5-2011

Molecular Mechanism By Which The Escherichia Coli Nucleoid Occlusion Factor, Slma, Keeps Cytokinesis In Check

Nam K. Tonthat

MOLECULAR MECHANISM BY WHICH THE *ESCHERICHIA COLI* NUCLEOID

OCCLUSION FACTOR, SLMA, KEEPS CYTOKINESIS IN CHECK

A

Dissertation

Presented to the Faculty of

The University of Texas

Health Science Center at Houston

and

The University of Texas

M. D. Anderson Cancer Center

Graduate School of Biomedical Sciences

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

by

Nam Ky Tonthat, BS, MS

Houston, Texas

May, 2011

Acknowledgements

I am fortunate enough to be an environment that fosters both academic and personal growth. Not only has my scientific acumen improved dramatically, but I have also developed emotionally and mentally. There are so many people who have contributed to my personal development. In particular, I would like to specifically thank Dr. Maria Schumacher for all the support and guidance I have received. Maria gave me the opportunity to naively navigate through my project and never contested my wild ideas. Her patience with me through my repeated "Nam-ism" is inspirational.

The 7th floor of the BSRB has been a great home for me, and all those on this floor have been a great source of companionship. I can always count on any member of the Schumacher, Brennan, Wilkinson, and Ladbury group for ideas, reagents, technical support, and even emotional encouragement. I would also like to thank all my classmates at the GSBS. In particular, I would like to express my appreciation for Brian Pickering. Brian and I meet serendipitously and quickly recognized each other's affinity to science and good conversation. He consequently became a very reliable collaborator and more importantly, a great friend.

Although my time in lab was critical to my ability to complete my doctoral degree, it could not have happened without continual support from my family and friends at home. I had two wonderful roommates, Robert Park and Carol Chuang, who supported my board-game addiction. Sadly, as life progress, people must move, and I genuinely thought that it would not be possible to find another roommate who can meet the same standard. Luckily, my good friend and now perpetual roommate, Kevin Chow is a "Mystical Warrior" who has proven himself to be exceedingly wonderful. I am also appreciative of Susanna Lai. Susanna has continuously provided me with patience and love through times when I did not deserve either.

Lastly, my mom, dad, sister, brother, and nephew were always supportive and ready to aid me in any endeavors. On a very personal note, I would like to acknowledge and dedicate this dissertation work to my dad. My dad was ferociously inquisitive and believed strongly in the need to obtain a thorough and thoughtful understand of all things life. My choice of profession reflects the dramatic impression he has left on me, and I am grateful for it. In conclusion, I would like to state that everyone has been exceedingly nice to me, more so than I deserve. To those who have shared their lives with me and I can only hope that I had positively impact on them as well.

Abstract

Cell division or cytokinesis is one of the most fundamental processes in biology and is essential for the propagation of all living species. In *Escherichia coli*, cell division occurs by ingrowth of the membrane envelope at the cell center and is orchestrated by the FtsZ protein. FtsZ self-assembles into linear protofilaments in a GTP dependent manner to form a cytoskeletal scaffold called the Z-ring. The Z-ring provides the framework for the assembly of the division apparatus and determines the site of cytokinesis. The total amount of FtsZ molecules in a cell significantly exceeds the concentration required for Z-ring formation. Hence, Z-ring formation must be highly regulated, both temporally and spatially. In particular, the assembly of Z-rings at the cell poles and over chromosomal DNA must be prevented. These inhibitory roles are played by two key regulatory systems called the Min and nucleoid occlusion (NO) systems.

In *E. coli*, Min proteins oscillate from pole to pole; the net result of this oscillatory process is the formation of a zone of FtsZ inhibition at the cell poles. However, the replicated nucleoid DNA near the midcell must also be protected from bisection by the Z-ring which is ensured by NO. A protein called SlmA was shown to be the effector of NO in *E. coli*. SlmA was identified in a screen designed to isolate mutations that were lethal in the absence of Min, hence the name SlmA (synthetic lethal with a defective Min system). Furthers SlmA was shown to bind DNA and localize to the nucleoid fraction of the cell. Additionally, light scattering experiments suggested that SlmA interacts with FtsZ-GTP and alters its polymerization properties. Here we describe studies that reveal the molecular mechanism by which SlmA mediates NO in *E. coli*. Specifically, we determined the crystal structure of SlmA, identified its DNA binding site specificity, and mapped its binding sites on the *E. coli* chromosome by chromatin immuno-precipitation experiments. We went on to determine the SlmA-FtsZ structure by small angle X-ray scattering and examined the effect of SlmA-DNA on FtsZ polymerization by electron microscopy. Our combined data show how SlmA is able to disrupt Z-ring formation through its interaction with FtsZ in a specific temporal and spatial manner and hence prevent nucleoid guillotining during cell division.

Table of Contents

List of Illustrations

List of Tables

Chapter 1. Introduction and background to bacterial cell division

1.1 Prokaryotic and eukaryotic cell division

Cell division is the fundamental process of life that allows single and multi-celled organisms to thrive and proliferate. The general cell division cycle requires the replication and segregation of genetic material followed by cytokinesis, which ends in the formation of two daughter cells. Prokaryotic cells are comparatively much simpler than their eukaryotic counterparts, and undergo cell division via a process called binary fission. In this process, the prokaryotic chromosome, a single circular DNA molecule, is replicated. As the chromosomes are segregated, a ring like structure forms at the mid-cell to initiate cytokinesis. Eukaryotic cells, on the other hand, undergo mitosis and cytokinesis to form two genetically equivalent daughter cells. After replication of the chromosomes, chromosomes, with two sister chromatids, are positioned at the mid-cell. A system of kinetochore microtubules "pulls" and segregates the chromosomes towards the poles of the cell. Then similarly to prokaryotic cells, a ring like structure forms at the mid-cell to carry out cytokinesis.

Although these two types of cells are vastly different in composition, the processes of cell division are analogous. In particular, both cell types share the primary concern of appropriately positioning the division plane, in order to ensure proper partition of cellular components and to maintain the integrity of its genome. Thus, the critical and irreversible step of cytokinesis requires tight spatial and temporal control. Studies performed in prokaryotic and eukaryotic systems have revealed several mechanisms that coordinate the assembly/disassembly of division machineries at the appropriate site. In yeast, as well as other animal cells, filamentous actin (F-actin), type II myosin, and several other proteins assemble into a ring like structure at the division plane (Figure 1A). The contraction of this ring consequently produces the necessary force for the cleavage of the cell. This contractile motion is well coordinated with chromosome segregation, membrane trafficking, and the generation of new membrane. Similarly, in bacteria, the FtsZ protein (a tubulin homolog) aggregates to form a ring-like structure, which then recruits other cell division proteins to form a mature Z-ring. The Z-ring then initiates cytokinesis and allows for the generation of two daughter cells (Figure 1B).

The fission and budding yeasts, *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae*, have provided great insight into the eukaryotic cell division process. Briefly, the cylindrical shaped *S. pombe* undergoes division that produces two daughter cells of approximately equal proportion. The placement of the division plane is determined by the

A

Eukaryotes

Actomyosin ring

Actomyosin ring

Fission yeast
Schizosaccharomyces pombe

Budding yeast
Saccharomyces cerevisiae

B

Prokaryotes

FtsZ ring

FtsZ ring

Gram negative
Escherichia coli

Gram positive
Staphylococcus aureus
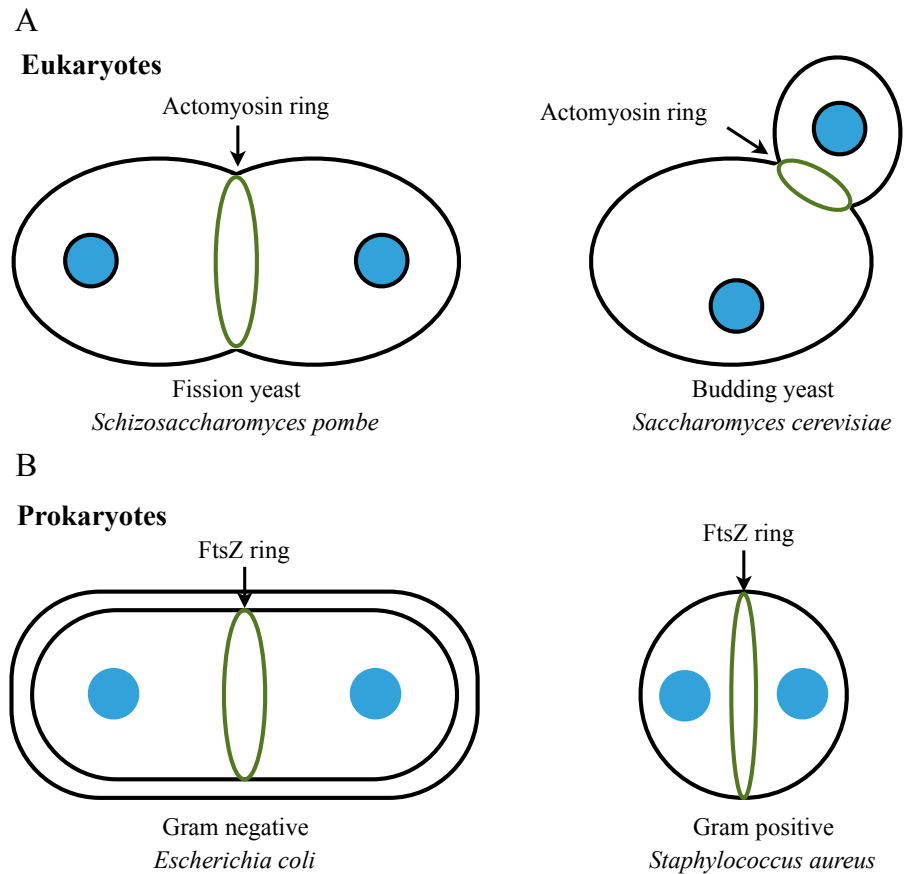
Figure 1. Contractile ring assemble at the division site. A) In eukaryotic cells such as fission and budding yeasts, a contractile ring composed of mainly F-actin and myosin provides the force necessary for cytokinesis. **B)** In prokaryotic bacteria, a ring formed primarily by FtsZ assembles at the division plane.

premitotic nucleus (1). Alteration in the position of the nucleus will also reposition the actomyosin ring and thus the cell division site (2, 3). The protein Pom1p localizes to the poles of the cylindrical cells and prevents the accumulation of Mid1p. Mid1p is therefore pooled toward the mid-cell and stimulates the assembly of the actomyosin ring at the cortex overlying the nucleus (4). Next, the mature actomyosin ring provides the contractile force to initiate cleavage.

The budding yeast, *S. cerevisiae*, in contrast, divides asymmetrically through budding. The division site is determined early in the cell cycle, at the G1/S stage. Cortical landmark proteins such as Bud3p, Bud4p, Bud10p localize to the site of bud assembly. These proteins then recruit Ras-related proteins, Ras1p, which leads to the recruitment of Cdc42p, GEF, and Cdc24p. These factors then recruit Gic1p and Gic2p, which allows for the assembly of the septin ring. The septin ring is important for the establishment of the division site because of its ability to promote actomyosin ring formation, which will carry out cytokinesis (5).

In prokaryotes, investigations of the *Escherichia coli* and *Bacillus subtilis* model systems have yielded a wealth of information regarding bacterial cell division. In bacteria, the structural equivalent to the eukaryotic actomyosin ring is the Z-ring. The Z-ring is formed predominantly by a self-polymerizing protein called FtsZ. FtsZ is one of the first molecules to localize to the future division site, and self-associates to form a ring-like structure, aptly named the Z-ring. This structure acts as a scaffold to recruit other cell division proteins such as FtsA, FtsK, and ZipA. The mature form of this ring is called the divisome. Given its importance in determining the plane of division, the spatial regulation of FtsZ is therefore critically regulated by multiple cellular processes. These processes modulate FtsZ polymerization and positioning. Proteins such as FtsA, ZipA, and SulA are regulators of Z-ring formation, while the Min and Nucleoid Occlusion system act to position the Z-ring in the appropriate location. These partially overlapping mechanisms of FtsZ regulation function in concert to ensure proper division (6-9).

1.2 Bacterial cell division

Although chromosome segregation and cell division are generally thought of as two separate processes, there is a considerable amount of overlap. The coordination of cell division with chromosome segregation ensures that both the correct proportion and the integrity of the genetic materials are passed on to each daughter cell. Thus, before delving into the functions of the cell division machineries, a primer on chromosome segregation is needed to provide context and background.

The bacterial chromosome takes the form of a single circular molecule of DNA. However, structurally, the chromosome can be partitioned into six domains: four macro-domains and two non-structured regions. The domains are the Ori, Ter, Left and Right macro-domains, and the Left and Right non-structured regions (Figure 2). These macro-domains are defined as regions that are spatially homogeneous and do not "collide" with each other; these conclusions were primarily based on fluorescent in situ hybridization experiments (10, 11). The Ori and Ter macro-domains were discovered first and were named based on the presence of the origin of replication site, OriC, and the termination site respectively. The Left and Right macro-domains were discovered later as regions that flank the Ter macro-domain. Similarly, the Left and Right non-structured regions flank the Ori macro-domain.

The factors responsible for the organization of the macro-domains are yet to be identified, with the exception of one. Mercier and colleagues demonstrated, in *E. coli*, the organization of the Ter macro-domain is carried-out by a single DNA binding protein called MatP. MatP is able to bind to a DNA motif termed *matS*, which is only located in the Ter region. The absence of MatP causes the Ter region to become less compacted and the mobility of the fluorescent DNA marker increases dramatically. Additionally, the depletion of MatP lead to the disruption of chromosome segregation, resulting a significant number of cells that display a filamentous and anucleated phenotype (12).

The organization of the chromosome into discrete domains is therefore necessary for chromosome segregation. As the chromosomes begin to segregate towards their respective poles, the Ori macro-domain and the two non-structured regions segregate concomitantly. Next, the Right and Left macro-domains segregates, leaving the Ter macro-domains at the mid-cell. Strikingly, the cell division machinery will assemble at the mid cell in the presence of the Ter macro-domains, and segregation does not occur until the moment before division (11, 13).

Bacterial cell division requires the formation of a large protein complex called the divisome at the division plane. This complex initiates the cytokinesis process and recruits a multitude of enzymes to synthesize the septum cell wall as it coordinates the invagination and inward growth of the membrane. Therefore the formation and placement of the divisome complex is critical, as it initiates the non-reversible process of cytokinesis. At the core of this complex is the Z-ring, formed by the self-polymerizing protein FtsZ. The FtsZ protein is thought to be the first protein to localize to the future division plane (14). The Z-ring extends around the circumference of the cell with the guidance of integral membrane proteins, such as ZipA, (15) and with membrane associated proteins such as FtsA (16). These proteins, along
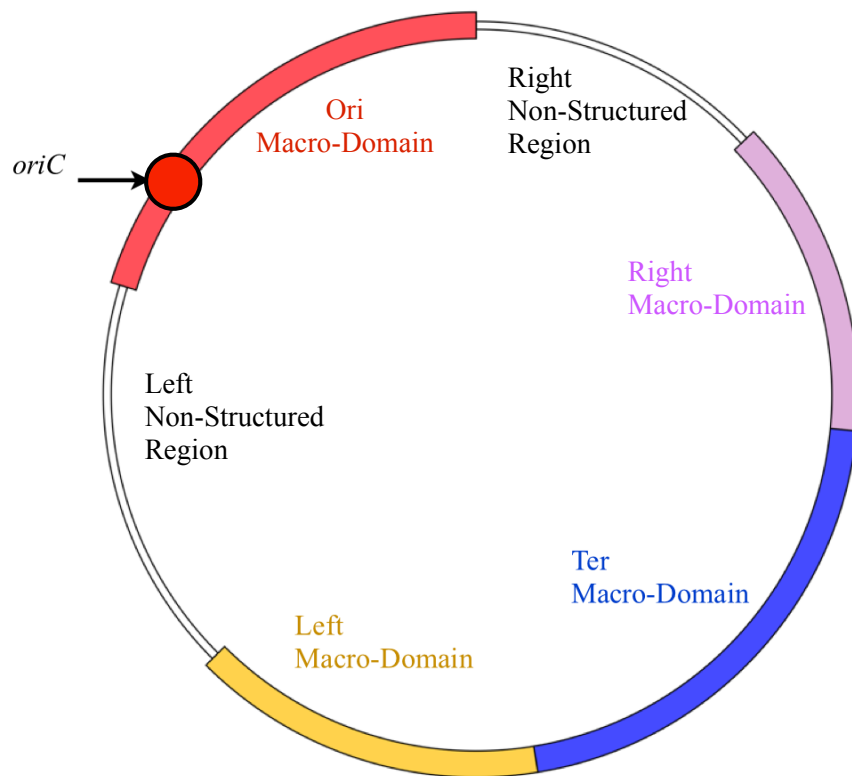
Figure 2. Representation of a bacterial chromosome. The chromosome can be partition in to six domains, comprising of four macro-domains and two non-structured region. The domains are: the Ori, Ter, Left and Right macro-domains, and the Left and Right non-structured region.

with other protein complexes and cofactors (17), anchor the Z-ring to the membrane, which is critical for the process of cytokinesis. The cytokinesis machinery also includes at least seven other proteins: FtsK, FtsQ, FtsB, FtsL, FtsW, FtsI, and FtsN (Figure 3A) (7). These proteins seem to be recruited in a linear fashion (Figure 3B) (7) with ZipA and FtsA being the initial complex that is required to recruit all other factors (18). After the recruitment of the necessary factors, this highly dynamic complex based from FtsZ (termed the divisome), will initiate the constriction and cleavage process.

The process of how FtsZ-ring contraction leads to cytokinesis is not fully understood but one inviting theory proposes that two sets of events are involved. In the first set of events, the Z-ring recruits proteins such as FtsI and FtsW, for the synthesis of the peptidoglycan cell wall, which stimulates the invagination and inward growth of the septum. The second set of events involves an actual mechanical force exerted by the Z-ring. Due to the rapid loss of FtsZ monomers, the FtsZ-ring contracts to produce the constricting force needed for cytokinesis (7), analogous to the force exerted by Dynamins on their membrane substrate (19). Finally, since the peptidoglycan layer is connected to the outer membrane via bridging lipoproteins, the outer membrane can follow the inward growth of the peptidoglycan (9).

1.3 Structure and function of the cell division protein FtsZ

The discovery of FtsZ has stimulated dramatic growth in the field of bacterial cell division. FtsZ is the most highly conserved protein in bacteria and plays a central role in the progression and regulation of division (20, 21). Schematically, FtsZ is composed of four segments: a variable N-terminal segment, a conserved core, a variable spacer, and a C-terminal tail (Figure 4). The conserved core region contains a GTPase and the C-terminal tail facilitates interactions with FtsA and ZipA (22). The function of the N-terminal and spacer segments, however, have not been elucidated.

The first structure of FtsZ was solved in 1998 by Löwe and Amos, from *Methanococcus jannaschii* (23). Since then, many other structures have been solved including FtsZ from: *Bacillus subtilis*, *Aquifex aeolicus*, and *Pseudomonas aeruginosa*. Structural comparison has shown that, as anticipated, these structures are highly congruent (21). The FtsZ core region has a two-domain architecture: an N-terminal nucleotide binding domain and a C-terminal domain that plays a role in forming the protofilament. These two domains are linked by a central helix (H7) (Figure 5A).
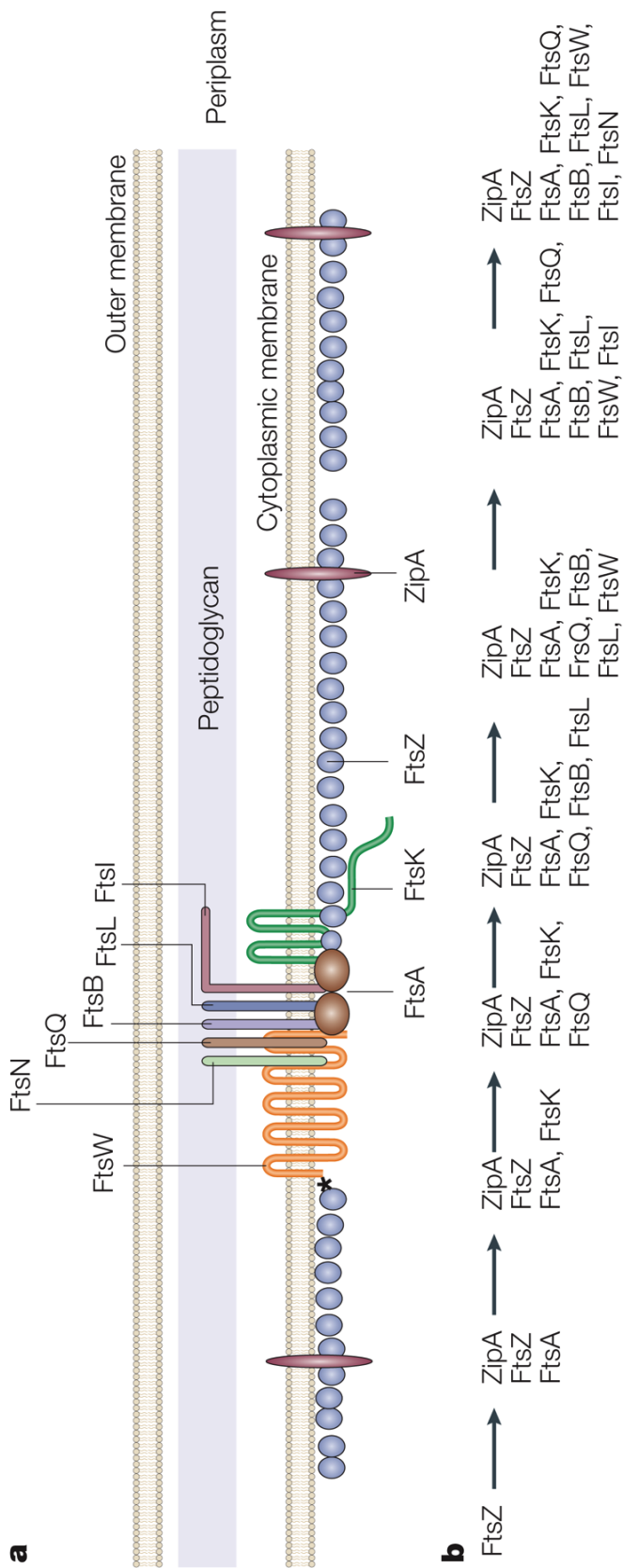
Figure 3. The cytokinesis machineries. A) A model of the *Escherichia coli* Z ring and its essential protein partners is shown in cross section. FtsZ is shown as a series of single protofilaments at the membrane, although the actual structure of FtsZ in the Z ring is unknown. Both ZipA and FtsA contact FtsZ as well as the membrane in *E. coli*. However, FtsZ contacts FtsW directly in *Mycobacterium tuberculosis* (asterisk), which lacks ZipA and FtsA. A single transmembrane subassembly associated with an FtsA dimer is shown, based on the low relative amounts of most of the integral membrane proteins that are essential for cell division. These membrane proteins include FtsQ, FtsB, FtsL, FtsI and FtsN, which are bitopic proteins that each have a single transmembrane and periplasmic domain, and FtsW and FtsK, which are polytopic proteins with multiple transmembrane and periplasmic domains. The network of protein–protein associations is implied by the proximity of the proteins in the diagram. Proteins implicated in stabilization of the ring structure are labelled below the cytoplasmic membrane lines, whereas proteins implicated in later functions in septum formation, such as septum synthesis, are labelled above the lines. **B)** The dependency order of recruitment of essential cell-division proteins to the Z ring, as deduced from the requirement of a given protein for another's localization to the Z ring.
Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Molecular Cell Biology (Margolin, 2005), copyright (2005).
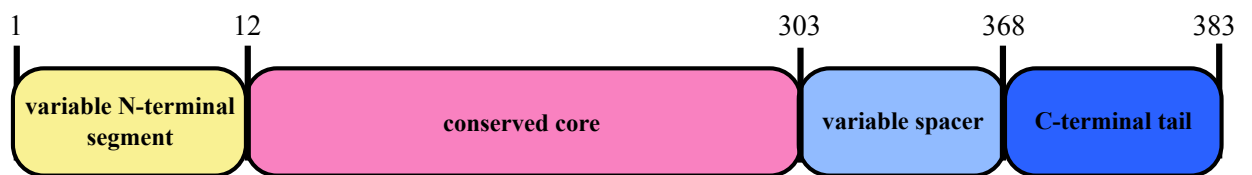
Figure 4. Schematic of FtsZ. FtsZ is composed of four segments: a variable N-terminal segment, a conserved core, a variable spacer, and a C-terminal tail. The number shown corresponds to the amino acid position from the *E. coli* FtsZ protein.
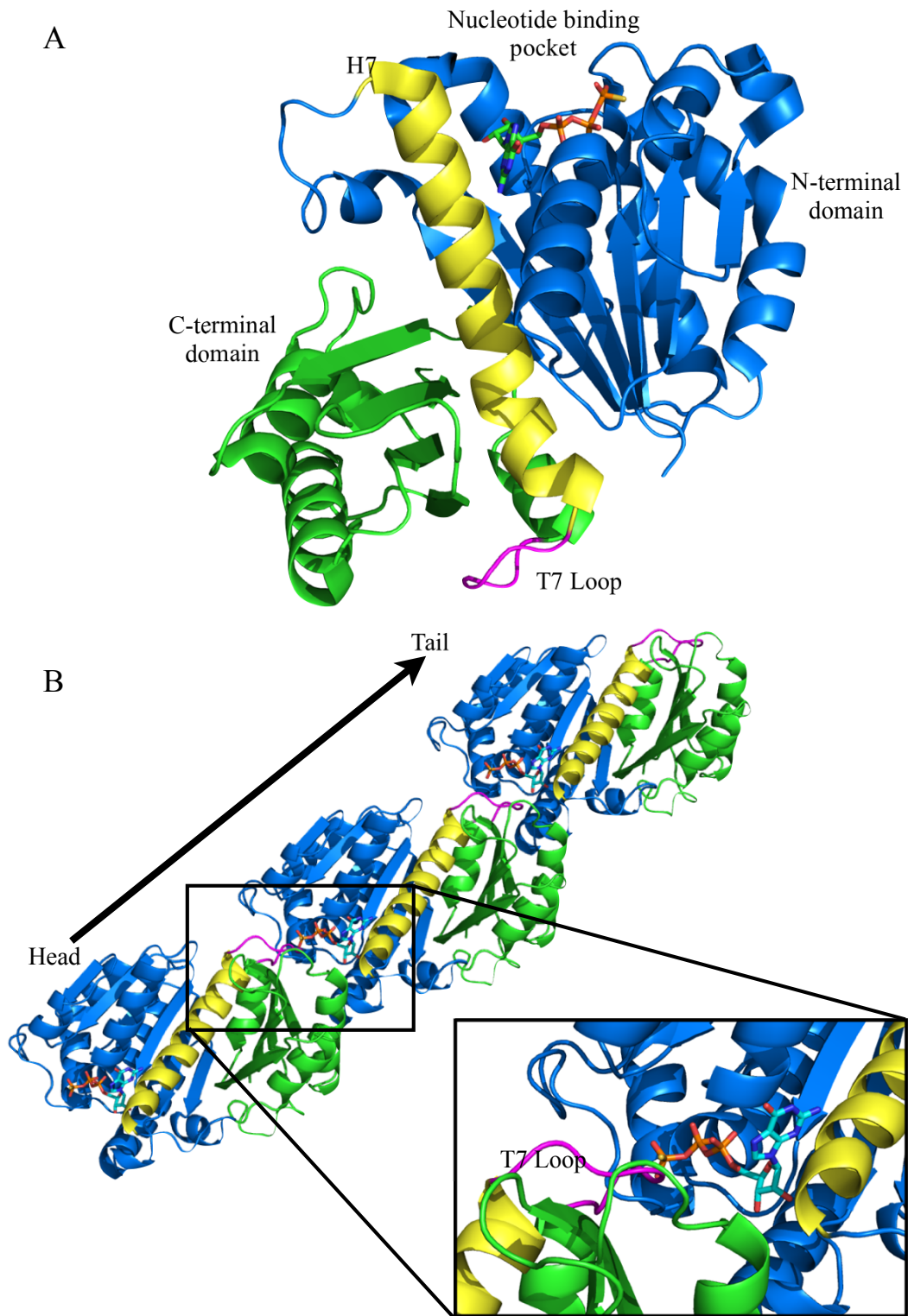
Figure 5. The structure of FtsZ and FtsZ protofilament. A) FtsZ has a two-domain architechure: an N-terminal nucleotide binding domain and a C-terminal domain that plays a role in forming the protofilament. These two domains are linked by a central helix (H7). **B)** FtsZ is able oligomerize into protofilaments. The GTPase active site is formed at the interface between monomers of FtsZ by insertion of the C-terminal domain's T7 loop into the nucleotide binding pocket of the preceeding monomer in the protofilament.

FtsZ is able to oligomerize into protofilaments, which are short strands of linearly interacting FtsZ monomers. Interestingly, the GTPase active site is formed at the interface between monomers of FtsZ by the insertion of the T7 loop of the preceding monomer's into the nucleotide binding pocket. Like tubulin, the FtsZ protofilament is considered to have a head to tail orientation (Figure 5B). However, there are some very distinct differences, the most significant of which concerns the nucleotide-binding pocket. In tubulin protofilaments, the nucleotide-binding pocket is occluded and dissociation is required for nucleotide exchange to occur. Thus tubulin filaments are formed with a distinct GTP cap followed by a GDP bound tail. However, in FtsZ, the nucleotide-binding pocket is solvent accessible, as shown by the crystal structure of FtsZ, and nucleotide exchange can happen freely. Unlike tubulin, the GDP-bound form of FtsZ is still able to form polymers, but it has been noted that these polymers exhibit a more curved conformation, comparative to the GTP-stabilized polymers (24). Given that the pool of available GTP *in vivo* is sufficient to saturate FtsZ with GTP, it is still unknown what significance the curved filament form may play. Nonetheless, it is clearly important that FtsZ is able to exchange nucleotides readily since GTP binding stimulates FtsZ to self-assemble into protofilaments (25). How the protofilaments arrangment *in vivo* is still a very open question. Nonetheless, additional factors such as Ca^{2+} (26) and macromolecular crowding (27) can cause these protofilaments to associate laterally to form bundles *in vitro* (28-31). A more thorough treatment regarding FtsZ polymerization and formation of the Z-ring can be found in Section 1.4.

Studies on Z-ring formation has demonstrated that the Z-ring is extremely dynamic. Several studies have revealed that the Z-ring can assemble and dissemble within less then 1 minute (32, 33). Furthermore, the turnover of FtsZ molecules in the ring is also extremely dynamic. The half-time for remodeling has been reported to be as low as 9 seconds (34). Thus the Z-ring scaffold is highly dynamic yet appears to be quite stable.

1.4 FtsZ polymerization and the formation of the Z-ring

FtsZ molecules are able to cooperatively interact with one another to form multiple polymeric states, ranging from short, single-stranded protofilaments to multi-filament bundles that can circumscribe the perimeter of a cell (35-37). Although the true nature of how FtsZ molecules organize themselves *in vivo* to form an active Z-ring is unclear, there is a wealth of *in vitro* data available (7, 8, 36). Additionally, recent *in vivo* investigations via cryo-electron

tomography and fluorescence photoactivation localization microscopy (FPALM) have contributed to the elucidation of this process (38, 39).

Currently, it is well accepted that the Z-ring is formed by short FtsZ protofilaments which are ~30 subunits long (34, 40, 41). *In vitro*, the formation of single FtsZ filaments can be observed at a concentration of approximately 0.5 to 1 µM (Figure 6A) (35, 42, 43). As FtsZ concentrations are increased to ~3 µM, the equilibrium shifts towards the formation of larger, multi-filament structures (discussed below) that take shape as rings (Figure 6B-D) (36, 43). Quantification of FtsZ concentration in most *E. coli* strains has shown that there are 5,000 to 7,000 FtsZ molecules per cell, which equates to a concentration of ~4 µM (44, 45). This concentration is well above the 1 µM concentration needed for protofilament formation, and suggests that, *in vivo*, FtsZ primarily exists as protofilaments, which are the fundamental units utilized for Z-ring assembly.

Two inviting models of how protofilaments can further assemble to form the Z-ring, have been suggested from the current literature. One possibility is that the protofilaments can anneal to one another to form a much longer filament that can be tethered to the inner membrane (Figure 7A). Support for this model includes atomic force microscopy studies which directly demonstrate that protofilaments can anneal to form long filaments, when absorbed on a 2D mica surface (Figure 8A) (46). Additionally, there is indirect evidence that protofilament annealing occurs in solution (47). However, this model does not fully address the rapid turnover rate of FtsZ molecules in a filament. Fluorescence recovery after photo-bleaching studies have demonstrated that FtsZ molecules in a protofilament have a turnover rate of approximately 9 seconds (34). This rapid exchange can lead to multiple breakages of the long filament and requires reannealing/assembly. A second model proposes that multiple protofilaments bundle in a staggered manner to form the Z-ring (Figure 7B). Cryo-tomography studies show that during the division process, the Z-ring is composed of short filaments. Interestingly, these short filaments are sparsely scattered around the circumference of the cell (Figure 8B) (38).

Consideration of these models raises a major question regarding the lateral interaction of the protofilaments. Specifically, are the lateral interactions between FtsZ molecules direct or indirect? A direct lateral interaction model requires that FtsZ molecules make specific contacts with each other, while indirect interactions are mediated by ions or even other macromolecules. In support of the lateral interaction model, fluorescence light microscopy data have shown the
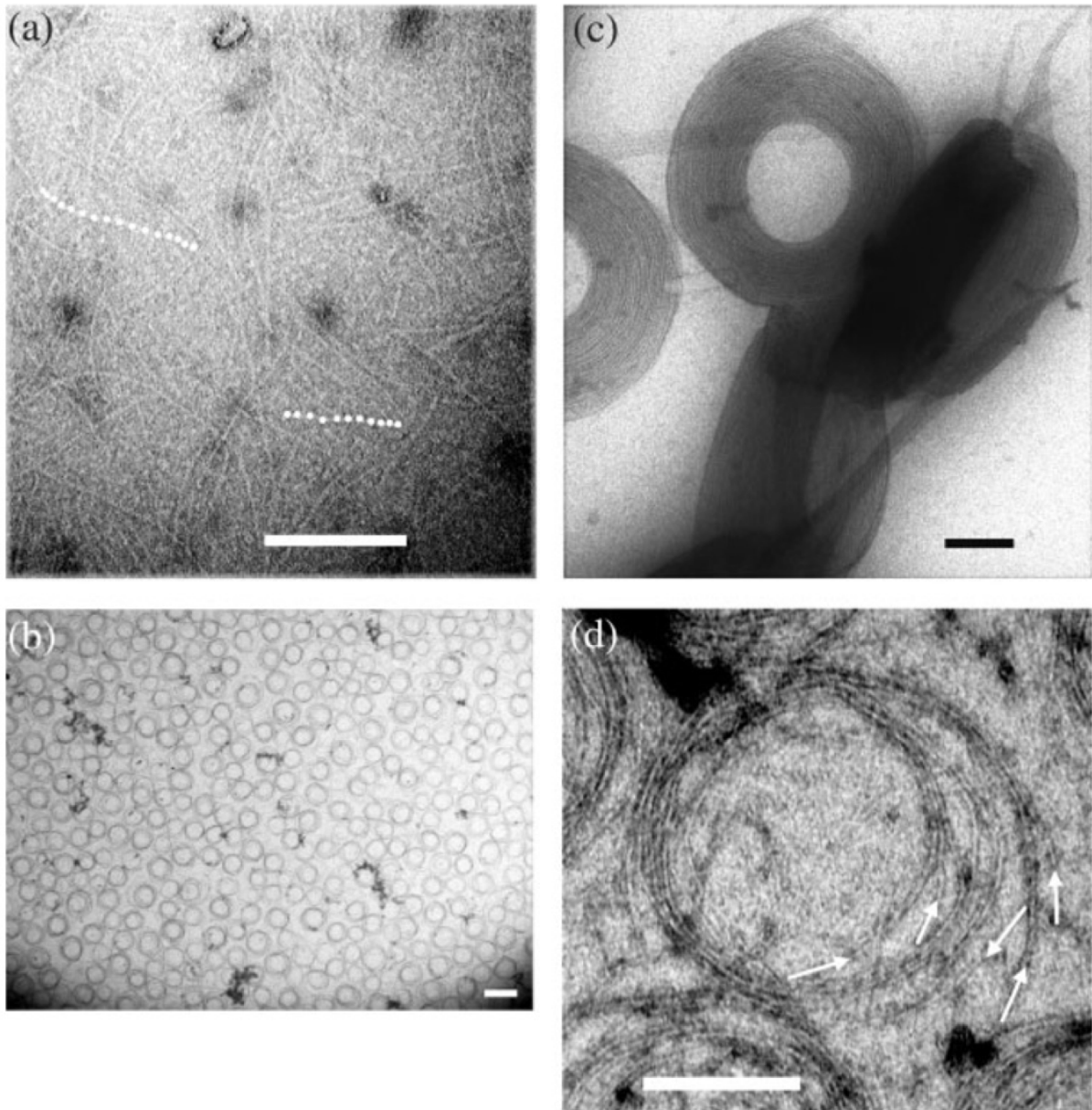
Figure 6. Representative images of FtsZ filaments. Results were pH independent (shown here pH 7.7, which is to the internal pH of *E. coli* in vivo). The appearance of the structures was similar with different crowding agents used, only the concentrations to induce the condensation phenomena differed between crowding agents. **A)** FtsZ filaments below a critical concentration of crowding agent were mainly single filaments. Shown are FtsZ-GTP filaments in the presence of 0.4% MC highlighted as dotted lines, scale bar 100 nm. **B)** Above the critical concentration, the equilibrium was shifted to rings consisting of several individual FtsZ filaments with an average diameter of about 220 nm. Shown are FtsZ-GTP filaments in the presence of 1.6% MC, scale bar 500 nm. **C)** Higher crowding agent concentrations (shown here FtsZ-GMPPNP in the presence of 8% PVA) condensed the structures into well defined toroids, scale bar 100 nm. **D)** A closer look at the architecture of rings, which just started to condense above the critical concentration (shown at 1% MC). Individual filaments which form lateral contacts to neighboring filaments can be seen and the ends of individual filaments are marked with an arrow. Most filaments appeared to be between 400 and 800 nm long. Most rings observed consisted of single filaments, scale bar 100 nm.
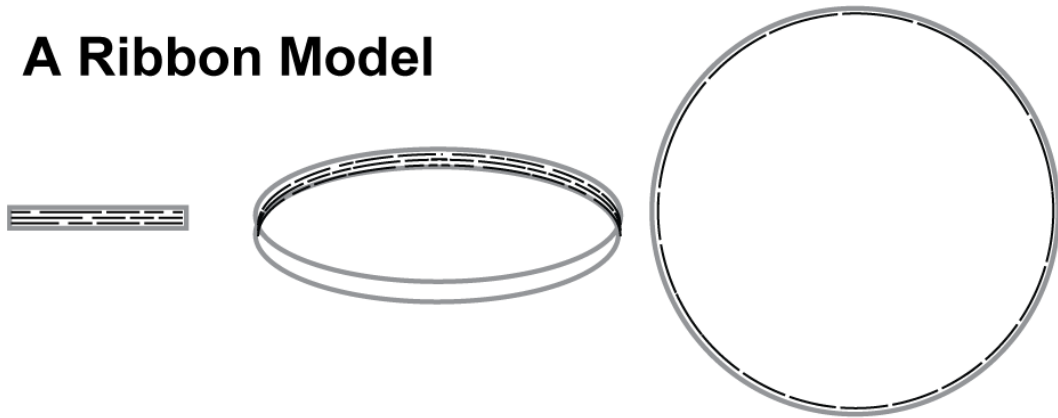
A Ribbon Model

B Bundle Model
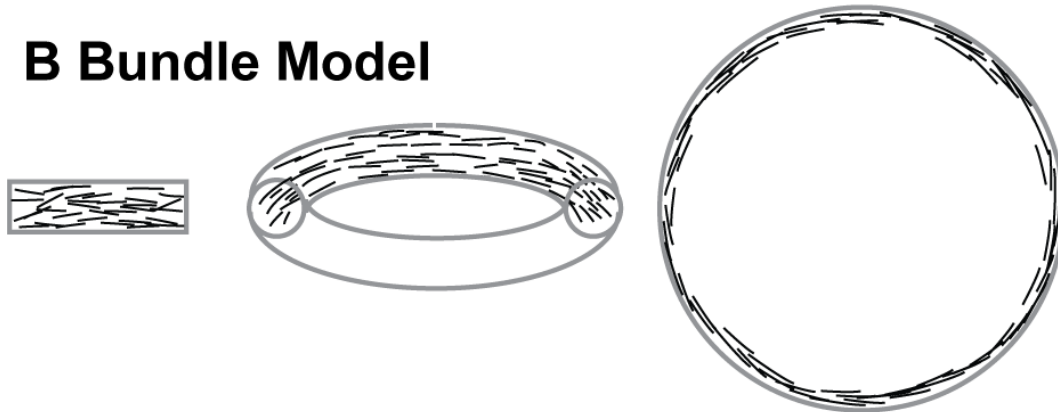
Figure 7. Two models of how FtsZ protofilaments can further assemble to the Z-ring. A) The Ribbon model proposes that FtsZ protofilaments can anneal to one another to form a much longer filament that can be tethered to the inner membrane. **B)** The Bundle model proposes that multiple FtsZ protofilaments bundle in a staggered manner to form the Z-ring.

Reprinted from: PLoS One (Fu et al, 2010), under the Creative Commons Attribution License.

