CHEMICAL BIOLOGY EVALUATION OF BACTERIAL TRANSCRIPTION INHIBITORS

CHAN SHU TING

MPhil

The Hong Kong Polytechnic University

2020

The Hong Kong Polytechnic University

Department of Applied Biology and Chemical Technology

Chemical Biology Evaluation of Bacterial Transcription Inhibitors

Chan Shu Ting

A thesis submitted in partial fulfilment of the requirements for the degree of Master of Philosophy

August 2019

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree of diploma, except where due acknowledge has been made in the text.

_____ _____ (Signed)

Chan Shu Ting (Name of student)

Abstract

Antibiotic resistance poses global health threat and it is urgent to identify new targets for novel antibacterial agent development. Bacterial transcription as a novel target has been identified and two new classes of compounds with antibacterial activities that target on different essential factors in bacterial transcription were synthesized. Herein, this thesis focuses on the biological evaluation on the two classes of bacterial transcription targeting inhibitors, i.e. the C3 derivatives and the MC4 derivatives. C3 derivatives are a novel class of transcription targeting inhibitors that can inhibit the bacterial transcription initiation by inhibition of the interaction between the clamp helix (CH) region of β' subunit of bacterial RNA polymerase (RNAP) and region 2.2 of σ factor through targeting the RNAP β'- CH. Compounds that are designed to target the necessary β'- CH to inhibit the formation of holoenzyme are potential to be developed into new antibacterial drugs. Biological evaluation on these compounds was performed. The fluorescence images of live bacterial *Bacillus subtilis* cells were captured to analyze the mechanism of action of the compounds. Compounds were found to have mechanism of action consistent to the inhibition of bacterial transcription initiation.

MC4 is a lead compound with antibacterial activity. The derivatives were designed to target the essential antitermination factor NusB. The interaction between NusB and NusE are critical for the formation of antitermination complex with bacterial RNAP. Compounds are designed to target NusB for the inhibition of NusB-NusE interaction and hence the rRNA synthesis in bacteria. Several biological evaluations of MC4 derivatives were performed to analyze their interaction with NusB. For instance, circular dichroism was measured to study the effect of compounds on the secondary structure of NusB. ITC assay was attempted to determine the dissociation constants of

compounds and native mass spectrometry as a complementary method for determination. Fluorescence microscopic images of treated live *B. subtilis* cells were captured to analyze the mechanism of action of the compounds.

Research publications

Journal papers

1. Yangyi Qiu, Shu Ting Chan, Lin Lin, Tsun Lam Shek, Tsz Fung Tsang, Nilakshi Barua, Yufeng Zhang, Margaret Ip, Paul Kay-sheung Chan, Nicolas Blanchard, Gilles Hanquet, Zhong Zuo, Xiao Yang, Cong Ma, "Design, synthesis and biological evaluation of antimicrobial diarylimine and –amine compounds targeting the interaction between the bacterial NusB and NusE proteins," *European Journal of Medicinal Chemistry*, vol. 178, pp. 214-231, 2019/09/15/ 2019.
2. Yangyi Qiu, Shu Ting Chan, Lin Lin, Tsun Lam Shek, Tsz Fung Tsang, Yufeng Zhang, Margaret Ip, Paul Kay-sheung Chan, Nicolas Blanchard, Gilles Hanquet, Zhong Zuo, Xiao Yang, Cong Ma, "Nusbiarylins, a new class of antimicrobial agents: Rational design of bacterial transcription inhibitors targeting the interaction between the NusB and NusE proteins," *Bioorganic Chemistry*, vol. 92, p. 103203, 2019/11/01/ 2019.

Acknowledgement

I would like to express my gratitude to my supervisor Dr. Ma, Cong for his guidance, time and endless effort in advising my research, Professor Zhao, Yanxiang as the Board of examination, Professor Xia, Jiang and Professor Zhu, Guangyu for their precious time and effort as external examiners in reviewing my thesis, Dr. Yang, Xiao for her effort as collaborator and provision of experimental materials, Professor Wu, Jianyong and Dr. Yao, Zhongping who led me to this path, State Key Laboratory of Chemical Biology and Drug Discovery and Department of ABCT for provision of research funding and resources, ULS and UCEA in provision of research facilities, scientific officers Dr. So, Pui-kin, Dr. Wong, Yee-man, Dr. Tsang, Chui-shan, Dr. Hung, Hiu-ling for their technical supports, my laboratory fellows Mr. Qiu, Yangyi, Mr. Jiqing Ye, Mr. Tsang, Tsz-fung on their effort in helping me with the experimental materials and process, Dr. Ng. Tommy, Ms Wang, Jianying and Mr. Wong, Tsz-fung for their professional knowledge on mass spectrometry, and my life partner Dr. Lee, Kui-fun who supports me without reservation and keeps me sane all the time. Finally, I would like to also thank everyone who are not listed above for their support and love which I could not live without.

Contents

List of abbreviations

RNAP	RNA polymerase
rRNA	Ribosomal RNA
Nus	N-utilization substances
NusB	N-utilization substance B, an antitermination factor
NusE	N-utilization substance E, an antitermination factor
nut	N-utilizing
B. subtilis	*Bacillus subtilis*
NusB-His	NusB with 6x His-tagging
eco	*Escherichia coli*
mtu	*Mycobacterium tuberculosis*
tma	*Thermotoga maririma*
MIC	Minimum inhibitory concentration
GFP	Green fluorescent protein
PPI	Protein-protein interaction
CD	Circular dichroism
ITC	Isothermal titration calorimetry
SAR	Structure-activity relationship

Chapter 0. Introduction

0.0 Chapter introduction

This thesis consists of two major chapters. Chapter 1 focuses on Transcription initiation and Chapter 2 focuses on Transcription antitermination. This chapter serves as a detailed introduction to the research background. Firstly, a general idea of antibiotic resistance and the needs to find novel targets for drug discovery will be discussed. Next, there will be a discussion on bacterial transcription and an in-depth description of bacterial transcription processes will be presented. Then, the new bacterial transcription targets and protein-protein interactions will be deliberated. After that, the bacterial transcription targeting inhibitors will be examined in detail including their mode of action and structure-activity relationship. Research gaps will be identified, and research questions will be raised followed by research methods and approaches. Finally, contribution of the research studies will be delivered.

0.1 Antibiotic resistance and needs of novel antibacterial targets

Antibiotics have been developed to treat infections and diseases caused by bacteria. Before the introduction of sulfonamides and penicillin in the 1930s and 1940s, bacterial infection was the major cause of death among children and adults[1]. The problem was eased with the development of antibiotics until the 1960s when research and development of new antibiotics became stagnant because of the assumption that people could effectively control all microbial infections. Nevertheless, the overuse of earlier antibiotics contributed to the increasing rate of development of antibiotic resistance within hospital and community. Because of natural selection, bacteria eventually develop resistance as a result of mutation of existing gene or horizontal gene transfer[2]. Even though the community is threatened by the spread of antibiotic-resistant bacteria, pharmaceutical research has been failing to meet the need of new antibiotics. Moreover, over the past two decades, only two new classes of antibiotics have been developed and clinically approved[3]. Miserably, new antibiotics released to the market are all derivatives of existing classes, resulting in cross resistance within microbiome[4]. Diseases caused by multidrug resistance in gram-positive and gram-negative bacteria are more difficult to be treated with conventional antibiotics and consequently there will be lack of therapy and preventive measures[5]. The emergence of antibiotic resistance rises public health concern and the problem is urged to be solved[6, 7]. One of the possible solutions to combat this global health challenge is to identify novel antibacterial targets with new modes of actions besides the renowned cell wall or protein synthesis[8].

Antibiotic resistance is acquired from spontaneous or induced mutation[6] and horizontal gene transfer[9] from resistant microorganisms. It is primarily caused by decreased membrane

permeability[10], increased efflux capacity[11, 12], enzymatic inactivation[13, 14] and direct mutation of binding site of drug targets[15]. Currently, majority of the clinically approved antibiotics target the growth or integrity of bacterial cell wall[16], translation and DNA replication[17] or segregation[18]. It seems to be true that there are limiting antibiotics that target specifically on bacterial transcription with approved mechanisms. Because of this, bacterial transcription is considered an underutilized target for novel antibacterial drug discovery.

0.2 Research gap

The problem resulted from antibiotic resistance causes urgent need to develop new antibiotics with novel targets. However, there are several difficulties that are hindering the development of new classes of antibiotics to overcome the situation. For example, the rate of developing resistant, multi-resistant and totally antibiotic-resistant infections has been faster than the development of new antibiotics. The new compounds released to the market are the derivatives of existing classes, such that cross-resistance often exists among the microbiomes. Besides, many compounds have been proven to be not suitable for further development[19].

Up to now there are only two commercially available transcription targeting antibiotics with proven mechanisms of actions, namely rifamycin and fidaxomicin (lipiarmycin). Considering the advantages of targeting bacterial transcription as the strategies to develop novel antibacterial agents, development on transcription targeting inhibitors has been limited. Although lots of research on transcription targeting inhibitors have been conducted, many of these inhibitors cannot pass the clinical trials because of various reasons. Nevertheless, transcription regulatory in bacteria is different from that in eukaryotes, which makes the bacterial transcription an excellent drug target.

Nevertheless, bacterial transcription is a valid, novel, and underutilized target for antibacterial drug development because bacterial transcription is an essential process in all prokaryotic lives. The bacterial RNA polymerase (RNAP) together with the transcription factors are highly conserved in all bacteria, allowing the potential development of broad-spectrum anti-transcriptional antibiotics[19]. Such antibiotics have a low potential cytotoxicity to human beings because the sequence and active sites of eukaryotic RNAP are not comparable with that of bacteria's and

therefore only bacterial RNAP will be targeted by these antibiotics. Eukaryotic RNAs are synthesised by different RNAP while there is only one RNAP for the synthesis of all classes of RNAs in bacteria so the RNAP-related transcription factors between bacterial and eukaryotic cells are also not conserved. They have different homology and the bacterial RNAP targeting antibiotics do not affect eukaryotic RNAPs. This could see the potential for the development of inhibitors targeting the interaction between RNAP and transcription factors in bacteria. Additionally, there is a number of data of high-resolution crystal and solution structures of bacterial RNAP available , allowing the design of structure-based drugs[20, 21].

0.3 Bacterial transcription

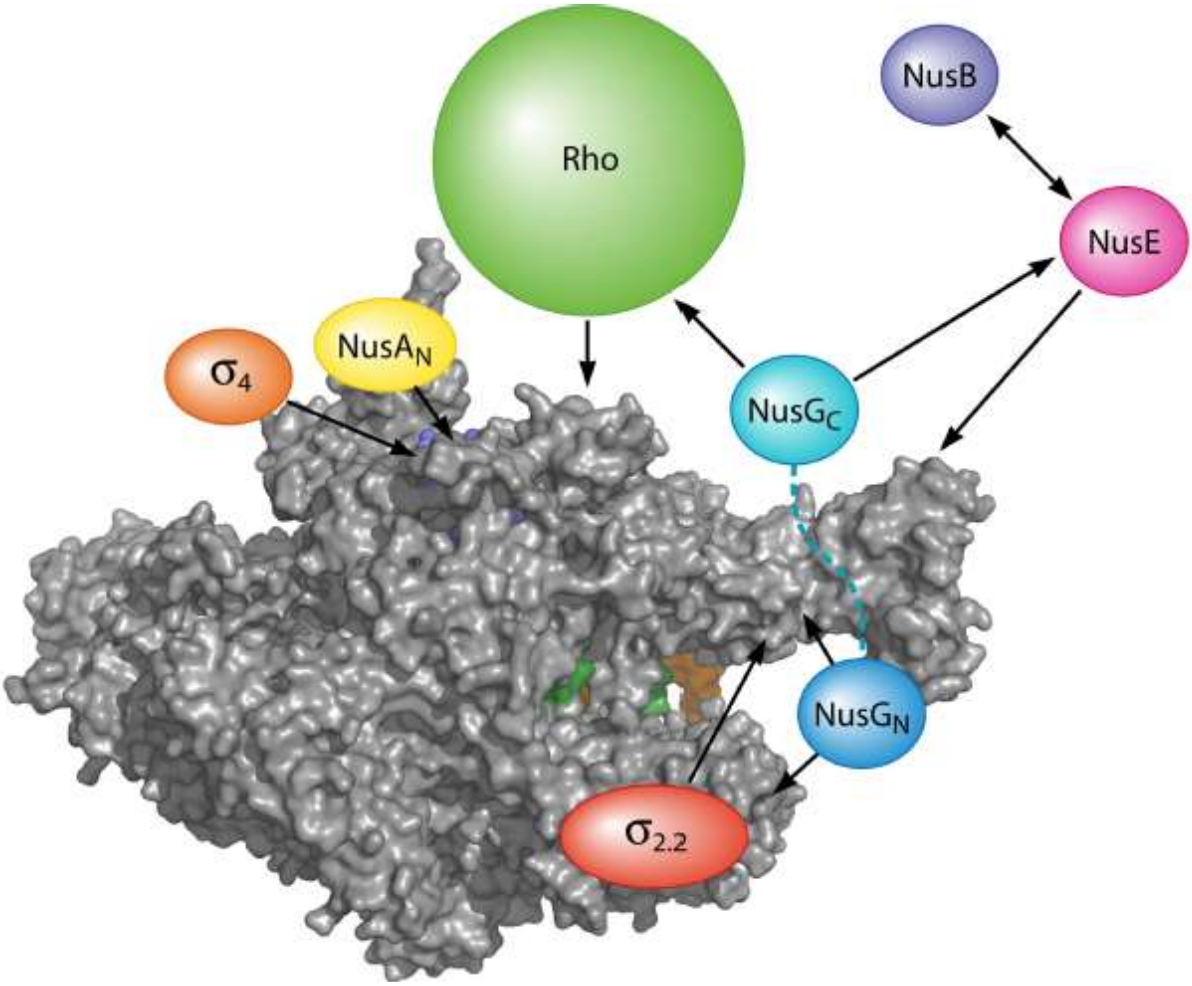


Figure 1 Bacterial transcription network showing the interaction between different transcription factors (e.g. σ factors, Rho, antitermination factors Nus A, NusB, NusE, NusG) with RNAP (gray). RNAP is the site of transcription, in which the initiation happens when the CH region of RNAP binds to $\sigma_{2.2.}$ As the transcript elongates with aids from Rho factor and NusA, it is regulated by NusB and NusE to form an antitermination complex for successful RNA synthesis.

Transcription is a process in which RNA is synthesized form its template DNA by enzyme RNA polymerase (RNAP)[19]. In eukaryotes, transcription is regulated by three types of RNAP (I, II and III) and many other transcription factors. It is more complicated than transcription in

prokaryotes, in which there is only one type of RNAP with regulating factors that regulate all transcription. RNA synthesis has been an intense research topic for over four decades, especially on bacterial transcription[22]. Transcription as the first step in bacterial genetic expression is considered a critical process in bacteria[23, 24] and is accomplished in DNA-dependent bacterial RNAP along with a series of transcription factors including transcription termination factor Rho, σ factor, N-utilization substances NusA, NusB, NusE and NusG etc[19, 23, 25] (figure 1). Nevertheless, there is only one RNAP for transcription of all classes of RNA in bacteria. Transcription consists of three major stages i.e. initiation, elongation and termination, where inhibition of initiation and inhibition of antitermination will be the two main foci in this research.

0.3.1 Bacterial RNA polymerase

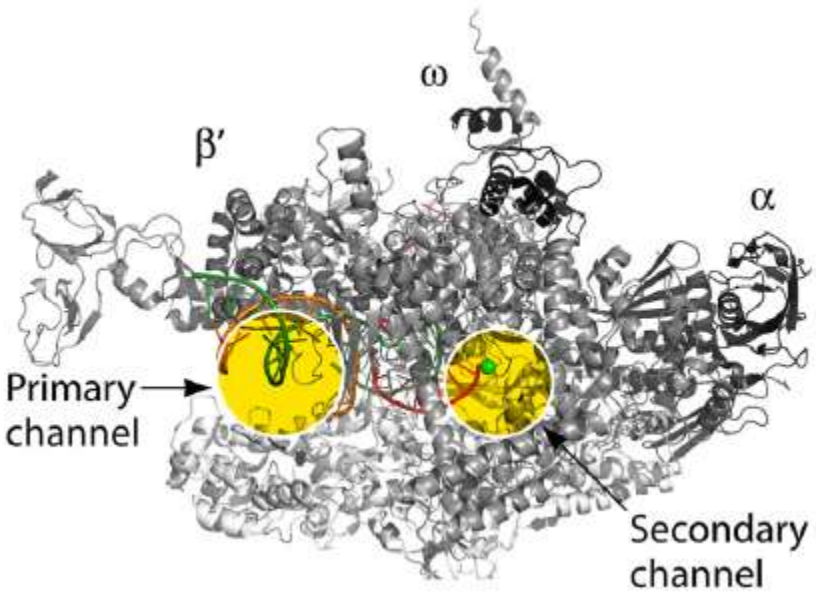


Figure 2 Structural model of bacterial RNAP, with relative position of subunits α, β', ω and primary channel and secondary channel.

Bacterial RNA polymerase (RNAP) is nucleotidyl transferase enzyme for synthesis of RNA copy of DNA or RNA template, and controlling initiation and termination of transcription in bacteria[26]. All RNAs in bacteria are synthesized in RNAP and of which, the DNA-dependent RNAP is the principal enzyme of transcription process. RNAP presents in prokaryotes, eukaryotes and many viruses but is conserved among bacteria. Bacterial RNAP is considered a valuable target for antibacterial drug discovery since it differs from eukaryotic RNAP by composing different subunits so bacterial transcription inhibitors will target specifically on bacterial RNAP without any action on eukaryotic RNAPs[19, 26, 27], indicating that the inhibitors will exhibit limiting cytotoxicity to human.

The core enzyme of bacterial RNAP (~350kDa) is composed of five subunits (two α, β, β' and ω) (figure 2). Each subunit has its own responsibility. Considering the overall shape of RNAP core enzyme as a crab claw, the two α subunits lie at the posterior and are required for assembly of the enzyme while β and β' occupy the enzyme intensively forming the gripper of the claw. β subunit is involved in chain initiation and elongation while β' subunit binds to the DNA templates. ω is an accessory unit which obliges β' subunit and assists its assembly into RNAP. α subunits are the transcription factors for transcription initiation of core enzyme by dimerization. β subunit is attached to the α dimer. β and β' subunits form the active centre of RNAP. ω binds to β' subunit and is supposed to support its correct folding and the assembly of β'ω with $\alpha_2\beta$[26]. The active centre of crab claw resembling RNAP is located at the bottom of the cleft between the two pincers of the claw, which composes of β and β' subunits. While RNAP comprises of two separate parts, namely the secondary channel (for nucleotide triphosphate 'NTP' substrates accessing to active center)[28, 29] and the RNA-exit channel[30], three aspartic acid residues together capture a

catalytic magnesium ion in the active center[31]. There is another magnesium ion bound to the incoming NTP substrate. On the other hand, the downstream of active center is formed by numerous movable elements (bridge helix) and trigger loops (TL) which are important for NTP loading, catalysis and translocation[26, 32, 33].

In addition, there is a σ factor for recognition and melting of the promoter regions[34-36]. It binds to RNAP during the initiation of transcription to form a holoenzyme[37-40]. Generally, bacterial RNAP is associated with many σ factors. Yet the most important 'housekeeping' σ factor in Gram-negative bacteria is designated as σ^{70}. Moreover, there are six other σ factors in *E. coli* (σ^{S}, σ^{32}, σ^{F}, σ^{E}, σ^{fecI}, and σ^{54}), which all bind to the core enzyme of RNAP to achieve global regulation of transcription[24, 41, 42]. The σ factor in Gram-positive bacteria is called σ^{A}. Transcription is further regulated by transcriptional activators, σ-binding anti- σ, and other small RNAs[22].

The RNAP holoenzyme, which consists of RNAP core enzyme binding to the σ factor, is formed for specific recognition of promoter DNA sequences in transcription initiation. Since the presence of holoenzyme is essential for initiation, compounds that are capable of inhibition of formation of holoenzyme can inhibit bacterial transcription initiation and are potential antibacterial drugs.

As the RNAP associates with different proteins in function of transcription, another interaction of RNAP is also interested as a novel target for development of antibacterial agents. For instance, protein NusB is another valuable target in this study. NusB is an antitermination factor that associates with RNAP for antitermination which ensures successful ribosomal RNA synthesis in bacteria. The presence of NusB-NusE interaction ensures the success of formation of

antitermination complex which in turn prevents any unnecessary termination. Therefore, interruption of this interaction will prevent the formation of antitermination complex and as a result antitermination will be inhibited. Therefore, NusB-NusE interaction has also become one of the novel targets.

0.3.2 Initiation

In transcription initiation, core enzyme of RNAP is only capable to initiate successfully in the presence of an important initiation factor (sigma, σ) to form a holoenzyme for specific DNA sequence (promoter) recognition, followed by the opening and synthesis of the first few nucleotides of the transcripts. The initiation of transcription will start following the formation of holoenzyme. Firstly, specific binding of the holoenzyme to two hexamers in the promoter region yields a closed promoter complex. Secondly, local DNA melting happens when double-stranded DNA of the complex unwinds to form an open promoter complex and hence the initiation of transcription in the presence of NTP[23, 24, 30, 43, 44]. In addition, the conserved region 3.2 of σ factor contributes to promoter escape by interaction with the nascent RNA, resulting in facilitation of σ dissociation[45]. The association of the first nucleotide of the nascent RNA chain with the secondary complex gives rise to the first ternary complex, and the incorporation of the secondary nucleotide of the RNA chain and formation of the phosphodiester bond yield an early ternary complex. This complex may lead to either the release of dinucleotide component (known as abortive initiation) or elongation of dinucleotide. After nine to eleven nucleotides being incorporated into the emerging RNA chain, the ternary complex is now stable and not prone to abortive initiation, which marks the end of initiation and the start of transcription elongation will follow[26].

A series of novel transcription initiation targeting inhibitors which target specifically the interaction of clamp helix (CH) region of β' subunit of bacterial RNAP and region 2.2 of σ factor in holoenzyme have been successfully synthesized. Since this protein-protein interaction is of absolute necessity in transcription initiation in bacteria, inhibition of the interaction is of high interest in this research study. The chemical biology evaluation of these inhibitors was done and will be discussed in detail in Chapter 1.

0.3.3 Elongation

Bacterial transcription elongation differs from initiation and termination by association with gene-scale DNA tracking system which involves bond formation at every base pair. Pausing is an important process that allows the coupling of transcription with translation and the binding of the regulatory factors to the elongation complex. In prokaryotes, transcription elongation is controlled by elongation factor NusG. It inhibits backtracking and is crucial for rapid elongation of rRNA[46]. Elongation begins when RNAP releases the initiation factors so that they escape from promoter sequence to form a stable transcribing complex. RNAP then translocates along the DNA without dissociation from the template or growing RNA sequence until it reaches a termination factor or perceives termination signal that ends the transcription cycle. This process is next followed by dissociation of transcribing complex and release of RNAP for a new cycle of transcription[26].

0.3.4 Termination and antitermination mechanism

Termination marks the end of transcription when the RNAP transcribes a terminator sequence. Yet, transcription antitermination is a unique mechanism which regulates the efficient RNA synthesis through the formation of antitermination complex between RNAP and associated antitermination factors. This complex supports the prevention of unnecessary transcription termination. Interaction

between protein NusB and NusE is of interest in this research for their importance and contribution in discovery of novel targets for transcription inhibition. Chemical biology evaluation on inhibitors targeting antitermination are discussed in detail in Chapter 2.

0.3 Bacterial transcription targeting inhibitors

Bacterial transcription is an underutilized target for developing new antibacterial agents. A number of transcription targeting inhibitors were developed previously[19], which target on different regions of RNAP, for instance the primary channel, the secondary channel and the switch region. Moreover, most of the earlier transcription targeting inhibitors were originated from microorganisms while others were synthetic compounds. A brief review on these compounds will be given below.

To date, rifamycin and fidaxomicin (lipiarmycin) are the only clinically approved bacterial transcription targeting inhibitors while no more new inhibitors of bacterial transcriptions have been developed for clinical use[47, 48]. This is because of the high rate of acquiring resistance in bacteria that hinders the development of new antibiotics. Furthermore, some of the inhibitors were not in use because of the unfavorable cyto-effects.

0.3.1 Primary channel inhibitors

Firstly, primary channel refers to the large conservable cleft formed by β and β' subunits of RNAP. Examples of primary channel inhibitors include rifamycins, sorangicin, GE23077, streptolydigin and CBR703.

0.3.1.1 Rifamycins

Rifamycins were the first class of inhibitors that target bacterial RNAP. Rifamycins bind within the cleft close to the active site of RNAP and thus provide steric hindrance to prevent RNA synthesis. They also interact with RNAP holoenzyme at β fork loop II. Other derivatives also show

similar binding pattern while some of them can further interact with σ subunit. Antibacterial activity of rifampin has proved the cleft close to active site of RNAP to be a target for discovery of transcription targeting inhibitors. Yet, the quick mutation of the cleft region has soon developed resistance against rifampin. To tackle the problem, possible solutions such as active site modification, combinatorial therapy and development of new delivery mechanism and derivatives have been suggested.

0.3.1.2 Sorangicin

Sorangicin binds to the same cleft region close to the active center in RNAP as rifampin and shows interaction with surrounding amino acid residues in a similar fashion. Sorangicin has the same mechanism of action and biological activity as rifampin. Even though sorangicin is structurally resembling rifampin, it is more flexible than rifampin that it fits better in RNAP. Moreover, sorangicin shows fewer resistance to mutant bacterial strains than rifampin does because of its higher flexibility.

0.3.1.3 GE23077

GE23077 is another antibiotic that binds to the i and i+1 sites of active center of RNAP, where are close to the active site that rifampin binds. Due to its hydrophilic nature, GE23077 has poor membrane permeability and thus poor antibacterial activity. Structural modification does not show improvement either. Later, construction of bipartite molecules by fusion of rifamycin SV and GE23077 provides antibiotics with outstanding antibacterial activity against resistant strains. The new compounds successfully bind to the rifampin binding site and GE23077 binding site. It is confirmed that these two sites can be considered a new combined target for drug discovery. There

are mobile elements surrounding the large cleft of primary channel such as bridge helix (BH) and trigger loop (TL) in β' subunit of RNAP. They involve in regulation of substrate loading and nucleotide addition by switching open or close the DNA binding cleft. These conformations are important in allowing RNA translocation and refolding of TL with BH in a stable form.

0.3.1.4 Streptolydigin

Streptolydigin is a broad-spectrum antibiotic that inhibits bacterial transcription initiation, elongation and pyrophosphorolysis. Co-crystal of this compound with RNAP shows that the compound binds to the neighboring region of RNAP active site where nucleotide addition cycle takes place. In fact, this compound stabilizes the translocation state by preventing conformational changes in BH and TL during nucleotide addition cycle. This sheds light on the development of inhibitors of RNAP by targeting nucleotide addition cycle.

0.3.1.5 CBR family

CBR family is another class of antibiotics that inhibits transcription elongation by stabilizing elongation complex isomerization and slowing down translocation. The compounds introduce allosteric effect to prevent TL folding and inhibit the movement of BH. One of the compounds in this family, namely CBR703 interacts with bacterial RNAP through β lobes, β flap loop, β' F-loop and BH of β' subunit. Confirmation of the binding site together with mutagenesis analysis propose that the binding of CBR compounds to bacterial RNAP could be a potential target for development of new antibiotics, and CBR703 could be an excellent starting point.

0.3.2 Secondary channel inhibitors

Secondly, there are secondary channel inhibitors that target the secondary channel in RNAP. Secondary channel refers to the funnel shaped structure formed by BH of β' spanning the DNA binding clamp immediately downstream of the active site. Several transcription factors regulate the activity of RNAP through interaction with the secondary channel. For instance, the RNA transcript in an elongation complex slides into the secondary channel and a transcription factor will penetrate the RNAP close to the active site and assist the intrinsic RNA cleavage of RNAP. The RNA transcript will then be released, and the transcription elongation complex will be liberated. Inhibitors targeting the secondary channel in RNAP will facilitate the inhibition of bacterial transcription. Herein, a series of secondary channel inhibitors that specifically target β' subunit will be discussed for further understanding on inhibition of the subunit.

0.3.2.1 Tagetitoxin

Tagetitoxin inhibits bacterial RNAP by targeting the TL in β' subunit and stabilizing the transcription elongation complex in inactive conformation. It binds to the site near active center in RNAP and specifically interacts with the arginine residue of the TL and hence inhibits RNAP translocation[49, 50]. Moreover, the crystal structure of tagetitoxin complexed to RNAP of *Thermus thermophilus* shows that it binds within the RNAP secondary channel through polar interactions with the β and β' subunits. There are two catalytic Mg^{2+} ions coordinated by the aspartic residues in β and β' subunits. The phosphate of tagetitoxin coordinates the third Mg^{2+} ion distinct from the catalytic ions. Tagetitoxin inhibits all RNAP catalytic reactions and proposes an inhibition mechanism through which the inhibitor bound Mg^{2+} ion has a crucial role in stabilization of the inactive transcription intermediate[51].

0.3.2.2 Microcin J25

Microcin J25 is a peptide inhibitor that is active against the DNA-dependent RNAP of gram-negative bacteria[52]. Microcin J25 consists of a lassoed tail and an internal lactam linkage while the tail passes through the lariat ring (eight-residue ring) with two amino acids straddling in each side. It inhibits abortive initiation and elongation of transcription by binding to RNAP secondary channel and hence prevents the traffic of NTP substrates to the catalytic center of the RNAP[53, 54]. The development of microcin J25 has been hindered since this antibacterial agent is only active to the RNAP of gram-negative bacteria. However, it highlights the possibility of secondary channel as a potential target for drug discovery.

0.3.3 Switch region inhibitors

Switch region within bacterial RNAP has been identified as a new drug target. Switch region is a structural motif that facilitates conformational changes and contacts that are required for RNAP to insert the DNA into the RNAP-active center cleft during transcription initiation. Switch region is located at the base of the RNAP clamp, which serves as the hinge on which the clamp swings in clamp opening and closing. Switch region targeting inhibitors inhibit bacterial RNAP by binding to the switch region and interfering the essential switch-region dependent conformation, DNA or RNA contact. Since switch region is a highly conserved element in bacterial RNAP, inhibitors exhibit broad-spectrum inhibition. Besides, inhibitors targeting the switch region should not be cross-resistant to rifamycin as the switch region does not overlap the binding site of rifamycin [55, 56].

Certain amount of research has shown that switch region is a novel target for inhibitors. Inhibitors that target the switch region within RNAP exhibit broad-spectrum inhibition as switch region is

present in most bacterial species. Currently, there are four identified switch region targeting inhibitors, namely myxopyronin, corallopyronin, ripostatin and lipiarmycin (fidaxomicin).

0.3.3.1 Fidaxomicin and Lipiarmycin

Fidaxomicin and lipiarmycin (also known as fidaxomicin tiacumicin B) are switch region inhibitors. They have structural similarity, and both target bacterial RNAP. Although there is no available structural data on complexes of these two compounds with bacterial RNAP, it has been shown that they both target on switch region of RNAP[57, 58]. Lipiarmycin inhibits promoter DNA melting, σ-dependent transcription, and template strand DNA binding to RNAP. A mutagenesis study shows that lipiarmycin targets on σ factor and switch region to prevent the formation of transcription open complex while fidaxomicin is a more promising compound than lipiarmycin that the deletion of σ_3 loop of RNAP does not result in resistance to fidaxomicin. In addition to the inhibition mode of lipiarmycin, it targets the region 3.2 of σ subunit and β' subunit switch-2 element of RNAP. This controls the clamping of promoter DNA in RNAP active-site cleft. Lipiarmycin then ends the isomerization of the closed-promoter complex to the transcriptionally competent open complex and hence blocks the σ-simulated RNA synthesis on 'promoter-less' DNA templates[57].

0.3.4 Limitations in novel antibiotics development

The transcription targeting inhibitors are not limited to those introduced above. Many antibacterial compounds have been developed but cannot pass the clinical trials. Even some of the antibiotics mentioned cannot be released to the market due to several reasons. In fact, out of all the transcription targeting inhibitors that have been isolated, only rifamycin and fidaxomicin/lipiarmycin have been approved for clinical use while no new transcription targeting

inhibitors are put in clinical trials. Others have been slowed down in development due to various reasons. For example, a high acquisition rate of resistance is the major problem that prevents the development of many compounds[59-61]. Although myxopyronins have proven their antibacterial activities, they are confiscated by serum albumin and are prevented from accessing their target. CBR703 has been proved to be unattractive due to significant cytotoxicity[62]. Compounds from the SB series have also been proved to have nonspecific mode of action.

