Stony Brook University

OFFICIAL COPY

The official electronic file of this thesis or dissertation is maintained by the University Libraries on behalf of The Graduate School at Stony Brook University.

© All Rights Reserved by Author.

Structural and Computational Study on the Interaction of Bacterial Cell Division Protein, FtsZ, with Its Inhibitors for New Antibacterial Drug Discovery

A Thesis Presented

by

Edwin Ottoniel Lazo

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Master of Science

in

Chemistry

Stony Brook University

May 2014

Stony Brook University

The Graduate School

Edwin Ottoniel Lazo

We, the thesis committee for the above candidate for the

Master of Science degree, hereby recommend

acceptance of this thesis.

Dr. Iwao Ojima – Thesis Advisor

Distinguished Professor of Chemistry and Director of ICB&DD

Dr. Jean Jakoncic – Thesis Advisor

Scientist at Brookhaven National Laboratory

Dr. Peter Tonge – Chair

Professor of Chemistry

Dr. Jarrod French – Third Member

Professor of Chemistry

This thesis is accepted by the Graduate School

Charles Taber

Dean of the Graduate School

Abstract of the Thesis

Structural and Computational Study on the Interaction of Bacterial Cell Division Protein, FtsZ, with Its Inhibitors for New Antibacterial Drug Discovery

by

Edwin Ottoniel Lazo

Master of Science

in

Chemistry

Stony Brook University

2014

Filamenting temperature-sensitive mutant Z (FtsZ) from *Mycobacterium tuberculosis* (Mtb) is an essential bacterial cell division protein that polymerizes into a structure called the "Z-ring". Here it has been targeted for drug discovery. Compounds that exhibit anti-tuberculosis activity have been synthesized and used in crystallization conditions. Four crystal structures of MtbFtsZ have been determined, of which two are similar to the published dimer (PDB 1RQ7) which exhibits lateral interactions; they belong to $P6_5$ space group and crystals diffracted to about 2.6Å. The other two structures had two trimers in the asymmetric unit. The crystals diffracted to approximately 3.5Å and the structures were refined to an R_{cryst} in the range of 0.24-0.28. These latter structures showed novel interactions including a hinge-opening mechanism in which Asn205, Asp207, and Asp210 from loop T7 of subunit A are within ~16Å of the nucleotide from subunit B. Additionally, subunit C from the trimer interacted with the central monomer B via a newly observed T9 loop interaction, where Glu231 from subunit B interacts

with Gly18 and Gly107 from the nucleotide binding pocket of subunit C. Since all these residues are conserved it is plausible that this novel T9 interaction could play a role in the biological process in bacterial cell division. As a result, virtual docking was conducted on these two newly observed interactions (the T7 region and the T9 loop) with two small molecules of interest, SB-P17G-A20, which is a tri-substituted benzimidazole and SB-RA-5001, which is a taxane. Simulations revealed that one of the compounds interacts in the T9 region with a binding energy better than -7.5kcal/mol.

Dedication Page

This work is dedicated to my beautiful wife (Paola A. Lazo), my mother (Marleny E. Lopez), and my grandmother (Clotilde Lopez) who have supported me throughout my entire academic career.

Frontispiece

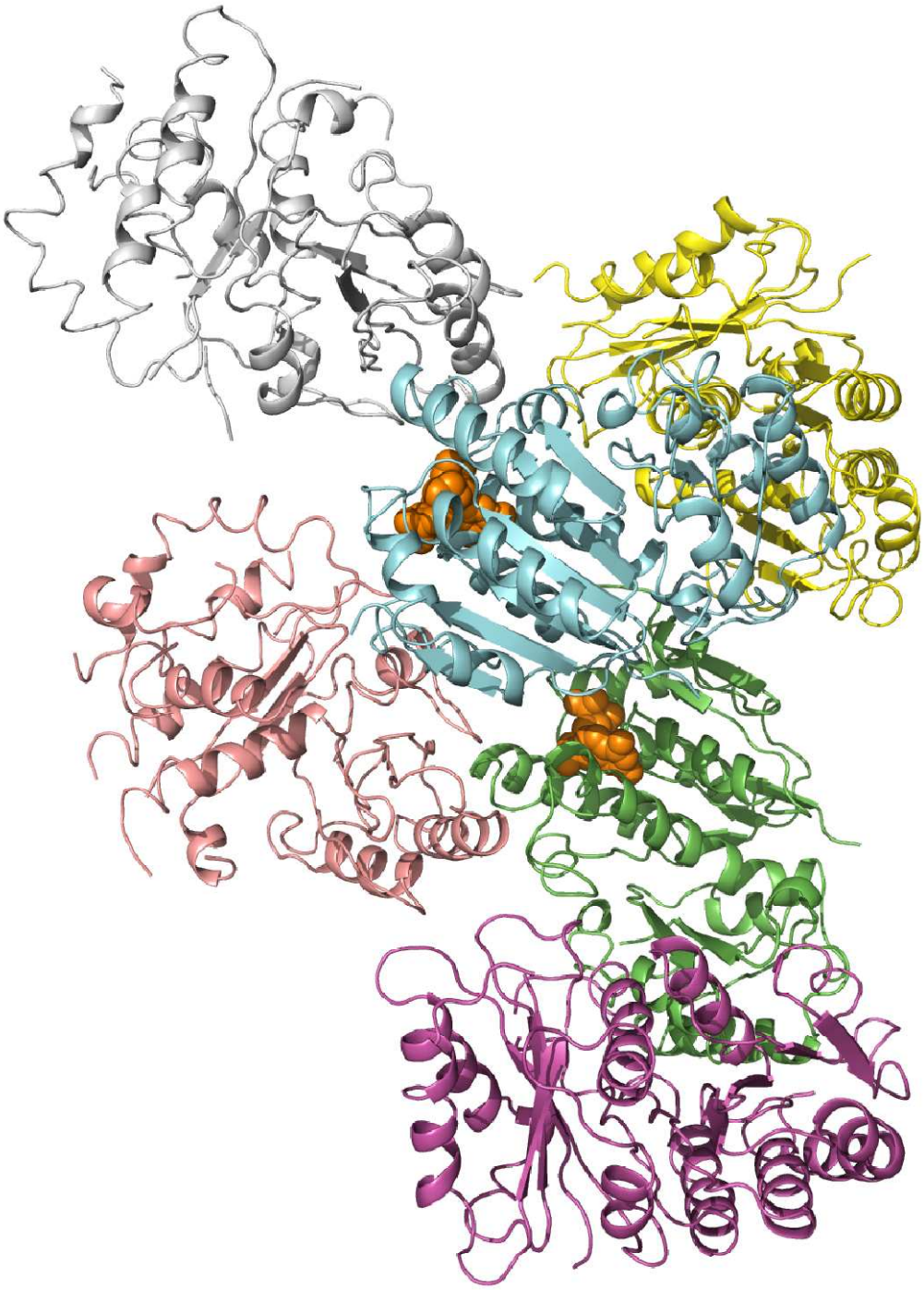


Table of Contents

Chapter 4: Discussion and Perspectives

List of Figures

List of Tables

List of Abbreviations

Ala	alanine
Asp	asparate
Asn	asparagine
ASU	asymmetric unit
Å	Angstroms
CCD	charge-coupled device
CCP4	collaborative computational project number 4
DNA	deoxyribonucleic acid
EcFtsZ	*E. coli* FtsZ
FtsZ	filamenting temperature-sensitive mutant Z
G	glycine
GAPS	GTPase-activating proteins
GDP	guanosine-diphosphate
Gln	glutamine
Glu	glutamate
Gly	glycine
GTP	guanosine 5'-triphosphate
GTPase	guanosine 5'-triphosphatase
GTPγS	guanosine 5'-O-[gamma-thio] triphosphate
H	horizontal
His	histidine
HIV	human immunodeficiency virus
Ile	isoleucine
ICB&DD	Institute of Chemical Biology and Drug Discovery
IPTG	isopropyl β-D-1-thioglactopyranoside
K	Kelvin
kcal	kilocalorie
KCl	potassium chloride
LB	Luria-Bertani
Leu	leucine
LN2	liquid nitrogen
LTB	latent infection of tuberculosis
MDR-TB	multidrug-resistant tuberculosis
MIC	minimum inhibitory concentration
Mg	magnesium
mg	milligram
MjanFtsZ	*Methanococcus jannaschii* FtsZ
ml	milliliter
mol	mole
MR	Molecular Replacement
Mtb	*Mycobacterium tuberculosis*
NaCl	sodium chloride
NaCitrate	sodium citrate
NH_4OAc	ammonium acetate

NMR	nuclear magnetic resonance
NP-40	Tergitol-type nonylphenoxypolyethoxylethanol-40
PDB	protein data bank
PDBQT	protein data bank partial charge (Q) atom type (T)
PEG	polyethylene glycol
[ppt]	precipitant concentration
ph	photon
rmsd	root mean square deviation
S	serine
s	second
SaFtsZ	*Staphylocucccus aureus* FtsZ
SBU	Stony Brook University
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
Ser	serine
sH2	switch Helix 2
T	threonine
TB	tuberculosis
Tris	tri(hydroxymethyl)aminomethane
V	vertical
XDR-TB	extensively drug-resistant tuberculosis
α	alpha
β	beta
γ	gamma
µm	micron

Acknowledgments

First of all I would like to thank my advisor, Distinguished Professor Iwao Ojima for accepting me into his group and pushing me through his questions, insights, and ideas. His strict academic structure provided a pathway for me to follow which allowed me to successfully complete my thesis program.

I want to thank Dr. Jean Jakoncic, my co-advisor at Brookhaven National Laboratory, for guiding me through my thesis program and for teaching me the theories and techniques that are involved in protein crystallography for without him I would not have been able to solve these structures and for accepting to be my co-advisor.

I also owe a big thank you to my laboratory mates, for their help and support throughout my research. In specific to Dr. Soumya Chowdhury for providing me the protein batch used in this study and teaching me the various steps required in protein purification; Divya Awasthi for providing me the latest lead compounds developed in Ojima's group and for fruitful discussions; and finally Longfei Wei for showing how to do computer simulations.

Lastly, I would like to thank my supervisor, Dr. Vivian Stojanoff, for allowing me a flexible schedule to go to class and conduct my research.

This research was carried out in part at the X6A beam-line, funded by the National Institute of General Medical Sciences, National Institute of Health under agreement GM-0080. The National Synchrotron Light Source, Brookhaven National Laboratory is supported by the U.S. Department of Energy under contract No.DE-AC02-98CH10886.

Data for this study were also measured at beam-line X25 of the National Synchrotron Light Source. Financial support comes principally from the Offices of Biological and Environmental Research and of Basic Energy Sciences of the US Department of Energy, and from the National Center for Research Resources (P41RR012408) and the National Institute of General Medical Sciences (P41GM103473) of the National Institutes of Health.

Chapter 1: Introduction

Tuberculosis

Tuberculosis (TB) is a deadly disease and the second leading cause of death from infectious disease worldwide[1]. *Mycobacterium tuberculosis* (Mtb) is the bacterium responsible for this potentially lethal sickness[2]. People with the human immunodeficiency virus (HIV) are especially vulnerable since Mtb becomes active when the immune system becomes weak making this the number one opportunistic infectious disease that causes death among this population[3, 4]. People with HIV are not the only population at high risk. Diabetes, alcohol abuse, and drug abuse are factors that increase the chances of infected people to develop TB[3, 5]. Mtb is quickly evolving and is out-pacing current treatments for the disease, causing a major health concern[6]. If action is not taken soon, it is estimated that by 2020 150 million people will be infected by Mtb and over 36 million will die[4].

TB is separated into latent infection of tuberculosis (LTB) and the disease TB. People with LTB have less Mtb in their bodies[6]. In LTB, bacteria are not active and cannot be spread[2]. Thus, infected people do not show any symptoms and are not ill. However, as they age, their immune system weakens and they can develop TB. In order to prevent this, people with latent TB are prescribed medication[7].

People with TB carry the active form of the bacteria[8]. This active form can be spread via the air when an infected individual speaks, sneezes or coughs[9]. The bacteria can survive in the air for several hours depending on environmental conditions[8]. The diseased individual also exhibits symptoms of TB which include a serious cough lasting 3 weeks, hemoptysis, chest pains, fatigue, weight loss, loss of appetite, chills, fever, and night sweats[10].

Current treatment for TB depends on the diagnosed type, either LTB or TB. People with

LTB have fewer bacteria and the treatment is straightforward. They are prescribed Isoniazid, Pyrazinamide, Ethambutol, and Rifampicin in various order and combinations; these are considered first-line treatments[11]. However, people with TB usually require extensive treatment, especially if they are infected with the multidrug-resistant tuberculosis (MDR-TB) strain. MDR-TB are resistant to Isoniazid and Rifampicin, two of the most powerful first-line drugs[12]. Medications used to combat MDR-TB include Pyrazinamide, Ethambutol, Thioamides, Cycloserine, Aminoglycosides, Cyclic peptides, PAS, and Fluoroquinolones[12]. Unfortunately, Mtb has evolved to become resistant even to these second-line treatments. This strain of Mtb is known as extensively drug-resistant tuberculosis (XDR-TB) and is resistant to Isoniazid, Rifampin, Fluoroquinolones, and to at least one of the injectable second-line drugs, which are Kanamycin, Capreomycin, and Amikacin[13]. This resistivity is one of the main driving forces of why it is important to find a new target to combat TB.

Filamenting temperature-sensitive mutant Z (FtsZ)

Filamenting temperature-sensitive mutant Z (FtsZ) polymerizes forming a structure called the Z-ring (figure 1). It undergoes a process called dynamic polymerization (figure 2)[14]. In this process, the cytoplasm of a cell ready to divide elicits nucleotide exchange with rapid equilibrium in favor of guanosine 5'-triphosphate (GTP) bound FtsZ (figure 2A). Once the critical concentration of GTP bound FtsZ is achieved, polymerization begins forming protofilaments (figure 2B). However, as polymerization is happening so is GTP hydrolysis in a process called "steady-state turnover" (figure 2C). Once the cell divides, regulation of GTP stops and guanosine diphosphate (GDP) bound polymers begin to disassemble, reverting back to GDP bound monomers (figure 2D).

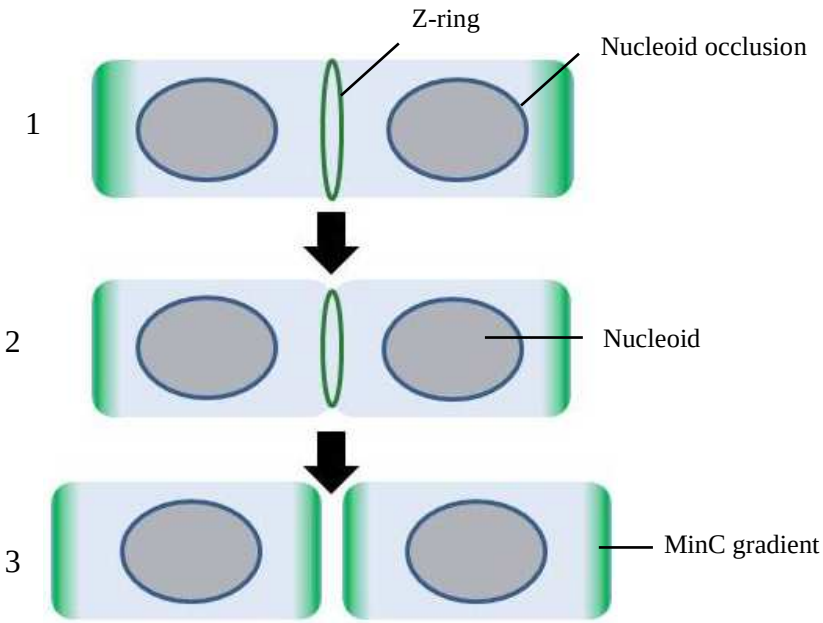


Figure 1: Schematic of Z-ring constriction. 1) Z-ring forms at mid-cell. The MinC gradient is part of the MinCDE system that prevents Z-ring formation at the poles. 2) Constriction of the Z-ring. 3) Formation of two identical daughter cells.

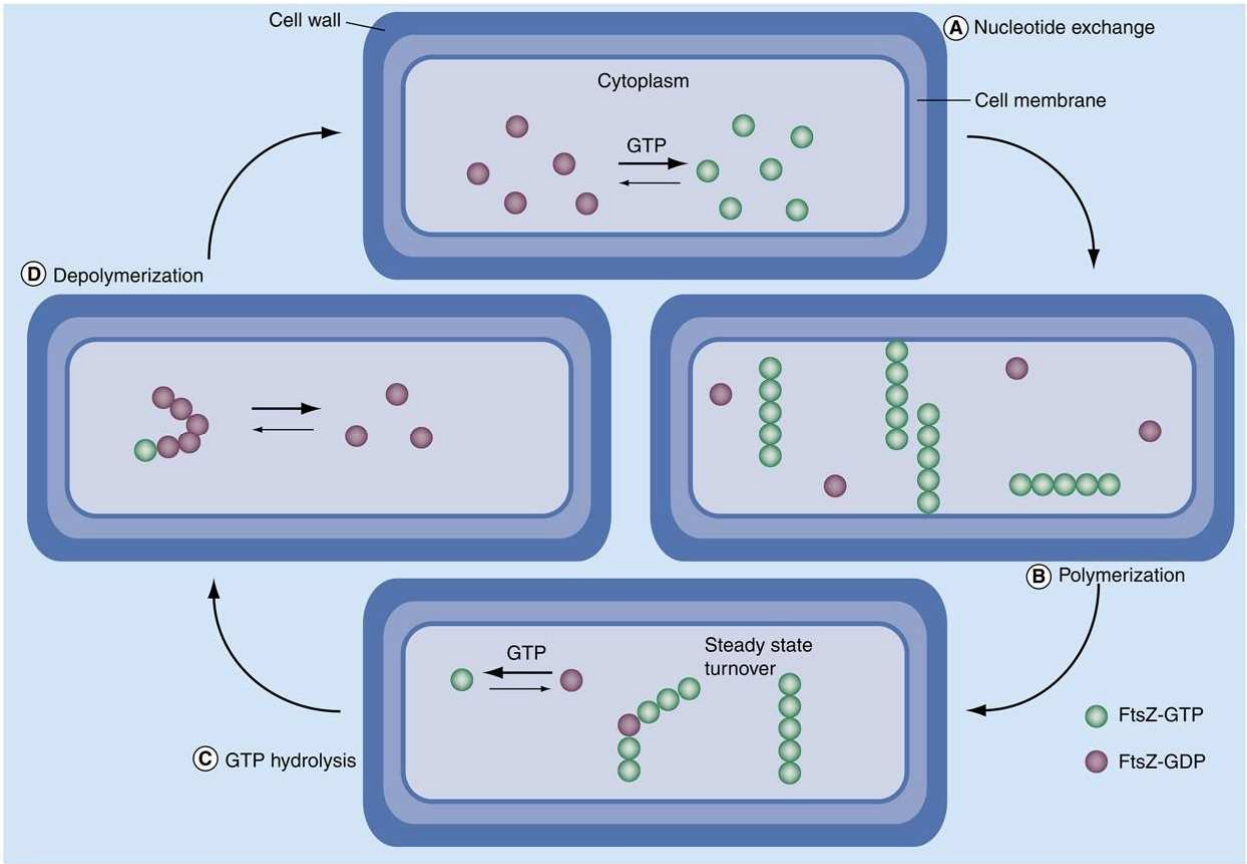


Figure 2: Schematic of Z-ring polymerization. Taken from Kumar *et al.* (2010)[15].