A

B1 B2 B3
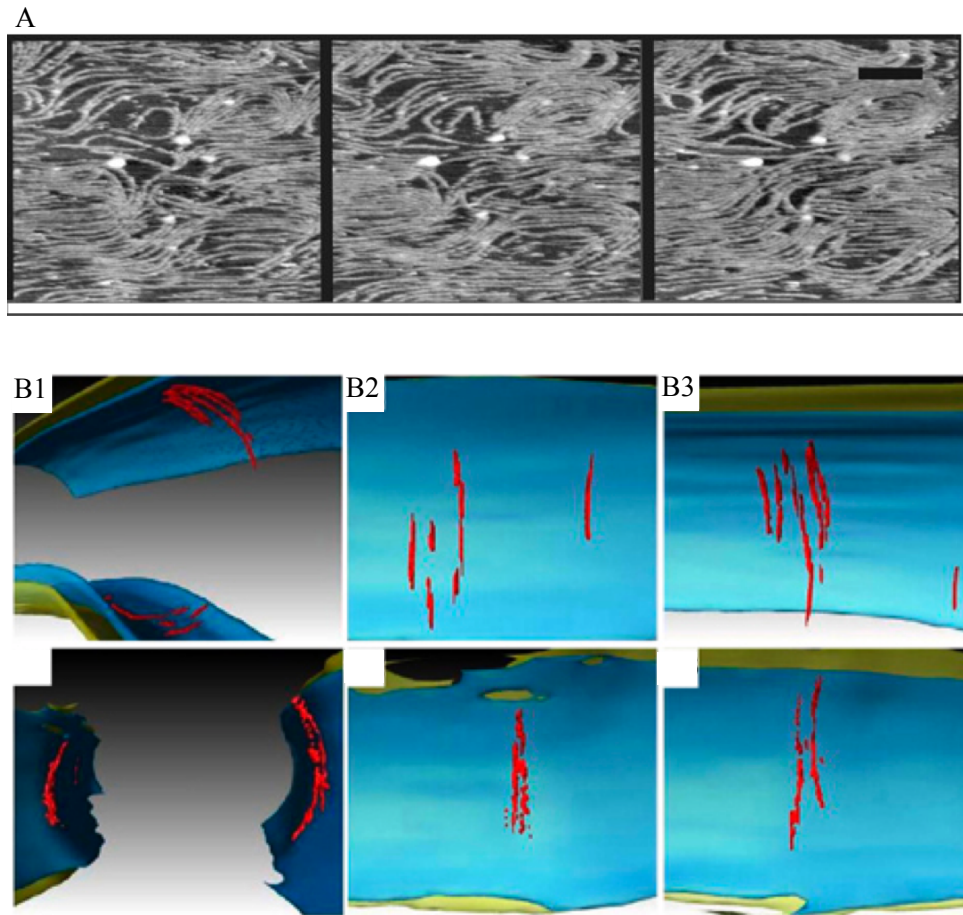
Figure 8. Visualization of the FtsZ filaments. A) Visualization of FtsZ filaments with atomic force microscopy. FtsZ filaments adsorbed on mica were observed while immersed in buffer containing 1 mm GTP over a period of 40 min. Images were taken every 2 min (the time needed to take an image). Scale bar, 200 nm. **B)** 3D reconstruction of the FtsZ filaments with cryo-tomography. 3-D segmentations of the division sites (B1). 'Face-on' views from the cytoplasm of the 'left' side of the cell wall (B2). 'Face-on' views of the 'right' side of the cell wall, again from the cytoplasm (B3).

Z-ring to be a thick, closed-ring structure compsed of sheets of protofilaments (48-50). These "thick filaments" were also seen in the crystal structure of FtsZ from *Methanococcus jannaschii* (23, 31, 49, 51). However, recent data have countered these observations with *in vivo* experiments that favor a model of indirect lateral interaction. As mentioned previously, cryo-tomography data show that FtsZ protofilaments are scattered sparsely with a spacing of 9.3 nm (38). This loose bundling of protofilaments is also observed via fluorescence photoactivation localization microscopy (39). Given that the FtsZ supra-structure *in vitro* can be affected by polymerization conditions, crowding agents, and the multitude of FtsZ binding proteins in the cell, the more recent *in vivo* cryo-tomography and FPALM studies may reflect the true nature of the Z-ring. Thus, the resolution of the matter of what structure is adopted by the Z-ring awaits further *in vivo* molecular studies.

1.5 Division Regulation: Regulation of FtsZ polymerization

1.5.1 Introduction

The spatial and temporal regulation of FtsZ is a major factor of where the FtsZ-ring will form and thus where the division site will be placed. There are several proteins which have been identified to interact directly and regulate Z-ring formation. These factors can be broadly categorized into two groups; those that affect FtsZ polymerization, and those that alter the positioning of FtsZ. They work in concert to prevent aberrant Z-ring formation at an inappropriate location and support the formation of a functional Z-ring at the appropriate location and time. In *E. coli*, these factors includes FtsA, ZipA, SulA, MinC, and SlmA. MinC and SlmA, which are a part of the Min and Nucleoid Occlusion systems, both regulate the position of the FtsZ ring and will be discussed further in section 1.5. ZipA, FtsA, and SulA all act to regulate FtsZ polymerization and will be discussed in turn.

1.5.2 ZipA

An integral element of Z-ring assembly is the association of FtsZ to the cell membrane, which is accomplished through the interaction of FtsZ to membrane-anchored proteins (38, 52, 53). In *E. coli*, there are two proteins that fulfill this function, ZipA and FtsA. Both proteins interact directly with FtsZ and are necessary for cell division, as the concerted deletion of both genes abrogates the formation of the Z-ring (15, 54). Additionally, although ZipA and FtsA appear to have overlapping functions, they also work collaboratively. When either is depleted, the resulting phenotype is an elongated filamentous cell with a non-functional Z-ring (15, 18).

While FtsA is conserved among bacteria, ZipA is not and is only present in γ-proteobacteria (20). The *E. coli* ZipA protein is composed of three domains: a short N-terminal membrane-anchored domain, a proline- and glutamine-rich central domain, and a C-terminal cytoplasmic domain. At the early stages of Z-ring assembly, ZipA is recruited and interacts directly with the C-terminal tail of FtsZ and anchors the FtsZ filaments to the membrane via the N-terminal domain of ZipA.

The C-terminal cytoplasmic domain of ZipA has been shown to be both necessary and sufficient for the formation of the Z-ring (29). Although, the *in vivo* interaction of ZipA and FtsZ is still unclear, *in vitro*, this cytoplasmic domain is capable of promoting the assembly of FtsZ into bundles which align laterally (28, 29). The crystal structure of the C-terminal domain of *E. coli* ZipA interacting with residues 367-383 (the last 17 amino acids) of FtsZ has revealed the atomic detail of this interaction (55). The structure revealed that the ZipA cytoplasmic domain is formed by the packing of three α-helices against a six-stranded anti-parallel β-sheet, with the topology β1-α1-β2-β3-β4-β5-α2-β6-α3 (Figure 9). The solvent exposed face of β1, β3, β4, β5, and β6 form a shallow hydrophobic cleft that allows the tail of FtsZ to bind. The ZipA-FtsZ interaction includes eleven residues of ZipA and seven of FtsZ (Asp370, Tyr371, Leu372, Ile374, Phe377, Leu378, and Gln381) (Figure 10) (55).

1.5.3 FtsA

FtsA, unlike ZipA, is very well conserved in bacteria and is a critical protein in cell division (17). Shortly after FtsZ localizes to the future site of division, FtsA and ZipA are recruited in order to facilitate the Z-ring formation. Both FtsA and ZipA function to anchor FtsZ to the membrane as well as stabilize and promote the assembly of the Z-ring in *E. coli*. FtsA shares a significant sequence similarity with the ATPase super-family which contains actin and Hsc70. The crystal structure of *Thermotoga maritima* FtsA reveals that FtsA is structurally related to actin, and consists of two major domains named relative to the domains of actin. Each domain can be further subdivided into two others yielding domains 1A, 1C, 2A, and 2B (Figure 11). Surprisingly, FtsA does not contain a domain 1B of actin. Instead, it has a domain located at a position on the opposite side that has no clear homology to known structures and is named domain 1C (56, 57). Although domain 1C has not been implicated in the interaction with FtsZ, this domain seems to play a role in the recruitment of division proteins to the FtsZ-ring (58, 59).
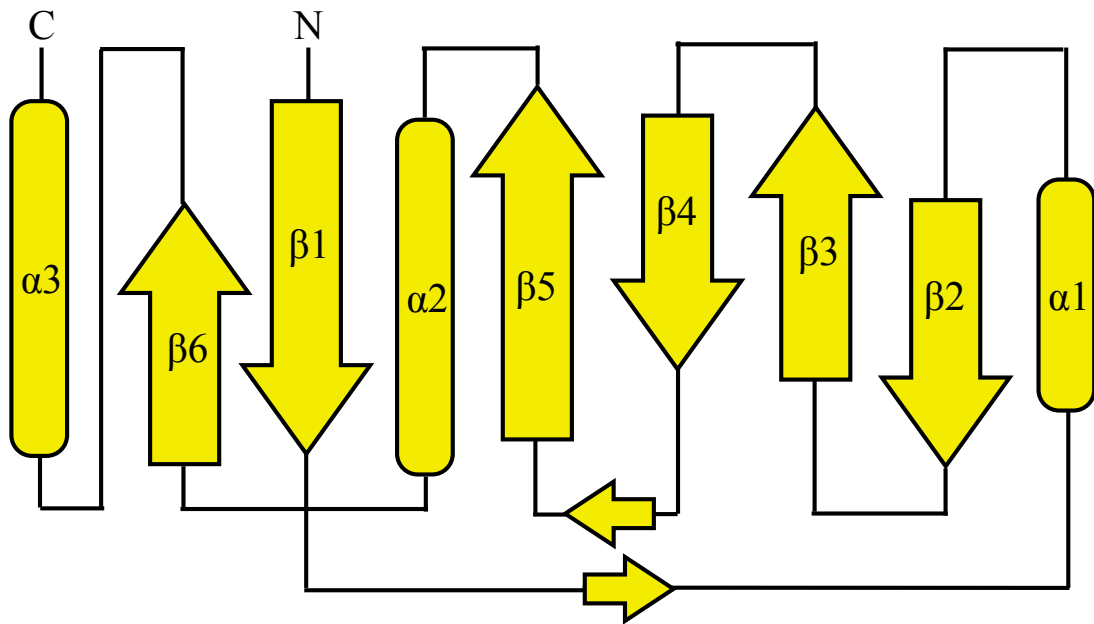
Figure 9. The structure of ZipA cytoplasmic domain. The structure revealed that the ZipA cytoplasmic domain is formed by the packing of three α-helices packing against a six-strand anti-parallel β-sheet, with the connective scheme of β1-α1-β2-β3-β4-β5-α2-β6-α3.
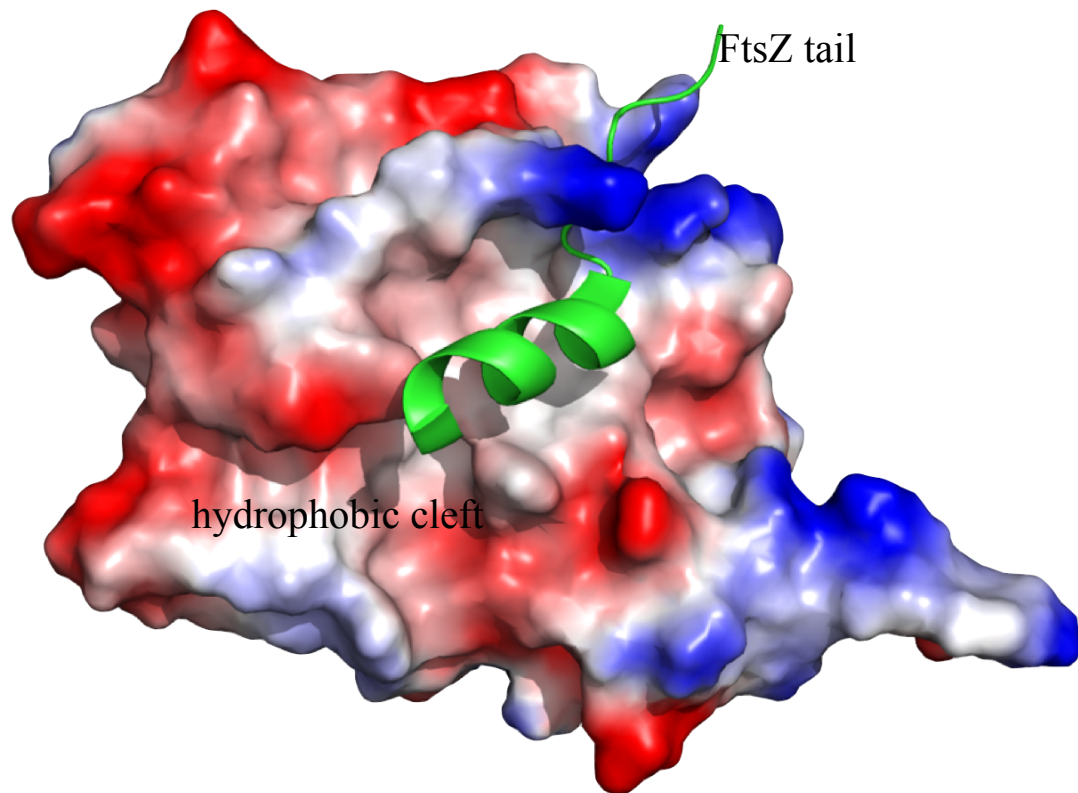
Figure 10. The crystal structure of the C-terminal domain of *E. coli* ZipA interacting with FtsZ C-terminal tail. The solvent exposed face of β1, β3, β4, β5, and β6 forms a shallow hydrophobic cleft that allows of the tail of FtsZ to bind. The ZipA-FtsZ interaction includes eleven residues of ZipA and seven of FtsZ.
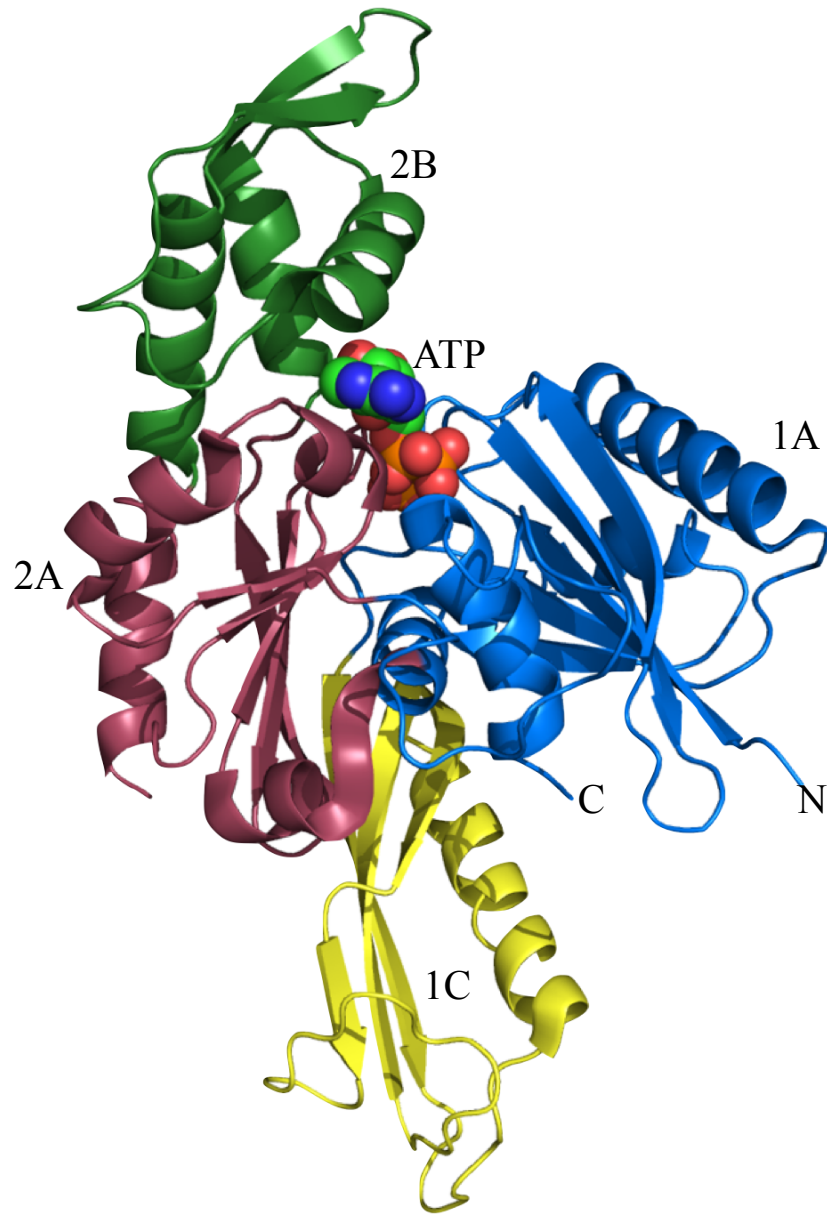
Figure 11. The crystal structure of *Thermotoga maritima* FtsA. FtsA consists of two major domains, and each domains can be subdivided into two others, yielding: domains 1A, 1C, 2A, and 2B.

Like actin, FtsA contains a nucleotide-binding pocket. The FtsA structure captured FtsA in an ATP bound form. This binding pocket is formed by subdomains 1A, 2A, and 2B (57). Although the role of nucleotide binding is still unclear, mutations of the nucleotide-binding pocket abolish the ability of FtsA to interact with itself and with FtsZ (60). Similarly to ZipA, FtsA functions to anchor FtsZ to the membrane, yet a membrane anchored domain has not been identified. However, a highly conserved motif at the C-terminus of FtsA (within subdomain 1A) contains a membrane-targeting amphipathic helix that is separated from the core protein by a flexible linker region (16). Additionally, like ZipA, FtsA also interacts with extreme C-terminus of FtsZ (15, 22, 54).

Interestingly, *E. coli* can bypass the need for ZipA with a single *ftsA* mutation, *ftsA**. This mutation stems from a single base change of C to T at position 856, which resulted in an Arg to Trp mutation at residue 286. Cells with *ftsA** assemble the Z-ring earlier in the cell-cyle and are able to tolerate higher levels of MinC than wild-type cells. FtsA* has a significantly stronger interaction with FtsZ than wild-type, and seems to be able to more effectively stabilize the Z-ring (61, 62). The location of the R286W mutation of FtsA* can be mapped to the β-strand S13 of domain 2B. Interestingly, this location appears to be well conserved in γ-proteobacteria (the subset of bacteria which contains ZipA homologs) (61). However, the molecular mechanism of how this particular residue is able to mitigate the need for ZipA remains unclear.

1.5.4 ZapA

ZapA is a small and well conserved protein consisting of 85 amino acids. ZapA is present and directly interacts with FtsZ in both *B. subtilis* and *E. coli*. This interaction promotes the assembly and stability of the Z-ring (30, 63, 64). The structure of ZapA is composed of 2 β-strands followed by 2 α-helices. The crystal packing revealed that ZapA can interact with itself to form a homo-tetramer formed by a pair of dimers interacting via a coiled-coil domain (Figure 12) (64). Although the exact mechanism of how ZapA is able to promote FtsZ-ring assembly is not clear, it does seem to be directly correlated with the ability of ZapA to inhibit the GTPase activity of FtsZ (64).

1.5.5 SulA

SulA is a cell division inhibitor which functions in the SOS response system. In particular, SulA is expressed in response to DNA damage (65-67). SulA interacts with FtsZ directly and is able to effectively disrupt the Z-ring. *In vitro* studies showed that SulA is able to increase the
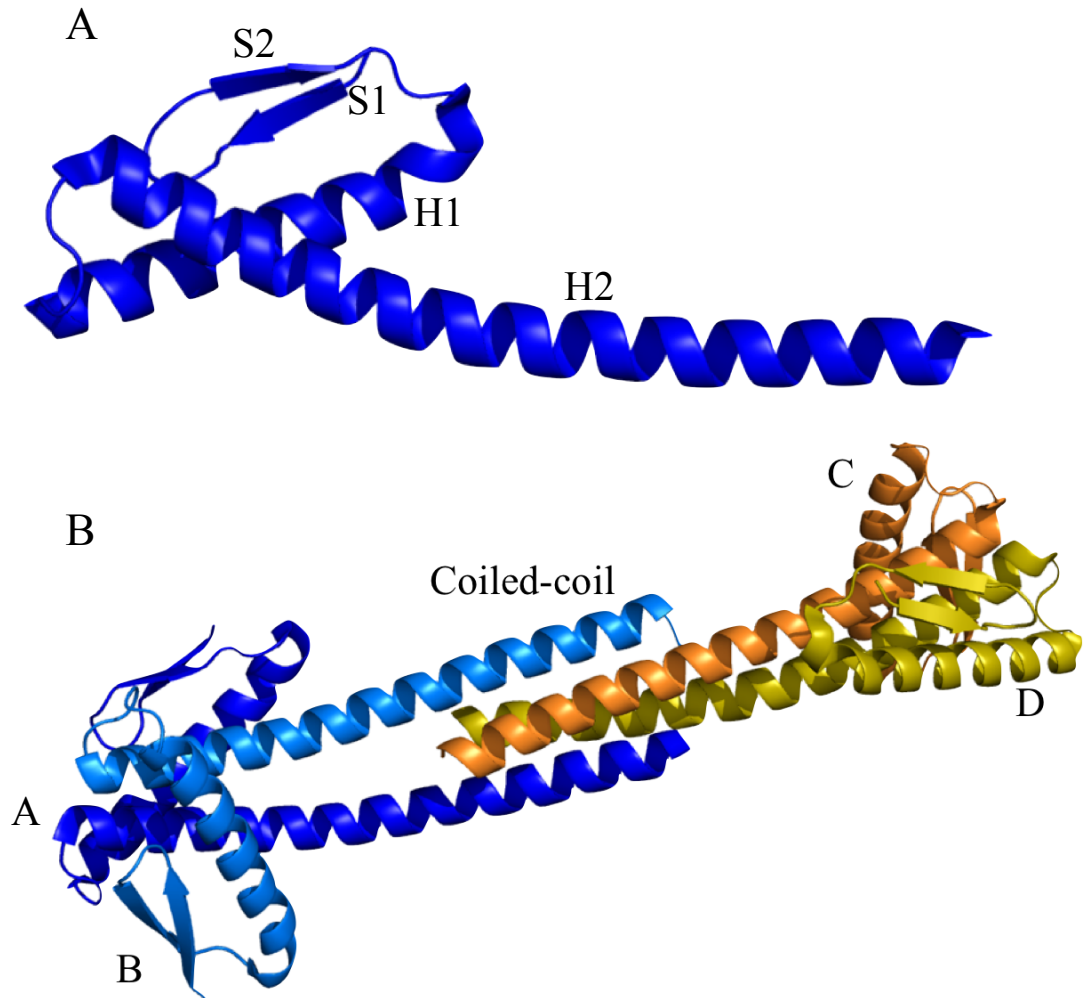
Figure 12. Crystal structure of ZapA. A) A ZapA subunit is composed of 2 β-strands followed by 2 α-helices. **B)** The crystal structure of ZapA revealed a homo-tetramer that is formed by a pair of dimers interacting via a coiled-coil domain.

critical polymerization concentration of FtsZ by five-fold (67). The crystal structure of SulA revealed that this 18 kDa protein consist of a central region of parallel β-strands sandwiched by four α-helices with the following topology: β1-α1-β2-α2-β3-α3-β4-α4-β5 (68) (Figure 13A). The structure of SulA bound to FtsZ demonstrates the mechanism of FtsZ de-polymerization. SulA interacts with the T7 loop of FtsZ and effectively "caps" one end of the FtsZ protofilament (Figure 13B). Additionally, SulA is able to reduce the GTPase activity of FtsZ through this interaction (68).

1.6 Division Regulation: Regulation of FtsZ position

1.6.1 Introduction

There are two major cell processes which spatially and temporally regulate FtsZ-ring position, the Min and Nucleoid Occlusion systems. These cellular processes work in conjunction with one another to ensure that the Z-ring forms at the appropriate position. Briefly, the Min system inhibits FtsZ polymerization at the cell poles, preventing the formation of anucleated mini-cells. The second system, Nucleoid Occlusion, prevents the assembly of the Z-ring over the nucleoid. To ensure that cytokinesis does not occur before chromosome segregation has been completed, as the consequence of premature division can cause shearing of the chromosomes and a deleterious phenotype.

1.6.2 The Min System

The cellular concentration of FtsZ has been quantified to be as high as 10 μM, which is much higher than the 1-2 μM concentration that is needed to induce protofilament formation *in vitro* (35). Regulation of Z ring assembly is therefore critical for cytokinesis. In *E. coli*, the Min system is composed of three proteins; MinC, MinD, and MinE. These proteins function in concert to destabilize the Z-ring. The MinCDE complex oscillates from pole to pole and creates a gradient of division inhibition at the two poles whilst leaving the mid-cell with the least amount of inhibitory signal (Figure 14) (8). MinC is a division inhibitor; it interacts directly with FtsZ and prevents the formation of FtsZ protofilaments (69). MinE act as the specificity factor that is responsible for the oscillation of the MinCDE complex and development of the gradient (70). MinD is a membrane protein responsible for membrane association of MinC and MinE (71, 72). The MinC and MinD structures have been solved individually and provide additionally insight in to their respective functions.

Figure 13. Structure of the SulA homo-dimer and the SulA-FtsZ complex. A) The crystal structure of SulA revealed that this protein forms a homo-dimer and consist of a central region of parallel β-sheets sandwiched by four α-helices; in this topology: S1-H1-S2-H2-S3-H3-S4-H4-S5. **B)** The structure of SulA bound to FtsZ demonstrated the binding mechanism of FtsZ polymerization. SulA interacts with the T7 loop of FtsZ.

Figure 14. The MinCDE oscillation cycle. The MinCDE polar zone begins assembling at a cell pole and grows towards midcell (1–2 and 5–6). The MinE ring then assembles at the leading edge of the polar zone (3 and 7). The polar zone then disassembles, releasing MinC, MinD and MinE molecules, shrinking back to the pole, and finally releasing MinE from the E-ring (4–5 and 8–1). Because of the rapid oscillation, a zone of division inhibition (dark blue shading) is present near the two ends of the cell for a large portion of the cell cycle.
Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Microbiology (Rothfield et al, 2005), copyright (2005).

The crystal structure of *Thermotoga maritima* MinC revealed that MinC consists of two domains connected by a short and flexible linker (Figure 15A). The C-terminal domain is a right-handed β-helix (Figure 15B) and is involved in MinC dimerization. The asymmetric unit contains two MinC dimers which demonstrates the flexibility of the linker region that connects the N and C-terminal domains. The N-terminal domain consists of two α-helices and five β-strands with the following topology: β1-β2-α1-β3-α2-β4-β5 (Figure 15C) (73). This domain is able to directly interact with the FtsZ C-terminal tail. Interestingly, MinC is able to bind to the same region as FtsA (74, 75).

The MinD structure exhibits a fold that is similar to those of other ATPases and has the following topology: β1-α1-β2-α2-α3-β3-β4-α4-α5-β5-α6-β6-α7-β7-α8-β8-α9-α10-α11 (Figure 16A) (76). Analysis of different nucleotide-bound states of MinD demonstrates that the hydrolysis of ATP does not appear to be directly coupled to a conformational change. Further structure-based site-directed mutagenesis illustrates that the residues around the MinD nucleotide binding pocket are important for its interaction with MinC. In particular, residue Lys11 occupies a central position in the protein-nucleotide binding network and mutation of this residue can disrupt MinC interaction. Other residues involved in the MinD-nucleotide interaction network include Glu144, Ser146 and Asp150. Not surprisingly, all these residues are highly conserved in other bacteria (Figure 16B) (77).

1.6.3 Noc mediated Nucleoid Occlusion in *Bacillus subtilis*

Woldringh *et al.*, first proposed the nucleoid occlusion effect almost two decades ago (78-80). The proposed model suggests that the effect of molecular crowding, stemming from the combined activities of transcription and translation, can have a local inhibitory effect on division in the region the nucleoid occupies. As the nucleoid segregates, the inhibitory effects are diminished in the region between the two nucleoids to allow for the resumption of division. Thus, the process of cytokinesis initiates after chromosome segregation to ensure the integrity of the genetic material (81).

Although the view of nucleoid occlusion as an indirect effect of cellular events is plausible, the effect of nucleoid occlusion is predictable and routinely observed, indicating a well-regulated process. This is indeed the case, as exciting news for the field came in 2004 when Wu and Errington reported the discovery of Noc as an effector of nucleoid occlusion in *Bacillus subtilis* (82). Noc was first identified as a ParB-like gene that has high similarity (35%
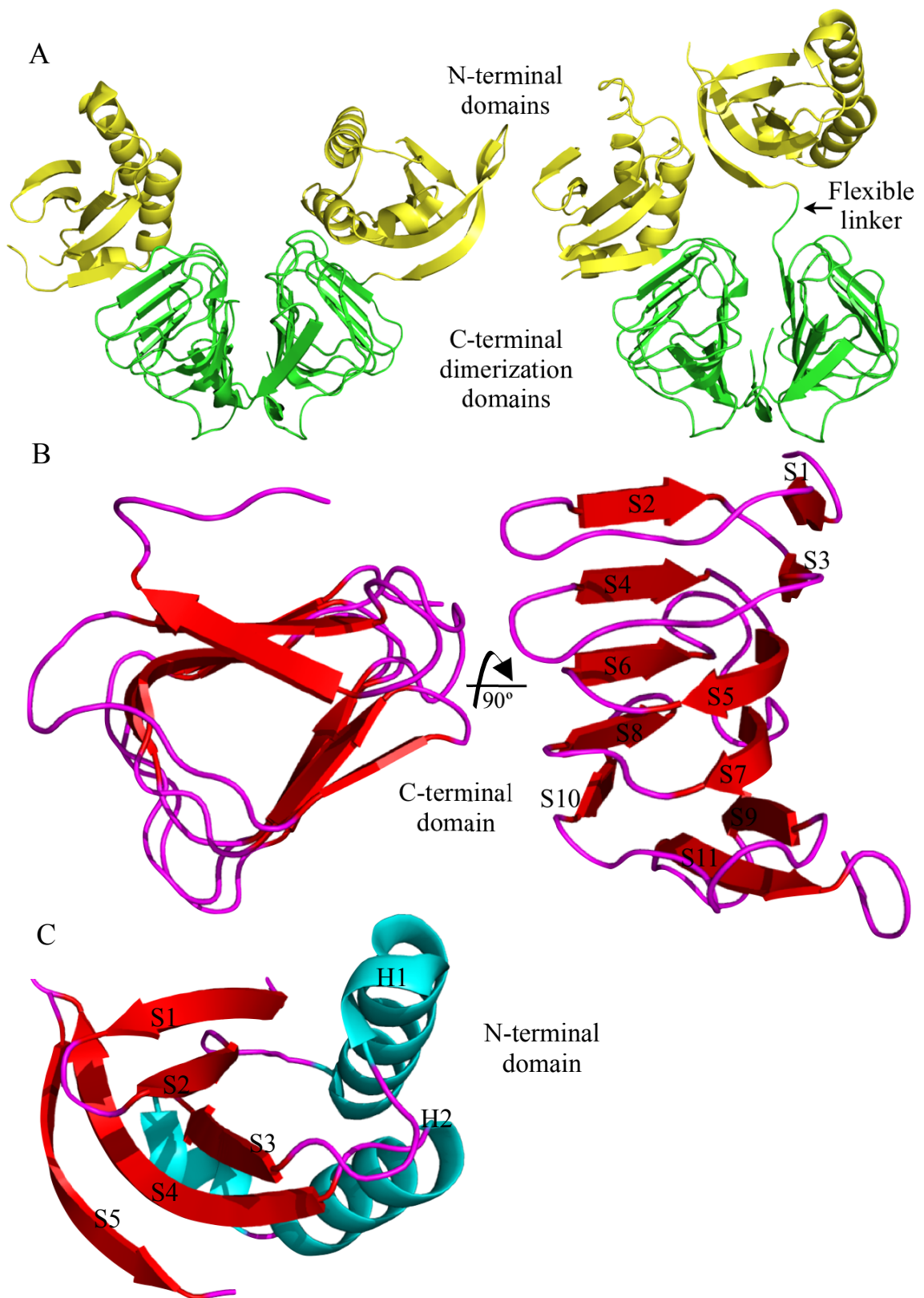
Figure 15. Structure of MinC. A) The crystal structure of *Thermotoga maritima* MinC revealed that MinC consists of two domains, N and C-terminal domains. The C-terminal domain dimerizes to form a homo-dimer. Two dimer forms were captured in the crystal structure, and comparison of these two dimers showed that the N and C-terminal domains are connected by a short and flexible linker. **B)** The C-terminal domain is a right-handed β-helix. **C)** The N-terminal domain is formed with two α-helices and five β-strands in the following topology: β1-β2-α1-β3-α2-β4-β5.
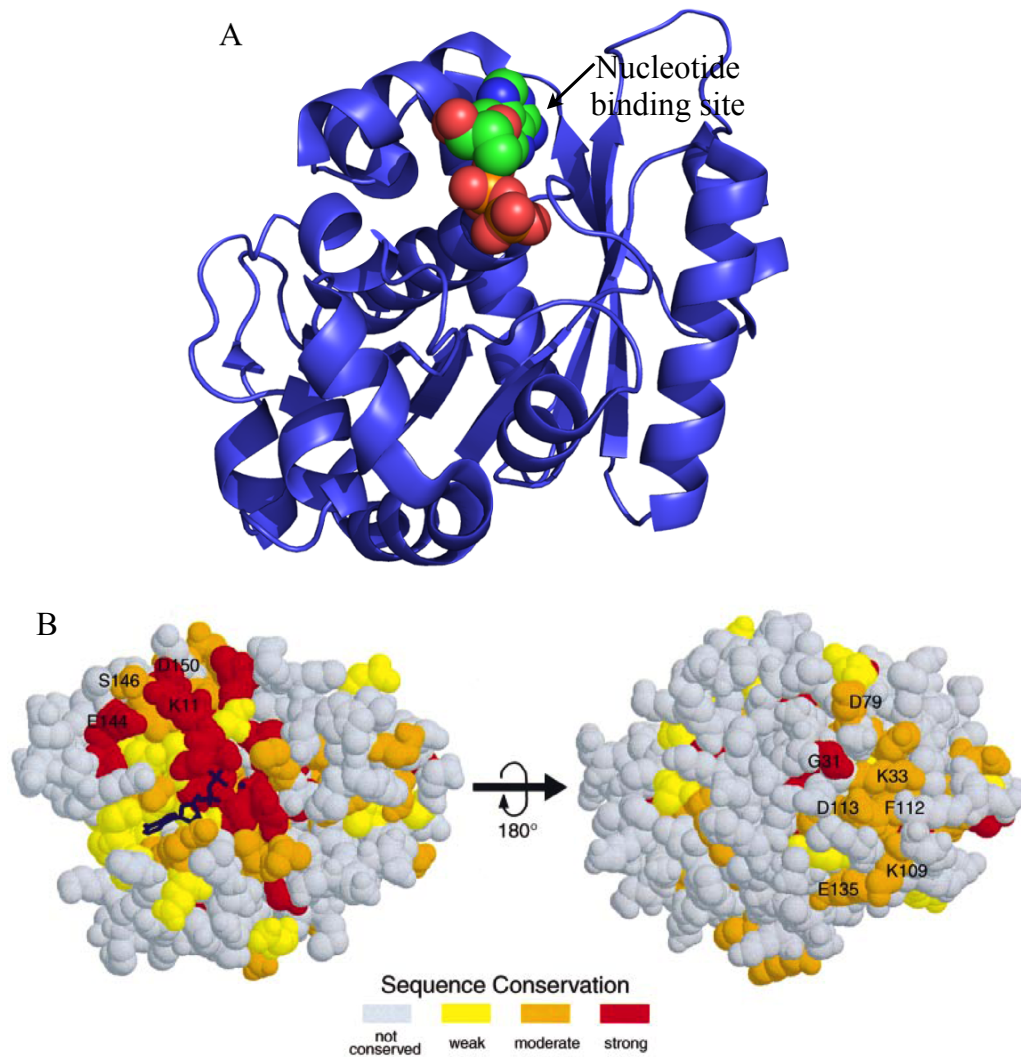
Figure 16. Structure of MinD. A) The MinD structure exhibit a fold that is similar to those of other ATPases and has the following topology: β1-α1-β2-α2-α3-β3-β4-α4-α5-β5-α6-β6-α7-β7-α8-β8-α9-α10-α11. **B)** Sequence conservation for MinD homologous proteins mapped on the structure. The alignment analysis was done with the seven MinD homologous proteins. Variable regions are colored in white, and increasing conservation is indicated with deepening red color. Two views of the protein are shown, including the bound nucleotide represented by thickened blue bonds. The conserved residues indicated in the right panel are located mostly at the N-termini of the β-strands.

identity) to Spo0J. However, Noc does not seem to play a role in chromosome segregation even though it has been observed to bind DNA and co-localized with the nucleoid (83).

Interestingly, cells with a defective *noc* gene display a wild-type phenotype. Those that over-expressed Noc are still viable, but display a phenotype of longer cells and delayed cell division. However, in conditions where the Noc-defective cells are perturbed by either the addition of HPUra (a chemical compound that inhibits replication of DNA) or depletion of DnaA (a critical replication initiation factor), a fraction of the cells still exhibited septation events through the nucleoid. These results indicated that although Noc is important in mediating nucleoid occlusion, there might be a Noc-independent system that can bias the division machinery away from the nucleoid. Additional evidence to suggest that Noc is important in cell division and division placement comes from the observation that in double mutant of Noc and MinD, the cells display a filament-like phenotype that can be attributed to an arrest in cell division (82).

Although a Noc interaction partner that is responsible for division inhibition is not known, the DNA binding capabilities has been identified as critical to its role as a nucleoid occlusion factor. Curiously, Noc is able to bind to a 14-bp long inverted repeat DNA sequence which is absent from the terminal region of the chromosome (Figure 17). The importance of this is clear when one considers that shortly after the completion of replication, the Ori, along with the Left and Right macro-domains of the chromosome segregate towards the cell poles, while the Ter region remains at the mid-cell well into the septation phase. The presence of these Noc binding sites allows for the concurrent localization of Noc protein away from the mid-cell as the chromosomes segregate, alleviating nucleoid occlusion at the mid-cell and signaling cell division machinery to assemble (84). This mechanism of coordinating chromosome segregation with cell division is also seen in *Caulobacter crescentus*. In the case of *C. crescentus*, MipZ is localized towards the origin region through its interaction with ParB. MipZ is thus able to form a gradient at the polar region that depolymerizes FtsZ filaments at the poles. Therefore MipZ is also regulated spatially by the segregation of the chromosome (85).

1.6.4 SlmA mediated Nucleoid Occlusion in *Escherichia coli*

Nucleoid occlusion is a safety mechanism that prevents the inappropriate formation of the Z-ring over the nucleoid. This mechanism is mediated by two unrelated nucleoid occlusion factors, Noc (in *B. subtilis*) and SlmA (in *E. coli*). Noc is discussed in section 1.5.3. Complementarily to the discovery of the nucleoid occlusion protein in gram-positive *B. subtilis*,
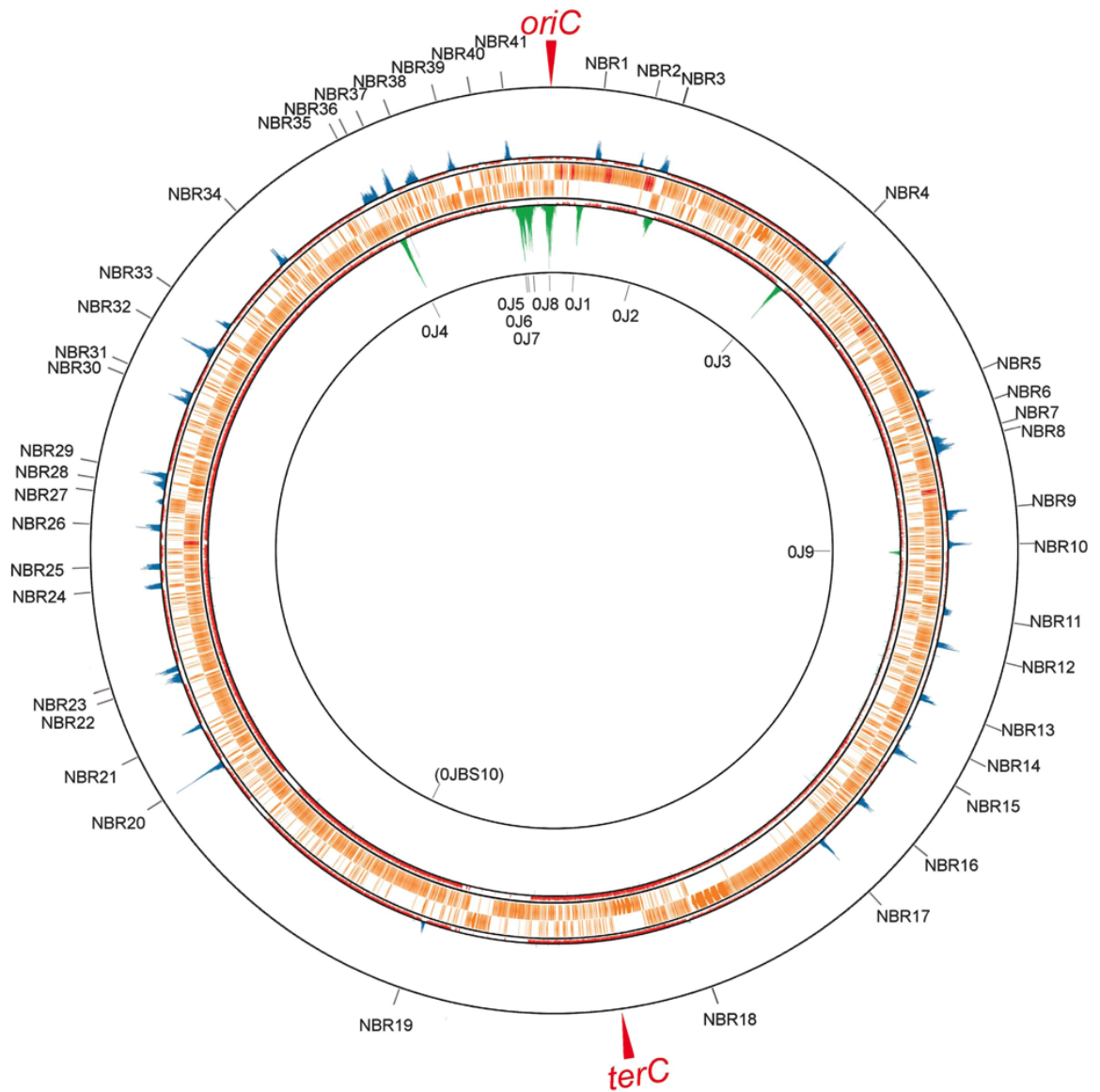
Figure 17. Genome-wide distribution of preferred NBRs mapped by ChAP-on-Chip. Noc (outer rings) and Spo0J (inner rings)-binding signals in wild-type strains (4704 and SI002), and shown at their corresponding genome coordinates. Top and bottom lines indicate signal intensities of 20 and 0, respectively. Middle lines exhibit threshold values used to define the binding regions of Noc (1.5) and Spo0J (1.8). Signals above and below the threshold values are shown as blue and pink lines, respectively. ORFs (orange bars), rRNA and tRNA (red bars) are also indicated between them. The IDs of Noc binding regions (NBRs) detected by our algorithm are shown at the outermost ring; 0J1–0J9 correspond to the Spo0J-binding sites. Reprinted by permission from Macmillan Publishers Ltd: The EMBO Journal (Wu et al, 2009), copyright (2009).

Bernhardt and de Boer discovered SlmA, which is used by gram-negative *E. coli*. Cells with a defective *slmA* gene exhibit a lethal phenotype when cell division is stalled via the repression of *dnaA*. In these cells, the septal ring clearly forms over the nucleoid resulting in the fragmentation of the nucleoid (Figure 18). Furthermore, SlmA functions as a cell division inhibitor; when over-expressed by 50 fold, cell division is completely blocked (7, 86).

Through sequence analyses, SlmA was predicted to contain two structural motifs: an N-terminal helix-turn-helix and a C-terminal coiled-coil. The N-terminal domain is responsible for mediating the association of SlmA to the nucleoid. Additionally, without this domain, SlmA is unable to effect nucleoid occlusion in the cell. Notably, through the use of light scattering experiments, it was observed that SlmA interacts with FtsZ *in vitro* in a GTP-dependent manner. Moreover, SlmA is localized to the nucleoid (Figure 19) and also recruits FtsZ to the nucleoid (86).

