8-26-2022

Nla28 Activator Modulates σ54-dependent Gene Regulation in Myxococcus xanthus

Muqing Ma
Syracuse University, mma109@syr.edu

ABSTRACT

The complex life cycle of *Myxococcus xanthus* makes it a model organism for studying multicellular developmental processes in bacteria. In response to environmental changes, *M. xanthus* cells collectively and dramatically adjust their morphology, physiological functions, and biosynthesis. Particularly, *M. xanthus* copes with starvation by producing fruiting bodies filled with dormant and stress-resistant spores. During secondary metabolism, it synthesizes polyketide (PK) and non-ribosomal peptide (NRP) natural products that are crucial sources of antibiotics. Bacteria including *M. xanthus*, require the alternative σ^{54} regulatory system to initiate the transcription of necessary genes for settling stress and completing developmental process. This dissertation introduces in detail how σ^{54} system integrates regulatory events in transcription by applying σ^{54}, targeted σ^{54} promoter DNA and corresponding activator proteins in mechanistic and genomic scales. Also, multicellular lifestyle of *M. xanthus* and natural products derived from myxobacteria are broadly reviewed.

My first project has better defined the gene regulatory networks of *M. xanthus* in response to starvation, which are modulated by σ^{54} system and a transcriptional activator, Nla28. Specifically, highly conserved DNA sequence where Nla28 targets has been characterized; Direct Nla28-mediated developmental genes have been identified and verified; Regulatory pathways and biofunctions associated with *M. xanthus* starvation response have been profiled and classified. My second project investigated the hypothesis that σ^{54} might be a common regulator of natural product genes in bacteria. The results suggest that *M. xanthus* employs σ^{54} system to regulate the

transcription of at least some of its natural product genes after analyzing its potential natural product σ^{54} promoters. Also, it shows that these natural product gene regulations in *M. xanthus* are modulated by Nla28.

Nla28 Activator Modulates σ^{54}-dependent Gene Regulation in *Myxococcus xanthus*

By

Muqing Ma

B.S., Jilin University, 2013

M.S., Jilin University, 2016

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Biology

Syracuse University

August 2022

ACKNOWLEDGEMENTS

I would like to acknowledge many people for providing me a lot of assistance with my life in Syracuse University. My achievements in research would not have been possible without their support.

I would especially like to thank my advisers, Dr. Anthony Garza, and Dr. Roy Welch, for their patient guidance, endless support, and encouragement. For me, they are great mentors and treasured friends. Also, I want to thank my research committee members, Dr. Eleanor Maine, Dr. Sarah Hall and Dr. Scott Erdman for their research advice and guidance.

My thanks extend to my excellent lab mates and friends, Jessica, Eddie, Trosporsha and Linnea. Their helps and supports have made a huge impact on my work and life over these last few years.

Lastly and mostly, I would like to thank my wife, Haiyan. Thank you for always being there with me. Without your love and support, I would not have the courage to keep studying and keep working hard.

TABLE OF CONTENTS

List of Figures

Chapter 2

Chapter 3

Chapter 4

Chapter 5

List of Tables

Chapter 4

Chapter 5

List of Abbreviations

DBD	DNA binding domain
EBP	Enhancer binding protein
EMSA	Electrophoretic mobility shift assay
EPS	Extracellular polysaccharide
FPLC	Fast protein liquid chromatography
HTH	Helix-turn-helix
IHF	Integration host factors
NRP	Non-ribosomal peptides
NRPS	Non-ribosomal peptide synthetase
NtrC	Nitrogen regulatory protein C
ORF	Open reading frames
PK	Polyketides
PKS	Polyketide synthase
(p)ppGpp	Guanine-5'(tri)di-phosphate 3'diphosphate
QS	Quorum sensing
RNAP	RNA polymerase
T4P	Type IV pili
TCS	Two component signal transductions
TSS	Transcription start site

Chapter 1

Introduction

1.1 Motivation and Research Goals

The σ^{54} system regulates the transcription of a wide variety of bacterial genes. One crucial component of this regulatory system is the alternative factor σ^{54}, which directs the binding of RNA polymerase (RNAP) to the conserved -12 and -24 regions of σ^{54} promoter sequence. Enhancer binding proteins (EBPs) functioning as transcriptional activators, are also crucial for the σ^{54} regulatory system. EBPs provide energy to open closed σ^{54}-RNAP complex, thus initiate transcription.

In response to starvation, *Myxococcus xanthus* undergoes multicellular development during which cells collectively and dramatically adjust their morphology, physiological functions, and biosynthesis. The early developmental pathway of *M. xanthus* relies heavily on EBPs. Nla28 protein is a response regulator-type σ^{54} EBP previously validated to regulate important signal production at early stage of development in *M. xanthus*. Also, gene expression analysis has demonstrated Nla28 is important for the expression of hundreds of *M. xanthus* developmental genes. However, little is known about the wide regulatory pathways that bacteria employ to modulate starvation response and natural product synthesis. Specifically, direct Nla28 regulons associated with starvation response and natural product synthesis *in M. xanthus* are yet identified and characterized.

There were three major research goals in my first project 'Nla28-mediated σ^{54}-dependent regulation of starvation response genes in *M. xanthus*': the first goal was to determine how Nla28 identifies its target σ^{54} promoters; the second goal was to use this information to identify the larger network of Nla28-targeted σ^{54} promoters, with an emphasis on promoters that show substantial developmental regulation; the final goal

was to analyze genes under direct Nla28-control to better understand the function of Nla28.

My second project 'Nla28-mediated σ^{54}-dependent regulation of natural product genes in *Myxococcus xanthus*' had two major research goals: the first goal was to verify the putative PK/NRP natural product σ^{54} promoters identified in our previous bioinformatics analysis; the second goal was to determine whether Nla28 serves as the EBP that modulates the σ^{54}-dependent regulation of these natural product genes.

1.2 Contributions to the Field

My first project characterized the highly conserved DNA sequence targeted by Nla28 protein, which is an important σ^{54} transcriptional activator regulating developmental genes in *M. xanthus*. It helps better understand how σ^{54} activator functions.

My first project also collected a pool of 140 developmental genes in *M. xanthus* potentially regulated by Nla28 in response to starvation. Furthermore, 12 of these genes associated with regulatory or signal transduction, motility, and defense mechanisms in *M. xanthus* are verified to be Nla28-regulons. These findings better define the early gene regulatory pathways involved in the starvation response of *M. xanthus*.

In my second project, bona fide σ^{54} promoters and corresponding Nla28-targeted sequences were identified and characterized in natural product genes in *M. xanthus*. It has uncovered the regulatory mechanisms underlying bacterial natural product synthesis, especially in the most complicated producers such as Myxobacteria. The

hypothesis that σ^{54} system might be a common regulator of natural product genes has been further validated in this project.

Most of the *M. xanthus* starvation response and natural product σ^{54} promoters that we characterized in these two projects are located in the coding sequences of genes (intragenic) instead of typically in intergenic regions. Taken together with many potential intragenic σ^{54} promoters in other bacteria identified in other genomic-wide studies, our finding has provided more evidence to a novel view that some of bacterial promoter elements are within intragenic regions, which deserves to be further investigated.

1.3 Dissertation Organization

The following dissertation is outlined as follows:

• **Chapter 2** – σ^{54} Background and Review

• **Chapter 3** – *Myxococcus xanthus* Background and Review

• **Chapter 4** – Nla28-Mediated σ^{54}-Dependent Regulation of Starvation Response Genes in *Myxoccocus xanthus*

• **Chapter 5** – Nla28-Mediated σ^{54}-Dependent Regulation of Natural Product Genes in *Myxococcus xanthus*

• **Chapter 6** – Conclusions and Future Directions

Chapter 2

σ^{54} Background and Review

2.1 Introduction

In bacteria, to accomplish the major cellular function of transcription, RNA polymerase (RNAP) must efficiently span a large fraction of the genome without preference for DNA sequence during the elongation phase of transcription. However, promoter recognition and initiation of transcription must be precisely controlled and tightly regulated. The finding that bacteria require a type of specialized protein, sigma (σ) factor associated with RNAP to recognize specific promoter DNA opened a new era in understanding regulated transcription which is essential for cellular adaption, differentiation, and growth. This finding introduced that bacterial transcription initiation is achieved when dissociable σ factors bind core RNAP and direct the resulting σ-RNAP holoenzyme to recognized promoter DNA [1]. Because bacterial σ factors implement all transcription initiation functions, they have become an ideal model for studying the mechanism under transcription initiation. All bacteria possess a housekeeping σ factor responsible for transcription of thousands of genes essential for growth from majority of promoters, and additional alternative σ factors that redirect RNAP to distinct promoters necessary for stress response or development across different species reflecting their lifestyle. Predominant σ^{70}-family named after the housekeeping σ^{70} of *Escherichia coli* and the alternative σ^{54} constitute the two distinct classes of σ factors in bacteria [2]. It is sufficient for promoter DNA melting double strands and initiating transcription when it binds to σ^{70}-RNAP holoenzyme. This transcription regulatory process has been known as open complex formation [3, 4]. Whereas the alternative σ^{54}-RNAP holoenzyme functions distinctly as it binds to promoter DNA forming a steady closed complex which is unable to activate transcription [5, 6]. Due to the vital regulatory roles of predominant

σ^{70}-family, it is not surprising that they have been extensively studied since their discovery, by contrast, there have been limited studies on σ^{54}. This section of dissertation is entirely and exclusively introducing the status of important broad trends in σ^{54}: how σ^{54} was discovered, identified, and classified; how σ^{54} performs its critical function of directing RNAP to promoters and initiating DNA-strand opening; how σ^{54} further integrate regulatory events in transcription. Overall, this chapter highlights the mechanistic, genomic, and regulatory roles of σ^{54} factor.

2.2 σ^{54} Promoter and RNAP-σ^{54}-DNA Complex Formation

In 1999, the study of Barrios et al. first demonstrated that unlike the predominant σ^{70} family, which recognizes the consensus -35 (TATAAT) and -10 (TTGACA) promoter DNA elements (-35 and -10 boxes), the alternative sigma factor σ^{54} binds to different consensus sequences that are more strongly conserved than those for σ^{70}. Binding occurs at the positions -24 (GG) and -12 (GC) relative to the +1 transcription start site (TSS) that are part of the wider consensus sequence YTGGCACGNNNNTTGCW (N is non-conserved, Y represents pyrimidines, and W is A or T). By compiling 186 -24/-12 promoter sequences reported in literature and generating an updated and extended consensus sequence, they successfully increased the probability of identifying genuine -24/-12 promoters [7]. Binding of σ^{54}-RNAP holoenzyme to σ^{54} promoter forms a stable closed promoter complex that is unable to spontaneously melt DNA and initiate transcripts, known as transcriptionally silent [5]. The activation of σ^{54}-RNAP occurs by signal transduction pathways using numerous and diverse activators. These signal transduction pathways have a common mechanism: the opening of closed transcription

complex involves the triggering of ATPase activity within an enhancer-binding activator protein (See **Fig. 2.1**). This ATPase is then used to overcome the block to DNA melting within the closed transcription complex and thereby allow transcription to initiate [8].

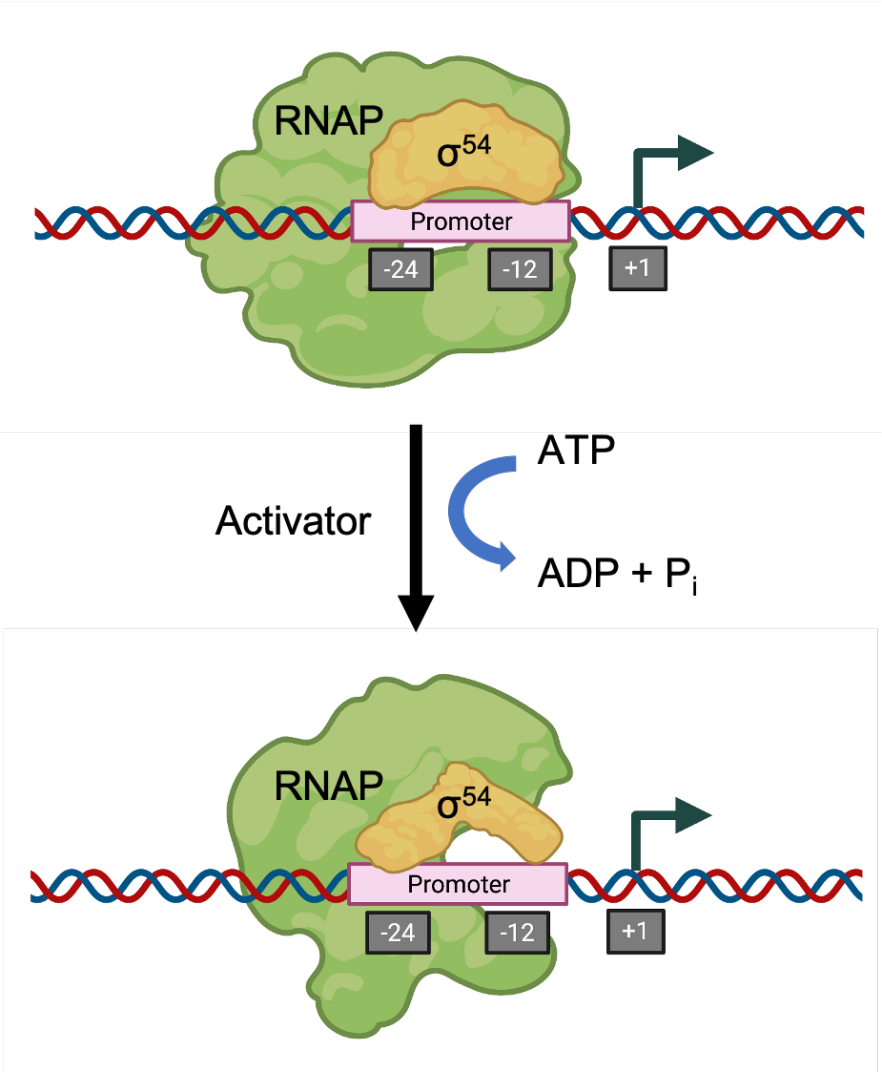


Figure 2.1 Model of activator-dependent σ^{54} transcriptional activation

σ^{54}-RNA polymerase holoenzyme binds to -24 and -12 consensus regions of σ^{54} promoters to form a stable closed complex which is transcriptionally silent. Activator-mediated ATP hydrolysis drives full open promoter complex formation.

σ^{54} factor contains a major RNAP core-binding domain (CBD), a flexible RpoN domain recognizing the -12 elements, Region III helix-turn-helix (RIII-HTH) domain recognizing the -24 elements, Region I (RI) domain interacting with activator protein and Region II (RII) domain [5]. σ^{54} structural domains are contained within four regions connected by long coils and loops that span a large area of core RNAP, which consists of two α, β, β' and ω subunits [9, 10]. σ^{54} Region I (RI, residues 1-56) forms a hook comprised of two α helices. Region II (RII) also contains two α-helices in addition to loops that are buried inside the RNAP. The major RNAP CBD of σ^{54} extends as a structural fold to residue 250 and consists of two α helical sub-domains. Following on from the CBD, the backbone extends back to connect to a loop region, before an extra-long α-helix (ELH) followed by the RIII-HTH domain, involved in interaction with the -12 promoter elements. The flexible RpoN domain, responsible for recognizing the -24 promoter elements, consists of a three-helical bundle [5].

At the beginning of the formation of RNAP-σ^{54} holoenzyme, the σ^{54} polypeptide chain convolves back and forth through its loop regions embedded in the RNAP. The CBD of σ^{54} binds the upstream face of the RNAP making extensive interactions with many functional modules within the RNAP including the β-flap, the C-terminus of the β subunit, the β' zipper/Zn-binding domain, the β'dock domain as well as the α subunit carboxyl terminal domain [11]. The RpoN domain of σ^{54} is the most conserved domain in σ^{54} among different bacteria, however, this domain extends flexibly from the main body of RNAP and does not contact other parts of RNAP-σ^{54}-DNA complex. Unlike the flexible RpoN domain, the RIII-HTH domain of σ^{54}, which binds to the -12 region of promoter, stably associates with the RNAP core through RI-RIII. Yang's study has

shown that -12 binding is a major functional determinant for σ^{54} holoenzyme promoter recognition as well as stable closed complex formation [12]. The σ^{54} RI domain plays an inhibitory role and contains contact sites for its cognate activator proteins [13, 14]. RI domain interacts with RIII, forming a structural module that lies along the cleft between β and β', where template strand DNA enters the active site cleft [15]. RII domain penetrates deeply into the DNA binding channel which occupies the space of the DNA template strand, indicating that this domain needs conformational change to permit template-strand DNA access into the RNAP active site and for transcription initiation [12].

2.3 Bacterial Enhancer Binding Proteins (EBPs)

As mentioned before, although σ^{54} factor interacts with the identical core RNAP enzyme with σ^{70} family members, the resulting holoenzymes initiate transcription via completely distinct mechanisms. The binding of σ^{70}-RNAP holoenzyme to promoter DNA forms an open complex which is sufficient for melting the duplex promoter DNA around the TSS and initiating transcription [16]. In contrast to the transcription initiation regulated by σ^{70} factor, the distinctive mechanism of σ^{54}- dependent regulation is that the binding of σ^{54}-RNAP holoenzyme to promoter DNA results in a stable closed promoter complex [6]. The stable closed complex is caused by the interaction of RI and RII domains of σ^{54} with a repressive fork junction DNA structure near -12 site, which prevents the binding of holoenzyme to non-template DNA strand. Therefore, it strictly requires assistance from a mechano-transcriptional activator to have a conformational change so that the thermodynamic and kinetic barriers restricting open complex

formation are overcome and transcription initiation is energized [12]. Such transcriptional activators typically bind to the sites 80 to 200 bp upstream of the promoter region which is known as enhancer or upstream activator sequences [17, 18], however, in some cases they can still stimulate transcription when moved thousands of nucleotides upstream or downstream from the transcription start and regardless of their orientation. This is similar to the binding of eukaryotic enhancer binding proteins (EBPs) that can still function when moved several kilobases away, therefore, activators of σ^{54}-dependent transcription are referred to as bacterial EBPs [19].

The conformational change which releases a mechanical force to remodel the closed complex into a transcriptional competent open complex requires energy generated from ATP hydrolysis catalyzed by EBPs. The γ-phosphate of the ATP is sensed and that ATP binding and hydrolysis lead to changes in σ^{54}-RNAP holoenzyme-protomer structure which are key in successful engagement of the closed complex and stimulating its conversion to an open complex [6]. EBPs are highly modular proteins which possess about 300 amino acids and are sufficiently similar in structure and function to be classified as members of the AAA+ (chaperone-like ATPases associated with various cellular activities) family of proteins [20]. They typically consist of three domains: an N-terminal regulatory domain involved in signal-responsive control, a central AAA+ domain which converts chemical energy from ATP hydrolysis into mechanical force on closed complex and a C-terminal HTH domain which mediates DNA binding, although natural variants lacking one or two of these domains exist [21]. EBPs are typically dimeric in their inactive state, however, after receiving specific stimulatory signals via N-terminal regulatory domain, the active forms of EBPs perform

higher order of oligomerization through assembling their central AAA+ domains into ring structures, usually hexamers [22] or heptamers [23], which in turn stimulates their ATPase activity and transcription activation. Studies on the EBP NorR in *E. coli* has also proved that multiple inactive EBP dimers can bind to adjacent enhancer sites which results in an increased local concentration of EBPs and facilitates oligomerization [24, 25]. Oligomerization is essential for ATPase activity of EBPs and transcription activation because the ATP-binding pocket is formed by adjacent dimers stabilized by arginine fingers. The resulting cyclic structure indicates that nucleotide binding and hydrolysis events that leads to conformational changes can be transmitted to adjacent dimers to give rise to mechanical motion.

Because EBPs bind relatively far upstream of promoter DNA (80 to 200 or even thousands of nucleotides upstream), another specific feature is that DNA-binding EBP must loop the DNA between itself and the promoter region for the EBP to directly interact with the σ^{54}-RNAP holoenzyme (See **Fig. 2.2**). In 1990, such DNA-looping structure was first discovered and visualized between NtrC and RNAP at the *glnA* promoter by electron microscopy [26]. Intriguingly, it has been demonstrated that compared with the promoter bound σ^{54}-RNAP closed complex, such type of DNA-looping must direct EBPs bind to the opposite face of the promoter DNA helix for efficient transcription initiation [27]. In this case, the process of DNA looping is commonly assisted by a class of small heterodimeric proteins, known as the integration host factors (IHF), which bind between the promoter and enhancer sites to bend the DNA up to 180° [28]. For example, at the *pspA* promoter, IHF has been shown to mediate architectural changes that aid the binding of the EBP PspF and increase

transcriptional output [29]. The EBP GAFTGA signature motif in the central AAA+ domain has been tested to be responsible for the major interaction between EBPs and the closed complex since mutations within this motif generally abolish interaction with σ^{54} and transcription initiation [30]. Once DNA looping and contact between EBP and the closed complex have occurred, the EBP performs ATP hydrolysis to facilitate conformational remodeling in the holoenzyme which converts the closed complex into an open complex for transcription initiation [6].

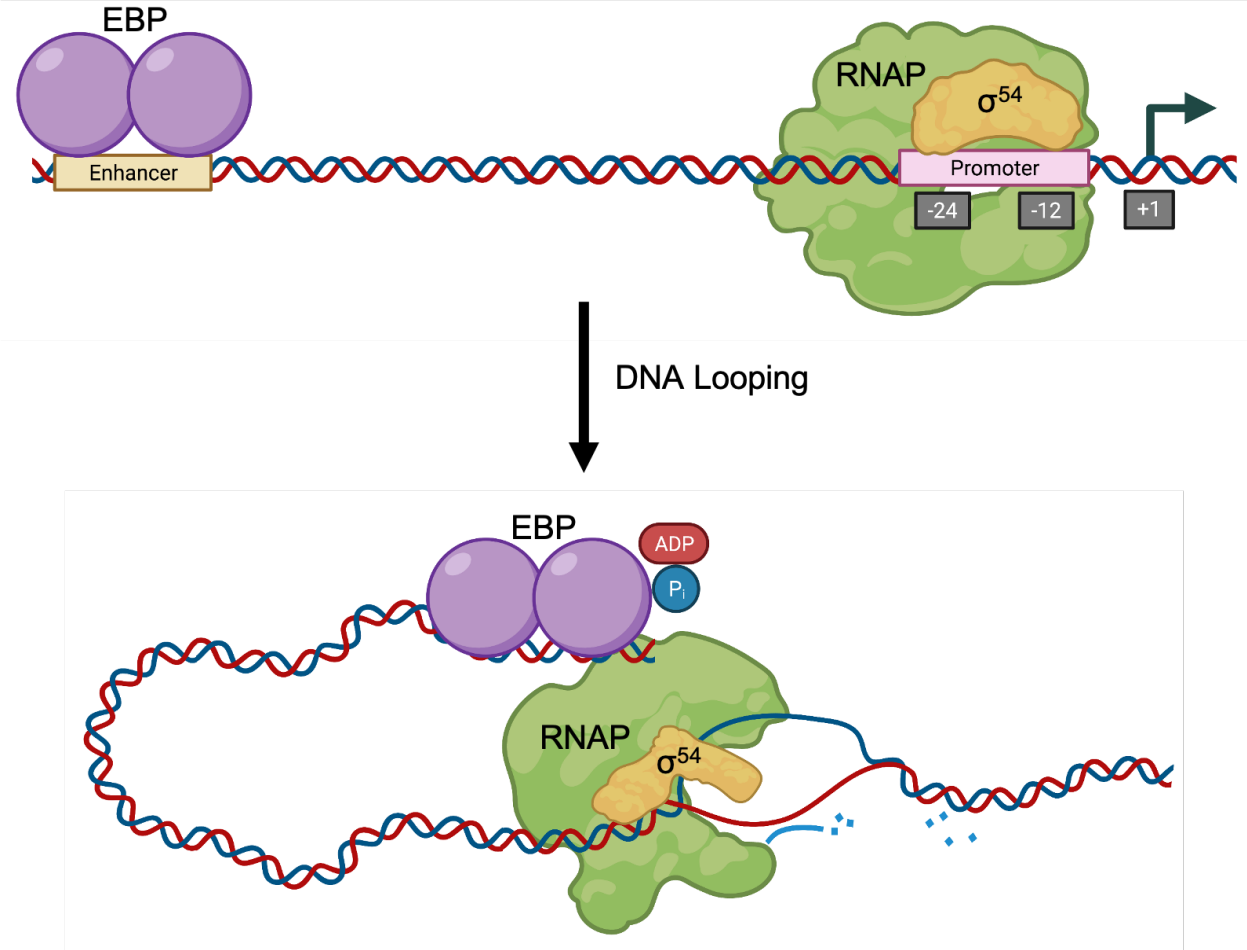


Figure 2.2 Model of EBP-mediated DNA looping and remodeling of closed complex for transcription initiation

EBPs recognize and bind to upstream enhancer sites, loop the DNA between the promoter and enhancer sites to directly interact with σ^{54}-RNAP holoenzyme, and perform ATP hydrolysis to facilitate conformational remodeling in the holoenzyme which convert the closed complex into an open complex for transcription initiation.

Until now, three essential roles that EBPs play in σ^{54}-dependent transcription regulation have been uncovered. First, EBPs must activate DNA melting at the -12 promoter site, as the transcriptions of EBPs knockout mutants initiate only if pre-melted DNA is present [14]. Second, EBPs must conformationally remodel region I and III of σ^{54}, which physically inhibits template DNA entering the active site of holoenzyme, to allow DNA access to the active site and facilitate open complex formation [31]. Third, EBPs must rearrange the DNA binding domains of σ^{54} to downstream DNA, since the -12 promoter site where DNA melting starts is located too far upstream from the active site of holoenzyme [31].