0.5 Bacterial transcription - an underutilized target for antibacterial drug discovery

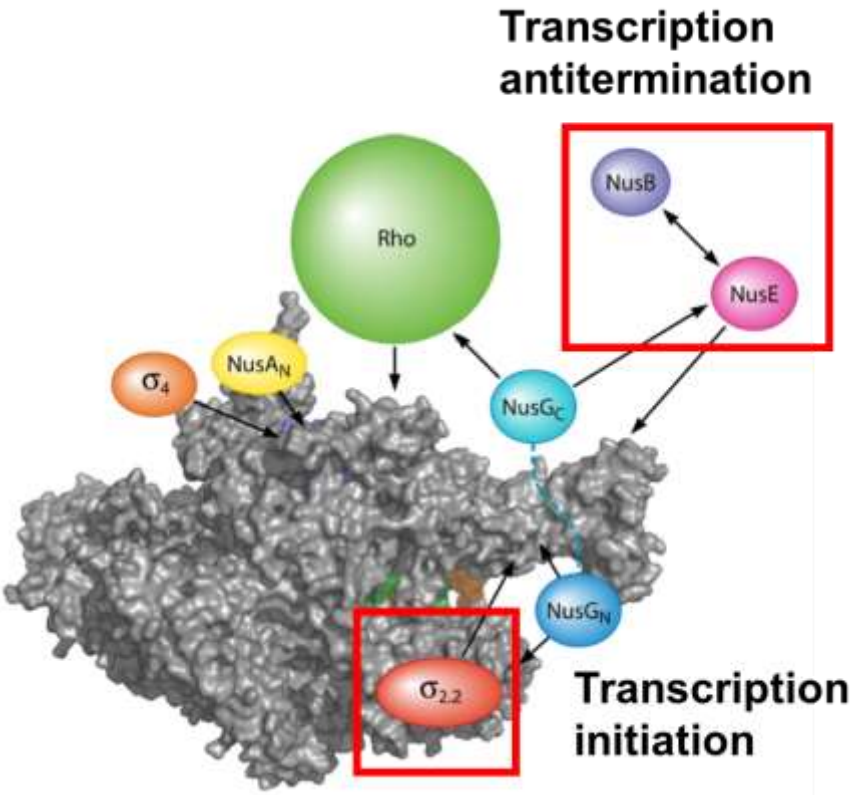
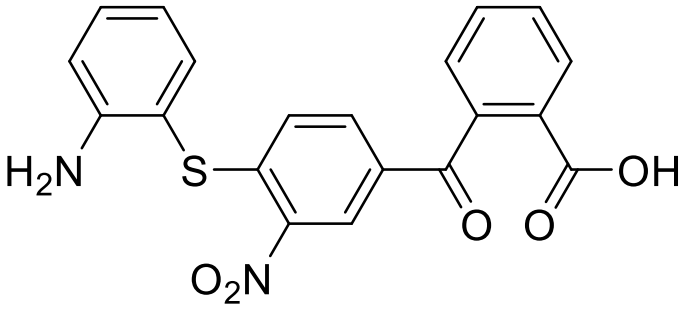


Figure 3 Bacterial transcription network with two new targets identified for novel antibacterial drug discovery: transcription initiation involving protein-protein interaction between initiation factor σ and β' subunit of RNAP and transcription antitermination involving interaction between antitermination factors NusB and NusE.

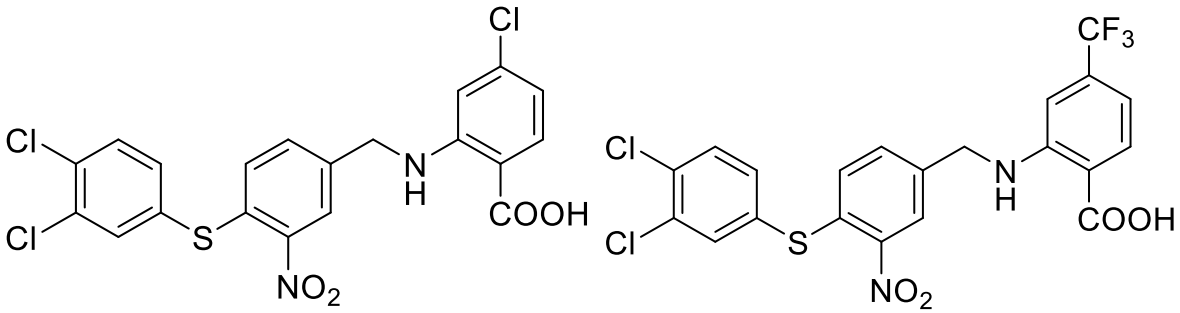
Because bacterial transcription is a valid and underutilized target, two targets have been identified for developing novel antibiotics (figure 3). The first one is targeting RNAP β'CH-$\sigma_{2.2}$ interaction in the formation of RNAP holoenzyme during transcription initiation (Chapter 1). The second one is targeting NusB-NusE interaction in the formation of antitermination complex during transcription antitermination (Chapter 2). To each of the target, a hit compound, namely C3 and MC4, has been identified with antimicrobial activities and found to target bacterial transcription.

0.5.1 C3 hit compound



C3

Figure 4 Structure of lead compound C3



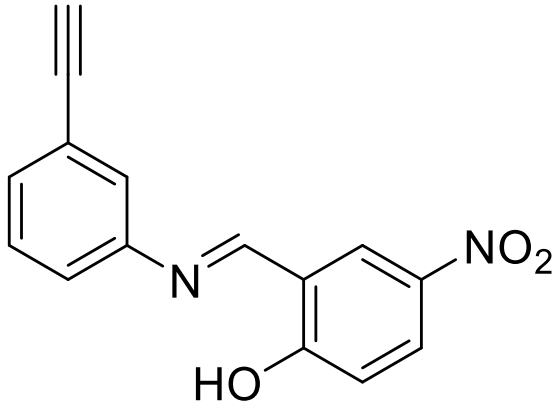
87 **89**

Figure 5 Structures of C3 derivatives

Lead compound C3 (figure 4) (MIC: 256 μg/ml) has been identified as a bacterial transcription inhibitor targeting the holoenzyme formation during transcription initiation with antibacterial activities. C3 targets RNAP β'CH-σ interaction to inhibit the formation of RNAP holoenzyme during transcription initiation. A series of C3 derivatives (figure 5) have been successfully synthesized with improved antibacterial activities. One of these compounds shows dramatically

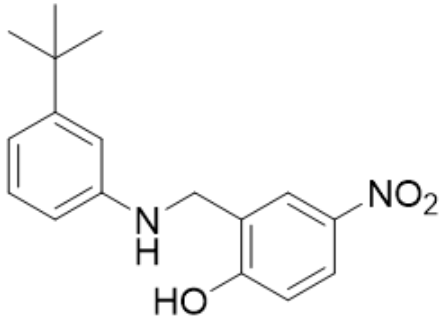
improved activities against pneumococci compared to the lead compound C3. This compound even diminishes the toxin release by *S. pneumoniae* more intensively than the existing antibiotics[63]. Chemical biological evaluation of C3 derivatives has been designed and will be discussed in Chapter 1.

0.5.2 Nusbiarylins

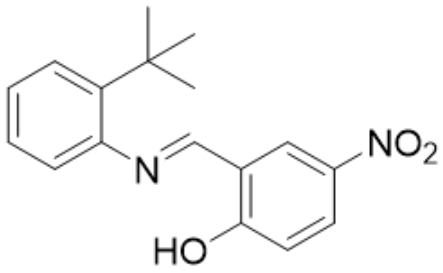


MC4

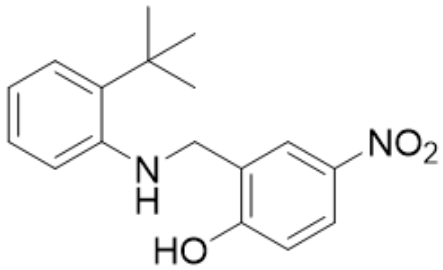
Figure 6 Structure of lead compound MC4



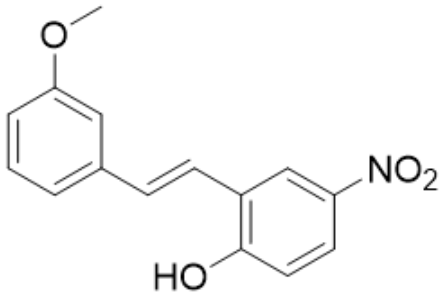
61



64



84



140

Figure 7 Structures of MC4 derivatives

Lead compound MC4 (figure 6) (K_d=1.45 μM ± 0.55 μM; MIC = 64 μg/ml) has previously been identified as a first-in-class inhibitor of bacterial ribosomal RNA (rRNA) synthesis with antibacterial activities by targeting on NusB-NusE interaction during transcription antitermination. MC4 has been shown to have antimicrobial activities again methicillin-resistant *S. aureus* (MRSA), and is able to be developed into antimicrobial agent against MRSA infections[64].

A series of MC4 derivatives (figure 7) have been rationally designed and successfully synthesized. The compounds are named Nusbiarylins[65] ("nus- biarylins") based on their target (protein NusB) and their structure (biaryl, two benzene rings). Nusbiarylins are a novel class of bacterial transcription inhibitors targeting on NusB-NusE interaction during transcription antitermination. Nusbiarylins have exhibited improved antimicrobial activities and improved binding with NusB (MIC: 1-2 μg /ml). The compounds are active against *S. aureus* including a panel of representative globally spread hospital-associated or community-associated MRSA strains. The MIC of some of the compounds are even lower than the common antibiotics that have already shown resistance to certain MRSA strains[65].

0.6 Research questions

It is hypothesized that the rationally designed small molecule inhibitors can target bacterial transcription and should have antibacterial activities since bacterial transcription is an essential biological process in bacterial cell lives. To prove the above hypothesis, a series of chemical biological evaluations have been designed to evaluate the inhibitors.

Chapter 1. Transcription initiation

1.0 Chapter introduction

This chapter focuses on chemical biology evaluation of C3 derivatives, the inhibitors targeting bacterial transcription initiation. General idea of bacterial transcription has been discussed in the previous chapter. Details of proteins involved in transcription initiation will be discussed here. These include the protein targets of interest, protein-protein interaction, and targeting the protein-protein interaction for structure-based design of novel antibacterial agent development. Contribution of targeting bacterial transcription initiation will also be presented. Next, literature on related research will be reviewed and criticized. Based on these research studies, research questions are raised. Then, methodology of this study will be presented, including the approaches on evaluation method, criticism on the chosen methods, methods and materials, procedures, and limitation. Results will then be presented, followed by finding and discussion. Finally, limitation of research and conclusion will be delivered.

1.1 Introduction

1.1.1 Bacterial transcription initiation

Transcription is the first step in genetic expression. It takes place in the RNAP in all organisms. The structure and composition of RNAP have been described in previous chapter. On the other hand, transcription initiation is the first stage in transcription and is a crucial genetic expression process in bacteria. It is essential in recognizing the promoter sequence, binding to promoter, opening and synthesis of the first few nucleotides of transcript.

Transcription initiation requires a catalytic RNAP core enzyme associating with a housekeeping σ factor to form a holoenzyme (PDB entry 4LJZ), where the σ factor is available for further competent specific binding to DNA promoter region only after its binding to RNAP. This promoter region consists of two hexanucleotides, namely -10 (TTGACA) and -35 (TATAAT) motifs, separated by 17 base pairs. Moreover, regulation of transcription initiation requires one housekeeping σ factor while other σ factors are responsible for transcribing other specific genes.

During initiation, RNAP scans the DNA until a specific region can be recognized by a specific region of the σ factor. The region 2.4 of σ factor will then bind to the -10 motif of promoter and the region 4.2 of σ factor to the -35 motif. This is followed by DNA melting of roughly 14 base pairs around the AT-rich -10 box. As a result, RNA synthesis is initiated at the +1-start site.

The key proteins to be focused on in this study of transcription initiation are the clamp-helix (CH) region on β' subunit of RNAP and the region 2.2 of housekeeping σ factor, which will be discussed in detail.

1.1.2 RNAP β'CH region

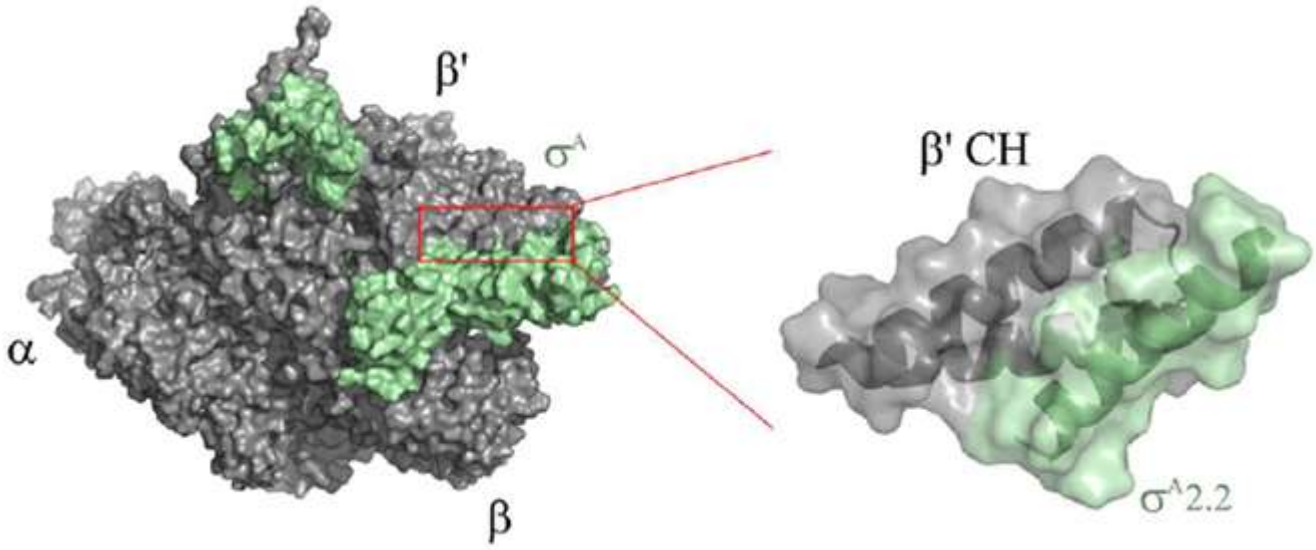


Figure 8 Homology model of *B. subtilis* RNAP holoenzyme where the RNAP core subunits are shown in gray and σ factor in green (left), and the zoom-in diagram showing the interaction between RNAP β'CH and $\sigma_{2.2}$ (right)[66].

Transcription initiation targeting inhibitors are designed to target the solvent exposed coiled-coil clamp helix (CH) region of β' subunit of bacterial RNAP. Figure 8 shows a homology model of RNAP holoenzyme consisting of interaction between β'CH and $\sigma_{2.2}$. The CH region is located at the edge of DNA binding cleft but distant away from the active site of the RNAP. The interhelical interaction is important for the retainment of the coiled-coil structure of β' subunit as the helical/coiled-coil structure in the CH region is necessary for the binding to the σ subunit. There are five subunits in bacterial RNAP core enzyme which is in contact with σ factors at various region. Yet, the most extensive contact is found between region 2.2 of σ factors and the CH region of β' subunit of RNAP.

1.1.3 σ factors

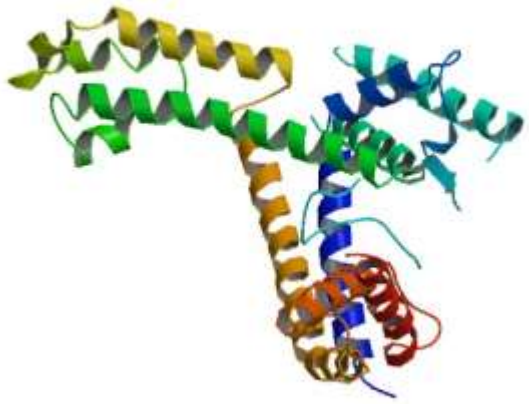


Figure 9 Crystal structure of *E. coli* σ factor from RNAP (PDB entry 1SIG)[67]

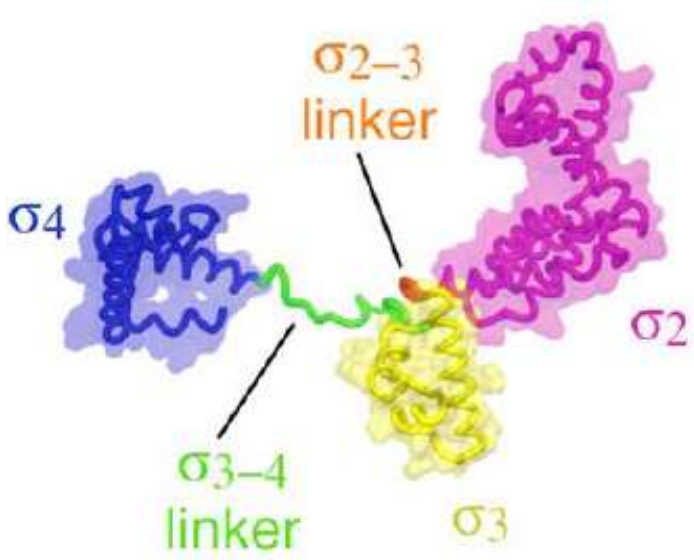


Figure 10 Structure of σ factor in holoenzyme with four conserved domains. The arrangement of σ domains and linkers as observed in the RNAP holoenzyme is shown in different colors: σ2 (pink), σ3 (yellow), σ4 (blue), σ2-3 (linker, orange), σ3-4 (linker, green)[43]

σ factors (figure 9), the bacterial transcription initiation factors, are a large family of proteins in bacterial cells that are essential and specific for transcription initiation by formation of transcription initiation complex with RNAP. The major function of σ factors is to associate with the catalytic core RNAP to guide through essential steps of initiation: the promoter recognition and the opening and synthesis of the first few nucleotides of the transcript in RNA synthesis. In most bacteria, there are a range of different σ factors for transcription of various genes. However, the highly conserved, housekeeping σ factor in *Bacillus subtilis* is designated as σ^A while that in *Escherichia coli* is σ^{70}. These σ factors are highly conserved in bacterial lives and consisted of four conserved regions, where the major binding region with β'CH is located at σ_2 and defined as region 2.2. The arrangement of the domains as observed in a RNAP holoenzyme is presented in figure 10.

Despite the huge amount of σ factors in eubacteria, bacterial transcription of housekeeping genes is dependent on highly conserved σ^{70} /σ^A factors[68]. In bacterial cell, the polar surface of region 2.2 of σ^{70}/σ^A factors interacts with RNAP through β' subunit at the CH region[69, 70]. Since the domains of the σ factors that interact with RNAP core enzyme are conservative, inhibitors of the interaction are expected to exhibit broad spectrum antibacterial activities[69].

1.1.4 RNAP β'CH-σ complex

1.1.4.1 RNAP β'CH-$\sigma_{2.2}$ interaction

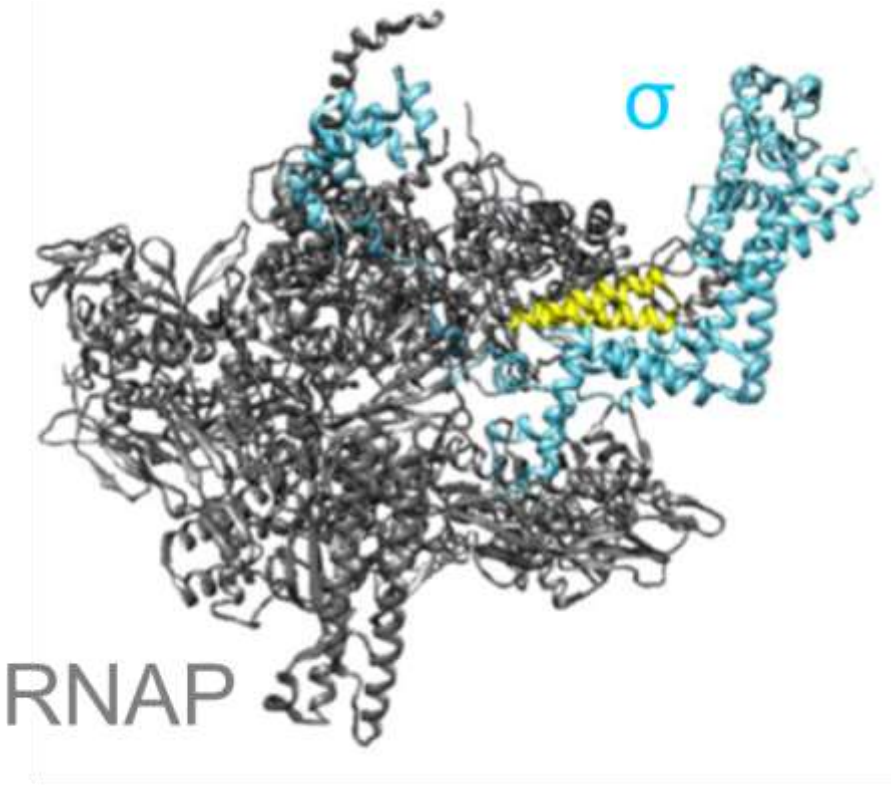


Figure 11 The crystal structure of RNAP holoenzyme with RNAP core enzyme (gray), CH region (yellow) and σ^{70} (blue)[63]

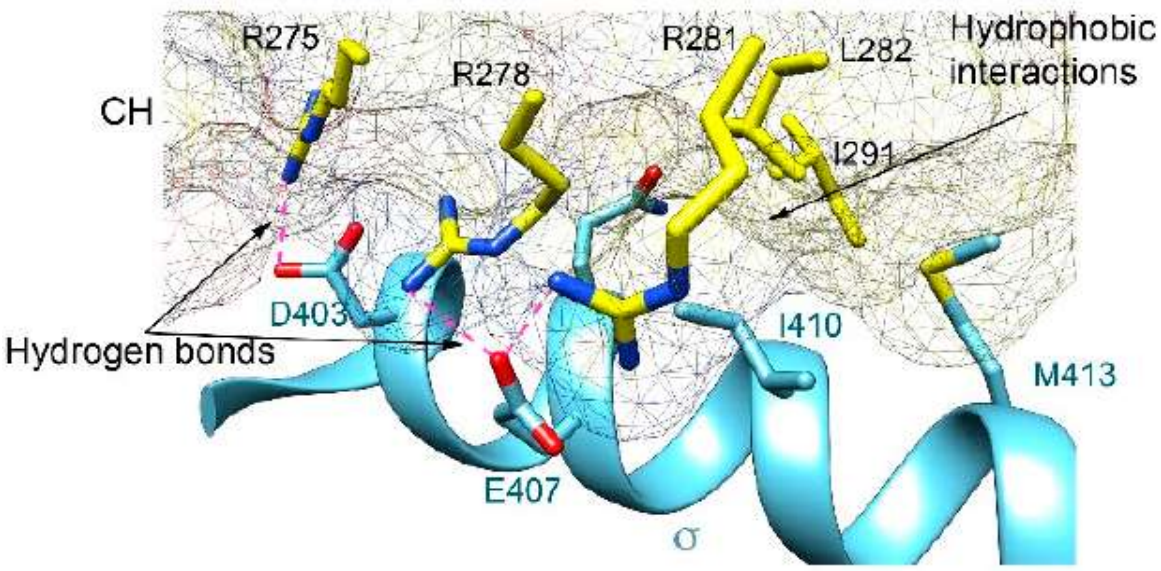


Figure 12 Zoom-in showing the interaction between RNAP CH region (mesh surface) and σ^{70} (blue)[63]

RNAP core enzyme alone is not able to carry out transcription initiation efficiently and specifically. To undergo transcription initiation, RNAP must bind to a σ factor (an initiation factor) to form a holoenzyme (HE) and hence to recognize promoter sequences. Although σ^{70}/σ^{A} interacts with RNAP β' through various location, the absolute necessary region on σ factors for binding to RNAP β' is designated as region 2.2 while that on RNAP β' is located at the solvent exposed CH region (figure 11) where the domain of σ in cyan represents the major binding site ($\sigma_{2.2}$) to the CH region (figure 12). This interaction is of absolute inevitability for the formation of transcription initiation complex[19].

The RNAP β'CH-$\sigma^{70}/\sigma^{A}{}_{2.2}$ interaction, however, is only seen when σ factors are bound to the core RNAP enzyme, meaning that σ factors alone are not able to bind to the β' subunit. Although σ factors are needed to recognize and bind to promoter DNA for transcription initiation, isolated σ

factors are not capable to do so. Instead, when σ factors bind to RNAP, the DNA binding region of the σ factor will be exposed for DNA recognition[24]. Furthermore, σ^{70} factors have shown to interact with isolated β' subunit but not β, indicating that the σ factors interact with the core RNAP through β' subunit and that there is a major binding site on β' subunit. It has been demonstrated that the major binding site of β' subunit is located at residue 260-309 for σ^{70} interaction[71].

The crystal structure of bacterial transcription initiation active RNAP HE from *Thermus thermophilus* has been reported[23]. The HE has shown an overall crab-claw like structure. The σ factor is located almost completely on the core RNAP surface except for region 313-342 which is inside the core enzyme. The binding of the σ factor to RNAP core reduces the space of important cavities of the enzyme that are supposed to carry the transcription bubble and the RNA product by having the following structure: (1) a N-terminal domain of the σ factor that bridges the β and β' pincers of the crab-claw on the upstream DNA of RNAP; (2) a hair-pin-like element of the σ factor protruding into the cleft of the active site; and (3) an extension of the hair-pin-like element that occupies the RNA-exit channel. In addition, many Mg^{2+} ions are found in the holoenzyme, coating on the protein surface. This suggests that these ions involve in binding and bending of the DNA molecules.

1.1.5 Inhibition of transcription initiation by targeting RNAP β'CH

Novel antibacterial strategies seek to inhibit transcription initiation holoenzyme complex formation by prevention of this protein-protein interaction. Targeting the critical bacterial RNAP β'CH-$\sigma_{2.2}$ interaction is able to inhibit transcription initiation. Since σ^{70}/ σ^{A} factors retain high structural similarity among bacteria, compounds that are designed to mimic the σ factors and target the CH region of β' subunit are believed to be able to inhibit the protein-protein interaction. As

σ^{70}/ σ^{A} factors are conserved only in bacteria, inhibitors targeting this interaction are unlikely to pose adverse effects on human cells[72]. Moreover, the development of pharmacophore models based on the bacterial β'CH homology allowed the structure-based design of initiation targeting inhibitors.

Fidaxomicin, a commercially available transcription targeting inhibitor, functions in inhibition of RNAP-σ interaction when the drugs are added prior to the interaction. Nevertheless, a series of bis-indole compounds called the GKL series have been developed and are successful in inhibiting bacterial transcription initiation by targeting RNAP β'CH-σ interaction. The compounds have been demonstrated to bind to the CH region of bacterial RNAP β' subunit in a competitive behavior, meaning that the compounds are able to inhibit the formation of transcription initiation complex after the occurrence of holoenzyme[69].

1.1.6 Contribution of structure-based design on antibiotics targeting RNAP β'CH

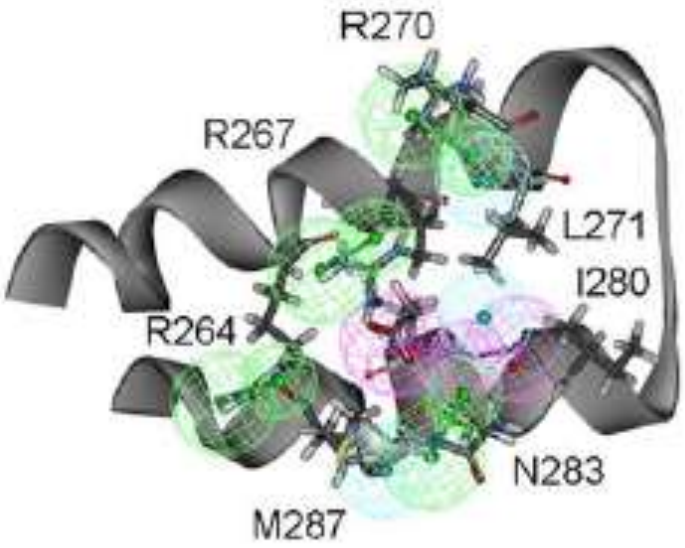


Figure 13 Pharmacophore model of β'CH, with hydrophobic groups (cyan spheres), hydrogen bond donors (pink spheres) and hydrogen bond acceptors (green spheres)[66]

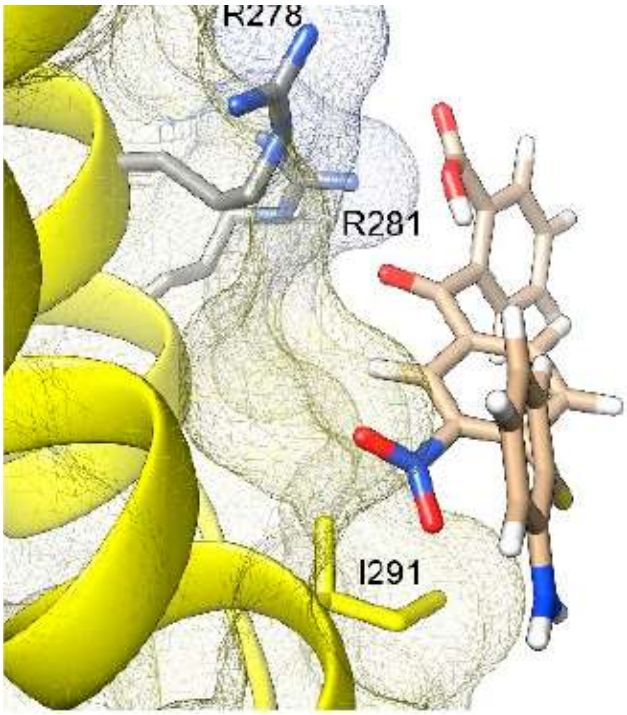


Figure 14 A docking model showing the interaction of C3 with CH region (yellow helix) and mesh surface[63]

Because initiation of bacterial transcription requires RNAP holoenzyme, inhibition of the necessary interaction between β'CH with $\sigma_{2.2}$ results in dissociation of RNAP core enzyme from the σ factor, which in turn prevents promoter recognition. Therefore, compounds that inhibit holoenzyme formation are supposed to inhibition bacterial transcription initiation and are potential antibacterial agents[66]. As β'CH is identified as a novel target, a pharmacophore model has been developed (figure 13). Novel compounds C3 derivatives have been designed to target the CH region of β' subunit of bacterial RNAP[63]. Docking C3 into the pharmacophore model of β'CH (figure 14) shows that there are interactions between the inhibitor and the protein three major amino acid residues. Since the bacterial RNA synthesis is controlled by solely one type of RNAP, compounds binding to this specific region in RNAP become critical to synthesis of all class of RNA in bacteria. Since C3 derivatives are designed to target the β'CH according to the docking model, the binding affinity of the compounds to the protein should be improved with the rationally designed structures. The compounds targeting the β'CH-$\sigma_{2.2}$ interaction should therefore inhibit the formation of RNAP holoenzyme and the promoter recognition for RNA synthesis in bacteria.

1.2 Literature review

The discovery, synthesis and introduction to therapeutic use of Rifampicin (Rifampin/Rif) can be dated back to the 1960s when tuberculosis (TB) were the major cause of death at that time[73-75]. It was one of the highly potent and broad spectrum yet relatively non-toxic antibiotics that was widely used as a treatment of TB[76] until the introduction of multidrug-resistant strains of *Mycobacterium tuberculosis* due to the frequent use of the drug, which caused a global health crisis according to the World Health Organization[77, 78].

Rifampicin is a bacterial transcription targeting inhibitor and has specific inhibition of bacterial RNAP through high-affinity binding[79]. High-resolution crystal structure of Rif-RNAP complex was reported along with other biochemical data[78]. This confirmed that Rifampicin binds to β subunit of bacterial RNAP and thus blocks the transcription initiation[78, 80].

Lipiarmycin is another bacterial transcription targeting inhibitor[81]. It was suggested that lipiarmycin achieved bacterial transcription inhibition by trapping an open-clamp state of RNAP. It only showed interaction with bacterial RNAP and had no inhibition on human RNAP[82]. It was later found that lipiarmycin inhibited bacterial RNAP by targeting σ^{70} factor at region 3.2 and β' subunit switch-2 element of RNAP, which controlled the clamping promoter DNA in RNAP active-site cleft, and are both essential for bacterial transcription initiation[57].