SlmA is able to mediate nucleoid occlusion by associating with the nucleoid and affecting the formation of the Z-ring through its interaction with FtsZ. Thus the ability of SlmA to interact with both DNA and FtsZ is critical for the occurrence of nucleoid occlusion. The current literature suggests two models for SlmA mediated nucleoid occlusion. DNA bound SlmA can competitively bind to FtsZ and out compete membrane bound septal ring components such as ZipA and FtsA. Thus, SlmA is able to inhibit Z-ring formation by passively localizing FtsZ near the nucleoid and away from other division proteins needed for Z-ring formation. In a second model, SlmA, in combination with an unknown factor, actively promotes the disassembly of FtsZ polymers (86).

1.6.5 Tetracycline repressor (TetR) family of transcriptional repressor proteins

Currently, approximately 86 TetR proteins have been characterized. These proteins are involved in a variety of cellular processes such as transcriptional control of multidrug efflux pumps, genes responsible for biosynthesis of antibiotics, osmotic stress, and pathogenicity (87). Notably, TetR proteins all function as transcriptional regulators, with the exception of SlmA. Contrary to previous sequence analysis, which suggested that SlmA contains a coiled-coil domain, this work clearly demonstrates that SlmA is a novel member of the TetR family. Moreover, SlmA is the first TetR protein that does not function as a transcriptional repressor.

In order to appreciate and understand how SlmA is able to effect nucleoid occlusion, it is necessary to consider the structural and functional aspects of TetR proteins. Specifically, we will discuss two prominent members of the family, TetR and QacR, to explore the mechanism
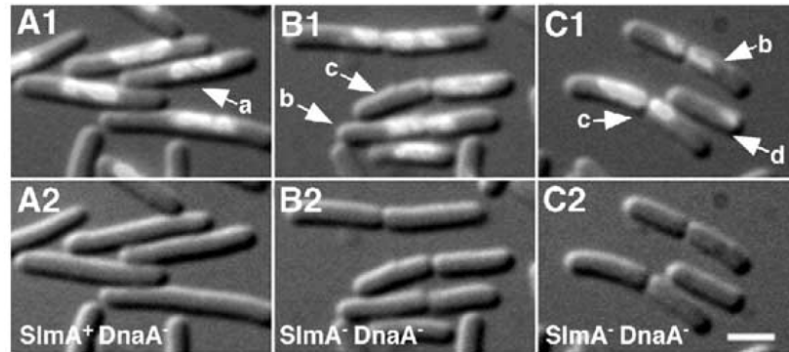
D	strain	total cells	cell length[a] (μm)	nucleoid length[a] (μm)	total septa	septal position[a,b]	nucleoid position[a,b]	nucleoid-free cells
	TB104 [λP$_R$::dnaA]	110	6.4(1.6)	2.7(0.8)	1	0.23	0.51(0.04)	2
	TB105 [λP$_R$::dnaA slmA::frt]	135	5.5(1.9)	2.7(0.8)	49	0.5(0.04)	0.52(0.1)	22

[a] Values are averages. Standard deviations are in parentheses.
[b] Values are given as fraction of cell length. The position of the nucleoid center is given.

Figure 18. Nucleoid cutting in SlmA⁻ DnaA⁻ *Escherichia coli* cells. (A) *E. coli* cells with *slmA* **(B and C)** E. coli cells without *slmA* were grown in LB for 3.5 hr at 30°C to deplete DnaA. Cells were fixed, stained with DAPI, and imaged with DAPI- and differential interference contrast (DIC) specific optics. A1, B1, and C1 show a digital overlay of the DIC and DAPI images, and A2, B2, and C2 show the DIC image only. Bar = 2 m. Several parameters of randomly selected cells from each strain were measured, and the results are summarized in **(D)**.

Reprinted from Molecular Cell. Vol 18/Issue 5, Bernhardt, T. G., and P. A. J. de Boer, SlmA, a nucleoid-associated, FtsZ binding protein required for blocking septal ring assembly over Chromosomes in *E. coli*, 555-564, Copyright (2005), with permission from Elsevier.
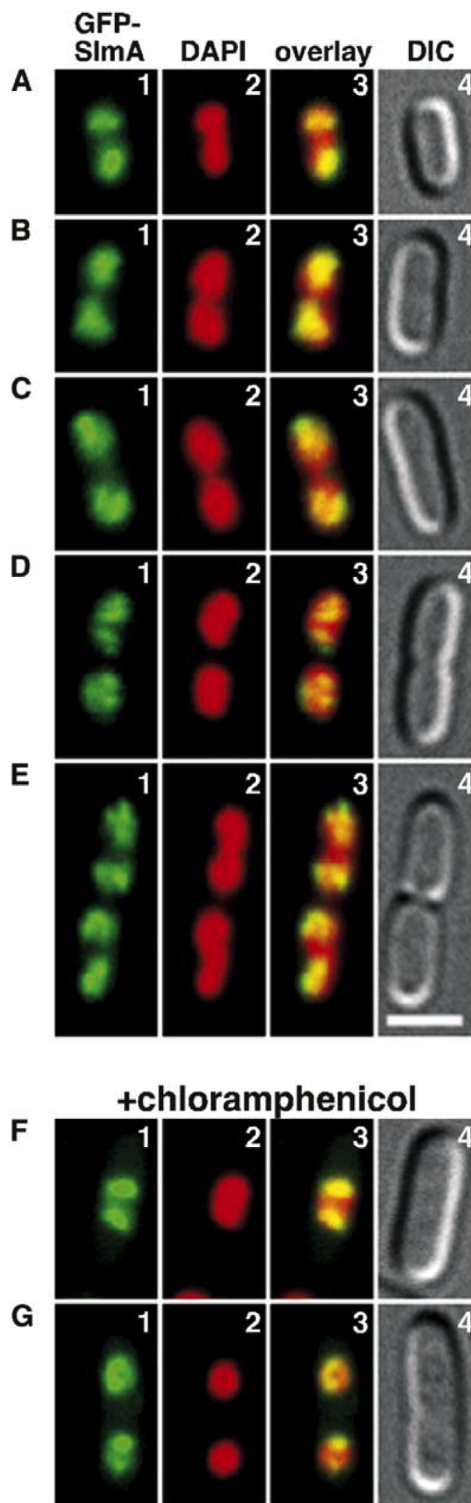
Figure 19. Distribution of GFP-SlmA on the Nucleoid. Shown are live cells of TB85(HKTB99) [_slmA::frt (Plac::gfp-slmA)] grown to OD600 = 0.5–0.6 at 30°C in LB with 250 μM IPTG. DAPI was added to 0.25 μg/ ml 30 min prior to imaging. Cells in (F) and (G) were treated with chloramphenicol (100 μg/ml) and grown for an additional 30 min prior to viewing. Panels show GFP-SlmA (1), DAPI (2), merged (3), and DIC (4) images. Bar = 2 μm.

Reprinted from Molecular Cell. Vol 18/Issue 5, Bernhardt, T. G., and P. A. J. de Boer, SlmA, a nucleoid-associated, FtsZ binding protein required for blocking septal ring assembly over Chromosomes in *E. coli*, 555-564, Copyright (2005), with permission from Elsevier.

of how TetR family members bind DNA and effector molecules. The TetR family of proteins was named after the founding member, TetR. This protein is a transcriptional repressor of the *tet* genes, whose products confer resistance to tetracycline. Specifically, TetR binds to the tet-operator site leading to repression of transcription of the *tetA* gene. However, in the presence of tetracycline, TetR binds to the antibiotic and is induced from the DNA leading to transcription of *tetA*. Subsequently, the TetA protein is responsible for the efflux of tetracycline out of the cell and thus confers resistance to the tetracycline (87).

TetR is an all-helical protein with 10 α-helices, which together forms two domains: an N-terminal DNA binding domain and a C-terminal dimerization domain. The N-terminal domain is formed by helices α1-α3. Helix α4 connects the N-terminal domain to the dimerization domain which is formed by helices α5-α10. The TetR dimer is stabilized by hydrophobic helix-to-helix contacts of helices α8-α10 packed against their dyadic-mates. The structure of TetR bound to a palindromic 15 bp *tet*-operator fragment shows a homodimer with each helix-turn-helix motif binding to two adjacent major grooves of the palindromic sequence (Figure 20A) (88). Helices α3 and α3' (of the second subunit) are know as the "recognition helices" and are responsible for making base-specific contacts, which allows TetR to bind to DNA in a sequence specific manner. Additionally, residues in these two helices are responsible for most of the DNA phosphate contacts (Figure 20B).

The TetR dimer contains two identical tetracycline binding pockets, and both are able to bind tetracycline concurrently. While a majority of the pocket is formed by helices α5, α8, and α10, the entrance to the pocket is "gated" by α9' of its dyad-mate. When bound, tetracycline makes contact with His100, Thr103, Arg104, and Pro105 (mediated by a Mg^{2+} ion). Critically, His64 of α4 acts as a pivot joint, which rotates as it interacts with tetracycline. This motion causes the recognition helix α3 to shift outward and disrupts its ability to bind to the major groove of the DNA (Figure 20C-D) (89).

A second prominent member of the TetR family is QacR, which is found on a number of nultidrug resistance plasmids harbored in *Staphylococcus aureus*. Interestingly, although the *qac* system is very similar to the *tet* system, the details of the mechanisms utilized by QacR and TetR reveal different modes of DNA and effector molecule binding. Unlike TetR, QacR is composed of 9 α-helices and contains an N-terminal domain formed by helices α1-α3, and a C-terminal domain formed by helices α4-α9. Functionally, QacR acts as a transcriptional repressor of the *qacA* multidrug transporter gene, whose product confers resistance to
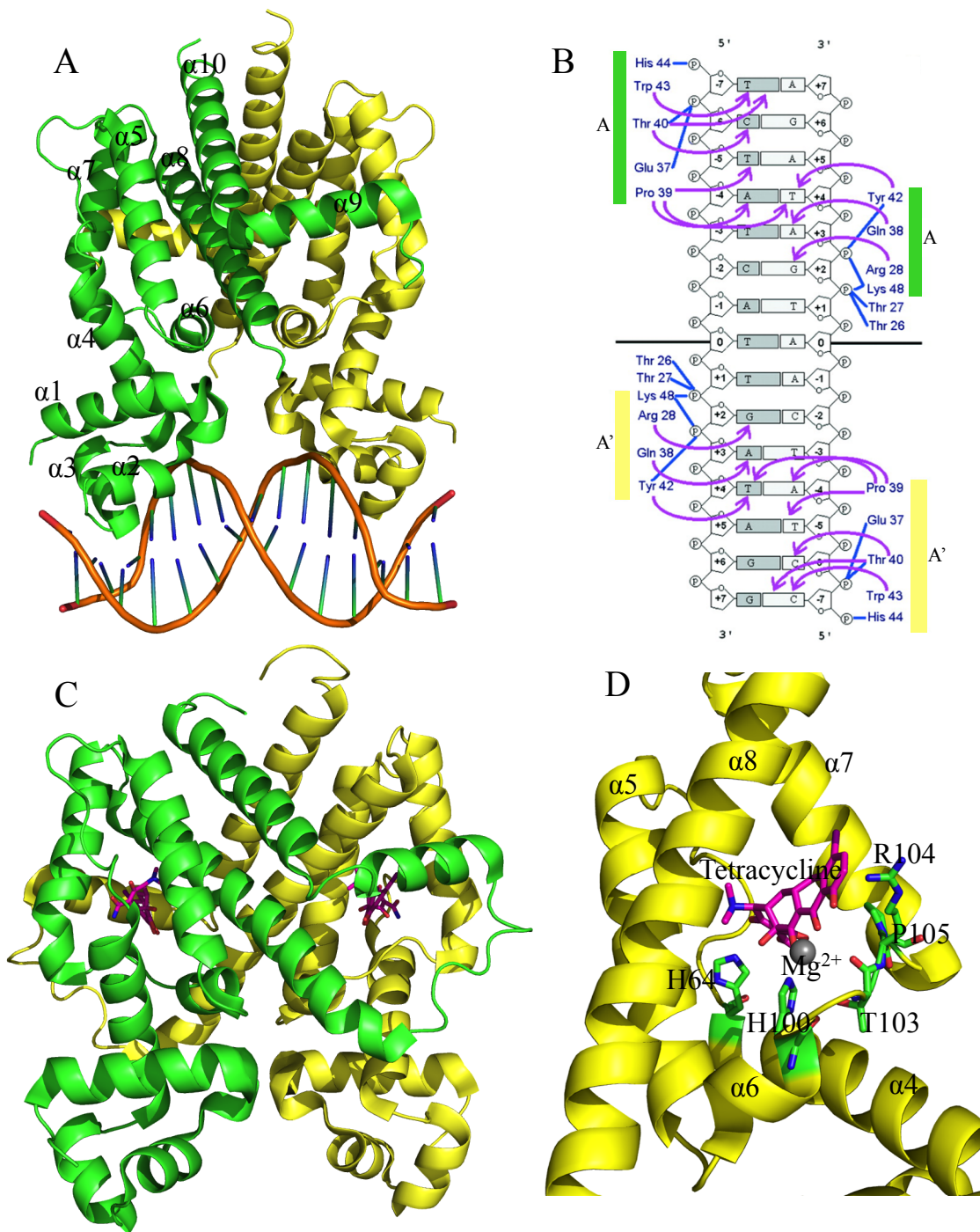
Figure 20. Structure of TetR bound to a fragment of the *tet*-operator site. A) TetR is an all-helical protein with 10 α-helices with two domains; an N-terminal DNA binding domain which contains a helix-turn-helix motif, and a C-terminal dimerization domain that also contain the binding pocket for its inducer-molecule(s). The structure of TetR bound to a palindromic, 15 bp *tet*-operator fragment shows that the biologically relevant state of TetR is a homodimer, and each helix-turn-helix motif binds to the major grooves of the palindromic sequence. **B)** Schematic representation of interactions between TetR and the 15 base pair operator. Hydrogen bonds are shown as blue lines, van der Waals interactions (3.5 Å) by purple arrows. **C)** Structure of TetR in complex with Tetracycline. **D)** Close-up of the binding pocket with the tetracycline making contact with His100 and Thr103, mediated by a Mg^{2+} ion, in 6 and induces a conformational change in loop 6. Additionally, Arg104 and Pro105 makes stabilizing contacts. Critically, His64 of helix α4 acts as a pivot joint which is rotated as it interacts with the tetracycline.

Reprinted by permission from Macmillan Publishers Ltd: Nature Structural Biology (Orth et al, 2000), copyright (2000). Reprinted by permission from The American Society for Microbiology: Microbiology and Molecular Biology Reviews (Ramos et al, 2005), copyright (2005).

monovalent and bivalent cationic lipophilic antiseptics and quaternary ammonium compounds (90). In the absence of these antibacterial drugs, QacR binds to an operator site that has two overlapping partial palindromes (Figure 21A-B). This is in contrast to the single symmetric operator site observed in the TetR system. Interestingly, unlike TetR, two QacR homodimers cooperatively bind to the operator site, with each dimer binding to consecutive majors grooves in the DNA (91). The recognition helix (α3) of each subunit is responsible for the majority of contacts established with the DNA. The critical difference in binding modes between TetR and QacR is observable through the conformation changes in the DNA. While TetR induces a 17° bend towards the protein during binding, QacR widens the major groove from 34 Å to 37 Å (91). This widening of the major groove, by the first QacR dimer, allows the DNA to accommodate the second dimer. Studies suggest that the two dimers must bind almost simultaneously and cooperatively in order to sustain the deformation of the DNA (92, 93). Interestingly, although these two proteins employ two different modes to bind DNA, they both function as transcriptional repressors (94).

The drug binding pocket of QacR is formed by helices α4-α9, and is able to "sense" a number of cationic, lipophilic drugs (95). Although, the pockets formed by each subunit are identical, equilibrium dialysis studies and isothermal titration calorimetry data have shown that only one subunit within the dimer actually binds the drug molecule (96). The crystal structures of QacR bound to different molecules have revealed remarkable insights into the versatility of this protein (Figure 22A-B). The QacR binding pocket is able to expand from a volume of ~400 Å^3 to ~1,100 Å^3 during drug binding. The pocket is rich in aromatic and acidic residues, which can accommodate positively charged drugs (96-101). Notably, the different structures revealed the presence of several "mini-pockets" within the larger binding pocket, which shows that QacR can tailor its binding site to different drugs. Additionally, it also suggests that multiple drugs may be able to bind simultaneously. In 2004, Schumacher *et al.* showed through crystallography and near-ultraviolet circular dichroism that QacR can, indeed, bind to two drugs, ethidium and proflavin, simultaneously. Each compound was bound in a separate mini-pocket. This was made possible by the malleability of the binding pocket (Figure 22C-D) (98).

After drug(s) binding, the QacR subunit undergoes a major conformational shift that leads to induction. Drug binding triggers a coil-to-helix transition which extends helix α5 and relocates helix α6 and the DNA binding domain. Altogether, this leads to a 9 Å translation and a 37° rotation of the DNA-binding domain and effectively releases QacR from the DNA (96).
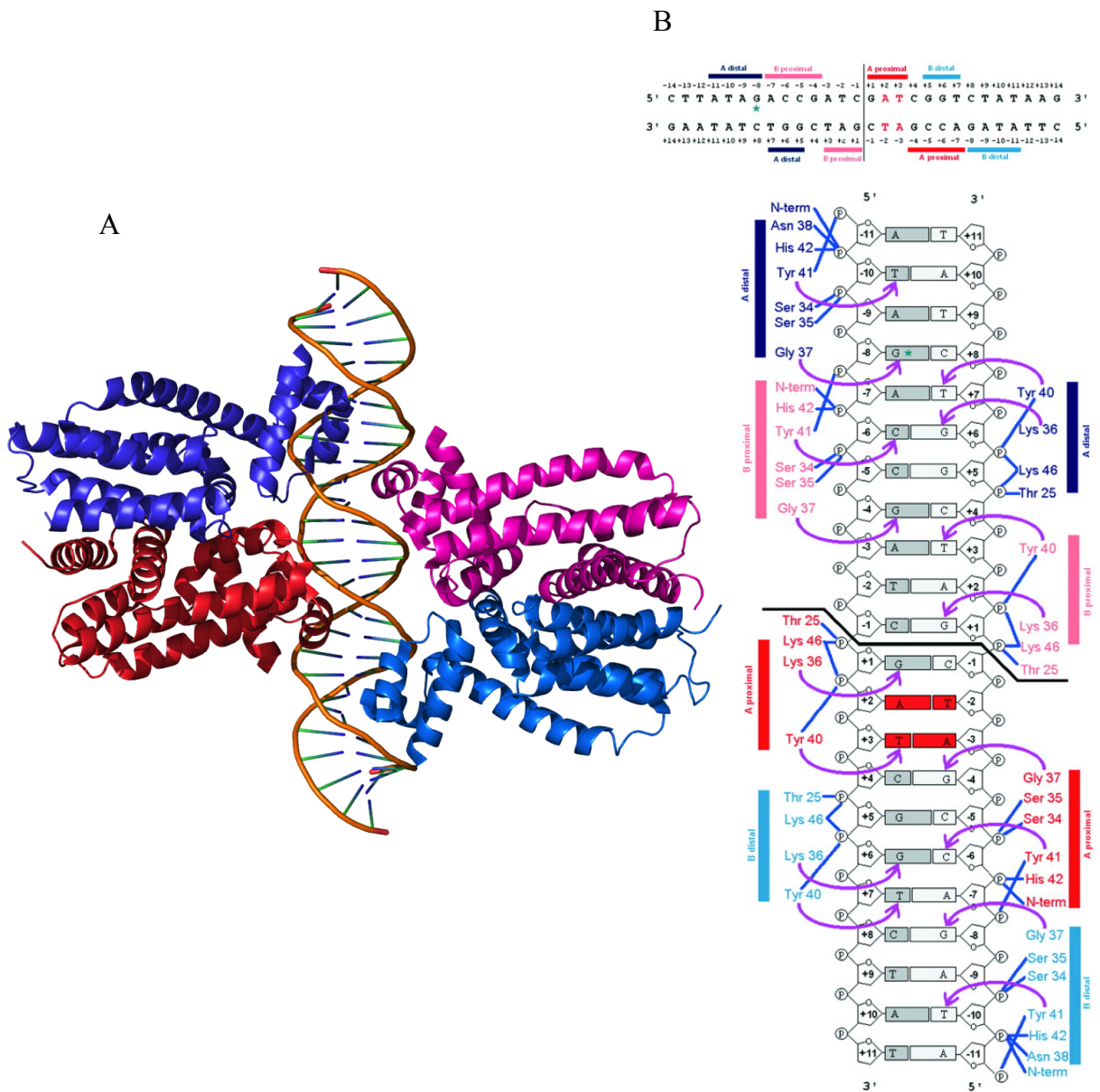
Figure 21. Structure of QacR bound to its operator site. A) QacR is an all-helical protein, made up of 9 α-helices, with an N-terminal DNA binding domain and a C-terminal dimerization domain. Two QacR homodimers cooperatively bind to the operator site that has two overlapping, partial palindromes. QacR dimers, labeled as proximal and distal, with respect to the position of the two-fold axis of symmetry of the operator. **B)** Schematic representation of interactions between QacR and the operator. Hydrogen bonds are shown as blue lines, van der Waals interactions (3.5 Å) by purple arrows. The respective subunit are colored as in (A).

Reprinted by permission from Macmillan Publishers Ltd: EMBO Journal (Schumacher et al, 2002), copyright (2002).

Reprinted by permission from The American Society for Microbiology: Microbiology and Molecular Biology Reviews (Ramos et al, 2005), copyright (2005).
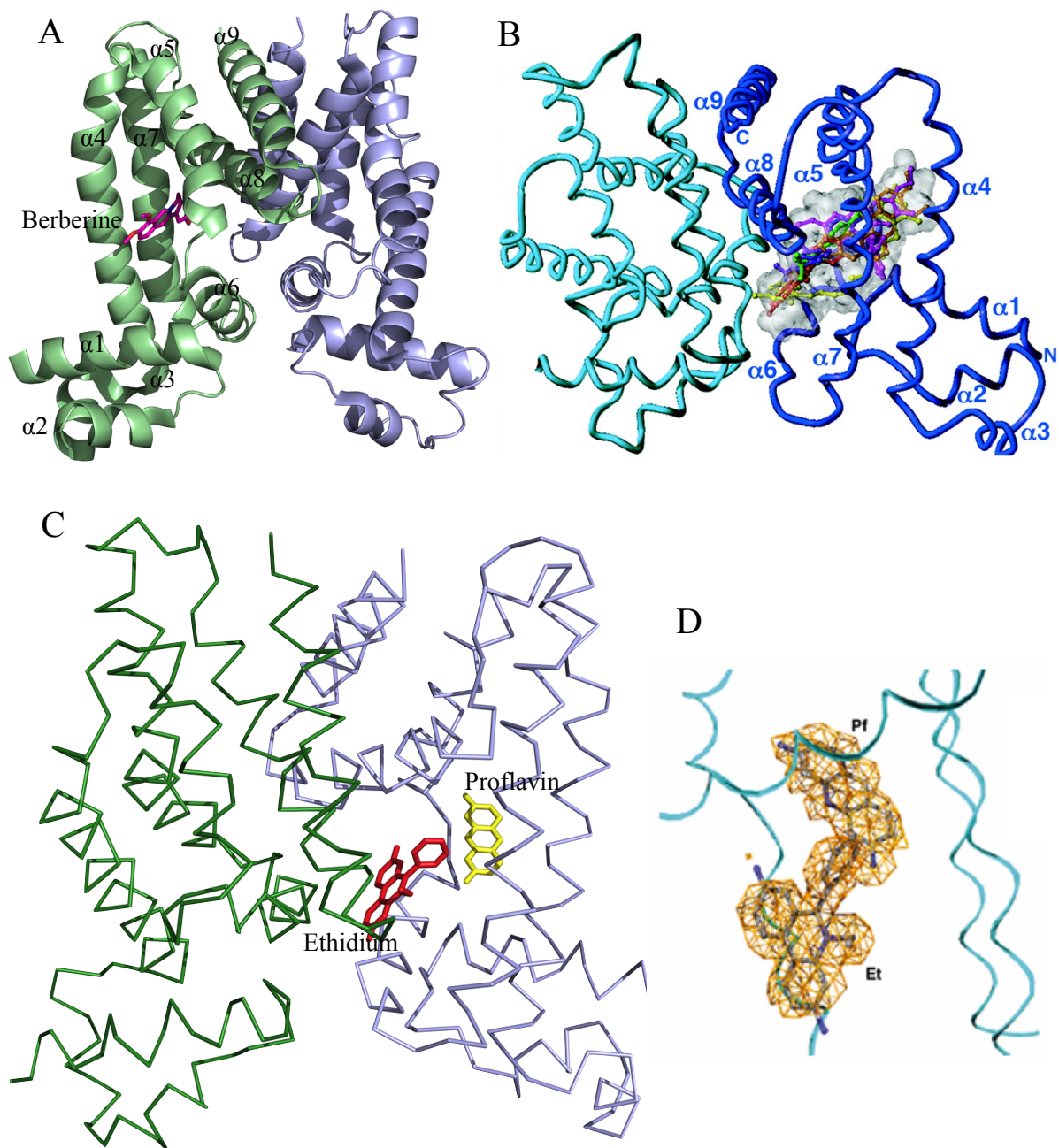
Figure 22. Structure of QacR bound to effector molecule(s). A) The effector molecule binding pocket of QacR is formed by helices 4-7. Only one subunit of the QacR dimer binds to the effector molecule. The QacR-berberine complex is shown with the subunit colored green and the berberine colored magenta. **B)** Ribbon diagram of the drug-bound QacR dimer looking down from the "top" of the dimer. The drug-bound subunit is colored dark blue and the other is cyan. The binding-site volume is depicted as a transparent surface. Shown within this volume as sticks are the drugs from all structures where rhodamine 6G is pink, ethidium is orange, dequalinium is light yellow, malachite green is green, crystal violet is violet, and berberine is dark yellow. **C)** Ribbon diagram of QacR in complex with both ethidium (red) and proflavin (yellow) simultaneously. **D)** Close-up view of (C) with the electron density for ethidium (Et) and proflavin (Pf) shown as a mesh.

From Schumacher MA, Miller MC, Grkovic S, Brown MH, Skurray RA, Brennan RG (2001) Structural mechanisms of QacR induction and multidrug recognition. *Science* **294:** 2158-2163. Reprinted with permission from AAAS.

Reprinted by permission from Macmillan Publishers Ltd: EMBO Journal (Schumacher et al, 2004), copyright (2004).

In summary, the main secondary structural elements in TetR proteins are α-helices, and their three-dimensional structures can be partitioned into an N- and C-terminal domain. The N-terminal domain contains a helix-turn-helix motif, which mediates DNA binding. Although all TetR proteins utilize this domain to bind DNA, there are multiple modes of DNA binding. TetR binds to a single palindromic DNA site as a dimer, while QacR binds cooperatively to an overlapping palindromic DNA site as a dimer of dimer. Furthermore, there are other TetR members such as EthR that binds to a 55 bp site cooperatively as an octamer (102).

There are two important functions mediated by the C-terminal domains of all TetR repressor proteins. The first is to mediate the formation of homodimers, which are the physiologically relevant forms. The second function is to allow TetR repressor proteins to bind ligands. Both QacR and TetR show remarkable ability to reform their binding pocket to accommodate diverse ligands (87). Related to these observations, whereas members of this family exhibit a high degree of sequence similarity in the N-terminal domain, the C-terminal domain is very diverse, and, generally, no significant level of sequence similarity can be observed within this domain. Given the widespread functions of TetR proteins, the diversity in the C-terminal domains may mirror the variety of effector molecules to which these domains can bind.

Even though TetR and QacR are a part of the same family of proteins and are structurally homologous, these proteins are able to implement different mechanisms of ligand interaction to invoke their respective functions. These observations reflect the versatility of TetR proteins in cellular processes. This recurring theme of functional adaptability is displayed by SlmA which directly interacts with cell division proteins to effect nucleoid occlusion. Interestingly, SlmA does not function as a repressor and its C-terminal domain does not bind a ligand. This invokes the possibility that there are other classes of TetR proteins that do not share the same features as TetR repressors, and is able to function in a completely novel way.

1.7 Usage of X-Ray Crystallography for structure determination

1.7.1 Overview of X-Ray Crystallography

X-Ray Crystallography is the major method used for the structural determination of macromolecules. Analogous to microscopy, X-ray crystallography also makes use of the scattered waves diffracted by the molecule and the "refocusing" of these reflected waves to de-convolute the "image" of the molecule (Figure 23). More specifically, the "image" is an
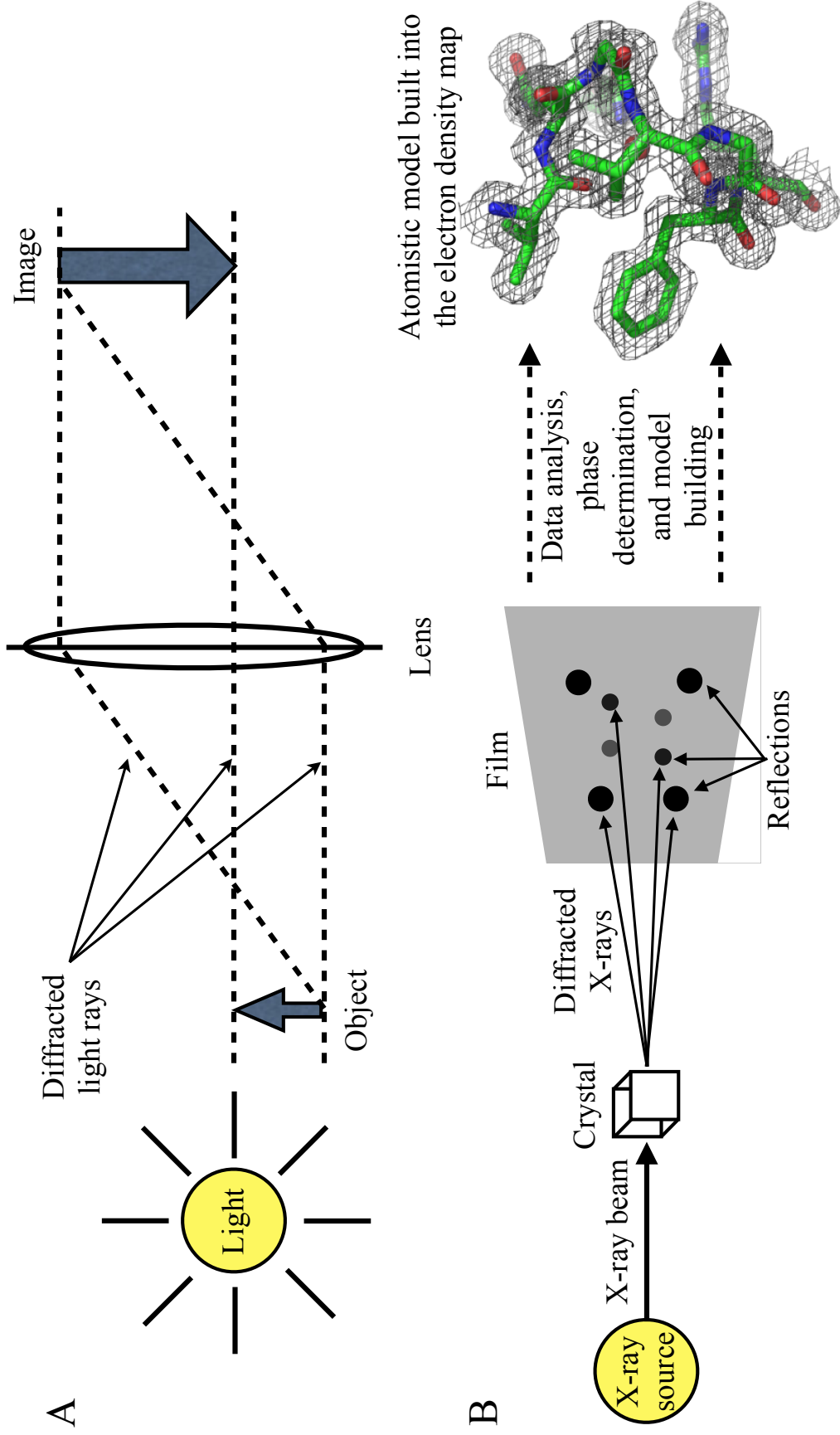
Figure 23. Overview of X-Ray Crystallography. A) An object will diffract light and these waves can be refocused into an image by a lens in our eyes or microscope. **B)** Upon exposure to direct X-rays beams, the crystal will diffract these X-rays and produce distinct "spots", or reflections, which can be captured on a film. These diffraction data can then interpreted as an electron density map that will allow the crystallographer to build a molecular model of what is in the crystal.

electron density map, which can be interpreted and allows for the development of an atomic model of the molecule. These atomistic models can reveal the inter- and intra-molecular interactions of macromolecules and provide critical insight into the molecular mechanisms of biological processes.

Practically, the steps involved in determining a structure via X-ray crystallography are as follows:

1. Grow high quality crystals from highly homogeneous samples

2. Determine the symmetry and space group of the crystal

3. Obtain X-ray intensity data of the crystals

4. Determine the phases of the diffraction data

5. Use the diffraction data and phase information to generate an electron density map

6. Build a model of the molecule into the electron density map

7. Refinement of the model

First, the molecule or molecules of interest must be crystallized. In certain conditions, macromolecules (such as proteins) can interact with one another in an orderly manner to form a crystal. The crystal can be mounted on an X-ray machine and X-rays diffracted by the crystal can be recorded at multiple angles. There are two major components to the diffracted X-rays, the amplitude and phase. The amplitude is captured and recorded as the intensity of the diffracted X-rays. However, the more critical information of phase is lost. Thus, since we cannot directly measure the phase information, we must derive the phases from indirect methods; two main methods are Molecular Replacement and Isomorphous Replacement. Once the phases are obtained, the Electron Density Equation can be solved to produce an electron density map of the molecules in the crystal. The crystallographer can then build a molecular model of the system which fits this electron density map. The following sections will go into further detail on three major aspect of this process: the nature of crystals (sections 1.7.2-1.7.3), diffraction data (sections 1.7.4-1.7.5), and the determination of phases (section 1.7.6).

1.7.2 Growing protein crystals

Crystallographers grow protein crystals by controlling the process of protein precipitation from the aqueous phase. A commonly used method for growing crystals is vapor diffusion (Figure 24). In this method, a droplet of highly homogeneous protein sample is mixed with a crystallization solution, which contains a precipitant. This droplet, along with a reservoir of crystallization solution is sealed in a compartment. Initially, the concentration of precipitant in
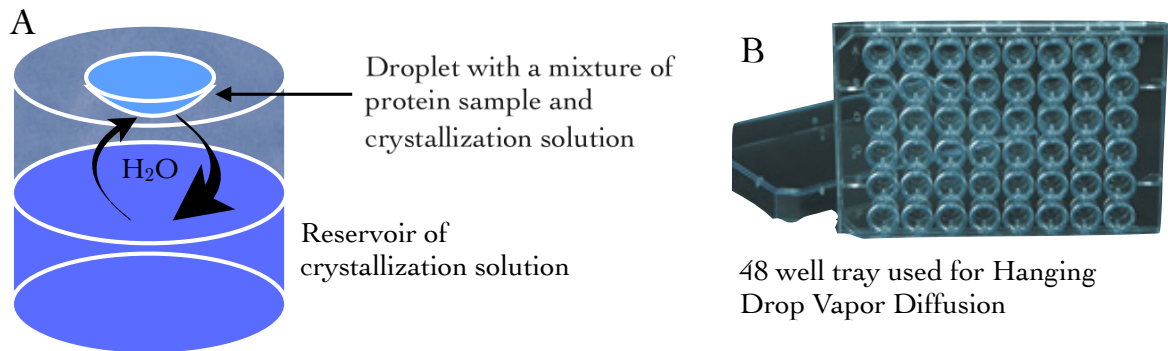
Figure 24. Hanging Drop Vapor Diffusion. A) In this method, a droplet of highly homogeneous protein sample is mixed with a crystallization solution, which contains a precipitant. This droplet, along with a reservoir of the crystallization solution is sealed in a compartment. Initially, the concentration of precipitant in the droplet is not sufficiently high to cause the protein to precipitate. However, as water diffuses from the droplet and equilibrates against the reservoir, the precipitant concentration increases. As this happens, the protein can precipitate out of solution, and either form crystals or amorphous precipitate. **B)** A 48 well tray that is commonly used to perform Hanging Drop Vapor Diffusion experiments. The experimented portrayed in (A) can be performed in one well of this tray.

the droplet is not sufficiently high to cause the protein to precipitate. However, as water diffuses from the droplet and equilibrates against the reservoir, the precipitant concentration increases. As this happens, the protein can precipitate out of solution, and either form crystals or amorphous precipitate. The ability to form crystals and not useless precipitate is dependent on multiple factors, including protein concentration, type of precipitant, pH, temperature, and ionic strength. A number of substances aid in this precipitation process and the search for a formulation that will yield high-quality crystals is a combinatorial problem that is intractable. Therefore, crystallographers generally take a heuristic approach by performing sparse matrix screens for conditions which are promising and eliminate those that are not. The formation of high quality crystals is contingent on many factors and is considered to be a non-deterministic process that is now the rate limiting step to the determination of structures via crystallography.

1.7.3 Crystal lattice and space groups

A crystal is formed by atoms arranged in a pattern that is repeated periodically in three dimensions. The pattern can be formed by a single atom, a group of atoms, a molecule, or a group of molecules. The key characteristic of a crystal is the periodicity of these patterns. This feature can be conceptualized by imagining that a very tiny creature is standing inside a crystal. If he notes all the atoms around him and walks in a straight line, he will eventually reach a point that looks identical to his starting point as he walks from one pattern to an adjacent pattern. As he continues to travel the same distance forward, he will continue to find more identical points. This concept also extends in the third dimensions, and this set of repeating and identical points constitutes a set of lattice points. If one were to connect these lattice points with straight lines, the crystal can be divided into repeating unit cells (Figure 25).

The unit cell has six variables to describe its dimensions. These variables are the angles α, β, and γ of the three independent edges **a**, **b**, and **c**. The angle α is the angle between edges **b** and **c**, β is between **a** and **c**, and γ is between **a** and **b** (Figure 26). Restrictions on these dimensions form seven crystal systems allowed for biological macromolecules. The ranking of the least to the most symmetrical crystal system is: triclinic, monoclinic, orthorhombic, tetragonal, rhombohedral, hexagonal, and cubic (Figure 27). Extending from this concept of a unit cell and symmetry, we can consider that a unit cell consists of asymmetric unit(s) that can be related to the unit cell by translation and/or rotation symmetry operators. The combination of symmetry operators that characterizes a crystal is called its space group.
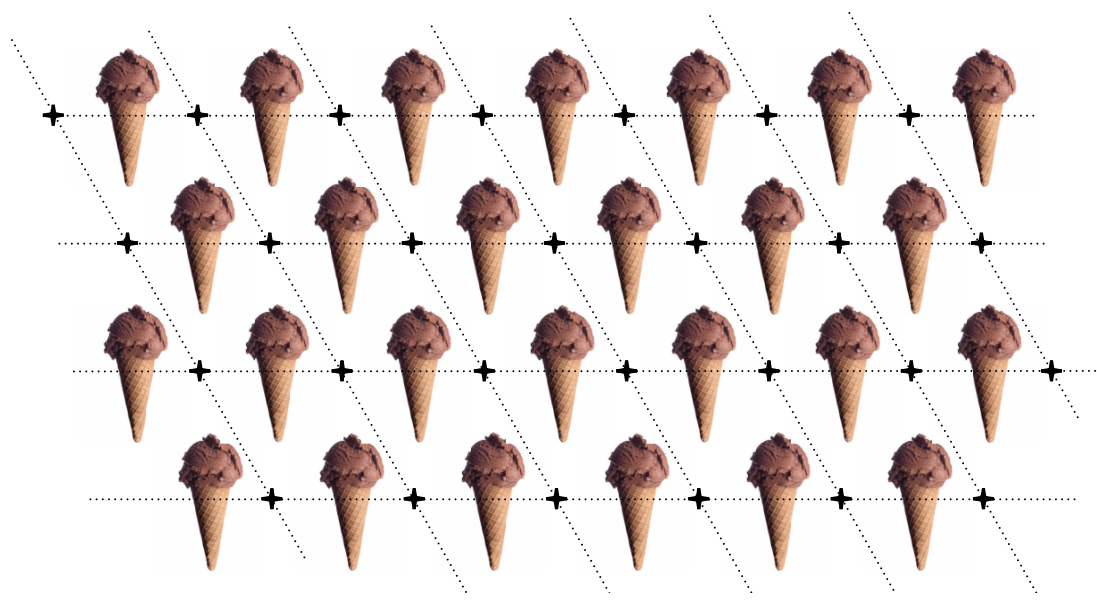
Figure 25. Two dimensional view of a crystal lattice. The crystal is made of repeating of repeating pattern. In this example, the repeating pattern is an ice-cream cone. The ✚ symbol represents lattice points, and the dotted lines demarcate the unit cells.
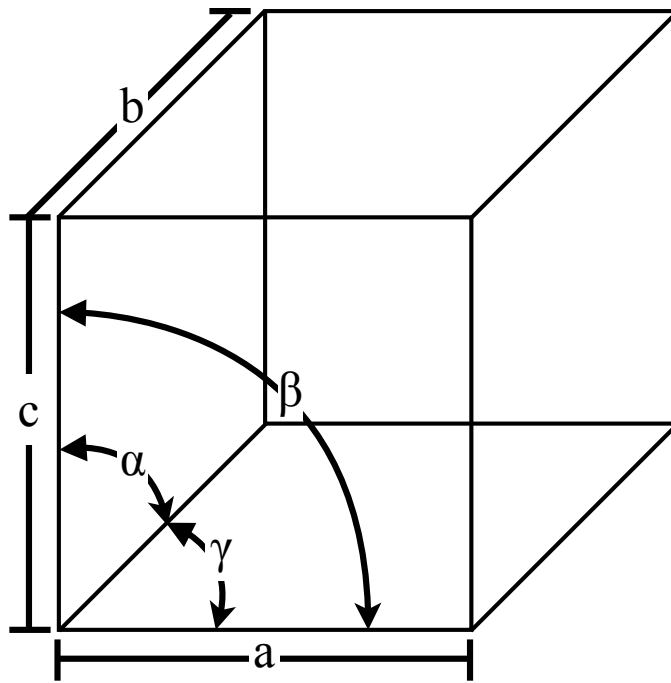
Figure 26. Dimensions of a unit cell. The unit cell has six variables to describe its dimensions. These variables are the angles α, β, and γ of the three independent edges a, b, and c. The angle α is the angle between edges b and c, β is between a and c, and γ is between a and b.
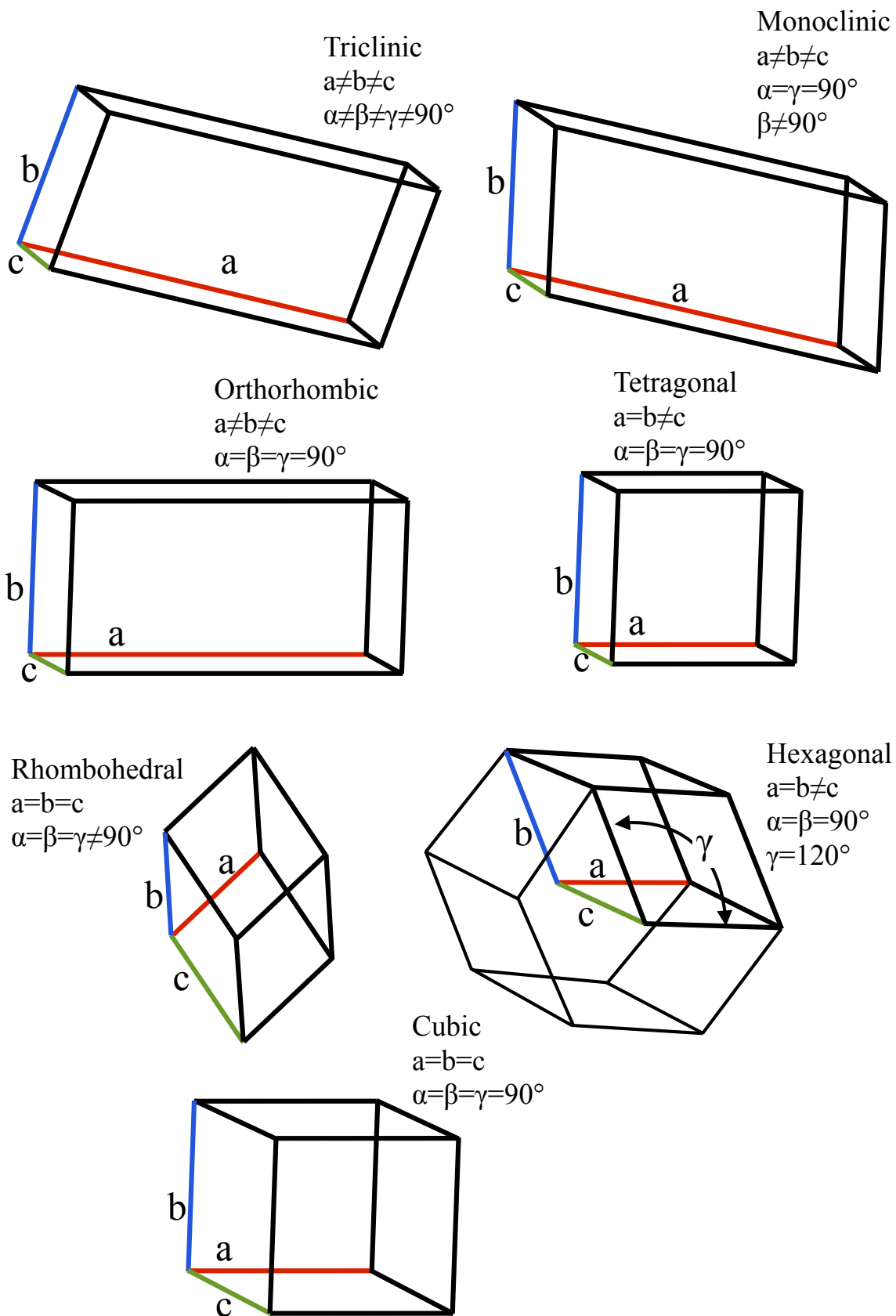
Figure 27. Descriptions of the seven crystal systems. The listing of crystal systems are in order of lowest to highest symmetry: triclinic, monoclinic, orthorhombic, tetragonal, rhombohedral, hexagonal, and cubic.
Adapted from: Rhodes G (2006) Crystallography made crystal clear: a guide for users of macromolecular models.

