions (hydrogen bonds and salt-bridges) also play important roles. When the proteins bind, the components of the complex bury part of their accessible surface area, which is inaccessible to the solvent due to protein-protein contact. This surface is referred as the *interface*. The surface area of the interface is usually large (approx. 1500-3000 $Å^2$) (6). This surface corresponds to the protein portion that is excluded from solvent upon binding, and is also known as *buried surface*. The interfaces is largely made up of hydrophobic amino acids like Tyr, His, Trp, Phe (7). This is especially evident in stable PPIs where protein interfaces have a composition very similar to protein interiors (8), (9), while in transient complexes this feature is less evident (7), (10). Charged residues are less abundant on the interfaces except Arg which is found in abundance (7).

1.2.2 Hot spots: The regions on the Protein-Protein interface responsible for binding

Experimentally, it has been observed that not all the residues that are present at the protein-protein interface contribute equally to the binding free energy. An approach called, Alanine Scanning has widely been used to predict those residues at the interface that contribute to the binding free energy. In this approach, the residues located at the

protein-protein interfaces are mutated into alanine and the difference in the binding free energy (ΔΔG) is measured between wild type and mutated complex. In fact, only a small fraction of residues at the interface, around 10%, shows ΔΔG values greater than 2 kcal/mol. (11), (12). This method will be discussed in detail in the following section. These small subsets of residues at the protein-protein interface that contribute to the binding energy are known as 'hot spots'. The interface is composed of a core and a rim. The residues in the core region are buried and are surrounded by the residues of the rim region. The core region is enriched with hydrophobic residues and the amino acid composition of rim region is same as that of rest of the protein (13). Keskin and coworkers have shown that there is a linear correlation between the numbers of hot-spots and the interface size (14). The residues around the hot spot regions are more tightly packed as compared to rest of the interface.

1.2.3 Types of Atoms at the protein-protein interface

Recognition sites at the interface are composed of various types of atoms: interface atoms, contact atoms and buried atoms (7)

(i) *Interface atoms:* All those atoms that lose solvent accessible surface on the formation of the complex are interface atoms. Type A, B and C shown in figure 2 are interface atoms. Not all the interface atoms make actual Van der Waal's (vdw) contact across the interface.

(ii) *Contact atoms:* These atoms are the subset of the interface atoms. These atoms actually make vdw contact across the interface. It can be further classified to fully buried contact atoms and partially buried contact atoms.

Fully Buried contact atoms: Those atoms that are accessible in the free components but have zero accessible surface area in the complex. (Type B shown in **Figure 2**)

Partially Buried contact atoms: Those atoms that are partially accessible to solvent and makes vdw contact across the interface. (Type C shown in **Figure 2**)

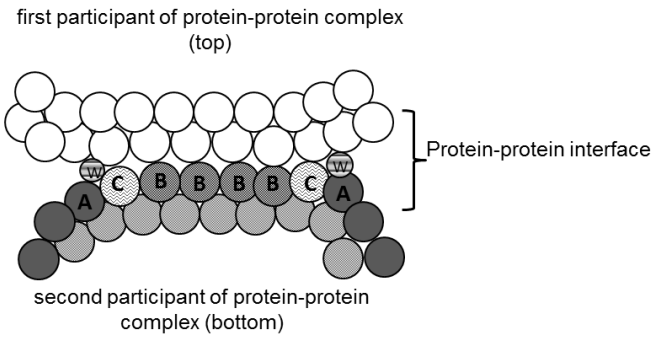

Figure 2: Figure represents the class of interface atoms in a protein-protein complex. One protein partner is shown at the top and the other is shown at the bottom. For clarity, atoms are specified for the bottom one only. Type B represents completely buried atoms. Type A and C retain partial solvent accessibility. Type B and C are contact atoms that make vdw contacts with the top protein partner. Type A represents the peripheral interface atoms. The sphere with W represents water molecule. The figure is reprinted from "The atomic structure of protein-protein recognition sites," L. Jo Conte et al, 1999, J. Mol. Biol.,285, p. 2182 (7)

One-half of the total interface atoms are contact atoms and one-third are buried atoms. The remaining atoms belong to the peripheral interface atoms of type A. The number of interface atoms scales linearly with the interface area (7). Bogan and Thorn, through experimental alanine scanning mutagenesis studies have shown that the hot spots are not homogenously distributed over the protein-protein interface, rather they are present in form of clusters. They have found that the hot spots are clustered in the center of the interface rather than at the rim (15). Hot spots have a tendency to couple across the

binding interfaces. **Figure 3** illustrates the hot spot regions on the protein-protein interface.

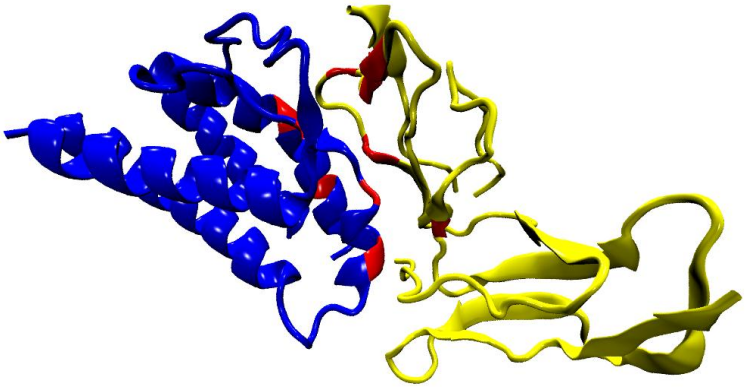

Figure 3: The figure represents the ribbon diagram of a protein-protein complex showing hot spots at the interface. Here, structure of the complex of interleukin-2 (blue) with its alpha receptor (yellow) is shown and the hot spots are represented in red.

1.3 Protein-Protein Interactions in drug design

1.3.1 PPIs as Drug Targets:

PPIs are an attractive emerging class of drug targets. Disruption or deregulation of the PPIs leads to a number of diseases like cancer and neurodegenerative diseases (16). This may happen through the loss of essential interaction, or through the formation of the protein complexes at an inappropriate time or location. Consequently, so many researches are going on, in order to study, how the proteins communicate with each other at the molecular basis. These researches are focused on the PPIs modulation (inhibition or stabilization) to develop novel therapeutic approaches especially in the anticancer field. There is a need to identify and assess useful and novel approaches for the design

and identification of "small molecule PPI modulators". These small molecule PPI modulators are a natural or synthetic product having a molecular weight of 500-900 Da, which has target selectivity and good binding affinity. They regulate PPI either through inhibition or through stabilization (17).

PPIs can have different extent of druggability. Druggability of PPI depends on the structure and physiochemical characteristics of the interface. London et al (18) demonstrated that the druggable PPIs are dominated by 'hot segments'. 'Hot segments' is an extension to the concept of hot spots. It can be explained as a continuous peptide stretch within a protein that contributes most to the binding (19).

It has been estimated that about 10% of the human genome is involved in some type of disease state, resulting in corresponding 3000 potential drug targets (20). Sugaya et al has developed a SVM (support vector machine) based method for the prediction of druggable PPIs (21)

1.3.2 PPIs: Challenges

PPIs are quite different from protein-ligand or enzyme-inhibitor interactions and designing small molecule PPI modulators is a challenging task. These challenges include the lack of binding pockets at the protein-protein interface. Interface is usually large (approx. 1500-3000 $Å^2$), flat and featureless. So, it is more challenging to obtain modulators for flat interfaces than for interfaces with well-defined pockets (6). Moreover, unlike enzymes, protein-protein interfaces are more diverse and the binding sites do not appear to be preserved. So, each protein interface is unique. (13), (22). Although, protein interfaces are larger in size, but the cluster of hot spots are distributed over the entire interface, making it possible to design low molecular weight compounds targeting hot-spot clusters. Small molecule discovered till now that inhibits/stabilizes PPI, have shown that PPIs can act as drug targets. For instance, proteins Bcl-2 and Bcl

XL (regulators of apoptotic cell death), and HDM2 (an oncoprotein with both p53-dependent and independent functions) are examples of signaling proteins where PPIs have been targeted and small molecules have entered clinical trials (23), (24).

It has been established that for the modulation of the PPI, a molecule is not required to bind with the large area of the protein's interface. In fact, it interacts with a small subset of amino acids.

1.3.3 Modulation of PPI

In principle, a protein-protein interaction can be exploited as a drug target either by reducing its strength, thus impairing protein complex formation, or by increasing its strength to stabilize the complex. It depends on the nature of the process mediated by the PPI and on the therapeutic aim whichever is the suitable strategy to be adopted in each case. So far, most efforts have been devoted to the design and development of PPIs inhibitors (25), (26) but recently, also PPIs stabilizing agents have gained an increasing attention (2).

PPI Inhibition: PPI inhibition is a process where a small molecule competes with the original protein partner of the PPI resulting in the disruption of the function. PPI inhibitors can target two sites: hot spots or allosteric sites. **Figure 4a** represents hot-spots approach to PPI inhibition. **Figure 4b** represents allosteric regulation mechanism of PPI inhibition. In allosteric mechanism, the inhibitor molecule binds at the distant site from the PP interface and induces structural modifications in the protein, resulting in inability of the binding of two protein partners (22), (27). Several studies have been performed on the direct and allosteric mechanism of PPI inhibition. For instance, P-arylthio cinnamides inhibits allosterically the interaction between Leukocyte function-associated antigen -1 (LFA-1) and intracellular adhesion molecule-1 (ICAM-1) (28). Drugs like Vinca alkaloids and colchicine directly inhibits microtubule polymerization by targeting tubulin-tubulin interaction (29). Generally, such small molecule inhibitors can cover an average of approximately 300-1000 $Å^2$ of protein-protein interaction surface. As the

surface area of the hot spot is approximately 600 $Å^2$, so these molecules can represent good starting points for developing better therapeutic agents (17).

1.3.4 Strategies to identify small molecule PPI modulators

There are three main classes of PPI modulators which can be identified: humanized monoclonal antibodies, small molecules and; peptides and peptidomimetics derived from PP interfaces. Antibodies are attractive modulators because of their high target selectivity. But they have drawbacks like poor cell permeability and they are not orally bioavailable. On the other hand, peptides derived from PP interface are better modulators as they are considered as the starting point for the development of small molecule PPI-modulators. However, they suffer from drawbacks like poor metabolic stability and low bioavailability. Therefore, there is a need of new modified types of peptides with improved stability, metabolic and pharmacokinetic properties (17). The identification approaches of the small molecule PPI modulators can be classified in the following categories:

1. Peptide and peptidomimetic approaches.
2. High throughput screening of chemical libraries and natural products.
3. Computational approaches based on X-Ray crystallography and NMR data.

The above approaches will be discussed in short. In peptide and peptidomimetic approaches, peptides with structures based on amino acid sequence at the PP interface is selected which acts as a lead for the development of novel and improved PPI modulators (30). High Throughput Screening is a selective approach for the discovery of lead PPI modulators (27), (31), (32), (33). To make HTS successful, two key things are important, first to maximize the chance of identifying suitable binding molecule (both in terms of selectivity and affinity) and the second key thing is the choice of assay (17). In the present time, computational approaches are promising strategy to identify and design

small- molecule PPI modulators (34), (35), (36), (37). Computational approaches assist various steps in drug discovery process since performing identification only through experimental methods is very time consuming and expensive. So, in the last ten years, there has been an increasing effort to develop computational strategies to screen PP interfaces (38), (12), (39). Several studies have documented the use of *in-silico* approaches with medicinal chemistry approaches in the discovery of PPI modulators (40), (41), (42), (38).

1.4 PPI stabilization

In this study, we focus on the stabilization of the PPI, where a small molecule strengthens the naturally occurring PPIs by binding to both of the protein partners. Stabilization can be accomplished either with an allosteric or a direct mechanism. In allosteric mechanism of PPI stabilization, stabilizer can bind to a single protein partner, increasing the mutual binding affinity of the protein partners by inducing conformational modifications. A typical example of allosteric mechanism of PPI stabilization is shown by paclitaxel. Paclitaxel is a small molecule modulator which binds to a hydrophobic pocket of the β subunit of polymerized tubulin and stabilizes polymerized microtubule structure (43), (44). In direct stabilization of PPI, the stabilizer binds to both of the protein partners, exerting a bridge like function which also increases the mutual binding affinity (2). Direct mechanism of PPI stabilization can be of two types. In first type, the stabilizing molecule first binds to one of the protein partners, thereby modifying the interaction surface for the second protein partner and stabilizes the complex. The effect of stabilizer is so strong that the complex is not stable in the absence of these stabilizers. In the second type of direct PPI stabilization, the stabilizer molecule directly binds to the rim of an already established PP interface and increases the binding affinity of the protein partners (2). **Figure 4c** shows the mechanism of direct stabilization.

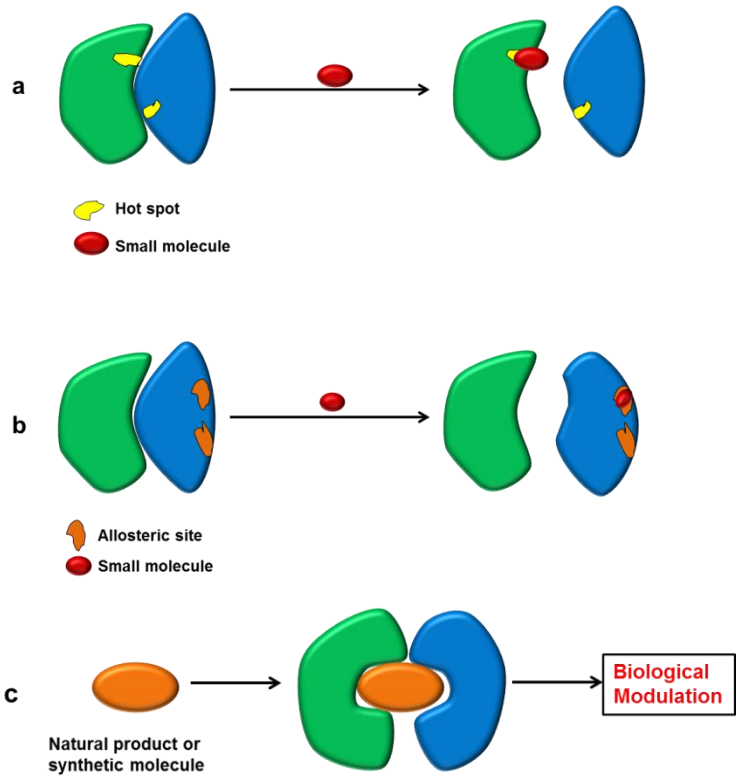

Figure 4: (a) Hot spot approach to PPI inhibition. (b) Allosteric approach to PPI inhibition. (c) PPI stabilization by direct mechanism. The two protein partners are represented in green and blue. Hot spot is represented in yellow. Reprinted from "Targeting protein-protein interactions for therapeutic intervention: a challenge for the future" by G. Zinzalla and D. E. Thurston, 2009. *Future Med. Chem.*, *1*, p.68. (17)

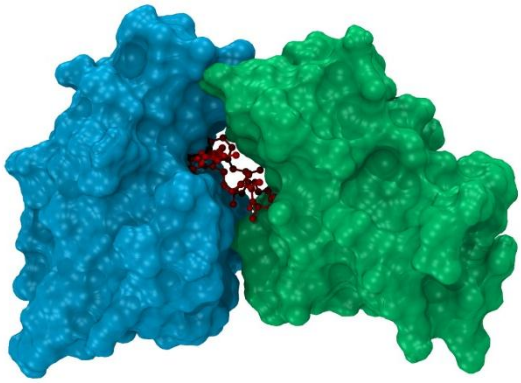

Figure 5: The structure of the ternary complex of FKBP12-rapamycin-FRB is shown representing PPI-stabilization. FKBP12 partner of the complex is represented in cyan whereas FRB partner in green. Rapamycin molecule is denoted in red. Image is rendered with Visual Molecular Dynamics (VMD) (45)

In this work, we have studied a typical case of first type of direct mechanism of PPI stabilization, the complex between FK506-binding protein (FKBP12) and FKBP12-rapamycin-binding domain of FRAP (FRB) with the small molecule stabilizer rapamycin (**Figure 5**). FRAP is FKBP-Rapamycin Associated Protein (46).

The approach of stabilization of PPI offers a valuable alternative for the discovery of novel bio-active drug-like molecules. Nature has provided us a number of compounds that acts as PPI stabilizers and they are good starting point in the drug discovery process. PPI stabilizers have several advantages. As they bind in a non-competitive way, their affinity does not need to be in low nanomolar range to trigger a strong physiological effect. Moreover, on the contrary to enzymes- inhibitors which are less specific because of strong homologies within enzyme families, PPI stabilizers are more specific because the PP interface of the complexes are more distinct (2). Despite of these advantageous features, few instances of PPI stabilizers resulting from target-oriented small molecule discovery have been reported yet. Nevertheless, the scientific community is making a great effort to pursue in this way towards drug discovery.

CHAPTER 2

Computational Methods I

2.1 Molecular dynamics: An overview

Molecular Dynamics (MD) is a computational method which describes equilibrium and biological properties of a biological system. MD helps to gain insight into processes taking place at atomic and molecular level, which is quite difficult to have purely from experiments because of lack of sufficient resolution. i.e., with MD, we can keep track of the behavior of each atom of a system. MD simulations provide information about the possible conformations, thermodynamic properties and dynamic behavior of a molecule.

In short, MD involves the solving of classical equations of motion for a group of atoms or molecules. For numerically solving those equations, MD requires three ingredients: I) A law which describes the mutual interaction between the atoms in a system from which potential energy and the forces on the atoms can be calculated given the position of atoms. II) An algorithm to integrate numerically the equations of motion for the atoms in the system. III) The integration scheme needs to be provided with initial conditions like initial positions and velocities of all atoms in the system to finally solve the equations. Apart from the above approximations, two more approximations are required. First, atoms should be considered as classical entities and obey Newton's equations of motion. Second, the model used for describing the interaction between the atoms of the system.

In MD, one explores the macroscopic properties (e.g. Pressure, energy, enthalpy) of the system through microscopic simulations (microscopic properties include atomic positions and velocities). For example, to calculate the change in binding free energy of a particular drug or to study the mechanism of conformational change of a molecule. The connection between microscopic simulations and macroscopic properties is made by statistical mechanics. In statistical mechanics, averages related to experiments are defined in terms of ensemble averages[1], so as to have good agreement with the experiments. With MD simulations, one can calculate time average of any observable of interest (expressed as a function of momentum p and position r) by integration. Since, MD performs by the time average of the observables and experimental observables are assumed to be ensemble averages, so to resolve this, a fundamental axiom of statistical mechanics is used in MD. This axiom states that statistical ensemble averages are equal to time averages of the system, and is known as ***Ergodic hypothesis*** (47).

Ensemble average = time average

$$\langle A \rangle_{ensemble} = \langle A \rangle_{time}$$

While performing MD, a system is allowed to evolve in time for enough duration so that the system passes through a large number of conformational states. Therefore, the aim of MD is to generate enough representative conformations such that the ergodic hypothesis is satisfied.

MD simulations calculate future positions and velocities of the atoms of a system, based on their current position and velocities (48), (49). MD is based on classical mechanics, i.e. on Newton's second law of motion. From the knowledge of the force on each atom, it

[1] Ensemble average is average taken over a large number of replicas of the system considered simultaneously.

is possible to determine the acceleration of each atom in the system, as the mass of the atoms is known. The equations of motions are then integrated to generate a trajectory that describes the positions, velocities and accelerations of the atoms with respect to the time. From this trajectory, one can determine the average values of properties. MD is a deterministic method; i.e. once the position and velocities of each atom are known, the state of the system can be predicted anytime in the future.

Newton's equation of motion:

$$F_i = m_i a_i$$

Where, F_i is the force exerted on the particle i.

m_i is the mass of the particle i.

a_i is the acceleration of the particle i.

The force can be expressed as the gradient of potential energy:

$$F_i = -\frac{dV}{dr_i}$$

Where, V is the potential energy function.

Combing the above two equations:

$$-\frac{dV}{dr_i} = m_i a_i$$

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

Thus, potential energy of the system can be related to the change in position of atoms with respect to the time.

Kinetic energy of the system is defined in terms of velocities:

$$K = \frac{1}{2}\sum_{i=1}^{N} m_i v_i^2$$

The total energy of the system is the sum of kinetic and potential energies. This is known as Hamiltonian and is expressed as follows:

$$H(r,p) = K(p) + V(r)$$

Where, *r* is a set of Cartesian coordinates and *p* is the momenta of the atoms (50).

Let us consider a simple case study where acceleration is constant.

$$F = m \cdot a \ = m \cdot \frac{dv}{dt} \ = m \cdot \frac{d^2r}{dt^2}$$

Since acceleration can be expressed as $\frac{dv}{dt}$, we get the velocity after integration.

$$v = v_0 + at$$

And, since velocity can be expressed as $v = \frac{dr}{dt}$, we get the position after integration.

$$r = r_0 + vt$$

By combining the above expressions of velocity (*v*) and position (*r*), we get the following expression which gives the value of *r* at time *t* as a function of acceleration *a*, initial velocity v_0 and initial position r_0

$$r = r_0 + at^2 + v_0 t$$

Acceleration can be expressed as a derivative of potential energy with respect to position:

$$a = -\frac{1}{m}\frac{dV}{dr}$$

The equations of motion are deterministic in nature, i.e., if the initial positions and velocities are known, then position and velocities at time t can be calculated. Therefore, for the calculation of the trajectory, the values which are needed are as follows:

a) Initial positions of the atoms: Through experimental structures predicted by X-ray crystallography or NMR spectroscopy, initial position of atoms can be obtained.

b) Initial distribution of the velocities: The initial velocities should obey the required temperature and corrected so there is no overall momentum (*P*), by setting the center of mass to zero. This prevents the system from drifting in space.

$$P = \sum_{i=1}^{N} m_i v_i = 0$$

Usually, the velocities v_i at a given temperature *T*, are chosen randomly by a Maxwell-Boltzmann distribution or Gaussian distribution. It gives the probability that an atom *i* has a velocity v_X in the direction *X* at temperature *T*.

$$pr(v_{iX}) = \left(\frac{m_i}{2\pi k_B T}\right)^{\frac{1}{2}} exp\left[-\frac{1}{2}\frac{m_i v_{iX}^2}{k_B T}\right]$$

Where, $v_{iX}$is the velocity of atom *i* in direction *X*.

$k_B$is the Boltzmann constant.

c) The acceleration: It is determined by the gradient of potential energy with respect to the position.

Temperature *T* can be calculated using velocities as follows:

$$T = \frac{1}{(3N)} \sum_{i=1}^{N} \frac{|p_i|}{2m_i}$$

Where, p_i is the momentum of the N is the atom *i* and N is the total number of atoms in the system.

The potential energy is the function of all the atoms' atomic positions (3N). Because of the complex nature of the function, they must be solved numerically using various numerical integration methods. Several numerical algorithms have been developed; for

e.g. Verlet algorithm, Leap-frog algorithm and Velocity Verlet algorithm. The studies have been performed with Amber package based on leap-frog algorithm for numerical integration which is discussed below.

Leap-frog Algorithm

This algorithm first calculates the velocity at time $(t + \frac{1}{2}\delta t)$ by the following expression

$$v\left(t + \frac{1}{2}\delta t\right) = v\left(t - \frac{1}{2}\delta t\right) + a(t)\delta t$$

The velocity at time $\left(t + \frac{1}{2}\delta t\right)$ is used to calculate the position *r* at time $(t + \delta t)$, which is as follows:

$$r(t + \delta t) = r(t) + v\left(t + \frac{1}{2}\delta t\right)\delta t$$

In the first time step δt, the velocities progresses from time $t = 0$ to $\left(t + \frac{1}{2}\delta t\right)$. In doing so, they "leap" over the position at time t, then the positions leap over the velocities. The expression for the calculation of velocity at time *t* is:

$$v(t) = \frac{1}{2}\left[v\left(t - \frac{1}{2}\delta t\right) + v\left(t + \frac{1}{2}\delta t\right)\right]$$

The acceleration is then determined and the cycle is renewed. Thus the velocities advances from $\left(t - \frac{1}{2}\delta t\right)$ to $\left(t + \frac{1}{2}\delta t\right)$ progressively and the resulting set of positions and velocities is the MD trajectory.

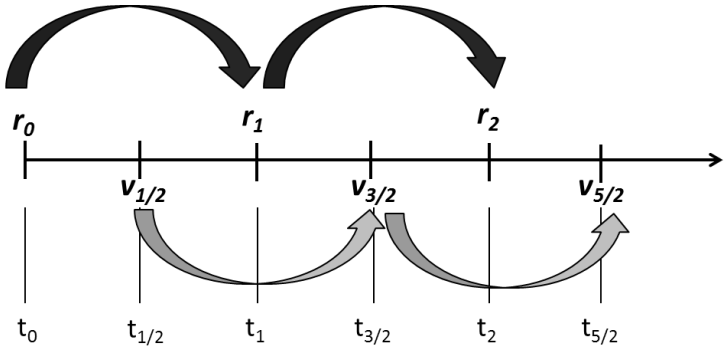

Figure 6: An illustration showing the principle of leap frog mechanism. Adapted from "Leap frog Integrator", Computational Physics by S. McMillan (51)

Force Fields

The basic idea behind MD simulations is the representation of the energy of the molecule as a function of its atomic coordinates. The conformations of the molecule are stable at low energy regions of the potential energy function, and the forces on the individual atoms are related to the gradient of this potential energy function. So, such functions are commonly known as "force fields" (52).

Although, the electronic structure calculations by means of quantum mechanics are highly accurate, it is not yet feasible to calculate energy surfaces for macromolecules by this method. So, most practical simulations use a set of simple classical functions to represent the energy, adjusting several parameters and optimizing them to be in agreement with experimental data (X-Ray and electron diffraction, NMR and IR spectroscopy) and with quantum mechanical calculations on smaller molecules. Thus in short, force field is a collection of simplified potential energy functions and all the related parameters to describe the interaction between atoms. Despite of its classical nature, force-fields can mimic the behavior of a system at atomic level with high accuracy in a fraction of time. Commonly used force fields are AMBER (53), CHARMM (54), GROMOS (55) and OPLS (56). Force fields must be chosen carefully

according to the nature of the system being simulated. In this study, we have used AMBER force field since this force field is appropriate for protein-protein and protein-ligand interaction studies.

Molecular mechanics assumes additivity of the energies:

$$E(r) = \underbrace{E_{str} + E_{bend} + E_{tors} + E_{imp}}_{bonded\ terms} + \underbrace{E_{vdw} + E_{elec}}_{non-bonded\ terms}$$

The potential energy expression for the AMBER force field is as follows:

$$\begin{aligned} E(r) = & \sum_{bonds} k_b(b - b_0)^2 \\ & + \sum_{angles} k_\theta(\theta - \theta_0)^2 \\ & + \sum_{dihedrals} k_\varphi[cos(n\varphi + \delta) + 1] \\ & + \sum_{atom\ i} \sum_{i>j} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \sum_i \sum_{i>j} \frac{q_i q_j}{4\pi\epsilon r_{ij}} \end{aligned}$$

Where, the first three summation terms are over bonds, angles and dihedrals. The fourth summation term describes the van der Waal's (Leonard-Jones 6-12 potential) and the final term represents the electrostatic potential.

2.2 Free energy calculations

The energetics provides the most important information about the relationship between the structure, function and stability of the biomolecules. Over the past decade, several computational approaches have been developed for the calculation of binding free energy, each with individual trade-offs between computational cost and accuracy. While choosing an approach for free energy calculations, two key points should be considered, accuracy and the computational cost of the method in reference to the available resources.

Most fast methods for prediction of binding free energies are the empirical or knowledge-based approaches, which are based on simple energy functions (57), (58). The main disadvantage of the above mentioned approach is the lack of the conformational sampling which makes these methods faster but at the cost of accuracy. On the other hand, the most time-consuming and accurate methods are based on molecular force fields and they use Molecular Dynamics (MD) and Monte Carlo (MC) simulations to generate ensemble averages. Free energy perturbation (59) and thermodynamic integration (60) are such methods which gives the accurate estimates but are computationally time consuming, thus, difficult to estimate binding free energies for large protein complexes. Another approach that relies on MD or MC simulation is Linear Interaction Energy approach (61). Since the above mentioned methods are computationally expensive, so more computationally affordable yet reliable methods have been developed to tackle this difficult task. These methods, referred to as MM–PBSA (Molecular Mechanics–Poisson–Boltzmann Surface Area) (62) or MM–GBSA (Molecular Mechanics–Generalized-Born Surface Area) (63), combine the molecular mechanical energies in gas phase, the Poisson–Boltzmann or the Generalized Born approach to evaluate solvation energy, and an empirical function to take into account the contribution of protein surface exposure to solvent. They perform *a posteriori* evaluation

of binding free energy on snapshots extracted from a molecular dynamics trajectory to evaluate the total binding free energy between two proteins forming a complex.

2.2.1 MM/GBSA approach

In this study, we have employed MM/GBSA approach for the calculation of binding free energy. MM/GBSA approach is combined Molecular Mechanics/ Generalized Born Surface Area to calculate the free energy of binding (64). This is faster method which is based on the force fields and computes free energy of binding from the difference between the two states (unbound and bound state) in solution. MM/GBSA method is very appropriate in case of protein-ligands, protein-protein and protein-peptide complexes. The advent of sophisticated computing-clusters and automated protein/ligand preparation routine makes the calculation of binding free energies by MM/GBSA model feasible. The net free energy change is treated as a sum of individual energy components, each with a physical basis. This approach uses a thermodynamic cycle (**Figure 7**) to calculate the binding free energy.