Pharmacophore model has been built to show $\sigma_{2.2}$ and RNAP β'CH interaction. An early mutational analysis of coiled-coil within β' subunit revealed its interaction with σ factors[71]. Crystal structure of bacterial transcription initiation active RNAP holoenzyme from *Thermus*

thermophilus was then reported, revealing the interface of $\sigma_{2.2}$ and solvent-exposed clamp-helix (CH) region of β' subunit of RNAP[23]. From this crystal structure, homology models of RNAP core and holoenzyme subunits were initiated and combined to create a pharmacophore model[70]. This model incorporated σ^A mutagenesis data of *B. subtilis* and identified the amino acid in $\sigma^A{}_{2.2}$ for major interaction between RNAP β' and σ^A. In addition, hydrogen bonding and hydrophobic interaction were involved in this interface. This revelation of structure sheds light on the structural organization of transcription initiation complex in bacteria. Further modeling based on the previous pharmacophore model was done by specific mutagenesis on amino acids in β'CH region and the key residues of amino acids in β' and σ^A were identified and thus a new pharmacophore model for inhibition of RNAP β'CH-σ was generated and used to design and screen RNAP β'CH-$\sigma_{2.2}$ interaction inhibitors[69].

The RNAP β'CH-$\sigma_{2.2}$ interaction has been one of the excellent targets for antibacterial drug discovery because of the uniqueness and high conservatory of σ factors in prokaryotic transcription initiation[70]. Although σ factors make multiple contacts with bacterial RNAP when forming transcription initiation complex, it is clear that $\sigma_{2.2}$ is of absolute necessity in prokaryotes for binding to clamp helix (CH) region of β' subunit during transcription initiation complex formation[83].

Several antibacterial agents targeting bacterial RNAP β'CH-$\sigma^{70}/\sigma^A{}_{2.2}$ interaction have been developed. Mielczarek et al. [68] synthesized several classes of bis-indole type inhibitors that can inhibit transcription initiation complex formation. These compounds were found active against RNAP β'CH-$\sigma^{70}/\sigma^A{}_{2.2}$ interaction and the inhibition of *B. subtilis* and *E. coli* was successful. The

structure-activity relationships were drawn that indole linkages and substituents at position of indole ring of the inhibitors were critical on the biological activity of the compounds. However, these compounds were restrained by their size, hence limiting solubility and ability to penetrate through the cell wall in gram-positive bacteria or the outer membrane in gram-negative bacteria[84]. Kandemir et al.[85] also synthesized a series of bis-indoles and indole-thiosemicarbazide compounds that were able to inhibit the transcription initiation complex formation and bacterial growth. Structure-activity relationship studies of these compounds suggested that hydrogen bond donors and a flexible linker were critical to antibacterial activity.

Later, Mielczarek et al.[84] synthesized another group of mono-indole inhibitors targeting this interaction. These low molecular weight compounds were also active against gram-positive and gram-negative bacteria. They were designed based on the bioactive moiety of indole in inhibition of transcription initiation complex formation and other benzofuran scaffolds to overcome the problem encountered by the large bis-indole inhibitors[68]. They were considered potential broad-spectrum antibacterial agents. Structure-activity relationship studies of these compounds suggested that the antibacterial activities arise from the low molecular weight and hydrophilicity-lipophilicity balance, and can be revealed from the mono-indoles having a greater solubility than the bis-indoles, retention of significant inhibition of β'CH-$\sigma^{70}/\sigma^{A}_{2.2}$ interaction and reduction in bacterial growth in culture.

A series of novel benzoic acid-based compounds synthesized by Wenholz et al. [72] were able to inhibit against RNAP β'CH-$\sigma^{70}/\sigma^{A}_{2.2}$ interaction. Compounds exhibited outstanding ability in inhibition of assembly of RNAP and σ factors and thus inhibition of bacterial growth. In addition,

possessing increasing hydrophilicity by incorporation of polar groups asserted them with improved aqueous solubility in significance compared to the previous bis-indole type inhibitors[68, 85] and mono-indole type inhibitors[84].

1.3 Methodology

1.3.1 Introduction

This section will start with the framework of research. Next, the approach of the studies with its justification will be discussed. Then, the methods of the studies, including the sampling techniques, data collection and procedures will be presented. These are followed by presenting the materials of the studies including laboratory equipment etc.

A novel class of C3 based bacterial transcription initiation targeting inhibitors targeting bacterial RNAP β'CH region has been synthesized and proven to be bioactive. Although many biology assays have been done on these compounds, chemical biology evaluation of the analogues have not been conducted. It has been predicted that the compounds inhibit formation of initiation complex by targeting CH region of RNAP β' subunit. Therefore, it is hypothesized that the compounds are active against initiation by preventing RNAP β'CH-$\sigma_{2.2}$ interaction.

The chemical biology evaluation of this chapter focuses on the study of effect of the compounds on bacterial cellular morphology. Localization of RNAP foci in live cell of *B. subtilis* BS1048 (*RpoC*-GFP) represents the site of RNA synthesis. Treatment of *B. subtilis* BS1048 with inhibitors may result in change in morphology and this will indicate the mode of action of the compounds. Herein, fluorescence microscopy will be employed to observe the cell morphology of compound treated cells for further understanding of their effects on bacterial transcription initiation.

1.3.2 Fluorescence microscopy

1.3.2.1 Introduction

Visual observation of cell morphology has become a powerful tool for screening bioactive compounds during morphology-based assay as a supplemental strategy for drug discovery[86]. While the bacterial transcription targeting inhibitors have been studied and their mechanisms of actions have been established, it is necessary to employ cell morphology imaging of bacteria carrying the targeted protein to understand the effect of the inhibitors on the proteins[87]. Elucidation of mode of action of antibacterial agents is a critical step in novel antibacterial drug development as the elucidation information allows anticipation of clinical safety or bacterial resistance-related problems[88], which can be achieved by observation of morphological change of treated bacterial cells in association with specific protein targets using advanced microscopic technology. This approach is straightforward in investigation of antibacterial mode of action and confirmation of suspected mechanism as it only requires the access to an electronic microscopy. Morphological changes in cell integrity, membrane permeability, cell structures and growth patterns can be observed upon the application of bacterial transcription targeting inhibitors through kill-time experiments. These allow in-depth understanding of mechanisms of actions of new drugs and molecular events in treated bacteria[88, 89].

Bacterial cell morphology is dependent on identity of antibacterial drug, concentration applied and exposure time. They also depend on the test microorganism and incubation conditions (growth medium, temperature etc)[90]. Other factors include the structure of bacterial cell walls, species and characteristics of test strains, inoculum density, growth phase which also affects the cell size, shape and cell wall thickness in the presence or absence of antibacterial drugs, and homogeneity of morphological changes of treated cells etc[88]. Reduction in distribution of RNAP signal is one

of the cell morphologies. After the application of protein inhibitors, delocalization of green fluorescent protein (GFP) signal as a result of RNAP delocalizing around nucleoid can be observed. This observable change serves as an indication of direct interaction between antibacterial drugs and the bacterial RNAP[88, 91, 92].

Fluorescence microscopy has been employed in the observation of morphological change of antibacterial agent treated bacterial cells. Confocal microscopy is employed because it is capable of generation of high resolution and 3-D reconstruction of cell[93]. It is also advantageous over traditional widefield optical microscopy because of its ability to control field depth by adjusting size of pinhole and removal of background signals. Observation of changes in fluorescence signal of GFP-tagged protein in bacterial cell allows the evaluation and confirmation of modes of actions of bacterial transcription inhibitors. Strains of *B. subtilis* containing GFP fusion to RNAP β' subunit was treated with bacterial transcription targeting inhibitors at different levels of MICs. The fluorescence signals of the treated cells were obtained, where the signals indicated the morphological changes of the corresponding transcription factors.

This method serves as a proof that C3 derivatives had the mechanism of action consistent with the inhibition of bacterial RNA synthesis through the interaction with RNAP β' CH region.

1.3.2.1 Materials

1.3.2.1.1 Chemicals

Cells of *B. subtilis* BS1048 (*RpoC-gfp*) was supplied in streak plates from the collaborator from The Chinese University of Hong Kong and was stored at -80 ℃. LB agar was supplemented with

0.5% chloramphenicol. Luria Broth (LB) was supplemented with 5 µg/L chloramphenicol. Stock solutions of C3 derivatives (**87**, **89**) were prepared in 10x MIC in acetonitrile. Microscopic plates were supplemented with 1.2 % (w/v) agarose pads prior to microscopic analysis.

1.3.2.1.2 Laboratory equipment

Fluorescence confocal microscope (Leica TCS SPE Confocal Microscope) with 63x/1.3 oil objectives, 488 nm laser, PMT detector, mercury metal-halide bulb and FITC filter was employed for fluorescence imaging.

1.3.2.2 Method

1.3.2.2.1 Procedure

Stock *B. subtilis* BS1048 was transferred onto LB agar and cultured at 37 °C for 16 to 20 hours. Single colony was picked and transferred in LB and grown at 37 °C until $O.D._{600}$ ~0.6. Stock solutions of C3 derivatives (**87**, **89**) were added to cell culture at different MICs. The culture was allowed to incubate for further 15 minutes. 2.5 µL of bacterial cell was added onto the agarose pad, covered with coverslip, and sent for microscopic work immediately.

1.3.2.3 Data treatment

Confocal fluorescence microscopy was controlled by LAS AF software and a motorized focus drive. The fluorescence images were then touched up with LAS X software for manual cropping and light contrasting.

1.4 Result

B. subtilis BS1048 contains a fusion of GFP to *RpoC* gene and represents the β' subunit of core RNAP, which serves to examine the localization of RNAP in bacterial cell and the effects of the inhibitors to the bacterial RNAP. Confocal fluorescence microscopy was employed to examine the cell morphology of the C3 derivatives **87** and **89** treated cells. Microscopic images of bacterial cells are presented below.

Figure 15 Control fluorescence image of *B. subtilis* BS1048 which fusion of GFP-tagged RNAP β’ subunit. The cells appeared to be long filaments in great amount. Localization of GFP signals that represent the RNAP in nucleoids can be observed.

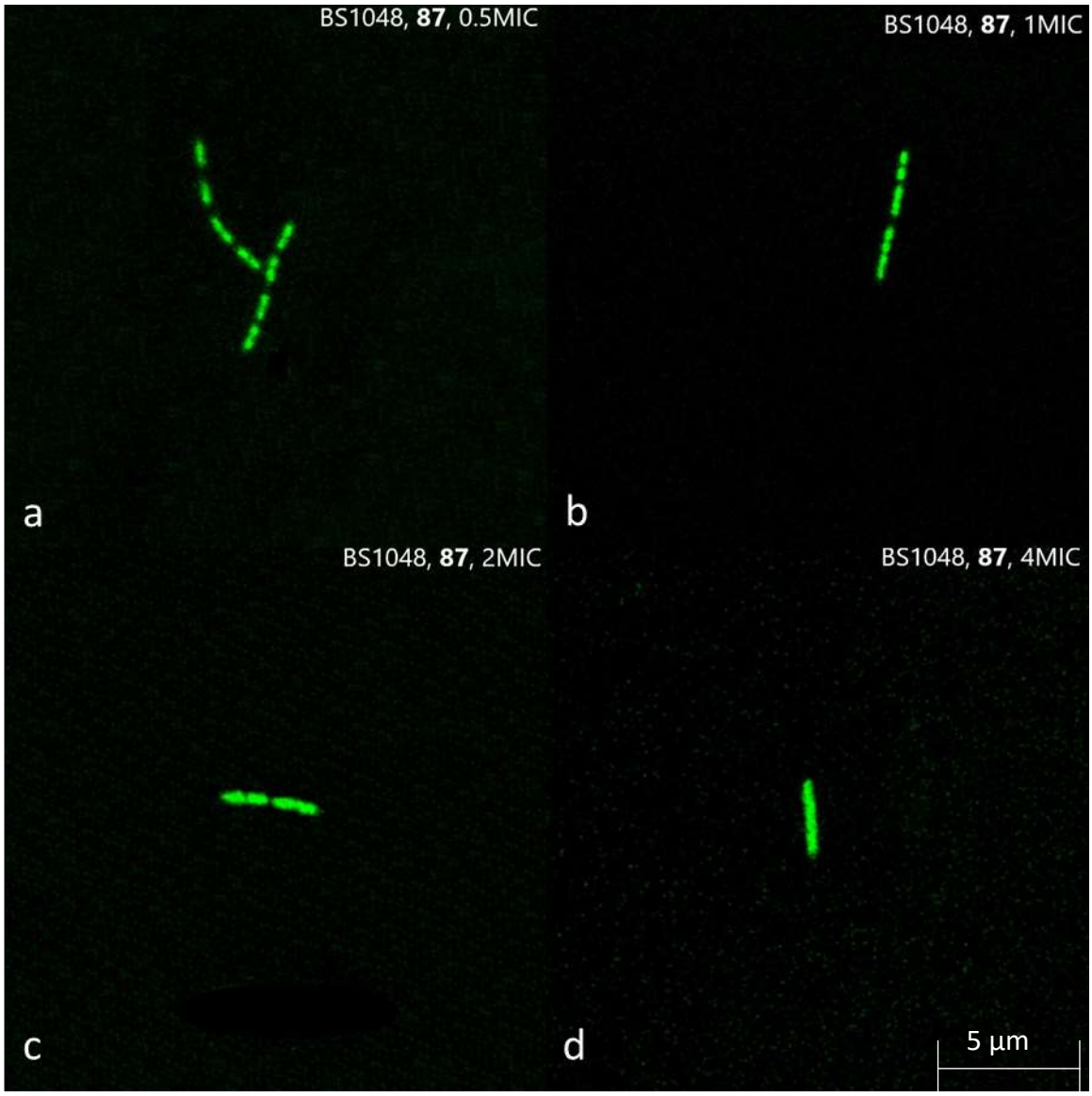


Figure 16 Fluorescence image of GFP-tagged RNAP β' subunit in *B. subtilis* BS1048 with compounds **87** applied at (a) 0. 5 MIC, (b) 1 MIC, (c), 2 MIC and (d) 4 MIC.

The control fluorescence image of *B. subtilis* BS1048 (figure 15) presents the distribution of GFP signal of RNAP without treatment of any inhibitor. It was found that RNAP was concentrated within the nucleoid and the two ends but towards center of the cell. Cells existed as long filaments. The location of GFP signals represents the RNAP β' subunit, the site of RNA synthesis. Figure

16(a) is a fluorescence image of *B. subtilis* BS1048 treated by C3 derivative **87** at 0.5 MIC. High quantity of localization of GFP signals in the nucleoids could be observed. There were only limited number of cells showing delocalization of GFP signals, indicating that the RNAP was slightly disrupted and dissociated around the nucleoids. Limited short filaments of cells had also started to appear. Figure 16(b) shows the image of *B. subtilis* BS1048 treated with 1 MIC of **87**. Amount of cell showing localization of GFP signal at nucleoids decreased while others showed dissociation of GFP signal. The number of short filaments increased under the treatment at this MIC. Figure 16(c) shows the image of cells treated with **87** at 2 MIC. High quantity of delocalization of GFP signals were observed while only small number of cells had remained in long filamentous form. The GFP signals in cells became fully delocalized after treatment of **87** at 4 MIC as shown in figure 16(d). Long filaments of cells no longer existed.

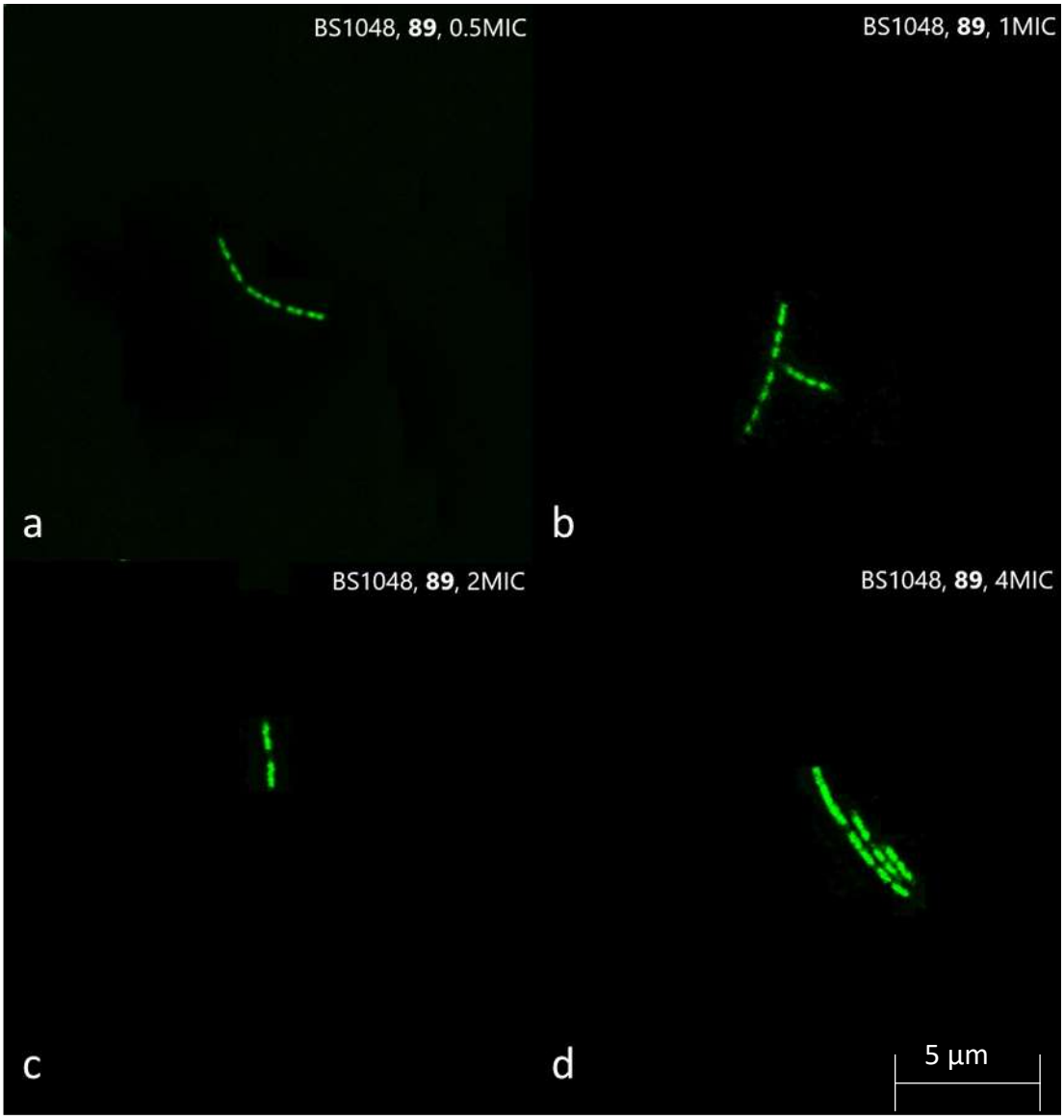


Figure 17 Fluorescence image of GFP-tagged RNAP β' subunit in *B. subtilis* BS1048 with compounds **89** applied at (a) 0. 5 MIC, (b) 1 MIC, (c), 2 MIC and (d) 4 MIC.

Figure 17(a) shows a fluorescence image of *B. subtilis* BS1048 treated with **89** at 0.5 MIC. High quantity of localization of GFP signals in nucleoids could be observed. The cells appeared to be all filamentous. There were increasing delocalized GFP signals at the nucleoids of *B. subtilis* BS1048 treated by **89** at 1 MIC as shown in figure 17(b). Filaments were also shortened. At 2 MIC

(figure 17c), most GFP signals in nucleoids were delocalized. Up to the treatment with 2 MIC, the cells had remained in long filamentous form while at 4 MIC (figure 17d) all cells showed full delocalization of GFP signals and the cells became short filaments.

1.5 Finding and discussion

The fluorescence imaging technique was employed to study the mechanism of action of C3 derivatives for their inhibition of bacterial transcription initiation. *RpoC* gene is a DNA-directed RNAP β' subunit. The fusion of GFP to *RpoC*-gene represented the β' subunit of bacterial core RNAP. Since the CH region of this subunit is the most prominent region for binding with $\sigma_{2.2}$, it was hypothesized that the $\sigma_{2.2}$ would form a holoenzyme with the core RNAP through the interaction with β' subunit at the CH region. As the C3 derivatives were designed to target specifically the CH region of β' subunit, it was therefore hypothesized that C3 derivatives would form a complex with the RNAP by binding to the CH region of β' subunit in order to prevent the formation of RNAP holoenzyme for transcription initiation.

Two C3 derivatives with promising bacterial inhibition activities were selected, namely **87** and **89**. It was intended to examine their mechanism of action in inhibition of the RNAP holoenzyme formation. In the control experiment, healthy *B. subtilis* BS1048 cells showed two distinct GFP foci at the nucleoids. This indicated the location of RNAP and was the site of promoter recognition for transcription initiation. Cells appeared to be all filamentous. In the treatment with **87**, the cells started to show a small amount of delocalization of GFP signals at 0.5 MIC. Limited amount of breakdown of long filaments was also observed. At this MIC, inhibitors were not concentrated enough to inhibit the RNAP. Therefore, only small proportion of cells were affected by the inhibitors while most of the cells remained unaffected. When the concentration increased to 1 MIC, amount of localization of GFP signal at nucleoids decreased while others showed dissociation of GFP signal. This delocalization represented that part of the cells interacted with the inhibitors and dissociation of RNAP-GFP around the cells was resulted. Moreover, some of the short filaments

were fully dissociated. The shortened cell of *B. subtilis* BS1048 might be a resulting morphology of the increasing MIC of **87**. The dissociation was resulted from the efflux of RNAP into the cell cytoplasm because of the elongation of nucleoid which finally disrupted and released the RNAP. At 2 MIC, amount of delocalization of GFP signals continued to increase. Delocalization could also be observed in some longer filaments. This is because most of the cells interacted with **87** and was disrupted. Finally, the cells were fully dissociated at 4 MIC and most of the cells because short filaments.

The delocalization of GFP signal in RNAP is a resulted of the interaction between **87** and RNAP. This might be caused by the binding of **87** to the β' subunit of RNAP because **87** was designed to specifically target the CH region of β' subunit. At as low as 0.5 MIC, the compound could interact with the RNAP, meaning that **87** had good binding with the β' subunit and good inhibition activity. A similar inhibition pattern was observed in the treatment of *B. subtilis* BS1048 with **89**. Small quantity of GFP was delocalized slight at 0.5 MIC. Increasing delocalization of GFP was observable at 1 MIC, and nearly half of GFP were delocalized at 2 MIC. The GFP signals of cell at 4 MIC were fully delocalized. The only difference between **87** and **89** was that there were still long filamentous cells after the treatment with 4 MIC of **89**. The phenomenon was raised from the fact that **87** had an even lower MIC and a better contact with the targeted CH region of β' subunit than **89** did for the inhibition of bacterial RNAP.

It was noticed that normal morphology of *B. subtilis* BS1048 was disrupted by both inhibitors at > 1 MIC. Delocalization of GFP was the response of the bacterial cell, that the β' subunit of RNAP would be released from the nucleoid and finally caused the elongation of nucleoid. This was similar

to a stringent response. The nucleoids of *B. subtilis* BS1048 in exponential growth showed confinement in spherical shape. Upon the treatment of low MIC, they elongated and gradually faded out toward center. The nucleoid elongation was a phenomenon of latter stage of exponential growth which was resulted from the effect of inhibitor and would usually occur more frequently when it approached the end of the exponential growth. The nucleoid elongation was observed when the inhibitors were applied at above the MIC. This indicated that the C3 derivatives were able to promote the end of exponential growth toward the stationary phase. Besides, when the cells approached their stationary phase, the physical status of the nucleoids changed by elongation such that the protein synthesis was inhibited. The elongation of nucleoids might be a result form the alternation in the number of DNA-binding proteins or the concentration of supercoils. Since the CH region targeting C3 derivatives bind to β' subunit of RNAP, the binding with the σ factor was prevented. Formation of the holoenzyme, promoter recognition and subsequent events were also hindered. This may explain the elongation of nucleoids in *B. subtilis* BS1048 [94]. Additionally, when cells were treated with **87** at 4 MIC, the cells became all shortened while at the same MIC, **89** only caused partial change in cell length.

To prove that the compounds bind to β' subunit of RNAP specifically and do not bind to other proteins, different proteins with GFP tagging was treated with the compounds. For example, *B. subtilis* BS23 is a live bacterial cell with a GFP fusion to ATP synthase. ATP synthase as GFP signal is localized to the cell membrane. The treatment of *B. subtilis* BS23 with colistin, a membrane targeting inhibitor, can cause observable membrane damage. However, when the cells were treated with C3 derivatives, no change in GFP localization was observed because C3 derivatives do not target cell membrane. Since colistin targets cell membrane, delocalization of

GFP signal in *B. subtilis* BS23 was resulted. On the other hand, when *B. subtilis* BS1048 was treated with C3 derivatives, GFP delocalization could be observed as the same morphology when treated with rifampicin, a transcription targeting inhibitor, because C3 derivatives and rifampicin both target the transcription. But when the cells were treated with colistin, no delocalization of GFP signal was resulted because colistin does not target β' subunit of RNAP. This proves that C3 derivatives are active against transcription and can inhibit bacterial transcription by targeting β' subunit of RNAP and have the same mode of action as the transcription targeting inhibitors[66].

It can be concluded that the disruption of GFP signal under the effect of C3 derivatives was due to the interaction of compounds with β' subunit of RNAP. Since **87** and **89** were designed to target the CH region, the experimental results also confirmed that the C3 derivatives inhibit transcription initiation by interaction with the CH region for inhibition of core RNAP to prevent the formation of RNAP holoenzyme and hence the promoter recognition in transcription initiation.

1.6 Limitation

This evaluation work is limited by the lack of compound availability and diversity of evaluation methods. Firstly, more compounds with different characteristics such as different linkers or flexibility to dock to protein target should be used to demonstrate the mechanism of action. Secondary, more studies on the protein target can be done for further understanding on the compounds and their mechanism of action or binding affinity etc to the protein.

1.7 Suggestion for future work

The evaluation work on C3 derivatives on their inhibition of RNAP β'CH-$\sigma_{2.2}$ is very limiting. For example, the dissociation constants of them have not been determined. This should be investigated in the future. The change of β' subunit under the influence of the inhibitors is also unknown and can be considered exploring in the future. Besides, the binding sites for the inhibitors on β' subunit are also not confirmed but only estimated from the pharmacophore models. X-ray crystallography can provide a powerful tool for studying the crystal structure of the complexes. The structural information can hopefully be useful for further modification on C3 derivatives for improvement on their binding affinity to the protein. Moreover, up to now there is no related data obtained by advanced mass spectrometry while mass spectrometry has been a widely used analyzing tool for studying protein-protein interaction. Advanced application such as the studying of binding site on the protein through HDX-LC-MS can be considered for a deeper and complementary understanding on the kinetics and mechanism of action of the inhibitors.

1.8 Conclusion

To conclude, the study of bacterial cellular morphology of *B. subtilis* BS1048 (*RpoC*-gfp) suggested that C3 derivatives could target RNAP β' subunit and have consistent mode of actions against formation of holoenzyme as transcription targeting inhibitor.

Chapter 2 Transcription antitermination

2.0 Chapter introduction

This chapter focuses on chemical biology evaluation of inhibitors targeting bacterial transcription antitermination (λN-dependent processive antitermination) and the interaction of antitermination factors NusB and NusE. The chapter starts with an introduction to bacterial transcription antitermination. Next, the importance of proteins NusB and NusE will be discussed. Then, the significance of antitermination complex in transcription antitermination, including the NusB-NusE interaction and the structure of this complex will be delivered. These are followed by the discussion on the inhibition of this complex by targeting the NusB-NusE interaction for the design of novel structure-based inhibitors. A literature review on previous research will then be presented and criticized. It is followed by the methodology chapter, presenting the approaches on evaluation methods and their criticism. Methods and materials, procedures and limitation will then be presented. Results, finding and discussion are followed. Finally, the limitation of the research, conclusion and perspectives will be deliberated.

2.1 Introduction

2.1.1 Transcription antitermination

After initiation of transcription at promoter, the transcript elongates eventually. Transcription termination is the final stage in bacterial genetic expression, which is caused by the pausing of transcription elongation complex. Generally, there are two types of termination, i.e. intrinsic (Rho-independent) termination and Rho-dependent termination. During transcription, a stable transcription elongation complex is formed, which is steadily associated with template DNA and RNA. Nucleotide is added to the transcript one at a time. Once an intrinsic transcription termination or Rho-dependent transcription termination signal is reached, the core RNAP is released from the template DNA, finishing the transcription cycle, and a new cycle of transcription initiation will start.

However, transcription termination is often not efficient, and the expression of downstream DNA can be controlled by altering the efficiency of terminator readthrough. Instead, it is regulated by elongation control in bacterial RNAP. There are two modes of elongation control. The first type involves inactivation of a single terminator by interaction with the upstream sequence in the transcript with a terminator-specific protein or with a translocating ribosome that follows closely behind RNAP. The second type is antitermination of phage λ early transcription (λN-dependent processive antitermination). This involves modification of RNAP into terminator-resistant form after leaving the promoter. The modified enzyme transcribes through sequential downstream terminators and is less sensitive to the pause site where the transcript elongation would usually be delayed. This research studies the proteins involved in λN-dependent processive antitermination. This step is necessary in regulating the successful synthesis of bacterial ribosomal RNA and gives

insight in development of novel antibacterial agents based on the inhibition of antitermination complex formation.

Transcription antitermination is a regulatory process for efficiency rRNA synthesis. This process bases on the formation of antitermination complex which consists of RNAP and antitermination factors to prevent transcription termination at otherwise terminator regions. During the process of transcription antitermination, RNAP, λN protein and other antitermination factors including NusA, NusB, NusE and NusG together form a large nucleoprotein complex. Since this protein is one of the largest bacterial transcription complexes, many researches focus on this structure and the interactions between the elements have been broadly studied[95, 96].

2.1.1.1 Mechanism of transcription antitermination

While the formation of antitermination complex is a regulatory step during transcription termination to ensure the transcription to proceed successfully, the mechanism of transcription antitermination is discussed here. Antitermination is an important event for transcription regulation in both eukaryotes and prokaryotes, which involves the interaction of protein host factors with RNA and RNAP transcription complex to allow transcription through early termination site. In λN-dependent antitermination, λN employs Nus factors (NusA, NusB, NusE and NusG) when recognizing mRNA transcript in N-utilization (*nut*) site. N protein and Nus factors then associate with *nut* site RNA and the RNAP complex to stabilize transcription. The λ *nut* sites consist of a short-conserved single-stranded 12-nucleotide called *boxA* RNA and a hair pin structure called *boxB* RNA. N protein recognizes the site of *boxB* RNA and associates with RNAP through NusA while NusG binds to NusA and RNAP and overrides defective antitermination caused by NusA

mutation. NusB interacts with *boxA* RNA in the *nut* site and NusE on the other hand interacts with NusB. NusB, NusE and *boxA* RNA together form a NusB-NusE-*boxA* RNA ternary complex while NusE is responsible for the complex stabilization.

These Nus factors participate in λN antitermination and ensure the longevity and long-range efficiency to the antitermination complexes. They form an associated antitermination system that triggers the transcriptional regulation of rRNA (*rrn*) operons. NusB binds to NusE and *boxA* RNA while NusE enhances this interaction. This forms a critical interaction NusB-NusE-*boxA* RNA which is vital to the transcription antitermination. The antitermination complexes rely highly on intricate protein-protein and protein-RNA interactions while NusB-NusE interaction and their association with *boxA* are the most dominant of all in bacteria. All these give insights into further understanding on gene regulation.

2.1.2 Antitermination factors

Proteins NusB and NusE are the two antitermination factors in focus in this research. A series of MC4 derivatives were successfully synthesized and it has been confirmed with antibacterial activities. These inhibitors mimic the interaction of NusE with NusB. Several means of chemical biology evaluation of these bacterial transcription targeting inhibitors that target on NusB were conducted in this research.

2.1.2.1 NusB

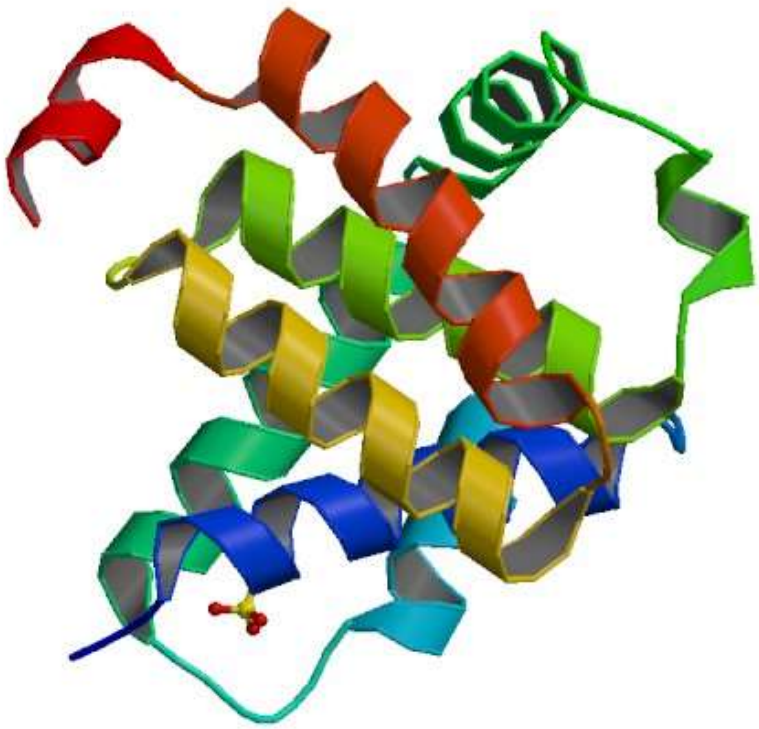


Figure 18 Crystal structure of *E. coli* NusB (PBD entry 1TZT)[97]

NusB (N-utilization substance B) is a bacterial transcription antitermination factor that is strictly required for the formation of antitermination complex which is used in transcription of ribosomal RNA (rRNA). NusB is a highly conserved essential bacterial protein. It has been shown that NusB presents all-helical fold. Crystal structures of NusB from *Escherichia coli* (*eco* NusB) (figure 18) suggest that it is monomeric and that from *Mycobacterium tuberculosis* (*mtu* NusB) is dimeric while the crystal structures of NusB from *Thermotoga maritima* (*tma* NusB) reveal that the protein shows a monomer/dimer equilibrium with a preference for the monomer. These studies on crystal

structures of NusB are valuable in determining its binding site on RNAP and its mechanisms on transcription antitermination.