In summary, the relationship between a lattice and unit cell is a prominent concept in crystallography. In a crystal, the lattice represents the translational periodicity of the unit cell. Mathematically, a lattice is a discrete and discontinuous function. Thus, it is zero everywhere except at specific, periodically distributed points, where it has a value of one. Therefore, the lattice is defined by the unit cell and vice versa. The crystal can be considered as a group of molecules which make up an asymmetric unit that can form a unit cell based on a series of symmetry operators (which defines the space group). The repetition of the unit cell is captured by the lattice.

1.7.4 Diffraction data

X-ray intensity data are recorded by exposing the crystal to an X-ray beam producing diffracting waves that can be captured on a detector. There are a near infinite number of X-rays which are diffracted, yet from an example image in Figure 28, the pattern of diffraction is discreet. The pattern is solely dependent on the crystal lattice, and not the type of molecules that form the crystal. The reason for this becomes clear when one considers how Bragg's law describes X-ray diffraction.

Bragg's law explains that coherent scattering from a crystal lattice occurs in discreet and repetitive distribution. Bragg's law can be expressed as: $2d\sin\theta = n\lambda$, where d is the spacing between planes in the lattice, θ is the angle between the incident and reflected rays, λ is the wavelength of the incident wave, and n is an integer. Figure 29A illustrate a condition for constructive interference. The black dots represent two planes of lattice points (the planes are demarcated as dash-lines) separated by a distance of d. The purple rays, R1 and R2, are reflected by the lattice points at angle θ. Since the triangle, ABC, is a right triangle, sine θ is equal to $BC/AB = BC/d \therefore BC = d\sin\theta$. Thus, since R2 traveled twice the distance of R1, $2BC = 2d\sin\theta$. If the difference in path length is an integral number of the wavelength, $n\lambda$, the rays will have the same phase and interfere constructively Figure 29B. Otherwise, they will interfere destructively.

Computationally, it is more convenient to work in reciprocal space instead of real space, and in the following example, we will see how reciprocal lattice points satisfy Bragg's Law. In Figure 29C, we will consider that the crystal is at point C and is at the center of a circle with a radius of $\frac{1}{\lambda}$, representing the wavelength of the X-ray beam in reciprocal space. The X-ray beam is depicted as the arrow XO, and impinges on the crystal, point C. As the crystal diffracts
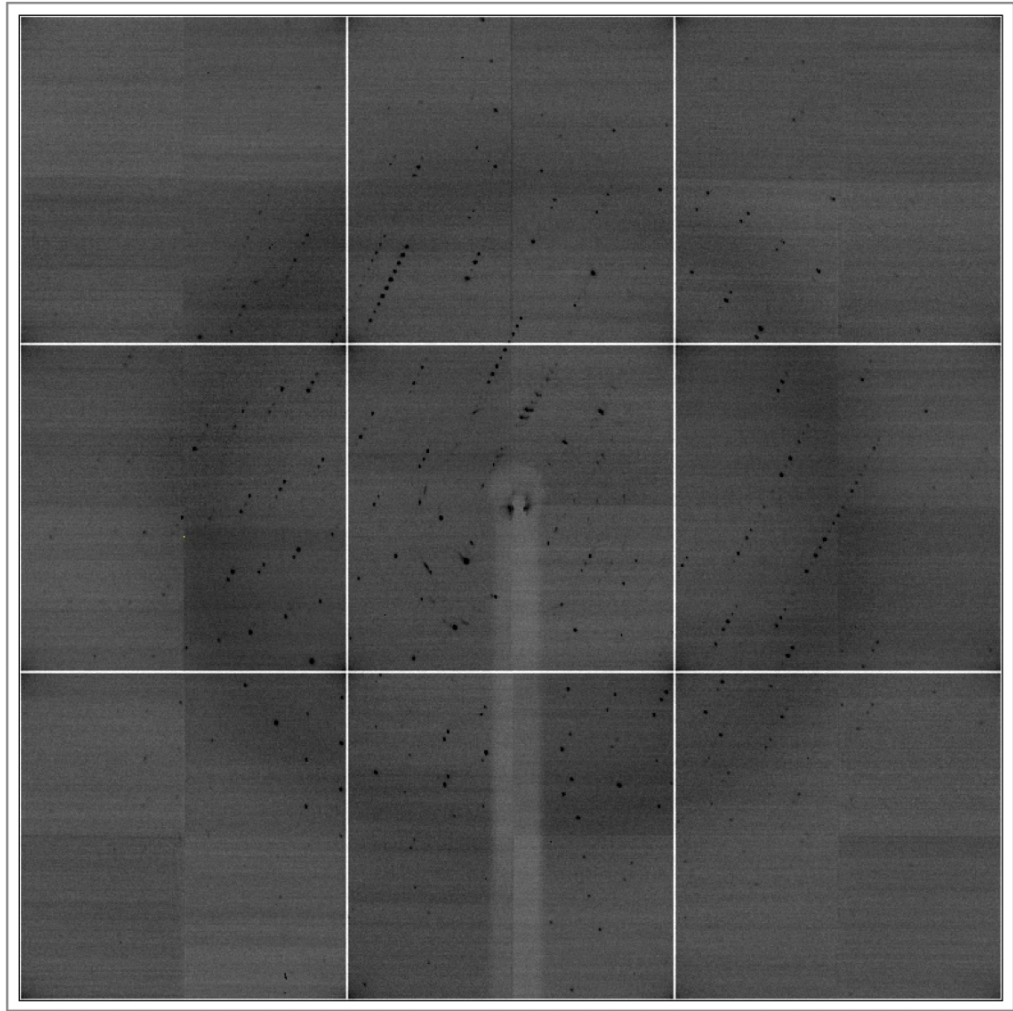
Figure 28. X-Ray diffraction. This is an X-ray diffraction image of the SlmA protein crystal, at one particular angle.
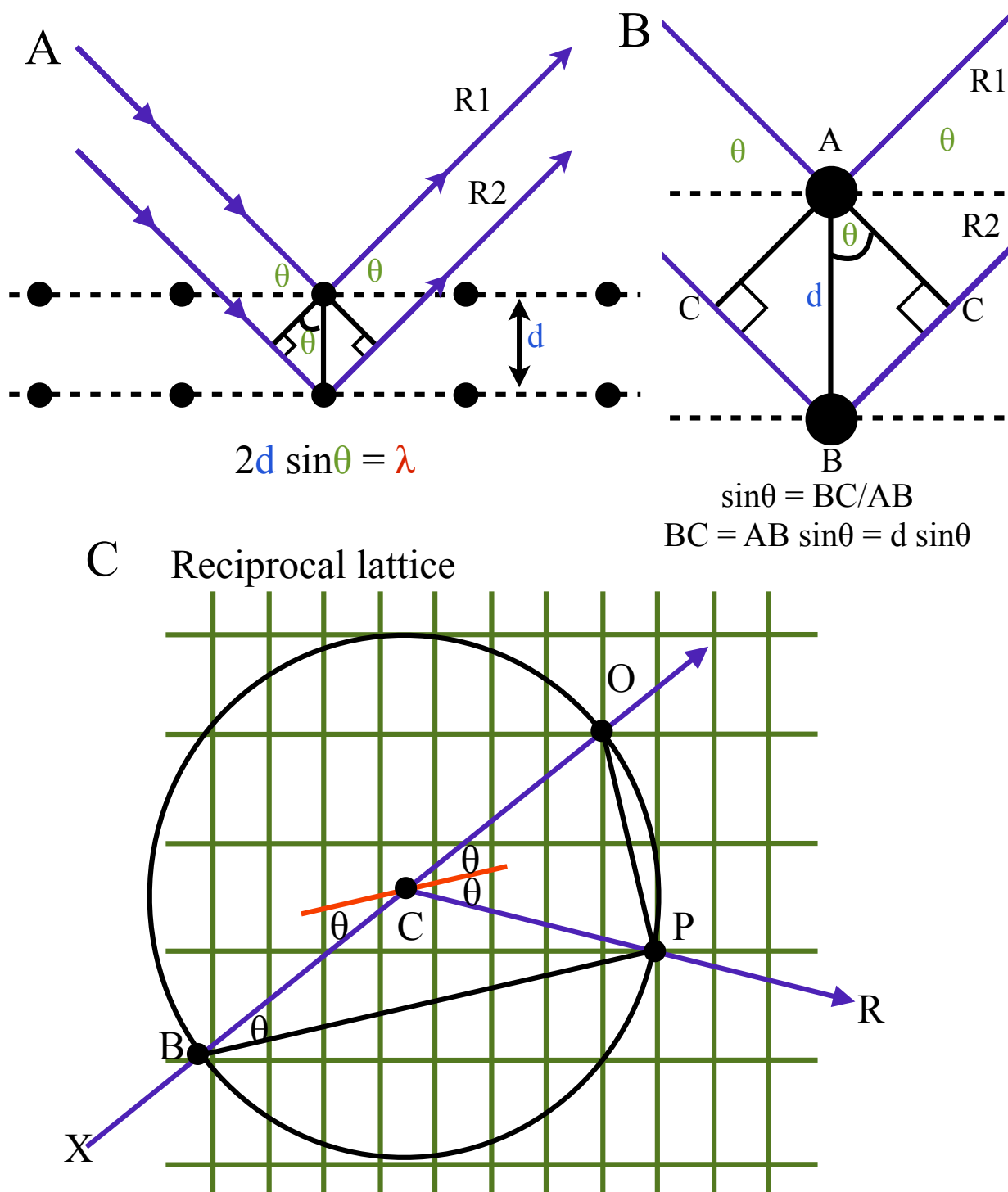
Figure 29. Bragg's Law describes the coherent scattering from a crystal lattice. A) . Bragg's law can be expressed as: $2d\sin\theta=n\lambda$, where d is the spacing between planes in the lattice, θ is the angle between the incident and reflected rays, λ is the wavelength of the incident wave, and n is an integer. **B)** The black dots represent two planes of lattice points (the planes are demarcated as dash-lines). The two planes are separated by a distance of d. The purple rays, R1 and R2, are reflected by the lattice points at angle θ. Since the triangle, ABC, is a right triangle, sine of θ is equal to BC/AB=BC/d therefore BC=$d$sinθ. Thus, since R2 traveled twice the distance, 2BC=$2d$sinθ. **C)** Diffraction in reciprocal space. Consider that the crystal is at point C and is at the center of a circle with a radius of $1/\lambda$. The X-ray beam is depicted as the arrow XO, and it impinges on point C. As the crystal diffracts X-rays, one ray, represented as the arrow from C to R, is reflected by a lattice point and is represented in reciprocal space as lattice point P. If we draw a line to connect point O to point P, we find that this triangle, PBO is a right triangle because it is inscribed in a semicircle. Only lattice points will result in these condition and satisfy Bragg's Law.

Adapted from: Rhodes G (2006) Crystallography made crystal clear: a guide for users of macromolecular models.

X-rays, one ray, represented as the arrow from C to R, is reflected by a lattice point and is represented in reciprocal space as lattice point P. If we draw a line to connect point O to point P, we find that this triangle, PBO is a right triangle because it is inscribed in a semicircle. Thus, $\sin\theta = \dfrac{OP}{OB} = \dfrac{OP}{2/\lambda}$ and this can be rearranged as $2\dfrac{1}{OP}\sin\theta = \lambda$. Since lattice planes are at a distance of d and both O and P are reciprocal lattice points, the length of OP is $\dfrac{1}{d}$. Therefore the equation can be rewritten as $2d\sin\theta = \lambda$. This can be repeated for any point in the crystal, yet only the lattice points will satisfy Bragg's law and exhibit constructive interference and produce a signal on the detector. For this reason, these "spots" are called "reflections". Therefore, the diffraction pattern on the detector is dependent on the lattice, while the intensity of each "reflection" is a contribution of all the atoms in the unit cell.

1.7.5 From diffraction data to electron density

A structure factor describes one diffracted X-ray, which produces one reflection, or spot, on the detector. A structure factor, F_{hkl}, can be treated as a Fourier sum of the individual terms that contributes to the reflection. A single term in a structure factor (F_{hkl}) is called an atomic structure factor f_{hkl}, and $f_{hkl} = f_j e^{2\pi i(hxj+kyj+lzj)}$, where f_j is the scattering factor of atom j. Each diffracted X-ray is a sum of all the scattering atoms and the structure factor for reflection F_{hkl} can be represented as $F_{hkl} = \sum\limits_{i=1}^{n} f_j e^{2\pi i(hxj+kyj+lzj)}$, where n is the number of atoms. Once more, the structure factor F_{hkl}, describes a reflection on the detector, and is a summation of the contribution by each f_{hkl}, which can be treated as a simple sphere of electron density. Each contribution, f_{hkl}, depends on 1) the amplitude (this depends on what the atom is and its scattering factor, f_j), and 2) the phase (this depends on its position in the unit cell, captured by the terms x_j, y_j, and z_j).

Moreover, the structure factor equation can be amended to account for a volume of electron density instead. Mathematically, the electron density of a volume centered at (x, y, z) is $\rho(x,y,z)$. We can make this term more and more accurate as was make the volume infinitely small. The resulting integral is as follow: $F_{hkl} = \iiint\limits_{xyz} \rho(x,y,z)e^{2\pi i(hx+ky+lz)}dxdydz$. The relationship that can be derived is that the electron density is the Fourier transformation of the structure factors. Therefore the electron density can be computed from structure factors as follow:

$$\rho(x,y,z) = \tfrac{1}{V} \sum\sum\sum F_{hkl} e^{-2\pi i(hx+ky+lz)} .$$

1.7.6 Phase determination

F_{hkl} is a periodic function and thus possesses amplitude, frequency, and phase. The amplitude is directly proportional to the square root of the reflection intensity captured on the detector. The three frequencies of this wave are h, k, and l. This means that the frequency of a structure factor is $\dfrac{1}{d_{hkl}}$ and the wavelength is the spacing between planes. Unlike the other two components, the phase is not directly measurable, and in order to compute the electron density map, $\rho(x,y,z)$, the phase of each diffracted X-ray must be determined. This is known as the "Phase Problem".

There two major methods used to obtain the phases of all the reflections, Molecular Replacement and Isomorphous Replacement. The Molecular Replacement method utilizes phases from structure factors of a known molecule as initial estimates. This method is useful when the model molecule is closely related structurally or the model molecule is a smaller portion of what is in the asymmetric unit. The second method, Isomorphous Replacement, is much more complicated experimentally and technically and will be discussed further below.

Recall that the contribution of an atom's intensity is greatest when the vertex of the crystal lattice intersects with the atom. Therefore some atoms will contribute strongly, weakly, or not at all to the intensity of the reflection. If we were to add a small number of atoms to identical positions in all the unit cells, we should see a change in the intensity of the reflections. In practice, heavy-atoms compounds can be soaked into native crystals as a mean of generating a crystal that is isomorphous to its un-soaked, or native, counterpart. In Figure 30 the native protein crystal is represented as $[P]$, and its heavy-atom derivative is represented as $[PH]$. Since the two crystals are isomorphous, the differences between the diffraction patterns will be due to the heavy-atoms, $[PH]-[P]=[H]$. This difference data allows the crystallographer to determine the position of the heavy-atom in the unit cell, utlizing a method called the Patterson function. With the location of the heavy-atoms, the crystallographer can then calculate the structure factor for the heavy-atoms, F_H. The Isomorphous Replacement method provides the structure factors of a heavy-atom derivative protein, F_{PH}, which contains contributions from the structure factor of the native protein, F_P, and the heavy-atom, F_H. So $F_{PH} = F_P + F_H \therefore F_P = F_{PH} - F_H$, which allows for the calculation of the structure factor for the native protein. The usage of the Patterson function to determine F_H and how a special case of

Protein [P]

Protein + Heavy Atom [PH]

[PH]-[P]=[H]

Interpret and locate H

with the Patterson function
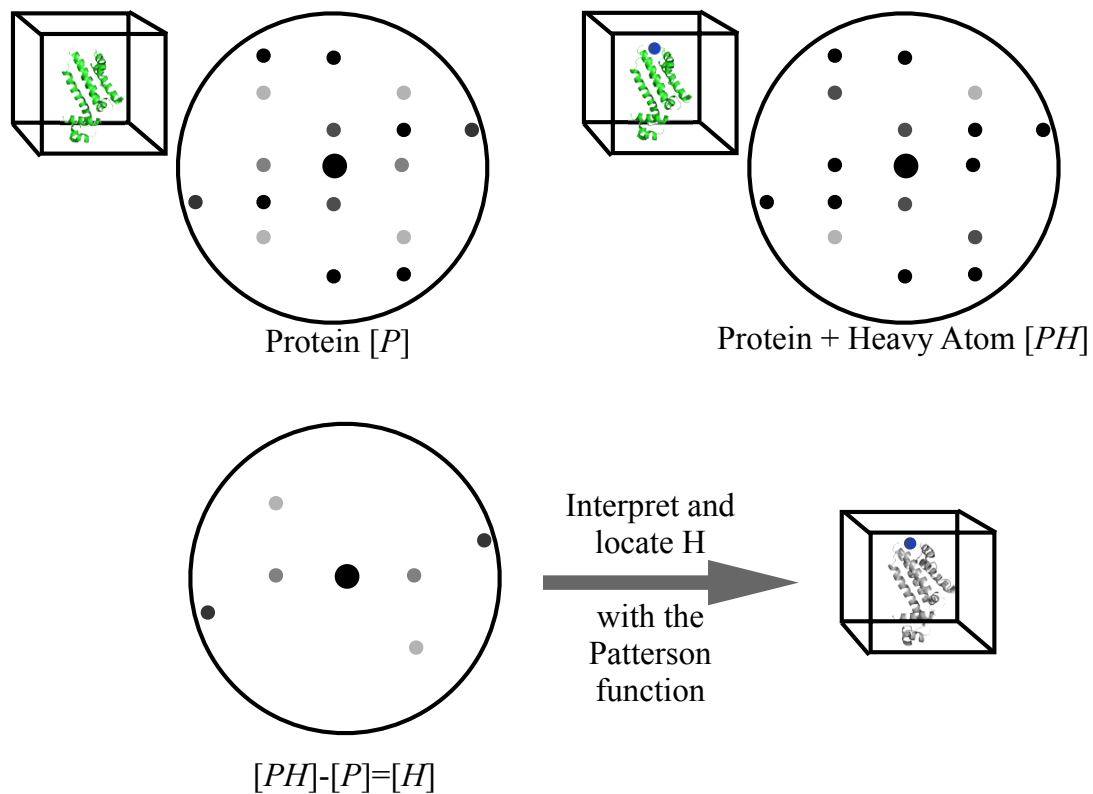
Figure 30. Isomorphous Replacement with heavy atom(s). The native protein crystal is represented as [P], and its heavy-atom derivative is represented as [PH]. Since the two crystals are isomorphous, the differences between the diffraction patterns will be due to the heavy-atoms, [PH]-[P]=[H]. This difference data allows the crystallographer to determine the position of the heavy-atom in the unit cell, with a method called the Patterson function. Then with the location of the heavy-atoms, the crystallographer can calculate the structure factor for the heavy-atoms and provides an initial estimate of phases for the protein.
Adapted from: Rhodes G (2006) Crystallography made crystal clear: a guide for users of macromolecular models.

Isomorphous Replacement call Multi-wavelength Anomalous Dispersion (MAD) can be used to determine F_{PH} will be discussed in sections 1.7.7 and 1.7.8, respectively.

1.7.7 The Patterson function

The Patterson function, $P(u,v,w)$, is a Fourier sum that is extremely useful in the determination of heavy-atom(s) coordinates. The coordinate system of (u,v,w) is used in Patterson space and is directly correlated to the (x,y,z) system of real space. The Patterson function does not contain phases and thus the amplitude of each term is the square of one structure factor, which is proportional to the intensity of the measured reflection. The Patterson function in general form can be written as:

$$P(u,v,w) = \frac{1}{V}\sum\sum\sum |Fhkl|^2 e^{-2\pi i(hu + kv + lw)}.$$

In order to obtain a Patterson function for only the heavy-atom(s) derivative crystal, a difference Patterson function must be constructed, where: $(\Delta F)^2 = (|F_{PH}|^2 - |F_P|^2)$. The difference Patterson function can then be written as:

$$\Delta P(u,v,w) = \frac{1}{V}\sum\sum\sum \Delta F_{hkl}^2 e^{-2\pi i(hu + kv + lw)}.$$

In contrast to the electron density contour map, $\rho(x,y,z)$, which shows peaks at the positions of atoms, the Patterson contour map, $P(u,v,w)$, shows peaks at location of vectors between atoms. Since there are more vectors between atoms than there are atoms, the Patterson map is more complicated than the electron density map. This combinatorial level of complexity limits the ability of the Patterson function to be applied to complicated systems. However, in simple systems with relatively few atoms, and hence vectors between atoms, the Patterson function is an extremely powerful tool.

Figure 31 is an example of a two-dimensional construction of a Patterson map. This example has three atoms represented as red dots (Figure 31A), and although only two vectors are drawn: $1 \rightarrow 3$ and $3 \rightarrow 2$, all six vectors should be considered. In Figure 31B, an origin is chosen (for simplicity it is usually $(0,0,0)$) and all vectors are redrawn with their tails at the origin. Next, in each unit cell, the Patterson atoms are reproduced to generate a complete Patterson map (the original structure is represented as red dots) (Figure 31C). Since there are three atoms in this structure, only the origin and two additional peaks represent the solution. To determine which peaks correspond to the structure, a trial and error approach is applied. A set
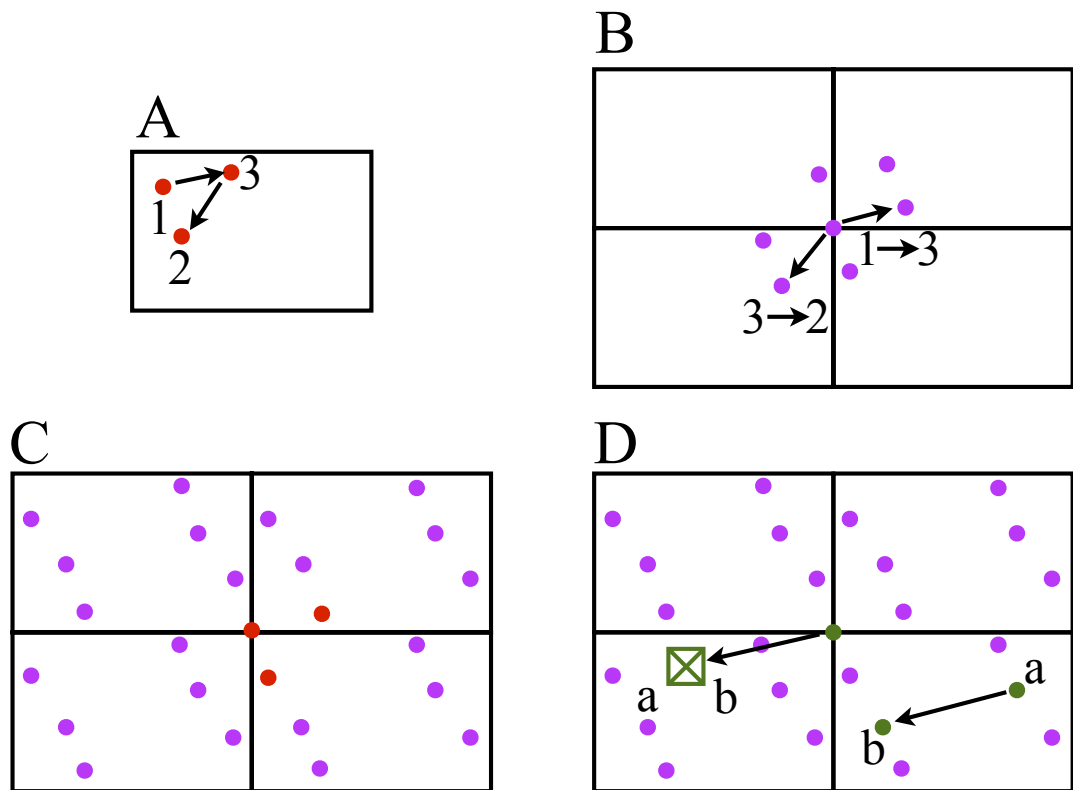
Figure 31. Example of a two-dimensional construction of a Patterson map. A) This example has three atoms represented as red dots, and although, only two vectors are drawn: 1→3 and 2→3, all six vectors should be considered. **B)** An origin is chosen (0,0,0) and all vectors are redrawn with their tails at the origin. **C)** In each unit cell, the Patterson atoms are reproduced to generate a complete Patterson map (the original structure is represented as red dots). **D)** An example of an incorrect solution. A set of peaks was chosen and when a Patterson map was constructed, the two maps did not match. The spot of incongruence is illustrated by a green box with an 'x'.

Adapted from: Rhodes G (2006) Crystallography made crystal clear: a guide for users of macromolecular models.

of peaks is chosen and a Patterson map is generated with these peaks. If the new map matches the original Patterson map, a solution is found, otherwise the search continues. Figure 31D illustrates an example of an incorrect solution. A set of peaks was chosen and when a Patterson map was constructed, the two maps did not match. The spot of incongruence is illustrated by a green box with an 'x'. The search will continue with the selection of another set of peaks until a solution is found.

The magnitude of scattering contributions by an element is roughly independent of the reflection angle and their scattering has been computed and can be found in the International Tables for Crystallography. The phase information is dependent on the location of the heavy-atom(s) in the unit cell, and thus once the position of the atom(s) are determined, via the Patterson function, the structure factor F_H can be solved. Next, the use of MAD to determine F_{PH} will be discussed.

1.7.8 Multi-wavelength Anomalous Diffraction

One particular method that is a subset of Isomorphous Replacement is called Multi-wavelength Anomalous Dispersion (MAD). MAD has become the predominant method of *de novo* phase determination over the past two decades. In order to discuss MAD as a method for phase determination, the absorption and emission of X-ray by atoms must be considered. An element absorbs X-rays and re-emits them. The ability of an element to absorb can drop sharply at a particular wavelength depending on the element. This sharp drop in absorption is called the absorption edge. This absorption edge is of interest because when the X-ray wavelength is on or near the absorption edge, a fraction of the radiation is absorbed by the atom and is re-emitted with altered phase causing an anomalous dispersion. In order to appreciate the usefulness of anomalous dispersion or scattering, we must consider Friedel's law. All reciprocal lattices possess a symmetry element called the point of inversion with respect to the origin. Inversion center reflections are annotated as hkl and $\overline{hkl}$. Friedel's law states that the intensity of reflection hkl is equivalent to $\overline{hkl}$, or stated as $I_{hkl} = I_{\overline{hkl}}$. However, when the X-ray wavelength is near the absorption edge, the element will exhibit anomalous dispersion and Friedel's law no longer holds. Therefore $I_{hkl} \neq I_{\overline{hkl}}$.

Heavy-atoms such as mercury, platinum, and selenium all exhibit anomalous dispersion in a wavelength range used in crystallography. However, the absorption edges for light-atoms such as carbon, nitrogen, and oxygen are not near the wavelengths used for crystallography.

These characteristics make these heavy-atoms very useful for MAD. MAD makes use of crystals with heavy-atoms, which can be achieved by making use of the existing heavy-atoms native to the molecule, soaking the crystal, or by substituting a heavy-atom derivative (such as selenomethionine for methionine) during protein expression.

The example in Figure 32A shows the structure factor of a heavy-atom derivative at wavelength 1, $F_{PH}^{\lambda 1}$, does not scatter anomalously. However, the same structure factor at a second wavelength near the absorption edge, $F_{PH}^{\lambda 2}$, does scatter anomalously. The differences in $F_{PH}^{\lambda 1}$ and $F_{PH}^{\lambda 2}$ has two contributors, the real and imaginary components, ΔF_r and ΔF_i. Therefore $F_{PH}^{\lambda 2} = F_{PH}^{\lambda 1} + \Delta F_r + \Delta F_i$. Figure 32B illustrates two Friedel pairs of a structure factor that are far and near the absorption edge: $F_{PH}^{\lambda 1+}$, $F_{PH}^{\lambda 1-}$, and $F_{PH}^{\lambda 2+}$, $F_{PH}^{\lambda 2-}$. Critically, it is within the disparity of $\left| F_{PH}^{\lambda 2+} \right| \neq \left| F_{PH}^{\lambda 2-} \right|$ that the phase can be extracted.

The anomalous scattering contributions ΔF_r and ΔF_i can be determine by determining the position of the heavy-atoms via the Patterson function. This information coupled with the disparity between intensities of Friedel pairs in the anomalous scattering data, $\left| F_{PH}^{\lambda 2+} \right| \neq \left| F_{PH}^{\lambda 2-} \right|$, will provide the necessary information for the determination of the structure factor of the data set that does not exhibit anomalous scattering, $F_{PH}^{\lambda 1}$. First, consider the following: $F_{PH}^{\lambda 1+} = F_{PH}^{\lambda 2+} - \Delta F_r^+ - \Delta F_i^+$. Figure 32B illustrates this equation as a vector diagram, with the vector solution to this equation is shown in Figure 33A as a Harker diagram. To solve this equation, first add the two vectors, $-\Delta F_r^+$ (purple) and $-\Delta F_i^+$ (cyan), and place them at the origin. Next, draw a circle (red) with radius of $\left| F_{PH}^{\lambda 2+} \right|$ (dotted red) with the head of $\sum -\Delta F_r^+ + \Delta F_i^+$ as the center. The circle represents the known amplitude (radius) but unknown phase information for $F_{PH}^{\lambda 2+}$. Then another circle (green) centered at the origin is drawn with $\left| F_{PH}^{\lambda 1+} \right|$ as the radius (dotted green). This circle represents the non-anomalous scattering data set collected at a wavelength further away from the absorption edge of the heavy-atom(s). The intersections, F_a and F_b, are two possible phases of the reflection $F_{PH}^{\lambda 1+}$. In order to distinguish between the two possible solutions, the Friedel partner comes in to play. A second vector equation can be written for the Friedel partner as: $F_{PH}^{\lambda 1+} = F_{PH}^{\lambda 2-} - \Delta F_r^+ - (-\Delta F_i^+)$. This equation can be solved in the same manner as the previous one and the solutions are shown in Figure 33B as a Harker diagram. While, the two solutions F_c and F_d are close to one another, F_c is closer to F_a and neither are close to F_b. Thus the disparity in intensities between Friedel
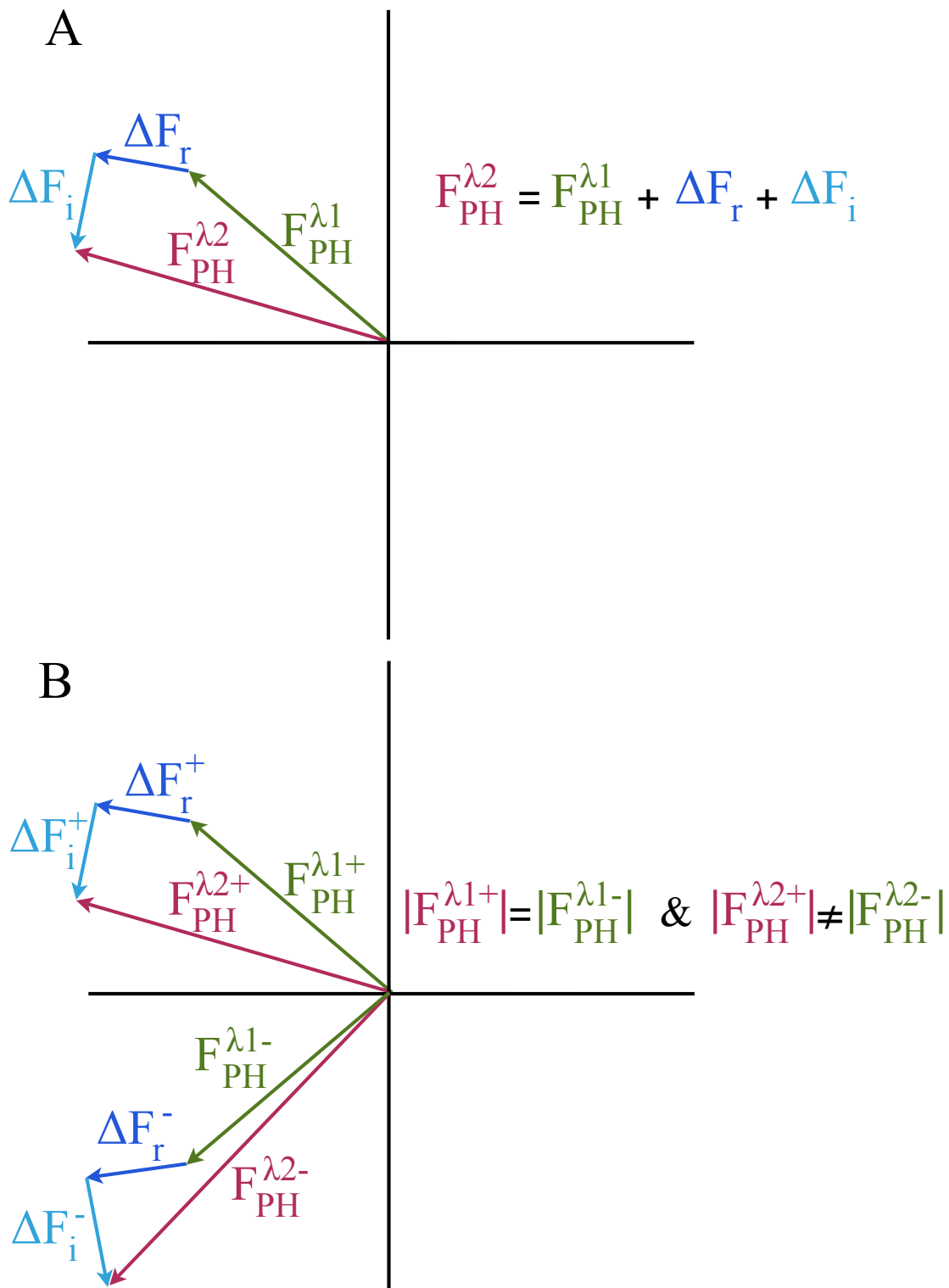
Figure 32. Vector representation of structure factors under anomalous and non-anomalous scattering conditions. A) The structure factor of a heavy-atom derivative at wavelength 1, $F_{PH}^{\lambda 1}$, does not scatter anomalously, and the same structure factor at a second wavelength near the absorption edge, $F_{PH}^{\lambda 2}$, where it does scatter anomalously. The two contributor to the differences in $F_{PH}^{\lambda 1}$ and $F_{PH}^{\lambda 2}$ are the real and imaginary components, ΔF_r and ΔF_i. **B)** Illustration of the anomalous scattering of a Friedel pair of structure factors, denoted as $F_{PH}^{\lambda 2+}$ and $F_{PH}^{\lambda 2-}$. The vectors and their components are drawn similarly to (A).

Adapted from: Rhodes G (2006) Crystallography made crystal clear: a guide for users of macromolecular models.
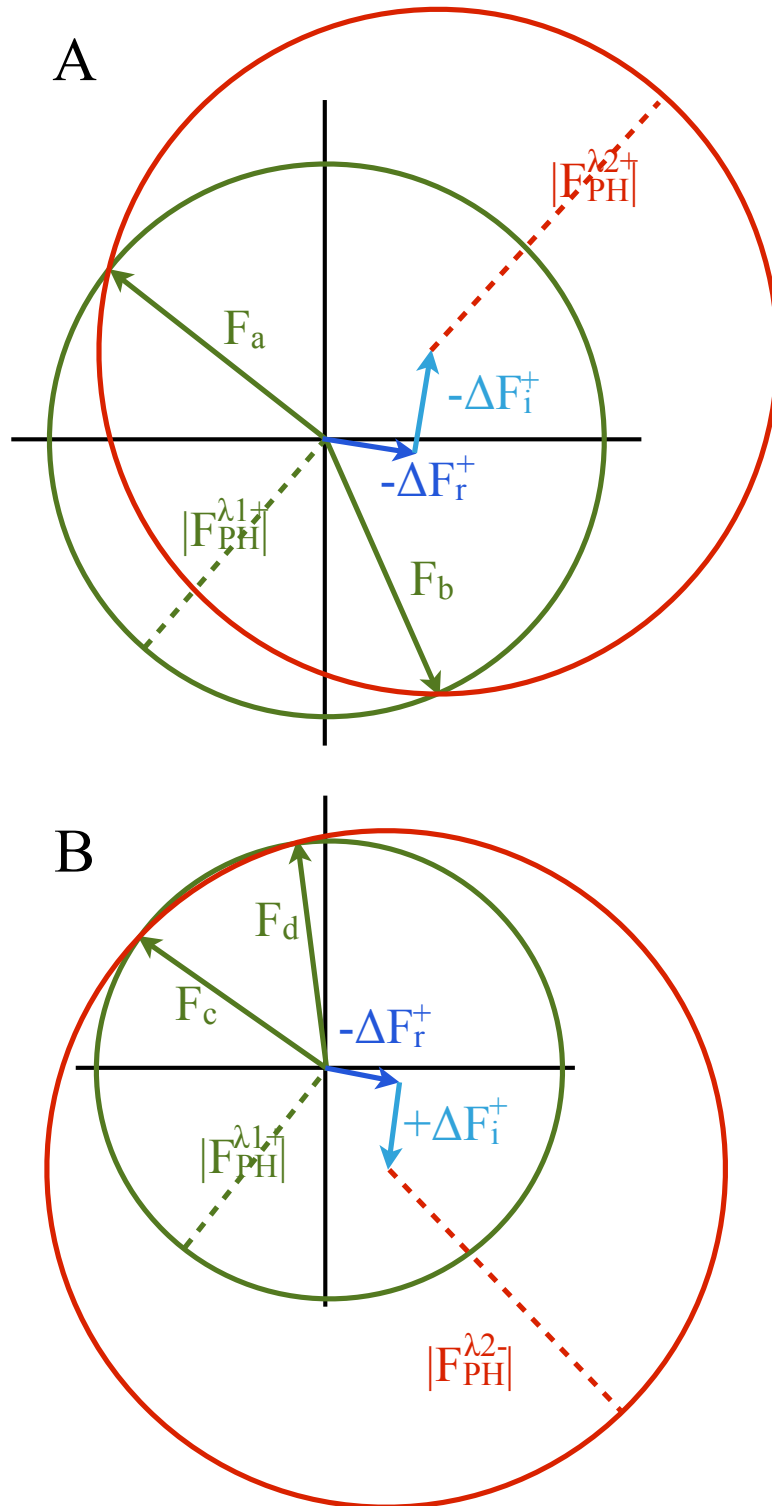
Figure 33. Usage of Friedel pairs under anomalous scattering condition to determine phase. A) The Harker diagram for the following equation: $F_{PH}^{\lambda1+} = |F_{PH}^{\lambda2+}| - \Delta F_r^+ - \Delta F_i^+$. **B)** The Harker diagram for the following equation: $F_{PH}^{\lambda1+} = |F_{PH}^{\lambda2-}| - \Delta F_r^+ - (-\Delta F_i^+)$. This equation uses the Friedel mate of the equation in (A). The vectors are drawn similarly to Figure 28.

Adapted from: Rhodes G (2006) Crystallography made crystal clear: a guide for users of macromolecular models.

mates, under anomalous scattering condition, provides the necessary constraints to approximate the phase for each reflection. In summation, the phase for each reflection is estimated via this methodology to determine all the structure factors for the heavy-atom derivative structure under non-anomalous scattering conditions. Thus, the amplitudes and phases for F_{PH} and F_H are known. With this information, the phase for the native data can be calculated via the equation $F_P = F_{PH} - F_H$.

