COMPUTATIONAL STUDY OF MICROTUBULE STABILITY AND ASSOCIATED CHEMOTHERAPEUTIC
AGENTS

by

Ahmed Taha Abd Elfattah Taha Ayoub

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy.

Department of Chemistry
University of Alberta

Abstract

Microtubules are cellular structures that are crucial to many cellular functions including maintenance of cell shape, vesicular transport, and cell division. The dynamic instability of microtubules is the basic feature which enables them to do their cellular functions. Their pivotal role in cell division makes them an important therapeutic target for cancer chemotherapeutic treatment. In this thesis, I have studied two major biological problems connected to microtubules: virtual screening for novel microtubule stabilizing agents, and energetic analysis of tubulin inter-dimer binding energies within microtubules.

In the virtual screening project a library of 33 million chemical compounds was screened against available microtubule stabilizing agents using similarity fingerprints and structure-based drug design techniques arriving at a novel scaffold predicted to bind at the taxol binding site. This novel scaffold shed light on the mechanism of antitumor action of lankacidin antibiotics, due to sharing a high degree of similarity, and was tested and partly confirmed computationally and experimentally.

In the microtubule energetic analysis project, various quantum chemical descriptors were tested and parameterized for the prediction of hydrogen bond energies. These descriptors and parameters were used in analyzing the strength of hydrogen bonds across the longitudinal inter-dimer interfaces through which tubulin dimers join head-to-tail to form protofilaments, and across the lateral inter-dimer interface through which protofilaments align side-by-side to form microtubule cylinders. As a continuation to this study, a molecular dynamics simulation of a complete microtubule was run, followed by a complete analysis and breakdown of MM/GBSA (Molecular Mechanics/Generalized Born-Surface Area) binding energies at lateral and longitudinal inter-dimer interfaces enabling thorough analysis of the contribution of each residue, domain, subunit, and dimer to the stability of a microtubule cylinder and shedding light on the driving force for microtubule disassembly and other important phenomena.

Preface

Some of the research conducted for this thesis forms part of an international research collaboration with Dr. K. Arakawa from Hiroshima University, Japan, T. Craddock from Nova Southeastern University, Florida and R. Ahmed from the University of Alberta, with Professor J.A. Tuszynski being the lead collaborator at the University of Alberta. In Chapter 3, isolation of lankacidin antibiotics was conducted by K. Arakawa, fluorescence quenching experiments were conducted by R. Ahmed, and computational predictions, research concept, hypothesis and plan of experiments were done by myself with the help of R. Ahmed, M. Klobukowski and J. Tuszynski.

Chapter 2 of this thesis has been published as A. T. Ayoub, M. Klobukowski, and J. Tuszynski, "Similarity-based virtual screening for microtubule stabilizers reveals novel antimitotic scaffold," *Journal of Molecular Graphics and Modelling*, vol. 44, 188-196, 2013. I was responsible for the simulations, modelling, data processing and analysis as well as the manuscript composition. M. Klobukowski and J. Tuszynski were the supervisory authors and were involved with concept formation and manuscript composition.

Chapter 4 of this thesis has been published as A. T. Ayoub, J. Tuszynski, and M. Klobukowski, "Estimating hydrogen bond energies: comparison of methods," *Theoretical Chemistry Accounts*, vol. 133, issue 8, 1520-1527, 2014. I was responsible for the calculations, data processing and analysis as well as the manuscript composition. J. Tuszynski and M. Klobukowski were the supervisory authors and were involved with concept formation and manuscript composition.

Chapter 5 of this thesis has been published as A. T. Ayoub, T. J. A. Craddock, M. Klobukowski, and J. Tuszynski, "Analysis of the strength of interfacial hydrogen bonds between tubulin dimers using quantum theory of atoms in molecules," *Biophysical Journal*, vol. 107, issue 3, 740-750, 2014. I was responsible for model building, simulations, data processing and analysis as well as the manuscript composition. T. Craddock assisted in model building and manuscript composition. M. Klobukowski and J. Tuszynski were the supervisory authors and were involved with concept formation and manuscript composition.

Chapter 6 of this thesis has been published as A. T. Ayoub, M. Klobukowski, J. Tuszynski, "Detailed Per-residue Energetic Analysis Explains the Driving Force for Microtubule Disassembly," *PLOS Computational Biology*, vol. 11, issue 6, e1004313, 2015. I was responsible for model building, simulations, data processing and analysis as well as the manuscript composition. J. Tuszynski and M. Klobukowski were the supervisory authors and were involved with concept formation and manuscript composition.

To my mother, my father, wife, kids and my beloved country Egypt

Acknowledgements

Contents

List of Tables

List of Figures

Acronyms

AIM Atoms in Molecules.

BCP Bond Critical Point.

FaSSIF Fasted-state Solubility in Simulated Intestinal Fluid.

GDP Guanosine Diphosphate.

GMPCPP Guanylyl-(alpha, beta)-Methylene-Diphosphonate.

GTP Guanosine Triphosphate.

HA Hydrogen bond Acceptor.

HD Hydrogen bond Donor.

LatA Lateral interface between two tubulin dimers in A-configuration.

LatB Lateral interface between two tubulin dimers in B-configuration.

LongAB Longitudinal interface between two tubulin dimers.

MD Molecular Dynamics.

MM/GBSA Molecular Mechanics/Generalized-Born Surface Area Method.

MM/PBSA Molecular Mechanics/Poisson-Boltzmann Surface Area Method.

MOC Mulliken Overlap Charge.

MSA Microtubule Stabilizing Agent.

MT Microtubule.

OWBO Overlap-Weighted Bond Order.

PDB Protein Data Bank.

Peff Effective Human Jejunal Permeability.

QTAIM Quantum Theory of Atoms in Molecules.

RMSD Root-Mean-Square Deviation.

SD Standard Deviation.

vdW van der Waals.

WBI Wiberg Bond Index.

Chapter 1

Introduction

In this thesis, we worked on microtubules as targets for cancer therapy and studied their stability and the effect of some drugs and factors on their cellular functionalities. In this chapter, some of the basics of the biology of cell division and the pathophysiology of cancer are introduced. The basics of microtubule structure and its dynamic instability are also covered. Finally, some basic ideas about the methodology employed and the basics of force fields and molecular dynamics as well as the quantum theory of atoms in molecules are given.

1.1 Cancer

Cancer is one of the leading causes of morbidity and mortality all over the world. In 2012, approximately 14 million new cancer cases and 8.2 million cancer related deaths were reported. According to the most recent statistics from the World Health Organization, the number of annual cancer cases is expected rise to 22 million within the next two decades [1]. Also known as malignant neoplasm, cancer is characterized by uncontrolled, unregulated cell growth within a certain tissue [2]. These malignant cells can invade nearby tissues or even distant tissues via the lymphatic system or the blood stream, a process called metastasis [3].

In 2000, Hanahan and Weinberg identified a list of six hallmarks of cancer [4], which was extended by the inclusion of other four hallmarks in 2011 [5]. The list of hallmarks include the following; 1) self-sufficiency in growth signals, 2) insensitivity to anti-growth signals, 3) evading apoptosis, 4) limitless replicative potential, 5) sustained angiogenesis, 6) tissue invasion and metastasis, 7) deregulated cellular energetics, 8) avoiding immune destruction, 9) genome instability and mutation, and 10) tumor-promoting inflammation.

Knowledge of these hallmarks of cancer is particularly important for designing effective treatment strategies. Since cancer is a cell growth disease, it is not possible to understand it or search for a cure for it without proper understanding of the cell growth and division process, mitosis.
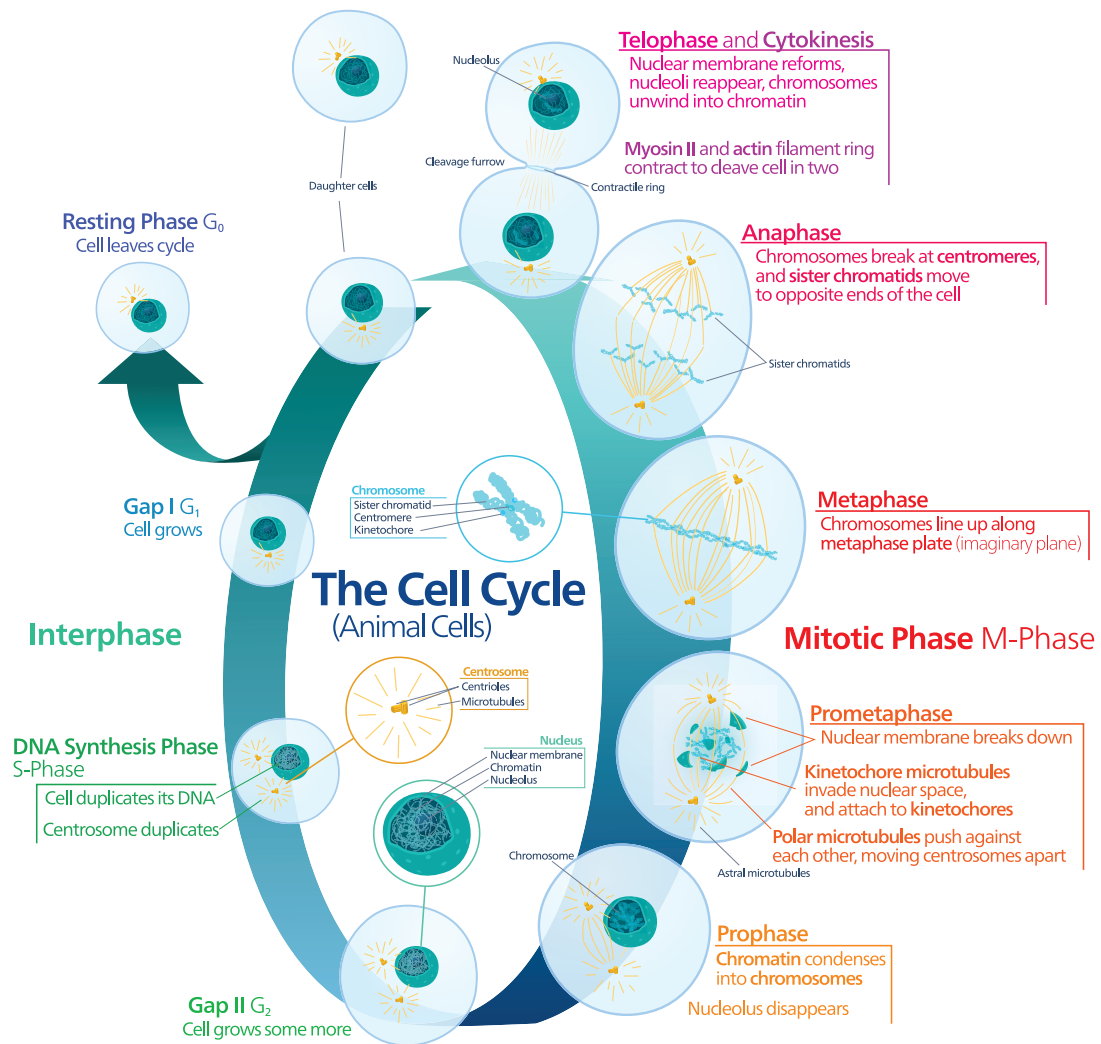
Figure 1.1: *Animal cell cycle and its different stages. Figure from [6].*

1.2 Mitosis

Mitosis is the biological process through which the cell divides its nuclear chromosomal content into two identical sets of chromosomes each in its own nucleus. This karyokinesis (nuclear division) is followed by cytokinesis where the cell membrane, organelles and cytoplasm are also equally divided between the two daughter cells [7].

Mitosis itself is one short stage in the cell cycle. As shown in Fig. 1.1, the cell cycle could be divided in four major phases; G_1 (first gap), S (synthesis), G_2 (second gap), the three making up the so-called interphase, and M (mitosis) phase. The cell grows in the G_1 phase and continues to grow through the S phase in which it also duplicates its chromosomes. The cell grows more in the G_2 phase as well and prepares for mitosis (M) [8]. Cells could also proceed to a resting phase (G_0) and stop dividing permanently or temporarily.

The M phase (Mitotic division) itself is subdivided into several stages. As shown in Fig. 1.1, it starts with the *prophase* in which the loosely packed nuclear chromatin condenses into discrete chromosomes and the nucleolus disappears. Centrosomes, the center for formation of microtubule spindle apparatus, duplicate and migrate to opposite sides of the cell. This is followed by the *prometaphase* where the nuclear envelope disintegrates and microtubules come into play. Polar microtubules from the opposing sides of the cell interact with one another forming the mitotic spindle and pushing the centrosomes farther apart. Kinetechore microtubules invade the nuclear space and attach themselves to the chromosomal kinetechores, located at the chromosomal centromeres. This structure is necessary for later separation of the sister chromatids in each chromosome. The cell then proceeds towards the *metaphase* where the coupling between the motor activity of kinetechores and the polymerization and depolymerization of microtubules provide enough force to pull chromosomes to the opposite poles of the cell. This causes the chromosomes to line up along the imaginary metaphase plate. In the subsequent phase of *anaphase*, sister chromatids start to separate and kinetechore microtubules start shortening and generating forces that pull chromatids apart to the opposite poles of the cell. Polar microtubules also push against one another, causing the cell to elongate. At the *telophase*, the cell elongates even more, via elongation of polar microtubules, and the nuclear membrane reforms around each set of the daughter chromosomes. Chromatin reforms and the nucleolus reappear. At this stage, mitosis is complete but the cell still did not divide. The remaining phase is called cytokinesis, where the two daughter cells are pinched off in the middle and the contractile ring forms a cleavage furrow which separates the two nuclei into two daughter cells.

1.3 Microtubules

In the previous overview of the mitotic cell division, it was clear that microtubules (MTs) are central players in its machinery. Without a proper functioning of the MT polymerization and depolymerization mechanisms, proper cell division may not be maintained. It is crucial, therefore, to understand the structure and function of MTs and their dynamic nature.
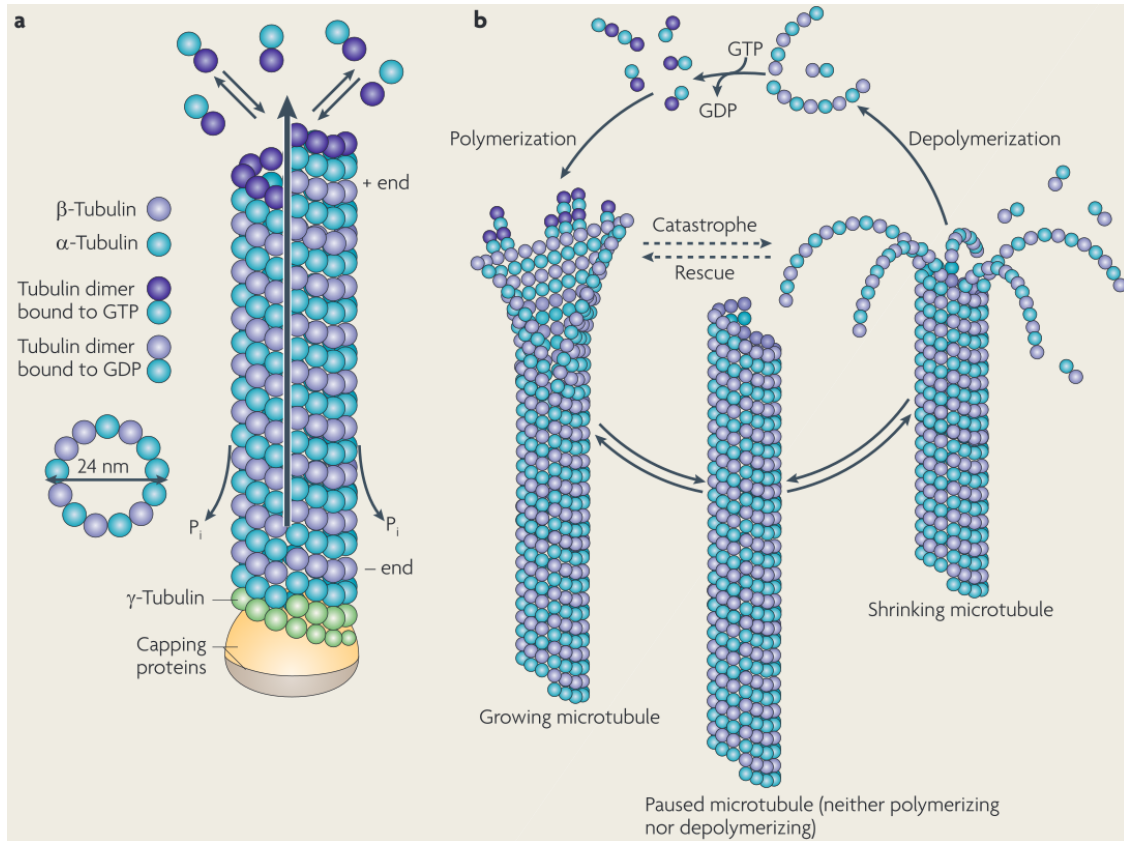
Figure 1.2: *Microtubule structure and its dynamic instability. Figure from [15].*

As Fig. 1.2a shows, the building block of MTs is the tubulin heterodimer which is a peanut-shaped protein composed of an α-subunit and a β-subunit [9]. The α-subunit is always bound to a non-exchangeable guanosine triphosphate (GTP) molecule at the N-site while the β-subunit is bound to an exchangeable GTP molecule at the E-site and it hydrolyzes to guanosine diphosphate (GDP) and inorganic phosphate shortly after assembly [10]. The GTP-bound $\alpha\beta$-tubulin heterodimers assemble longitudinally head-to-tail to produce a protofilament, 13 of which join laterally to produce an MT cylinder [11]. Due to the nature of this assembly, one end of MTs has β-subunits exposed, called the (+) end, while the other end has α-subunits exposed, called the (−) end. Growth usually occurs at the (+) end, where GTP-bound tubulin dimers attach and hydrolyze their GTP molecules shortly after assembly [12]. This results in a GTP cap always present at the (+) end of each growing microtubule. Mitchison and Kirschner proposed the so-called GTP-cap model which states that as long as the plus end of an MT is capped with GTP, it continues to grow. However, if GTP hydrolysis is sufficiently fast to catch up to the growing tip of the MT, rapid shrinkage, called a catastrophe, results [13]. Upon binding to an MT, however, some pharmacological agents such as taxol or epothilone stabilize the system and inhibit shrinkage [14]. The depolymerization process is characterized by protofilaments bending outward and displaying a structure that looks like ram's horns (see Fig. 1.2b). GTP-bound tubulin can begin adding to the tip of the MT again providing a new cap that protects against disassembly, a process termed rescue.

MTs generate forces within cells via the stochastic switching between growth and shrinkage, also known as dynamic instability [16]. These forces contribute to several biological processes within cells including mitotic spindle formation and chromosome segregation during cell division, maintenance of cellular morphology and cytoskeleton, and intracellular transport. MTs have also been implicated in playing direct or indirect roles in signaling, information processing, and consciousness and they are linked to cognitive diseases such as Alzheimer's disease.

The role of MTs in cell division makes them important targets for cancer treatment. Since the rapidly dividing tumor cells depend on MT during cell division, disruption of the subtle MT stability balance, either by stabilization or destabilization, should hinder the cell division process. Several available chemotherapeutic agents act by stabilizing MT structure, such as taxanes and epothilones, or destabilizing it and preventing tubulin polymerization, such as colchicine and vinca alkaloids. In both cases, cell division is arrested and the cells die by apoptosis. This cytotoxicity is not only limited to cancer cells, but also affects other rapidly dividing cells such as intestinal epithelium and bone marrow, leading to the known side effects of chemotherapy.

1.4 Computational Methodology

In this thesis, we simulated several MT stabilizing agents and studied their effects on tubulin and their free energies of binding. We also studied the stability of MTs and their binding energetics at lateral and longitudinal inter-dimer interfaces. These stability studies related to the MT dynamic instability and mechanisms of disassembly. In these studies, we employed different computational techniques including; similarity-based virtual screening, molecular docking, molecular dynamics simulations, Molecular Mechanics Poisson-Boltzmann (Generalized Born) Surface Area MM/PB(GB)SA free energy calculations and Density Functional Theory (DFT) calculations followed by population analysis employing the quantum theory of atoms in molecules (QTAIM). In the remaining part of this section, we will cover some of the basics of the methods employed throughout the thesis. The references that were primarily used in this section include the books by Cramer [17] and Bader [18] and the online CHARMM tutorial [19]. Other useful reviews could be found in [20] and [21].

1.4.1 Molecular Mechanics Force Fields

[1] A force field is a collection of parameters and equations used to calculate the potential energy of a system of atoms in molecular mechanics simulations and geometry optimizations. In calculating the potential energy of a system of atoms or molecules, several terms contribute to the overall energy of the system. In molecular mechanics, those terms represent either bonded interactions or non-bonded interactions. Bonded interactions include atoms which are separated by three bonds or less, including bond stretching terms, angle bending terms and torsional or dihedral terms. Non-bonded interactions include atoms that are either non-bonded or farther than three bonds apart

---
[1]This section is based on section 2.2 from the textbook by Cramer [17]

like van der Waals interactions and electrostatic interactions. With these contributions summed up over all interacting atoms, the overall force field equation is:

$$E_{\text{total}} = \sum E_{\text{bond}} + \sum E_{\text{angle}} + \sum E_{\text{torsion}} + \sum E_{\text{electrostatic}} + \sum E_{\text{vdW}} \qquad (1.1)$$

We will briefly introduce each term and discuss the corresponding formula.

1.4.1.1   Bond Stretching

In molecular mechanics force fields, the atoms are treated as solid spheres attached by springs where the potential energy can simply be calculated using the parameters of the springs and Hooke's law. The bond stretching term is evaluated according to the following equation:

$$E_{\text{bond}} = \sum_{bonds} K_r \left(r - r_{eq} \right)^2, \qquad (1.2)$$

where the sum runs through all the bonds in the system, K_r is the force constant, r_{eq} is the equilibrium bond length specific for each bond, and r is the distance between the two bonded atoms. Values of the force constant specific for each bond is obtained from either quantum mechanics calculations or from experimental data such as infrared stretching frequencies. Equilibrium bond lengths could be obtained from high resolution crystal structure and microwave spectroscopy. Although this harmonic potential is less accurate than the Morse potential in describing bond stretching terms, it is much cheaper in terms of computations and performs reasonably well around equilibrium bond lengths where the bond deformations are small.

1.4.1.2   Angle Bending

Following a similar principle, the angle bending term of the potential energy function could be evaluated according to the following harmonic potential:

$$E_{\text{angle}} = \sum_{angles} K_\theta \left(\theta - \theta_{eq} \right)^2, \qquad (1.3)$$

where K_θ and θ_{eq} are the force constant and equilibrium angle associated with the three atoms in question, respectively, and θ is the actual angle between the three bonded atoms. The sum runs through all the angles between bonded atoms in the system and assigns a penalty for each deviation from equilibrium angles.

1.4.1.3   Torsions

A torsional or dihedral angle ϕ between bonded atoms ABCD is defined as the angle between bonds AB and CD when they are projected into the plane bisecting the BC bond. The rotation of the BC bond changes the steric interactions between A and D, altering the potential energy of the system. This potential is periodic and is often expressed as a cosine function like the following form:

$$E_{\text{torsion}} = \sum_{torsions} \frac{V_n}{2} \left[1 + \cos(n\phi - \gamma) \right], \qquad (1.4)$$
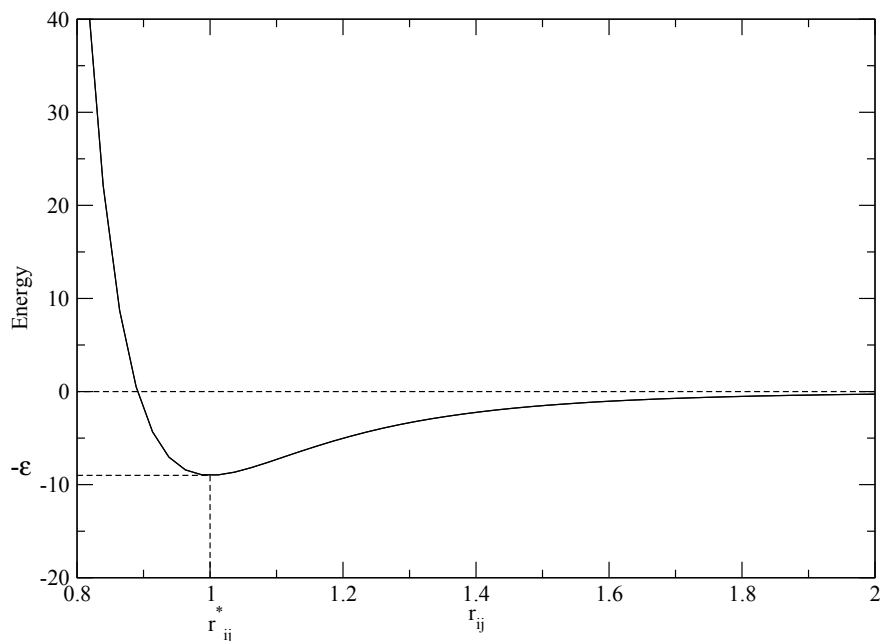
Figure 1.3: *Lennard-Jones potential curve showing optimum distances r_{ij}^* and well depth ε. Units on axes are arbitrary.*

where V_n is the term amplitude, n is the periodicity, and γ is the phase angle. The parameters of the previous bonded potentials are usually obtained from spectroscopic data of small model compounds aided with *ab inito* quantum calculations.

1.4.1.4 van der Waals Interactions

This type of non-bonded interactions has two components, an attractive component and a repulsive one. The attraction comes from fluctuations in the electron cloud around an atom giving rise to an instantaneous dipole which induces a dipole in a nearby atom giving rise to attractive forces. The repulsion comes from overlap of electron clouds at short distances. The two competing forces have different dependence on distance giving rise to a minimum in the potential energy curve. The curve is best described via the Lennard-Jones potential shown in Fig. 1.3. The total potential energy due to the van der Waals term is expressed in the following equation:

$$E_{vdW} = \sum_{i<j} \varepsilon \left[\left(\frac{r_{ij}^*}{r_{ij}} \right)^{12} - 2 \left(\frac{r_{ij}^*}{r_{ij}} \right)^6 \right], \tag{1.5}$$

where ε is the well-depth and r_{ij}^* is the optimum distance specific for each pair of atoms, and r_{ij} is the distance between the atoms. Because this pairwise energy function scales as N^2, where N is the number of atoms in the system , a typical cut-off distance of 10 Å is usually applied which reduces the number of calculations considerably. Since abrupt decline of energy at the cut-off distance may introduce discontinuities, a smooth switching function is usually applied.

7

1.4.1.5   Electrostatic Interactions

The second type of non-bonded interactions is the electrostatic interaction between pairs of atoms. If ϵ is the dielectric constant of the medium and r_{ij} is the distance separating two atoms having charges q_i and q_j, then the electrostatic potential energy becomes;

$$E_{ele} = \sum_{i<j} \frac{q_i q_j}{\epsilon r_{ij}} \tag{1.6}$$

Such a pairwise interaction would scale as N^2 which would become very demanding especially in periodic systems. The application of cut-off distances is not as an attractive solution as in the case of van der Waals interactions since electrostatic interactions decay in a much slower fashion. In other words, the long-range component of the Lennard-Jones potential decays proportionally to r^{-6} while electrostatic potential decays proportionally to r^{-1}. A better alternative for evaluating long-range electrostatic potential in infinite periodic systems is the Ewald sum technique. Through a reciprocal-space technique for treatment of long-range interactions, the total electrostatic interaction can be calculated to given level of accuracy with a scaling that is as low as $N\log N$ in the most favorable case (i.e. Particle-Mesh-Ewald).

1.4.2   Molecular Dynamics Simulations

[2] Molecular dynamics (MD) simulations belong to a group of computational methods often employed in the study of behavior and properties of biological molecules. Molecular dynamics simply provides a means, based on a potential energy function like the one described above, of simulating the evolution of the phase-space trajectory of a molecular system with time. It is a deterministic method, meaning that provided a point on the phase space trajectory of a molecular system, the "next" point in time can be determined. Statistical mechanics is at the core of molecular dynamics simulations where the microscopic information provided through the simulations is converted into macroscopic properties.

Molecular dynamics simulations generate a sequence of points in the phase space as a function of time. These points are different conformations of the system at a particular thermodynamic state, belonging to a certain thermodynamic ensemble. Since most of chemical processes take place under constant temperature and pressure, the isothermal-isobaric ensemble is often used for our molecular dynamics simulations. In a certain statistical ensemble, the ensemble average of any property A which is dependent on positions $\mathbf{p}$ and momenta $\mathbf{p}$ is expressed as follows:

$$\langle A \rangle_{\text{ensemble}} = \iint A(\mathbf{p}, \mathbf{q}) \rho(\mathbf{p}, \mathbf{q}) d\mathbf{p} d\mathbf{q} \tag{1.7}$$

where $\rho(\mathbf{p}, \mathbf{q})$ is the probability density function of the system being at a particular phase space point. This probability is dependent on the energy $E(\mathbf{p}, \mathbf{q})$ associated with the phase space point according to

$$\rho(\mathbf{p}, \mathbf{q}) = \frac{1}{Q} e^{-E(\mathbf{p}, \mathbf{q})/k_B T} \tag{1.8}$$

---

[2]This section is based on chapter 3 from the textbook by Cramer [17]

where k_B is the Boltzmann's constant, T is the temperature and Q is the partition function

$$Q = \iint e^{-E(\mathbf{p},\mathbf{q})/k_B T} d\mathbf{p} d\mathbf{q} \tag{1.9}$$

which could be viewed as a normalization constant for the probability density function ρ. Since the evaluation of these integrals requires knowledge about all the possible microstates of a system, it is extremely difficult to calculate. In a molecular dynamics simulation, however, we are provided with sequential points in time along the phase-space trajectory. In this case we can calculate the time average of the property of interest

$$\langle A \rangle_{time} = \lim_{\tau \to \infty} \frac{1}{\tau} \int_{t=0}^{\tau} A(t) dt \tag{1.10}$$

where $A(t)$ is the instantaneous value of the property A given the positions and momenta at time t, and τ is the simulation time. In the limit of an infinite simulation time, we can assume that the system will pass through all the possible microstates within a certain thermodynamic ensemble. This leads to one of the most fundamental hypotheses of statistical mechanics, i.e the ergodic hypothesis. It states that the time average equals the ensemble average:

$$\langle A \rangle_{ensemble} = \langle A \rangle_{time} \tag{1.11}$$

However, it is impossible to run a simulation for an infinitely long time and hence the limit in Eq. 1.10 can not be reached. This limitation can be circumvented by realizing that many points in the phase space are negligible. Eq. 1.8 shows that conformers having very high energies, for instance due to substantial molecular deformations, will have near-zero probabilities and hence the integrand in Eq. 1.7 will also be near-zero (as long as the property A does not approach infinity). This integrand will thus contribute negligibly to the property of interest and can be ignored as it does not contribute to the property expectation value. Starting with a reasonable, preferably experimentally-determined, structure, molecular dynamics provides a suitable prescription for sampling important phase space points (i.e. points having low energy and high probability) which contribute most to the property of interest. For a finite number of phase space points, the expression in Eq. 1.10 approximates to

$$\langle A \rangle_{time} \approx \frac{1}{M} \sum_{i=1}^{M} A(t_i) \tag{1.12}$$

where M is the number of times the property is sampled during the simulation, which reflects the number of important conformations sampled from the trajectory.

Given the positions of atoms in a system which are preferably obtained from experiment, initial distribution of velocities is assigned conforming to a certain temperature T to guarantee no overall momentum, i.e.

$$P = \sum_{i=1}^{N} m_i v_i = 0 \tag{1.13}$$

where the velocities v_i are assigned randomly through the Maxwell-Boltzmann distribution at a given temperature. The probability of an atom i having a velocity v_{ix} in the x direction is

$$p(v_{ix}) = \sqrt{\frac{m_i}{2\pi k_B T}} e^{-mv_{ix}^2/2k_B T} \tag{1.14}$$

and the overall temperature of the system can be obtained from the relationship

$$T = \frac{1}{3Nk_B} \sum_{i=1}^{N} \frac{|p_i|^2}{m_i} \tag{1.15}$$

where N is the number of atoms. Given this initial set of positions and momenta, the trajectory is traced and positions and momenta are re-calculated after every time step Δt until the simulation time is covered. Time steps are chosen so they are shorter than the fastest motion in the system, typically the heavy-atom-hydrogen bond having a period of 10^{-14} s. However, with the application of a constraint such that these bonds remain at a constant length, using for instance the SHAKE algorithm, a time step as long as 2 fs can be used. The next challenge in molecular dynamics is how the positions an momenta can be updated after each time step. This is achieved through either of the following integration algorithms.

1.4.2.1   Verlet Algorithm

The position after a time step Δt can be expressed using the Taylor series expansion:

$$\mathbf{q}(t + \Delta t) = \mathbf{q}(t) + \mathbf{v}(t)\Delta t + \frac{1}{2!}\mathbf{a}(t)(\Delta t)^2 + \frac{1}{3!}\frac{d^3\mathbf{q}(\tau)}{dt^3}\bigg|_{\tau=t}(\Delta t)^3 + ... \tag{1.16}$$

where velocity $\mathbf{v}$ and the acceleration $\mathbf{a}$ represent the first and second time derivatives of the position vector $\mathbf{q}$. Evaluating the Taylor expansion corresponding to a reverse time step,

$$\mathbf{q}(t - \Delta t) = \mathbf{q}(t) - \mathbf{v}(t)\Delta t + \frac{1}{2!}\mathbf{a}(t)(\Delta t)^2 - \frac{1}{3!}\frac{d^3\mathbf{q}(\tau)}{dt^3}\bigg|_{\tau=t}(\Delta t)^3 + ... \tag{1.17}$$

Summing the two equations and truncating at the second order (which is equivalent to truncating at the third order since it has a coefficient of zero) we get:

$$\mathbf{q}(t + \Delta t) = 2\mathbf{q}(t) - \mathbf{q}(t - \Delta t) + \mathbf{a}(t)(\Delta t)^2 \tag{1.18}$$

Hence, the Verlet algorithm updates the position using only the position and acceleration at time t and the position at time $(t - \Delta t)$ without any explicit reference to velocities. This makes the algorithm straightforward and less computationally demanding, but it also lacks velocity information which is often required at least for temperature control.

1.4.2.2   The Leapfrog Algorithm

As an alternative to the Verlet algorithm which lacks explicit velocity information, the leapfrog algorithm uses a Taylor series expansion about $(t + \Delta t/2)$ truncated at the second order as follows:

$$\mathbf{q}\left(t + \frac{1}{2}\Delta t + \frac{1}{2}\Delta t\right) = \mathbf{q}\left(t + \frac{1}{2}\Delta t\right) + \mathbf{v}\left(t + \frac{1}{2}\Delta t\right)\frac{1}{2}\Delta t + \frac{1}{2!}\mathbf{a}\left(t + \frac{1}{2}\Delta t\right)\left(\frac{1}{2}\Delta t\right)^2 \tag{1.19}$$

and

$$\mathbf{q}\left(t + \frac{1}{2}\Delta t - \frac{1}{2}\Delta t\right) = \mathbf{q}\left(t + \frac{1}{2}\Delta t\right) - \mathbf{v}\left(t + \frac{1}{2}\Delta t\right)\frac{1}{2}\Delta t + \frac{1}{2!}\mathbf{a}\left(t + \frac{1}{2}\Delta t\right)\left(\frac{1}{2}\Delta t\right)^2 \quad (1.20)$$

Subtracting Eq. 1.20 from Eq. 1.19 we obtain

$$\mathbf{q}\left(t + \Delta t\right) = \mathbf{q}(t) + \mathbf{v}\left(t + \frac{1}{2}\Delta t\right)\Delta t \quad (1.21)$$

A similar expansion for $\mathbf{v}$ gives

$$\mathbf{v}\left(t + \frac{1}{2}\Delta t\right) = \mathbf{v}\left(t - \frac{1}{2}\Delta t\right) + \mathbf{a}\left(t\right)\Delta t \quad (1.22)$$

In this algorithm, the velocities are calculated at $(t + \Delta t/2)$ and are then used to calculated the positions at $(t + \Delta t)$ and thus the velocities leap over positions and then the positions leap over the velocity, hence the name. Other integration algorithms also exist such as the velocity Verlet and the Beeman's algorithm.

1.4.3 Energy Calculations with MM/PB(GB)SA Algorithm

This approach was first described by Kollman et al. in 2000 [22]. After a molecular dynamics simulation of a ligand bound to a receptor protein is run in a solvent and in the presence of counterions, an equilibrated system of the ligand-receptor complex is obtained. A set of uncorrelated snapshots from the trajectory is then post-processed by removal of the solvent and counterions. The average free energy of each species, ligand, receptor or complex, is calculated according to the following equation:

$$\bar{G} = \bar{E}_{\mathrm{MM}} + \bar{G}_{\mathrm{PB(GB)SA}} - TS_{\mathrm{MM}} \quad (1.23)$$

where $\bar{E}_{\mathrm{MM}}$ is the average molecular mechanical energy calculated according to Eq. 1.1. The term $\bar{G}_{PBSA}$ refers to the solvation free energy obtained from the Poisson-Boltzmann (or generalized Born) equation for implicit solvent model and an estimate for the non-polar free energy via a surface area term, vide infra. The term $-TS_{MM}$ represents the entropic component which involves translational, rotational and vibrational contributions. Vibrational contributions can be calculated via normal mode analysis where different conformations are minimized and the Hessian matrix is diagonalized, then vibrational entropies can be estimated from vibrational frequencies. After calculating the average free energies for the ligand, the receptor and the complex, the total average free energy of binding in solution equals:

$$\Delta G = \bar{G}_{complex} - \bar{G}_{receptor} - \bar{G}_{ligand} \quad (1.24)$$

We will give a brief introduction about the solvent effects in the continuum model PB, GB and the surface area based method. Extended discussions on continuum models can be found in chapter 11 of the text book by Cramer [17].

1.4.3.1 Poisson-Boltzmann Electrostatics

The Poisson-Boltzmann equation describes the electrostatic environment of a solute present in an ion-containing solvent. It has the following form:

$$\vec{\nabla} \cdot \left[\epsilon(\vec{r}) \vec{\nabla} \Psi(\vec{r}) \right] = -\rho^f(\vec{r}) - \sum_i c_i^\infty z_i q \lambda(\vec{r}) e^{-z_i q \Psi(\vec{r})/k_B T} \tag{1.25}$$

where $\epsilon(\vec{r})$ represents the position-dependent dielectric, $\Psi(\vec{r})$ represents the electrostatic potential, $\rho^f(\vec{r})$ represents the charge density of the solute, c_i^∞ represents the concentration of ion i at a distance of infinity from the solute, z_i is the valence of the ion, q is the charge of a proton, k_B is the Boltzmann constant, T is the temperature, and $\lambda(\vec{r})$ is a factor for the position-dependent accessibility of position r to the ions in solution. If the potential is not large, the equation can be linearized to be solved more efficiently.

1.4.3.2 Generalized Born Electrostatics

If the solute is treated as a set of spheres with an internal dielectric constant that is different from the external solvent, the generalized Born model can offer an approximation to the exact linearized Poisson-Boltzmann equation with the following form:

$$G_s = \frac{1}{8\pi} \left(\frac{1}{\epsilon_0} - \frac{1}{\epsilon} \right) \sum_{i,j}^N \frac{q_i q_j}{f_{GB}} \tag{1.26}$$

where

$$f_{GB} = \sqrt{r_{ij}^2 + a_{ij}^2 e^{-D}} \tag{1.27}$$

and

$$D = \left(\frac{r_{ij}}{2a_{ij}} \right)^2, \qquad a_{ij} = \sqrt{a_i a_j} \tag{1.28}$$

In the equation, ϵ_0 is the permitivity of free space while ϵ is the solvent dielectric constant. q_i is the charge on atom i and r_{ij} is the distance between atoms i and j. The quantity a_i is the effective Born radius, reflecting its degree of burial within the solute.