MtbFtsZ was selected as the target for drug discovery efforts because of its role in cell division[16]. Studies identified that one of the key steps in Mtb proliferation is cell division which is triggered by FtsZ[17, 18]. In a study published by Respicio and co-workers, residue aspartate 210 in MtbFtsZ was mutated to a glycine which significantly reduced polymerization (100-fold) and lowered GTP hydrolysis (5-fold) when compared to the wild type, which led to poorer viability of the FtsZ merodiploid strain[18].

FtsZ, like tubulin, forms protofilaments which make up the Z-ring[19]. Ojima and co-workers have synthesized small molecules that exhibit anti-TB activity[15, 17, 20]. These novel compounds have two modes of action: trisubstituted benzimidazoles which reduce polymerization and taxanes that stabilized the FtsZ polymers[15, 17, 20]. Albendazole and thiabendazole are fungicides and parasiticides that inhibit septum formation in tubulin[16]. Studies conducted by White, Reynolds, and co-workers identified FtsZ inhibitors which contain the pyridopyrazine and pteridine pharamocophores, which are similar to the benzimidazole moiety[21, 22]. Based on this knowledge, it was hypothesized that the benzimidazole scaffold was a good starting point for novel MtbFtsZ inhibitors[15]. Paclitaxel (Taxol®) is a microtubule-stabilizing anticancer agent that was co-crystallized in αβ-tubulin (PDB 1JFF). It was discovered to show modest activity against drug-resistant strains of Mtb[15]. This served as a "launching pad" to develop taxanes that have been shown to exhibit significant anti-TB activity[15].

Current Structural Model of FtsZ

As previously mentioned, FtsZ forms protofilaments that evolve into the Z-ring. The mechanism in which FtsZ accomplishes this is by inserting the T7 loop into the nucleotide binding site in a head-to-tail manner (figure 3)[23]. The T7 loop contains amino acids Asn205, Asp207, and Asp210 which are necessary for GTP hydrolysis[24-26]. GTP hydrolysis is not

required for FtsZ to assemble into protofilaments, but is responsible for the conformational changes that protofilaments undergo during FtsZ assembly[27, 28]. GTP bound FtsZ forms straight protofilaments while GDP bound favors the curved conformation (figure 4)[28].

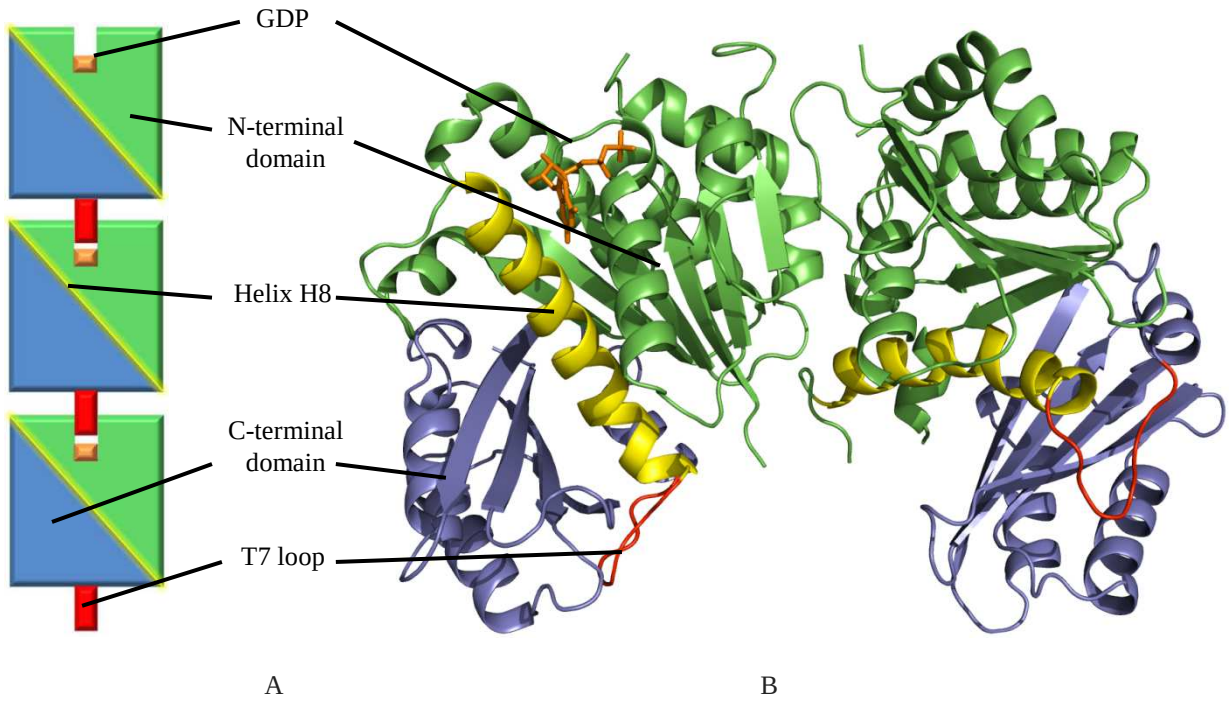


Figure 3: Polymerization of the FtsZ in a head-to-tail fashion. A) T7 loop (red) inserts into the nucleotide binding site. Green represents N-terminal and blue is the C-terminal domain. GDP is represented in orange. B) Dimeric crystal structure of MtbFtsZ (structure A this study). The two globular domains are separated by the helix H8 shown in yellow. Same color scheme as in figure 4A.

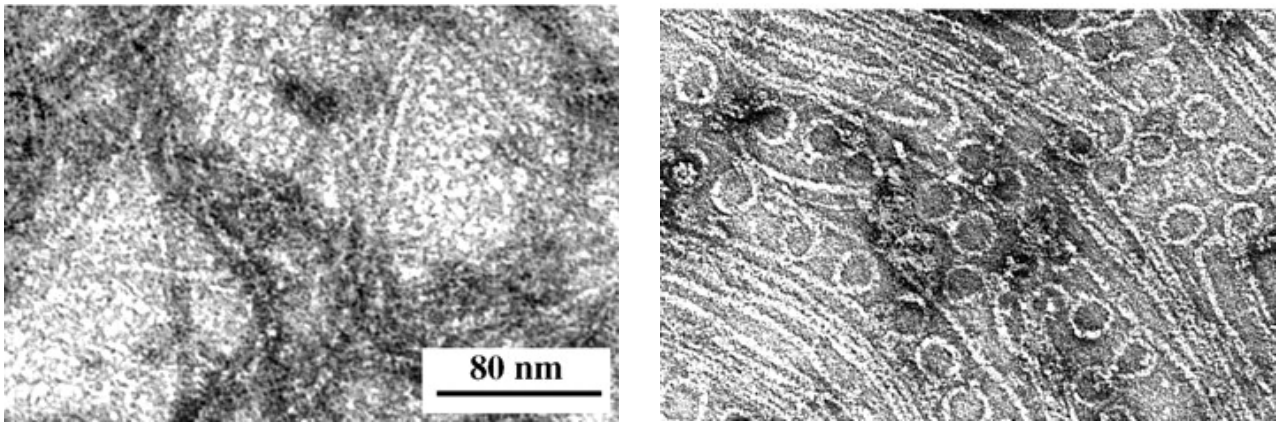


Figure 4: Electron micrograph of straight protofilaments formed with GTP (left) and minirings assembled with GDP (right). Scale bar applies to both. Figures adapted from Lu *et al.* (2000) [28].

FtsZ is a GTPase activated protein that contains two globular domains separated by a helix called H8 (figure 3) and has the tubulin signature motif 103-GGGTGSG-109 (figure 5), with one substitution where S (serine) is replaced with T (threonine) in FtsZ, which can bind nucleotides such as GTP, guanosine-diphosphate (GDP), and citrate[24, 29]. Switch I and switch II have been identified in MtbFtsZ to be helix sH2 (s denotes switch) and the T3 loop, respectively (figure 6).

Switch I in MtbFtsZ is spatially analogous to the G-protein switch I. It contains the highly conserved Asn41, Thr42, and Asp43 residues which are required for GTPase activity[26, 30]. Mutagenesis studies of the corresponding residues, Asn43 and Asp45, in *E. coli* FtsZ (EcFtsZ) showed a significant reduction in GTPase activity[30]. These three residues form an intricate hydrogen bond network that is determined by the ligation state of FtsZ (figure 7)[29]. When GTP is bound, the side-chain oxygen of Asn41 is directed to the β and γ -phosphate oxygen atoms by a water molecule. Thr42 oxygen from the side-chain coordinates with Thr106 side-chain oxygen which interacts with γ -phosphate oxygen through a hydrogen bond. The side-chain of Asp43 is directed away and interacts with Ala46 nitrogen atom. When the switch is in the OFF position, the side-chains of these residues flip their orientation, breaking the bridge. Once this happens, the entire sH2 region takes the form of a beta sheet or a loop and in rare cases remains a helix.

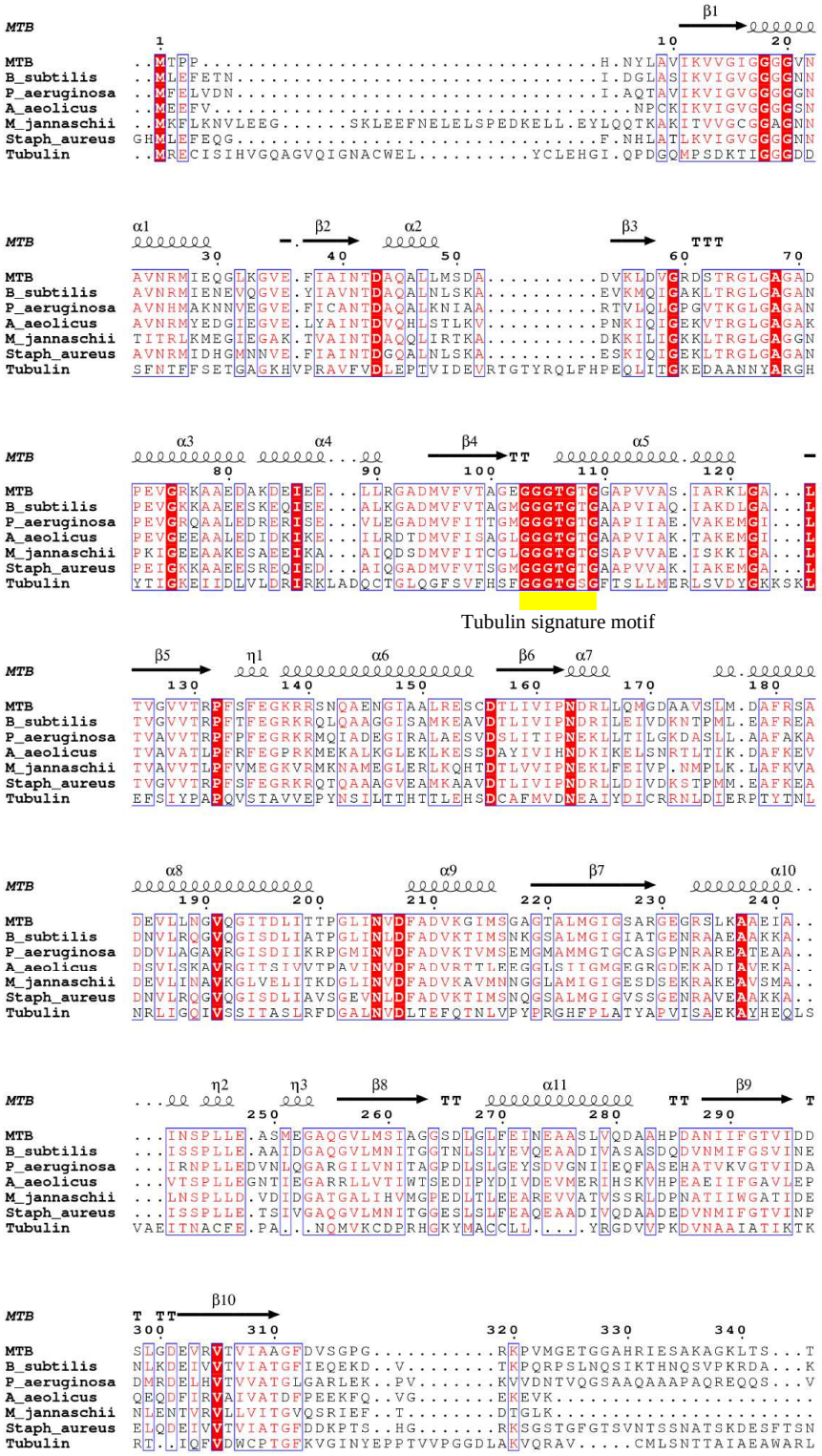


Figure 5: Sequence alignment of Ftsz from different organism and tubulin. Note that the tubulin signature motif is conserved (underlined by a yellow bar).

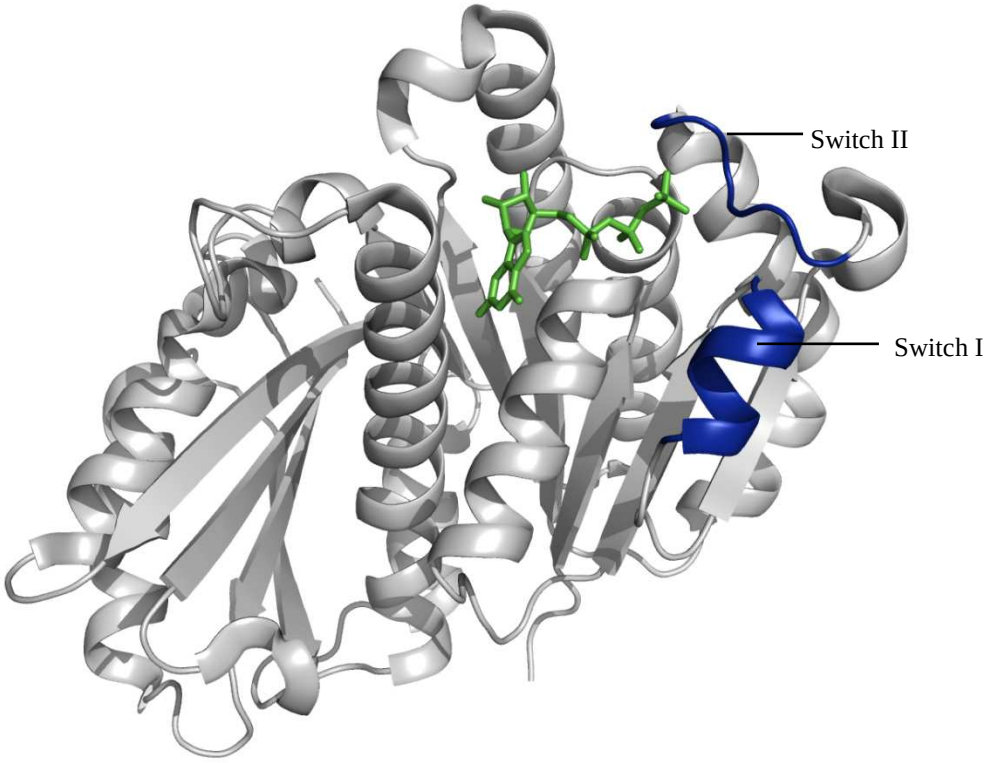


Figure 6: Crystal structure of MtbFtsZ chain A bound to GTPγS (shown in green). Blue represents switch I (sH2) and switch II (T3 loop) in their ON position. OFF position for sH2 would take a beta sheet or loop conformation and the T3 loop would become disordered. PDB 1RLU.