A wide variety of σ^{54} EBPs exist in a large number of bacterial species. The best studied σ^{54} EBPs are nitrogen regulatory protein C (NtrC) and nitrogen fixation-specific activator A (NifA), which stimulate the expression of genes required for nitrogen assimilation and nitrogen fixation, respectively, in a number of free-living and symbiotic bacteria [32]. Some bacteria possess only one EBP. For example, a number of *Chlamydia trachomatis* genes are regulated by chlamydial NtrC-activated σ^{54}-mediated transcription initiation [33]. The functional mannose phosphotransferase system in *Lactobacillus plantarum* is controlled by σ^{54} in concert with the σ^{54}-activator ManR [34].

However, most species have more variants of EBPs. *E. coli* has been reported to have twelve EBPs that regulate genes associated with nitrogen assimilation, carbon source utilization, certain fermentation pathways, flagellar synthesis, stress response and virulence [35]. In *Bacillus subtilis*, BkdR and YplP both regulate the cold shock response, and AcoR, LevR, and RocR are involved in the activation of acetoin, carbohydrate, and amino acid metabolism, respectively [36]. Particularly, in *M. xanthus* genome 53 sequence homologs of σ^{54} EBPs associated with bacterial development have been identified [37].

2.4 σ^{54}-Regulated Bacterial Functions

σ^{54} factor was first described in 1978 in a groundbreaking study on the regulation of nitrogen metabolism. In this study, mutation in a site, *glnF*, on the chromosome of *Klebsiella pneumoniae* induced a type of glutamine auxotrophy caused by insufficient synthesis of glutamine synthetase [38]. Study on *E. coli* in 1985 demonstrated that the gene *glnF*, now known as *rpoN*, encodes a new σ factor specifically required for the transcription of nitrogen-regulated and of nitrogen-fixation promoters. In specific, it revealed that the purified glnF product of *E. coli* binds to core RNA polymerase and these proteins together initiate the transcription at the nitrogen-regulated promoter glnAp2 on a supercoiled template by recognizing consensus sequence C-T-G-G-Y-A-Y-R-N4-T-T-G-C-A [39]. In another study in 1985, the purified new factor was named gpntrA and was proved to bind core RNA polymerase in vitro and to initiate transcription of the *glnA* operon in enteric bacteria for glutamine metabolism during nitrogen limitation as a sigma subunit for RNA polymerase. Also, it firstly confirmed that the transcription of

glnA is dependent on a DNA-binding protein, NtrC product, which binds to five sites in the *glnA* promoter-regulatory region and appears to activate transcription initiation [40].

Over subsequent years, a large number of studies on σ^{54} factor (RpoN) have prominently extended the number and types of σ^{54}-controlled genes, the products of which have diverse physiological roles in bacteria including carbon utilization, fermentation, pathogenesis, stress resistance and so on. For example, in 1987 Ronson reported the identification and cloning of *rpoN* gene of *Rhizobium meliloti* and uncovered that σ^{54}-holoenzyme transcribes *dctA* gene of Rhizobia which encodes a transport component for dicarboxylic acids. Their results suggested that the DctB plus DctD-mediated transcriptional activation of *dctA* in Rhizobia may be mechanistically similar to NtrB plus NtrC-mediated activation of *glnA* in *E. coli* [41]. Dixon's study in 1986 and Inouye's study in 1987 demonstrated that the promoter and regulatory genes in *xyl* operon located on the toluene plasmid of *Pseudomonas putida* activate the transcriptions of *xyl* genes required for the catabolism in the presence of m-toluate, m-xylene or m-methylbenzyl alcohol [42, 43]. In addition, genes encoding two of the components of formate-hydrogen-lyase (formate degradation) pathway in *E. coli* were also identified in 1987. Specifically, the regulatory elements involved in expression of the gene *fdhF* for the selenopolypeptide of formate dehydrogenase and of *hyd* specifically responsible for the formation of the gas-evolving hydrogenase in *E. coli* were investigated [44]. In the meanwhile, genes controlled by σ^{54} that encode hydrogenases responsible for the oxidation of molecular hydrogen in *Alcaligenes eutrophus* and *Pseudomonas facilis* have been identified in a couple of studies [45-47].

In 1987, a set of positively regulated flagellar and cell cycle regulation gene promoters with sequence homology to the σ^{54}-controlled *nif* gene promoters of *K. pneutnoniae* were mapped in *Caulobacter crescentus* [48, 49]. Importantly, these flagellar and cell cycle regulation gene promoters were demonstrated to be recognized and bound by *E. coli* σ^{54}-RNA polymerase holoenzyme in vitro [50]. σ^{54}-controlled genes that encode pilins in *Pseudomonas aeruginosa* were identified in 1989: the promoter region of the *P. aeruginosa* pilin gene has a high degree of similarity to the nitrogen-regulated promoters of enteric bacteria. By showing that σ^{54} bind to *P. aeruginosa* pilin gene promoters in vitro and transcription did not occur in vivo in mutant strains that lack σ^{54}, it confirmed that these promoters are recognized by σ^{54} and the *P. aeruginosa* pilin gene is transcribed by the σ^{54}-containing RNAP [51]. σ^{54}-controlled pilus genes that encode essential hook proteins allowing pathogen to adhere to human epithelial cells have been sequenced and characterized in *Neisseria gonorrhoeae* [52, 53]. The studies of Weiner reported that the phase shock protein (*psp*) genes encode essential proteins in response to heat, ethanol, hyperosmotic shock, infections by filamentous, nutritional deficiency and oxidative phosphorylation in *E. coli.* They also demonstrated that the transcription of *psp* operon is regulated by a network of positive and negative regulatory factors and that transcription in response to all inducing agents is directed by σ^{54} [54, 55]. In 1998, a couple of genes important for normal fruiting body formation and aggregation in *M. xanthus* were firstly verified to be regulated by σ^{54}. In specific, Gorski and Kaiser showed that of seven σ^{54} activator gene-inactivated *M. xanthus* strains, four exhibited defects in the development of fruiting bodies; one failed

to develop in submerged culture; three displayed arrested development of fruiting bodies, each at a morphologically different stage of aggregation [56].

In recent decades, a number of studies have proved that the presence of σ^{54} among different species is also highly relevant to motility, and mutation at *rpoN* locus commonly results in bacterial phenotypes including deficient in biofilm formation, luminescence and colonization. For example, it has been demonstrated that σ^{54} controls flagellar biogenesis, biofilm development, and bioluminescence in *Vibrio fischeri*. Also, σ^{54} plays a requisite role initiating the symbiotic association of *V. fischeri* with juveniles of squid [57]. In 2009, Leang revealed that the *Geobacter sulfurreducens* σ^{54} controls genes involved in a wide range of cellular functions including fumarate/succinate exchange, ammonia assimilation, central metabolism enzymes as well as cytochromes involved in extracellular electron transfer to Fe [58]. In 2011, a comparative genome analysis uncovered the commonality of σ^{54} in Gram-positive and Gram-negative bacteria: 522 of 842 sequenced species encode σ^{54} whose presence was nearly universal in Gram-negative species and was also found in many Gram-positive species belonging to *Bacillus, Clostridia, Enterococcus, Listeria, and Pediococcus* genera. Importantly, all corresponding σ^{54} activators were identified and annotated in this study [59].

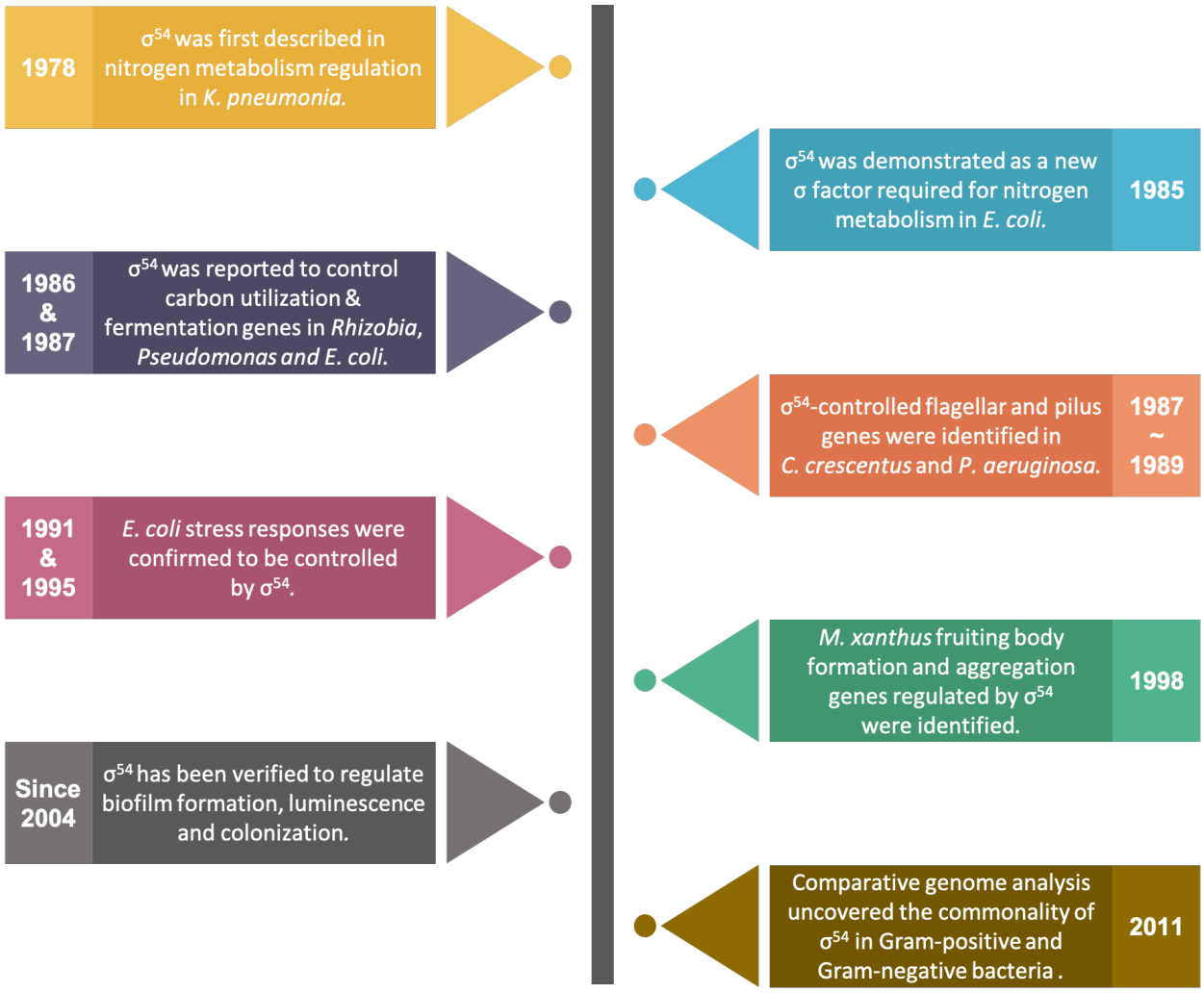


Figure 2.3 Key moments in identifying σ^{54}-regulated bacterial functions

2.5 Genome-scale Search for σ^{54} Regulation

Over recent decades, the availability of complete genome sequences of bacteria offers an opportunity to extend our understanding of EBP-mediated σ^{54}-dependent transcriptional regulation. Many genome sequencing projects have identified the open reading frames (ORFs) encoding σ^{54}, σ^{54} promoters and activators, and new functions have been predicted for this special form of transcription in diverse bacteria by analyzing genome sequences. In 2000, Studholme and Buck identified σ^{54} gene in the

genome sequences of a couple of bacteria including *Aquifex aeolicus*, *B. subtilis*, *Chlamydia spp.* and *Borrelia burgdorferi*, which proved the biological importance of σ^{54} [60]. In particular, six σ^{54}-dependent promoters were found in the extreme thermophile *A. aeolicus* genome and three of them were demonstrated performing sequence-specific σ^{54}-binding [61]. The σ^{54}-RNAP-regulated functions in *A. aeolicus* have been predicted to include sulfur respiration, nitrogen assimilation, nitrate reductase and nitrite reductase activity. The candidate σ^{54}-RNAP-dependent promoter found in *B. burgdorferi*, the causative agent of Lyme disease, has been verified to precede the *rpoS* ORF, encoding the alternative σ factor σ^{38} which is the primary regulator of stationary phase genes in response to general stress [62].

Several genomic comparative studies have been performed on σ^{54} and σ^{54} EBPs-mediated regulation. In 2001, an in-depth comparative analysis was made for *E. coli*, which first demonstrated that nitrogen assimilation was one of the main processes connecting the transcription regulated by σ^{54}: a reasonably complete list of 30 σ^{54}-dependent operons were identified in *E. coli* genome by computer analysis combined with a DNA microarray analysis, and about half are involved in nitrogen assimilation and metabolism. The physiological relationship between these identified σ^{54}-dependent genes was predicted to be based on the fact that nitrogen assimilation consumes energy and intermediates of central metabolism [35]. Another study developed a new algorithm that increased the sensitivity of traditional methods by taking advantage of the well-established requirement of EBPs of σ^{54} promoters, and of the sequence features of the binding sites for σ^{54} factors. They predicted 46 σ^{54} promoters in the *P. putida* genome and linked *P. putida* σ^{54} to carbon metabolism and flagellar biosynthesis [63]. In

2005, Kill constructed a profile hidden Markov model to identify the genes for σ^{54} throughout 240 bacterial genomes based on experimentally verified sequences from UNIPROT database. They found σ^{54} gene are widely distributed among *Firmicutes*, *Proteobacteria* and *Spirochetes*, and the number of σ^{54} genes per genome relates to the environmental variation allowing growth for a given species [64].

In order to identify and annotate true σ^{54} promoters and σ^{54} EBPs, a range of comparative genome analyses was conducted to identify the presence of σ^{54} and its associated EBPs for *Proteobacteria, Fermicutes, Acidobacteria, Nitrospirae, Aquificae, Chlamydiae, Spirohetes, Bacteriodetes* et al. in 2011. Their results uncovered that there is a phylum-dependent distribution that is suggestive of an evolutionary relationship between σ^{54} and lipopolysaccharide and flagellar biosynthesis. Additionally, it was proved that σ^{54} is associated with phosphotransfer-mediated signaling, the transport and assimilation of carboxylates and nitrogen containing metabolites was substantiated. Furthermore, a large number of σ^{54} promoters were localized using a straightforward scoring strategy that was formulated to identify similar motifs. Importantly, these analyses directly implicate σ^{54} as a central player in the control over the processes that involve the physical interaction of an organism with its environment like in the colonization of a host (virulence) or the formation of biofilm [59].

In recent years, several new methods have been developed to predict σ^{54} promoters in bacterial genomes. For example, a sequence-based predictor for identifying σ^{54} promoters in bacteria named iPro54-PseKNC was developed in 2014. In the predictor, the samples of DNA sequences were formulated by a novel feature vector called ‘pseudo k-tuple nucleotide composition’, which was further optimized by the

incremental feature selection procedure for rapidly and effectively identifying σ^{54} promoters. Meanwhile, they also discovered through an in-depth statistical analysis that the distribution of distances between the transcription start sites and the translation initiation sites were governed by the gamma distribution, which may provide a fundamental physical principle for studying the σ^{54} promoters [65]. Another new method reasonably integrates motif finding and machine learning strategies to capture the intrinsic features of σ^{54} promoters. Their experiments on *E. coli* benchmark test set showed that this method has good capability to distinguish σ^{54} promoters from surrounding or randomly selected DNA sequences. Also, the applications of this method on another three bacterial genomes indicate its potential robustness and applicable power on a large number of bacterial genomes [66].

2.6 Summary

The alternative sigma factor σ^{54}, which is employed to mediate gene-specific expression in response to a variety of environmental conditions, has no significant sequence similarity to σ^{70} and instead recognizes -24 and -12 promoter regions relative to transcription start site. Unlike σ^{70}, the σ^{54}-RNAP closed complex is unable to spontaneously isomerize to an open complex but requires ATP dependent activator proteins, known as EBPs, bound remotely upstream from the promoter site to activate transcription. Much progress has been made in the last 40 years in our understanding of molecular mechanisms of the σ^{54}-dependent transcription regulation biochemically and genetically: structural studies on σ^{54}-RNAP holoenzyme, the closed and EBP-bound intermediate complexes, and the open and the initial transcribing complex have

provided a wealth of information on how σ^{54} inhibits transcription and how EBPs act to first relieve this inhibition and then drive DNA opening; the great progress of gene sequencing techniques and computational sequence comparative analysis has accelerated the identifications of σ^{54}-controlled genes, σ^{54} promoters and EBPs in diverse bacteria which help us describe and understand the pattern of functions regulated by this form of activator-dependent transcription.

In any case, the existence of alternative sigma factors such as σ^{54} indicates that evolution might have developed various ways to engineer site-specific transcript initiation, and probably many of them are yet to be discovered. It can be speculated that the existence of σ^{54} factor despite the dominance of the σ^{70} family is due to its near absolute dependence on a EBP activator, and such strict, almost hash, regulation must have conferred the advantage necessary for their survival. One of the major benefits for bacteria to engage σ^{54}-controlled site-specific transcript initiation for any gene or gene cluster is that it sets the stage for the evolution of gene-specific regulation, involving activators. The most important thing that we have learned from studying the functions of different transcriptional factors in bacteria is that how distinct mechanisms evolved in response to different steps of gene transcriptional process, based on the specific sigma factor that is involved.

2.7 References

1. Haugen, S.P., W. Ross, and R.L. Gourse, *Advances in bacterial promoter recognition and its control by factors that do not bind DNA.* Nature Reviews Microbiology, 2008. **6**(7): p. 507-519.
2. Gruber, T.M. and C.A. Gross, *Multiple sigma subunits and the partitioning of bacterial transcription space.* Annual review of microbiology, 2003. **57**: p. 441.
3. Busby, S. and R.H. Ebright, *Transcription activation by catabolite activator protein (CAP).* Journal of molecular biology, 1999. **293**(2): p. 199-213.
4. Ishihama, A., *Functional modulation of Escherichia coli RNA polymerase.* Annual review of microbiology, 2000. **54**: p. 499.
5. Buck, M., et al., *The bacterial enhancer-dependent ς54 (ςN) transcription factor.* Journal of bacteriology, 2000. **182**(15): p. 4129-4136.
6. Zhang, X., et al., *Mechanochemical ATPases and transcriptional activation.* Molecular microbiology, 2002. **45**(4): p. 895-903.
7. Barrios, H., B. Valderrama, and E. Morett, *Compilation and analysis of sigma(54)-dependent promoter sequences.* Nucleic Acids Res, 1999. **27**(22): p. 4305-13.
8. Wedel, A. and S. Kustu, *The bacterial enhancer-binding protein NTRC is a molecular machine: ATP hydrolysis is coupled to transcriptional activation.* Genes & development, 1995. **9**(16): p. 2042-2052.
9. Cramer, P., *Multisubunit RNA polymerases.* Current opinion in structural biology, 2002. **12**(1): p. 89-97.
10. Werner, F. and D. Grohmann, *Evolution of multisubunit RNA polymerases in the three domains of life.* Nature Reviews Microbiology, 2011. **9**(2): p. 85-98.

11. Wigneshweraraj, S.R., et al., *Regulated communication between the upstream face of RNA polymerase and the β' subunit jaw domain.* The EMBO journal, 2004. **23**(21): p. 4264-4274.

12. Yang, Y., et al., *Structures of the RNA polymerase-σ54 reveal new and conserved regulatory strategies.* Science, 2015. **349**(6250): p. 882-885.

13. Arora, S.K., et al., *A transcriptional activator, FleQ, regulates mucin adhesion and flagellar gene expression in Pseudomonas aeruginosa in a cascade manner.* Journal of bacteriology, 1997. **179**(17): p. 5574-5581.

14. Chaney, M., et al., *Binding of transcriptional activators to sigma 54 in the presence of the transition state analog ADP–aluminum fluoride: insights into activator mechanochemical action.* Genes & development, 2001. **15**(17): p. 2282-2294.

15. Zuo, Y. and T.A. Steitz, *Crystal structures of the E. coli transcription initiation complexes with a complete bubble.* Molecular cell, 2015. **58**(3): p. 534-540.

16. Browning, D.F. and S.J. Busby, *The regulation of bacterial transcription initiation.* Nature Reviews Microbiology, 2004. **2**(1): p. 57-65.

17. Studholme, D.J., M. Buck, and T. Nixon, *Identification of potential σN-dependent promoters in bacterial genomes.* Microbiology, 2000. **146**(12): p. 3021-3023.

18. Morett, E. and L. Segovia, *The sigma 54 bacterial enhancer-binding protein family: mechanism of action and phylogenetic relationship of their functional domains.* Journal of bacteriology, 1993. **175**(19): p. 6067-6074.

19. Rappas, M., D. Bose, and X. Zhang, *Bacterial enhancer-binding proteins: unlocking σ54-dependent gene transcription.* Current opinion in structural biology, 2007. **17**(1): p. 110-116.

20. Neuwald, A.F., et al., *AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes.* Genome research, 1999. **9**(1): p. 27-43.

21. Studholme, D.J. and R. Dixon, *Domain architectures of σ54-dependent transcriptional activators.* Journal of bacteriology, 2003. **185**(6): p. 1757-1767.

22. Sallai, L. and P.A. Tucker, *Crystal structure of the central and C-terminal domain of the σ54-activator ZraR.* Journal of structural biology, 2005. **151**(2): p. 160-170.

23. Batchelor, J.D., *Structure and regulatory mechanism of sigma54 activator proteins*. 2009: University of California, Berkeley.

24. D'Autréaux, B., et al., *A non-haem iron centre in the transcription factor NorR senses nitric oxide.* Nature, 2005. **437**(7059): p. 769-772.

25. Tucker, N., et al., *DNA binding properties of the Escherichia coli nitric oxide sensor NorR: towards an understanding of the regulation of flavorubredoxin expression*. 2005, Portland Press Ltd.

26. Su, W., et al., *DNA-looping and enhancer activity: association between DNA-bound NtrC activator and RNA polymerase at the bacterial glnA promoter.* Proceedings of the National Academy of Sciences, 1990. **87**(14): p. 5504-5508.

27. Huo, Y.X., et al., *Protein-induced DNA bending clarifies the architectural organization of the σ54-dependent glnAp2 promoter.* Molecular microbiology, 2006. **59**(1): p. 168-180.

28. Arfin, S.M., et al., *Global Gene Expression Profiling in Escherichia coliK12: THE EFFECTS OF INTEGRATION HOST FACTOR* 210.* Journal of Biological Chemistry, 2000. **275**(38): p. 29672-29684.

29. Jovanovic, G. and P. Model, *PspF and IHF bind co-operatively in the psp promoter-regulatory region of Escherichia coli.* Molecular microbiology, 1997. **25**(3): p. 473-481.

30. Joly, N., et al., *Trapping of a transcription complex using a new nucleotide analogue: AMP aluminium fluoride.* Journal of molecular biology, 2008. **375**(5): p. 1206-1211.

31. Bose, D., et al., *Organization of an activator-bound RNA polymerase holoenzyme.* Molecular cell, 2008. **32**(3): p. 337-346.

32. Merrick, M., *Regulation of nitrogen fixation genes in free-living and symbiotic bacteria.* Biological nitrogen fixation, 1992.

33. Mathews, S.A. and P. Timms, *Identification and mapping of sigma-54 promoters in Chlamydia trachomatis.* Journal of bacteriology, 2000. **182**(21): p. 6239-6242.

34. Stevens, M.J., et al., *σ54-mediated control of the mannose phosphotransferase sytem in Lactobacillus plantarum impacts on carbohydrate metabolism.* Microbiology, 2010. **156**(3): p. 695-707.

35. Reitzer, L. and B.L. Schneider, *Metabolic Context and Possible Physiological Themes of ς54-Dependent Genes in Escherichia coli.* Microbiology and Molecular Biology Reviews, 2001. **65**(3): p. 422-444.

36. Wiegeshoff, F., et al., *Sigma L is important for cold shock adaptation of Bacillus subtilis.* Journal of Bacteriology, 2006. **188**(8): p. 3130-3133.

37. Jakobsen, J.S., et al., *σ54 enhancer binding proteins and Myxococcus xanthus fruiting body development.* Journal of bacteriology, 2004. **186**(13): p. 4361-4368.

38. Gaillardin, C.M. and B. Magasanik, *Involvement of the product of the glnF gene in the autogenous regulation of glutamine synthetase formation in Klebsiella aerogenes.* Journal of Bacteriology, 1978. **133**(3): p. 1329-1338.

39. Hunt, T.P. and B. Magasanik, *Transcription of glnA by purified Escherichia coli components: core RNA polymerase and the products of glnF, glnG, and glnL.* Proceedings of the National Academy of Sciences, 1985. **82**(24): p. 8453-8457.

40. Hirschman, J., et al., *Products of nitrogen regulatory genes ntrA and ntrC of enteric bacteria activate glnA transcription in vitro: evidence that the ntrA product is a sigma factor.* Proceedings of the National Academy of Sciences, 1985. **82**(22): p. 7525-7529.

41. Ronson, C.W., et al., *Deduced products of C4-dicarboxylate transport regulatory genes of Rhizobium leguminosarum are homologous to nitrogen regulatory gene products.* Nucleic acids research, 1987. **15**(19): p. 7921-7934.

42. Dixon, R., *The xylABC promoter from the Pseudomonas putida TOL plamid is activated by nitrogen regulatory genes in Escherichia coli.* Molecular and General Genetics MGG, 1986. **203**(1): p. 129-136.