1.8 Theory and usage of Small Angle X-ray Scattering for structure analysis

X-ray crystallography can provide deep structural understanding of a biological system, but there are great limitations to achieving these data. In particular, for macromolecular systems that are highly flexible or intrinsically disordered, the crystallization process can be very difficult and limits the acquisition of data. Therefore, there is a need for a technique which can provide structural insights on macromolecular assemblies in solution. Facilitating this need is small angle X-ray scattering (SAXS). SAXS is able to complement other experimental approaches by providing information regarding the radius of gyration (R_G), shape, oligomerization states, and allow for the discrimination between mono-disperse and aggregated samples.

In a SAXS experiment, an X-ray beam is applied to a sample in solution and the intensities of the scattered waves are collected on a detector (Figure 34). In contrast to protein crystallography where the ordered macromolecules produces diffraction intensities, the SAXS profile, *I(q)*, represents the simultaneous scattering of molecules in all orientations which inherently reduces the resolution.

Before we discuss how the SAXS profile can be interpreted, we will first consider how the SAXS profile is obtained from a SAXS experiment. From Figure 34, we can see that the scattering intensity, $q = \left|\vec{q}\right| = \left|\vec{k_s} - \vec{k_i}\right|$, is collected as a function of the scattering angle, θ, so $q = \dfrac{4\pi \sin\theta}{\lambda}$, where $\vec{k_i}$ is the incident wave, $\vec{k_s}$ is the scattered wave, θ is the scattering angle, and λ is the wavelength. The SAXS profile, *I(q)*, is directly proportional to the product of the amplitude of scattering and its complex conjugate, *A(q)* and *A(q)**, respectively. This relationship can be written as $I(q) \propto A(q) \bullet A(q)^*$. The amplitude of scattering is further related to the electron density distribution of a specific volume, and the relationship can be expressed as $A(q) = \int \Delta\rho(r)e^{iqr}dr$, where *A(q)* is the amplitude of scattering, $\Delta\rho(r)$ is the electron density

$$q = |\vec{q}| = |\vec{k_s} - \vec{k_i}|$$

$$q = \frac{4\pi\ sin\theta}{\lambda}$$
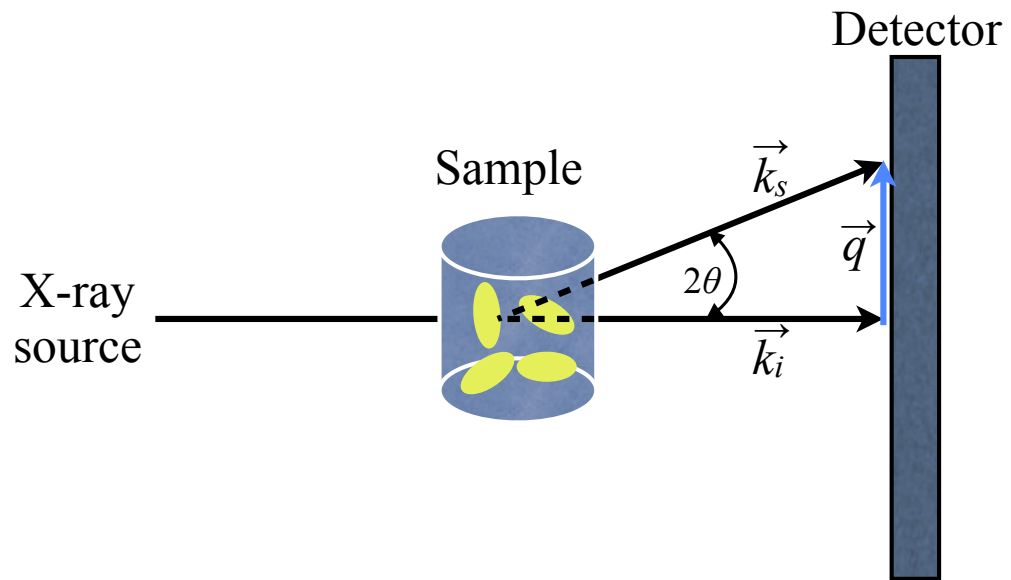
Figure 34. The components of a small angle scattering experiment. The X-ray source is used to produce an X-ray beam which is scattered by the sample and the scattering intensities are captured on a detector. The equations shows the relationship between a scattered wave, $\vec{ks}$, and the incident wave, $\vec{ki}$, and how the scattering intensity, q, can be calculated as dependent of the scattering angle, θ.

distribution, and q is the scattering intensity. These relationships describe how the intensity data that are collected are related to the electron density of the molecule.

The SAXS profile can be analyzed and transformed to perform three different types of analyses; the Guinier analysis that can provide the R_G, the Kratky analysis can qualitatively assess the "folded-ness", and the *P(r)* distribution that can be used to generate SAXS envelopes of the molecule (103). First, the Guinier analysis makes use of the observation (by Andre Guinier in 1939) that in the low-resolution range, the Debye factor can be reduced to $\ln[I(q)] = \ln[I(0)] - \dfrac{q^2 R_G^3}{3}$, where q is the scattering intensity, *I(q)* is the SAXS profile, *I(0)* is the extrapolated intensity at zero scattering angle, and R_G is the radius of gyration. The only unknown in this equation is R_G, which can be calculated directly from the data. Additionally, the plot of *ln[q(I)]* vs. $q^2(Å^{-2})$ (roughly, the plot of the intensity versus resolution) can indicate whether the sample is homogeneous or is likely to be aggregated. At the low resolution, a curve which is non-linear indicates the presence of aggregation (Figure 35B), while a homogeneous sample will display a linear curve in the qR_G limit of ≤ 1.3 (Figure 35A,C) (103).

The Kratky analysis allows for the qualitative assessment of the "folded-ness" of the sample. This analysis is based on Porod's law which states that when two media are separated by a sharp interface, the scattering intensity will display an asymptote in the higher-resolution range. A Kratky plot can be generated by plotting the weighted intensity versus the resolution, $q^2 I(q)$ vs. $q(Å^{-1})$. Since less structured molecules will not have a sharp interface between protein and solvent, and well-folded proteins do, the plot can suggest the level of "folded-ness" of a protein (Figure 36) (103).

A third type of analysis of SAXS data utilizes the pair-distance distribution function, *P(r)*. The *P(r)* describes all the inter-atomic vectors within the molecule. The function can be written as: $P(r) = \dfrac{r}{2\pi^2} \displaystyle\int_0^\infty I(q)q\sin(qr)dq$, where *I(q)* is the SAXS profile, q is the scattering intensity, and r is the inter-atomic vector. Although it is possible to calculate a theoretical distribution of pair-distance vectors from an atomic model, the inverse is not true. Therefore the atomic details of the molecule cannot be determined from the SAXS profile alone. The *P(r)* can be very useful, and one such usage is in observing structural shifts within a molecule. Since all paired-distances are included in the function, relatively small structure shifts are noticeable in the *P(r)*.
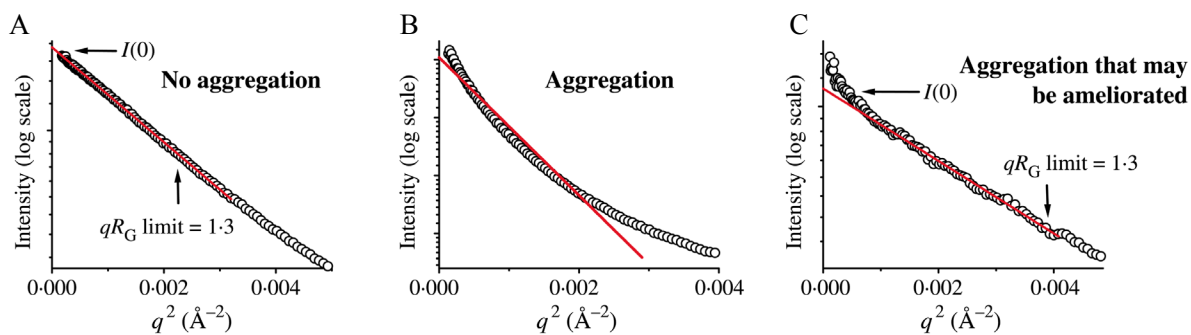
Figure 35. Calculated R_G and $I(0)$ from the Guinier plot. A) a linear dependence of $ln[I(q)]$ *vs.* q^2 suggests that the sample is homogeneous. **B)** A nonlinear dependence of $ln[I(q)]$ *vs.* q^2 indicates the presence of aggregation. Scattering from aggregated samples strongly influences the entire data set and no further processing should be performed. **C)** Aggregation in the sample can be reduced or eliminated by varying buffer conditions, centrifugation, and filtration.

Reprinted by permission from Cambridge University Press: Quarterly Reviews of Biophysics (Putnam *et al*, 2007), copyright (2007).
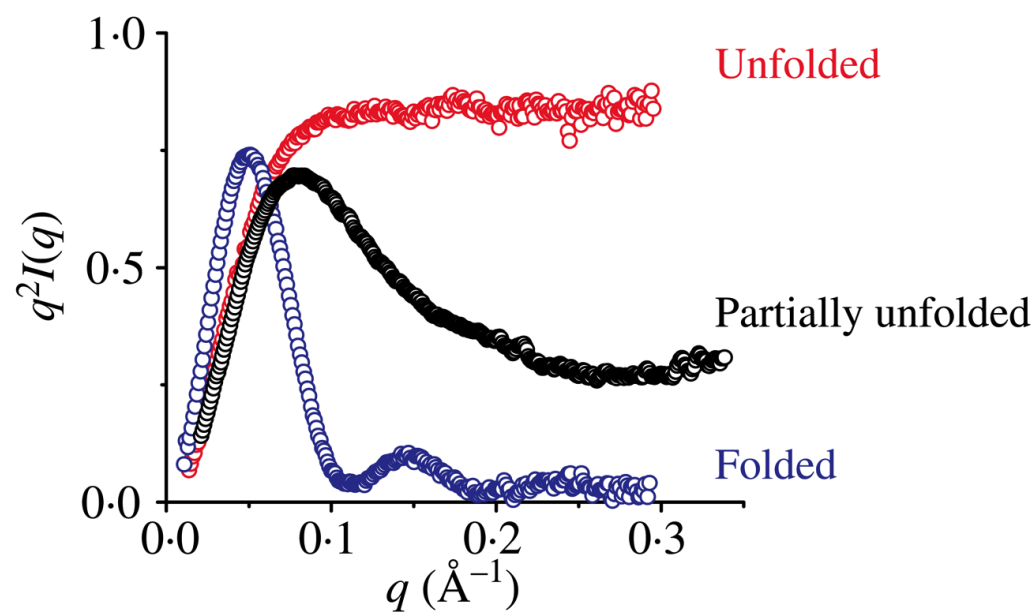
Figure 36. Kratky plot approximate the "folded-ness" of the molecule. This analysis is based on Porod's law which states that when two media are separated by a sharp interface, the scattering intensity will display an asymptote in the higher-resolution range.

Reprinted by permission from Cambridge University Press: Quarterly Reviews of Biophysics (Putnam et al, 2007), copyright (2007).

Figure 37A shows an example of a structural shift in the T7 Promoter complex and the resultant *P(r)*. The difference between the two *P(r)* distributions is evident. Furthermore, several computational algorithms, such as DAMMIN, GASBOR, and DALAI_GA have been developed to make use of the *I(q)* and *P(r)* to generate an envelope of the molecule. Figure 37B shows the results from six independent GASBOR runs. Figure 37C shows that the atomic model of the protein fits well within the consensus envelope (103, 104). A SAXS envelope can provide information regarding conformational changes and the spatial organization of proteins in a complex. Altogether, SAXS analysis can provide great insight into macromolecular complexes and is a good complement to other biochemical and structural analyses.
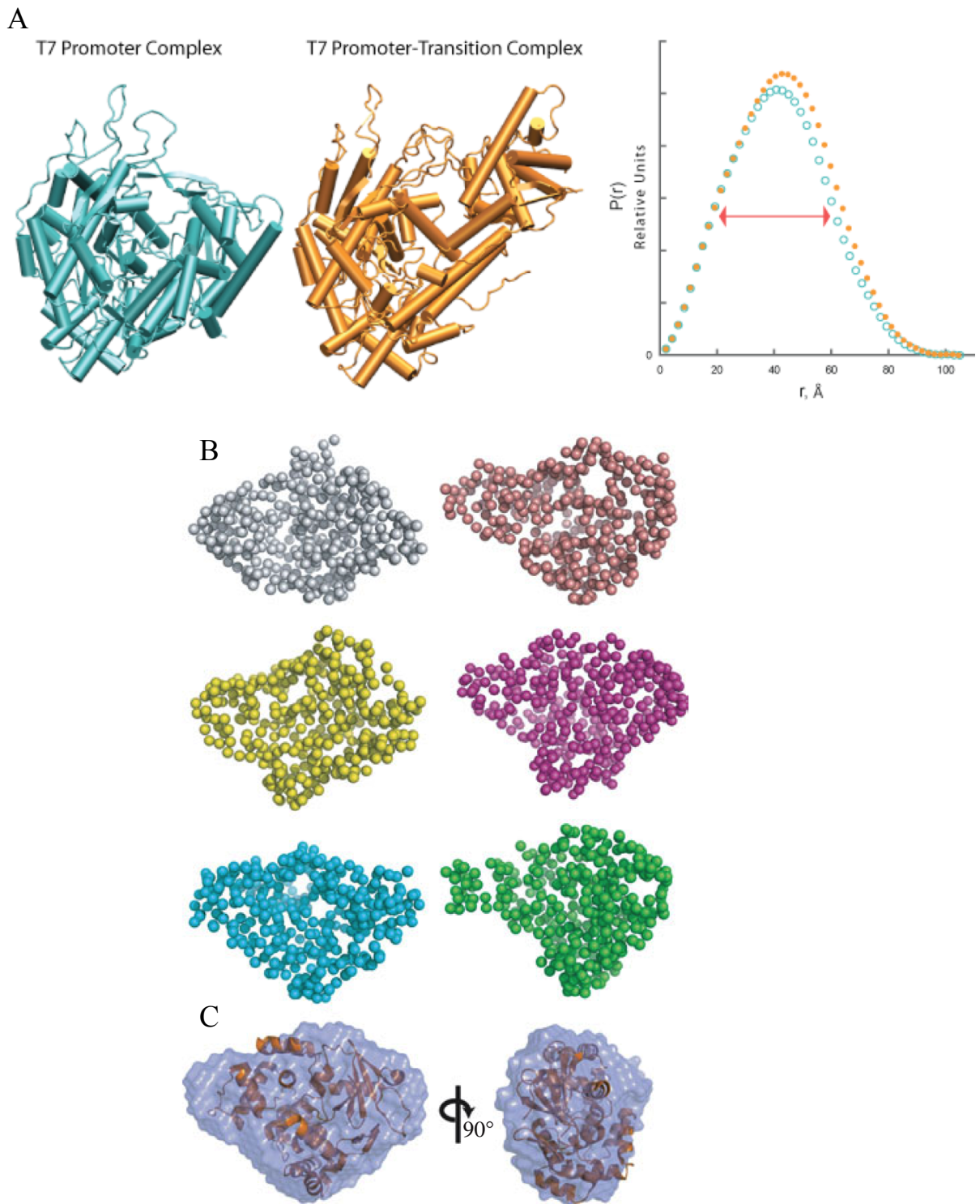
Figure 37. Pair-distance distribution function describes the shape of the molecule. A) The two structural states of the T7 Promoter complex can be observed via the P(r) distribution. **B)** Six independent runs of GASBOR shows similar results. **C)** The crystal structure of OGG1 fits well within the averaged *ab-initio* envelope from (B).

Reprinted by permission from Cambridge University Press: Quarterly Reviews of Biophysics (Putnam *et al*, 2007), copyright (2007).

Chapter 2. Experimental Procedures

2.1 Fluorescence Anisotropy/Polarization

Fluorescence polarization (FP) is a technique that can be used to obtain binding constants for macromolecular interactions. A powerful use of the method is the measurement of protein binding to a fluorescein-tagged oligonucleotide. In such measurements, the rotational motion of the fluorescein-tagged oligonucleotide is slowed by protein binding, increasing the fluorescence emission anisotropy value for the tagged DNA. Proteins may also be fluorescently tagged but DNA is easier to label in a manner that does not interfere with binding. Additionally, because of the rod-like geometry of DNA, protein binding to an oligonucleotide generally has a greater effect on rotational motion, leading to larger changes in fluorescence anisotropy (Figure 38). FP assays were performed with a PanVera Beacon 2000 fluorescence polarization system. Samples were excited at 490 nm, and fluorescence emission was measured at 520 nm. All oligonucleotides (oligos) used in these assays contain a 5′ fluorescein-tag. Each assay was carried out with 1 nM oligo in the binding buffer (200 mM NaCl, 25 mM Tris-HCl pH 7.5, 2.5 mM $MgCl_2$). Either SlmA or FtsZ was titrated into the reaction mixture. The polarization data was analyzed with KaleidaGraph and fitted to a simple bimolecular binding model by nonlinear regression (105).

2.2 Restriction Endonuclease Protection, Selection, and Amplification (REPSA)

REPSA selection utilizes a library of DNA oligos, which has static ends and a variable mid-section. The static end encodes for two different type IIS restriction enzyme sites, FokI and BpmI. These sites act to recruit the restriction enzyme to bind and cleave the DNA in the variable mid-section. This will allow for the digestion of all DNA, with the exception for those that are bound by SlmA. SlmA can bind to a specific site and protect it from the restriction enzyme. Then these sites are selected for and amplify through PCR (Figure 39). The process continues until a desirable level of enrichment has occurred.

The REPSA experiment was conducted as previously described (106). Briefly, 40 μM SlmA was bound to 4 ng of REPSA selection template in binding buffer (10 mM Tris-HCl pH 7.9, 50 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT) for 30 minutes (min) at 37 °C. The cleavage reaction was then performed with either 0.5 unit of FokI or BpmI restriction enzyme for 5 min at 37 °C. Products bound by SlmA and consequently protected from endonuclease digestion were amplified by PCR. Resulting PCR products were subjected to additional rounds of
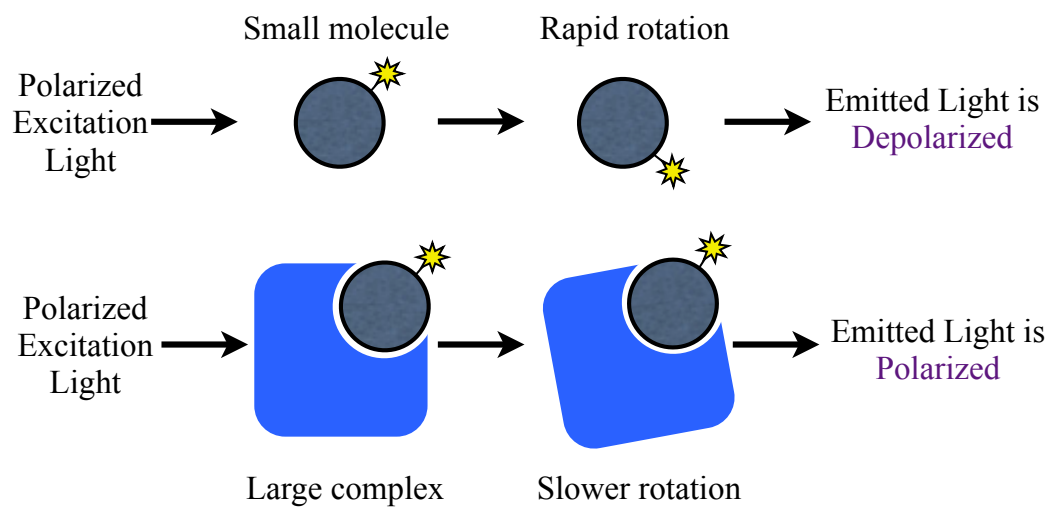
Figure 38. Fluorescence polarization theory diagram. This technique can be used to observe molecular interactions. This method makes use of a fluorescently labeled molecule and the polarization of this molecule is measured as it interacts with a second, unlabeled, molecule. Once excited by polarized light, a small, labeled, molecule can rotate rapidly and the emitted light is depolarized. If this molecule is bound by a larger protein, the larger complex will rotate slower and the emitted light will be polarized, in comparison.
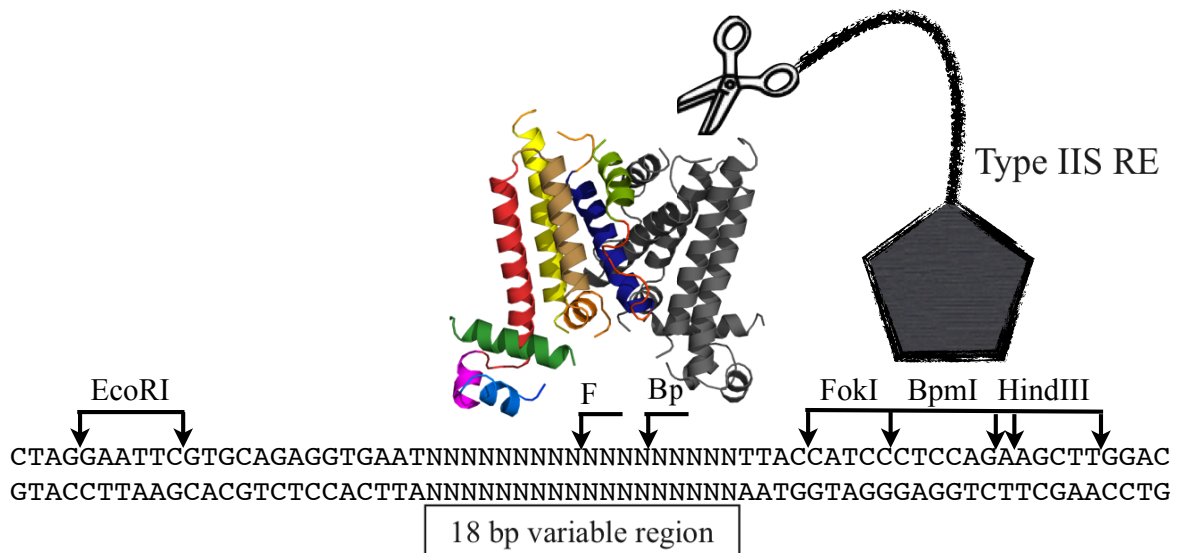
Figure 39. Restriction Endonuclease Protection, Selection, and Amplification (REPSA).
Identification of the SlmA DNA Binding Sequence. REPSA selection makes utilizes a library of DNA oligos which has static ends and a variable mid-section. The static end encodes for two different type IIS restriction enzyme sites, FokI and BpmI. These sites act to recruit the restriction enzyme to bind and cleave the DNA in the variable mid-section (F and Bp, respectively). This will allow for the digestion of all DNA, with the exception for those that are bound by SlmA. SlmA can bind to a specific site and protect it from the restriction enzyme. Then these sites are selected for and amplify through PCR.

selection until convergence, as detected by DNA sequencing. The resulting DNA sequences were analyzed by the multiple expectation maximum for motif elicitation program (MEME) (107). Default parameters were used to search for palindromic motifs. The position specific scoring matrix from the MEME analysis was input into FIMO (Find Individual Motif Occurrence) (108) with default parameters. The *Escherichia coli* strain K-12 sub-strain MG1655 (GenBank ID: U00096) was used as the sequence input.

MEME is a computational tool used for discovering sequence motifs from a group of DNA or protein sequence. The methodology and web-based tool are both described by Bailey and colleagues in (107). Briefly, MEME uses statistical modeling to search for repeated and ungapped sequence patterns. These motifs are represented as position-dependent probability matrices. A multi-fasta file containing the 43 unique predicted SlmA binding sequences is inputted into the MEME web-based program. The default settings were used. The output is a motif represented as a matrix of probabilities. An example of the output is shown in Figure 40.

2.3 Chromatin Immunoprecipitation coupled with Polymerase Chain Reaction

slmA with an N-terminal FLAG tag was inserted into a pDSW210 vector between XbaI and PstI, and this construct was then placed in a *slmA- E. coli* cells to produce strain WM3363 WM3363 cells were grown in LB with 50 μg/ml ampicillin and 1 mM IPTG for 3 hrs to an OD_{600} of ~0.5 OD. Cells were harvested, washed twice with PBS and resuspended in 10 ml PBS. Crosslinking was accomplished by adding formaldehyde to a final concentration of 1% for 20 min followed by quenching with 0.25 M glycine for 10 min at RT. Cells were washed twice with 10 ml PBS and resuspended in 1 ml lysis buffer (10 mM Tris pH 8.0, 20% sucrose, 50 mM NaCl, 10 mM EDTA, 1 mg/ml lysozyme) and incubated at 37 °C for 1 hr. Lysates were sonicated to an average size of 300-1000 bp and cellular debris removed by centrifugation. Lysates were diluted 1:2 in immunoprecipitation buffer (50 mM Hepes-KOH pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% sodium dodecyl sulfate [SDS], 1 mM phenylmethylsulfonyl fluoride) to a final volume of 800 μl and 1 μg of FLAG M2 antibody was incubated overnight at 4 °C followed by the addition of 50 μl protein A/G plus agarose for 2 hrs at 4 °C. Following immunoprecipitation, the samples were washed once in 1 ml immunoprecipitation buffer, twice with immunoprecipitation buffer plus 500 mM NaCl, once is wash buffer (10 mM Tris-HCl pH 8.0, 250 mM LiCl, 1 mM EDTA, 0.5% Nonidet-P40, 0.5% sodium deoxycholate) and once with Tris-EDTA pH 7.5.

Figure 40. Excerpt of MEME output. MEME uses statistical modeling to search for repeated and ungapped sequence patterns. The alignment of DNA sequences is used to determine a consensus motif.

Immunoprecipitated complexes were then removed from the beads by treatment with 100 µl elution buffer (50 mM Tris-HCl pH 7.5, 10 mM EDTA, 1% SDS) at 65 °C for 10 min. Samples were uncross-linked by incubation for 2 hrs at 42 °C and 6 hrs at 65° C in 0.5 X elution buffer plus 0.5 mg/ml proteinase K. DNA was purified using a minElute kit and eluted in a final volume of 250 µl of 10 mM Tris pH 8.0.

Following DNA purification, 1 µl of sample was analyzed by PCR using 500 nM primer and the reaction allowed to proceed for 28 cycles before 10 µl of product was analyzed on a 2% agarose gel. All binding sites were validated independently in triplicate.

2.4 Chromatin Immunoprecipitation coupled with Sequencing

Immunoprecipation and subsequent purification of SlmA associated DNA was prepared as described in section 2.3. Sequencing of the immunoprecipitated products were done on an Illumina Solexa Genome Analyzer II. Peaks are determined based on the enrichment of tags in a 200-bp window in ChIP sample of FLAG-tagged *slmA* comparing to the control. In the ChIP samples, the tags in the 200-bp window are also compared with immediate vicinity of regions surrounding the window to ensure that the peak is enriched in comparison to its genome background.

Following Zhang *et al.* (MACS), we shift the tags toward upstream for tags mapped to '+' strand and toward downstream for tags mapped to '−' strand---in order to improve the signal for detecting binding event. The distance shifted is determined by the maximum of the correlation function $C(d) = \sum_i s_+(i)s_-(i+d)$ where $s_+(i)$ and $s_-(i)$ are the number of tags mapped to genome location i for the + and − strands. The local maximum of the correlation function of $C(d)$ near the estimated DNA fragment length is taken to be d. For the ChIP sample, $d = 100$ and for the control, $d = 116$. The distance shifted is $\frac{d}{2}$.

Given the number of shifted tags in a 200-bp window in ChIP sample and in control sample, and given ratio of tags determined from the total number of tags sequenced in both samples, we can compute a *p*-value of ChIP sample enrichment using a binomial distribution. Call this value p_c. Similarly to ascertain that the tags are enriched in the 200-bp window relative to surroundings in the ChIP sample, we compare the number of shifted tags in the 200-bp window to the number of shifted tags in the 500-bp windows upstream and downstream of

the 200-bp window. A binomial distribution and an expected ratio of $p = 0.2$ is used to compute the probability p_s.

To determine the threshold, we have tried three schemes.

(1) Without using p_s; adjust the threshold p_c to achieve a pre-determined false discovery rate (FDR).

(2) Selecting only peaks that $p_s < 10^{-5}$; adjust p_c to achieve a pre-determined FDR.

(3) Setting $p_c = p_s = p$, adjust p to achieve a pre-determined FDR.

The FDR is computed by dividing the peaks found in control vs ChIP by that from ChIP vs control. At the same FDR level, the best scheme has the largest fraction of peaks containing a REPSA motif. Using this criterion, scheme (3) is the best.

When $p_c = p_s = 10^{-43}$, the FDR formally reaches zero (no peak is found in control vs. ChIP). At this threshold, ChIP vs control has 62 peaks. Some of these peaks consist of several 200-bp windows that are significant. Each of the 62 peaks is distinct not directly connect to another peak.

We examined all 62 peaks with the UCSC genome browser. 10 of 62 peaks are not real. 8 of the 10 are in an rRNA cluster. Because several copies of rRNA exist in the *E. coli* genome, these regions contain repetitive sequences that can cause problem in mapping. For the remaining two false peaks, one covers the gene *slmA*; the other covers the gene *lacI*. For both genes, the tags are enriched relative to the background in both the ChIP sample and in the control suggesting that the FLAG antibody binds to the protein.

2.5 Crystallization and structure determination of SlmA

The *slmA* gene was purchased from Genscript Corporation (Piscataway, NJ, USA); Web: www.genscript.com. The gene was subcloned into pET15b such that an N-terminal hexa-histidine tag was expressed and the protein was purified using Ni-NTA chromatography. SlmA protein was concentrated to 6 mg/ml and crystallized in 100 mM Tris pH 8.5, 10% PEG 400, 58 mM LiSO$_4$ by hanging-drop vapor diffusion. Data were collected at the Advanced Light Source (ALS) beamline 8.2.1 and processed with MOSFLM and SCALA. The SlmA structure was solved by MAD using crystals grown with selenomethionine-substituted protein. MAD data were collected and the selenium sites were located using SOLVE (109). Model building was carried out using Coot (110) and refinement with CNS (111). The SlmA structure contains one molecule per asymmetric unit, and has R$_{work}$/R$_{free}$ values of 22.4%/26.5% to 2.5 Å

resolution. The oligomeric states of SlmA and SlmA-DNA were determined by size exclusion chromatography on a Superdex 200 26/60 column, using appropriate standards. Coordinates and structure factor amplitudes for the SlmA structure have been deposited with the Protein Data Bank under the accession code 3NXC.

2.6 Small Angles X-ray Scattering

SlmA, FtsZ, and FtsZ-GFP proteins were expressed in *E. coli* strain BL21 (DE3). Each protein was purified to at least 95% purity. Samples were stored in a buffer containing 25 mM Tris-HCl pH 7.5, 200 mM NaCl, 1 mM MgCl2, and 5% glycerol. Sample concentration was determined via A280 measurement. In order to obtain a blank, during the process of concentrating each sample, the flow-through were collected and used for the purpose of buffer subtraction. Samples of SlmA, FtsZ, FtsZ-GFP, SlmA-FtsZ (1:1 stoichiometric mixture), and SlmA-FtsZ-GFP (1:1) were concentrated to the maximal level of 6 mg/ml, 10 mg/ml, 10 mg/ml, 5 mg/ml, and 6 mg/ml, respectively. SAXS was conducted on each of these samples at concentrations range from 1 mg/ml to its maximal concentration.

SAXS data were collected at the ALS beamline 12.3.1 at a wavelength of 1 Å and a temperature of 10 °C (Lawrence Berkeley National Laboratory, Berkeley, CA) (112). SAXS data were collected for protein samples over a range of concentration, and the profiles were evaluated for aggregation using Guinier analyses (113). The radius of gyration (R_G) was derived by the Guinier approximation $I(q) = I(0)\ exp(-q^2 R_G{}^2/3)$ with the limits $qR_G < 1.3$. The program GNOM (114) was used to compute the pair distance distribution functions, $P(r)$. The overall shapes were calculated from the experimental data using the program DAMMIN (115) or GASBOR (116). The models generated by BUNCH (117) were evaluated for q of ranges (0.020–0.40 Å^{-1}).

2.7 Negative stain electron microscopy

All samples (FtsZ, SlmA and their complexes with and without DNA) were in a buffer consisting of 25 mM Hepes pH 7.4, 100 mM potassium glutamate, 300 mM potassium acetate, 5 mM magnesium acetate and 2 mM GTP. The concentration of FtsZ was 3 mM and SlmA was 0.6 mM. The DNA (GCAGTGAGTACTCACTGC; top strand) was 1 mM. Samples were placed on 100 mesh formvar coated copper grids treated with poly-L-Lysine for 1 hr. Excess samples were blotted with filter paper, then stained with filtered 2% uranyl acetate for 1 min.

Stain was blotted dry from the grids with filter paper and samples were allowed to dry. Samples were then examined in a JEM 1010 transmission electron microscope (JEOL, USA, Inc., Peabody, MA) at an accelerating voltage of 80 Kv. Digital images were obtained using the AMT Imaging System (Advanced Microscopy Techniques Corp., Danvers, MA).

2.8 Size Exclusion Column Chromatography

The oligomeric state of SlmA was determined by size exclusion chromatography. This was performed as described in Kumaraswami *et al.*, 2009 (118). 400 nmol of SlmA alone or with an equal molar amount of the extended SlmA DNA binding sequence (SBS), GCAGTGAGTACTCACTGC, was injected into a Superdex 200 26/60 column (Amersham Pharmacia Biotech) with a mobile phase of 25 mM Tris–HCl pH 7.5, and 200 mM NaCl. The elution volume was then plotted against a standard curve to determine the relative molecular weight of the sample.

The standard curve was generated a Superdex 200 26/60 column (Amersham Pharmacia Biotech) with a mobile phase containing 20 mM Tris–HCl pH 7.5, 200 mM NaCl and 2.5% glycerol. The void volume was determine with the use of Blue dextran (Sigma). Four protein samples were used to calibrate the column: RNase A (Mr 13 700 kDa), carbonic anhydrase (Mr 29 000 kDa), bovine serum albumin (Mr 66 000 kDa) and alcohol dehydrogenase (Mr 150 000 kDa). The $K_{Average}$ (K_{ave}) was calculated as follow: $K_{ave}=(V_E - V_O)/(V_T - V_O)$, where V_T, V_E and V_O are the total column volume, elution volume and void volume of the column, respectively. The standard plot was generated by graphing the logarithm of the molecular weight (Mr) against the K_{ave}. The K_{ave} of each marker as well as the experimental samples were the average value of three experiments.

2.9 Expression and purification of FtsZ

Full length FtsZ and C-terminal truncated FtsZ, FtsZ(1-360), from *E. coli* were produced as previously described with minor modifications (42). Specifically, an extra 25% ammonium sulfate precipitation was performed and the precipitant was solubilized in storage buffer (50 mM Tris pH 8, 100 mM NaCl, 1 mM MgSO₄). Proteins were polymerized with 10 mM MgSO₄, 1 M monosodium glutamate, and 1 mM GTP at 37 °C for 30 min. The resulting pellets were then redissolved in storage buffer. The FtsZ(1-316)-GFP fusion protein was generated by cloning a C-terminal truncated version of *E. coli* FtsZ (from residues 1-316) along with GFP in

the pET15b vector. The protein was expressed in BL21 (DE3) cells and purified in one step using Ni-NTA chromatography.

2.10 SlmA dimer-interface mutant

The SlmA dimer-interface mutant has three mutations: L171R, Q175R, and F179R (referred to as the LQF SlmA mutant). This mutant was generated with single step QuikChange site-direct mutagenesis (Stratagene, La Jolla, CA) using the following primers:

```
NT_L171R_Q175R_F179R_f: GATGAAACCCTG CgG GCCAGC CgG ATTCTGGCG cgT TGCGAAGGC
NT_L171R_Q175R_F179R_r: GCCTTCGCA Acg CGCCAGAAT CcG GCTGGC CcG CAGGGTTTCATC
```

The template used was the same pET15b *slmA* plasmid that has been used in previous experiments. The LQF plasmid was transformed into the BL21 (DE3) strain. Expression of both wild-type and LQF SlmA was done at 37 °C. Expression of the protein was induced with 1 mM IPTG at $OD_{600} = 0.6$. After 4 hours of induction, 1 mL of cells containing the wild-type and LQF SlmA was harvested. The cells were lysed with B-Per solution (Pierce, Rockford, IL) for 15 minutes at room temperature. The solution was subsequently pelleted and the supernatant and pellet were separate and subjected to SDS PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis).

Chapter 3. Introduction to the SlmA mediated nucleoid occlusion system

Cell division or cytokinesis is one of the most fundamental processes in biology and is essential for the propagation of all living species. In *Escherichia coli*, cell division occurs by ingrowth of the membrane envelope at the cell center and is orchestrated by the FtsZ protein (7, 36, 119). FtsZ has a tubulin-like fold and self-assembles into linear protofilaments in a GTP dependent manner by the interaction of the plus end of one subunit with the minus end of another subunit, resulting in a head-to-tail geometry. While FtsZ and tubulin protofilaments are similar, the higher order polymers they form are notably different. Specifically, tubulin protofilaments interact to produce microtubules while FtsZ protofilaments combine to form a cytoskeletal scaffold called the Z-ring (25, 120). The Z-ring provides the framework for the assembly of the division apparatus and determines the site of cytokinesis (25, 36). Several studies have suggested that the functional unit of FtsZ used in Z-ring formation consists of parallel interacting FtsZ protofilaments, which have been termed "thick filaments" (49, 50). However, the precise arrangement of FtsZ protofilaments within the Z-ring is currently unknown. The total amount of FtsZ molecules in a cell significantly exceeds the concentration required for Z-ring formation and this concentration remains constant during the cell cycle. Hence, Z-ring formation must be highly regulated, both temporally and spatially. In particular, the assembly of Z-rings at the cell poles and over chromosomal DNA must be prevented. These inhibitory roles are played by two key regulatory systems called the Min system and the nucleoid occlusion (NO) system, respectively (121).

The Min system has been extensively studied and, in *E. coli*, is comprised of the FtsZ inhibitor, MinC, a membrane associated ATPase called MinD and MinE, a factor that binds and spatially organizes the MinCD complex (122-124). MinC, which interacts with MinD, inhibits FtsZ polymerization by preventing lateral interactions required for Z-ring formation. MinE binds MinCD and oscillates from pole to pole (71, 72). The net result of this oscillatory process is the formation of a zone of FtsZ inhibition at the cell poles. However, the replicated nucleoid DNA near the midcell must also be protected from bisection by the Z-ring and this is ensured by NO. In contrast to the Min system, the mechanisms responsible for NO have been unclear. Indeed, although the process of NO was proposed over 20 years ago by Woldringh and coworkers, it took until 2004 for Wu and Errington to identify a factor, Noc, that is responsible for NO in *Bacillus subtilis* (80-82). Cells lacking Noc had no obvious cell division phenotype, but inhibiting DNA replication, in a Min mutant background, resulted in aberrant formation of

cell division septa over unpartitioned nucleoids and subsequent nucleoid guillotining. Furthermore, Noc localized to nucleoids and excess Noc inhibited division. These findings established Noc as a *bona fide* NO factor. Subsequently, it was shown that Noc binds to specific DNA sites with the consensus, 5′-ATTTCCCGGGAAAT-3′ in the *B. subtilis* chromosome (84). However, the mechanism by which Noc prevents Z-ring formation over the nucleoid is still unclear as it does not appear to bind FtsZ or any regulator of cell division.

Following the discovery of Noc, a 198 residue protein called SlmA was shown to be the effector of NO in *E. coli* (86). SlmA was identified similarly to Noc, in a screen designed to isolate mutations that were lethal in the absence of Min, hence the name SlmA (synthetic lethal with a defective Min system). Like Noc, SlmA was shown to bind DNA and localized to the nucleoid fraction of the cell. However, SlmA and Noc show no sequence homology and belong to different families of DNA binding proteins. While Noc is a ParB-family member, SlmA contains a putative N-terminal HTH motif and a predicted C-terminal coiled-coil (86, 125). Light scattering experiments suggested that SlmA interacts with FtsZ-GTP and alters its polymerization properties. However, this interaction appeared to enhance rather than disrupt polymer formation, leaving in question how it could be involved in NO. Here we describe studies that reveal the molecular mechanism by which SlmA mediates NO in *E. coli*. Specifically, we determined the crystal structure of SlmA, identified its DNA binding site specificity and mapped its binding sites on the *E. coli* chromosome by ChIP experiments. We went on to determine the SlmA-FtsZ structure by small angle X-ray scattering (SAXS) and examined the affect of SlmA-DNA on FtsZ polymerization by electron microscopy. Our combined data show how SlmA is able to disrupt Z-ring formation through its interaction with FtsZ in a specific temporal and spatial manner and hence prevent nucleoid guillotining during cell division.

Chapter 4. The X-ray crystal structure of SlmA

To gain insight into the function of SlmA, we first crystallized (Figure 41) and determined its crystal structure to 2.50 Å resolution by multiple wavelength anomalous diffraction (MAD) (Table 1). The final SlmA structure consists of residues 9-25, 32-113, 120-148, 150-198, contains 14 solvent molecules and has R_{work}/R_{free} values of 22.4%/26.5%. The structure shows that SlmA is comprised of nine helices (α1-α9) and can be divided into two domains, a small N-terminal domain (residues 1-53) and a C-terminal domain (residues 54-198). The N-terminal domain is formed by the first three helices (α1-α3). Helices 2 and 3 form a canonical helix-turn-helix (HTH) motif, suggesting that this domain functions in DNA binding. Helices α4-α9 form the C-terminal domain and crystal packing analyses reveal that this region mediates dimerization (Figure 42A). The dimer interface buries an extensive 2640 Å^2 of protein surface from solvent. Typical dimer interfaces bury on the order of 1000 Å^2 and the dimerization or protein-protein binding energy has been shown to be directly related to the buried hydrophobic surface area (Janin *et al*, 1988). The SlmA dimer interface is unusual compared to most oligomer interfaces in that it is almost entirely hydrophobic. Residues that are involved in dimerization include Leu171, Ala178, Phe179, and Met183 on the internal face of helix α8, and Ala202 and Ala209 from helix α9 (Figures 42B). Mutation of three hydrophobic residues, Leu171Arg, Gln175Arg and Phe179Arg, resulted in insoluble protein that was found in inclusion bodies, underscoring the important role these residues play in dimerization and hence proper protein folding (Figure 43A-B). Size exclusion chromatography analyses, which resulted in a calculated mass of 48 kDa, support that SlmA is dimeric (Figure 42C).