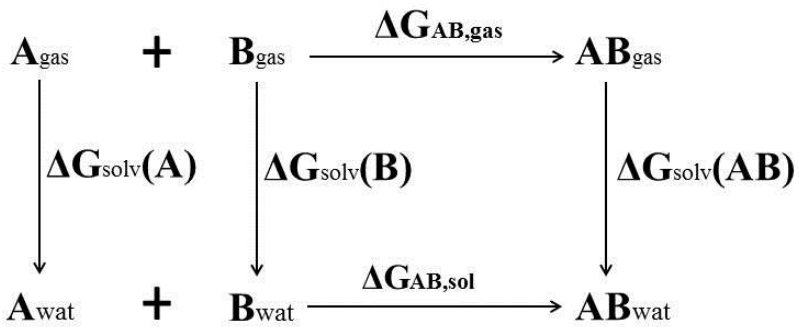

Figure 7: The thermodynamic cycle to calculate the binding free energy of the complex in solution ($\Delta G_{AB,sol}$) The binding energy is decomposed into binding free energy in vacuum ($\Delta G_{AB,gas}$) and solvation free energies of the reacting species.

The solvation contribution of a species X can be expressed as:

$$X_{gas} \xrightarrow{\Delta G_{solv}(X)} X_{sol}$$

$\Delta G_{solv}(X)$ is the free energy of solvation of *X*. where *X=A,B* in this case.

$$\Delta G_{solv}(X) = G(X)_{sol} - G(X)_{gas}$$

$G(X)_{sol}$ is free energy of *X* in solvent. $G(X)_{gas}$ is free energy of *X* in vacuum.

Therefore, the free energy of *X* in solvent can be written as:

$$G(X)_{sol} = \Delta G_{solv}(X) + G(X)_{gas}$$

The binding energy of the complex *AB* is :

$$\Delta G_{AB,sol} = G(AB)_{sol} - [G(A)_{sol} + G(B)_{sol}]$$

$$\Delta G_{AB,sol} = G(AB)_{sol} - G(A)_{sol} - G(B)_{sol}$$

The above equation can be written as:

$$\Delta G_{AB,sol} = [\Delta G_{solv}(AB) + G(AB)_{gas}] - [\Delta G_{solv}(A) + G(A)_{gas}] - [\Delta G_{solv}(B) + G(B)_{gas}]$$

$$= \Delta G_{solv}(AB) + G(AB)_{gas} - \Delta G_{solv}(A) - G(A)_{gas} - \Delta G_{solv}(B) - G(B)_{gas}$$

$$= \Delta G_{solv}(AB) - \Delta G_{solv}(A) - \Delta G_{solv}(B) + [G(AB)_{gas} - (G(A)_{gas} + G(B)_{gas})]$$

$$= \Delta G_{solv}(AB) - \Delta G_{solv}(A) - \Delta G_{solv}(B) + \Delta G_{AB,gas}$$

$$\boldsymbol{\Delta G_{AB,sol} = \Delta G_{AB,gas} + \underbrace{\Delta G_{solv}(AB) - \Delta G_{solv}(A) - \Delta G_{solv}(B)}_{\Delta\Delta G_{solv}}}$$

The above equation is the decomposition of $\Delta G_{AB,sol}$ into vacuum term $\Delta G_{AB,gas}$ and a solvation term $\Delta\Delta G_{solv}$ according to the thermodynamic cycle of **Figure 7**. This equation is the basic equation which calculates the binding free energy on the basis of type of solvent model used.

$$\boldsymbol{\Delta G_{AB,sol} = \Delta G_{AB,gas} + \Delta\Delta G_{solv}}$$

Dealing with the snapshots extracted from a MD trajectory, the equation can be written as:

$$\boldsymbol{\Delta G_{AB,sol} = \langle \Delta E_{MM} \rangle + \Delta\Delta G_{solv} - T\Delta S}$$

Where, $\langle \Delta E_{MM} \rangle$ is the average molecular mechanics (MM) energy calculated during MD simulation.

$T\Delta S$ is the entropic term. *T* is the absolute temperature (in Kelvin).

Molecular mechanics energy contribution

The molecular mechanics energy ΔE_{MM} is expressed as the sum of the intermolecular (ΔE_{inter}) and intramolecular (ΔE_{intra}) energies.

$$\Delta E_{MM} = \Delta E_{intra} + \Delta E_{inter}$$

The intramolecular energy is composed of the energies related to bonds, angles and dihedrals. The intermolecular energy has two components electrostatic and van der Waal's component.

$$\boldsymbol{\Delta E_{MM} = \Delta E_{bonds} + \Delta E_{angles} + \Delta E_{dihedrals} + \Delta E_{elec} + \Delta E_{vdw}}$$

ΔE_{MM} is the molecular mechanics energy which is calculated during MD/MC simulation.

Entropic contribution

$\boldsymbol{T\Delta S}$ is the entropic term which is challenging to compute. It can thus be approximated to ΔS_{vib}, i.e. the contribution due to the internal vibrations, whose calculation is usually time-consuming and can be affected by large statistical uncertainty, so it can be performed only on a part of a protein complex (65). Some of the researchers advocates its usage (66), while some believes to neglect it because it could be a reason of major source of error (67).

Entropy is calculated as a combination of change in translational (ΔS_{tr}), rotational (ΔS_{rot}) and vibrational (ΔS_{vib}) entropies, related with corresponding degrees of freedom. Usually, in literatures, the term 'configurational entropy' (ΔS_{conf}) is used.

$$\Delta S_{conf} = \Delta S_{tr} + \Delta S_{rot} + \Delta S_{vib}$$

In the case of small molecules binding to the large receptors, the first two terms ($\Delta S_{tr} + \Delta S_{rot}$) might prevail on ΔS_{vib}. But in case of protein-protein complex, ($\Delta S_{tr} + \Delta S_{rot}$) is generally neglected and the change in vibrational entropy (ΔS_{vib}) is more relevant. ΔS_{vib} is usually estimated by using Normal mode analysis of MD trajectories.

Solvation contribution

In MM/GBSA approach, the solvation term $\Delta\Delta G_{solv}$ is computed as:

$$\Delta\Delta G_{solv} = \Delta G_{GB} + \Delta G_{SA}$$

Where, ΔG_{GB} refers to the electrostatic component of the solvation while ΔG_{SA} refers to the non-electrostatic component of the solvation, called as cavitation energy that corresponds to the energy needed to create the cavity occupied by the solute in the

solvent (68). The solvation term is evaluated *'a posteriori'* using "*implicit* solvent"[2] calculations. Here, it should be noted that for MD simulations, "*explicit* solvent"[3] model is usually used while it have been changed to implicit solvent model for free energy calculations by MM/GBSA method. Thus, the solvent model and therefore the Hamiltonian are different in simulation and free energy calculation steps (69).

Electrostatic component of the solvation term

The first term ΔG_{GB} can be calculated using the modified Generalized Born (GB) model developed by Onufriev et al (70).

The electrostatic component of the solvation energy is related with the creation of a charge q_i in a spherical cavity of radius a_i in a solvent having dielectric constant ε. For N charges located at r_{ij} distance, the defining equation employed for electrostatic contribution to the solvation free energy with MM/GBSA model would be as follows:

$$\Delta G_{GB} = -\frac{1}{2}\left(1-\frac{1}{\varepsilon}\right)\sum_{i,j\ (i\neq j)}^{N}\frac{q_i q_j}{f\left(r_{ij}, a_{ij}\right)}$$

where, $$f\left(r_{ij}, a_{ij}\right) = \sqrt{r_{ij}^2 + a_{ij}^2 e^{-D}}$$

$$a_{ij} = \sqrt{a_i a_j}$$

$$D = \frac{r_{ij}^2}{\left(2a_{ij}\right)^2}$$

[2] Implicit solvent is representation of the solvent as continuum medium with the dielectric properties. Using implicit water model is computationally affordable. Free energy by MM/GBSA or MM/PBSA approaches can only be estimated by using implicit continuum models (223).

[3] Explicit solvent is where a biological molecule is surrounded by a large number of solvent molecules. This representation is physiologically more realistic but is computationally expensive because of the necessity of periodic boundary conditions (224).

Non-electrostatic component of the solvation term

The second non-electrostatic term is computed as a function of solvent accessible surface area (71) which is as follows:

$$\Delta G_{SA} = \gamma A + b$$

where, A is the total solvent accessible surface area. γ and b are constants.

Combining the different contributions, total free energy of binding can be expressed as:

$$\Delta \boldsymbol{G_{AB,sol}} = \langle \Delta \boldsymbol{E_{MM}} \rangle + \Delta \boldsymbol{G_{GB}} + \Delta \boldsymbol{G_{SA}} - \boldsymbol{T} \Delta \boldsymbol{S}$$

2.3 Computational Alanine Scanning

Computational alanine scanning (CAS) is a computational technique for the prediction of residues which are energetically important for binding of the protein partners in a protein-protein complex. CAS is analogous to experimental alanine mutagenesis. On the protein-protein interface, there are few residues that are energetically important; these are called *"Hot Spots"* by Clackson and Wells (72) in their work on the binding of human growth hormone to its receptor, where only a small fraction of interfacial residues contributes to the binding free energy. And this is the general property of the most protein-protein complexes (15), (73). CAS technique has been first used by P. Kollman to study PPIs (74). They have reported an extension to the MM/GBSA free energy calculation approach used by Srinivasan et al (75).

The calculation is performed by evaluating the binding free energy between the protein subunits making up the complex (ΔG_{WT}) and the binding energy upon mutation

of each interfacial generic amino acid X into alanine ($\Delta G_{MUT,X}$), so, for each residue X, a $\Delta\Delta G$ value can be obtained as

$$\Delta\Delta G_X = \Delta G_{MUT,X} - \Delta G_{WT}$$

Mutating all interfacial amino acid and the calculation of $\Delta\Delta G_X$ in the above mentioned way yields a map predicting which amino acids are important for PPI and which are not. These amino acids tend to be grouped into small clusters on the interface and they are classified on the basis of difference in binding free energy between the wild type and mutant type as *hot spots* ($\Delta\Delta G$ > 4 kcal mol^{-1}), *warm spots* (2 kcal mol^{-1} < $\Delta\Delta G$ < 4 kcal mol^{-1}) and *cold spots or null spots* ($\Delta\Delta G$ < 2 kcal mol^{-1}) (76). These hot spots are correlated with conserved residues at specific locations. The fundamental principle of CAS is based on experimental alanine mutagenesis. Despite of advances in molecular biology, experimental alanine mutagenesis is more time consuming, difficult and expensive for performing high throughput screening of PP interfaces. Thus, CAS plays important role in predicting hot spots in modeled or experimentally determined PP complexes. CAS along with structural and thermodynamic studies has enabled the discovery of important protein-protein interactions. **Figure 8** represents the flowchart for the calculation of difference in free energy of binding ($\Delta\Delta G$) by CAS approach.

There are some assumptions that were made in the calculation of $\Delta\Delta G$ by CAS approach which are as follows:

1. The terms in the energy function used for the calculation of $\Delta\Delta G$ (refer to **section 2.2.1**) are pairwise additive. This assumption is necessary for the fast computational evaluation of $\Delta\Delta G$.
2. Next assumption is that in CAS approach, cofactors, metal ions and water molecules participating in hydrogen bonding at the protein interface are not taken into account.

3. Another assumption is that the symmetry in a protein-protein complex is not taken into account. During CAS, only one residue at a time is mutated to alanine. This point is valid for the protein homodimer complexes. According to CAS, one residue is mutated at a position in interface of a partner, and the corresponding residue on the other partner of the dimer is considered as wild type (77).
4. Proline residue is not mutated to alanine and thus not considered for CAS. This is because its backbone conformation is different from that of alanine and the mutation would lead to significant conformational changes (78). While the mutation of other residues to alanine causes only local changes and the overall conformation of the complex is preserved.

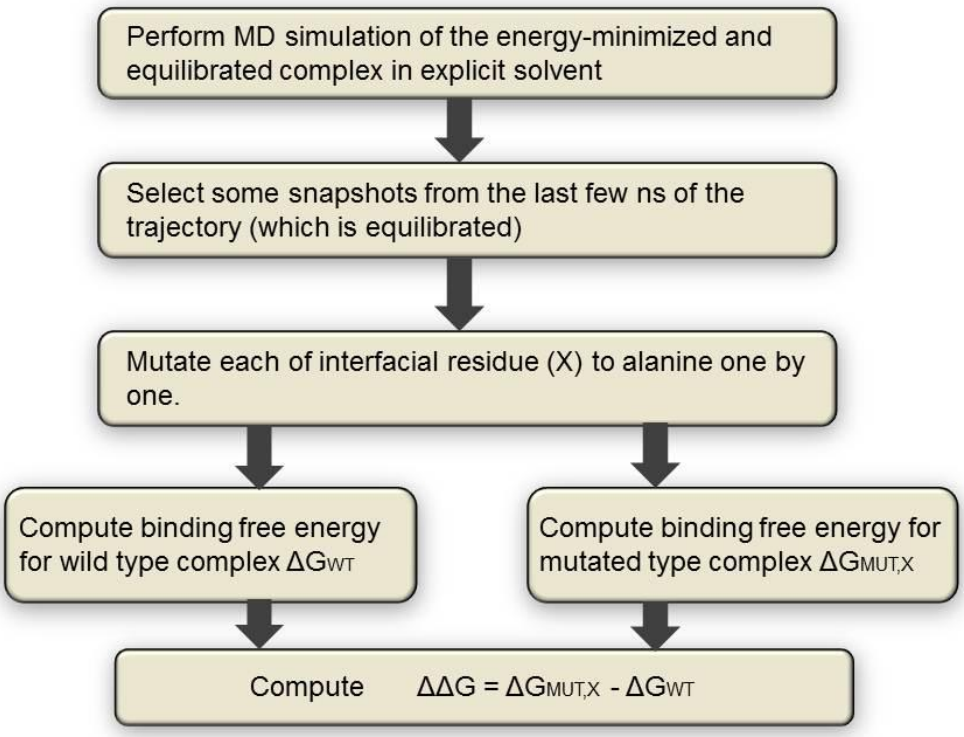

Figure 8: Flowchart illustrating the procedure of Computational Alanine Scanning approach

CHAPTER 3

TERNARY COMPLEX:

FKBP12-RAPAMYCIN-FRB

Abstract

Modulating Protein-Protein Interactions is a viable therapeutic strategy to fight many different diseases. Several studies have been performed and provided evidence for the development of small molecules that modulate PPIs which has opened new opportunities in the field of drug discovery. In principle, PPI can be exploited as a drug target either by inhibiting or stabilizing the complex. Until now, most research has been dedicated to the design and development of PPIs inhibitors, but in recent times, several researches are also going on PPIs stabilizing agents. In this study, we have analyzed a typical case of protein-protein interaction stabilization: the complex between FKBP12 and FRB with rapamycin. We have analyzed the stability of the complex and characterized its interactions at the atomic level by performing free energy calculations and computational alanine scanning. It is shown that rapamycin stabilizes the complex by acting as a bridge between the two proteins forming the biologically active ternary complex. The main objective of this study is to characterize the interaction network responsible for the binding of rapamycin with FKBP12 and FRB and to get deeper understanding of rapamycin stabilization of the above mentioned complex.

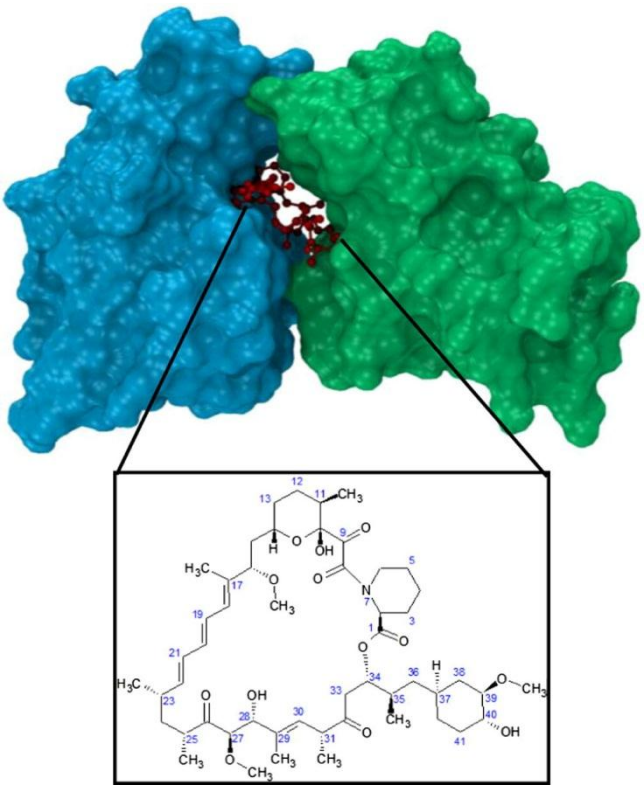

Figure 9: The structure of the ternary complex of FKBP12–rapamycin–FRB is shown (upper panel) along with the chemical structure of rapamycin (lower panel). FKBP12 partner of the complex is represented in cyan whereas FRB partner in green. Rapamycin molecule is denoted in red.

3.1 Biological background

The biologically active ternary complex is composed of two proteins FKBP12 and FRB and a small molecule rapamycin. FKBP12 is FK506-binding protein. FRB is FKBP12-Rapamycin-binding domain of FRAP. FRAP is FKBP Rapamycin associated protein. **Figure 9** represents the structure of the complex.

Rapamycin is an antifungal agent produced by *Streptomyces hygroscopicus* which was isolated from Easter Island soil sample (79). This compound is a white crystalline solid which is insoluble in aqueous solutions but soluble in organic solvents (80), (81). It is a 31-membered lipophilic macrolide compound. Studies have shown that rapamycin has antitumor and immunosuppressive activity (82), (83). It shows antifungal activity especially against *candida albicans* and other filamentous fungi (84). Rapamycin has received more focused consideration as an immunosuppressant. Thus, it is now a FDA-approved drug and is used clinically as an immunosuppressant for organ transplant patients (85). Several experiments have been performed to demonstrate the anti-proliferative action of Rapamycin (86), (87), (88). All these activities have been studied and it has been observed that they share a common mechanism (89). Rapamycin binds to specific family of immunophilins (cytosolic binding proteins), known as FK506 binding proteins (FKBPs). The most relevant protein of this family for the immunosuppressive effects of rapamycin is FKBP12. FKBP12 is a 12 kDa protein and functions as *cis*/*trans* peptidylpropyl isomerases that catalyzes the cis to trans isomerization of Xaa-proline peptide bonds in short synthetic peptides (90), (91), (92). Rapamycin binds to FKBP12 and inhibits its isomerase activity, but this drug–immunophilin complex is insufficient to mediate immunosuppressive effect of rapamycin. Rapamycin becomes biologically active only when FKBP12–rapamycin complex binds with a small 11 kDa hydrophobic binding domain (FRB domain) present on the 289 kDa protein FRAP. This FRB domain is the specific intracellular target of FKBP12–rapamycin complex (93). FRAP is also referred by terms like mTOR (mammalian target of Rapamycin), RAFT1 (Rapamycin FKBP12 target) and RAPT1 (Rapamycin target) (94), (95), (96). The inhibition of this protein blocks signal transduction pathways preventing cell cycle progression from G1 to S phase in various cell types, allowing rapamycin to exploit its action (89).

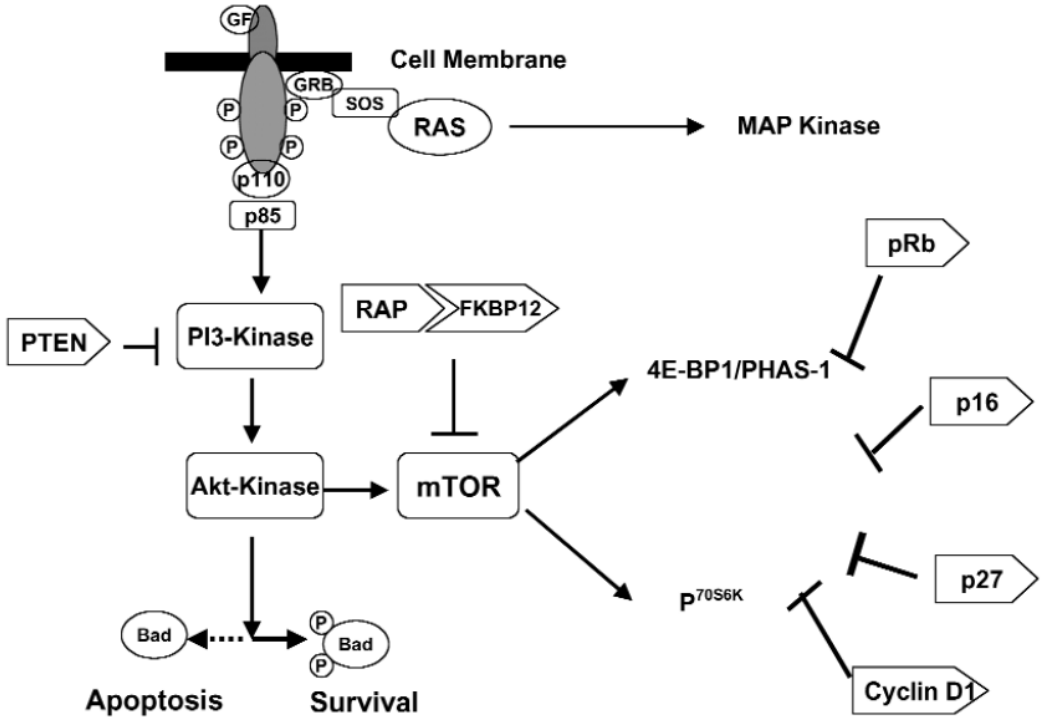

Figure 10: The signal transduction pathway showing the involvement of Rapamycin. Rapamycin binds to immunophilin FKBP12. The Rapamycin-FKBP12 complex binds to FRAP (represented here as mTOR). The ternary complex inhibits the phosphorylation and activation of downstream translational regulators 4E-BP1/ PHAS-1 and $p70^{s6k}$. These in turn, decrease the translation of mRNA of specific proteins essential for G1 to S phase progression. Rapamycin is represented as RAP. Adapted from "The molecular target of rapamycin (mTOR) as a therapeutic target against cancer" by M.M. Mita et al, 2003, *Cancer Biology & Therapy*, *2*, p. S171 (97)

3.1.1 Mechanism of action of Rapamycin

Rapamycin interferes with the elements of the signal transduction pathways, especially FRAP (mTOR) pathway resulting in the cell cycle progression arrest from G1 to S-phase. As previously mentioned, rapamycin first binds to FKBP12 and the resultant complex FKBP12-Rapamycin inhibits kinase activity of FRAP. This in turn blocks the activity of two downstream signaling elements $p70^{s6k}$ (40s ribosomal protein s6 kinase) and 4E-BP1/PHAS-1 [Eukaryotic initiation factor 4E binding protein-1(4E-BP1), also

known as PHAS-1 (Phosphorylated heat and acid stable protein-1)]. These signaling elements will be discussed later in this section. The inhibition of these signaling elements results in the inhibition of the synthesis of proteins important for cell-cycle proliferation. Thus, rapamycin acts as an anti-tumor agent. **Figure 10** represents the mechanism of action of rapamycin in FRAP (mTOR) pathway.

Upstream of FRAP

FRAP belongs to PIKK (phosphoinositide 3-kinase related kinase) family of protein kinases which is involved in many regulatory cellular functions related to cell cycle progression, DNA damage and repair; and DNA recombination. PIKK family members possess a common carboxy-terminal catalytic domain that bears sequence homology to lipid kinase domains of PI3K (phosphatidylinositol 3-kinase). FRAP contains three domains, namely: FRB domain, catalytic domain and heat domain. FRB domain is evolutionary conserved.

Although, the precise mechanism of upstream pathway of FRAP is not completely understood, some researches have shown that PI3K/ protein kinase B(PI3K/Akt) appears to be the key factors in upstream pathway (**Figure 10**) (98), (99). PI3K is a lipid enzyme that plays key role in cell proliferation and has been demonstrated to be upregulated in malignant cells. PI3K/Akt can phosphorylate BAD (Bcl-2-related apoptosis regulatory protein) (100). Akt pathway is particularly important because of its role in inhibiting apoptosis and promoting cell proliferation. Akt affects the phosphorylation state of BAD (**Figure 10**).

The activation and downstream signaling of PI3K and Akt is inhibited by tumor suppressor gene PTEN. FRAP mediates many signals arising from PI3K and/or Akt like signals associated with tumorigenesis (101). Therefore, studying these pathways in detail may lead to develop therapeutic strategies against Cancer.

Downstream of FRAP

FRAP functions downstream of PI3K/Akt pathway. FRAP plays a role in modulating two downstream signaling elements as mentioned earlier: $p70^{s6k}$ and 4E-BPI/PHAS-1 (102), (103). There are many evidences which indicate that activation of either PI3K or Akt- expressing cancer cells, and/or PTEN mutations leads to the phosphorylation of both $p70^{s6k}$ and 4E-BPI/PHAS-1which is mediated through the activation of FRAP (104), (105). When PI3K- or Akt- expressing cancer cells and cells lacking PTEN suppressor function are treated with rapamycin, it blocks the phosphorylation of $p70^{s6k}$ and 4E-BPI/PHAS-1 and thus inhibiting the cell proliferation. Downstream targets $p70^{s6k}$ and 4E-BPI/PHAS-1 are described below:

1. **4E-BPI/PHAS-1**

It is a low molecular weight protein that represses the protein translation initiation. When treated with rapamycin, the phosphorylation of FRAP is blocked and thus the phosphorylation of 4E-BPI/PHAS-1. In its dephosphorylated state, 4E-BPI/PHAS-1 binds to eIF-4E, thus inhibits the translation of mRNAs to important regulatory proteins like growth factors, onco-proteins and other cell-cycle regulators (106).

2. **$p70^{s6k}$**

This serine-threonine kinase is another target downstream of FRAP. Following rapamycin treatment, the phosphorylation of FRAP is inhibited, thus the activation of $p70^{s6k}$ protein is blocked which leads to reduced translation of 5'-terminal oligopyrimidine (5' - TOP) mRNAs that encode for ribosomal protein, elongation factors and insulin growth factor-II. This results in rapid decrease in protein synthesis (107), (108).

Summarizing the whole mechanism, Rapamycin first binds to FKBP12 and then the resultant complex binds to FRAP. Thus, inhibiting the activity of FRAP which in turn

results in the decrease in the activity of downstream translational regulators $p70^{s6k}$ and 4E-BPI/PHAS-1. This results in inhibition of protein synthesis and G1 to S phase traverse. Due to these effects, Rapamycin inhibits the proliferation of cancerous cells in vitro and in vivo.

3.2 Structural details of the complex

The X-ray crystallographic structure of the ternary complex of FKBP12–rapamycin–FRB has been solved at 2.7 Å resolution by Choi et al. in 1996 (46). Ternary complexes of rapamycin derivatives were then crystallized and their structures were solved at 1.85 and 2.2 Å resolution by Liang et al. in 1999 (109). The structure of the complex is shown in **Figure 9**. FKBP12 protein consists of a β sheet made of five antiparallel β strands. A short α helix is also present. Rapamycin binds in a hydrophobic pocket formed between the α helix and β sheet (46). FRB domain of FRAP is composed of a bundle of four α helices with rapamycin binding to a hydrophobic pocket formed by helices α1 and α4 (46). Chemical structure of rapamycin is shown in the **Figure 9** (lower panel). It is evident that rapamycin interacts with both receptor proteins. Moreover, experimental evidences are available which demonstrates that the FKBP12 protein is unable to bind with FRB in the absence of rapamycin (85), (110). To get the deeper insight in the interaction network responsible for the binding of rapamycin with FKBP12 and FRB, we have performed molecular dynamics (MD) simulations and free energy calculations on this ternary system. In particular, we have calculated binding free energies between the components of the FKBP12–rapamycin–FRB ternary system and performed computational alanine scanning (CAS) to evaluate the contribution of each of the amino acids at the protein–protein and protein–rapamycin interface to the binding energy.

CHAPTER 4

Stabilization of FKBP12-FRB complex by Rapamycin

4.1 Materials and methods

The initial structure of the ternary complex, FKBP12-rapamycin-FRB for our simulation was retrieved from Protein Data Bank (Protein Data Bank Id: 1FAP) determined at 2.7 Å resolution (46). It corresponds to a structure of the complex of human FKBP12 and FRB domain of human FRAP with immunosuppressant drug rapamycin. Molecular dynamics simulation was performed with Amber 11 package (111) using explicit solvent and periodic boundary conditions. Two parameter sets were used- Amber99SB force field (112) for the two proteins and Generalized Amber Force Field (GAFF) (53) for rapamycin molecule. The system was solvated with TIP3P water (113) molecules. SHAKE algorithm (114) was employed to constrain all covalent bonds

involving hydrogen atoms to their equilibrium bond lengths, allowing a time step of 1 fs. The system was first submitted to 10000 steps of minimization (1000 steps using steepest descent method followed by 9000 steps using conjugate gradient method) to remove the close contacts in the crystal structure. The systems was then equilibrated for 100 ps with the number of particles, system volume and temperature constant (NVT ensemble) in order to equilibrate the temperature of the system, and subsequently to 100 ps with number of particles, system pressure (1 atm) and temperature (300 K) constant (NPT ensemble) in order to equilibrate the density of the system with a 500 kcal mol^{-1} $Å^{-2}$ restraint on the positions of C_α atoms. A 20 ns production phase of molecular dynamics simulation was then carried out in the NPT ensemble at the same temperature and pressure with a time step of 1 fs. Langevin thermostat (115) with a collision frequency of 5 ps^{-1} has been used to keep the temperature constant. C_α atoms were not restrained in the production phase. A 10 Å cut-off for non-bonded Van der Waals interactions was applied and Particle Mesh Ewald algorithm (116) was employed to calculate long-range electrostatic interactions.

The free energy of binding between the components of the ternary complex FKBP12-rapamycin-FRB was evaluated using the MM-GBSA approach (117), (118). The entropic contribution to binding free energy was evaluated with NMODE module of the package AMBER 11. Due to the dimension of the system under consideration, only residues at the protein-protein interface and protein-rapamycin interface were considered in normal mode analysis, as previously suggested (63), (65). The entropic contribution was evaluated on 50 snapshots extracted from last 5ns of the simulation. CAS was carried out over snapshots taken at 10 ps intervals from the last 5 ns of the trajectory (500 snapshots in total). Each amino acid at the interface was mutated into alanine removing all side chain atoms beyond Cβ and replacing it with a hydrogen atom, thus obtaining an alanine side chain (119). The interface was defined as the set of amino acids with a non-zero difference in solvent accessible surface area (SASA) in the complex versus the free proteins. SASA was calculated with the NACCESS software (120). The

free energy of binding is calculated using MM/GBSA approach for the wild type and the mutated complex.