43. Inouye, S., A. Nakazawa, and T. Nakazawa, *Expression of the regulatory gene xylS on the TOL plasmid is positively controlled by the xylR gene product.* Proceedings of the National Academy of Sciences, 1987. **84**(15): p. 5182-5186.

44. Birkmann, A., et al., *Factors affecting transcriptional regulation of the formate-hydrogen-lyase pathway of Escherichia coli.* Archives of microbiology, 1987. **148**(1): p. 44-51.

45. Römermann, D., et al., *Pleotropic mutants from Alcaligenes eutrophus defective in the metabolism of hydrogen, nitrate, urea, and fumarate.* Archives of microbiology, 1988. **149**(5): p. 471-475.

46. Römermann, D., et al., *An rpoN-like gene of Alcaligenes eutrophus and Pseudomonas facilis controls expression of diverse metabolic pathways, including hydrogen oxidation.* Journal of bacteriology, 1989. **171**(2): p. 1093-1099.

47. WARRELMANN, J. and B. FRIEDRICH, *Mutants of Pseudomonas facilis defective in lithoautotrophy.* Microbiology, 1986. **132**(1): p. 91-96.

48. Minnich, S.A. and A. Newton, *Promoter mapping and cell cycle regulation of flagellin gene transcription in Caulobacter crescentus.* Proceedings of the National Academy of Sciences, 1987. **84**(5): p. 1142-1146.

49. Mullin, D., et al., *A set of positively regulated flagellar gene promoters in Caulobacter crescentus with sequence homology to the nif gene promoters of Klebsiella pneumoniae.* Journal of molecular biology, 1987. **195**(4): p. 939-943.

50. Ninfa, A., et al., *Escherichia coli sigma 54 RNA polymerase recognizes Caulobacter crescentus flbG and flaN flagellar gene promoters in vitro.* Journal of bacteriology, 1989. **171**(1): p. 383-391.

51. Ishimoto, K.S. and S. Lory, *Formation of pilin in Pseudomonas aeruginosa requires the alternative sigma factor (RpoN) of RNA polymerase.* Proceedings of the National Academy of Sciences, 1989. **86**(6): p. 1954-1957.

52. Meyer, T.F., et al., *Pilus genes of Neisseria gonorrheae: chromosomal organization and DNA sequence.* Proceedings of the National Academy of Sciences, 1984. **81**(19): p. 6110-6114.

53. Watt, P.J. and M.E. Ward, *Adherence of Neisseria gonorrhoeae and other Neisseria species to mammalian cells*, in *Bacterial adherence*. 1980, Springer. p. 251-288.

54. Weiner, L., J. Brissette, and P. Model, *Stress-induced expression of the Escherichia coli phage shock protein operon is dependent on sigma 54 and modulated by positive and negative feedback mechanisms.* Genes & development, 1991. **5**(10): p. 1912-1923.

55. Weiner, L., et al., *Analysis of the proteins and cis-acting elements regulating the stress-induced phage shock protein operon.* Nucleic acids research, 1995. **23**(11): p. 2030-2036.

56. Gorski, L. and D. Kaiser, *Targeted mutagenesis of ς54 activator proteins in Myxococcus xanthus.* Journal of Bacteriology, 1998. **180**(22): p. 5896-5905.

57. Wolfe, A.J., et al., *Vibrio fischeri σ54 controls motility, biofilm formation, luminescence, and colonization.* Applied and Environmental Microbiology, 2004. **70**(4): p. 2520-2524.

58. Leang, C., et al., *Genome-wide analysis of the RpoN regulon in Geobacter sulfurreducens.* BMC genomics, 2009. **10**(1): p. 1-19.

59. Francke, C., et al., *Comparative analyses imply that the enigmatic Sigma factor 54 is a central controller of the bacterial exterior.* BMC genomics, 2011. **12**(1): p. 1-21.

60. Studholme, D.J. and M. Buck, *The biology of enhancer-dependent transcriptional regulation in bacteria: insights from genome sequences.* FEMS microbiology letters, 2000. **186**(1): p. 1-9.

61. Studholme, D.J., et al., *Functionality of Purified ςN (ς54) and a NifA-Like Protein from the Hyperthermophile Aquifex aeolicus.* Journal of bacteriology, 2000. **182**(6): p. 1616-1623.

62. Studholme, D.J. and M. Buck, *Novel roles of σN in small genomes.* Microbiology (Reading), 2000. **146**(1): p. 4-5.

63. Cases, I., D.W. Ussery, and V. De Lorenzo, *The σ54 regulon (sigmulon) of Pseudomonas putida.* Environmental microbiology, 2003. **5**(12): p. 1281-1293.

64. Kill, K., et al., *Genome update: sigma factors in 240 bacterial genomes.* MICROBIOLOGY-READING-, 2005. **151**(10): p. 3147.

65. Lin, H., et al., *iPro54-PseKNC: a sequence-based predictor for identifying sigma-54 promoters in prokaryote with pseudo k-tuple nucleotide composition.* Nucleic acids research, 2014. **42**(21): p. 12961-12972.

66. Liu, B., et al., *Computational prediction of sigma-54 promoters in bacterial genomes by integrating motif finding and machine learning strategies.* IEEE/ACM transactions on computational biology and bioinformatics, 2018. **16**(4): p. 1211-1218.

Chapter 3

Myxococcus xanthus Background and Review

3.1 Introduction

Myxobacteria are social bacteria that manifest a complex and fascinating multicellular life cycle. *Myxococcus xanthus*, is a non-pathogenic gram-negative Myxobacterium. It grows in top layer of soil and can be easily cultured under standard laboratory conditions. *M. xanthus* cells collectively and dramatically adjust their morphology, physiological functions, and biosynthesis in response to changes in nutrient availability and environmental conditions. A large number of studies have validated that multicellular behaviors of *M. xanthus* require complicated and tight regulations of gene expression at individual cell level. Based on recent findings on the mechanisms of *M. xanthus* development, a couple of signal-responsive σ^{54}-dependent gene regulatory networks governing aggregation initiation and motility have been uncovered. In addition, myxobacteria are known for their extraordinary capability to produce novel natural products which are promising sources of new antibiotics. Although little is known about the regulatory mechanisms of bacterial natural product synthesis, a couple of genomic analyses and molecular experiments have indicated that σ^{54} may play an important role in regulating natural product gene expression in *M. xanthus*. Taken together, *M. xanthus* has become a powerful model to elucidate how individual bacterial cells cooperate their multicellular behaviors and secondary metabolism for survival under severe conditions. In this chapter, we review the multicellular development and natural product synthesis in *M. xanthus*.

3.2 Multicellular Lifestyle of *M. xanthus*

M. xanthus employs complex and varied multicellular behaviors throughout its life cycle involving collective movement and predation in a pattern of swarming, as well as the formation of complex multicellular biofilms, where cells differentiate into dormant spores under starvation condition. This feature has made *M. xanthus* an ideal model for exploring the molecular and genetic mechanisms underlying the formation of bacterial multicellular communities. In particular, the development of many tools for the genetic manipulation of *M. xanthus* provide strong support for investigating the regulations of bacterial multicellular behaviors [1].

When preys or other nutrient sources are sufficient, *M. xanthus* undergoes vegetative growth (See **Fig. 3.1**). During growth, multicellular swarms of *M. xanthus* conduct strictly coordinated 'wolf-pack' predation [2] to obtain nutrient by killing and degrading prey microorganisms or other sources with costly extracellular digestive enzymes or toxins [3]. This coordinated and cooperative characteristic of predation was first evidenced by a study showing that the growth of *M. xanthus* on insoluble protein sources, such as casein, is density dependent [4]. However, when *M. xanthus* confronts severe conditions with low or depleted nutrient, starving cells undergo the process of multicellular development [4, 5]. During development, originally dispersed *M. xanthus* cells begin to aggregate toward common centers and organize into mound-like intermediate structure [6]. Eventually, cells aggregate into groups of around 100 thousand individuals and construct elevated three-dimensional structures, known as fruiting bodies, in which stress-resistant spores are formed and are capable of dispersing to unexploited resource patches [7, 8]. During the developmental process, *M.*

xanthus cells divide into three subpopulations: only ~10% of cells finally differentiate into spores in the fruiting body. These spores are highly resistant to harsh environmental conditions, and are able to regerminate to a new population when nutrients become available thus protecting the swarm's future from temporary nutrient depletion [9]; around 30% of the cells differentiate into peripheral rods, a cell type distinct from either vegetatively growing cells or spores. Peripheral rods remain outside aggregates and surround the fruiting body keeping in a persister-like, or dormant state due to lack of nutrient [9]; the rest of population, approximately 60% of cells, undergo programmed cell death, during which cell lysis releases and provides nutrient to other cell types. The occurrence of this process spans from slightly before aggregation process to sporulation [10]. Although some of the signal-responsive gene regulatory network governing *M. xanthus* development process have been uncovered in a couple of studies (reviewed in [11]), how starving cells are partitioned into subpopulations is yet known.

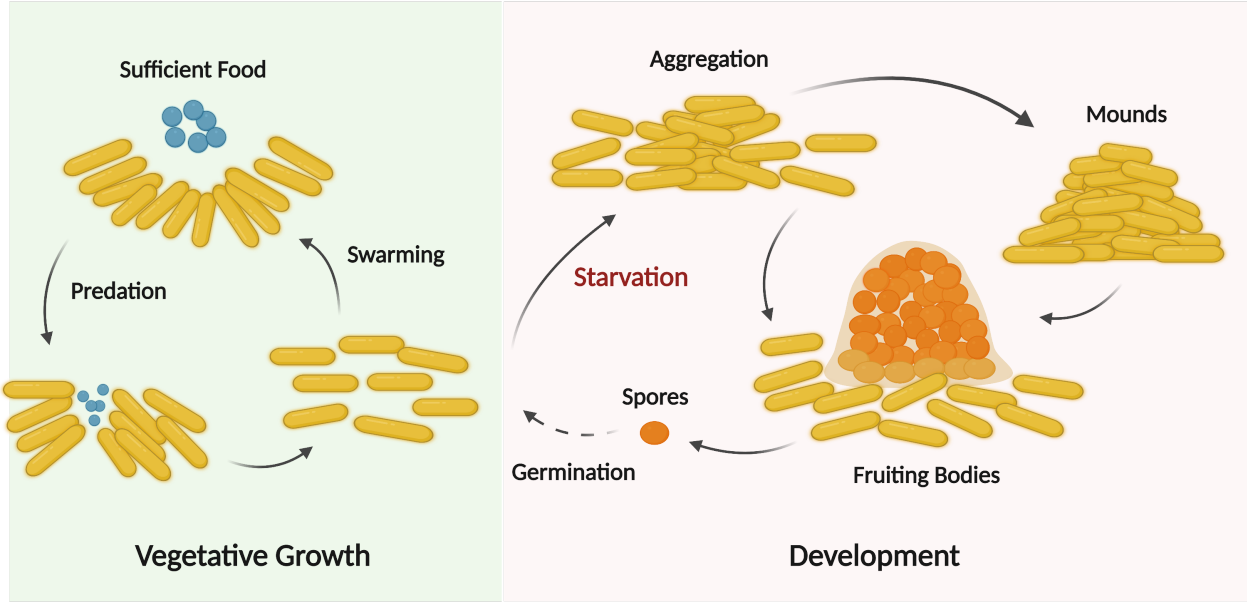


Figure 3.1 *Myxococcus xanthus* multicellular life cycle

(Left) When food sources are sufficient, *M. xanthus* conducts predation in swarms and undergoes vegetative growth; (Right) when nutrients are limited, *M. xanthus* undergoes multicellular development during which starving cells begin to aggregate toward common centers and organize into mound-like intermediate structure, and eventually construct elevated three-dimensional fruiting bodies filled with stress-resistant spores. When nutrients become available, *M. xanthus* spores can regerminate to normal cells.

The very beginning of multicellular development in *M. xanthus* is regulated by the alarmone Guanine-5'(tri)di-phosphate-3'diphosphate, (p)ppGpp, serving as an intracellular starvation signal in individual cells [12]. When nutrients like carbon sources, phosphates or amino acids are limited, stringent response of *M. xanthus* is induced during which (p)ppGpp starts production and accumulation. (p)ppGpp functions as a unique transcriptional regulator that binds directly to the β, β' subunit of the RNA polymerase to downregulate the transcription of genes required for growth while upregulate the genes for response to starvation [13]. It has been verified that production of (p)ppGpp is necessary and sufficient to initiate the expression of genes essential for the earliest stage of development in *M. xanthus* [12].

In the meantime, the initiation of multicellular development in *M. xanthus* is also regulated by quorum sensing (QS). QS is a communication mechanism between bacteria that allows specific processes to be controlled, such as biofilm formation, production of natural products and stress responses. It involves the production and secretion of small molecules or peptides, named autoinducers, that are diffusible and can be shared by all members of a bacterial population [14]. Every individual in bacterial population produces and secretes autoinducers. Therefore, upon detecting sufficiently

high concentration of autoinducers around indicating high population density, bacteria may respond by regulating the expression of their genes [15]. Under nutrient-limited condition, the rapid accumulation of (p)ppGpp in *M. xanthus* results in the expression of the QS autoinducer, A-signal, consisting of a mixture of peptides, amino acids, and proteases [16]. The produced A-signal is subsequently exported from cells and accumulated to a sufficiently high density, causing cells begin to aggregate [17]. Another autoinducer, known as C-signal encoded by the gene *csgA*, drives the aggregation process of *M. xanthus* [18]. As *M. xanthus* cells begin to aggregate, the C-signal originated from one cell directly interacts with a second adjacent cell, which promotes the production of C-signal in the second cell. Importantly, once the C-signal on the second cell is presented to the first cell, additional C-signal will be produced in the first cell. This set of interactions serve as a positive feedback loop that rapidly increases the amount of C-signal in *M. xanthus* population and promotes aggregation [19]. In the meanwhile, the accumulated C-signal also promotes the production of FruA protein which is initially produced in response to A-signal accumulation [20]. Phosphorylated FruA sequentially activates a series of downstream targets promoting further aggregation and eventual sporulation [21, 22]. It has been demonstrated that C-signal functions in a concentration dependent manner: intermediate concentrations of C-signal induce aggregation, whereas high concentration of C-signal triggers sporulation [23].

Motility is essential for the multicellular life style of *M. xanthus* both during vegetative growth and development, including swarming, predation, fruiting body formation, and sporulation [8, 24]. In specific, without the ability to move on the surface

of environment, *M. xanthus* would be unable to prey on other microorganisms and food sources by swarming. Also, motility allows *M. xanthus* individuals to aggregate towards centers to form multicellular fruiting bodies and sporulate. Motility mutants in *M. xanthus* were first studied by Burchard in 1970 after he observed mutants that were nonmotile as individual cells but became transiently motile when in apposition to other cells [25]. Several years later, an extensive mutagenesis study isolated a large number of motility mutants and revealed that motility of *M. xanthus* on solid surfaces is driven by two different types of motility systems: the social motility (S motility) and the adventurous motility (A motility), respectively controlled by two sets of genes [26]. *M. xanthus* employs S motility system, powered by the extension and retraction of type IV pili (T4P), to coordinately move in swarms (See **Fig. 3.2**) [27-29]. *M. xanthus* S motility system dominates on softer, wetter substrates [30]. It functions by extending a T4P from the leading end of a cell, tethering it to another cell that it encounters, and then retracting that cell toward the leading cell [31, 32]. PilA protein is the major component of the *M. xanthus* T4P [33], and the extension and retraction of the T4P are mediated by the cytoplasmic ATPase PilB and its homologue PilT, respectively [34]. The production of extracellular polysaccharide (EPS) tracks across the surface where *M. xanthus* cells move is also required for S motility (**Fig. 3.2**). These EPS tracks are thought to form channels, funneling cells into large streams, and tightening cell-cell interactions. It is also assumed that the T4P recognizes EPS as a receptor allowing retraction of pilus to pull the cell forward [35]. Intriguingly, T4P activates EPS secretion via Dif signaling pathway, which has typical features of bacterial chemotaxis pathways allowing cells to move towards chemical sources, such as nutrient gradients [36-38].

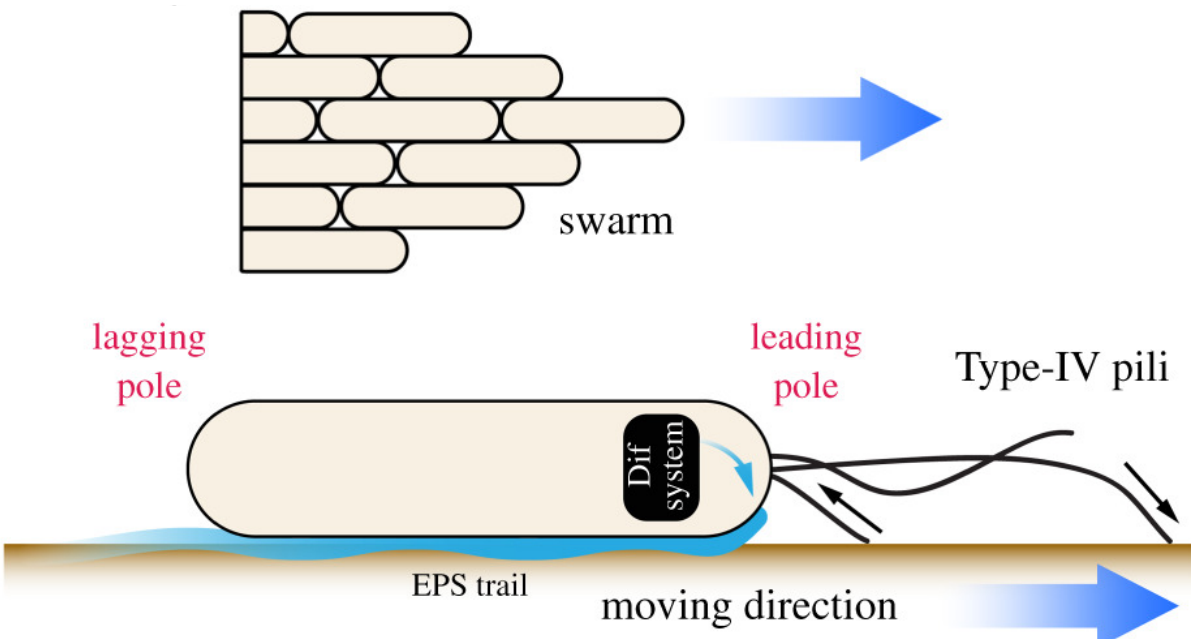


Figure 3.2 Model of S motility system powered by type IV pili

The S motility system promotes coordinated group motility of *M. xanthus* and involves the extension and retraction of a Type-IV pili at the leading pole that pulls cells forward. EPS secretion is activated by the Dif pathway, forming tracks across surfaces that are proposed to channel cells into larger streams.

Distinct from S motility system functioning in multiple cells, A motility allows *M. xanthus* to move across surfaces individually and typically dominates on stiffer, drier substrates [30]. A number of genes and encoded proteins essential for A motility have been identified, nevertheless, what drives A motility in *M. xanthus* is still under debate [24]. To date, there have been three models assumed regulating A motility: slime secretion [39], helical tracks [40], and the best proven model so far, focal adhesion complex (See **Fig. 3.3**) [41]. According to focal adhesion complex model, a couple of Agl and Glt motor proteins that connect to cytoskeleton and envelope-spanning focal adhesion complexes, assemble at the leading cell pole, and move toward the lagging

cell pole. This movement drives the cells forward by pushing against the focal adhesion complexes. Thus, the distributed focal adhesion complexes appear stationary as cells move forward.

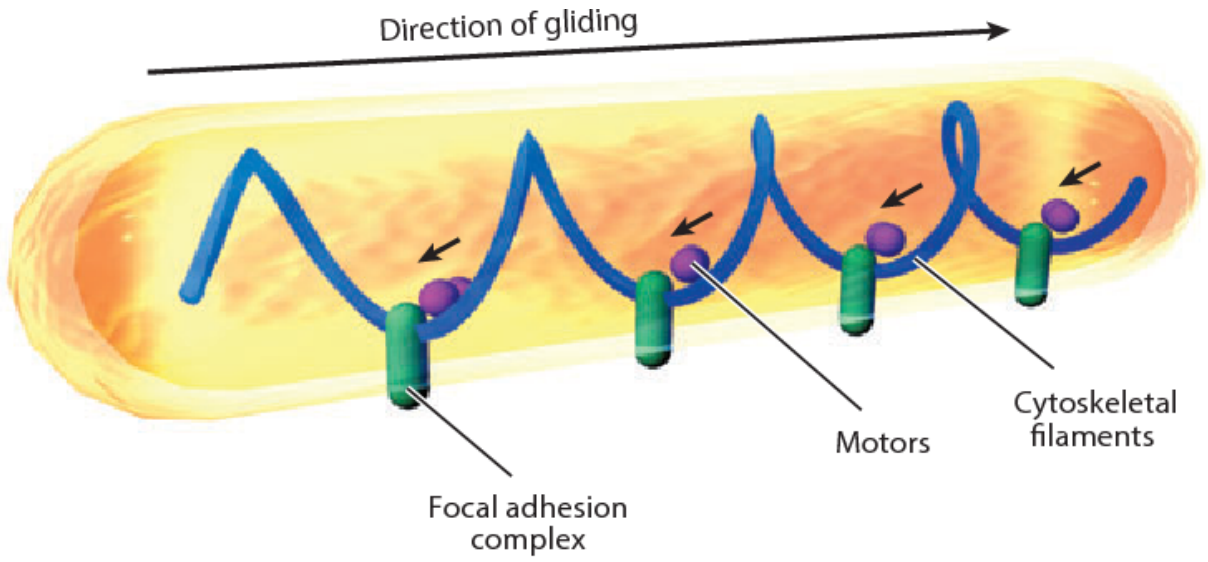


Figure 3.3 Model of A motility system powered by the focal adhesion complex
The A motility system promotes individual motility of *M. xanthus*. In this model, large focal adhesion complexes penetrate the cell envelope, stick to the substratum at one end, and connect to cytoskeletal filaments at the other end. Motor proteins push backward (marked by small arrows) against those focal adhesion complexes, pushing the cells forward.

3.3 σ^{54} in *M. xanthus*

M. xanthus requires tight coordination of gene regulations to accomplish its complicated multicellular life cycle, which is achieved by the vast repertoire of signal transduction networks harbored in the *M. xanthus* genome [42]. Despite the diversity of signals associated with multicellular behaviors that are detected by and transduced within *M. xanthus* cells, one of its major signal transduction networks belong to two component signal transduction systems (TCS). Typically, TCS consists of a sensor

histidine kinase protein which detects an environmental signal and passes phosphates, and a response regulator protein which receives the phosphates from the sensor protein and modulates cellular activities according to environmental changes. TCSs are crucial for signal transduction and gene regulations *of M. xanthus* both during vegetative growth and during development, though it is the one-component systems that dominate the signal transduction networks in *M. xanthus*, and in prokaryotic genomes generally [42, 43].

As described in the chapter of σ^{54}, unlike σ^{70}, the σ^{54}-RNAP holoenzyme-protomer structure requires energy generated from ATP hydrolysis catalyzed by activator proteins, known as EBPs, to isomerize into a transcriptional competent open complex for initiating transcription [44]. σ^{54}-dependent transcription is regulated by TCS signal transduction networks in which EBPs detect intra and extracellular signals and regulate the activity of σ^{54} promoters in response to environmental cues [45]. The best characterized EBPs are response regulators belonging to the NtrB/NtrC TCS family that target genes with σ^{54} promoters. The NtrC-like EBPs have an N-terminal regulatory domain to receive signal, a central AAA+ domain which provides energy from ATP hydrolysis to closed complex, and a C-terminal HTH domain which mediates DNA binding [46]. Upon being activated by their cognate histidine kinases, NtrC-like EBPs bind to their target DNA sequence known as an enhancer, which is typically found upstream of the σ^{54} promoter region. After that, they directly interact with σ^{54}-RNAP holoenzyme, and perform ATP hydrolysis to facilitate conformational remodeling in the closed complex for transcription initiation.

The larger number of genes encoding σ^{54} activators in the genomes of δ proteobacteria than in other classes of bacteria indicates that these bacteria have a greater capacity to regulate gene expression by σ^{54} in response to environmental changes [45]. Specifically, the importance of σ^{54}-dependant gene regulation in *M. xanthus* has been further demonstrated by the high number of σ^{54} activator genes encoded in its genome [42]. σ^{54} is not only vital for the growth of *M. xanthus* [47], though it is not essential for other bacteria, it also plays important role in regulating developmental gene expression in *M. xanthus* in response to starvation. Many of the σ^{54} activators have been evidenced to regulate *M. xanthus* development. For example, the NtrC-like activator Nla29 encoded by gene MXAN1077 is an important sensor protein that inhibits development under nutrient-sufficient condition. Mutations in MXAN1077, also known as *spdR*, lead to high level of expression of developmental genes in vegetative cells when nutrient is sufficient, and finally to fruiting body formation and sporulation [48, 49].

Initiation of *M. xanthus* development in response to starvation is regulated by accumulation of (p)ppGpp and stringent response [50]. The *M. xanthus* strain containing an *nla4* mutation failed to produce (p)ppGpp under starvation indicating that the stringent response in *M. xanthus* is regulated by the σ^{54} activator, Nla4 [51]. Expression analysis has found out that mutation in *nla4* drastically downregulate the expression of the synthase gene of (p)ppGpp, relA, in *M. xanthus*. And because a putative σ^{54} promoter has been detected in the upstream region of the *relA* gene and Nla4 indeed binds to the promoter DNA fragment of the *relA* gene in vitro, it has been verified that the Nla4 EBP directly regulates the expression of *relA* associated with the transition

from vegetative growth to development in *M. xanthus* under starvation [51]. In addition to Nla4, another σ^{54} EBP MXAN355 appears to also regulate the initiation of *M. xanthus* development in response to starvation. Mutation analysis has demonstrated that compared to wild type strain, an MXAN3555 gene mutant aggregates at least 48 hours later and its sporulation efficacy decreases more than 90% under stringent starvation conditions, although it forms fruiting bodies and sporulates efficiently [52, 53]. Intriguingly, the development of MXAN3555 mutant processes normally under the condition that all nutrients are absent, while fails to process when limited amounts of nutrients are present, indicating that MXAN3555 EBP regulates development of *M. xanthus* in response to nutrient-depletion rather than nutrient-absence.