1.4.3.3 Accessible Surface Area Based Method

Contrary to the PB or GB models which only estimate the enthalpic components of free energy, the accessible surface area method gives the total free energy of solvation. This is calculated as follows:

$$\Delta G_{\text{SA}} = \sum_{i=1}^N \sigma_i \gamma_i \tag{1.29}$$

where γ_i is the accessible surface area of atom i and σ_i is the solvation parameter of the same atom. The solvation parameter is defined as the contribution of the atom to the free energy of solvation per unit surface area. The free energy of solvation mounts to the change in free energy upon transfer of the atom/molecule from the solvent to vacuum (or other solvents) which is determined from the partition coefficients of the compounds between different media.
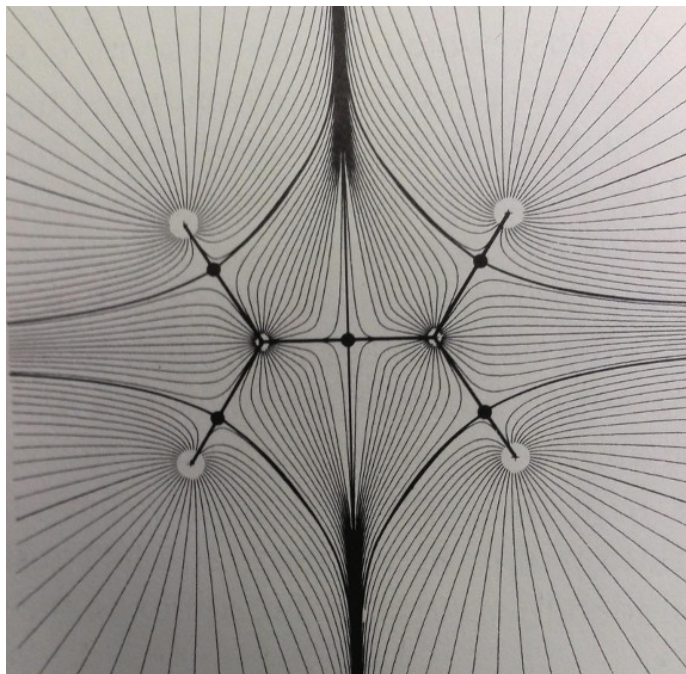
Figure 1.4: *A map of the electron density gradient vector field for the plane containing the nuclei of the ethene molecule. Each line represents a trajectory traced out by the vector $\boldsymbol{\nabla}\rho(\mathbf{r})$. Image adapted from [23].*

1.4.4   Quantum Theory of Atoms in Molecules

The quantum theory of atoms in molecules (QTAIM) is a theory that was developed by Richard Bader and his research group at McMaster University over the course of decades. The theory characterizes atoms and bonds in a molecular structure based on the topology of the electron density. A molecular structure is characterized by stationary points of the electron density together with the gradient paths of the electron density that originate and terminate at these points.

According to quantum mechanics, the nuclei of atoms in molecular structures act as point attractors immersed in a cloud of negative charge. The distribution of the negative charge in space around the nuclei is described by the electron density distribution function $\rho(\mathbf{r})$ where $\mathbf{r}$ is the space coordinates of a single electron. The charge density is calculated through the following equation:

$$\rho(\mathbf{r}) = N \int d\tau \, \psi^*(\mathbf{x})\psi(\mathbf{x}) \tag{1.30}$$

where N is the number of electrons, τ represents the spin coordinates of all electrons and the Cartesian coordinates of all electrons but one, and ψ is the state function expressed in $\mathbf{x}$ which is the collection of electronic space and spin coordinates.

Each topological feature of the charge density function, including maxima, minima, and saddle points, is characterized by a point in space called a critical point, denoted by the vector $\mathbf{r_c}$. This is a point in which the first derivative of the density function is zero, i.e. $\boldsymbol{\nabla}\rho(\mathbf{r}) = 0$. The gradient

is evaluated as follows:

$$\boldsymbol{\nabla}\rho = \mathbf{i}\frac{\partial \rho}{\partial x} + \mathbf{j}\frac{\partial \rho}{\partial y} + \mathbf{k}\frac{\partial \rho}{\partial z} \tag{1.31}$$

Whether the function has a maximum or a minimum at a stationary point is dictated by the second derivative, i.e. the curvature, at that point. The second derivative of the charge density function being the trace of the Hessian matrix is expressed as follows:

$$\boldsymbol{\nabla}^2\rho = \boldsymbol{\nabla}.\boldsymbol{\nabla}\rho = \frac{\partial^2 \rho}{\partial x^2} + \frac{\partial^2 \rho}{\partial y^2} + \frac{\partial^2 \rho}{\partial z^2} \tag{1.32}$$

The principal axes and their corresponding curvatures at the critical points are obtained as the eigenvectors and the corresponding eigenvalues in the diagonalization of the Hessian matrix of $\rho(\mathbf{r_c})$. A critical point could thus be characterized by two values; the rank (ω) and the signature (σ). The rank is the number of non-zero eigenvalues or curvatures of ρ at the critical point, and the signature is the algebraic sum of the signs of the eigenvalues. There are four different combinations of (ω,σ) for four different types of critical points:

- $(3,-3)$: where all curvatures are negative, signifying a local maximum in ρ. This usually indicates the position of the nuclei or the atoms, thus called atomic critical point.

- $(3,-1)$: where two curvatures are negative indicating two maxima and one is positive indicating a minimum. This occurs between two neighboring atoms and defines a bond between them, thus called a bond critical point.

- $(3,+1)$: where two curvatures are positive indicating two minima and one is negative indicating a maximum. This point occurs in the middle of several bonds forming a ring, thus called a ring critical point.

- $(3,+3)$: where all curvatures are positive indicating minima. This is found between several points that form a cage and thus called cage critical point.

Fig. 1.4 shows a map of the gradient vector field of the electron density $\boldsymbol{\nabla}\rho(\mathbf{r})$ at the plane of the nuclei in the ethene molecule. The figure shows the trajectory lines that terminate at the nuclei (the small circles). The set of trajectories that terminate at a given nucleus represent the basin of that nucleus. The figure also shows the sets of trajectories which terminate and originate at bond critical points (denoted by dots).

1.5   Scope of Thesis

In the next chapters, our analysis of the dynamic instability of microtubules and associated chemotherapeutic agents is presented. In Chapter 2, the results of virtual screening for novel microtubule stabilizing agents based on similarity fingerprints and structure-based design are presented. In Chapter 3, it is shown how the novel hits lead to the prediction of the mechanism of antitumor action of lankacidin group antibiotics and how this was confirmed experimentally. In Chapter 4,

analysis of different methods of estimation of hydrogen bond energies is performed and the parameters and descriptors needed for estimation are provided. These parameters and descriptors are used in Chapter 5 to estimate hydrogen bonds between tubulin dimers across lateral and longitudinal interfaces. Finally, Chapter 6 presents an atomistic simulation of a complete microtubule model followed by a detailed per-residue MM/GBSA analysis which unravels the driving force for microtubule disassembly.

Chapter 2

Similarity-Based Virtual Screening for Microtubule Stabilizers Reveals Novel Antimitotic Scaffold

2.1 Summary

Microtubules are among the most studied and best characterized cancer targets identified to date. Many microtubule stabilizers have been introduced so far that work by disrupting the dynamic instability of microtubules causing mitotic block and apoptosis. However, most of these molecules, especially taxol and epothilone, suffer absorption, toxicity and/or resistance problems. Here we employ a novel similarity-based virtual screening approach in the hope of finding other microtubule stabilizers that perform better and have lower toxicity and resistance. Epothilones, discodermolide, eleutherobin and sarcodictyin A have been found to compete with taxanes for the β-tubulin binding site, which suggests common chemical features qualifying for that. Our approach was based on similarity screening against all these compounds and other microtubule stabilizers, followed by virtual screening against the taxol binding site. Some novel hits were found, together with a novel highly rigid molecular scaffold. After visual manipulations, redocking and rescoring of this novel scaffold, its affinity dramatically increased in a promising trend, which qualifies for biological testing.

2.2 Introduction

Cancer is characterized by uncontrolled growth of cells within certain tissues that could spread to other tissues through metastasis. One of the most important cancer drug targets is the structural protein tubulin, the building block of microtubules. Microtubules are long filamentous tube-shaped protein polymers required in cells, among numerous other roles, for transport of vesicles and struc-

---

tural support [24]. During the anaphase of the mitotic division, microtubules play the role of separating chromosomes to the poles of cell preceding telophase. Proper dynamics of microtubules, in terms of the delicate balance between catastrophe and rescue phases, are crucial for these mitotic phases to complete [24]. Taxol and other anticancer microtubule stabilizing agents (MSAs) work by stabilizing the microtubules and thus disrupting their required dynamic instability, which leads to mitotic arrest followed by apoptosis [14,25]. The binding site of taxol was correctly identified to be in the β-tubulin of the $\alpha\beta$-tubulin heterodimer and was elucidated via electron crystallography and photoaffinity labeling [26–28]. Among the recognized MSAs are also epothilone, discodermolide, sarcodictyin and eleutherobin, all of which have been shown to compete with taxol for the β-tubulin binding site [29–32]. Also, laulimalide, which is an MSA, was shown to have affinity toward the taxol binding site [33,34] but its primary binding location has been shown to differ from that for taxol [35]. Fig. 2.1 shows the two-dimensional structures of these molecules. Most of these MSAs, however, suffer from some solubility, toxicity and/or drug resistance problems [36–39] revealed in pre-clinical and clinical trials. This strongly motivates the search for novel tubulin-binding agents that may be able to provide a better pharmacokinetic profile and evade the resistance caused by mutations or cellular efflux pumps. Several studies have been conducted to find out a common pharmacophore for the MSAs that compete with taxol for the binding site and some models have been presented [40–42]. However, the lack of definitive data regarding the bioactive conformation of some of these ligands, besides the high structural complexity of most of them poses some questions regarding the reliability of these pharmacophore models in virtual screening. Therefore, we employed another approach in this study, which is the similarity-based approach, for filtering large sets of molecules, instead of the pharmacophore approach. We also employed docking and virtual screening techniques supported by Tanimoto similarity [43] in the hope of finding new scaffolds for tubulin-binding agents that can strongly bind to the taxol binding site whilst having better molecular properties. For a pair of similar molecules A and B, Tanimoto similarity coefficient $J(A, B)$ is defined as

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \tag{2.1}$$

which is the ratio between the intersection and the union of the fingerprints of both molecules.

2.3 Computational Methods

2.3.1 Filtering PubChem Compound Library

As explained earlier, paclitaxel, epothilone A, discodermolide, sarcodictyin A, eleutherobin and laulimalide all have affinities toward the taxol binding site [29–34]. Based on that, we considered the ensemble of chemical features represented in these molecules as being representative of most of the chemical features required for binding to the taxol binding site. Hence, similarity with any of these six compounds was used as a criterion for filtering the compound database. The PubChem compound library, which is composed of nearly 33 million chemical entities, was screened against

(a) *Epothilone A*   **(b)** *Paclitaxel*   **(c)** *Eleutherobin*

(d) *Discodermolide*   **(e)** *Sarcodictyin A*   **(f)** *Laulimalide*
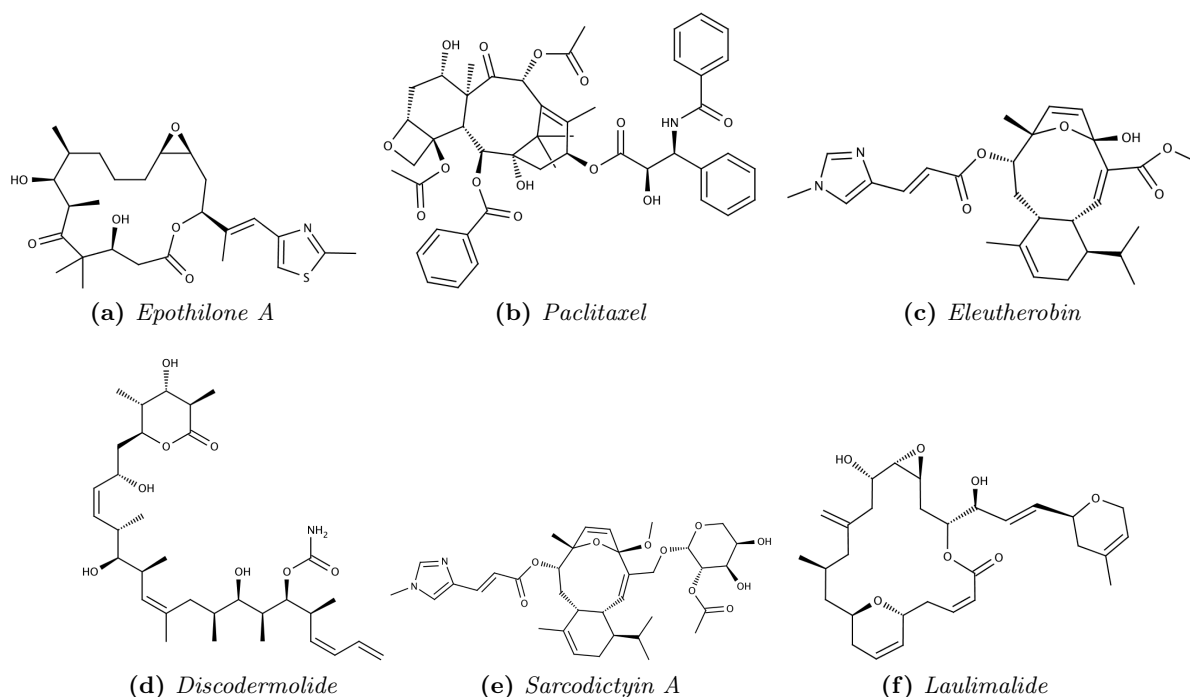
Figure 2.1: *Structures of microtubule stabilizing agents that have affinity towards the taxol binding site on β-tubulin.*

each of these six compounds separately to filter all the molecules that are at least 80% similar to the query molecules based on the Tanimoto coefficient similarity metric [43]. This threshold was chosen based on a trade-off between the need to retain as large a number as possible to avoid false negatives and the need to filter out as large a number as possible to increase the efficiency and speed. The filtered molecules were then subjected to another filtering step based on Lipinski's rule of five that is widely accepted as a criterion from a pharmacological usefulness viewpoint [44]. The rule states that orally active drugs should have no more than one violation of the following criteria: 1) no more than 5 hydrogen bond donors, 2) no more than 10 hydrogen bond acceptors, 3) a molecular mass less than 500 daltons, and 4) an octanol water partition coefficient ($\log P$) of less than 5.

The filtered molecules based on the six query structures were then combined together to give a library of 1591 molecules. Duplicates were deleted, the library was clustered and a diversity subset was selected based on 90% Tanimoto similarity to avoid having to evaluate molecules that are almost identical in structure. We ended up creating a library of 645 molecules that satisfy all of the above criteria and are ready for docking. We have added epothilone A as positive control. The 646 molecules were then docked into the taxol binding site using AutoDock 4.2 [45]. Details of the filtration process are depicted in Fig. 2.2. The Lamarckian genetic algorithm, which is a hybrid genetic algorithm with local search, was used for docking, with a maximum of 3 million runs of energy evaluations in a population size of 500 and a maximum of 50,000 generations. Docking poses were clustered based on a root-mean-square tolerance of 2.0 A and the results were ranked
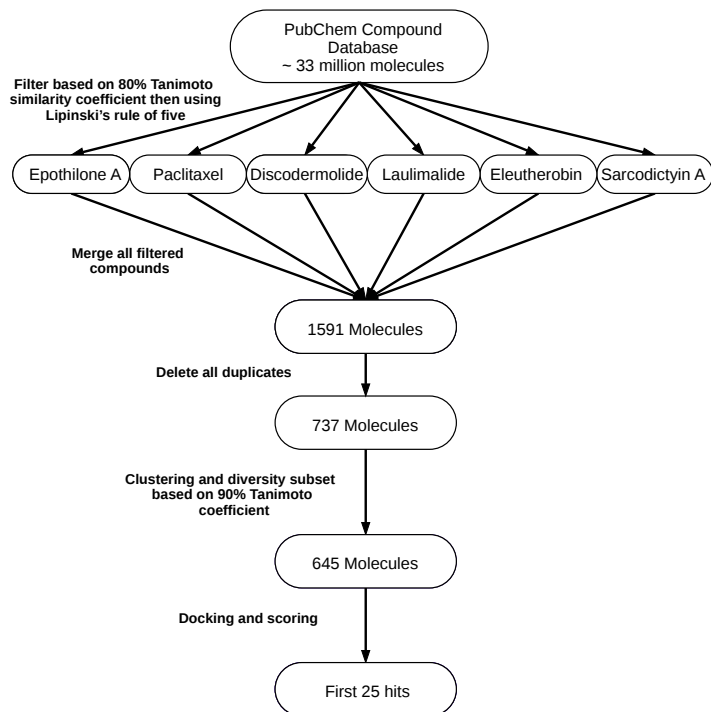
Figure 2.2: *Filtration scheme of the PubChem Compound Library of ∼ 33 million molecules preceding docking.*

based on the lowest energy in the largest cluster.

The top 25 hits were assessed for novelty of the scaffold and the best candidate was modified visually. MarvinSketch 5.10.3, 2012 (ChemAxon) was used for drawing, displaying and characterizing chemical structures. Avogadro 1.0.3 was also used to assist in visualization and for energy minimization of the structures prior to docking [46]. The best candidate from the virtual screening run was modified and several functional groups were added to different sites in the molecule to optimize the ligand-receptor complex based on visual inspection of the complex structure. The modified molecules were then redocked into the taxol binding site. The visual modification and redocking steps were repeated five times until the derivatives showed highest affinity and no further improvements were being generated. Hybridization between the best scoring hits in each run was utilized to optimize the results in the subsequent run.

2.3.2 Preparing Receptor for Docking

The receptor, tubulin, was obtained from the crystal structure 1TVK from the Protein Data Bank. 1TVK was preferred over 1JFF and 1TUB because of its high resolution (2.89 Å as compared to 3.50 Å and 3.70 Å, respectively). The molecule was minimized using AMBER 12 [47] to relieve any unfavorable interactions that may result from crystal packing. The receptor was parameterized according to the AMBERff99SB force field [48]. Hydrogen atoms were added to the protein based on the built-in AMBER residue templates, missing fragments were capped and residue His227 was

kept in the deprotonated state [49]. The Antechamber module of AMBER 12 package was used for parameterizing the ligands (epothilone, GDP and GTP) via the general AMBER force field [50] and atomic partial charges were assigned using the AM1-BCC method [51]. The complex was built using the *tleap* module of AMBER which was also used to neutralize the complex by addition of 18 sodium ions followed by extra 22 sodium and chloride ions to adjust ion concentration to 100 mM. The complex was solvated through a truncated octahedral box of the TIP3P water model that extended 12 Å in each direction. The model was minimized using AMBER 12 in a series of steepest descent and conjugate gradient steps. The minimized structure was then processed for docking; the α-tubulin subunit was deleted together with the solvent and all the ligands as they play no role in binding to the identified location on β-tubulin. Only the β-subunit was used for docking, at the epothilone binding site, which is the same as the taxane binding site. AutoDock Tools 4.2 was used to process the receptor [45]. The grid box for docking was centered around the crystallized epothilone structure in the binding site of 1TVK and was given the dimensions of $58 \times 58 \times 58$ grid points and was ready for docking. The software VMD was used for visualizing molecular dynamics (MD) trajectories [52].

2.3.3   Rescoring the best hits using MM/PBSA

Five of the six query molecules that were used for the similarity-based filtering (see Fig. 2.2) plus three of the highest-scoring hits that were obtained from the consecutive modification/redocking runs had their binding energies rescored using molecular mechanics/Poisson-Boltzmann surface area method (MM/PBSA) approach that was introduced by Kollman et al. [22]. This approach has been applied successfully in predicting the binding energies of many protein-ligand complexes [53–57]. Building on the success of these previous trials, we first rescored the five test molecules that interact with the taxol binding site with available experimental binding energies, namely discodermolide, eleutherobin, epothilone A, paclitaxel and sarcodictyin A, to verify the predictive power of the method through a linear fit. We neglected laulimalide because there was no available experimental data regarding its binding to the taxol binding site. Then we ran the same protocol with the highest-scoring hits to confirm the ranking as well as the binding energies predicted by the docking algorithm.

All the ligand-receptor complexes were subjected to molecular dynamics simulation to obtain an equilibrated system. We used only the β-tubulin subunit for the simulations because the α-subunit is not involved in binding and to save computational time. The simulations were performed using AMBER 12 package [47] following the same procedures as described in Sec. 2.3.2. The GDP molecule that is normally attached to β-tubulin in the 1TVK structure was also parameterized following the same approach as other ligands and was docked into its binding site when forming the complexes to make the model as realistic as possible. The ligand-receptor complexes were built using the *tleap* module of AMBER and were solvated with a truncated octahedral box of TIP3P water model extending 12 Å in each direction. The charge was neutralized by the addition of 18 sodium ions, and then 22 sodium and chloride ions were added to adjust the ion concentration to

100 mM which mimics cellular conditions. Each complex was then minimized without SHAKE on hydrogen and with very strong restraints on the protein through 500 steepest descent runs followed by 500 conjugate gradient steps. This was to relieve any unfavorable hydrogen contacts caused by the addition of hydrogen by *tleap*. Then, the complexes were minimized once more without restraints or SHAKE through 2000 steps of steepest descent followed by 2000 steps of conjugate gradient to remove any other unfavorable contacts. The complexes were then subjected to 20 ps of heating using Langevin thermostat to 300 K at constant volume under SHAKE and restraints on the protein. This was followed by 200 ps of density equilibration at constant pressure with SHAKE on hydrogen but without restraints on the protein. Each complex was then taken through a production phase of 11 ns under constant pressure with the same settings as the previous density equilibration step. All the simulations were performed under periodic boundary conditions using Particle-Mesh-Ewald method for treating long-range electrostatics and Langevin thermostat for temperature control. Coordinates were recorded every 2 ps in the production run to ensure that the snapshots are uncorrelated. Total energy, temperature, density and root-mean-square deviation (RMSD) of the complexes were then checked for equilibration. Complexes were then subjected to an MM/PBSA calculation in AMBER 12 to calculate the binding enthalpy under an ionic strength of 100 mM and an internal and external dielectric constants of 1.0 and 80.0, respectively. Although computationally very demanding, the entropies of the solutes were estimated using normal mode analysis in order to be able to calculate the total ΔG of the molecules and compare them to available experimental values as a check of the predictive power of the method in reproducing experimental data and hence being reliable in predicting the binding energy of the newly proposed ligands.

2.3.4 Prediction of Physicochemical Properties

The compounds that showed high affinities during the docking/rescoring runs were checked for their physicochemical parameters. ADMET Predictor software [58] was used for calculation of such parameters. This software uses predictive neural networks ensemble models with 2D and 3D descriptors. Parameters such as partition coefficient (log P), distribution coefficient (log D), effective human jejunal permeability (Peff), and fasted-state solubility in simulated intestinal fluid (FaSSIF) were calculated. Potential toxicity was also assessed by the same software.

2.4 Results and Discussion

2.4.1 Results of Virtual Screening
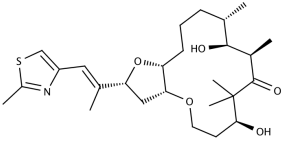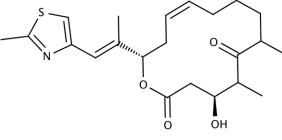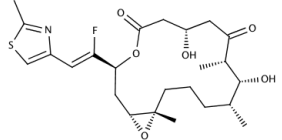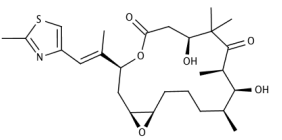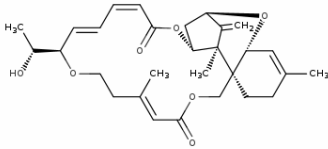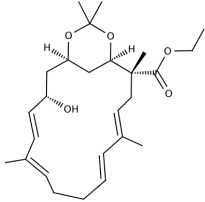
As described previously in Fig. 2.2, the similarity-based filtering of the ~ 33 million molecules yielded 645 molecules. These molecules, together with epothilone A used as a positive control, were docked with β-tubulin as described in Sec. 2.3.1. After ranking the docking results based on the lowest energy of the largest cluster, the best-scoring 25 molecules were examined. Before listing the results, it is worth mentioning that most of the highest scoring molecules are expected to be well-known and probably patented molecules. The approach that we are following, i.e. the

similarity approach, retains molecules that are similar to active ones, including their patented analogs. If these analogs occur on top of the list, this is expected. Our main objective it to identify novel scaffolds, among these top hits, that can be modified to generate better ligands. The virtual screening algorithm predicted the first 25 hits to be strong tubulin binding agents spanning a binding energy from -10.03 kcal/mol for the best hit (ligand PubChem CID of 10718437) to -9.02 kcal/mol for the worst one (ligand CID of 643649). To prove the efficiency of the docking algorithm in predicting active molecules, we will discuss the first five hits. From SciFinder, we found that the best hit, with CAS Registry Number of 198475-16-0, is an epothilone analog that was shown by Su et al. to have efficacy against cancer cell lines as a microtubule stabilizer [59]. The second hit, CAS 219990-00-8, also has an activity against tubulin, similar to that of epothilone, and was patented by Vite et al. (patent number WO 9902514) and Francis Lee (WO 2002066038) for treatment of hyperproliferative cellular disease and refractory tumors, respectively. The third hit, which is a laulimalide analog with CAS of 1049737-16-7, was shown by Mooberry et al. to possess potent antiproliferative activity due to microtubule stabilization [60]. The fourth hit, CAS 220889-57-6, is an epothilone analog proven active by Lam et al. [61]. Finally, the fifth hit, CAS 252917-46-7, was shown by Hardt et al. to be a potent epothilone analog [62]. In fact, all the top 25 hits, except for five novel ligands, represent compounds that have experimentally been shown to have tubulin stabilizing activities, which proves the reliability of our docking algorithm in determining active molecules.

The five novel hits found in virtual screening are presented in Tab 2.1. and compared to the positive control, epothilone A. The first of these five molecules, molecule 53 (CID 44333743), was shown by Sefkow et al. to have weak cytotoxic activity against the L 929 mouse fibroblast cell line, but was not patented [63]. Molecule 70 (CID 23242270) and 84 (CID 20822803) were not yet shown, experimentally, to have binding affinity toward tubulin. However, these three molecules have obvious similarity to the epothilone scaffold (molecule 44) and hence represent no new scaffold for us to adopt. Molecule 118 (CID 10625150) is 12,13-deoxyroridin E which is produced by the marine-derived fungus *Myrothecium roridum* and was shown by Namikoshi et al. to have cytotoxicity against human (HL-60) and murine (L1210) leukemia cell lines but until 2013 it has not yet been patented as an anticancer tubulin-binding agent [64]. This all shows that the protocol was strong enough in separating active from non-active molecules, with most of the strongly active molecules appearing on top of the list. The most attractive molecule is molecule 168 (CID 10477941) for it represents a completely new scaffold in this category of compounds. Not only is the scaffold entirely new, but its structure is much simpler than many tubulin binding agents and appears to be very rigid. Due to its promising characteristics, we decided to undertake further optimization runs for this molecule through modification and redocking.

As described in Sec. 2.3.1, we modified molecule 168 by addition of several functional groups to different positions in the molecule. The modification was based on visual inspection of the binding mode of molecule 168 to tubulin and on trying to optimize binding and the hydrogen bond network. We ran this modification/redocking step four consecutive times till the derivatives showed

Table 2.1: *Binding energies of the five novel hits obtained in the virtual screening run, plus the positive control, epothilone A.*

Rank	ID[a]	Molecule	ΔE^b
6	53		-9.58
17	70		-9.24
21	84		-9.08
22[c]	44		-9.07
23	118		-9.06
24	168		-9.03

[a] These are the IDs that we gave to the molecules; [b] Binding energy from AutoDock4.2 in kcal/mol; [c] The positive control, epothilone A.
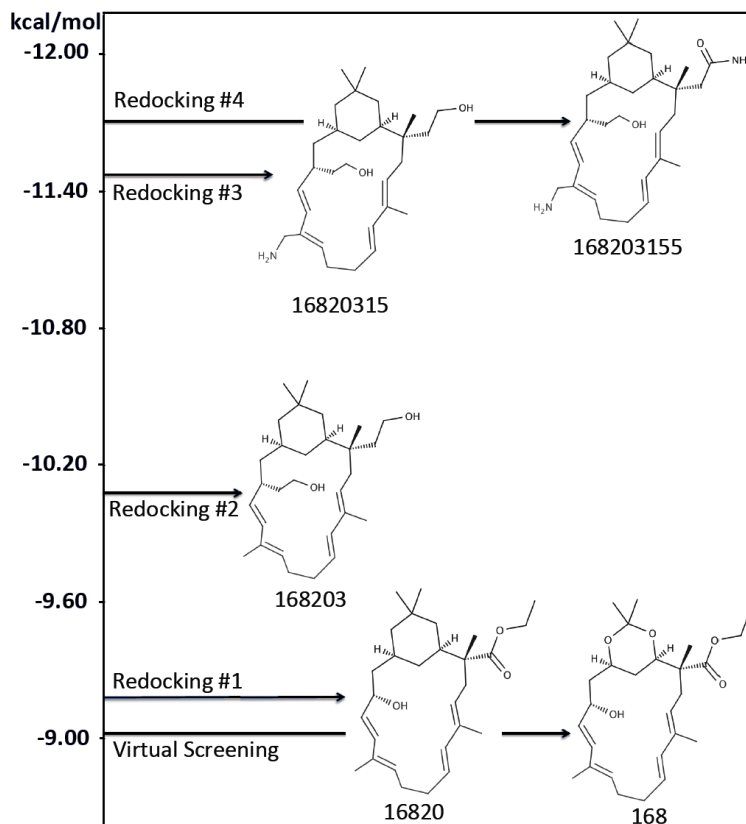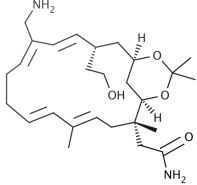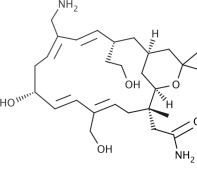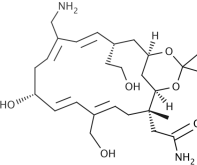
Figure 2.3: *Quantitative improvement of binding energies (on the y-axis) of the top hits obtained in each of the consecutive modification/redocking runs on molecule 168.*

best affinity and no further improvements were obtained. Fig. 2.3 shows the top hits from each run and shows the trend of the improvement of the binding affinity following our modification/redocking protocol. It is clear that this protocol managed to bring molecule 168 from a binding affinity of -9.03 kcal/mol to a binding affinity of -11.76 kcal/mol (molecule 168203155) which is much higher than the top hit obtained in the first virtual screening run (molecule 38 with a binding energy of -10.03 kcal/mol). However, these hits that we obtained appear to be highly lipophilic and their water solubility, and absorption could be questioned. These are serious drawbacks in the context of chemotherapy applications. Therefore, we decided to run another modification/redocking run on the best hit, molecule 168203155, with the purpose of only increasing the hydrophilicity of the derivatives, while still maintaining a good affinity.

Table 2.2 shows the result of this fifth run with the corresponding binding energies of each derivative. It shows that we could not get any derivative to be any better than the parent molecule, molecule 168203155, in terms of binding energy. However, their binding energies are still in the top range and we managed to increase their hydrophilicity by the inclusion of hydrophilic functional groups. It is worth mentioning that all the derivatives of molecule 168 are very rigid and there is only one mode of binding available for these molecules in the tubulin binding site. Within an RMSD of 2.00 Å, nearly all the binding poses would group in one cluster. This rigidity can be

Table 2.2: *Binding energies in the fifth modification/redocking run of the derivatives of molecule 168203155.*

Rank	ID	Molecule	ΔE [a]
1	1682031551		-11.47
2	1682031559		-11.22
3	1682031553		-11.06

[a] Binding energy calculated by AutoDock 4.2 in kcal/mol.

exploited in tailor designing analogs that are more selective toward certain tubulin isoforms over the others, thus increasing the specificity of the treatment.

Fig. 2.4(a) shows the binding of molecule 1682031551 to β-tubulin where the ligand forms five hydrogen bonds with Thr274, Arg282, Arg318, Asp26 and Ala231. Such a strong network of hydrogen bonds should help fixing the ligand in position and enhancing its affinity. This could be compared to the hydrogen bond network of epothilone A in β-tubulin, Fig. 2.4(b), where residues Thr274, Arg276, His227 and Arg282 form hydrogen bonds with the ligand as also pointed out by Nettles et al. [49]. Thr274 and Arg282 are residues in common for the two ligands, but Asp26, Arg318 and Ala231 are specific to our new molecule. This could foretell that the molecule might be effective in epothilone-resistant cell lines.

2.4.2 Results of Rescoring via MM/PBSA

Molecular dynamics simulations were run for the all complexes (five test molecules and three highest-scoring hits). The five test molecules, discodermolide, eleutherobin, epothilone A, paclitaxel and sarcodictyin A, were used as a test for the predictive power of the method. Before using the MM/PBSA script, the five complexes were checked for equilibration of total energy, temperature, density and RMSD and we made sure that they were well-equilibrated. Fig. 2.5a shows the RMSD equilibration for the five complexes, each of which was equilibrated for at least the last 5 ns of the simulation which was then post-processed using MM/PBSA for binding energy calcula-

(a) *molecule 1682031551*

(b) *Epothilone*
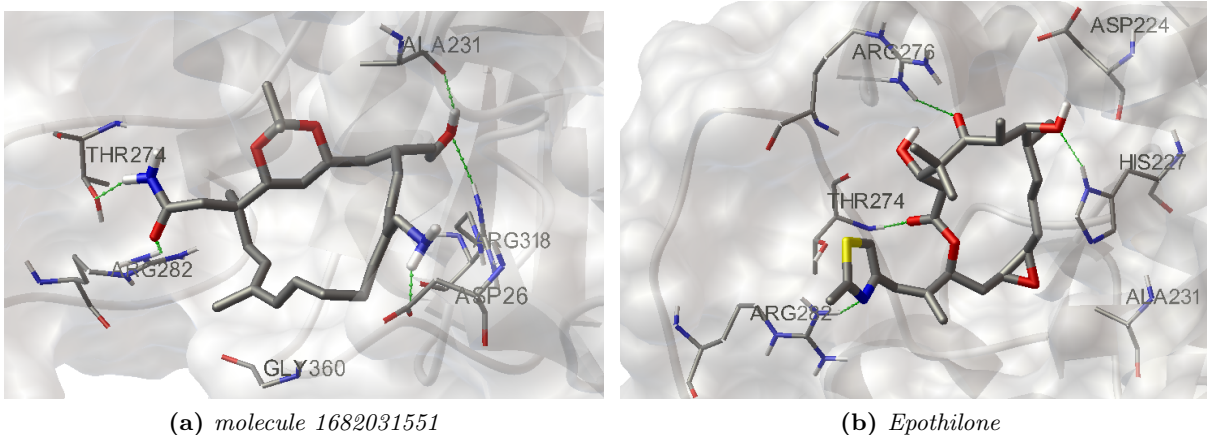
Figure 2.4: *Ligand poses in taxol binding site.*

tion. Enthalpic contribution was calculated using the last 500 evenly spaced snapshots extracted from the last 5 ns. This ensures that each two consecutive snapshots are 10 ps apart and hence uncorrelated. The entropic contribution was calculated using normal mode analysis with 48 evenly spaced snapshots also extracted from the last 5 ns. We could not use more snapshots because normal mode analysis is highly demanding in terms of computational power. The experimental binding constants, K_b, of the five test molecules to the taxol binding site were determined by Buey et al. through displacement of the fluorescent taxoid Flutax-2 [32,65]. These values are presented in Table 2.3, together with the binding free energies, ΔG^o, calculated from the binding constants via the following thermodynamic relationship:

$$\Delta G^o = -RT \ln K_b \tag{2.2}$$

where R is the universal gas constant, $R = 1.987$ cal mol^{-1}K^1, and T is the absolute temperature. The experimentally derived binding energy, ΔG^o, for each molecule was plotted against the MM/PBSA predicted value as shown in Fig. 2.6. The figure shows the linear fit of the data which produces an excellent linear correlation with a coefficient of determination, R^2, of 0.99, a slope of 0.2 and an intercept of -4.6 kcal/mol. This proves that the MM/PBSA protocol that we followed here is a reliable method in predicting the binding energies of the tubulin ligands binding at the taxol binding site and hence can be carried forward and used to predict the binding energies of the novel hits that we obtained from the screening protocol. Thus, the same MM/PBSA protocol was applied for the three highest-scoring hits shown previously in Table 2.2 , namely molecules 1682031551, 1682031559 and 1682031553. As was done previously with the test molecules, we checked the complexes of these three ligands for equilibration of the total energy, temperature, density and RMSD and made sure that they were all well-equilibrated.

Fig. 2.5b shows the equilibration of RMSD of the three complexes and confirms that the three complexes were equilibrated at least in the last 5 ns. We then post-processed these 5 ns in the same way as the previous complexes and used 500 evenly spaced snapshots for enthalpy calculation and 48 evenly-spaced ones for normal mode analysis. The results of this prediction, after extrapolation
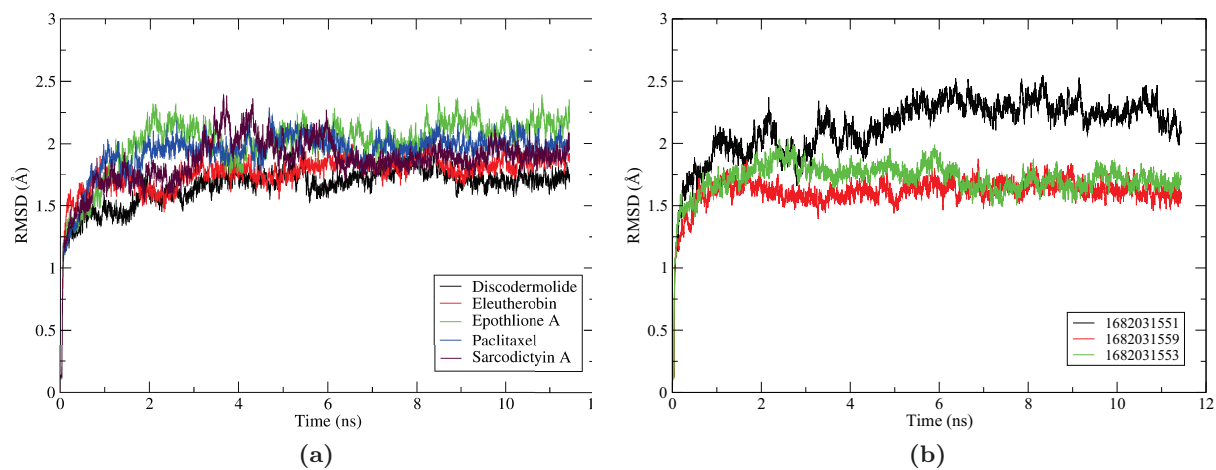
Figure 2.5: *RMSD equilibration of the ligand-receptor complexes of the five ligands with experimental data (a) and of the five model kits (b).*
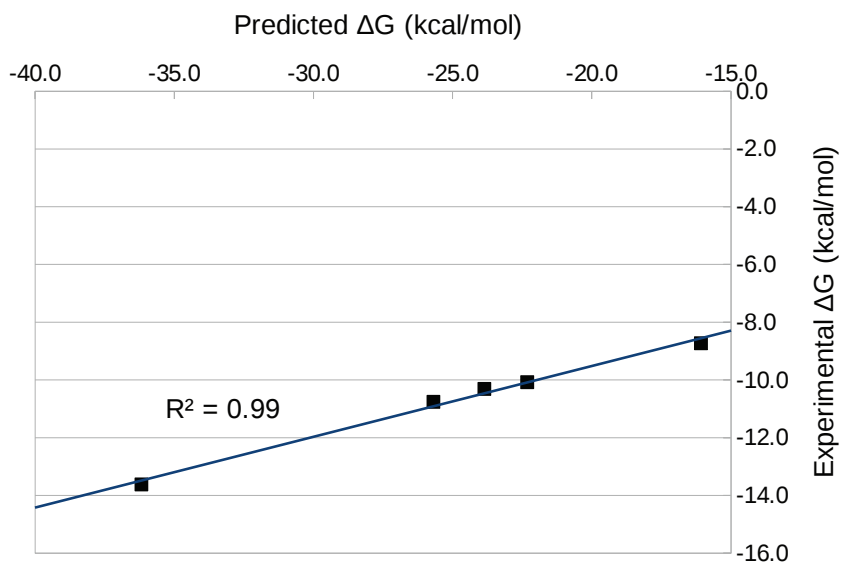
Figure 2.6: *linear regression plot of the MM/PBSA-predicted binding energies versus the experimentally derived values.*

Table 2.3: *Experimental binding constants, K_b, of the five test molecules and the corresponding values of the binding free energies, ΔG^o.*

Ligand	K_b $(10^7 M^{-1})^a$	ΔG^o (kcal/mol)^a
Discodermolide	837 ± 77	-13.6 ± 0.1
Eleutherobin	3.26 ± 0.25	-10.3 ± 0.05
Epothilone A	6.94 ± 1.08	-10.8 ± 0.1
Paclitaxel	2.19 ± 0.05	-10.1 ± 0.01
Sarcodictyin A	0.23 ± 0.09	-8.7 ± 0.2

^{a} Data $\pm$ standard error; average of at least four independent measurements [32, 65].

through the linear fit presented Fig. 2.6, are shown in Table 2.4. The results in the table clearly confirm the outcome of the docking experiments and predict that the novel hits may have strong binding energies to the taxol binding site. Molecule 1682031551 shows the lowest predicted binding free energy, -14.7 kcal/mol, which is much lower than all of the test molecules shown in Table 2.3, all of them being strong tubulin stabilizers. Molecule 1682031553 is also predicted to have a strong affinity of -12.5 kcal/mol toward the taxol binding site which surpasses all of the test molecules except for discodermolide. Finally, molecule 1682031559 is predicted to have an affinity of -10.8 kcal/mol which is still better than paclitaxel, eleutherobin and Sarcodictyin A. In general, these results meet the aim of the study, confirm our findings in the docking runs and show that our novel scaffold could act as a good anticancer tubulin-binding agent.