Database searches using the Dali server (126) revealed that the SlmA structure is most similar to that of the QacR protein, thus establishing SlmA as a new member of the TetR family. The DNA binding domains of all TetR proteins show sequence homology, however, their C-terminal domains do not. Despite this, all TetR members whose structures have been solved possess C-terminal domains that are similar structurally. In TetR proteins the C-terminal domain mediates dimerization and all TetR proteins are dimers (Ramos *et al*, 2005). A multiple sequence alignment of SlmA with TetR members that have been structurally characterized showed that the most conserved region between the proteins lies within the HTH, which overall shows 23% sequence similarity compared to the 6% sequence correspondence found in the comparison of their C-terminal domains (Figure 44). Despite the lack of sequence similarity, structural superimpositions of SlmA with TetR members QacR and TetR, reveal that SlmA has
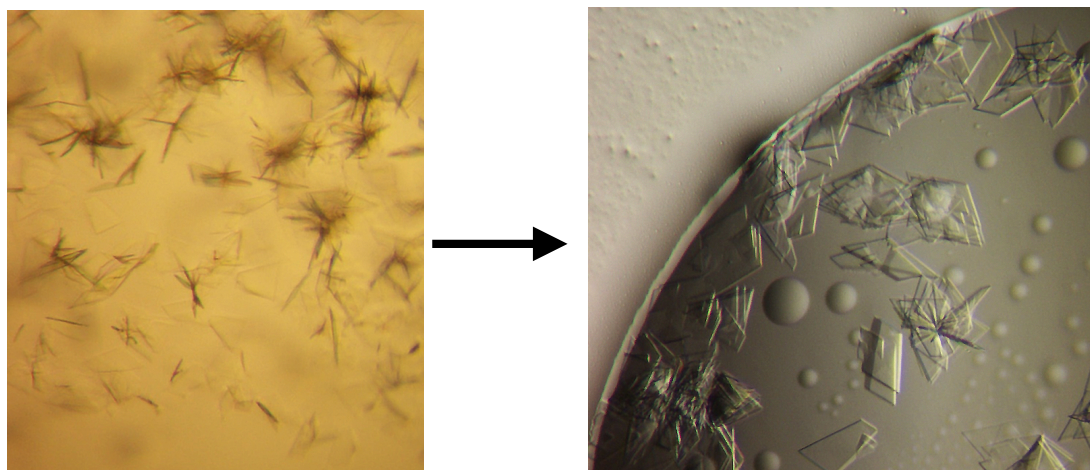
Crystallization condition:
100 mM Tris pH 8.5
10% PEG 400
58 mM LiSO$_4$
5 mg/ml
1:2 ratio of protein to ppt
12°C for 7 days
cryo condition: 25% PEG 400
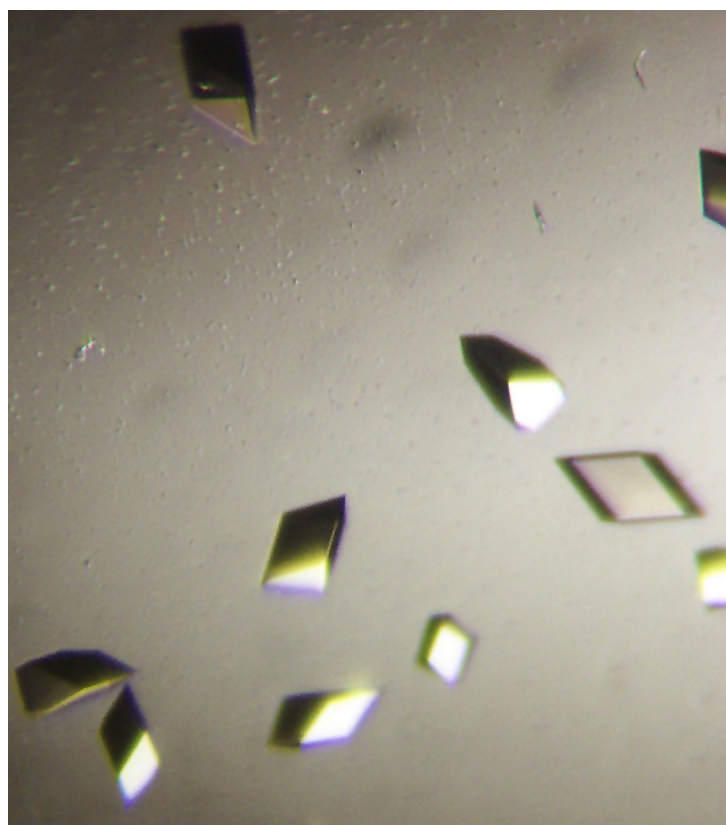
Figure 41. Crystallization of SlmA. SlmA protein was grown in 100 mM Tris pH 8.5, 10% PEG 400, and 58 mM LiSO$_4$. The hanging drop vapor diffusion method was used. SlmA was mixed to the crystallization solution at a concentration of 5 mg/ml in a ratio of 1:2 (protein:ppt). The crystals appeared within seven days. The optimization of SlmA crystals is shown.

Table 1. Statistics of X-Ray data and refinement.

Data Collection	SeMet SlmA λ1	SeMet SlmA λ2	SeMet SlmA λ3
Space group	P3$_2$21	P3$_2$21	P3$_2$21
a, b, c (Å)	50.38, 50.38, 121.32	50.38, 50.38, 121.32	50.38, 50.38, 121.32
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 120
Wavelength (Å)	0.9795	0.97953	0.95372
Resolution (Å)	60.86-2.5 (2.63-2.5)	60.86-2.5 (2.64-2.5)	60.86-2.5 (2.57-2.5)
Rmerge (%)	10.9 (42.1)	10.6 (41.3)	11.9 (48)
I/σI	19.4 (4.1)	20.2 (4.1)	18.0 (3.6)
Completeness (%)	99.8 (100)	99.7 (99.7)	99.7 (99.9)
Multiplicity	6.6 (7.0)	6.6 (7.0)	6.6 (7.0)
Total reflections (#)	44415	44378	47741
Unique reflections (#)	6683	6678	7191
MAD Phasing			
Figure of merit	0.55		
Refinement			
Resolution (Å)	2.5		
R$_{work}$/R$_{free}$ (%)	22.4% / 26.5%		
Ramachandran Analysis			
Most favored (%)	89.2%		
Additionally allowed (%)	7.6%		
Generously allowed (%)	3.2%		
Disallowed (%)	0%		
RMSD			
Bond lengths (Å)	0.008		
Bond angles (Å)	1.47		
B main-chain atoms (Å^2)	1.358		

[a]$R_{merge} = \Sigma_{hkl}\Sigma_i |I_i(hkl) - \overline{I(hkl)}|/\Sigma_{hkl}\Sigma_i I_i(hkl)$, where $I_i(hkl)$ is observed intensity and $\overline{I(hkl)}$ is the final average value of intensity.

[b]Values in parentheses are for the highest resolution shell.

[c]Figure of Merit = $<|\Sigma P(\alpha)e^{i\alpha}/P(\alpha)|>$, where is the phase and P(α) is the phase probability distribution.

[d]$R_{work} = \Sigma||F_{obs}| - |F_{calc}||/\Sigma|F_{obs}|$ and $R_{free} = \Sigma||F_{obs}| - |F_{calc}||/|F|_{obs}|$, where all reflections belong to a test set of 10% data randomly selected by CNS.
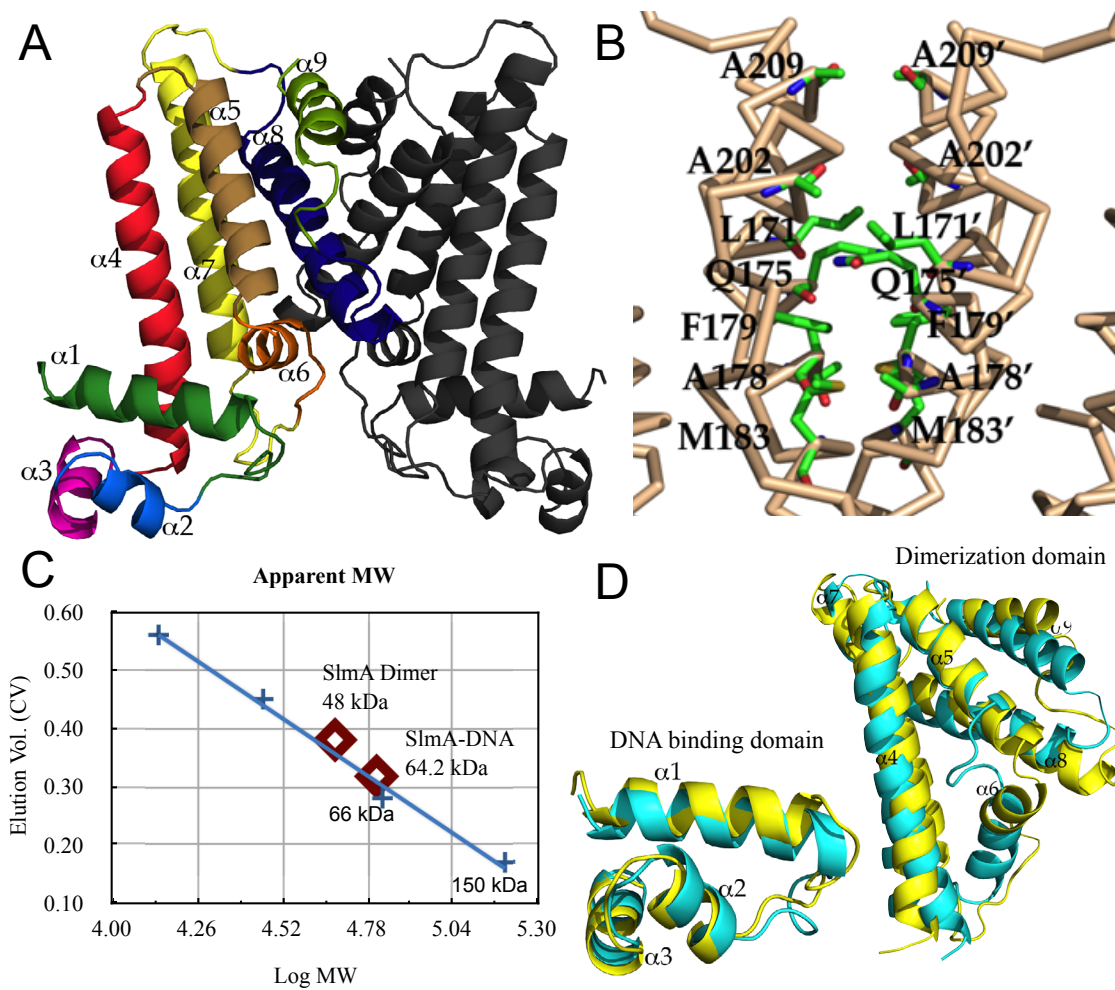
Figure 42. Overall structure of SlmA and analyses of the SlmA dimerization domain. A) Ribbon diagram of the SlmA dimer. In the left subunit, the helices are colored differently and each helix is labeled. **B)** SlmA dimer interface. The dimer was generated from crystallographic symmetry. Residues on α8 and α9 that mediate dimerization are shown as green sticks and labeled. **C)** SlmA and the SlmA-DNA stoichiometries as determined by size exclusion chromatography. The graph illustrates that both are dimeric. **D)** Overlay of the SlmA (yellow) and QacR (cyan) DNA binding domains (helices α1 to α3), and dimerization domains (helices α4 to α6).
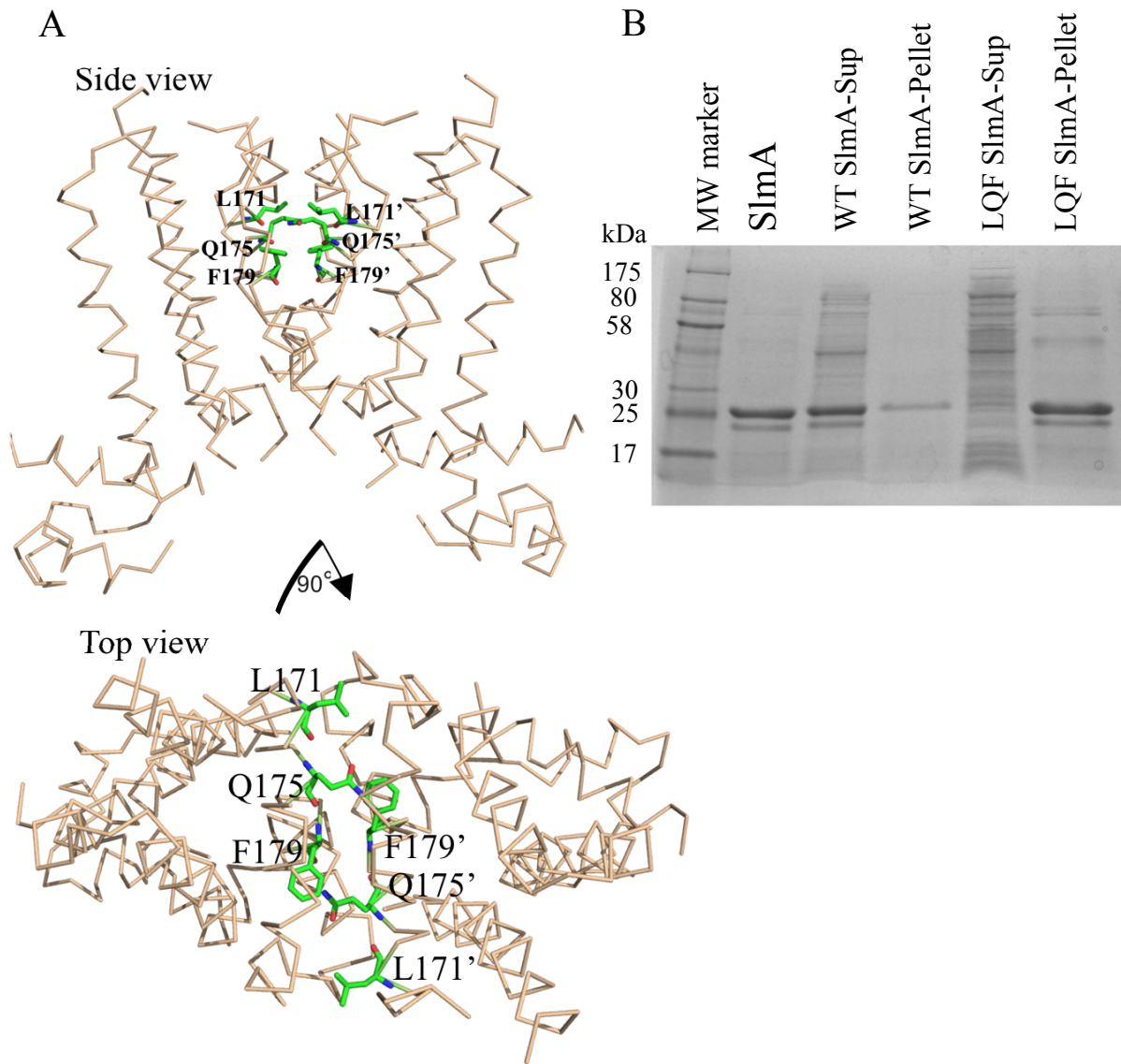
Figure 43. Disruption of SlmA dimer interface leads to insoluble protein. A) A side and top view of the SlmA protein with the three residues which were mutated are shown. The residues Leu171, Gln175, and Phe179 were all mutated to Arg, and will be referred to as LQF SlmA. **B)** SDS PAGE of SlmA (~25kDa), supernatant (soluble fraction) and pellet (insoluble fraction) of cells expressing WT and LQF SlmA are shown.

```
SlmA                                    ━━━━━━━━━━  ━━━ α1 ━━━  ━━━━━━━━━━          ━━ α2 ━━  ■
SlmA    ----------  ---MAEKQTA  KRNRREEILQ  SLALMLESSD  GSQRITTAKL  AASVGVSEAA   47
EthR    MTTSAASQAS  LPRGRRTARP  SGDDRELAIL  ATAENLLEDR  PLADISVDDL  AKGAGISRPT   60
QacR    ----------  ----------  --MNLKDKIL  GVAKELFIKN  GYNATTTGEI  VKLSESSKGN   38
TetR    ----------  ----------  MSRLDKSKVI  NSALELLNEV  GIEGLTTRKL  AQKLGVEQPT   40
CgmR    ----------  ----------  MRTSKKEMIL  RTAIDYIGEY  SLETLSYDSL  AEATGLSKSG   40
TtgR    ----------  ---MVRRTKE  EAQETRAQII  EAAERAFYKR  GVARTTLADI  AELAGVTRGA   47
                                     .      :          *             :   .: .              .

SlmA    ■ α3 ━━━   ━━━━━━━━━━  ━━━━ α4 ━━━━  ━━━━━━━━━━       ━━━━━ α5 ━━━━  -■
SlmA    LYRHFPSKTR  MFDSLIEFIE  DSLITRINLI  LKD-EKDTTA  RLRLIVLLLL  G---FGERNP  103
EthR    FYFYFPSKEA  VLLTLLDRVV  NQADMALQTL  AENPADTDRE  NMWRTGINVF  F-----ETFG  115
QacR    LYYHFKTKEN  LFLEILNIEE  SKWQEQWKKE  QIKCKTNREK  FYLYNELSLT  T-----EYYY   93
TetR    LYWHVKNKRA  LLDALAIEML  DRHHTHFCPL  EGESWQDFLR  NNAKSFRCAL  LSHRDGAKVH  100
CgmR    LIYHFPSRHA  LLLGMHELLA  DDWDKELRDI  TRD-PEDPLE  RLRAVVVTLA  EN--------   91
TtgR    IYWHFNNKAE  LVQALLDSLH  ETHDHLARAS  ESEDEVDPLG  CMRKLLLQVF  NELVLDARTR  107
        :    :.  .:       :.    :              .                  .

SlmA    ■ α6 ━  ━━━━━━━━━━━  ━━━━━━━━━ α7 ━━━━━━━━━━  ━━━━━━━━━━        ■■■■■
SlmA    GLTRILTGHA  LMFEQ-DRLQ  GRINQLFERI  EAQLRQV--L  REKRMREGEG  YATDETLLAS  160
EthR    SHKAVTRAGQ  AARATSVEVA  ELWSTFMQKW  IAYTAAV--I  DAERDRGAAP  RTLPAHELAT  173
QacR    PLQNAIIEFY  TEYYKTNSIN  EKMNKLENKY  IDAYHVI--F  KEGNLNGEWC  IN-DVNAVSK  150
TetR    LGTRPTEKQY  ETLENQLAFL  CQQGFSLENA  LYALSAVGHF  TLGCVLEDQE  HQVAKEERET  160
CgmR    -VSRPELLLL  IDAPSHPDFL  NAWRTVNHQW  IPDTDDL--E  NDAHKR----  -------AVY  137
TtgR    RINEILHHKC  EFTDDMCEIR  QQRQSAVLDC  HKGITLA--L  ANAVRRGQLP  GELDAERAAV  165
                                  .

SlmA    ■■■■ α8 ■■■■  ━━━━━━━━━━  ━━ α9 ━━
SlmA    QILAFCEG-M  LSRFVRSEFK  YRPTDDFDAR  WPLIAAQLQ-  --------    198
EthR    ALNLMNERTL  FASFAGEQPS  VPEARVLDTL  VHIWVTSIYG  ENR-----    216
QacR    IAANAVNG--  IVTFTHELNI  NERIKLMNKF  SQIFLNGLSK  --------    188
TetR    PTTDSMPP-L  LRQAIELFDH  QGAEPAFLFG  LELIICGLEK  QLKCESGS    207
CgmR    LVQLAADGLF  VHDYIHDDVL  SKSKR--QAM  LETILELIPS  QT------    177
TtgR    AMFAYVDG-L  IRRWLLLPDS  VDLLGDVEKW  VDTGLDMLRL  SPALRK--    210
                     .                                 :
```

Figure 44. Multiple sequence alignment of SlmA and other TetR family members. Multiple sequence alignment. Secondary structural elements are represented above the sequence and colored as in figure 42A. Amino acid residues considered to be similar are highlighted, and the degree of similarity is represented below the sequence: '*' represents perfect identity, ':' represents highly similar, and '.' represents moderately similar.

the same structural topology as these TetR proteins. In particular, comparison of the DNA binding and dimerization domains of SlmA to QacR yielded a root mean squared deviation (RMSD) of 1.9 Å and 3.9 Å, respectively (Figure 42D) (88, 94, 96, 127).

The biological functions of 85 TetR members have been elucidated (87). Notably, all these proteins function as transcriptional regulators. The genes they regulate encode products involved in diverse pathways such as multidrug resistance, catabolism, antibiotic biosynthesis, osmotic stress, and the pathogenicity of Gram-negative and Gram-positive bacteria. To carry out their regulatory functions, TetR proteins respond to small molecule ligand sensors (87, 88, 96, 101, 127). Indeed, the notable lack of sequence homology within the C-terminal dimerization domains of TetR proteins reflects the fact that, in addition to dimerization, this domain also functions as a ligand binding domain. Ligand binding leads to structural changes that cause the proteins to dissociate from their DNA sites, allowing transcription. Although SlmA is clearly a member of this family of transcriptional regulators, it has a very different function, which is NO (86). Consistent with this distinction, analysis of the SlmA structure reveals that unlike canonical TetR proteins, the SlmA dimerization domain contains only a small cavity with a volume of ~360 Å^3. Moreover, there is also no clear entrance to this potential pocket as it is occluded by helix α8′ from the other subunit in the dimer (Figure 45). Thus, although SlmA is a structural member of the TetR family of regulators it is unique among these proteins in that it does not function in transcription and it also lacks an obvious ligand binding site within its C-terminal dimerization domain.
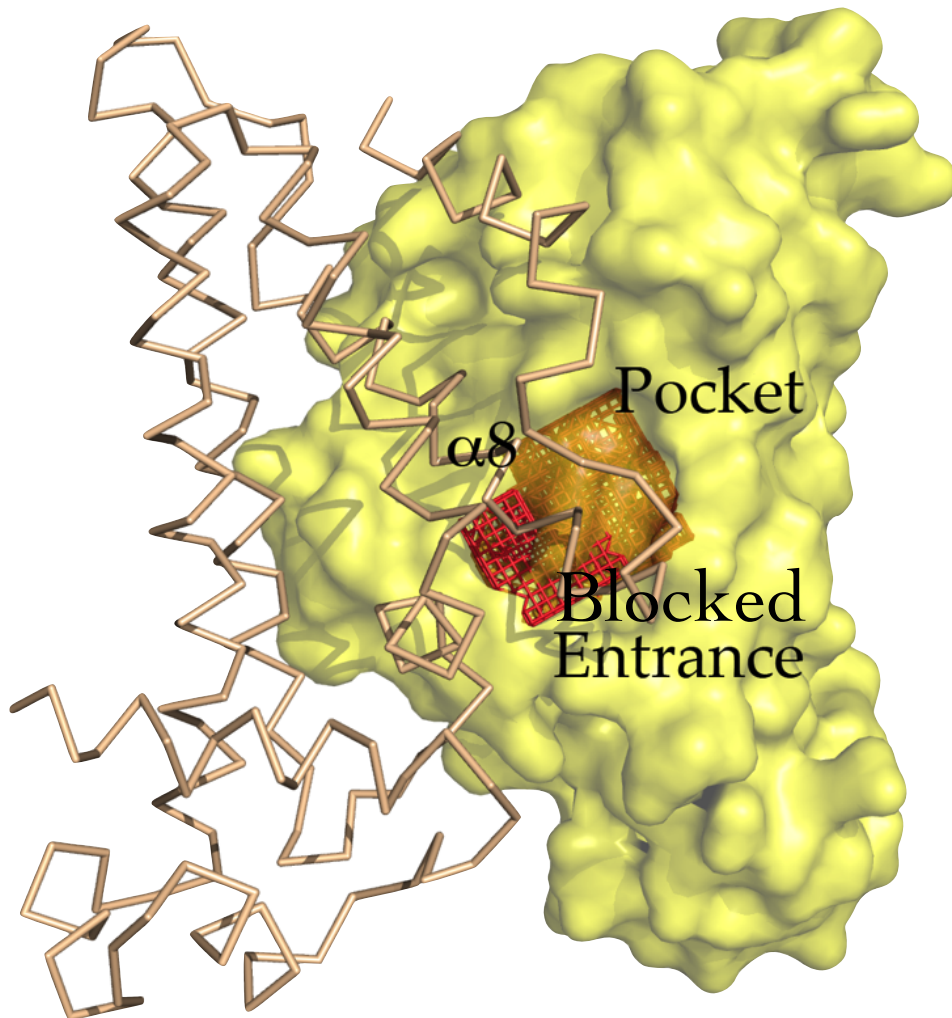
Figure 45. Possible cavity in the SlmA dimerization domain. One subunit of SlmA is represented as a surface (yellow) and the other, a ribbon. The pocket volume is represented as an orange mesh and the solvent exposed region, a red mesh. Helix α8 of the second subunit blocks the only observed entrance to the pocket.

Chapter 5. DNA binding preference of SlmA

Previous studies showed that the ability of SlmA to associate with the nucleoid is essential for its NO function. In particular, cells containing an N-terminal truncation of residues 1-64 were not functional in NO (86). Our SlmA structure, showing that it contains a N-terminal HTH and is a TetR member, suggested that it may bind the DNA major grooves of a palindromic DNA site as a homodimer, in a manner similar to other TetR proteins (88). With this *a priori* assumption, we went on to determine if SlmA displays DNA binding specificity by conducting a restriction endonuclease protection, selection, and amplification (REPSA) experiment (106). The 43 unique possible binding sequences identified via REPSA were analyzed with the sequence motif discovery program, Multiple Expectation Maximum for Motif Elicitation (MEME) (Figure 46) (107). The results revealed that SlmA binds in a specific manner to DNA duplexes containing a 12-bp palindromic site with the consensus, 5´-GTGAGTACTCAC-3´, herein called the SlmA DNA binding sequence (SBS).

To determine the affinity of SlmA for the SBS and further dissect its DNA binding preferences, we performed a series of fluorescence polarization (FP) assays (105). These analyses showed that SlmA binds the SBS with a K_d of ~50 nM. By contrast, SlmA showed no detectable binding to DNA containing randomized sequences (Figure 47A). Next, each of the 6 corresponding positions of the palindromic 12-bp SBS were systematically mutated and their binding affinities for SlmA determined (Figure 47B, 48A-B). These results showed that there is a strong preference for a G, T, A, and G at positions 1, 2, 4, and 5, respectively. Mutation of these bases significantly impaired SlmA binding. However, SBS position 3 is able to accommodate any purine nucleotide, as mutation of the guanine at this position to an adenine yielded a K_d similar to the consensus SBS of ~60 nM. Lastly, position 6 is the most flexible in terms of nucleotide specificity. Any pyrimidine in this position allowed high affinity binding to SlmA, and mutation to guanine allowed binding but with reduced affinity. The DNA sequence preferences for SlmA binding to these double mutants are summarized as a sequence logo in Figure 48B.

Subsequently, we systematically mutated each position in the first half-site of the SBS, while leaving the second half-site unchanged. The results showed that unlike the double mutations, these single mutations still permitted SlmA binding, albeit with decreased affinity. Double mutations at positions 3&10, 4&9, or 5&8, abrogated SlmA-DNA binding, but mutations at only one site (3, 4, 5, 8, 9, or 10) allows SlmA to bind to DNA with K_ds ranging

A

```
01. ACTTACTGACCCGCCCCT
02. CAGAATGTGAGTGAACAC
03. GCTTGCTAACTGCCTCTT
04. AGCTGCGAATGAGAAGCA
05. CCCAATGTGAGTACTCAC
06. CCTATAGTGTGTACTCAC
07. TCCTGTGTAAGTCCTCAC
08. ACTCACACACTGCCAATG
09. CCCAACGTAAGTGAGTAC
10. CCGCAAGTAAGTACTCAC
11. GCGACAGTGAGTCCACAC
12. GCGTCTCATTCAGACATG
13. CCCAACGTGAGCAACCAC
14. CACAACGTGAGCACTCAC
15. CCTATAGTGTGTACTCAC
16. GGGGTTTACACAGGATGC
17. CCCTTTGAATGATACGCA
18. CCCTTTGAATGATACGCA
19. CGCAACGTGAGTGAGCAC
20. ACGTATGTAAGTGCTCAC
21. CGGGGATGCGCAAAGCAG
22. GGACATGTAAGTGGTTAC
23. CTTAATGTAAGTGAGTGC
24. AAGAATCACTCGCGGGAC
25. CACGGTGTAAGCACTAAC
26. GCTGGATGCTTATTTTTG
27. AGCTGGGAATGAGACGCA
28. CGCAACGTGAGTGAGCAC
29. ACATGAGTAAGTGAGTGC
30. GCACCAGTAAGTGACCAC
31. GCTTGCTAACTGCCTCTT
32. ACTTACTAACTACGCGTT
33. GTGTGTGTGAGTACTCAC
34. CATCCAAACCCGGAAAAA
35. GCTCACTTACCCCGTGTG
36. GCGTAAGTTAGCGCTTAC
37. CGAGGATGGGCCAGCAGG
38. GGCTTTGTAAGTGCCCAC
39. GAGACGCATTCAAGGACC
40. GCGTCTCATTCAGACATG
41. CCTATGTGAGCAAGTGC
42. GTCGAGGTGAGTGTTCAC
43. GCATCCAGCCTTTCCGCA
```

B

Name	Strand	P-value	Sites
33	+	5.91E-08	GTGTGT**GTGAGTACTCAC**
5	+	5.91E-08	CCCAAT**GTGAGTACTCAC**
14	–	2.96E-07	**GTGAGTGCTCAC**GTTGTG
10	+	2.96E-07	CCGCAA**GTAAGTACTCAC**
20	+	6.52E-07	ACGTAT**GTAAGTGCTCAC**
15	+	1.69E-06	CCTATA**GTGTGTACTCAC**
6	+	1.69E-06	CCTATA**GTGTGTACTCAC**
7	+	7.52E-06	TCCTGT**GTAAGTCCTCAC**
38	–	1.09E-05	**GTGGGCACTTAC**AAAGCC
28	–	1.09E-05	**GTGCTCACTCAC**GTTGCG
19	–	1.09E-05	**GTGCTCACTCAC**GTTGCG
25	–	1.26E-05	**GTTAGTGCTTAC**ACCGTG
2	–	1.5E-05	**GTGTTCACTCAC**ATTCTG
42	–	1.59E-05	**GTGAACACTCAC**CTCGAC
36	+	2.01E-05	GCGTAA**GTTAGCGCTTAC**
13	–	2.01E-05	**GTGGTTGCTCAC**GTTGGG
9	–	2.86E-05	**GTACTCACTTAC**GTTGGG
11	+	3.02E-05	GCGACA**GTGAGTCCACAC**
30	–	3.24E-05	**GTGGTCACTTAC**TGGTGC
22	–	4.35E-05	**GTAACCACTTAC**ATGTCC
41	–	1.43E-04	**GCACTTGCTCAC**ATAGG
29	–	1.98E-04	**GCACTCACTTAC**TCATGT
23	–	1.98E-04	**GCACTCACTTAC**ATTAAG

Figure 46. Identification of the SlmA DNA binding sequence (SBS). A) A listing of the 43 unique sequences identified by REPSA. **B)** A ranking of sites identified by MEME that contain the SBS motif. The name of each site correspond to the list in (A). The DNA strand is specified, '+' or '-', corresponding to the sense or anti-sense strand. The p-value of a site is computed from the match score of the site with the position-dependent probability matrix for the motif. The sites are shown aligned with each other.
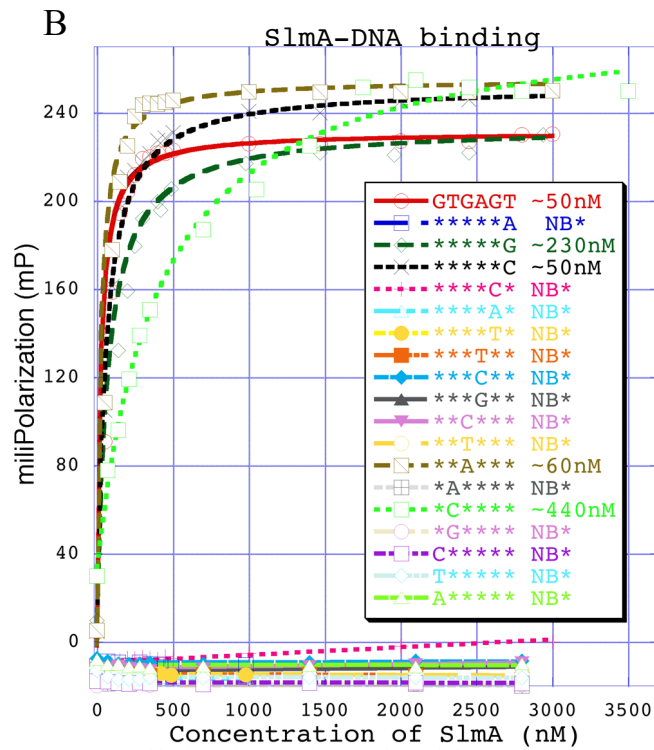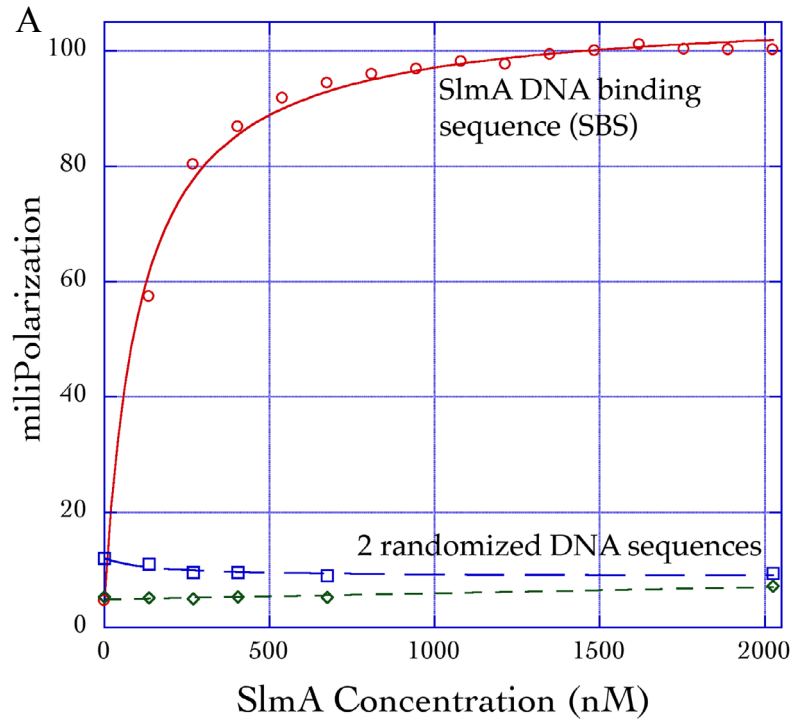
Figure 47. SlmA-DNA binding preference. A) FP binding curves of SlmA to SBS double mutant. Each position was mutated combinatorially. **B)** FP results summarizing the affect of SBS double mutations on SlmA binding. * Indicates that the nucleotides were not changed from the top strand and the nucleotides that were mutated are shown as letters.

A

Binding site	Kd
GTGAGTACTCAC	53±11 nM
*****AT*****	No binding
*****GC*****	Non-specific
*****CG*****	52±6 nM
****C**G****	No binding
****A**T****	No binding
****T**A****	No binding
T*A***	No binding
C*G***	No binding
G*C***	No binding
C****G**	No binding
T****A**	No binding
A****T**	62±8 nM
*A********T*	No binding
*C********G*	444±23 nM
*G********C*	No binding
C**********G	No binding
T**********A	No binding
A**********T	No binding

C

Binding site	Kd
GTGAGTACTCAC	53±11 nM
*****A******	No binding
*****G******	1250±43 nM
*****C******	88±11 nM
****A*******	697±23 nM
****T*******	1735±57 nM
****C*******	2319±46 nM
T*****	640±31 nM
G*****	964±52 nM
C*****	1801±61 nM
C*******	1838±54 nM
T*******	904±43 nM
A*******	56±12 nM
*A**********	No binding
*C**********	712±36 nM
*G**********	820±41 nM
C***********	No binding
A***********	No binding
T***********	No binding

B

D

Figure 48. Probing SlmA-DNA binding specificity. A) FP results summarizing the affect of SBS double mutations on SlmA binding. **B)** Sequence logo summarizing the preference of SlmA for DNA based on the FP studies in (A). **C)** FP results on the affect of single mutations in the SBS on SlmA binding. **D)** Sequence logo summarizing the preference of SlmA for DNA based on the FP studies in (C). This logo illustrates the ability of SlmA to bind to a highly variable half-site, while the second half-site conforms to the high affinity sequence.
* Indicates that the nucleotides were not changed from the top strand and the nucleotides that were mutated are shown as letters.

from high nM to low μM (650 nM to 2.3 μM) (Figure 48C-D). Similar mutations of the nucleotides in the other half-site abrogated DNA binding. These combined results were also consistent with our REPSA experiment, which showed that the first half-site of the selected DNA species was highly conserved, with the consensus, 5´-GTGAGT-3´, while the other half-site was less conserved. In summary, the results indicate that SlmA binds DNA in a sequence specific manner, but with a degree of flexibility (Figure 48D).

Chapter 6. Distribution of SlmA binding sites on the E. coli chromosome

We hypothesized that the sequence specific yet relaxed DNA binding capability of SlmA likely plays a role in its NO function. Thus, to efficiently identify all possible SlmA binding sites, we performed a ChIP (chromatin immunoprecipitation) followed by DNA sequencing (ChIP-Seq). After mapping the tag sequences onto the *E. coli* chromosome, 52 peaks were identified to be statistically significant (Figure 49A-B) (128). ChIP followed by polymerase chain reaction (ChIP-PCR) experiments were conducted on these sites confirm the positive signals (Figure 50A-B). Moreover, the Motif Alignment and Search Tool (MAST) revealed that 50 of the 52 sites conform to the SBS motif shown in Figure 48B, indicating that the SBS identified by REPSA is the specific sequence recognized by SlmA *in vivo* (108). Examination of the location of the SBS sites revealed that they are primarily clustered in specific defined regions of the chromosome called macrodomains (MD). Studies have demonstrated that the bacterial chromosome is organized into four ordered macrodomains, the Ori, Ter, Right and Left MDs and two less structured regions (so-called nonstructured regions) (11). These parts of the chromosome form compact regions and are concentrated in the same cellular space. The Ori MD contains the origin of replication and is located opposite the Ter MD, which contains the replication terminus site. On either side of the Ter domain are the Left and Right MD, while the Ori MD is flanked by the two nonstructured regions. The SBS sites cluster within the Ori MD and nonstructured regions and notably, none of these sites is located in promoter regions, consistent with previous data indicating that SlmA does not exert its NO function via transcription regulation (Figure 51A) (86). In addition, we see no evidence of spreading of SlmA along the DNA as has been observed for Noc and other ParB proteins (Wu & Errington, 2004). Perhaps the most significant finding, however, was that SBS sites are essentially absent in the Ter MD and largely absent from the MDs that surround the Ter, most notably the Right MD (Figure 51A). Multiple sequence alignments of SlmA proteins show that the region corresponding to the recognition helix is completely conserved among these proteins in Gram-negative bacteria and γ-proteobacteria (Figure 52A-B). This indicates that these proteins all likely bind DNA sites with the same or similar sequences. Hence, we used the Find Individual Motif Occurrence program (FIMO) to map the putative SBS sites on the chromosomes of the uropathogenic *E. coli* strain 536 (GB: CP000247), enterotoxigenic *E. coli* strain E24377A (GB: CP000800), avian pathogenic *E. coli* strain APEC O1 (GB: CP000468), *Salmonella typhimurium* (GB: AE006468), *Klebsiella pneumoniae* (GB: CP000647) and *Enterobacter*
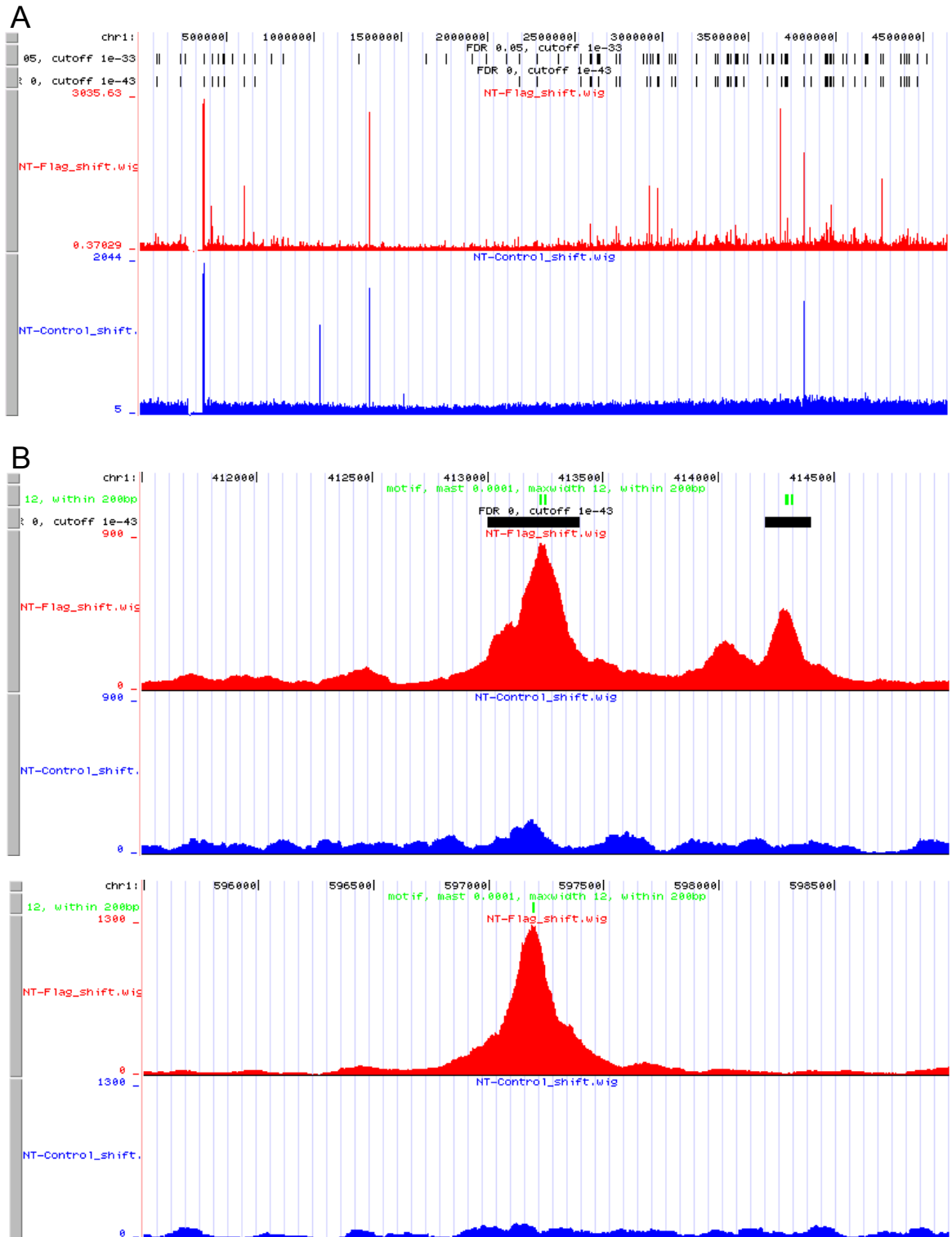
Figure 49. Global view and two representative peaks from ChIP-Seq analysis. A) A global view of ChIP-Seq data. Red bar graph represents tags sequenced from samples containing SlmA bound DNA, and the blue bar graph represents a negative control. The data were analyzed as stated in the supplemental experimental procedure section. **B)** Two representative peaks are shown. The data are represented as in (A). There is a clear peak in the experimental samples (red) in comparison to the control (blue).