In early development, the quorum sensing A-signal in *M. xanthus* helps to assess starvation and induce preaggregation. It has been evidenced that the σ^{54} EBP Nla28 is involved in regulating the production of A-signal [17]. Specifically, A-signal mutant *M. xanthus* displays defect in development which cannot be rescued by mutation in *nla28* gene [54]. Also, Nla28 regulates *asgA*, of which the expression is essential for A-signal production. Furthermore, Nla28 appears to regulate the detection of nutrient availability and cell density as a nutrient sensor [55]. The QS C-signal helps to pattern cell movement and shape the fruiting body during development, the production of which is also verified to be regulated by a σ^{54} activator, ActB. An insertion mutation in *actB* gene produces significantly less C-signal than wild type *M. xanthus*, though it produces wild type levels of A-signal [19, 56]. In addition, mutation in *actB* shows phenotypes of wild type-like aggregation under starvation but is not able to produce viable spores [57].

Taken together, EBP Nla28-dependent σ^{54} regulation involves in the production of A-signal and C-signal associated with QS in response to starvation.

EBP-mediated gene regulation by σ^{54} also plays important roles in A and S motility of *M. xanthus*, which are essential for normal functions both during vegetative growth and development under starvation. A motility of *M. xanthus* has been demonstrated to be regulated by a σ^{54} EBP, Nla24, since the strain with *nla24* insertion mutation shows severe defect in gliding motility [58]. Although the exact role of Nla24 in A motility is yet known to date, gene expression analysis on *nla24* mutant using qPCR has reported that two important genes for A motility, *aglU* and *cglB* [59], are significantly downregulated in this mutant suggesting that they are regulated by Nla24 [58]. As mentioned before, S motility of *M. xanthus* is powered by polar type IV pili [60]. It has been proved that σ^{54} EBP PilR regulates the expression of *pilA* gene encoding the major component of type IV pilus, the PilA protein, as mutation in *pilR* severely downregulates the expression of *pilA* [61]. Biosynthesis of EPS which is essential for *M. xanthus* S motility is regulated by Nla24 as well. *M. xanthus* strains containing a *nla24* mutation are defective in S motility because they fail to produce EPS, and hence are S motility deficient [58, 62].

Importantly, a cascade of σ^{54} EBPs has been uncovered that regulate expression of genes that play important role in processing the multicellular development of *M. xanthus* [63]. Six coregulating EBPs, Nla18, Nla4, Nla6, Nla28, ActB and MXAN4899, are activated chronologically during the early stages of development including transition from vegetative growth, preaggregation and aggregation. The results of expression studies, in vivo promoter mutational analyses, and in vitro promoter binding assays

indicate that this regulatory cascade of transcription is propagated by each EBP directly activating the expression of the EBP required for the next stage of the developmental cycle. Specifically, the cascade is initiated by the Nla18 and Nla4 EBPs, both of which are important for growth and development of *M. xanthus*. Insertion mutation in *nla18* displays the phenotypes of slower growth, delay in aggregation and no sporulation, while *nla4* insertion mutants delay in growth and aggregation as well and show significant reduction in sporulation than wild type *M. xanthus* [51, 54, 64]. Upon starvation, Nla4 and Nla18 regulate the transition from growth to development after detecting the external signal and presumably interacting with a serine/threonine kinase. Both Nla4 and Nla18 EBPs activate the expression of the preaggregation stage EBP gene *nla6* encoding EBP Nla6. The expression of Nla6 is activated around 2 hours post starvation and it sequentially activates the expression of the nla28 EBP. Nla6 and Nla28 belong to TCSs that regulate a wide range of cellular processes during development including fruiting body formation, and they both begin functioning in the preaggregation stage by regulating the activation of EBP ActB. ActB is activated about 6 hours post starvation during the aggregation stage of development, and activates the terminal EBP of the cascade MXAN4899 [63].

3.4 Natural Product in Myxobacteria

Natural products originating as secondary metabolites from a myriad of organisms, including plants, animals, and microorganisms, have been an important source of new antibiotics for decades [65]. Myxobacteria are known for their extraordinary capability to produce novel classes of structurally diverse secondary

metabolites. Particularly, as a single strain usually produces several different compound families, myxobacteria are regarded as outstanding producers of natural products and an important source for the development of new lead structures for drug discovery processes [66, 67]. For example, epothilone, a drug derived from a myxobacterial natural product, has been used in clinic as antitumor drugs which acts via microtubule stabilization. Importantly, in comparison to a number of alternative drugs with a similar mode of bioaction, epothilones have significant advantages, like very highest activity in the nanomolar range and lowest susceptibility towards multidrug resistance [68]. Furthermore, many other natural products have been characterized from myxobacteria, especially from *Sorangium*, including sorangicins, potent eubacterial RNA polymerase inhibitors, bactericidal sorangiolides, the antifungal chivosazoles, and the sulfangolides [69-72].

Natural products have evolved to offer myxobacteria increased fitness over other competing organisms in their natural habitats by exhibiting various biological activities, mostly against fungi (54%) and bacteria (29%) [73]. Most of the early-identified antifungal natural products in myxobacteria (See **Fig. 3.4**), like stigmatellin, inhibit the electron flow within the mitochondrial respiratory chain in a unique way compared to the natural products isolated from other bacterial sources [74]. However, a good number of myxobacterial natural products possess antifungal activity via other specific mechanisms (**Fig. 3.4**). For example, soraphen functions by inhibiting acetyl CoA carboxylase activity in eukaryotic cells [75]; leupyrrin shows good biological activity against various fungi and eukaryotic cells by interrupting DNA replication and

transcription [76]; jerangolid and ambruticin kill yeast with high efficacy through a transient but substantial disruption of the integrity of the cell membrane [77, 78].

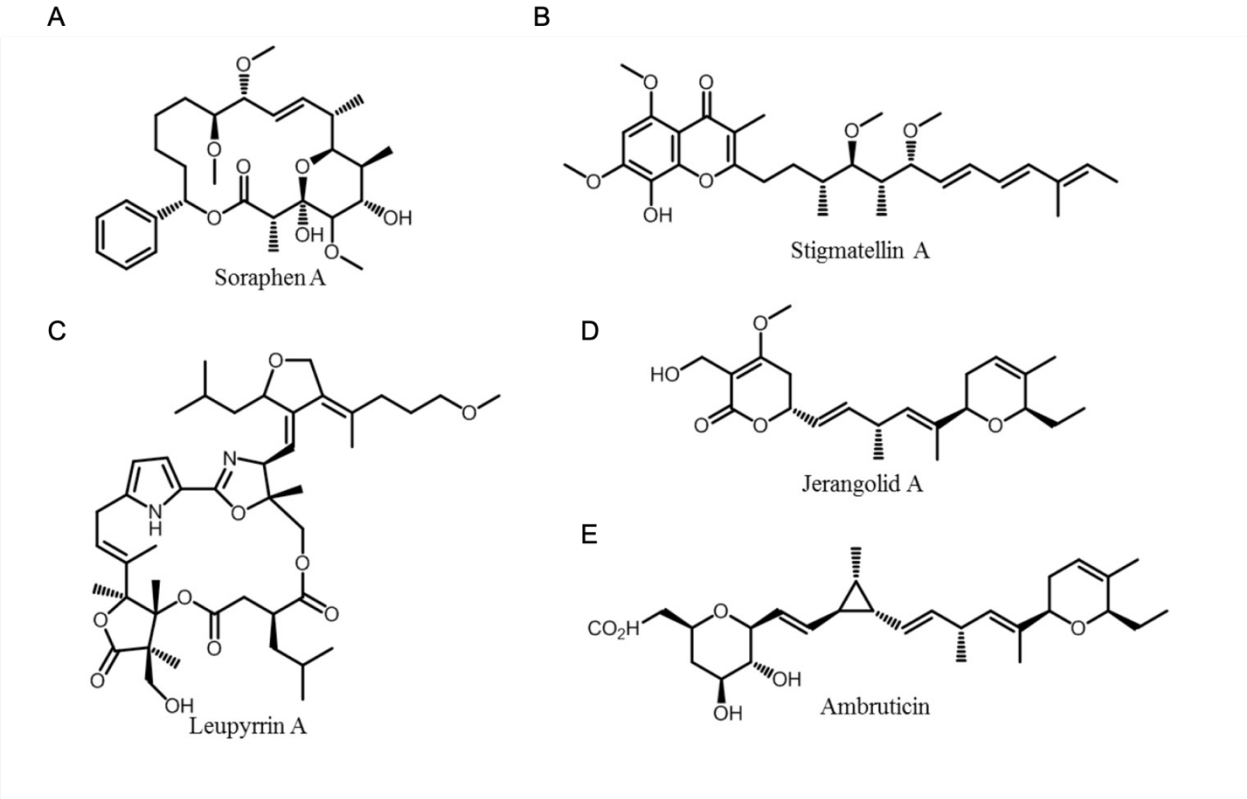


Figure 3.4 Myxobacterial natural products with antifungal activities
(A, C-E) Soraphen A, Jerangolid A, Leupyrrin A and Ambruticin from *Sorangium cellulosum*; (B) Stigmatellin A from *Stigmatella aurantiaca*.

As for antibacterial natural products (See **Fig. 3.5**), corallopyronin A is known for its potent biological activity of which the mode of action involves the selective inhibition of bacterial DNA-dependent RNA polymerase [79]; Myxovirescins function as bactericidal by interfering cell wall synthesis in many Gram-negative and Gram-positive bacteria [80]; Thuggacins act on the electron transport system of *Mycobacterium tuberculosis* making this compound an interesting potential candidate for anti-

tuberculosis therapy [81]; Myxobacteria-derived cystobactamids are new topoisomerase inhibitors with activity against Gram-negative pathogens inhibiting gyrase [82].

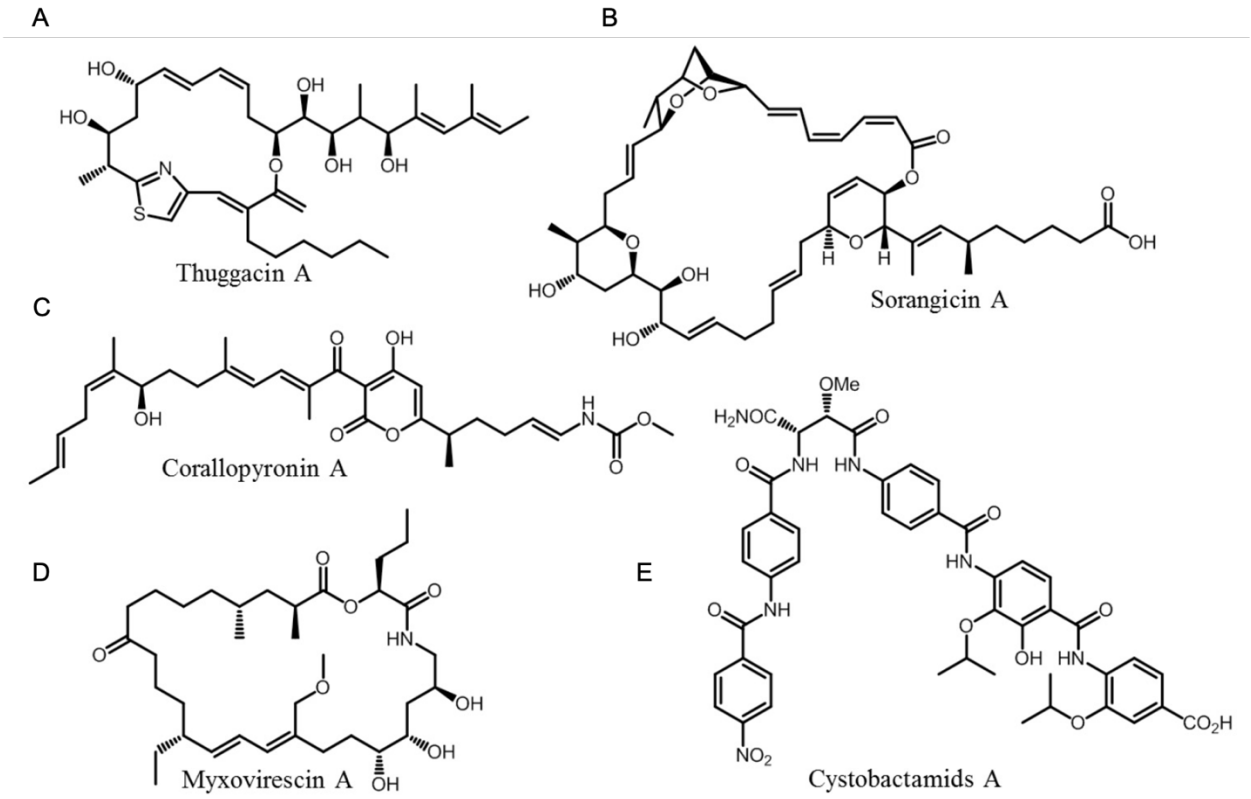


Figure 3.5 Myxobacterial natural products with antibacterial activity
(A and B) Thuggacin A and Sorangicin A from *Sorangium cellulosum*; (C) Corallopyronin A from *Corallococcus coralloides*, (D) Myxovirescin A from *Myxococcus virescens*, (E) Cystobactamids A from *Cystobacter* sp.

Myxobacteria harbor the largest known genomes among bacteria and contain numerous natural product biosynthetic genes which are usually organized into clusters within the genome [83]. Importantly, these biosynthetic genes are usually organized in clusters, which are located on genomic regions spanned from 20 to 200 kb in size [84]. In recent decades, a lot of researchers have focused on identifying and characterizing

natural product gene clusters from a variety of myxobacterial strains and that has contributed to a deeper understanding of the complex biosynthetic machinery in natural products production. For example, scanning of the *M. xanthus* DK1622 genome sequence for the presence of natural product encoding genes revealed the presence of at least 18 biosynthetic gene clusters [72]. Similar to *M. xanthus*, a large number of natural product biosynthetic gene loci are encoded in the genome, analysis of the 10.2 Mbp genome of *Stigmatella aurantiaca* DW4/3-1 revealed the presence of over 8000 genes, a significant number of which are involved in natural products synthesis [85]. After *in silico* identification of biosynthetic gene clusters and pathways, a couple of studies on gene expression and regulation have been performed, and the functionality of these natural product genes and the encoded biosynthetic enzymes have been analyzed. From 2006 to 2008, a series of postgenomic examination on *M. xanthus* extracts by performing HPLC-MS, combined with mutagenesis experiments and structure analysis, led to the identification of five myxobacterial natural product families: myxovirescines, myxalamides, myxochelines, myxochromides A and DKxanthenes [72, 86-89]. Furthermore, the studies on the correlation between these natural products and their respective biosynthetic gene clusters has indicated that first, most natural product genes are inactive under standard laboratory conditions, commonly referred to as a ‘silent gene’ or ‘silent biosynthetic gene cluster’; Second, *M. xanthus* harbors some redundant secondary metabolite biosynthesis genes that are non-functional or missing biosynthetic pathways; Third, extremely low production yields make the identification of natural products difficult and might thus be the reason why some of the gene cluster-identified natural products have escaped being detected so far [72, 86-89].

To date, the majority of characterized and analyzed natural products from myxobacteria are polyketides (PK), non-ribosomal peptides (NRP) or hybrids of these two structural types. PKs are derived from sequential condensations of acetate, propionate, and other selected short-chain carboxylic acid precursors and NRPs are composed of amino acids, both proteinogenic and nonproteinogenic, hybrid PK/NRP natural products are composed of a mixture of both short-chain carboxylic acids and amino acids [90]. However, the number of PK and NRP natural products identified from myxobacteria is much less compared with those found in actinomyces and fungi. The relatively less identifications are probably attributed to the lack of efficient methods for screening PK and NRP genes in the past. Importantly, it has been demonstrated that most of these natural products are synthesized on gigantic molecular assembly lines composed of polyketide synthase (PKS) and non-ribosomal peptide synthetase (NRPS) multi-enzymes [91, 92]. Taking advantage of the fact that PKS and NRPS are composed of highly conserved regions, a couple of approaches have been described to identify PK and NRP gene clusters within bacterial genome both *in silico* and experimentally [72, 84]. Therefore, since the complete genome sequences of the producer myxobacteria became available, the screening for new PK and NRP natural products and their biosynthetic gene clusters have been considerably facilitated.

Recently, to increase the chance for the identification of new PK and NRP natural products from biosynthetic pathways, deeper analyses on the secondary metabolite profile of *M. xanthus* are being conducted. For example, to detect a wide range of chemical diversity in microbial extracts, an advanced screening platform has been developed. It functions by combining ultra-performance liquid chromatography and high-

resolution mass spectrometry for fast chromatographic compound separation, with statistical tools for data examination, principal component analysis, to identify and characterize diverse patterns of bacterial natural products [93]. In another study, 37 novel natural products were identified in addition to compounds already known from myxobacteria after screening the secondary metabolite profiles of 98 *M. xanthus* isolates [94]. It has indicated a tight correlation between the large number of novel natural products detected in this screen and the large number of predicted natural product genes in *M. xanthus* genome of which no specific functions have been characterized so far. This high level of diversity also suggests that a significant number of natural products in *M. xanthus* have yet to be identified. Taken together, these studies have demonstrated that *M. xanthus*, and myxobacteria in general, are and will continue to be promising sources for natural products, especially PKs and NRPs in the future.

3.5 Summary

As σ^{54} is essential for both the vegetative growth and development of *M. xanthus*, σ^{54}-dependent gene regulation appears to play a more prominent role in *M. xanthus* than in other bacteria. The importance of σ^{54} regulated gene expression in *M. xanthus* is further evidenced by the large number of genes encoding σ^{54} EBPs and putative σ^{54} promoters in its genome. As σ^{54} regulates transcription in response to environmental signals, it is reasonable that *M. xanthus* uses σ^{54} system to regulate developmental gene expression when nutrient availability is changed. A large number of characterized σ^{54} EBP have been verified to be important for *M. xanthus* development. They play

crucial roles in all stages of the *M. xanthus* multicellular life cycle. In addition, many important developmental gene clusters have putative σ^{54} promoters and σ^{54} EBP binding sites in their upstream regions, suggesting the importance of σ^{54}-dependent gene regulation in response to harsh environmental changes.

Myxobacteria such as *M. xanthus* show a remarkable potential for producing novel natural products. The speed at which these microbial cell factories can be analyzed and engineered has been enormously accelerated by the availability of genome sequences. Although a number of technologies to identify natural product synthetic gene clusters and new natural products were exploited, it is desirable to continuously develop novel approaches, such as identifying and characterizing regulatory factors involved in natural product syntheses by powerful bioinformatic tools or gene sequencing, manipulating or transferring natural product biosynthetic gene clusters into heterologous hosts to obtain a better yield and produce new analogs, and figuring out improved solutions for the difficulties in genetic engineering of myxobacteria.

3.6 References

1. Hartzell, P., *Genetic tools for dissecting motility and development of Myxococcus xanthus.* Modern microbial genetics. New York: Wiley-Liss, 2002: p. 289-322.
2. Berleman, J.E. and J.R. Kirby, *Deciphering the hunting strategy of a bacterial wolfpack.* FEMS microbiology reviews, 2009. **33**(5): p. 942-957.
3. Shimkets, L.J., *Social and developmental biology of the myxobacteria.* Microbiological reviews, 1990. **54**(4): p. 473-501.
4. Rosenberg, E., K.H. Keller, and M. Dworkin, *Cell density-dependent growth of Myxococcus xanthus on casein.* Journal of Bacteriology, 1977. **129**(2): p. 770-777.
5. Dworkin, M., *Nutritional regulation of morphogenesis in Myxococcus xanthus.* Journal of Bacteriology, 1963. **86**(1): p. 67-72.
6. Wireman, J.W. and M. Dworkin, *Morphogenesis and Developmental Interactions in Myxobacteria: These prokaryotes demonstrate complex cellular interactions unique among bacteria.* Science, 1975. **189**(4202): p. 516-523.
7. Wireman, J., *Developmental induction of Myxococcus xanthus myxospores.* Journal of Bacteriology, 1979. **140**(1): p. 147-153.
8. Zhang, Y., et al., *From individual cell motility to collective behaviors: insights from a prokaryote, Myxococcus xanthus.* FEMS microbiology reviews, 2012. **36**(1): p. 149-164.
9. O'Connor, K.A. and D.R. Zusman, *Development in Myxococcus xanthus involves differentiation into two cell types, peripheral rods and spores.* Journal of Bacteriology, 1991. **173**(11): p. 3318-3333.

10. Muñoz-Dorado, J., et al., *Myxobacteria: moving, killing, feeding, and surviving together.* Frontiers in microbiology, 2016. **7**: p. 781.
11. Kroos, L., *Highly signal-responsive gene regulatory network governing Myxococcus development.* Trends in genetics, 2017. **33**(1): p. 3-15.
12. Singer, M. and D. Kaiser, *Ectopic production of guanosine penta-and tetraphosphate can initiate early developmental gene expression in Myxococcus xanthus.* Genes & development, 1995. **9**(13): p. 1633-1644.
13. Artsimovitch, I., et al., *Structural basis for transcription regulation by alarmone ppGpp.* Cell, 2004. **117**(3): p. 299-310.
14. Pena, R.T., et al., *Relationship between quorum sensing and secretion systems.* Frontiers in microbiology, 2019. **10**: p. 1100.
15. Miller, M.B. and B.L. Bassler, *Quorum sensing in bacteria.* Annual review of microbiology, 2001. **55**(1): p. 165-199.
16. Kaplan, H.B. and L. Plamann, *A Myxococcus xanthus cell density-sensing system required for multicellular development.* FEMS microbiology letters, 1996. **139**(2-3): p. 89-95.
17. Kuspa, A., L. Plamann, and D. Kaiser, *A-signalling and the cell density requirement for Myxococcus xanthus development.* Journal of Bacteriology, 1992. **174**(22): p. 7360-7369.
18. Shimkets, L.J., R.E. Gill, and D. Kaiser, *Developmental cell interactions in Myxococcus xanthus and the spoC locus.* Proceedings of the National Academy of Sciences, 1983. **80**(5): p. 1406-1410.

19. Gorski, L., T. Gronewold, and D. Kaiser, *A ς54 activator protein necessary for spore differentiation within the fruiting body of Myxococcus xanthus.* Journal of Bacteriology, 2000. **182**(9): p. 2438-2444.

20. Ellehauge, E., M. Nørregaard-Madsen, and L. Søgaard-Andersen, *The FruA signal transduction protein provides a checkpoint for the temporal co-ordination of inter-cellular signals in Myxococcus xanthus development.* Molecular microbiology, 1998. **30**(4): p. 807-817.

21. Licking, E., L. Gorski, and D. Kaiser, *A common step for changing cell shape in fruiting body and starvation-independent sporulation of Myxococcus xanthus.* Journal of bacteriology, 2000. **182**(12): p. 3553-3558.

22. Søgaard-Andersen, L. and D. Kaiser, *C factor, a cell-surface-associated intercellular signaling protein, stimulates the cytoplasmic Frz signal transduction system in Myxococcus xanthus.* Proceedings of the National Academy of Sciences, 1996. **93**(7): p. 2675-2679.

23. Kim, S.K. and D. Kaiser, *C-factor has distinct aggregation and sporulation thresholds during Myxococcus development.* Journal of bacteriology, 1991. **173**(5): p. 1722-1728.

24. Mauriello, E.M., et al., *Gliding motility revisited: how do the myxobacteria move without flagella?* Microbiology and Molecular Biology Reviews, 2010. **74**(2): p. 229-249.

25. Burchard, R.P., *Gliding motility mutants of Myxococcus xanthus.* Journal of Bacteriology, 1970. **104**(2): p. 940-947.

26. Hodgkin, J. and D. Kaiser, *Genetics of gliding motility in Myxococcus xanthus (Myxobacterales): two gene systems control movement.* Molecular and General Genetics MGG, 1979. **171**(2): p. 177-191.

27. Sun, H., D.R. Zusman, and W. Shi, *Type IV pilus of Myxococcus xanthus is a motility apparatus controlled by the frz chemosensory system.* Current Biology, 2000. **10**(18): p. 1143-1146.

28. Bischof, L.F., et al., *The type IV pilus assembly ATPase PilB of Myxococcus xanthus interacts with the inner membrane platform protein PilC and the nucleotide-binding protein PilM.* Journal of Biological Chemistry, 2016. **291**(13): p. 6946-6957.

29. Wu, S.S. and D. Kaiser, *Genetic and functional evidence that type IV pili are required for social gliding motility in Myxococcus xanthus.* Molecular microbiology, 1995. **18**(3): p. 547-558.

30. Shi, W. and D.R. Zusman, *The two motility systems of Myxococcus xanthus show different selective advantages on various surfaces.* Proceedings of the National Academy of Sciences, 1993. **90**(8): p. 3378-3382.

31. Clausen, M., et al., *High-force generation is a conserved property of type IV pilus systems.* Journal of bacteriology, 2009. **191**(14): p. 4633-4638.