NusB can bind to *boxA* alone but the interaction is significantly enhanced by the presence of NusE. NusB initially interacts with transcription factor NusE (to be discussed below) to form NusB-NusE heterodimer which then binds to RNAP for rRNA synthesis in bacteria. Once the protein-protein interaction is formed between NusB and NusE, the complex binds to *boxA* RNA through NusB, resulting in a NusB-NusE-*boxA* RNA ternary antitermination complex. It was suggested that the interaction between NusB and *boxA* may be of certain importance because it prevents the inhibition of an antitermination inhibitor. Yet only NusB was found essential in antitermination.

2.1.2.2 NusE

NusE is another transcription antitermination protein. It is otherwise known as S10 of the 30S ribosome subunit. It is only partially folded in the absence of ribosome and has limited solubility. This protein plays several roles in different processes including transcription and translation mutually exclusively[98]. NusE interacts with NusB to form heterodimer. It also increases the affinity of *boxA* RNA to NusB in the heterodimer. Other than that, it stabilizes the NusB-NusE-*boxA* RNA complex for their further association with RNAP through NusE[40]. However, it was confirmed that the interaction between NusE and *boxA* RNA in absence of NusB is nonspecific. The interaction between NusB and NusE have been intensively studied and the MC4 derivatives that were used in this chapter mimic the interaction of NusE to NusB for the formation of a NusB-inhibitor complex for inhibition of formation of NusB-NusE heterodimer.

2.1.3 Antitermination complex

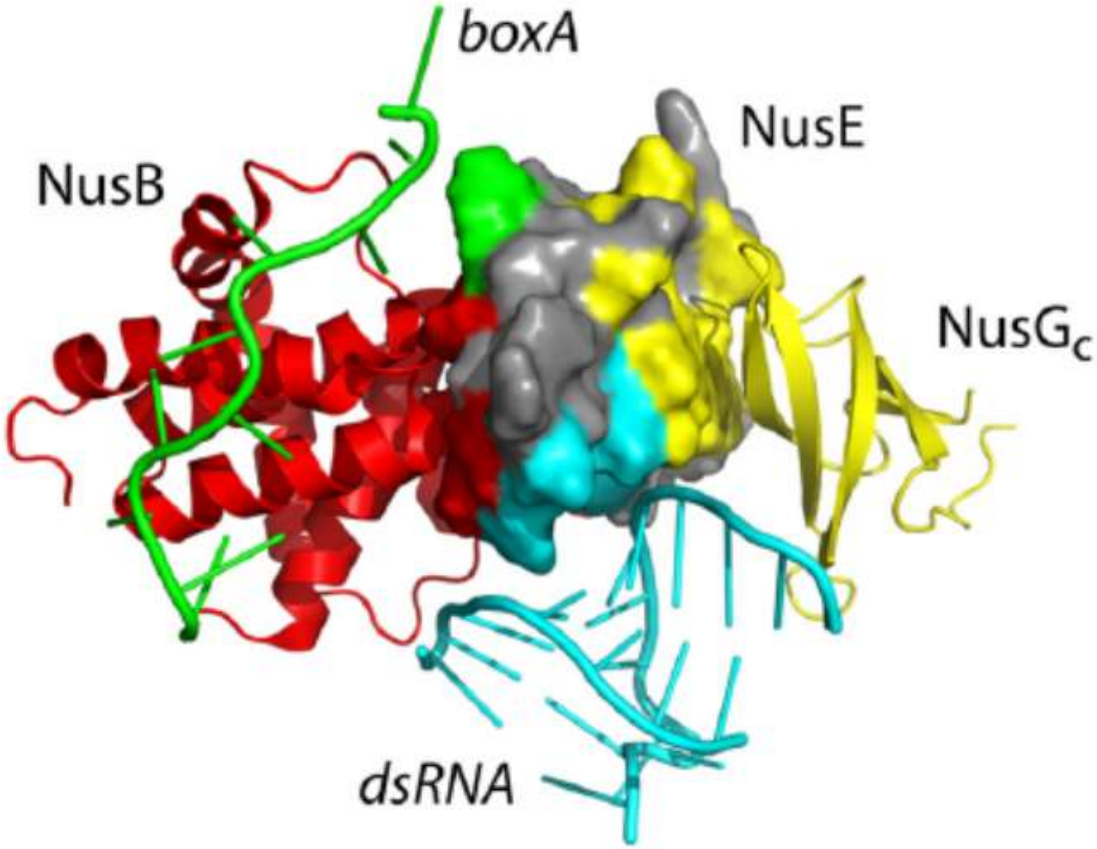


Figure 19 Crystal structure of antitermination complex. NusB (red) and NusE (grey) form heterodimer and then bind to *boxA* RNA (green) to form a ternary complex[19]

Processive transcription antitermination requires the construction of a complete antitermination complex (figure 19). This complex is indeed initiated by the formation of ternary NusB-NusE-*boxA* RNA complex. Firstly, NusB and NusE bind to form a heterodimer, which then bind with *boxA* RNA to form a ternary complex. The λ antitermination complex composes of phage N protein, N-utilization (*nut*) RNA control sequences *boxA* and *boxB*, N-utilization substances (also known as host proteins including NusA, NusB, NusE and NusG). At the beginning of the formation of

antitermination complex, on one hand, NusB and NusE bind together, yielding a hosting heterodimer. *boxA* RNA and *boxB* RNA facilitate the antitermination through protein-RNA interaction. As NusB and NusE interaction is formed, *boxA* RNA (a newly transcribed conserved RNA sequence in the rRNA operon leader sequence) binds to the heterodimer through NusB and a ternary complex is resulted. On the other hand, stem loop of *boxB* RNA binds to the phage N protein. The two RNA sequences are separated by a spacer to which NusA binds.

The construction of λ antitermination complex is initiated when NusB and *boxA* RNA associated as the *boxA* sequence is transcribed. The occurrence of NusB-NusE heterodimer facilitates and strengthens their binding to other elongation components while that of NusE increases the affinity of RNA to NusB in the complex for efficient antitermination.

2.1.3.1 NusB-NusE interaction

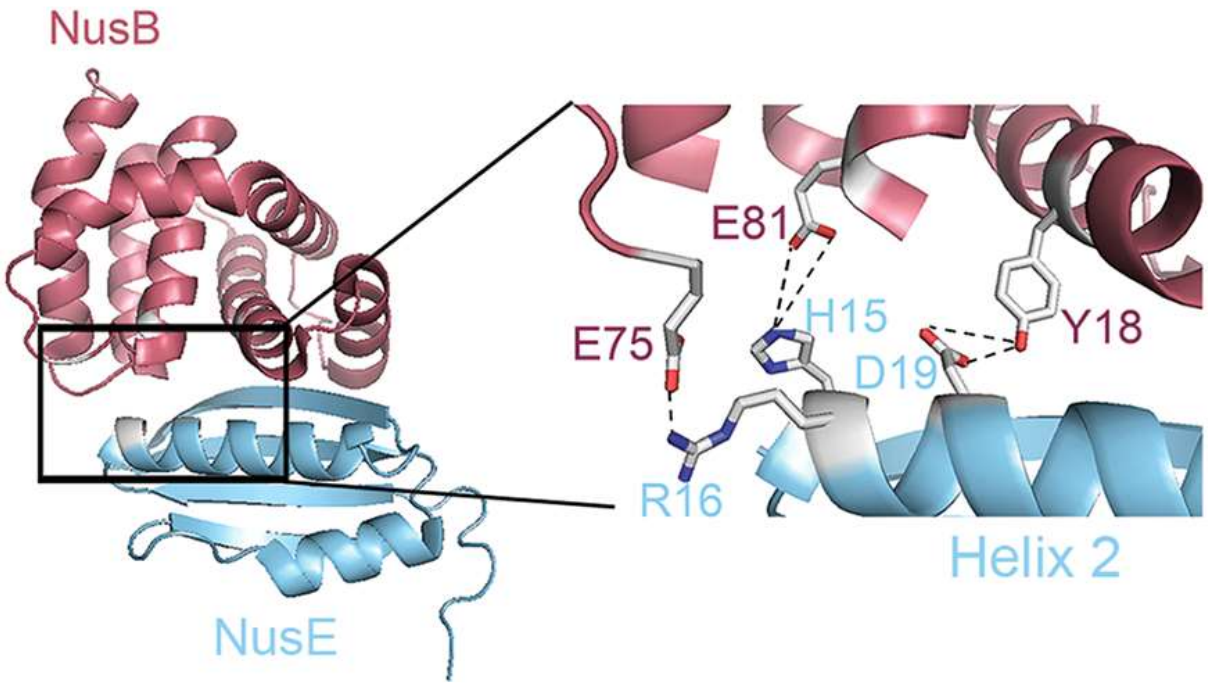


Figure 20 Crystal structure of NusB-NusE(S10) heterodimer (PDB entry 3D3B) (left) and zoom-in showing the conserved interface between the two proteins (right) [40]

NusB-NusE interaction is crucial for the formation of antitermination complex for the regulation of stable bacterial RNA transcription. Figure 20 shows a NusB-NusE heterodimer with the interaction between the two proteins and from the structural data a pharmacophore model was built. According to the published crystal structures of *E. coli* NusB-NusE heterodimer, it has revealed that NusE contains 18% α-helix and it binds with NusB mainly through helix 2[64]. This structure also shows some important hydrogen bonding between NusB and NusE, which is highly conserved in bacteria. Structural information about crystal structure of NusB-NusE heterodimer from *E. coli* is published in Protein Data Bank (PDB) 3D3B, which reveals that NusB contains all helical fold with two perpendicular three-helix bundles (I, II) while NusE contains a four-stranded antiparallel β sheet with two α-helices one at the back of each side. α1 and β2 of NusE connect the two helical

bundles of NusB. The contact regions of the two helix bundle I and II of NusB are located at the flank of the first three helical bundle I and on a tip of the second three helix bundle II. Inside the heterodimer there is a complementary electrostatic surface between the two proteins, which comprises a combined surface area upon the complex formation. NusB and NusE interacted through hydrophobic and hydrophilic interactions. For instance, hydrogen bonds are formed between Asp19-Arg72 ion pair of NusE with Tyr18 of NusB and hence positioning Try18 between Pro39 and Pro41 of proline element on β2 of NusE[99].

The interruption of the NusB-NusE interaction with a NusB-targeting small molecule results in mutations and thus the reduction of protein-protein binding affinity and the formation of antitermination complex. The rate of bacterial rRNA synthesis is therefore reduced. A dissociation constant K_d = 1.1 ± 0.1 μM of NusB-NusE interaction was reported and this titration data could be fitted to a single binding site model[40].

2.1.4 Inhibition of transcription antitermination targeting NusB-NusE interaction

Antibiotic resistance has been a global health problem due to the high rate of bacteria acquiring resistance against existing antibiotics so the impossibility to effective treatment of bacterial infections is foreseeable. It has been predicted that over ten million of deaths will be caused by antibiotic-resistant infections without effective treatment if there is no intensive effort to novel antibacterial drug discovery and development by year 2050[100, 101]. The fundamental role of bacterial RNAP has been an attractive drug target for development of new antibiotics. Even though only rifamycin can enter clinic use, there are many other RNAP targeting inhibitors being discovered and waiting to be investigated. The NusB-NusE interaction is a valuable novel target for antibacterial drug discovery because it involves two transcription factors that are essential for

formation of a critical antitermination complex in rRNA synthesis, without which essential rRNA synthesis is inhibited in bacteria.

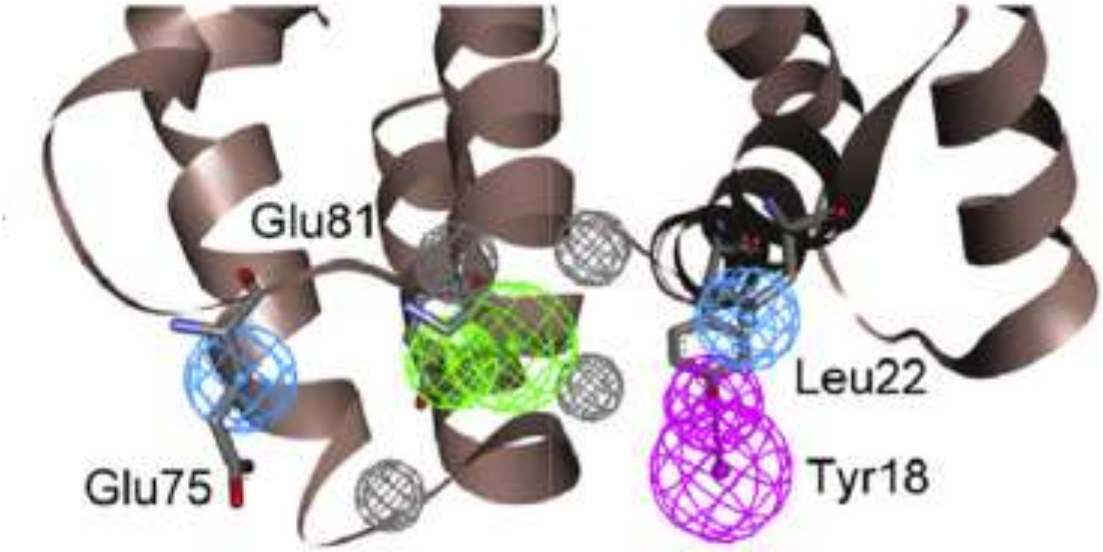


Figure 21 Pharmacophore model of NusB constructed based on the interface [99]

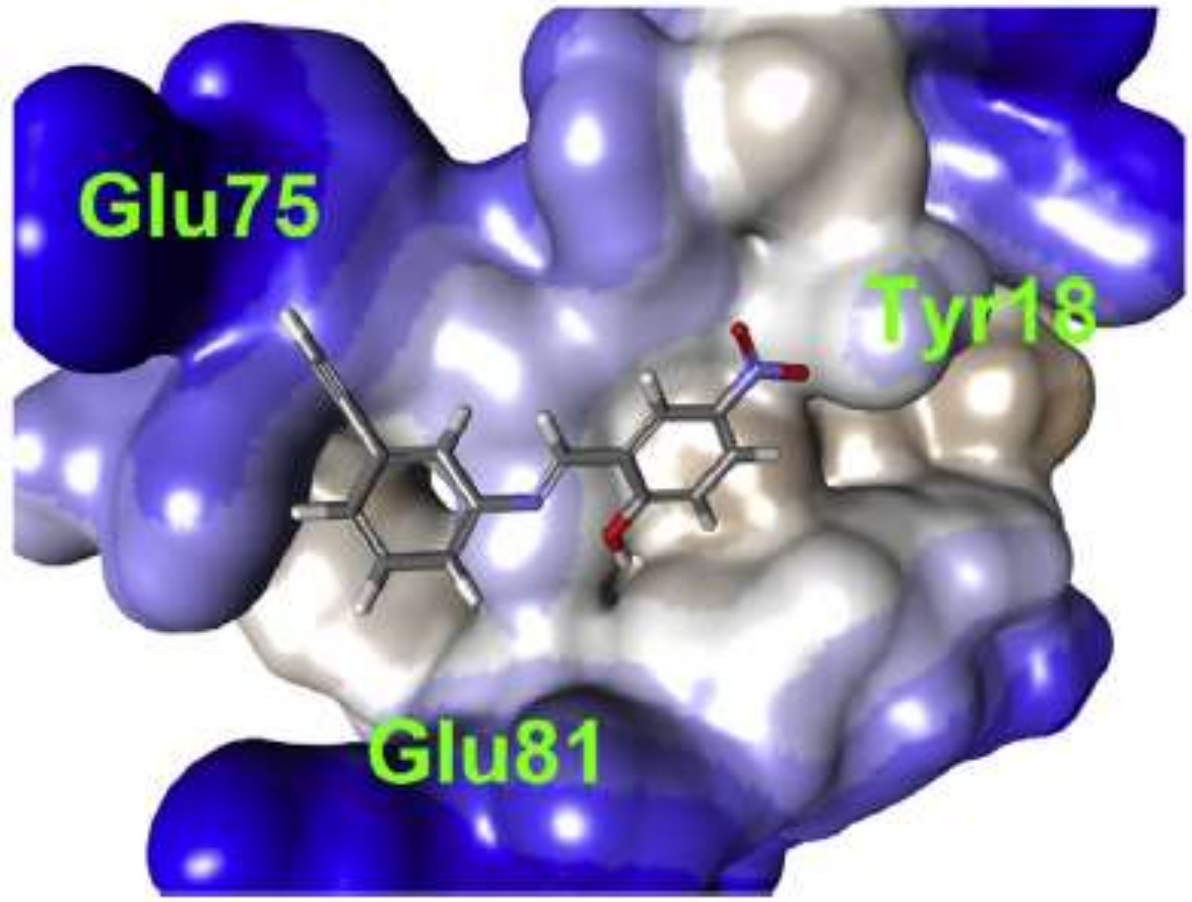


Figure 22 A docking model showing interaction between lead compound MC4 and NusB surface. The hydrophilicity of NusB is shown in blue.

As NusB-NusE interaction has been identified as a new target for antibiotic drug discovery, a pharmacophore model has been developed based on the interface of NusB (figure 21). The docking model (figure 22) shows important interaction between NusB and MC4. Because NusE binds to NusB with a one site binding mode, the inhibitors will compete with NusE for binding and that a complex of protein-ligand will be resulted. Once a NusB-inhibitor complex is formed, the number of NusE binding to NusB is supposed to be reduced and hence the reduction of number of NusB-NusE complex will be resulted. It has been demonstrated that the disruption of NusB-NusE heterodimer formation or interaction with *boxA* or inhibition of NusB-NusE-*boxA* are potential

target for novel antibiotics development. While the inhibitors disrupt NusB-NusE interaction, the events following the formation of NusB-NusE heterodimer cannot occur normally, for example, cooperation with NusG in formation of antitermination complex for connecting transcription and translation in bacteria. Therefore, it is hypothesized that the small molecules have antibacterial activities and can target the NusB-NusE interaction. The small molecules binding to the NusB interface can interrupt the NusB-NusE interaction and cause mutation to the protein-protein interaction and hence interrupt the formation of antitermination complex and reduce the rate of rRNA synthesis in bacteria.

2.1.5 Contribution of structure-based design and biochemical evaluation on antibiotics targeting NusB-NusE interaction

Because of the fast acquisition of resistance in bacteria, new targets must be identified for novel antibiotic development. NusB-NusE interaction is an underutilized target for development of new antibiotics[19]. There are no known antibiotics that target on this interaction in the market yet. Research and available data on the compounds that target this interaction are also limited. Since the prokaryotic RNAP is highly conserved and the synthesis of all RNA in bacteria requires only one RNAP, inhibitors with antibacterial activities that target specifically on this interaction can be potentially developed into new broad-spectrum antibiotics.

A series of MC4 derivatives have been successfully synthesized[65, 102], which mimic the NusE interaction to NusB, and some of them show relatively outstanding bioactivities. These derivatives have been selected and applied to the protein NusB to evaluate their activities with the protein. It is proposed that the compounds have good binding to NusB and that they form NusB-inhibitor complexes. Studies on these complex and the change of protein would be of importance in further

understanding the effect of this transcription targeting inhibitor and mechanism of action of the inhibitors on NusB.

2.2 Literature review

This chapter reviews the previous studies on crystal structures of NusB protein as well as related studies on NusB-NusE interaction. Reviews on inhibitors that are able to target and inhibit NusB are also done here.

Crystal structures of transcription antitermination factor *eco*NusB in monomeric form from *Escherichia coli*[38] and *mtu*NusB in dimeric form from *Mycobacterium tuberculosis*[103] have been resolved previously. Five crystal structures of NusB from *Thermotoga maritima* (*tma*) were determined and their corresponding crystallization conditions were described. Three out of five structures were monomeric while the others were dimeric. Solution studies of *tma* NusB showed that monomeric form was preferred. From the crystal structures, it was hypothesized that NusB was inactivated through dimerization until it was available for antitermination complex formation[97].

The crystal structure of transcription-active NusB-NusE antitermination complex was also obtained in form of NusB-S10 complex. In this complex, all helical folding with two perpendicular three-helix bundles was observed in NusB while a fur-stranded antiparallel β sheet backed by two α helices was observed in S10 (NusE), in which helix α1 and irregular β2 strand connect the two helix bundles of NusB. Structural data revealed that NusB-S10 complex exhibited enhanced

affinity for *boxA* RNA. Moreover, globular structure of NusB and NusE were conserved in complex and they adjusted by local induced fit during complex formation.[99].

NusB-NusE heterodimer sequentially binds to *boxA* RNA sequence for completing the ternary NusB-NusE-*boxA* RNA antitermination complex. The crystal structure of this complex was obtained and demonstrated that the interaction between NusB and *boxA* proved the significance of nucleotide-sequence specificity. Since the direct interaction of NusE to *boxA* was restricted, it shed lights on the importance of NusE in initiation of antitermination in which additional antitermination complex components are required. The binding site of *boxA* RNA on NusB-NusE heterodimer was also determined[98].

Furthermore, to confirm the interaction between antitermination factors and RNAP, Drogemuller et al. performed 1D HSQC. Addition of RNAP and β to NusB:^{15}N-NusE resulted in loss of signals in 1D [^{1}H, ^{15}N]-HSQC, indicating that RNAP and β bind to NusE. Addition of RNAP to ^{15}N-NusB: NusE resulted in unchanged spectrum, indicating that NusB was not released upon binding of NusE to RNAP and confirmed that NusB could not bind to RNAP alone. The resulted concluded that NusB-NusE complex binds to bacterial RNAP through NusE with β subunit of RNAP being the major target[104].

In a structure-guided study of interaction of NusB and NusE, crystal structure of λN–NusA–NusB–NusE–*nut* RNP (RNA-protein complex) was obtained to probe the stability of individual interfaces and importance of contact-mediating residues for RNAP. It was concluded that NusE alone would

assemble with NusA and λN on *nut* RNA while NusB in the absence of NusE lost its assembly. This result indicated that NusE is required for stable integration of NusB[105].

In view of the importance of NusB-NusE antitermination complex for bacterial transcription, mimicking the interaction between NusB and NusE highlights the development of novel bacterial transcription inhibitor with new modes of action. The discovery of bacterial ribosomal RNA synthesis inhibitors with specific antibacterial activity against methicillin-resistant *Staphylococcus aureus* (MRSA) was reported, in which a pharmacophore model based on the protein-protein interaction between bacterial rRNA transcription factors NusB and NusE was constructed and devoted for screening to identify the lead compounds. It was reported that (E)-2-(((3-ethynylpheny) imino) methyl)-4-nitrophenol (MC4) demonstrated antibacterial activity against several *S. aureus* strains with minimum inhibitory concentration (MIC) valued 8 μg/mL and 16 μg/mL against MRSA with low mammalian cytotoxicity and the rRNA level in bacteria was successfully reduced by MC4. MC4 inhibited NusB-NusE heterodimer formation at molecular level with high specificity, with IC_{50} ~35 μM. Moreover, analogues of MC4 revealed the three functional groups on the inhibitors that targeted interactions between NusB E81 and NusE H15, NusB Y18 and NusE D19, and NusB E75 and NusE R16 were necessary for the inhibition of NusB-NusE interaction. The binding affinity of MC4 to NusB was reported (1.45 ±0.55 μM) with a one-site binding mode. In addition, MC4 was unable to bind with NusE nor could it bind with NusB variants with the three important amino acid residues accountable for NusE interaction being substituted with alanine. This confirmed that there is specific interaction between NusB and MC4. The results validated the bacterial rRNA transcription machinery as a novel antibacterial target which could be valuable for development of novel bacterial transcription inhibitors targeting other transcription

factors. Pharmacophore model of MC4 docking revealed that the model comprised of two hydrogen donors and one hydrogen acceptor to mimic the major hydrogen bonding between NusB and NusE. Additionally, there was one conserved hydrophobic interaction between NusB and NusE. Exclusion zones also existed to minimize steric clashes between shallow pockets that formed binding site of NusB. Small molecules were theoretically capable to dock into this pharmacophore model and demonstrate inhibitory effects as the model was constructed based on the significant amino acid residues on NusE responsible for binding to NusB[64].

The NusB-NusE protein-protein interaction was also studied by Cossar et al[37]. Another series of compounds with that target NusB-NusE interaction were synthesized. One of the compounds was found to have MIC value lower than 3 μg/mL against gram-positive *Streptococcus pneumonia* and MRSA, and lower than 51 μg/mL against gram-negative *Pseudomonas aeruginosa* and *Acinetobacter baumannii*. An epifluorescence study on strain *B. subtilis* BS61 (NusB-GFP with a signal restricted to the sub-nucleoid foci that represent the site of rRNA synthesis) was done to confirm this compound had a mechanism of action coherent with the inhibition of rRNA transcription through targeting NusB-NusE interaction by treating the live bacteria cell. The treatment of *B. subtilis* BS61 with this compound caused a significant delocalization of NusB-GFP signal around the cell. This signal indicated the loss of rRNA transcription in bacterial cell. In contrast, another strain *B. subtilis* BS23, which contains GFP fusion to α subunit of the membrane-localized ATP synthase, was also treated with the compound but no effect on ATP synthase localization was observed. The study revealed that the compound inhibited NusB-NusE PPI as proposed but not targeting α subunit nor other cell components. The result confirmed that the compound had a mechanism of action consistent with the inhibition of rRNA transcription through

targeting NusB-NusE interaction. This compound has become a potent lead for a novel antibacterial target.

The latest derivatives of MC4, new analogues targeting NusB-NusE interaction, were synthesized by Qiu et al[102] to study the structure-activity relationship of MC4. Inhibitory activities of these compounds were tested and IC_{50} values were measured. Circular dichroism spectroscopy was also employed to detect the influence of representative derivatives on NusB folding. Binding of these compounds to NusB did not show change of secondary structure, as the spectra of the complexes and protein were of a typical α-helical structure. Besides, the CD spectra showed that the compounds caused conformational changes of NusB in similar fashion, which suggested that the compounds may bind to the same binding site on NusB. One of the representative compounds even showed the best MIC as 1 μg/mL. With other biological evaluation of the novel MC4 derivatives, it was concluded that these compounds represented a new class of antimicrobial agents against an underutilized PPI in bacteria with a different mechanism of action to current antibiotics.

The structurally modified derivatives of MC4 were synthesized based on the NusB-NusE interaction but their interaction with NusB was not studied. Although the crystal structure of NusB, NusB-NusE and NusB-NusE-*boxA* antitermination complexes were available and the bacterial transcription machinery was extensively studied, protein crystal of NusB-inhibitor (MC4 or derivatives) complex has not been obtained, and no information on structure of the complex is available for understanding of the interaction of inhibitors with NusB. The binding site for the inhibitors on NusB is also yet to be confirmed. Therefore, it is essential to obtain the protein crystal

of NusB-MC4/derivative complexes to complete the comprehensive information on the interaction between NusB and the transcription targeting inhibitors.

Although the crystal structures of NusB and NusB-NusE antitermination complex have been obtained and hence the binding site on NusB is estimated, the binding site of NusB is yet to be confirmed because of the lack of crystal structures of inhibitors binding to NusB. Moreover, the modes of action of MC4 derivatives are yet to be confirmed. It also lacks the data on binding affinity of the compounds to NusB. These will be addressed in this research.

2.3 Methodology

2.3.1 Introduction

In this section, the framework of research will firstly be introduced together with the research questions. Next, the approach of the studies will be discussed, following by the justification of the approach. Then, the methods of the studies, including the sampling techniques, data collection and procedures will be presented. Finally, materials of the studies including laboratory equipment will be given.

A novel class of bacterial transcription targeting inhibitors targeting bacterial transcription antitermination subunit NusB has been synthesized and proven to be bioactive. Although the developed techniques of synthesis and structural optimization of these drug analogues have been employed to achieve the optimal activities against bacterial targets and the bioactivities of them have been tested, chemical biology evaluation of the analogues are yet to be done for further development of the compounds into potential drug candidates. A novel class of bacterial transcription targeting compounds is synthesized based on MC4 which is a lead compound that targets bacterial NusB-NusE protein-protein interaction. While the bioactivity test from previous studies showed that analogues **61** and **84** of MC4 showed the lowest minimum inhibitory concentration (MICs), the binding affinities of these drug analogues are not known. The effect of drugs on bacterial cellular morphology are not studied. Even though computer simulation allowed the chemical docking of compounds into target proteins and structural optimization has been done based on the MIC and docking model, the actual binding sites on the proteins of interest remain undetermined. Moreover, it is anticipated to obtain the protein crystal structures of the protein-drug complexes.

Herein, several chemical biology evaluations of the synthetic drug compounds are exercised for a complete understanding of their effects on bacterial transcription targets. These include the determination of binding affinities of the compounds, effects of compounds on the secondary structures of protein, observation of bacterial cellular morphology of compound treated cells and acquirement of protein-drug complex crystals for the determination of binding site of the proteins.

2.3.2 Protein overproduction and purification

B. subtilis NusB-His was used in circular dichroism analysis, ITC assay and mass spectrometry. *B. subtilis* BS61 contains a green fluorescent protein (GFP) fusion to NusB. *Thermotoga maritima* (*tma*) NusB was used in protein crystallization. The following shows the general overproduction procedure of the NusB.

B. subtilis NusB

B. subtilis NusB was overproduced and purified using the following approach: *E. coli* BL21 was transformed with plasmid DNA containing ampicillin resistant gene and NusB with 6x His-tagging. Bacterial cultures were grown in autoinduction medium for three days at 20 ^{o}C. The cell was lysed with B-per complete reagent followed by clarification. The His-tagged NusB was purified using a 1ml HisTrap HP column (GE Healthcare). Purified protein was dialyzed into first dialysis buffer and then into second dialysis buffer[37].

Lysis buffer for NusB-His: Buffer A with0.5mg/ml lysozyme, 0.05% Triton X-100, 0.1% protease inhibitor cocktail, and autoinduction medium in 1:1.

Purification Buffer A (binding): 20 mM sodium phosphate, 0.5 M sodium chloride, 20 mM imidazole, pH 8.0.

Purification Buffer B (elution): 20 mM sodium phosphate, 0.5 M sodium chloride, 200 mM imidazole, pH 8.0.

First dialysis buffer: 1x PBS with 0.1 mM DTT and 0.1 mM EDTA.

Secondary dialysis buffer: 1x PBS with 0.1 mM DTT and 30% glycerol.

tma NusB

tma NusB was provided by Dr. Yang group, the collaborator from The Chinese University of Hong Kong.

2.3.3 Binding affinity

One of the important characterizing methods for drug discovery is to evaluate the strength of protein-ligand interaction by measuring the binding affinity of the protein of interest and drug. Binding affinity is the measure of strength of interaction between a macromolecule such as protein and a ligand such as drug or inhibitor. Binding affinity is represented by equilibrium dissociation constant (K_d) which marks the strength of interaction between the two molecules, where a strong interaction is represented by a small K_d value and weak by large K_d value. It is influenced by non-covalent intermolecular interaction such as hydrogen bonding, hydrophobic interaction, electrostatic interaction and Van der Waals' forces between two molecules. This interaction can

also be affected by the presence of another molecules. Two approaches in quantitation of interaction of biomolecular system are taken in this research, namely isothermal titration calorimetry (ITC) and mass spectrometry (MS).

2.3.3.1 ITC

2.3.3.1.1 Introduction

ITC is a premier tool to quantitate the biomolecular interaction by measuring the enthalpy change during a binding reaction. This method can be applied to different biological systems such as protein-protein, protein-ligand, protein-small molecule, protein-lipid, protein-receptor and antibody-antigen etc. Biophysical quantification of interaction between analogue **61** of MC4 and NusB was done using ITC method.