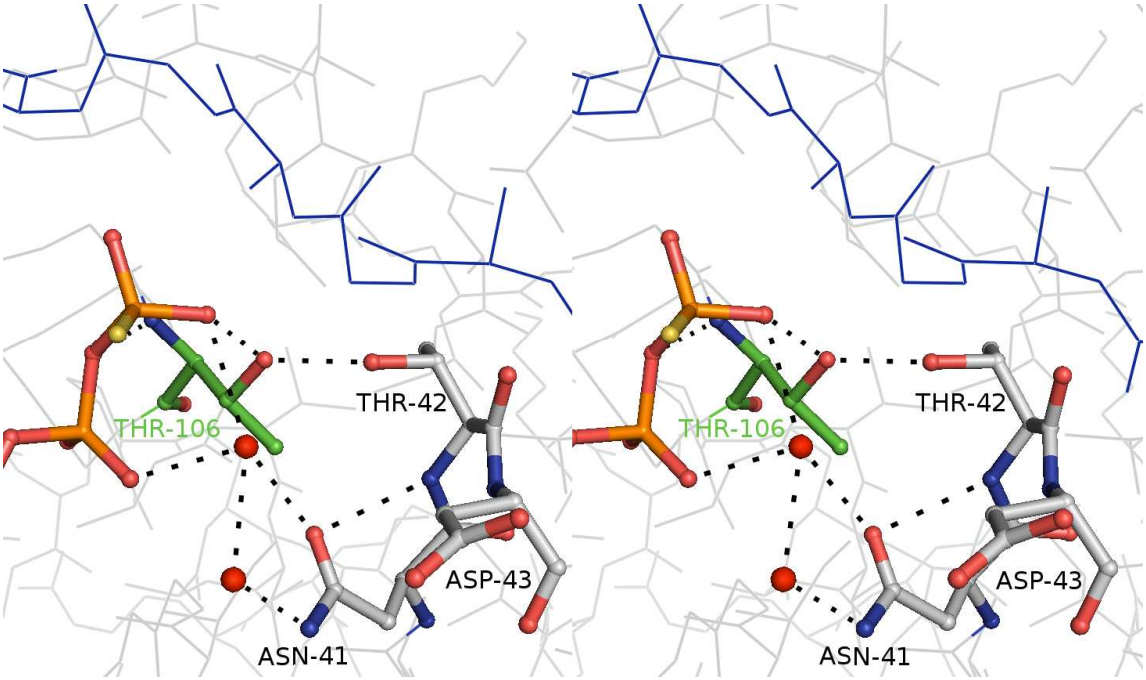
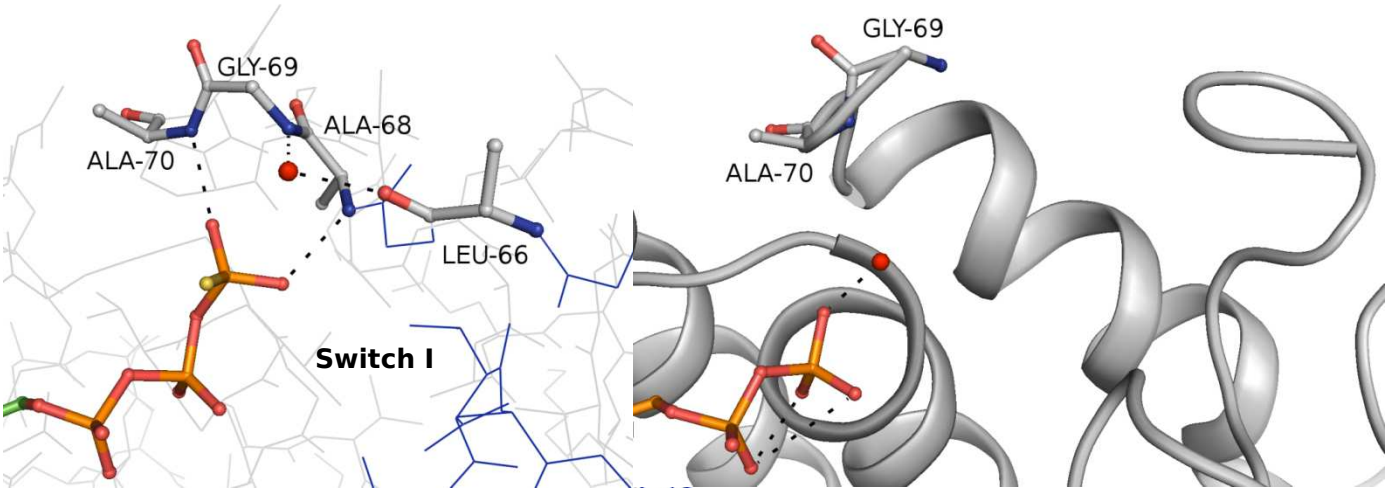


Figure 7: Stereo-view of hydrogen network between the γS-phosphate (orange) and switch I residues Asn41, Thr42, and Asp43. The blue segment (top) represents switch II. Red spheres represent water molecules and black dash lines represent hydrogen bonding. PDB 1RLU.

Switch II needs to be in the OFF position for switch I to change conformations from a helix to a beta sheet or loop[29]. The T3 loop becomes ordered when GTP(γ)S is the nucleotide and is disordered in the presence of GDP, citrate, or when no nucleotide is present (figure 8)[29, 31]. Residues from the T3 loop, Leu66, Gly69, Ala68, and Ala70 become rigid when the nucleotide is GTP(γ)S[29]. The T3 loop collapses inward when it interacts with these residues. Ala68 and Ala70 coordinate directly to the γ-phosphate and Leu66 and Gly69 are stabilized by a water molecule (figure 8A)[29]. After GTP hydrolysis, this system is broken and switch II is in its OFF position, which allows switch I to take its beta sheet or loop form (figure 8B)[29].



A B

Figure 8: Comparison of the T3 loop (switch II) region. A) GTPγS bound to FtsZ (PDB 1RLU). Switch II is in the ON position. B) GDP bound to FtsZ (structure A this study). Switch II is in the OFF position. The red spheres represent water molecules and the black dash lines represent hydrogen bonding.

In *Methanococcus jannaschii* FtsZ, GTP hydrolysis requires residues Asp235 and Asp 238, which are analogous to Asp207 and Asp210, respectively[23, 29]. Asp235 and Asp238 polarize the attacking water molecule (figure 9). Mg^{2+} polarizes the γ-phosphate with Gln75, several water molecules, and the β and γ-phosphates (figure 9).

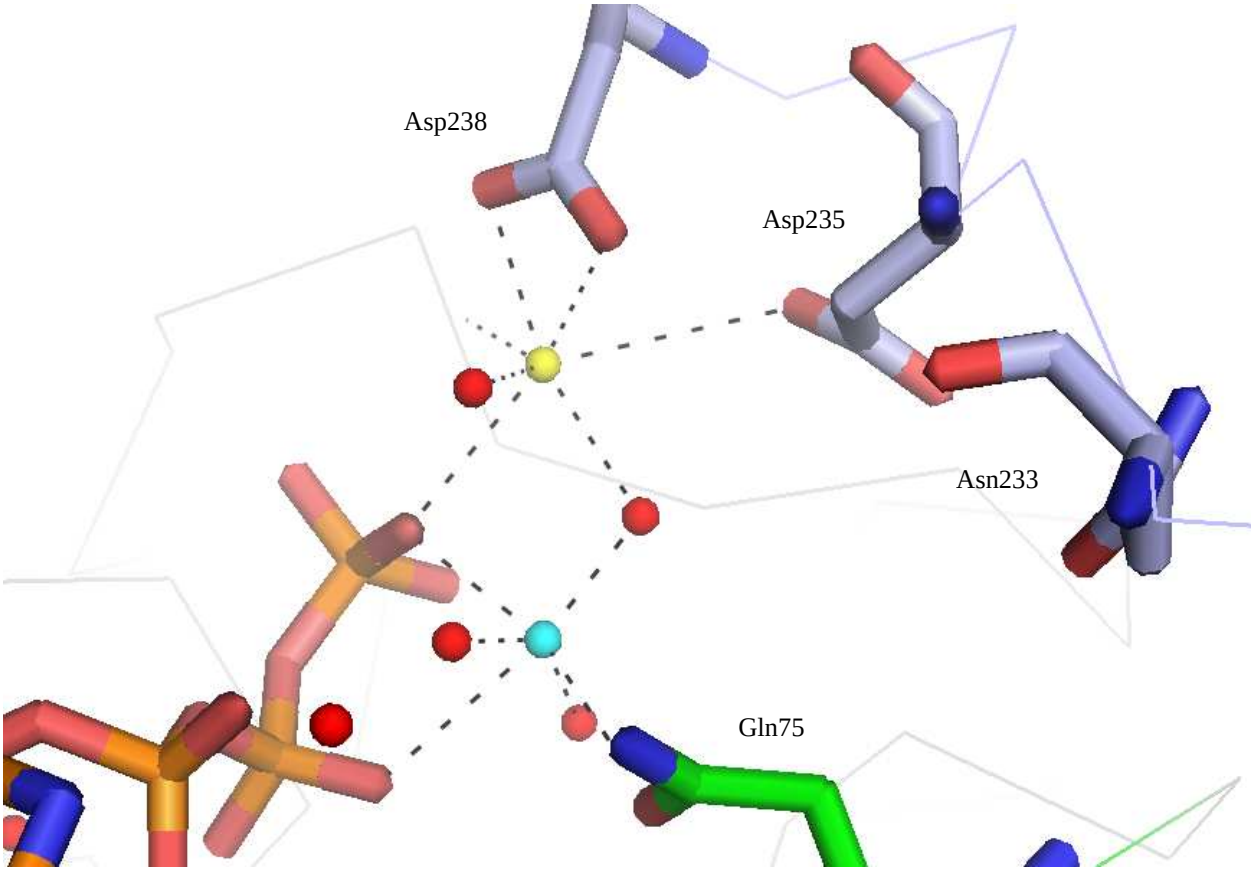


Figure 9: MjanFtsZ crystal structure exhibits the complete GTPase active site by bringing into close proximity Asp235 and Asp238 from the T7 loop (PDB 1W5A). Residues Asp235 and Asp238 polarized the attacking water molecule (yellow) and the Mg^{2+} (cyan) coordinates with Gln 75 (green) and polarizes the γ-phosphate. Asn233 is analogous to Asn205 of MtbFtsZ, which is necessary for GTP hydrolysis. Black dash lines depict hydrogen bonding and red spheres represent water molecules.

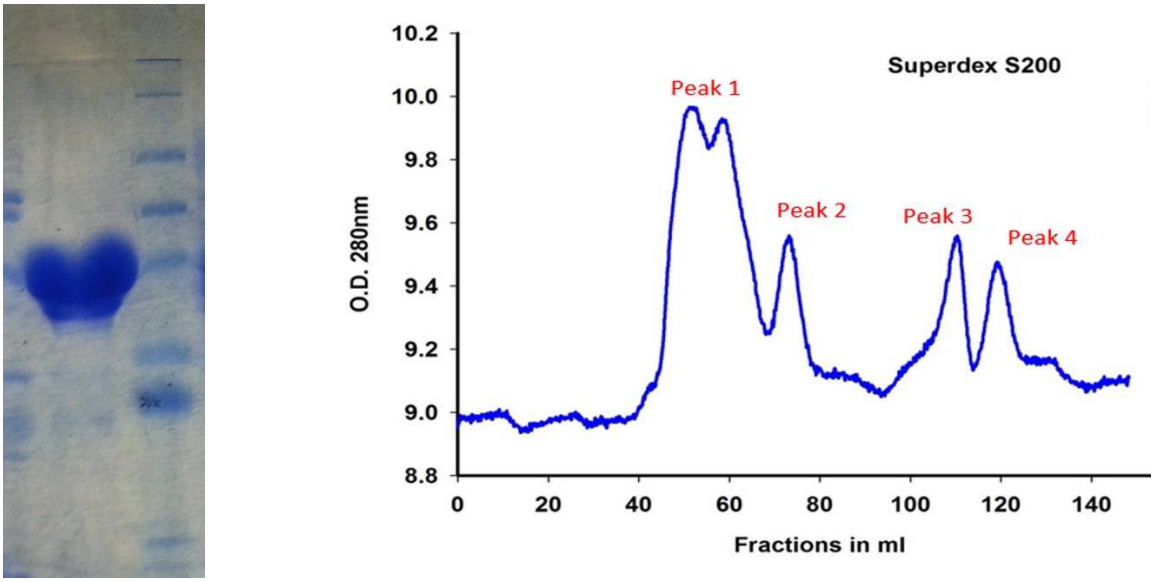
Here we present a set of crystal structures of MtbFtsZ that will aid in new drug development. They are the basis for computational studies to develop novel anti-TB agents.

Chapter 2: Materials and Methods

Protein Purification

Protein purification was carried out by Dr. Souyma Chowdhury, who was part of the Ojima group of the Institute of Chemical Biology and Drug Discovery (ICB&DD) at Stony Brook University (SBU) and who also kindly provided the protein for these crystallization trials and taught me the various steps in protein purification, as reported in White *et al.*[32]. *E. coli* expression plasmids constructs that carry the FtsZ gene from Mtb were used to grow bacteria expressing MtbFtsZ. In order to grow these colonies, selective Luria-Bertani (LB) media was used. Protein expression was induced with 1mM IPTG (isopropyl β-D-1-thiogalactopyranoside) and cells were pelleted and flash frozen for storage. Cells were later thawed and suspended in 50mM Tris pH 7.5, 500mM NaCl, 100mM KCl, 0.1% NP-40 and lysed in a cell disrupter. Centrifugation was performed and clear lysate was added to Ni^{2+} His-bind resin. Protein was washed in 50mM Tris pH 7.5, 300mM NaCl, 100mM KCl, 0.1% NP-40, and 10mM Imidazole and eluted in 50mM Tris pH 7.5, 500mM NaCl, 100mM KCl, and 500mM imidazole. The protein was concentrated to 3mg/ml and checked using Bradford assay. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was conducted to determine protein purity (figure 10A). The N-terminal 6xHis tag affinity was removed with thrombin treatment (0.25 units biotinylated thrombin/ mg tagged FtsZ protein). To remove biotinylated thrombin, uncut FtsZ protein, and free cut off affinity tag were removed by successive passes through streptavidin agarose and fresh Ni^{2+}charged His-bind resin. Further filtration was performed by size exclusion chromatography using an Akta driven Superdex S200 60/16 column in 50mM Tris pH 7.8, 200mM NaCl, 100mM KCl storage buffer (figure 10B). The protein was then flash

frozen and taken to Brookhaven National Laboratory (BNL) for crystallization trials and diffraction experiments.



A B

Figure 10: A) SDS-PAGE gel that shows most of the protein is 40kDa with very few contaminants. B) FPLC profile of MtbFtsZ without the His tag. Peak4 was the peak used in this study and represents a dimeric form of 80kDa. The rest of the peaks represent higher order structural forms (protomer weight of MtbFtsZ is ~40kDa) that are greater than 80kDa. Figures kindly provided by Souyma R. Chowdry.

Protein X-ray Crystallography

Protein crystallography is a major "tool" enabling 3 dimensional structure solutions from a crystal[33]. The major steps in protein X-ray crystallography are, protein purification, crystallization, irradiation to X-rays, phase determination, followed by model building and refinement (figure 11)[33]. It is the most often used technique to determine structures; in fact over 88% of all structures deposited in the protein data bank (PDB) have been solved by X-ray crystallography and 10.5% by nuclear magnetic resonance (NMR)[34]. The remaining <0.5% are solved by techniques such as SAXS, electron microscopy, and neutron diffraction[35].

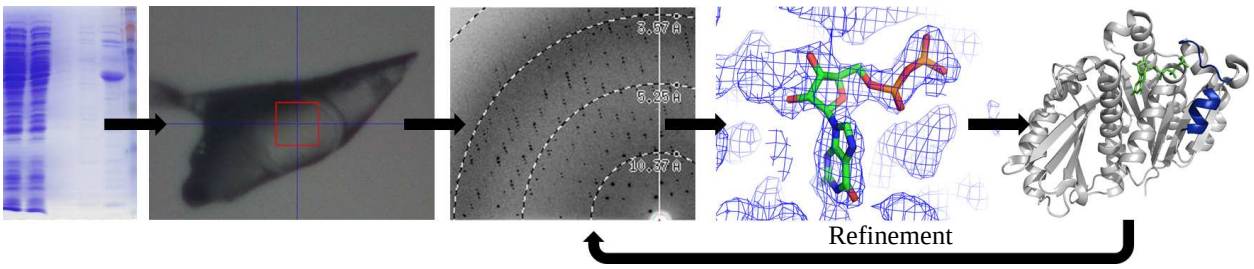


Figure 11: The major steps in protein X-ray crystallography. From left to right: protein purification, crystallization, irradiation to X-rays, phase determination, model building, and cycles of refinement.

X-ray crystallography is a technique used to determine the position of atoms or molecules in a crystal[33]. Monochromatic X-rays in the range of 5 to 20keV are used to irradiate a sample[36]. At these energies, more than 90% of the X-rays do not interact with the sample[33]. The majority of X-rays that interact with the sample is through the photoelectric effect, which deposits energy and is a direct cause of radiation damage[37]. In order to extend the lifetime of the sample, data are usually collected at 100 K[38]. The remaining X-rays interact by Compton scattering which causes background noise and Rayleigh scattering[37]. In the latter, electrons interact with the incident photons and are scattered by planes (*hkl*) to produce a diffraction pattern[37]. The relationship between electron density and a diffracted wave is represented by equation 1, which includes the structure factor amplitude, $|F(hkl)|$ and the phase, [33]. The intensity of the diffracted beam, I ("spots" observed in the diffraction pattern) of a rotating crystal in the incident X-ray beam is measured and thus $|F(hkl)|$ is obtained (equation 2)[36]. However, is lost during the experiment and indirect methods are required; this is known as the phase problem[39].

Equation (1) –

In the above equation, is electron density at position and V is unit cell volume[33].

Equation (2) $I = \left(\frac{\lambda}{wV^2}\right)^3 \left(\frac{e^2}{mc^2}\right) V_{\mathrm{cr}} I_0 LPT|F(hkl)|^2$

In equation 2, λ is the wavelength (energy) of the incident wave, *e* the electronic charge, *m* the electronic mass, V_{cr} the volume of the crystal and V the volume of the unit cell; *L* and *P* are the Lorentz and polarization correction factors and *T* refers to the proportion of X-rays that are not absorbed by the crystal[36].

Among the different methods used to solve the phases, molecular replacement (MR) is the most often used and it was the method used to determine the structures presented here[34].

MR uses a known molecular model to aid in solving the unknown crystal structure[40]. The known model usually has a sequence identity greater than 25% and the α-carbons typically have a root mean square deviation (rmsd) of 2Å or less[39]. To perform MR the model structure needs to be rotated and then translated in the unknown unit cell. To rotate the model structure, three angles need to be specified (α, β, γ) and to translate; three vectors need to be provided (a, b, c). Thus, if there is one molecule in the asymmetric unit cell, which is the smallest unit of volume that contains all structural information that can reproduce the unit cell, MR becomes a 6 dimensional problem[39-41]. Programs split this 6-dimensional problem to two 3-dimensional problems, rotation and then translation[42]. The solution is the best match between the predicted structure factors (calculated model) and the observed structure factors (unknown model). Then the initial model is refined against the observed structure factor combined with the calculated phases until convergences of the R-factor and R_{free}. The R-factor also known as the R_{work} is a statistic that is used to determine the error in a data set[43]. The R_{free} is used in conjunction with R_{work} to make sure there is no bias towards the model that was used to determine the initial phases[44].

Crystallization

Protein crystallization was developed in the 19th century to provide a means for purification of specific proteins at a time when there were few other alternatives[45]. Protein crystallization served as a test to verify that the sample was purified and was a laboratory curiosity[46]. There are several factors affecting protein crystallization such as protein concentration, precipitant type/concentration, pH, temperature, time, and many more including the method used; this is why crystallization is the "bottle neck" in protein crystallography[45]. Vapor diffusion is a common method used to crystallize macromolecular molecules (figure 12)[47]. This method has a closed system which includes the reagent solution and the drop, composed of protein and reagent solution in various quantities and combinations. The concept behind vapor diffusion is rather simple. Water leaves the drop and travels to the reagent solution by passive diffusion. This, in essence, increases the protein concentration and precipitant concentration, and if the nucleation zone is reached, crystals may begin to grow[45].

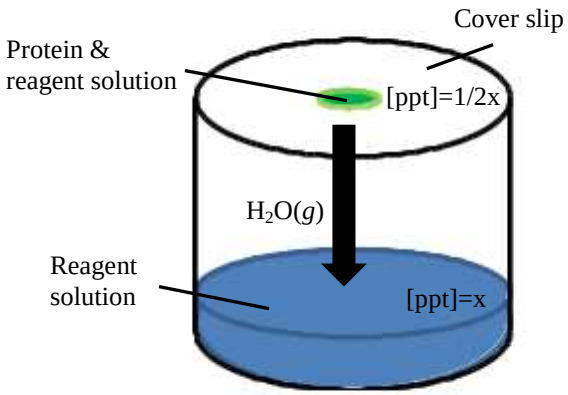


Figure 12: Hanging drop vapor diffusion method. The drop contains 1:1 protein to reagent solution. The concentration of the precipitant [ppt] is one half in the drop with respect to the reagent solution. Water vapor travels from the drop to the reagent solution to equilibrate the [ppt], which results in supersaturation of the protein. This forces the protein to interact and nucleation begins.