32. Dinet, C., et al., *Linking single-cell decisions to collective behaviours in social bacteria.* Philosophical Transactions of the Royal Society B, 2021. **376**(1820): p. 20190755.

33. Li, Y. and L. Plamann, *Purification and in vitro phosphorylation of Myxococcus xanthus AsgA protein.* Journal of bacteriology, 1996. **178**(1): p. 289-292.

34. Jakovljevic, V., et al., *PilB and PilT are ATPases acting antagonistically in type IV pilus function in Myxococcus xanthus.* Journal of bacteriology, 2008. **190**(7): p. 2411-2421.

35. Li, Y., et al., *Extracellular polysaccharides mediate pilus retraction during social motility of Myxococcus xanthus.* Proceedings of the National Academy of Sciences, 2003. **100**(9): p. 5443-5448.

36. Zusman, D.R., et al., *Chemosensory pathways, motility and development in Myxococcus xanthus.* Nature Reviews Microbiology, 2007. **5**(11): p. 862-872.

37. Black, W.P., Q. Xu, and Z. Yang, *Type IV pili function upstream of the Dif chemotaxis pathway in Myxococcus xanthus EPS regulation.* Molecular microbiology, 2006. **61**(2): p. 447-456.

38. Xu, Q., et al., *Independence and interdependence of Dif and Frz chemosensory pathways in Myxococcus xanthus chemotaxis.* Molecular microbiology, 2008. **69**(3): p. 714-723.

39. Kaiser, D., *Are there lateral as well as polar engines for A-motile gliding in myxobacteria?* Journal of bacteriology, 2009. **191**(17): p. 5336-5341.

40. Nan, B., et al., *Bacteria that glide with helical tracks.* Current Biology, 2014. **24**(4): p. R169-R173.

41. Mignot, T., et al., *Evidence that focal adhesion complexes power bacterial gliding motility.* Science, 2007. **315**(5813): p. 853-856.

42. Goldman, B., et al., *Evolution of sensory complexity recorded in a myxobacterial genome.* Proceedings of the National Academy of Sciences, 2006. **103**(41): p. 15200-15205.

43. Ulrich, L.E., E.V. Koonin, and I.B. Zhulin, *One-component systems dominate signal transduction in prokaryotes.* Trends in microbiology, 2005. **13**(2): p. 52-56.

44. Zhang, X., et al., *Mechanochemical ATPases and transcriptional activation.* Molecular microbiology, 2002. **45**(4): p. 895-903.

45. Shingler, V., *Signal sensory systems that impact σ54-dependent transcription.* FEMS microbiology reviews, 2011. **35**(3): p. 425-440.

46. Studholme, D.J. and R. Dixon, *Domain architectures of σ54-dependent transcriptional activators.* Journal of bacteriology, 2003. **185**(6): p. 1757-1767.

47. Keseler, I.M. and D. Kaiser, *σ54, a vital protein for Myxococcus xanthus.* Proceedings of the National Academy of Sciences, 1997. **94**(5): p. 1979-1984.

48. Hager, E., H. Tse, and R.E. Gill, *Identification and characterization of spdR mutations that bypass the BsgA protease-dependent regulation of developmental gene expression in Myxococcus xanthus.* Molecular microbiology, 2001. **39**(3): p. 765-780.

49. Tse, H. and R.E. Gill, *Bypass of A-and B-signaling requirements for Myxococcus xanthus development by mutations in spdR.* Journal of bacteriology, 2002. **184**(5): p. 1455-1457.

50. Harris, B.Z., D. Kaiser, and M. Singer, *The guanosine nucleotide (p) ppGpp initiates development and A-factor production inMyxococcus xanthus.* Genes & development, 1998. **12**(7): p. 1022-1035.

51. Ossa, F., et al., *The Myxococcus xanthus Nla4 protein is important for expression of stringent response-associated genes, ppGpp accumulation, and fruiting body development.* Journal of bacteriology, 2007. **189**(23): p. 8474-8483.

52. Giglio, K.M., J. Eisenstatt, and A.G. Garza, *Identification of enhancer binding proteins important for Myxococcus xanthus development.* Journal of bacteriology, 2010. **192**(1): p. 360-364.

53. Kim, D.-H., et al., *Operon required for fruiting body development in Myxococcus xanthus.* Journal of microbiology and biotechnology, 2009. **19**(11): p. 1288-1294.

54. Caberoy, N.B., et al., *Global mutational analysis of NtrC-like activators in Myxococcus xanthus: identifying activator mutants defective for motility and fruiting body development.* Journal of Bacteriology, 2003. **185**(20): p. 6083-6094.

55. Sarwar, Z. and A.G. Garza, *The Nla28S/Nla28 two-component signal transduction system regulates sporulation in Myxococcus xanthus.* Journal of bacteriology, 2012. **194**(17): p. 4698-4708.

56. Gronewold, T.M. and D. Kaiser, *act operon control of developmental gene expression in Myxococcus xanthus*. 2002, Am Soc Microbiol.

57. Gorski, L. and D. Kaiser, *Targeted mutagenesis of ς54 activator proteins in Myxococcus xanthus.* Journal of Bacteriology, 1998. **180**(22): p. 5896-5905.

58. Lancero, H., et al., *Characterization of a Myxococcus xanthus mutant that is defective for adventurous motility and social motility.* Microbiology, 2004. **150**(12): p. 4085-4093.

59. Youderian, P., et al., *Identification of genes required for adventurous gliding motility in Myxococcus xanthus with the transposable element mariner.* Molecular microbiology, 2003. **49**(2): p. 555-570.

60. Wall, D. and D. Kaiser, *Type IV pili and cell motility.* Molecular microbiology, 1999. **32**(1): p. 01-10.

61. Wu, S.S. and D. Kaiser, *Regulation of expression of the pilA gene in Myxococcus xanthus.* Journal of bacteriology, 1997. **179**(24): p. 7748-7758.

62. Lu, A., et al., *Exopolysaccharide biosynthesis genes required for social motility in Myxococcus xanthus.* Molecular microbiology, 2005. **55**(1): p. 206-220.

63. Giglio, K.M., et al., *A cascade of coregulating enhancer binding proteins initiates and propagates a multicellular developmental program.* Proceedings of the National Academy of Sciences, 2011. **108**(32): p. E431-E439.

64. Diodati, M.E., et al., *Nla18, a key regulatory protein required for normal growth and development of Myxococcus xanthus.* Journal of bacteriology, 2006. **188**(5): p. 1733-1743.

65. Chin, Y.-W., et al., *Drug discovery from natural sources.* The AAPS journal, 2006. **8**(2): p. E239-E253.

66. Reichenbach, H., *Myxobacteria, producers of novel bioactive substances.* Journal of Industrial Microbiology and Biotechnology, 2001. **27**(3): p. 149-156.

67. Gerth, K., et al., *Myxobacteria: proficient producers of novel natural products with various biological activities—past and future biotechnological aspects with the focus on the genus Sorangium.* Journal of biotechnology, 2003. **106**(2-3): p. 233-253.

68. Altmann, K.-H., et al., *The epothilones: an outstanding family of anti-tumor agents: from soil to the clinic*. Vol. 90. 2009: Springer Science & Business Media.

69. Irschik, H., et al., *The sorangicins, novel and powerful inhibitors of eubacterial RNA polymerase isolated from myxobacteria.* The Journal of antibiotics, 1987. **40**(1): p. 7-13.

70. Irschik, H., et al., *Disorazol A, an efficient inhibitor of eukaryotic organisms isolated from myxobacteria.* The Journal of Antibiotics, 1995. **48**(1): p. 31-35.

71. Jansen, R., et al., *Antibiotics from gliding bacteria, LXXX. Chivosazoles A–F: novel antifungal and cytotoxic macrolides from Sorangium cellulosum (Myxobacteria).* Liebigs Annalen, 1997. **1997**(8): p. 1725-1732.

72. Bode, H.B. and R. Müller, *Secondary metabolism in myxobacteria.* Myxobacteria: Multicellularity and differentiation, 2007: p. 259-282.

73. Weissman, K.J. and R. Müller, *A brief tour of myxobacterial secondary metabolism.* Bioorganic & medicinal chemistry, 2009. **17**(6): p. 2121-2136.

74. Thierbach, G., et al., *The mode of action of stigmatellin, a new inhibitor of the cytochrome b-c1 segment of the respiratory chain.* Biochimica et Biophysica Acta (BBA)-Bioenergetics, 1984. **765**(2): p. 227-235.

75. Shen, Y., et al., *A mechanism for the potent inhibition of eukaryotic acetyl-coenzyme A carboxylase by soraphen A, a macrocyclic polyketide natural product.* Molecular cell, 2004. **16**(6): p. 881-891.

76. Bode, H.B., et al., *The Leupyrrins: A Structurally Unique Family of Secondary Metabolites from the Myxobacterium Sorangium c ellulosum#.* Journal of natural products, 2003. **66**(9): p. 1203-1206.

77. Knauth, P. and H. Reichenbach, *On the mechanism of action of the myxobacterial fungicide ambruticin.* The Journal of antibiotics, 2000. **53**(10): p. 1182-1190.

78. Gerth, K., et al., *The Jerangolids: A Family of New Antifungal Compounds from Sorangium cellulosum (Myxobacteria) Production, Physico-chemical and*

Biological Properties of Jerangolid A. The Journal of Antibiotics, 1996. **49**(1): p. 71-75.

79. Erol, Ö., et al., *Biosynthesis of the myxobacterial antibiotic corallopyronin A.* ChemBioChem, 2010. **11**(9): p. 1253-1265.

80. Gerth, K., et al., *The myxovirescins, a family of antibiotics from Myxococcus virescens (Myxobacterales).* The Journal of antibiotics, 1982. **35**(11): p. 1454-1459.

81. Steinmetz, H., et al., *Thuggacins, macrolide antibiotics active against Mycobacterium tuberculosis: isolation from myxobacteria, structure elucidation, conformation analysis and biosynthesis.* Chemistry–A European Journal, 2007. **13**(20): p. 5822-5832.

82. Baumann, S., et al., *Cystobactamids: myxobacterial topoisomerase inhibitors exhibiting potent antibacterial activity.* Angewandte Chemie International Edition, 2014. **53**(52): p. 14605-14609.

83. Wenzel, S.C. and R. Müller, *Myxobacteria—'microbial factories' for the production of bioactive secondary metabolites.* Molecular BioSystems, 2009. **5**(6): p. 567-574.

84. Bode, H.B. and R. Müller, *The impact of bacterial genomics on natural product research.* Angewandte Chemie International Edition, 2005. **44**(42): p. 6828-6846.

85. Ronning, C.M. and W.C. Nierman, *The genomes of Myxococcus xanthus and Stigmatella aurantiaca.* Myxobacteria: Multicellularity and differentiation, 2007: p. 283-298.

86. Meiser, P., et al., *DKxanthene biosynthesis—Understanding the basis for diversity-oriented synthesis in myxobacterial secondary metabolism.* Chemistry & biology, 2008. **15**(8): p. 771-781.

87. Simunovic, V., et al., *Myxovirescin A Biosynthesis is Directed by Hybrid Polyketide Synthases/Nonribosomal Peptide Synthetase, 3-Hydroxy-3-Methylglutaryl–CoA Synthases, and trans-Acting Acyltransferases.* ChemBioChem, 2006. **7**(8): p. 1206-1220.

88. Wenzel, S.C., et al., *Nonribosomal peptide biosynthesis: point mutations and module skipping lead to chemical diversity.* Angewandte Chemie International Edition, 2006. **45**(14): p. 2296-2301.

89. Meiser, P., H. Bode, and R. Müller, *DKxanthenes: Novel secondary metabolites from the myxobacterium Myxococcus xanthus essential for sporulation.* Proc. Natl. Acad. Sci. USA, 2006. **103**: p. 19128-19133.

90. Rath, C.M., et al., *NRPS/PKS hybrid enzymes and their natural products.* Comprehensive natural products II: chemistry and biology. Oxford: Elsevier, 2010: p. 453-492.

91. Staunton, J. and K.J. Weissman, *Polyketide biosynthesis: a millennium review.* Natural product reports, 2001. **18**(4): p. 380-416.

92. Finking, R. and M.A. Marahiel, *Biosynthesis of nonribosomal peptides.* Annual review of microbiology, 2004. **58**: p. 453.

93. Krug, D., et al., *Efficient mining of myxobacterial metabolite profiles enabled by liquid chromatography–electrospray ionisation-time-of-flight mass spectrometry and compound-based principal component analysis.* Analytica chimica acta, 2008. **624**(1): p. 97-106.

94. Krug, D., et al., *Discovering the hidden secondary metabolome of Myxococcus xanthus: a study of intraspecific diversity.* Applied and environmental microbiology, 2008. **74**(10): p. 3058-3068.

Chapter 4

Nla28-Mediated σ^{54}-Dependent Regulation of Starvation Response Genes in *Myxoccocus xanthus*

4.1 Author Contributions

Muqing Ma designed and conducted all experiments involving mutagenesis on plasmids, strain constructions, Nla28 DBD protein expression and purification, in vitro protein-DNA binding assays and in vivo promoter activity assays; assisted with experiments involving putative promoter identification and qPCR; collected, analyzed, and interpreted the data with Anthony Garza; David Lemon performed phenotypic assays and qPCR. Eduardo Caro, Linnea Ritchie assisted with qPCR. Charles Ryan, Victoria Spearing, Kimberly Murphy assisted with data collection. Anthony Garza and Muqing Ma wrote the manuscript. Anthony Garza and Roy Welch conceived and supervised the research.

4.2 Abstract

Myxococcus xanthus copes with starvation by producing fruiting bodies filled with dormant and stress-resistant spores. Here, we aimed to better define the gene regulatory network associated with Nla28, a transcriptional activator/enhancer binding protein (EBP) and a key regulator of the early starvation response. Previous work showed that Nla28 directly regulates EBP genes that are important for fruiting body development. However, the Nla28 regulatory network is likely to be much larger, as hundreds of starvation-induced genes are downregulated in a *nla28* mutant. To identify candidates for direct Nla28-mediated transcription, we analyzed the downregulated genes using bioinformatics; nine potential Nla28 target promoters (29 genes) were uncovered. The results of in vitro promoter binding assays, coupled with in vitro and in vivo mutational analyses, suggested that the nine promoters, along with the three previously identified EBP gene promoters, are indeed in vivo targets of Nla28. These results also suggest that Nla28 uses tandem, imperfect repeats of an 8-bp sequence for promoter binding. Interestingly, eight of the new Nla28 target promoters are predicted to be intragenic. Based on mutational analyses, the newly identified Nla28 target loci contain at least one gene that is important for starvation-induced development. Most of these loci contain genes predicted to be involved in metabolic or defense-related functions. Using the consensus Nla28 binding sequence, bioinformatics and expression profiling, 58 additional promoters and 102 genes were tagged as potential Nla28 targets. Among these

putative Nla28 targets, functions such as regulatory, metabolic and cell envelope biogenesis were assigned to many genes.

4.3 Introduction

Bacteria typically have numerous transcription factors to activate and/or repress transcription of genes that are important for development. In the soil bacterium *Myxococcus xanthus*, the early developmental pathway relies heavily on enhancer binding proteins or EBPs [1, 2], which are transcriptional activators that help σ^{54} -RNA polymerase initiate transcription at σ^{54}-type promoters [3-5]. Indeed, a cascade of EBPs modulates changes in developmental gene transcription during sequential stages of early development [6]. The aim of this study was to identify and characterize the developmental gene targets of an early-functioning EBP called Nla28.

During development, *M. xanthus* forms a biofilm that contains a mat of rod-shaped cells, known as peripheral rods [7, 8], and multicellular fruiting bodies that house about 100,000 dormant and stress-resistant spores [9]. Starvation triggers the developmental process, and much is known about the subsequent morphological changes that yield spore-filled fruiting bodies. Namely, cells migrate into an aggregation center, the cell aggregate becomes larger and eventually develops the dome-shaped appearance of a fruiting body, the rod-shaped cells in the newly formed fruiting body differentiate into spherical cells, and the spherical cells mature into stress-resistant spores.

In standard developmental assays, at least two starvation-induced signaling events must occur before cells begin building fruiting bodies. The first event is the RelA-mediated accumulation of the intracellular starvation signal (p)ppGpp [10, 11]. The second event, which depends on relatively high levels of

(p)ppGpp [11], is the accumulation of an extracellular signal called A-signal. A-signal is thought to be a cell-density signal composed of amino acids and perhaps peptides [12-14]. Thus, it has been suggested that (p)ppGpp accumulation indicates cells are starving, and the subsequent accumulation of A-signal indicates that enough starving cells are present to build a fruiting body.

Our previous work linked four of the 53 *M. xanthus* EBPs to the accumulation of early developmental signals [15-19]. The EBP Nla28, which is the focus of this study, was linked to A-signal production via extracellular complementation assays [15]. Nla28 is a response regulator-type EBP that partners with the membrane-bound histidine protein kinase Nla28S to form a two component signal transduction system [20, 21]. Two pieces of data led to the suggestion that the Nla28S/Nla28 signal transduction system might respond to A-signal, in addition to being important for A-signal production [21]. First, *nla28S* and *nla28* form a two-gene operon and Nla28 is involved in autoregulation [6]. Secondly, A-signal is important for full developmental expression of the *nla28S* gene and presumably *nla28S-nla28* operon [20].

In addition to being connected to A-signal production, *nla28* is known to be expressed early in development and important for expression of early developmental genes; inactivation of *nla28* impairs or abolishes expression of many genes that are induced 1-2 hours post-starvation [6]. This finding led to the suggestion that the Nla28S/Nla28 signal transduction system targets starvation-associated or stress-responsive genes [19, 21]. Indeed, the results of a recent study suggest that Nla28 directly modulates the activities of three natural product

promoters during the transition into stationary phase and during development; both events are associated with nutrient depletion [22].

Although Nla28 is important for expression of hundreds of *M. xanthus* developmental genes and for the developmental process [6, 15], only two classes of developmental genes have been linked to direct Nla28-mediated regulation: regulatory genes and natural product-associated genes [6, 22]. One goal of this study was to determine how Nla28 identifies its target promoters. A second goal was to use this information to identify the larger network of Nla28-targeted promoters, with an emphasis on promoters that show substantial developmental regulation. The final goal was to analyze genes under direct Nla28 control to better understand the function of Nla28.

Since EBPs modulate transcription at σ^{54} promoters [23, 24], our search for direct Nla28 targets focused on known or putative σ^{54} promoters. Namely, we searched for known or putative σ^{54} promoters upstream of genes that showed substantial developmental regulation and at least 2-fold downregulation in a *nla28* mutant. Further analyses suggested that the candidate σ^{54} promoters and genes are indeed in vivo targets of Nla28 and that Nla28 uses tandem 8-bp sequences for promoter binding. Our results also indicated that one or more genes in each Nla28 target loci is important for production of stress-resistant spores during starvation-induced development. Additional candidates for direct Nla28 regulation were identified using the consensus Nla28 binding sequence; the Nla28 regulon is likely to be well over 100 genes. Known and predicted

functions of Nla28 target genes are discussed in the context of *M*. *xanthus* development.

4.4 Materials and Methods

4.4.1 Bacterial strains and plasmids

The strains, plasmids and primers used in this study are shown in Tables 4.1 and 4.2. Plasmids generated for this study were analyzed via DNA sequence analysis. *M. xanthus* strains were confirmed as described below. Plasmid insertions in Nla28 target genes in wild-type strain DK1622 were generated as previously described [15, 25] and confirmed via PCR and DNA sequencing.

Table 4.1 Strains and plasmids

Strain or Plasmid	Relevant Characteristics	Reference
E. coli strains		
BL21 (DE3)	F^- *ompT gal dcm lon hsdS*$_B$($r_B^-m_B^-$) λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB*$^+$]$_{K\text{-}12}$(λ^S)	[26]
TOP10	F- *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ *lac*X74 *rec*A1 *ara*D139 Δ(*araleu*)7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G	Invitrogen
M. xanthus strains		
DK1622	Wild-type (parental strain)	[27]
AG328	pNBC28::*nla28* (insertion in the *nla28* gene)	[15]
MM502	pMM502::*nla28* locus (*lacZ* fused to wild-type *nla28* promoter)	This study
MM503	pMM503::*nla28* locus (*lacZ* fused to mutated *nla28* promoter)	This study
MM504	pMM504::*nla28* locus (*lacZ* fused to mutated *nla28* promoter)	This study
MM505	pMM505::*nla28* locus (*lacZ* fused to mutated *nla28* promoter)	This study

MM506	pMM506::*nla28* locus (*lacZ* fused to mutated *nla28* promoter)	This study
MM507	pMM507::*actB* locus (*lacZ* fused to wild-type *actB* promoter)	This study
MM508	pMM508:: *actB* locus (*lacZ* fused to mutated *actB* promoter)	This study
MM509	pMM509:: *actB* locus (*lacZ* fused to mutated *actB* promoter)	This study
MM510	pMM510:: *actB* locus (*lacZ* fused to mutated *actB* promoter)	This study
MM511	pMM511::*actB* locus (*lacZ* fused to mutated *actB* promoter)	This study
MM512	pMM512::*pilA* locus (*lacZ* fused to wild-type *pilA* promoter)	This study
MM513	pMM513::*pilA* locus (*lacZ* fused to mutated *pilA* promoter)	This study
MM514	pMM514::*pilA* locus (*lacZ* fused to mutated *pilA* promoter)	This study
MM515	pMM515::*pilA* locus (*lacZ* fused to mutated *pilA* promoter)	This study
MM516	pMM516::*pilA* locus (*lacZ* fused to mutated *pilA* promoter)	This study
MM517	pMM517::MXAN5040 locus (*lacZ* fused to wild-type MXAN5040 promoter)	This study
MM518	pMM518::MXAN5040 locus (*lacZ* fused to mutated MXAN5040 promoter)	This study
MM519	pMM519::MXAN5040 locus (*lacZ* fused to mutated MXAN5040 promoter)	This study
MM520	pMM520::MXAN5040 locus (*lacZ* fused to mutated MXAN5040 promoter)	This study
MM521	pMM521::MXAN5040 locus (*lacZ* fused to mutated MXAN5040 promoter)	This study
DK2161	A^-S^-	[28]

DK10603	D*actB*	[29]
MXAN881	pKAM1001::MXAN881 (insertion in the MXAN881 gene)	This study
MXAN989	pKAM1002::MXAN989 (insertion in the MXAN989 gene)	This study
MXAN2510	pKAM1003::MXAN2510 (insertion in the MXAN2510 gene)	This study
MXAN5040	pKAM1004::MXAN5040 (insertion in the MXAN5040 gene)	This study
MXAN6732	pEC101::MXAN6732 (insertion in the MXAN6732 gene)	This study
MXAN7147	pKAM1005::MXAN7147 (insertion in the MXAN7147 gene)	This study
MXAN7279	pKAM1006::MXAN7279 (insertion in the MXAN7279 gene)	This study
DK10410	D*pilA*	[30]
SW2802	D*mrpB*	[31]
Plasmids		
pKAM1001	pCR2.1TOPO (Invitrogen) containing a 319-bp fragment of the MXAN881 gene	This study
pKAM1002	pCR2.1TOPO (Invitrogen) containing a 459-bp fragment of the MXAN989 gene	This study
pKAM1003	pCR2.1TOPO containing a 231-bp fragment of the MXAN2510 gene	This study
pKAM1004	pCR2.1TOPO containing a 364-bp fragment of the MXAN5040 gene	This study
pKAM1005	pCR2.1TOPO containing a 349-bp fragment of the MXAN7147 gene	This study
pKAM1006	pCR2.1TOPO containing a 377-bp fragment of the MXAN7279 gene	This study
pMAL-c5x	Amp^r (Maltose-binding protein fusion vector for Nla28-DBD protein expression)	New England Biolabs

pMM302	Amp^r (Fragment for Nla28-DBD expression in pMAL-c5x)	[22]
pREG1727	Amp^r Kan^r (attP for integration at the chromosomal Mx8 attachment site and promoterless lacZYA genes)	[32]
pMM500	pCR-Blunt (Invitrogen) containing the 500-bp nla28 promoter fragment	This study
pMM501	Amp^r Kan^r (a derivative of pREG1727 with attP removed)	This study
pMM502	pMM501 with a 500-bp nla28 promoter fragment that contains the three Nla28 half binding sites (HS1-3) and the σ^{54} promoter	This study
pMM503	pMM502 with a 2-bp substitution in HS2 (CTCCGCAG to AACCGCAG)	This study
pMM504	pMM502 with a 4-bp substitution in HS2 (CTCCGCAG to AAAAGCAG)	This study
pMM505	pMM502 with a 6-bp substitution in HS2 (CTCCGCAG to AAAAAAAG)	This study
pMM506	pMM502 with an 8-bp substitution in HS2 (CTCCGCAG to AAAAAATA)	This study
pMM507	pMM501 with a 500-bp actB promoter fragment that contains the two Nla28 half binding sites (HS1 and 2) and the σ^{54} promoter	This study
pMM508	pMM507 with a 2-bp substitution in HS1 (CGCCGCGG to AACCGCGG)	This study
pMM509	pMM507 with a 4-bp substitution in HS1 (CGCCGCGG to AAAAGCGG)	This study
pMM510	pMM507 with a 6-bp substitution in HS1 (CGCCGCGG to AAAAAAGG)	This study
pMM511	pMM507 with a 8-bp substitution in HS1 (CGCCGCGG to AAAAAAAA)	This study

pMM512	pMM501 with a 500-bp pilA promoter fragment that contains the two Nla28 half binding sites (HS1 and 2) and the σ^{54} promoter	This study
pMM513	pMM512 with a 2-bp substitution in HS1 (GTGCGCAC to AAGCGCAC)	This study
pMM514	pMM512 with a 4-bp substitution in HS1 (GTGCGCAC to AAAAGCAC)	This study
pMM515	pMM512 with a 4-bp substitution in HS1 (GTGCGCAC to AAAAAAAC)	This study
pMM516	pMM512 with a 4-bp substitution in HS1 (GTGCGCAC to AAAAAATA)	This study
pMM517	pMM501 with a 500-bp MXAN5040 promoter fragment that contains the two Nla28 half binding sites (HS1 and 2) and the σ^{54} promoter	This study
pMM518	pMM517 with a 2-bp substitution in HS1 (TGGCGCAG to AAGCGCACG)	This study
pMM519	pMM517 with a 4-bp substitution in HS1 (TGGCGCAG to AAAAGCAG)	This study
pMM520	pMM517 with a 4-bp substitution in HS1 (TGGCGCAG to AAAAAAAG)	This study
pMM521	pMM517 with a 4-bp substitution in HS1 (TGGCGCAG to AAAAAATA)	This study