Table 2.4: *Extrapolated binding energies of novel hits as predicted by the MM/PBSA protocol.*

Ligand	ΔG (kcal/mol)^a
1682031551	-14.7 ± 0.9
1682031559	-10.8 ± 0.6
1682031553	-12.5 ± 0.7

^{a} Extrapolated value $\pm$ standard error based on a 95% confidence interval.

2.4.3   Results of Physicochemical Predictions

After predicting that our new molecules should have high affinity toward the taxol binding site, we checked the physicochemical and toxicological properties of them and compared them to known MSAs to see how our now candidates will compare in terms of drug pharmacokinetics. Table 2.5 shows the different parameters calculated by ADMET Predictor software. The partition coefficient ($\log P$) is a measure of lipophilicity. It shows that our molecules span a wide range of lipophilicity from 5.33 to 0.94. This is useful since lipophilicity is required for some drugs acting on the brain so that they can pass through the blood-brain barrier which may help with brain tumors. However, too lipophilic compounds might not be soluble in the intestinal fluids and less lipophilic compounds are preferred in this case. The distribution coefficient ($\log D$) is different from $\log P$ for all our four new

Table 2.5: *Predicted physicochemical properties of the novel hits compared to epothilone A*

Ligand	log P	log D	P_{eff} (μm/s)	FaSSIF (mg/mL)
168203155[a]	5.33	3.52	1.35	0.0615
1682031551	3.21	1.55	0.47	0.501
1682031559	1.81	0.35	0.28	1.89
1682031553	0.94	-0.45	0.14	3.99
Epothilone A	3.07	3.07	1.46	0.0433

[a] The parent compound of the fifth run. It is listed here to show how the fifth run affected the lipophilicity of the compounds. See Table 2.2.

compounds which shows that they are all ionizable. The effective human jejunal permeability (P_{eff}) decreases with decreasing lipophilicity and spans a wide range from 1.35 to 0.14 μm/s. Interestingly, the solubility in simulated intestinal fluid in fasted state (FaSSIF) is much higher for our new compounds when compared to epothilone A. The solubility of compound 1682031553, for example, is almost 100-fold better than epothilone A, our positive control. Achieving molecules with better solubility was one of the main aims of the study since paclitaxel, for example, is insoluble in water and this was one of its drawbacks [38]. Therefore, the physicochemical profiles of our molecules do achieve one of the aims of this study, that is; finding new tubulin binding candidates with better pharmacokinetic profiles. As to the predicted toxicity of these compounds, all of them showed potential hepatotoxicity, with 1682031559 and 1682031553 showing potential SGOT and SGPT enzyme elevation. Only compound 168203155 showed potential acute rat toxicity. We compared this profile to the predicted profile of our positive control, epothilone A, which showed potential hepatotoxicity and carcinogenicity in rats. Therefore, the toxicological profiles of our molecules and epothilone A are not identical and when carcinogenicity is considered, our molecules may do better than epothilone A as they are predicted to be devoid of this toxic side effect. However, all these are data predicted from simulations, and toxicological assays will indeed be needed to confirm such findings.

2.5 Conclusion

The similarity-based virtual screening study reported here predicted five active molecules that have not been patented as tubulin-binding agents before. One of them, molecule 168, is very promising and possesses a novel scaffold. Upon modifying this molecule, optimizing its binding to the receptor and redocking its derivatives in subsequent runs, we managed to obtain several derivatives that are superior to the parent compound, molecule 168, in terms of calculated binding affinity to the target and physicochemical properties. They also showed superiority in binding energy over known tubulin binding agents. Rescoring of the top hits using the MM/PBSA protocol confirmed our findings and showed that at least three of our top hits possess higher binding affinities than the five tubulin stabilizers that were used as query molecules in the similarity-based filtration step. The

predicted physicochemical properties of the new compounds also show superiority to the known tubulin stabilizers especially regarding solubility in the intestinal fluid. The toxicological profile predicts comparable toxicity to known MSAs except for carcinogenicity which our new molecules were devoid of. As a follow-up to this study, the next step should be synthesizing these molecules and testing their activities experimentally to confirm the computational findings. It is hope that this work will lead to the development of new and improved cancer chemotherapy compounds.

Chapter 3

Unravelling the Mechanism of Action of Antitumor Lankacidin

3.1 Summary

Lankacidin group antibiotics are known to have antitumor activities with unclear mechanism of action. Based on a previous virtual screening study for taxol-like drugs which produced hits very close in structure to lankacidin, we suggest that the cytotoxic effect of lankacidin antibiotics is due to microtubule stabilization though a taxol-like mechanism. This suggestion was confirmed both computationally using molecular dynamics simulations and binding energy calculations and experimentally using fluorescence quenching assays.

3.2 Introduction

Lankacidin-group antibiotics (T-2636) are fermentation products that are produced by the organism *Streptomyces rochei* [66]. First isolated in 1969, the antibiotic group was later fully characterized and the structure of lankacidin was elucidated [67–70]. The parent of this group, lankacidin C Fig. 3.1a, is a 17-membered macrocyclic tetraene which shows strong antimicrobial activity against various Gram-positive bacteria, including many strains that are resistant to macrolide antibiotics, and is used in veterinary medicine [71].

Besides its antimicrobial activity, lankacidin C and some of its derivatives also displayed considerable *in vivo* antitumor activity against L1210 leukemia, B16 melanoma and 6C3 HED/OG lymphosarcoma [72, 73]. In addition to these pharmacological activities, being well-tolerated *in vivo*, displaying low toxicities in mice [69] and having a complete synthetic pathway available [74] makes the molecule very attractive for drug research.

The antimicrobial mechanism of action of lankacidin has been attributed to interference with peptide bond formation during protein synthesis by binding at the peptidyl transferase center of the eubacterial large ribosomal subunit [75, 76]. It is unclear, however, whether the antitumor activity of lankacidins is related to their interference with protein synthesis or not [73]. Hence, the mechanism of antitumor activity of lankacidins is yet to be elucidated.
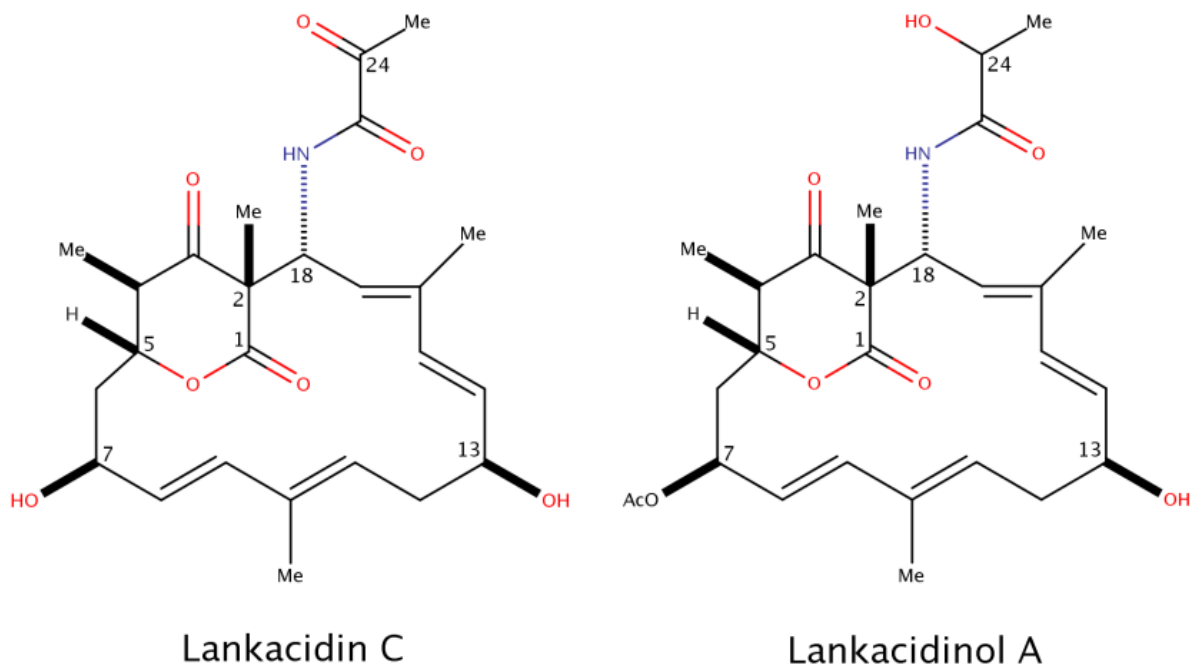
Figure 3.1: *Structure of lankacidin C and lankacidinol A*

In a previous virtual screening study for taxol-like microtubule stabilizers, we have identified a novel scaffold predicted to bind to tubulin at the taxol binding site [77]. Due to difficulties synthesizing this novel compound, we tried to search for available molecular structures that resemble it. It turned out that the molecular framework of our novel hit very closely resembles that of lankacidin [77]. The striking similarity between the two compounds suggests a hypothesis that lankacidin C could also bind to tubulin at the taxol binding site which could explain its unknown antitumor mechanism of action. In addition to explaining their mechanism of action, if this hypothesis is tested and proven true, this would widen our understanding of microtubule stabilization by taxol-like agents and would also open the door for structural modifications of the low-toxicity synthetically-available lankacidins towards better binding with tubulin. This may help provide a promising cancer treatment. On the computational side, if the hypothesis is proven, this would provide a good example of the power of computational predictions and their aid in explaining chemical and pharmacological processes.

3.3 Methodology

3.3.1 Computational Simulations

Computational simulations of lankacidin C bound to the taxol binding site in β-tubulin were carried out according to the procedures outlined in previous work [77]. Briefly, we parameterized GDP-bound β-tubulin subunit using AMBERff99SB forcefield for the protein [48] and Meagher et al. parameter set for GDP molecule [78]. Lanakcidin C was parameterized using the GAFF force

field [50]. Prior to molecular dynamics simulations, the crystal structure of lankacidin C from Protein Data Bank (PDB) ID:3JQ4 was docked to the taxol binding site using AutoDock 4.2 as explained in [77]. The complex was then neutralized, solvated, heated to 300K, and taken through a molecular dynamics simulation for density equilibration followed by a production phase for nearly 11 ns using the Amber package [47]. Evenly-spaced snapshots from the last 5 ns of the simulation were used for an MM/PBSA binding energy calculation in which both the enthalpy (500 snapshots) and entropy (48 snapshots) were estimated, the latter using normal mode analysis. Using the linear fitting that was developed in [77], we extrapolated the MM/PBSA-calculated binding energy to get a more realistic value that could be compared to binding energies of available microtubule stabilizers.

3.3.2 Isolation of Lankacidin

Streptomyces rochei strain 51252 [79] was cultured in YM liquid medium (0.4% yeast extract, 1.0% malt extract, and 0.4% D-glucose, pH 7.3) at 28°C for 2 days. The culture broth was extracted twice with equal volume of ethyl acetate. The combined organic phase was dried (Na_2SO_4), filtered, and concentrated to dryness. The crude residue was purified by Sephadex LH20 (GE Healthcare) with methanol. The fractions containing lankacidin C and lankacidinol A were combined, and further purified by silica gel chromatography with a mixture of chloroform-methanol (50:1-20:1). All the spectral data for lankacidin C and lankacidinol A were identical to the reported data [80] .

3.3.3 Fluorescence Quenching Assays

Reagents were purchased from Sigma-Aldrich Canada Ltd. (Oakville, Ontario, Canada) and Fisher Scientific Company (Ottawa, Ontario, Canada). In a 96-well microplate, equimolar mixtures of recombinant human tubulin TUB-BI or porcine cytoskeleton tubulin protein (purchased from Cytoskeleton Inc; Cat. # T240-DX) and the buffer (10 mM sodium phosphate, 10 mM $MgCl_2$, 1 mM guanosine 5'-tri-phosphate (GTP), 0.5% DMSO, 250 mM sucrose, pH = 7.0) were mixed to reach a final tubulin (dimer) concentration of 2 μM. GTP was added to the samples to a final concentration of 1 mM. The microplate was incubated on ice for 10 min to allow for the formation of the tubulin dimer. The calculated amounts of stock solution of the compounds in DMSO were added to the protein samples to obtain final ligand concentrations of 5, 10, 20, 40, 60, 80 and 100 μM. The control was ligand-free, and the total sample volume was 100 μL. A glass bead was inserted into each well, and the microplate was covered with a protective film, sealed with a lid, and incubated for 30 min at 25°C. After that time, the microplate was transferred to a rotating platform and vigorously rotated for 1 h at room temperature. From each well, 80 μL of samples and control was transferred to a 1-cm fluorescence cell. Fluorescence spectra were collected on a PTI MODEL-MP1 spectrofluorometer using 10 mm path length cell at 295 nm (excitation wavelength), and the scan range was 310–400 nm. Spectral data were collected using fluorescence software, and data analysis was performed using ORIGIN 6.1 software (OriginLab, Northampton, MA).

Isolation of lankacidin was carried out by K. Arakawa, Hiroshima University, Japan.

3.3.4 Determination of Kinetic Parameters

Data from the fluorescence experiments were used to determine the apparent binding constant of ligands to tubulin dimers by using the Stern–Volmer equation:

$$\frac{F_0 - F}{F} = K_a[L] \tag{3.1}$$

here F_0 and F are the fluorescence intensities in the absence and in the presence of quencher, respectively, K_a is the formation constant of the donor–acceptor (quencher–fluorogen) complex, and $[L]$ is the concentration of the ligand added. Excitation and emission slits were set at 4 nm. All spectra were collected with samples having final optical densities (1 cm) < 0.3 at maximum absorbance of added ligand and were corrected for the inner filter effect according to the following equation:

$$F_{\text{corr}} = 10^{(A_{ex}+A_{em})/2} F_{\text{obs}} \tag{3.2}$$

where F_{corr} is the corrected fluorescence, F_{obs} is the measured fluorescence, A_{ex} is the absorption value at the excitation wavelength (295 nm), and A_{em} is the absorption value at the emission wavelength (336 nm). From the slope of the linear plot of $(F_0 - F/F)$ versus $[L]$, binding constant values were estimated. The results are expressed as mean values $\pm$ standard deviation (SD) (n=4).

3.4 Results

3.4.1 Computational Predictions

The equilibrated trajectory of the complex of lankacidin C and β-tubulin was post-processed by removal of water and ions from the last 5 ns and a MM/PBSA and normal mode analysis were run. The MM/PBSA returned a binding enthalpy of -34.5 kcal/mol while the normal mode analysis returned an entropic contribution ($T\Delta S$) of -23.4 kcal/mol. Both values yield a calculated binding free energy of -11.1 kcal/mol. Utilizing the linear fitting that was developed in previous work [77], this value can be extrapolated using the following relationship

$$\Delta G_{\text{predicted}} = 0.2 \, \Delta G_{\text{calculated}} - 4.7 \tag{3.3}$$

to give the predicted binding energy of lankacidin C to the taxol binding site. The equation yields a predicted binding energy of -7.4 ± 0.9 kcal/mol. When this value is compared to known taxol-like microtubule stabilizers upon which the linear fit was built, it is closest to sarcodityin A which has a predicted binding energy of -8.55 ± 0.7 kcal/mol [77]. Since it is established in previous studies that the antitumor activity of lankacidin C is relatively weak [72, 73], the comparison of our predicted binding energy of lankacidin C to the taxol binding site versus that of sarcodictyin A seems plausible. Based on our predicted binding free energy, we predict a dissociation constant of

---

Fluorescence quenching experiments were carried out by R. Ahmed, Department of Biochemistry, University of Alberta.
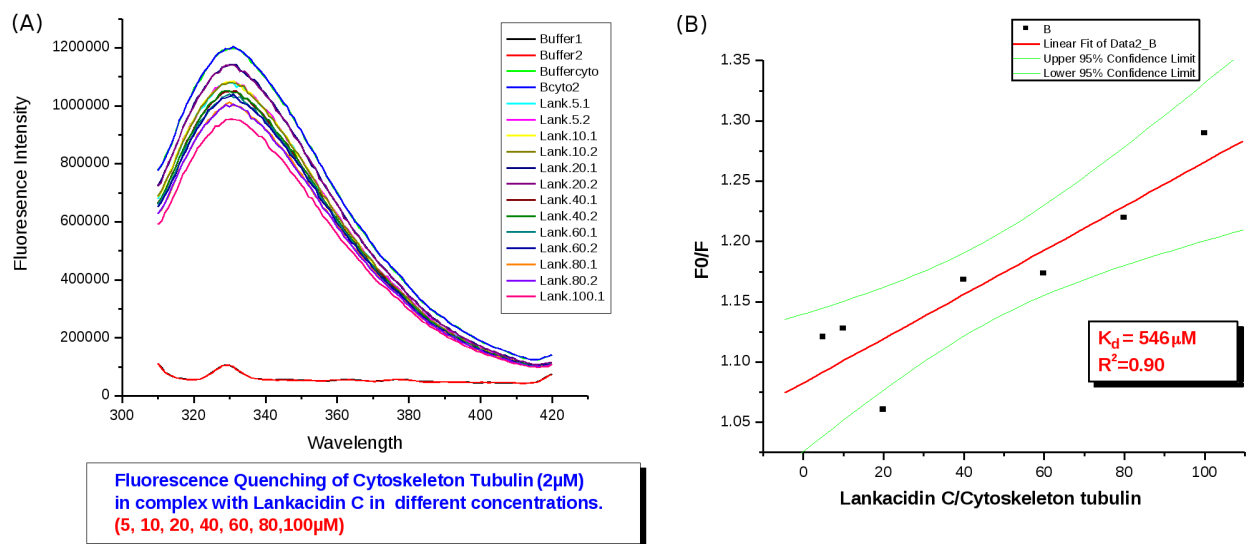
Figure 3.2: *Effect of Lankacidin C on fluorescence of the porcine cytoskeleton tubulin using fluorescence quenching assays.*

lankacidin C from the taxol binding site of $4.6\,\mu$M, as compared to a value of $0.6\,\mu$M for sarcodictyin A. Therefore, our computational predictions support our hypothesis that the antitumor activity of lankacidin antibiotics could be due to binding at the taxol binding site which probably incurs stabilization of microtubules.

3.4.2 Fluorescence Quenching Assays

Lankacidins show binding affinities when tested with fluorescence quenching assays towards porcine cytoskeleton tubulin with K_d value of $546\,\mu$M and 1.1 mM for lankacidin C and lankacidinol A, respectively. The binding induces a conformational change in tubulin as illustrated in Fig. 3.2 and Fig. 3.3.

We also tested whether lankacidins directly bind to the recombinant purified TUB-BI dimers, which would help us decide as to which tubulin subunit lankacidins bind. Interestingly, we found that both lankacidins have binding affinities towards recombinant purified TUB-BI. The characteristic tubulin fluorescence emission spectrum was significantly quenched by lankacidinol A with a K_d value of $654\,\mu$M versus 1.06 mM for lankacidin C. Both are affecting the TUB-B1 conformation in a concentration-dependent manner (see Fig. 3.4 and Fig. 3.5). These findings indicate that lankacidins have moderate effects causing conformational changes in tubulin upon binding. The findings also indicate that lankacidins bind to the β-tubulin subunit, which narrows down the search for proving our hypothesis.

3.5 Discussion and Conclusions

The present work shows an example of how computational predictions aid in the understanding of mechanisms of drug actions. Based on similarity to a previous hit that was predicted to bind
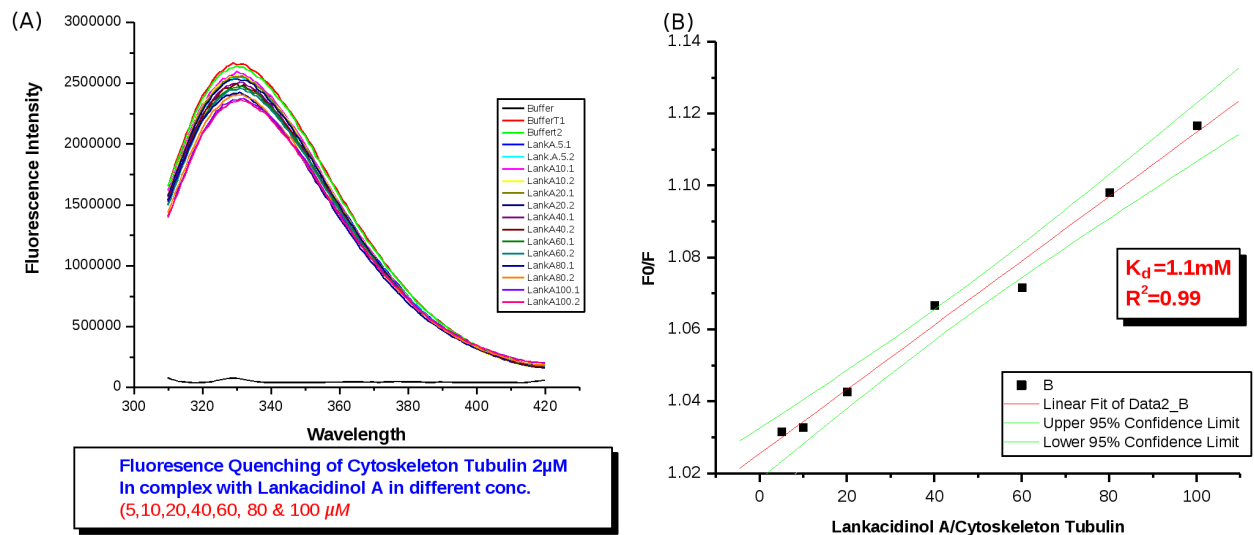
Figure 3.3: *Effect of lankacidinol A on fluorescence of the porcine cytoskeleton tubulin using fluorescence quenching assays.*
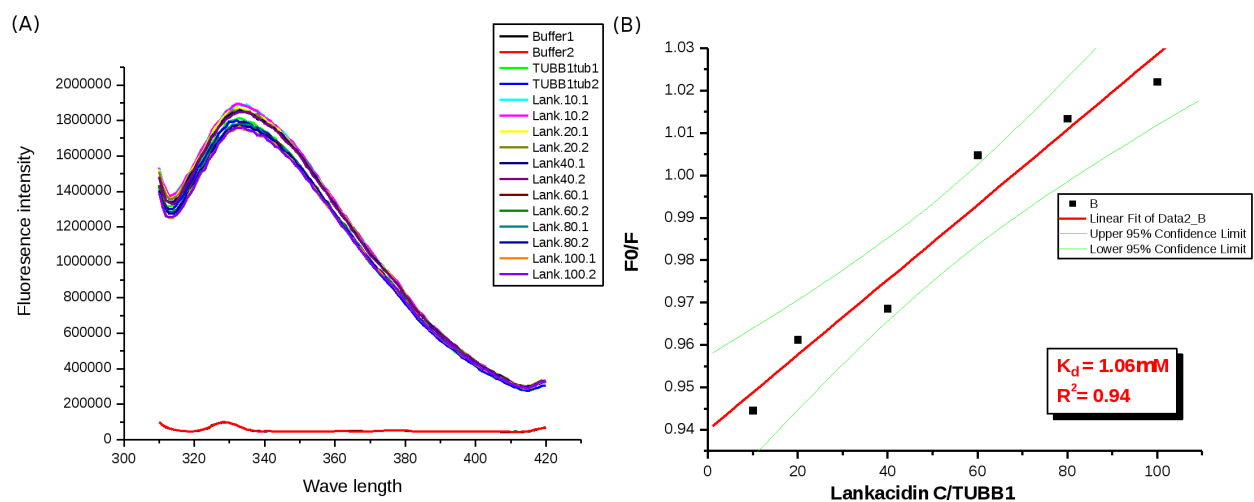
Figure 3.4: *Effect of lankacidin C on fluorescence of purified recombinant TUB-BI using fluorescence quenching assays.*
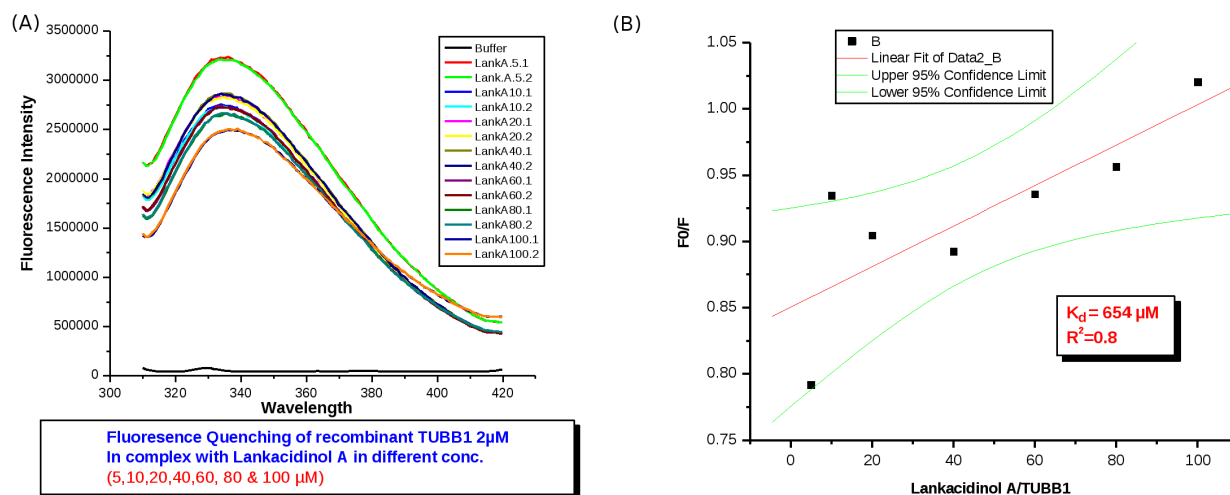
Figure 3.5: *Effect of lankacidinol A on fluorescence of purified recombinant TUB-BI using fluorescence quenching assays.*

to the tubulin taxol binding site, lankacidins were also predicted to bind to the same site, which was suggested to be the mechanism of their antitumor activity. Computational estimation of lankacidin C binding energy to the taxol binding site returned values that reflect relatively weak affinity, comparable to that of sarcodictyin A, which agrees with its relatively weak antitumor activity. Fluorescence quenching experiments confirmed the computational predictions and indicated that lankacidins bind to tubulin and induce conformational changes in a concentration-dependent manner. When fluorescence quenching experiments were conducted using recombinant TUB-BI, a similar effect was noticed which narrows down our search to the β-tubulin subunit where the taxol binding site is, rather than the α-subunit.

Because this still does not prove the hypothesis, we need to do more experiments which should prove if lankacidin actually binds at the β-tubulin taxol binding site or not. Currently, we are carrying out a competitive binding assay in which the competition of lankacidins with fluorescent taxoid (Flutax-2) for the taxol binding site is assessed. The results of these experiments should shed light on whether our hypothesis regarding lankacidin antitumor mechanism of action is true or not.

Chapter 4

Estimating Hydrogen Bond Energies: Comparison of Methods

4.1 Summary

Hydrogen bonds are among the most important non-bonded interactions found in molecules. Different methods of estimating the strength of hydrogen bonds have been proposed to date. In this work, we present a comparison between methods of estimating hydrogen bond energies that are based on several electron density descriptors based on the quantum theory of atoms in molecules, the natural bond orbital theory, and Mulliken population analysis. The results indicate that the most powerful approach is based on the quantum theory of atoms in molecules, followed by the one employing the natural bond orbital theory. The Mulliken population analysis performed very poorly. The effect of including dispersion correction was also studied. Parameters for predicting hydrogen bond energies are presented.

4.2 Introduction

The hydrogen bond is an important type of non-bonded interactions, which comprises a strong electromagnetic dipole-dipole interaction between polar molecules. The interaction takes place between a hydrogen bond donor, which is a hydrogen atom attached to a highly electronegative atom, and a hydrogen bond acceptor, which is another highly electronegative atom. Electronegative atoms involved in hydrogen bonding are usually oxygen, nitrogen, and fluorine [81–83]. Hydrogen bonds are also important in forming the secondary and tertiary structures of proteins and the stability between subunits in multimeric proteins [84–86]. Therefore, a facile way of estimating the strength of hydrogen bonds could be very useful. The electron density distribution around the atoms involved in the formation of a hydrogen bond, and other bonds like covalent bonds, have

been shown to correlate with the bond strength [87–89], and be a strong predictor of hydrogen bond strength [90–98].

There are several methods of characterizing the electron density distribution in hydrogen-bonded systems. Among these methods are the quantum theory of atoms in molecules (QTAIM) [18, 99], the natural bond orbital population analysis [100], and the Mulliken population analysis [101]. In QTAIM, the electron density at the bond critical point (ρ) is a useful measure of bond strength. A bond critical point is a point along the bond coordinate that has a zero gradient of the electron density distribution [18, 99]. This quantity, ρ, has been shown to correlate with bond strengths [94–98, 102, 103]. In natural bond orbital population analysis [100], there are two quantities that are similar in concept and are also expected to correlate with the strength of the bonds. These two quantities are Wiberg bond index (WBI) [104] and the overlap-weighted bond order (OWBO) [105, 106] between the hydrogen atom and the hydrogen bond acceptor atom, both calculated in natural atomic orbital basis [100]. Finally, in the Mulliken population analysis [101], the Mulliken overlap charge (MOC) between the interacting atoms was also shown to correlate with bond strengths [89]. In the present study, all these quantities, namely ρ, WBI, OWBO, and MOC, were assessed for relationship with strength of hydrogen bonds to establish which one leads to a better correlation. Dispersion correction and basis set effects were also characterized.

Two quantum chemical approaches are available to calculate the strength of a hydrogen bond between two molecules, the supermolecular approach and intermolecular perturbation theory [107, 108]. In the supermolecular approach, the one adopted here, the interaction energy between two molecules A and B, is calculated as

$$E_I\left(\mathbf{r_A}, \mathbf{r_B}\right) = E_{AB}\left(\mathbf{r_A}, \mathbf{r_B}\right) - E_A\left(\mathbf{r_A}\right) - E_B\left(\mathbf{r_B}\right) \tag{4.1}$$

where $\mathbf{r_A}$ and $\mathbf{r_B}$ denote the coordinates of the atoms in molecules A and B, respectively. E_{AB} is the energy of the complex while E_A and E_B are the energies of molecules A and B, respectively. Such direct evaluation of the hydrogen bond interaction energy is not always possible: either the interacting systems are too large or more than just hydrogen bonding interactions are present. It is possible to get around these obstacles by designing an interpolation formula that depends only on the electron density between the atoms involved in a hydrogen bond, that is, the hydrogen atom and the hydrogen bond acceptor atom. Parameters in such a fitting formula could be derived using known hydrogen bond energies in systems that interact only via a single hydrogen bond. This approach was followed by Thar and Kirchner who used a training set of complexes satisfying these criteria [91]. They used these molecules in designing an interpolation formula based on the two-center shared electron number in order to detect hydrogen bonds [91]. We used a subset of their training set in drawing our correlation.

4.3 Methodology

All the isolated hydrogen-bonded complexes, 45 in total, were optimized using the density functional theory with the B3LYP functional [109–111]. A primary basis set used in this work was the triple-

zeta valence polarized basis set (TZVP) [112, 113]. Since hydrogen bond interaction energies E_I were calculated in a supermolecular approach, all the energies of the isolated molecules needed to be counterpoise corrected for the basis set superposition error [114, 115]. We employed the counterpoise correction in the calculation of the interaction energy from the optimized structure but not during the optimization process itself. The optimized structure was then used for QTAIM analysis, natural bond orbital population analysis, and Mulliken population analysis to extract the values of the descriptors ρ, WBI, OWBO, and MOC introduced before. The hydrogen bond energies in all the complexes were plotted against each of these four descriptors and the correlation was assessed. Case studies of a test set were also performed to investigate the ability of each descriptor to predict the hydrogen bond energies. In order to assess a possible effect of dispersion correction, we studied the same 45 hydrogen-bonded complexes that were already minimized at B3LYP/TZVP level of theory, using several DFT functionals as well as the Møller-Plesset perturbation theory, namely B97D functional [116], wB97xD functional [117, 118], MP2 [119–121], and the previously used B3LYP. We also checked the effect of augmenting the basis set: diffuse functions from the aug-cc-pVDZ basis set [122, 123] were added to the TZVP basis set (the augmented TZVP basis set was denoted aug-TZVP). The software used in the calculations was Gaussian 09 [124]. Avogadro software was used for building and viewing the molecules [46].

4.4   Results and Discussion

4.4.1   Performance of Descriptors

The energies of the hydrogen bonds in the 45 hydrogen-bonded complexes as well as the values of the four hydrogen bond descriptors under study (ρ, WBI, OWBO, and MOC) were calculated. These data are included in Table S1 in the Appendix. The optimized coordinates of all complexes are available online in [125]. The results of the linear regression analysis are included in Fig. 4.1, and linear fitting parameters are included in Table 4.1. As Fig. 4.1a shows, the best correlation is obtained for the density at the bond critical point (ρ) evaluated in QTAIM, with a coefficient of determination (R^2) of 0.96, which indicates a very strong correlation. We should point out that we used a heterogeneous training set that contains different types of atoms as acceptors and donors including oxygen, nitrogen, fluorine, chlorine, bromine, sulfur, phosphorous, and carbon (see Appendix). Consequently, the descriptor ρ calculated in QTAIM analysis proves to strongly correlate with the hydrogen bond energy regardless of the types of the atoms involved in the hydrogen bond. As Table 4.1 shows, the correlation with the descriptor ρ has a slope (m) of -979 kJ/mol and an intercept (b) of 4.59 kJ/mol. The standard deviation is relatively very low (σ=2.69 kJ/mol). Therefore, these regression parameters could be used in predicting the hydrogen bond energies of more complex cases only by calculating the descriptor ρ, or any other descriptor, and substituting into the equation:

$$E_I = m\mathcal{D} + b \tag{4.2}$$

where m is the slope, $\mathcal{D}$ is the descriptor value, and b is the intercept.
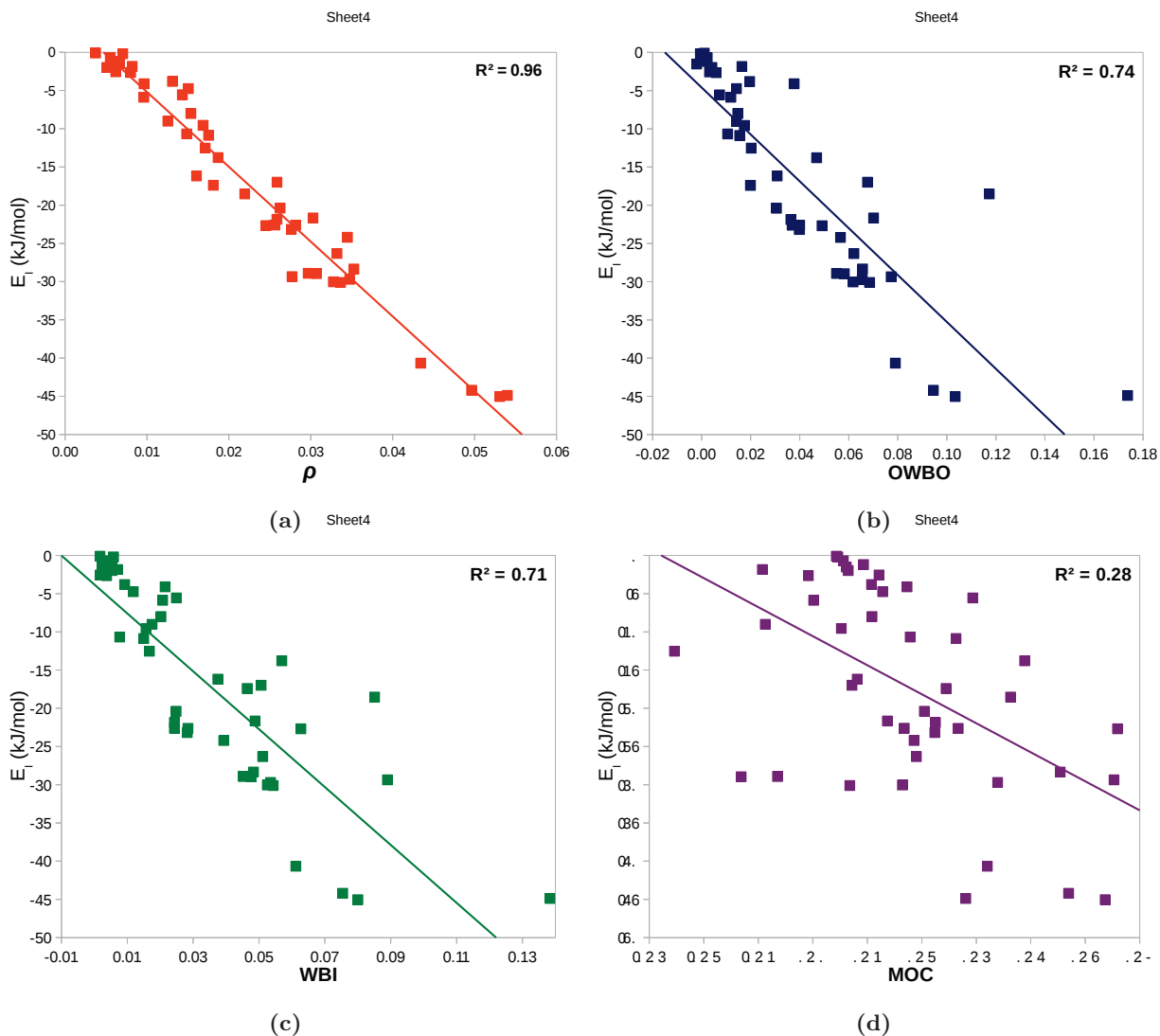
Figure 4.1: *Regression plots of the hydrogen bond energy E_I versus the four descriptors studied.*

Table 4.1: *Linear fitting parameters at B3LYP/TZVP level of theory for the plots of hydrogen bond energies versus different descriptors.*

Descriptor	R^2	m	b	σ
ρ	0.96	−979	4.59	2.69
OWBO	0.74	−307	−4.61	6.64
WBI	0.71	−379	−3.80	6.97
MOC	0.28	−380	−10.56	10.98

R^2 is the coefficient of determination. m, b, and σ are the slope, intercept, and standard deviation measured in kJ/mol, respectively.

The three remaining plots in Fig. 4.1 demonstrate that the three other approaches to estimating hydrogen bond energies exhibit a much lower degree of correlation than the one based on ρ. The overlap-weighted bond order (OWBO) from natural population analysis displays a correlation with an R^2 of 0.74 (Fig. 4.1b) similar to the Wiberg bond index (WBI) correlation with an R^2 value of 0.71. Thus, the natural population analysis still provides good descriptors for predicting hydrogen bond energies, although not as good as QTAIM. The fitting parameters of the linear regression for OWBO and WBI have very similar values. However, despite the lower quality of the correlation, the values of m and b derived for natural bond orbital descriptors may be useful in predicting hydrogen bond energies according to Eq.4.2. Fig. 4.1d shows that the Mulliken overlap charge (MOC) exhibits the poorest correlation with the hydrogen bond energies, with an R^2 value of only 0.28. Moreover, the standard deviation σ is very high, 10.98 kJ/mol. Therefore, we do not recommend using the descriptor MOC in predicting the hydrogen bond energies. It should be mentioned that MOC values are sometimes negative, devoid of a physical sense. This happens even for hydrogen bonds that are actually very strong (see Fig. 4.1d and Table S1). This behavior can also be seen in the case of OWBO, but only in fairly rare situations for very weak hydrogen bonds (see Fig. 4.1b and Table S1). However, the values of ρ and WBI are always positive. In particular, the values of ρ would either be positive, when a hydrogen bond was present, even weak ones, or they would be zero in the absence of hydrogen bonds. Therefore, we recommend that the QTAIM-based descriptor ρ be used for the best predictions of hydrogen bond energies. It should be mentioned that Parthasarathi et al. also established a strong correlation between ρ and hydrogen bond energy [97]. However, the training set they used was relatively small (28 hydrogen bonding systems), and it spanned a huge range of energies up to nearly -200 kJ/mol. We should point out that the inclusion of such very strong hydrogen bonds in the correlation analysis weakens its predictive power and increases the deviations, especially for moderate or weak hydrogen bonds. Since in this study we aimed at developing a reliable method for predicting the energies of biologically relevant hydrogen bonds, we used a larger training set and focused on hydrogen bonds with weak to moderate strength.

Inspection of Fig. 4.1b reveals that most of the data are clustered nicely in a straight line with a much better correlation. However, some of the data points deviate from that line. As shown in Fig. 4.2, all the hydrogen-bonded complexes that had a highly polarizable hydrogen bond donor, namely sulfur, phosphorous, bromine, and chlorine, appeared to deviate largely from the other data points. We carried out a separate regression analysis for this group. The results show that, within the applied level of theory, the hydrogen-bonded complexes with non-polarizable hydrogen bond donors have a strong correlation with OWBO with an R^2 of 0.93, while those with highly polarizable hydrogen bond donors have a correlation with a respectable R^2 value of 0.86 with hydrogen bond energy. Linear fitting parameters for the two cases are shown in Table 4.2. We tried to use the same approach with WBI but in that case the data points are scattered more randomly than in the OWBO plot. Similarly, the MOC correlation could not be improved since data points are very scattered.