MtbFtsZ was crystallized by the hanging drop vapor diffusion method (figure 12)[48, 49]. Crystallization condition reported by Leung *et al.* was used as the initial condition[50]. Several crystallization conditions were screened. The crystals presented here were obtained in 0.1M NaCitrate pH 5.6, 0.3M NH_4OAc, 15% PEG 4000 with a volume of 500µL and the protein, which had a concentration of 3mg/ml, was incubated with the small molecules of interest which are SB-P17G-A20, SB-P17G-C2, and SB-RA-5001 (although these crystals contained GDP, no GDP was added to the crystallization conditions). The way in which these small molecules where incubated was prior to forming the drop, a small volume of drug was added to the protein from a stock solution to vary the concentration of drug; the concentration range was from 0.1mM to 5mM (table 1). This was repeated for all three molecules. It is worth mentioning that it was not until the addition of these small molecules was that crystals began to diffract. Soaking was also attempted for all these drugs. However, none of the crystals that were soaked had good data statistics. SB-P17G-A20, SB-P17G-C2 were synthesized as described in Awasthi *et al.*[20] and kindly provided by Divya Awasthi in Ojima's laboratory. Briefly, these compounds were synthesized by a series of reactions including substitution, acylation and cyclization. The starting material was 2,4-dinitro-5-fluoroaniline, a commercially available compound. The final products were derived from an intermediate and purified via chromatography. SB-RA-5001 was also provided by Divya Awasthi and purified according to Huang *et al.*[17]. The drop was a 1:1 reagent solution to protein ratio. Crystallization was performed at 20ºC and placed in incubators to keep the temperature within +/-1ºC. Initial crystals, which were grown in solution without drug, grew to ~100µm but did not diffract (figure 13A). Additional screening was performed using seeding techniques (figure 13B) and additive screening solutions from Hampton Research (HR2-420)[51, 52]. These techniques aided in obtaining crystals of "higher diffraction quality" (figure 14). The

crystals were soaked in mother liquor containing 30% glycerol and flash cooled for diffraction studies at the National Synchrotron Light Source located in BNL.

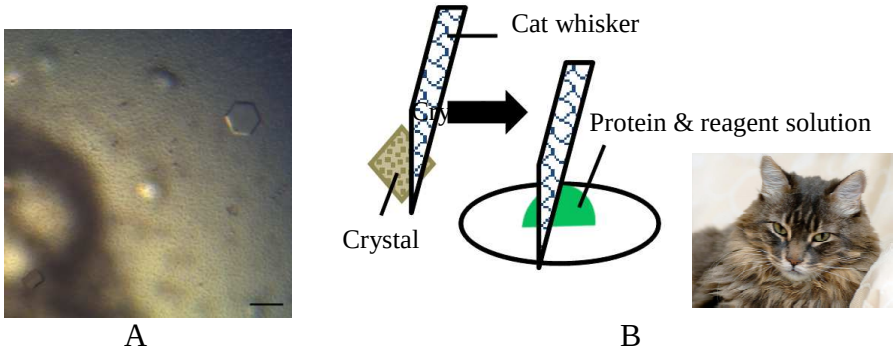


Figure 13: A) Initial crystals of MtbFtsZ that did not diffract grown in 0.1M NaCitrate pH 5.6, 0.3M NH_4OAc, and 20% PEG 4000. Bar scale represents 100µm. B) Seeding technique in which a cat whisker was taken and gently rubbed against the initial crystals and then swiped through a fresh protein drop in an attempt to obtain "higher quality" crystals. Since crystallization is mostly a two-step process (nucleation and growth), seeding allows for the control of nucleation.

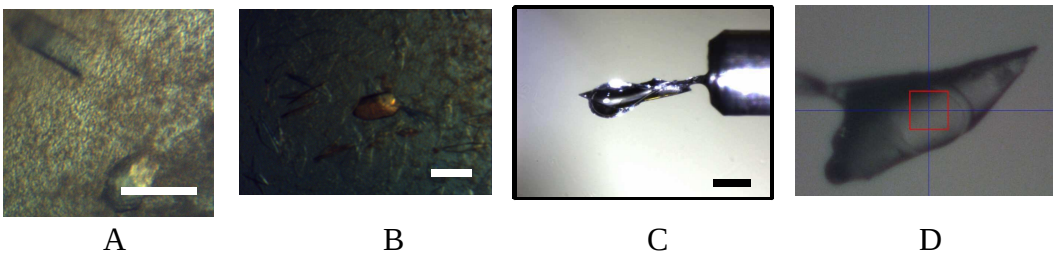


Figure 14: Crystals of MtbFtsZ. All crystals grew in 0.1M NaCitrate pH 5.6, 0.3M NH_4OAc, 15% PEG 4000 with minor differences. The same protein batch was used (no His tag) and the protein concentration was 3mg/ml. Protein with His tag produced poorly diffracting crystals. A) Protein was incubated in 5mM SB-P17-A20. B) Protein from crystal B was incubated in 0.1mM SB-P17-A20. C) Protein from crystal C was incubated in 0.5mM SB-P17G-C2. D) Protein from crystal D was incubated with 0.5mM SB-P17G-A20. Scale bar for A and B represent 100µm, C 150µm. The red box is 100µm in D.

Flash cooling is a technique used in protein crystallography where crystals are submerged in liquid nitrogen (LN2) very quickly and stored until data collection. However, since water forms ice, it is necessary to first transfer the crystal from its crystallization drop to a cryo-solution drop usually made of 30% glycerol[38, 53]. If this is done, data is collected at cryo-

temperatures (100K), which was the case in this study; this also mitigates radiation damage to the biological samples[53].

Data collection strategy

Data for all three crystals A, B, and C were collected at beam-line X6A; a bending magnet beam-line equipped with an Area Detector Systems Corporation (ADSC) Q270 CCD detector [54]. Data were collected at the Selenium K edge (0.9795Å) with a 200µm beam (flux on the order of $1x10^9$ ph/s). The first crystals that grew were crystallized without any compound and were screened manually (~15 samples/ 3hours); none diffracted. Thus, the automounter was used, enabling the testing of several hundreds of crystals (with a screening time of 16 samples/22 minutes) (figure 15). Of all the crystals screened, few gave data of sufficient quality to enable structure solution. Crystals A, B, C, and D represent structures A, B, C, and D, respectively (figure 14).

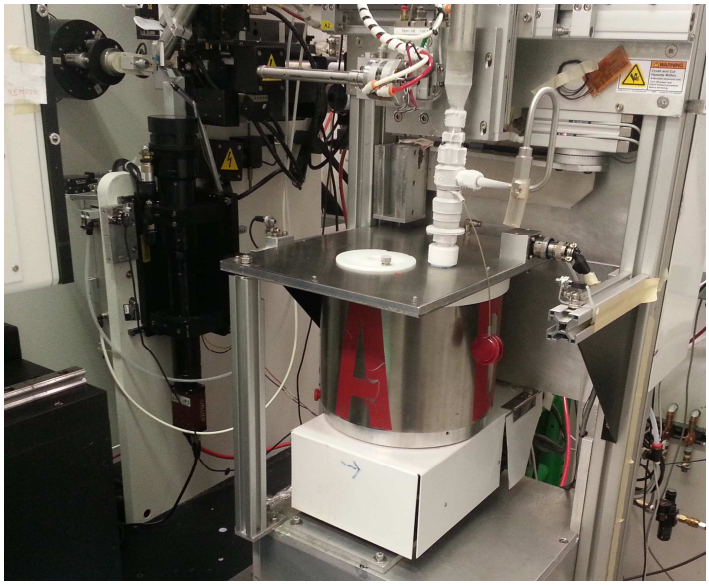


Figure 15: Automounter at the X6A protein crystallography beamline. The automounter's dewar is capable of housing four pucks, meaning it can hold up to 64 samples.

Diffraction data from crystal D were collected at the NSLS on the X25 beam-line. The major differences between X6A and X25 are the detector and the X-ray source. X25 is equipped with a Pilatus detector; it is a silicon pixel array detector working in photon counting mode. This newer technology allows for a faster read-out time and reduced noise, as opposed to the more traditional phosphor CCD detectors. X25 uses an undulator X-ray source producing a brighter beam and as a result X25 has a flux of $4.6x10^{11}$ph/s at 11.5KeV (1.1Å) with a 100µm beam.

The fact that each crystal has different diffraction characteristics is the reason why a crystallographer must use different collection strategies, and in some cases different beam-lines to collect data[45]. Table 1 shows the data collection strategies used for each crystal.

Table 1: Ligands used and collection strategies for each crystal.

Crystal	A	B	C	D
Ligand¥	SB-P17-A20	SB-P17-A20	SB-P17-C2	SB-P17-A20
Method	Co-crystallization	Co-crystallization	Co-crystallization	Co-crystallization
Concentration (mM)	5.0	0.1	0.5	0.5
Wavelength (Å)	1.0781	1.0781	1.0781	1.1000
Detector Distance (mm)	300	260	400	550
Oscillation range (□)	0.5	0.2	0.5	0.2
Beam (µm)	125	150	150x125 (HxV)	100
Exposure time (s)	30	20	30	1

¥ Ligands used in crystallization trials but none were observed.

Data analysis

Data were processed with HKL2000 and statistics are shown in table 2[55]. Crystals diffracted to ~3.5Å and as high as 2.3Å (figure 16). Crystals A and B have better statistics than C and D and have the unit cell parameters of the published structure (PDB 1RQ7). Crystals C and D have different packing as shown by their unit cell parameters and poorer diffracting power (table 2)[56].

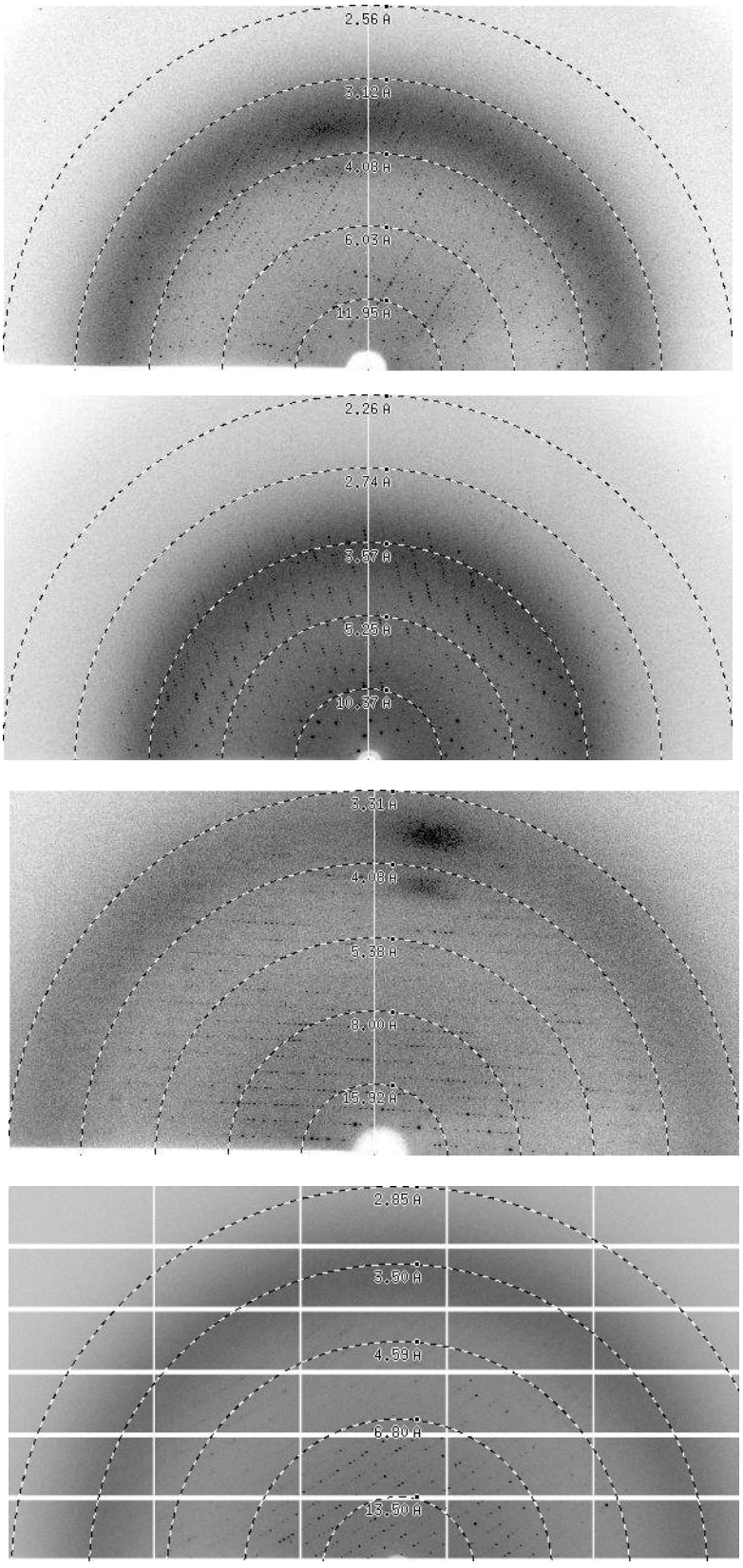


Figure 16: Diffraction of crystals. A, B, C, D (from top to bottom) represents their respective crystal.

Table 2: Crystallization Conditions, Data and Refinement statistics from MtbFtsZ.

Crystal	A	B	C	D
Space group	$P6_5$	$P6_5$	$P2_1$	$P2_12_12_1$
Unit cell parameters (Å, □)	$a = b = 88.84$, $c = 180.1$, $\alpha = \beta = 90.0$, $\gamma = 120.0$	$a = b = 89.16$, $c = 179.8$, $\alpha = \beta = 90.0$, $\gamma = 120.0$	$a = 123.9$, $b = 72.85$, $c = 162.5$, $\alpha = \gamma = 90.0$, $\beta = 99.6$	$a = 73.08$, $b = 181.0$, $c = 220.2$, $\alpha = \beta = \gamma = 90.0$
Matthews coefficient*	2.5	2.6	3.0	3.0
Solvent content (%)*	51.1	52.0	58.3	59.4
Molecules in ASU*	2	2	6	6
Resolution (Å)	30-2.55 (2.59-2.55)	50-2.34 (2.38-2.34)	50-3.57 (3.63-3.57)	50-3.50 (3.56-3.50)
No. of unique reflections	26241	34200	34564	38133
Completeness (%)	98.7 (93.6)	99.3 (98.8)	98.7 (99.0)	99.6 (95.8)
Mosaicity (□)	0.65	1.0	0.95	0.90
Multiplicity	3.1 (3.0)	16 (15.4)	3.6 (3.5)	5.8 (4.1)
R_{linear}† (%)	12.8 (73.3)	9.5 (95.9)	15.7 (79.7)	17.0 (79.1)
□□ *I* □/σ(□ *I* □)□‡	12.7 (1.9)	31.2 (3.5)	9.1 (1.7)	8.9 (1.8)
Refinement statistics				
Resolution (Å)	38.87-2.55 (2.65-2.55)	44.59-2.34 (2.41-2.34)	44.01-3.57 (3.70-3.57)	49.79-3.49 (3.60-3.49)
No. of reflections (working/test)	25839/1318	32058/1624	28159/1422	34660/1730
R_{work}§	17.9 (27.0)	16.9	24.4	27.5
R_{free}¶	23.4 (31.6)	21.8	31.3	35.8
Wilson *B* factor (Å^2)	39.9	22.8	94.4	63.0
Average *B* factor (Å^2)				
Protein atoms	40.6	29.0	120.4	69.2
Water molecules	37.4	28.5	0	0
R.m.s. deviations from ideal††				
Bond lengths (Å)	0.007	0.026	0.012	0.011
Bond angles (□)	1.10	2.68	1.63	1.63
Ramachandran plot‡‡				
Favored regions (%)	98.6	96.6	73.8	74.9
Generously allowed (%)	1.2	3.2	19.2	18.7
Disallowed regions (%)	0.2	0.2	7.0	6.4
Rotamers‡‡				
Poor rotamers (%)	3.85	10.3	19.9	21.9

* Are statistics from Matthews Probability Calculator [56, 57] † $R_{linear} = \sum_{hkl}\sum_i |I_i(hkl) - < I(hkl) > |/\sum_{hkl}\sum_i(hkl)$ [58]. ‡ □□ *I* □/σ(□ *I* □)□ is the mean *I*(*hkl*) over the standard deviation of the mean *I*(*hkl*) averaged over all reflections in a resolution shell. § $R_{work} = \sum_{hkl} ||F_{obs}| - |F_{calc}||/\sum_{hkl}|F_{obs}|$, where $|F_{obs}|$ is the observed structure-factor amplitude and $|F_{calc}|$ is the calculated structure-factor amplitude. ¶ R_{free} is the *R*-factor based on 5% of the data which were excluded from refinement. †† As described in Engh and Huber, (1991) [59]. ‡‡ Are statistics from *MolProbity* [60].

Molecular Replacement

Phases for crystal A were solved by MR using Phaser from the ccp4 suite with PDB 2Q1Y as the search model[42, 61]. The structure from crystal B was also determined by MR using Phaser and structure A was used as the search model. Phenix was used for the refinement of structures A and B[62]. Crystal C needed more analysis and proved to be challenging. Several phasing programs were utilized which included, Phaser, Balbes, and Molrep[34, 42, 61, 63]. Phenix's automated MR program solved the phases and final refinement was done by Phenix[62]. Structure from crystal D was solved by MR using Phaser and structure A was used as the search model as well.

Docking

Docking is a method in which computer simulations predict the preferred orientation of one molecule to another to form a stable complex[64]. When the preferred orientation is known, one can use this to predict binding affinity between these two molecules based on scoring functions. One can think of docking as a "hand-in-glove" analogy where the hand represents the ligand and the glove the receptor[65]. Docking attempts to find the overall best fit. It is often used in drug discovery[66].