4.2 Free energy calculations

Rapamycin acts as a stabilizing agent that allows the association of FKBP12 and FRB. From Solvent Accessible Surface Area calculations, we have evaluated the protein surface buried upon complex formation, which amounts to 1422 $Å^2$. Approximately 50% of the buried surface is occupied by rapamycin, which is placed between the two proteins.

In order to understand the interaction network at the detailed atomic level, binding free energy calculations and CAS studies were performed. To analyze the interface of the FKBP12-rapamycin-FRB complex, a 20 ns MD simulation with explicit solvent and periodic boundary conditions has been performed starting from the crystal structure.

We have used MM-GBSA method for the evaluation of binding free energies. They perform an *a posteriori* evaluation of binding free energy on snapshots extracted from the MD trajectory to evaluate the binding free energy. In principle, three molecular dynamics simulations should be run to calculate binding free energy, one for the complex and one for each of the monomers with rapamycin in order to sample their conformations. Actually, if the conformations of the isolated monomers do not differ too much from their conformations in the complex, a single MD trajectory of the complex can be computed, and conformations of the monomers can be extracted from the same trajectory. This approach, when the structural modification upon binding are not extensive, is accurate and has the advantage to be less computationally expensive and provides results with a lower statistical uncertainty (121). Please refer to the section **2.2.1** for detailed theory behind free energy calculation by MM-GBSA method.

When performing binding free energy calculations with the MM-GBSA method, it is customary to define a receptor and a ligand unit in the system, between which the binding energy is evaluated. In this work, we have three molecules, two proteins and rapamycin, so two sets of free energy calculation and CAS studies have been performed, in the first case considering rapamycin as a part of the FKBP12 unit, the second one including rapamycin in the FRB unit. Binding free energy of each of the proteins with rapamycin and between the two proteins in the absence of rapamycin were also calculated using single trajectory MM/GBSA method (**Table 1**). Entropic corrections have also been evaluated. While the molecular mechanics energy term can be easily obtained from the results of a molecular dynamics simulation, the entropic term is often difficult to achieve. It can thus be approximated to ΔS_{vib}, i.e. the contribution due to the internal vibrations, whose calculations are nonetheless usually time-consuming and can be affected by large statistical uncertainty (65). The relative contribution of the change in conformational entropy to the $\Delta\Delta G$ is considered to be negligible for the mutational studies, since it is supposed to cancel up when calculating it in the native and in the mutated complex, as in CAS studies (74). For the details of entropic calculations, please refer to **section 2.2.1**.

The entropy corrected value of average binding free energy referred to the first case (rapamycin considered as a part of the FKBP12 unit) is estimated to be -28.59 ± 2.42 kcal mol^{-1} (value ± S.E.M.), and for the second case (rapamycin considered as part of the FRB unit), it is found to be -29.30 ± 2.32 kcal mol^{-1}. Entropy values calculated with normal mode analysis are affected by large errors and can be considered as semi-quantitative, it is anyway interesting to notice that the entropy corrected average binding free energy between FKBP12 and FRB in the absence of rapamycin turned out to be +25.45 ± 2.79 kcal mol^{-1}, a positive value suggesting the impossibility of the formation of a binary complex between the two proteins. The value of the binding free energy for the ternary complex (FKBP12.RAP.FRB) can be obtained by adding the binding free energy for the first case (FKBP12.RAP + FRB) and binding free energy between FKBP12 and rapamycin, thus resulting to be -55.30 ± 4.26 kcal mol^{-1}. Likewise, adding

ΔG for the second case (FKBP12 + RAP.FRB) and ΔG between FRB and rapamycin, we get an almost identical value -55.72 ± 3.91 kcal mol^{-1}. These data suggest that the ternary complex is stable only in the presence of rapamycin, that is in agreement with the experimental observations which report that the formation of the FKBP12-FRB complex is strictly rapamycin dependent. (85), (110).

It is worth noting that, even if only half of the protein surface buried upon ternary complex formation is occupied by rapamycin (see above), the binding free energy between FRB and rapamycin is almost equivalent to the binding free energy between the FKBP12-rapamycin complex and FRB, implying that the direct protein-protein interaction is negligible and the binding is completely mediated by rapamycin. The same observation is valid for the FKBP12-rapamycin interaction, which is, considering statistical uncertainty, almost identical to the interaction between the FKBP12 itself and the FRB-rapamycin complex (see **Table 1**). All the data obtained from free energy evaluations show that rapamycin can be thought to act as a bridge between FKBP12 and FRB. The details of this bridging action will now be further investigated.

Systems	ΔG (without entropy) (kcal mol^{-1}) Mean ΔG ± S.E.M.	ΔG (with entropy) (kcal mol^{-1}) Mean ΔG ± S.E.M.
FKBP12.RAP + FRB	-69.69 ± 0.19	-28.59 ± 2.42
FKBP12 + RAP.FRB	-76.58 ± 0.20	-29.30 ± 2.32
FKBP12 + FRB	-21.35 ± 0.16	+25.45 ± 2.79
FKBP12 + RAP	-54.91 ± 0.15	-26.71 ± 1.84
FRB + RAP	-48.02 ± 0.09	-26.42 ± 1.59

Table 1: Binding free energy values (given in kcal mol^{-1}) calculated with MM-GBSA method for all the cases considered. RAP stands for rapamycin.

4.3 Computational Alanine Scanning Studies

The calculation of the binding free energy allows to perform a CAS analysis of the protein-protein interface, in order to highlight the residues that contributes significantly to protein-protein binding (hot-spots, for details refer to the **section 1.2.2**) and to analyze the interaction network of the ternary complex at the atomic level. The calculation is performed by evaluating the binding free energy between the protein subunits making up the complex (ΔG_{WT}) and the binding energy upon mutation of each interfacial amino acid into an alanine ($\Delta G_{mut,X}$), so, for each residue, a $\Delta\Delta G$ value can be obtained as

$$\Delta\Delta G_X = \Delta G_{mut,X} - \Delta G_{WT}$$

Also in this case, in principle two MD trajectories should be computed, one for the wild type and other one for the mutated complex. This would require performing a different MD simulation for each of the point mutations to be sampled in the complex, leading to a substantial number of simulations for a large surface. It is possible to make the hypothesis that a single point mutation does not significantly affect the structure of a protein or of a protein-protein complex. This is a strong approximation, because at least locally the protein structure may be modified by the substitution of an amino acid, but in general it is rather accurate. It is therefore possible to use the so-called single-trajectory approach to perform the computational alanine scanning. According to this procedure, a single molecular dynamics simulation is performed on the protein complex, then a set of snapshots is extracted from the trajectory for the ΔG_{WT} evaluation. Subsequently on the same snapshots point mutations are introduced one at a time and the $\Delta G_{mut,X}$ is calculated for each of the amino acids at the protein-protein interface, and $\Delta\Delta G_X$ is easily obtained (74). The single trajectory approach proved to be rather accurate, but it tends to overestimate $\Delta\Delta G_X$ for charged residues. This problem can be limited by associating a

dielectric constant greater than 1 (usually between 2 and 4) to the protein, mimicking the dielectric properties of protein interiors (117)

We have performed CAS using MM-GBSA approach. Also for CAS, we have considered rapamycin as belonging either to the FKBP12 or to the FRB unit, thus obtaining two ΔΔG values for each amino acid. CAS was carried out over snapshots taken at 10 ps intervals from the last 5 ns of the MD trajectory thus resulting in 500 snapshots in total. For the implementation of CAS, only amino acids located at the protein-protein interface were considered. Thus, the so defined interface comprises 38 residues in total, 21 residues on the FKBP12 and 17 on the FRB protein. From the analysis of the CAS results, it is clear that the binding free energy is not distributed uniformly on the protein-protein interface, but it is concentrated to some critical residues (**Figure 11**). These amino acids tend to be grouped into small clusters on the interface and they are classified on the basis of difference in binding free energy between the wild type and mutant type as hot spots ($\Delta\Delta G > 4$ kcal mol^{-1}), warm spots (2 kcal $mol^{-1} < \Delta\Delta G < 4$ kcal mol^{-1}) and cold spots ($\Delta\Delta G < 2$ kcal mol^{-1}).

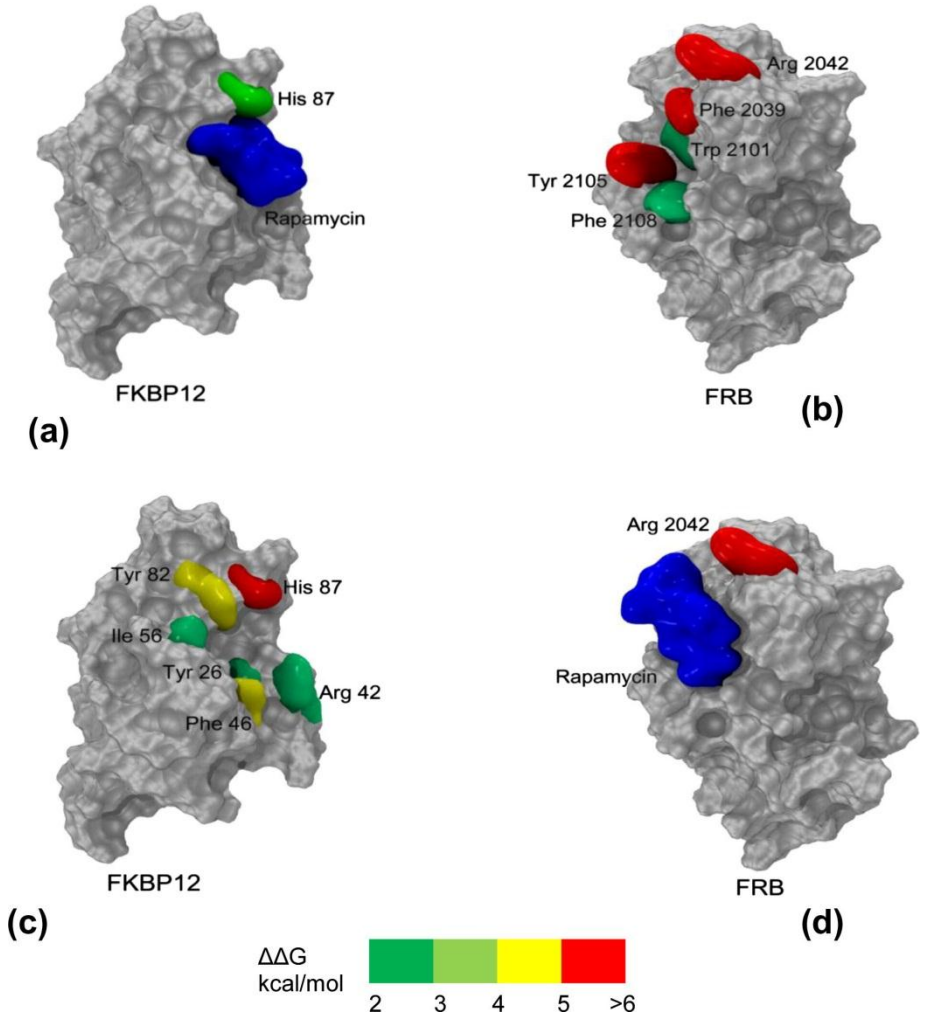

Figure 11: Hot and warm spot distribution at protein-protein interface. Different colors from green to red refer to the variation in the binding energy of complex upon mutation of each of the residues at interface. Hot and warm spots thus identified are shown. Figure 11a and 11b represents the distribution of hot and warm spots when rapamycin is considered as a part of FKBP12. Figure 11c and 11d represents the distribution of hot and warm spots when rapamycin is considered as a part of FRB. Rapamycin is shown in blue.

It is clear from **Figure 11a** that, despite rapamycin occupies only half of the buried surface of FKBP12, the protein shows only one warm spot, whereas on its binding partner FRB (**Figure 11b**), there are several hot and warm spots distributed at the

interface. Similarly, when rapamycin is considered as a part of the FRB unit (**Figure 11d**) and covers a similar proportion of the protein buried surface, FRB shows only one hot spot while protein FKBP12 exhibits a few hot and warm spots (**Figure 11c**). The two sets of calculated ΔΔG values for the 38 residues building up the interface are reported in **Figure 12**.

It can be observed in **Figure 12** (ΔΔG values are given in **Table 2**)that most of the residues exhibited a very large difference in their ΔΔG values depending on which protein (FKBP12 or FRB) rapamycin is associated to in the CAS procedure (red and blue bars respectively). This indicates that the residue which is estimated as hot/warm spot when rapamycin is considered a part of the FKBP12, are not found to be a hot/warm spot when it is a part of the FRB unit and vice versa. Thus, it implies that most of the amino acids at the interface interact with rapamycin. In very few residues, there is a minor reduction in ΔΔG values for the two cases considered, which suggests that these residues directly interact with the protein partner. Interestingly, out of 15 hot/warm spots found at the protein interface, only two, namely His_{87} and Arg_{2042} are found in both sets of data, thus mediating a direct protein-protein interaction.

Residues	FKBP12.RAP + FRB		FKBP12 + RAP.FRB	
	ΔΔG (Kcal mol^{-1})	S.E.M.	ΔΔG (Kcal mol^{-1})	S.E.M.
Thr_{21}	0	0.001	0.01	0.001
Tyr_{26}	0.1	0.003	**2.75**	0.084
Phe_{36}	0.09	0.003	1.79	0.086
Asp_{37}	0.17	0.014	1.43	0.176
Arg_{42}	1.72	0.163	**2.56**	0.189
Lys_{44}	0.27	0.064	0.26	0.064
Phe_{46}	1.05	0.068	**4.37**	0.112
Lys_{47}	0.38	0.045	0.41	0.047
Phe_{48}	0.07	0.001	0.42	0.011
Gln_{53}	-0.03	0.003	0.28	0.027
Glu_{54}	-0.18	0.016	1.39	0.103
Val_{55}	0.06	0.001	1.72	0.069
Ile_{56}	0.08	0.003	**2.54**	0.077
Trp_{59}	0.08	0.001	**3.4**	0.074
Tyr_{82}	0.69	0.037	**4.91**	0.196
Thr_{85}	0.65	0.069	0.65	0.069
His_{87}	**3.07**	0.147	**5.04**	0.123
Pro_{88}	1.5	0.089	1.58	0.090
Ile_{90}	0.61	0.061	1.69	0.097
Ile_{91}	0.07	0.001	0.68	0.064
Phe_{99}	0.05	0.001	**2.05**	0.079
Leu_{2031}	0.4	0.019	0.02	0.001
Glu_{2032}	1.05	0.084	0.14	0.027
Ser_{2035}	0.82	0.099	0.04	0.001
Arg_{2036}	0.91	0.053	0.02	0.006
Tyr_{2038}	1.25	0.063	0.86	0.061
Phe_{2039}	**6.83**	0.116	1.07	0.056
Arg_{2042}	**6.22**	0.232	**6.12**	0.230
Val_{2094}	1.4	0.092	1.37	0.092
Lys_{2095}	0.81	0.172	0.8	0.170
Thr_{2098}	1.5	0.092	0.43	0.034
Trp_{2101}	**2.62**	0.082	0.15	0.006
Asp_{2102}	1.17	0.128	0.91	0.123
Tyr_{2104}	0.17	0.004	0.02	0.001

Tyr_{2105}	**6.17**	0.136	1.97	0.094
His_{2106}	0.15	0.021	0.13	0.021
Phe_{2108}	**2.87**	0.108	0.14	0.012
Arg_{2109}	0.83	0.095	0.63	0.099

Table 2: ΔΔG values for the residues at the FKBP12-FRB interface in both cases (FKBP12.RAP + FRB) and (FKBP12 + RAP.FRB) calculated by Computational Alanine Scanning. Hot and warm spots are reported in bold characters.

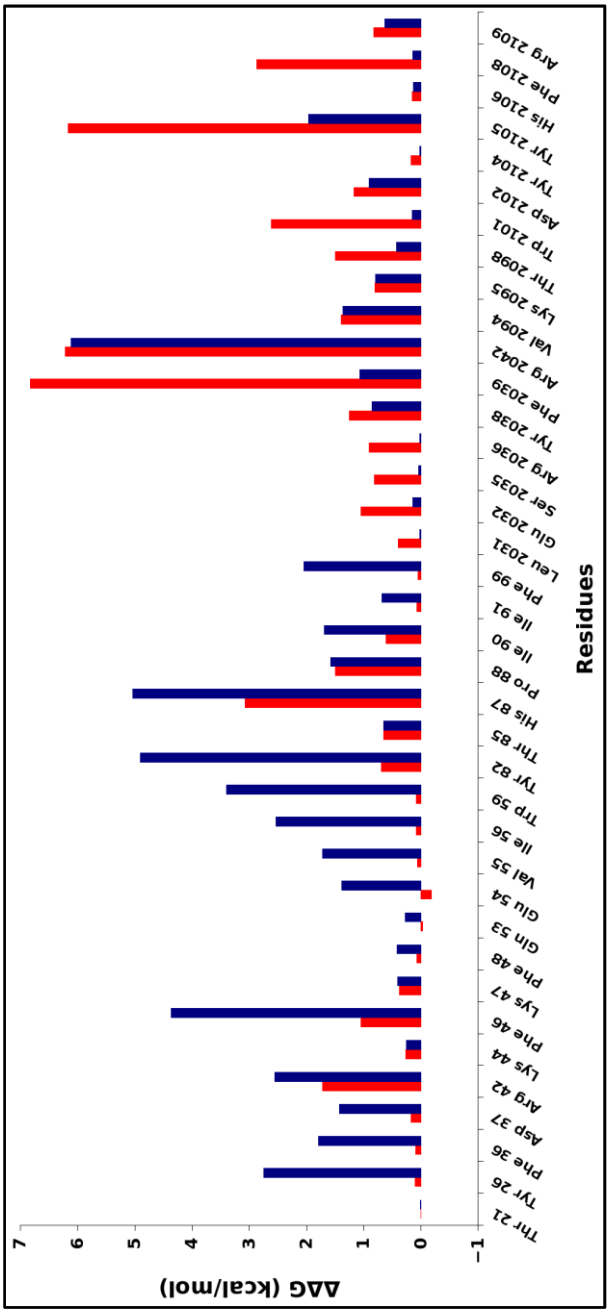

Figure 12: ΔΔG values obtained from CAS considering rapamycin as a part of FKBP12 (red bars) or as a part of FRB (blue bars).

4.4 Analysis of the Interaction Network

On analyzing the environment of CAS estimated hot and warm spots during the simulation, it is found that they participates in a network of interactions including hydrogen bonding, hydrophobic contacts and salt bridges. Rapamycin binding site on FKBP12 is delimited by a cluster of residues: Trp_{59}, Phe_{99}, Tyr_{26} and Phe_{46}. These residues are estimated to be hot or warm spots by our CAS studies and they interact with pipecolinyl ring (C2 to N7, **Figure 9**) of rapamycin hydrophobically (**Figure 13c**). This agrees with the crystallographic structure of the ternary complex reported by Choi et al (46) and also with the crystal structure of the FKBP12-rapamycin binary complex (122).

The most prominent hot spots belonging to the FKBP12 protein are Phe_{46}, Tyr_{82} and His_{87}. These residues were characterized by high $\Delta\Delta G$ values (4.37 kcal mol^{-1}, 4.91 kcal mol^{-1} and 5.04 kcal mol^{-1} respectively). Phe_{46} hydrophobically interacts, as we pointed out, with rapamycin, while Tyr_{82} interacts via a stable hydrogen bond. Hydroxyl group of side chain of Tyr_{82} forms a hydrogen bond (average bond length 2.65 Å) with carbonyl oxygen at C8 of rapamycin that is populated for 82% of the analyzed MD trajectory (**Figure 13a**), which is in agreement with crystallographic structure (46), (109). His_{87} which is predicted to be a hot spot (refer to **Table 2**), interacts with pyranose ring of rapamycin hydrophobically. Interestingly, His_{87} has been experimentally found to interact with the pyranose ring of the ligand FK506 which is structurally similar to rapamycin (123). In addition, backbone nitrogen of Ile_{56}, a warm spot ($\Delta\Delta G$ 2.54 kcal mol^{-1}) interacts with carbonyl oxygen at C1 of rapamycin via a further hydrogen bond (**Figure 13b**). The average bond length of this hydrogen bond is 2.9 Å and is found to be populated for 73% of the analyzed trajectory. This interaction agrees with the crystallographic structure (46), (109). Three more hydrogen bond interactions have been observed between FKBP12 and rapamycin, namely Gln_{53} main chain carbonyl with C40 hydroxyl of rapamycin, Glu_{54} main chain carbonyl with C28 hydroxyl and Asp_{37} carboxyl oxygen with C10 hydroxyl. These interactions are predicted by CAS to give

only a little contribution to the binding of FKBP12 and rapamycin, considering that the aforementioned residues exhibit ΔΔG values around 1 kcal mol^{-1} in our study.

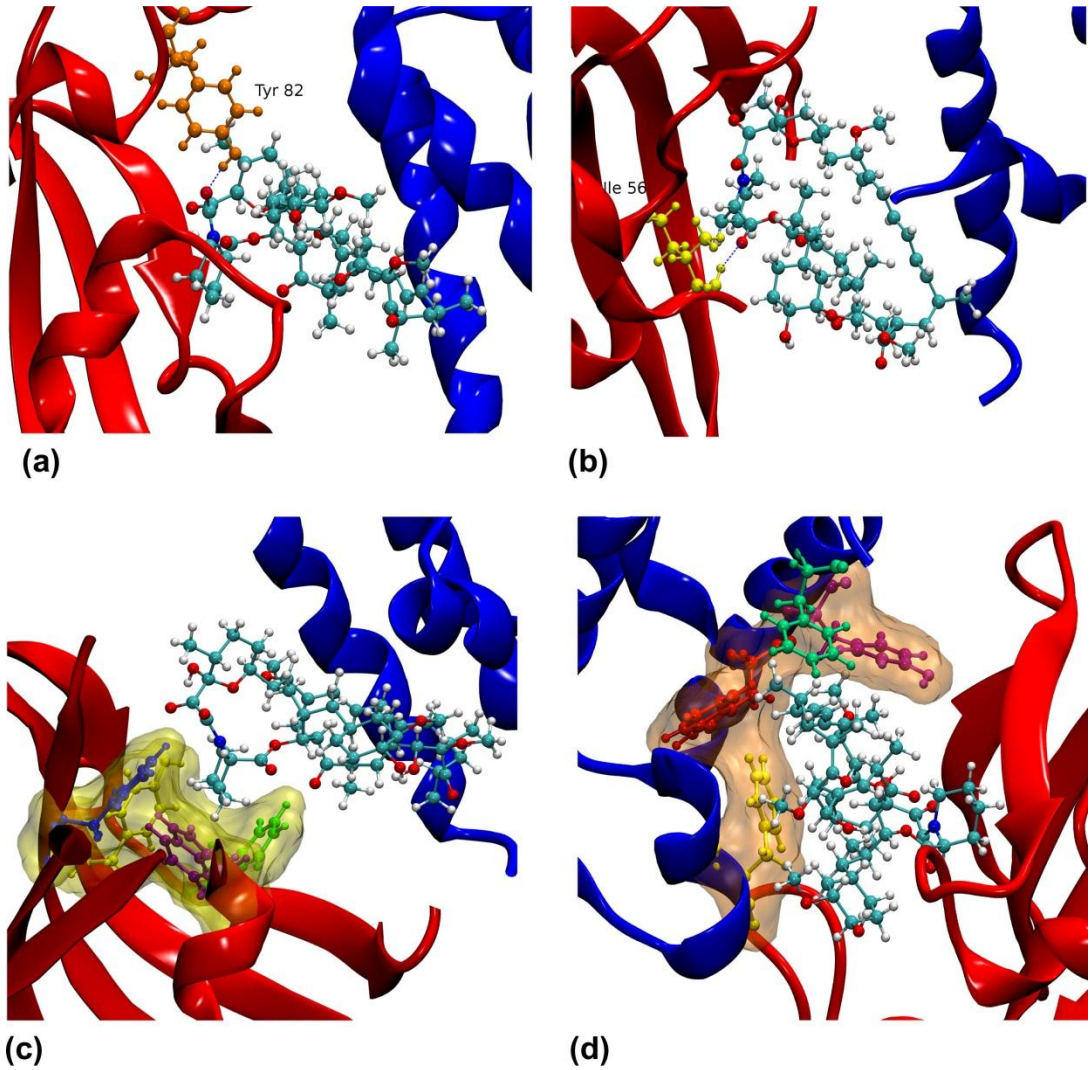

Figure 13: Details of network of interactions between FKBP12 and rapamycin (a,b and c)and between FRB and rapamycin (d) are displayed. Key hydrogen bonds are highlighted between FKBP12 and rapamycin (a and b). Hydrogen bonds are represented with blue dotted lines. A network of residues are involved in hydrophobic contacts between rapamycin and FKBP12 (c) where Trp_{59} is represented in purple, Tyr_{26} is represented in yellow, Phe_{46} is represented in green

and Phe_{99} is represented in blue. A network of residues involved in hydrophobic contacts between rapamycin and FRB are shown (d) where Phe_{2039} is represented in yellow, Trp_{2101} in red, Tyr_{2105} in purple and Phe_{2108} in green. FKBP12 protein is shown in red cartoon representation. FRB protein is shown in blue cartoon representation. Key residues are shown in ball and stick model. Rapamycin is shown in ball and stick representation in standard atom colors with Carbon atoms colored in cyan, nitrogen in blue, oxygen in red and hydrogen in white.

The most important hot spots on the interface of FRB are Phe_{2039} and Tyr_{2105} as represented by their high $\Delta\Delta G$ values 6.83 kcal mol^{-1} and 6.17 kcal mol^{-1} respectively. Other important residues that are estimated to be warm spots are Trp_{2101} and Phe_{2108}. These aromatic residues, that lined a hydrophobic cavity on the surface of FRB, hydrophobically interact with triene arm of rapamycin (C16 to C23, **Figure 9**). This cavity is shallower than the rapamycin binding pocket on FKBP12 and the interactions between FRB and rapamycin are completely hydrophobic in nature (**Figure 13d**).

The results obtained in this study have shown that there are extensive interactions between FKBP12 and rapamycin and between FRB and rapamycin, but very limited interactions are observed between the two proteins. Only two significant direct interactions between FKBP12 and FRB have been observed. One of them is the hydrogen bond between the side chain of Arg_{2042} of FRB and main chain carbonyl oxygen of Gly_{86} of FKBP12 (**Figure 14a**). This hydrogen bond is populated for 81% of the analyzed trajectory and the average bond length is 2.83 Å. Arg_{2042} is estimated to be a hot spot with $\Delta\Delta G$ value of 6.22 kcal mol^{-1} by CAS studies. This hydrogen bond interaction was not observed in previous crystallographic studies (46), (109). In addition, side chain of Arg_{42} on FKBP12 formed a stable salt bridge with side chain of Asp_{2102} during the whole trajectory, (109) (**Figure 14b**). Analysis of the CAS results and interaction network analysis of this ternary complex are in agreement with the experimental observation by L.A. Banaszynski et al (85). With the help of fluorescence polarization and surface plasmon resonance techniques, they have shown that the complex is stable only in the presence of rapamycin since key residues on the interface of the two proteins strongly interact with rapamycin and make few interactions with one

another. This is also coherent with the bridging action of rapamycin suggested by binding free energy calculations described in this work.

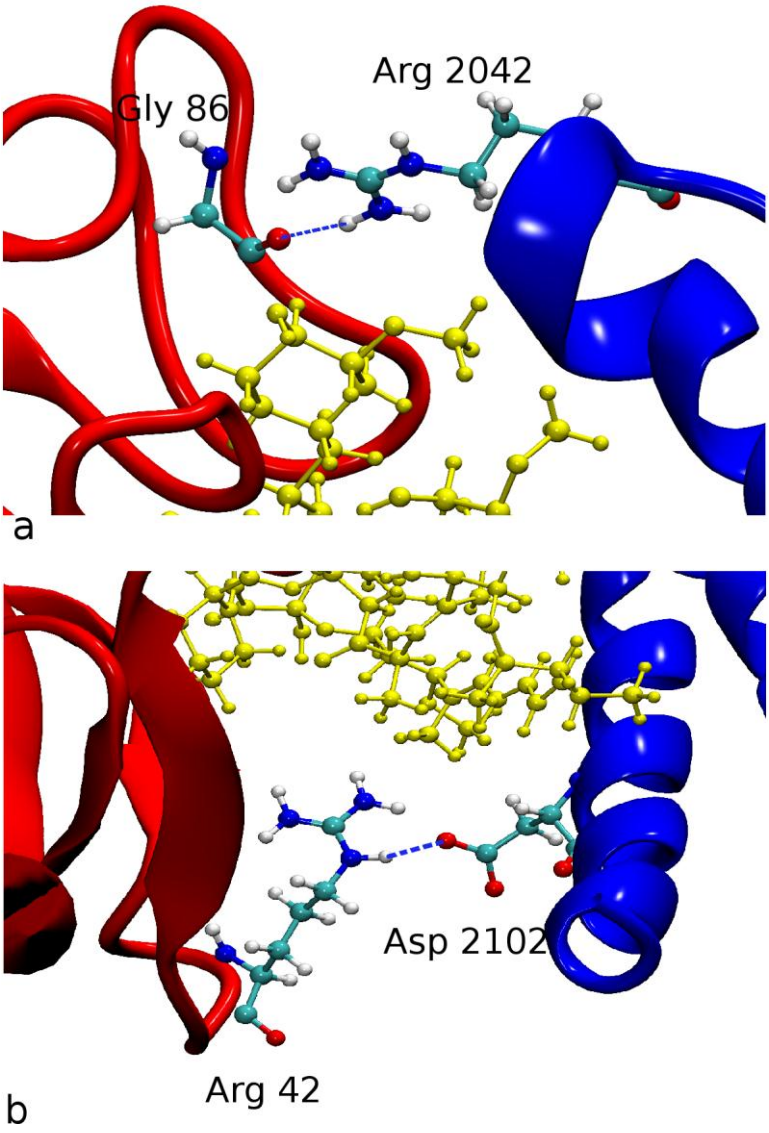

Figure 14: Details of interaction network between FKBP12 and FRB. Hydrogen bond is highlighted between Gly_{86} of FKBP12 and Arg_{2042} of FRB (a). Salt bridge between Arg_{42} of FKBP12 and Asp_{2102} of FRB is depicted (b). Rapamycin is represented in yellow.