Table 4.2 Primers

	Sequence	Amplicon size
qPCR		
actB foward *actB* reverse	5'-CTCCAGGACGAGGAGTTCTTCCG-3' 5'-GCGATTTCCTTCTCCAGGTCGC-3'	101 bp
mrpB forward *mrpB* reverse	5'-GCCAGCCTGATTCCCACTTT-3' 5'-ACCGTACTTCTGGAGCTTGC-3'	144 bp
nla6 forward	5'-GCTCATCGAGTCCGAGCTG-3'	113 bp

nla6 reverse	5’-ATTCGTCCAGGAAGAGCGTG-3’	
nla28 forward *nla28* reverse	5’-GTGTTGCAGGAGGGCGAAATCC-3’ 5’-CGCGTCTTGAGGTCCTTGTTCG-3’	98 bp
nla28S (MXAN1166) forward *nla28S* (MXAN1166) reverse	5'-CGCTCGGAAGAAGGAAAGGG-3' 5'-CGGAAGCAGTCGTTGTCTGA-3'	100 bp
pilA forward *pilA* reverse	5'-GTCGTTGGAGATGCACGGAACG-3' 5'-CAACCGTTACGGCTACCGTGTG-3'	106 bp
rpoD forward *rpoD* reverse	5’-GACGTCTTGGAGCGAGAGCTGTC-3’ 5’-CTCCATCATCTTGGTCCGGAGGTC-3’	102 bp
MXAN881foward MXAN881reverse	5’-ATGATGGGGATGGCTTGGTG-3’ 5’-GCAGCGCGGATAATCGTTTT-3’	104 bp
MXAN989 forward MXAN989 reverse	5’-GGAGAAGAACGTGGTGGAGG-3’ 5’-GGTCCTGCTCGTACACATCC-3’	107 bp
MXAN2511 forward MXAN2511 reverse	5’-TGGCGGAATGATGAACGACA-3’ 5’-TTGCTGGGGACCGTACAAC-3’	118 bp
MXAN5040 forward MXAN5040 reverse	5’-GTGGACCAACTGCTACCACA-3’ 5’-TGACCAGCAGGTTCTTCGTC-3’	128 bp
MXAN6732 forward MXAN6732 reverse	5’-ATGACTGCGCCTTCCTGAAA-3’ 5’-GTTGACCGTGATGGGAGAGG-3’	117 bp
MXAN7147 forward MXAN7147 reverse	5’-TTCACCGTGTCCAACCACAC-3’ 5’-TGGACATTCCCAAAGCCAAGA-3’	154 bp
MXAN7279 forward MXAN7279 reverse	5'-GCCACGGGTACCTTCAACAA-3' 5'-GGCCTTGCCTCCCAGTTG-3'	105 bp
Promoter fragments		
actB forward *actB* reverse	5’-CCGCCGTCGTGGAGTC-3’ 5’-CGGATTTAGCAATGGTTGTGCCAC-3’	210 bp
dev forward *dev* reverse	5’-ACGTTGCAGACGGGGTGAG-3’ 5’-CCTCGTACTTCGACTTCCGAAGAG-3’	180 bp
mrpB forward *mrpB* reverse	5’-TGAGGCCGGTGTTCCGGG-3’ 5’-TCCGCCAACCTCGGCGG-3’	220 bp
nla28 forward	5’-GATGACGCGCGCAGCTTGC-3’	200 bp

nla28 reverse	5’-GAGATTGCGCCGTCCCAGC-3’	
pilA forward	5’-GAGCGCTTCGGATGCGTAGG-3’	200 bp
pilA reverse	5’-TCCTCAGAGAAGGTTGCAACGG-3’	
MXAN881 forward	5’-CGCGCTTCTCCACGTCCTTG-3’	196 bp
MXAN881 reverse	5’-GACGTCATGCTGGAGATTTCCGG-3’	
MXAN989 forward	5’-GTAATCGTCGCCGTGGCCTG-3’	201 bp
MXAN989 reverse	5’-TCCTTCCAGGGTGTCCTGCG-3’	
MXAN2511 forward	5’-CTCCGCTGGTGTACCACATGG-3’	200 bp
MXAN2511 reverse	5’-CAGAATCGGCATCAGGATGGAGA-3’	
MXAN5040 forward	5’-GTCCGTTGCGCGAGCTGGA-3’	200 bp
MXAN5040 reverse	5’-CGCGACAGGTGCGTCGC-3’	
MXAN6732 forward	5’-GGGGCTGTCTCCGTGACGC-3’	201 bp
MXAN6732 reverse	5’-CGCAACGGGCATGCCGA-3’	
MXAN7147 forward	5’-AGGTGATGGCGAAGCGGGC-3’	197 bp
MXAN7147 reverse	5’-CCTTCTCGAAGTGTTCCTGGACCA-3’	
MXAN7280 forward	5’-CAGTGCTGGTGGCGAAGCAC-3’	200 bp
MXAN7280 reverse	5’-CTCAGCCAGCTCGGGCG-3’	
Insertions		
MXAN881 forward	5’-CTGCACGTCCCTGCCGGAC-3’	319 bp
MXAN881 reverse	5’-GTTGCAGGCGCCGCAGTGG-3’	
MXAN989 forward	5’-CTGCTGGCGCGTCAGAGC-3’	459 bp
MXAN989 reverse	5’-GCGGACGTCCGCGTACATG-3’	
MXAN2510 forward	5’-GAAGCAGCGCCGCAACCG-3’	231 bp
MXAN2510 reverse	5’-GCCTCCACCAGCGCTTGC-3’	
MXAN5040 forward	5’-GAGCAGACGCCCTCCAGC-3’	364 bp
MXAN5040 reverse	5’-GCTCGCTGATGAGGGAGCGC-3’	
MXAN6732 foward	5’-GGATGGCCAACCCGGCGA-3’	418 bp
MXAN6732 reverse	5’-AAGATCCAGTCGTGCGCCTCC-3’	
MXAN7147 forward	5’-GCCGAGCGTCAACTGGCC-3’	349 bp
MXAN7147 reverse	5’-TCGACTCCACCTGCGCGC-3’	
MXAN7279 forward	5’-CGAAGGCGGGTCCGGAGG-3’	377 bp
MXAN7279 reverse	5’-CTCCGCCTCCGGCAGCAG-3’	

4.4.2 Growth and development

E. coli strains were grown at 37°C in LB broth containing 1.0% tryptone, 0.5% yeast extract, and 0.5% NaCl or on plates containing LB broth and 1.5% agar. LB broth and LB plates were supplemented with 40 μg of kanamycin sulfate/ml, 100 μg of ampicillin/ml or 10 μg of oxytetracycline/ml as needed.

M. xanthus strains were grown at 32°C in CTTYE broth [1.0% Casitone, 0.5% yeast extract, 10 mM Tris-HCl (pH 8.0), 1 mM KH_2PO_4, and 8 mM $MgSO_4$] or on plates containing CTTYE broth and 1.5% agar. CTTYE broth and plates were supplemented with 40 μg of kanamycin sulfate/ml or 10 μg of oxytetracycline/ml as needed.

Fruiting body development was induced by placing *M. xanthus* cells on plates containing TPM buffer [10 mM Tris-HCl (pH 8.0), 1 KH_2PO_4, and 8 mM $MgSO_4$] and 1.5% agar or in 6-well microtiter plates containing MC7 buffer (10 mM MOPS, 1 mM $CaCl_2$, final pH 7.0), and incubating the plates at 32°C. Briefly, *M. xanthus* strains were grown in flasks containing CTTYE broth and the cultures were incubated at 32°C with vigorous swirling. The cells were pelleted when the cultures reached a density of about 5 x 10^8 cells/ml, the supernatants were removed, and the cells were resuspended in TPM buffer or MC7 buffer to a final density of 5 x 10^9 cells/ml. Aliquots of the cells in TPM buffer were spotted onto TPM agar plates and aliquots of the cells in MC7 buffer were placed in 6-well microtiter plates containing MC7 buffer. The cells were incubated at 32°C and development was monitored as previously described [15, 19]. Cells were harvested at various times during development and prepared for quantitative

PCR (qPCR), β-Galactosidase assays or sporulation assays.

To examine the sporulation efficiency of each *M. xanthus* strain, developing cells were harvested from TPM agar plates after 5 days and the cells were placed in 400 μl of TPM buffer. The resuspended cells were first dispersed by a 10-second burst of sonication using a model 100 Sonic Dismembrator (Fisher Scientific) that was set at an intensity of 1.5. Aliquots of the dispersed cells were placed on a Petroff-Hausser counting chamber and phase-contrast microscopy was used to determine the number of spherical-shaped cells that were present. Other aliquots of the cell suspension were subjected to three 10-second bursts of sonication using an intensity setting of 4, and the sonication-treated cells were incubated at 50°C for 2 hours. The number of heat- and sonication-resistant spores that germinated into colonies was determined by placing heat- and sonication-treated cells in liquified CTT soft agar [1.0% Casitone, 10 mM Tris-HCl (pH 8.0), 1 mM KH_2PO_4, 8 mM $MgSO_4$ and 0.7% agar], pouring the soft agar onto CTTYE agar plates and incubating the plates at 32°C for 5 days.

4.4.3 Motility assays

Motility assays were performed as previously described [15, 25]. Briefly, *M. xanthus* cells were grown to a density of about 5×10^8 cells/ml in CTTYE broth. The cells were pelleted by centrifugation, the supernatant was removed, and the cells were resuspended in CTTYE broth to a density of 5×10^9 cells/ml. Aliquots (3 μl) of the cell suspensions were spotted onto CTTYE plates containing 1.5 or 0.4% agar, the spots were allowed to dry, and the plates were

placed at 32°C. After 3 days of incubation, five swarms of each strain were measured, and their mean diameter was normalized to the mean diameter of five swarms formed by wild-type strain DK1622.

4.4.4 Plasmid transfer to M. xanthus

Plasmids containing internal fragments of Nla28 target genes or wild-type or mutant copies of the *nla28* promoter were electroporated into wild-type strain DK1622 as described previously [33]. Kan^r electroporants that contained a plasmid integrated in the *nla28* locus or in a *nla28* target gene by homologous recombination, or containing a plasmid integrated in the chromosomal Mx8 phage attachment site (*attB*) by site-specific recombination were identified via PCR and DNA sequencing. Kan^r electroporants carrying a single plasmid insertion were assayed for development, motility or promoter activity as needed.

4.4.5 Standard DNA procedures

Chromosomal DNA from *M. xanthus* strains was extracted using a ZYMO Research gDNA extraction kit. Oligonucleotides used in PCR reactions were synthesized by Integrated DNA Technologies (IDT) and are listed in Table 4.2. Plasmid DNA was extracted using the Promega Nucleic Acid Purification kit. Amplified and digested DNA fragments were purified using the Gel Extraction Minipreps kit from Bio Basic. For all kits, the manufacturer's protocols were used. The compositions of all plasmids and promoter fragments were confirmed by DNA sequencing (Genewiz).

4.4.6 In vivo mutational analysis of the putative Nla28 binding site in the nla28 promoter

Wild-type and mutant *actB*, *nla28*, *pilA* and MXAN5040 promoter fragments were cloned into the promoterless *lacZ* expression vector pREG1727 to create *lacZ* transcriptional fusions [32]. The plasmids were introduced into strain DK1622 or a derivative of strain DK1622 carrying an insertion in the *nla28* gene, and cells carrying a plasmid integrated at the Mx8 phage attachment site in the chromosome were identified via PCR and DNA sequence analysis. The in vivo activities of wild-type and mutant promoters were determined by measuring the specific activities of β-galactosidase in cells developing in submerged cultures for 1, 2, 6, 8, 12, 24 or 48 hours [22, 34]. β-Galactosidase-specific activity is defined as nanomoles of o-nitrophenol (ONP) produced per minute per milligram of protein. Wild-type and mutant promoters were analyzed in triplicate at each time point using independent biological samples.

4.4.7 Expression and purification of Nla28-DBD

A fragment of the *nla28* gene corresponding to the Nla28 DNA Binding Domain (Nla28-DBD) [6, 22] was PCR amplified using gene-specific primers (**Table 4.2**) and then cloned into plasmid pMAL-c5x to create an N-terminal maltose binding protein (MBP) fusion to Nla28-DBD. The Nla28-DBD expression plasmid was introduced into *E. coli* strain BL21 (DE3) using electroporation. Cells containing Nla28-DBD expression plasmids were grown in rich LB broth to a density of approximately 2×10^8 cells/ml. Protein expression was induced by the addition of 0.3 mM Isopropyl β-D-1 thiogalactopyranoside (IPTG) to the culture

and the subsequent incubation of the culture for 12 hours at 15°C. Cells were pelleted via centrifugation and resuspended in 25 ml column buffer (20 mM Tris-HCl, 200 mM NaCl, 1mM EDTA, 5 U/ml DNase I, 1 mM DTT, pH 7.4) per liter of culture. The resuspended cells were lysed by a combination of freeze-thawing and sonication and pelleted by centrifugation. The crude extract (supernatant) containing Nla28-DBD was diluted by adding 125 ml of cold column buffer to every 25 ml aliquot of crude extract. The diluted crude extract containing Nla28-DBD was loaded onto 5 ml MBPTrap HP columns (GE Healthcare) at a flow rate of 5 ml/min and washed with 600 ml cold column buffer at a flow rate of 10 ml/min on an ÄKTA Fast Protein Liquid Chromatography (FPLC) system (GE Healthcare). Nla28-DBD was eluted using 100 ml cold column buffer containing 10 mM maltose; the flow rate was 5 ml/min and 20 fractions containing 5 ml were collected. The presence of eluted Nla28-DBD was detected by UV absorbance at 280 nm. Nla28-DBD-containing fractions were pooled and incubated with 1 mg of Factor Xa at 4°C overnight to cleave the MBP tag. Nla28-DBD was separated from MBP and concentrated to about 1 mg/ml using Amicon Ultra centrifugal filter units (EMD Millipore). SDS-PAGE and Bradford assays were used to determinate the purity and concentration of Nla28-DBD.

4.4.8 Electrophoretic mobility shift assays (EMSAs)

The PCR-generated fragments of the putative Nla28 target promoters contained approximately 180-220 bp of DNA upstream of the σ^{54}-RNA polymerase binding sites, which were identified experimentally [6, 19, 34, 35] or using a bioinformatics tool (PromScan) that was specifically developed to find

such sites in the sequences of bacterial DNA [36]. For use in electrophoretic mobility shift assays, the promoter fragments were PCR amplified using 5'Cy5-labeled primers synthesized by IDT (Table 4.2). Binding reactions contained EMSA buffer (25 mM Tris/acetate [pH 8.0], 8.0 mM magnesium acetate, 10 mM KCl, 1.0 mM DTT), 2.0 mM, 1.5 mM, 1.0 mM, 0.75 mM, 0.5 mM, or 0.25 mM of Nla28-DBD and 5 ng of a 5' Cy5-labeled promoter fragment. The binding reactions were allowed to proceed for 30 minutes at 30°C, and EMSAs were performed under 100V for 60 minutes using native (non-denaturing) PAGE (3% to 12%, Bis-Tris, Invitrogen). The binding reactions were imaged and analyzed by a Bio-Rad imager.

4.4.9 Quantitative PCR (qPCR)

To examine expression of Nla28 target genes, wild-type cells and cells with an inactivated *nla28* gene were harvested during vegetative growth (0 hours) and during development in submerged cultures. Total cellular RNA was isolated from developmental cells using the RNAprotect Bacteria Reagent (Qiagen) and the RNeasy Mini Kit (Qiagen) as described in the manufacturer's protocols. To help lyse developmental cells, 0.1 mm diameter glass beads were added after the lysis buffer and the cell suspensions were subjected to vigorous shaking using a VWR DVX-2500 multi-tube vortexer. For each time point (0, 1, 2, 8, 12 or 24 hours), total RNA was isolated from seven independent biological replicates of the wild-type strain and from seven independent biological replicates of the *nla28* mutant. The wild-type RNA samples from each time point were pooled and the *nla28* mutant RNA samples from each time point were pooled,

and the pooled samples were subsequently used to generate cDNA as previously described [19, 37, 38]. The CFX Connect real-time PCR detection system (Bio-Rad) was used to perform the qPCR analysis [19]. Each time point for the wild-type and *nla28* mutant was analyzed in triplicate (*ie*., we analyzed three technical replicates of each pooled sample). Relative fold-changes in mRNA levels were calculated using the reference gene *rpoD* or 16S rRNA and the $\Delta\Delta$CT method as previously described [19, 37].

4.4.10 Identifying potential targets of Nla28

Our search for direct targets of Nla28 started with a set of *M*. *xanthus* genes that showed substantial changes in expression during development; between any two time points in the developmental time course, expression of the genes increased at least 2-fold [our DNA microarray data on Gene Expression Omnibus, accession # **GSE13523**] [6]. We then searched for genes that showed at least a 2-fold decrease in expression when comparing any developmental time point in the *nla28* mutant and wild-type cells [6]. Fifty-one genes with substantial developmental regulation met this criterion and hence were classified as Nla28-dependent. We should note that the genes in the *actB*, *nla6* and *nla28* operons were not included in this analysis, since previous work showed that certain regions of the *act* and *nla* promoters are bound by Nla28-DBD; these promoters were already known to be good candidates for direct in vivo regulation by Nla28 [6]. These promoters were, however, included in subsequent studies to identify the Nla28 promoter binding sites.

The DNA region upstream of the 51 Nla28-dependent genes were scanned for putative σ^{54} promoters using the algorithm developed for PromScan, a bioinformatics tool that identifies potential σ^{54}-RNA polymerase binding sites using conserved nucleotides in the -12 bp and -24 bp regions of σ^{54} promoters [36]. Because of the distant location (in the primary DNA structure) of the σ^{54} promoter elements of the *nla6* and *nla28* operons, we searched for putative σ^{54} promoters in the 500-bp region upstream of each Nla28-dependent developmental gene. Intergenic and intragenic regions were evaluated since previous mutational analyses placed the putative σ^{54}-RNA polymerase binding sites in the *actB* and *nla6* promoters in intragenic regions and the putative σ^{54}-RNA polymerase binding site in the *nla28* promoter in an intergenic region (**Fig. 4.1**). It is notable that previous PromScan trial runs with known σ^{54} promoters and non-σ^{54} promoters yielded a false positive rate of about 4% and a false negative rate of about 19% [6]. Therefore, we anticipated similar false assignments in our search.

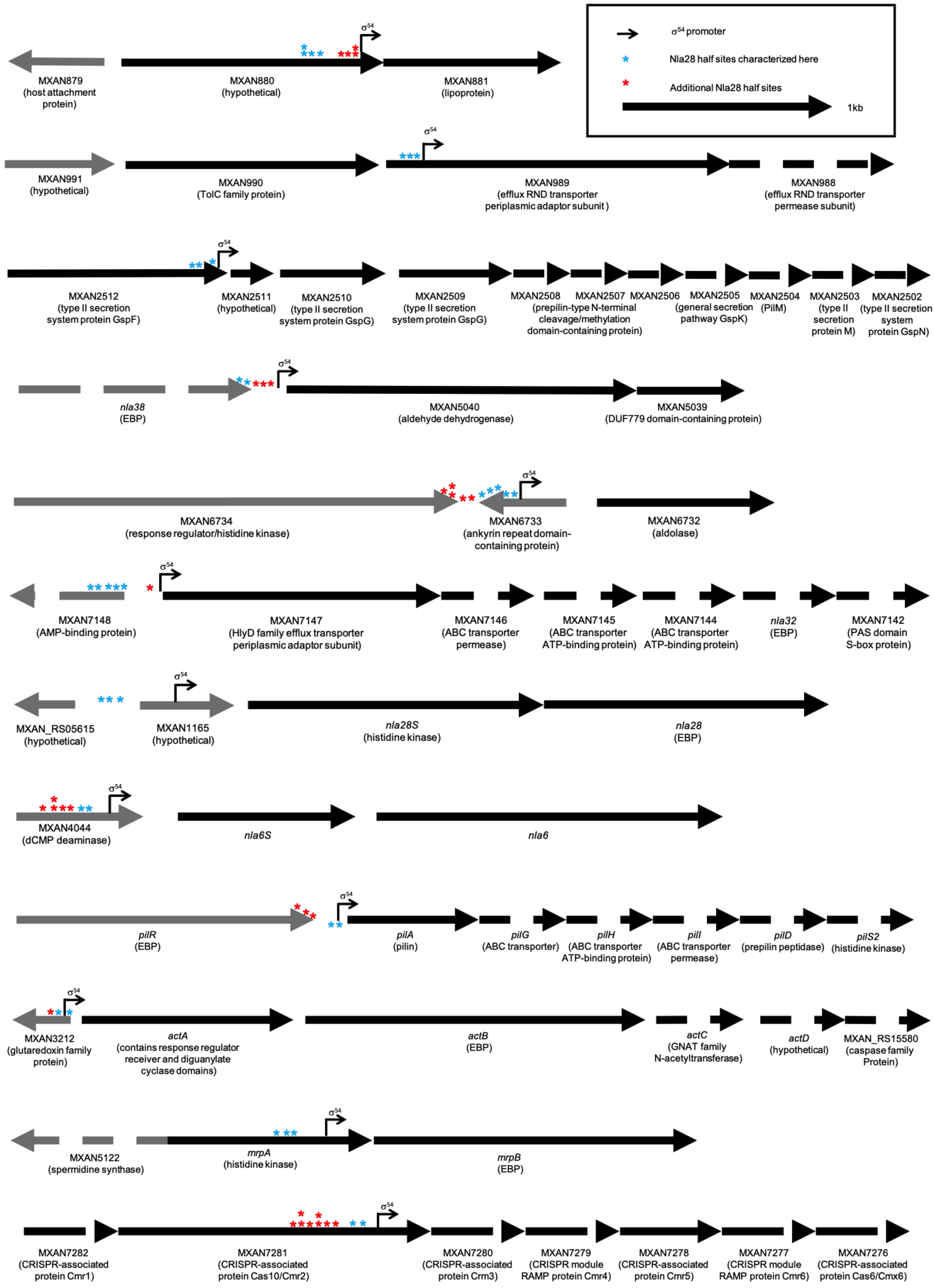

σ54 promoter
Nla28 half sites characterized here
Additional Nla28 half sites
1kb
MXAN879 (host attachment protein)
MXAN880 (hypothetical)
MXAN881 (lipoprotein)
MXAN991 (hypothetical)
MXAN990 (TolC family protein)
MXAN989 (efflux RND transporter periplasmic adaptor subunit)
MXAN988 (efflux RND transporter permease subunit)
MXAN2512 (type II secretion system protein GspF)
MXAN2511 (hypothetical)
MXAN2510 (type II secretion system protein GspG)
MXAN2509 (type II secretion system protein GspG)
MXAN2508 (prepilin-type N-terminal cleavage/methylation domain-containing protein)
MXAN2507
MXAN2506
MXAN2505 (general secretion pathway GspK)
MXAN2504 (PilM)
MXAN2503 (type II secretion protein M)
MXAN2502 (type II secretion system protein GspN)
nla38 (EBP)
MXAN5040 (aldehyde dehydrogenase)
MXAN5039 (DUF779 domain-containing protein)
MXAN6734 (response regulator/histidine kinase)
MXAN6733 (ankyrin repeat domain-containing protein)
MXAN6732 (aldolase)
MXAN7148 (AMP-binding protein)
MXAN7147 (HlyD family efflux transporter periplasmic adaptor subunit)
MXAN7146 (ABC transporter permease)
MXAN7145 (ABC transporter ATP-binding protein)
MXAN7144 (ABC transporter ATP-binding protein)
nla32 (EBP)
MXAN7142 (PAS domain S-box protein)
MXAN_RS05615 (hypothetical)
MXAN1165 (hypothetical)
nla28S (histidine kinase)
nla28 (EBP)
MXAN4044 (dCMP deaminase)
nla6S
nla6
pilR (EBP)
pilA (pilin)
pilG (ABC transporter)
pilH (ABC transporter ATP-binding protein)
pilI (ABC transporter permease)
pilD (prepilin peptidase)
pilS2 (histidine kinase)
MXAN3212 (glutaredoxin family protein)
actA (contains response regulator receiver and diguanylate cyclase domains)
actB (EBP)
actC (GNAT family N-acetyltransferase)
actD (hypothetical)
MXAN_RS15580 (caspase family Protein)
MXAN5122 (spermidine synthase)
mrpA (histidine kinase)
mrpB (EBP)
MXAN7282 (CRISPR-associated protein Cmr1)
MXAN7281 (CRISPR-associated protein Cas10/Cmr2)
MXAN7280 (CRISPR-associated protein Crm3)
MXAN7279 (CRISPR module RAMP protein Cmr4)
MXAN7278 (CRISPR-associated protein Cmr5)
MXAN7277 (CRISPR module RAMP protein Cmr6)
MXAN7276 (CRISPR-associated protein Cas6/Cmx6)

Figure 4.1 Gene arrow models depicting locations of characterized and additional putative Nla28 half binding sites in confirmed targets of Nla28 in *M. xanthus*.