In this study, docking was conducted with small molecules (ligand) and our structure D (receptor). The computer calculations were done with the software Autodock4 using the Lamarckian genetic algorithm for conformational searching[67]. A genetic algorithm is a heuristic search that emulates the natural selection process[68]. A Lamarckian genetic algorithm has an added feature which allows individual conformations to search their local conformational space, finding local minima, and passing this information to the next generation[69]. The PDB files of the macromolecule (structure D) and the ligand were preprocessed (converted from PDB to

PDBQT). Hydrogen atoms were added to the structure D file and the number of torsions for the ligand was set to most atoms. The grid box size was set to 126x126x126. The x, y, z, center for the T7 loop region was set to 8.316, 43.334, and 24.257, respectively; which is GDP. For the T9 interaction the grid box center was set to 9.565, 18.275, and -3.123 for x, y, and z, respectively; which is Glu231.

Chapter 3: Results

Structures A, B, C, and D

Structures A and B both crystallized as a dimer. Structure A was superimposed onto structures 2Q1Y, 2Q1X, and 1RQ7 and they exhibited α-carbon rmsd of 0.197, 0.221, 0.176 Å, respectively[35]. There was no structural difference observed except for the rigidity of the T3 loop when compared to 1RQ7, which contained GTPγS. The T3 loop (Switch II) becomes rigid when the nucleotide is GTPγS[29]. When there is an extra phosphate group, a hydrogen network is able to stabilize the T3 loop as described previously (figure 8). This confirms the findings of Leung *et al.*[29].

Subunit A has clear differences from subunit B from structures A and B (figure 17). The most noticeable is the lack of a nucleotide in subunit B. This leads to structural changes of several secondary structures. In subunit B, of structures A and B, the T4 loop collapses inward into the nucleotide binding pocket and helix H8 bends outward ~5Å. Arg140 located in helix H6 has a large movement of 13.7Å. Many of the classic GTPase-activating proteins (GAPS) have an 'arginine finger' that reaches into the active site[70]. It has been suggested that this is to stabilize the charge produced during nucleotide hydrolysis[70]. This movement of the arginine is the reason why helix H6 becomes disordered in subunit B of structures A and B. The switch I region of subunit B from structures A and B is more disordered and takes the form of a loop. According to Leung *et al.*, the switch I region should take the form of a beta sheet and in some cases retain its alpha helix shape as mentioned earlier[29]. Despite this loop form from subunit B in structures A and B, switch I still retains its functionality because loop T3 cannot close when switch I takes the form of a loop (figure 18). These major changes seem to be affected by the absence or presence of GDP.

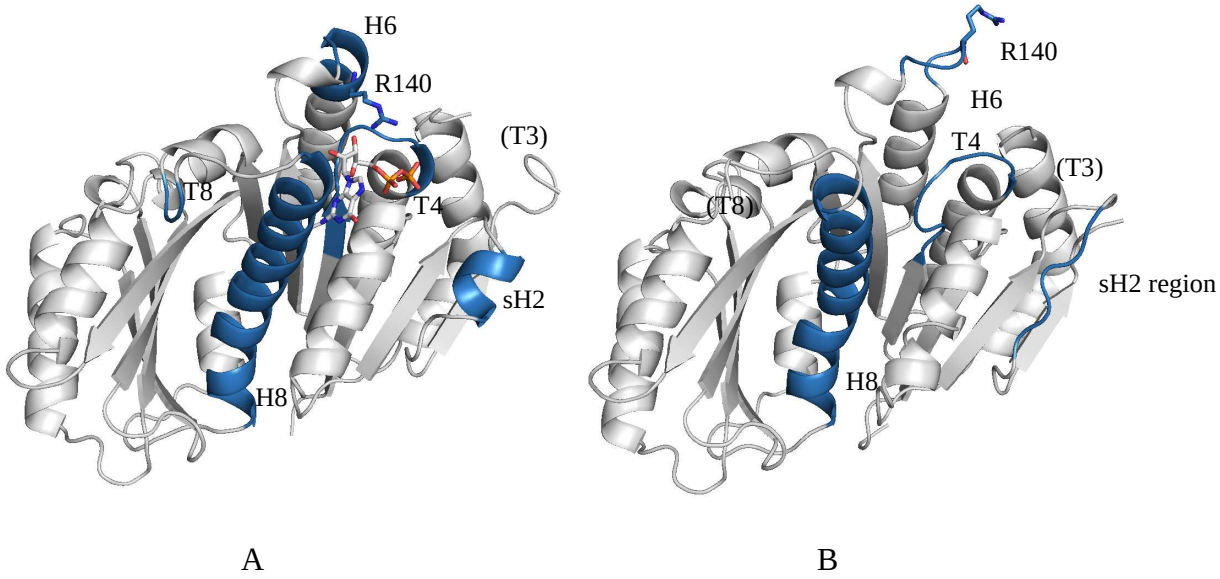


Figure 17: Subunits A (A) and B (B) from crystal A. Structural differences between the subunits shown in blue. Notice the absences of the T8 loop in B and the exaggerated movement of Arg140. The T3 loop is absence in both subunits.

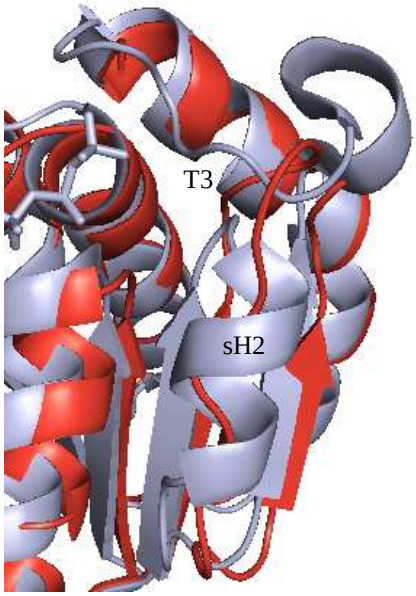


Figure 18: Superimposition of subunit A from 1RLU (light blue) and subunit B (red) of structure A from this study. In subunit B, the sH2 region becomes a loop and the T3 loop needs to be in it's OFF position to avoid clashing.

Structure B was superimposed onto structure A and had an α-carbon rmsd of 0.392 Å, indicating that the structures were virtually the same. To date, structure B is the highest resolution structure of MtbFtsZ with bound GDP (table 2, figure 16). It was crystallized in the same conditions as A, except that it was incubated with 1mM of SB-P17G-A20 instead of 5mM. SB-P17G-A20 is a tri-substituted benzimidazole that has an MIC of 0.16µg/ml against the H37Rv strain of Mtb[20].

Structures C and D exhibited similar interactions as one another. Structure C was crystallized in similar conditions as in A and B except that it was incubated with 0.5mM SB-P17G-C2. SB-P17G-C2 is also a tri-substituted benzimidazole that exhibited an MIC of 0.06µg/ml against the H37Rv strain of Mtb. The first set of crystals were "fished" and diffracted to ~8Å. The drop was clear and 2 weeks later crystals appeared. These new crystals were irradiated and one diffracted to ~3.5Å (figures 14C, 16C), which was adequate to obtain a structure. When structures C and D were superimposed, the α-carbon rmsd was 1.065Å, which is an indication that the structures are very similar. A noticeable difference is that in structure C only one of the central protomers contains GDP (central protomers in structure C and D are B and E). In structure D, both central protomers from each of the trimers contain GDP. It is noteworthy to mention that structure D crystallized in $P2_12_12_1$ and structure C crystallized in $P2_1$. Just like in the drop from structure C, the first set of crystals diffracted to ~8Å.

Initially, structure C was thought to be a dodecamer and was analyzed in the P1 space group. After the determination of structure D, it was clear that structure C had similar interactions. Superimposition to structure D revealed that this was in fact true. Structure C was indexed as a $P2_1$ and a model was built with the aid of structure D (figure 19); it was two trimers just as in structure D.

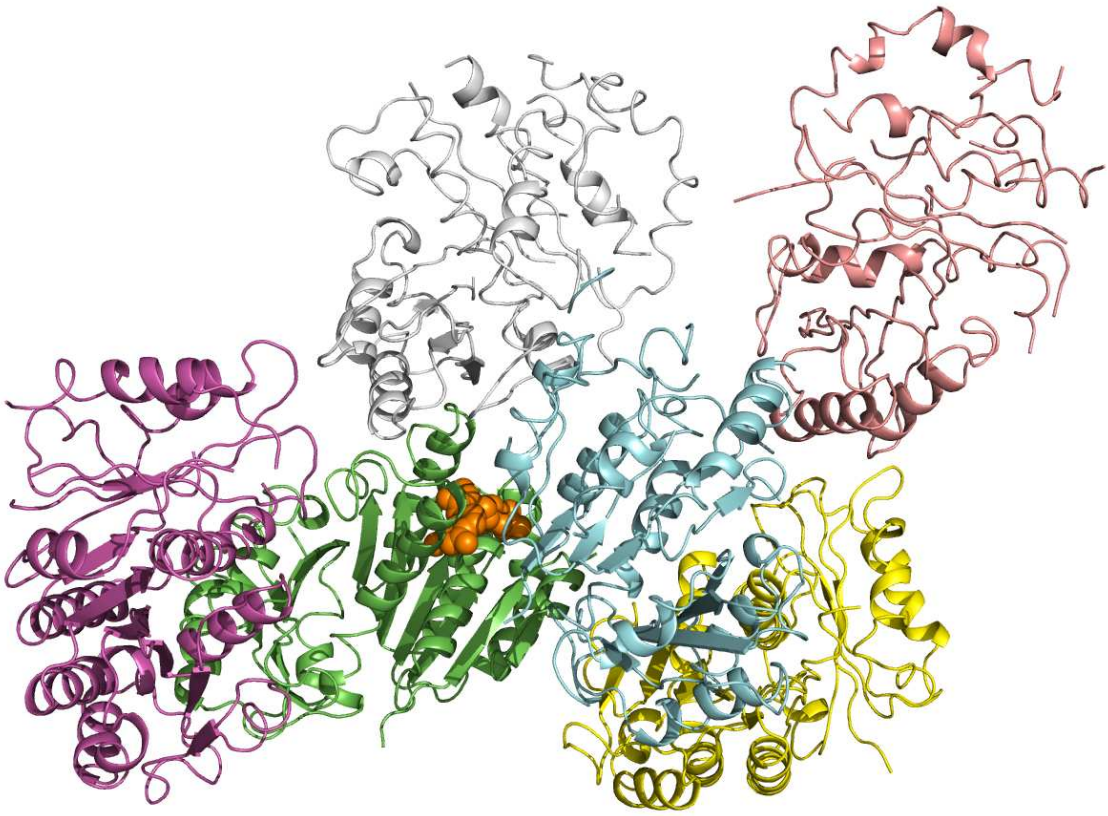


Figure 19: A) Two trimers, ABC and DEF in the asymmetric unit from crystal C. GDP is represent as orange spheres. The six subunits are represented as follows: A is salmon, B is green, C is purple, D is Gray, E is cyan, and F is yellow.

In figure 19, there are two trimers, ABC and DEF, which exhibit similar interactions. These trimers are held together by interactions involving the switch I region in structure C. The highly conserved amino acid Gln45 from chain E interacts with Lys33 from chain A (figure 20). Chain B interacts with chains E and F (figure 21). Gln45 from chain B interacts with Ser298 from chain E. Arg60 from chain B forms polar contacts with Asp51 from chain F and Leu48 from chain B hydrogen bonds to Gln45 from chain F. Given the fact that the highly conserved residue Gln45 from switch I is involve in these interactions, is indicative that they could possess some biological relevance. However, Ser 298 and Arg60 are not conserved. Additional studies will be needed in order to further investigate the relevance of these interactions, including site directed mutagenesis.

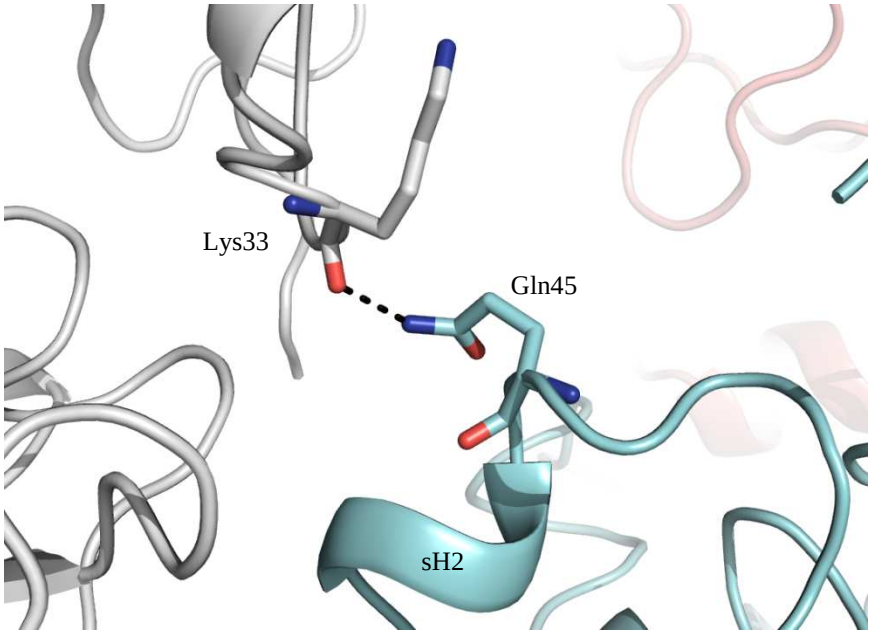


Figure 20: Interactions between the trimers where Gln45 from chain E (cyan) hydrogen bonds to Lys33 from chain A (gray).

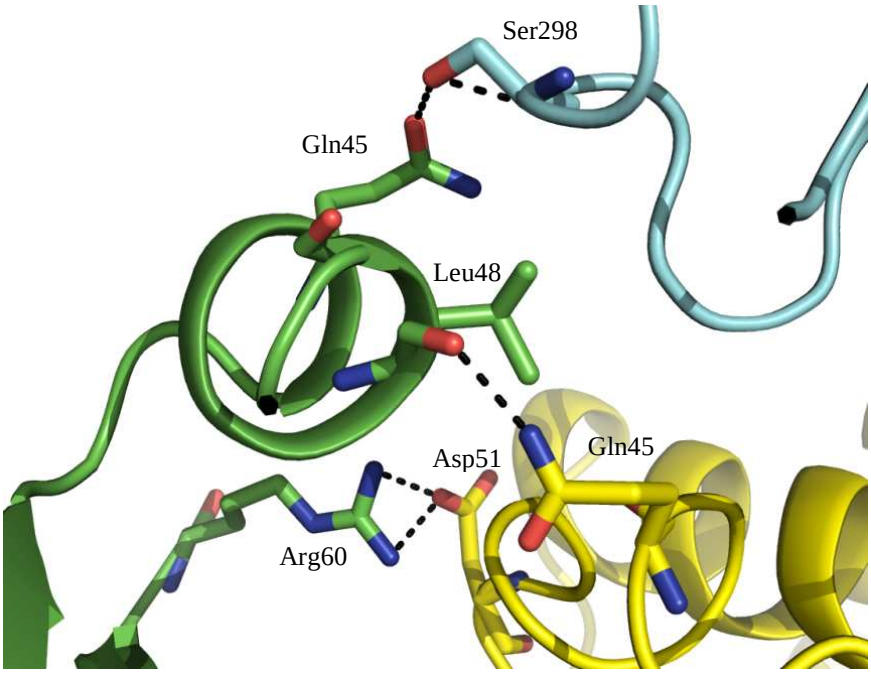


Figure 21: Interactions between the trimers chain B (green), chain E (cyan), chain F (yellow).

The four trimers from structure C and D revealed novel interactions. It was not until recently, that one of them has been reported[71]. Figure 22 represents how trimer ABC from structure D exhibits these interactions. The T7 loop from chain A brings Asn205, Asp207 and Asp210 within 16Å of GDP from chain B (figure 23). These subunits are part of the same trimer. It has been well documented that these residues are required for GTP hydrolysis[23, 24, 29, 72]. Helix H11 from chain A "sits" on helices η1, H7 and loop T6 from chain B. There are no polar contacts involved in this interaction and it resembles a hinge-opening mechanism for the insertion of the T7 loop into the nucleotide binding pocket. This interplay appears to be a conformation following GTP hydrolysis because the nucleotide is GDP. Between subunits B and C, Glu231 from subunit B inserts itself into the nucleotide binding pocket of C (figure 24). In the T9 interaction, chain B inserts the T9 loop located in the C-terminal domain into the empty binding pocket of chain C. Glu231 from the T9 loop fits into the binding pocket and interacts with Gly18 and Gly107. Gly18 is highly conserved throughout all FtsZ proteins and tubulin. Gly107 is part of the highly conserved tubulin signature motif. Glu231 is semi-conserved. This is indicative that this interaction could possess biological validity. Studies have shown that regions from the C-terminal domain do interact with other proteins[73].

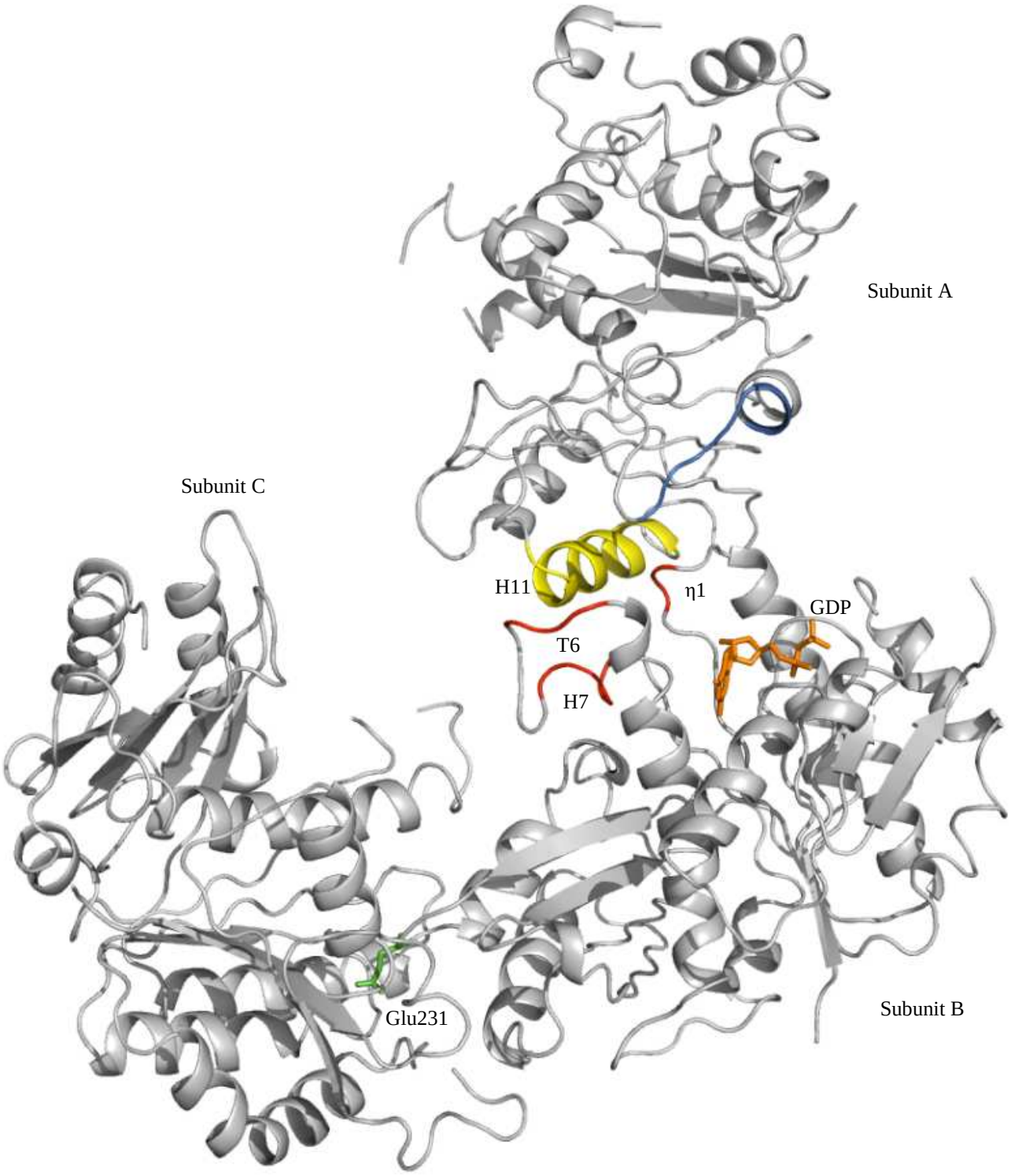


Figure 22: A) Trimer ABC from structure D showing how helix H11 (yellow) "sits" on helices η1, H7 and loop T6 (all shown in red), the T7 loop (blue) within 16Å of GDP (orange), and the T9 loop interaction in which Glu231 (green) is inserted into the nucleotide binding pocket.