Doyle[106] pointed that one strength of ITC that makes it an accurate and powerful method is that it does not require chemical modification of the macromolecules because physiochemical manipulations may disturb the precise balance and chemical functions of a biological system. Although it can be physically separated by washing and filtering, separation process however would bother the reaction equilibrium. On other hand, ITC allows measurement of binding affinity of the native and unmodified macromolecules in solution, and binding and non-binding molecules in equilibrium. Because of these, ITC became the most precise method for accurate characterization of binding affinity of macromolecular interactions.

2.3.3.1.2 Materials

2.3.3.1.2.1 Laboratory equipment

Malvern MicroCal PEAQ-ITC Automated Ultrasensitive Isothermal Titration Calorimeter was used in ITC analysis.

2.3.3.1.2.2 Chemicals

B. subtilis NusB-His was used in ITC. The protein overproduction and purification were in 2.3.2.

2.3.3.1.3 Method

2.3.3.1.3.1 Procedure

ITC buffer (50 mM KH_2PO_4, 150 mM NaCl, pH 7.4) was prepared. A stock solution of compound (50 mM in DMSO) was diluted to 500 μM ITC buffer so that there was 1% DMSO in ITC buffer. The solvent system of the experiments maintains the same. Protein *B. subtilis* NusB-His was prepared without addition of glycerol in second dialysis buffer during purification and dialyzed into 1% DMSO in ITC buffer. MC4 was titrated against 50 μM *B. subtilis* NusB-His using 1% DMSO in ITC buffer as control. 150 μL sample of NusB or buffer was used in cell while 400 μL sample of compounds or buffer was used in syringe. Blank experiments of buffer against NusB and compounds against buffer were first performed.

2.3.3.1.4 Data treatment

All data obtained from ITC experiment were analyzed with the software Malvern MicroCal PEAQITC.

2.3.3.2 Mass Spectrometry

2.3.3.2.1 Introduction

Mass Spectrometry (MS) has been a powerful analytic tool for characterization of a wide range of biological and chemical systems and its application in drug discovery has been matured and varied. Not only does it allow the measurement of mass of the biological system, it provides means to characterize even more precise analysis of protein such as conformation change, peptide sequencing, binding mechanism etc. Native MS analysis allows the measurement of binding affinity of protein in its native form. Small molecule binding to a protein can be indicated by the increase in mass-to-change (*m/z*) ratio of the protein. The binding affinity of the small molecules to the protein can be interpreted as dissociation constant (K_d) which can be calculated from the ratio of small molecules to protein.

MC4 analogues interact with NusB through hydrogen bonding. When MC4 analogues were added to *B. subtilis* NusB (17121Da) molar ratio e.g. 10:1, the mass of the complex should increase, which indicates the binding of small molecules onto NusB. Since they interact through hydrogen bonding, nanoelectrospray ionization mass spectrometry (nano-ESI-MS) provides a native MS detection method to the complex, where the analytes are detected in their natural environment.

Yet another approach was taken due to the limitations in native MS analysis. Derivatives that are photoactive were synthesized to purposely bind to the protein by creating covalent bond through photo-crosslinking. Since covalently bound protein-ligand complex is stable, non-native MS approach can be adopted. Herein, LC-MS can simply be used for detecting the photo-crosslinked protein-ligand complex.

2.3.3.2.2 Materials

2.3.3.2.2.1 Laboratory equipment

Water Synapt G2-Si High Definition Ion Mobility Mass Spectrometer equipped with nanoelectrospray ionization (nano ESI) source was employed in the native MS analysis. Agilent 6540 Liquid Chromatography – Electrospray Ionization Triple Quadrupole Mass Spectrometer was used in the detection of photo-crosslinked protein-ligand complex.

2.3.3.2.2.2 Chemicals

B. subtilis NusB-His was used in mass spectrometry. The protein overproduction and purification were in presented 2.3.2. MC4 derivatives were synthesized by the research group.

2.3.3.2.3 Method

2.3.3.2.3.1 Procedure

B. subtilis NusB-His was exchanged into 20 mM ammonium acetate prior to MS analysis. Stock solution of inhibitor was prepared in 10x concentration and added to NusB. The mixture was incubated in 37℃ for 20 minutes prior to analysis.

Photo-crosslinking experiment was performed by preparing inhibitors in DMSO and adding to *B. subtilis* NusB in PBS (3:1 molar ratio) in 5% final DMSO concentration. The protein/ligand mixture was transferred into a quartz cuvette (4 x 1 x 1 cm) and irradiated in 365 nm UV for 0, 10, 20, and 30 minutes. Supernatant and precipitate were separated and treated discretely for MS analysis. For instance, supernatant was added formic acid to increase its ionic strength for ESI or

LC-MS. Precipitate was dissolved by urea followed by buffer exchange. Sample was then sent for LC-MS analysis.

2.3.3.2.4 Data treatment

All data obtained from MS experiments were analyzed with the software Masslynx V4.1 and Qualitative Analysis B.06.00. for protein deconvolution and peak recognition.

2.3.4 Secondary structures of protein

2.3.4.1 Introduction

Interaction of proteins with small molecules takes place at specific binding sites. Change in secondary structure of a protein can be observed when the protein is complexed with ligands. It was pointed out that many proteins interacted as a quasi-rigid body and such structure showed very limited conformational change when complexed[107]. Yet, protein structure rearrangement would occur significantly when the binding occurred at the interface or at site away from interface of a protein to yield a stable complex[108].

Studying the secondary structure of protein NusB allowed the understanding of its conformational change when complexed with inhibitors. Since it was predicted that the binding site of NusB was located on the surface[64], such study could also verify this surface binding. Circular dichroism spectroscopy was taken as the approach to study the conformational change of secondary structure of NusB under the effects of ligands. It was assumed that NusB interacted with NusE through the surface so that the small molecule would also bind to NusB through the binding site on the protein surface. It was also assumed that the small molecules bound to NusB in solution during circular

dichroism spectrometry analysis and such method was capable to differentiate the NusB in bound and unbound forms.

2.3.4.2 Circular dichroism

2.3.4.2.1 Introduction

Circular dichroism (CD) spectrometry is a powerful tool used for secondary structure, folding and binding determination of biopolymer such as purified proteins and nucleic acids. It has especially wide application in determination of conformational change of secondary structure of a protein under the effects of temperature, inhibitions, mutations, heat, denaturants, and protein-protein interactions[109]. Insignificant changes in a CD spectrum can be observed when the complexed protein has a rigid structure, and vice versa[107]. Changes can also be observed from different wavelength which indicates the presence of secondary structure components in varying quantity. Therefore, CD can be served as a tool to monitor the change in secondary structure of a complexed protein.

In the analysis of protein secondary structures, CD signals arise from far UV region (240 nm to 190 nm) because of the chromophores of polypeptide backbone. When the chromophores of the amides of polypeptide backbone of the proteins align in arrays, their optical transitions will shift due to excitation transitions[109], i.e. $n \rightarrow \pi^*$ at 222 nm and $\pi \rightarrow \pi^*$ at 208 nm and 190 nm[110]. This electric excitation gives rise to proteins with different structures having their characteristic CD spectra. For example, the spectrum of an α-helical protein has negative peak at 222 nm and 208 nm and positive peak at 190 nm. Well-defined antiparallel β-pleated sheet proteins have negative peak at 218 nm and positive peak at 195 nm. Disordered proteins have negative peak at 195 nm but low ellipticity at above 210 nm [109, 111].

When a spectrum of purified NusB is compared with that of a complex, change in peaks can be compared and secondary structure alternation can be observed. Interpretation of the spectrum provides information for estimation of secondary structural composition of the proteins. Difference in spectra of NusB under the effect of different inhibitors can also be compared.

Nonetheless, CD analysis is only a complementary analyzing tool that provides limiting structural characterization data and cannot replace X-ray crystallography, NMR spectrometry or mass spectrometry analysis of proteins because CD analysis only allows analysis of change of secondary structures but not binding sites nor protein structures. Despite the limitation, it is advantageous over the mentioned techniques because CD can be carried out in a wide range of experimental solution conditions such as temperature and concentration[110]. However, the accuracy of CD of proteins of different structures varies. For instance, 97% and 75% of accuracy can be obtained for α-helices and β-sheets respectively but only 50% for β-turns[111].

2.3.4.2.2 Materials

2.3.4.2.2.1 Laboratory equipment

Circular Dichroism (CD) Spectrometer (Jasco J-810-450S) equipped with fluorescent lamp was used in the CD measurement. A quartz cuvette (4x1x1cm) was used to contain all sample for measurement. A quartz cuvette (4x1x1cm) was used to carry the sample for analysis.

2.3.4.2.2.2 Chemicals

B. subtilis NusB-His was overproduced and purified with the procedure described above. 1x PBS aided with 5% acetonitrile was used as sample buffer. All inhibitors were dissolved in 100% acetonitrile and added to 1xPBS in 50 times dilution before analysis.

2.3.4.2.3 Method

2.3.4.2.3.1 Procedure

A blank with no organic content was first measured by CD spectrometer to confirm the solvent system. *B. subtilis* NusB-His was dissolved in 1xPBS with 5% acetonitrile as control. Circular dichroism spectroscopy of NusB with inhibitors were performed with quartz cuvette and experiment was triplicated.

2.3.4.2.3.2 Data collection

A blank spectrum of PBS in different organic solvents was obtained to test the signal of solvent system. It was found that PBS in 5% acetonitrile gave the best CD signals. A spectrum of *B. subtilis* NusB in PBS in 2% acetonitrile was then obtained. All experiments were triplicated, and all spectrum were obtained after averaging the three spectra.

2.3.4.2.4 Data treatment

BeStSel (http://bestsel.elte.hu) is a freely accessible web server that provides Beta Structure Selection method to analyze CD spectra recorded by conventional or synchrotron radiation CD spectrometers based on the orientation and twist of β-sheet proteins or fibrils. This method provides estimation of β-sheet proteins, membrane proteins, protein aggregates and amyloid, and

α-helix with improved accuracy than previously available methods. Moreover, this method allows provision of detailed information on β-sheets, and differentiation of parallel and antiparallel β-sheets with three different twists (left-hand twisted, relaxed and right-handed twisted). Prediction of protein fold in topology level is also possible with CATH fold classification by BeStSel. Theoretical reliability of BeStSel in protein fold prediction has been shown to be the highest compared to other search engines[112, 113].

CD spectrum of NusB and complex of NusB-inhibitors were obtained and data were processed through BeStSel to obtain a component report generated from the server which provided information on the quality and quantity of secondary structure components of NusB and complexes.

2.3.5 Protein crystal and binding site determination

2.3.5.1 Introduction

Even though NusB has been extensively studied and the protein structure has been determined, the interaction between NusB and MC4 derivatives (i.e. binding site on NusB) has only been predicted but not yet confirmed. Furthermore, only crystal structures of NusB and NusB-NusE complex was obtained previously but not the complexed structure of NusB and MC4 derivatives. In order to evaluate the bacterial transcription inhibitors and have more in-depth understanding on the protein-inhibitor complex, it is obligatory to obtain the crystal structure of the complexes. One approach that was taken to obtain crystal structure was protein crystallization of NusB followed by soaking of the single protein crystal in compounds. As the crystalline complex was able to show diffraction patterns of monochromatic X-ray beams, X-ray crystallography provides means to determine the crystal structures of proteins or complexes.

Protein-ligand complex structure could be determined by X-ray crystallography. The X-ray diffraction of crystal provides information on the binding site of protein and allows the confirmation of protein-protein interaction. Therefore, it is intended to obtained complexes of MC4 derivatives with NusB from *Thermotago maritima*.

2.3.5.2 Materials

2.3.5.2.1 Chemicals

Thermotoga maritima (*tma*) NusB was provided by the collaborator. MC4 derivatives were chosen based on their MIC, inhibitory activity, and binding affinity.

2.3.5.2.2 Laboratory equipment

Protein crystal diffraction system (Rigaku MicroMax 007HF Imaging Plate X-ray Diffraction System) was employed to perform protein X-ray diffraction.

Protein crystal plates with siliconized cover slips were used in hanging-drop vapor diffusion plates. All protein crystallization plates were put into 16 ℃ incubator for protein crystallization. Conventional widefield microscopy was used to monitor crystal growth.

2.3.5.3 Method

2.3.5.3.1 Procedure

All crystallization experiments were performed using hanging-drop vapor diffusion method with drop volume 4 μL and 500 μL reservoirs. Crystallization conditions for *tma* NusB were initially screened according to literature[97] and Hampton Research Index Screening at 16 ℃.

Crystallization condition Index 17 (1.26 M sodium phosphate monobasic monohydrate and 0.14 M potassium phosphate dibasic, pH 5.6) from Hampton Research Index was chosen as the crystallization condition. *tma* NusB crystals were obtained at optimized condition after refinement of pH and concentrations of precipitant.

2.3.6 Bacterial cellular morphology

2.3.6.1 Introduction

Details about bacterial cellular morphology have been discussed in section 1.5.1. Reduction in quantity of ribosomes is one of the cell morphologies that will be observed. After the application of protein inhibitors, decrease in number of ribosomes in bacterial cells can be resulted. This observable change serves as an indication of direct interaction between antibacterial drugs and ribosomes, and indirect effects of antibacterial drugs against bacterial cell growth or defect on cell envelope that causes efflux of ribosomes[88, 91].

Strains of *B. subtilis* containing GFP fusion to bacterial transcription factors were treated with bacterial transcription targeting inhibitors at different levels of MIC. The fluorescence signals of the treated cells were obtained, where the signals indicated the morphological changes of the corresponding transcription factors.

Observation of change in fluorescence signal of GFP-tagged protein in bacterial cell serves as a proof that MC4 derivatives have mechanism of action consistent with the inhibition of rRNA transcription through targeting NusB-NusE interaction, and allows the evaluation and confirmation of modes of actions of bacterial transcription inhibitors.

2.3.7.2 Materials

2.3.7.2.1 Chemicals

Cells of *B. subtilis* BS61 (*NusB-gfp*) was supplied in streak plates from the collaborator from The Chinese University of Hong Kong and was stored at -80 ℃. LB agar was supplemented with 0.5% chloramphenicol. Luria Broth (LB) was supplemented with 5 μg/L chloramphenicol. Stock solutions of MC4 derivatives (**61**, **123**, **134**) were prepared in 10x MIC in acetonitrile. Microscopic plates were supplemented with 1.2 % (w/v) agarose pads prior to microscopic analysis.

2.3.7.2.2 Laboratory equipment

Fluorescence confocal microscope (Leica TCS SPE Confocal Microscope) with 63x/1.3 oil objectives, 488 nm laser, PMT detector, mercury metal-halide bulb and FITC filter was employed for fluorescence imaging.

2.3.7.3 Method

2.3.7.3.1 Procedure

Stock *B. subtilis* BS61 was transferred onto LB agar and cultured at 37 ℃ for 16 to 20 hours. Single colony was picked and transferred in LB and grown at 37 ℃ until $O.D._{600}$ ~0.6. Stock solutions of MC4 derivatives (**61**, **123**, **134**) were added to cell culture at different MIC. The culture was allowed to incubate for further 15 minutes. 2.5 μL of bacterial cell was added onto the agarose pad, covered with coverslip and sent for microscopic work immediately.

2.3.7.3.2 Data treatment

Confocal fluorescence microscopy was controlled by LAS AF software and a motorized focus drive. The fluorescence images were then touched up with LAS X software for manual cropping and light contrasting.

2.4 Result

2.4.1 Introduction

This chapter presents the data of circular dichroism analysis, fluorescence microscopy, ITC assay, mass spectrometry and X-ray diffraction of *B. subtilis* NusB or *tma* NusB and complexes of NusB with MC4 derivatives. Data shown in this section were treated with data processors such as Excel, or the original software of the instruments used.

2.4.2 ITC assay

Performing isothermal titration calorimetry (ITC) allows the determination of thermodynamic parameters of interactions in solution. Such technique is useful in study in binding of small molecules (ligand/inhibitor) to macromolecules (protein). MC4 derivatives compound **61** was selected to perform ITC assay against protein *B. subtilis* NusB for estimation of the compound's binding affinity and calculate its dissociation constant. Firstly, the buffer system was tested. This was followed by the testing of protein/buffer and buffer/ligand systems. Finally, the ITC of protein/ligand system was performed. The ITC thermograms of experiments are presented below.

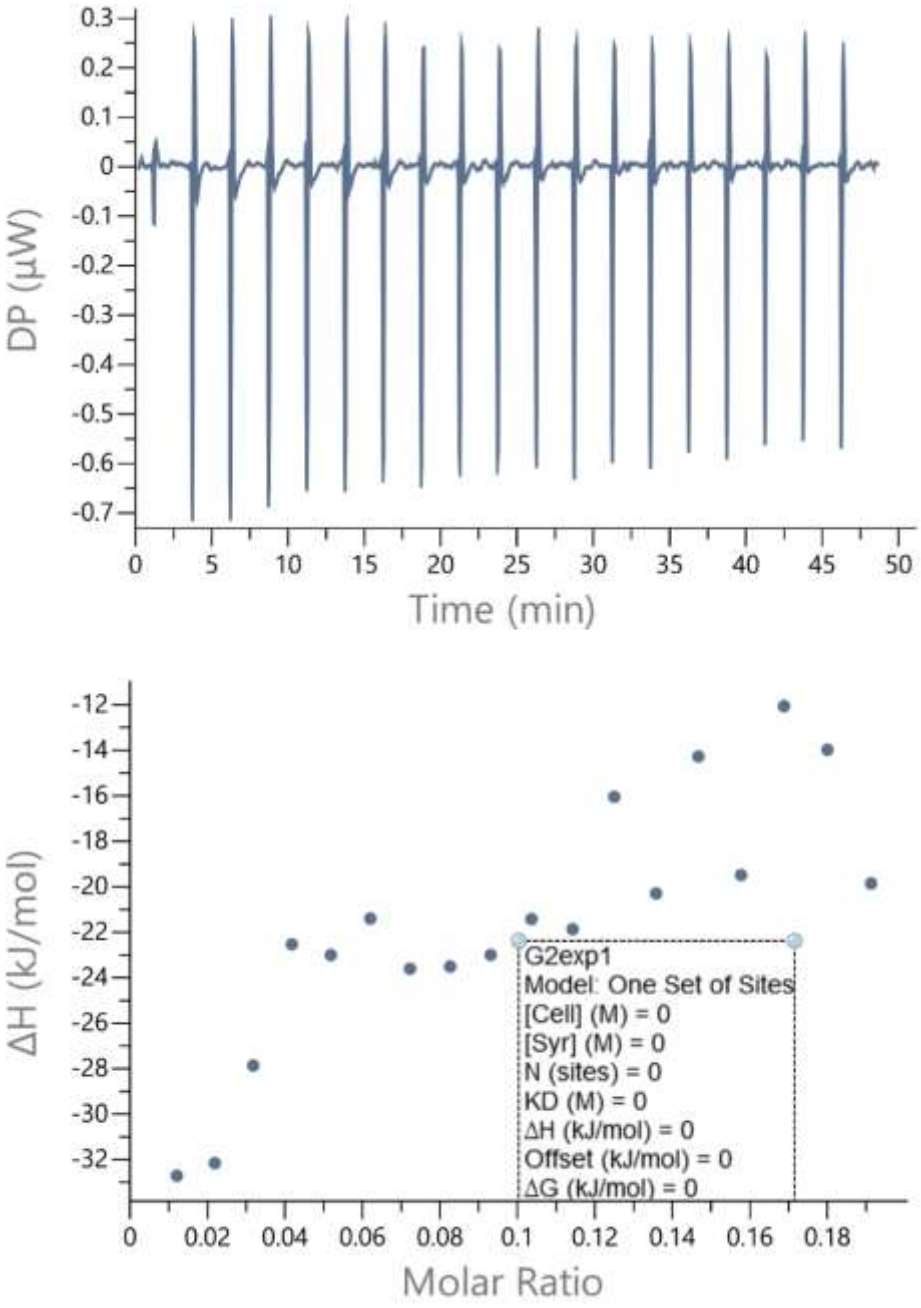


Figure 23 The ITC thermogram of buffer/buffer with 1% DMSO (pH 7.4) (above) and the dependence of released heat in each injection vs. the ratio between total NusB (ligand) concentration and total buffer (syringe) concentration.

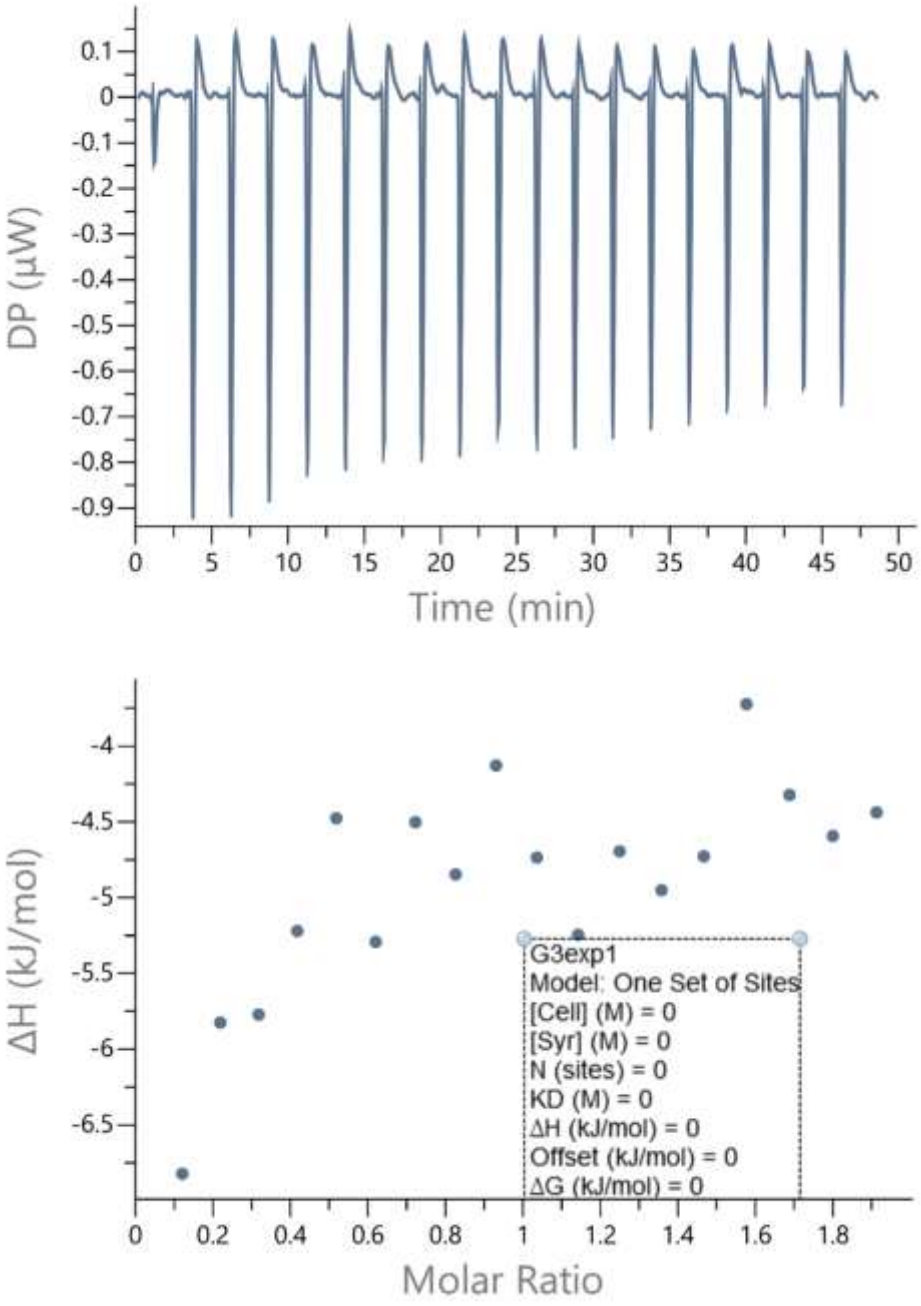


Figure 24 The ITC thermogram of buffer/buffer with 1% DMSO (pH 7.4) (above) and the dependence of released heat in each injection vs. the ratio between total buffer (ligand) concentration and total **61** (syringe) concentration.

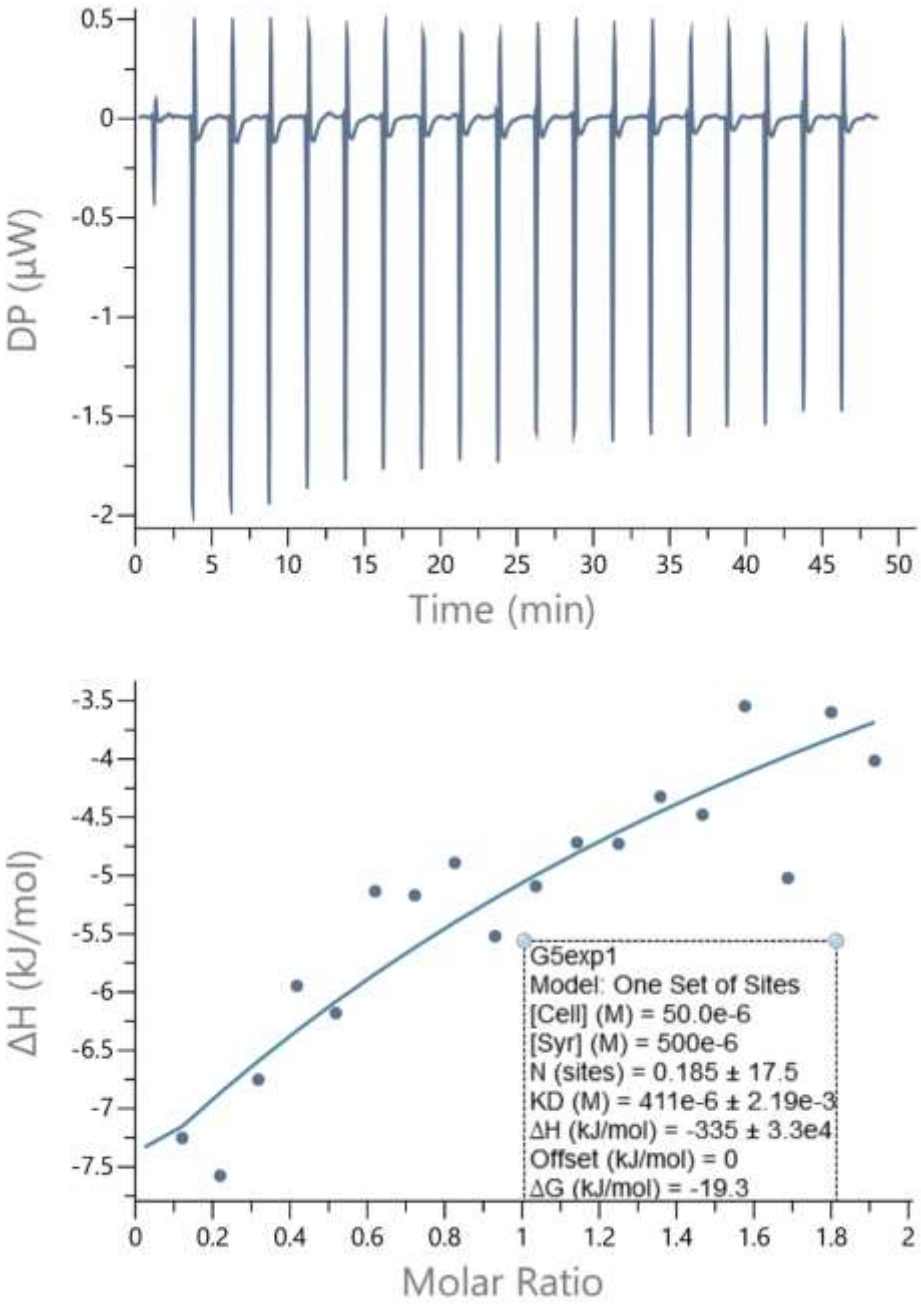


Figure 25 The ITC thermogram of buffer/buffer with 1% DMSO (pH 7.4) (above) and the dependence of released heat in each injection vs. the ratio between total NusB (ligand) concentration and total **61** (syringe) concentration.

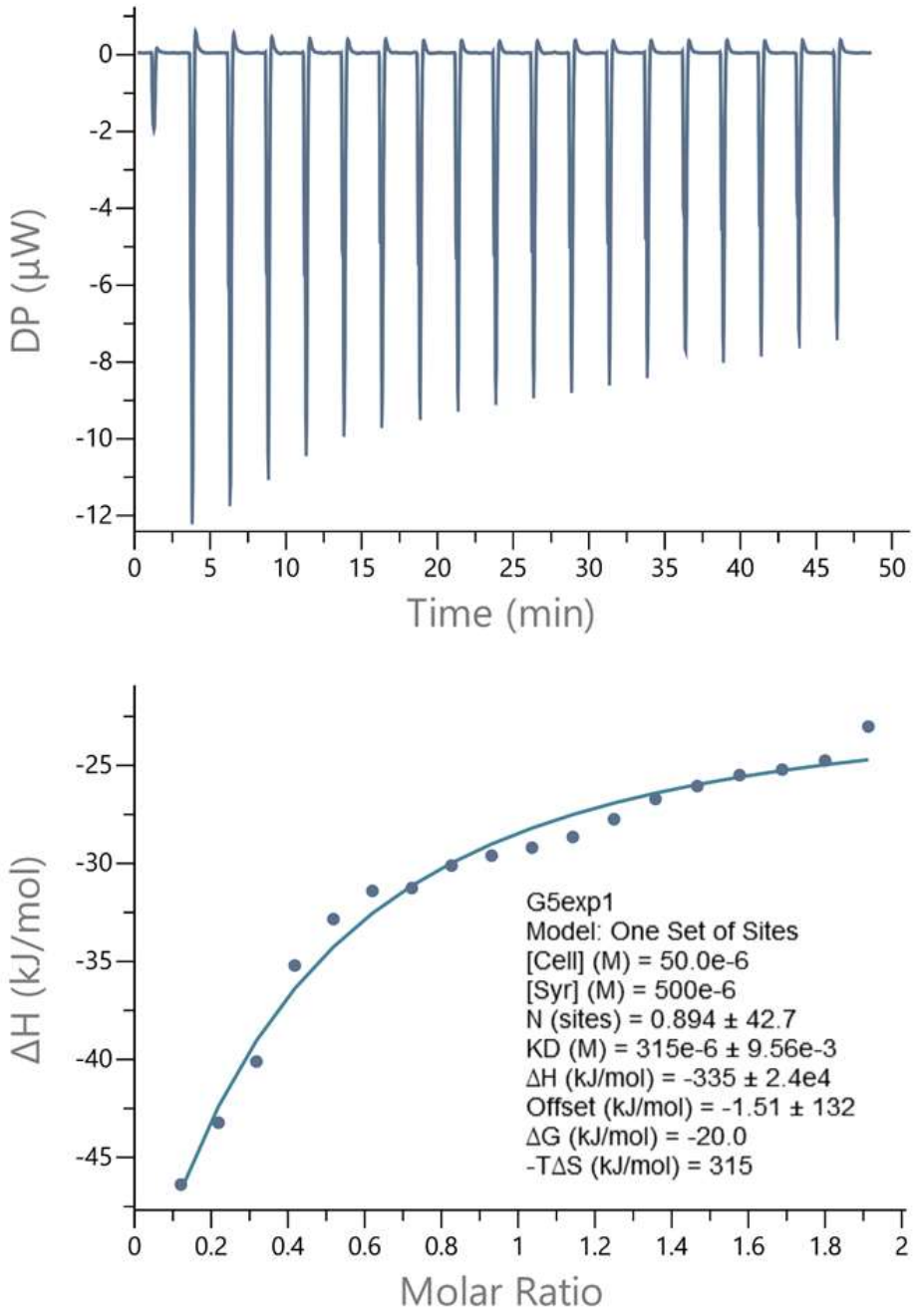


Figure 26 The ITC thermogram of NusB/**61** with 1% DMSO (pH 7.4) and the dependence of released heat in each injection vs. the ration between total NusB (ligand) concentration and total **61** (syringe) concentration

Figures 23 to 25 show the ITC thermograms of NusB against compound **61**. Figure 23 and 24 show no binding between NusB/buffer or buffer/**61**. There was binding between NusB and **61** and the dissociation constant K_d of the inhibitor was 411 μM ± 2.19e-3 (figure 25). Figure 26 shows a thermogram of the second attempt on the ITC of NusB against **61**, which shows that there was binding between NusB and **61**.

2.4.3 Mass spectrometry

It was intended to obtain a NusB-inhibitor complex physically and native mass spectrometry was employed to calculate the dissociation constant of the compounds based on the complexes obtained. Since MC4 derivatives bind to NusB through non-covalent hydrogen bonding, native mass spectrometry does not break these interactions and allows the measurement of the dissociation constant of a biomolecule in its native environment.

However, signal of complex was not captured successfully, and the result presented below only represented the protein. Further investigation will be needed.