Conclusions

In the present work, we described how a small molecule can induce the formation of an otherwise unstable protein–protein complex by a direct, bridging mechanism, even if it covers only a limited portion of the protein–protein contact surface. The reported results and the good performance of standard molecular modeling techniques in describing our model system can be interesting not only in the design and development of improved molecules acting as FKBP12–FRB protein interaction stabilizers, but also in the somehow neglected study of PPIs stabilizers in general.

PART II

Computational Modeling of the Application of Mechanical Force to the Biomolecules

CHAPTER 5

COMPUTATIONAL METHODS II

5.1 Steered Molecular Dynamics

Some biological processes are slow and inaccessible at natural time scale. They can be made accessible to the molecular dynamics time scale by Steered Molecular Dynamics (SMD) technique. For instance, the unbinding of ligand from the receptor, conformational change of the biomolecules, protein folding and unfolding, study of dissociation pathway of a protein dimer and explanation of mechanics of biopolymers. Some of the applications will be discussed later in this section.

Despite an abundance of computational methods for the protein/enzyme–ligand interaction and protein-protein interaction, little is known about their pathways; i.e how a ligand binds to a protein or how the two proteins interact with each other. Presently, the main point of view concerning molecular simulations describing ligand binding and determining binding affinities of ligands, is to try for the ideal of *reversibility*, which is realized in umbrella sampling and free energy calculations (124), (125), (126). But reaching this ideal is a slow process and requires expensive simulations. A new computational method which accepts *irreversibility* is SMD, is useful in this case (127).

In short, SMD accelerates the processes to simulation time scales (ns) with the help of time-dependent external force applied on the system.

The fundamental idea behind SMD is inspired from the experimental techniques which are based on the principle of application of mechanical forces to single molecule in small assemblies. These techniques are Atomic Force Microscopy (AFM) (128), optical tweezers (129) and biomembrane force probe (130). The principle of AFM is described in **Figure 15.** These techniques have been applied to the study of binding properties of biomolecules and their response to external mechanical stimulations (131).

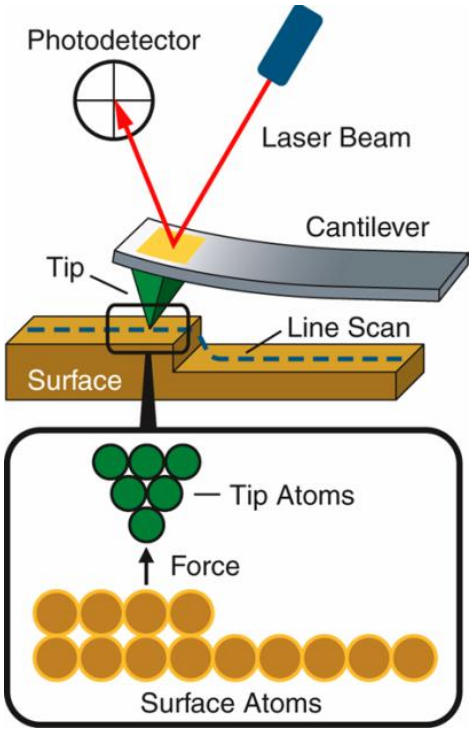

Figure 15: An illustration of the working principle of AFM. This figure is adapted from (132)

In SMD, external force is applied to the system to perform the desired process. In the case of dissociation of a complex comprising two interacting species, a series of configurations are generated along a reaction coordinate, between two interacting species. One of the species is considered as reference, and the other species is placed at increasing center of mass (COM) distance from the reference (133). SMD simulations or COM pulling can be used to bias the behavior of a system towards a particular phenomenon that is not accessible on the time scale of conventional MD. The external force applied to one of the interacting species is required to restrain that species to a point in space (restraint point) by an external potential (e.g. harmonic potential). The restraint point is then shifted in a selected direction, forcing that species to move from initial position and explore the new positions in the dissociation path (134), (135), (136). In the case of protein/peptide unfolding, external potential is applied on a terminal atom of the structure while the other terminal atom is considered fixed.

Assume a reaction coordinate *x*, and an external potential *U*,

$$\boldsymbol{U = K\frac{(x - x_0)^2}{2}}$$

Where, *K* is the stiffness of the restraint, $\boldsymbol{x_0}$ is the initial position of the restraint point moving with a constant velocity *v*, the external force applied on the system can be expressed as:

$$\boldsymbol{F = K\,(x_0 + vt - x)}$$

F represents the force by which the interacting species is pulled by a harmonic spring of spring constant *K* with its end moving with velocity *v*.

Forces apart from harmonic potential can also be used according to the processes concerned, e.g. constant forces, surface tension or torque. SMD simulation requires selection of the path, i.e. a series of directions of the applied force. For some cases, path could be a straight line and for some other cases, the path is not straight. In study of the unbinding of avidin-biotin complex (137) and unfolding of titin immunoglobulin domains (138) the path of dissociation and unfolding respectively is straight. In the case of biomolecular systems where the ligand is situated at the bottom of the binding cleft, e.g. In pulling of a single bacteriorhodopsin out of a membrane (139) and unbinding of substrate from prostaglandin H_2 synthase (140), the direction of force is not straight and same throughout the unbinding process, it should be changed to avoid distortion of the surrounding protein. The direction of the force can be chosen randomly according to the structural information. Then the force is applied in the selected direction. The force applied is accepted or rejected based on the factors like conservation of the secondary structure of the protein, the magnitude of the applied force, velocity by which the ligand moves etc. (135), (127)

SMD: Basic steps

1. Generate hypothesis of the conformational transition. Select the reaction coordinate.
2. Apply forces that induce the hypothetical process during reasonable time.
3. Analyze results to test the working hypothesis.

Applications of SMD on biomolecules

Application area of SMD is vast which includes unbinding of the ligand from the protein, protein-protein interactions, extraction of lipids from membranes (141) and describing the elastic properties of the proteins. Application of an external force causes the displacement in the simulated system, which allows for the calculation of work (a

path dependent quantity). In the case of a complex, the analysis of the applied force and system's position with respect to the time; and the details about the interactions between the two interacting species gives important information about the structure-function relationship of the interacting species. On the other hand, in the case of unfolding of a peptide, analysis of the applied force and position of the 'pulled' atom (on which external force is applied) with respect to time gives insight into the stability of the peptide. SMD not only yields qualitative information, but also quantitative information about the binding free energy or the unfolding free energy depending upon the case. Two applications are discussed briefly as follows:

1. ***Unbinding of biotin from Avidin-biotin Complex*** (137): Avidin-Biotin is a protein-ligand system which has been studied extensively by Atomic Force Microscopy experiments. In 1997, Israilev et al have studied the microscopic detail of unbinding of the ligand biotin from protein avidin by using SMD technique with force constants close to those of AFM cantilevers. They have compared the results of SMD simulation with that of AFM experiments. Avidin is a tetrameric glycoprotein and Biotin is a vitamin molecule which acts as an activated CO_2 carrier in some physiological pathways. Avidin can bind upto four molecules of biotin due to its tetrameric structure. The binding affinity of biotin to avidin is reported enormously high (binding free energy of about 20 kcal/mol (142), (143) and binding constant $k_d = 10^{-15}$ (144). In this simulation, dissociation was performed by means of harmonic restraint (K ranging from 0 to 120pN/Å). The force grows almost linearly with time, similar to the results reported in AFM experiments (145). They have shown that during the process of unbinding, biotin move in discrete steps due to the formation and rupture of hydrogen bonds at each step. The results found were consistent with the experimental studies.

2. ***Unfolding of I1 domain of Titin*** (146): One of the applications of SMD is to explore the energetic barriers to mechanical unfolding of proteins, which is in principle similar to the mechanical pulling of a single molecule in AFM. Although MD simulations are widely applied for the study of protein dynamics, the commonly available computational power

can only simulate systems for hundreds of nanoseconds. Unfortunately, the biological processes which includes large conformational changes, takes place on timescales of milliseconds to seconds. Unfolding of titin is such a process and it has been studied by implementing SMD technique, where a force is applied to the starting conformation and new conformations were generated by along a reaction coordinate. In this study, authors have reported SMD simulations for the unfolding of I1 domain of titin. They have performed the studies on both oxidized and reduced I1 domain of titin, latter form is devoid of a disulfide bridge between two β-strands. Previously, I27 domain of titin (an immunoglobin domain) has been studied well experimentally exploiting AFM technique and computationally using SMD. (147). Mechanical pulling of I27 domain on the microscopic level is also explored using SMD (148) (149).

Due to the limitation of currently available computational resources, the timescale accessible to SMD simulation, i.e. nanoseconds, is shorter than that of AFM experiments, which is in milliseconds, by six orders of magnitude. This gap in the timescale requires a pulling velocity used in simulations about six order of magnitude faster than in experiments, which leads to the discrepancy in the unfolding forces as shown in previous studies (150), (149). This problem can be solved by using reduced pulling velocities close to experimental values, which however requires expensive computational simulations.

5.2 Umbrella Sampling and Weighted Histogram Analysis Method

The major challenge in simulations is that of sampling of the conformational space. Molecular Dynamics are just not sufficient to sample the entire conformational space unless run for a very long time. And due to the limitation of computational power available, sampling is very difficult using conventional MD. In order to measure the entire available space in larger systems, extensive sampling is needed. That is why Umbrella sampling method is used by which one can bias or force the system to sample a particular region(s) based on some reaction coordinate (ξ). (151)

The canonical partition function (Q) for a given reaction coordinate (ξ) is given by:

$$Q(\xi) = \frac{\int \delta[\xi(\mathrm{r}) - \xi] \exp[(-\beta \mathrm{E}) \mathrm{d}^{\mathrm{N}}\mathrm{r}]}{\int \exp[(-\beta \mathrm{E})\, \mathrm{d}^{\mathrm{N}}\mathrm{r}]}$$

Where, E is the potential energy.

$\beta = \frac{1}{k_b T}$ $\mathrm{k_b}$ is the Boltzmann's constant

The above equation allows for the calculation of potential of mean force (PMF) $A(\xi)$ along a reaction coordinate.

$$A(\xi) = -\frac{1}{\beta} \ln Q(\xi)$$

PMF can be described as the potential that gives an average force over all the configurations of a system. In computer simulations, if every point in phase space is visited during the simulation process, PMF can be directly calculated from molecular dynamics simulation by monitoring the distribution of the system along the reaction coordinate. But simulations run for finite lengths of time and hence exact ensemble

average cannot be achieved. Regions in conformational space around a minimum are sampled well whereas regions of higher energy are sampled rarely. To obtain a complete profile for PMF calculation these high energy regions are required.

Umbrella sampling or biased molecular dynamics is an accelerated sampling technique used to calculate free energy. This sampling technique was developed by Torrie and Valleau (152). Here, bias potentials along a reaction coordinate drive a system from one thermodynamic state to another. The bias ensures an efficient sampling along the whole reaction coordinate. The effect of this bias potential in connecting energetically separated regions gave rise to the name 'umbrella sampling'. The intermediate states are covered by windows and independent molecular dynamics simulations are performed in each window. For a better sampling, a proper choice of the reaction coordinate is more important. If the reaction coordinate misses important structural changes, it can lead to artificial change in the results. A much higher barrier may result in unfavorable path being taken. A much lower barrier may be the result of discontinuities in the path.

Unbiased free energy is calculated by

$$A_i(\xi) = -(1/\beta) ln P_i^b(\xi) - w_i(\xi) + F_i$$

Where, $w_i(\xi)$ is the bias potential for window *i* and is analytically given.

$P_i^b(\xi)$ is the normalized frequency of finding the system in the vicinity of a given value of ξ

$F_i = -(1/\beta) \ln < \exp[-\beta w_i(\xi)] >$ and is independent of ξ.

Thus, with the assumption that the sampling in each window is sufficient and appropriate choice of umbrella potential is made, the PMF can be measured as long as the window spans the whole range of ξ and unbiased PMF can be calculated.

Bias potential choice in ideal condition is such that sampling along the reaction coordinates is uniform.

i.e. $w_{opt} = -A(\xi)$.

w_{opt} is the bias potential at optimal condition.

As A(ξ) is not known *a priori*, various methods have emerged for the estimation of bias potential.

Harmonic bias potential is a commonly used bias potential for a series of windows. In each window a bias function is applied to keep the system close to the reference point. A simple harmonic bias of strength K is used:

$$w_i(\xi) = K/2\ (\xi - \xi_i^{ref})^2$$

The choice of K is a critical decision and has to be made before the simulation run. Overall K has to be large enough to drive the system over the barrier. If K is too large then configurations with high energies will be overrepresented.

Adaptive bias potential starts with an initial guess of w(ξ) and iteratively improves it to achieve a uniform distribution. Other specialized umbrella potentials have also been used.

The function F_i, a component in the calculation of the unbiased free energy potential, has to be calculated for global PMF. However this function cannot be directly obtained from sampling. A number of methods have been used to analyze the umbrella sampling simulations i.e. to combine the results of different windows in umbrella sampling; the promising one being Weighted Histogram Analysis Method (WHAM).

Weighted Histogram Analysis Method (WHAM) is an extension of the Ferrenberg and Swendsen's Multiple histogram technique (153) which was first used on biomolecules by Kumar and coworkers (154). WHAM is based on the use of multiple histograms generated from NVT simulation. Here, a histogram is a probability distribution of the total energy of the system. In such technique, histograms at different parameters or one histogram at a particular value of the specific parameter are combined

to form a single histogram. It has been implemented in various studies to determine the free energy of a system.

The main advantages of this method are

(a) built-in estimate of sampling errors,

(b) yields the best value of free energies by taking into account all simulations, thus minimizing statistical errors and

(c) allows multiple overlaps of probability distributions for obtaining better estimates of free energy differences.

The global distribution is calculated by a weighted average of the distributions of individual windows:

$$P^u(\xi) = \sum_{i}^{windows} p_i(\xi) P_i^u(\xi)$$

Where the weight p_i is chosen to minimize the statistical error, leading to

$$p_{i=\frac{a_i}{\sum_j a_j}}$$

and $a_i(\xi) = N_i \exp[-\beta w_i(\xi) + \beta F_i]$ where N_i is total number of steps sampled for window i.

F_i is calculated as:

$$\exp(-\beta F_i) = \int P^u(\xi) \exp[-\beta w_i(\xi)] d\xi$$

Since in the above equations both F_i and P^u are present, these have to be iterated until a convergence is achieved.

CHAPTER 6

MODELLING THE EFFECT OF OSMOLYTES ON PEPTIDE MECHANICAL UNFOLDING

Abstract

In this work, we have studied about how the osmolyte molecules affect the mechanical unfolding of a peptide. Mechanical unfolding has been performed by using Steered Molecular Dynamics and free energy of a model peptide has been obtained from umbrella sampling and weighted histogram analysis method. The effect of four different osmolytes on the free energy difference between the folded and the denaturated state have been calculated. The observed trend mirrors the expected behavior of the studied osmolytes and unfolding pathways analysis allows an insight in the osmolyte action mechanism.

6.1 Osmolytes and Protein Stability

Osmolytes are small soluble molecules found in the cellular fluids which maintain the cellular volume and fluid balance by the effect of osmosis. Osmolyte molecules can be classified in two different groups with respect to their effect on protein stability: osmoprotectants and denaturants. Osmoprotectants stabilize the protein or peptide in their folded state. On the other hand, denaturants favour the denatured state of the peptides or proteins. Nature developed a large number of strategies in order to allow plants and animals to survive in extremely severe condition such as extreme temperatures, high pressures or high salinity. One of these strategies is the accumulation of small, highly soluble molecules in cellular fluids in such resistant organisms, because of these molecules, the osmotic stress of cells is reduced; such molecules are osmoprotectants. Several barophilic bacteria, as *Psychromonas haladis*, proliferate optimally at 50 MPa pressure (155), while *Methanopyrus kandleri*, a hyperthermophilic organism can survive at temperature as high as 120 ºC. (156) Some angiosperms, for example, the famous resurrection plants, which shows drought tolerance and are able to 'die' in the dry season and to 'resurrect' in the rainy one (157), (158).

Osmoprotectants, also called compatible solutes due to the high concentration they can reach in living organisms without toxic effects, are small organic molecules, typically amino acids or polyols (159). These molecules have also shown activity in protein stabilization against either thermal or chemical denaturation (160). Besides their biological function, osmoprotectants have a key role in very different agricultural and industrial fields. For example, in agriculture the development of plants capable of synthesizing this type of compounds could lead to the development of species more resistant to drought and requiring lower amounts of water for irrigation, with all of the benefits easily imaginable. In the cosmetic field, osmoprotectants are used as components of products aimed at protecting the epithelial from dehydration and drying. Moreover, the ocular pharmacology could be cited as a field where osmoprotectants have

recently found application in the treatment of dry eye, to compensate extracellular hyperosmolarity without interfering with cellular metabolic processes (161).

Although the osmoprotectants are widely used in different fields, their exact molecular action to protein stability is still unknown. They can hinder both thermal and chemical denaturation, the latter induced by molecules such as urea and guanidinium chloride that acts as protein denaturants. Two different hypotheses have been developed to rationalize the effect of osmoprotectants and denaturants. One tries to explain osmoprotection and denaturation invoking a direct mechanism in which the osmolyte directly interact with the protein backbone and side-chain varying its stability (162), (163), (164), (165). The second hypothesis focus its attention to the structure and role of the solvent: denaturants weaken water structure allowing a better solvation of hydrophobic residues while osmoprotectants strengthen hydrogen bond network in the solvent achieving the opposite effect (166), (167), (168), (169), (170), (171). There are a lot of computational and experimental data supporting each theory, but which is the most probable driving force is still a matter of debate (160), (166).

The work described here is about the computational determination of the mechanical unfolding free energy difference for a reference system, a small peptide in pure water, and in the presence of different osmolytes, both osmoprotectants and denaturants. The reference peptide used is the β-hairpin fragment of G protein B1 domain (**Figure 17a**) (172). This peptide has been chosen because its folded state is stable in water at biological condition and it is largely studied both computationally and experimentally (173), (174), (175). Osmoprotectants used are glycine betaine and ectoine that should contrast the denaturation process, increasing the free energy difference, while the denaturants used are urea and guanidinium chloride which should decrease the free energy difference making denaturation easier than in pure water. Structural formula of each osmolyte is shown in **Figure 16**.

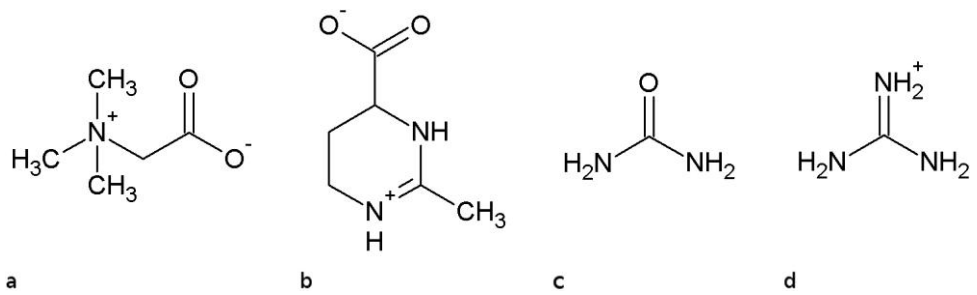

Figure 16: Structural formulae of the four used osmolytes: (a) glycine betaine (GBE); (b) ectoine (ECT); (c) urea (URE); (d) guanidinium (GUA).

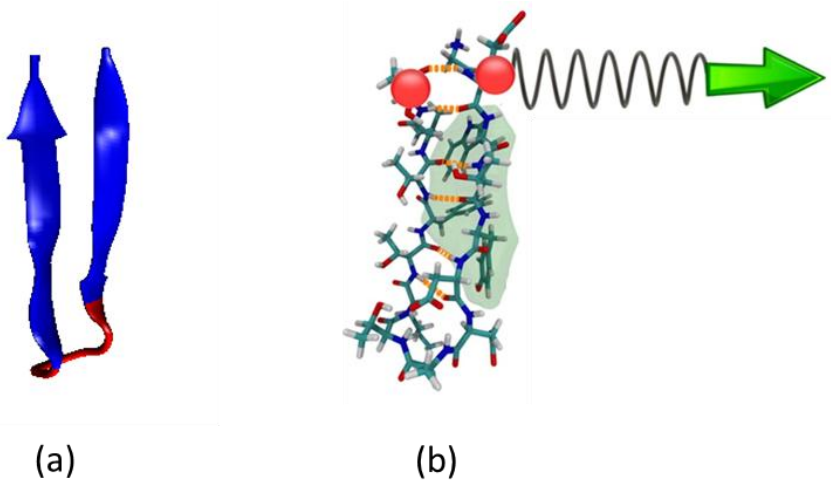

Figure 17: (a) Structure of the β-hairpin showing two antiparallel β-sheets connected by a loop. (b) The illustration representing the application of mechanical force on the peptide.

The calculated mechanical unfolding free energy corresponds to the free energy difference between an initial state, corresponding to the folded state of the peptide and a

fully extended final state. The reaction coordinate chosen to describe the process is the distance between the α carbon of the first and α carbon of the last amino acid of the peptide. In the initial state the distance is about 0.4 nm, while in the final state the distance from first and final alpha carbon is maximum, about 4 nm. The elongation process has been simulated via steered molecular dynamics (SMD) (refer to the **section 5.1**). Since, the steering is mechanical, so for this reason the obtained potential of mean force (PMF) could not be interpreted as the thermally unfolding free energy profile, but as a mechanical free energy along a specific pathway. It's worth pointing out that from SMD it is not possible to get the PMF, unless equilibrium dynamics are performed. Thus, different sampling algorithms have to be used, algorithms allowing a better sampling of the whole pathway, which results in a more accurate estimate of the PMF.

The way found to sample the entire range of distance, the reaction coordinate ξ, is the umbrella sampling method (refer to **section 5.2**) originally proposed by Torrie and Valleau (176). In this method a set of dynamics is performed, restraining the reaction coordinate to successive values of a reference distance ξ_i. The restraining potential usually employed is harmonic:

$$w_i(\xi) = \frac{K_i}{2}(\xi - \xi_i)^2$$

Where, K_i is the strength of the potential and ξ_i is the reference position of the restrain. The aim of all the dynamics is to significantly sample the complete reaction path by using the sampling algorithm, 'Umbrella sampling'. From each run a histogram is recorded as a function of ξ and the PMF is calculated based on such distribution (177). However, the calculated PMF curves are biased by the $w_i(\xi)$ potential and so the obtained histograms have to be reweighted to remove the bias. The weighted histogram analysis method (WHAM) proposed by Ferrenberg and Swendsen (153), and later applied to biological systems by Kumar (154) has been used. This method have the great advantage of providing the best possible PMF making use of all data obtained by simulations;

overlap between contiguous windows are not discarded as in other methods, but are efficiently used to improve the PMF.

6.2 Materials & Methods

The model system chosen is the β-hairpin fragment of G protein B1 domain (Protein Data Bank Id: 2GB1) corresponding to residues from GLY_{41} to GLU_{56}. The β-hairpin has been simulated in a water box of $5 \times 5 \times 8\ nm^3$ using periodic boundary condition. The largest dimension of the box corresponds to the z direction along which the β-hairpin is unfolded. The system is solvated using explicit water with the TIP4P model (113), which has been reported to contribute to a better description of the conformation of peptides in water (178). Three sodium ions were added to assure system neutrality. Electrostatic interactions were simulated using Particle Mesh Ewald (179) with a cutoff of 1.4 nm; the same cutoff was used for van der Waals interactions. This is referred to as the reference system since no osmolytes were added.

Four more systems have been set up with the same methodology. In each of them 241 molecules of osmolyte were randomly placed inside the box in order to obtain a 2 M solution. The osmolytes used are: glycine betaine (GBE) and ectoine (ECT) as osmoprotectants, urea (URE) and guanidinium chloride (GUA) as denaturants. To achieve system neutrality where guanidinium cations are used, for each molecule of GUA, a chlorine ion is added. AMBER03 (180) force field has been used to describe the protein while the osmolytes were described with the GAFF force field (53). All the molecular dynamics simulations were performed with the GROMACS molecular dynamics package version 4.5.4 (181), (182)

The first step performed was the mechanical unfolding of the β-hairpin in order to have the required different initial configurations for umbrella sampling (US). This has

been performed via steered molecular dynamics (SMD) keeping the alpha carbon of GLY_{41} fixed and pulling the alpha carbon of GLU_{56} restraining it with an harmonic potential with a force constant of 1000 kJ/(mol nm^2) to a virtual point moved linearly in the z direction with a velocity of 1.0 nm/ns. This is intended to generate structures to use in the next US stage where the potential of mean force is calculated. An illustration of SMD on the β-hairpin peptide is shown in **Figure 17b**.

For umbrella sampling, initial configurations are chosen equally spaced along the reaction coordinate. From each initial configuration, a dynamic in the NPT ensemble has been performed at 300 K and 1 bar. Temperature and pressure were kept constant to their reference values using the Nosé–Hoover thermostat (183) and the Parrinello–Rahman barostat respectively (184). The PMF profile is obtained using the weighted histogram analysis method (WHAM) and the error estimation were performed using the bayesian bootstrapping (185). The free energy difference ΔG for the mechanical unfolding processes is calculated as difference between G_{max}, the maximum value of the free energy along the unfolding pathway, and G_{min} which corresponds to the minimum. The standard deviation σ of ΔG is calculated from the sum of variances.

In order to find the optimal set of simulation parameters to have a converged PMF, preliminary tests were performed on the reference system, composed of the β-hairpin in pure water. An initial set of 20 windows of 5 ns each was used and then both the number of windows and the simulation length were increased. It has been found that increasing the number of windows from 20 to 37 (keeping unchanged the length of the MD associated to each window) caused a large variation on the ΔG of unfolding, while a further increase from 37 to 73 windows made no significant changes. Subsequently, using 73 windows, the simulation time for each MD was increased up to 15 ns. Non converged values were observed until 10 ns, while from 10 to 15 ns a quite robust trend toward convergence was observed (**Figure 18**). This set of parameters (73 windows with 15 ns long MDs) was then used to simulate all the systems considered, with an analogous convergence trend (**Figure 18**). The value of the force constant K_i of the restraining

harmonic potential was set to 1000 kJ/(mol nm^2). This value should be high enough to overcome any energy barrier that could limit the sampling. The chosen value grants a good sampling, as it can be inferred from population histograms. The population histogram for the reference system is represented in **Figure 19**. In the histogram, it is possible to notice that distributions obtained from contiguous windows well overlap, which should guarantee a good sampling along the whole reaction coordinate and thus the obtained free energy profiles converges and the error estimation is good (185). This is required to be confident in the calculated PMF (133), yet additional criteria have been recently discussed by Zhu et al about WHAM (186).

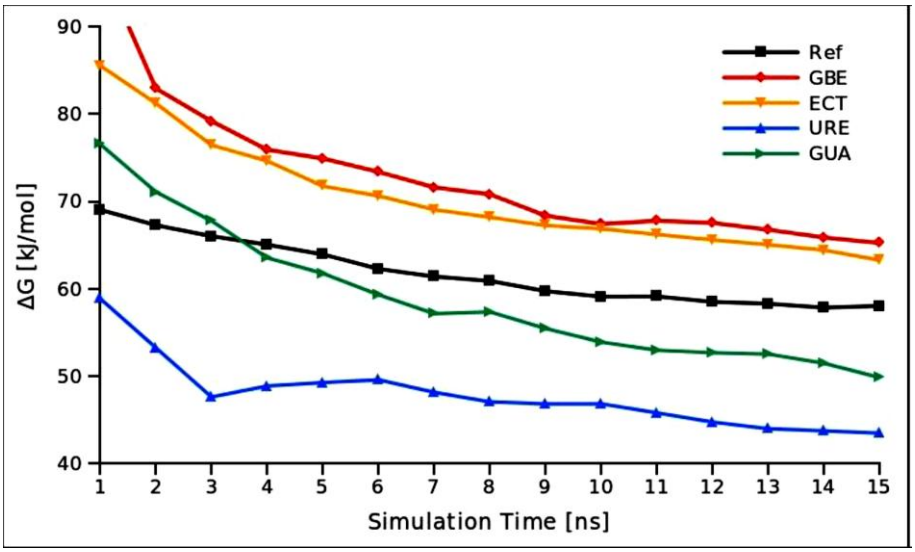

Figure 18: Profile of the unfolding free energy as a function of the simulation time. This graph shows the effect of the length of each window on the unfolding free energy. This corresponds to the set composed of 73 windows. The robust trend of convergence can be observed.

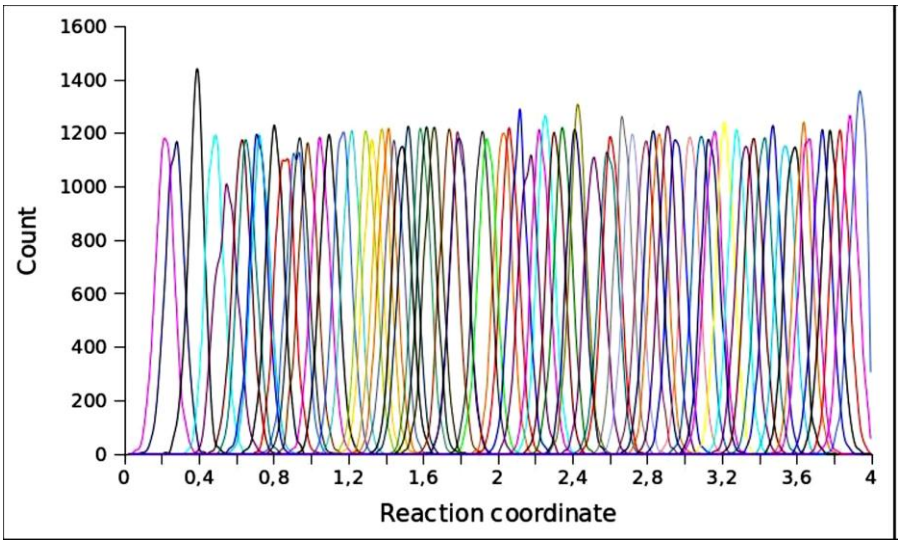

Figure 19: The histogram obtained for the reference system in umbrella sampling. It shows the number of sampled conformations in each window as a function of the reaction coordinate. Such graphs are useful to evaluate if the reaction coordinate is fully covered by multiple histograms, if all the curves are overlapping properly and to evaluate if the force constant used in the biased potential is strong enough.