To test the predictive power of these models, we carried out several case studies for more
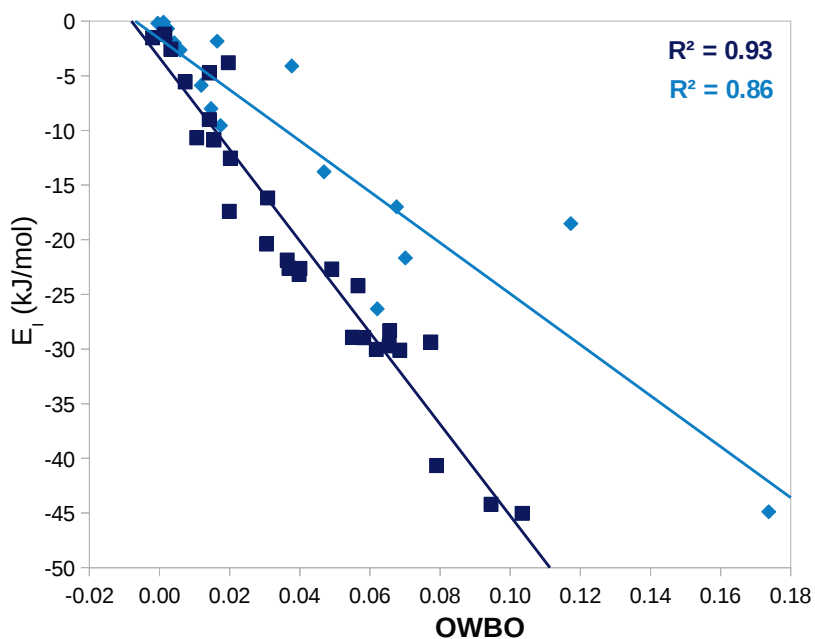
Figure 4.2: *Improvement of OWBO descriptor performance after subdivision of training set based on polarizability of the hydrogen bond donor.*

Table 4.2: *Linear fitting parameters of improved OWBO descriptor versus hydrogen bond energies with polarizable (P) and non-polarizable (NP) hydrogen bond donors.*

Descriptor	R^2	m	b	σ
OWBO (NP)	0.93	−419	−3.36	3.16
OWBO (P)	0.86	−233	−1.61	4.49

m, b, and σ are given in kJ/mol.

43

complex systems. We calculated the hydrogen bond energies using the supermolecular approach and compared them to the energies predicted via the four quantum mechanical descriptors studied here. We used the linear fitting parameters in Tables 4.1 and 4.2 together with Eq. 4.2. The results of the case studies are shown in Table 4.3. It should be noted that "OWBO$_{imp}$" refers to the energies predicted through the parameters of the improved OWBO fit as shown in Table 4.2. Comparison of the results shows that the most accurate descriptor is ρ from the QTAIM method. The results obtained with this descriptor are closest to the energies obtained via the supermolecular approach, and they lead to the lowest mean absolute error. The descriptors OWBO, WBI, and MOC perform poorly in predicting hydrogen bond energies. However, the descriptor MOC is not as bad as could be expected from its R^2 value in Fig. 4.1d. Its performance is comparable to that of OWBO and WBI. This behavior of the MOC descriptor becomes clear after comparing the plots for OWBO and WBI in Fig. 4.1. The two plots show that most of the data points are below the regression line, especially in OWBO. In OWBO, the points that are below the regression line belong to the systems that have non-polarizable hydrogen bond donors. However, the systems that have polarizable hydrogen bond donors are far above the regression line. This is manifested by most of the OWBO predictions being underestimated (see Table 4.3), since they involve non-polarizable hydrogen bond donors. Only the last two cases (HSH-Pyridine and HCl-Pyridine) are exceptions, which are predicted to be grossly overestimated since they involve polarizable hydrogen bond donors. It is, however, clear that the OWBO$_{imp}$ descriptor resolved this overestimation and made predictions very close to the E_I values. This underscores the importance of the distinction between polarizable and non-polarizable hydrogen bond donors in the OWBO$_{imp}$ descriptor in Table 4.2. This makes the OWBO$_{imp}$ descriptor the second most accurate one after ρ in the QTAIM method. It should however be made clear that all the descriptors may start to fail in systems that have hydrogen bond energies outside the training set range. This is clear through the large errors introduced in the predictions of the hydrogen bond energy in the HCl-Pyridine system, which is outside the training set energy range. A good application of this hydrogen bond energy prediction scheme was found in an unpublished study that we performed using these parameters. The QTAIM parameters of the descriptor ρ were used to assess the strength of hydrogen bonds that bring tubulin subunits together in microtubule cylinders. The study provided an important insight into the energetics of tubulin assembly [126]. The study is presented in Chapter 5.

4.4.2 Effect of Dispersion

We have only considered one level of theory (B3LYP density functional) and one basis set (TZVP) because we aimed at a simple yet reliable level of theory that is applicable for large systems, such as proteins [126]. However, it is prudent to compare this level of theory to models that include dispersion as well as compare the TZVP basis set with one that includes diffuse functions. Tables S2 to S5 in the Appendix show the values of the hydrogen bond energies and the descriptors for the 45 hydrogen-bonded complexes under the different levels of theory. Table 4.4 shows the linear fitting parameters for the different levels of theory for every descriptor. Comparison of data in

Table 4.3: *Case studies implementing the linear fit parameters for the different descriptors in predicting the hydrogen bond energies.*

Complex	E_I	ρ	WBI	MOC	OWBO	$OWBO_{imp}$
MeOH–MeF	–13.5	–12.8 (0.7)	–8.1 (5.4)	–17.2 (–3.7)	–10.9 (2.6)	–11.9 (1.6)
EtOH–EtOH	–22.2	–22.5 (–0.3)	–14.4 (7.8)	–20.2 (2.0)	–16.7 (5.5)	–19.8 (2.3)
HOH–Pyridine	–27.1	–25.0 (2.1)	–18.9 (8.2)	–18.8 (8.3)	–19.7 (7.3)	–24.0 (3.1)
MeOH–Pyridine	–26.8	–25.6 (1.2)	–19.3 (7.5)	–19.7 (7.1)	–20.7 (6.2)	–25.3 (1.6)
HCOOH–HCOOH[a]	–37.1	–40.2 (–3.1)	–35.3 (1.9)	–34.2 (2.9)	–33.3 (3.8)	–42.5 (–5.4)
AcOH–AcOH[a]	–39.0	–41.9 (–3.0)	–36.4 (2.6)	–35.2 (3.7)	–34.5 (4.4)	–44.2 (–5.3)
Uracil–Uracil[a]	–22.0	–21.9 (0.1)	–17.4 (4.6)	–24.6 (–2.6)	–18.4 (3.6)	–22.2 (–0.2)
HSH-Pyridine	–14.9	–17.3 (–2.5)	–17.6 (–2.8)	–14.4 (0.5)	–20.8 (–5.9)	–13.9 (1.0)
HCl-Pyridine	–48.4	–60.6 (–12.2)	–64.2 (–15.8)	–29.7 (18.7)	–61.6 (–13.2)	–44.9 (3.5)
MAE [b]	NA	2.6	5.5	4.9	5.4	2.9

$OWBO_{imp}$ is the improved fit of the OWBO descriptor. Energies in kJ/mol, and errors in parenthesis.
[a] Energies per hydrogen bond are shown since these molecules have more than one hydrogen bond.
[b] Mean absolute error

Tables 4.1 and 4.4 reveals that regardless of the level of theory, the descriptors vary similarly: the linear fitting parameters do not change significantly by augmenting the basis set or changing the level of theory. Specifically, comparison of the B3LYP results in Tables 4.1 and 4.4 reveals the effect of larger basis set on the descriptors. It is interesting to note that the MOC results change the most, reflecting known dependence of the results from Mulliken population analysis on quality of the basis sets, in particular on the presence of diffuse functions.

These results support the conclusion that the use of the economical B3LYP/TZVP should be preferred over the other levels of theory, since B3LYP/TZVP is the fastest and simplest and yet has the highest R^2 value. However, when checking the absolute hydrogen bond energies (Tables S1-S5), we see that methods that include dispersion (such as B97D, wB97XD, and MP2) tend to bring about larger hydrogen bond energies especially for weak bonds. This is not surprising, since weak hydrogen bonds are dispersion-dominated while stronger ones are electrostatically dominated and the strongest ones are covalently dominated [83]. Consequently, one may prefer to use one of the methods that include dispersion, especially if the system of interest is small and weakly bound. Regarding the performance of different descriptors at different levels of theory, while ρ, OWBO, and WBI do not significantly differ, MOC does. At the B3LYP/TZVP level of theory (Table 4.1), MOC had a negative slope, while under all other methods in Table 4.4 the slope is actually positive. A positive slope means that most of the MOC values were negative, which is physically meaningless as MOC refers to the overlap population between two atoms.

4.5 Conclusion

We have assessed the performance of different quantum mechanical descriptors in predicting hydrogen bond energies. Three methods with different descriptors were assessed including quantum theory of atoms in molecules (QTAIM) with the descriptor ρ which represents the density and the

Table 4.4: *Linear fitting parameters at different levels of theory for the plots of the hydrogen bond energies versus different descriptors.*

Method	Descriptor	R^2	m	b	σ
B3LYP aug–TZVP	ρ	0.95	−978	4.77	2.8
	OWBO	0.67	−274	−5.29	7.4
	WBI	0.70	−374	−3.37	7.0
	MOC	0.25	92	−11.12	11.2
B97D aug–TZVP	ρ	0.91	−904	0.94	3.7
	OWBO	0.72	−271	−7.31	6.7
	WBI	0.74	−353	−5.68	6.4
	MOC	0.32	102	−12.95	10.3
wB97XD aug–TZVP	ρ	0.93	−1005	1.80	3.6
	OWBO	0.66	−297	−8.41	7.8
	WBI	0.71	−408	−6.45	7.2
	MOC	0.27	107	−14.67	11.4
MP2 aug–TZVP	ρ	0.93	−857	2.70	3.0
	OWBO	0.60	−245	−6.89	7.1
	WBI	0.67	−377	−4.84	6.5
	MOC	0.37	73	−8.86	9.0

R^2 is the coefficient of determination. m, b, and σ are the slope, intercept, and standard deviation measured in kJ/mol, respectively.

bond critical point, the natural bond orbital population analysis method with two descriptors, the Wiberg bond index (WBI) and the overlap-weighted bond order (OWBO), and the Mulliken population analysis method with the Mulliken overlap charge (MOC) as a descriptor. The descriptor q derived from the QTAIM method is the most accurate, followed by the improved OWBO descriptor. In contrast to the other methods, QTAIM was shown to be invariant to the type of atoms involved in the hydrogen bond. Therefore, we recommend the use of QTAIM and the descriptor ρ in the hydrogen bond energy calculations. Linear fitting parameters necessary to the prediction of hydrogen bond energies according to different descriptors have been presented. One limitation of the present approach is that very strong hydrogen bonds that are outside the training set energy range are more difficult to predict and the errors may start to grow larger as the interactions get stronger. The change in the basis set and the inclusion of dispersion correction in the quantum mechanical calculations did not significantly affect the performance of the descriptors, except MOC. MOC results are sensitive to the basis set used. Due to this behavior of the MOC descriptor, we do not recommend its use in hydrogen bond analysis. Absolute hydrogen bond energies are also sensitive to the inclusion of dispersion especially in weak hydrogen bonds.

Chapter 5

Analysis of the Strength of Interfacial Hydrogen Bonds between Tubulin Dimers Using Quantum Theory of Atoms in Molecules

5.1   Summary

Microtubules are key structural elements that, among numerous biological functions, maintain the cytoskeleton of the cell and have a major role in cell division, which makes them important cancer chemotherapy targets. Understanding the energy balance that brings tubulin dimers, the building blocks of microtubules, together to form a microtubule is especially important for revealing the mechanism of their dynamic instability. Several studies have been conducted to estimate various contributions to the free energy of microtubule formation. However, the hydrogen-bond contribution was not studied before as a separate component. In this work, we use concepts such as the quantum theory of atoms in molecules to estimate the per-residue strength of hydrogen bonds contributing to the overall stability that brings subunits together in pair of tubulin heterodimers, across both the longitudinal and lateral interfaces. Our study shows that hydrogen bonding plays a major role in the stability of tubulin systems. Several residues that are crucial to the binding of vinca alkaloids are shown to be strongly involved in longitudinal microtubule stabilization. This indicates a direct relation between the binding of these agents and the effect on the interfacial hydrogen-bonding network, and explains the mechanism of their action. Lateral contacts showed much higher stability than longitudinal ones (-462 ± 70 vs. -392 ± 59 kJ/mol), which suggests a dramatic lateral stabilization effect of the GTP cap in the b-subunit. The role of the M-loop in lateral stability in absence of taxol was shown to be minor. The B-lattice lateral hydrogen bonds

are shown to be comparable in strength to the A-lattice ones (-462 ± 70 vs. -472 ± 46 kJ/mol). These findings establish the importance of hydrogen bonds to the stability of tubulin systems.

5.2 Introduction

There are two different geometrical configurations of microtubules, as Fig. 5.1 shows, namely the A-lattice and B-lattice. In the A-lattice configuration, the α-tubulin subunits are lying almost beside the β-subunits in neighboring protofilaments, producing a continuous pattern of alternating α- and β-subunits. In B-lattice, the α-subunits are lying almost beside the α-subunits in neighboring protofilaments (and α beside α). Having 13 protofilaments in a cylinder, the B-lattice would always include a discontinuous seam, one lateral domain where adjacent dimers are in the A-configuration [127, 128]. There is evidence showing that the B-lattice is the dominant form both in vitro and in vivo [129–133]. However, the universality of the B-lattice was revisited because the formation of microtubules in vitro in the presence of End- Binding Protein 1 showed that A-lattice contacts are more favorable under these circumstances [127, 134, 135]. Because End-Binding Protein 1 is present in cells during polymerization of microtubules, the same effect is also expected in vivo [127, 134, 135]. In a 2003 study that considered the contribution of the solvation energy in terms of solvent-accessible surface area energy as well as the Poisson-Boltzmann electrostatic energy, Sept et al. [136] showed computationally that the B-lattice configuration is slightly more stable than the A-lattice one, providing an explanation for the B-lattice predominance. Along the same lines, Drabik et al. [137] calculated the potential of mean force between lateral interfaces of tubulin dimers in a microtubule and arrived at the same conclusion, i.e., that the B-lattice is more stable than the A-lattice configuration. Erickson and Pantaloni [138] also calculated, in 1981, the entropic contribution to the total energy profile.

The motivation for our work stems from noticing that the studies regarding tubulin interfacial energetics have so far not considered the hydrogen-bond energy contribution as a separate component. There is no doubt that hydrogen bonds play a very important role in protein energetics, especially in the stability between subunits in multimeric proteins [85, 86]. Therefore, studying the effect of hydrogen bonds on the energetics of interfacial interactions between tubulin heterodimers is essential for understanding the proper thermodynamics and kinetics of assembly. As shown in Fig. 5.1, we studied different interfaces of tubulin-tubulin interactions. Regarding the B-lattice, we studied the lateral interface between two tubulin dimers in two adjacent protofilaments and we called it the "LatB" interface. Regarding the A-lattice, we studied the equivalent lateral interface, calling it the "LatA". We assumed that the longitudinal interactions between tubulin heterodimers in the same protofilament are identical between the B-lattice and A-lattice cases, as the geometry of the protofilament is not expected to be affected by lateral contacts, at least over the simulation time range. Therefore, we called them both the "LongAB" interface. The three different interfaces were studied and the total as well as per-residue hydrogen-bond energies were calculated. The calculations were performed using molecular dynamics (MD) and quantum mechanics (QM) calculations followed by electron density analysis using Bader's theory of atoms in molecules (AIM) [18, 99] in
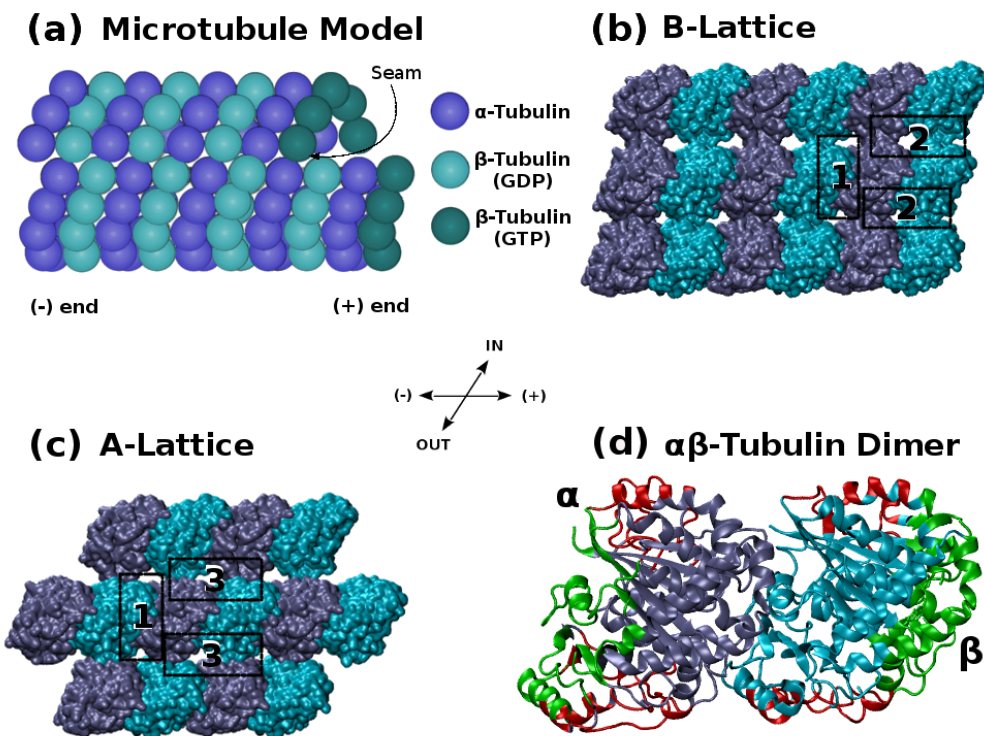
Figure 5.1: *Microtubule lattice and interfaces, with (dark blue) α-subunits and (cyan) β-subunits. (a) A model of a microtubule cylinder. (b) A model of the B-lattice configuration showing only nine tubulin dimers. (c) A model for the A-lattice configuration showing only seven tubulin dimers. In panels b and c, the three different interfaces between tubulin dimers that we studied are highlighted. These are 1) Longitudinal interface between two tubulin dimers (LongAB); 2) Lateral interface between two tubulin dimers in B-configuration (LatB); and 3) Lateral interface between two tubulin dimers in A-configuration (LatA). (d) A more detailed model of the $\alpha\beta$-tubulin heterodimer showing the domains that make lateral contacts (red) and the domains that make longitudinal contacts (green).*

relation to the hydrogen-bond strength.

5.3 Methods

5.3.1 Energy Calculation using AIM Approach

A hydrogen bond is a bond that involves three atoms: a hydrogen atom (H) attached covalently to an electronegative atom, such as N, O, or F, as one partner and an electronegative atom as another partner. The former electronegative partner is called the hydrogen bond donor (HD) whereas the latter electronegative partner is called the hydrogen bond acceptor (HA). Bader's AIM theory is a very attractive and successful method of characterizing bond strengths based on properties of critical points [18, 99]. Several successful studies that characterized hydrogen bonds based on topological properties of electron density at the bond critical points have been reported [94, 97, 139, 140]. In a previous study,we built a strong linear correlation between the density at the bond critical point

(BCP) located between the hydrogen atom and the acceptor atom, ρ_{H-A}, and the strength of the hydrogen bond obtained from a supermolecular approach [125]. The relationship had a coefficient of determination, r^2, of 0.96. This relationship was true for all kinds of hydrogen bonds that span the range of $0 - 60$ kJ/mol, as reflected by the heterogeneous training set used. Using this relationship, we obtained the parameters necessary for calculating the strength of hydrogen bonds by knowing only the value of the electron density, ρ_{H-A}, at the BCPs [125]. This relationship is given as:

$$E_{HB} = m\,\rho_{HA} + b \tag{5.1}$$

where E_{HB} is the energy calculated from a supermolecular approach, and m and b is the slope and the intercept of the linear correlation obtained, respectively. The parameters m and b were used to calculate the strength of hydrogen bonds in the tubulin interfaces.

5.3.2   Molecular Dynamics Simulations

Toward calculating the energies of hydrogen bonds in our system using this method, we obtained the Protein Data Bank (PDB) [141] crystal structure of bovine brain tubulin PDB:1JFF [142] and repaired it via basic homology modeling by adding missing residues from PDB:1TUB [26] using the software MODELER 9V6 [143]. The repaired PDB:1JFF structure was optimized using energy minimization via a conjugate gradient method over 40,000 time steps in an MD simulation in a neutralized water box using the NAMD program [144]. Using this minimized structure, the microtubule A- and B-lattice structures based on the microtubule geometry described in Li et al. [145] and Sept et al. [136] were built using an in-house PYTHON script in the software PY-MOL0.99rc6 [146] . Lateral orientation of the B-lattice was verified by overlaying a pair of lateral tubulin heterodimers from our model to the model prepared by Wells and Aksimentiev [147]. A root mean-square deviation (RMSD) of only 3.4 Å was reported, which is actually smaller than the resolution of the PDB:1JFF structure itself (3.5 Å).

Subsequently, a pair of interacting $\alpha\beta$-tubulin heterodimers were separated from each lattice to be used to study the hydrogen bonds. Specifically, a pair of longitudinal neighbors from the B-lattice model was separated to study the longitudinal interface (LongAB), a pair of lateral neighbors was separated from the B-lattice model to study the lateral B interface (LatB), and a third pair of lateral neighbors from the A-lattice was separated to study the lateral A interface (LatA). All the interfaces as well as the interacting pairs of $\alpha\beta$-tubulin heterodimers that were separated are shown in Fig. 5.1. Hence, we investigated three distinct systems, each containing a pair of $\alpha\beta$-tubulin heterodimers. For each $\alpha\beta$-tubulin pair system, we ran an MD simulation to obtain an equilibrated system. In detail, we added the cofactors, GTP and GDP, to their binding sites with the help of SWISS PDBVIEWER 4.1 [148]. Taxol, or any other stabilizer, was not added to the system.

As stated earlier, terminal β-subunits that are not capped with GTP are unstable and prone to depolymerization. Therefore, the terminal β-subunits in the three systems were all capped with GTP instead of GDP. The magnesium ion at the α-subunit GTP binding site was also included

---

Model building was done with the assistance of T. Craddock, Nova Southeastern University, Florida.

to stabilize the complex. C-termini were capped with n-methylamide residues. The C-terminal tails were not simulated because they are not available in PDB structures, and they are highly mobile and variable among tubulin isotypes. The C-terminal tail is also far away from lateral and longitudinal interfaces, and hence is not expected to have any direct contribution to lateral interactions. Moreover, the inclusion of this tail would require the usage of a very large water box that would significantly increase the computational load. We parameterized the protein system using the AMBERff12SB force field [48,149]. We parameterized the cofactors using the parameter set developed by Meagher et al. [78]. Ionization states were assigned using the PROPKA server [150–153]. Each system was solvated with a TIP3P water box extending 10 Å in each direction and neutralized by the addition of 72 Na$^+$ ions.

Additional ion pairs of Na$^+$ and Cl$^-$ were added to bring the ion concentration to 100 mM to mimic cellular conditions. Although the initial coordinates were obtained from a 13-protofilament microtubule model, we are effectively simulating a free pair of tubulin heterodimers in each system, given the geometry of the water box used. Then, the AMBER MD package [47] was used to minimize the complex through a series of 2000 steepest-descent and conjugate gradient steps with strong restraints on the protein heavy atoms (500 kcal mol^{-1} Å^{-2}). This was intended to relieve any hydrogen contacts caused by the addition of hydrogens using AMBER residue templates. Another 4500-step minimization was done to bring the whole system to the nearest local energy minimum. Then the system was heated, over 20 ps under constant volume, to a temperature of 310 K by the Langevin thermostat using restraints on the protein (10 kcal mol^{-1} Å^{-2}).

The restraints were then released gradually through a 100-ps run under constant pressure and temperature, and a production phase of 30-45 ns was run under the same conditions to attain RMSD equilibration. This production step was performed using GPU cores on the PharmaMatrix Cluster (University of Alberta) through the AMBER GPU-accelerated code [154–156]. All simulations were performed using periodic boundary conditions where the particle-mesh Ewald method was used for treating long-range electrostatics with a cutoff of 8.0 Å. When considering RMSD equilibration, we gave more attention to the interfacial residues than the residues that are distant from the interface. We clustered the snapshots that correspond to 20 ns extracted from the equilibrated region in the trajectories based on RMSD of the interfacial residues in eight clusters, using the average linkage algorithm [157] through the CPPTRAJ module of AMBER [158].

The centroid of the each cluster was considered to be the most representative structure of the cluster, and was processed further. Therefore, we used eight snapshots for every system. Each snapshot was processed using the CPPTRAJ module of AMBER to detect all hydrogen bonds. This is done by AMBER by listing all possible hydrogen-bond donors (HD) and acceptors (HA) in the system and then analyzing the distances between them as well as the angle made by HA-H-HD atoms. A cutoff of 3.0 Å for distance and 135° for angle are used as default criteria by AMBER for hydrogen bonds. Although these values are reasonable [159], we relaxed the strictness of our criteria to a distance of 3.3 Å and an angle of 120°, and then depended on the AIM method to confirm the presence of each hydrogen bond (a bond is present if there is a nonzero electron density

51

at the bond critical point).

These relaxed criteria were used to prevent missing any possible hydrogen bonds, i.e., to prevent false-negatives. The hydrogen bonds detected by AMBER were analyzed and all the bonds that are not interfacial in nature, i.e., not binding the two $\alpha\beta$-tubulin heterodimers together, were ignored. Other energetic contributions such as van der Waals and electrostatic interactions were also estimated using the molecular mechanics/generalized-Born surface area method (MM/GBSA) method as implemented in AMBER [22]. The program VMD 1.9.1 was used for viewing the MD trajectories [52]. After that, every hydrogen-bonding residue pair was analyzed individually using the AIM method.

5.3.3 Quantum Mechanical Calculations

Each pair of interacting amino-acid residues was characterized in a separate QM single-point calculation. The QM region was specified as the parts of the two residues making the hydrogen-bond contact, and we avoided cutting at polar or saturated bonds. The rest of the $\alpha\beta$-tubulin pair system, i.e. the MM region, was treated, together with the solvent, using electronic embedding, which incorporates the partial charges of the MM region into the quantum-mechanical Hamiltonian. This technique provides a better description of the electrostatic interaction between the QM region and the MM region (because it is treated at the QM level) and allows the QM wavefunction to be polarized. Table 5.1 lists the bonds that were cut to separate the QM region from MM region. The QM region was treated using density functional theory with the density functional B3LYP [109–111] and the basis set TZVP [112, 113]. This functional and basis set were chosen to match the ones that we used to develop the parameters [125].

All the QM calculations were done using the software GAUSSIAN 09 [124]. Subsequently, an AIM analysis was carried out using GAUSSIAN 09 and the electron densities at the BCPs were obtained. Difficult cases, i.e., cases that did not converge in Gaussian, were treated using the software suite AIMPAC which is more stable [160]. The hydrogen-bond energy was calculated using the parameters that we had developed in Ayoub et al. [125]. This QM calculation was applied to each instance of hydrogen bonding occurring between any pair of residues. Hence, we built several BASH scripts to automate all these procedures. As stated earlier, we used eight different representative snapshots for every system, and hence all these calculations were repeated for every snapshot. The total hydrogen-bond energies, as well as the per-residue energies, for each of the three main systems used were then obtained and analyzed.

5.4   Results and Discussion

The MD simulation runs were continued until RMSD equilibration of the interfacial residues was attained (interfacial residues are residues that have at least one atom within 8 Å from the neighboring tubulin heterodimer). Other residues distant from the interface were not considered, as they do not contribute to interfacial hydrogen bonding. Fig. S1 in the Appendix shows the RMSD equilibration plot of the backbone atoms of interfacial residues relative to the starting structure. The

Table 5.1: *Cutting bonds for QM/MM interface. The table lists the bonds that were cut to separate the QM from the MM region for each hydrogen-bonded residue. In all side-chain hydrogen bonds, the side chain, after the cut indicated in the table, was treated with QM. In backbone hydrogen bonds, atoms C, O, N, H, CA, and HA were treated with QM in all residues except Gly where the side chain was also included in the QM region. Amber atom names are used in the table (i.e. A, B, C, ... stand for $\alpha, \beta, \gamma, ...,$ respectively).*

Residue	Bond cut	Residue	Bond cut
Glu	CB–CG	ARG	CG–CD
Asp	CA–CB	GLN	CB–CG
Lys	CD–CE	Asn	CA–CB
His	CA–CB	Thr	CA–CB
Ser	CA–CB	Trp	CA–CB
Tyr	CA–CB		

LatB and LongAB systems were equilibrated early in the simulations. For LatA, we pursued the simulations a bit longer to make sure that the system was well equilibrated. The trajectory of the equilibrated region of each system was clustered in eight clusters, as explained before. These eight representative snapshots were analyzed for hydrogen bonds, and each pair of interacting residues was then subjected to a QM calculation and AIM analysis. The results of each system are listed in Tables 5.2–5.4. Each table shows the average total hydrogen-bond energies as well as the average per-residue energies over the eight different representative snapshots that were processed.

It should be noted that hydrogen bonds are highly dynamic in nature, which means that they keep forming and breaking over the course of the MD simulation. Therefore, we expect to see large variations in the per-residue hydrogen-bond energies, and this is why the standard deviation (SD) can sometimes be very high. In this case, high SD would represent highly dynamic bonds, whereas low SD would represent bonds that are persistent over the course of the MD trajectory. SD, in this case, does not reflect statistical errors in the calculations; instead, it reflects the transitory nature of each individual bond. However, the total hydrogen-bond energy values are expected to have a relatively smaller SD because bonds that are broken over the course of the trajectory are usually replaced by other bonds that are forming simultaneously.

Hence, we should have more precise values for the over- all hydrogen-bond energies. These variations could, however, be compensated for by other binding interactions, such as electrostatic interactions or van der Waals interactions, which were not included in this study. It is also important to note that the per-residue energies listed in the tables include all hydrogen-bond instances between the residue pairs. Therefore, the energy could be due to more than one hydrogen bond between the interacting pair. The tables only list hydrogen bonds that are stronger than -10 kJ/mol. Other weak bonds are included in the Appendix. Residue numbering follows the same scheme as in PDB:1JFF [142].

5.4.1 Longitudinal Interactions

The longitudinal (LongAB) interface (see Fig. 5.1b) is particularly important as it contributes to the building of a protofilament and happens to accommodate an important class of anticancer agents. This class includes the microtubule destabilizers known as vinca alkaloids [161, 162]. Understanding the interactions at this interface could give us insight into the mechanism of action for vinca alkaloids. An all-atom model of the LongAB system with subunit assignment can be found in Fig. S2. All the energies of hydrogen bonds in the LongAB interface can be found in Table S6 in the Appendix. The energies of the strongest bonds are listed in Table 5.2. Analyzing these results, we find that the total hydrogen-bond energy is -392 ± 59 kJ/mol. The table also shows the per-residue hydrogen-bond energies between the α-subunit of heterodimer 1 and the β-subunit of heterodimer 2. As shown in this table, the strongest bond network is the one between αArg2 and βGlu71 with an average energy of -40.1 kJ/mol. This bond is also persistent along the MD trajectory, as shown by the relatively low SD. αArg2 is, in fact, the very first residue in tubulin after the first methionine, and it makes an energetically significant bond. The second partner of this strong bond, βGlu71, is present in the S2-H2 loop of the β-tubulin, which shows the importance of this loop to longitudinal stability.

Another important residue is βArg401, which is present in the H11-H11' loop. This residue alone makes several strong hydrogen-bond networks with αGlu434 (-32.6 kJ/mol), αTyr262 (-32.1 kJ/mol), and αAsp438 (-23.9 kJ/mol), summing up to a total of nearly -90 kJ/mol on average, which is nearly one-fourth of the total binding energy. Thus, residue βArg401 could be described as the longitudinal glue of microtubules. Two of the partners that bind to this residue, namely αGlu434 and αAsp438, belong to the C-terminal tail of the α-subunit. Residues from 176 to 181 in β-tubulin, which correspond to the S5-H5 loop, make several hydrogen-bond networks with α-residues that belong to helix H10 through loop S9-S10. These bonds collectively make up nearly -90 kJ/mol, which comprises nearly one-fourth of the overall stability, and again reflects the importance of the S5-H5 loop in longitudinal stability.

The bonds involving S5-H5 loop are not only collectively strong, but they are also relatively persistent during the MD simulation, as indicated by their relatively low SD values. In fact, it has been shown that the S5-H5 loop in β-tubulin, particularly residues from 174 to 179, are very important for the vinca alkaloid binding and they comprise part of the vinca binding site [161, 163]. Considering that, as shown in this work, the same region contributes significantly to the longitudinal stability indicates that it is very likely that the binding of these anticancer agents destabilizes microtubules simply by disrupting the longitudinal hydrogen-bond networks between adjacent heterodimers along a protofilament. This could be verified by performing another equivalent study in the presence of one of these agents, then comparing the results.

The bonds made by βGly100 and βGly98 on one side and αThr257 and αGln133 on the other side, respectively, are also strong and steady with relatively low SD, suggesting a strong and persistent stabilization due to the S3-H3' loop of β-tubulin. It is also noticeable that GDP has no contribution to longitudinal stability when hydrogen bonds are considered; it does not appear in

Table 5.2: *Energy of hydrogen bonds in the LongAB interface. This structure refers to two αβ–tubulin heterodimers aligned longitudinally (TUB 1 and TUB 2). Average energy and standard deviation (SD) values are taken from eight different representative snapshots. Energies are expressed in kJ/mol.*

TUB 1–α	TUB 2–β	E_{average}	SD
Arg2	Glu71	−40.1	14
Glu434	Arg401	−32.6	17
Tyr262	Arg401	−32.1	17
Arg243	Asp76	−29.9	14
Thr349	Val181	−25.8	12
Asp438	Arg401	−23.9	25
Val260	His406	−22.1	11
Gln133	Gly98	−21.9	8
Thr257	Gly100	−21.4	7
Lys352	Thr180	−18.7	9
Asn249	Gln11	−18.6	11
Asn329	Lys176	−15.3	10
Lys163	Glu411	−14.7	19
Asn258	Val181	−13.9	8
Other Bonds		−61.3	—
Total Energy		−392	59

our list, although it binds close to the longitudinal interface. It is worth mentioning that residues making hydrogen-bond networks with more than one residue, such as βArg401, usually have a relatively high SD. This is not surprising because during the MD simulation, such a residue may break its bonds with one residue and soon form other bonds with another residue to maintain the longitudinal stability. This behavior raises the SD calculated over the eight representative snapshots for each bond. All the major hydrogen bonds in the longitudinal interface are shown in Fig. 5.2. It is obvious that major strong hydrogen bonds are well distributed along the width and the length of the longitudinal interface, which imparts even more stability to the protofilaments because they act as pillars for the protofilament structure. Residue βArg401 and its strong and persistent hydrogen-bond network are also shown in the figure.

5.4.2 Lateral B Interactions

The LatB interface represents the lateral interface between two tubulin dimers in two adjacent protofilaments in the B-configuration (see Fig. 5.1b). An all-atom model of the LatB system with subunit assignment can be found in Fig. S3. This interface is especially important not only because it brings protofilaments together to form a microtubule cylinder, but because it is also very close to the taxane binding site. The taxane binding site is the binding site for many microtubule-stabilizing antimitotic drugs such as taxol, epothilone, discodermolide, eleutherobin, and sarcodictyin [29–32, 126]. Analyzing the lateral interface could give us an insight into the

Figure 5.2: *Major hydrogen bonds at the longitudinal interface. It is clear that they are distributed all over the width and length of the interface to provide stronger support to the protofilament structure.*
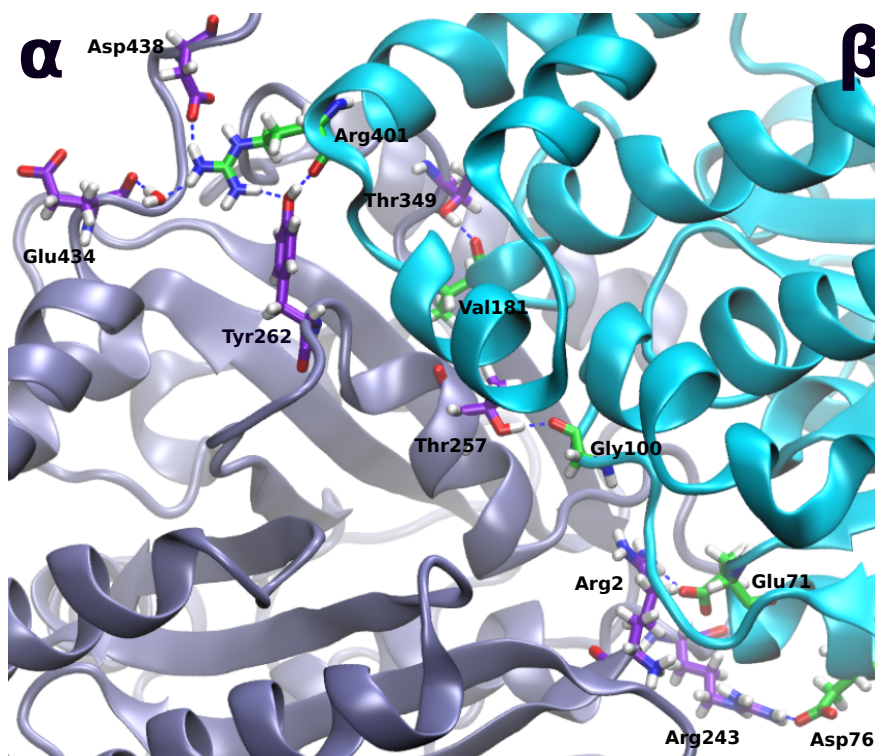
detailed mechanism of action of such agents. All the energies of hydrogen bonds in the LatB interface can be found in Table S7 in the Appendix. The energies of the strongest bonds are listed in Table 5.3. The first and most obvious observation in the table is that, in a 95% confidence interval, lateral hydrogen bonds are significantly stronger than longitudinal ones, -462 ± 70 vs. -392 ± 59 kJ/mol.

This is apparently counterintuitive, because we know, a priori, that lateral contacts break before longitudinal ones and this is why depolymerizing microtubules display transient structures that look like ram's horns [136, 164]. There could be two justifications for getting such results: the first one is that we only calculated the respective hydrogen-bond energies. Other sources of energies could balance out this energy difference. Particularly, electrostatic interactions and dipole-dipole interactions are expected to be more destabilizing in the lateral orientation than in the longitudinal one. This is because the similarly-charged subunits are packed closer together in the lateral orientation than in the extended longitudinal one. We estimated the van der Waals as well as electrostatic contributions using the MM/GBSA calculation and found that the longitudinal interactions in a protofilament are ~ 130 kJ/mol stronger than lateral interactions, which supports our justification.

The other possible reason for this difference is that in the LatB simulations we actually modeled GTP-capped tubulin dimers, as explained in Methods. Terminal tubulin dimers capped with

Table 5.3: *Energy of hydrogen bonds in the LatB interface. This structure refers to two αβ–tubulin heterodimers (TUB 1 and TUB 2) aligned laterally in the B–configuration. Average energy and standard deviation (SD) values are taken from eight different representative snapshots. Energies are expressed in kJ/mol.*

α–α **Interactions**					β–β **Interactions**			
TUB 1–α	TUB 2–α	E_{average}	SD		TUB 1–β	TUB 2–β	E_{average}	SD
Arg215	Glu90	−42.7	18		Arg308	Asp116	−44.8	20
Lys338	Asp127	−28.5	8		Glu290	Arg88	−38.9	12
Glu297	Arg121	−24.6	19		Arg308	Asp120	−36.7	14
Glu297	Lys124	−23.6	12		Lys299	Asp90	−32.6	12
Glu284	Ser54	−22.3	14		Asp297	Lys124	−22.3	19
Gln372	Glu55	−18.4	9		Tyr342	Asp120	−18.4	18
Tyr282	Ser48	−11.7	11		Ser280	Arg88	−16.6	16
His283	Phe49	−10.5	8		Lys338	Lys124	−14.8	10
					Lys338	Ser126	−12.0	11
Other Bonds		−27.5	—		Other Bonds		−14.8	—
Subunit Energy		−210	35		Subunit Energy		−252	65
Total Energy							−462	70

GTP stabilize microtubule structures more strongly than those capped with GDP, which is why depolymerization usually happens after hydrolysis of the terminal GTP [13]. Hence, lateral contacts in our case are expected to be enhanced by the presence of GTP, and this could be the reason why they are stronger than longitudinal ones. Preliminary results from other simulations being presently performed support this explanation, because we found out that in the presence of GDP instead of GTP, the two heterodimers are significantly more weakly connected to one another. The data in Table 5.3 also show that the average contribution of the β-β interactions is comparable to the average contribution of the α-α interactions, namely −252 ± 65 vs. −210 ± 35 kJ/mol, at a 95% the confidence interval. However, if the simulations were run in presence of taxol, the relative contributions of the two subunits could have been different.