Genes are represented by black (Nla28-targeted locus) or gray (other locus) arrows facing the direction of transcription. Dashed arrows represent extremely large genes that are unable to show full length. Elbow arrows represent σ^{54} promoters. Blue asterisks indicate characterized Nla28 half binding sites. Red asterisks indicate additional putative Nla28 half binding sites. Blue horizontal bar indicates the size in nucleotides of 1kb.

Of the 51 Nla28-dependent genes analyzed as described above, only nine had a putative σ^{54}-RNA polymerase binding site in the 500-bp region. These nine promoters, which includes the previously characterized promoter in the *pilA* locus [34], and the promoters of the *actB*, *nla6* and *nla28* operons (**Table 4.3**) were all considered candidates for direct in vivo regulation by Nla28.

Table 4.3 Potential developmental targets of the EBP Nla28

Locus	Genes	Notable functions or putative functions of protein product(s)	Activation time of development[a]	Putative σ^{54}-RNA polymerase binding site (-12 and -24 promoter regions)[b]	Putative Nla28 half binding sites[d]
actB	4	ActB EBP and ActA signal transduction protein	4 hours	TGGCACA-N4-TTGCT[c]	CGCCGCGG-N54-CTGCCCAG
nla6	2	Nla6S histidine kinase and Nla6 EBP	1 hour	TGGTGCG-N4-GTGTT[c]	CTGCGTGG-N9-ACGCGGAG
nla28	2	Nla28S histidine kinase and Nal28 EBP	1 hour	TGGAGCG-N4-CTGCT[c]	GTGGGGAG-N3-CTCCGCAG-N25-CTCCCCAG
mrpB	1	MrpB EBP	1 hour	TGGCCCA-N4-CTGCT	CGGTGCAG-N47-CTCCTCGG-N9-CGGTGCAG
pilA	6	Type IV pili biogenesis	-	TGGCATG-N4-GTGCT	GTGCGCAC-N10-ATGCGTAG

MXAN881	1	Putative lipoprotein	1 hour	TGGCGTG-N4-GTGCT	GTCCGGCG[e]-N10-CTCCCGAT-N37- CTCAGAAG or CGCCGGAT[e]-N4-CTCCCGAT-N37-CTCAGAAG
MXAN989	2	Copper/silver cation efflux proteins (CusA and CusB)	1 hour	CCGCACG-N4-TTGCT	CAGCGGCG-N6-CGCCGCTG-N10-CCCTGGAG
MXAN2511	10	Type II protein secretion systems	1 hour	GGGCGCG-N4-TTGCA	CTCCGGCC-N14-CTCACCAG-N67-CTCCTCAT
MXAN5040	2	Aldehyde dehydrogenase	-	TGGCACG-N4-CTGCT	TGGCGCAG-N12-CTGGGCAT
MXAN6732	1	Class II aldolase	12 hours	CGGCATG-N4-TTGCG	CTCAGCCG-N36-GTCCGCCG-N17-TTGCGCAG-N19-GTCAGCAG or CTCCGCCA-N28-GTCCGCCG-N17-TTGCGCAG-N19-GTCAGCAG
MXAN7147	2	RND family efflux transporter and ABC-type efflux transport protein	1 hour	TGGCACG-N4-TCGCT	CTCCCCCG-N8-CCGCGCTG-N41-GTCCGTAG-N21-GTCAGCAG-N6-CTCGGGCG
MXAN7280	5	Putative CRISPR-associated proteins	1hour	TGGCACC-N4-TCGCG	CTGTGGAC-N52-CTGTCGAG

[a] The indicated times are based on the qPCR data presented here or on previously published microarray data [6].

[b] The bold and underlined nucleotides match the nucleotides in the consensus σ^{54}-RNA polymerase binding sequence, which is TGGCACG-4N-TTGC(T/A). N4 indicates any four nucleotides.

[c] The -12 and -24 regions of the putative σ^{54} promoter have been characterized via mutational analysis.

[d] The bold and underlined nucleotides match the nucleotides in the Nla28 consensus half binding site [CT(C/G)CG(C/G)AG]. N indicates the number of nucleotides in the spacer region between half sites.

[e] Indicates overlapping sites found in the same promoter region.

Electrophoretic mobility shift assays were subsequently used to examine whether Nla28-DBD was capable of binding to at least one DNA fragment in the nine newly identified promoter regions; the fragments corresponded to DNA flanking the putative σ^{54}-RNA polymerase binding sites in the -12 and -24 bp-regions. Nla28-DBD was consistently positive for binding to at least one fragment of each promoter region. All fragments that were positive for Nla28-DBD binding contained at least two 8-bp sequences that were similar to one another and similar to 8-bp sequences in other promoter fragments (**Table 4.3**). Furthermore, these sequences were similar to 8-bp sequences in *actB*, *nla6* and *nla28* promoter fragments that were positive for Nla28-DBD binding [6]. Since dimers of other EBPs are known to bind to tandem repeat sequences [23, 24], we speculated that at least two of the direct, imperfect repeats that we identified in each promoter fragment serve as sites for in vivo Nla28 dimer binding. The consensus repeat sequence and presumed half binding site that we identified for Nla28 was CT(C/G)CG(C/G)AG.

Once we had a potential direct repeat binding sequence for Nla28 in hand, we examined whether the 300-bp of DNA located upstream of the putative -12 and -24-bp promoter regions contained additional sequences closely matching the consensus half binding site of Nla28. Additional (putative) half binding sites for Nla28 were found in many of the promoter regions (**Fig. 4.1**).

In a subsequent search for additional Nla28 targets, the *M. xanthus* genome sequence [16] was scanned for close matches to the consensus Nla28 half binding site. In particular, we looked for 8-bp sequences with no more than

two mismatches relative to the consensus Nla28 half binding site. Since EBP dimers bind to DNA, we also looked for potential Nla28 half sites arranged in tandem; tandem sites were defined as putative Nla28 half sites separated by no more than 25 bp, which was the mean number of base pairs between putative Nla28 half sites in the 12 promoter regions we characterized. We decided to use this relatively conservative upper limit on the putative spacer region between repeats/putative half binding sites because all of Nla28 binding sites that we had characterized, with the exception of the putative binding site in the *actB* promoter, had shorter spacer regions [22]. Therefore, we felt more confident that the tandem repeats we selected might be Nla28 binding sites.

We then searched for putative σ^{54}-RNA polymerase binding sites located within 500-bp downstream of the tandem sequences. Since we were primarily interested in Nla28's developmental gene targets and also developmental genes whose expression shows a strong dependence on Nla28, we further curated the list of potential Nla28 targets by doing the following: 1) we identified single genes and operons located downstream of the putative σ^{54}-RNA polymerase binding sites in the -12 and -24-bp regions; 2) using DNA microarray data [6], we determined which of the single genes and operons showed at least a 2-fold increase in expression during the development of wild-type cells; 3) using DNA microarray data [6], we also identified the single genes and operons that showed at least a 2-fold decrease in expression during the development of *nla28* mutant cells. The single genes and operons that met the above criteria are shown in Table 4.4.

Table 4.4 Additional potential targets of Nla28 and expression levels in WT/nla28 $^{-}$ cells

Locus	Numbers of Genes	Peak expression time point in WT cells (h)	Fold increase of expression at peak time point in WT cells	Fold decrease of expression in *nla28* $^{-}$ cells
MXAN162	1	12	2.5	8.3
MXAN179	3	24	6.1	3.7
MXAN255	1	12	2.2	4.2
MXAN419	2	12	2.1	5.6
MXAN496	2	12	3.1	3.7
MXAN542	2	12	2.3	3.7
MXAN562	1	1	2.3	3.6
MXAN854	1	6	4.1	3.0
MXAN859	1	12	3.1	2.5
MXAN909	2	12	2.4	3.3
MXAN934	1	6	2.6	3.7
MXAN1043	2	12	2.1	6.7
MXAN1124	1	12	2.4	2.4
MXAN1404	1	12	2.2	4.3
MXAN1412	1	6	3.1	3.6
MXAN1433	1	1	2.1	3.3
MXAN1752	2	6	2.7	3.3
MXAN1783	1	12	2.0	2.9
MXAN1968	1	1	3.2	1.7
MXAN2159	2	12	2.4	4.3
MXAN2359	2	24	2.4	4.8

MXAN2362	1	12	2.8	2.0
MXAN2363	1	12	2.8	2.4
MXAN2447	7	24	2.4	2.9
MXAN2515	5	24	3.1	2.9
MXAN2735	1	12	2.0	5.3
MXAN2752	2	12	2.3	2.6
MXAN2777	1	12	2.2	7.1
MXAN2907	1	12	2.3	2.6
MXAN2947	1	12	2.7	7.7
MXAN2952	5	6	2.2	1.8
MXAN3026	2	6	3.7	5.9
MXAN3149	2	6	9.0	5.3
MXAN3167	1	12	2.1	3.8
MXAN3735	2	12	2.1	2.2
MXAN3778	1	12	2.2	4.3
MXAN3824	1	12	2.6	2.7
MXAN3962	2	12	2.8	2.3
MXAN3975	1	12	2.6	3.1
MXAN3979	2	12	2.3	8.3
MXAN4007	1	1	2.9	2.3
MXAN4155	1	12	2.2	5.9
MXAN4198	4	18	2.0	3.8
MXAN4236	1	6	2.3	1.7
MXAN4687	1	1	7.5	4.5
MXAN4784	2	6	2.0	2.8
MXAN4877	1	2	2.6	2.1
MXAN5543	6	6	4.3	2.3
MXAN5615	1	6	2.1	5.0
MXAN5715	1	6	3.3	100.0
MXAN6209	1	1	6.1	5.3
MXAN6223	5	6	4.3	2.1
MXAN6297	1	12	2.5	4.0
MXAN6437	1	6	2.6	2.2

MXAN6788	1	6	3.2	33.3
MXAN7157	1	18	2.0	2.9
MXAN7225	2	12	2.4	5.0
MXAN7435	1	1	4.6	2.0

4.4.11 Statistical analysis

Individual sample sizes are specified in each figure legend. Comparisons between groups were assessed by two-way analysis of variance (ANOVA) with Tukey's multiple comparisons *post hoc* tests, as appropriate. The significance level was set at $p<0.05$ or lower. Prism (GraphPad) v9.2 was used for all analyses.

4.5 Results

4.5.1 Identifying a set of potential developmental promoter targets of Nla28

In our initial studies, we aimed to identify and characterize a collection of developmental regulated promoters that are likely to be directly regulated by Nla28 in vivo. As described in the *Materials and Methods*, we searched for known or putative σ^{54} promoters upstream of genes that showed substantial developmental regulation and at least 2-fold downregulation in a *nla28* mutant. Using this strategy, nine developmentally regulated promoters were identified and classified as potential targets of Nla28 (**Table 4.3 and Fig. 4.1**). Of these promoters, only *pilA*'s had been previously characterized; a potential σ^{54} promoter was identified via primer extension analysis [34].

To confirm the DNA microarray data, which indicated that the nine newly identified promoters are Nla28 dependent in vivo, we used quantitative PCR (qPCR). In particular, we examined whether developmental expression of one gene downstream of each promoter is downregulated in a *nla28* mutant relative to the wild-type strain; we compared the mRNA levels of the genes in the *nla28* mutant and wild-type cells at various developmental time points. The developmental expression patterns of the genes in wild-type and *nla28* mutant cells are shown in **Fig. 4.2** and a summary of how the *nla28* mutation impacts their developmental expression is summarized in **Table 4.5**.

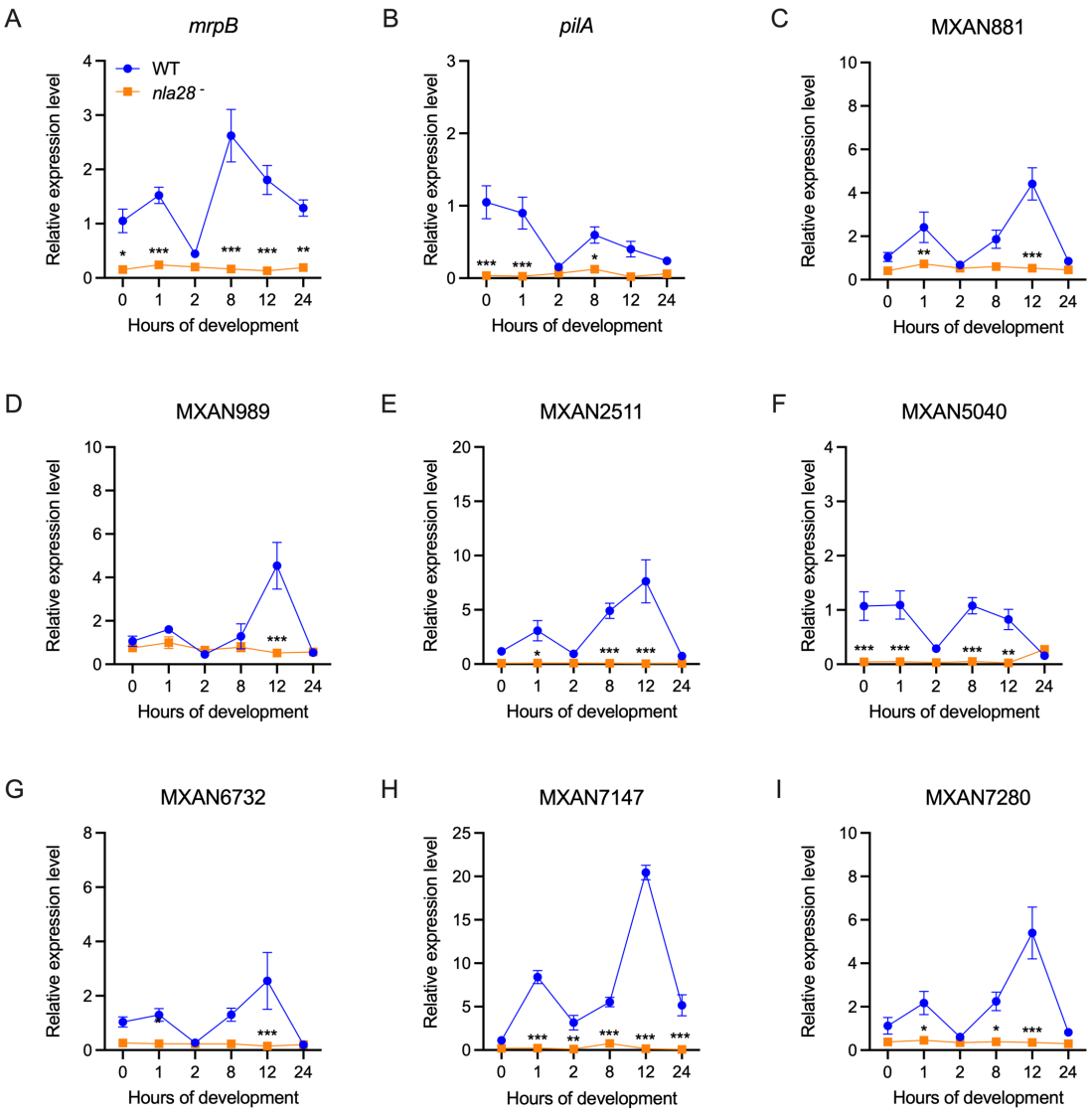


Figure 4.2 Quantitative real-time PCR measurement of the developmental expression levels and patterns of Nla28 target genes in wild-type and *nla28* mutant strains.

The developmental mRNA levels of nine Nla28 target genes in wild-type (WT) and *nla28* mutant (*nla28* $^{-}$) strains were determined using qPCR. Wild-type and *nla28* mutant cells were harvested at 0, 1, 2, 8, 12 and 24 hours of development for RNA isolation and qPCR analysis. N = 3 technical replicates of pooled RNA samples at each time point. Error bars are standard deviations of the means. The data was analyzed using two-way analysis of variance (ANOVA) and Tukey's

multiple comparisons *post hoc* tests; ***p < 0.001; **p < 0.01; *p < 0.05 for mRNA levels in *nla28* mutant versus wild-type cells.

Table 4.5 Developmental expression levels of putative Nla28 target genes in *nla28*$^-$ mutant.

Gene	Time of peak expression in WT	Fold decrease in *nla28*$^-$ relative to WT		
		At the time of peak expression in WT	Largest	Mean
mrpB	8 hours	15.8***	15.8***	8.6***
pilA	0 hours	29.6***	33.4***	15.4***
MXAN881	12 hours	8.4***	8.4***	3.4***
MXAN989	12 hours	8.7***	8.7***	2.5***
MXAN2511	12 hours	187.0***	187.0***	54.8***
MXAN5040	0 hours	22.8***	32.8**	18.2***
MXAN6732	12 hours	17.2***	17.2***	5.7***
MXAN7147	12 hours	125.2***	125.2***	48.0***
MXAN7280	12 hours	15.4***	15.4***	5.6***

** represents a significant difference ($p<0.01$); *** represents a significant difference ($p<0.001$)

Seven of the mRNAs that we examined showed significant increases during the development of wild-type cells; peak developmental expression levels ranged from about 2.5 to 20.5-fold higher than the baseline vegetative (0 hours) levels (**Fig. 4.2A, C-E and G-I**). In contrast, the levels of *pilA* mRNA did not rise above the vegetative baseline during the wild-type developmental time course (**Fig. 4.2B**). However, we should note that our qPCR analysis lacked the

developmental time point at which *pilA* mRNA levels peaked in the microarray studies [6]. Perhaps this resulted the qPCR's failure to detect a burst of *pilA* mRNA in developing cells. The qPCR analysis of wild-type cells also failed to detect an increase in MXAN5040 mRNA above the vegetative baseline (**Fig. 4.2F**), a finding that agreed with the results of the microarray analysis [6]. This led us to examine the MXAN5040 mRNA expression pattern more closely, which revealed why MXAN5040 was classified as a locus that shows a substantial increase in expression during development. In particular, a locus was defined as such when it showed a 2-fold expression increase between any two time points in the developmental time course; MXAN5040 mRNA levels increased more than 2-fold between 2 and 8 hours of development but never rose above vegetative levels due to the sharp decrease in expression between 1 and 2 hours (**Fig. 4.2F)**.

In the *nla28* mutant, the vegetive baseline levels of some mRNAs were reduced (**Fig. 4.2A-I**), suggesting that Nla28 is important for expression of these mRNAs during vegetative growth and supporting the previous assertion that Nla28 might be generally involved in starvation-induced gene regulation rather than fruiting body-specific regulation [22]. As for the mRNA levels in developing *nla28* mutant cells, expression was substantially lower at the wild-type peak and at least one time point but typically at multiple time points in development (**Table 4.5 and Fig. 4.2A-I**). These findings indicate that Nla28 is important for the normal expression of the mRNAs during development.

4.5.2 Predicted locations of the putative Nla28 target promoters

The locations and sequences of the putative σ^{54}-RNA polymerase binding sites in the -12- and -24-bp regions of the 9 newly identified promoters are shown in **Fig. 4.1** and **Table 4.3**, respectively. The locations of the putative -12- and -24-bp regions of the *actB*, *nla6* and *nla28* promoters are also shown, since previous data indicated that Nla28 binds to these promoters and they are likely to be σ^{54}-type promoters [6, 19]. It is noteworthy that the 9/12 of the putative σ^{54}-RNA polymerase binding sites or core σ^{54} promoter regions are located within genes and not in intergenic sequences. This finding is consistent with previous studies that placed many core σ^{54} promoter regions in the coding sequences of *M. xanthus* genes [6, 16, 19, 22] and raises the possibility that intragenic σ^{54} promoters might be common in *M. xanthus* and in bacteria in general. As in previous analyses of *M. xanthus* σ^{54} promoters, some of the intragenic -12- and -24-bp regions are located in the protein coding sequence of an upstream gene (upstream promoters) and some are located in the protein coding sequence of one gene in an operon (internal promoters) [22]. The implications of these findings are addressed in the *Discussion*.

4.5.3 Promoter fragments that are positive for in vitro Nla28-DBD binding have similar 8-bp repeats

In our initial electrophoretic mobility shift assays (EMSAs), we examined whether purified Nla28-DBD (Nla28 DNA binding domain) is capable of binding to the 9 newly identified candidates for direct Nla28 regulation; DNA fragments

flanking the putative -12- and -24-bp promoter regions were used in the assays. At least one fragment of each promoter region was consistently positive for Nla28-DBD binding. All the promoter fragments except the negative control *dev* promoter fragment, which is a fragment of a non-σ^{54} promoter, produced at least one shifted complex (**Fig. 4.3A**). The positive control, which is a fragment of *nla28* promoter, also produced a shifted complex as predicted (**Fig. 4.3A**).

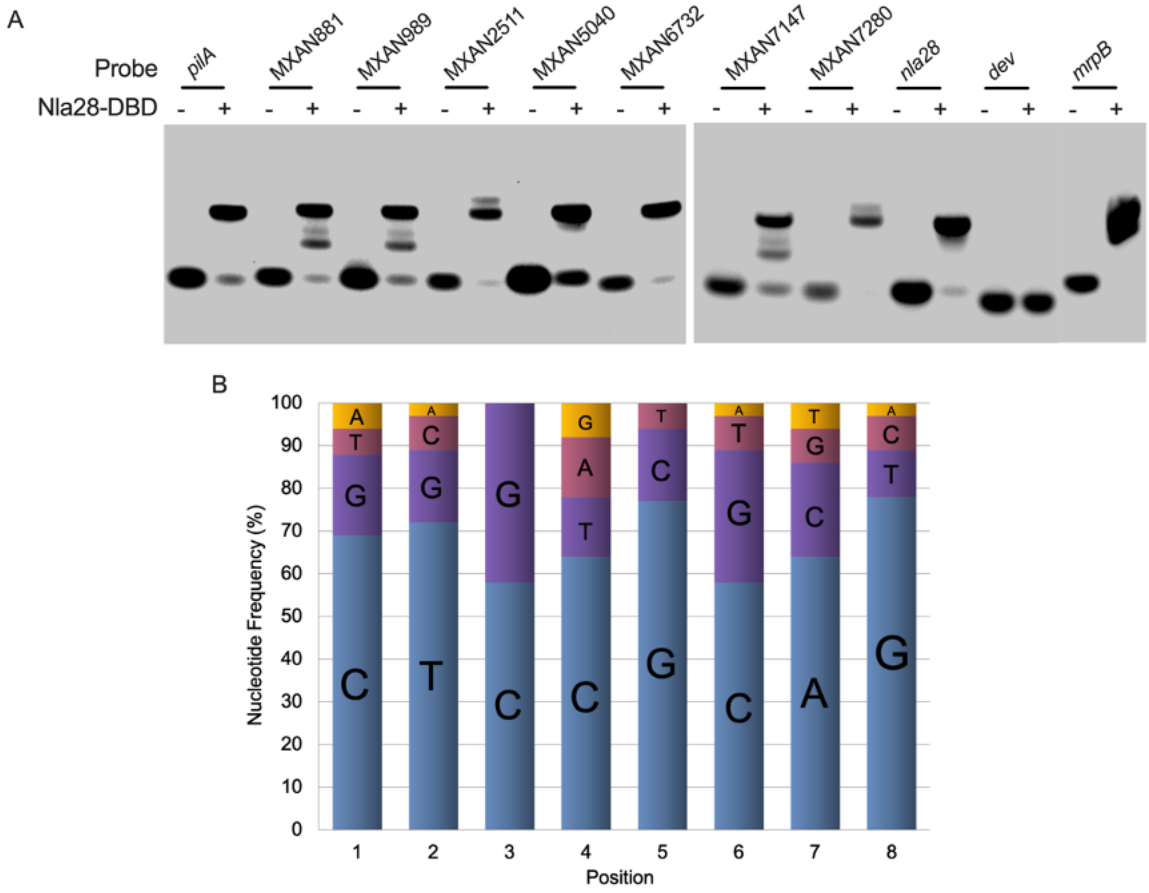


Figure 4.3 EMSAs with Nla28-DBD and fragments of putative target promoters

(A) Purified Nla28-DBD binds to fragments of the 9 newly identified target promoters. EMSAs performed with purified Nla28-DBD and an *pilA*, MXAN881,

MXAN989, MXAN2511, MXAN5040, MXAN6732, MXAN7147, MXAN7280, *nla28* (positive control), *dev* (negative control) or *mrpB* promoter fragment containing putative Nla28 binding sites. Binding reactions were performed with (+) or without (−) 2 μM of purified Nla28-DBD and 5 ng of Cy5 5' end-labeled promoter fragments. Similar 8-bp sequences (putative Nla28 half binding sites) were found in all promoter fragments that were positive for Nla28 binding. (B) Nucleotide frequency at each position of the 8-bp sequence. The consensus 8-bp sequence or consensus Nla28 half binding site derived from the frequency data is CT(C/G)CG(C/G)AG.

When we searched the promoter fragments that were positive for Nla28-DBD binding, we identified similar 8-bp sequences (**Table 4.3**). Each promoter fragment tested here contained at least two, but often 3-5 of these 8-bp sequences (**Table 4.3**), as did the *actB*, *nla6*, and *nla28* promoter fragments that were previously shown to be positive for Nla28-DBD binding [6, 19]. Since EBP dimers typically bind to tandem repeat sequences or tandem half binding sites [23, 24], we speculated that at least two of the 8-bp sequences in each promoter fragment serve as Nla28-DBD half binding sites. We also speculated that Nla28-DBD might be able to form different binding complexes when the promoter fragments have more than one half site pair; this might explain why promoter fragments such as MXAN881, MXAN989 and MXAN7147 yielded more than one shifted complex in EMSAs (**Fig. 4.3A**). We should note that many of the promoter regions contain a second putative cluster of Nla28 half sites (**Fig. 4.1**); we identified these sites using the consensus Nla28 half binding site 5'-CT(C/G)CG(C/G)AG-3' (**Fig. 4.3B**).