The same protocol has been used to calculate mechanical unfolding ΔG for the reference β-hairpin in pure water and for the four systems in osmolytes 2M solution. To deal with the huge amount of computation to be done, about 30 years of total CPU time on a 8 cores Intel Xeon E5620 CPU @ 2.40 GHz, a BOINC server was installed and the work was distributed over a total of 386 volunteers which offer more than one thousand computers (187).

Once all the PMF have been obtained, four $\Delta\Delta G$ were calculated as $\Delta\Delta G = \Delta G - \Delta G_{ref}$, where ΔG is the free energy difference for the system with osmolytes and ΔG_{ref} is the reference ΔG for the β-hairpin in pure water. A positive $\Delta\Delta G$ is expected for osmoprotectants, since osmoprotectants stabilize proteins against

denaturation, while a negative $\Delta\Delta G$ is expected for denaturants. To evaluate how much these values significantly differ from zero the *erf* error function is used:

$$P = \frac{1}{2}\left[1 + erf\left(\frac{\Delta\Delta G}{\sigma\sqrt{2}}\right)\right]$$

Where σ is the standard deviation of $\Delta\Delta G$ and P is the probability that $\Delta\Delta G$ is greater than zero; the probability that $\Delta\Delta G$ is lower than zero is, obviously, the complementary to one.

6.3 Free Energy Calculation of Mechanical Unfolding Of the Peptide

The free energy profile of the mechanical unfolding process as a function of the distance between the α carbon of the first and of the last amino acid of the peptide, corresponding to the reaction coordinate ξ, were analyzed to evaluate similarities and differences between the five systems under consideration. The five profiles are shown in **Figure 20**. To make all the curves comparable, the minimum value of the free energy was arbitrarily put to zero in each set of data. It is observable that, taken as reference, the system with pure water, the two systems containing denaturants, urea and guanidinium chloride, always exhibit a lower free energy profile with respect to control, while osmoprotectants, glycine betaine and ectoine, have a significant increase in free energy in the range between 2.0 and 3.0 nm. This behaviour is in agreement with expectations: osmoprotectants protect the peptide making denaturation more difficult, and this result in the increase of the unfolding free energy. On the other hand, denaturants exhibit the opposite effect decreasing the unfolding free energy, as it is clearly observable in **Figure 20** where curves relative to denaturants are always below the reference.

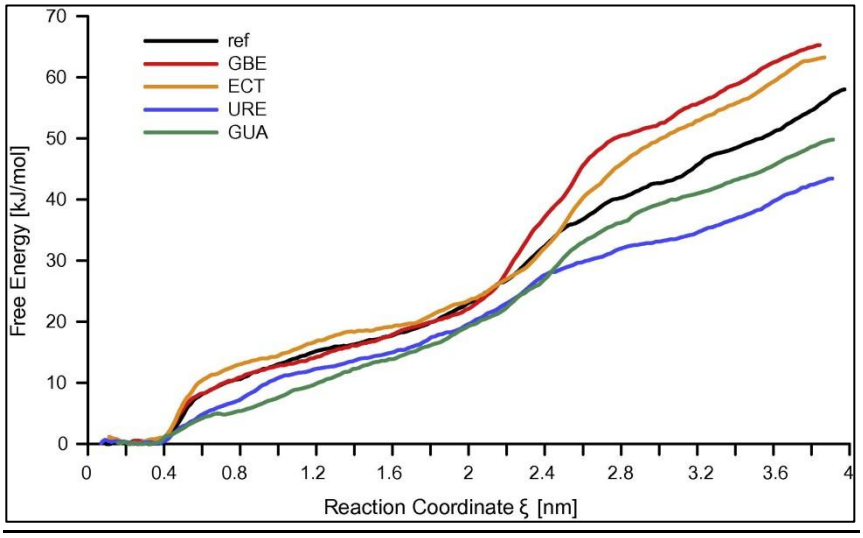

Figure 20: Free energy profile obtained for each analyzed systems as a function of the reaction coordinate ξ, i.e. the distance between the α carbon of the first and of the last amino acid of the peptide. It is worth noting that osmoprotectants curves show a significant rise in the region between 2 and 3 nm, while denaturants curves maintain a lower trend along the whole profile.

Comparing the values at the end of the mechanical denaturation pathway, for the osmoprotectants the free energy difference is much higher than the reference, while for denaturants it is noticeable lower. Focusing attention on the quantitative $\Delta\Delta G$ and its standard deviation, it is interesting to see if osmoprotectants show a significant positive $\Delta\Delta G$, while if for denaturants it is negative. The calculation of the $\Delta\Delta G$ and of its probability of being positive or negative was performed as stated in the Materials and Methods (**section 6.2**). Both osmoprotectants show a positive $\Delta\Delta G$, +7.20 ± 3.50 kJ/mol for glycine betaine and +5.24 ± 3.86 kJ/mol for ectoine, with a related probability of being positive of 98.0% and 91.3% respectively. Even better results for denaturants: guanidinium chloride show a negative $\Delta\Delta G$ of −8.21 ± 3.92 kJ/mol with a probability of being negative greater than 98% and urea obtained −14.6 ± 4.88 kJ/mol with a probability of 99.9%.

To better emphasize the similarities and differences between free energy profiles, a numerical differentiation has been performed to get the derivative of the free energy with respect to the reaction coordinate. The obtained graphs are shown in **Figure 21**. Two prominent peaks can be distinguished: the first one, very sharp, around 0.5 nm, the second, with a broader shape, between 2.0 and 3.0 nm. These peaks matches the regions along the denaturation pathway where the process involves rapid variations of the free energy and can be hypothesized to be in correspondence with regions where some significant structural transition along the mechanical unfolding path occur.

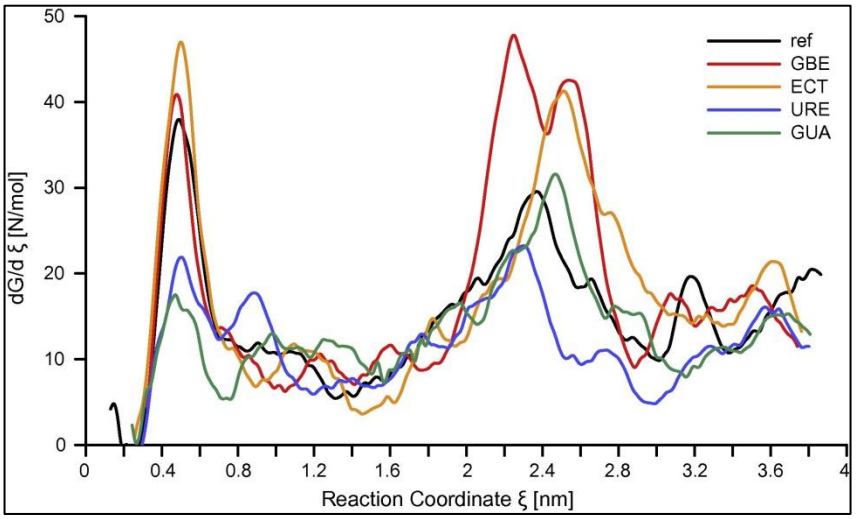

Figure 21: Derivative of free energy with respect to the reaction coordinate for the five analyzed systems as obtained by numerical differentiation of the PMF. Peaks indicates large variations in the free energy profile and focus attention towards particular sections of the reaction coordinate interesting for the unfolding process.

6.4 Structural Changes during Mechanical Unfolding of the Peptide

Although the employed methodology aims at the calculation of free energy differences, from this data it is possible to get an insight into the unfolding mechanism of the β-hairpin. In **Figure 22,** some snapshot along the reaction coordinate showing the unfolding pathway are depicted. These conformations are obtained from non-equilibrium SMD of the reference system. Similar structures are found in the SMD simulations of the peptide in osmolyte solutions, thus these snapshot are intended to be indicative of a generic mechanical unfolding process along the selected coordinate. In **Figure 22a**, the initial folded structure is shown with the six HBs highlighted in orange and a hydrophobic cluster made by residues Trp_{43}, Tyr_{45} and Phe_{52}. In **Figure 22b,** is drawn the conformation assumed immediately after the first peak in the free energy derivative profile, while in **Figure 22c** and 22**d**, the conformations assumed before and after the second peak are shown respectively. At the end, in the last snapshot, **Figure 22e**, completely steered structure is represented.

Analysing the folded structure (**Figure 22a**), it is observable that it is characterised by the presence of six hydrogen bonds (HB) between the backbones of the two strands. In addition to these strong, directional, polar interactions, residues Trp_{43}, Tyr_{45} and Phe_{52} form a little hydrophobic cluster. Looking at the conformations assumed by the peptide along the reaction coordinate, in the windows corresponding to the first peak, the breaking of the first two HBs is observed (**Figure 22b**). HBs break occurs in a very decisive manner and at 1 nm distance between peptide extremes the two HBs are already broken. This is in agreement with the sharp shape of the peak (**Figure 21**, first peak). From 2 to 3 nm instead, in correspondence with the second diffuse peak, the slow rupture of the remaining four HBs is observed, rupture concomitant with the breaking-up of the non-directional hydrophobic interaction between Trp_{43}, $Tyr_{45}$and Phe_{52} (**Figure 22c** and **22d**). This peak has a broader shape, which is coherent with the interplay between directional HBs and non-directional hydrophobic interactions (**Figure 21**, second peak).

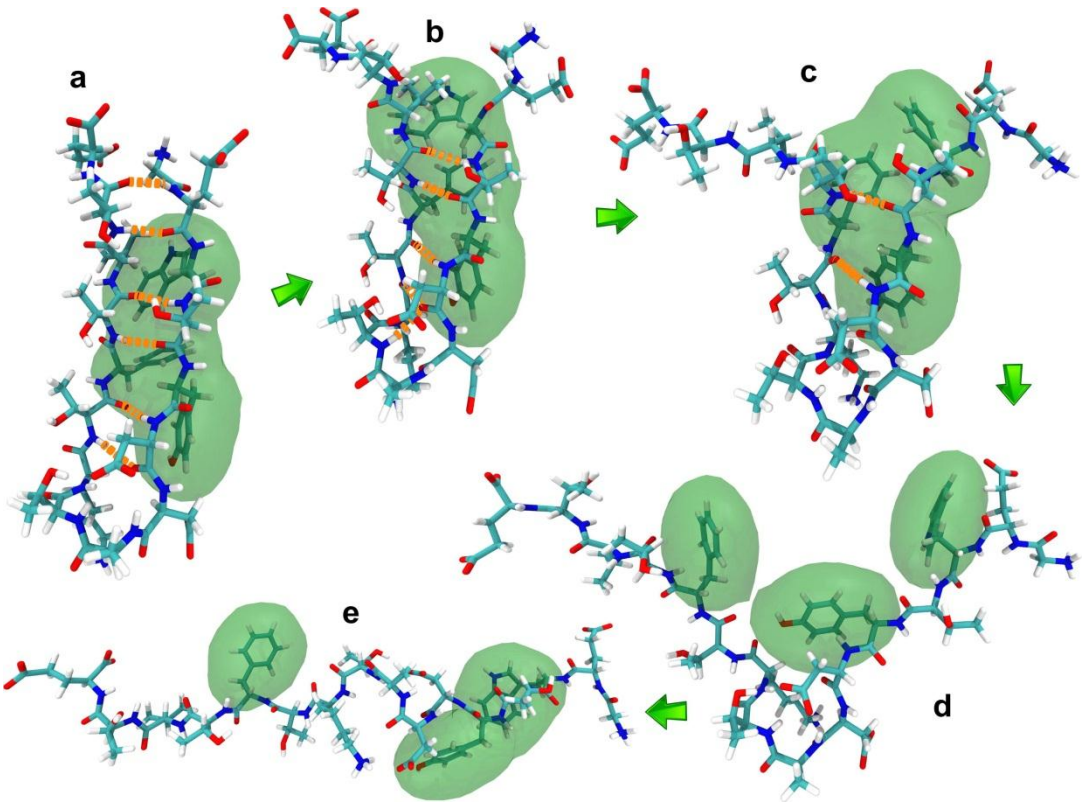

Figure 22: Some representative snapshots of the unfolding pathway. The value of the reaction coordinate in each of the snapshot in figure is: (a) 0.5 nm: the folded stated; (b) 1.3 nm: immediately after the first peak in free energy derivative profile plot; (c) 2.1 nm: before the second peak; (d) 3.0 nm: after the second peak; (e) 4.0 nm: the completely steered conformation. Figure is rendered using VMD software (45).

The differences in the behaviour of osmoprotectants and denaturants with respect to the reference in (**Figure 21**) can be associated to the steps constituting the unfolding process, which allows to get some more insight into the qualitative effect of the osmolytes on peptide unfolding. At first glance, it is observable that curves referring to osmoprotectants in **Figure 21** show higher peaks indicating larger energy barriers,

compared to denaturants curves, which shows lower profiles, sign of reduced energy barriers. Focusing on the first peak, both the reference and osmoprotectants show a high value of the free energy derivative. This could means the transition associated with the first peak, the breaking-up of the first two HBs, is a difficult task to achieve for both the system in pure water and with osmoprotectants. The curves referred to denaturants are completely different, where the maximum value of the free energy derivative corresponding to the first peak is halved, indicating that here denaturants make the transition much easier.

An opposite behaviour is observed in correspondence of the second, broader peak. Here the free energy derivative plots in the reference system and in the simulation in 2 M denaturants have very similar profiles, showing no significant effects of denaturants in the unfolding process. On the other hand the profile for osmoprotectants 2 M solutions are very different in the region of the second peak. Glycine betaine shows two high superimposed peaks, while ectoine shows a large complex peak. For both, the area subtended the peak, that is the free energy, is much larger as compared to the reference or to denaturants. This could mean in that region of the reaction coordinate, where the four innermost HB are break together with the hydrophobic core, the analysed osmoprotectants exert their primary action making this transition much more difficult.

To have a further check about this two-peak behaviour, the non-equilibrium force applied during the SMD is studied as a function of the reaction coordinate. A representative graph is shown in **Figure 23**. It is qualitatively observable that how the first peak is there immediately between 0.5 and 1.0 nm, and a large and complex second peak between 1.7 and 3.0 nm. It goes without saying that the reaction coordinate intervals in the non-equilibrium SMD force and in the free energy profile are not exactly the same, but it is worth noting how the qualitative behaviour of the non-equilibrium SMD force very well fit with the quantitative results obtained from free energy calculations.

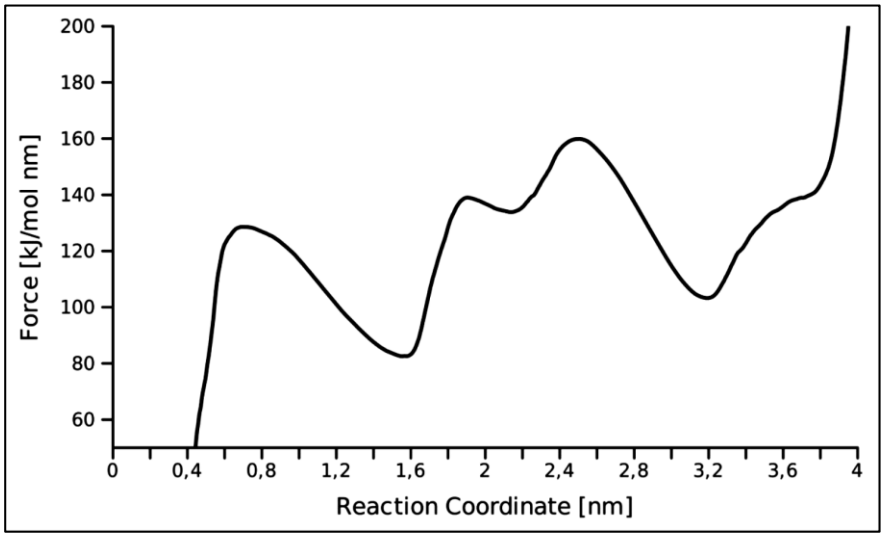

Figure 23: Profile of the force applied to the β-hairpin as a function of the reaction coordinate obtained from SMD of the reference system in pure water. The virtual point pulled is the α carbon of THR_{55}, pulled at a constant velocity of 10nm/ns. This graph is reported to make a qualitative matching with the derivative of the free energy with reference to the reaction coordinate. In both graphs, the two-peak behaviour is visible.

Conclusions

In this study, it is shown that how it is possible to reproduce the expected trend in the mechanical unfolding ΔG of a model system in mixed solutions containing either osmoprotectants or denaturants using free energy calculations. The phases of the mechanical denaturation process have been observed and the actions of different osmolytes capable of increasing or decreasing the unfolding ΔG have been discussed in term of free energy derivatives. Umbrella sampling and WHAM here demonstrate their potentiality to get powerful insight about protein stability. The results from PMF, and its derivative with respect to the reaction coordinate, are in very good agreement with the non-equilibrium SMD force and, more interestingly, with structural processes and events where the different actions of denaturants and osmoprotectants are related to hydrogen bonds breaking and hydrophobic interactions disruption along the reaction pathway.

CHAPTER 7

PROTEIN-PROTEIN INTERACTIONS AMONG TUBULIN HETERODIMERS: AN *in-silico* STUDY.

Abstract

Microtubules (MTs) constitute a key component in the cytoskeleton of the living organisms and has fundamental role in the process of cell division. Since MTs are crucial for cell division, so, they are responsible for the diseases characterized by aberrant or unregulated cell duplication, such as cancer. Therefore microtubules act as targets for the drugs against such diseases. There are several MT-targeting drugs presently available, but because of the problems related to presently available molecules, there is a need for the development of new improved MT-targeting drugs. These MT-targeting drugs target lateral and longitudinal interactions within MT which encourages us for the detailed study about MTs lateral and longitudinal interaction network at the molecular level.

In this work, we presented a study of the molecular basis of the longitudinal interactions as well as lateral interactions between the adjacent tubulin heterodimers

using an *in silico* approach steered molecular dynamics (SMD). Our study has allowed us to identify the amino acids and the critical interactions responsible for the binding of tubulin heterodimers. The observations that we have obtained are important for the design of compounds that target those interactions and acts as inhibitors by interrupting the native interactions between α subunit and β subunit of different heterodimer.

7.1 Microtubules

7.1.1 Structure and Dynamics of MT

Microtubules (MTs) are cytoskeletal structures found in all eukaryotic cells. They are involved in diverse functions like cell movement, vesicle transportation and chromosome segregation during mitosis. Mitotic spindle apparatus is made up of MTs (44). MTs are hollow cylindrical structures with an outer diameter of 25 nm. They are made up of repeating α-β tubulin heterodimers that binds head to tail to form protofilaments, which in turn bind laterally to form sheets and are gradually rolled up into a cylindrical structure (**Figure 24**). Specific number of protofilaments in turn associates in parallel to make the MT wall. MTs exist in vivo with varying number of pfs (12-16). But under physiological conditions, microtubule with 13 pfs is found most commonly (188). To the wall of MT, a variety of Microtubule associated proteins (MAPs) and motor proteins binds (189). α and β tubulin subunits are highly homologous and composed of 450 amino acids each with high sequence similarity (about 40 %). Each subunit is a compact ellipsoid of approximate dimension of 46 x 40 x 65 Å, composed of three domains: a N-terminal domain hosting a GTP (guanosine 5'-triphosphate) / GDP (guanosine 5'-diphosphate) binding site, a central small domain and a C-terminal helical domain. Both subunits contain nucleotide binding sites, where one molecule of GTP/GDP binds. The final MT structure is organized in a polar manner such that the α-tubulin subunit is exposed to one end, while the β- tubulin subunit is exposed at the

other. The two ends of MT are not equivalent: the β-capped one, known as *plus-end*, is more dynamic, faster growing and faster shrinking than the other, *minus-end* (190). Protofilament elongation, from the so-called minus-end to the plus-end, is closely related to the binding and hydrolysis of GTP by tubulin. Both α and β tubulin bind GTP, although GTP bound to α tubulin, at the interface between two monomers in the same dimer, is nonexchangeable, whereas the nucleotide bound to β tubulin, at the interface between two adjoining dimers, is exchangeable. The association between two tubulin dimers leads to the hydrolysis of the β tubulin-bound GTP, after a short lag. This β tubulin-bound GTP is buried at dimer-dimer interface and becomes nonexchangeable too. As a result, the microtubule body is made up of GDP-bound tubulin units, whereas, at the growing end, a layer of GTP-bound tubulins (the so-called GTP-cap) is present. A polymer made up of just GDP bound is unstable and would lead to the formation of curved protofilaments which eventually depolymerize (191). The GTP-cap has a great importance in stabilizing a straight protofilament, thus promoting a regular microtubule growth.

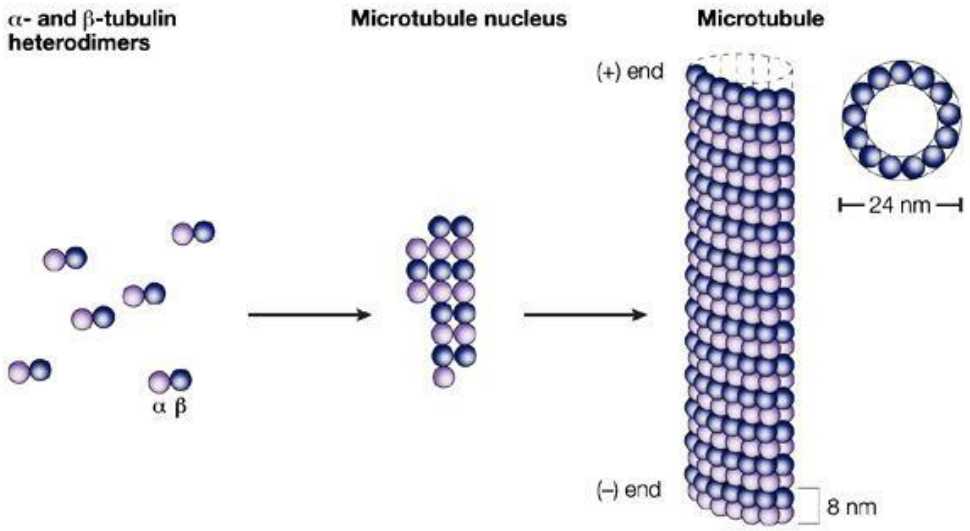

Figure 24: The figure illustrating the structure of microtubule and its polymerization process. Microtubule polymerization occurs through α-β tubulin head-to-tail association. Adapted from "Microtubule as a target for anticancer drugs" by M.A. Jordan and L. Wilson, 2004, *Nat. Rev. cancer*, 4, p. 254. (192)

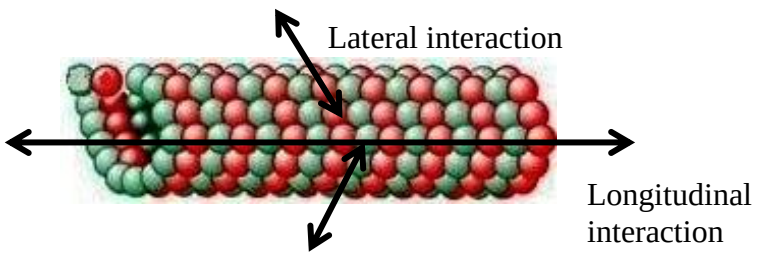

Figure 25: The illustration of longitudinal and lateral interactions. Longitudinal interactions run along a protofilament and lateral interactions are those between parallel protofilaments.

Dynamic Instability

Microtubule polymerization is a complex mechanism which involves nucleation and elongation processes. MTs are highly dynamic and can switch stochastically between growing and shrinking phases. This non-equilibrium behaviour is known as Dynamic Instability (**Figure 26**) (193), (194). Growth and shrinking of MT are cooperative processes which involves lateral and longitudinal structural transitions. When the GDP-β-tubulin terminus is exposed, as a consequence of the loss of the GTP-cap, a conformational change is triggered, resulting in rapid depolymerization of the MT, this event is called *catastrophe* (195). On the other hand, the formation of a new layer of GTP-bound β-tubulins can restore the protective GTP cap and switch off depolymerization, allowing the MT to grow again. This event is known as *rescue*.

Protofilament polymerization occurs by a delay between the head-to-tail association and the GTP hydrolysis. In this way, during the growing phase, a GTP molecule can always be present at the plus-end of the protofilaments, thus stabilizing the GDP-rich body and allowing the formation of straight microtubules. On the other hand, when the GTP cap is lost, protofilaments no longer grow in a straight conformation and quickly disassemble.

The polymerization process involves two types of contacts between tubulin subunits: longitudinal and lateral. Longitudinal interaction is among the monomers that run along the length of the MT whereas lateral interactions occur between parallel pfs that complete the MT wall (**Figure 25**). Longitudinal contacts are much stronger than lateral contacts (196). Lateral interactions are either between α-α, β-β or α-β or β-α. This depends upon the type of MTs, i.e on the number of protofilaments they are composed of.

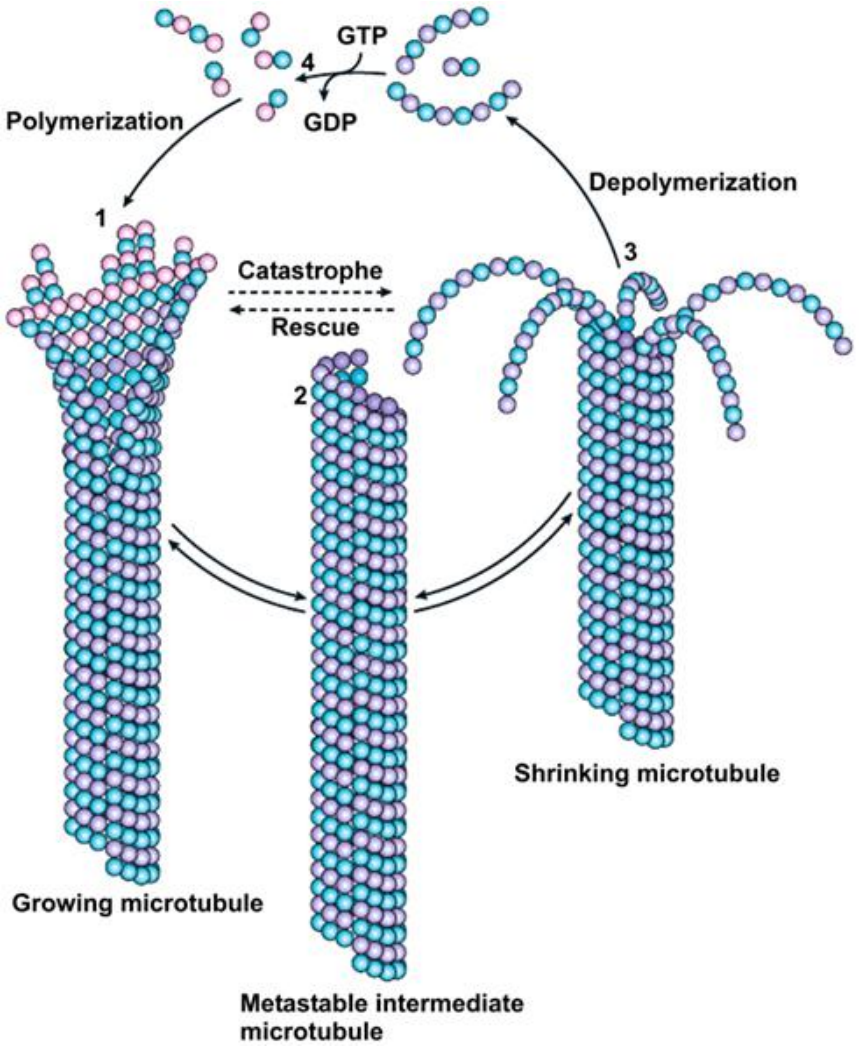

Figure 26: Illustration of dynamic instability in microtubules. MT dynamics is characterized from alternative growing and shrinking phases depending on the presence of a GTP-cap at the plus end. Adapted from "Tracking the ends: A dynamic protein network controls the fate of microtubule tips" by A. Akhmanova and M. O. Steinmetz. 2008, *Nat. Rev. Mol. Cell Bio.* 9, p. 310. (197)

7.1.2 Biological role of Microtubules

MTs are the critical component of a cell which is responsible for maintaining the cellular shape, transportation of organelles and vesicles, movement of cytoplasm within cells, but especially for the cell division (198), (189). During the process of the cell division, mitosis and meiosis, MTs forms '*spindle apparatus*' which refers to the cellular structure that segregates chromosomes between daughter cells. It is called as *mitotic spindle* during mitosis and *meiotic spindle* during meiosis. The spindle apparatus is composed of Microtubule organizing center (MTOC), kinetochore, astral microtubules, kinetochore microtubules and polar microtubules.

During the interphase in animal cells (phase of the cell cycle in which cell prepares itself for the cell division), usually just one MTOC is present, known as *centrosome* which serves as a central hub for all MTs in the cell (**Figure 27**). A centrosome is comprised of two units, called as *centrioles*. The two centrioles disengage from each other and replicate themselves during S-phase, and then during M-phase, they separate to form opposite poles of the cell. Now there are two MTOCs and each MTOC would belong to new daughter cell. A new network of microtubules (spindle apparatus) can thus grow from the newly generated organelles, extending throughout the cytoplasm. During prometaphase and metaphse, MTs are said to fall into three categories:

- Astral MT: These MTs points outwardly towards the cortex of the cell, in order to anchor the whole spindle apparatus.
- Kinetochore MTs: These MTs attach to the kinetochore of the chromatids.
- Polar MTs: These MTs are oriented parallel to each other but in opposite directions. They are crucial for pushing the spindle apparatus apart during mitosis.