Examining the β-β subunit interactions, we find out that the strongest hydrogen-bond network is the one between βArg308 (from the H9' helix) and βAsp116, with a strength of −44.8 kJ/mol and a relatively moderate SD, signifying persistence of the bonds over the MD trajectory. βArg308 still makes another strong and largely persistent bond with βAsp120, at a strength of −36.7 kJ/mol. Thus, βArg308 contributes, in total, nearly −80 kJ/mol to lateral stability, which is nearly one-third of the overall β-β stabilization. βGlu290 and βArg88 contribute a largely persistent hydrogen-bond network of −38.9 kJ/mol. The H2-S3 loop, which involves βArg88 and βAsp90, is extensively involved in lateral stabilization, making bonds that sum up to nearly −90 kJ/mol, which is nearly one-third of the overall β-β hydrogen- bond energy. The H3 helix, which involves residues βAsp116, βAsp120, βLys124, and βSer126 and others, is responsible for most of the stabilization occurring between β-subunits. Interactions involving these residues sum up to a total of nearly −165 kJ/mol, which is approximately two-thirds of the overall β-β stabilization.
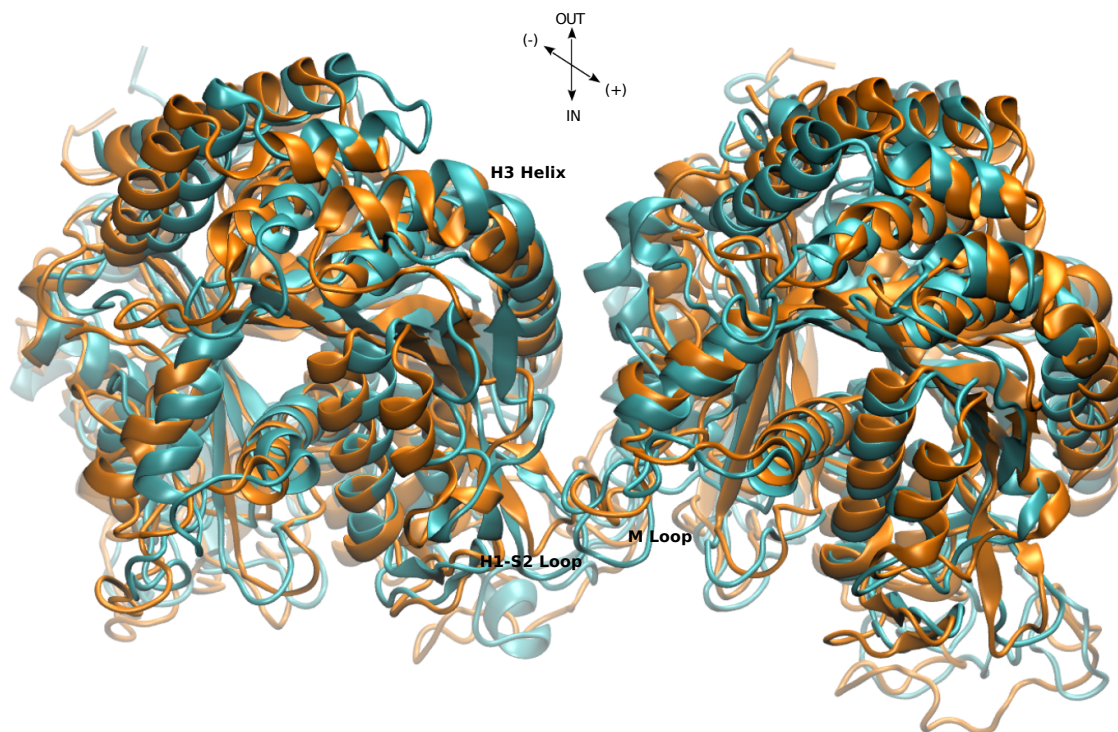
Thus, H2-S3 loop is responsible for one-third of β-β hydrogen-bond energy, and the H3 helix is responsible for the remaining two-thirds. On the other hand, the contribution of the M-loop from the opposite β-subunit (TUB 1-β in Table 5.3) is relatively small compared to the H3 helix contribution from TUB 2-β, amounting to only -16.6 kJ/mol on average, via bonds that involve βSer280 from the M-loop. The H1-S2 loop in TUB 2-β has no contribution to lateral hydrogen-bonding in B-lattice. This is contrary to the conclusions that were drawn by Li et al. [145], who argued that lateral stability is mostly due to interactions between M-loop and H1-S2 loop rather than being due to the H3 helix. However, these authors stated that this result itself is contrary to a previous conclusion they reached, which attributed most of the lateral stability to the H3 helix rather than the H1-S2 loop. This shows some discrepancies that could be attributed to the fact that their conclusions were not based on a study of the energetics of lateral contacts, but only on geometric criteria. It could also be attributed to the fact that we considered hydrogen bonds only, in this study. Other energetic components could still come into play. However, considering the conditions of our simulations, the difference between the results of Li et al. [145] and our findings may be understood because microtubules stabilized by taxol were used in their experiments but not in ours. It is known that taxol restructures the M-loop in a way that stabilizes lateral contacts [145, 165–167]. Because our simulations did not include taxol or any other stabilizer, we do not expect our results to match the results of Li et al. [145] with regards to the role of the M-loop in imparting lateral stability. Moreover, Li et al. argued that the role of the H3 helix becomes

Figure 5.4: *Major hydrogen bonds in the LatB system at (a) β–β interface and (b) α–α interface.*
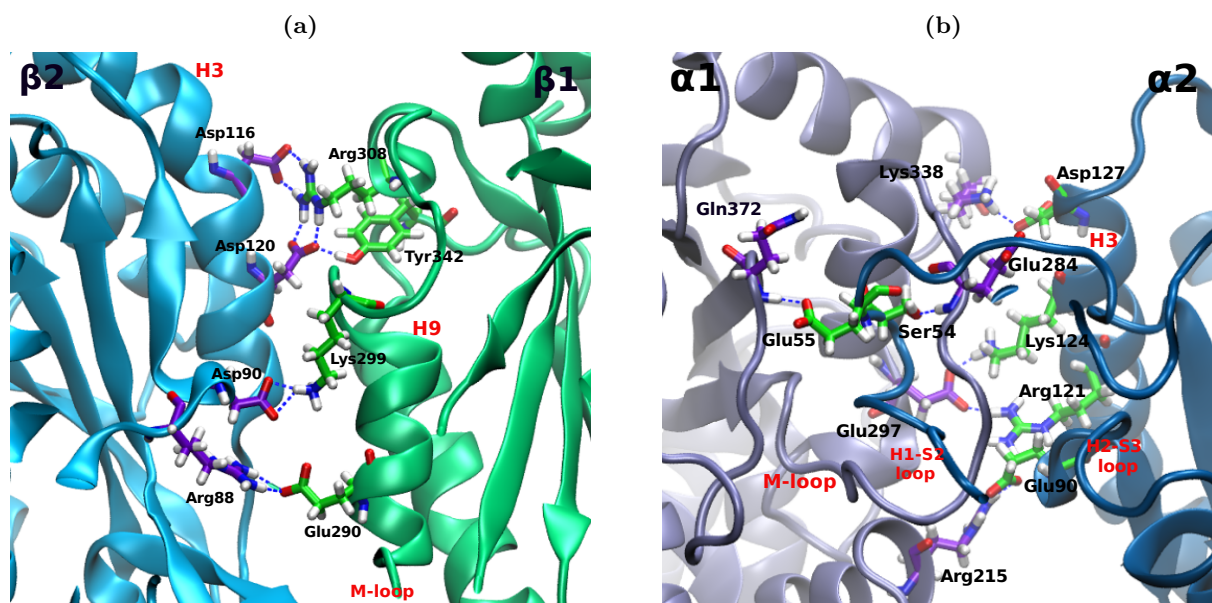
more pronounced when the number of protofilaments in a microtubule increases, inasmuch as this decreases the angle between laterally adjacent protofilaments and brings the H3 helix closer to the neighboring heterodimer. This could be more similar to our simulated system that included only one pair of heterodimers instead of a complete microtubule, and thus could rearrange during the course of the MD simulations and draw the subunits closer to generate more H3 helix contacts. Fig. 5.3 shows the LatB system before (orange) and after (cyan) the simulations. It is clear that after the simulations, the two heterodimers have rotated inward, coming closer to each other and creating more interactions with the H3 helix at the expense of breaking interactions between the M-loop and the H1-S2 loop.

Comparing the residues on our list to the residues that compose the taxol-binding site, we find that residues βGlu290 and βSer280 are in common. These two residues make hydrogen-bond networks with the neighboring β-tubulin subunit that sum up to -55.5 kJ/mol, most of which comes from the bonds between βGlu290 and βArg88, which alone make up -38.9 kJ/mol. Comparing the role of H3 and the M-loop in lateral stability in our study and considering the finding that H3 is much more involved in stabilization than the M-loop, it appears more likely that the binding of taxol may make the M-loop more involved in lateral stability. Addition of taxol could rearrange this domain and favor stronger lateral contacts. (This could be verified by performing another simulation in the presence of taxol and estimating the per-residue contribution of all these residues.) It is worth mentioning that a conformational change of the M-loop into a short helix upon binding of stabilizers was recently confirmed by Prota et al. [168]. However, they did not address the energetic effects of this restructuring on lateral contacts. The major hydrogen-bond networks in β-β interactions are shown in Fig. 5.4a.

The α-α interactions are similar to the β-β ones in that they extensively involve helix H3, with the H2-S3 loop on one side and the H9-S8 loop on the other side. However, the α-α interactions are different because they also extensively involve interactions between the H1-S2 loop and the M-loop. These interactions involve the bonds between αSer54, αSer48, and αPhe49 on one side and αGlu284, αTyr282, and αHis283 on the other side. These, and other bonds between the M-loop and the H1-S2 loop shown in Table S7 in the Appendix, add up to an average total of nearly -63 kJ/mol. It is also apparent that the hydrogen-bond network between αArg215 and αGlu90 is the strongest in the system, with an average energy of -42.7 kJ/mol. Fig. 5.4b shows the major hydrogen-bond networks that bring α-subunits together in lateral orientation. An interesting phenomenon that is noticed from the figure is the intertwining of the M-loop and the N-terminal H1-S2 loop. This structure was conserved over the entire length of the MD simulation, which reflects its stability.

The lateral interface is highly populated with oppositely charged residues as compared to the longitudinal interface; hence, we also expect a stabilizing electrostatic contribution between these charged residues.

5.4.3   Lateral A Interactions

This LatA interface represents the lateral interface between protofilaments in an A-lattice configuration, shown in Fig. 5.1c. An all-atom model of the LatA system with subunit assignment can be found in Fig. S4. The A-lattice configuration is less significant than the B-lattice because the latter has been empirically observed to be much more predominant. It is worth mentioning that to simulate an A-lattice, three, rather than two, tubulin dimers must be included in the simulation because of the subunit offset. Because this is computationally very demanding, we simulated the relevant subunits only. That is, we used α- and β-subunits from dimer 1, a β-subunit from dimer 2, and an α-subunit from dimer 3, discarding the α-subunit of dimer 2 and the β-subunit of dimer 3. This is acceptable because, as illustrated in Fig. 5.1c, the discarded subunits have no contacts with the studied interfaces and are far away from them.

All the energies of hydrogen bonds in the LatA interface can be found in Table S8 in the Appendix. The energies of the strongest bonds are listed in Table 5.4. It shows that, in a 95% confidence interval, the average overall hydrogen-bonding in the A-lattice is not significantly different from the B-lattice, with energies of -472 ± 46 vs. -462 ± 70 kJ/mol, respectively. Sept et al. [136] also studied the difference between B-lattice and A-lattice energetics considering solvation energy only, but they found that the B-configuration, corresponding to a subunit rise of 8-9 Å, is more stable than the A-configuration, corresponding to a subunit rise of 52 Å. Drabik et al. [137] also found a similar effect when comparing the potential of mean force in the two configurations. Therefore, in light of our findings, the difference in stability between B-lattice and A-lattice configurations could be attributed to solvation energy and other energetic components rather than to hydrogen bonds. It is worth mentioning that the A-lattice configuration is not exclusive to the A-lattice; it is also a part of the B-lattice that appears only at the seam, as depicted in Fig. 5.1a.

In Table 5.4, we differentiate between α-β interactions and β-α interactions, because, due to

Table 5.4: *Energy of hydrogen bonds in the LatA interface. This structure refers to one $\alpha\beta$–tubulin heterodimer (TUB 1) aligned laterally in the A–configuration with an α- and β-subunit (TUB 2β and TUB 3α). Energies are expressed in kJ/mol.*

α–β **Interactions**				β–α **Interactions**			
TUB 1–α	TUB 2–β	E_{average}	SD	TUB 1–β	TUB 3–α	E_{average}	SD
Asp47	Arg284	−42.2	15	Arg88	Glu279	−35.2	15
Lys124	Asp297	−31.7	9	Lys124	Glu284	−27.0	12
Gln85	Ser280	−23.6	10	Ile86	Tyr282	−23.0	12
Asp46	Arg278	−23.0	13	Asp90	Lys280	−18.6	15
Asp127	Asn334	−22.6	5	Asn54	Glu284	−14.4	13
Asp120	Lys338	−22.4	14	Glu127	Thr334	−13.8	18
Gln128	Gln293	−18.6	12	Asp90	Ala281	−12.9	11
Asp47	Gln282	−16.5	14	Glu127	Thr337	-10.3	14
Glu55	Arg284	−14.8	18				
Arg121	Asp297	−13.1	24				
Asp47	Arg278	−11.9	13				
Other Bonds		−58.2	—	Other Bonds		−17.8	—
Subunit Energy		−299	23	Subunit Energy		−173	44
Total Energy						−472	46

differences between α- and β-subunits, they are not identical. In our notation, α-β interactions represent the half of the system in which the N-terminal H1-S2 loop, helix H3, and the H2-S3 loop of the α-subunit interact with the M-loop and other domains of the β-subunit. However, β-α interactions represent the other half of the system in which the opposite is true. An interesting observation is that β-α interactions are much stronger than β-α interactions, as manifested by an energy value of -299 ± 23 vs. -173 ± 44 kJ/mol, respectively. This suggests that the involvement of the M-loop of the β-subunit, rather than the α-subunit, in lateral contacts greatly enhances the stability of the system by interacting with N-terminal H1-S2 loop of the opposite subunit.

In particular, residues βArg284, βArg278, and βGln282 and others make lateral hydrogen bonds with the N-terminal loop of the adjacent α-subunit that add up to nearly -120 kJ/mol. Most of these bonds are absent in the β-α interaction half-system, and the contribution of M-loop H1-S2 loop interactions is nearly -30 kJ/mol. The M-loop of the β-α interaction half-system prefers to bind with H2-S3 loop and H3 helix. This is illustrated in Fig. 5.5 a and b, which shows the major hydrogen bonds in the two half-systems. Based on this we can also expect taxol, which is hypothesized to induce M-loop lateral interactions, to impart stability to the system via this mechanism. This can even be extrapolated to the B-lattice because Table 5.3 does not record any major contribution of the β M-loop, especially residue βArg284, to the overall stability. Inclusion of taxol in the simulation could alter this behavior and enhance the role of the M-loop.

Finally, the comparison of the top-ranking residue pairs in Tables 5.2 to 5.4 to the residues that are conserved throughout different $\alpha\beta$-tubulin isotypes [163,169] showed that there is considerable agreement. In other words, residues important for interfacial stability are highly conserved among

Figure 5.5: *Major hydrogen bonds in the latA system at (a) α–β interface and (b) β–α interface.*
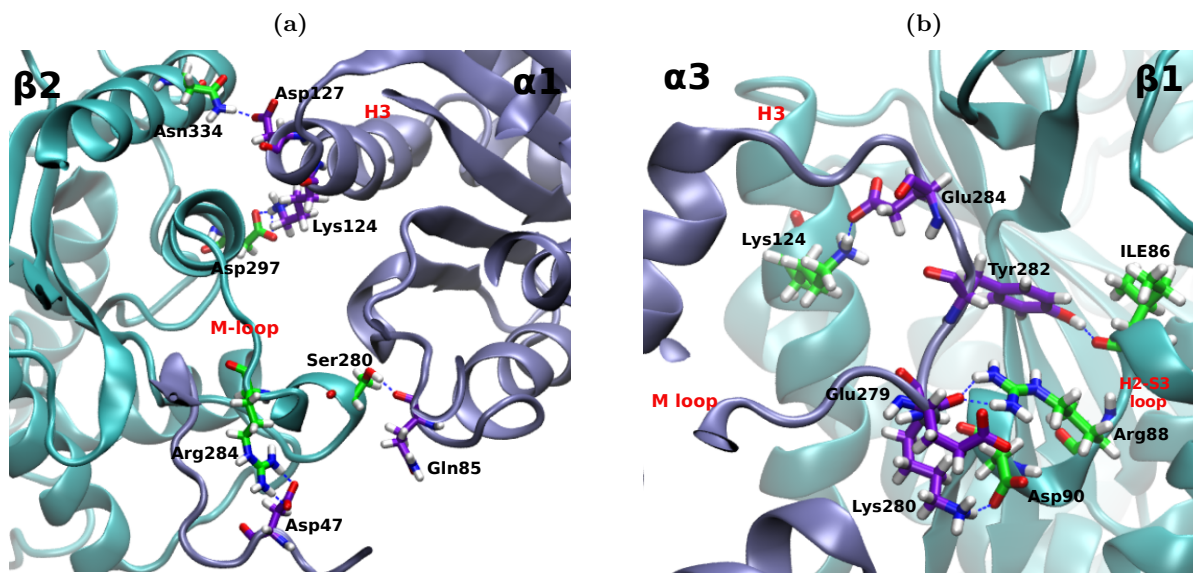
different tubulin isotypes.

5.5   Conclusion

The concept of the density at the bond critical point obtained from the AIM analysis is very useful in the calculation of hydrogen-bond energies. In this article, we have implemented a seemingly new technique for the application of this method to macromolecules, namely tubulin dimer-dimer systems. The systems were equilibrated by MD simulations and then studied by QM calculations employing density functional theory followed by an AIM analysis. The three different interfaces studied, longitudinal interface as well as lateral interfaces in B- and A-lattice configurations, revealed that hydrogen bonding is an important player in the stability of tubulin systems. One limitation of this study is the fact that we used only eight representative snapshots from the trajectory of every system. Running relatively long simulations, ensuring clustering of the trajectories, and choosing the centroid of each cluster, should alleviate this limitation. Analyzing the overall hydrogen-bond energies in different interfaces showed that lateral contacts are stronger than longitudinal ones, which was attributed to the stabilization imparted by the GTP cap on β-tubulin subunits.

The contribution of the β-β interactions to the overall lateral stability in the B-configuration was shown to be comparable to that of the α-α interactions in a 95% confidence interval. Running the same simulations in the presence of taxol could give different results and offer more insight into this aspect. The study also showed that the stability of the B-lattice configuration is comparable to the A-lattice when hydrogen bonds are concerned. This suggests that other energetic contributions could be responsible for the observed difference in predominance between the two lattice forms.

Per-residue hydrogen-bond analysis was found to be in agreement with empirical data regarding residues critical to longitudinal stability and residues involved in the binding of vinca alkaloids. This suggests the mechanism of action of vinca alkaloids could be in the alteration of the conformations of interfacial residues upon binding, which disrupts the interfacial hydrogen-bond network and destabilizes the microtubule.

The β M-loop was shown to have a weak contribution to the stability of the LatB system, contrary to its large contribution to the stability of the LatA system. The weak contribution of the M-loop to the stability of the LatB system was attributed to the absence of taxol or any other microtubule stabilizer in our simulation that causes the M-loop to drift away and be replaced by helix H9 and the H9-S8 loop interacting with helix H3 in lateral contacts. Further elucidation of the role of anticancer agents would require running the simulations in the presence of vinca alkaloids, taxol, and GDP to reach a final conclusion regarding the mechanisms of stabilization or destabilization of microtubules. Most of the residues that contributed significantly to stability of tubulin-tubulin interactions were also found to be highly conserved among different tubulin isotypes.

Chapter 6

Detailed Per-residue Energetic Analysis Explains the Driving force for Microtubule Disassembly

6.1 Introduction

Microtubules (MTs) are cellular organelles that participate in major cellular processes such as mitosis, cell shape maintenance, cell motility and motor protein transport and constitute a major target for a wide range of drugs, most notably anti-mitotic chemotherapy agents such as paclitaxel. Due to their importance in cell biology, MTs have been the topic of active research into their structure and function for several decades [170]. The pivotal role of MTs in cell division, by forming the mitotic spindle that segregates chromosomes, makes them an important target for antimitotic cancer chemotherapy drugs [24, 171]. A detailed description of the MT structure and dynamics instability was presented in Sec. 1.3.

Several studies have been conducted to determine which specific structural transitions that accompany GTP hydrolysis or taxol binding are responsible for their effect on MT stability, especially the transition of the tubulin dimer between its straight and curved states [172–176]. In the most recent of these studies, Alushin et al. found that GTP hydrolysis leads to a compaction around the E-site nucleotide which is reversed upon taxol binding [176]. This compaction was proposed to generate a strain which is powered by the energy of GTP hydrolysis and is believed to be released only through outward curving of protofilaments, initiating disassembly [177]. A missing component in these studies, however, is the quantification of the free energy changes that accompany these structural transitions. Due to the difficulties related to its experimental measurement, many simulations have been conducted to study detailed MT energetics [136–138, 164, 178, 179]. In a recent study we have analyzed the strength of hydrogen bonds that bring and hold tubulin subunits
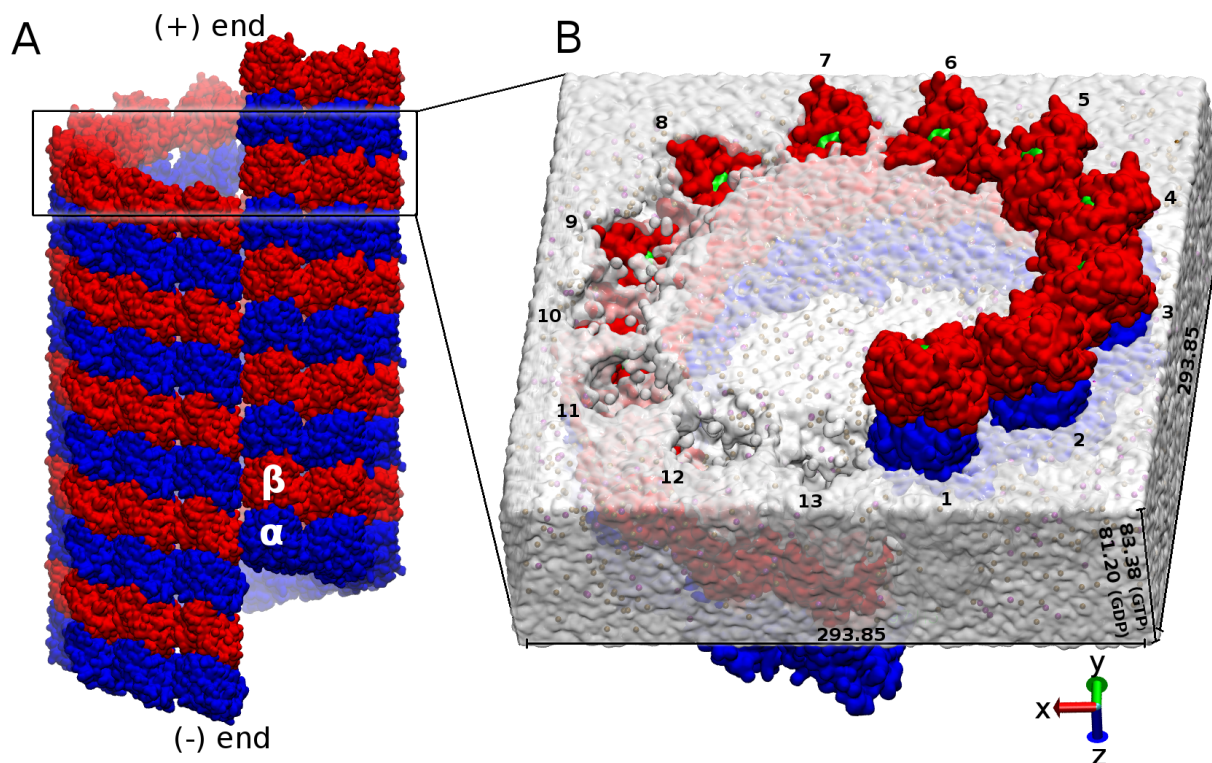
Figure 6.1: *Model of MT structure*
(A) A model of an MT lattice showing α (blue) and β (red) tubulin subunits. It shows the plus and minus end as well as the seam.
(B) A model of the system used in the molecular dynamics simulations. Tubulin dimers are numbered from 1 to 13, GDP (or GTP) cofactor is shown in green within β-tubulin, the second GTP cofactor is buried between α and β subunits, water is represented by the white box, within which purple spheres represent Cl⁻ and brown spheres represent Na⁺. Periodic box dimensions in units of Å are also shown.

together within different lattice configurations [126]. However, in all of these simulations, several factors were still missing. Most importantly, the full energetics of a complete MT model, which is essential to understanding the thermodynamics of tubulin assembly, has not been estimated yet due to the high computational price associated with such analyses. A detailed energy balance involving contributions due to each residue, domain or subunit, to the best of our knowledge, was never considered.

As a result of recent advances in computational technology, GPU-based computations can now be implemented to perform very demanding calculations in a reasonable amount of time. With this technology readily available, we simulated two complete all-atom MT models and studied in detail their energetics. The models studied are: (a) an MT with GDP in the E-site (GDP-Model) and (b) an MT with GTP in the E-site (GTP-Model). We did not need to look for a non-hydrolyzable analogue of GTP as hydrolysis is not a problem in molecular dynamics simulations, in contrast to experimental procedures [177]. The MT model that we used was initially built by Wells and Aksimentiev [147] utilizing sophisticated theoretical techniques to combine experimental structural

information from a cryo-electron microscopy map of MT at 8 Å resolution [145] and electron crystallography structure of tubulin at 3.5 Å resolution [142]. We combined this model with the recently published crystal structures [176] in order to generate an atomistic representation involving an infinite number of infinitely long MTs. This is possible due to the use of periodic boundary conditions.

6.2 Methodology

6.2.1 Building the Models

The recent structures for guanylyl-(alpha, beta)-methylene-diphosphonate (GMPCPP) and GDP bound MTs at resolutions of 4.7 and 4.9 Å respectively [176], represented an excellent starting point for building the models presented here. The 3×3 lattice PDB structures of 3J6E (with GMPCPP) and 3J6F (with GDP) were processed using MOE software [180] by the addition of hydrogens and prediction of ionization states. The central tubulin dimer of the 3×3 lattice in each case was separated and was repaired by addition of the missing residues (Residue 1 in β-tubulin and residues 1,39 to 48, 440 in α-tubulin) from the PDB structure 1TUB [26], using MOE. We modified GMPCPP into GTP since in our simulations there is no need to use the nonhydrolyzable GTP analogue as hydrolysis is not expected in MD simulations. Next, for both GTP and GDP systems, the repaired tubulin was superimposed over the 13 tubulin dimers in the complete MT model built by Wells and Aksimentiev [147], producing a hybrid complete MT model for both systems. Thus, we produced two models, the GTP-Model and the GDP-Model, by combining the helical structural configuration developed by Wells and Aksimentiev with the lattice tubulin coordinates obtained from Alushin's model. Several clashes existed at lateral interfaces between tubulin dimers and were resolved through a short minimization using the Generalized Born (GB) continuum model in Amber [47].

Each model, as shown in Figure 6.1B, has 13 tubulin dimers in an MT orientation. For the GDP-Model, each tubulin has GTP, Mg^{2+} and four coordinating water molecules at the α-tubulin N-site, and GDP at the β-tubulin E-site. For the GTP model, there was GTP, Mg^{2+} and four coordinating water molecules at both the N-site and the E-site. Solvation was carried out using box of dimensions $293.85 \times 293.85 \times 83.38$ (or 81.20) Å^3 for the GTP- and GDP-Models, respectively. The z-component was obtained from Alushin's lattice structure [176] and ensures perfect longitudinal alignment of tubulin dimers in both systems (see Figure 6.1B). Both x and y components were obtained from Wells' structure [147]. A total of 181,000 TIP3P water molecules were added in the solvation box. This number was obtained based on several optimization trials which guaranteed consistency in box dimensions and density throughout the simulations. A total number of 442 Na^+ ions was needed for neutralizing the GTP-Model, versus 455 for the GDP-Model. An extra 327 Na^+ and Cl^- ions were added to bring the salt concentration to 0.1 M.

During the addition of water and ions, we made sure that no atoms were placed in positions which will be occupied by the periodic images of our system in both the positive and negative z

direction (see the gaps in the water box of Figure 6.1B) . Thus, exploiting the periodic boundary conditions, the replication of our nearly 720,000-atom system in all directions should effectively result in an infinite number of infinitely long MTs. The AMBER Molecular Dynamics package was used for solvation, ionization, and dynamics [47].

6.2.2 Parameterization and Dynamics

The all-atom forcefield AMBERff12SB was used to parameterize the protein [48, 149]. Cofactors were parameterized utilizing the parameter set developed by Meagher et al. [78]. Each of the two systems was then minimized through nearly 1000 steps of the steepest descent algorithm followed by about 6000 steps of the conjugate gradient algorithm. Then, the systems were heated, with restraints of 10 kcal mol^{-1} Å^{-2} on the protein, to a temperature of 310 K using the Langevin thermostat over 20 ps under constant volume. This was followed by 200 ps of density equilibration under constant temperature and pressure, in which the restraints were eliminated gradually, followed by a production phase of 50 ns for each system. Simulations were performed using NVIDIA Tesla K20X GPU cards on the PharmaMatrix Cluster (University of Alberta) through AMBER GPU-accelerated code [154–156]. All simulations were performed using periodic boundary conditions employing the particle-mesh Ewald method for treating long-range electrostatics with a non-bonded cut off of 10.0 Å under constant pressure with anisotropic volume rescaling.

6.2.3 Trajectory Analysis

The 50-ns trajectory of each system was analyzed for several structural and conformational aspects. Most of the analysis was done utilizing the CPPTRAJ module in AMBER [158], MM/GBSA implementation in AMBER [181] plus several scripts that we designed to facilitate data analysis. The software VMD 1.9.1 was also used for viewing trajectories and image rendering [52].

Data analysis included calculating the total as well as the per-residue MM/GBSA binding energies [22] between pairs of tubulin dimers in lateral and longitudinal orientations. These calculations involved all the 13 heterodimers included in the simulations and would always give the energy per MT ring (Figure 6.1B). Hence, energetic contributions were assessed via the equation:

$$E_x = \epsilon_x(R_{13}L_1) + \sum_{k=1}^{12} \epsilon_x(R_k L_{k+1}) \tag{6.1}$$

for lateral systems, and the equation:

$$E_x = \sum_{k=1}^{13} \epsilon_x(R_k L'_k) \tag{6.2}$$

for longitudinal systems. In both equations, E_x represents an energetic contribution of a given residue, domain or subunit x per MT ring of 13 tubulin dimers shown in Figure 6.1B. In Equation 6.1, $\epsilon_x(R_k L_{k+1})$ is the energetic contribution of the same entity x in a subsystem composed of only tubulin k, treated as a "receptor", and tubulin $k + 1$, treated as a "ligand". $\epsilon_x(R_{13}L_1)$ does the

same but at the lateral seam, taking into account the flip between α- and β-subunits. In Equation 6.2, $\epsilon_x(R_k L'_k)$ carries the same concept except that the ligand in a longitudinal subsystem is simply the periodic image of the receptor, hence the prime. Therefore, we ended up investigating 12 lateral subsystems plus one lateral subsystem at the seam and 13 longitudinal subsystems, for each model. An illustration of each subsystem is shown in Figure 6.2. This means that the dimer whose M-loop is involved in lateral interactions was termed "receptor" in lateral subsystems, and the dimer whose α-tubulin was involved in longitudinal interactions was termed as "receptor" in longitudinal subsystems. This distinction was necessary since we noticed that the energetic contributions can vary between tubulin dimers acting as receptors and those acting as ligands.
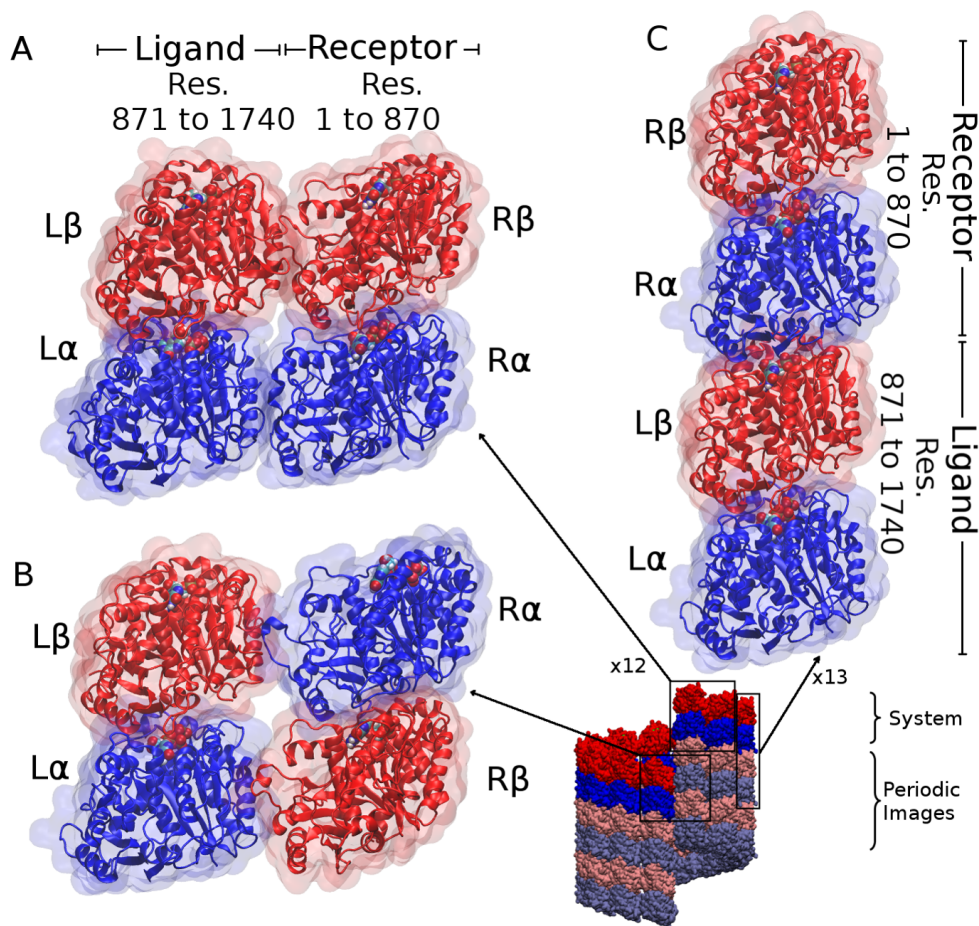
Figure 6.2: *Tubulin subsystems used for MM/GBSA calculations.*
(A) A subsystem of two lateral tubulin dimers extracted from MT simulations where the receptor (R) is tubulin k and the ligand (L) is tubulin $k+1$, where k runs from 1 to 12,
(B) Same as (A) but at the seam. I.e. the receptor is tubulin 13 and the ligand is tubulin 1. Residue (Res.) numbering in (A) and (B) are the same,
(C) A subsystem of two longitudinal tubulin dimers where the receptor is tubulin k and the ligand is its periodic image, where k runs from 1 to 13.
Total number of residues may differ slightly between the GDP and GTP models.

All the energy calculations were performed on 200 evenly-spaced snapshots from the last 10

ns of the molecular dynamics trajectory where equilibration was confirmed. A solvent and solute dielectric constant of 80 and 1, respectively, were used for electrostatics in the Amber MM/GBSA implementation.

6.3 Results and Discussion

6.3.1 Molecular Dynamics Equilibration

A 50-ns MD trajectory was analyzed for several equilibration aspects, the first of which is the root-mean-square deviation (RMSD) of the backbone atoms relative to the starting structure. In addition, two nearly perpendicular MT cylinder diameters, namely Dx and Dy, were also calculated along the trajectory. Referring to the tubulin dimer numbering in Figure 6.1B, the diameter Dx was defined as the distance between the center of mass of dimer 4 and the center of mass of dimer 10 and 11, while Dy was defined as the distance between the center of mass of dimer 1 and the center of mass of dimer 7 and 8. In both diameters, only the distance projection on the x-y plane was considered as this is what gives the cylinder diameter. Plots showing the change in RMSD of the backbone atoms, Dx and Dy over simulation time for the GDP- and GTP-Models are shown in Figures 6.3A and 6.3B. The two diagrams indicate a strong correlation between fluctuations in RMSD and in diameters which indicates that most of RMSD fluctuations are due to changes in the circular shape of MT cylinders rather than the rearrangement of domains. The two diagrams also show the flexibility of MT cylinders as they deform spontaneously from a circular to an oval shape and vice versa.

Since our particular interest is in the MT energetics, we used the overall MT energy across lateral and longitudinal inter-dimer interfaces as an indication of whether the system is equilibrated or not. Hence, we calculated these energies using MM/GBSA and the formulae in Equations 6.1 and 6.2 and plotted the total energy per MT ring versus simulation time (Figures 6.3C and 6.3D). Both plots indicate that the overall lateral and longitudinal energies in both the GDP- and GTP-Models have already equilibrated at least before the last 20 ns of the MD simulation time. The plots also show that the large fluctuations in RMSD or Dx and Dy hardly affect the MT energetics at either of the two interfaces, which is a good indication of the energetic stability of our models.

6.3.2 Lateral Energetics in the GDP-Model

Total breakdown of the predicted energy contributions enabled us to perform the analysis for different residues, domains, subunits, and dimers across both lateral and longitudinal inter-dimer interfaces. Before listing the results, it should be noted that energies calculated via the MM/GBSA method do not necessarily reflect absolute energy values. Rather, they are used for relative comparison within the same model [182]. It should also be noted that all energies listed here are calculated per MT ring, unless otherwise specified.

The overall energy of interaction across the 13 lateral tubulin interfaces (see Figure 6.1B), E_{tot}^{lat}, was found to be -411 ± 29 kcal/mol, nearly 60% of which is due to α-α interactions and the rest is
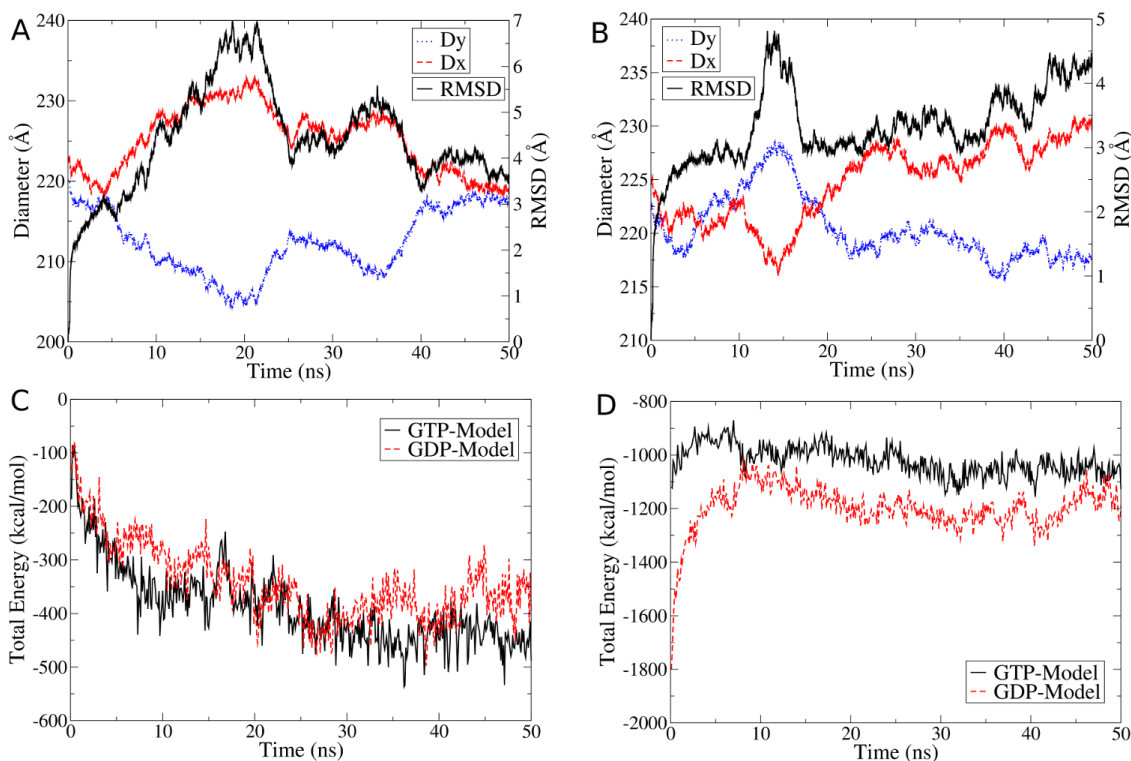
Figure 6.3: *Equilibration plots.*
Plots showing the change in RMSD of protein backbone atoms and the two nearly perpendicular diameters Dx *and* Dy *over simulation time in (A) GTP-Model and (B) GDP-Models.*
Equilibration of total sum of interaction energies versus simulation time across (C) lateral and (D) longitudinal inter-dimer interfaces.

due to β-β interactions. On the other hand, the contribution of the dimer acting as a receptor (see Figures 6.2A and 6.2B), E_R^{lat}, was about 54% of the overall energy while the rest was attributed to the ligand, E_L^{lat}, with the difference entirely attributed to solvation effects rather than direct interactions. It should be noted, however, that the α subunit of the ligand (Lα) and the β subunit of the receptor (Rβ) together contribute -312 ± 29 kcal/mol which is nearly 75% of E_{tot}^{lat}, with the Lα contribution slightly larger than that due to Rβ. The contribution of Lβ and Rα was found to be much smaller, only 25% of E_{tot}^{lat}. Upon structural inspection, this 50% difference, being almost entirely due to electrostatic interactions, was attributed to diagonal interactions between subunits; although the interface between Lα and Rβ is dominated by oppositely-charged residues and thus stabilizing the interaction, the opposite is true at the destabilizing interface between Rα and Lβ which has, for example, residues Rα/Glu220 and Lβ/Asp130 destabilizing the lateral interface by 12 ± 1 and 10 ± 2 kcal/mol, respectively.