A

| Target: | Observed site in Ori MD | Observed site in Left MD | Observed site in Left NS | Observed site in Right MD |

Sample: Input Beads FLAG | Input Beads FLAG | Input Beads FLAG | Input Beads FLAG

Result:

amplification   amplification   amplification   amplification

B

| Target: | random 300 bp region in the Ter MD | random 300 bp region in the Ter MD | random 300 bp region in the Ter MD | random 300 bp region in the Ter MD |

Sample: Input Beads FLAG | Input Beads FLAG | Input Beads FLAG | Input Beads FLAG

Result:

no amplification   no amplification   no amplification   no amplification

Figure 50. Verification of SlmA binding sites on the *E. coli* chromosome. A) ChIP-PCR were conducted on observed sites from the ChIP-Seq analysis. Shown are representative results for sites within the Ori macrodomain, Left macrodomain, Left nonstructured region, and Right nonstructured region. **B)** Ten random region in the Ter macrodomain were tested to identify potential SlmA binding sites. The representative results are shown. Each experiments were done in triplicate.
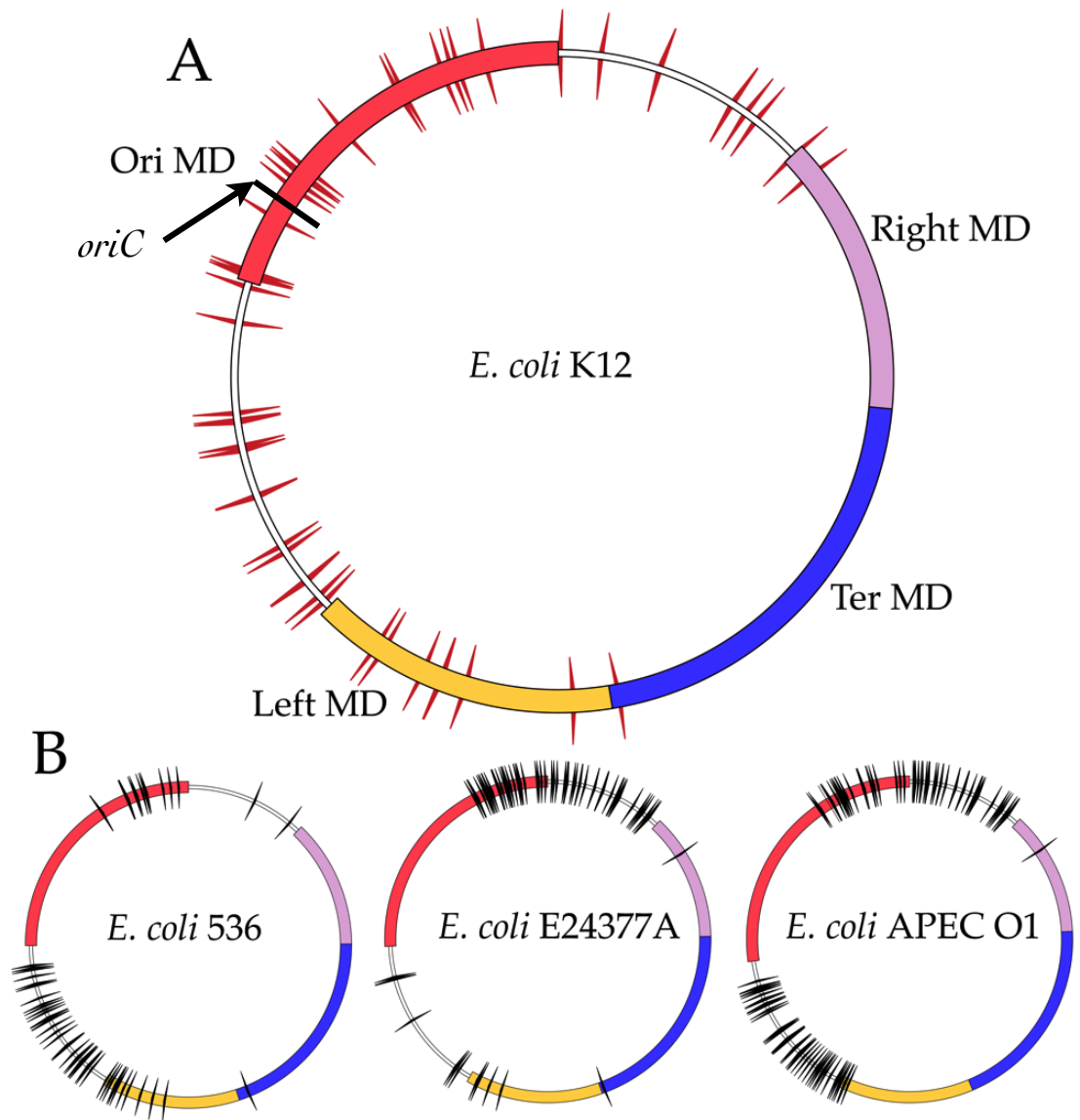
Figure 51. Mapping of putative SlmA binding sites on the *E. coli* chromosome. A) SlmA binding sites as determined by ChIP-Seq are represented as red triangular ticks. The four *E. coli* chromosomal macrodomains: Ori (red), Right (purple), Ter (blue), Left (yellow), are shown as blocks. **B)** The location of SlmA binding sites on the chromosome of three *E. coli* strains: uropathogenic *E. coli* strain 536 (GB: CP000247), enterotoxigenic *E. coli* strain E24377A (GB: CP000800), and avian pathogenic *E. coli* strain APEC O1 (GB: CP000468). The sites are represented as black triangular ticks and the MD are represented as in (A).
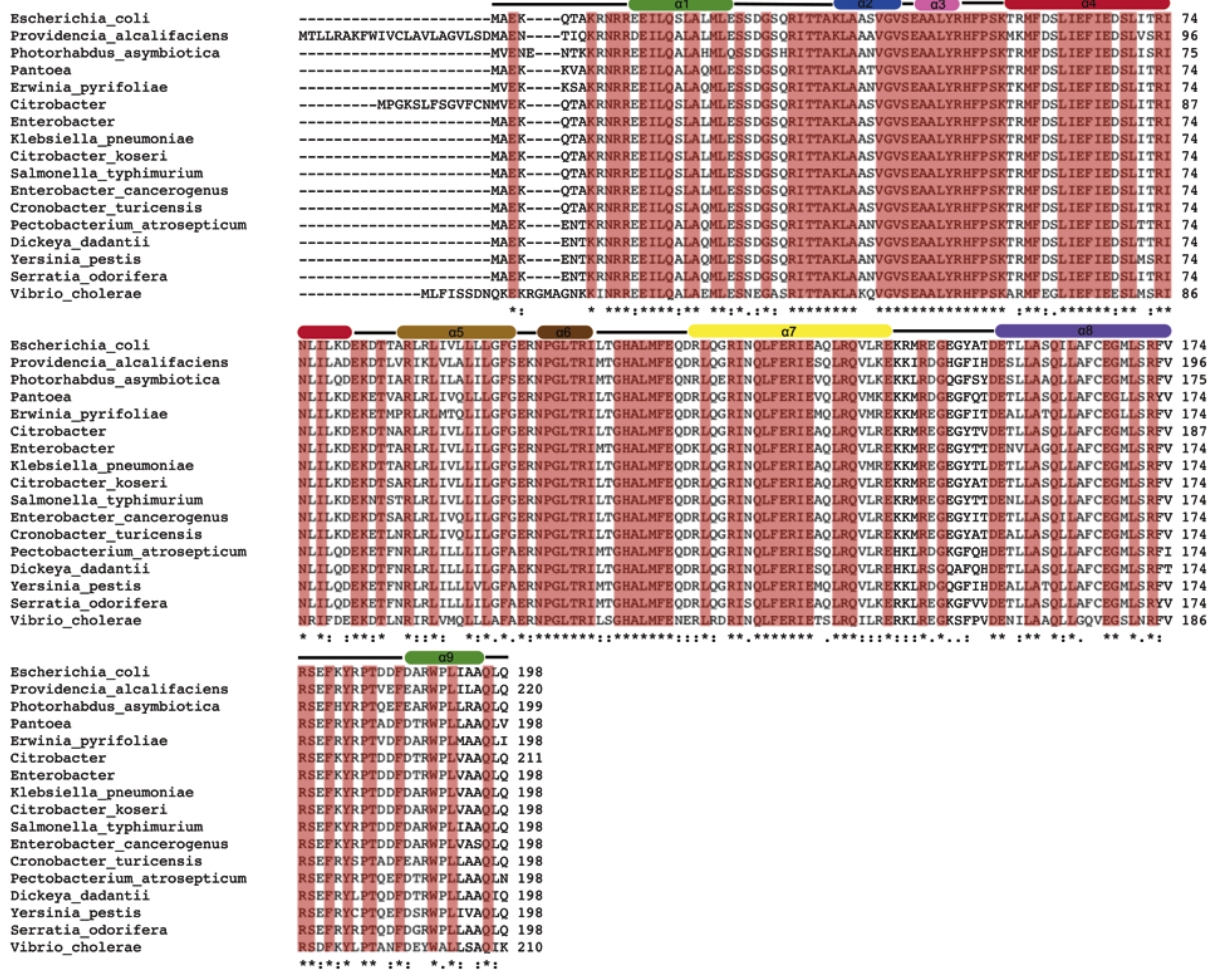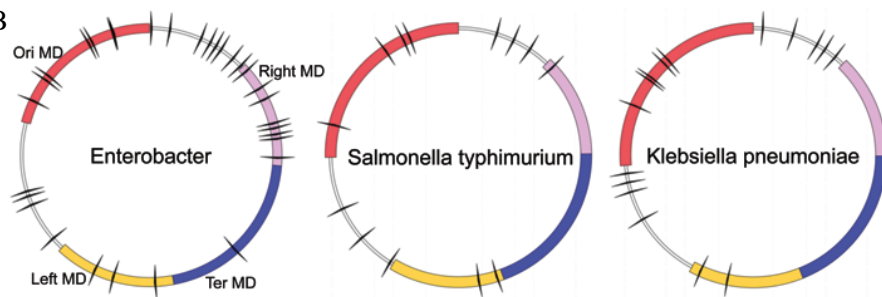
Figure 52. Alignment of the SlmA-like proteins in other γ-proteobacteria, and the predicted SlmA binding sites in *Enterobacter*, *Salmonella typhimurium*, and *Klebsiella pneumoniae*. A) Identical residues are colored red and indicated by an asterisk under the alignment. The secondary structural elements are shown above and colored according to Figure 42A. **B)** The predicted SlmA binding sites for *Enterobacter*, *Salmonella typhimurium*, and *Klebsiella pneumoniae* are represented as chromosome maps. The macro-domains are labeled and the predicted SlmA binding sites are shown are black ticks.

(GB: CP000653) (Bailey and Gribskov, 1998). Strikingly, these analyses showed that, in all cases, the SBS sites are largely dispersed over the chromosomes with the exception of the Ter MD (Figure 51B, 42B).

The finding that SlmA sites are clustered in specific regions of the chromosome was intriguing, as it has been speculated that the formation of specific MDs may play roles in certain cellular processes, key among them, cell division (11). In particular, previous studies have indicated that Z-ring assembly appears to be coordinated with chromosome segregation (129). These studies showed that after replication, the Ori MD abruptly migrates towards the cell poles, with the other MDs following. The Ter MD is the last to migrate and its segregation coincides with the onset of cell division (13). SlmA binding to non-Ter DNA to prevent Z-ring formation at these regions and not the Ter region is consistent with these events. The relative lack of SBS sites in the Right and, to a lesser extent Left, MDs may serve as a buffer to ensure that septation does not occur at the Ter MD. This mechanism may work in concert with FtsK, which pumps DNA to the correct cellular compartments and protects the Ter regions (130). It appears that *B. subtilis* and *Caulobacter crescentus* similarly coordinate DNA segregation with cell division. Specifically, the *B. subtilis* chromosome was also shown to lack binding sites in its Ter MD for its NO factor, Noc although it is still unknown how Noc interacts with the division machinery to inhibit cell division (84). In the case of *C. crescentus*, the MipZ protein interacts with ParB, localizing it near the Ori region, and interferes with Z-ring assembly to restrict its formation to the midcell (85).

Chapter 7. The SlmA-DNA-FtsZ complex

Our data show that the location of SBS sites on the chromosome optimally positions SlmA to act as a negative regulator of cell division. One way in which SlmA could effect such inhibition is via interactions with proteins involved in divisome assembly. Perhaps, most effective would be an interaction with FtsZ, as it initiates cell division. In fact, previous data suggested that SlmA and FtsZ may interact. However, these studies, based on light scattering, implied that polymerization is not inhibited by SlmA, which appeared to be contrary to the mechanism of NO (86). Thus, to investigate whether SlmA interacts with FtsZ and, importantly, whether SlmA can interact with DNA and FtsZ simultaneously, we used FP (105). Similar to previous FP studies, SlmA was titrated into SBS mixtures until saturation was reached. Then, increasing concentrations of FtsZ were added to the same reaction mixture. A clear second binding event was observed upon FtsZ addition (Figure 53A). As a control for molecular crowding, BSA was titrated in the place of FtsZ and revealed no second binding event. Also, when FtsZ was titrated in to a reaction tube with only labeled SBS, there was no appreciable change in polarization, showing that FtsZ alone does not bind the SBS (Figure 53A). The titration curve for the second binding event of FtsZ to the SlmA-DNA complex was used to calculate an apparent K_d of ~120 nM. Notably, the interaction of FtsZ with SlmA-DNA did not require GTP nor was it affected by guanine nucleotides; binding assays performed in the presence of GTP, GTPγS, GDP, and buffer alone yielded apparent affinities of: 142 ± 9 nM, 130 ± 21 nM, 205 ± 5 nM, and 119 ± 11 nM, respectively (Figure 53B).

FtsZ interacts with a number of proteins involved in cell division or its regulation. Most of these interactions have been shown to be mediated by the extended C-terminal tail of FtsZ, including its binding to FtsA and ZipA (22, 55, 131, 132). Interestingly, the TetR protein EthR binds extended ligands such as ethionamide in its C-terminal pocket and TetR itself can bind peptides, which act as tetracycline agonists (100, 101). This suggested that the SlmA C-terminal domain might similarly bind the FtsZ C-tail as although its C-domain pocket appears inaccessible, structural alterations may allow entrance and binding of the FtsZ tail. To test this possibility, a FtsZ truncation mutant, FtsZ(1-360), was used in binding assays with SlmA-DNA. The FP analyses revealed that FtsZ(1-360) bound SlmA-DNA with an apparent affinity that was essentially equal to wild type FtsZ (148 ± 17 nM compared to ~120 nM) (Figure 53B). Thus, these combined data show that SlmA can bind DNA and FtsZ simultaneously and that
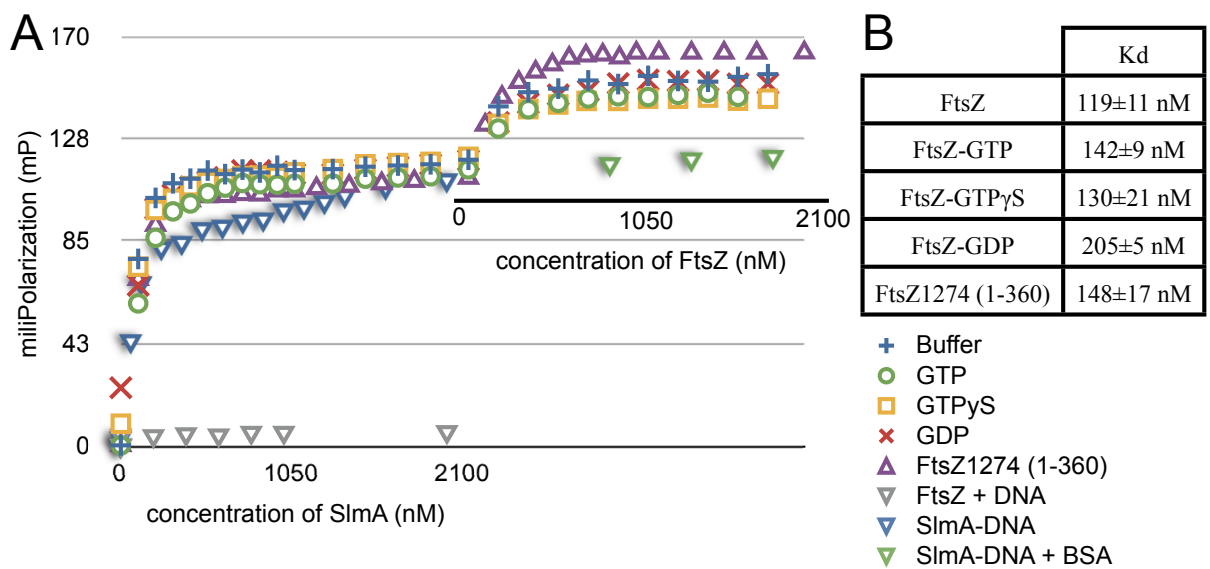
Figure 53. Analyses of FtsZ binding to SlmA-DNA. A) FtsZ binding to SlmA-DNA as measured by FP. SlmA was initially titrated into DNA until saturation and then FtsZ was added to assess binding to the SlmA-DNA complex. These assays were done in different buffer conditions: binding buffer (✚), 1 mM GTP (O), 1 mM GTPγS (□), 1 mM GDP (✖). Additionally, a C-terminal truncation variant of FtsZ (FtsZ(1-360)) was used, represented by Δ. **B)** Table of the binding affinity of SlmA-DNA binding to FtsZ. The second binding curve from FP assays in (A) was used to calculate the apparent K_d.

guanine nucleotides are not required for this interaction. Moreover, SlmA does not interact with the FtsZ C-terminal tail.

Chapter 8. SAXS reveals the structure of the SlmA-FtsZ complex

The finding that SlmA-DNA interacts with FtsZ provides a direct link between an NO factor and the key cell division protein. However, to ascertain how this interaction might lead to NO necessitates a molecular understanding of the SlmA-FtsZ complex and how it may impinge on Z-ring formation. Thus, to gain insight into the molecular interactions between FtsZ with SlmA, we employed small angle X-ray scattering (SAXS) (103). SAXS analyses were carried out on SlmA, FtsZ and the SlmA-FtsZ complex. The SAXS profile and R_G (radius of gyration) of SlmA alone over a concentration range of 1 to 6 mg/ml indicated that the protein is homogeneous (Figure 54A-E). The experimental R_G of SlmA was 31.4 ± 0.01 Å, which agrees well with the value of 28.8 Å, derived from our crystal structure. By contrast, the SAXS profile of FtsZ in the presence or absence of guanine nucleotides shows that it is prone to aggregation, which is expected as FtsZ is known to form protofilaments and other polymer structures (25, 36). Guinier analysis of FtsZ samples at low concentrations, 1 to 2 mg/ml, yielded a R_G of 75.9 ± 0.97 Å, and a rod analysis yielded an R_G of 26.3 ± 0.10 Å for the cross-section. The R_G estimated for the cross-section agrees well with the calculated R_G of 26.2 Å for a FtsZ monomer suggesting that, at this concentration, FtsZ exists largely as protofilament-like structures (Figure 54B).

Interestingly, compared to the behavior of FtsZ alone, SAXS profiles of the SlmA-FtsZ complex, at a concentration range of 1 to 5 mg/ml, revealed it to be aggregation free (Figure 54C). Thus, these data were used to calculate *ab initio* SAXS envelopes for the SlmA-FtsZ complex. Multiple calculations of independent models with the *ab initio* shape determination programs, DAMMIN and GASBOR yielded consistent SAXS envelopes with only small variations between runs (115, 116). The overall shape of the envelope can be described as a symmetric ellipsoid. A homology model of the *E. coli* FtsZ protein along with our atomic model of SlmA were used in the protein-protein docking servers, ClusPro and PatchDock (133-135). These predictions were then used as inputs for the multi-domain modeling program BUNCH (117). The best-fit models from BUNCH (Figure 55) is a structure with a 1:1 SlmA:FtsZ ratio with one SlmA dimer sandwiched between two FtsZ subunits (Figure 56A). The overall fit of the model was quite good except for the presence of unaccounted for density near the FtsZ molecules. However, this unaccounted for portion of the envelope could be explained by the large number of missing residues (residues 317-383) from the FtsZ structure that was used to model the SlmA-FtsZ complex, which contains the full length FtsZ protein.

A SlmA SAXS profile

B FtsZ SAXS profile

C SlmA-FtsZ SAXS profile

D SlmA-FtsZ-GFP SAXS Profile
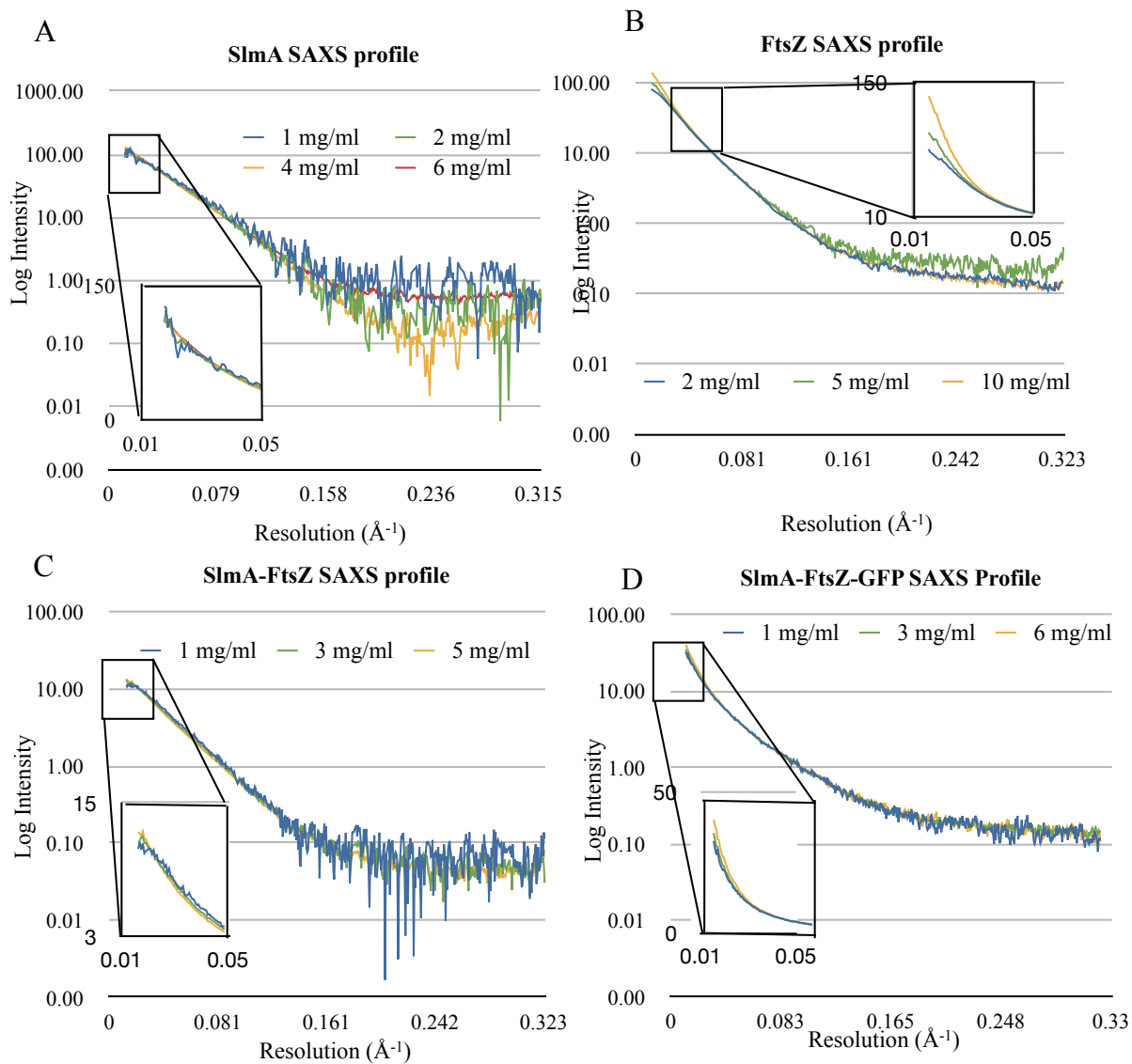
E

	Experimental R_G	Theoretical R_G
SlmA	31.4 ± 0.01 Å	28.79 Å
FtsZ	75.90 ± 0.97 Å*	26.20 Å
	26.30 ± 0.10 Å* (cross-section)	
SlmA-FtsZ	46.00 ± 0.40 Å	45.61 Å
SlmA-FtsZ_GFP	90.18 ± 8.06 Å	

*These values were calculated with data collected from an FtsZ sample at 2 mg/ml

Figure 54. SAXS analyses of SlmA, FtsZ, and the SlmA-FtsZ complex. The SAXS profiles of **A)** SlmA, **B)** FtsZ, **C)** SlmA-FtsZ, and **D)** SlmA-FtsZ-GFP over a range of concentrations display the multimerization behavior of the samples. A more prominent upward bend of the SAXS curves at very low q values (see magnification boxes) with increasing concentration is indicative of concentration-dependent aggregation/polymerization. **E)** The experimental R_G of SlmA, FtsZ, SlmA-FtsZ and the SlmA-FtsZ-GFP complex were calculated from the experimental scattering profile from (A), (B), (C), and (D) respectively. The R_G were approximated from Guinier plots and the cross-section value for FtsZ sample was derived from a rod approximation. All calculations were performed with Primus (Konarev et al, 2003). The theoretical R_G values were calculated by CRYSOL (Svergun et al, 1995) from the available models.
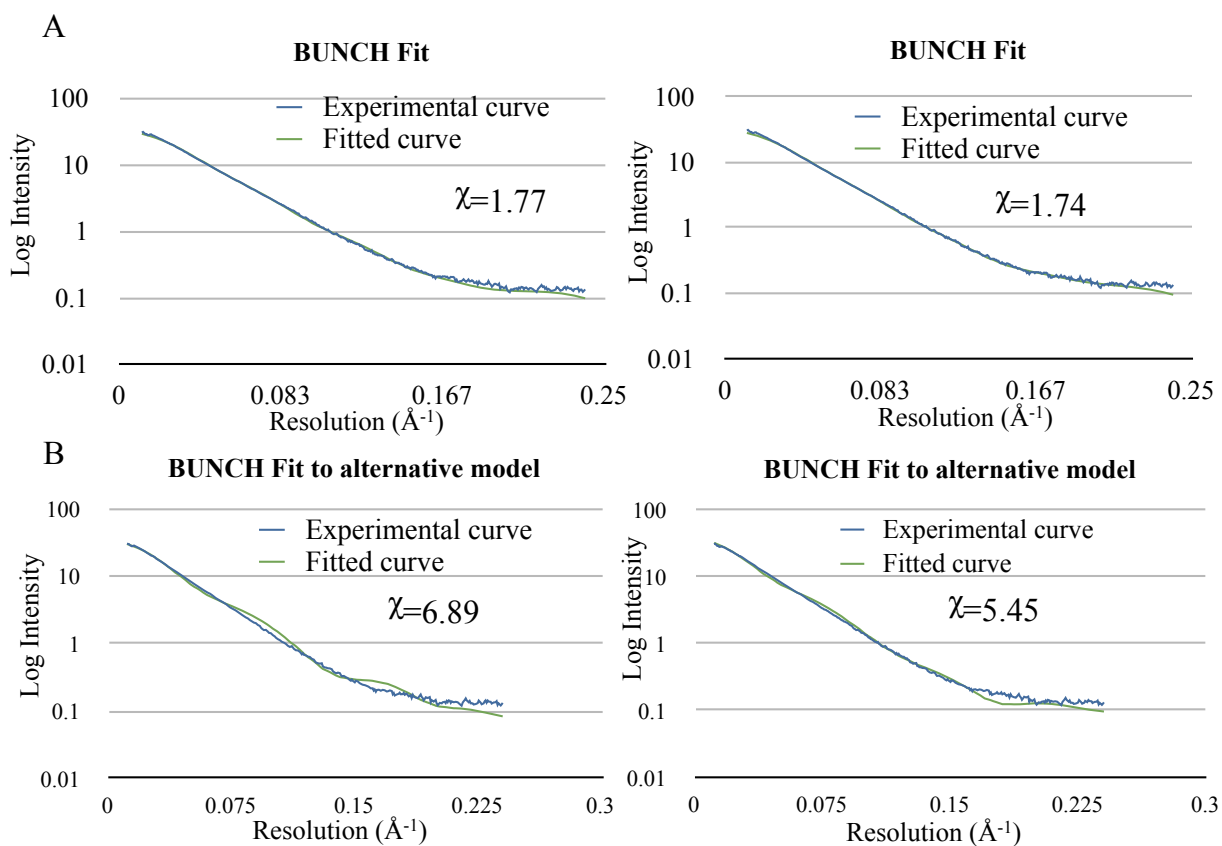
Figure 55. Modeling of the SlmA-FtsZ complex with BUNCH. A) Rigid body docking of SlmA and FtsZ were used as initial inputs for BUNCH. Displayed are two representative results from BUNCH for the final best model of SlmA-FtsZ. BUNCH was run ten times with this docking model, and the results were highly congruent. The experimental curves of the SlmA-FtsZ complex are in green and the BUNCH curves are in blue. The chi values are 1.77 and 1.74 respectively. **B)** The two results shown are from alternative models. The results have higher chi values than those shown in (A). More importantly, visual inspection of the fitted curve conveys the inaccuracy of these models. Models with similar characteristics were disregarded.
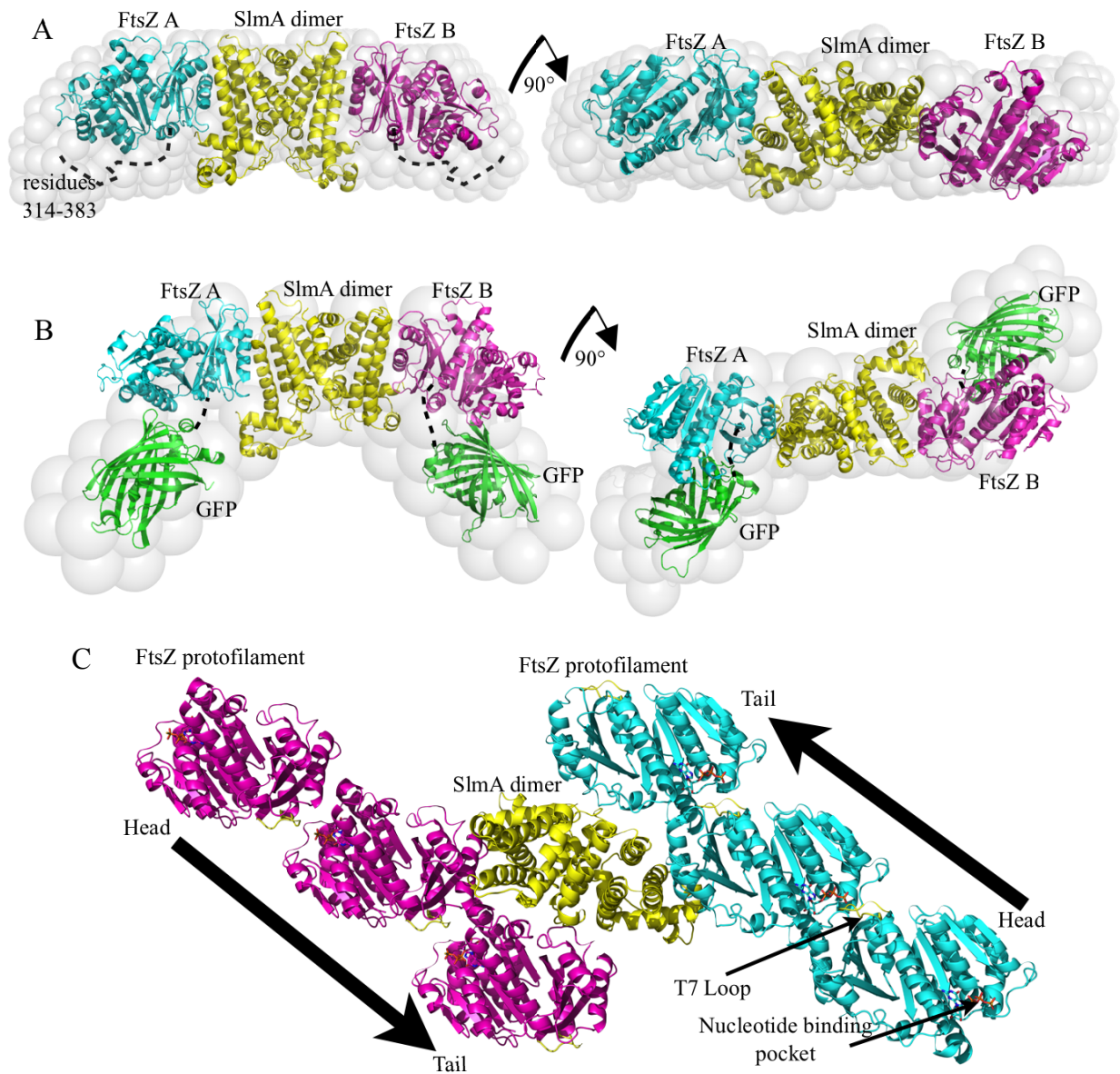
Figure 56. SAXS structure of the SlmA-FtsZ complex. A) The average SlmA-FtsZ envelope, as determined by GASBOR, is displayed as gray spheres (Svergun, 1999; Svergun *et al*., 2001, Kozin and Svergun, 2001). The model of the SlmA-FtsZ complex was calculated by BUNCH (Petoukhov & Svergun, 2005). In this model, a SlmA dimer (yellow) is flanked by two FtsZ molecules (turquoise and magenta). **B)** The SAXS envelope of the SlmA-(FtsZ-GFP) complex was calculated as in (A). Compared to the SlmA-FtsZ envelope (A) additional density was clearly observed which corresponded to the GFP fused at the C-terminus of FtsZ(1-316). **C)** SAXS structure of SlmA-FtsZ showing that when bound to SlmA, FtsZ protofilaments can form but emanate in opposite directions relative to each other. The two FtsZ oligomers (cyan and magenta) in the structure flank the SlmA dimer (yellow).

Examination of the best-fit model shows that the last visible C-terminal residue of FtsZ lies next to this extra density, suggesting how the C-terminal residues may extend into the envelope (Figure 56A). Hence, the model is consistent with our biochemical data showing that the FtsZ C-tail does not bind SlmA. Also consistent with the model is the fact that the calculated R_G of the model, 45.6 Å, compares remarkably well with the experimentally calculated R_G for the complex of 46.0 Å (Figure 54E). While the SlmA dimer can be docked in the envelope, the precise orientation of the FtsZ proteins was more ambiguous due to its spherical shape (Figure 56A). Thus, to obtain additional constraints on the FtsZ orientation in the envelope, SAXS analyses were carried out on a SlmA-FtsZ complex containing a FtsZ fusion protein in which GFP was attached after FtsZ residue 316. The presence of the GFP protein was evident from the calculated SAXS envelope of the complex and confirmed the previously obtained orientation (Figure 56B). The structure shows that SlmA helices α4 and α7, which contain several basic residues, from each subunit interact with helices on the surface exposed face of each FtsZ C-terminal domain, which contain multiple glutamate residues.

Chapter 9. SlmA-SBS complex alters higher order polymer assembly by FtsZ

Studies indicate that subtle changes in FtsZ polymer assembly can result in large changes in Z-ring formation (35). Therefore, the intercalation of SlmA between FtsZ protofilaments, as observed in the SlmA-FtsZ SAXS structure, may be enough to disfavor the formation of a functional Z-ring (Figure 56C). However, the structure reveals a further means by which SlmA binding would affect Z-ring disruption, which is alteration of the assembly of FtsZ polymers. Specifically, in the structure, the SlmA dimer interacts with helices on the surface exposed face of each FtsZ C-terminal domain and not the GTP-binding domain. As a result, the FtsZ GTP-binding pockets and T7 loops, which are required for protofilament formation, remain exposed in the SlmA-FtsZ complex (Figure 57A-B). This suggests that SlmA binding would not prevent the linear polymerization of FtsZ. Indeed, modeling indicates that FtsZ protofilament formation would still be possible when bound to SlmA (Figure 57A-B). Strikingly, examination of the model of SlmA-DNA bound to FtsZ protofilaments shows that when bound to the SlmA-DNA, FtsZ protofilaments would be forced to grow in anti-parallel directions relative to each other. This would prevent the formation of parallel thick filaments, which have been proposed to be involved in FtsZ Z-ring formation (Figure 56C) (Löwe & Amos, 1999; Oliva *et al*, 2003). However, to further address the affect of SlmA on FtsZ protofilament interactions we performed negative stain electron microscopy (EM) experiments on SlmA and its complexes with DNA and FtsZ. As previously observed by others, our EM images show that FtsZ forms filament bundles in the presence of GTP/Mg^{2+} (Erickson *et al*, 1996) (Figure 58A). The addition of SBS DNA, SlmA or SlmA with non SBS DNA had no affect on the appearance of these bundles (Figure 58B-D). By contrast, addition of SlmA and SBS-DNA prevented FtsZ-GTP/Mg^{2+} from forming long bundles and instead led to the creation of ordered helical-like structures, of a fairly uniform size (typical lengths of ~150-200 nm) (Figure 58E-F). The filamentous structures within the helices resemble the FtsZ protofilament bundles but are packed in a side-by-side orientation (Figure 58E-F). Although the resolution restricts a detailed description of the EM structures, the close packing of the two filamentous structures is consistent with the idea that SlmA-DNA enforces an antiparallel arrangement of FtsZ polymers.