Mitosis is divided into five phases based on the physical state of the chromosomes and spindle apparatus. These are prophase, prometaphase, metaphase anaphase and telophase (**Figure 27**). (199). During *prophase*, chromosomes begin to undergo condensation process that will continue till metaphase. The spindle apparatus begin to form as the two pairs of centrioles move to opposite poles and MTs begin to polymerize from the duplicated centrosomes. During *prometaphase* which is the extremely dynamic part of the cell cycle, MTs rapidly assemble and disassemble as they grow out of the centrosomes and attaches at chromosome kinetochores. Here, chromosomes are pulled in opposite directions by MTs growing out from both poles but the sister chromatids do not break apart in this phase. During *metaphase*, chromosomes assumed their most compact form and the centromeres of the cells chromosomes line up at the equator. At this phase, cells can be arrested with mitotic poisons such as colchicine. A complex checkpoint mechanism of the cell ensures that only those cells would enter anaphase which has properly assembled spindles. During *anaphase*, sister chromatids separates from each other. The kinetochore MT shortens and the chromosomes started moving towards the opposite poles. During *Telophase*, chromosomes reach the poles and they decondense as the nuclear envelop reforms around the two daughter nuclei and; finally by cytokinesis, the cell is divided into two separate daughter cells.

In short, the function of MTs during cell division is the segregation of the chromosomes. MTs plus end can grow freely in the cytoplasm reaching a length of typically 5-10μm (200) and attaches to the chromosomes. On the other hand, minus end of MTs is anchored to the centrosomes. Once a chromosome is bound to one microtubule, it then attaches to another one coming from the opposite pole. The tug-of-war generated by the microtubules coming from the opposite poles, forces the chromosomes to align along the cell equatorial plane. At that stage, microtubules depolymerization pulls the chromatids to opposite sides of the cell, so allowing the anaphase to take place and finally the formation of two daughter cells.

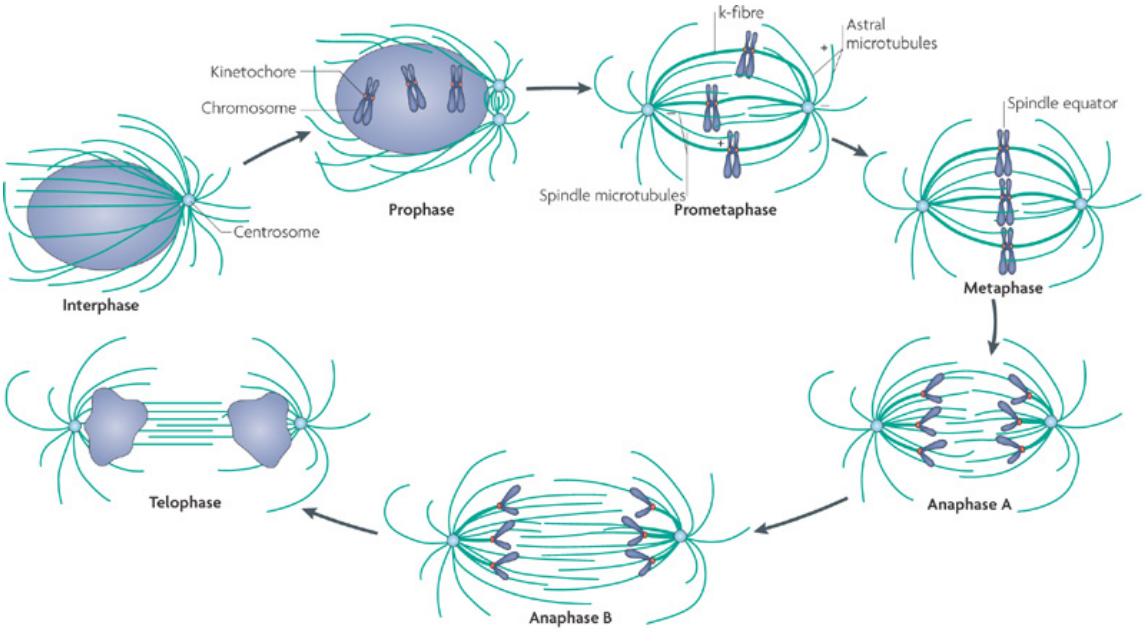

Figure 27: Illustration of the phases in the process of mitosis. Adapted from "Mechanisms of chromosome behavior during mitosis" by C.E. Walczak et al, 2010, Nat. Rev. Mol. Cell. Bio., 11, p. 93. (201)

7.1.3 Microtubules as a drug-target

As already been discussed, MTs have a fundamental role in mitosis during which the duplicated chromosomes of a cell are separated into two identical daughter cells. They are responsible for the diseases characterized by aberrant or unregulated cell duplication, such as cancer. Therefore, MTs acts as an important target for anticancer therapy (202), (200), (203). In depth study of lateral and longitudinal interactions in MT is thus essential for the development of MT-targeting molecules. These molecules interfere with the dynamics of MT arresting the cell division process in tumor cells (29).

MTs are the targets of a chemically diverse group of antimitotic drugs which binds to tubulin at different binding sites and have been used with great success in the

treatment of cancer (204), (205). MT-targeting therapeutic agents can be classified into two main categories:

1. *Microtubule destabilizing agents*: These agents inhibits MT polymerization causing loss of cellular MTs. e.g. Vinca alkaloids (vinblastine, vincristine) and cochicinoids.
2. *Microtubule stabilizing agents*: These agents promotes MT polymerization resulting in increased density of cellular MTs. e.g. taxanes (Paclitaxel) and epothilones.

The effects of MT-destabilizing drugs inhibit mitosis and are classified as *antimitotics* (200). Vinca alkaloids and colchicinoids, at high concentration, blocks MT dynamics and thus mitosis, leading to cell death by apoptosis. The other class of agents which includes paclitaxel acts by stabilizing the MT, so avoiding the transport of the chromosomes to the daughter cells.

Current mitotic agents target different binding sites on tubulin and that's why they have different mechanisms. Their mechanism of action is strongly coupled to the mechanism of MT dynamics, especially structural features that affect nucleotide binding, GTP hydrolysis and stabilization of longitudinal and lateral protofilament contacts. MT-targeting drugs target lateral and longitudinal interactions and thus any alterations in those interactions suppresses the MT dynamics. Current antimitotic agents are found to target three different sites in microtubules structure (**Figure 28**):

- The **vinca domain**: It is located at the interface between two tubulin dimers;
- The **taxol site:** It is located on the β-tubulin surface;
- The **colchicine site:** It is located at the interface between two monomers.

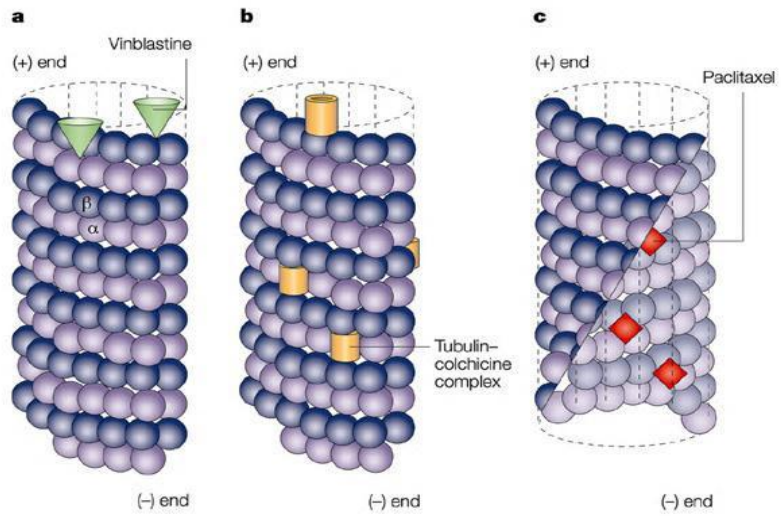

Figure 28: Representation of the binding regions of the microtubule-targeting drugs (vinca alkaloids, colchicinoids and taxanes) interfering either longitudinal interactions or lateral interactions or both. (a) Vinblastine binds to sites at MT-plus end and suppresses MT dynamics. (b) Colchicine form complexes with tubulin dimers and copolymerizes into MT lattice, suppressing MT dynamics. (c) Paclitaxel binds to the interior surface of the MT and suppresses MT dynamics. Adapted from "Microtubules as a target for anticancer drugs" by M.A. Jordan & L. Wilson, 2004, *Nat. Rev. Cancer,* 4, p. 261. (200)

Natural compounds taxanes and epothilones binds to MT at the M loop and stabilize lateral contacts between protofilaments (206), (207). Specifically, taxanes like paclitaxel binds to taxol site of the tubulin and clinically this drug is widely used against ovarian, breast and lung cancer. Vinca alkaloids alters both longitudinal and lateral interactions with the preference of suppressing the dynamics at the plus-end. They bind to different regions at the large interdimer surface (208). Colchicine site is the target for colchicinoids (209), (210).

7.2 Study of Longitudinal Interactions Among Tubulin Heterodimers

Steered Molecular Dynamics (SMD) has been used to study at atomic level the interactions required for the assembly of tubulin heterodimers. To study the longitudinal interactions among tubulin heterodimers, we have considered a portion of a microtubule protofilament composed of two adjacent tubulin heterodimers. It is done by fixing first tubulin heterodimer and pulling the other heterodimer by means of a hypothetical spring by using specialized module implemented in Gromacs (182). Then the trajectory was analyzed to predict the key interactions responsible for binding longitudinally and to describe how these interactions evolve during the course of simulation time as a function of the force applied.

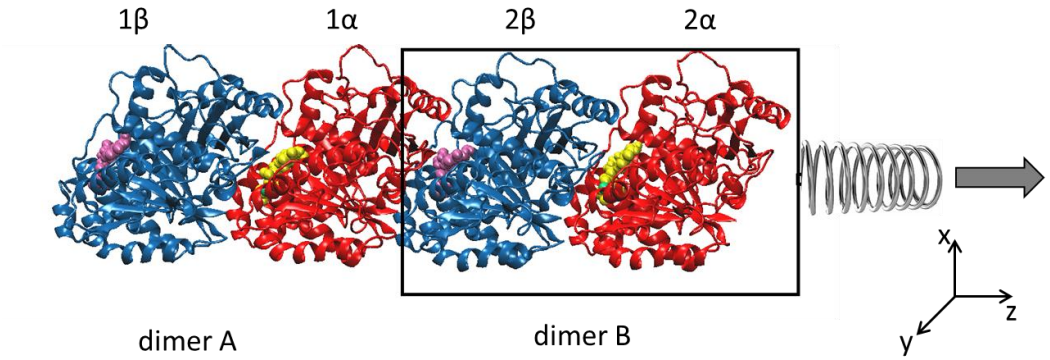

Figure 29: An illustration representing the SMD applied on a system of tubulin tetramer for the study of longitudinal interactions

7.2.1 Materials and Methods

The initial structure of tubulin heterodimer for our simulation is retrieved from Protein Data Bank (PDB id 1TUB) (211) corresponding to the structure of pig tubulin α/β heterodimer which is composed of anti-parallel protofilaments forming a planar sheet, resulting in a model specifically suitable for the study of longitudinal interactions. Then, the tetrameric structure (composed of two adjacent heterodimers) was built by adding to the crystal structure a second dimer with 80 Å translation (on the basis of crystallographic information) of the tubulin unit along the protofilament axis which corresponds to the x-axis in the PDB file.

The system is composed of two tubulin heterodimers, i.e a tetramer including GTP (Guanosine triphosphate) in each α subunit and GDP (Guanosine diphosphate) in each β subunit. Simulations were carried out with Gromacs 4.5.3 package (182), (212) and parameters from Amber 03 force field (111) were applied to the system. Explicit solvent and periodic boundary conditions were used. Short-range non-bonded interactions were cut off at 1.4 nm, with long range electrostatics using the Particle Mesh Ewald (PME) algorithm (116), (179). Firstly, topology and coordinate files were generated, then a unit cell is defined. Special consideration has been taken while defining an unit cell since it must allow enough space in the pulling direction so that it doesn't interact with the periodic image of the system. A box of size 15 X 15 X 40 nm is used. The system was solvated with TIP4P water molecules and neutralized by adding 76 Na^+ ions. The system was then energy minimized for 100,000 steps using steepest descent algorithm. Then equilibration has been performed for 100 ps with the number of particles, system volume and temperature constant (NVT ensemble) in order to equilibrate the temperature of the system. Protein and non-protein atoms were coupled to separate temperature coupling baths, and temperature was maintained at 310 K using v-

rescale coupling method (213). V-rescale is a method of temperature coupling using velocity rescaling with a stochastic term. This is similar to berendsen thermostat, but the stochastic term ensures that a proper canonical ensemble is generated. Subsequently, equilibrium of 100 ps is performed with number of particles, system pressure (1 atm) and temperature (300 K) constant (NPT ensemble) in order to equilibrate the density of the system and a restraint on the positions of C_α atoms. In this case, Berendsen weak coupling method (214) is used to maintain pressure isotropically at 1.0 bar. This structure was used as starting configuration for pulling simulation (SMD). The pulling simulation has been performed for 650 ps and snapshots were saved every 2 ps. For this, one dimer is considered as fixed (reference group, dimer A) (**Figure 29**) and the adjacent dimer is being pulled (pull group, dimer B) (**Figure 29**) with the application of an external force along a reaction coordinate z. The dimer B is pulled with a velocity 0.01 nm/ps using a force constant of 1000 kJ mol^{-1} nm^{-2}. To distinguish the dimers, we have referred α and β subunit of the first dimer as 1α and 1β and those belonging to second dimer as 2α and 2β. Nosé-Hoover thermostat (215), (216) is used to maintain temperature, and Parrinello-Rahman barostat (184) is used to isotropically regulate pressure. This combination of thermostat and barostat ensures that a true NPT ensemble is sampled. During pulling simulation, restraints were removed from the system except dimer A. It was used as an immobile reference for the pulling simulation. The main idea behind pulling simulation is to capture enough configurations along the reaction coordinate to obtain regular spacing of the umbrella sampling windows, in terms of center of mass distance between pull group and reference group. Since performing umbrella sampling calculations for this system is computationally very expensive, and it is very difficult for the present computational power that we have, that is why we had to limit the analysis at this point. But, it could be possible in the future with the availability of higher computational power.

7.2.2 Results and Discussion

Output data of SMD simulation consists of positions along z-axis at 0.001 ps intervals, force applied and center of mass (COM) distance between the dimers. By pulling the COM of dimer B away from dimer A, force builds up until a breaking point is reached, at which time critical interactions are disrupted, allowing dimer B to dissociate from dimer A. **Figure 30** represents some snapshots of the system which shows how the two heterodimers dissociate.

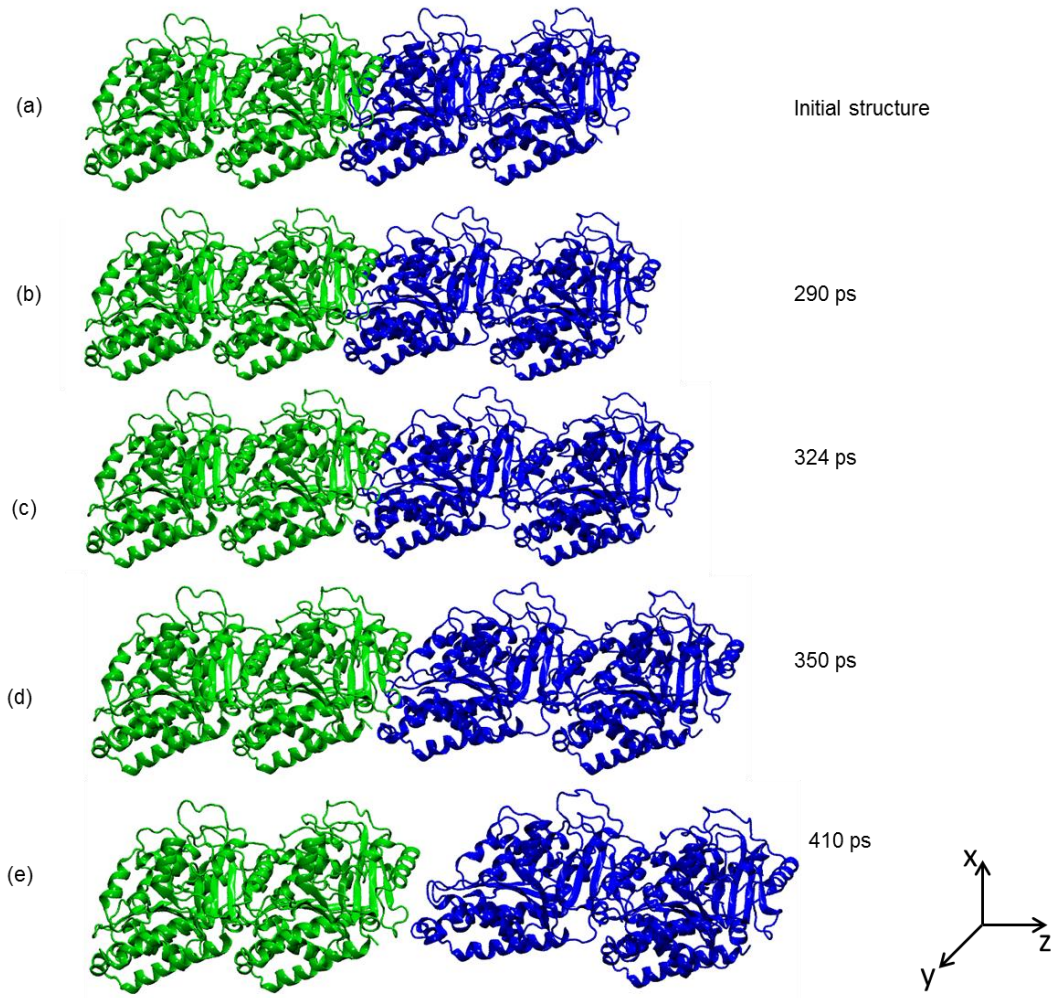

Figure 30: Representation of the snapshots of the tubulin tetramer system at different time frames during the pulling simulation. Dimer A (reference group) dimer is shown in green and Dimer B (pull group) is shown in blue. (a) Starting conformation. (b) At 290 ps. (c) At 324 ps: the time corresponding to the maximum force applied. (d) At 350 ps. (e) At 410 ps.

Force applied is plotted as a function of simulation time (**Figure 31a**). It can be clearly observed that initially the force rises till the point of maximum force (2276 kJ/mol nm) which corresponds to time 324 ps, after that point the force diminishes showing the separation of dimers. If we focus on the second plot (**Figure 31b**)

representing the profile of force applied as a function of COM distance, we see that from the initial structure (COM distance 8 nm), dissociation is very slow till COM distance of about 9 nm, which corresponds to the point of maximum force and after that there is a sharp decrease in the force applied. This is because the dimers are dissociated completely just within few ps of 324 ps, the time corresponding to the maximum force.

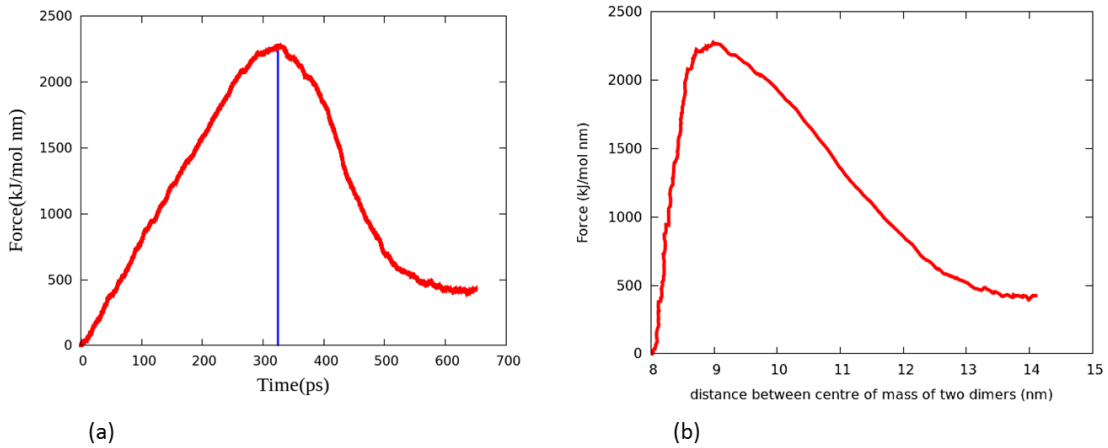

Figure 31: (a) The 'applied force' profile with respect to the time during the dissociation of the two heterodimers. The point of maximum force (2276 kJ/mol nm) is represented by blue line which is at 324ps (b) Force-extension profile: the profile of the force applied with respect to the center of mass distance between the two dimers during the dissociation process.

A deeper insight into the network of interactions between the residues at the interface that contribute to microtubule stability can be obtained by analysis of the simulation trajectory relating it with the force applied. Throughout this chapter, all the amino acids on α-tubulin subunit will be referred by their position and those on β-tubulin subunit will be referred with a '**'**' over the residue number. In this study of dissociation of two heterodimers, the point of maximum force corresponds to the instant just before the breaking of few critical interactions between the two heterodimers. One of them is the hydrogen bond between Val 353 of α subunit of first heterodimer and Asp 177' of β subunit of second heterodimer breaks. This hydrogen bond is between the backbone

Nitrogen of Val 353 and backbone Oxygen of Asp 177'. **Figure 32b** represents the dissociation of this interaction just after the point of maximum force. To support our findings, we have compared them with the observations from CAS studies previously performed by our group (30). Peraccini et al have stated the presence of the above mentioned interaction. Moreover, In the same study, CAS study of the interface residues of the tubulin tetramer has predicted Val 353 as a warm spot ($\Delta\Delta G$ = 3.33 kcal/mol) and Asp 177' as a good hot spot with high $\Delta\Delta G$ = 15.23 kcal/mol, implying both of the above mentioned amino acids are important for the longitudinal interaction among tubulin heterodimers. Also, Asp 177' belongs to the T5 loop (according to the tubulin secondary structure classification proposed by Lowe et al (44), Val 353 is a part of S9 strand; and according to the MT models (190), T5 loop and S9 strand were proposed to interact with each other on the geometrical basis.

On the basis of thorough analysis of the various interdimer interactions, we have found that after reaching the maximum force at 324 ps (**Figure 31a**), many interactions were disrupted within only 14 ps. This observation highlight the importance of those interactions in stabilizing the microtubule.

The interactions that we have observed are discussed as follows. Thr 349 of 1α subunit interact with Val 175' of 2β subunit via a hydrogen bond (Hbond). This is in agreement with the observations from the previous studies by our group by Pieraccini et al. Moreover, according to CAS study both of these residues were predicted to be hot-spot and warm spot with $\Delta\Delta G$ of 4.60 kcal/mol and 2.47 kcal/mol respectively (**Figure 33**). Backbone oxygen of Gln 256 on 1α subunit interact with sidechain nitrogen of Trp 397' on 2β subunit through a Hbond. Another Hbond is observed between sidechain oxygen of Asn 258 and backbone nitrogen of Val 179'. Sidechain oxygen of Glu 254 interacts via Hbond with sidechain nitrogen of Asn 99'. We have also observed some hydrophobic key interactions which are responsible for the stability of microtubules. Pro 325 interact with Tyr 222' and Tyr 208' via hydrophobic interactions. This is in agreement with the previous study by our group (217). $\Delta\Delta G$ values obtained from CAS studies supports these observations. The tetramer is stabilized by hydrogen-bonding and

hydrophobic interactions. The use of COM pulling allows monitoring of the evolution of these interactions over time as dimer B is pulled away from dimer A. This gives qualitative assessment of the specific interactions between the two heterodimers which stabilizes the microtubule longitudinally.

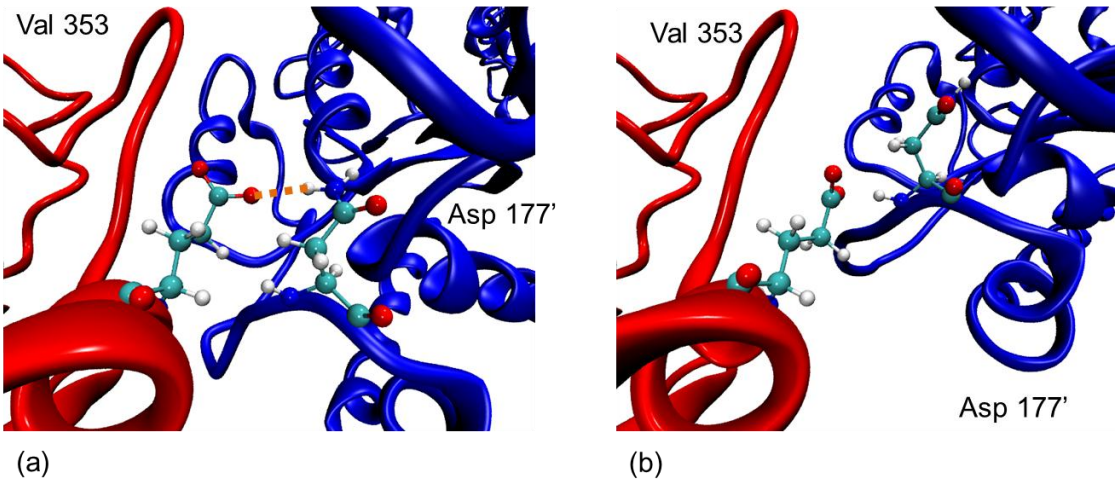

Figure 32: (a) Initial configuration of the tetramer showing Hydrogen bond between Val353 of α subunit of dimer A and Asp 177' of β subunit of dimer B. (b) Snapshot showing the breaking of Hydrogen bond between Val353 and Asp177' just after the point of maximum force (at 324 ps).

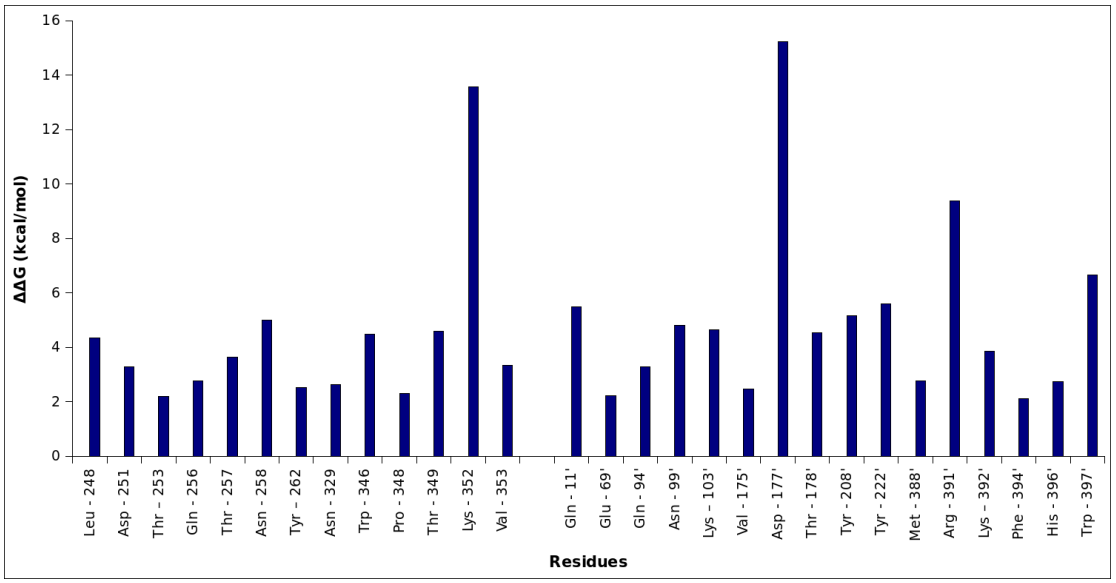

Figure 33: ΔΔG values obtained from CAS study for the residues located at the tubulin-tubulin longitudinal interface. The residues with a prime symbol (‘) belongs to β-tubulin subunit (30)

An *in-silico* analysis of the protein-protein interface was carried out with Steered Molecular Dynamics studies. We have used a harmonic potential in order to make qualitative observations about the dissociation pathway corresponding to the separation of tubulin heterodimers. The harmonic potential allows the force to vary according to the nature of the interactions of dimer A with dimer B. That is, the force will build up until certain critical interactions are broken. We were able to monitor how the interactions evolve with respect to the time during the dissociation process. This study could be valuable in better understanding of the longitudinal interactions along the protofilament that stabilize the microtubule.

7.3 Study of Lateral Interactions among Tubulin Heterodimers

MTs possess lateral flexibility to enable bending with high curvature and it is based on interprotofilament interaction or lateral interaction; and is crucial for maintaining cellular shape, cellular transport and cell division. We have studied the interaction network responsible for stabilizing the MT laterally using an in-silico approach: SMD. To study lateral interactions among MT protofilaments, we have considered a portion of MT comprising two nearby parallel protofilaments. Two tubulin heterodimers belonging to different protofilament were chosen for the study (**Figure 34**). This study helps in having a detailed view of the interactions responsible for MT stabilization and also in exploring time evolution of the crucial interactions.