As to the energetic breakdown according to interaction types, the contribution of the van der Waals and non-polar solvation energy, $E(\text{vdW}+\text{SA})$, to the overall energy is largely stabilizing with an average value of -1476 kcal/mol, 85% of which is due to the van der Waals (vdW) interactions. This stabilization is opposed by destabilization due to electrostatic interactions; the average sum

70

of electrostatic and the polar solvation energy, $E(\text{ele} + \text{GB})$, is 1065 kcal/mol. This is expected since tubulin dimers are highly negatively charged and tend to repel each other.

Regarding the detailed energy contributions per individual residues, the most important residue across the lateral interface was found to be Rβ/Tyr283 followed by Rα/His283 and Lα/His88, with overall stabilization energies of -90 ± 5, -47 ± 5 and -42 ± 3 kcal/mol per MT ring, respectively. Rβ/Tyr283 alone supplies more than 20% of lateral stability most of which is due to the vdW interactions. In fact, most of the stabilizing residues on top of our list were neutral ones with a strong stabilizing vdW component. On the other hand, almost all of the destabilizing residues were charged ones with a strong electrostatic component, most destabilizing of which is Lβ/Lys124 with an energy of 22 ± 7 kcal/mol. A complete list of the different energetic contributions of each residue in the ligand and receptor per MT ring is available in [183].

Domain contributions to the overall energy per MT ring were also calculated and Figures 6.4A and 6.4B shows the most relevant of them. The contribution of the M-loop in both α and β subunits is by far the largest, with values of -112 ± 10 and -159 ± 10 kcal/mol, respectively, making up about two thirds of the energy of the overall lateral interactions. This agrees well with previous predictions, although precise values of their energetic contributions were never calculated [145, 165, 166]. Other less important domains are the Lα/N-terminal loop, Lα/H2-S3 loop, Lα/H3 helix and Lα/H9 helix at the α interface with a stabilization of -72 ± 6, -62 ± 6, -57 ± 10 and -16 ± 7 kcal/mol, respectively [145, 165]. Lβ/H3 helix at the β-β interface, however, has a strongly destabilizing effect of 37 ± 8 kcal/mol. This supports previous predictions based on structural analysis by Li et al. and Nogales et al; however, these authors did not specify if these interactions are stabilizing or not [145,165]. Additionally, Lβ/H2" helix and Lβ/H1'-S2 loop also have relatively strong stabilizing contributions of -52.6 ± 7 and -43 ± 5 kcal/mol, respectively.

6.3.3 Lateral Energetics in the GTP-Model

The overall interaction energy across the lateral interface in the GTP-Model, E_{tot}^{lat}, was found to be -482 ± 29 kcal/mol, nearly 60% of which is due to the α-α interactions. This average overall energy is 71 kcal/mol (nearly 20%) more stable than the overall energy of the GDP-Model which explains the role of GTP in stabilizing MTs as will be shown later. Nearly 90% of this difference in stability is solely attributed to enhancement of the contribution of the ligand, both α- and β-subunits, rather than the receptor. As was noticed in the GDP-Model, Lα and Rβ are also responsible for most of the lateral stabilization in the GTP-Model, -338 ± 22 kcal/mol (70% of E_{tot}^{lat}).

Upon breakdown of the interaction energy to its individual components, we find that in the GTP-Model, the $E(\text{vdW} + \text{SA})$ contribution becomes -1432 kcal/mol while $E(\text{ele} + \text{GB})$ becomes 950 kcal/mol. Comparing this to the GDP-Model, it turns out that GTP destabilizes the vdW and non-polar solvation interactions by 44 kcal/mol and stabilizes electrostatic and polar solvation interactions by 115 kcal/mol, which results in the net stabilization of 71 kcal/mol as mentioned earlier. This difference becomes clear by analyzing Figures 6.4A and 6.4B for domain contributions and Figure 6.5 for residual contributions. It is apparent from Figure 6.4A that GTP strengthens the
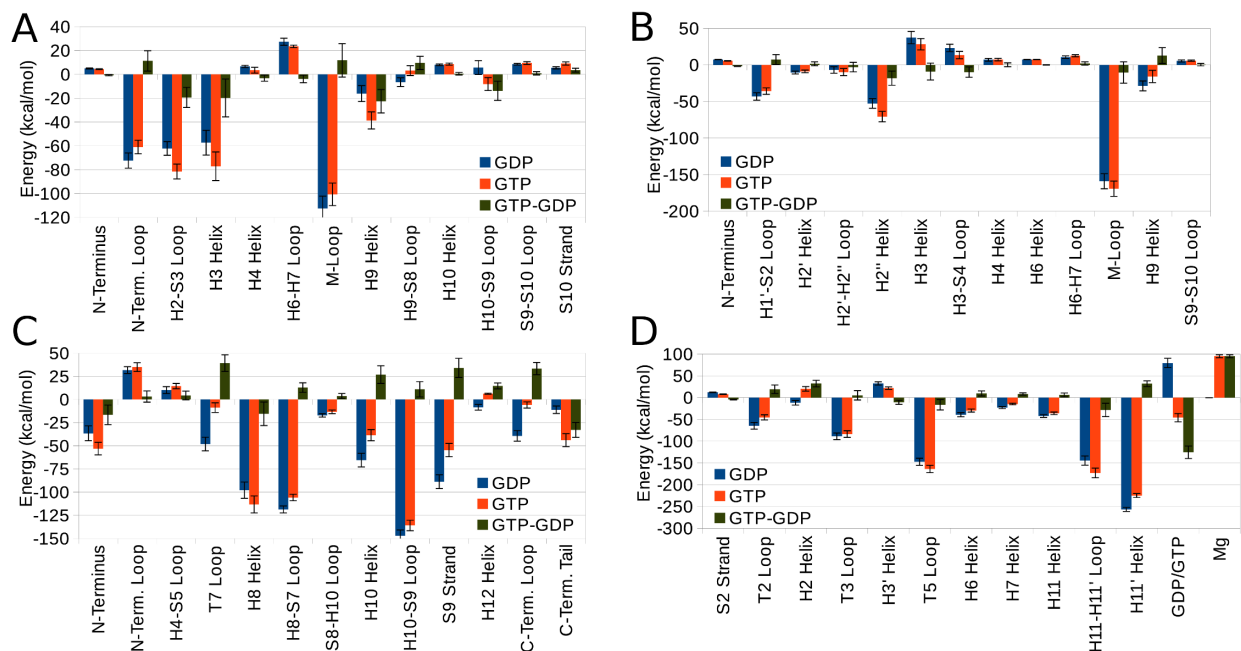
Figure 6.4: *Domain contributions to overall energy.*
Energetic contributions of important domains across lateral interface in (A) α and (B) β subunits and across longitudinal inter-dimer interface in (C) α and (D) β subunits. Data are shown for GDP- and GTP-Model as well as the difference between them (GTP-GDP). On the x-axis of (A) and (B), domains H4 helix and before occur at lateral interface of the ligand while domains after that occur at receptor lateral interface. In (C), all domains belong to receptor while all the domains in (D) belong to ligand. See ligand and receptor definitions in Figure 6.2.

contributions of the Lα/H3 helix and Rα/H9 helix by 23 ± 10 and 20 ± 16 kcal/mol, respectively. Most of this helix stabilization can be attributed to interactions involving Rα/Glu290 (residue number in Figure 6.5, i, is 290), residue Lα/Asp127 (i=998), and residue Lα/Arg123 (i=994). These three residues stabilize the GTP-Model over the GDP-Model by energy values of 31, 20 and 19 kcal/mol, respectively, mostly due to electrostatic interactions. Upon structural analysis it is apparent that GTP slightly rotates the dimer acting as a ligand toward the one acting as a receptor, thus allowing stronger interactions between H3 and H9 with oppositely-charged residues. GTP also enhances the stability imparted by the Lα/H2-S3 loop and the Rα/H10-S9 loop, although it moderately decreases the role of the Lα/N-terminal loop as well as the Rα/M-loop in the overall MT stability.

Similar conclusions are reached in regard to the β-subunit and the effect of the Lβ/H2" helix through residue Lβ/Asp90 (i=1401) and the Rβ/M-loop through residue Rβ/Arg284 (i=724). Both domains are stabilized in the GTP-Model by extra 18 ± 10 and 10 ± 15 kcal/mol compared to the GDP-Model, respectively. The charged nature of all these residues explains why most of GTP stabilization is manifested in $E(\text{ele} + \text{GB})$ not $E(\text{vdW} + \text{SA})$. Figure 6.4B also shows that GTP reduces the destabilization caused by the Lβ/H3 helix and the Lβ/H3-S4 loop. On the other hand, GTP reduces stability imparted by the Lβ/H1'-S2 loop and the Rβ/H9 helix. Details of the
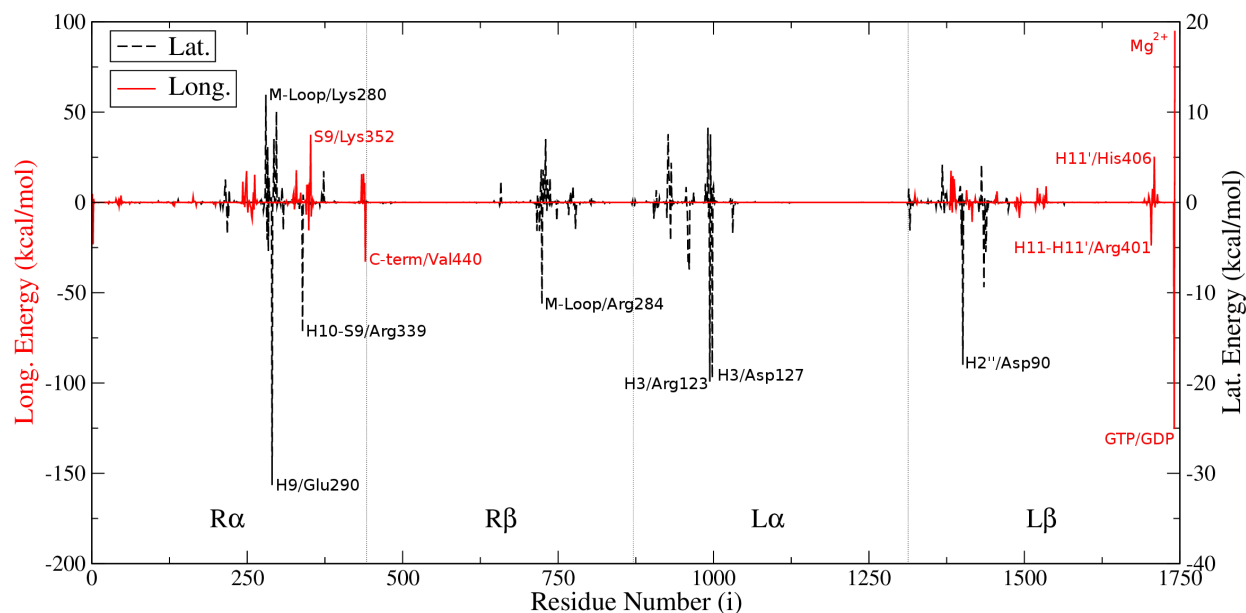
Figure 6.5: *Energetic contributions of residues. Difference between overall residual contributions per MT ring in GTP- and GDP-Model; ($E_i^{GTP} - E_i^{GDP}$), where i is the residue number running from 1 to 1742. Different energy axes are used due to differences in magnitude of interactions at both interfaces. Important residues are labeled.*

contribution of each residue in the GTP-Model is available in [183]. .

6.3.4 Longitudinal Energetics in the GDP-Model

Analysis of the strength of interactions across the longitudinal inter-dimer interface in the GDP-Model yielded an overall energy of -1240 ± 32 kcal/mol per MT ring, which is nearly three times the lateral interaction energy. This is in agreement with structural observations [165]. Due to the orientation of tubulin dimers at the longitudinal inter-dimer interface, the contributions of Lα and Rβ are essentially zero and will always be neglected here. On the other hand, the contribution of Lβ is 54% of the total value, and the remainder is contributed by Rα. The breakdown of this energy yields an average $E(\text{vdW} + \text{SA})$ of -2668 kcal/mol which is almost twice as large as the value across the lateral interface. This is obviously due to the tighter packing of the residues here as opposed to looser packing at the lateral interface. The average $E(\text{ele} + \text{GB})$ across the longitudinal inter-dimer interface is 1428 kcal/mol and it is 34% larger than its value at the lateral interface.

Per-residue energy analysis reveals the most important residues to longitudinal stability, the first of which is Lβ/Arg401 from the H11-H11' loop which alone supplies -101 ± 7 kcal/mol (nearly 10%) [126]. After that come residues Lβ/Phe404 and Lβ/Trp407 from the H11' helix both of which support longitudinal stability by contributing -91 ± 3 and -78 ± 3 kcal/mol, respectively. This makes the two former domains, which constitute part of the tubulin C-terminal domain, the most critical for longitudinal stability in the β-subunit (Figure 6.4D). The figure also shows that the following domains: the T5 loop, T3 loop, and T2 loop are also very important for longitudinal

stability. The role of the GDP cofactor appears quite influential at the longitudinal inter-dimer interface, in contrast to the lateral one. It is primarily destabilizing with a large contribution of 79 ± 11 kcal/mol due mainly to a strong electrostatic repulsion with the highly negative environment, despite its strong salt bridge with Rα/Lys352. Residual analysis of the Rα subunit also shows some relatively less important residues; Rα/Trp346, Rα/Tyr262 and Rα/Lys352 with energy contributions of nearly -60 kcal/mol for each of them. These and other residues are responsible for the following domains: the H10-S9 loop, H8-S7 loop, and the S9 strand being the top stabilizers in Figure 6.4C. The H8 and H10 helices are also relatively important for longitudinal stability. Both the C- and N-terminal domains are important as well, with the N-terminal loop being a destabilizer, in contrast to its role at the lateral interface.

6.3.5 Longitudinal Energetics in the GTP-Model

The overall interaction energy across the longitudinal inter-dimer interface in the GTP-Model was found to be -1098 ± 30 kcal/mol per MT ring, which is 141 kcal/mol (10%) less stable than the GDP-Model system. This difference is attributed to a 7% decrease in the Rα and 3% decrease in the Lβ interactions. Upon energetic breakdown we see that GTP destabilizes the vdW and non-polar solvation energy by nearly 250 kcal/mol, while stabilizing electrostatic and polar solvation energy by nearly 110 kcal/mol. This could be due to the longstanding observation that GTP leads to an expansion in the E-site and lengthening of the tubulin dimers. That is, axial dimer repeat changes from 81.20 Å in the GDP-Model to 83.38 Å in the GTP-Model [173, 176]. This reduces the packing of atoms at the interface and hence lowers both the vdW attraction and electrostatic repulsion, the former being affected most due to its stronger dependence on distance.

Looking into domain contributions in Figures 6.4C and D we see how GTP destabilization of longitudinal interactions can be subdivided. The most pronounced difference between the GDP- and the GTP-Model appears in regard to the cofactors at the E-site. Although GDP was largely destabilizing in the GDP-Model, GTP becomes relatively largely stabilizing, with an energy change from the GDP-Model of nearly 125 ± 14 kcal/mol. However, this change should not be considered without taking into account the effect of the Mg^{2+} ion that accompanies GTP. This magnesium ion introduces an instability of 95 ± 4 kcal/mol to the GTP-Model. Hence, the overall effect of replacing GDP by GTP and a magnesium ion is a stabilization of 30 kcal/mol on average. Other causes of the lack of stability in the GTP-Model Lβ include the decrease in the contribution of the H11' helix because GTP offsets interactions by Lβ/His406 (i=1709) by as much as 25 kcal/mol. This is because this histidine is protonated in the GTP-Model and neutral in the GDP-Model and therefore behaves differently in both cases. Being charged in the GTP-Model, it is distracted from the strong attractive vdW interactions it makes with the Rα/H8-S7 loop by electrostatic and hydrogen bonds with other residues in the Lβ subunit. GTP also causes longitudinal stabilization due to the domains: the H2 helix and the T2 loop whose contributions decline while causing stabilization due to the H11-H11' loop and the T5 loop whose contributions rise. As to the Rα-subunit (Figure 6.4C), stabilization due to several domains declines in the GTP-Model. These domains include the

T7 loop, the S9 strand, the C-terminal loop, the H10 helix, the H12 helix, the H8-S7 loop, and the H10-S9 loop. In short, the GTP-Model is longitudinally less stable than the GDP-Model in most of the domains occurring at the longitudinal inter-dimer interface. An exception to this rule is the increased stabilization due to the C-terminal tail, the N-terminus and the H8 helix, Figure 6.4C shows the extent of stabilization or destabilization imparted by GTP on each domain. We should also mention that the strong attraction of the Rα/T7 loop emerging after GTP hydrolysis (Figure 6.4C) could explain the proposed compaction of the E-site after GTP hydrolysis [176]. In fact, the overall increase in longitudinal dimer-dimer attraction after GTP hydrolysis, as shown by the different values of E_{tot}^{long} in both models, explains the driving force for this E-site compaction.

Among other important residues, Rα/Lys352 (i=352) of the domain S9 strand has a largely reduced contribution in the GTP-Model, as shown in Figure 6.5, which is 37 kcal/mol less stabilizing than in the GDP-Model. While having comparable vdW contributions in the two models, this residue suffers strong repulsion probably due to a nearby Mg^{2+} ion in the GTP-Model. Another important residue is Rα/Val440, located in the C-terminus of the α-subunit in our model. GTP enhances the stabilization caused by this residue by nearly 33 kcal/mol over the GDP-Model. Additional important residues and their contributions are available in [183].

6.3.6 Energy Profile Explains the MT Disassembly Mechanism

Depolymerizing MTs display protofilaments that peel into "ram's horns" formations under high magnesium buffer conditions. The ends of MTs become frayed, however, under physiological concentrations of magnesium [172]. The energy profile throughout the longitudinal inter-dimer interface provides a clear explanation for the disassembly mechanism, its driving force, and its relation to Mg^{2+} concentration. We characterized each residue in the longitudinal subsystems by its radial distance from the MT lumen in Å, which was plotted on the x-axis. The interaction energies of residues, per MT ring, over half-closed intervals of $[x, x+3)$ were summed up and plotted on the y-axis to produce the radial energy profiles in Figures 6.6A, 6.6B and 6.6C.

The diagram in Figure 6.6A leads to a striking observation that the energy distribution throughout the longitudinal inter-dimer interface is not even, with the outward portion ($x > 30$ Å) largely outweighing the inward portion ($x < 30$ Å), with the center of mass of tubulin being at $x \approx 30$ Å. To mention specific values, in the GTP-Model, the outward portion provides nearly -956 kcal/mol while in the GDP-Model it provides -982 kcal/mol, both values being larger than 80% of the overall longitudinal interaction energy. This uneven distribution of energy, or forces of attraction, is proposed to yield a strong torque that tends to curl MT protofilaments outwardly, breaking lateral bonds and promoting disassembly as illustrated in Figure 6.7A.

Radial energy profiles of different components of the interaction energy are also shown in Figures 6.6B and 6.6C, where electrostatic interactions cause very strong repulsion through the inward portion and attraction only at the periphery where the H11-H11' loop and particularly residue Lβ/Arg401 are located. We propose a pivotal role for this residue, and for the entire C-terminal domain, in regulating dynamic instability. Electrostatic repulsion by the inner domains and at-
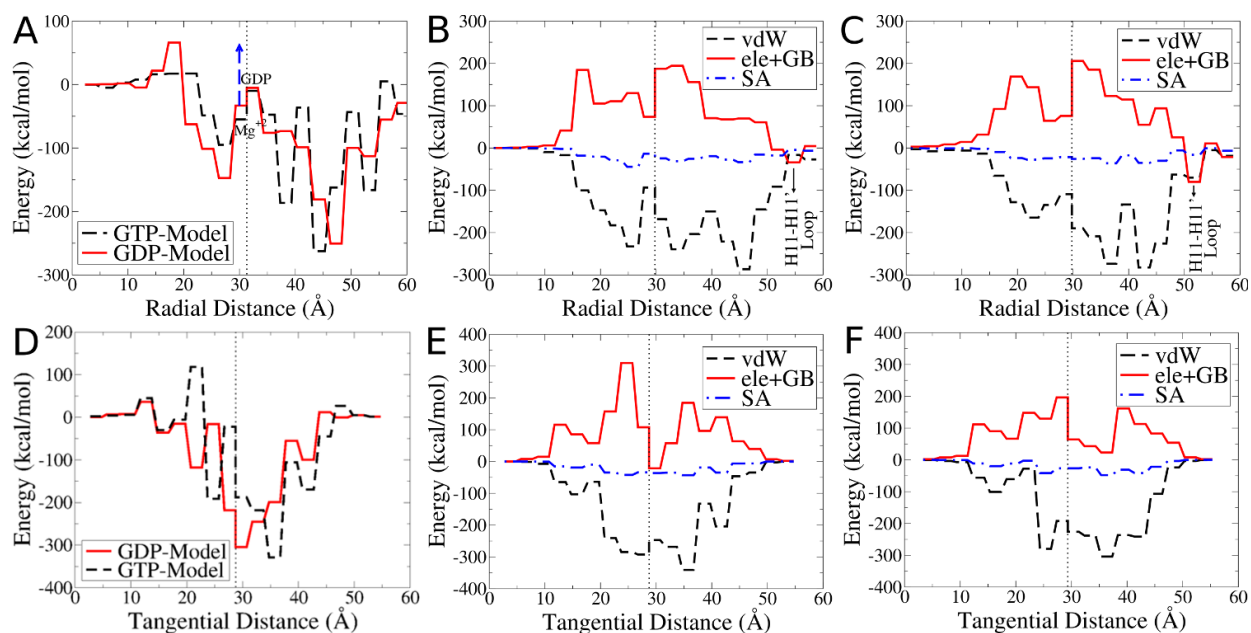
Figure 6.6: *Energy profiles at longitudinal inter-dimer interface.*
The figures shows the sum of energetic contributions of residues located at distance intervals of 3 Å apart, plotted against the radial distance between these residues and the MT lumen (A,B,C) or the tangential distance between the residues and laterally adjacent dimer (D,E,F) in GTP- and GDP-Model. Dotted lines represent the center of mass of tubulin.
(A) Radial distribution of total energy in both models. Blue dashed arrow shows how destabilizing the effect of Mg^{2+} is on the GDP-Model if it remains after GTP hydrolysis. (B) Radial distribution of the energy components $E(vdW)$, $E(ele+GB)$ and $E(SA)$ in the GDP-Model and in (C) the GTP-Model, (D) Tangential distribution of total energy in both models. On the x-axis, $x < 30$ is the intermediate domain and $x > 30$ is the nucleotide binding domain. (E) Tangential distribution of the energy components $E(vdW)$, $E(ele+GB)$ and $E(SA)$ in the GDP-Model and in (F) the GTP-Model.

76

traction by the outer C-terminal domain is the recipe for outward curling and disassembly in MTs. The vdW distribution will also work, as shown in Figure 6.6B for the GDP-Model and 6.6C for the GTP-Model, to curl protofilaments outward until the vdW contacts, and other components, are balanced out.

The largely destabilizing Mg^{2+} ion (see Figure 6.4D) also plays an important role. Even though GDP at the E-site has low affinity for Mg^{2+} [184], it may still bind to Mg^{2+} if it is present in high concentrations or Mg^{2+} may stay in the E-site after GTP hydrolysis. This largely destabilizes the inner portion of the protofilament (blue dashed arrow in Figure 6.6A), allowing outward forces to pull tubulin out with even less resistance from the other side, thus promoting outward curling and MT disassembly. This explains why large Mg^{2+} concentrations promote ram's horns formations [185] and increase the rate of disassembly [186, 187], while its low concentrations produce frayed ends and lower rates of disassembly [172].

To explain MT disassembly from a free energy perspective, Figure 6.7A shows an illustration of the analyzed situation. As already established, uneven distribution of attractive interactions along the longitudinal inter-dimer interface favors outward curling. In the GTP-Model, outward curling is favored by -956 kcal/mol of interaction energy outwardly with respect to the center of mass of tubulin, as compared to -982 kcal/mol in the GDP-Model. These curl-favoring energies/forces are opposed by the lateral interaction energies which tend to pull protofilaments back from both sides, i.e. double the effect. The magnitude of this effect is $2 \times E_{tot}^{lat}$, giving -964 kcal/mol in the GTP-Model which is much larger than -822 kcal/mol in the GDP-Model, all energies given per MT ring. We propose that this lateral inward pull balances out the longitudinal outward push in case of the GTP-Model. That being said, the presence of a GTP cap at the tip of the MT would prevent outward curling and thus provide stability for the entire MT structure. After GTP hydrolysis reaches the cap, however, lateral bonds become weaker and longitudinal outward push manages to break the lateral contacts, causing outward curling and MT disassembly. High concentrations of Mg^{2+} may also increase outward curling and the disassembly rate, as explained earlier.

Similar observations could be made about the tangential energy profiles at the longitudinal inter-dimer interface. Figures 6.6D, 6.6E and 6.6F show the tangential energy profiles with the x-axis showing the distance from the laterally adjacent protofilament. On the x-axis, $x < 30$ is the tubulin intermediate domain while $x > 30$ is the nucleotide binding domain with $x \approx 30$ being at the center of mass (see Figure 6.7B). Figure 6.6D shows that in The GTP-Model, the distribution is also uneven with right-side portion being -1023 kcal/mol (nearly 93% of the total) as compared to -887 kcal/mol (71% of the total) in the GDP-Model. This means that in the GTP-Model, there is a strong force tilting it sideways. However, after GTP hydrolysis and rearrangement of domains at the longitudinal inter-dimer interface, that force largely decreases and the uneven distribution starts to balance out, as shown in Figure 6.6D, decreasing the strain on lattice integrity. This is in perfect agreement with the recent findings of Alushin et al. [176] They observed that GTP hydrolysis and the release of an inorganic phosphate group leaves a hole within the longitudinal inter-dimer interface between tubulin dimers producing a strain which results in sideway tilting
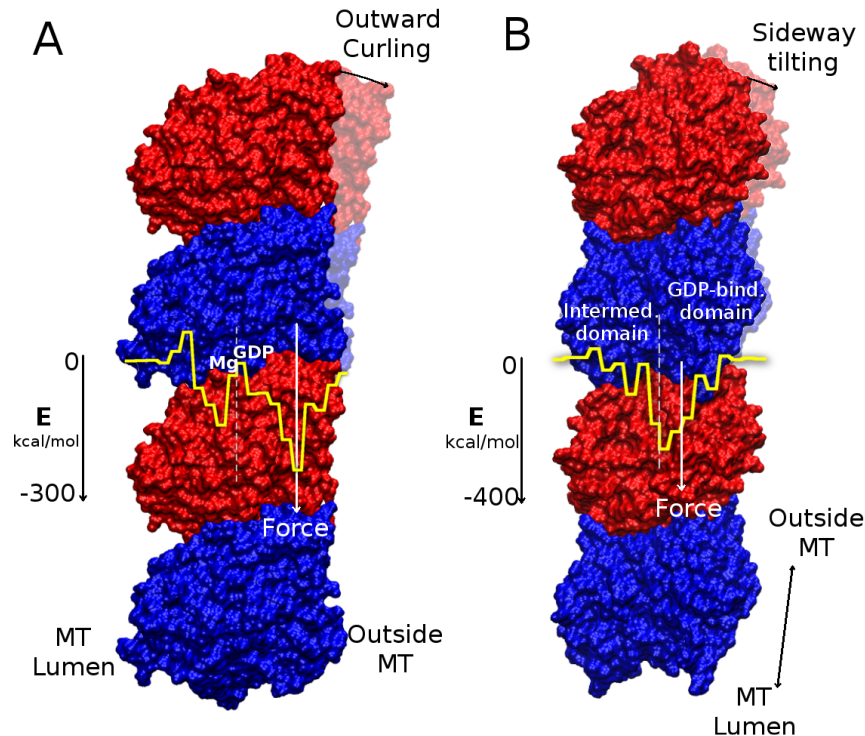
Figure 6.7: Mechanism of MT disassembly.
(A) Radial energy distribution of the GDP-Model at the longitudinal inter-dimer interface is superposed on a protofilament to show how uneven the energy distribution is. This produces a torque that leads to outward curling of the protofilament. (B) Tangential energy distribution of the GDP-Model showing slight sideway tilting due to the slightly uneven distribution of energy. α-subunits are colored blue while β-subunits are red.

in the same direction [177]. In the present work we show that this tilting is also driven by the uneven energy distribution along the same direction as in the work of Alushin et al. [176] (see Figure 6.7B). However, this sideway tilting should not be considered as the the driving force of disassembly since it is orthogonal to the outward curling. Combining the two effects together, we conclude that uneven distribution at the longitudinal inter-dimer interface generally leads to a large outward and slight sideway tilting of protofilaments, the former of which is responsible for disassembly of GDP-bound MTs.

6.3.7 Energy Distribution around the Microtubule Ring

As mentioned in the Methods section, the MT ring was divided into 13 subsystems of laterally adjacent tubulin dimers and another 13 subsystems of longitudinally adjacent tubulin dimers (see Figure 6.2). All of the energies presented earlier were expressed per MT ring, meaning that they were summed over the 13 subsystems. In this section, however, we focus on the interaction energy in each subsystem. Figures 6.8A and 6.8B shows energy diagrams for lateral and longitudinal interactions superposed over the MT ring. We first note that the shape of the lateral interactions (Figure 6.8A) in the GDP-Model is very distorted with several "kinks" of very low energy. When

compared to the GTP-Model, its shape is much less distorted. This could come as a straightforward consequence of the fact that GTP-Model is laterally more stable than the GDP-Model and hence suffers less "deformations".
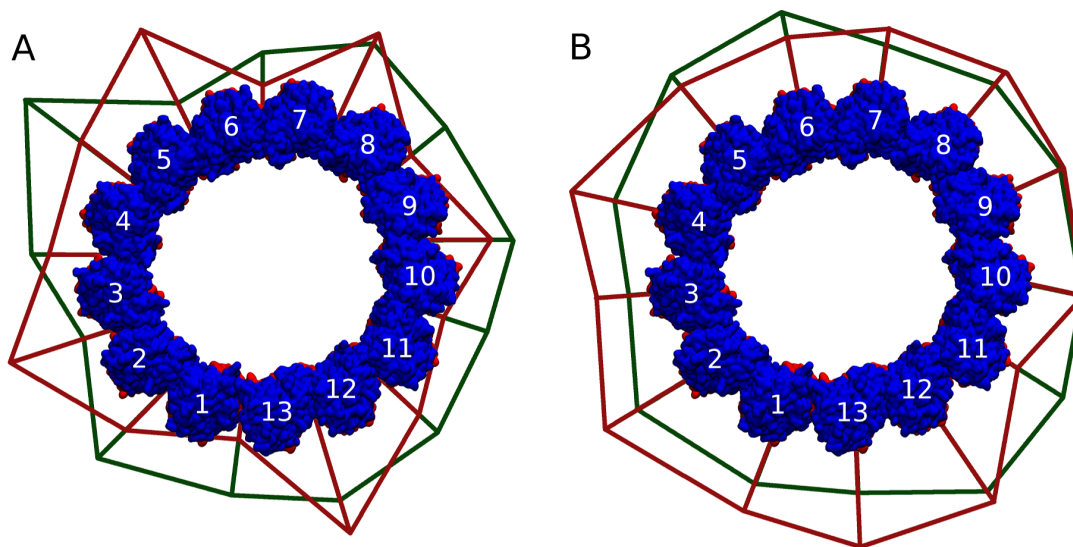
Figure 6.8: *Energy diagrams of the complete MT ring.*
The diagram shows the magnitude of interaction energies at each interface between two tubulin dimers, whether at (A) the lateral interface, or at (B) the longitudinal inter-dimer interface. The magnitude of the interactions is proportional to the swelling at each interface with swellings in (A) being exaggerated to aid viewing. Green represents GTP-Model while red represents GDP-Model.

It is worth mentioning that the deepest of the kinks in the GDP-Model energy diagram, i.e. the interface with the weakest interactions, is the one occurring at the seam (between dimer 13 and dimer 1), in contrast to its strength in the GTP-Model. It has an energy of -9 ± 7 kcal/mol which is very low compared to the one at the interface between dimer 12 and 13, for example, which has an energy of interaction equal to -57 ± 9 kcal/mol. We predict that protofilaments number 1 and 13 having very strong longitudinal contacts antagonized by very weak lateral contacts at the seam, will be the first to dissociate laterally and curve outwards. This should open the MT cylinder which should then trigger disassembly. Therefore, MT energetics suggest that the seam is the most labile part of the MT and could act as a trigger point for disassembly. This is precisely what was reported recently [188].

The energy diagrams at the longitudinal inter-dimer interfaces (Figure 6.8B) appear to be more even than at the lateral interfaces. However, we see no major difference in the pattern between the GTP-Model and the GDP-Model except that longitudinal interactions in the GDP-Model are stronger, which was established earlier.

6.4 Conclusions

We used sophisticated all-atom molecular dynamics simulations to produce accurate MT models, combined with high resolution cryo-electron microscopy maps, to generate an infinite number of infinitely long MT representations. The MM/GBSA energy analysis that followed the simulations enabled an estimate of the contributions of individual residues, domains, subunits and dimers toward the lateral and longitudinal stability of a complete MT ring. We found that longitudinal interactions are about two to three times stronger than lateral interactions explaining the greater stability of the MT structure along its axis than radially. This finding agrees with previous structural observations [165] and computational estimations [164,179]. We also found that interactions are not evenly distributed radially along the longitudinal inter-dimer interface. That is, attractive interactions are largely concentrated away from the MT lumen, producing a force that curls protofilaments outward and eventually causing MT disassembly. The GTP-Model was laterally more stable than the GDP-Model and the opposite was true for the longitudinal inter-dimer interface. Since lateral forces oppose outward curling while longitudinal forces support it, we expect the GTP-Model to be less prone to disassembly than the GDP-Model. With its lateral forces being strong enough to prevent outward curling caused by longitudinal forces, the GTP-cap at the plus end can stabilize an entire MT cylinder. After GTP hydrolysis reaches the cap, lateral forces are too weak to prevent outward curling, especially at the seam which has the weakest lateral contacts. This results in outward curling and microtubule disassembly.

We also confirmed that the MT seam is most likely to act as a trigger point for MT disassembly by being the most labile interface in the MT cylinder [188]. Magnesium ion was demonstrated to be an influential factor in MT stability. Being present at the inner portion of the longitudinal inter-dimer interface, the largely destabilizing Mg^{2+} ion repels the inward portion and enhances outward curling, the formation of ram's horns structures and rapid disassembly, which is consistent with key experimental findings [172]. This action of Mg^{2+} at the E-site of tubulin is suppressed by GTP in GTP-capped MTs. As we showed earlier, the ensemble of Mg^{2+} and GTP at the E-site is collectively stabilizing. However, hydrolysis of GTP and release of inorganic phosphate create a gap at the longitudinal inter-dimer interface and leave the largely destabilizing ensemble of GDP and Mg^{2+} which rapidly promotes outward curling to fill this gap. This happens only at large Mg^{2+} concentrations since GDP at the E-site has low affinity for Mg^{2+} [184]. At low Mg^{2+} concentrations, disassembly becomes slower and outward curling becomes less pronounced [172].

Tangential energy profiles at the longitudinal inter-dimer interface were also shown to be uneven and confirmed the hypothesis that GTP hydrolysis produces a strain which promotes sideway titling [176,177]. However, much of this strain could be tolerated within the lattice constraints and its orthogonality to the direction of outward curling rules out its role in disassembly.

We also identified the most important residues and domains with respect to MT stability at both interfaces and their energetic contributions. At the lateral interface, the α/M-loop, β/M-loop, α/H3 helix, α/N-terminal loop and the α/H2-S3 loop were shown to be most stabilizing while the β/H3 helix was actually destabilizing. This supports predictions based on structural

studies [145,165]. Residue α/Tyr283 was shown to form a very strong network of vdW interactions with neighboring residues and to provide the largest amount of stability at the lateral interface. At the longitudinal inter-dimer interface, the β/C-terminal domain was found to be of paramount importance not only to stability but also to the mechanism of MT disassembly. In particular, residues β/Arg401, β/Phe404, and β/Trp407 of the C-terminal H11 helix and the H11-H11' loop were shown to provide more than 20% of longitudinal stability in both the GTP- and GDP-Models. The complete breakdown of MT energetics per every single residue was further analyzed in order to provide crucial insights into many aspects of MT dynamic instability. Of highest importance is the calculation of the amount of force generated through outward curling due to uneven longitudinal interactions. This could help unravel many aspect of the molecular machinery of cell division, in particular the force generation requirement for chromosome segregation.

Chapter 7

Conclusion and Future Work

In this thesis, I reported the study of MT structure and stability as well as the effect of drugs on it. I carried out a virtual screening study based on similarity of molecular fingerprints to six available MSAs, namely paclitaxel, epothilone A, eleutherobin, discodermolide, sarcodictyin A, and laulimalide. The library was then docked to the taxol binding site and the hits were analyzed for novelty. The most novel hit was optimized visually to increase its binding affinity as well as its pharmacokinetic profile. Rescoring of the binding energies confirmed my predictions about the affinities of the proposed molecules. The proposed molecules were predicted to have much higher affinities to the taxol binding site than the available MSAs, with reasonable pharmacokinetic profiles.

Due to the difficulty in synthesizing these novel hits, I looked for the available molecules which have structures that resemble my novel hits, which turned out to be the antibiotic family of lankacidin. Based on similarity to my novel hits, lankacidin C was hypothesized to bind to the taxol binding site. Docking and MM/PBSA rescoring of the binding of lankacidin C to the taxol binding site confirmed this hypothesis computationally. Together with my collaborators, we made an experimental design to prove this hypothesis using laboratory experiments. Fluorescence quenching experiments showed that both lankacidin C and lankacidinol A affect the conformation of tubulin, specifically that of TUB-B1, in a concentration-dependent manner, thus confirming the binding of lankacidin antibiotics to the β-tubulin subunit of the tubulin dimer. This sheds the light on the unknown antitumor mechanism of action of lankacidin antibiotics and strengthens the possibility that its antitumor action is separate from its antimicrobial action, the latter being associated with interference with protein synthesis. As a future plan, the binding of lankacidin to the taxol binding site will be tested through the displacement of fluorescent taxoid, Flutax-2, which will prove whether lankacidin binds to the taxol binding site of β-tubulin or not. Also, tubulin polymerization assays are now being performed to study the effect of lankacidin on the stability and polymerization rate of microtubules. If our hypothesis is established, lankacidin structure may be further optimized for perfect fit in the binding pocket to improve its affinity as well as pharmacokinetic parameters.

In another direction, I also investigated the stability of MT structure and the thermodynamic

aspects of the binding between tubulin dimers within an MT lattice. I first started by investigating the hydrogen bonds between tubulin dimers. This required the development of a robust and affordable methodology for estimating the hydrogen bond energies. For that purpose, I investigated the utility of quantities like Mulliken overlap population, Wiberg bond index, overlap-weighted bond orbital, and electron density at bond critical point in estimating hydrogen bond energies following a DFT calculation and population analysis. The effect of diffuse functions in the basis sets used as well as the effect of dispersion were also assessed. I determined that the most robust and affordable method for estimating hydrogen bond energies was the use of the descriptor ρ which is the electron density at the bond critical point following a QTAIM analysis. The functional B3LYP as well as the basis set TZVP were chosen as the most suitable level of theory. Linear fitting parameters for each of the four studied descriptors were provided and they could be used directly to estimate hydrogen bond strengths.