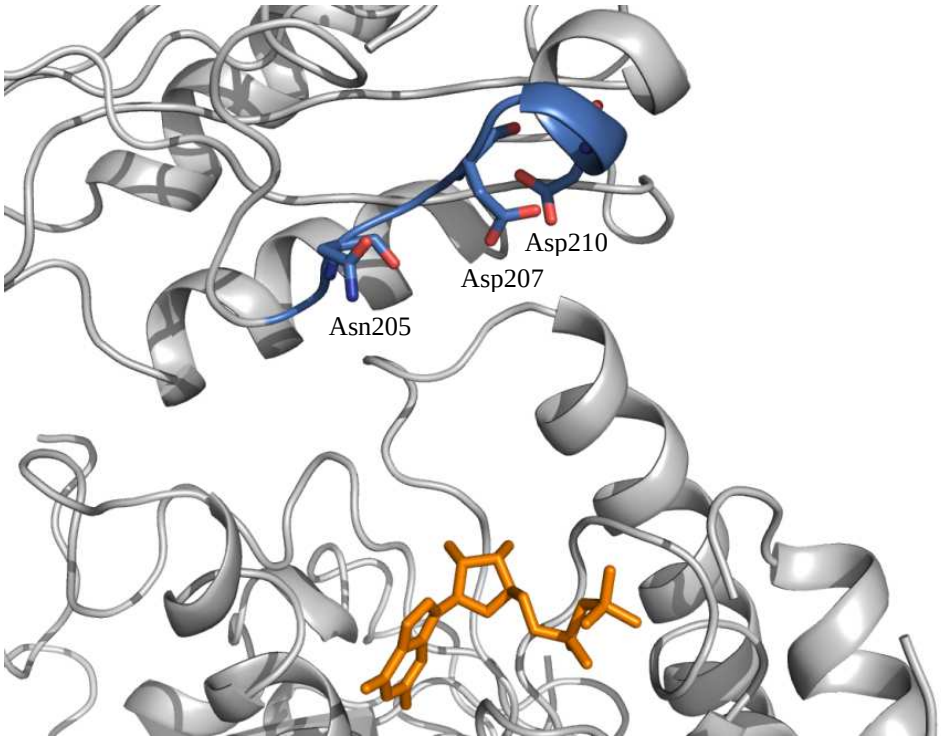


Figure 23: Subunits A and B from structure D, with a close up of the T7 loop interaction. Asn205, Asp207, and Asp210 (blue) from subunit A are within ~16Å from GDP (orange) of subunit B.

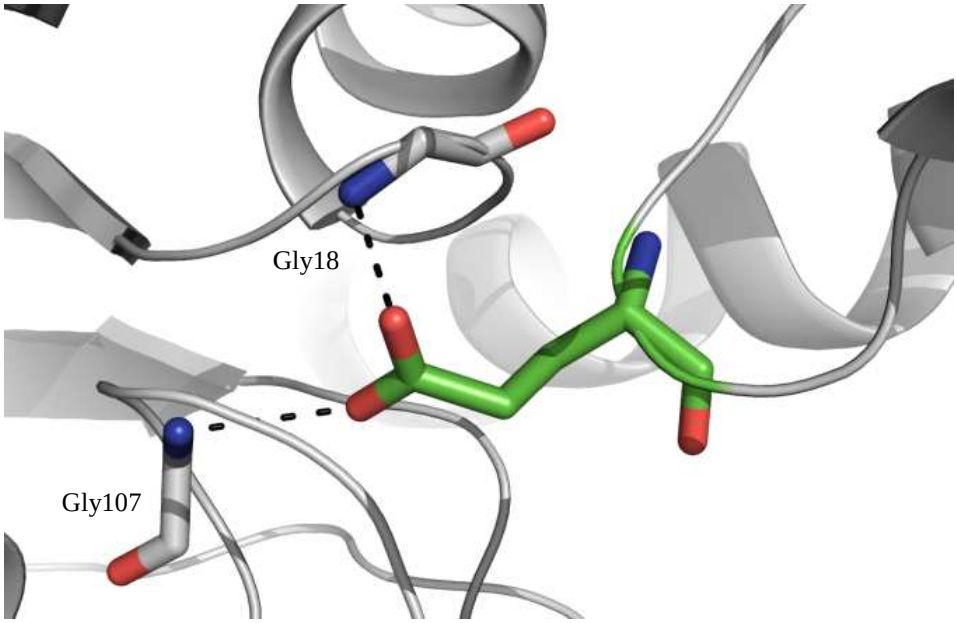


Figure 24: Glu231 (green) from subunit B interacts with Gly18 and Gly107 from subunit C in structure D.

Subunits B and C exhibit some structural differences in structure D (figure 25). The T3 loop (residues 62-68) from subunit C was rigid when compared to B. Residues 170 and 171 from the T6 loop are rigid in subunit A as electron density was clearly observed. The N-terminal end of helix H8 (residues 172-180) is ~ 3Å away from the nucleotide area in subunit C. There is an rmsd greater than 3Å for residues 264-266 from the T10 loop between subunits B and C. This could be explained by the flexibility of the T10 loop in subunit C, where no density was observed.

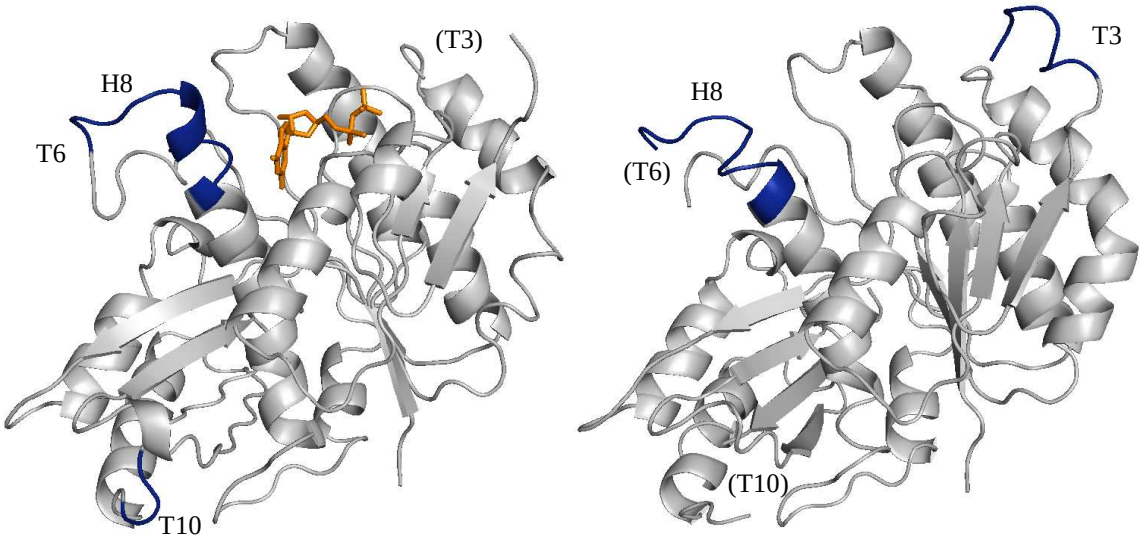


Figure 25: Differences of subunit B (left) and subunit C (right) from structure D shown in blue.

Subunit A from structure D had many missing features due to crystal packing. However, the data was enough to obtain a model. When superimposed to subunit B, there were regions of interest where the rmsd was greater than 3Å. Most notably were the highly conserved Gly18 which has an rmsd of 3.8Å and Gly105 of the tubulin signature motif that has an rmsd of 3.2Å.

Subunits A and C were very similar. They only had one residue in which the rmsd was greater than 3Å; Gly104, which has an rmsd of 3.2Å. The overall structure has an α-carbon rmsd of 1.0Å. This is less than subunits A and B, which have α-carbon rmsd of 1.05Å, respectively.

This indicates that subunits A and C are very similar, but subunit B is the least similar when compared to subunits A and C.

Docking Results

Computational studies using the Autodock program were performed focusing on two regions of structure D. The first region was the novel T9 interaction which involves subunits B and C. The second is the T7 region which is between subunits A and B. The simulations revealed that compound SB-P17G-A20 is nestled within helices H3, H5, and H6 in the T9 region and forms a hydrogen bond with Asn142 from chain C with a binding energy of -7.73kcal/mol (figure 26). In the T7 region, the compound interacts with Leu166 from helix H8 with a binding energy of -7.65kcal/mol (figure 27). Compound SB-RA-5001, which is a taxane that exhibits anti-TB activity, was also docked in these two regions[17]. Crystallization was attempted with this compound but no crystals were of high diffraction quality. In the T9 area, SB-RA-5001 interacts with Ile225 and Asn189 with a binding energy of -6.97kcal/mol (figure 28). In the T7 area, SB-RA-5001 forms no polar contacts and lays near GDP. It has a low binding energy of -5.43kcal/mol (figure 29).

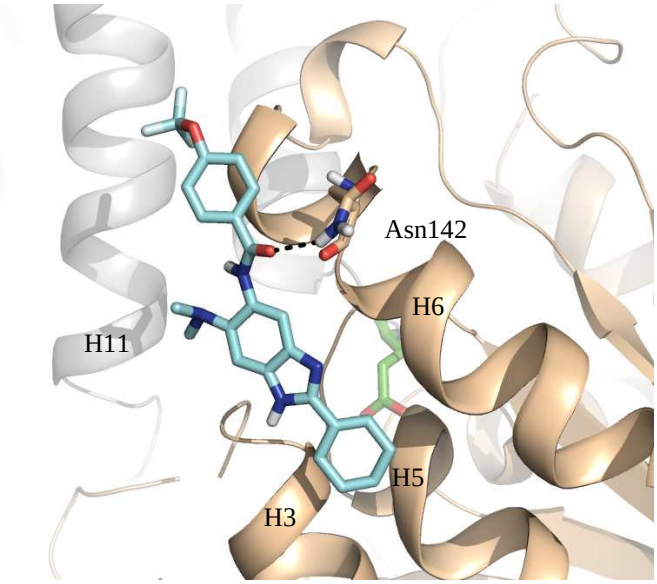


Figure 26: Results from Autodock. SB-P17G-A20 (cyan) interacts with Asn142 of chain C (wheat) from structure D in the T9 region. Glu231 from chain B (gray) is shown in green where it is inserted into the nucleotide binding pocket of chain C.

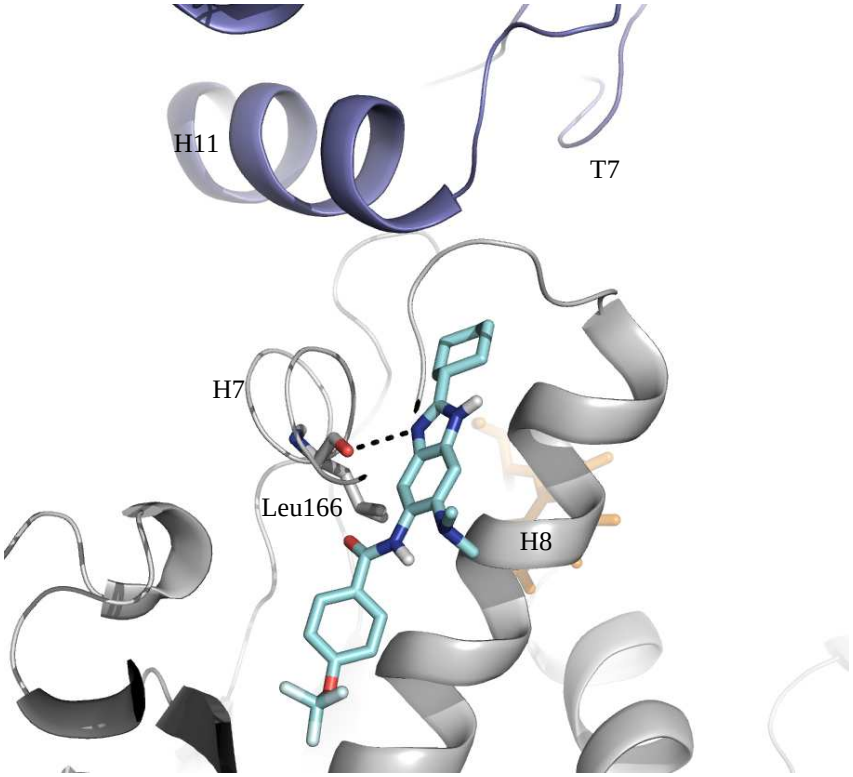


Figure 27: SB-P17G-A20 (cyan) interacts with Leu166 from chain B (gray) in the T7 region, which is between chain A (blue) and chain B. GDP is shown in orange in the nucleotide binding pocket.

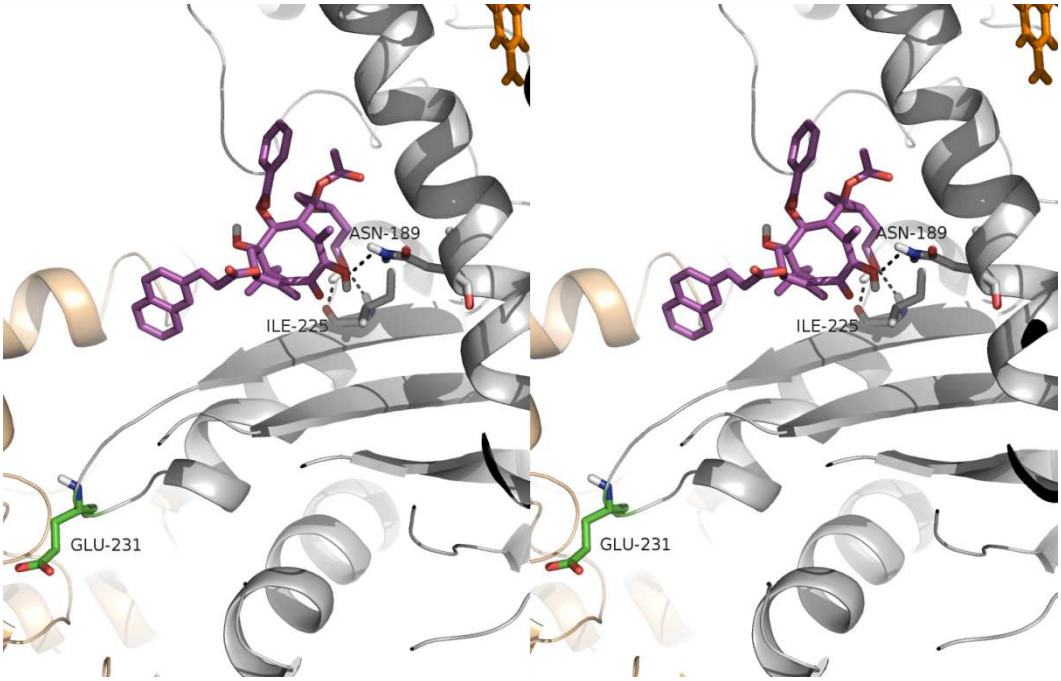


Figure 28: Stereo-view of Autodock results of SB-RA-5001 (purple) interacting with Asn189 and Ile225 from helix H8 and beta sheet β7, respectively in the T9 region. Glu231 is shown in green, GDP is in orange, chain B is in gray, chain C is in wheat.

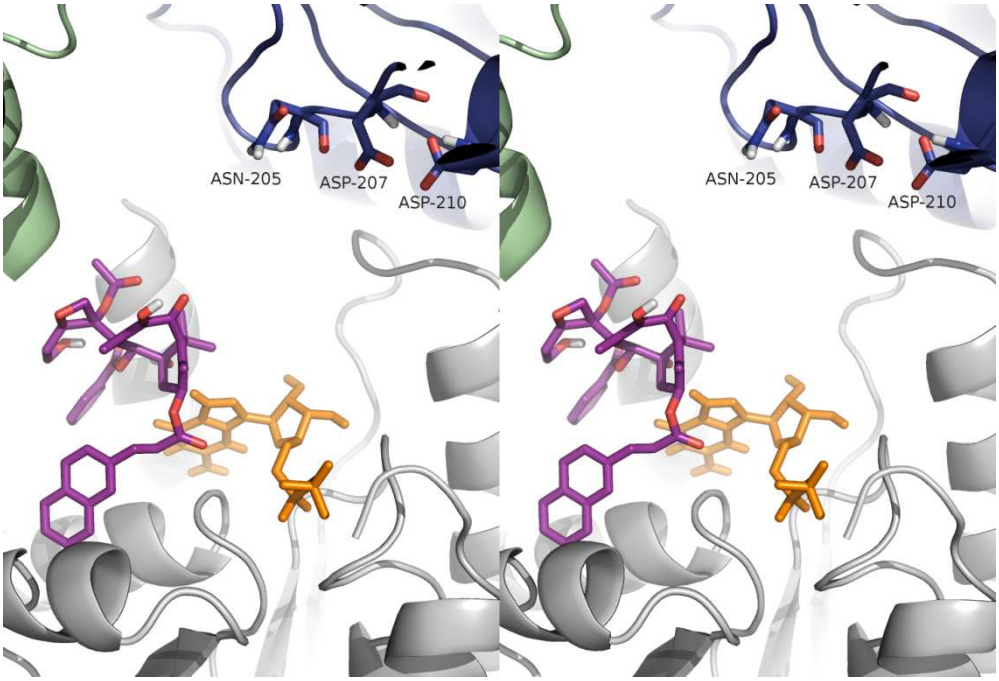


Figure 29: Stereo-view of docking results of SB-RA-5001 (purple) in the T7 region. SB-RA-5001 lays near GDP (orange) and forms no polar contacts.