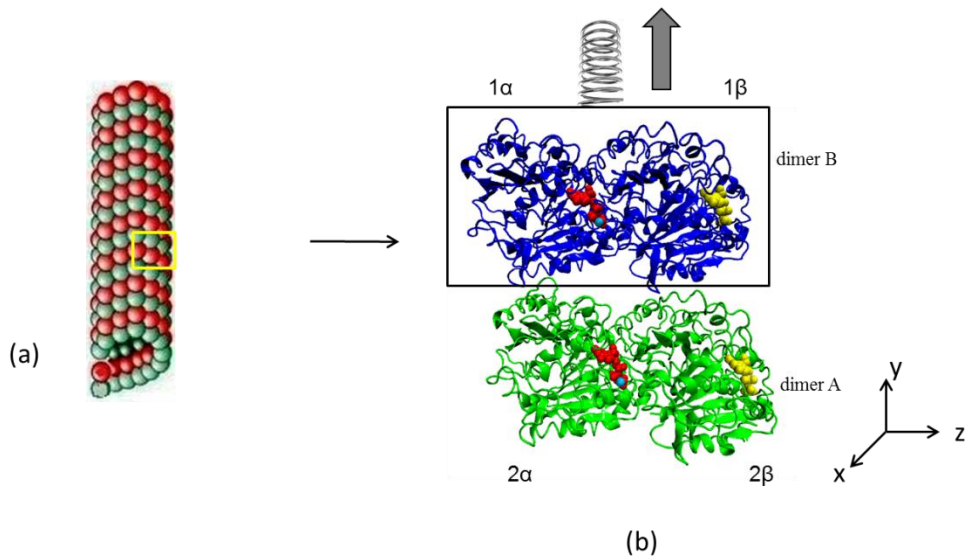

Figure 34: (a) Structure of microtubule showing the two tubulin heterodimers (in box) considered for the study. (b) An illustration representing the SMD applied on a system of tubulin tetramer for the study of lateral interactions in MT.

7.3.1 Materials and methods

The initial structure of tubulin heterodimer for our simulation is retrieved from Protein Data Bank (PDB id 1TUB) (211) corresponding to the structure of pig tubulin α/β heterodimer. Then, the system to be studied (composed of two heterodimers) was built by adding to the crystal structure a second dimer (on the basis of crystallographic information) from the adjacent protofilament.

The system is composed of two tubulin heterodimers from two adjacent protofilaments, which includes GTP (Guanosine triphosphate) in each α subunit and GDP (Guanosine diphosphate) in each β subunit. Simulations were carried out with Gromacs 4.5.3 package and parameters from Amber 03 force field were applied to the system. All the parameters used are similar to that of the longitudinal interaction study. Special consideration has been taken while defining a unit cell since it must allow enough space in the pulling direction so that it doesn't interact with the periodic image of the system. A box of size 16 X 35 X 16 nm is used. The system was solvated with TIP4P water molecules and neutralized. Energy minimization and equilibration has been performed. The structure obtained after equilibration was used as starting configuration for pulling simulation (SMD). The pulling simulation has been performed for 650 ps and snapshots were saved every 2 ps. For this, one dimer is considered as fixed (reference group, dimer A) (**Figure 34b**) and the other dimer is being pulled (pull group, dimer B) (**Figure 34b**) with the application of an external force along a reaction coordinate y. The dimer B is pulled with a velocity 0.01 nm/ps using a force constant of 1000 kJ mol^{-1} nm^{-2}. We have referred α and β subunit of the first dimer as 1α and 1β and those of second dimer as 2α and 2β. During pulling simulation, restraints were not applied on the system except dimer A which is considered as immobile reference for the pulling simulation.

7.3.2 Results and Discussion

By pulling the COM of dimer B away from dimer A, force builds up until a breaking point is reached, at which time critical interactions are disrupted, allowing dimer B to dissociate from dimer A. **Figure 35** represents some snapshots of the system which shows how the two dimers have separated.

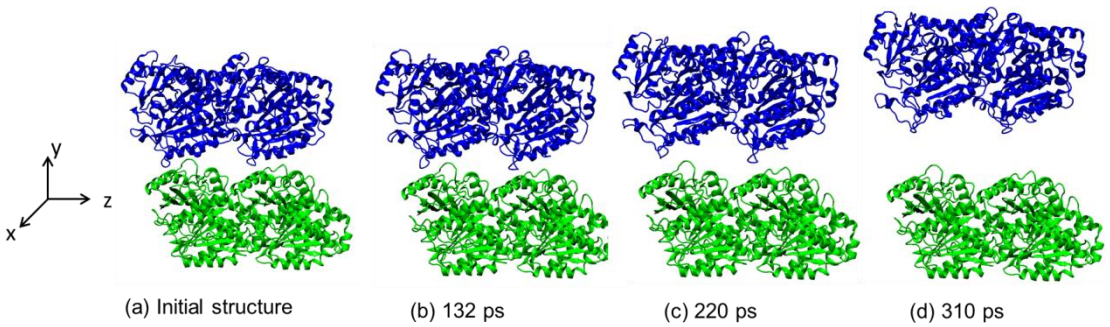

Figure 35: Representation of the snapshots of the tubulin tetramer system at different time frames during the pulling simulation. Dimer A (reference group) is shown in green and Dimer B (pull group) is shown in blue. (a) Initial structure. (b) Snapshot taken before the point of maximum force, at 132 ps. (c) Snapshot at the time of maximum force, 220 ps. (d) Snapshot taken after the point of maximum force, 310 ps.

Forces applied is plotted as a function of simulation time (**Figure 36a**). It can be clearly observed that initially the force rises till the point of maximum force (1535 kJ/mol nm) which corresponds to time 220 ps, after which the force diminishes showing the separation of dimers. If we focus on the second plot (**Figure 36b**) representing the profile of force applied as a function of COM distance, we see that from the initial structure (COM distance 5.2 nm), dissociation is very slow till COM distance of about 5.8 nm, which corresponds to the point of maximum force and after that there is a sharp decrease in the force applied. This is because the dimers are dissociated completely just within few ps of 220 ps, the time corresponding to the maximum force.

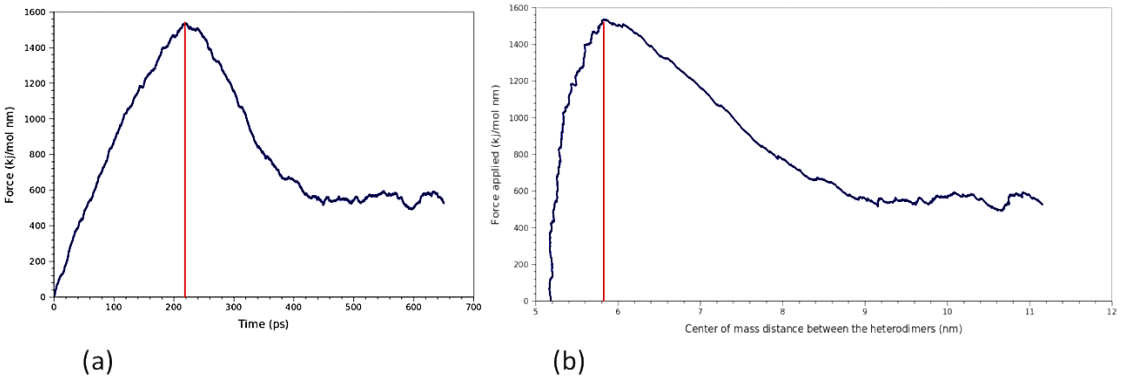

Figure 36: (a) The 'applied force' profile with respect to the time during the lateral dissociation of the two heterodimers. The point of maximum force (1535 kJ/mol nm) is represented by red line which is at 220ps (b) Force-extension profile: the profile of the force applied with respect to the center of mass distance between the two dimers during the lateral dissociation process.

A deeper insight into the network of interactions between the residues at the interface that contribute to microtubule stability can be obtained by analysis of the simulation trajectory in relation with the force-time and force-com profile obtained. In this study of dissociation of two heterodimers, the point of maximum force corresponds to the instant just before the breaking of few critical interactions between the two heterodimers. One of the important interaction among them is the hydrogen bond between Arg 86' of 1β subunit of first heterodimer and Gln 279' of 2β subunit of second heterodimer breaks. This hydrogen bond is between the sidechain Nitrogen of Arg 86' of 1β and sidechain Oxygen of Gln 279' of 2β. **Figure 37b** represents the dissociation of this interaction just after the point of maximum force. In another study about the lateral interactions between protofilaments, our group has predicted the residues that are crucial for lateral binding of tubulin subunits using CAS. In that study, CAS analysis of the interface residues of the same system has predicted Arg 86' as a hot spot ($\Delta\Delta G$ = 5.75 kcal/mol) and Gln 279' as a warm spot with $\Delta\Delta G$ = 2.46 kcal/mol, implying both of the

above mentioned amino acids are important for the lateral interaction among tubulin heterodimers. (**Figure 38**)

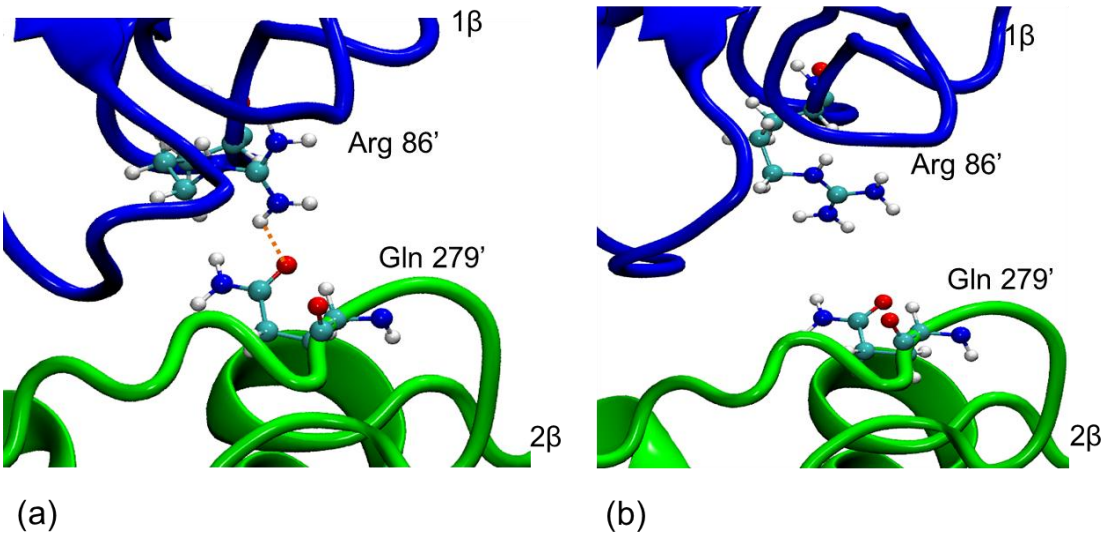

Figure 37: (a) Initial configuration showing Hydrogen bond between Arg 86' of 1β and Gln 279' of 2β. (b) Snapshot taken just after the point of maximum force showing which corresponds to the breaking of the Hydrogen bond between Arg 86' of 1β and Gln 279' of 2β.

On the basis of thorough analysis of the various interdimer lateral interactions, we have found other interactions that play a role in stabilization of the MT. Backbone oxygen of Thr 56 of 1α subunit interact with backbone nitrogen of Glu 284 of 2α subunit via a Hbond. According to CAS study Thr 56 was predicted to be a warm-spot with ΔΔG of 2.25 kcal/mol (**Figure 38**). Another Hbond is observed between sidechain nitrogen of Gln 85 on 1α and backbone oxygen of Glu 279 on 2α. Sidechain oxygen of Glu 125' on 1β interacts via Hbond with sidechain nitrogen of Lys 336' on 2β. The use of COM pulling allows monitoring of the evolution of these interactions over time as dimer B is pulled away from dimer A. This gives qualitative assessment of the specific interactions

between the two heterodimers which sare responsible for the stability and dynamics of the microtubule.

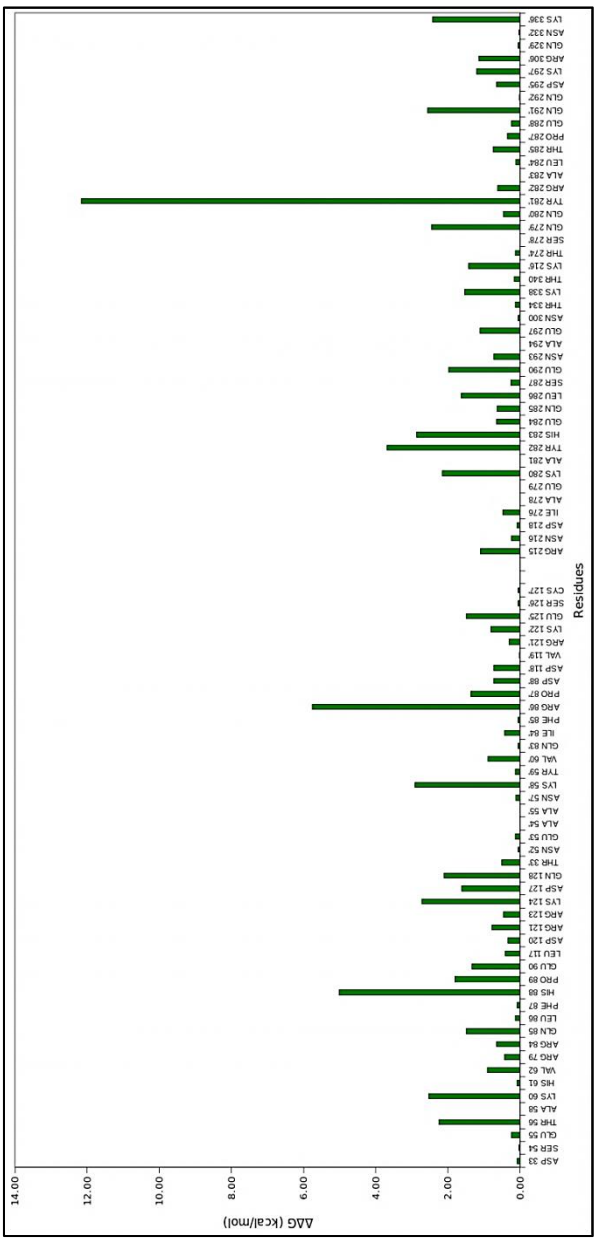

Figure: ΔΔG values obtained from CAS study performed by our research group, for the residues located at the tubulin-tubulin lateral interface. The residues with a prime symbol (‘) belongs to β-tubulin subunit

Conclusions & Future perspectives

This work underscores how Steered Molecular Dynamics studies can provide detailed information in terms of key interactions involved in protein-protein interface of tubulin heterodimers that are important for the stability of microtubules. Thus, we can propose that the compounds designed to interfere with the observed interactions between the two heterodimers, may act as viable therapeutic candidates to destabilize or inhibit the tubulin-tubulin interaction. *In silico* analysis of protein- protein interface can in perspective be useful in development of drug like compounds targeting protein-protein interactions, i.e. in targeted drug design.

Because of the limitation of the computational power, further analysis of the longitudinal and lateral interface could not be performed. Although, in future with enhanced computational power, it would be possible to calculate the free energy of binding of dimers by implementation of Umbrella sampling method.

Bibliography

1. **Sokolowski, B. and Orchard, S.** Exploiting Protein Interaction Networks to Unravel Complex Biological Questions. X. Xia. *Selected Works in Bioinformatics.* 2011.

2. *Small molecule stabilization of protein-protein interactions: an underestimated concept in drug discovery?* **Thiel, P., Kaiser, M. and Ottmann, C.** 2012, Angew. Chem. Int., Vol. 51, pp. 2012-2018.

3. **Jung, W., H.H., Jeong and Lee, K.Y,.** Protein Interactome and Its Application to Protein Function Prediction. Dr. W Cai. *Protein-Protein Interactions - Computational and Experimental Tools.* 2012.

4. *Modulating protein-protein interactions with small molecules: The importance of binding hot-spots.* **Thangudu, R. R., et al.** 2012, J. Mol. Biol., Vol. 415, pp. 443-453.

5. **Wendt, M.D.** *Protein-protein Interactions as drug targets, Protein-protein Interactions.* 2012.

6. *Protein-protein interactions in receptor activation and intracellular signaling.* **Blundell, T. L., et al.** 2005, Biol Chem., Vol. 381, pp. 955-959.

7. *The atomic structure of protein-protein recognition sites.* **Lo Conte, L., Chothia, C. and Janin, J.** 1999, J. Mol. Biol., Vol. 285, pp. 2177-2198.

8. *Studies of protein-protein interfaces: A statistical analysis of the hydrophobic effect.* **Tsai, C. J., et al.,** 1997, Protein Sci., Vol. 6, pp. 53-64.

9. *Structural motifs at protein-protein interfaces: protein cores versus two state and three state model complexes.* **Tsai, C. J., Xu, D. and Nussinov, R.** 1997, Protein Sci., Vol. 6, pp. 1793-1805.

10. *Statistical analysis and prediction of protein-protein interfaces.* **Bordner, A. J. and Abagyan, R.** 2005, Proteins, Vol. 60, pp. 353-366.

11. *ASEdb: A database of alanine mutations and their effects on the free energy of binding in protein interactions.* **Thorn, K. S. and Bogan, A. A.** 2001, Bioinformatics, Vol. 17, pp. 284-285.

12. *Hot spots- A review of the protein- protein interface determinant amino acid residues.* **Moreira, I. S., Fernandes, P. A. and Ramos, M. J.** 2007, Proteins, Vol. 68, pp. 803-812.

13. *Drugs targeting protein-protein interactions.* **Chene, P.** 2006, Chem. Med. Chem., Vol. 1, pp. 400-411.

14. *Hot regions in protein-protein interactions: The organization and contribution of structurally conserved hot spot residues.* **Keskin, O., Ma, B. and Nussinov, R.** 2005, J. Mol. Biol., Vol. 345, pp. 1281-1294.

15. *Anatomy of hot-spots in protein interfaces.* **Bogan, A. A. and Thorn, K. S.** 1998, J. Mol. Biol., Vol. 280, pp. 1-9.

16. *Protein-protein interactions in human diseases.* **Ryan, D. P. and Matthews, J. M.** 2005, Curr. Opin. Struct. Biol., Vol. 15, pp. 441-446.

17. *Targeting protein-protein interactions for therapeutic intervention: a challenge for the future.* **Zinzalla, G. and Thurston, D. E.** 2009, Future Med. Chem., Vol. 1, pp. 65-93.

18. *Druggable protein-protein interactions: from hot spots to hot segments.* **London, N., Raveh, B. and Schueler-Furman, O.** 2013, Curr. Opin. In Chem. Biol., Vol. 17, pp. 952-959.

19. *Can self-inhibitory peptides be derived from the interfaces of globular protein-protein interactions.* **London, N., et al.** 2010, Proteins, Vol. 78, pp. 3140-3149.

20. *The Druggable genome.* **Hopkins, A. L. and Groom, C. R.** 2002, Nat. Rev. Drug Discov., Vol. 1, pp. 727-730.

21. *Dr. PIAS 2.0: An update of a database of predicted druggable protein-protein interactions.* **Sugaya, N., Kanai, S. and Furuya, T.** 2012, Database (Oxford) 2012, pp. 1-3.

22. *Small molecule inhibitors of protein-protein interactions: Progressing towards the dream.* **Arkin, M. R. and Wells, J. A.** 2004, Nat. Rev. Drug Discov., Vol. 3, pp. 301-317.

23. *Small-Molecule Inhibitors of the MDM2-p53 Protein-Protein Interaction to Reactivate p53 Function: A Novel Approach for Cancer Therapy.* **Shangary, S. and Wang, S.** 2009, Annu. Rev. Pharmacol. Toxicol., Vol. 49, pp. 223-241.

24. *Bcl-2 inhibitors: small molecules with a big impact on cancer therapy.* **Vogler, M., et al.** 2009, Cell Death Differ., Vol. 16, pp. 360-367.

25. *The road less traveled: modulating signal transduction enzymes by inhibiting their protein–protein interactions.* **Arkin, M. R. and Whitty, A.** 2009, Curr. Opin. Chem. Biol., Vol. 13, pp. 284-290.

26. *Protein–protein interaction inhibitors get into the groove.* **Mullard, A.** 2012, Nat. Rev. Drug. Discov., Vol. 11, pp. 173-175.

27. *Strategies for targeting protein-protein interactions with synthetic agents.* **Hang, Y. and Hamilton, A. D.** 2005, Angew. Chem., Vol. 44, pp. 4130–4163.

28. *Novel p-arylthio cinnamides as antagonists of leukocyte function-associated antigen-1/intracellular adhesion molecule-1 interaction. 2. Mechanism of inhibition and structure-based improvement of pharmaceutical properties.* **Liu, G., et al.** 2001, J. med. Chem., Vol. 44, pp. 1202-1210.

29. *Mechanism of action of antitumor drugs that interact with microtubules and tubulin.* **Jordan, M. A.** 2002, Curr. Med. Chem. Anti-Cancer agents, Vol. 2, pp. 1-17.

30. *In-silico design of tubulin-targeted antimitotic peptides.* **Pieraccini, S., Saladino G., et al.** 2009, Nat. Chem., Vol. 1, pp. 642-648.

31. *Modulation of protein-protein interactions with small molecules.* **Berg, T.** 2003, Angew. Chem. Int. Ed. Engl., Vol. 42, pp. 2462-2481.

32. *Protein-protein interaction inhibitors: small molecules from screening techniques.* **Fletcher, S. and Hamiltonian, A. D.** 2007, Curr. Top. Med. Chem., Vol. 7, pp. 922-927.

33. *Discovery of inhibitors of protein-protein interactions from combinatorial libraries.* **Vincent, M. J., Perez-Paya, E. and M., Orzaez.** 2007, Curr. Top. Med. Chem., Vol. 7, pp. 83-95.

34. *Inhibiting protein-protein interactions as an emerging paradigm for drug discovery.* **Gerrard, J. A., Hutton, C. A. and Perungini, M. A.** 2007, Mini. Rev. Med. Chem., Vol. 7, pp. 151-157.

35. *Reaching for high hanging fruit in drug discovery at protein-protein interfaces.* **Wells, J. A. and McClendon, C. L.** 2007, Nature, Vol. 450, pp. 1001-1009.

36. *Strategies to search and design stabilizers of protein-protein interactions: a feasibility study.* **Block, P., Weskamp, A., Wolf, A., Klebe, G.** 2007, Proteins, Vol. 68, pp. 170-186.

37. *Computational identification of inhibitors of protein-protein interactions.* **Zhong, S., Macias, A. T. and Mackerell, A. D. J.** 2007, Curr. Top. Med. Chem., Vol. 7, pp. 63-82.

38. *Targetting protein-protein interactions with small molecules: challenges and perspectives for computational binding epitope detection and ligand finding.* **Gonzalez-Ruiz, D. and Gohlke, H.** 2006, Curr. Med. Chem., Vol. 13, pp. 2607-2625.

39. *Automated clustering of probe molecules from solvent mapping of protein surfaces: new algorithms applied to hot spot mapping and structure-based drug design.* **Lerner, M., Meagher, K. and Carlson, H.** 2008, J. Comp.-Aided Mol. Des., Vol. 22, pp. 727-736.

40. *Protein-protein interactions in drug discovery.* **Fischer, P. M.** 2005, Drug Des. Rev., Vol. 2, pp. 179-207.

41. *Targetting protein-protein interactions: lessons from p53/Mdm2.* **Murray, J. K. and Gellman, S. H.** 2007, Peptide Sci., Vol. 88, pp. 657-686.

42. *Computationally designed peptide inhibitors of protein-protein interactions in membranes.* **Caputo, G. A., et al.** 2008, Biochem., Vol. 47, pp. 8600-8606.

43. *Structure of tubulin at 6.5 Å and location of the taxol-binding site.* **Nogales, E., et al.** 1995, Nature, Vol. 375, pp. 424-427.

44. *Refined structure of αβ-tubulin at 3.5 Å resolution.* **Lowe, J., et al.** 2001, J. Mol. Biol., Vol. 313, pp. 1045-1057.

45. *VMD - Visual Molecular Dynamics.* **Humphrey, W., Dalke, A. and Schulten, K.** 1996, J. Molec. Graphics, Vol. 14, pp. 33-38.

46. *Structure of the FKBP12-Rapamycin Complex Interacting with Binding Domain of Human FRAP.* **Choi, J., et al.** 1996, Science, Vol. 273, pp. 239-242.

47. **Roland, S.** Molecular Dynamics Simulations Charmm. [Online] Embnet.ch tutorial, October 26, 1999. [Cited: August 10, 2013.] http://www.ch.embnet.org/MD_tutorial/.

48. *Computer simulation of molecular dynamics – methodology, applications, and perspectives in chemistry.* **Van Gunsteren, W.F., Berendsen, H.J.C.** 1990, Angew. Chem. Int. Ed. Engl., Vol. 29, pp. 992-1023.

49. *Molecular dynamics simulations in biology.* **Karplus, M. and Petsko, G. A.** 1990, Nature, Vol. 347, pp. 631-639.

50. **Hypercube, Incorporated.** Molecular Dynamics. *HyperChem Computational Chemistry.* 1994.

51. **McMillan, S.** "Leap frog Integrator", Computational Physics. [Online] http://einstein.drexel.edu/courses/Comp_Phys/Integrators/leapfrog/.

52. *Force fields for protein simulations.* **Ponders, J. W. and Case, D. A.** 2003, Adv. Protein. Chem., Vol. 66, pp. 27-85.

53. *Development and testing of a general Amber force field.* **Wang, J., et al.** 2004, J. Comput. Chem., Vol. 25, pp. 1157-1174.

54. *CHARMM: A program for macromolecular energy, minimization, and dynamics calculations.* **Brooks, B. R., et al.** 983, J Comp Chem., Vol. 4, pp. 187–217.

55. *A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6.* **Oostenbrink, C., et al.** 2004, J. Comput. Chem., Vol. 25, pp. 1656-1676.

56. *Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids.* **Jorgensen, W. L., Maxwell, D. S. and Tirado-Rives, J.** 1996, J. Am. Chem. Soc., Vol. 118, pp. 11225–11236.

57. *The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known three-dimensional structure.* **Bohm, H. J.** 3, 1994, J. Comput.-Aided Mol. Des., Vol. 8, pp. 243-256.

58. *Scoring noncovalent protein-ligand interactions: A continuous differentiable function tuned to compute binding affinities.* **Jain, A. N.** 5, 1996, J. Comput.-Aided Mol. Des., Vol. 10, pp. 427-440.

59. *Free energy calculations: Applications to chemical and biochemical phenomena.* **Kollman, P. A.** 7, 1993, Chem. Rev., Vol. 93, pp. 2395–2417.

60. *Free energy calculations for theophylline binding to an RNA aptamer: Comparison of MM-PBSA and thermodynamic integration methods.* **Gouda, H., et al.** 2003, Biopolymers, Vol. 68, pp. 16–34.

61. *A new method for predicting binding affinity in computer-aided drug design.* **Åqvist, J., Medina, C. and Samuelsson, J. E.** 3, 1994, Protein Eng., Vol. 7, pp. 385-391.

62. *The Poisson–Boltzmann equation for biomolecular electrostatics: a tool for structural biology.* **Fogolari, F., Brigo, A. and Molinari, H. J.** 2002, J. Mol. Recognit., Vol. 15, pp. 377-392.

63. *Comparison between computational alanine scanning and per-residue binding free energy decomposition for protein–protein association using MM-GBSA: Application to the TCR-p-MHC complex.* **Zoete, V. and Michielin, O.** 2007, Proteins, Vol. 67, pp. 1026-1047.

64. *Application of a pairwise generalized born model to proteins and nucleic acids: inclusion of salt effects.* **Srinivasan, J., et al.** 1999, Theor. Chem. Acc., Vol. 101, pp. 426-434.

65. *Binding of a Diverse Set of Ligands to Avidin and Streptavidin: An Accurate Quantitative Prediction of Their Relative Affinities by a Combination of Molecular Mechanics and Continuum Solvent Models.* **Kuhn, B. and Kollman, P. A.** 2000, J. Med. Chem., Vol. 43, pp. 3786-3791.

66. *Compensating enthalpic and entropic changes hinder binding affinity optimization.* **Lafont, V., et al.** 6, 2007, Chem. Biol. & Drug Des., Vol. 69, pp. 413-422.

67. *Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA.* **Rastelli, G., et al.** 4, 2010, J. Comput. Chem., Vol. 31, pp. 797-810.

68. *Semianalytical treatment of solvation for molecular mechanics and dynamics.* **Still, W. C., et al.** 1990, J. Am. Chem. Soc., Vol. 112, pp. 6127-6129.

69. *Comparison of MM/GBSA calculations based on explicit and implicit solvent simulations.* **Godschalk, F., et al.** 2013, Phys. Chem. Chem. Phys., Vol. 15, pp. 7731-7739.

70. *Exploring protein native states and large-scale conformational changes with a modified generalized Born model.* **Onufriev, A., Bashford, D. and Case, D.A.** 2004, Proteins, Vol. 55, pp. 383–394.

71. *The interpretation of protein structures: Estimation of static accessibility.* **Lee, K. and Richards, F. M.** 1971, J. Mol. Biol., Vol. 55, pp. 379-400.

72. *A hot spot of binding energy in a hormone-receptor interface.* **Clackson, T. and Wells, J. A.** 1995, Science, Vol. 267, pp. 383–386.

73. *Protein-protein interaction hotspots carved into sequences.* **Ofran, Y. and Rost, B.** 7, 2007, PLoS. Comput. Biol., Vol. 3, p. e119.

74. *Computational Alanine Scanning to Probe Protein-Protein Interactions: A Novel Approach to Evaluate Binding Free Energies.* **Massova, I. and Kollman, P. A.** 36, 1999, J. Am. Chem. Soc., Vol. 121, pp. 8133–8143.

75. *Continuum Solvent Studies of the Stability of DNA, RNA, and Phosphoramidate−DNA Helices.* **Srinivasan, J., et al.** 1998, J. Am. Chem. Soc., Vol. 120, pp. 9401-9409.

76. *Energetic analysis of an antigen/antibody interface: Alanine scanning mutagenesis and double mutant cycles on the HyHEL-10/lysozyme interaction.* **Pons, J., Rajpal, A. and Kirsch, J. F.** 1999, Protein Sci., Vol. 8, pp. 958–968.

77. *Computational Alanine Scanning of Protein-Protein Interfaces.* **Kortemme, T., Kim, D. E. and Baker, D.** 2004, Sci. STKE, pp. 1-8.

78. *Structural and functional analysis of the 1:1 growth hormone:receptor complex reveals the molecular basis for receptor affinity.* **Clackson, T., et al.** 1998, J. mol. Biol., Vol. 277, pp. 1111-1128.

79. *Rapamycin (AY-22,989), a new antifungal antibiotic. II. Fermentation, isolation and characterization.* **Sehgal, S. N., Baker, H. and Vezina, C.** 10, 1975, J. Antibiot. (Tokyo), Vol. 28, pp. 727-732.