The strength of the hydrogen bonds bringing tubulin dimers together laterally and longitudinally was assessed utilizing the aforementioned parameters and descriptors. The overall as well as the pairwise hydrogen bond energies were calculated. The study revealed the importance of hydrogen bonds in general in tubulin energetics. The study showed that hydrogen bonds at the lateral interface are stronger than that at the longitudinal interface although when the overall MM/GBSA energy is considered, the opposite is true. The β-β interactions were comparable to the α-α interactions in the B-lattice in a 95% confidence interval. The study also showed that the stability of the B-lattice configuration is comparable to that of the A-lattice when hydrogen bonds are concerned. This suggests that other energetic contributions could be responsible for the observed difference in predominance between the two lattice forms. Pairwise hydrogen bond energies were in good agreement with experimental data and could be used in many several analysis regarding the energetics of tubulin dimer-dimer interactions.

Finally, a molecular dynamics simulation of a complete MT model was carried out utilizing periodic boundary conditions to simulate an infinite number of infinitely long MT cylinders. The trajectory provided through these simulations were analyzed to calculate the overall as well as per-residue MM/GBSA binding energies between residues at lateral or longitudinal interfaces all along the MT cylinder. The energetic analysis revealed the stability contribution of each residue, domain, subunit, and dimer to the overall stability of an MT cylinder. Since two systems were modeled, GDP- and GTP-Model, the comparison of the two gave insight regarding the role of GTP-hydrolysis in MT stability and disassembly. The study also allowed to propose a detailed explanation of the driving force behind MT disassembly, which is the uneven distribution of binding energy along the longitudinal interface. With the outer portion largely outweighing the inner one, a torque is generated curling protofilaments outward and only opposed by the lateral bonds between protofilaments. Since lateral bonds are stronger in the GTP-Model than in the GDP-Model, outward curling and hence disassembly happens often after GTP hydrolysis and not before, confirming the GTP-cap model.

It would be prudent, however, to carry out a simulation of a free protofilament in order to find

out about the effect of uneven longitudinal energy distribution on the extent of outward curling. By comparing the energy of a free protofilament to the energy of a protofilament constrained within our MT model, we can predict the amount of free energy released by outward curling and additional light could be shed on the mechanism and driving forces in MT disassembly and force generation due to microtubule shortening. In this way, chromosome segregation could be better understood. Also, simulating a GDP-Taxol case would be necessary to understand the molecular mechanisms by which taxol bound to an MT prevents outward curling and MT disassembly. This constitutes a plan for our future work in this project.

Bibliography

[1] B. Stewart and C. P. Wild. *Word Cancer Report*. World Health Organization, 2014.

[2] Defining cancer. `http://www.cancer.gov/cancertopics/what-is-cancer`. Accessed: 17 March 2015.

[3] Metastatic cancer. `http://www.cancer.gov/cancertopics/what-is-cancer/metastatic-fact-sheet`. Accessed: 17 March 2015.

[4] D. Hanahan and R. A. Weinberg. The hallmarks of cancer. *Cell*, 100(1):57 – 70, 2000.

[5] D. Hanahan and R. A. Weinberg. Hallmarks of cancer: The next generation. *Cell*, 144(5):646 – 674, 2011.

[6] The Cell Cycle. `http://en.wikipedia.org/wiki/Mitosis#/media/File:Animal_cell_cycle.svg`. Accessed: 24 March 2015.

[7] N. A. Campbell, J. B. Reece, and L. G. Mitchell. *Biology*. Benjamin Cummings Publ. Co., Inc., Menlo Park, CA, 5th edition, 1999.

[8] J. J. Blow and T. U. Tanaka. The chromosome cycle: coordinating replication and segregation. *EMBO Rep.*, 6(11):1028–1034, 2005.

[9] R. C. Weisenberg. Microtubule formation in vitro in solutions containing low calcium concentrations. *Science*, 177(4054):1104–1105, 1972.

[10] T. J. Mitchison. Localization of an exchangeable GTP binding site at the plus end of microtubules. *Science*, 261(5124):1044–1047, 1993.

[11] P. Dustin. *Microtubules*. Springer-Verlag, New-York, 1984.

[12] R. A. Walker, E. T. O'Brien, N. K. Pryer, M. F. Soboeiro, W. A. Voter, H. P. Erickson, and E. D. Salmon. Dynamic instability of individual microtubules analyzed by video light microscopy: rate constants and transition frequencies. *J. Cell Biol.*, 107(4):1437–1448, 1988.

[13] T. Mitchison and M. Kirschner. Dynamic instability of microtubule growth. *Nature*, 312(5991):237–242, November 1984.

[14] P. B. Schiff and S. B. Horwitz. Taxol stabilizes microtubules in mouse fibroblast cells . *Proc. Natl. Acad. Sci. USA*, 77(3):1561–1565, March 1980.

[15] C. Conde and A. Cáceres. Microtubule assembly, organization and dynamics in axons and dendrites. *Nat. Rev. Neurosci.*, 10(5):319–32, May 2009.

[16] M. F. Carlier and D. Pantaloni. Kinetic analysis of guanosine 5'-triphosphate hydrolysis associated with tubulin polymerization. *Biochemistry*, 20(7):1918–1924, 1981.

[17] C. J. Cramer. *Essentials of Computational Chemistry: Theories and Models*. John Wiley & Sons Ltd. (Canada), 2004.

[18] R. F. W. Bader. *Atoms in Molecules – A Quantum Theory*. Oxford University Press, Oxford, U.K., 1990.

[19] Molecular dynamics tutorial. `http://www.ch.embnet.org/MD_tutorial/`. Accessed: 24 March 2015.

[20] S. A. Adcock and J. A. McCammon. Molecular dynamics: survey of methods for simulating the activity of proteins. *Chem. Rev.*, 106(5):1589–1615, 2006.

[21] M. Karplus and J. A. McCammon. Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.*, 9(9):646–652, 2002.

[22] P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case, and T. E. Cheatham. Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc. Chem. Res.*, 33(12):889–897, October 2000.

[23] The Theory of Atoms in Molecules. `http://www.chemistry.mcmaster.ca/aim/aim_1.html`. Accessed: 16 April 2015.

[24] J. H. Hayden, S. S. Bowser, and C. L. Rieder. Kinetochores capture astral microtubules during chromosome attachment to the mitotic spindle: direct visualization in live newt lung cells. *J. Cell Biol.*, 111(3):1039–1045, 1990.

[25] P. B. Schiff, J. Fant, and S. B. Horwitz. Promotion of microtubule assembly in vitro by taxol. *Nature*, 277(5698):665–667, February 1979.

[26] E. Nogales, S. G. Wolf, K. H. Downing, R. A. Gray, and A. M. Pertsov. Structure of the $\alpha\beta$ tubulin dimer by electron crystallography. *Nature*, 393(6681):191, May 1998.

[27] S. Rao, G. A. Orr, A. G. Chaudhary, D. G. I. Kingston, and S. B. Horwitz. Characterization of the Taxol Binding Site on the Microtubule . *J. Biol. Chem.*, 270(35):20235–20238, September 1995.

[28] S. Rao, L. He, S. Chakravarty, I. Ojima, G. A. Orr, and S. B. Horwitz. Characterization of the Taxol Binding Site on the Microtubule . *J. Biol. Chem.*, 274(53):37990–37994, December 1999.

[29] D. M. Bollag, P. A. McQueney, J. Zhu, O. Hensens, L. Koupal, J. Liesch, M. Goetz, E. Lazarides, and C. M. Woods. Epothilones, a New Class of Microtubule-stabilizing Agents with a Taxol-like Mechanism of Action. *Cancer Res.*, 55(11):2325–2333, June 1995.

[30] E. Hamel, D. L. Sackett, D. Vourloumis, and K. C. Nicolaou. The Coral-Derived Natural Products Eleutherobin and Sarcodictyins A and B: Effects on the Assembly of Purified Tubulin with and without Microtubule-Associated Proteins and Binding at the Polymer Taxoid Site. *Biochemistry*, 38(17):5490–5498, April 1999.

[31] R. J. Kowalski, P. Giannakakou, S. P. Gunasekera, R. E. Longley, B. W. Day, and E. Hamel. The Microtubule-Stabilizing Agent Discodermolide Competitively Inhibits the Binding of Paclitaxel (Taxol) to Tubulin Polymers, Enhances Tubulin Nucleation Reactions More Potently than Paclitaxel, and Inhibits the Growth of Paclitaxel-Resistant Cells. *Mol. Pharmacol.*, 52(4):613–622, October 1997.

[32] R. M. Buey, I. Barasoain, E. Jackson, A. Meyer, P. Giannakakou, I. Paterson, S. Mooberry, J. M. Andreu, and J. F. Díaz. Microtubule interactions with chemically diverse stabilizing agents: thermodynamics of binding to the paclitaxel site predicts cytotoxicity. *Chem. Biol.*, 12(12):1269–79, December 2005.

[33] O. Pineda, J. Farràs, L. Maccari, F. Manetti, M. Botta, and J. Vilarrasa. Computational comparison of microtubule-stabilising agents laulimalide and peloruside with taxol and colchicine. *Bioorg. Med. Chem. Lett.*, 14(19):4825–9, October 2004.

[34] M. Khrapunovich-Baine, V. Menon, C.-P. H. Yang, P. T. Northcote, J. H. Miller, R. H. Angeletti, A. Fiser, S. B. Horwitz, and H. Xiao. Hallmarks of molecular action of microtubule stabilizing agents: effects of epothilone B, ixabepilone, peloruside A, and laulimalide on microtubule conformation. *J. Biol. Chem.*, 286(13):11765–78, April 2011.

[35] M. J. Bennett, K. Barakat, J. T. Huzil, J. Tuszynski, and D. C. Schriemer. Discovery and characterization of the laulimalide-microtubule binding mode by mass shift perturbation mapping. *Chem. Biol.*, 17(7):725–34, July 2010.

[36] G. A. Orr, P. Verdier-Pinard, H. McDaid, and S. B. Horwitz. Mechanisms of Taxol resistance related to microtubules. *Oncogene*, 22(47):7280–95, October 2003.

[37] K. Natarajan and S. Senapati. Understanding the Basis of Drug Resistance of the Mutants of $\alpha\beta$-Tubulin Dimer via Molecular Dynamics Simulations. *PLoS ONE*, 7(8):e42351, January 2012.

[38] A. K. Singla, A. Garg, and D. Aggarwal. Paclitaxel and its formulations. *Int. J. Pharm.*, 235(12):179–192, March 2002.

[39] E. Hopper-Borge, X. Xu, T. Shen, Z. Shi, Z.-S. Chen, and G. D. Kruh. Human multidrug resistance protein 7 (ABCC10) is a resistance factor for nucleoside analogues and epothilone B. *Cancer Res.*, 69(1):178–84, January 2009.

[40] I. Ojima, S. Chakravarty, T. Inoue, S. Lin, L. He, S. B. Horwitz, S. D. Kuduk, and S. J. Danishefsky. A common pharmacophore for cytotoxic natural products that stabilize microtubules. *Proc. Natl. Acad. Sci. USA*, 96(8):4256–61, April 1999.

[41] L. He, P. G. Jagtap, D. G. Kingston, H. J. Shen, G. A. Orr, and S. B. Horwitz. A common pharmacophore for Taxol and the epothilones based on the biological activity of a taxane molecule lacking a C-13 side chain. *Biochemistry*, 39(14):3972–8, April 2000.

[42] P. Giannakakou, R. Gussio, E. Nogales, K. H. Downing, D. Zaharevitz, B. Bollbuck, G. Poy, D. Sackett, K. C. Nicolaou, and T. Fojo. A common pharmacophore for epothilone and taxanes: molecular basis for drug resistance conferred by tubulin mutations in human cancer cells. *Proc. Natl. Acad. Sci. USA*, 97(6):2904–9, March 2000.

[43] N. Nikolova and J. Jaworska. Approaches to Measure Chemical Similarity- a Review. *QSAR Comb. Sci.*, 22(910):1006–1026, December 2003.

[44] C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.*, 46(13):3 – 26, 2001.

[45] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, and A. J. Olson. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.*, 30(16):2785–2791, 2009.

[46] M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek, and G. R. Hutchison. Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminform.*, 4(1):17, August 2012.

[47] D. Case, T. Darden, T. Cheatham III, C. Simmerling, J. Wang, R. Duke, R. Luo, R. Walker, W. Zhang, K. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A. Goetz, I. Kolossváry, K. Wong, F. Paesani, J. Vanicek, R. Wolf, J. Liu, X. Wu, S. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M. Hsieh, G. Cui, D. Roe, D. Mathews, M. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P. Kollman. Amber 12, 2012. University of California, San Francisco.

[48] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, and C. Simmerling. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Bioinf.*, 65(3):712–725, November 2006.

[49] J. H. Nettles, H. Li, B. Cornett, J. M. Krahn, J. P. Snyder, and K. H. Downing. The Binding Mode of Epothilone A on α,β-Tubulin by Electron Crystallography. *Science*, 305(5685):866–869, August 2004.

[50] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case. Development and testing of a general amber force field. *J. Comput. Chem.*, 25(9):1157–74, July 2004.

[51] A. Jakalian, D. B. Jack, and C. I. Bayly. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.*, 23(16):1623–41, December 2002.

[52] W. Humphrey, A. Dalke, and K. Schulten. VMD: Visual molecular dynamics. *J. Mol. Graphics*, 14(1):33 – 38, 1996.

[53] W. Wang and P. A. Kollman. Computational study of protein specificity: The molecular basis of HIV-1 protease drug resistance. *Proc. Natl. Acad. Sci. USA*, 98(26):14937–14942, December 2001.

[54] K. Barakat, J. Mane, D. Friesen, and J. Tuszynski. Ensemble-based virtual screening reveals dual-inhibitors for the p53-MDM2/MDMX interactions. *J. Mol. Graph. Model.*, 28(6):555–68, February 2010.

[55] M. Lepšík, Z. Kříž, and Z. Havlas. Efficiency of a second-generation HIV-1 protease inhibitor studied by molecular dynamics and absolute binding free energy calculations. *Proteins: Struct., Funct., Bioinf.*, 57(2):279–293, November 2004.

[56] S. P. Brown and S. W. Muchmore. High-Throughput Calculation of Protein-Ligand Binding Affinities: Modification and Adaptation of the MM-PBSA Protocol to Enterprise Grid Computing. *J. Chem. Inf. Model.*, 46(3):999–1005, February 2006.

[57] T. Hou and R. Yu. Molecular dynamics and free energy studies on the wild-type and double mutant HIV-1 protease complexed with amprenavir and two amprenavir-related inhibitors: mechanism for binding and drug resistance. *J. Med. Chem.*, 50(6):1177–88, March 2007.

[58] ADMET Predictor, Simulations Plus. www.simulations-plus.com.

[59] D.-S. Su, A. Balog, D. Meng, P. Bertinato, S. J. Danishefsky, Y.-H. Zheng, T.-C. Chou, L. He, and S. B. Horwitz. Structure-activity relationship of the epothilones and the first in vivo comparison with paclitaxel. *Angew. Chem. Int. Edit.*, 36(19):2093–2096, 1997.

[60] S. L. Mooberry, M. K. Hilinski, E. A. Clark, and P. A. Wender. Function-oriented synthesis: Biological evaluation of laulimalide analogues derived from a last step cross metathesis diversification strategy. *Mol. Pharm.*, 5(5):829–838, 2008.

[61] E. T. Lam, S. Goel, L. J. Schaaf, G. F. Cropp, A. L. Hannah, Y. Zhou, B. McCracken, B. I. Haley, R. G. Johnson, S. Mani, and M. A. Villalona-Calero. Phase I dose escalation study of KOS-1584, a novel epothilone, in patients with advanced solid tumors. *Cancer Chemoth. Pharm.*, 69(2):523–531, 2012.

[62] I. H. Hardt, H. Steinmetz, K. Gerth, F. Sasse, H. Reichenbach, and G. Höfle. New natural epothilones from sorangium cellulosum, strains So ce90/b2 and So ce90/d13: isolation, structure elucidation, and SAR studies. *J. Nat. Prod.*, 64(7):847–856, 2001.

[63] M. Sefkow, M. Kiffe, and G. Höfle. Derivatization of the C12C13 functional groups of epothilones A, B and C. *Bioorg. Med. Chem. Lett.*, 8(21):3031–3036, November 1998.

[64] M. Namikoshi, K. Akano, S. Meguro, I. Kasuga, Y. Mine, T. Takahashi, and H. Kobayashi. A New Macrocyclic Trichothecene, 12,13-Deoxyroridin E, Produced by the Marine-Derived Fungus Myrothecium roridum Collected in Palau. *J. Nat. Prod.*, 64(3):396–398, February 2001.

[65] R. M. Buey, J. F. Díaz, J. M. Andreu, A. O'Brate, P. Giannakakou, K. C. Nicolaou, P. K. Sasmal, A. Ritzén, and K. Namoto. Interaction of Epothilone Analogs with the Paclitaxel Binding Site: Relationship between Binding Affinity, Microtubule Stabilization, and Cytotoxicity. *Chem. Biol.*, 11(2):225–236, February 2004.

[66] S. Harada, E. Higashide, T. Fugono, and T. Kishi. Isolation and structures of T-2636 antibiotics. *Tetrahedron Lett.*, 10(27):2239 – 2244, 1969.

[67] S. Harada, T. Kishi, and K. Mizuno. Studies on T-2636 antibiotics. II. Isolation and chemical properties of T-2636 antibiotics. *J. Antibiot.*, 24(1):13 – 22, 1971.

[68] S. Harada, S. Tanayama, and T. Kishi. Studies on lankacidin-group (T-2636) antibiotics. 8. Metabolism of lankacidin c 14-propionate in rats and mice. *J. Antibiot.*, 26(11):658 – 668, 1973.

[69] S. Harada, T. Yamazaki, K. Hatano, K. Tsuchiya, and T. Kishi. Studies on lankacidin-group (T-2636) antibiotics. VII. Structure-activity relationships of lankacidin-group antibiotics. *J. Antibiot.*, 26(11):647 – 657, 1973.

[70] S. Harada and T. Kishi. Studies on lankacidin-group (T-2636) antibiotics. V. Chemical structures of lankacidin-group antibiotics. 1. *Chem. Pharm. Bull.*, 22(1):99 – 108, 1974.

[71] K. Tsuchiya, T. Yamazaki, Y. Takeuchi, and T. Oishi. Studies on T-2636 antibiotics. IV. In vitro and in vivo antibacterial activity of T-2636 antibiotics. *J. Antibiot.*, 24(1):29 – 41, 1971.

[72] K. Ootsu and T. Matsumoto. Effects of lankacidin group (T2636) antibiotics on the tumor growth and immune response against sheep erythrocytes in mice. *Gann.*, 64(5):481 – 492, 1973.

[73] K. Ootsu, T. Matsumoto, S. Harada, and T. Kishi. Antitumor and immunosuppressive activities of lankacidin-group antibiotics: structure-activity relationships. *Cancer Chemother. Rep.*, 59(5):919 – 928, 1975.

[74] A. S. Kende, K. Liu, I. Kaldor, G. Dorey, and K. Koch. Total synthesis of the macrolide antitumor antibiotic lankacidin c. *J. Am. Chem. Soc.*, 117(31):8258–8270, 1995.

[75] T. Auerbach, I. Mermershtain, C. Davidovich, A. Bashan, M. Belousoff, I. Wekselman, E. Zimmerman, L. Xiong, D. Klepacki, K. Arakawa, H. Kinashi, A. S. Mankin, and A. Yonath. The structure of ribosome-lankacidin complex reveals ribosomal sites for synergistic antibiotics. *Proc. Natl. Acad. Sci. U.S.A.*, 107(5):1983–1988, 2010.

[76] M. J. Belousoff, T. Shapira, A. Bashan, E. Zimmerman, H. Rozenberg, K. Arakawa, H. Kinashi, and A. Yonath. Crystal structure of the synergistic antibiotic pair, lankamycin and lankacidin, in complex with the large ribosomal subunit. *Proc. Natl. Acad. Sci. U.S.A.*, 108(7):2717–2722, 2011.

[77] A. T. Ayoub, M. Klobukowski, and J. Tuszynski. Similarity-based virtual screening for microtubule stabilizers reveals novel antimitotic scaffold. *J. Mol. Graph. Model.*, 44(0):188 – 196, 2013.

[78] K. L. Meagher, L. T. Redman, and H. A. Carlson. Development of polyphosphate parameters for use with the AMBER force field. *J. Comput. Chem.*, 24(9):1016–1025, 2003.

[79] H. Kinashi, E. Mori, A. Hatani, and O. Nimi. Isolation and characterization of linear plasmids from lankacidin-producing streptomyces species. *J. Antibiot.*, 47(12):1447 – 1455, 1994.

[80] K. Arakawa, F. Sugino, K. Kodama, T. Ishii, and H. Kinashi. Cyclization mechanism for the synthesis of macrocyclic antibiotic lankacidin in streptomyces rochei. *Chem. Biol.*, 12(2):249 – 256, 2005.

[81] E. Arunan, G. R. Desiraju, R. A. Klein, J. Sadlej, S. Scheiner, I. Alkorta, D. C. Clary, R. H. Crabtree, J. J. Dannenberg, P. Hobza, H. G. Kjaergaard, A. C. Legon, B. Mennucci, and D. J. Nesbitt. Definition of the hydrogen bond (IUPAC Recommendations 2011). *Pure Appl. Chem.*, 83(8):1637–1641, July 2011.

[82] E. Arunan, G. R. Desiraju, R. A. Klein, J. Sadlej, S. Scheiner, I. Alkorta, D. C. Clary, R. H. Crabtree, J. J. Dannenberg, P. Hobza, H. G. Kjaergaard, A. C. Legon, B. Mennucci, and D. J. Nesbitt. Defining the hydrogen bond: An account (IUPAC Technical Report). *Pure Appl. Chem.*, 83(8):1619–1636, July 2011.

[83] G. R. Desiraju. Hydrogen bridges in crystal engineering: Interactions without borders. *Accounts Chem. Res.*, 35(7):565–573, 2002.

[84] S. Deechongkit, H. Nguyen, E. T. Powers, P. E. Dawson, M. Gruebele, and J. W. Kelly. Context-dependent contributions of backbone hydrogen bonding to β-sheet folding energetics. *Nature*, 430(6995):101–105, July 2004.

[85] G. D. Rose, P. J. Fleming, J. R. Banavar, and A. Maritan. A backbone-based theory of protein folding. *Proc. Natl. Acad. Sci. U.S.A.*, 103(45):16623–16633, 2006.

[86] M. Hellgren, C. Kaiser, S. Haij, Å. Norberg, and J.-O. Höög. A hydrogen-bonding network in mammalian sorbitol dehydrogenase stabilizes the tetrameric state and is essential for the catalytic power. *Cell Mol. Life Sci.*, 64(23):3129–3138, 2007.

[87] E. R. Davidson. Electronic population analysis of molecular wavefunctions. *J. Chem. Phys.*, 46(9):3320–3324, 1967.

[88] K. R. Roby. Quantum theory of chemical valence concepts. *Mol. Phys.*, 27(1):81–104, 1974.

[89] C. Ehrhardt and R. Ahlrichs. Population analysis based on occupation numbers II. relationship between shared electron numbers and bond energies and characterization of hypervalent contributions. *Theor. Chim. Acta*, 68(3):231–245, 1985.

[90] M. Reiher, D. Sellmann, and B. A. Hess. Stabilization of diazene in fe(II)–sulfur model complexes relevant for nitrogenase activity. I. a new approach to the evaluation of intramolecular hydrogen bond energies. *Theor. Chem. Acc.*, 106(6):379–392, 2001.

[91] J. Thar and B. Kirchner. Hydrogen bond detection. *J. Phys. Chem. A*, 110(12):4229–4237, 2006.

[92] M. Schmidt, S. Zahn, M. Carella, O. Ohlenschläger, M. Görlach, E. Kothe, and J. Weston. Solution structure of a functional biomimetic and mechanistic implications for nickel superoxide dismutases. *ChemBioChem*, 9(13):2135–2146, 2008.

[93] S. Schenk, B. Le Guennic, B. Kirchner, and M. Reiher. First-principles investigation of the schrock mechanism of dinitrogen reduction employing the full HIPTN$_3$N ligand. *Inorg. Chem.*, 47(9):3634–3650, 2008.

[94] S. J. Grabowski. Ab initio calculations on conventional and unconventional hydrogen bonds – study of the hydrogen bond strength. *J. Phys. Chem. A*, 105(47):10739–10746, 2001.

[95] R. W. Gora, S. J. Grabowski, and J. Leszczynski. Dimers of formic acid, acetic acid, formamide and pyrrole-2-carboxylic acid: An ab initio study. *J. Phys. Chem. A*, 109(29):6397–6405, 2005.

[96] R. Parthasarathi, V. Subramanian, and N. Sathyamurthy. Hydrogen bonding in phenol, water, and phenol–water clusters. *J. Phys. Chem. A*, 109(5):843–850, 2005.

[97] R. Parthasarathi, V. Subramanian, and N. Sathyamurthy. Hydrogen bonding without borders: An atoms-in-molecules perspective. *J. Phys. Chem. A*, 110(10):3349–3351, 2006.

[98] C. Gatti, V. R. Saunders, and C. Roetti. Crystal field effects on the topological properties of the electron density in molecular crystals: The case of urea. *J. Chem. Phys.*, 101(12):10686–10696, 1994.

[99] R. F. W. Bader. A quantum theory of molecular structure and its applications. *Chem. Rev.*, 91(5):893–928, 1991.

[100] E. D. Glendening, C. R. Landis, and F. Weinhold. Natural bond orbital methods. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2(1):1–42, January 2012.

[101] R. S. Mulliken. Electronic population analysis on LCAO–MO molecular wave functions. *J. Chem. Phys.*, 23(10):1833–1840, 1955.

[102] S. Grabowski. BeH$_2$ as a proton–accepting molecule for dihydrogen bonded systems–ab initio study. *J. Mol. Struct.*, 553(1–3):151–156, 2000.

[103] I. Alkorta and J. Elguero. Non–conventional hydrogen bonds. *Chem. Soc. Rev.*, 27:163–170, 1998.

[104] K. Wiberg. Application of the Pople–Santry–Segal {CNDO} method to the cyclopropyl-carbinyl and cyclobutyl cation and to bicyclobutane. *Tetrahedron*, 24(3):1083–1096, 1968.

[105] A. E. Reed and F. Weinhold. Natural bond orbital analysis of near–Hartree–Fock water dimer. *J. Chem. Phys.*, 78(6):4066–4073, 1983.

[106] A. E. Reed, R. B. Weinstock, and F. Weinhold. Natural population analysis. *J. Chem. Phys.*, 83(2):735–746, 1985.

[107] S. Scheiner. *Calculating the Properties of Hydrogen Bonds by ab Initio Methods*, pages 165–218. John Wiley & Sons, Inc., 2007.

[108] B. Jeziorski and K. Szalewicz. *Intermolecular Interactions by Perturbation Theory*, pages 1376–1398. John Wiley & Sons, Ltd, 2002.

[109] A. D. Becke. Density–functional thermochemistry. III. the role of exact exchange. *J. Chem. Phys.*, 98(7):5648–5652, 1993.

[110] C. Lee, W. Yang, and R. G. Parr. Development of the Colle–Salvetti correlation–energy formula into a functional of the electron density. *Phys. Rev. B*, 37:785–789, Jan 1988.

[111] S. H. Vosko, L. Wilk, and M. Nusair. Accurate spin–dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.*, 58(8):1200–1211, 1980.

[112] A. Schäfer, H. Horn, and R. Ahlrichs. Fully optimized contracted gaussian basis sets for atoms Li to Kr. *J. Chem. Phys.*, 97(4):2571–2577, 1992.

[113] A. Schäfer, C. Huber, and R. Ahlrichs. Fully optimized contracted gaussian basis sets of triple zeta valence quality for atoms Li to Kr. *J. Chem. Phys.*, 100(8):5829–5835, 1994.

[114] S. Boys and F. Bernardi. The calculation of small molecular interactions by the differences of separate total energies. some procedures with reduced errors. *Mol. Phys.*, 19(4):553–566, 1970.

[115] S. Simon, M. Duran, and J. J. Dannenberg. How does basis set superposition error change the potential surfaces for hydrogen–bonded dimers? *J. Chem. Phys.*, 105(24):11024–11031, 1996.

[116] S. Grimme. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.*, 27(15):1787–1799, 2006.

[117] J.-D. Chai and M. Head-Gordon. Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.*, 128(8):–, 2008.

[118] J.-D. Chai and M. Head-Gordon. Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Phys. Chem. Chem. Phys.*, 10:6615–6620, 2008.

[119] C. Møller and M. S. Plesset. Note on an approximation treatment for many-electron systems. *Phys. Rev.*, 46:618–622, Oct 1934.

[120] M. Head-Gordon, J. A. Pople, and M. J. Frisch. MP2 energy evaluation by direct methods. *Chem. Phys. Lett.*, 153(6):503 – 506, 1988.

[121] M. J. Frisch, M. Head-Gordon, and J. A. Pople. A direct MP2 gradient method. *Chem. Phys. Lett.*, 166(3):275 – 280, 1990.

[122] T. H. Dunning. Gaussian basis sets for use in correlated molecular calculations. I. the atoms boron through neon and hydrogen. *J. Chem. Phys.*, 90(2), 1989.

[123] R. A. Kendall, T. H. Dunning, and R. J. Harrison. Electron affinities of the firstrow atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.*, 96(9), 1992.

[124] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox. Gaussian 09 Revision D.01. Gaussian Inc. Wallingford CT 2009.

[125] A. T. Ayoub, J. Tuszynski, and M. Klobukowski. Estimating hydrogen bond energies: comparison of methods. *Theor. Chem. Acc.*, 133(8):1520–1526, 2014.

[126] A. T. Ayoub, T. J. Craddock, M. Klobukowski, and J. Tuszynski. Analysis of the strength of interfacial hydrogen bonds between tubulin dimers using quantum theory of atoms in molecules. *Biophys. J.*, 107(3):740–750, December 2014.

[127] J. R. McIntosh, M. K. Morphew, P. M. Grissom, S. P. Gilbert, and A. Hoenger. Lattice structure of cytoplasmic microtubules in a cultured mammalian cell. *J. Mol. Biol.*, 394(2):177–182, 2009.

[128] E. M. Mandelkow, R. Schultheiss, R. Rapp, M. Müller, and E. Mandelkow. On the surface lattice of microtubules: helix starts, protofilament number, seam, and handedness. *J. Cell. Biol.*, 102(3):1067–1073, 1986.

[129] C. Cohen, D. DeRosier, S. C. Harrison, R. E. Stephens, and J. Thomas. X-ray patterns from microtubules. *Ann. N. Y. Acad. Sci.*, 253(1):53–59, 1975.

[130] E.-M. Mandelkow, E. Mandelkow, N. Unwin, and C. Cohen. Tubulin hoops. *Nature*, 265(5595):655–657, February 1977.

[131] C. Wais-Steider, N. White, D. Gilbert, and P. Eagles. X-ray diffraction patterns from microtubules and neurofilaments in axoplasm. *J. Mol. Biol.*, 197(2):205–218, 1987.

[132] E. Mandelkow, Y. H. Song, and E. M. Mandelkow. The microtubule lattice–dynamic instability of concepts. *Trends Cell Biol.*, 5(7):262–266, July 1995.

[133] Y. H. Song and E. Mandelkow. Recombinant kinesin motor domain binds to beta-tubulin and decorates microtubules with a b surface lattice. *Proc. Natl. Acad. Sci. U.S.A.*, 90(5):1671–1675, 1993.

[134] A. des Georges, M. Katsuki, D. R. Drummond, M. Osei, R. A. Cross, and L. A. Amos. Mal3, the Schizosaccharomyces pombe homolog of EB1, changes the microtubule lattice. *Nat. Struct. Mol. Biol.*, 15(10):1102–1108, October 2008.

[135] B. Vitre, F. M. Coquelle, C. Heichette, C. Garnier, D. Chrétien, and I. Arnal. EB1 regulates microtubule dynamics and tubulin sheet closure in vitro. *Nat. Cell Biol.*, 10(4):415–421, April 2008.

[136] D. Sept, N. A. Baker, and J. A. McCammon. The physical basis of microtubule structure and stability. *Protein Sci.*, 12(10):2257–2261, 2003.

[137] P. Drabik, S. Gusarov, and A. Kovalenko. Microtubule stability studied by three-dimensional molecular theory of solvation. *Biophys. J.*, 92(2):394–403, January 2007.

[138] H. P. Erickson and D. Pantaloni. The role of subunit entropy in cooperative assembly. nucleation of microtubules and other two–dimensional polymers. *Biophys. J.*, 34(2):293–309, 1981.

[139] P. L. A. Popelier. Characterization of a dihydrogen bond on the basis of the electron density. *J. Phys. Chem. A*, 102(10):1873–1878, 1998.

[140] S. Scheiner, S. J. Grabowski, and T. Kar. Influence of hybridization and substitution on the properties of the cho hydrogen bond. *J. Phys. Chem. A*, 105(46):10607–10612, 2001.

[141] F. C. Bernstein, T. F. Koetzle, G. J. Williams, E. F. M. Jr., M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi. The protein data bank: A computer-based archival file for macromolecular structures. *Arch. Biochem. Biophys.*, 185(2):584–591, 1978.

[142] J. Löwe, H. Li, K. Downing, and E. Nogales. Refined structure of $\alpha\beta$-tubulin at 3.5 Å resolution. *J. Mol. Biol.*, 313(5):1045–1057, 2001.

[143] A. Šali and T. L. Blundell. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.*, 234(3):779–815, 1993.

[144] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten. Scalable molecular dynamics with NAMD. *J. Comput. Chem.*, 26(16):1781–1802, 2005.

[145] H. Li, D. J. DeRosier, W. V. Nicholson, E. Nogales, and K. H. Downing. Microtubule structure at 8 Å resolution. *Structure*, 10(10):1317–1328, 2002.

[146] W. L. DeLano. The PyMOL molecular graphics system. San Carlos, CA: Delano Scientific, 2002.

[147] D. B. Wells and A. Aksimentiev. Mechanical properties of a complete microtubule revealed through molecular dynamics simulation. *Biophys. J.*, 99(2):629 – 637, 2010.

[148] N. Guex and M. C. Peitsch. SWISS-MODEL and the Swiss-PDB Viewer: An environment for comparative protein modeling. *Electrophoresis*, 18(15):2714–2723, 1997.

[149] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.*, 117(19):5179–5197, 1995.

[150] H. Li, A. D. Robertson, and J. H. Jensen. Very fast empirical prediction and rationalization of protein pKa values. *Proteins: Struct., Funct., Bioinf.*, 61(4):704–721, 2005.

[151] D. C. Bas, D. M. Rogers, and J. H. Jensen. Very fast prediction and rationalization of pKa values for protein–ligand complexes. *Proteins: Struct., Funct., Bioinf.*, 73(3):765–783, 2008.

[152] M. H. M. Olsson, C. R. Søndergaard, M. Rostkowski, and J. H. Jensen. PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions. *J. Chem. Theory Comput.*, 7(2):525–537, 2011.

[153] C. R. Søndergaard, M. H. M. Olsson, M. Rostkowski, and J. H. Jensen. Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pKa values. *J. Chem. Theory Comput.*, 7(7):2284–2295, 2011.

[154] R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand, and R. C. Walker. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent Particle Mesh Ewald. *J. Chem. Theory Comput.*, 9(9):3878–3888, 2013.

[155] A. W. Götz, M. J. Williamson, D. Xu, D. Poole, S. Le Grand, and R. C. Walker. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized Born. *J. Chem. Theory Comput.*, 8(5):1542–1555, 2012.

[156] S. L. Grand, A. W. Götz, and R. C. Walker. SPFP: Speed without compromise - a mixed precision model for GPU accelerated molecular dynamics simulations. *Comput. Phys. Commun.*, 184(2):374 – 380, 2013.

[157] J. Shao, S. W. Tanner, N. Thompson, and T. E. Cheatham. Clustering molecular dynamics trajectories: 1. characterizing the performance of different clustering algorithms. *J. Chem. Theory Comput.*, 3(6):2312–2334, 2007.

[158] D. R. Roe and T. E. Cheatham. PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.*, 9(7):3084–3095, 2013.

[159] K. Wendler, J. Thar, S. Zahn, and B. Kirchner. Estimating the hydrogen bond energy. *J. Phys. Chem. A*, 114(35):9529–9536, 2010.

[160] R. F. W. Bader. *AIMPAC, Suite of programs for the Theory of Atoms in Molecules*. McMaster University, Hamilton, Ontario, Canada, 1994.

[161] B. Gigant, C. Wang, R. B. G. Ravelli, F. Roussi, M. O. Steinmetz, P. A. Curmi, A. Sobel, and M. Knossow. Structural basis for the regulation of tubulin by vinblastine. *Nature*, 435(7041):519–522, May 2005.

[162] R. J. Toso, M. A. Jordan, K. W. Farrell, B. Matsumoto, and L. Wilson. Kinetic stabilization of microtubule dynamic instability in vitro by vinblastine. *Biochemistry*, 32(5):1285–1293, 1993.

[163] J. T. Huzil, R. F. Ludueña, and J. Tuszynski. Comparative modelling of human β tubulin isotypes and implications for drug binding. *Nanotechnology*, 17(4):S90–S100, 2006.

[164] V. VanBuren, D. J. Odde, and L. Cassimeris. Estimates of lateral and longitudinal bond energies within the microtubule lattice. *Proc. Natl. Acad. Sci. U.S.A.*, 99(9):6035–6040, 2002.

[165] E. Nogales, M. Whittaker, R. A. Milligan, and K. H. Downing. High-resolution model of the microtubule. *Cell*, 96(1):79–88, 1999.

[166] J. P. Snyder, J. H. Nettles, B. Cornett, K. H. Downing, and E. Nogales. The binding conformation of taxol in β-tubulin: A model based on electron crystallographic density. *Proc. Natl. Acad. Sci. U.S.A.*, 98(9):5312–5316, 2001.

[167] L. A. Amos and J. Löwe. How Taxol stabilizes microtubule structure. *Chem. Biol.*, 6(3):R65–R69, March 1999.

[168] A. E. Prota, K. Bargsten, D. Zurwerra, J. J. Field, J. F. Díaz, K.-H. Altmann, and M. O. Steinmetz. Molecular mechanism of action of microtubule-stabilizing anticancer agents. *Science*, 339(6119):587–590, 2013.

[169] M. Sirajuddin, L. M. Rice, and R. D. Vale. Regulation of microtubule motors by tubulin isotypes and post-translational modifications. *Nat. Cell Biol.*, 16(4):335–344, April 2014.

[170] J. S. Hyams and C. W. Lloyd. *Microtubules*. New York: Wiley-Liss, 1993.

[171] C. Dumontet and M. A. Jordan. Microtubule-binding agents: a dynamic field of cancer therapeutics. *Nat. Rev. Drug Discov.*, 9(10):790–803, October 2010.

[172] E. M. Mandelkow, E. Mandelkow, and R. A. Milligan. Microtubule dynamics and microtubule caps: a time-resolved cryo-electron microscopy study. *J. Cell Biol.*, 114(5):977–991, 1991.

[173] A. A. Hyman, D. Chrétien, I. Arnal, and R. H. Wade. Structural changes accompanying GTP hydrolysis in microtubules: information from a slowly hydrolyzable analogue guanylyl-(alpha,beta)-methylene-diphosphonate. *J. Cell Biol.*, 128(1):117–125, 1995.

[174] E. Nogales and H.-W. Wang. Structural mechanisms underlying nucleotide-dependent self-assembly of tubulin and its relatives. *Curr. Opin. Struc. Biol.*, 16(2):221 – 229, 2006.

[175] P. Barbier, A. Dorléans, F. Devred, L. Sanz, D. Allegro, C. Alfonso, M. Knossow, V. Peyrot, and J. M. Andreu. Stathmin and interfacial microtubule inhibitors recognize a naturally curved conformation of tubulin dimers. *J. Biol. Chem.*, 285(41):31672–31681, 2010.

[176] G. M. Alushin, G. C. Lander, E. H. Kellogg, R. Zhang, D. Baker, and E. Nogales. High-resolution microtubule structures reveal the structural transitions in $\alpha\beta$-tubulin upon GTP hydrolysis. *Cell*, 157(5):1117–1129, December 2014.

[177] T. J. Mitchison. The engine of microtubule dynamics comes into focus. *Cell*, 157(5):1008 – 1010, 2014.

[178] J. R. André, M.-J. Clément, E. Adjadj, F. Toma, P. A. Curmi, and P. Manivet. The state of the guanosine nucleotide allosterically affects the interfaces of tubulin in protofilament. *J. Comput. Aid. Mol. Des.*, 26(4):397–407, 2012.

[179] O. Kononova, Y. Kholodov, K. E. Theisen, K. A. Marx, R. I. Dima, F. I. Ataullakhanov, E. L. Grishchuk, and V. Barsegov. Tubulin bond energies and microtubule biomechanics determined from nanoindentation in silico. *J. Am. Chem. Soc.*, 136(49):17036–17045, 2014.

[180] Molecular Operating Environment (MOE), 2012. Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7.