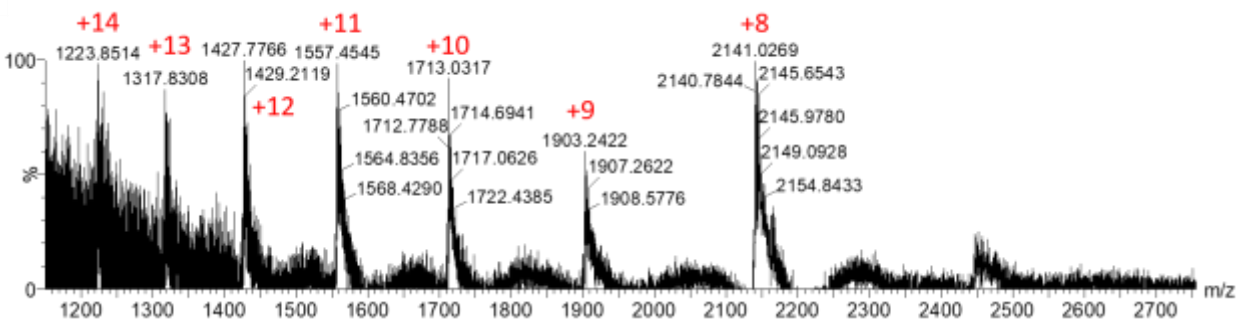


Figure 27 Native ESI MS spectrum of 40 μM *B. subtilis* NusB in 20 mM Ammonium acetate, MW=17120 ± 0.8 Da.

Figure 27 represents a mass spectrum of *B. subtilis* NusB (40 μM, MW = 17120 ± 0.8 Da) obtained by nano-ESI MS. The protein was kept in a native state as presented by the narrow charge state distribution. However, no mass spectrum of protein-ligand complex was successfully obtained. Because the native protein-ligand complex could not be captured by the native ESI MS successfully, an approach of observing an intact protein-ligand complex was adopted, in which a complex was obtained by covalently binding the inhibitors to NusB through photo-crosslinking method.

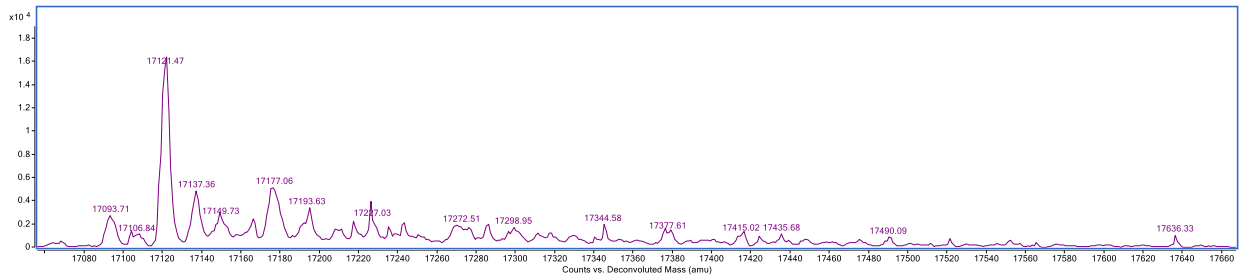


Figure 28a Deconvolution of complex of NusB with inhibitor (-N_3-H).

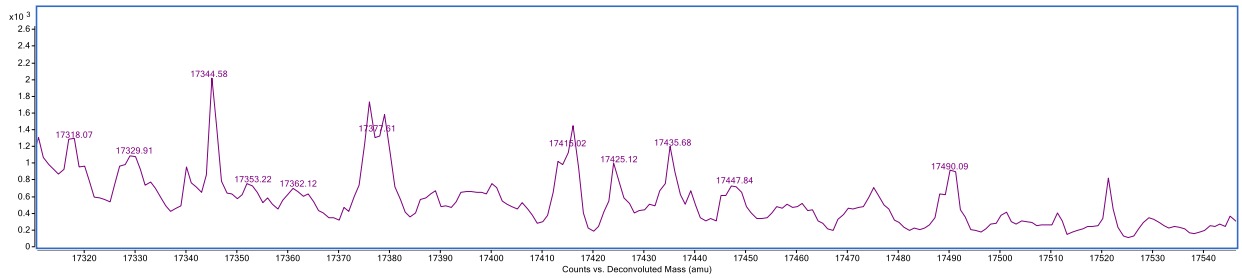


Figure 28b Zoom-in spectrum of figure 28a, deconvolution of complex of NusB with inhibitor (-N_3-H).

Figure 28a is a deconvoluted spectrum of complex of NusB (MW = 17121) with MC4 derivative that carries a photoactive functional group -N_3-H (MW = 256 Da) and figure 28b is a zoom-in of figure 28a showing the presence of peak with MW = 17377 Da. In this experiment, 60 μM *B.*

subtilis NusB was added 180 μM inhibitors (-N_3-H). This complex was supposed to covalently bind to NusB only through photo-crosslinking. However, a complex peak with MW 17377 Da, which represents NusB-inhibitor complex, was obtained after adding to NusB for 20 minutes without treatment of UV irradiation. It was confirmed that the peak was not a background signal but the presence of a complex at comparatively lower quantity than NusB.

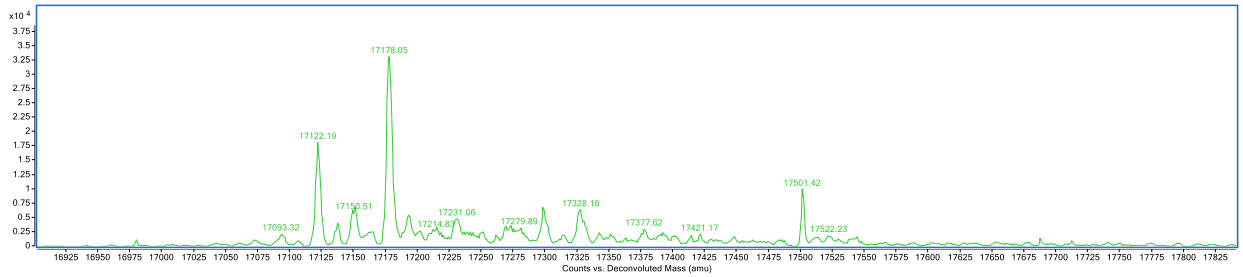


Figure 29 Deconvolution of complex of NusB with inhibitor (-N_3).

A similar result was obtained from treatment of NusB with inhibitor with functional group (-N_3). Figure 29 shows the deconvolution results of the complex. The only difference of this result from that shown in figure 28a-b is that the above spectrum is resulted from treatment of photo-crosslinking for 5 minutes. The major peak (MW = 17178 Da) and a peak corresponding to NusB (MW = 17121 Da) were resulted. The added mass of the major peak did not match any compound and it probably arose from unexpected solvent binding. A peak of MW = 17377 Da was observed again, which corresponds to complex of NusB with inhibitor. It was also confirmed that the peak was not a background signal but the presence of a complex at comparatively lower quantity than NusB.

2.4.4 Protein crystal structure

Process of obtaining protein crystal of *tma*NusB is on-going.

2.4.5 Secondary structure analysis by circular dichroism

The circular dichroism spectrum of NusB and complexes with three MC4 derivatives are shown below. The CD spectrum measures the CD signals in [mdeg] arise from far UV region 190 nm to 250 nm.

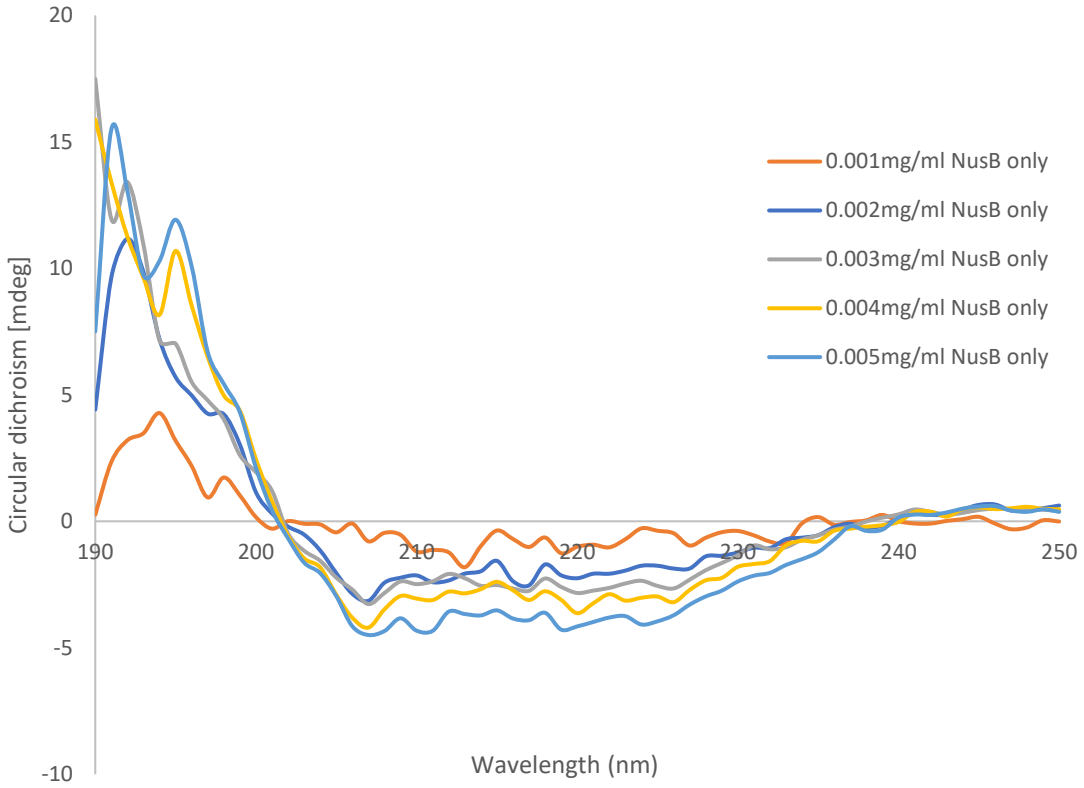


Figure 30 CD spectrum of pure B. subtilis NusB at different concentrations with buffer 5% acetonitrile in PBS (pH 7.4). All CD spectrum of B. subtilis NusB showed a similar pattern. They all have a maximum peak at 194nm, two minimum peaks at 208 nm and 220 nm. This is a typical pattern of an all-α helical protein. The concentration of B. subtilis NusB to be used is selected according to the high tension (HT) voltage of the corresponding CD spectrum. Spectrum of 0.002 mg/ml B. subtilis NusB gave the best HT voltage of all and was used in the CD analysis of complexes.

The CD signals of pure *B. subtilis* NusB of different concentration at different wavelength are shown in figure 30. The CD spectrum *B. subtilis* NusB at different concentrations was obtained by dissolving the protein in 5% acetonitrile in phosphate buffered saline (pH 7.4). Different magnitudes were o*bs*erved in spectrum of different concentration. Despite observing spectrum of different magnitude, HT voltage was the major parameter that determined if the spectrum of the

corresponding concentration of protein could be adopted for further investigation. 0.002 mg/ml of *B. subtilis* NusB in 5% acetonitrile was chosen as the optimum concentration of the sample because the HT voltage of the spectrum given by this concentration did not exceed 650 V but only the signal at 190 nm and the spectrum was well defined, which had better performance than other conditions.

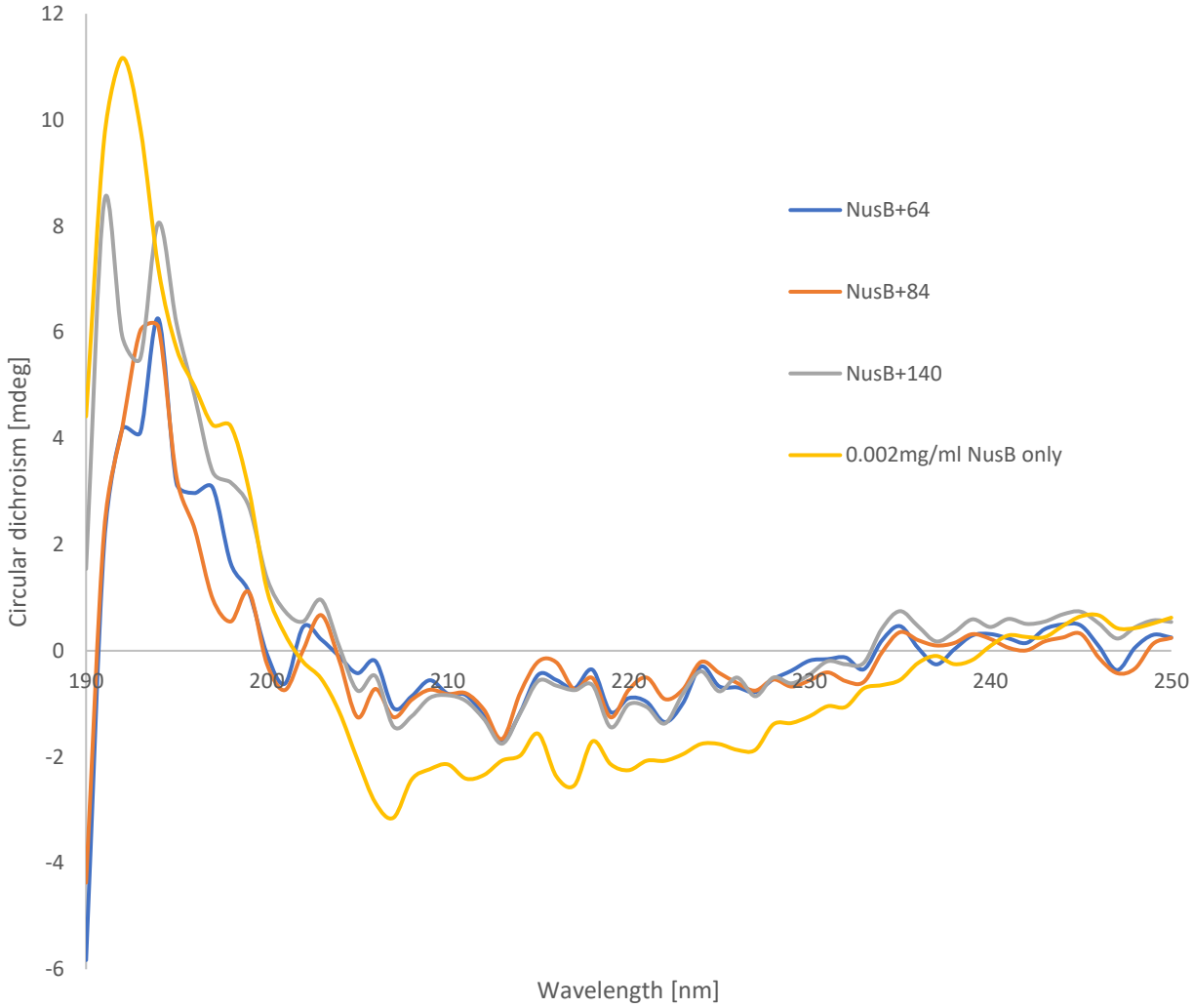


Figure 31. CD spectrum of B. subtilis NusB complexed with MC4 derivatives **64**, **84** and **140**, with CD spectrum of pure B. subtilis NusB. All four spectrum showed a maximum peak at 194 nm and a minimum peak at 208 nm. The peak at 222 nm was not obvious enough to be observed.

The CD spectra of NusB and complexes of *B. subtilis* NusB with MC4 derivatives **64**, **84** and **104** were shown in figure 31. Samples of 0.002 mg/ml *B. subtilis* NusB and complexes were dissolved in 5% acetonitrile in phosphate buffered saline (pH 7.4) while all inhibitors were added to the sample in 1:1 ratio. All these spectra showed a similar pattern by having a maximum peak at around 194 nm and two minimum peaks at around 208 nm and 220 nm. The magnitude of the spectra of complexes reduced slight compared to that of *B. subtilis* NusB alone. There is a significant difference between CD spectra of complexes at around 190 nm but almost no difference between 200 nm to 240 nm in complexes. The only difference observed between 200 nm to 240 nm will be dissected into two parts. Firstly, comparing to spectrum of *B. subtilis* NusB at around 208 nm, compound **64** decreased the most in magnitude, followed by compound **140** and the least was seen in compound **84**. Secondly, comparing to spectrum of *B. subtilis* NusB at around 220 nm, compound **84** decreased the most in magnitude while there was no change observed in the spectra of compounds **64** and **140**. Other wavelength was not compared in this study because the mentioned wavelength range was where the peaks could be observed in CD spectrum of a typical helical protein.

2.4.6 Cell morphology by fluorescence microscopy

Fluorescence microscopic images of green fluorescent protein (GFP) – tagged NusB strain *B. subtilis* BS61 cell were obtained by a fluorescence confocal microscope. MC4 derivatives **61**, **123** and **134** were added at different MIC for observation of the change of cell morphology. The fluorescence images of the cell provide means to confirm the mechanism of action of inhibitors. Fluorescence images of compound **61**, **123** and **134** treated *B. subtilis* NusB cells will be presented sequentially. Cells were treated with different compounds at different MIC.

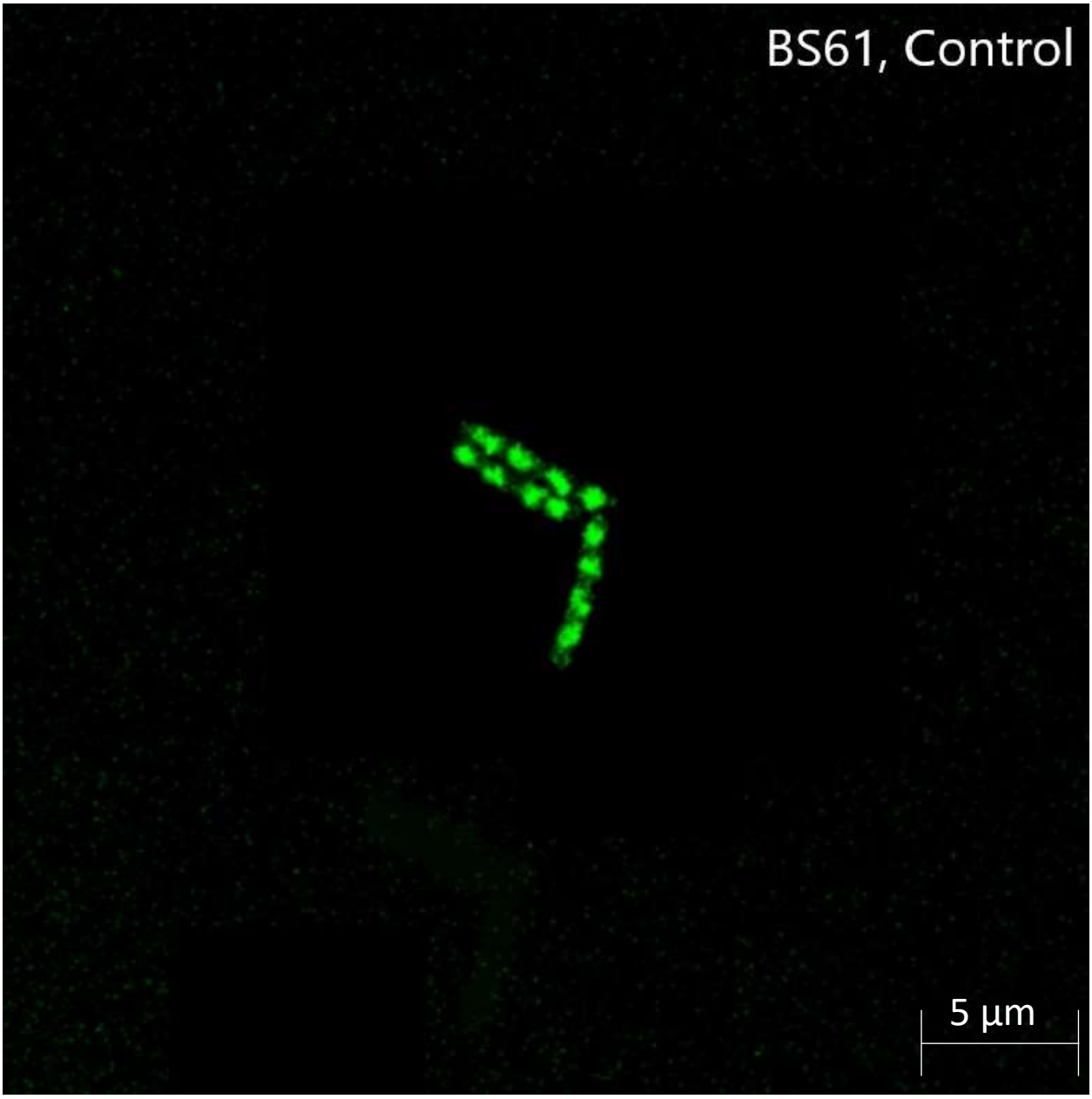


Figure 32 Control fluorescence image of GFP-tagged NusB in B. subtilis BS61. Localization of NusB in nucleoids can be observed in each cell, indicating a health morphology of B. subtilis BS61.

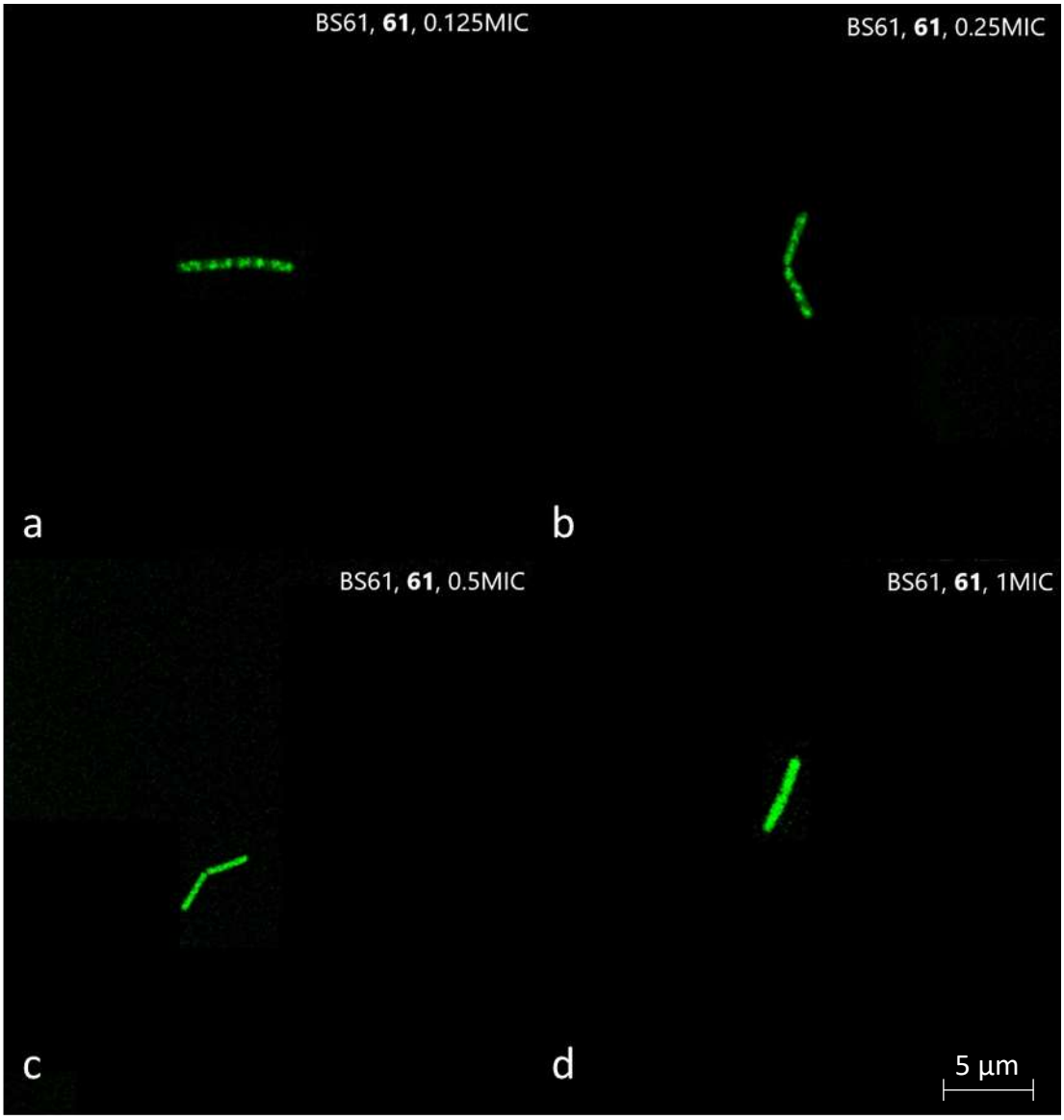


Figure 33 Fluorescence image of GFP-tagged NusB in B. subtilis BS61 with compounds **61** applied at (a) 0.125 MIC, (b) 0.25 MIC, (c) 0.5 MIC, and (d) 1 MIC.

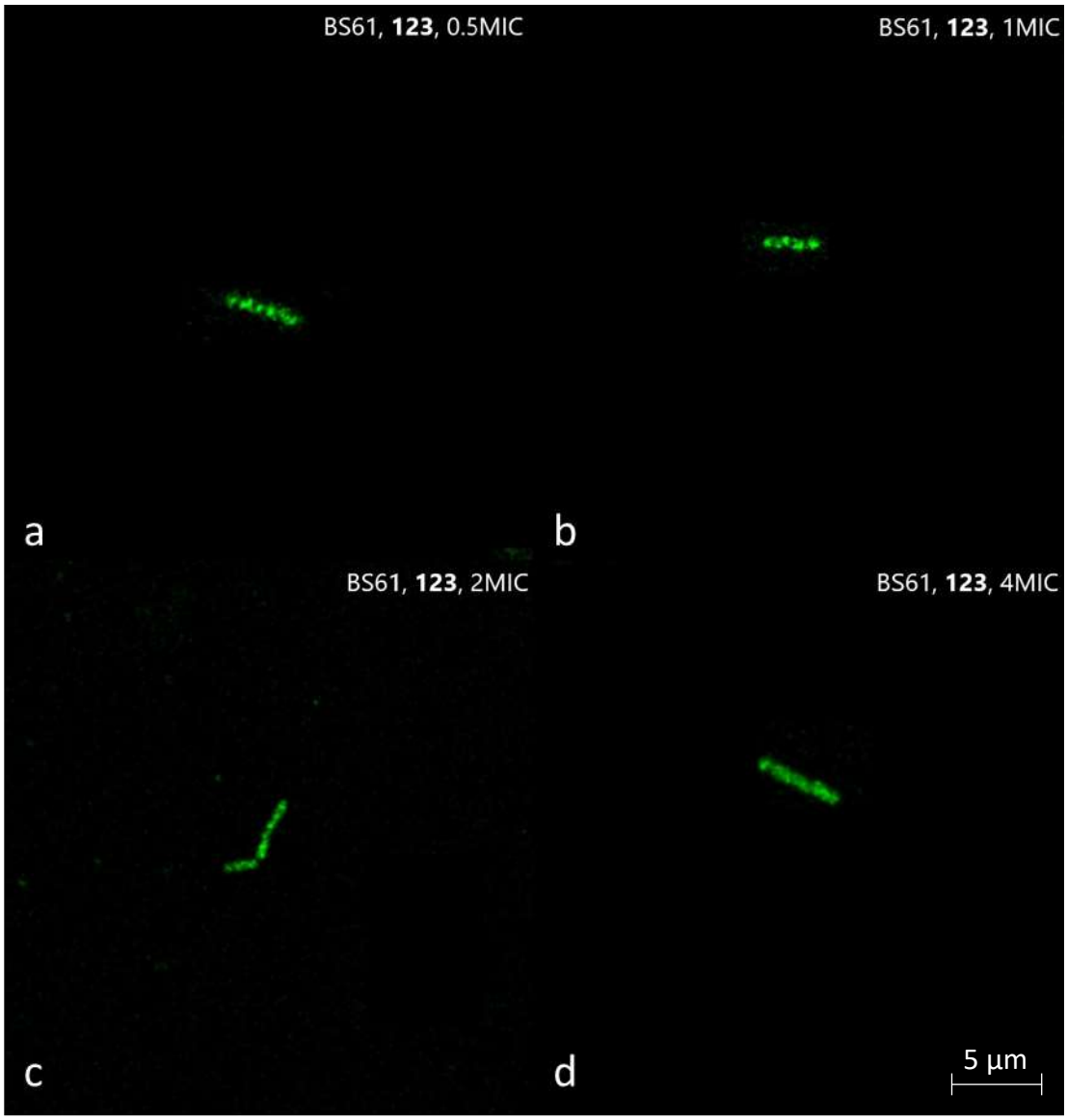


Figure 34 Fluorescence image of GFP-tagged NusB in *B. subtilis* BS61 with compounds **123** applied at (a) 0.5 MIC, (b) 1 MIC, (c) 2 MIC, and (d) 4 MIC.

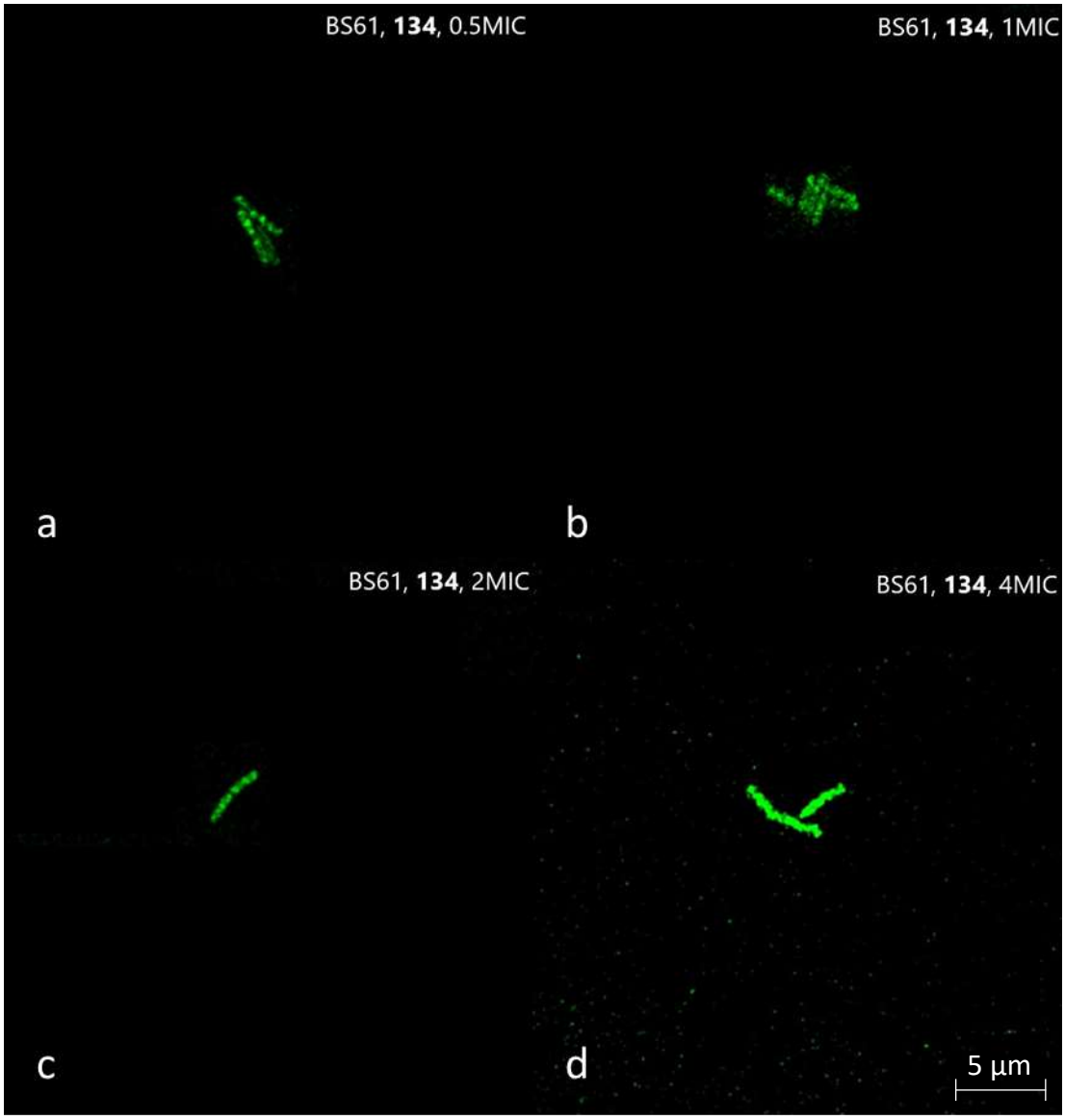


Figure 35 Fluorescence image of GFP-tagged NusB in *B. subtilis* BS61 with compounds **134** applied at (a) 0.5 MIC, (b) 1 MIC, (c) 2 MIC, and (d) 4 MIC.