80. *Rapamycin: A new antifungal antibiotic I. taxonomy of the producing streptomycete and isolation of active principle.* **Vezina, C., Kudelski, A. and Sehgal, S. N.** 1975, J. Antibiot., Vol. 28, pp. 721-726.

81. *Rapamycin a new antifungal antibiotic III. In vitro and in vivo evaluation.* **Baker, H., et al.** 1978, J Antibiot., Vol. 31, pp. 539-545.

82. *Sirolimus: its discovery, biological properties, and mechanism of action.* **Sehgal, S. N.** 2003, Transplant. Proc., Vol. 35, pp. S7-S14.

83. *Sirolimus: a ten-year perspective.* **Kahan, B. D.** 2004, Transplant. Proc., Vol. 36, pp. 71-75.

84. *Rapamycin. A new antifungal antibiotic IV. Mechanism of action.* **Singh, K., Sun, S. and Vezina, C.** 1979, J. Antibiot. (Tokyo), Vol. 32, pp. 630-645.

85. *Characterization of the FKBP•Rapamycin•FRB Ternary Complex.* **Banaszynski, L. A., Liu, C. W. and Wandless, T. J.** 2005, J. Am. Chem. Soc., Vol. 127, pp. 4715-4721.

86. *Activity of rapamycin against transplanted tumors.* **Eng, C. P., Sehgal, S. N. and Vezina, C.** 1984, J. Antibiot. (Tokyo), Vol. 37, pp. 1231-1237.

87. *Rapamycin inhibits constitutive P70s6k phosphorylation, cell proliferation and colony formation in small lung cancer cells.* **Seufferlein, T. and Rozengurt, E.** 1996, Cancer Res., Vol. 56, pp. 3895-3897.

88. *Rapamycin, a potent immunosuppressive drug causes programmed cell death in B lymphoma cells.* **Muthukkumar, S., Ramesh, T. M. and Bondada, S.** 3, 1995, Transplantation, Vol. 60, pp. 264-70.

89. *Rapamune® (RAPA, rapamycin, sirolimus): mechanism of action immunosuppressive effect results from blockade of signal transduction and inhibition of cell cycle progression.* **Sehgal, S. N.** 1998, Clin. Biochem., Vol. 31, pp. 335-340.

90. *A cytosolic binding protein for the immunosuppressant FK506 has peptidyl-prolyl isomerase activity but is distinct from cyclophilin.* **Siekierka, J. J., et al.** 1989, Nature, Vol. 341, pp. 755-757.

91. *Targets of cell cycle arrest by the immunosuppressant rapamycin in yeast.* **Heitman, J., Movva, N. R. and Hall, M. N.** 1991, Science, Vol. 253, pp. 905-909.

92. *Rapamycin sensitivity in Saccharomyces cerevisiae is mediated by a peptidyl propyl cis trans isomerase related to human FK506-binding protein.* **Koltin, Y., et al.** 1991, Mol. Cell. Biol, Vol. 11, pp. 1718-1723.

93. *Identification of an 11-kDa FKBP12-rapamycin-binding domain within the 289-kDa FKBP12-rapamycin-associated protein and characterization of a critical serine residue.* **Chen, J., et al.** 1995, Proc. Natl. Acad. Sci. USA, Vol. 92, pp. 4947-4951.

94. *Raft1 : A mammalian protein that binds to fkbp12 in a rapamycin dependent fashion and is homologous to yeast TORs.* **Sabatini, D. M., et al.** 1994, Cell, Vol. 78, pp. 35-43.

95. *A mammalian protein targeted by G1-arresting-rapamycin receptor complex.* **Brown, E. J., et al.** 1994, Nature, Vol. 369, pp. 756-758.

96. *Rapt1, a mammalian homolog of yeast Tor, interacts with FKBP12/rapamycin complex.* **Chui, M. I., Katz, H. and Berlin, V.** 1994, Proc. Natl. Acad. Sci. USA, Vol. 91, pp. 12574-12578.

97. *The molecular target of rapamycin (mTOR) as a therapeutic target against cancer.* **Mita, M.M., Mita, A. and Rowinsky, E.K.** 2003, Cancer Biology & Therapy, Vol. 2, pp. 169-177.

98. *Evidence of insulin-stimulated phosphorylation and activation of mTOR mediated by a protein kinase B signaling pathway.* **Scott, P. H., et al.** 1998, Proc. Natl. Acad. Sci. USA, Vol. 95, pp. 7772-7777.

99. *Mammalian target of rapamycin is a direct target for protein kinase B: identification of a convergence point for opposing effects of insulin and amino-acid deficiency on protein translation.* **Nave, B., et al.** 1999, Biochem. J., Vol. 344, pp. 427-431.

100. *Mechanisms and consequences of activation of protein kinase B/Akt.* **Downward, J.** 1998, Curr. Opin. Cell. Biol., Vol. 10, pp. 262-267.

101. *A direct linkage between the phosphoinositide 3-kinase-Akt signaling pathway and the mammalian target of rapamycin in mitogen-stimulated and transformed cells.* **Sekulic, A., et al.** 2000, Cancer Res., Vol. 60, pp. 3504-3514.

102. *Phosphorylation of the translational repressor PHAS-1 by the mammalian Target of Rapamycin.* **Brunn, G. J., et al.** 1997, Science, Vol. 277, pp. 99-101.

103. *4E-BPI, a repressor of mRNA translation, is phosphorylated and inactivated by the Akt(PKB) signaling pathway.* **Gingras, A. C., et al.** 1998, Genes Dev., Vol. 12, pp. 502-513.

104. *Protein Kinase B(c-Akt) in phosphatidylinositol-3-OH kinase signal transduction.* **Burgering, B. M. T. and Coffer, P. J.** 1995, Nature, Vol. 376, pp. 599-602.

105. *PDGF and insulin dependent pp70s6k activation mediated by phosphatidylinositol-3-OH kinase.* **Chung, J., et al.** 1994, Nature, Vol. 370, pp. 71-75.

106. *eIF4 initiation factors: Effectors of mRNA recruitment to ribosomes and regulators of translation.* **Gingras, A. C., Raught, B. and Sonenberg, N.** 1999, Annu. Rev. Biochem., Vol. 68, pp. 913-963.

107. *Control of p70s6k kinase by kinase activity of FRAP in vivo.* **Brown, E. J., et al.** 1995, Nature, Vol. 377, pp. 441-446.

108. *Raft 1 phosphorylation of the translational regulators p70s6k kinase and 4E-BPI.* **Burnett, P. E., et al.** 1998, Proc. Natl. Acad. Sci. USA, Vol. 95, pp. 1432-1437.

109. *Refined structure of the FKBP12-rapamycin-FRB ternary complex at 2.2 Å resolution.* **Liang, J., Choi, J. and Clardy, J.** 1999, Acta. Crystallogr. D, Vol. D55, pp. 736-744.

110. *Two distinct signal transmission pathways in T lymphocytes are inhibited by complexes formed between an immunophilin and either FK506 or rapamycin.* **Bierer, B. E., Mattila, P. S., Standaert, R.F., et al.** 23, 1990, Proc. Natl. Sci. USA, Vol. 87, pp. 9231-9235.

111. *AMBER 11.* **Case, D. A., Darden, T. A., Cheatham, T. E., et al.** 2010, University of California, San Francisco.

112. *Improved side-chain torsion potentials for the Amber ff99SB protein force field.* **Lindorff-Larsen, K., et al.** 2010, Proteins, Vol. 78, pp. 1950-1958.

113. *Comparison of simple potential functions for simulating liquid water.* **Jorgensen, W.L., et al.** 1983, J. Chem. Phys., Vol. 79, pp. 926-935.

114. *Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes.* **Ryckaert, J. P., Ciccotti, G. and Berendsen, H.J.C.** 1977, J. Comput. Phys., Vol. 23, pp. 327-341.

115. *Langevin dynamics of peptides: The frictional dependence of isomerization rates of N-acetylalanyl-N'-methylamide.* **Loncharich, R. J., Brooks, B. R., Pastor R.W.** 1992, Biopolymers, Vol. 32, pp. 523-535.

116. *Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems.* **Darden, T., York, D. and L., Pedersen.** 1993, J. Chem. Phys., Vol. 98, pp. 10089-10095.

117. *Unraveling the importance of protein-protein interaction: Application of a computational alanine-scanning mutagenesis to the study of the IgG1 streptococcal protein G (C2 Fragment) complex.* **Moreira, I.S., Fernandes, P.A. and Ramos, M.J.** 22, 2006, J. Phys. Chem B, Vol. 110, pp. 10962-10969.

118. *Computational alanine scanning mutagenesis - An improved methodological approach.* **Moreira, I.S., Fernandes, P.A. and Ramos, M.J.** 3, 2007, J. Comput. Chem., Vol. 28, pp. 644-654.

119. *Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding.* **Massova, I. and Kollman, P.A.** 2000, Perspect. Drug Discovery Des., Vol. 18, pp. 113-135.

120. *Naccess.* **Hubbard, S. J. and Thornton, J. M.** 1, 1993, Computer Program, Department of Biochemistry and Molecular Biology, University College London, Vol. 2.

121. *Computational alanine scanning of the 1:1 human growth hormone-receptor complex.* **Huo, S., Massova, I. and Kollman, P.A.** 1, 2002, J. Comp. Chem., Vol. 23, pp. 15-27.

122. *Atomic structure of the rapamycin human immunophillin FKBP-12 complex.* **Van Duyne, G.D., et al.** 1991, J. Am. Chem. Soc., Vol. 113, pp. 7433-7434.

123. *FK506 binding protein mutational analysis. Defining the surface residue contributions to stability of the calcineurin co-complex.* **Futer, O., et al.** 1995, J. Biol. Chem., Vol. 270, pp. 18935-18940.

124. *Dynamics of proteins and nucleic acids.* **McCammon, J. A. and Harvey, S. C.** 1, 1987, Cambridge University Press, Vol. 13, pp. 35-36.

125. *The statistical-thermodynamic basis for computation of binding affinities: a critical review.* **Gilson, M.K., et al.** 3, 1997, Vol. 72, pp. 1047-1069.

126. *Computational Methods to Predict Binding Free Energy in Ligand-Receptor Complexes.* **Ajay and Murcko, M. A.** 26, 1995, J. Med. Chem., Vol. 38, pp. 4953–4967.

127. *Steered Molecular Dynamics. Computational Molecular Dynamics: Challenges, Methods, Ideas.* **Izrailev, S., et al.** 1999, Lecture Notes in Computational Science and Engineering, Vol. 4, pp. 39-65.

128. *Atomic Force Microscope.* **Binnig, G., Quate, C. F. and Gerber, Ch.** 1996, Phys. Rev. Lett., Vol. 56, pp. 930–933.

129. *Optical trapping and manipulation of viruses and bacteria.* **Ashkin, A. and Dziedzic, J.M.** 1987, Science, Vol. 235, pp. 1517–1520.

130. *Sensitive force technique to probe molecular adhesion and structural linkages at biological interfaces.* **Evan, E., Ritchie, K. and Merkel, R.** 1995, Biophys J, Vol. 68, pp. 2580–2587.

131. *Biological Applications of Optical Forces.* **Svoboda, K. and Block, S.M.** 1994, Annual Review of Biophysics and Biomolecular Structure, Vol. 23, pp. 247-285.

132. AFMSTM. [Online] http://www.vub.ac.be/META/toestellen_AFMSTM.php?m=xpand.

133. *Assessing the stability of Alzheimer's Amyloid Protofibrils Using Molecular Dynamics.* **Lemkul, J. and Bevan, D.R.** 2010, J Phys Chem. B., Vol. 114, pp. 1652-1660.

134. *Ligand binding: Molecular mechanics calculation of the streptavidin-biotin rupture force.* **Grubmuller, V., Heymann, B. and Tavan, P.** 1996, Science, Vol. 271, pp. 997-999.

135. *Binding pathway of retinal to bacterio-opsin: a prediction by molecular dynamics simulations.* **Isralewitz, B., Izrailev, S. and Schulten, K.** 1997, Biophysical journal, Vol. 73, pp. 2972-2979.

136. *Extraction of Lipids from Phospholipid Membranes by Steered Molecular Dynamics.* **Stepaniants, S., Izrailev, S. and Schulten, K.** 1997, J. Mol. Model., Vol. 3, pp. 473-475.

137. *Molecular dynamics study of unbinding of the avidin-biotin complex.* **Izrailev, S., et al.** 4, 1997, Biophysical Journal, Vol. 72, pp. 1568–1581.

138. *Unfolding of Titin Immunoglobulin Domains by Steered Molecular Dynamics Simulation.* **Lu, H., et al.** 2, 1998, Biophysical Journal, Vol. 75, pp. 662–671.

139. *Pulling single bacteriorhodopsin out of a membrane: Comparison of simulation and experiment. Biochimica et Biophysica Acta (BBA).* **Cieplak, M., et al.** 2006, Biomembranes, Vol. 1758, pp. 537-544.

140. *Simulated unbinding and binding of fatty acid substrates in the cyclooxygenase site of prostaglandin H2 synthase-1.* **Molnar, F., Norris, L. S. and Schulten, K.** 2000, Prog. Reaction Kinetics and Mechanisms, Vol. 25, pp. 263-298.

141. *Steered Molecular Dynamics.* **Izrailev, S., et al.** 1998, Springer-Verlag: Berlin, Vol. 4.

142. *Thermodynamic analysis of biotin binding to avidin. A high sensitivity titration calorimetric study.* **Swamy, M. J.** 1, 1995, Biochemistry & Molecular Biology International, Vol. 36, pp. 219-225.

143. *Thermochemistry of the Avidin-Biotin Reaction.* **Suurkuusk, J. and WadsÖ, I.** 1972, European journal of biochem, Vol. 28, pp. 438-441.

144. *Spectrophotometric determination of avidin and biotin.* **Green, N. M.** 1970, Methods in Enzymology, Vol. 18A, pp. 418-424.

145. *Adhesion forces between individual ligand-receptor pairs.* **Florin, E.L., Moy, V.T. and Gaub, H.E.** 1994, Science, Vol. 264, pp. 415-417.

146. *Smd studies of titin I1 domain unfolding.* **Gao, M., Wilmanns, M. and Schulten, K.** 2002, Biophysical journal, Vol. 83, pp. 3435-3445.

147. *Reversible unfolding of individual titin immunoglobulin domains by AFM.* **Rief, M., et al.** 1990, Science, Vol. 276, pp. 1109-1112.

148. *Unfolding of titin immunoglobulin by Steered Molecular Dynamics simulation.* **Lu, H., et al.** 1998, Biophysical journal, Vol. 75, pp. 662-671.

149. *Steered Molecular Dynamics simulation of changes of immunoglobulin domain I27 interpreted atomic force microscopy observations.* **Lu, H. and Schulten, K.** 1999, Chemical physics, Vol. 247, pp. 141-145.

150. *Binding pathway of retinal to bacterio-opsin: A prediction by Molecular Dynamics simulations.* **Isralewitz, B., Izrailev, S. and Schulten, K.** 1997, Biophysical journal., Vol. 73, pp. 2972-2979.

151. *Umbrella sampling.* **Kästner, J.** 2011, WIREs Comput Mol Sc, Vol. 1, pp. 932–942.

152. *Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling.* **Torrie, G.M. and Valleau, J.P.** 2, 1977, Journal of Computational Physics, Vol. 23, pp. 187–199.

153. *Optimized Monte Cralo data Analysis.* **Ferrenberg, A.M. and Swendsen, R.H.** 1989, Phys. Rev. Lett., Vol. 63, pp. 1195-1198.

154. *The weighted histogram analysis method for free energy calculations on biomolecules I. The method.* **Kumar, S., et al.** 1992, J. Computational Chem., Vol. 13, pp. 1011-1021.

155. *Isolation of a Deep-Sea Barophilic Bacterium and Some of Its Growth Characteristics.* **Yayanos, A. A., et al.** 4408, 1979, Science, Vol. 205, pp. 808-810.

156. *Cell proliferation at 122 degrees C and isotopically heavy CH4 production by a hyperthermophilic methanogen under high-pressure cultivation.* **Takai, K., et al.** 31, 2008, Proc Natl Acad Sci U S A., Vol. 105, pp. 10949-54.

157. *High level transcription of a member of a repeated gene family confers dehydration tolerance to callus tissue of Craterostigma plantagineum.* **Furini, A., et al.** 1997, EMBO J., Vol. 16, pp. 3599-3608].

158. *Characterization of five novel dehydration-responsive homeodomain leucine zipper genes from the resurrection plant Craterostigma plantagineum.* **Deng, X., et al.** 2002, Plant Mol. Biol., Vol. 49, pp. 601-610.

159. *Water and Life – Comparative Analysis of Water Relationships at the Organismic, Cellular, and Molecular Levels.* **Somero, G.N., Osmond, C.B. and Bolis, C.L.** 1992, Springer.

160. **Beck, D.A.C., et al.** Simulations of Macromolecules in Protective and Denaturing Osmolytes: Properties of Mixed Solvent Systems and Their Effects on Water and Protein Structure and Dynamics. Sies H. Häussinger D. *Methods in Enzymology.* 2007, pp. 373-396.

161. *Living with water stress: Evolution of osmolyte systems.* **Yancey, P.H., et al.** 1982, Science (New York, N.Y.), Vol. 217, pp. 1214-1222.

162. *Backbone additivity in the transfer model of protein solvation.* **Hu, C.Y., et al., et al.** 2010, Protein Sci., Vol. 19, pp. 1011-1022.

163. *Interaction of urea with amino acids: Implications for urea-induced protein denaturation.* **Stumpe, M.C. and Grubmüller, H.** 2007, J. Am. Chem. Soc., Vol. 129, pp. 16126-16131.

164. *Urea denaturation by stronger dispersion interactions with proteins than water implies a 2-stage unfolding.* **Hua, L., et al.** 2008, Proc. Natl. Acad. Sci., Vol. 105, pp. 16928-16933.

165. *Anatomy of energetic changes accompanying urea-induced protein denaturation.* **Auton, M., Holthauzen, L.M.F. and Bolen, D.W.** 2007, Proc. Nat. Acad. Sci., Vol. 104, pp. 15317-15322.

166. *Effects of urea, tetramethyl urea, and trimethylamine n-oxide on aqueous solution structure and solvation of protein backbones: A molecular dynamics simulation study.* **Wei, H., Fan, Y. and Gao, Y.Q.** 2010, J. Phys. Chem. B, Vol. 114, pp. 557-568.

167. *Counteraction of urea-induced protein denaturation by trimethylamine N-oxide: A chemical chaperone at atomic resolution.* **Bennion, B.J. and Daggett, V.** 2004, Proc. Natl. Acad. Sci. USA, Vol. 101, pp. 6433-6438.

168. *The molecular mechanism of stabilization of proteins by TMAO and its ability to counteract the effects of urea.* **Zou Q., Bennion B.J., Daggett V., Murphy K.P.** 2002, J. Am. Chem. Soc., Vol. 124, pp. 1192-1202.

169. *Hydrophobic interactions in urea-trimethylamine-N-oxide solutions.* **Paul, S. and Patey, G.N.** 2008, J. Phys. Chem. B, Vol. 112, pp. 11106-11111.

170. *A simple mechanism underlying the effect of protecting osmolytes on protein folding.* **Saladino, G., et al.** 2011, J. Chem. Theory Comput., Vol. 7, pp. 3846-3852.

171. *Atomic level description of the protecting effect of osmolytes against thermal denaturation of proteins.* **Pieraccini, S., et al.** 2007, Chem. Phys. Lett., Vol. 438, pp. 298-303.

172. *A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G.* **Gronenborn, A.M., et al.** 1991, Science, Vol. 253, pp. 657-661.

173. *Folding dynamics and mechanism of β-hairpin formation.* **Muñoz, V., et al.** 1997, Nature, Vol. 390, pp. 196-199.

174. *Understanding the key factors that control the rate of β-hairpin folding.* **Du, D., et al.** 2004, Proc. Natl. Acad. Sci. USA, Vol. 101, pp. 15915-15920.

175. *The mechanism of β-hairpin formation.* **Dyer, R.B., et al.** 2004, Biochemistry, Vol. 43, pp. 11560-11566.

176. *Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling.* **Torrie, G.M. and Valleau, J.P.** 1977, J. Comput. Phys., Vol. 23, pp. 187-199.

177. *The calculation of the potential of mean force using computer simulations.* **Roux, B.** 1995, Comput. Phys. Commun., Vol. 91, pp. 275-282.

178. *Classical and quantum mechanical/molecular mechanical molecular dynamics simulations of alanine dipeptide in water: Comparisons with IR and vibrational circular dichroism spectra.* **Kwac, K., et al.** 2008, J. Chem. Phys., Vol. 128.

179. *A smooth particle mesh Ewald method.* **Essmann, U., et al.** 1995, J. Chem. Phys., Vol. 103, pp. 8577-8593.

180. *A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations.* **Duan, Y., et al.** 2003, J. Comput. Chem., Vol. 24, pp. 1999-2012.

181. *GROMACS: A message-passing parallel molecular dynamics implementation.* **Berendsen, H.J.C., Van der Spoe, l D. and Van Drunen, R.** 1995, Comput. Phys. Commun., Vol. 91, pp. 43-56.

182. *GROMACS: Fast, flexible, and free.* **Van Der Spoel, D., et al.** 2005, J. Comput. Chem., Vol. 26, pp. 1701-1718.

183. *The Nose-Hoover thermostat.* **Evans, D.J. and Holian, B.L.** 1985, J. Chem. Phys., Vol. 83, pp. 4069-4074.

184. *Polymorphic transitions in single crystals: A new molecular dynamics method.* **Parrinello, M. and Rahman, A.** 1981, J. Appl. Phys., Vol. 52, pp. 7182-7190.

185. *G-whams-a free Weighted Histogram Analysis implementation including robust error and autocorrelation estimates.* **Hub, J.S., De Groot, B.L. and Van der Spoel, D.** 2010, J. Chem. Theory Comput., Vol. 6, pp. 3713-3720.

186. **Zhu, F. and Hummer, G.** 2012, J. Comput. Chem., Vol. 33, pp. 453-465.

187. *BOINC: A System for Public-Resource Computing and Storage.* **D.P., Anderson.** 2004, Grid Computing, IEEE/ACM International Workshop.

188. *Arrangement of subunits in flagellar microtubules.* **Amos, L. and Klug, A.** 1974, Journal of Cell Science, Vol. 14, pp. 523-549.

189. *Structural insights into Microtubule function.* **Nogales, E.,.** 2000, Annu. Rev. Biochem., Vol. 69, pp. 277-302.

190. *High resolution model of the microtubule.* **Nogales, E., et al.** 1999, Cell, Vol. 96, pp. 34-79.

191. *Intrinsic Bending and Structural Rearrangement of Tubulin Dimer: Molecular Dynamics Simulations and Coarse-Grained Analysis.* **Gebremichael, Y., Chu, J.W. and Voth, G.A.** 2008, Biophys. J., Vol. 85, pp. 2487-2499.

192. *Microtubules as a target for anticancer drugs.* **Jorda, n M.A. and Wilson, L.** Nature Reviews Cancer, Vol. 4, pp. 253-265.

193. *Dynamic Instability of Microtubule growth.* **Mitchison, T. and Kirschner, M.** 1984, Nature, Vol. 312, pp. 237-242.

194. *Microtubule polymerization dynamics.* **Desai, A. and Mitchison, T. J.** 1997, Annual Review of Cell and developmental Biology, Vol. 13, pp. 83-117.

195. *Dynamic instability of individual microtubules analyzed by video light microscopy: rate constants and transition frequencies.* **Walker, R. A., et al.** 1988, The Journal of Cell Biology, Vol. 107, pp. 1437-1448.

196. *Microtubule structure at 8 Å resolution.* **Li, H., et al.** 2002, Structure, Vol. 10, pp. 1317-1328.

197. *Racking the ends: a dynamic protein network controls the fate of microtubule tips.* **Akhmanova, A. and Steinmetz, M.O.** 2008, Nat. Rev. Mol. Cell Bio., Vol. 9, pp. 309-322.

198. *Spindle microtubules in flux.* **Rogers, G.C., Rogers, S.L. and Sharp, D.J.** 2005, J. Cell. Sci., Vol. 118, pp. 1105-1116.

199. *Cell Division: Stages of Mitosis.* **O'Connor, C.** 1, 2008, Nature Education, Vol. 1, p. 188.

200. *Microtubules as a target for anticancer drugs.* **Jordan M.A., Wilson L.** 2004, Nat. Rev., Vol. 4, pp. 253-265.

201. *Mechanisms of chromosome behaviour during mitosis.* **Walczak, C.E., Cai, S. and Khodjakov, A.** 2010, Nature Reviews Molecular Cell Biology, Vol. 11, pp. 91-102.

202. *Antimitotic natural products and their interactions with tubulin.* **E., Hamel.** 1996, Med. Res. Rev., Vol. 16, pp. 207-231.

203. *Effects of vinblastine, podophyllotoxin and nocodazole on mitotic spindles. Implications for the role of microtubule dynamics in mitosis.* **Jordan, M.A., Thrower, D. and Wilson, L.** 1992, J. Cell. Sci., Vol. 102, pp. 401-416.

204. *Mechanisms of Action of and Resistance to Antitubulin Agents: Microtubule Dynamics, Drug Transport, and Cell Death.* **Dumontet, C. and Sikic, B.I.** 1999, J. Clin. Oncol., Vol. 17, pp. 1061-1070.

205. *Review: Tubulin Function, Action of Antitubulin Drugs, and New Drug Development.* **Pellegrini, F. and Budman, D.R.** 2005, Cancer Investigation, Vol. 23, pp. 264-73.

206. *Txol binds to cellular microtubules.* **Manfredi, J.J., Parness, J. and Horwitz, S.B.** 1982, J. Cell. Biol., Vol. 94, pp. 688-696.

207. *Identification and Characterization of an Intermediate Taxol Binding Site Within Microtubule Nanopores and a Mechanism for Tubulin Isotype Binding Selectivity.* **Freedman, H., et al.** 2009, J. Chem. Inf. Model, pp. 424-436.

208. *Antimicrotubular drugsbinding to vinca domain of tubulin.* **Gupta, S. and Bhattacharyya, B.** 2003, Mol. Cell. Biochem., Vol. 253, pp. 41-47.

209. *Interactions of colchicine with tubulin.* **Hastie, S.B.** 1991, Pharmacol. Ther., Vol. 512, pp. 377-401.

210. *Colchicine binding to tubulin monomers: a mechanistic study.* **S., Banerjee, G., Chakrabarti and B., Bhattacharyya.** 1997, Biochemistry, Vol. 36, pp. 5600-5606.

211. *Structure of the alpha beta tubulin dimer by electron crystallography.* **Nogales, E., Wolf, S.G. and Downing, K. H.** 1998, Nature, Vol. 391, pp. 199-203.

212. *GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation.* **Hess, B., et al.** 2008, J. Chem. Theory Comp, Vol. 435.

213. *Canonical sampling through velocity rescaling.* **Bussi, G., Donadio, D. and Parrinello, M.** 2007, J. Chem. Phys, Vol. 126, pp. 014-101.

214. *Haak Molecular dynamics coupled to an external bath.* **Berendsen, H. J. C., et al.** 1984, J. R. J. Chem. Phys., Vol. 81.

215. *A unified formulation of the constant temperature molecular dynamics methods.* **S., Nosé.** 1984, J. Che. Phys., Vol. 81, p. 511.

216. *Canonical dynamics: Equilibrium phase-space distributions.* **Hoover, W.J.** 1985, Phys. Rev. A: At. Mol. Pot. Phys., Vol. 31, p. 1695.

217. *Vinblastine perturbation of tubulin protofilament structure: A computational insight.* **Rendine, S., Pieraccini, S. and Sironi, M.** 2010, Phys. Chem. Chem. Phys., Vol. 12, pp. 15530-15536.

218. *Protein-protein interactions in receptor activation and intracellular signaling.* **Sibanda, B. L.** 2005, Biol Chem., Vol. 381, pp. 955-959.

219. *Binding of a diverse set of ligands to avidin and streptavidin: An accurate quantitative prediction of their relative affinities by a combination of molecular mechanics and continuum solvent models.* **Kuhn B., Kollman P.A.** 2000, J. Med. Chem., Vol. 43, pp. 3786-3791.

220. *Two distinct signal transmission pathways in T lymphocytes are inhibited by complexes formed between an immunophilin and either FK506 or rapamycin.* **Bierer, B.E., et al.** 1990, Proc.i. USA Natl.Acad. Sc, Vol. 87, pp. 9231-9235.

221. *Binding of a diverse set of ligands to avidin and streptavidin: An accurate quantitative prediction of their relative affinities by a combination of molecular mechanics and continuum solvent models.* **Kuhn, B. and Kollman, P.A.** 2000, J. Med. Chem., Vol. 43, pp. 3786-3791.

222. *Molecular Target of Rapamycin (mTOR) as a therapeutic target against cancer.* **Mita, M. M., Mita, A. and Rowinsky, E. K.** 2003, Cancer biology and therapy, Vol. 2, pp. 169-177.

223. *Solvation free energies estimated from macroscopic continuum theory: An accuracy assessment.* **Simonson, T. and Brunger, A.** 1994, J. Phys. Chem., Vol. 98, pp. 4683–4694.

224. *Discrimination between native and intentionally misfolded conformations of proteins: ES/IS, a new method for calculating conformational free energy that uses both dynamics simulations with an explicit solvent and an implicit solvent continuum model.* **Vorobjev, Y. N., Almagro, J. C. and Hermans, J.** 1998, Proteins, Vol. 32, pp. 399-413.

225. *Development and testing of a general Amber force field.* **Wang, J., et al.** 2004, J. Comput. Chem., Vol. 25, pp. 1157-1174.

226. *Comparison between computational alanine scanning and per-residue binding free energy decomposition for protein-protein association using MM-GBSA: Application to the TCR-p-MHC complex.* **Zoete, V. and Michielin, O.** 2007, Proteins: Struct. Funct. Bioinf., Vol. 67, pp. 1026-1047.