[181] B. R. Miller, T. D. McGee, J. M. Swails, N. Homeyer, H. Gohlke, and A. E. Roitberg. MMPBSA.py: An efficient program for end-state free energy calculations. *J. Chem. Theory Comput.*, 8(9):3314–3321, 2012.

[182] T. Hou, J. Wang, Y. Li, and W. Wang. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J. Chem. Inf. Model.*, 51(1):69–82, January 2011.

[183] A. T. Ayoub, M. Klobukowski, and J. A. Tuszynski. Detailed Per-residue Energetic Analysis Explains the Driving Force for Microtubule Disassembly. *PLoS Comput. Biol.*, 11(6):e1004313, June 2015.

[184] J. J. Correia, L. T. Baty, and R. C. Williams. Mg^{2+} dependence of guanine nucleotide binding to tubulin. *J. Biol. Chem.*, 262(36):17278–17284, 1987.

[185] R. P. Frigon and S. N. Timasheff. Magnesium-induced self-association of calf brain tubulin. II. Thermodynamics. *Biochemistry*, 14(21):4567–4573, 1975.

[186] V. Gal, S. Martin, and P. Bayley. Fast disassembly of microtubules induced by Mg^{2+} or Ca^{2+}. *Biochem. Bioph. Res. Co.*, 155(3):1464 – 1470, 1988.

[187] E. T. O'Brien, E. D. Salmon, R. A. Walker, and H. P. Erickson. Effects of magnesium on the dynamic instability of individual microtubules. *Biochemistry*, 29(28):6648–6656, 1990.

[188] M. Katsuki, D. R. Drummond, and R. A. Cross. Ectopic A-lattice seams destabilize microtubules. *Nat. Commun.*, 5:3094, January 2014.

Appendix

Table S1: *Details of the binding energies and quantum mechanical descriptor values for the 45 complex systems studied with B3LYP/TZVP level of theory.*

Complex	MOC	ρ	OWBO	WBI	E_I (kJ/mol)
HSH–FNF2	0.0043	0.0037	0.0011	0.0017	–0.10
ClH–FOF	0.0046	0.0070	–0.0005	0.0058	–0.19
BrH–FNF2	0.0056	0.0055	0.0023	0.0035	–0.69
H3CH–OMe2	0.0093	0.0061	0.0015	0.0023	–1.20
H3CH–NHMe2	0.0061	0.0066	–0.0020	0.0053	–1.51
MeHPH–NH3	–0.0093	0.0082	0.0164	0.0070	–1.84
HSH–BrH	0.0065	0.0051	0.0042	0.0054	–1.98
HOPh–FMe	0.0121	0.0062	0.0032	0.0018	–2.55
H2PH–OH2	–0.0008	0.0080	0.0059	0.0037	–2.64
FH–FNF2	0.0108	0.0131	0.0196	0.0092	–3.80
BrH–BrH	0.0173	0.0096	0.0377	0.0215	–4.11
FH–FOF	0.0128	0.0150	0.0142	0.0118	–4.75
FH–ClOCl	0.0294	0.0143	0.0073	0.0249	–5.55
HSH–PH2Me	0.0002	0.0096	0.0119	0.0207	–5.87
MeSH–NH2Ph	0.0109	0.0153	0.0147	0.0202	–8.00
MeOH–PH3	–0.0087	0.0125	0.0142	0.0175	–9.00
MeSH–OHMe	0.0052	0.0168	0.0174	0.0157	–9.55
MeHNH–OMe2	0.0179	0.0148	0.0106	0.0077	–10.67
MeHNH–NMe3	0.0263	0.0175	0.0155	0.0150	–10.87
MeHNH–NH3	–0.0254	0.0171	0.0202	0.0167	–12.53
ClH–SH2	0.0389	0.0187	0.0469	0.0569	–13.78
PhOH–PH2Me	0.0082	0.0160	0.0308	0.0375	–16.17
HSH–NH2Me	0.0072	0.0259	0.0676	0.0506	–16.99
MeOH–PMe3	0.0245	0.0181	0.0199	0.0464	–17.40
ClH–PH2Me	0.0363	0.0219	0.1173	0.0851	–18.52
FH–FH	0.0205	0.0263	0.0305	0.0248	–20.38
BrH–OH2	0.0137	0.0303	0.0701	0.0488	–21.67
MeOH–OH2	0.0225	0.0259	0.0364	0.0243	–21.86
HOH–OH2	0.0168	0.0256	0.0369	0.0243	–22.60
MeOH–OHMe	0.0267	0.0281	0.0400	0.0284	–22.62
FH–SH2	0.0559	0.0245	0.0491	0.0627	–22.68
HOH–OHEt	0.0224	0.0276	0.0398	0.0282	–23.17
FH–OHCl	0.0186	0.0345	0.0566	0.0393	–24.20
ClH–OH2	0.0190	0.0332	0.0621	0.0511	–26.32
HOH–NMe3	0.0454	0.0353	0.0656	0.0483	–28.34
HOH–NH3	–0.0065	0.0297	0.0550	0.0452	–28.92
MeOH–NH3	–0.0131	0.0307	0.0582	0.0476	–28.96
FH–PH2Me	0.0553	0.0277	0.0773	0.0890	–29.35
HOH–NHMe2	0.0339	0.0347	0.0655	0.0535	–29.71
HOH–NH2Me	0.0165	0.0328	0.0618	0.0525	–30.03
MeOH–NH2Me	0.0068	0.0336	0.0685	0.0543	–30.11
FH–OH2	0.0320	0.0434	0.0790	0.0612	–40.66
FH–OHMe	0.0469	0.0496	0.0945	0.0754	–44.21
ClH–NH3	0.0281	0.0540	0.1737	0.1383	–44.89
FH–OMe2	0.0536	0.0531	0.1034	0.0800	–45.03

Table S2: *Details of the binding energies and quantum mechanical descriptor values for the 45 complex systems studied with B3LYP/aug–TZVP level of theory.*

Complex	MOC	ρ	OWBO	WBI	E_I (kJ/mol)
HSH-FNF2	-0.002	0.00372	0.0031	0.0016	0.0
ClH-FOF	-0.005	0.00702	-0.0040	0.0079	0.0
BrH-FNF2	0.001	0.00551	0.0064	0.0037	-0.5
H3CH-OMe2	-0.047	0.00600	-0.0031	0.0039	-1.1
H3CH-NHMe2	-0.037	0.00649	-0.0098	0.0076	-1.3
MeHPH-NH3	-0.057	0.00784	0.0188	0.0071	-1.3
HSH-BrH	0.002	0.00511	0.0035	0.0048	-1.9
HOPh-FMe	-0.031	0.00612	0.0013	0.0025	-2.5
H2PH-OH2	-0.056	0.00774	0.0079	0.0041	-2.3
FH-FNF2	0.013	0.01307	0.0074	0.0084	-3.7
BrH-BrH	0.016	0.00966	0.0132	0.0213	-4.0
FH-FOF	0.008	0.01499	0.0092	0.0118	-4.6
FH-ClOCl	0.040	0.01436	0.0055	0.0232	-5.4
HSH-PH2Me	-0.030	0.00957	0.0079	0.0187	-5.8
MeSH-NH2Ph	-0.092	0.01518	0.0188	0.0213	-7.8
MeOH-PH3	0.014	0.01254	0.0116	0.0188	-8.9
MeSH-OHMe	-0.053	0.01664	0.0175	0.0164	-9.1
MeHNH-OMe2	-0.090	0.01477	0.0099	0.0086	-10.5
MeHNH-NMe3	-0.162	0.01747	0.0368	0.0212	-10.8
MeHNH-NH3	-0.156	0.01668	0.0189	0.0169	-11.4
ClH-SH2	0.021	0.01866	0.0756	0.0584	-13.5
PhOH-PH2Me	0.087	0.01607	0.0245	0.0390	-16.0
HSH-NH2Me	-0.098	0.02569	0.0750	0.0513	-16.4
MeOH-PMe3	0.029	0.01812	0.0871	0.0606	-17.6
ClH-PH2Me	0.026	0.02199	0.1178	0.0836	-18.7
FH-FH	-0.021	0.02606	0.0289	0.0233	-19.3
BrH-OH2	-0.089	0.03005	0.0787	0.0503	-20.0
MeOH-OH2	-0.130	0.02558	0.0346	0.0246	-20.3
HOH-OH2	-0.135	0.02530	0.0349	0.0237	-21.7
MeOH-OHMe	-0.124	0.02795	0.0407	0.0284	-21.9
FH-SH2	0.043	0.02452	0.0445	0.0632	-22.9
HOH-OHEt	-0.106	0.02753	0.0346	0.0276	-23.4
FH-OHCl	-0.128	0.03439	0.0499	0.0373	-23.9
ClH-OH2	-0.090	0.03304	0.0839	0.0522	-24.9
HOH-NMe3	-0.184	0.03527	0.0848	0.0566	-29.1
HOH-NH3	-0.150	0.02949	0.0527	0.0453	-28.5
MeOH-NH3	-0.151	0.03050	0.0614	0.0494	-27.9
FH-PH2Me	0.043	0.02782	0.0717	0.0888	-30.1
HOH-NHMe2	-0.152	0.03472	0.0569	0.0536	-30.2
HOH-NH2Me	-0.139	0.03265	0.0556	0.0521	-30.1
MeOH-NH2Me	-0.130	0.03350	0.0667	0.0551	-29.6
FH-OH2	-0.135	0.04334	0.0790	0.0606	-39.8
FH-OHMe	-0.123	0.04962	0.0896	0.0737	-44.4
ClH-NH3	-0.079	0.05415	0.1821	0.1423	-44.7
FH-OMe2	-0.112	0.05311	0.0920	0.0799	-45.9

Table S3: *Details of the binding energies and quantum mechanical descriptor values for the 45 complex systems studied with B97D/aug–TZVP level of theory.*

Complex	MOC	ρ	OWBO	WBI	E_I (kJ/mol)
HSH-FNF2	-0.002	0.00360	0.0030	0.0018	-2.2
ClH-FOF	-0.002	0.00697	-0.0039	0.0090	-2.7
BrH-FNF2	0.001	0.00540	0.0063	0.0041	-3.5
H3CH-OMe2	-0.044	0.00583	-0.0026	0.0039	-3.9
H3CH-NHMe2	-0.015	0.00639	-0.0099	0.0081	-4.8
MeHPH-NH3	-0.044	0.00779	0.0198	0.0086	-4.0
HSH-BrH	0.004	0.00502	0.0039	0.0055	-5.1
HOPh-FMe	-0.037	0.00598	0.0057	0.0030	-7.2
H2PH-OH2	-0.052	0.00761	0.0082	0.0047	-3.8
FH-FNF2	0.012	0.01313	0.0078	0.0093	-4.7
BrH-BrH	0.020	0.00967	0.0145	0.0240	-8.3
FH-FOF	0.008	0.01515	0.0097	0.0132	-4.7
FH-ClOCl	0.038	0.01449	0.0063	0.0249	-6.7
HSH-PH2Me	-0.030	0.00956	0.0085	0.0208	-9.6
MeSH-NH2Ph	-0.075	0.01523	0.0214	0.0238	-14.8
MeOH-PH3	0.002	0.01258	0.0327	0.0200	-11.7
MeSH-OHMe	-0.043	0.01671	0.0200	0.0184	-11.7
MeHNH-OMe2	-0.092	0.01470	0.0107	0.0091	-14.8
MeHNH-NMe3	-0.146	0.01753	0.0371	0.0219	-20.3
MeHNH-NH3	-0.153	0.01680	0.0189	0.0187	-13.9
ClH-SH2	0.023	0.01889	0.0776	0.0637	-16.3
PhOH-PH2Me	0.081	0.01617	0.0256	0.0411	-20.5
HSH-NH2Me	-0.075	0.02614	0.0780	0.0578	-21.6
MeOH-PMe3	0.016	0.01826	0.0869	0.0622	-22.0
ClH-PH2Me	0.021	0.02224	0.1195	0.0887	-22.5
FH-FH	-0.021	0.02648	0.0302	0.0254	-16.8
BrH-OH2	-0.078	0.03065	0.0833	0.0561	-21.3
MeOH-OH2	-0.133	0.02596	0.0357	0.0269	-20.1
HOH-OH2	-0.140	0.02568	0.0365	0.0259	-20.4
MeOH-OHMe	-0.128	0.02836	0.0423	0.0307	-23.5
FH-SH2	0.029	0.02490	0.0463	0.0661	-23.2
HOH-OHEt	-0.110	0.02795	0.0363	0.0298	-24.0
FH-OHCl	-0.135	0.03508	0.0519	0.0401	-22.6
ClH-OH2	-0.081	0.03365	0.0876	0.0575	-25.1
HOH-NMe3	-0.193	0.03581	0.0866	0.0591	-35.8
HOH-NH3	-0.154	0.03004	0.0549	0.0489	-29.2
MeOH-NH3	-0.156	0.03107	0.0631	0.0534	-30.1
FH-PH2Me	0.021	0.02815	0.0734	0.0907	-30.7
HOH-NHMe2	-0.152	0.03532	0.0597	0.0572	-35.4
HOH-NH2Me	-0.145	0.03326	0.0574	0.0560	-33.0
MeOH-NH2Me	-0.137	0.03413	0.0686	0.0592	-34.4
FH-OH2	-0.140	0.04431	0.0815	0.0646	-37.6
FH-OHMe	-0.132	0.05073	0.0922	0.0778	-43.7
ClH-NH3	-0.060	0.05528	0.1852	0.1516	-49.7
FH-OMe2	-0.131	0.05422	0.0946	0.0834	-46.5

Table S4: *Details of the binding energies and quantum mechanical descriptor values for the 45 complex systems studied with wB97XD/aug–TZVP level of theory.*

Complex	MOC	ρ	OWBO	WBI	E_I (kJ/mol)
HSH-FNF2	-0.005	0.00354	0.0025	0.0014	-1.5
ClH-FOF	-0.006	0.00682	-0.0036	0.0072	-1.3
BrH-FNF2	-0.002	0.00529	0.0057	0.0033	-2.3
H3CH-OMe2	-0.041	0.00588	-0.0034	0.0035	-3.6
H3CH-NHMe2	-0.035	0.00643	-0.0096	0.0068	-4.9
MeHPH-NH3	-0.045	0.00776	0.0185	0.0065	-4.8
HSH-BrH	0.004	0.00501	0.0030	0.0041	-4.7
HOPh-FMe	-0.031	0.00596	0.0017	0.0022	-6.9
H2PH-OH2	-0.047	0.00764	0.0080	0.0038	-4.3
FH-FNF2	0.014	0.01286	0.0065	0.0079	-4.3
BrH-BrH	0.017	0.00949	0.0291	0.0186	-6.9
FH-FOF	0.010	0.01479	0.0087	0.0111	-4.3
FH-ClOCl	0.038	0.01412	0.0049	0.0206	-5.9
HSH-PH2Me	-0.011	0.00949	0.0070	0.0167	-9.4
MeSH-NH2Ph	-0.100	0.01502	0.0160	0.0195	-15.5
MeOH-PH3	0.027	0.01242	0.0308	0.0169	-11.7
MeSH-OHMe	-0.048	0.01648	0.0136	0.0154	-13.5
MeHNH-OMe2	-0.077	0.01468	0.0091	0.0081	-16.1
MeHNH-NMe3	-0.147	0.01737	0.0363	0.0197	-20.7
MeHNH-NH3	-0.120	0.01658	0.0175	0.0160	-15.2
ClH-SH2	0.021	0.01834	0.0723	0.0531	-16.1
PhOH-PH2Me	0.103	0.01594	0.0227	0.0361	-21.0
HSH-NH2Me	-0.094	0.02537	0.0722	0.0477	-22.3
MeOH-PMe3	0.038	0.01800	0.0197	0.0457	-22.2
ClH-PH2Me	0.037	0.02175	0.1135	0.0780	-21.9
FH-FH	-0.016	0.02596	0.0281	0.0227	-18.9
BrH-OH2	-0.084	0.02968	0.0746	0.0478	-23.4
MeOH-OH2	-0.115	0.02541	0.0336	0.0238	-22.9
HOH-OH2	-0.118	0.02517	0.0341	0.0228	-23.4
MeOH-OHMe	-0.111	0.02774	0.0393	0.0272	-25.9
FH-SH2	0.046	0.02421	0.0426	0.0577	-23.5
HOH-OHEt	-0.095	0.02734	0.0337	0.0263	-26.6
FH-OHCl	-0.122	0.03420	0.0490	0.0354	-24.7
ClH-OH2	-0.081	0.03275	0.0821	0.0504	-28.0
HOH-NMe3	-0.161	0.03491	0.0579	0.0471	-36.9
HOH-NH3	-0.119	0.02930	0.0510	0.0431	-31.5
MeOH-NH3	-0.119	0.03027	0.0592	0.0473	-32.1
FH-PH2Me	0.058	0.02767	0.0690	0.0829	-31.1
HOH-NHMe2	-0.120	0.03438	0.0553	0.0509	-36.6
HOH-NH2Me	-0.107	0.03237	0.0535	0.0495	-34.6
MeOH-NH2Me	-0.097	0.03318	0.0641	0.0526	-35.6
FH-OH2	-0.126	0.04308	0.0780	0.0587	-41.0
FH-OHMe	-0.115	0.04921	0.0879	0.0711	-46.6
ClH-NH3	-0.067	0.05355	0.1800	0.1376	-49.5
FH-OMe2	-0.101	0.05257	0.0902	0.0768	-49.3

Table S5: *Details of the binding energies and quantum mechanical descriptor values for the 45 complex systems studied with MP2/aug–TZVP level of theory.*

Complex	MOC	ρ	OWBO	WBI	E_I (kJ/mol)
HSH-FNF2	-0.005	0.00368	-0.0012	0.0012	-1.1
ClH-FOF	-0.018	0.00696	-0.0024	0.0066	-0.8
BrH-FNF2	-0.002	0.00538	0.0064	0.0027	-1.8
H3CH-OMe2	-0.052	0.00601	-0.0048	0.0034	-2.6
H3CH-NHMe2	-0.052	0.00656	-0.0096	0.0065	-2.7
MeHPH-NH3	-0.079	0.00776	0.0068	0.0050	-2.0
HSH-BrH	-0.006	0.00500	0.0006	0.0030	-3.1
HOPh-FMe	-0.035	0.00610	-0.0006	0.0024	-5.2
H2PH-OH2	-0.070	0.00768	0.0096	0.0031	-2.6
FH-FNF2	0.003	0.01274	0.0032	0.0060	-2.9
BrH-BrH	0.004	0.00934	0.0254	0.0146	-4.7
FH-FOF	-0.001	0.01461	0.0044	0.0085	-3.1
FH-ClOCl	0.023	0.01385	0.0032	0.0161	-3.3
HSH-PH2Me	-0.043	0.00941	0.0062	0.0132	-6.0
MeSH-NH2Ph	-0.206	0.01513	0.0125	0.0172	-12.3
MeOH-PH3	0.004	0.01243	0.0130	0.0143	-9.1
MeSH-OHMe	-0.086	0.01643	0.0084	0.0127	-10.4
MeHNH-OMe2	-0.117	0.01468	0.0244	0.0109	-13.2
MeHNH-NMe3	-0.263	0.01757	0.0026	0.0185	-15.6
MeHNH-NH3	-0.190	0.01648	0.0201	0.0132	-11.0
ClH-SH2	-0.006	0.01804	0.0694	0.0434	-10.5
PhOH-PH2Me	0.062	0.01593	0.0283	0.0313	-17.8
HSH-NH2Me	-0.153	0.02525	0.0604	0.0404	-15.0
MeOH-PMe3	0.028	0.01793	0.0750	0.0473	-17.5
ClH-PH2Me	0.005	0.02142	0.1043	0.0660	-14.9
FH-FH	-0.033	0.02533	0.0220	0.0176	-15.8
BrH-OH2	-0.124	0.02922	0.0675	0.0406	-18.1
MeOH-OH2	-0.175	0.02519	0.0339	0.0199	-18.7
HOH-OH2	-0.175	0.02484	0.0304	0.0190	-19.2
MeOH-OHMe	-0.177	0.02762	0.0348	0.0232	-21.6
FH-SH2	0.029	0.02383	0.0408	0.0488	-18.2
HOH-OHEt	-0.162	0.02714	0.0291	0.0224	-22.1
FH-OHCl	-0.172	0.03381	0.0469	0.0302	-21.6
ClH-OH2	-0.129	0.03222	0.0736	0.0419	-21.2
HOH-NMe3	-0.317	0.03511	0.0408	0.0490	-30.1
HOH-NH3	-0.177	0.02899	0.0489	0.0370	-25.4
MeOH-NH3	-0.180	0.03008	0.0576	0.0409	-25.8
FH-PH2Me	0.037	0.02723	0.0645	0.0727	-24.2
HOH-NHMe2	-0.233	0.03438	0.0534	0.0456	-29.8
HOH-NH2Me	-0.181	0.03220	0.0523	0.0435	-27.9
MeOH-NH2Me	-0.178	0.03314	0.0639	0.0467	-28.8
FH-OH2	-0.174	0.04236	0.0721	0.0502	-34.6
FH-OHMe	-0.174	0.04867	0.0827	0.0624	-39.6
ClH-NH3	-0.132	0.05295	0.1728	0.1216	-37.8
FH-OMe2	-0.188	0.05225	0.0882	0.0684	-42.1

Figure S1: *RMSD equilibration of the backbone atoms of the interfacial residues relative to the starting structure in the three systems; LongAB, LatB and LatA.*
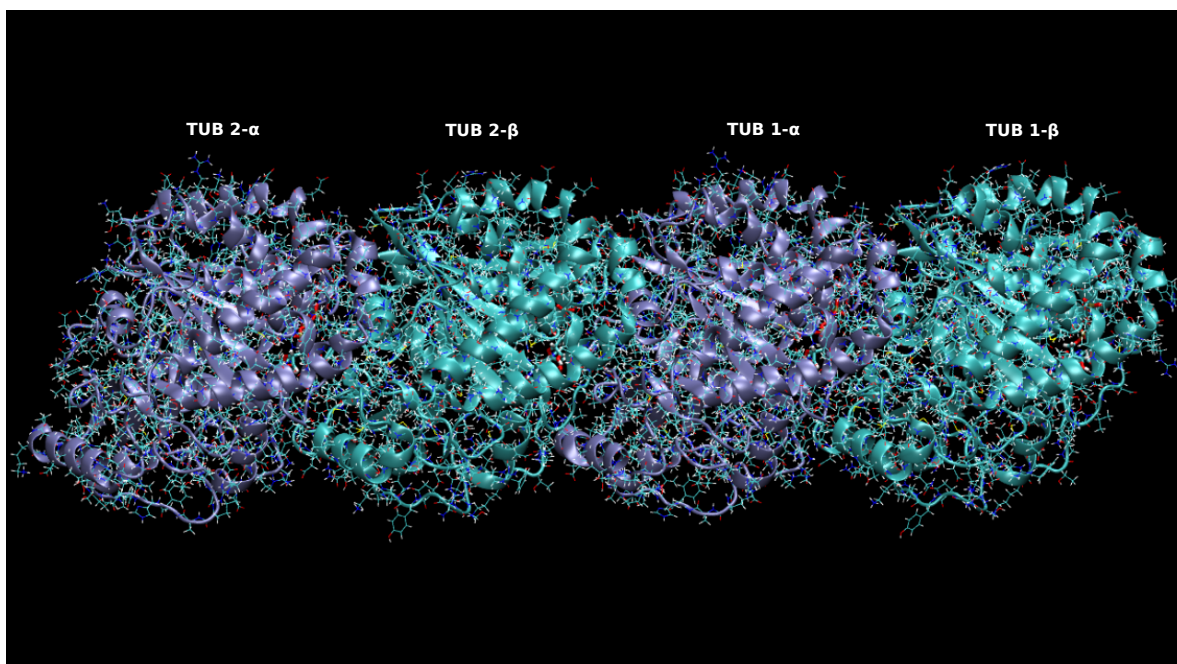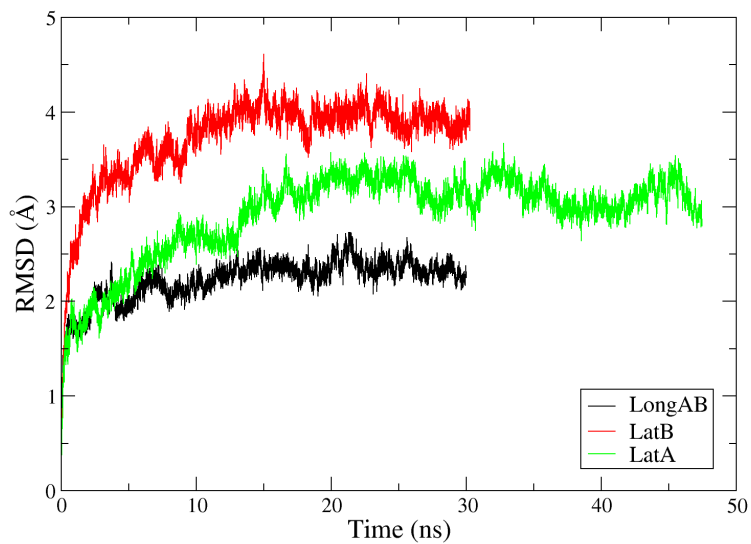
Table S6: *Energy of hydrogen bonds in the LongAB interface in kJ/mol. SS# is snapshot number.*

TUB 1–α	TUB 2–β	SS1	SS2	SS3	SS4	SS5	SS6	SS7	SS8	$E_{average}$	SD
Arg2	Glu71	-62.5	-33.4	-42.6	-42.3	-38.3	-20.1	-58.7	-23.2	-40.1	15
Glu434	Arg401	-30.0	-17.5	-39.3	-59.0	-32.4	-44.3	-38.1	0.0	-32.6	18
Tyr262	Arg401	-46.7	-45.0	-32.7	-45.5	0.0	-13.3	-25.6	-47.6	-32.0	18
Arg243	Asp76	-37.3	-17.3	-39.4	-35.7	0.0	-24.4	-42.8	-42.1	-29.9	15
Thr349	Val181	-44.6	-34.6	-13.8	-14.3	-43.6	-21.5	-21.0	-13.0	-25.8	13
Asp438	Arg401	-26.3	-55.1	-57.1	0.0	0.0	0.0	0.0	-52.7	-23.9	27
Val260	His406	-31.7	-34.1	-12.9	-8.8	-12.1	-41.6	-15.5	-20.0	-22.1	12
Gln133	Gly98	-16.8	-39.1	-20.7	-22.8	-21.2	-28.2	-10.7	-16.0	-21.9	9
Thr257	Gly100	-22.7	-35.5	-17.6	-27.2	-16.3	-13.2	-14.3	-24.2	-21.4	8
Lys352	Thr180	-18.5	-16.8	-22.1	0.0	-30.4	-16.3	-32.1	-12.9	-18.7	10
Asn249	Gln11	-22.6	-20.8	-27.5	-23.1	0.0	-28.6	0.0	-26.3	-18.6	12
Asn329	Lys176	-26.4	-18.2	-26.1	-19.4	-24.4	0.0	0.0	-7.6	-15.3	11
Lys163	Glu411	-34.2	0.0	-40.1	0.0	0.0	-43.7	0.0	0.0	-14.8	21
Asn258	Val181	0.0	-14.8	-15.6	-29.7	-8.8	-9.4	-22.4	-10.2	-13.9	9
Lys352	Asp179	0.0	0.0	0.0	-20.0	-12.2	-33.7	-12.9	0.0	-9.9	12
Asn249	Glu71	0.0	0.0	-31.3	0.0	0.0	-22.8	0.0	0.0	-6.8	13
Asp345	Arg400	0.0	0.0	0.0	0.0	0.0	-28.0	-23.6	0.0	-6.4	12
Asn258	Asn101	0.0	-20.6	0.0	0.0	-14.3	0.0	-12.9	0.0	-6.0	9
Arg2	Gln96	-3.5	0.0	0.0	-15.2	0.0	0.0	0.0	-23.8	-5.3	9
Val260	Trp407	0.0	-9.5	0.0	-4.3	-17.7	-6.7	0.0	0.0	-4.8	6
Gln133	Ser97	0.0	0.0	-4.7	0.0	0.0	0.0	0.0	-30.4	-4.4	11
Leu132	Gln96	0.0	0.0	0.0	0.0	0.0	-18.5	-14.8	0.0	-4.2	8
Asn258	Val182	0.0	-13.8	-13.5	0.0	0.0	0.0	0.0	0.0	-3.4	6
Lys326	Tyr210	0.0	0.0	0.0	0.0	0.0	-17.9	0.0	0.0	-2.2	6
Thr130	Gln96	0	0	0	0	-17	0	0	0	-2.1	6
Asn258	Thr180	0.0	0.0	0.0	-16.1	0.0	0.0	0.0	0.0	-2.0	6
Val437	Arg401	0.0	0.0	0.0	0.0	0.0	-11.6	0.0	0.0	-1.4	4
Val353	Asp179	0.0	-5.0	-4.4	0.0	0.0	0.0	0.0	0.0	-1.2	2
Lys336	Lys176	0.0	0.0	-2.3	0.0	0.0	0.0	0.0	-4.2	-0.8	2
Glu254	Asn101	0.0	0.0	0.0	0.0	0.0	0.0	-3.5	0.0	-0.4	1
Total Energy		-424	-431	-464	-384	-289	-444	-349	-354	-392	59

Figure S3: *An all-atom model of the LatB system with subunit assignment.*
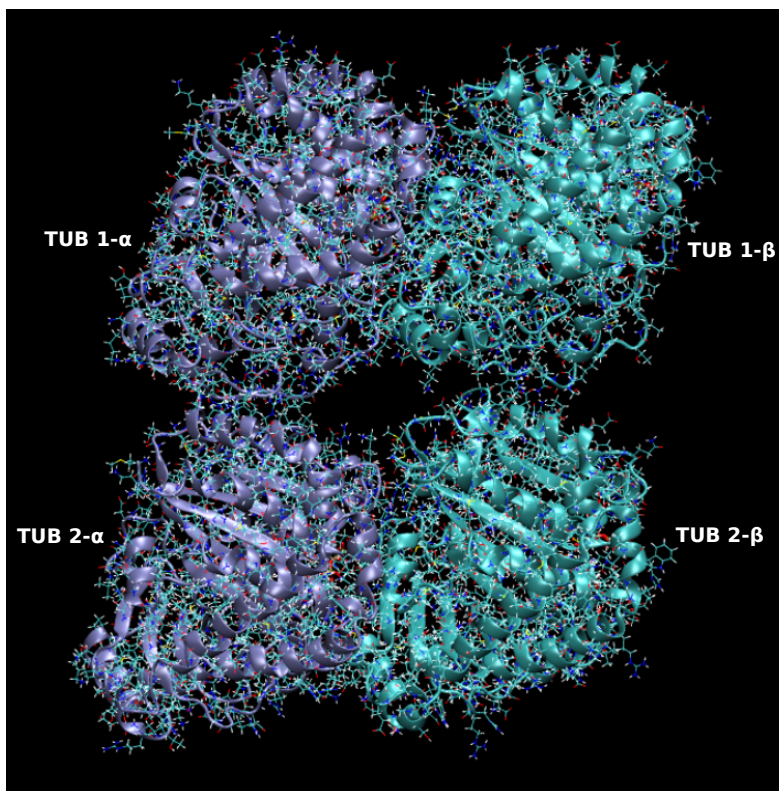
Table S7: *Energy of hydrogen bonds in the LatB interface in kJ/mol. SS# is snapshot number.*

						$\alpha-\alpha$ **Interactions**					
TUB 1–α	TUB 2–α	SS1	SS2	SS3	SS4	SS5	SS6	SS7	SS8	$E_{average}$	SD
Arg215	Glu90	0.0	-44.0	-36.7	-49.3	-64.7	-51.4	-40.6	-55.3	-42.7	19
Lys338	Asp127	-24.5	-13.8	-33.6	-34.1	-26.2	-41.9	-28.9	-24.7	-28.5	8
Glu297	Arg121	-26.1	-21.9	-40.0	0.0	-62.8	0.0	-23.6	-22.2	-24.6	20
Glu297	Lys124	-20.0	-29.2	-15.9	-38.3	0.0	-32.1	-35.8	-17.1	-23.5	13
Glu284	Ser54	0.0	-32.6	-24.3	0.0	-33.4	-24.5	-38.7	-24.8	-22.3	15
Gln372	Glu55	-28.0	0.0	-13.7	-28.0	-15.9	-20.9	-16.9	-24.2	-18.4	9
Tyr282	Ser48	0.0	-31.8	-19.3	-20.9	-8.1	0.0	0.0	-13.2	-11.7	12
His283	Phe49	-7.5	-8.8	-28.2	0.0	-12.8	-5.8	-7.9	-13.5	-10.5	8
Ala278	Asn50	-20.3	0.0	0.0	-10.0	-14.9	0.0	0.0	0.0	-5.7	8
Arg373	Ser54	-8.1	0.0	-25.4	0.0	0.0	0.0	0.0	0.0	-4.2	9
Glu284	Lys60	0.0	0.0	0.0	0.0	0.0	-8.5	-23.6	0.0	-4.0	8
Leu286	Ser54	0.0	-2.9	-7.6	0.0	-9.3	-4.7	0.0	-6.1	-3.8	4
Gln285	Gly57	-15.1	0.0	0.0	0.0	0.0	0.0	-11.9	0.0	-3.4	6
Gln372	Thr56	0.0	0.0	0.0	0.0	0.0	-24.6	0.0	0.0	-3.1	9
His283	Asn50	0.0	-11.3	0.0	0.0	0.0	0.0	0.0	0.0	-1.4	4
Glu290	Gln128	0	0	0	0	-9	0	0	0	-1.1	3
Lys370	Thr51	0.0	0.0	0.0	0.0	0.0	0.0	-7.1	0.0	-0.9	3
Total Subunit Energy		-149	-196	-245	-180	-257	-214	-235	-201	-210	35
						$\beta-\beta$ **Interactions**					
TUB 1–β	TUB 2–β	SS1	SS2	SS3	SS4	SS5	SS6	SS7	SS8	$E_{average}$	SD
Arg308	Asp116	-46.7	-58.9	0.0	-37.5	-69.8	-57.4	-54.3	-33.7	-44.8	22
Glu290	Arg88	-51.6	-35.3	-32.2	-55.4	-22.6	-30.3	-30.3	-53.9	-38.9	13
Arg308	Asp120	-29.2	-26.9	-27.3	-25.6	-37.7	-27.6	-61.5	-58.2	-36.7	15
Lys299	Asp90	-11.0	-37.8	-25.5	-34.6	-31.0	-34.7	-31.3	-55.0	-32.6	12
Asp297	Lys124	-31.8	0.0	-31.3	0.0	-32.1	0.0	-29.0	-54.4	-22.3	20
Tyr342	Asp120	0.0	-9.9	0.0	-10.8	-25.1	-8.1	-41.7	-51.8	-18.4	19
Ser280	Arg88	-14.5	-21.7	-6.3	-22.9	0.0	-51.6	0.0	-16.1	-16.6	17
Lys338	Lys124	-6.4	-20.8	-15.5	-24.9	-25.2	0.0	0.0	-25.4	-14.8	11
Lys338	Ser126	-24.0	-23.6	-7.2	0.0	-14.0	0.0	0.0	-27.5	-12.0	12
Lys338	Arg123	0.0	0.0	0.0	0.0	0.0	-35.9	-19.9	-4.1	-7.5	13
Asn334	Glu127	0.0	-19.0	0.0	-32.7	0.0	0.0	0.0	0.0	-6.5	12
Asn335	Glu128	0.0	-6.8	0.0	0.0	0.0	0.0	0.0	0.0	-0.9	2
Total Subunit Energy		-215	-261	-145	-244	-257	-246	-268	-380	-252	65
Total Energy		-365	-457	-390	-425	-514	-460	-503	-581	-462	70

Table S8: *Energy of hydrogen bonds in the LatA interface in kJ/mol. SS# is snapshot number.*

α–β Interactions										
TUB 1–α	TUB 2–β	SS1	SS2	SS3	SS4	SS5	SS6	SS7	SS8	$E_{average}$	SD
Asp47	Arg284	-44.6	-17.9	-41.5	-40.4	-32.1	-54.1	-71.1	-36.0	-42.2	15.7
Lys124	Asp297	-24.7	-19.4	-39.2	-42.9	-29.9	-25.5	-27.4	-44.8	-31.7	9.4
Gln85	Ser280	-31.4	-23.8	0.0	-29.4	-34.9	-23.1	-22.4	-23.4	-23.6	10.6
Asp46	Arg278	-26.7	-10.3	-39.2	-39.6	-14.3	-19.8	-33.9	0.0	-23.0	14.4
Asp127	Asn334	-25.0	-24.3	-27.9	-16.4	-21.9	-21.8	-28.5	-15.0	-22.6	4.9
Asp120	Lys338	-39.2	0.0	-20.4	-36.6	-26.5	0.0	-25.3	-31.4	-22.4	15.1
Gln128	Gln293	-6.5	-20.3	-27.8	-36.5	-10.6	-28.8	-18.1	0.0	-18.6	12.3
Asp47	Gln282	-32.0	-8.2	-17.8	0.0	-29.7	0.0	-8.6	-36.0	-16.5	14.5
Glu55	Arg284	-39.8	0.0	-45.9	0.0	0.0	-19.8	-13.0	0.0	-14.8	18.9
Arg121	Asp297	0.0	-69.3	0.0	0.0	-35.3	0.0	0.0	0.0	-13.1	25.9
Asp47	Arg278	0.0	-18.5	0.0	0.0	0.0	-34.0	-16.8	-25.5	-11.9	13.7
Ser54	Lys372	0.0	0.0	0.0	0.0	-21.1	-18.6	-13.5	-24.3	-9.7	10.8
Lys124	Gln293	0.0	-12.8	-17.0	0.0	-15.6	0.0	0.0	-9.9	-6.9	7.7
His88	Gln281	0.0	-9.3	-19.8	-16.2	0.0	0.0	0.0	0.0	-5.7	8.3
Ser54	Arg284	-11.1	0.0	0.0	0.0	-32.0	0.0	0.0	0.0	-5.4	11.4
Arg123	Asn334	0.0	-15.8	0.0	-5.3	0.0	-1.9	-16.1	0.0	-4.9	7.1
Thr56	Gly370	-7.7	-16.7	-4.4	-7.5	0.0	0.0	0.0	0.0	-4.5	6.0
Gly59	Arg284	0.0	0.0	-19.2	-16.3	0.0	0.0	0.0	0.0	-4.4	8.2
Phe53	Ser374	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-31.1	-3.9	11.0
Asn50	Gln282	0.0	0.0	0.0	-9.6	0.0	-20.4	0.0	0.0	-3.8	7.5
Phe87	Ser280	0.0	0.0	-24.5	0.0	0.0	0.0	0.0	0.0	-3.1	8.7
Gln128	Thr287	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-18.8	-2.4	6.7
Ser54	Met373	0.0	0.0	0.0	0.0	0.0	-15.1	0.0	0.0	-1.9	5.3
Gln128	Glu290	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-14.6	-1.8	5.2
Total Subunit Energy		-289	-267	-344	-297	-304	-283	-295	-311	-299	23
β–α Interactions										
TUB 1–β	TUB 3–α	SS1	SS2	SS3	SS4	SS5	SS6	SS7	SS8	$E_{average}$	SD
Arg88	Glu279	-45.6	-26.5	-36.3	-47.6	-40.4	-40.1	-45.3	0.0	-35.2	15.7
Lys124	Glu284	-25.6	0.0	-36.6	-32.4	-30.2	-38.7	-35.2	-17.2	-27.0	12.9
Ile86	Tyr282	-7.5	-29.9	0.0	-31.9	-38.9	-23.1	-25.2	-27.1	-22.9	12.9
Asp90	Lys280	-36.6	0.0	-27.4	-19.7	-39.1	-25.8	0.0	0.0	-18.6	16.5
Asn54	Glu284	-37.0	0.0	-24.6	0.0	0.0	-19.8	-14.5	-19.5	-14.4	13.6
Glu127	Thr334	0.0	-34.3	0.0	-33.4	-43.0	0.0	0.0	0.0	-13.8	19.3
Asp90	Ala281	-17.4	-14.8	0.0	0.0	-22.6	-17.9	-30.9	0.0	-12.9	11.7
Glu127	Thr337	0.0	-22.9	0.0	-39.0	0.0	0.0	0.0	-20.4	-10.3	15.2
Glu55	Gln285	-5.4	-13.1	-11.9	-9.3	-6.7	-20.2	-9.1	0.0	-9.5	6.0
Thr33	His283	0.0	0.0	-18.3	-5.2	0.0	0.0	0.0	-11.6	-4.4	7.0
Asp120	Lys338	-18.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-2.3	6.5
Arg88	Ala281	0.0	0.0	0.0	-8.4	0.0	-5.0	0.0	0.0	-1.7	3.2
Total Subunit Energy		-193	-141	-155	-227	-221	-191	-160	-96	-173	44
Total Energy		-482	-408	-499	-523	-525	-474	-455	-406	-472	46

Figure S4: *An all-atom model of the LatA system with subunit assignment.*