Chapter 4: Discussion and Perspectives

Discussion

As previously mentioned, initial crystals were irradiated with X-rays and did not produce a diffraction pattern. To make sure this was not an isolated case, approximately 30 crystals were screened, which all yielded the same result. It was not until compounds SB-P17G-A20 and SB-P17G-C2 were included in the crystallization condition did crystals produce diffraction patterns. This is indicative that despite not being observed, these potential drugs may interact with the protein.

Structures A and B are consistent with previous studies[29]. Both support the fact that the T3 loop is a switch involved in GTP hydrolysis because in these structures, no electron density is observed for the T3 loop when GDP is bound (subunits A) and when there is no nucleotide (subunits B), indicating high flexibility. However, if GTP is the nucleotide, there is electron density observed for the T3 loop as reported by Leung *et al.* (2004)[29].

The other two structures show conformational differences. The hinge-opening observed in structures C and D is consistent with a recently deposited structure (PDB 4KWE). The T9 loop interaction is however a new interaction, never reported before.

Lu *et al*., (2000) proposed that FtsZ could provide the motile force for constriction of the Z-ring [28]. In their study, they determined that GTP bound FtsZ preferred a straight conformation and GDP bound FtsZ a curved conformation (figure 3). Li *et al.*, (2013) further supports Lu *et al.*, findings[71]. They determined the crystal structure of a single curved MtbFtsZ protofilament (PDB 4KWE). Superimposition demonstrates that structures C, D and 4KWE exhibit similar hinge-opening interaction (figure 30). Mutagenesis studies determined that this interaction is biologically relevant and that Leu269 is a key residue that serves as a pivot. In their model

(figure 31A), the T3 loop of GDP bound MtbFtsZ takes a “relaxed” (R) conformation while GTP bound MtbFtsZ has a “tense” (T) conformation. When the T state monomers assemble, the T7 loop “cuts” the γ-phosphate and coordinates with the T3 loop. The T3 loop is now in the R state and MtbFtsZ pivots at Leu269. MtbFtsZ is bound to the membrane via FtsA and is able to pull the membrane inward (figure 31B). Free energy calculations show that this mechanism generates enough force for constriction.

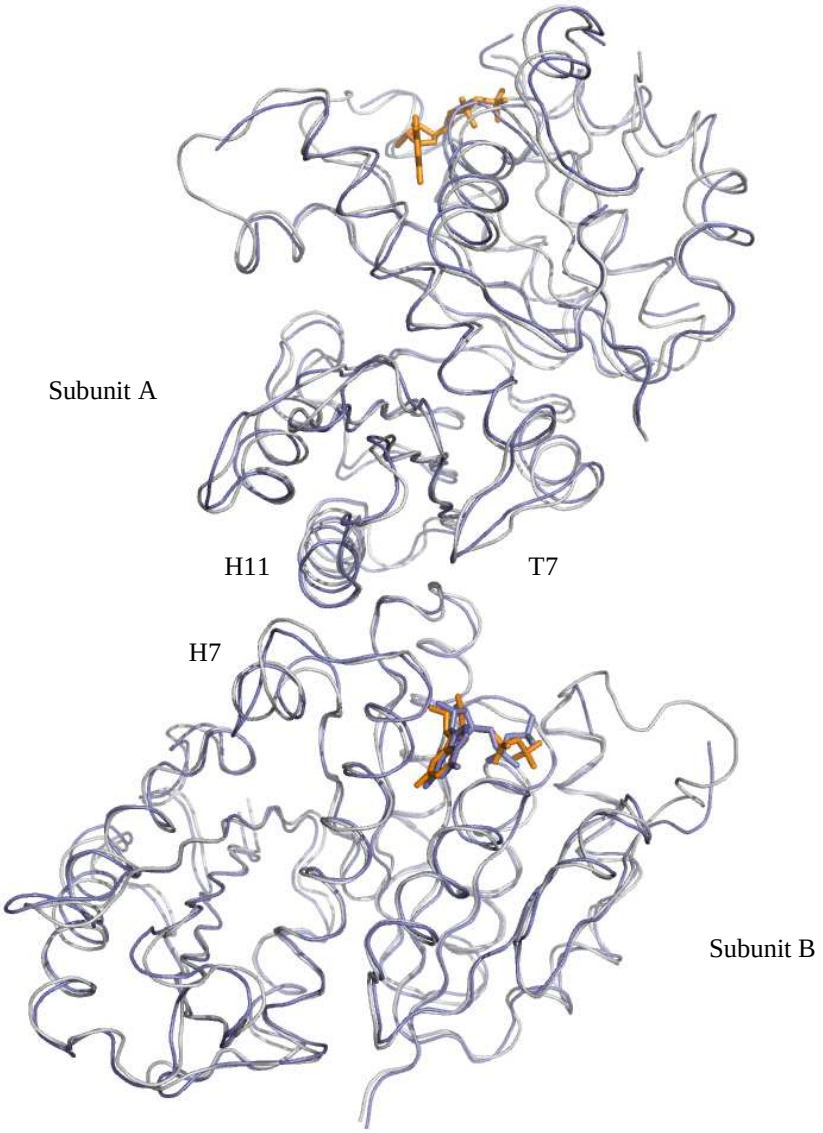
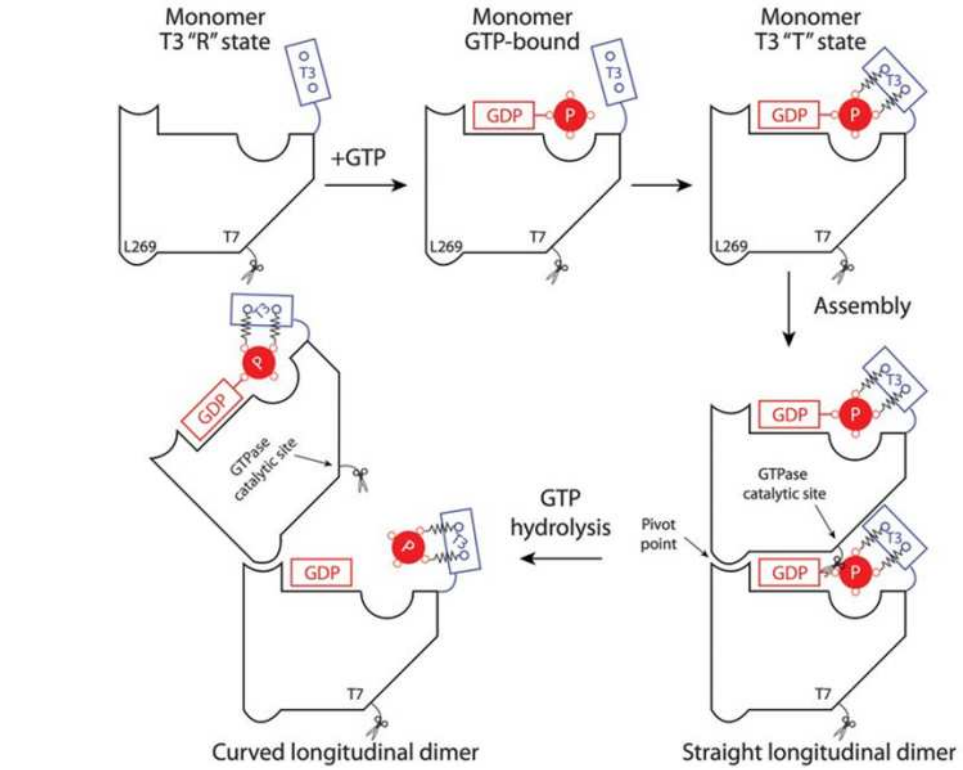
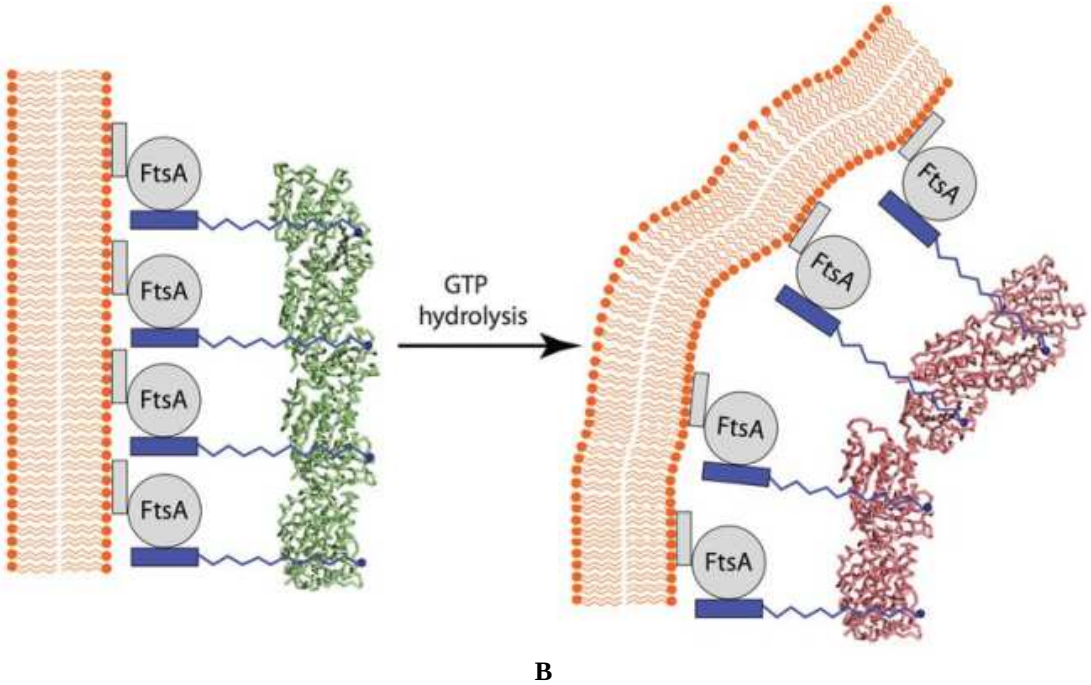


Figure 30: Superimposition of structure D subunits A and B (blue) onto subunits A and B of PDB 4KWE (gray). GDP is shown in orange for 4KWE and blue for subunit B from structure D.



A



B

Figure 31: A) Schematic representation of the steady-state turnover in MtbFtsZ. B) Model representation of straight protofilaments pulling on the cell membrane when they curve. Figures taken from Li *et al.* (2013) [71].

Perspectives

Subunits A and B from structure D and C exhibit the T9 loop interaction (figures 24). FtsZ has been shown to interact with other proteins, including itself[23, 29, 73, 74]. A crystal structure of FtsZ from *M. jannaschii* revealed the insertion of the T7 loop into the nucleotide binding pocket of another FtsZ monomer completing the GTPase active site[23]. MtbFtsZ was crystallized as a dimer that revealed a lateral interaction and a SulA-FtsZ complex in which SulA interacts with the T7 loop to prevent polymerization was also solved[29, 74]. SulA inhibits FtsZ polymerization as a result of an SOS signal. The SOS signal is a mechanism that inhibits cell division due to DNA (deoxyribonucleic acid) damage and begins to repair the damaged DNA[74]. In *E. coli*, ZipA is a protein that anchors FtsZ to the cell membrane[73]. A ZipA-FtsZ complex revealed that this is mediated by 17 residues from the C-terminal domain of FtsZ, residues 367-383[73]. These residues bind to the ZipA/M185 cavity[73]. Recently, MtbFtsZ was crystallized with the T7 loop in relatively close proximity (~16Å) to the nucleotide binding site by Li *et al.* (2013) and us[71]. However, the insertion of the T9 loop into the nucleotide binding pocket is a novel interaction. Gln231 inserts into the nucleotide binding pocket and represents the classical lock-and-key model (figure 24). Residue Gly18 and the nucleotide binding pocket residue Gly107 are highly conserved through all FtsZ proteins and tubulin. Gln231 is conserved throughout all FtsZ proteins. This may be an indication that this interaction has biological validity. To date, not much is known about the lateral interactions of protofilaments from MtbFtsZ. This new interaction may help with understanding lateral interactions between protofilaments as this new interaction could be how protofilaments link to one another. According to Stricker *et al.* (2002), the Z-ring is comprised of protofilaments that have an average length of 80 subunits[75]. In one of the model proposed by Stricker *et al.* (2002), the Z-ring has protofilaments of lengths 40, 80, and 160

subunits. Since the Z-ring acts as a single structure, these protofilaments must interact with one another. From protofilaments, we are proposing a model on Z-ring formation via the newly described T9 interaction (figure 32). An FtsZ monomer interacts with Glu231 from the T9 loop from a protofilament. Another protofilament can now interact with this monomer via the T9 loop interaction again, linking the protofilaments.

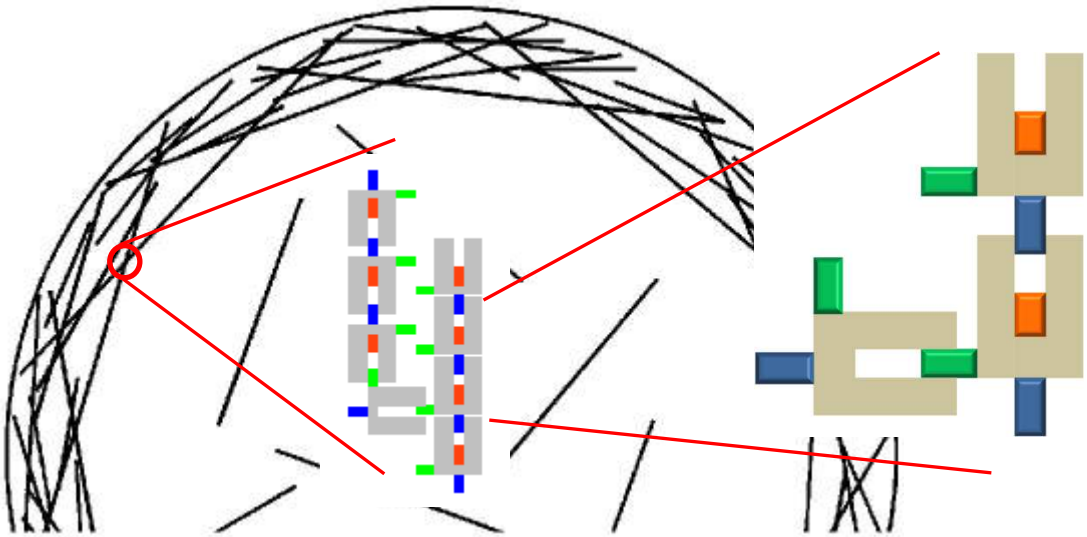


Figure 32: The semi-circle is a transverse cross-section of 200nm at the division, which represents the model proposed by Stricker *et al.* (2002) [75]. The black lines around the semi-circle represent FtsZ protofilament distribution in the Z-ring. The schematic is a close-up of our proposed model (left) which is MtbFtsZ protofilaments interacting via the T9 loop and to the left is a schematic of our trimer from structures C and D. Glu231 from T9 is shown in green, the T7 loop is shown in blue, and GDP is shown in orange.

Docking studies were conducted and interactions that involved conserved residues were of particular interest because conserved residues usually are involved in biological processes. The interaction Asn142 and SB-P17G-A20 from docking calculations does not seem biologically relevant because Asn142 is not conserved (figure 4). On the other hand, the T7 interaction between this compound and Leu166 from helix H7 is semi-conserved (figure 4). Despite Leu166 being semi-conserved, this interaction with SB-P17G-A20 does not seem promising because SB-

P17G-A20 inhibits polymerization. According to docking studies conducted by Wei, L. from Ojima's laboratory, SB-P17G-A20 should interact near the T7 loop region, in close proximity to Asn205, Asp207, and Asp210, which are the residues necessary for GTP hydrolysis. The taxane SB-RA-5001 interacting with Ile225 and Ser244 could be of interest as both these residues are highly conserved (figure 4). Compound SB-RA-5001 also has a different mode of action from SB-P17G-A20 and SB-P17G-C2 in which it promotes polymerization by stabilizing the protofilament structure. The simulation in the T7 region for this compound seems to not be relevant because this taxane does not form any polar contacts and has a low binding energy of -5.43kcal/mol further supporting that the interaction of SB-RA-5001 between Ile225 and Ser244 could have biological validity.

Currently, only two structures have been solved with compounds that stabilized the protofilament which are PDB 3VOB and 1JFF[76, 77]. 3VOB is *Staphylocuccus aureus* FtsZ (SaFtsZ) with PC190723 and 1JFF is tubulin, which is the eukaryotic homologue of FtsZ, with taxol. In SaFtsZ, a novel hydrophobic cleft was found near the T7 loop, in which PC190723 binds to (figure 33). The compound interacts with several residues from the T7 loop, helix H7, and beta sheets β7, β8, β9, and β10. The mechanism in which this compound stabilizes the protofilament is by stabilizing the T7 loop in such a way that the residues necessary for GTP hydrolysis are not in position to allow the reaction to proceed. In the 1JFF structure, the taxol pocket is above the T7 loop and next to helix H7 in the C-terminal domain (figure 34). These structures where superimposed to the docking study which involved the taxane SB-RA-5001, Ileu225, and Asn189 (figure 35). The docking revealed that SB-RA-5001 is in relative close proximity to PC190723 and Taxol. All of the molecules are located above the T7 loop and next to helix H7 (in MtbFtsZ case helix H8) in the C-terminal domain and given the fact that all these

molecules have the same mode of action is a good indication that this preliminary docking study has elucidated a promising base for future computer simulations.