The control fluorescence image of *B. subtilis* BS61 (figure 32) presents the distribution of GFP signal of NusB without treatment of any inhibitor. It was found that NusB was concentrated within the nucleoid with distinct localization at the two ends towards the center of the cells. The location of GFP signals represents the NusB where the site of rRNA synthesis is located. Figure 33(a)

shows a fluorescence image of *B. subtilis* BS61 treated by MC4 derivative **61** at 0.125 MIC. High quantity of localization of GFP signals in nucleoids and no change in cell morphology can be observed at this level. In figure 33(b), there was only a limited number of cells showing delocalization of GFP signals, indicating that the NusB is slightly disrupted and dissociated around nucleoids at 0.25 MIC. Figure 33(c) shows the image of *B. subtilis* BS61 treated with 0.5 MIC of **61**. Amount of cell showing localization of GFP signal at nucleoids decreased while others showed dissociation of GFP signal. Figure 33(d) shows the image of 1 MIC treated cells. High quantity of delocalization of GFP signals were observed. Although the nucleoids were elongated and diffused, it was obvious that some GFP was still concentrated at the two ends of the cells.

Figure 34(a) shows a fluorescence image of *B. subtilis* BS61 treated with 0.5 MIC of compound **123**. There was no observable change in morphology at this level. It started to show delocalization at 1MIC (figure 34b), partially delocalized at 2 MIC (figure 34c) and became fully dissociated at 4 MIC (figure 34d). *B. subtilis* BS61 morphology was changed in the same fashion for **134** treated cells, where at 0.5 MIC (figure 35a) the cells were not affected by the compound. Small amount of delocalization could be observed at 1 MIC (figure 35b). The GFP signals in cells were partially delocalized at 2 MIC (figure 35c) and became fully dissociated at 4 MIC (figure 35d).

2.5 Finding and discussion

2.5.1 ITC assay

The dissociation constant (K_d) of **61** to NusB was 411 μM ± 2.19e-3. However, the lead compound MC4 has a one-site binding mode to NusB and its K_d is 1.45 ± 0.55 μM[114]. The result obtained here was far from the literature value. One possible reason for this phenomenon was that there was a high amount of heat release from the buffer system. However, this may not be the major reason in this case. Firstly, as *B. subtilis* NusB has an isoelectric point of 6.1, protein at this value will be neutral and stable. The buffer had a pH 7.4 which might not be stable for the protein. Secondly, 1% of DMSO might be the major source of heat because DMSO as an organic solvent might not be decent for the protein with high hydrophilicity. Furthermore, addition of compound in DMSO might further promote the heat. The compound has a high hydrophobicity which might not be compatible to the protein and could further encourage precipitation of NusB. Considering ITC assay measures the heat release in a system, ITC in this research indeed was not an ideal method to evaluate the compounds and a more physical method such as the native mass spectrometry should be adopted.

The heat profile from ITC showed that there was abnormal heat release during the experiment. This is due to the protein precipitation when compounds are added to it. Since the NusB is very insoluble, a lot of it has been precipitated during purification. Addition of compounds further precipitates the protein. This is the major reason for observing a large amount of heat release during the titration. Nevertheless, the heat profile in figure 26 shows that there are binding activities between NusB and the compound. Yet this will need further investigation.

2.5.2 Mass spectrometry

Since the past three decades, electrospray ionization mass spectrometry has become a powerful tool in analysis of biomolecules. Mass spectrometry allows the determination of precise molecular masses of molecules. Dissociation constant (K_d) and stoichiometry are physical parameters that indicate the interaction of a protein-ligand complex. Recent advance in nano ESI even allows the detection of protein-ligand complex without the interruption of its non-covalent interaction by providing a gentle desolvation of complex through increasing vacuum pressure in mass spectrometer[115]. This method also provides information on the relative binding strength and topology of the complex[116].

The approach of obtaining dissociation constant of MC4 derivatives form the protein-ligand complex in this research has not been successful yet and further investigation will be required. The general procedure of detection of dissociation constant of a complex is to first exchange the *B. subtilis* NusB into 20 mM ammonium acetate and obtain a native mass spectrum. Then the same condition of protein was adopted but with addition of inhibitors in 1:10 ratio in 10% organic solvents. It resulted in obtaining the mass spectrum of NusB in success but there was no signal of complex detected at all. Even if there were signals obtained in a complex sample, the mass increase did not equal the protein with inhibitor, meaning the inhibitor was not detected in the 'complex' obtained. A few explanations were proposed to this situation.

Under the native condition, the stoichiometry of a protein-ligand complex can be determined from the mass shift when ligand is added to protein and 1:1 molar ratio of protein-ligand complex is obtained. The mass shift represents the mass of ligand.

Dissociation constant on the other hand can be determined by the dose-response or competition experiment[115], where a fixed protein concentration is added various concentration of ligand or vice versa.

Although many successful data on analysis of protein-ligand complex have been published, the detection of these complexes has been challenging for various reasons. For example, it was suggested that the analysis of these complexes requires aqueous buffered solution to avoid dissociation of complexes prior to analysis. This deviates from the standard operation of ESI by having the need to generate a stable spray from the solvent and at the same time to maintain effective ion desolvation. The desolvation requires a harsh interface condition but labile complexes may be destroyed under this condition. To tackle this, collisional energy will have to be applied but under this situation desolvation is incomplete and as a result ion signal intensity is low. Consequently, the peaks of the intact complexes are broad because of the adduct formation with solvent molecules or salt or buffer ion. Besides, low ion intensity of hydrophilic complexes could be observed compared to proteins because of the effect of sample surface activity in ESI as surface charge will be built up in non-neutral solution. Analytes enriched in the droplet surface were preferred within the offspring droplet cascade while the analytes in the interior of droplet will be lost[117].

This explanation is consistent with the result obtained from nano ESI in this research as the ion intensity of pure *B. subtilis* NusB had been very low at a relatively high concentration. Once the inhibitors were added, almost no ion signal was obtained. This may be because on one hand the protein ion signal can only be retained at high concentration and on the other hand high

concentration of protein-ligand complex is prone to aggregation before any ion signal can be detected. Another possible reason for this phenomenon is that the protein and complex may require different solvent system. Ammonium acetate (pH 6-8) is a commonly used solvent for ESI detection of intact protein and complexes. It has been demonstrated by Hernandez et al.[116] that at a lower ammonium acetate concentration there was no signal of complex detected while at a high concentration, peak assignment was possible but with insufficient resolution of charge states for mass measurement. Using a higher ammonium acetate concentration may be a solution yet this may rise another problem in nano ESI. Since nano ESI is a very sensitive detection method, a small amount of salt can easily be detected, and the salt signal will effortlessly mask the peaks of samples. Ammonium acetate can easily vaporize in ESI, but the high concentration of it should be avoided due to the ease of saturation of mass analyzer. Moreover, the signal responses of protein-ligand complex and intact protein are not always the same as indicated by Ishii et al.[115] because the ion emission efficiency, transmission efficiency and detector efficiency may be different for complex and protein. This is true for the case of native NusB and complex. The signal of NusB was attainable only at an exceptionally high concentration for regular protein detection by nano ESI. While at the same concentration of NusB, the complex with inhibitor at molar ratio > 1:1 with 10% organic solvent, no signal of complex nor protein was detected.

Isoelectric point of the protein may be of another concern. The isoelectric point of protein is the pH at which the protein carries zero net charge. At this pH, protein is prone to precipitation due to the balanced positive and negative charge on protein surface. There is reducing repulsive electrostatic force between protein and the attractive forces dominate. As a result, protein is subject to aggregation and precipitation. One of the disadvantages of precipitation at isoelectric point is

that the denaturation is irreversible[118]. This may explain the situation of the protein-ligand complex. If the pH of the system is close to the isoelectric point of the complex, the complex is subject to precipitation. The isoelectric point of NusB is 6.2 and the pH of buffer solution (ammonium acetate, 20 mM) of the experiment has stably been ~7.0. At this pH, NusB was stable and native mass spectrum could be obtained. However, at the same buffer condition, complex of NusB-MC4 derivatives precipitated. This indicated that the pH of buffer might be close to the isoelectric point of the complex and caused precipitation. Since there was no data on the isoelectric point of protein-ligand complex available, further investigation on this will be needed to confirm the deduction.

Because of the protein being insoluble, addition of compounds further precipitates the protein, and this hindered the capture of complex signal. Another possible explanation to the precipitation of complex is that the complex formation was too quick that it may require quenching within very short period. As the experiment was designed to add excess molar ratio of MC4 derivatives to NusB to allow complex binding, the concentration of MC4 derivatives exceeded the IC_{50} that all the NusB was inhibited.

Because of the high chance of precipitation, photo-crosslinking between protein and inhibitors was taken as another approach to obtain a covalently bound protein-ligand complex which was analyzed with liquid chromatography mass spectrometry. The aim of this was to obtain a more stable and covalently bonded complex to overcome the weak interaction between NusB and MC4 derivatives. MC4 derivatives with either a $-N_3$ group or $-N_3H$ group could be photo-activated to create covalent bonds with NusB. Figure 28a-b and 29 both were resulted from a compound

binding to NusB because of the presence of corresponding peak. However, the unexpected results of obtaining a complex without any UV-irradiation (figure 28b), though at extremely low quantity, could possibly explain the compound had certain level of binding to NusB even with very low binding affinity. The peaks of complex showing after UV-irradiation was as expectation yet in very low abundance. This experiment indicated that MC4 derivatives could bind to NusB. It could be argued that the low abundance of the peak could be background signal and should be ignored. However, the peaks were essentially confirmed to be peaks of complex because of their presence after background signal removal. While the peaks were too short compared with that of NusB, it could conclude that the binding of ligand to NusB was not complete. In addition, because NusB is insoluble, addition of compounds further precipitates the protein. Very low intensity of signal of complex was captured while that of NusB was relatively high. This is because most complexes have precipitated in the mobile phase while not all the compound could successfully bind to NusB before precipitation. Another point to be added is that not all experiment condition (irradiation time) gave the same result, i.e. the presence of a complex. It could be possible that the complex was formed but somehow precipitated as it reached its isoelectric point. The treatment could further be modified for improved binding.

2.5.3 Circular dichroism

The purpose of considering circular dichroism spectroscopy was to examine the influence of derivatives to the NusB folding, and to confirm that the altered binding is due to the change in the NusB interface and not due to protein misfolding. To achieve this, spectra of circular dichroism of *B. subtilis* NusB and complexes were compared. It was hypothesized that complexes of *B. subtilis* NusB with inhibitors were formed and the change of *B. subtilis* NusB folding under the effect of inhibitors could be reflected by CD signals.

The ratio of compounds to protein is 1:1. This is because there is only one binding site of NusE on NusB and the binding activities between NusB and NusE is of one-site binding mode[64], and there is only one NusE binding to one NusB specifically.

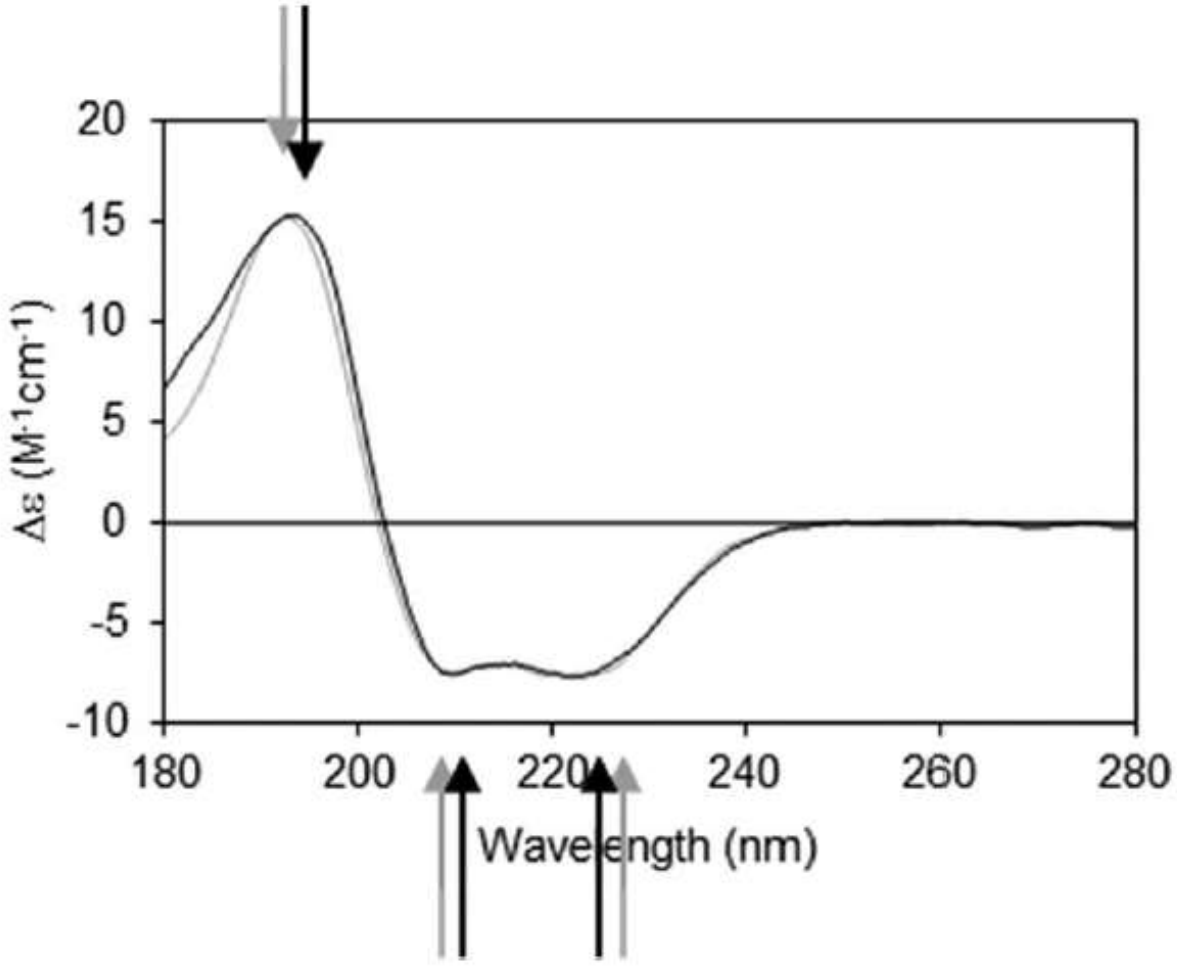


Figure 36 CD spectra of typical predominant helical membrane proteins (blank spectrum) and soluble proteins (grey spectrum). The arrows indicate the position of peaks of the spectra of corresponding proteins

Figure 36 shows typical CD spectra of predominant helical membrane proteins and soluble protein[110]. The CD spectra of helical membrane protein and helical soluble protein are almost the same except there is a slightly difference at 180nm, but this difference is ignorable because wavelength range 190 nm to 250 nm is selected for measurement of CD of secondary structure. Both spectra have a maxima at 190 nm and two minima at 208 nm and 220 nm Standard CD spectrum of an α-helical protein has one positive peak at 194 nm and two negative peaks at 208 nm and 220 nm[112, 119].

In figure 30 (section 2.4.5), the spectra of *B. subtilis* NusB at different concentration differ significantly from 190 nm to 200 nm but were similar in shape from 210 nm onwards. The HT voltage applied to the photomultiplier tube of the detector is a critical parameter to determine whether the CD spectrum is good enough for analysis. Generally, a spectrum with voltage higher than 650V should be ignored. 2 μg/ml *B. subtilis* NusB in 5% acetonitrile in neutral phosphate buffer gave a well-defined spectrum and the HT voltage of the corresponding spectrum was acceptable. This concentration was considered optimum for the sample and was adopted for further analysis. The *B. subtilis* NusB spectrum had one maximum peak at 190nm, and two minimum peaks are 208 nm and 220 nm. This is a typical CD spectrum of all-α helical protein. The result is also consistent to the multiple α-helical structure as observed by X-ray crystallography. This confirmed that the *B. subtilis* NusB consisted of all-α helical structure.

In figure 31 (section 2.4.5), CD spectra of three MC4 derivatives, **64**, **84** and **140**, binding to *B. subtilis* NusB were shown and compared to that of *B. subtilis* NusB. 2 μg/ml *B. subtilis* NusB in the same buffer was added these compounds in 1:1 ratio. **64** and **84** were two representative

compounds of imine and amine structures, respectively. They were found to have outstanding performance in interaction with NusB. Being compounds with representative structures and low MICs, they were chosen to detect their influence on NusB folding because it was suspected that these two different linkers might affect the unique position of the two benzene rings in the hit compound MC4 and their binding affinity to NusB. **140** had the lowest MIC of all. This compound was supposed to have the best binding activity as it showed more inhibition to NusB.

Firstly, the spectra of *B. subtilis* NusB after binding with inhibitors presented in a similar fashion. They all showed a maximum at around 194 nm and minima at around 208 nm and 220 nm, which are a typical α-helical character. Secondly, changes on spectra of complexes were observed and the magnitude of the spectra of complexes reduced compared to that of only *B. subtilis* NusB. This change on spectra was due to the protein-ligand binding, in which the chiral character of NusB decreased due to binding to derivatives. Thirdly, although there is slight difference in magnitude between different complexes, the overall shape of spectra did not change. This indicates that similar structural changes of NusB after binding to derivatives were resulted.

CD signals arise from far UV region (240 nm to 190 nm) because of the chromophores of polypeptide backbone. When the chromophores of the amides of polypeptide backbone of the proteins align in arrays, their optical transitions will shift due to excitation transitions[109], i.e. $n \rightarrow \pi^*$ at 222 nm and $\pi \rightarrow \pi^*$ at 208 nm and 190 nm[110]. This electric excitation gives rise to proteins with different structures having their characteristic CD spectra. For example, the spectrum of an α-helical protein has negative peak at 222 nm and 208 nm and positive peak at 190 nm. Well-defined antiparallel β-pleated sheet proteins have negative peak at 218 nm and positive peak at

195 nm. Disordered proteins have negative peak at 195 nm but low ellipticity at above 210 nm [109, 111].

$\pi \rightarrow \pi^*$ transition can be referred to the carbonyl bond (C=O) of peptide. As the compounds bind to NusB through hydrogen-bonding, such interaction could possibly be able to pull the bonds at certain extend that is enough to cause the change in chirality of the chromophores. The wavelength also corresponds to the increase in hydrophobic interaction of binding site after the docking of compounds into NusB and a more folded conformation was resulted in the area, leading to decrease in chirality of chromophores. This explains the major change in magnitude of peak at 194 nm.

The slight difference in magnitude from 208 nm to 220 nm was due to the conformational flexibility of protein, allowing ligands binding to the protein without significant change in conformation. This region refers to the peptide backbone. When ligands bind to the protein causing no change at 208 nm to 220 nm, this indicates the peptide backbone was not altered. However, there is a significant increase in magnitude at 190 nm upon ligand binding. Firstly, this could be due to the decrease in amount of α-helix at this region. Peak at 190 nm refers to the peptides around the active site of NusB. Due to the significant change of chiral properties of peptide backbone around the active site when ligands interact with peptide through hydrogen-bonding. The CD spectra were also in agreement with a significant increase at 190 nm but slight gain at 208 nm and 220 nm because of dissociation of protein since the ionic strength of the protein solution changed in ligand binding. Ionic strength increased at 190 nm resulted in increase in magnitude[109, 120, 121].

Additionally, CD spectrum of proteins with higher molecular weight retain the same features but slightly shifted wavelength (194 nm to 192 nm). On other hand, a distorted α-helical protein shows similar features but significant red shift of positive peak (194 nm to 186 - 193 nm) or negative peak (208 nm to 199 – 209 nm) and blue shift (222nm to 227 – 228 nm). Major reduction of magnitude can also be observed[119]. In the spectra of complexes, no significant change in magnitudes of all peaks was observed. Peaks also retained the same position. The slight change from 205 nm to 225 nm was due to slight change of interface[70]. The relatively small difference between different CD indicates that the inhibitors had no effect on change of global NusB structure[122].

To conclude, *B. subtilis* NusB was stable because the NusB retained an α-helical secondary structure after protein-ligand binding. The slight change in CD spectra of complexes did not imply the change of shape of *B. subtilis* NusB. Such change resulted from the decreased in chiral character of NusB after binding and was a result of protein-ligand binding. Change of binding interface also caused the change in spectra. Yet, the secondary structure of NusB was not altered and the protein did not misfold. The three derivatives of MC4 were able to cause the same conformational change on NusB. This suggests that these compounds may bind to the same binding site on NusB.

2.5.4 Fluorescence images

Fluorescence microscopic images of *B. subtilis* NusB cells was captured with a confocal fluorescence microscopy. Fusion of GFP-tagged NusB into strain *B. subtilis* BS61 was employed for observation of localization of NusB in cell. Cells of *B. subtilis* BS61 were then treated with MC4 derivatives **61**, **123** and **134** at different MICs for observation of cell morphology. The purpose of capturing the fluorescence images of bacterial cells was to confirm the mechanism of action of the compounds.

The MICs of compounds **61**, **123** and **134** are 2 μg/ml, 1 μg/ml and 0.0625 μg/ml respectively. **134** is by far the inhibitor with the best MIC. The levels of MICs applied were determined by observation on morphology of *B. subtilis* NusB cells i.e. starting with treatment with 1 MIC, if the morphology of treated cell was not observed, the level of MIC was increased. Conversely, if the treated cells changed their morphology, the level of MIC was reduced.

Firstly, *B. subtilis* BS61 was treated with **61** at 0.125, 0.25, 0.5 and 1 MIC. 0.125 MIC of **61** did not caused any change in the cells as localization of NusB in nucleoids can be observed clearly. Slight change in morphology started to be observed at 0. 25 MIC, where a limited number of delocalization can be observed but slight to almost no change in morphology was observed in the majority of cells. Delocalization of fluorescence signals started to be observable at 0.5 MIC treated cells and the fluorescence signals were fully dissociated at 1 MIC. Secondly, *B. subtilis* BS61 was treated with **123** at 0.5, 1, 2 and 4 MIC. Treatment with 0.5 MIC did not show any change in morphology but it started show delocalization at 1 MIC and were fully dissociated at 4 MIC. *B. subtilis* BS61 was finally treated with **134** at 0.5, 1, 2 and 4 MIC. Treatment with 0.5 MIC did not show any effect to cell morphology. Localization of nucleoids could be seen clearly at this MIC.

Cell morphology started to change slightly at 1 MIC but there was only a limited number of cells showing delocalization of nucleoids. Treatment of 2 MIC increased the number of nucleoid elongation while 4 MIC could dissociate most of the GFP signals.

In the control experiment, *B. subtilis* BS61 showed a localization of GFP at the nucleoid where represents the site of rRNA synthesis. The nucleoids were confined to a spherical shape with defined edge. This is a morphology of *Bacillus subtilis* at mid-exponential growth. Yet, *B. subtilis* BS61 delocalization was observed at different levels of MIC under the treatment with different compounds. The delocalization came along with elongation of nucleoids, in which two slightly enlarged nucleoids connected in the center and elongated into a dumb bell shape. This morphology was a typical late stage of exponential growth in *Bacillus subtilis*. It was suggested that the nucleoids would elongate with greater frequency when the cell approached the end of exponential growth. When the cell reached the stationary phase, the nucleoid changed physically upon inhibition of protein synthesis. One of the physical change observables is the elongation. This may raise from the placement of ATP interaction with the cell envelope or the decrease of supercoiling between exponential phase and early stationary phase. The elongation of nucleoids was induced by the addition of antitermination targeting inhibitors. During sporulation, two nucleoids elongate due to the reduction in protein synthesis and fused to form axial filament. Because of the change in morphology upon treatment with inhibitors, it may deduce that MC4 derivatives can promote the end of exponential growth upon inhibition of protein synthesis.

Treatment of cells with **134** only required 0.25 μg/mL which is equivalent to 4 MIC to fully delocalize GFP-NusB. This is probably because of the excellent binding activity of this compound

compared to others and that this compound has good interaction with NusB against NusB-NusE interaction.

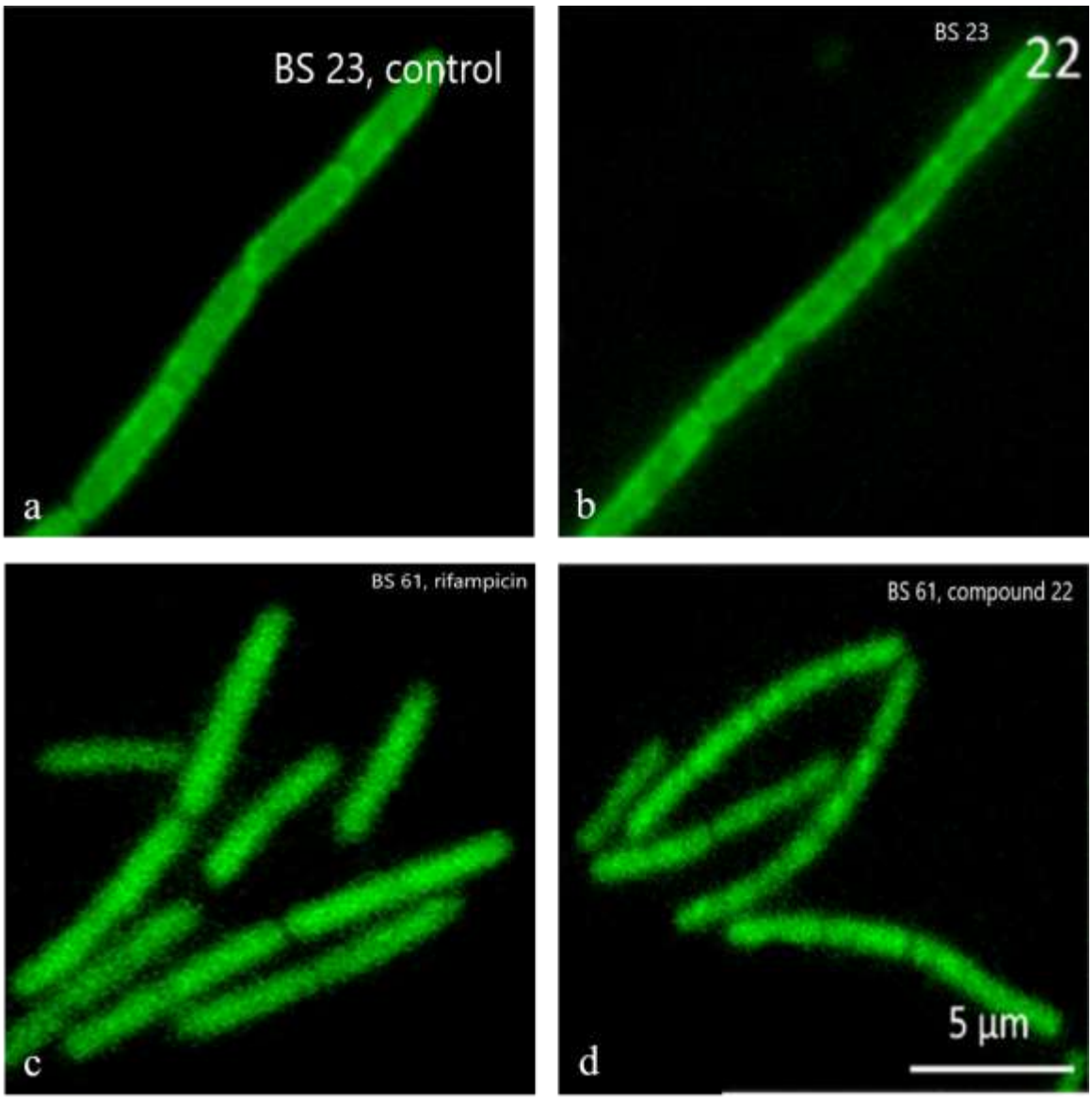


Figure 37 Epifluorescence image showing the morphology of (a) untreated B. subtilis BS23 (ATP-gfp), (b) B. subtilis BS23 treated with a NusB-targeting inhibitor 22, (c) B. subtilis BS61 treated with rifampicin, a transcription targeting inhibitor that interact with bacterial RNAP, and (d) *B. subtilis* BS61 treated with **22** that targets NusB-NusE interaction. Delocalization of nucleoids was observed in (c) and (d) while the same morphology as untreated cell was obtained in (b), indicating that the compound did not interact with the ATPase[37].

To demonstrate the mode of action of MC4 derivatives being consistent with the inhibition of interaction of NusB and NusE, which finally results in significant reduction in bacterial rRNA synthesis, the morphology of inhibitor-treated *B. subtilis* BS61 was compared with that of treated *B. subtilis* BS61 with another NusB-NusE targeting inhibitor with confirmed mode of action. *B. subtilis* BS23 contains a GFP fusion to α subunit of the ATP synthase. Because α subunit does not take part in antitermination, inhibitors with interaction with NusB should not affect this protein. In *B. subtilis* BS23, ATP synthase as GFP signal is localized to the cell membrane. The treatment of BS23 with colistin, a membrane targeting inhibitor, can cause observable membrane damage. However, when the cells were treated with a transcription targeting inhibitor, no change in GFP localization was observed because it does not target cell membrane. Since colistin targets cell membrane, delocalization of GFP signal in BS23 was resulted. On the other hand, when BS61 was treated with MC4 derivatives or transcription targeting inhibitor, GFP delocalization could be observed as the same morphology when treated with rifampicin, a transcription targeting inhibitor, because the compounds and rifampicin both target RNAP. But when the cells were treated with colistin, no delocalization of GFP signal was resulted because colistin does not target NusB. This proves that MC4 derivatives are active against transcription and can inhibit bacterial transcription by targeting NusB and have the consistent mode of action as the transcription targeting inhibitor[37]

Figure 37a shows an image of morphology of untreated cells of *B. subtilis* BS23. It was reported that compound **22**, a potent lead compound targeting NusB-NusE PPI and a non-MC4 derivative, did not show any effect on morphology of *B. subtilis* BS23 as shown in figure 37b, in which there was no delocalization of ATP synthase indicative of damage of membrane. On the other hand, if a membrane targeting inhibitor (e.g. colistin) is employed to treatment of *B. subtilis* BS23, ATP

synthase localization occurs because cell membrane is targeted and damaged[37]. Yet *B. subtilis* BS61 treated with rifampicin (figure 37c) and **22** (figure 37d) showed delocalization and resulted in the same morphology of MC4 derivative treated *B. subtilis* BS61 cells as shown in Results section. Treatment of the selected compounds **61**, **123** and **134** resulted in delocalization of nucleoids which represents the localization of rRNA synthesis. The delocalization of NusB-GFP signal was coherent with the loss of rRNA transcriptional activity, which was comparable to the induction of stringent response in which major reduction of rRNA synthesis was resulted from a bacterial starvation response. Moreover, significant delocalization of NusB-GFP signals were similar to that shown in figure 37c and d, indicating that there was interaction between MC4 derivatives and NusB, and that MC4 derivatives target on NusB for inhibition of rRNA synthesis. This confirmed the mode of action of MC4 derivatives.

To conclude, MC4 derivatives **61**, **123** and **134** can target NusB-NusE interaction and inhibit rRNA synthesis in bacteria. Their mechanism of action is the same as proposed that MC4 derivatives are transcription targeting inhibitor, targeting NusB-NusE interaction for inhibition of rRNA synthesis.

2.6 Limitation

The limitations of the biochemical evaluation of MC4 derivatives are listed in this section. Firstly, the evaluation methods adopted are not complete and more evaluations are needed. For instance, the use of circular dichroism is to predict the secondary structures of complexes. No information on binding site of protein is given. It only predicts that the compounds have similar binding activity and that the secondary structure of NusB is maintained after binding with compounds, and only it shows the conformation change of NusB due to surface binding. CD provides only qualitative measurement on whether there are binding activities between the protein and compounds, and this cannot serve as a mean to proof that MC4 derivatives specifically bind to NusB.

Secondly, more crystallography data on compounds of different linkers will be needed as different linkers or flexibility affect the binding affinity of compounds to NusB. The possibility of different types of compounds binding to different sites should not be ignored.

Thirdly, the dissociation constant of the compounds could not be obtained directly through native mass spectrometry due to the high probability of precipitation of NusB with compounds of high hydrophobicity. Since NusB binds to NusE to form a stable complex, it was suspected that NusB could only be stabilized in the presence of NusE. Therefore, the interaction between NusB and NusE may be favorable over NusB with inhibitors. In the presence of inhibitors, NusB may be prone to precipitation into a stable form. Since the surface charge of protein is stabilized at its isoelectric point, precipitation of protein or complex can be observed at this pH level. From the ITC assay, it was obvious that there was confirmation change of NusB after binding to inhibitors. Since the interaction occurs in the protein surface, the change in surface, though not significant,

may cause change in surface charge of protein in certain extend. This change in surface charge may be notable enough to change the isoelectric point of complex of protein-ligand complex and cause precipitation. If a complex is precipitated, it can be deduced that it possibly reaches its isoelectric point. However, there is no available data of the isoelectric point of complex. It is also not feasible to predict or measure it at this phase. The situation will need to be further explored. In addition, while NusB is hydrophilic, application of hydrophobic compounds may result in buffer incompatibility and increase the chance of precipitation.