In the experiments, the ratio of FtsZ to SlmA used was 5:1 in an effort to establish conditions close to the physiological state. The typical filament bundles formed by FtsZ-GTP (Figure 58A) were never observed in SlmA-DNA-FtsZ samples. Indeed, these samples
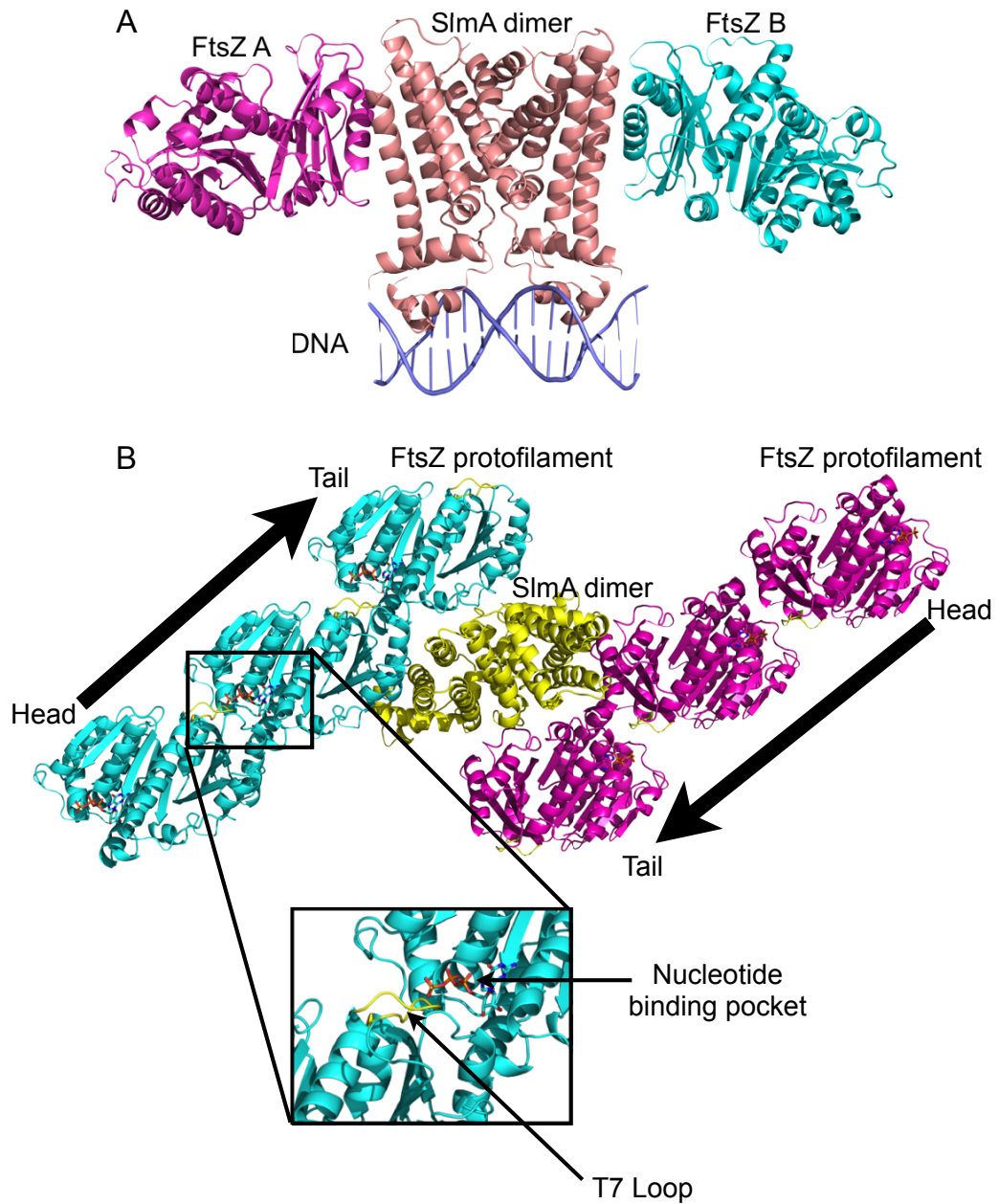
Figure 57. Model of the SlmA-DNA-FtsZ complex; Implications for FtsZ polymerization and simultaneous FtsZ and DNA binding by SlmA. A) Model of the SlmA-FtsZ-DNA complex. The SlmA-FtsZ SAXS structure was used and SlmA HTH elements were inserted into consecutive major grooves as observed in TetR-DNA and QacR-DNA structures. This model shows how SlmA can bind DNA and FtsZ simultaneously. **B)** FtsZ protofilaments form in a head to tail fashion. In the SlmA-FtsZ structure, the FtsZ T7 loops are exposed and can insert into the GTP-binding pocket of an adjacent FtsZ molecule without steric clash. This shows that when in complex with SlmA, FtsZ is still able to form protofilaments.
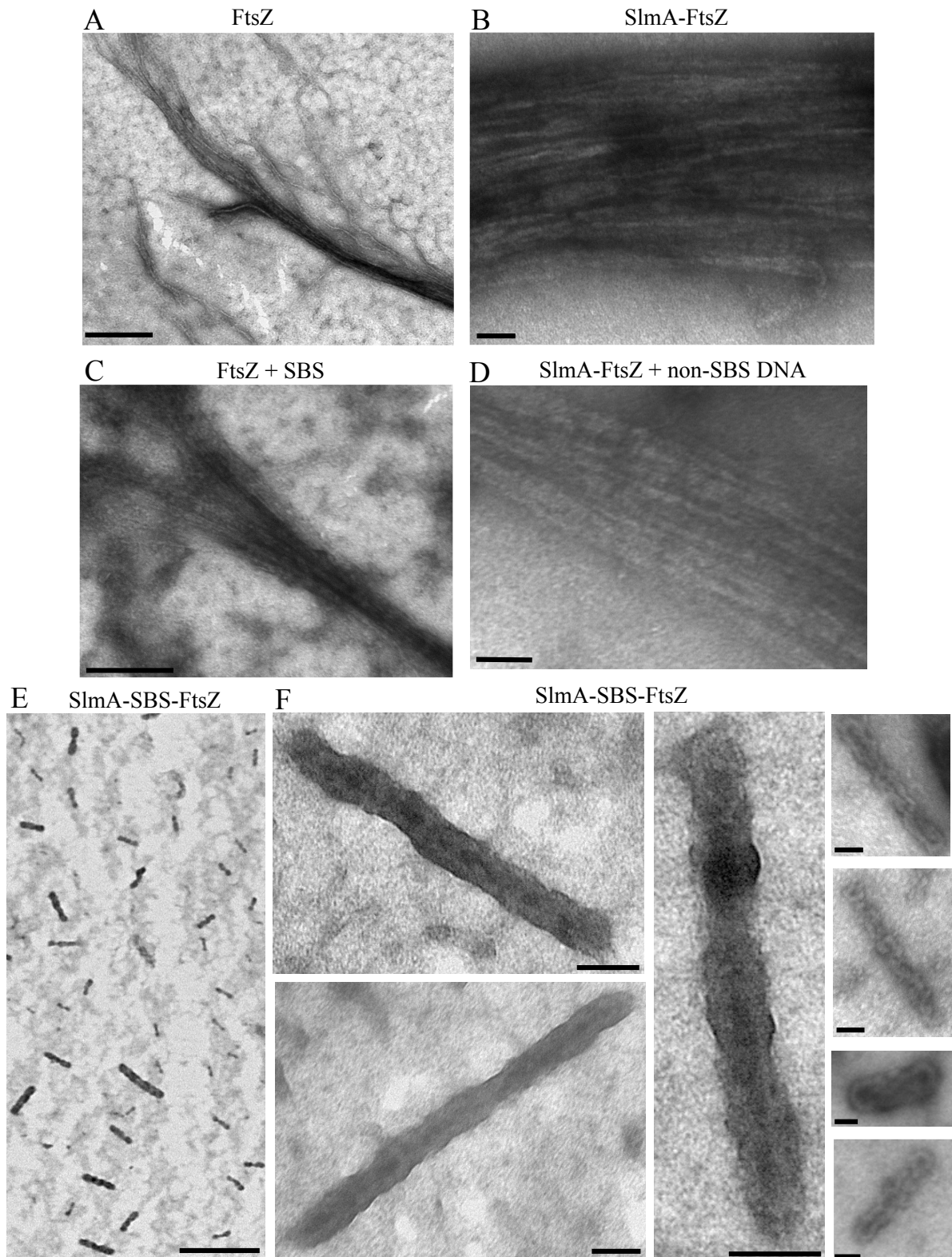
Figure 58. SlmA-DNA complex induces formation of helical FtsZ bundles. A) Negative-stain EM of 3 μM FtsZ under polymerizing condition, causes FtsZ filaments to bundle in to higher-ordered structures. **B)** Negative-stain EM of 3 μM FtsZ, and 0.6 μM SlmA. The bundling of FtsZ protofilaments can be observed. **C)** Negative-stain EM of 3 μM FtsZ, and SBS DNA. **D)** Negative-stain EM of 3 μM FtsZ, 0.6 μM SlmA, and 1 μM non-SBS DNA. **E)** Negative-stain EM of 3 μM FtsZ, 0.6 μM SlmA, and 1 μM SBS DNA. Unusual helical structures are observed throughout the sample. **F)** Higher magnification of helical-like structures seen in (C). These structures range from containing a single to multiple spirals with closed loops. Scale bar represents 500 nm in (A), (C), and (E), and 100 nm in (B), (D), and (F).

consistently showed only the uniform helical structures as shown in Figure 58E-F. This suggests that a small amount of SlmA-DNA is sufficient to inhibit the formation of functional FtsZ bundles and further indicates that SlmA-DNA act as nucleation sites to promote the growth of a non-functional FtsZ helices, which can propagate several hundred nm. Interestingly, SlmA must be bound to SBS DNA to impart this effect as EM samples with FtsZ and SlmA alone or SlmA and non SBS DNA failed to affect FtsZ polymer assembly. Because our SAXS structure was obtained using a 1:1 ratio of SlmA to FtsZ, it cannot address how the SlmA DNA binding domain may impact the polymerization properties of a growing FtsZ protofilament attached to SlmA-DNA. Like other TetR proteins, the SlmA DNA binding domains are flexible and likely only become fixed upon cognate DNA binding. It seems probable that the precise orientation of the DNA bound form of the SlmA DNA binding domains and the DNA itself may be necessary in steering the growing FtsZ protofilaments into the specific helical structures we observe. The inability of SlmA alone to affect FtsZ polymer assembly could also function as a failsafe measure to prevent unwanted perturbation of cytosolic FtsZ polymers where Z-ring assembly is desired. However, it is likely that there is little SlmA present in the cytosol. In fact, previous studies showed that SlmA is localized entirely within the nucleoid fraction of the cell (86). Moreover, data suggest that DNA binding proteins that are not bound to their cognate sites interact non-specifically and slide along the DNA or are engaged in rapid dissociation/reassociation from/onto DNA (136, 137). Thus, the DNA bound form of SlmA is the physiologically relevant form.

Chapter 10. Conclusions on the molecular model for SlmA mediated NO

Our combined data provide a molecular mechanism for SlmA mediated NO. First, ChIP-seq analyses revealed that the SlmA binding sites are dispersed on non-Ter regions of the chromosome. SlmA has been shown to be present at a constant concentration of ~400 molecules per cell, which is enough to bind all the identified sites on the chromosome (86). The fact that the Ter MD regions are not bound by SlmA and therefore do not exhibit NO is consistent with the finding that the Z-ring formation occurs nearly concomitantly with replication of the Ter region (13). In this regard, the ability of SlmA to bind DNA and FtsZ simultaneously is crucial for NO as it localizes SlmA specifically to the non-Ter MD. Consistent with this mechanism, FP studies demonstrated that SlmA can bind DNA and FtsZ at the same time and our SlmA-FtsZ structure shows that the SlmA N-terminal HTH domains do not interact with FtsZ and, thus can bind DNA simultaneously (Figure 57A). Studies indicate that subtle changes in FtsZ polymer assembly can result in large changes in Z-ring formation (35). Therefore, the intercalation of SlmA between FtsZ protofilaments, as observed in the SlmA-FtsZ structure, may be enough to disfavor the formation of a functional Z-ring (Figure 56C). Specifically, in the SAXS structure, the SlmA dimer interacts with helices on the surface exposed face of each FtsZ C-terminal domain and not the GTP-binding domain (Figure 57B). As a result, the FtsZ GTP-binding pockets and T7 loops, which are required for protofilament formation, remain exposed in the SlmA-FtsZ complex (Figure 57B). This suggests that SlmA binding would not prevent the linear polymerization of FtsZ. Indeed, modeling shows that FtsZ protofilament formation is still possible when bound to SlmA.

Strikingly, examination of the structure shows that the FtsZ molecules bound to SlmA are oriented in opposite directions. As a result, when bound to the SlmA dimer, FtsZ protofilaments would be forced to grow in anti-parallel directions relative to each other, which would prevent the formation of parallel thick filaments (Figure 58E-F). This is also consistent with and suggests an explanation for previous data showing that FtsZ can form larger polymers when bound to SlmA, but that the polymers appear to be distinct from those normally formed by FtsZ (86). Intercalation between FtsZ protofilaments and prevention of proper lateral interactions for Z-ring construction provide a means by which SlmA-DNA impede Z-ring formation. The effects can be observed from our EM experiments demonstrating that SlmA-DNA severely affects the higher order assembly of FtsZ filaments. Thus, unlike other FtsZ regulators that inhibit Z-ring formation by preventing polymerization, SlmA derails Z-ring

formation by preventing the correct formation and bundling of FtsZ protofilaments, instead the SlmA-DNA complex can induce the growth of nonfunctional helical FtsZ structures (Figure 59). The sequestration of FtsZ molecules by SlmA could also play a role in Z-ring inhibition and combined these mechanisms would provide multiple levels of protection against nucleoid bisection. Given the high conservation of SlmA in Gram-negative bacteria, we propose that this NO mechanism is likely utilized by all bacteria that harbor a SlmA protein.
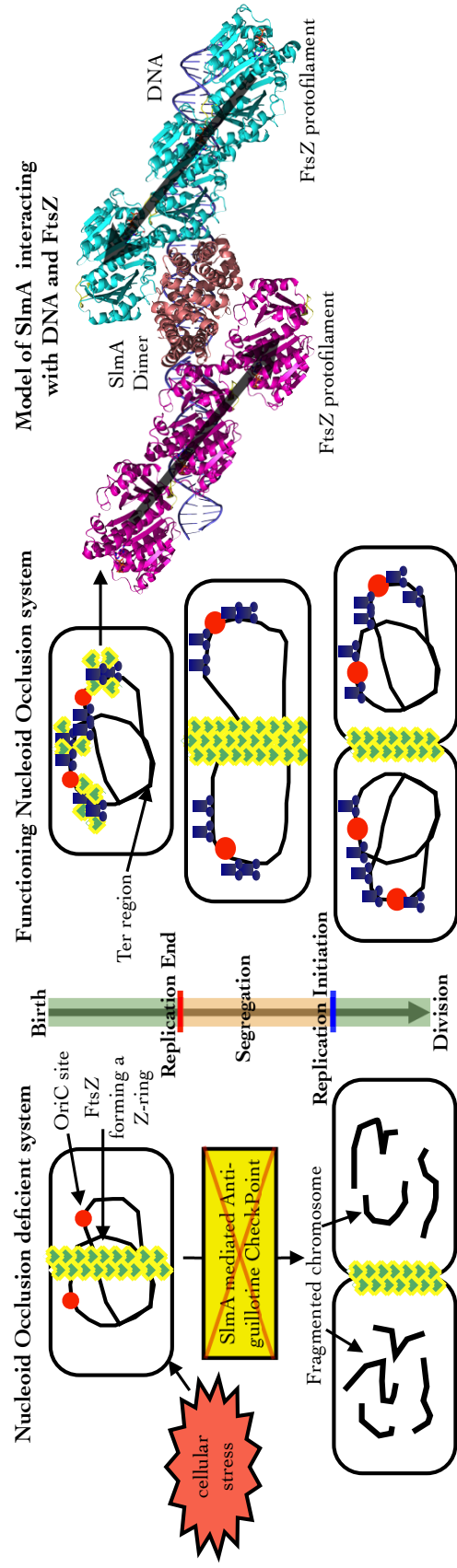
Figure 59. Model for SlmA-mediated Nucleoid Occlusion. Schematic representation of SlmA-mediated NO. OriC sites, opposite the Ter region, are shown as red circles. FtsZ molecules are green arrows with yellow outline. Left shows that without functional SlmA, FtsZ polymerizes to form a premature Z-ring, and in the presence of cellular stress caused by the inhibition of chromosome partition, DNA replication, or a non-functional Min system, this leads to nucleoid guillotining and DNA fragmentation. Middle, in the presence of SlmA (blue), the Z-ring cannot form over the DNA until replication of the Ter MD because SlmA does not bind the Ter MD. Cell division is thus coordinated with DNA segregation as the cell division machinery assembles right before the commencement of DNA segregation. Close up of the SlmA-FtsZ interaction, illustrating that SlmA disrupts Z-ring formation by reorienting the FtsZ molecules and by inserting between FtsZ protofilaments in an antiparallel manner.

Chapter 11. Future directions

The insights drawn from these studies suggest that in order to gain a deeper understanding of the molecular mechanism, the formation of the SlmA-SBS-FtsZ spiral structures needs to be investigated further. There are three major point of interest that arises. First, what are the structural changes occur during SlmA-DNA binding that can dramatically affect FtsZ filament interaction and form the spiral structures? The data show that FtsZ does not associate with SBS DNA. Furthermore, SlmA is able to bind to FtsZ and DNA independently. This indicates that the DNA and FtsZ binding interfaces are independent of one another and are not likely to be altered during the formation of the SlmA-SBS-FtsZ complex. However, SlmA requires SBS DNA to be present in order to constrain FtsZ filaments to form the spiral structures. Therefore a structural understanding of SlmA in comparison to the SlmA-SBS complex will reveal the necessity of DNA in the nucleoid occlusion process.

Another question that needs to be addressed is: how are the SlmA, FtsZ, and DNA molecules organized in the spiral structure? The spiral structures are all formed with regularity and consistency, and displays two interesting structural features. The FtsZ filament is able to curve and form a spiral that crossover itself. Secondly, the FtsZ filament displays a striking turn at the ends of the spiral structure to form a closed loop. This suggests that SlmA-SBS must bind to FtsZ in a specific location and orientation in the spiral structure. Further investigation can aid to identify the position of each type of molecules in the structure. This information will shed light on where the SlmA-SBS complex is critically place in order to constrain the FtsZ filament to form the spiral structure.

The crucial role SlmA plays as an effector of nucleoid occlusion makes it an inviting target for the development antibacterial therapeutic. Thus the third unexplored point of interest is the therapeutic potential of a SlmA inhibitor molecule. Our data clearly demonstrate that SlmA-DNA binding is a necessary interaction for nucleoid occlusion. Fascinatingly, other characterized TetR family members have the capability to bind various inducer molecules, and after doing so, the protein is unable to bind to DNA. Although the putative binding pocket of SlmA is observed to be relatively small and inaccessible, other TetR family members have shown that the binding pocket has great plasticity can accommodate a variety of molecules (refer to section 1.6.5). This opens the opportunity to develop a molecule that can inhibit the ability of SlmA to bind DNA, which can potentially be exploited as an antibiotic.

Bibliography

1. Paoletti, A., and F. Chang. 2000. Analysis of mid1p, a protein required for placement of the cell division site, reveals a link between the nucleus and the cell surface in fission yeast. Molecular Biology of the Cell 11:2757-2773.

2. Daga, R. R., and F. Chang. 2005. Dynamic positioning of the fission yeast cell division plane. Proceedings of the National Academy of Sciences, USA 102:8228-8232.

3. Tolic-Nørrelykke, I. M., L. Sacconi, C. Stringari, I. Raabe, and F. S. Pavone. 2005. Nuclear and division-plane positioning revealed by optical micromanipulation. Current Biology 15:1212-1216.

4. Motegi, F., M. Mishra, M. K. Balasubramanian, and I. Mabuchi. 2004. Myosin-II reorganization during mitosis is controlled temporally by its dephosphorylation and spatially by Mid1 in fission yeast. Journal of Cell Biology 165:685-695.

5. Oliferenko, S., T. G. Chew, and M. K. Balasubramanian. 2009. Positioning cytokinesis. Genes and Development 23:660-674.

6. Goehring, N. W., and J. Beckwith. 2005. Diverse paths to midcell: assembly of the bacterial cell division machinery. Current Biology 15:R514-526.

7. Margolin, W. 2005. FtsZ and the division of prokaryotic cells and organelles. Nature Reviews: Molecular Cell Biology 6:862-871.

8. Rothfield, L., A. Taghbalout, and Y.-L. Shih. 2005. Spatial control of bacterial division-site placement. Nature Reviews: Microbiology 3:959-968.

9. Weiss, D. S. 2004. Bacterial cell division and the septal ring. Molecular Microbiology 54:588-597.

10. Niki, H., Y. Yamaichi, and S. Hiraga. 2000. Dynamic organization of chromosomal DNA in *Escherichia coli*. Genes and Development 14:212-223.

11. Valens, M., S. Penaud, M. Rossignol, F. Cornet, and F. Boccard. 2004. Macrodomain organization of the *Escherichia coli* chromosome. EMBO Journal 23:4330-4341.

12. Mercier, R., M.-A. Petit, S. Schbath, S. Robin, M. El Karoui, F. Boccard, and O. Espéli. 2008. The MatP/matS site-specific system organizes the terminus region of the *E. coli* chromosome into a macrodomain. Cell 135:475-485.

13. Espeli, O., R. Mercier, and F. Boccard. 2008. DNA dynamics vary according to macrodomain topography in the *E. coli* chromosome. Molecular Microbiology 68:1418-1427.

14. Bi, E. F., and J. Lutkenhaus. 1991. FtsZ ring structure associated with division in *Escherichia coli*. Nature 354:161-164.

15. Pichoff, S., and J. Lutkenhaus. 2002. Unique and overlapping roles for ZipA and FtsA in septal ring assembly in *Escherichia coli*. EMBO Journal 21:685-693.

16. Pichoff, S., and J. Lutkenhaus. 2005. Tethering the Z ring to the membrane through a conserved membrane targeting sequence in FtsA. Molecular Microbiology 55:1722-1734.

17. Rothfield, L., S. Justice, and J. García-Lara. 1999. Bacterial cell division. Annual Review of Genetics 33:423-448.

18. Hale, C. A., and P. A. de Boer. 1997. Direct binding of FtsZ to ZipA, an essential component of the septal ring structure that mediates cell division in *E. coli*. Cell 88:175-185.

19. Praefcke, G. J. K., and H. T. McMahon. 2004. The dynamin superfamily: universal membrane tubulation and fission molecules? Nature Reviews: Molecular Cell Biology 5:133-147.

20. Margolin, W. 2000. Themes and variations in prokaryotic cell division. FEMS Microbiology Reviews 24:531-548.

21. Oliva, M. A., D. Trambaiolo, and J. Löwe. 2007. Structural insights into the conformational variability of FtsZ. Journal of Molecular Biology 373:1229-1242.

22. Ma, X., and W. Margolin. 1999. Genetic and functional analyses of the conserved C-terminal core domain of Escherichia coli FtsZ. Journal of Bacteriology 181:7531-7544.

23. Löwe, J., and L. A. Amos. 1998. Crystal structure of the bacterial cell-division protein FtsZ. Nature 391:203-206.

24. Lu, C., M. Reedy, and H. P. Erickson. 2000. Straight and curved conformations of FtsZ are regulated by GTP hydrolysis. Journal of Bacteriology 182:164-170.

25. Erickson, H. P., D. W. Taylor, K. A. Taylor, and D. Bramhill. 1996. Bacterial cell division protein FtsZ assembles into protofilament sheets and minirings, structural homologs of tubulin polymers. Proceedings of the National Academy of Sciences, USA 93:519-523.

26. Yu, X. C., and W. Margolin. 1997. Ca2+-mediated GTP-dependent dynamic assembly of bacterial cell division protein FtsZ into asters and polymer networks in vitro. EMBO Journal 16:5455-5463.

27. González, J. M., M. Jiménez, M. Vélez, J. Mingorance, J. M. Andreu, M. Vicente, and G. Rivas. 2003. Essential cell division protein FtsZ assembles into one monomer-thick ribbons under conditions resembling the crowded intracellular environment. Journal of Biological Chemistry 278:37664-37671.

28. RayChaudhuri, D. 1999. ZipA is a MAP-Tau homolog and is essential for structural integrity of the cytokinetic FtsZ ring during bacterial cell division. EMBO Journal 18:2372-2383.

29. Hale, C. A., A. C. Rhee, and P. A. de Boer. 2000. ZipA-induced bundling of FtsZ polymers mediated by an interaction between C-terminal domains. Journal of Bacteriology 182:5153-5166.

30. Gueiros-Filho, F. J., and R. Losick. 2002. A widely conserved bacterial cell division protein that promotes assembly of the tubulin-like protein FtsZ. Genes and Development 16:2544-2556.

31. Marrington, R., E. Small, A. Rodger, T. R. Dafforn, and S. G. Addinall. 2004. FtsZ fiber bundling is triggered by a conformational change in bound GTP. Journal of Biological Chemistry 279:48821-48829.

32. Addinall, S. G., C. Cao, and J. Lutkenhaus. 1997. Temperature shift experiments with an ftsZ84(Ts) strain reveal rapid dynamics of FtsZ localization and indicate that the Z ring is required throughout septation and cannot reoccupy division sites once constriction has initiated. Journal of Bacteriology 179:4277-4284.

33. Rueda, S., M. Vicente, and J. Mingorance. 2003. Concentration and assembly of the division ring proteins FtsZ, FtsA, and ZipA during the Escherichia coli cell cycle. Journal of Bacteriology 185:3344-3351.

34. Anderson, D. E., F. J. Gueiros-Filho, and H. P. Erickson. 2004. Assembly dynamics of FtsZ rings in Bacillus subtilis and *Escherichia coli* and effects of FtsZ-regulating proteins. Journal of Bacteriology 186:5775-5781.

35. Romberg, L., and P. A. Levin. 2003. Assembly dynamics of the bacterial cell division protein FTSZ: poised at the edge of stability. Annual Review of Microbiology 57:125-154.

36. Adams, D. W., and J. Errington. 2009. Bacterial cell division: assembly, maintenance and disassembly of the Z ring. Nature Reviews: Microbiology 7:642-653.

37. Caplan, M. R., and H. P. Erickson. 2003. Apparent cooperative assembly of the bacterial cell division protein FtsZ demonstrated by isothermal titration calorimetry. Journal of Biological Chemistry 278:13784-13788.

38. Li, Z., M. J. Trimble, Y. V. Brun, and G. J. Jensen. 2007. The structure of FtsZ filaments in vivo suggests a force-generating role in cell division. EMBO Journal 26:4694-4708.

39. Fu, G., T. Huang, J. Buss, C. Coltharp, Z. Hensel, and J. Xiao. 2010. *In vivo* structure of the *E. coli* FtsZ-ring revealed by photoactivated localization microscopy (PALM). PLoS ONE 5:e12682.

40. Chen, Y., and H. P. Erickson. 2005. Rapid in vitro assembly dynamics and subunit turnover of FtsZ demonstrated by fluorescence resonance energy transfer. Journal of Biological Chemistry 280:22549-22554.

41. Stricker, J., P. Maddox, E. D. Salmon, and H. P. Erickson. 2002. Rapid assembly dynamics of the *Escherichia coli* FtsZ-ring demonstrated by fluorescence recovery after photobleaching. Proceedings of the National Academy of Sciences, USA 99:3171-3175.

42. Romberg, L., M. Simon, and H. P. Erickson. 2001. Polymerization of Ftsz, a bacterial homolog of tubulin. is assembly cooperative? Journal of Biological Chemistry 276:11743-11753.

43. Popp, D., M. Iwasa, A. Narita, H. P. Erickson, and Y. Maéda. 2009. FtsZ condensates: an in vitro electron microscopy study. Biopolymers 91:340-350.

44. Kubitschek, H. E. 1986. Increase in cell mass during the division cycle of *Escherichia coli* B/rA. Journal of Bacteriology 168:613-618.

45. Reshes, G., S. Vanounou, I. Fishov, and M. Feingold. 2008. Cell shape dynamics in *Escherichia coli*. Biophysical Journal 94:251-264.

46. Mingorance, J., M. Tadros, M. Vicente, J. M. González, G. Rivas, and M. Vélez. 2005. Visualization of single *Escherichia coli* FtsZ filament dynamics with atomic force microscopy. Journal of Biological Chemistry 280:20909-20914.

47. Chen, Y., and H. P. Erickson. 2009. FtsZ filament dynamics at steady state: subunit exchange with and without nucleotide hydrolysis. Biochemistry 48:6664-6673.

48. Löwe, J., and L. A. Amos. 2000. Helical tubes of FtsZ from *Methanococcus jannaschii*. Biological Chemistry 381:993-999.

49. Löwe, J., and L. A. Amos. 1999. Tubulin-like protofilaments in Ca2+-induced FtsZ sheets. EMBO Journal 18:2364-2371.

50. Oliva, M. A., S. Huecas, J. M. Palacios, J. Martín-Benito, J. M. Valpuesta, and J. M. Andreu. 2003. Assembly of archaeal cell division protein FtsZ and a GTPase-inactive mutant into double-stranded filaments. Journal of Biological Chemistry 278:33562-33570.

51. Oliva, M. A., S. C. Cordell, and J. Löwe. 2004. Structural insights into FtsZ protofilament formation. Nature Structural & Molecular Biology 11:1243-1250.

52. Errington, J., R. A. Daniel, and D.-J. Scheffers. 2003. Cytokinesis in bacteria. Microbiology and Molecular Biology Reviews 67:52-65, table of contents.

53. Erickson, H. P. 2001. The FtsZ protofilament and attachment of ZipA--structural constraints on the FtsZ power stroke. Current Opinion in Cell Biology 13:55-60.

54. Hale, C. A., and P. A. de Boer. 1999. Recruitment of ZipA to the septal ring of Escherichia coli is dependent on FtsZ and independent of FtsA. Journal of Bacteriology 181:167-176.

55. Mosyak, L., Y. Zhang, E. Glasfeld, S. Haney, M. Stahl, J. Seehra, and W. S. Somers. 2000. The bacterial cell-division protein ZipA and its interaction with an FtsZ fragment revealed by X-ray crystallography. EMBO Journal 19:3179-3191.

56. Löwe, J., F. van den Ent, and L. A. Amos. 2004. Molecules of the bacterial cytoskeleton. Annual Review of Biophysics and Biomolecular Structure 33:177-198.

57. van den Ent, F., and J. Löwe. 2000. Crystal structure of the cell division protein FtsA from *Thermotoga maritima*. EMBO Journal 19:5300-5307.

58. Corbin, B. D., B. Geissler, M. Sadasivam, and W. Margolin. 2004. Z-ring-independent interaction between a subdomain of FtsA and late septation proteins as revealed by a polar recruitment assay. Journal of Bacteriology 186:7736-7744.

59. Rico, A. I., M. García-Ovalle, J. Mingorance, and M. Vicente. 2004. Role of two essential domains of *Escherichia coli* FtsA in localization and progression of the division ring. Molecular Microbiology 53:1359-1371.

60. Pichoff, S., and J. Lutkenhaus. 2007. Identification of a region of FtsA required for interaction with FtsZ. Molecular Microbiology 64:1129-1138.

61. Geissler, B., D. Elraheb, and W. Margolin. 2003. A gain-of-function mutation in ftsA bypasses the requirement for the essential cell division gene zipA in *Escherichia coli*. Proceedings of the National Academy of Sciences, USA 100:4197-4202.

62. Geissler, B., D. Shiomi, and W. Margolin. 2007. The *ftsA** gain-of-function allele of *Escherichia coli* and its effects on the stability and dynamics of the Z ring. Microbiology 153:814-825.

63. Small, E., R. Marrington, A. Rodger, D. J. Scott, K. Sloan, D. Roper, T. R. Dafforn, and S. G. Addinall. 2007. FtsZ polymer-bundling by the Escherichia coli ZapA orthologue, YgfE, involves a conformational change in bound GTP. Journal of Molecular Biology 369:210-221.

64. Low, H. H., M. C. Moncrieffe, and J. Löwe. 2004. The crystal structure of ZapA and its modulation of FtsZ polymerisation. Journal of Molecular Biology 341:839-852.

65. Huisman, O., and R. D'Ari. 1981. An inducible DNA replication-cell division coupling mechanism in *E. coli*. Nature 290:797-799.

66. Huisman, O., R. D'Ari, and S. Gottesman. 1984. Cell-division control in *Escherichia coli*: specific induction of the SOS function SfiA protein is sufficient to block septation. Proceedings of the National Academy of Sciences, USA 81:4490-4494.

67. Dajkovic, A., A. Mukherjee, and J. Lutkenhaus. 2008. Investigation of regulation of FtsZ assembly by SulA and development of a model for FtsZ polymerization. Journal of Bacteriology 190:2513-2526.

68. Cordell, S. C., E. J. H. Robinson, and J. Lowe. 2003. Crystal structure of the SOS cell division inhibitor SulA and in complex with FtsZ. Proceedings of the National Academy of Sciences, USA 100:7889-7894.

69. Hu, Z., A. Mukherjee, S. Pichoff, and J. Lutkenhaus. 1999. The MinC component of the division site selection system in *Escherichia coli* interacts with FtsZ to prevent polymerization. Proceedings of the National Academy of Sciences, USA 96:14819-14824.

70. Raskin, D. M., and P. A. de Boer. 1997. The MinE ring: an FtsZ-independent cell structure required for selection of the correct division site in *E. coli*. Cell 91:685-694.

71. Raskin, D. M., and P. A. de Boer. 1999. Rapid pole-to-pole oscillation of a protein required for directing division to the middle of *Escherichia coli*. Proceedings of the National Academy of Sciences, USA 96:4971-4976.

72. Raskin, D. M., and P. A. de Boer. 1999. MinDE-dependent pole-to-pole oscillation of division inhibitor MinC in Escherichia coli. Journal of Bacteriology 181:6419-6424.

73. Cordell, S. C., R. E. Anderson, and J. Löwe. 2001. Crystal structure of the bacterial cell division inhibitor MinC. EMBO Journal 20:2454-2461.

74. Hu, Z., and J. Lutkenhaus. 2000. Analysis of MinC reveals two independent domains involved in interaction with MinD and FtsZ. Journal of Bacteriology 182:3965-3971.

75. Justice, S. S., J. García-Lara, and L. I. Rothfield. 2000. Cell division inhibitors SulA and MinC/MinD block septum formation at different steps in the assembly of the *Escherichia coli* division machinery. Molecular Microbiology 37:410-423.

76. Cordell, S. C., and J. Löwe. 2001. Crystal structure of the bacterial cell division regulator MinD. FEBS Letters 492:160-165.

77. Hayashi, I., T. Oyama, and K. Morikawa. 2001. Structural and functional studies of MinD ATPase: implications for the molecular recognition of the bacterial cell division apparatus. EMBO Journal 20:1819-1828.

78. Mulder, E., and C. L. Woldringh. 1989. Actively replicating nucleoids influence positioning of division sites in Escherichia coli filaments forming cells lacking DNA. Journal of Bacteriology 171:4303-4314.

79. Woldringh, C. L. 2002. The role of co-transcriptional translation and protein translocation (transertion) in bacterial chromosome segregation. Molecular Microbiology 45:17-29.

80. Woldringh, C. L., E. Mulder, J. A. Valkenburg, F. B. Wientjes, A. Zaritsky, and N. Nanninga. 1990. Role of the nucleoid in the toporegulation of division. Research in Microbiology 141:39-49.

81. Woldringh, C. L., E. Mulder, P. G. Huls, and N. Vischer. 1991. Toporegulation of bacterial division according to the nucleoid occlusion model. Research in Microbiology 142:309-320.

82. Wu, L. J., and J. Errington. 2004. Coordination of cell division and chromosome segregation by a nucleoid occlusion protein in *Bacillus subtilis*. Cell 117:915-925.

83. Sievers, J., B. Raether, M. Perego, and J. Errington. 2002. Characterization of the parB-like yyaA gene of Bacillus subtilis. Journal of Bacteriology 184:1102-1111.

84. Wu, L., S. Ishikawa, Y. Kawai, T. Oshima, N. Ogasawara, and J. Errington. 2009. Noc protein binds to specific DNA sequences to coordinate cell division with chromosome segregation. EMBO Journal.

85. Thanbichler, M., and L. Shapiro. 2006. MipZ, a spatial regulator coordinating chromosome segregation with cell division in *Caulobacter*. Cell 126:147-162.

86.   Bernhardt, T. G., and P. A. J. de Boer. 2005. SlmA, a nucleoid-associated, FtsZ binding protein required for blocking septal ring assembly over Chromosomes in *E. coli*. Molecular Cell 18:555-564.

87.   Ramos, J. L., M. Martínez-Bueno, A. J. Molina-Henares, W. Terán, K. Watanabe, X. Zhang, M. T. Gallegos, R. Brennan, and R. Tobes. 2005. The TetR family of transcriptional repressors. Microbiology and Molecular Biology Reviews 69:326-356.

88.   Orth, P., D. Schnappinger, W. Hillen, W. Saenger, and W. Hinrichs. 2000. Structural basis of gene regulation by the tetracycline inducible Tet repressor-operator system. Nature Structural Biology 7:215-219.

89.   Orth, P., W. Saenger, and W. Hinrichs. 1999. Tetracycline-chelated Mg2+ ion initiates helix unwinding in Tet repressor induction. Biochemistry 38:191-198.

90.   Brown, M. H., and R. A. Skurray. 2001. Staphylococcal multidrug efflux protein QacA. Journal of Molecular Microbiology and Biotechnology 3:163-170.

91.   Schumacher, M. A., M. C. Miller, S. Grkovic, M. H. Brown, R. A. Skurray, and R. G. Brennan. 2002. Structural basis for cooperative DNA binding by two dimers of the multidrug-binding protein QacR. EMBO Journal 21:1210-1218.

92.   Grkovic, S., M. H. Brown, and R. A. Skurray. 2001. Transcriptional regulation of multidrug efflux pumps in bacteria. Seminars in Cell and Developmental Biology 12:225-237.

93.   Grkovic, S., K. M. Hardie, M. H. Brown, and R. A. Skurray. 2003. Interactions of the QacR multidrug-binding protein with structurally diverse ligands: implications for the evolution of the binding pocket. Biochemistry 42:15226-15236.

94.   Schumacher, M. A., and R. G. Brennan. 2002. Structural mechanisms of multidrug recognition and regulation by bacterial multidrug transcription factors. Molecular Microbiology 45:885-893.

95.   Grkovic, S., M. H. Brown, N. J. Roberts, I. T. Paulsen, and R. A. Skurray. 1998. QacR is a repressor protein that regulates expression of the *Staphylococcus aureus* multidrug efflux pump QacA. Journal of Biological Chemistry 273:18665-18673.

96.   Schumacher, M. A., M. C. Miller, S. Grkovic, M. H. Brown, R. A. Skurray, and R. G. Brennan. 2001. Structural mechanisms of QacR induction and multidrug recognition. Science 294:2158-2163.

97.   Murray, D. S., M. A. Schumacher, and R. G. Brennan. 2004. Crystal structures of QacR-diamidine complexes reveal additional multidrug-binding modes and a novel

mechanism of drug charge neutralization. Journal of Biological Chemistry 279:14365-14371.

98. Schumacher, M. A., M. C. Miller, and R. G. Brennan. 2004. Structural mechanism of the simultaneous binding of two drugs to a multidrug-binding protein. EMBO Journal 23:2923-2930.

99. Kisker, C., W. Hinrichs, K. Tovar, W. Hillen, and W. Saenger. 1995. The complex formed between Tet repressor and tetracycline-Mg2+ reveals mechanism of antibiotic resistance. Journal of Molecular Biology 247:260-280.

100. Luckner, S. R., M. Klotzsche, C. Berens, W. Hillen, and Y. A. Muller. 2007. How an agonist peptide mimics the antibiotic tetracycline to induce Tet-repressor. Journal of Molecular Biology 368:780-790.

101. Frénois, F., J. Engohang-Ndong, C. Locht, A. R. Baulard, and V. Villeret. 2004. Structure of EthR in a ligand bound conformation reveals therapeutic perspectives against tuberculosis. Molecular Cell 16:301-307.

102. Engohang-Ndong, J., D. Baillat, M. Aumercier, F. Bellefontaine, G. S. Besra, C. Locht, and A. R. Baulard. 2004. EthR, a repressor of the TetR/CamR family implicated in ethionamide resistance in mycobacteria, octamerizes cooperatively on its operator. Molecular Microbiology 51:175-188.

103. Putnam, C. D., M. Hammel, G. L. Hura, and J. A. Tainer. 2007. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. Quarterly Reviews of Biophysics 40:191-285.

104. Rambo, R. P., and J. A. Tainer. 2010. Bridging the solution divide: comprehensive structural analyses of dynamic RNA, DNA, and protein assemblies by small-angle X-ray scattering. Current Opinion in Structural Biology.

105. Lundblad, J. R., M. Laurance, and R. H. Goodman. 1996. Fluorescence polarization analysis of protein-DNA and protein-protein interactions. Molecular Endocrinology 10:607-612.

106. Van Dyke, M. W., N. Van Dyke, and G. Sunavala-Dossabhoy. 2007. REPSA: general combinatorial approach for identifying preferred ligand-DNA binding sequences. Methods 42:118-127.

107. Bailey, T. L., N. Williams, C. Misleh, and W. W. Li. 2006. MEME: discovering and analyzing DNA and protein sequence motifs. Nucleic Acids Research 34:W369-373.

108. Bailey, T. L., and M. Gribskov. 1998. Combining evidence using p-values: application to sequence homology searches. Bioinformatics 14:48-54.

109. Terwilliger, T. C., and J. Berendzen. 1999. Automated MAD and MIR structure solution. Acta Crystallographica. Section D: Biological Crystallography 55:849-861.

110. Emsley, P., and K. Cowtan. 2004. Coot: model-building tools for molecular graphics. Acta Crystallographica. Section D: Biological Crystallography 60:2126-2132.

111. Brünger, A. T., P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, and G. L. Warren. 1998. Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallographica. Section D: Biological Crystallography 54:905-921.

112. Hura, G. L., A. L. Menon, M. Hammel, R. P. Rambo, F. L. Poole, S. E. Tsutakawa, F. E. Jenney, S. Classen, K. A. Frankel, R. C. Hopkins, S.-J. Yang, J. W. Scott, B. D. Dillard, M. W. W. Adams, and J. A. Tainer. 2009. Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). Nature Methods 6:606-612.

113. Koch, M. H., P. Vachette, and D. I. Svergun. 2003. Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution. Quarterly Reviews of Biophysics 36:147-227.

114. Svergun, D. 1992. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. Journal of Applied Crystallography 25:495-503.

115. Svergun, D. I. 1999. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. Biophysical Journal 76:2879-2886.

116. Svergun, D. I., M. V. Petoukhov, and M. H. Koch. 2001. Determination of domain structure of proteins from X-ray solution scattering. Biophysical Journal 80:2946-2953.

117. Petoukhov, M. V., and D. I. Svergun. 2005. Global rigid body modeling of macromolecular complexes against small-angle scattering data. Biophysical Journal 89:1237-1250.

118. Kumaraswami, M., J. T. Schuman, S. M. Seo, G. W. Kaatz, and R. G. Brennan. 2009. Structural and biochemical characterization of MepR, a multidrug binding transcription regulator of the *Staphylococcus aureus* multidrug efflux pump MepA. Nucleic Acids Research 37:1211-1224.

119. Osawa, M., D. E. Anderson, and H. P. Erickson. 2008. Reconstitution of contractile FtsZ rings in liposomes. Science 320:792-794.

120. Nogales, E. 2000. Structural insights into microtubule function. Annual Review of Biochemistry 69:277-302.

121. Yu, X. C., and W. Margolin. 1999. FtsZ ring clusters in min and partition mutants: role of both the Min system and the nucleoid in regulating FtsZ ring localization. Molecular Microbiology 32:315-326.

122. Hu, Z., and J. Lutkenhaus. 2001. Topological regulation of cell division in E. coli. spatiotemporal oscillation of MinD requires stimulation of its ATPase by MinE and phospholipid. Molecular Cell 7:1337-1343.

123. Hu, Z., E. P. Gogol, and J. Lutkenhaus. 2002. Dynamic assembly of MinD on phospholipid vesicles regulated by ATP and MinE. Proceedings of the National Academy of Sciences, USA 99:6761-6766.

124. Shih, Y.-L., T. Le, and L. Rothfield. 2003. Division site selection in *Escherichia coli* involves dynamic redistribution of Min proteins within coiled structures that extend between the two cell poles. Proceedings of the National Academy of Sciences, USA 100:7865-7870.

125. Schumacher, M. A. 2008. Structural biology of plasmid partition: uncovering the molecular mechanisms of DNA segregation. Biochemical Journal 412:1-18.

126. Holm, L., S. Kääriäinen, P. Rosenström, and A. Schenkel. 2008. Searching protein structure databases with DaliLite v.3. Bioinformatics 24:2780-2781.

127. Hinrichs, W., C. Kisker, M. Düvel, A. Müller, K. Tovar, W. Hillen, and W. Saenger. 1994. Structure of the Tet repressor-tetracycline complex and regulation of antibiotic resistance. Science 264:418-420.

128. Kent, W. J., C. W. Sugnet, T. S. Furey, K. M. Roskin, T. H. Pringle, A. M. Zahler, and D. Haussler. 2002. The human genome browser at UCSC. Genome Research 12:996-1006.

129. Den Blaauwen, T., N. Buddelmeijer, M. E. Aarsman, C. M. Hameete, and N. Nanninga. 1999. Timing of FtsZ assembly in Escherichia coli. Journal of Bacteriology 181:5167-5175.

130. Bigot, S., V. Sivanathan, C. Possoz, F.-X. Barre, and F. Cornet. 2007. FtsK, a literate chromosome segregation machine. Molecular Microbiology 64:1434-1441.

131. Liu, Z., A. Mukherjee, and J. Lutkenhaus. 1999. Recruitment of ZipA to the division site by interaction with FtsZ. Molecular Microbiology 31:1853-1861.

132. Haney, S. A., E. Glasfeld, C. Hale, D. Keeney, Z. He, and P. de Boer. 2001. Genetic analysis of the *Escherichia coli* FtsZ.ZipA interaction in the yeast two-hybrid system. Characterization of FtsZ residues essential for the interactions with ZipA and with FtsA. Journal of Biological Chemistry 276:11980-11987.

133. Arnold, K., L. Bordoli, J. Kopp, and T. Schwede. 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22:195-201.

134. Schneidman-Duhovny, D., Y. Inbar, R. Nussinov, and H. J. Wolfson. 2005. PatchDock and SymmDock: servers for rigid and symmetric docking. Nucleic Acids Research 33:W363-367.

135. Comeau, S. R., D. W. Gatchell, S. Vajda, and C. J. Camacho. 2004. ClusPro: a fully automated algorithm for protein-protein docking. Nucleic Acids Research 32:W96-99.

136. Dowd, D. R., and R. S. Lloyd. 1990. Biological significance of facilitated diffusion in protein-DNA interactions. Applications to T4 endonuclease V-initiated DNA repair. Journal of Biological Chemistry 265:3424-3431.

137. van Noort, S. J., K. O. van der Werf, A. P. Eker, C. Wyman, B. G. de Grooth, N. F. van Hulst, and J. Greve. 1998. Direct visualization of dynamic protein-DNA interactions with a dedicated atomic force microscope. Biophysical Journal 74:2840-2849.

Vita

Nam Tonthat was born in Vietnam on December 23rd, 1983. After immigrating to the United States at the age of 10, he and his family settled in Houston, Texas. He attended Baylor University in 2001 and graduated in 2004 with the degree of Bachelor of Science in Informatics. He then attended the University of Texas at El Paso and earned the degree of Master of Science in Bioinformatics in 2006. In that same year, he matriculated into the Graduate School of Biomedical Science at the University of Texas Health Science Center in Houston. He is currently a member of the Schumacher lab in the Department of Biochemistry and Molecular Biology at the UT MD Anderson Cancer Center. This work was completed under the guidance of Dr. Maria Schumacher, and was done is in partial fulfillment for the degree of Doctor of Philosophy.