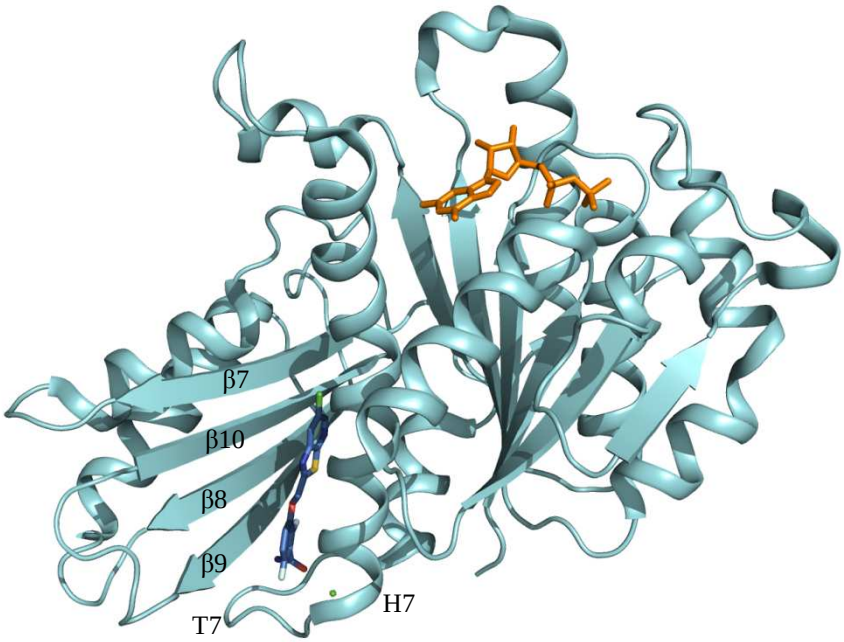


Figure 33: Crystal sturcture of SaFtsZ with bound PC190723 (dark blue) and Mg ion is shown as green sphere. PDB 2VOB (cyan).

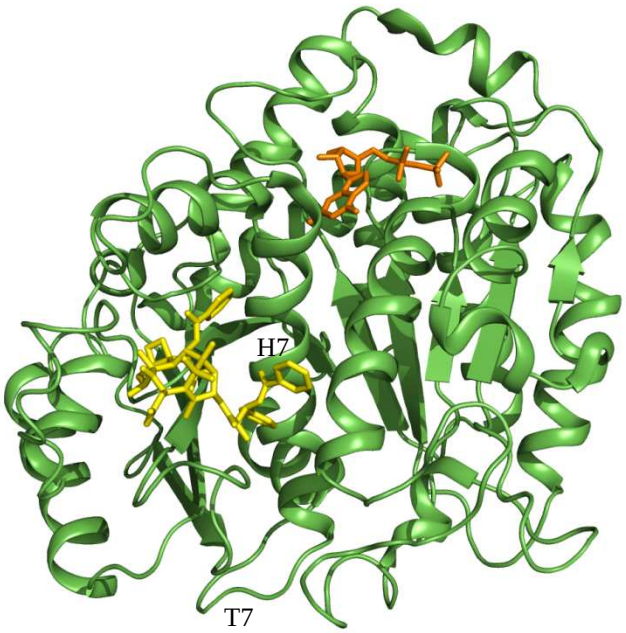


Figure 34: Crystal structure of tubulin with bound Taxol (yellow). PDB 1JFF (green).

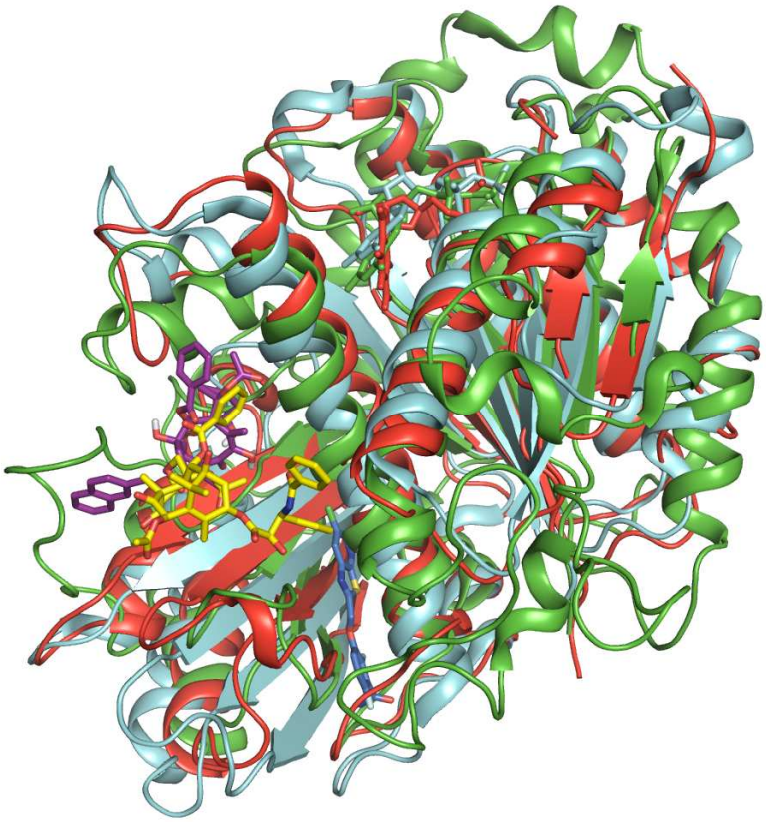


Figure 35: Superimposition of MtbFtsZ (red) this study with SB-RA-5001 (purple), 2VOB, and 1JFF (same color scheme as figure 33 and 34, respectively). GDP is shown as sticks.

In conclusion, this study has revealed new interactions, in which the T9 loop interacts with the nucleotide binding pocket in a lock-and-key fashion and the T7 loop residues: Asn205, Asp207, and Asp210 are in close proximity to the nucleotide. Docking study has elucidated a promising base for future computer simulations. Combined with the novel interactions discovered, this has provided a possible foundation for novel drug development to combat TB.

References:

1. Bloom BR, Murray CJL. Tuberculosis: Commentary on a Reemergent Killer. *Science.* 1992;257(5073):1055-1064.

2. Kumar V. *Robbins basic pathology*. Philadelphia Pa.: Saunders Elsevier; 2007.

3. Lawn SD, Zumla AI. Tuberculosis. *The Lancet.* 2011;378(9785):57-72.

4. WHO | Tuberculosis.

5. Restrepo BI. Convergence of the tuberculosis and diabetes epidemics: renewal of old acquaintances. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America.* 2007;45(4):436-438.

6. Gomez JE, McKinney JD. M. tuberculosis persistence, latency, and drug tolerance. *Tuberculosis.* 2004;84(1):29-44.

7. CDC | TB | Treatment for TB Disease.

8. CDC | TB | Basic TB Facts.

9. Cole EC, Cook CE. Characterization of infectious aerosols in health care facilities: An aid to effective engineering controls and preventive strategies. *American Journal of Infection Control.* 1998;26(4):453-464.

10. Gibson AMP. *Evidence-based respiratory medicine*. Malden: BMJ Books/Blackwell Pub.; 2005.

11. Ackah AN, Digbeu H, Daillo K, Greenberg AE, Coulibaly D, Coulibaly IM, Vetter KM, de Cock KM. Response to treatment, mortality, and CD4 lymphocyte counts in HIV-infected persons with tuberculosis in Abidjan, Côte d'Ivoire. *The Lancet.* 1995;345(8950):607-610.

12. Becerra MC, Bayona J, Freeman J, Farmer PE, Kim JY. Redefining MDR-TB transmission ‘hot spots’ [Counterpoint]: International Union Against Tuberculosis and Lung Disease.

13. Raviglione MC, Smith IM. XDR Tuberculosis — Implications for Global Public Health. *New England Journal of Medicine.* 2007;356(7):656-659.

14. Romberg L, Levin PA. Assembly dynamics of the bacterial cell division protein FtsZ: Poised at the edge of stability. *Annual Review of Microbiology.* 2003;57:125-154.

15. Kumar K, Awasthi D, Berger WT, Tonge PJ, Slayden RA, Ojima I. Discovery of anti-TB agents that target the cell-division protein FtsZ. *Future medicinal chemistry.* 2010;2(8):1305-1323.

16. Slayden RA, Knudson DL, Belisle JT. Identification of cell cycle regulators in Mycobacterium tuberculosis by inhibition of septum formation and global transcriptional analysis. *Microbiology (Reading, England).* 2006;152(Pt 6):1789-1797.

17. Huang Q, Kirikae F, Kirikae T, Pepe A, Amin A, Respicio L, Slayden RA, Tonge PJ, Ojima I. Targeting FtsZ for antituberculosis drug discovery: noncytotoxic taxanes as novel antituberculosis agents. *Journal of medicinal chemistry.* 2006;49(2):463-466.

18. Respicio L, Nair PA, Huang Q, Anil B, Tracz S, Truglio JJ, Kisker C, Raleigh DP, Ojima I, Knudson DL, Tonge PJ, Slayden RA. Characterizing septum inhibition in Mycobacterium tuberculosis for novel drug discovery. *Tuberculosis (Edinburgh, Scotland).* 2008;88(5):420-429.

19. Li H, DeRosier DJ, Nicholson WV, Nogales E, Downing KH. Microtubule Structure at 8 Å Resolution. *Structure.* 2002;10(10):1317-1328.

20. Awasthi D, Kumar K, Knudson SE, Slayden RA, Ojima I. SAR studies on trisubstituted benzimidazoles as inhibitors of Mtb FtsZ for the development of novel antitubercular agents. *J Med Chem.* 2013;56(23):9756-9770.
21. White EL, Suling WJ, Ross LJ, Seitz LE, Reynolds RC. 2-Alkoxycarbonylaminopyridines: inhibitors of Mycobacterium tuberculosis FtsZ. *J Antimicrob Chemother.* 2002;50(1):111-114.
22. Reynolds RC, Srivastava S, Ross LJ, Suling WJ, White EL. A new 2-carbamoyl pteridine that inhibits mycobacterial FtsZ. *Bioorg Med Chem Lett.* 2004;14(12):3161-3164.
23. Oliva MA, Cordell SC, Löwe J. Structural insights into FtsZ protofilament formation. *Nature structural & molecular biology.* 2004;11(12):1243-1250.
24. Oliva MA, Trambaiolo D, Löwe J. Structural insights into the conformational variability of FtsZ. *Journal of molecular biology.* 2007;373(5):1229-1242.
25. Scheffers D-J, de Wit JG, den Blaauwen T, Driessen AJM. GTP Hydrolysis of Cell Division Protein FtsZ: Evidence that the Active Site Is Formed by the Association of Monomers †. *Biochemistry.* 2002;41(2):521-529.
26. Wang X, Huang J, Mukherjee A, Cao C, Lutkenhaus J. Analysis of the interaction of FtsZ with itself, GTP, and FtsA. *J. Bacteriol.* 1997;179(17):5551-5559.
27. Huecas S, Andreu JM. Polymerization of nucleotide-free, GDP- and GTP-bound cell division protein FtsZ: GDP makes the difference. *FEBS letters.* 2004;569(1-3):43-48.
28. Lu C, Reedy M, Erickson HP. Straight and curved conformations of FtsZ are regulated by GTP hydrolysis. *J Bacteriol.* 2000;182(1):164-170.
29. Leung AKW, Lucile White E, Ross LJ, Reynolds RC, DeVito JA, Borhani DW. Structure of Mycobacterium tuberculosis FtsZ reveals unexpected, G protein-like conformational switches. *Journal of molecular biology.* 2004;342(3):953-970.
30. Lu C, Stricker J, Erickson HP. Site-specific mutations of FtsZ--effects on GTPase and in vitro assembly. *BMC Microbiol.* 2001;1:7.
31. Diaz JF, Kralicek A, Mingorance J, Palacios JM, Vicente M, Andreu JM. Activation of cell division protein FtsZ. Control of switch loop T3 conformation by the nucleotide gamma-phosphate. *J Biol Chem.* 2001;276(20):17307-17315.
32. White EL, Ross LJ, Reynolds RC, Seitz LE, Moore GD, Borhani DW. Slow Polymerization of Mycobacterium tuberculosis FtsZ. *Journal of Bacteriology.* 2000;182(14):4028-4034.
33. Helliwell JR. *Macromolecular crystallography with synchrotron radiation*. Cambridge England ; New York, NY, USA: Cambridge University Press; 1992.
34. Long F, Vagin AA, Young P, Murshudov GN. BALBES: a molecular-replacement pipeline. *Acta crystallographica. Section D, Biological crystallography.* 2008;64(Pt 1):125-132.
35. RCSB Protein Data Bank - RCSB PDB.
36. Girard E, Legrand P, Roudenko O, Roussier L, Gourhant P, Gibelin J, Dalle D, Ounsy M, Thompson AW, Svensson O, Cordier MO, Robin S, Quiniou R, Steyer JP. Instrumentation for synchrotron-radiation macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr.* 2006;62(Pt 1):12-18.
37. Nave C. Radiation-Damage in Protein Crystallography. *Radiation Physics and Chemistry.* 1995;45(3):483-490.
38. Juers DH, Matthews BW. Cryo-cooling in macromolecular crystallography: advantages, disadvantages and optimization. *Quarterly Reviews of Biophysics.* 2004;37(2):105-119.

39. Taylor G. The phase problem. *Acta Crystallographica Section D Biological Crystallography.* 2003;59(11):1881-1890.
40. Evans P, McCoy A. An introduction to molecular replacement. *Acta Crystallogr D Biol Crystallogr.* 2008;64(Pt 1):1-10.
41. Rossmann MG. The molecular replacement method. *Acta Crystallographica Section A Foundations of Crystallography.* 1990;46(2):73-82.
42. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *Journal of applied crystallography.* 2007;40(Pt 4):658-674.
43. McRee DE. *Practical protein crystallography*. San Diego: Academic Press; 1993.
44. Brunger AT. Free R-Value - a Novel Statistical Quantity for Assessing the Accuracy of Crystal-Structures. *Nature.* 1992;355(6359):472-475.
45. McPherson A. Introduction to protein crystallization. *Methods.* 2004;34(3):254-265.
46. Mcpherson A. A Brief-History of Protein Crystal-Growth. *Journal of Crystal Growth.* 1991;110(1-2):1-10.
47. Chayen NE. Comparative studies of protein crystallization by vapour-diffusion and microbatch techniques. *Acta Crystallographica Section D-Biological Crystallography.* 1998;54:8-15.
48. Luft JR, DeTitta GT. HANGMAN: a macromolecular hanging-drop vapor-diffusion technique. *Journal of Applied Crystallography.* 1992;25(2):324-325.
49. Wlodawer A, Hodgson KO. Crystallization and crystal data of monellin. *Proceedings of the National Academy of Sciences of the United States of America.* 1975;72(1):398-399.
50. Leung AKW, White EL, Ross LJ, Borhani DW. Crystallization of the Mycobacterium tuberculosis cell-division protein FtsZ. *Acta Crystallographica Section D Biological Crystallography.* 2000;56(12):1634-1637.
51. Cudney R, Patel S, Weisgraber K, Newhouse Y, McPherson A. Screening and optimization strategies for macromolecular crystal growth. *Acta crystallographica. Section D, Biological crystallography.* 1994;50(Pt 4):414-423.
52. Stura EA, Wilson IA. Applications of the streak seeding technique in protein crystallization. *Journal of Crystal Growth.* 1991;110(1):270-282.
53. Garman EF, Owen RL. Cryocooling and radiation damage in macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr.* 2006;62(Pt 1):32-47.
54. X6A.
55. Otwinowski Z, Minor W. [20] Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology.* 1997;276:307-326.
56. Kantardjieff KA, Rupp B. Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein science : a publication of the Protein Society.* 2003;12(9):1865-1871.
57. Matthews BW. Solvent Content of Protein Crystals. *Journal of Molecular Biology.* 1968;33(2):491-&.
58. Weiss MS, Hilgenfeld R. On the use of the merging R factor as a quality indicator for X-ray data. *Journal of Applied Crystallography.* 1997;30:203-205.
59. Engh RA, Huber R. Accurate Bond and Angle Parameters for X-Ray Protein-Structure Refinement. *Acta Crystallographica Section A.* 1991;47:392-400.
60. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. MolProbity: all-atom structure validation for

macromolecular crystallography. *Acta Crystallographica Section D-Biological Crystallography.* 2010;66:12-21.

61. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS. Overview of the CCP4 suite and current developments. *Acta crystallographica. Section D, Biological crystallography.* 2011;67(Pt 4):235-242.

62. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography.* 2010;66(Pt 2):213-221.

63. Vagin A, Teplyakov A. MOLREP : an Automated Program for Molecular Replacement. *Journal of Applied Crystallography.* 1997;30(6):1022-1025.

64. Lengauer T, Rarey M. Computational methods for biomolecular docking. *Curr Opin Struct Biol.* 1996;6(3):402-406.

65. Wei BQ, Weaver LH, Ferrari AM, Matthews BW, Shoichet BK. Testing a flexible-receptor docking algorithm in a model binding site. *J Mol Biol.* 2004;337(5):1161-1182.

66. Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov.* 2004;3(11):935-949.

67. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem.* 2009;30(16):2785-2791.

68. Mitchell M. *An introduction to genetic algorithms*. Cambridge, Mass.: MIT Press; 1996.

69. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *Journal of Computational Chemistry.* 1998;19(14):1639-1662.

70. Scheffzek K, Ahmadian MR, Kabsch W, Wiesmuller L, Lautwein A, Schmitz F, Wittinghofer A. The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants. *Science.* 1997;277(5324):333-338.

71. Li Y, Hsin J, Zhao L, Cheng Y, Shang W, Huang KC, Wang HW, Ye S. FtsZ protofilaments use a hinge-opening mechanism for constrictive force generation. *Science.* 2013;341(6144):392-395.

72. Scheffers DJ, de Wit JG, den Blaauwen T, Driessen AJ. Substitution of a conserved aspartate allows cation-induced polymerization of FtsZ. *FEBS Lett.* 2001;494(1-2):34-37.

73. Mosyak L, Zhang Y, Glasfeld E, Haney S, Stahl M, Seehra J, Somers WS. The bacterial cell-division protein ZipA and its interaction with an FtsZ fragment revealed by X-ray crystallography. *the The European Molecular Biology Organization Journal.* 2000;19(13):3179-3191.

74. Cordell SC, Robinson EJ, Lowe J. Crystal structure of the SOS cell division inhibitor SulA and in complex with FtsZ. *Proc Natl Acad Sci U S A.* 2003;100(13):7889-7894.

75. Stricker J, Maddox P, Salmon ED, Erickson HP. Rapid assembly dynamics of the Escherichia coli FtsZ-ring demonstrated by fluorescence recovery after photobleaching. *Proc Natl Acad Sci U S A.* 2002;99(5):3171-3175.

76. Matsui T, Yamane J, Mogi N, Yamaguchi H, Takemoto H, Yao M, Tanaka I. Structural reorganization of the bacterial cell-division protein FtsZ from Staphylococcus aureus. *Acta Crystallographica Section D-Biological Crystallography.* 2012;68:1175-1188.
77. Lowe J, Li H, Downing KH, Nogales E. Refined structure of alpha beta-tubulin at 3.5 A resolution. *Journal of Molecular Biology.* 2001;313(5):1045-1057.