Finally, because NusB is prone to precipitation for stability, the thermogram obtained from ITC assay may contain artifacts. As a tiny amount of heat release was even detected from the buffer/buffer system, it could be expected that the control experiment may introduce a certain amount of heat to the NusB/inhibitor system and an unexpected amount of heat other than the actual heat from binding activity may be observed. Because of this, it is possible that the dissociation constant calculated from the thermogram of ITC is higher than the actual value. It was determined to obtain the dissociation constant of the compounds through a physical technique by obtaining an actual complex as ITC may introduce artifacts while native mass spectrometry was a reliable way to examine a protein/inhibitor complex in its native form. A stable protein/inhibitor complex is yet to be obtained. A solvent completable system for compounds is needed to be developed while at the same time the hydrophilicity of NusB should also be considered when developing the compounds. For instance, the isoelectric point of *Bacillus subtilis* NusB is 6.2 while the pK_a of most MC4 derivatives are above 7. The slightly acidic nature of NusB may be the key that affects the stability of the complex.

Furthermore to the possible artifacts in ITC to detect a complex, in a recent publication[123], the IC_{50} of MC4 derivatives were calculated from the luminescence signal emitted by NusB-NusE complex against MC4 derivatives based on protein complementation technique. Herein an argument is raised on the availability of protein-ligand complex.

2.7 Suggestion for future work

A few recommendations on extending the biochemical evaluation on transcription targeting inhibitors are suggested below. The transcription targeting inhibitors are designed to inhibit the interaction between NusB-NusE antitermination so when the inhibitors are applied to the NusB protein, bacterial transcription should be prevented and decreased. As reported in previous publication, the rRNA synthesis in bacteria was decreased. Therefore, it is possible to further extend the study by studying the transcription level of the complex through different biological assay such as Western blotting assay and real-time PCR technology. In a real-time PCR analysis, small molecules binding to a protein target is measured by the melting temperature shift of the protein.

Hydrogen-deuterium exchange liquid chromatography mass spectrometry (HDX-LC-MS) has been a powerful analyzing tool to study the protein binding sites. Global binding and local binding. This method has also been applied widely on pharmaco-kinetic study on the mechanism of protein inhibition and proteomics. Through extremely short time of inhibition and quenching experiment, the mechanisms of inhibitors and the binding location on peptides can be measured. This also complements the X-ray crystallography.

To further prove the compounds having specific binding to the protein targets, alternatives on binding assays are listed below. Firstly, there is a specific binding site on NusB with three conserved amino acid residues for its interaction with NusE. One can determine whether there is specific binding activity between NusB and MC4 derivatives by replacing the three amino acid residues on NusB. If there were no binding activities, it could conclude that the three amino acid residues are critical for the specific binding. Secondly, MC4 derivatives are competitive inhibitors which compete with NusE for the same binding site on NusB. As MC4 derivatives are rationally designed to have improved binding activities with NusB, competitive binding assay of the MC4 derivatives with labelled (e.g. fluorescence label) NusB should provide more information on the binding affinities of MC4 derivatives.

2.8 Conclusion

To conclude, the study of bacterial cellular morphology of *B. subtilis* BS61 (NusB-GFP) suggested that MC4 derivatives are active against NusB and have consistent mode of actions against formation of transcription antitermination complex as transcription targeting inhibitors. Circular dichroism of *BS* NusB and complexes showed that MC4 derivatives do not cause change of global conformation of the protein. The retainment of secondary structure of NusB suggested that MC4 derivatives have similar binding activities with NusB and they do not cause conformational change of NusB.

Reference

[1] E. H. Sultan, D. W. Jenkins, and W. C. Cutting, " Sulfonamide compounds and penicillin: the effect of combined therapy on experimental infections in mice," *JAMA Internal Medicine,* vol. 76, no. 3, pp. 161-162, 1945.

[2] J. E. Gomez *et al.*, "Ribosomal mutations promote the evolution of antibiotic resistance in a multidrug environment," *eLife,* vol. 6, p. e20420, 2017/02/21 2017.

[3] E. Tacconelli *et al.*, "Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis," *The Lancet Infectious Diseases,* vol. 18, no. 3, pp. 318-327, 2018.

[4] E. C. Cole *et al.*, "Investigation of antibiotic and antibacterial agent cross-resistance in target bacteria from homes of antibacterial product users and nonusers," *J Appl Microbiol,* vol. 95, no. 4, pp. 664-76, 2003.

[5] R. Gaynes, "The Discovery of Penicillin—New Insights After More Than 75 Years of Clinical Use," *Emerging Infectious Diseases,* vol. 23, no. 5, pp. 849-853, 2017.

[6] N. Woodford and M. J. Ellington, "The emergence of antibiotic resistance by mutation," *Clinical Microbiology and Infection,* vol. 13, no. 1, pp. 5-18, 2007/01/01/ 2007.

[7] M. Frieri, K. Kumar, and A. Boutin, "Antibiotic resistance," *Journal of Infection and Public Health,* vol. 10, no. 4, pp. 369-378, 2017/07/01/ 2017.

[8] M. F. Chellat, L. Raguž, and R. Riedl, "Targeting Antibiotic Resistance," *Angewandte Chemie International Edition,* vol. 55, no. 23, pp. 6600-6626, 2016/06/01 2016.

[9] N. A. Lerminiaux and A. D. S. Cameron, "Horizontal transfer of antibiotic resistance genes in clinical environments," *Can J Microbiol,* vol. 65, no. 1, pp. 34-44, Jan 2019.

[10] J. C. Nguyen Van and L. Gutmann, "Resistance to antibiotics caused by decrease of the permeability in gram-negative bacteria," *Presse Med,* vol. 23, no. 11, pp. 522, 527-31, Mar 19 1994. Resistance aux antibiotiques par diminution de la permeabilite chez les bacteries a gram negatif.

[11] M. A. Webber and L. J. V. Piddock, "The importance of efflux pumps in bacterial antibiotic resistance," *Journal of Antimicrobial Chemotherapy,* vol. 51, no. 1, pp. 9-11, 2003.

[12] S. M. Soto, "Role of efflux pumps in the antibiotic resistance of bacteria embedded in a biofilm," *Virulence,* vol. 4, no. 3, pp. 223-229, 2013.

[13] G. De Pascale and G. D. Wright, "Antibiotic resistance by enzyme inactivation: from mechanisms to solutions," *Chembiochem,* vol. 11, no. 10, pp. 1325-34, Jul 5 2010.

[14] A. M. Egorov, M. M. Ulyashova, and M. Y. Rubtsova, "Bacterial Enzymes and Antibiotic Resistance," *Acta naturae,* vol. 10, no. 4, pp. 33-48, Oct-Dec 2018.

[15] J. M. Munita and C. A. Arias, "Mechanisms of Antibiotic Resistance," *Microbiology spectrum,* vol. 4, no. 2, pp. 10.1128/microbiolspec.VMBF-0016-2015, 2016.

[16] K. Bush, "Antimicrobial agents targeting bacterial cell walls and cell membranes," *Rev Sci Tech,* vol. 31, no. 1, pp. 43-56, Apr 2012.

[17] W. Hong, J. Zeng, and J. Xie, "Antibiotic drugs targeting bacterial RNAs," *Acta pharmaceutica Sinica. B,* vol. 4, no. 4, pp. 258-265, 2014.

[18] H. S. Misra, G. K. Maurya, R. Chaudhary, and C. S. Misra, "Interdependence of bacterial cell division and genome segregation and its potential in drug development," *Microbiological Research,* vol. 208, pp. 12-24, 2018/03/01/ 2018.

[19] C. Ma, X. Yang, and P. J. Lewis, "Bacterial Transcription as a Target for Antibacterial Drug Development," *Microbiology and Molecular Biology Reviews,* vol. 80, no. 1, pp. 139-160, 2016.
[20] P. Villain-Guillot, L. Bastide, M. Gualtieri, and J.-P. Leonetti, "Progress in targeting bacterial transcription," *Drug Discovery Today,* vol. 12, no. 5, pp. 200-208, 2007/03/01/ 2007.
[21] K. B. Arnvig and F. Werner, "A new spanner in the works of bacterial transcription," *eLife,* vol. 3, p. e02840, 2014/04/22 2014.
[22] R. R. Burgess and L. Anthony, "How sigma docks to RNA polymerase and what sigma does," *Current Opinion in Microbiology,* vol. 4, no. 2, pp. 126-131, 2001/04/01/ 2001.
[23] D. G. Vassylyev *et al.*, "Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution," *Nature,* Article vol. 417, p. 712, 05/08/online 2002.
[24] A. Feklístov, B. D. Sharon, S. A. Darst, and C. A. Gross, "Bacterial Sigma Factors: A Historical, Structural, and Genomic Perspective," *Annual Review of Microbiology,* vol. 68, no. 1, pp. 357-376, 2014/09/08 2014.
[25] B. M. Burmann and P. Rösch, "The role of E. coli Nus-factors in transcription regulation and transcription:translation coupling: From structure to mechanism," *Transcription,* vol. 2, no. 3, pp. 130-134, May-Jun 2011.
[26] R. Mariani and S. I. Maffioli, "Bacterial RNA Polymerase Inhibitors: An Organized Overview of their Structure, Derivatives, Biological Activity and Current Clinical Development Status," *Current Medicinal Chemistry,* vol. 16, no. 4, pp. 430-454, 2009.
[27] K. Turecka and K. Waleron, "Inhibitors of bacterial transcription are compounds for potent antimicrobial drugs," *Curr Pharm Biotechnol,* vol. 14, no. 15, pp. 1275-86, 2013.
[28] B. E. Nickels and A. Hochschild, "Regulation of RNA polymerase through the secondary channel," *Cell,* vol. 118, no. 3, pp. 281-4, Aug 6 2004.
[29] J. Symersky, A. Perederina, M. N. Vassylyeva, V. Svetlov, I. Artsimovitch, and D. G. Vassylyev, "Regulation through the RNA polymerase secondary channel. Structural and functional variability of the coiled-coil transcription factors," *J Biol Chem,* vol. 281, no. 3, pp. 1309-12, Jan 20 2006.
[30] D. Pupov, I. Kuzin, I. Bass, and A. Kulbachinskiy, "Distinct functions of the RNA polymerase σ subunit region 3.2 in RNA priming and promoter escape," *Nucleic acids research,* vol. 42, no. 7, pp. 4494-4504, 2014.
[31] V. Sosunov *et al.*, "The involvement of the aspartate triad of the active center in all catalytic activities of multisubunit RNA polymerase," *Nucleic acids research,* vol. 33, no. 13, pp. 4202-4211, 2005.
[32] V. Mekler, L. Minakhin, S. Borukhov, A. Mustaev, and K. Severinov, "Coupling of downstream RNA polymerase-promoter interactions with formation of catalytically competent transcription initiation complex," *Journal of molecular biology,* vol. 426, no. 24, pp. 3973-3984, 2014.
[33] E. Nudler, "RNA polymerase active center: the molecular engine of transcription," *Annual review of biochemistry,* vol. 78, pp. 335-361, 2009.
[34] Y. L. Juang and J. D. Helmann, "A promoter melting region in the primary sigma factor of Bacillus subtilis. Identification of functionally important aromatic amino acids," *J Mol Biol,* vol. 235, no. 5, pp. 1470-88, Feb 4 1994.

[35] J. Chen, S. A. Darst, and D. Thirumalai, "Promoter melting triggered by bacterial RNA polymerase occurs in three steps," *Proceedings of the National Academy of Sciences,* vol. 107, no. 28, p. 12523, 2010.
[36] K. Fredrick and J. D. Helmann, "RNA polymerase sigma factor determines start-site selection but is not required for upstream promoter element activation on heteroduplex (bubble) templates," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 94, no. 10, pp. 4982-4987, 1997.
[37] P. J. Cossar *et al.*, "Small-Molecule Inhibitors of the NusB–NusE Protein–Protein Interaction with Antibiotic Activity," *ACS Omega,* vol. 2, no. 7, pp. 3839-3857, 2017/07/31 2017.
[38] A. S. Altieri *et al.*, "The structure of the transcriptional antiterminator NusB from Escherichia coli," *Nat Struct Biol,* vol. 7, no. 6, pp. 470-4, Jun 2000.
[39] J. Drogemuller, M. Strauss, K. Schweimer, B. M. Wohrl, S. H. Knauer, and P. Rosch, "Exploring RNA polymerase regulation by NMR spectroscopy," *Sci Rep,* vol. 5, p. 10825, Jun 4 2015.
[40] R. Das *et al.*, "Structural biophysics of the NusB:NusE antitermination complex," *Journal of molecular biology,* vol. 376, no. 3, pp. 705-720, 2008.
[41] M. J. Kazmierczak, M. Wiedmann, and K. J. Boor, "Alternative sigma factors and their roles in bacterial virulence," *Microbiology and molecular biology reviews : MMBR,* vol. 69, no. 4, pp. 527-543, 2005.
[42] M. S. Paget, "Bacterial Sigma Factors and Anti-Sigma Factors: Structure, Function and Distribution," *Biomolecules,* vol. 5, no. 3, pp. 1245-1265, 2015.
[43] R. A. Mooney, S. A. Darst, and R. Landick, "Sigma and RNA Polymerase: An On-Again, Off-Again Relationship?," *Molecular Cell,* vol. 20, no. 3, pp. 335-345, 2005.
[44] V. M Cook and P. DeHaseth, *Strand opening-deficient Escherichia coli RNA polymerase facilitates investigation of closed complexes with promoter DNA - Effects of DNA sequence and temperature*. 2007, pp. 21319-26.
[45] I. Petushkov, D. Esyunina, V. Mekler, K. Severinov, D. Pupov, and A. Kulbachinskiy, "Interplay between sigma region 3.2 and secondary channel factors during promoter escape by bacterial RNA polymerase," *Biochem J,* vol. 474, no. 24, pp. 4053-4064, Dec 1 2017.
[46] M. Imashimizu, N. Shimamoto, T. Oshima, and M. Kashlev, "Transcription elongation. Heterogeneous tracking of RNA polymerase and its biological implications," *Transcription,* vol. 5, no. 1, pp. e28285-e28285, 2014.
[47] C. L. Ventola, "The antibiotic resistance crisis: part 1: causes and threats," *P & T : a peer-reviewed journal for formulary management,* vol. 40, no. 4, pp. 277-283, 2015.
[48] R. I. Aminov, "A brief history of the antibiotic era: lessons learned and challenges for the future," *Frontiers in microbiology,* vol. 1, pp. 134-134, 2010.
[49] I. Artsimovitch, V. Svetlov, S. M. Nemetski, V. Epshtein, T. Cardozo, and E. Nudler, "Tagetitoxin inhibits RNA polymerase through trapping of the trigger loop," *The Journal of biological chemistry,* vol. 286, no. 46, pp. 40395-40400, 2011.
[50] H. Yamada, M. Adachi, and T. Nishikawa, "Stereocontrolled synthesis of the oxathiabicyclo[3.3.1]nonane core structure of tagetitoxin," *Chemical Communications,* 10.1039/C3CC46949B vol. 49, no. 95, pp. 11221-11223, 2013.
[51] D. G. Vassylyev *et al.*, "Structural basis for transcription inhibition by tagetitoxin," *Nature structural & molecular biology,* vol. 12, no. 12, pp. 1086-1093, 2005.

[52] K.-A. Wilson *et al.*, "Structure of Microcin J25, a Peptide Inhibitor of Bacterial RNA Polymerase, is a Lassoed Tail," *Journal of the American Chemical Society*, vol. 125, no. 41, pp. 12475-12483, 2003/10/01 2003.
[53] J. Yuzenkova *et al.*, "Mutations of bacterial RNA polymerase leading to resistance to microcin j25," *J Biol Chem*, vol. 277, no. 52, pp. 50867-75, Dec 27 2002.
[54] J. Mukhopadhyay, E. Sineva, J. Knight, R. M. Levy, and R. H. Ebright, "Antibacterial Peptide Microcin J25 Inhibits Transcription by Binding within and Obstructing the RNA Polymerase Secondary Channel," *Molecular Cell*, vol. 14, no. 6, pp. 739-751, 2004/06/18/ 2004.
[55] A. Srivastava *et al.*, "New target for inhibition of bacterial RNA polymerase: 'switch region'," *Current Opinion in Microbiology*, vol. 14, no. 5, pp. 532-543, 2011/10/01/ 2011.
[56] J. Mukhopadhyay *et al.*, "The RNA Polymerase "Switch Region" Is a Target for Inhibitors," *Cell*, vol. 135, no. 2, pp. 295-307, 2008/10/17/ 2008.
[57] A. Tupin, M. Gualtieri, J.-P. Leonetti, and K. Brodolin, "The transcription inhibitor lipiarmycin blocks DNA fitting into the RNA polymerase catalytic site," *The EMBO journal*, vol. 29, no. 15, pp. 2527-2537, 2010.
[58] I. Artsimovitch, J. Seddon, and P. Sears, "Fidaxomicin Is an Inhibitor of the Initiation of Bacterial RNA Synthesis," *Clinical Infectious Diseases*, vol. 55, no. suppl_2, pp. S127-S131, 2012.
[59] N. Jackson, L. Czaplewski, and L. J. V. Piddock, "Discovery and development of new antibacterial drugs: learning from experience?," *J Antimicrob Chemother*, vol. 73, no. 6, pp. 1452-1459, Jun 1 2018.
[60] M. Gajdács, "The Concept of an Ideal Antibiotic: Implications for Drug Design," *Molecules (Basel, Switzerland)*, vol. 24, no. 5, p. 892, 2019.
[61] S. Höjgård, "Antibiotic resistance - why is the problem so difficult to solve?," *Infection ecology & epidemiology*, vol. 2, p. 10.3402/iee.v2i0.18165, 2012.
[62] W. Zhu, J. Haupenthal, M. Groh, M. Fountain, and R. W. Hartmann, "New Insights into the Bacterial RNA Polymerase Inhibitor CBR703 as a Starting Point for Optimization as an Anti-Infective Agent," *Antimicrobial Agents and Chemotherapy*, vol. 58, no. 7, p. 4242, 2014.
[63] J. Ye, A. J. Chu, L. Lin, X. Yang, and C. Ma, "First-In-Class Inhibitors Targeting the Interaction between Bacterial RNA Polymerase and Sigma Initiation Factor Affect the Viability and Toxin Release of Streptococcus pneumoniae," *Molecules*, vol. 24, no. 16, Aug 9 2019.
[64] X. Yang *et al.*, "First-In-Class Inhibitor of Ribosomal RNA Synthesis with Antimicrobial Activity against Staphylococcus aureus," *Biochemistry*, vol. 56, no. 38, pp. 5049-5052, 2017/09/26 2017.
[65] Y. Qiu *et al.*, "Nusbiarylins, a new class of antimicrobial agents: Rational design of bacterial transcription inhibitors targeting the interaction between the NusB and NusE proteins," *Bioorganic Chemistry*, vol. 92, p. 103203, 2019/11/01/ 2019.
[66] C. Ma, X. Yang, and P. J. Lewis, "Bacterial Transcription Inhibitor of RNA Polymerase Holoenzyme Formation by Structure-Based Drug Design: From in Silico Screening to Validation," *ACS Infectious Diseases*, vol. 2, no. 1, pp. 39-46, 2016/01/08 2016.
[67] A. Malhotra, E. Severinova, and S. A. Darst, "Crystal structure of a sigma 70 subunit fragment from E. coli RNA polymerase," *Cell*, vol. 87, no. 1, pp. 127-36, Oct 4 1996.

[68] M. Mielczarek *et al.*, "Synthesis and biological activity of novel bis-indole inhibitors of bacterial transcription initiation complex formation," *Organic & Biomolecular Chemistry*, 10.1039/C4OB00460D vol. 12, no. 18, pp. 2882-2894, 2014.
[69] C. Ma *et al.*, "Inhibitors of Bacterial Transcription Initiation Complex Formation," *ACS Chemical Biology*, vol. 8, no. 9, pp. 1972-1980, 2013/09/20 2013.
[70] E. B. Johnston, P. J. Lewis, and R. Griffith, "The interaction of Bacillus subtilis sigmaA with RNA polymerase," *Protein Sci*, vol. 18, no. 11, pp. 2287-97, Nov 2009.
[71] T. M. Arthur, Anthony, Larry C., Burgess, Richard R., "Mutational Analysis of b*260–309, a s70 Binding Site Located on
Escherichia coli Core RNA Polymerase," *Journal of Biological Chemistry*, vol. 275, no. 30, pp. 23113-23119, 2000.
[72] D. S. Wenholz *et al.*, "Small molecule inhibitors of bacterial transcription complex formation," *Bioorganic & medicinal chemistry letters*, vol. 27, no. 18, pp. 4302-4308, 2017/09// 2017.
[73] P. Sensi, A. M. Greco, and R. Ballotta, "Rifomycin. I. Isolation and properties of rifomycin B and rifomycin complex," *Antibiot Annu*, vol. 7, pp. 262-70, 1959.
[74] P. Sensi, "History of the Development of Rifampin," *Reviews of Infectious Diseases*, vol. 5, no. Supplement_3, pp. S402-S406, 1983.
[75] V. Lorian and M. Finland, "In vitro effect of rifampin on mycobacteria," *Applied microbiology*, vol. 17, no. 2, pp. 202-207, 1969.
[76] L. Hyde, "Rifampin in the treatment of pulmonary tuberculosis," *California medicine*, vol. 117, no. 6, pp. 18-21, 1972.
[77] M. C. Raviglione, D. E. Snider, Jr, and A. Kochi, "Global Epidemiology of Tuberculosis: Morbidity and Mortality of a Worldwide Epidemic," *JAMA*, vol. 273, no. 3, pp. 220-226, 1995.
[78] E. A. Campbell *et al.*, "Structural Mechanism for Rifampicin Inhibition of Bacterial RNA Polymerase," *Cell*, vol. 104, no. 6, pp. 901-912, 2001/03/23/ 2001.
[79] G. Hartmann, K. O. Honikel, F. Knüsel, and J. Nüesch, "The specific inhibition of the DNA-directed RNA synthesis by rifamycin," *Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis*, vol. 145, no. 3, pp. 843-844, 1967/10/16/ 1967.
[80] S. Kumar and L. Jena, "Understanding Rifampicin Resistance in Tuberculosis through a Computational Approach," *Genomics & informatics*, vol. 12, no. 4, pp. 276-282, 2014.
[81] M. Talpaert, F. Campagnari, and L. Clerici, "Lipiarmycin: An antibiotic inhibiting nucleic acid polymerases," *Biochemical and Biophysical Research Communications*, vol. 63, no. 1, pp. 328-334, 1975/03/03/ 1975.
[82] W. Lin *et al.*, "Structural Basis of Transcription Inhibition by Fidaxomicin (Lipiarmycin A3)," *Mol Cell*, vol. 70, no. 1, pp. 60-71.e15, Apr 5 2018.
[83] T. M. Arthur, Burgess, Richard R., "Localization of a ς70 Binding Site on the N Terminus of the Escherichia coli RNA Polymerase β´ Subunit," *Journal of Biological Chemistry*, vol. 273, no. 47, pp. 31381-31387, 1998.
[84] M. Mielczarek *et al.*, "Synthesis and biological activity of novel mono-indole and mono-benzofuran inhibitors of bacterial transcription initiation complex formation," *Bioorganic & Medicinal Chemistry*, vol. 23, no. 8, pp. 1763-1775, 2015/04/15/ 2015.
[85] H. Kandemir *et al.*, "Synthesis and biological evaluation of 2,5-di(7-indolyl)-1,3,4-oxadiazoles, and 2- and 7-indolyl 2-(1,3,4-thiadiazolyl)ketones as antimicrobials," *Bioorganic & Medicinal Chemistry*, vol. 22, no. 5, pp. 1672-1679, 2014/03/01/ 2014.

[86] Y. Futamura *et al.*, "Morphobase, an Encyclopedic Cell Morphology Database, and Its Use for Drug Target Identification," *Chemistry & Biology,* vol. 19, no. 12, pp. 1620-1630, 2012/12/21/ 2012.
[87] L. Marcellini, M. Giammatteo, P. Aimola, and M. L. Mangoni, "Fluorescence and Electron Microscopy Methods for Exploring Antimicrobial Peptides Mode(s) of Action," in *Antimicrobial Peptides: Methods and Protocols,* A. Giuliani and A. C. Rinaldi, Eds. Totowa, NJ: Humana Press, 2010, pp. 249-266.
[88] T. P. T. Cushnie, N. H. O'Driscoll, and A. J. Lamb, "Morphological and ultrastructural changes in bacterial cells as an indicator of antibacterial mechanism of action," *Cellular and Molecular Life Sciences,* vol. 73, no. 23, pp. 4471-4492, 2016/12/01 2016.
[89] M. B. Sadiq, J. Tarning, T. Z. Aye Cho, and A. K. Anal, "Antibacterial Activities and Possible Modes of Action of Acacia nilotica (L.) Del. against Multidrug-Resistant Escherichia coli and Salmonella," *Molecules (Basel, Switzerland),* vol. 22, no. 1, p. 47, 2017.
[90] M. C. F. van Teeseling, M. A. de Pedro, and F. Cava, "Determinants of Bacterial Morphology: From Fundamentals to Possibilities for Antimicrobial Targeting," *Frontiers in microbiology,* vol. 8, pp. 1264-1264, 2017.
[91] V. Lorian, L. D. Sabath, M. Simionescu, and D. W. Watson, "Decrease in ribosomal density of Proteus mirabilis exposed to subinhibitory concentrations of ampicillin or cephalothin," *Proc Soc Exp Biol Med,* vol. 149, no. 3, pp. 731-5, Jul 1975.
[92] P. J. Lewis, S. D. Thaker, and J. Errington, "Compartmentalization of transcription and translation in Bacillus subtilis," *The EMBO journal,* vol. 19, no. 4, pp. 710-718, 2000.
[93] S. J. Wright and G. Schatten, "Confocal fluorescence microscopy and three-dimensional reconstruction," *J Electron Microsc Tech,* vol. 18, no. 1, pp. 2-10, May 1991.
[94] J. E. Bylund, M. A. Haines, P. J. Piggot, and M. L. Higgins, "Axial filament formation in Bacillus subtilis: induction of nucleoids of increasing length after addition of chloramphenicol to exponential-phase cultures approaching stationary phase," *J Bacteriol,* vol. 175, no. 7, pp. 1886-90, Apr 1993.
[95] R. A. Weisberg and M. E. Gottesman, "Processive Antitermination," *Journal of Bacteriology,* vol. 181, no. 2, p. 359, 1999.
[96] G. Wegrzyn, K. Licznerska, and A. Wegrzyn, "Phage lambda--new insights into regulatory circuits," *Adv Virus Res,* vol. 82, pp. 155-78, 2012.
[97] I. Bonin *et al.*, "Crystal structures of the antitermination factor NusB from Thermotoga maritima and implications for RNA binding," *The Biochemical journal,* vol. 383, no. Pt. 3, pp. 419-428, 2004.
[98] A. S. Altieri *et al.*, "Structural basis for RNA recognition by NusB and NusE in the initiation of transcription antitermination," *Nucleic Acids Research,* vol. 39, no. 17, pp. 7803-7815, 2011.
[99] X. Luo *et al.*, "Structural and Functional Analysis of the E. coli NusB-S10 Transcription Antitermination Complex," *Molecular Cell,* vol. 32, no. 6, pp. 791-802, 2008/12/26/ 2008.
[100] M. E. A. de Kraker, A. J. Stewardson, and S. Harbarth, "Will 10 Million People Die a Year due to Antimicrobial Resistance by 2050?," *PLOS Medicine,* vol. 13, no. 11, p. e1002184, 2016.
[101] A. Cassini *et al.*, "Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the EU and the European Economic Area in

2015: a population-level modelling analysis," *The Lancet Infectious Diseases,* vol. 19, no. 1, pp. 56-66, 2019.
[102] Y. Qiu *et al.*, "Design, synthesis and biological evaluation of antimicrobial diarylimine and –amine compounds targeting the interaction between the bacterial NusB and NusE proteins," *European Journal of Medicinal Chemistry,* vol. 178, pp. 214-231, 2019/09/15/ 2019.
[103] B. Gopal *et al.*, "The crystal structure of NusB from Mycobacterium tuberculosis," *Nat Struct Biol,* vol. 7, no. 6, pp. 475-8, Jun 2000.
[104] J. Drögemüller, M. Strauß, K. Schweimer, B. M. Wöhrl, S. H. Knauer, and P. Rösch, "Exploring RNA polymerase regulation by NMR spectroscopy," *Scientific Reports,* Article vol. 5, p. 10825, 06/04/online 2015.
[105] N. Said *et al.*, "Structural basis for λN-dependent processive transcription antitermination," *Nature Microbiology,* Article vol. 2, p. 17062, 04/28/online 2017.
[106] M. L. Doyle, "Characterization of binding interactions by isothermal titration calorimetry," *Current Opinion in Biotechnology,* vol. 8, no. 1, pp. 31-35, 1997/02/01/ 1997.
[107] A. Kondo, F. Murakami, and K. Higashitani, "Circular dichroism studies on conformational changes in protein molecules upon adsorption on ultrafine polystyrene particles," *Biotechnology and Bioengineering,* vol. 40, no. 8, pp. 889-894, 1992/10/20 1992.
[108] L. S. Swapna, S. Mahajan, A. G. de Brevern, and N. Srinivasan, "Comparison of tertiary structures of proteins in protein-protein complexes with unbound forms suggests prevalence of allostery in signalling proteins," *BMC Struct Biol,* vol. 12, p. 6, May 3 2012.
[109] N. J. Greenfield, "Using circular dichroism spectra to estimate protein secondary structure," *Nature protocols,* vol. 1, no. 6, pp. 2876-2890, 2006.
[110] A. J. Miles and B. A. Wallace, "Circular dichroism spectroscopy of membrane proteins," *Chemical Society Reviews,* 10.1039/C5CS00084J vol. 45, no. 18, pp. 4859-4872, 2016.
[111] B. Ranjbar and P. Gill, "Circular dichroism techniques: biomolecular and nanostructural analyses- a review," *Chem Biol Drug Des,* vol. 74, no. 2, pp. 101-20, Aug 2009.
[112] A. Micsonai *et al.*, "Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy," *Proceedings of the National Academy of Sciences,* vol. 112, no. 24, p. E3095, 2015.
[113] A. Micsonai *et al.*, "BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra," *Nucleic Acids Research,* vol. 46, no. W1, pp. W315-W322, 2018.
[114] X. Yang *et al.*, "First-In-Class Inhibitor of Ribosomal RNA Synthesis with Antimicrobial Activity against Staphylococcus aureus," *Biochemistry,* vol. 56, pp. 5049-5052, 2017.
[115] K. Ishii, M. Noda, and S. Uchiyama, "Mass spectrometric analysis of protein-ligand interactions," *Biophysics and physicobiology,* vol. 13, pp. 87-95, 2016.
[116] H. Hernandez and C. V. Robinson, "Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry," *Nat Protoc,* vol. 2, no. 3, pp. 715-26, 2007.
[117] A. Schmidt, U. Bahr, and M. Karas, "Influence of pressure in the first pumping stage on analyte desolvation and fragmentation in nano-ESI MS," *Anal Chem,* vol. 73, no. 24, pp. 6040-6, Dec 15 2001.
[118] P. S. Ciboroswki, J., "Protein Extration and Precipitation," in *Proteomic Profiling and Analytical Chemistry The Crossroads*2nd ed.: Elsevier, 2016, pp. 56-57.

[119] N. Sreerama, S. Y. Venyaminov, and R. W. Woody, "Estimation of the number of alpha-helical and beta-strand segments in proteins using circular dichroism spectroscopy," *Protein science : a publication of the Protein Society,* vol. 8, no. 2, pp. 370-380, 1999.
[120] S. M. Kelly and N. C. Price, "The use of circular dichroism in the investigation of protein structure and function," *Curr Protein Pept Sci,* vol. 1, no. 4, pp. 349-84, Dec 2000.
[121] W. C. Johnson, "Secondary Structure of Proteins Through Circular Dichroism Spectroscopy," *Annual Review of Biophysics and Biophysical Chemistry,* vol. 17, no. 1, pp. 145-166, 1988/06/01 1988.
[122] B. M. Burmann, X. Luo, P. Rösch, M. C. Wahl, and M. E. Gottesman, "Fine tuning of the E. coli NusB:NusE complex affinity to BoxA RNA is required for processive antitermination," *Nucleic acids research,* vol. 38, no. 1, pp. 314-326, 2010.
[123] T. F. Tsang, Y. Qiu, L. Lin, J. Ye, C. Ma, and X. Yang, "Simple Method for Studying in Vitro Protein–Protein Interactions Based on Protein Complementation and Its Application in Drug Screening Targeting Bacterial Transcription," *ACS Infectious Diseases,* vol. 5, no. 4, pp. 521-527, 2019/04/12 2019.