4.5.4 Two 8-bp sequences in the actB promoter and in the nla28 promoter are crucial for in vitro Nla28-DBD binding

Since all the promoter fragments that were positive for Nla28-DBD binding in EMSAs contained similar 8-bp sequences, we asked if these sequences are important for Nla28-DBD binding. We initially focused on the *actB* promoter fragment, which has only two 8-bp sequences (i.e., two putative half sites) for the presumed binding of a Nla28-DBD dimer. Three *actB* promoter fragments were generated for the in vitro Nla28-DBD binding analysis. One of the *actB* promoter fragments had wild-type half sites 1 and 2, one fragment had a wild-type half site 1 and a half site 2 that was converted to all A nucleotides, and one fragment had wild-type half site 2 and a half site 1 that was converted to all A nucleotides (**Fig. 4.4A**). EMSAs were performed with 2 μM of Nla28-DBD and 5' Cy5-labeled wild-type or mutant *actB* promoter fragments. Nla28-DBD was capable of binding to the *actB* promoter fragment carrying two wild-type half sites as in previous EMSAs. However, no binding was detected when either half site 1 or half site 2 contained all A nucleotides (**Fig. 4.4B**). This finding is consistent with our proposal that the two 8-bp sequences serve as half binding sites for Nla28-DBD dimers and that both sites are important for Nla28-DBD binding.

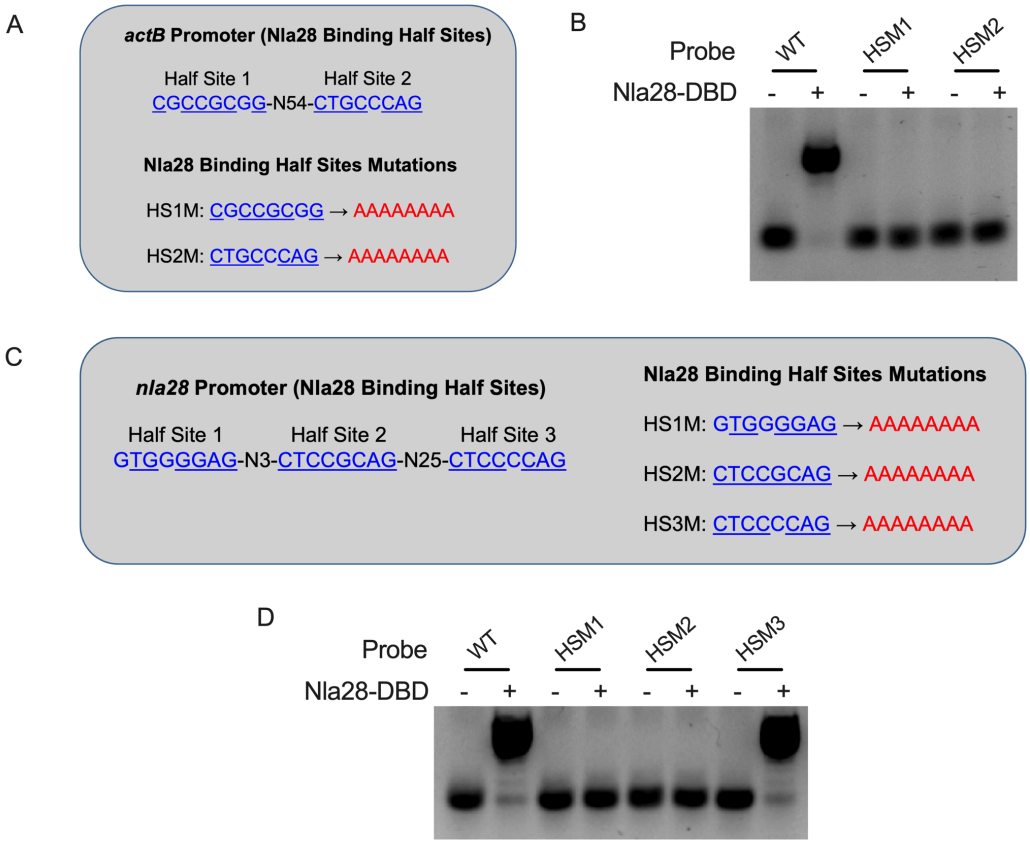


Figure 4.4 EMSAs with Nla28-DBD and wild-type or mutant fragments of the *actB* and *nla28* promoters

Two 8-bp sequences (Half Site 1 and 2, shown in blue) that closely match the consensus Nla28 half binding were identified in the wild-type *actB* promoter fragment. Underlined sequences represent nucleotides that match the consensus Nla28 half binding site. Mutations (HS1M and HS2M, shown in red) in these putative Nla28 half binding sites were generated for in vitro Nla28-DBD binding analysis. (B) EMSAs were performed with (+) or without (−) 2 μM of purified Nla28-DBD and 5 ng Cy5 5' end-labeled *actB* promoter fragment containing two wild-type (WT) Nla28 half binding sites, a mutated half site one (HS1M) or a mutated half site two (HS2M). (C) Three 8-bp sequences (Half Site 1, 2 and 3, shown in blue) that closely match the consensus Nla28 half binding were identified in the wild-type *nla28* promoter fragment. Underlined sequences represent nucleotides that match the consensus Nla28 half binding site. Mutations (HS1M, HS2M and HS3M, shown in red) in these putative Nla28 half

binding sites were generated for in vitro Nla28-DBD binding analysis. (D) EMSAs were performed with (+) or without (−) 2 μM of purified Nla28-DBD and 5 ng Cy5 5' end-labeled *nla2B* promoter fragment containing three wild-type (WT) Nla28 half binding sites, a mutated half site one (HS1M), a mutated half site two (HS2M) or a mutated half site three (HS3M).

In the next experiment we examined Nla28-DBD binding to the *nla28* promoter fragment, which contains three 8-bp sequences or potential Nla28-DBD half binding sites. Four *nla28* promoter fragments were generated for the in vitro binding assays. One of the *nla28* promoter fragments contained wild-type half sites 1-3 and each of the remaining *nla28* promoter fragments contained two wild-type half sites and one half site converted to all A nucleotides (**Fig. 4.4C**). EMSAs were performed with 2 μM of Nla28-DBD and 5' Cy5-labeled wild-type or mutant *nla28* promoter fragments. Nla28-DBD binding was detected when the *nla28* promoter fragment contained all wild-type half sites, as previously noted. Nla28-DBD binding was also detected when the *nla28* promoter fragment contained wild-type half sites 1 and 2, and a half site 3 that was converted to all A nucleotides. However, no Nla28-DBD binding was detected when half site 1 or a half site 2 was converted to all A nucleotides, even though the remaining two half sites were wild type (**Fig. 4.4D**). Thus, it seems that half sites 1 and 2, but not half site 3 in the *nla28* promoter are crucial for in vitro Nla28-DBD binding. This is not the predicted result based solely on a putative half site's similarity to the consensus, as half sites 1, 2 and 3 have 2, 0 and 1 mismatch(es), respectively, relative to the consensus Nla28 half binding site (**Fig. 4.4C and Table 4.3**). Perhaps some nucleotide positions in a half site are more important

for Nla28-DBD binding than others and/or the spacing between one half site and its neighboring half site is important for the efficient binding of Nla28-DBD in vitro. Of course, in an in vivo setting it is likely that Nla28's confirmation and oligomerization states, which are presumably influenced by Nla28S-mediated phosphorylation [20], and perhaps the intrinsic curvature of the promoter region would also influence binding.

4.5.5 Mutations in putative Nla28 binding sites substantially reduce promoter activity in developing cells

The in vitro binding and in vivo expression data presented here and in previous work [6] suggested that Nla28 might directly regulate the nine newly identified promoters, as well as the *actB*, *nla6* and *nla28* promoters. To further examine this idea, we selected four Nla28 target promoters for an in vivo mutational analyses. We started with fragments of the *nla28*, *actB*, *pilA* and MXAN5040 promoter regions; each fragment contained the putative σ^{54}-RNA polymerase and Nla28 binding sites. Next, fragments with a/an 2-bp substitution, 4-bp substitution, 6-bp substitution or 8-bp substitution in Nla28 half site 1 or 2 were generated (**Fig. 4.5A-H**); the native nucleotides were converted to A nucleotides in all cases.

Wild-type and mutant *nla28* promoter fragments were cloned into pREG1727, a plasmid that creates transcriptional fusions between cloned promoters and the *lacZ* gene [32]. The *lacZ* fusion plasmids were introduced into wild-type strain DK1622, and Kan^r isolates carrying a plasmid integrated at the Mx*8* phage attachment site in the *M*. *xanthus* chromosome were identified via

PCR and analyzed via DNA sequencing. Strains carrying a wild-type or mutant promoter fragment were allowed to develop for various amounts of time and β-Galactosidase assays were used to infer developmental promoter activities.

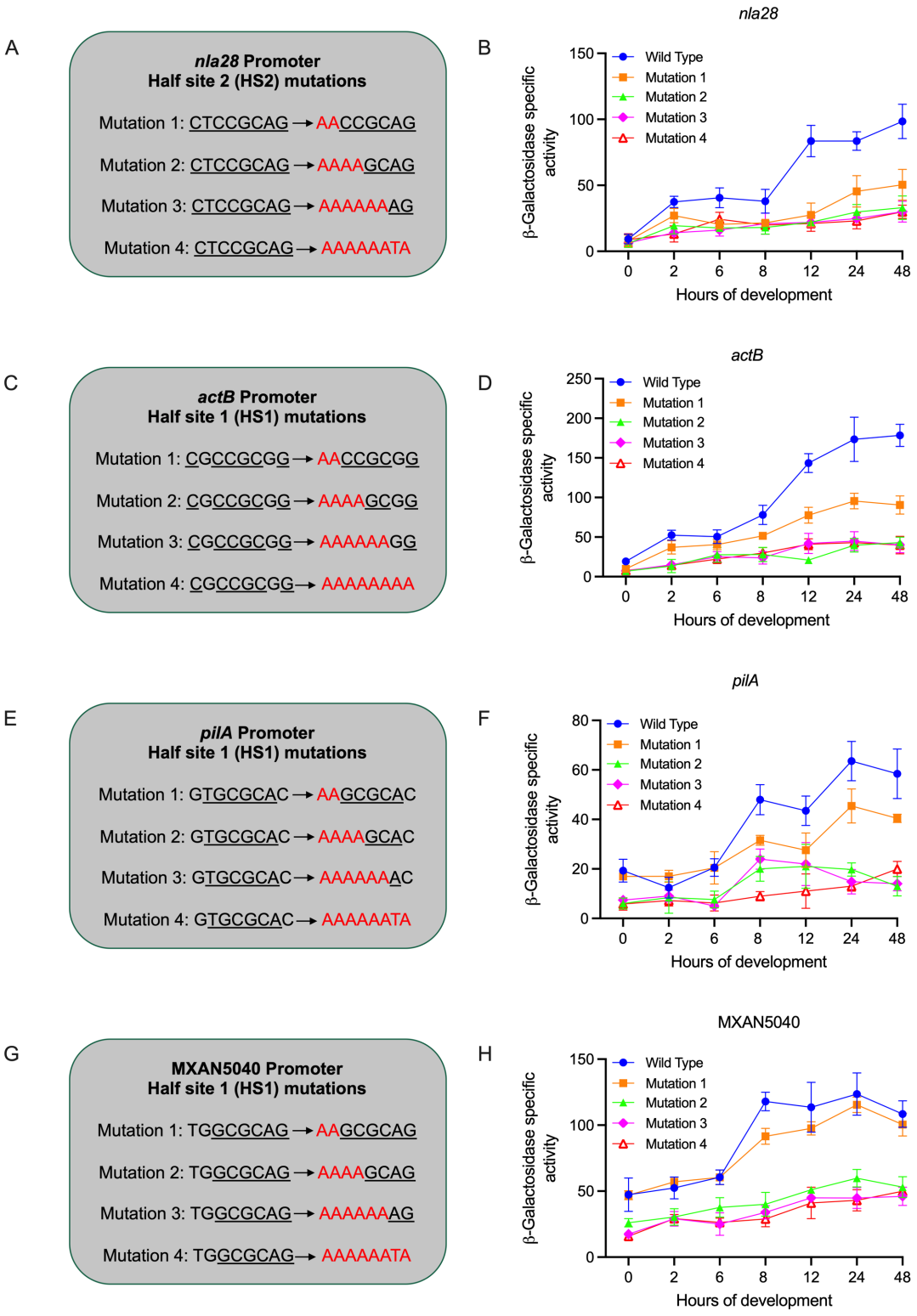


Figure 4.5 In vivo activities of wild-type and mutant promoter targets of Nla28

(A, C, E and G) *nla28*, *actB*, *pilA* and MXAN5040 promoter fragments with a 2-bp substitution (Mutation 1), 4-bp substitution (Mutation 2), 6-bp substitution (Mutation 3) or 8-bp substitution (Mutation 4) in Nla28 half site 2 were generated by site-directed mutagenesis. Substituted nucleotides are shown in red. (B, D, F and H) Wild-type and derivatives of the *nla28*, *actB*, *pilA* and MXAN5040 promoters carrying Mutation 1, Mutation 2, Mutation 3 or Mutation 4 were cloned into a *lacZ* expression vector and transferred to the wild-type *M. xanthus* strain DK1622. In vivo activities of wild-type and mutant promoters were inferred from β-galactosidase-specific activities (defined as nanomoles of ONP produced per minute per milligram of protein) at various time points (0, 2, 6, 8, 12, 24 and 48 h) during development. N = 3 biological replicates at each of the indicated time points. Error bars are standard deviations of the means.

As shown in **Fig. 4.5A-B,** a 4-bp, 6-bp or 8-bp substitution in putative Nla28 half site two substantially reduced the peak in vivo activity of the *nla28* promoter relative to that of the wild-type. Similar results were observed with the *actB*, *pilA* and MXAN5040 promoter fragments that contained a 4-bp, 6-bp or 8-bp substitution in putative Nla28 half site one, as shown in **Fig. 4.5C-H**. We also observed substantially reduced peak activities when the *nla28* and *actB* promoters contained a 2-bp substitution in putative Nla28 half site two and one, respectively (**Fig. 4.5A-D**). In contrast, a 2-bp substitution had only a slight impact on the peak activity of the *pilA* promoter (**Fig. 4.5E-F**) and no impact on the peak activity of the MXAN5040 promoter (**Fig. 4.5G-H**). In the case of the MXAN5040 promoter, it is perhaps not surprising that the two native nucleotides are not crucial for in vivo activity because they do not match the nucleotides in the consensus Nla28 half binding site. Regardless, the results of our in vivo and in vitro assays provide strong evidence that the 8-bp sequences that we

identified are important for Nla28 binding and for the in vivo activities of Nla28 target promoters.

4.5.6 Nla28 target genes are important for development

The results of our in vitro and in vivo expression studies indicate that the loci shown in Table 4.3 are likely to be in vivo targets of Nla28 during development. Four of the loci (*actB*, *nla6*, *nla28* and *mrpB*) contain characterized EBP genes and the corresponding EBPs are known to be key components in *M. xanthus* developmental regulatory pathways [15, 19, 29, 31, 39]. Another locus that has been characterized is *pilA*; this locus is important for production of type IV pili, social motility and normal development [30, 40, 41].

To examine whether the remaining Nla28 targets listed in Table 4.3 are important for development, we generated an insertion in one gene in each locus. Namely, we generated an insertion in the single gene or in one gene in the operon located downstream of the putative core σ^{54} promoter element. Subsequently, we determined whether the insertions affected the formation of aggregates of cells and/or sporulation (**Table 4.6**). Indeed, all of the insertions affected the timing of aggregate formation and some also affected the final shape of aggregates; however, none of the insertions completely blocked the formation of aggregates. In addition, the insertions substantially reduced the number of sonication- and heat-resistant spores that were able to germinate into colonies, as indicated by the viable spore numbers. Interestingly, the insertions had relatively minor impacts on spore numbers; spore number refers to the number of developing cells that were able make the shape change (rod to spherical shape)

associated with the early differentiation stage of sporulation (**Table 4.6**). Thus, it seems that the insertions had a stronger impact on the maturation of spores into stress resistant cells and/or on spore germination than on spore differentiation.

Table 4.6 Developmental phenotypes of wild-type and mutant cells

Strain[a]	Aggregation[b]	Fruiting body sporulation[c]	
		Spore number (%)	Viable spore number (%)
DK1622	+	100±12.6	100±8.5
nla28 [-]	+/-	108.1±7.6	2.8±1.2
MXAN881 [-]	+/-	72.2±0.7	0.5±0.5
MXAN989 [-]	+/-	91.1±5.0	<0.1
MXAN5040 [-]	+/-	55.5±7.3	1.2±0.8
MXAN7147 [-]	+/-	78.5±8.1	7.9±2.5
MXAN7279 [-]	+/-, -	44.0±7.8	5.4±2.1
MXAN2510 [-]	+/-, -	35.0±8.6	<0.1
MXAN6732 [-]	-	47.4±3.7	6.3±3.3

[a] -, the mutant strain of *M. xanthus* containing an insertion in the indicated gene.

[b] +, produced normal-looking aggregates; +/-, aggregates failed to compact or darken as quickly as wild-type aggregates; -, aggregates (after 5 days of development) had abnormal shapes.

[c] The mean values (± standard deviations) for the spore assays are shown as percentages of *M. xanthus* DK1622 (wild type). The number of fruiting body spores produced by wild-type cells ranged from 6×10^6 to 4×10^7. Means (± standard deviations) derived from three independent experiments are shown.

4.5.7 Insertions in some Nla28 target genes affect swarm expansion

As noted above, one of the characterized targets of Nla28 is *pilA*, a gene that is important for type IV pili-based social motility, surface spreading and normal development [30, 40, 41]. Since many of the remaining Nla28 targets have yet to be tested for roles in *M. xanthus* motility, we examined whether mutations in Nla28 target genes affect colony spreading on agar surfaces.

Specifically, cells carrying mutations in Nla28 target genes were placed on the surface of 1.5% and 0.4% agar plates, the plates were incubated at 32°C for 3 days, and the colony diameters were compared to those produced by wild-type cells and the negative control cells, which are unable to actively spread on agar surfaces due to mutations that inhibit social motility and adventurous motility (*i.e.*, the two motility systems that *M. xanthus* uses for surface spreading) [28, 42].

The results indicated that most of the mutations in Nla28 target genes have little or no impact on surface spreading (**Table 4.7**). Insertions in MXAN2510 and MXAN7279 were notable exceptions, as a strain carrying an insertion in either gene showed substantial reductions in surface spreading compared to the wild-type strain. In the case of the MXAN7279 insertion mutant, a reduction in surface spreading was observed on 0.4% agar plates, which provide a soft and wet surface that favors social motility, but not on 1.5% agar plates, which provide a relatively firm and dry surface that favors A-motility [43]. As for the MXAN2510 insertion mutant, surface spreading on 1.5% agar plates was slightly reduced and surface spreading on 0.4% agar plates was substantially reduced. This phenotype is reminiscent of the *pilA* (social motility) mutant, which displays reduced surface spreading on both agar surfaces, but the reduction is most dramatic on 0.4% agar (**Table 4.7**).

Table 4.7 Swarm diameters of wild-type and mutant cells on 0.4% and 1.5% agar[a]

Strain[b]	Mean swarm diameter (percentage of wild type)	
	Soft agar (0.4%)	Hard agar (1.5%)
DK1622 (wild type)	100 ± 10	100 ± 5
DK 2161 (A^- S^-)[c]	38 ± 3	37 ± 3
mrpB -	94 ± 8	100 ± 5
nla28 -	96 ± 5	93 ± 6
pilA -	27 ±10	72 ± 6
MXAN881 -	101 ± 10	96 ±11
MXAN989 -	88 ± 11	81 ±10
MXAN2510 -	48 ± 3	82 ± 7
MXAN5040 -	89 ± 5	82 ± 4
MXAN6732 -	108 ± 12	108 ± 4
MXAN7147 -	96 ± 7	94 ± 4
MXAN7279 -	41 ± 7	110 ± 3

[a] The mean diameter (± standard deviation) of five swarms produced by each mutant strain was determined and normalized to the mean diameter of five swarms produced by wild-type strain DK1622.

[b] -, the mutant strain of *M. xanthus* containing an insertion in the indicated gene.

[c] A^- represents defect in adventurous motility (A-motility); S^- represents defect in social motility (S-motility).

4.6 Discussion

4.6.1 The likely binding site of Nla28 is an 8-bp direct repeat

One goal of this study was to better understand how Nla28 identifies its target developmental promoters. Using a collection of potential promoter targets (**Table 4.3**), in vitro promoter binding assays (**Fig. 4.3**) and in vitro mutational analyses (**Fig. 4.4**), similar 8-bp sequences [consensus CT(C/G)CG(C/G)AG] were implicated in Nla28 binding. Our in vivo studies further suggest that the 8-bp sequences (**Fig. 4.5)**, as well as the Nla28 protein (**Table 4.5 and Fig. 4.2)**, are important for the activity of Nla28 target promoters in developing *M. xanthus* cells. Assuming Nla28 binds to DNA as a dimer, which would be consistent with other characterized EBPs (23, 24), we propose that the DNA binding sites of Nla28 are tandem, imperfect 8-bp repeats. Based on scans of the 300 bp upstream of putative -12- and -24-bp regions, most of the Nla28 target promoters have clusters of 3-5 imperfect repeats (**Fig. 4.1**). In fact, seven promoters are predicted to have two such clusters. Thus, most of the promoters have multiple putative sites for the binding of Nla28 dimers. One possible explanation for this arrangement is that the clustering of binding sites helps sequester Nla28 at the promoters, increasing local concentrations. A relatively high local concentration of Nla28 would in turn facilitate transcription, even when the individual binding sites in the promoter are low affinity sites [44-46].

4.6.2 Potential Nla28 binding sites are found upstream of many developmental genes

Since we were interested in getting a more global view of the numbers and types of developmental genes that Nla28 directly regulates, we used the *M. xanthus* genome sequence, the consensus Nla28 binding site, bioinformatics and expression data to identify additional promoter/gene targets of Nla28 (see *Materials and Methods*). As shown in Table 4.4, an additional 58 putative σ^{54} promoters and 102 genes were tagged as potential targets of Nla28 using the consensus Nla28 binding sequence, bringing the total number of candidates for direct Nla28 regulation to 70 promoters and 140 genes. Interestingly, many of the putative promoters that we identified have multiple tandem 8-bp sites and hence multiple potential sites for the binding of Nla28 dimers, which is consistent with the initial 12 promoters that we examined. The types of genes that were discovered in this analysis are discussed below.

4.6.3 The vast majority of potential Nla28 target promoters are intragenic

In a recent study of natural product gene regulation in *M. xanthus* [22], 89% of experimentally confirmed and putative σ^{54}-RNA polymerase and Nla28 binding sites were localized to natural product genes and not to intergenic sequences. In our analysis of 12 developmental promoter targets of Nla28, which is described here, we obtained similar results (75% are located in protein coding sequences, as summarized in **Fig. 4.1**). This also holds true for the putative developmental promoter targets of Nla28 listed in Table 4.4. With these findings and previous data linking the majority of the σ^{54} promoter targets of Nla6 to

intragenic regions [19], a pattern of intragenic σ^{54} promoter usage in *M. xanthus* is starting to emerge. However, we are still unable to make sweeping conclusions regarding the genomic locations of *M. xanthus* promoters, as large-scale promoter analyses are lacking. This aside, it is worth reminding the reader that in vivo mutational analyses have been performed on several intragenic promoter targets of Nla6 and Nla28 [6, 22, 35] and the results indicate that the promoters have the signature properties of σ^{54} promoter elements and are crucial for developmental and/or growth-phase related gene expression. Thus, it seems these intragenic targets of Nla6 and Nla28 are bona fide σ^{54} promoters, as predicted. Given the experimental confirmation of these intragenic promoters, we suggest that having a promoter in the coding sequences of genes might have advantages in some cases. For example, some of these putative σ^{54} promoters are in the coding sequence of a gene that is far upstream of the single gene or operon they are predicted to regulate. This distant location would produce a relatively long 5' untranslated region (UTR) in the mRNA and perhaps the 5' UTR provides an additional layer of regulation for natural product or developmental gene expression. Other intragenic σ^{54} promoters appear to be internal promoters; promoters located within a gene in an operon instead of upstream of the first gene in the operon. In such cases, the promoter is predicted to yield an mRNA corresponding to a subset of an operon's genes, providing flexibility if only some of the genes are needed in a particular environment [47].

4.6.4 Most of the confirmed Nla28 targets have regulatory functions or are predicted to have defense functions

As noted above, the molecular/cellular functions of some of the confirmed developmental targets of Nla28 were identified prior to this work. Regulatory or signal transduction functions dominate this group of Nla28 targets, as all but one of these loci fall into this category. This includes the *actB*, *nla6*, *nla28* and *mrpB* loci, which contain genes that code for EBPs; these EBPs modulate transcription of different types of developmental genes [15, 16, 29, 31, 39, 48-50] (**Table 4.3 and Table 4.4)**. The EBP loci also contain genes for signal transduction proteins such as the histidine protein kinase partners of Nla6 and Nla28, which are called Nla6S and Nla28S, respectively [16, 19-21, 37]. The one confirmed and previously characterized Nla28 target that does not have a regulatory function is *pilA*; the *pilA* locus contains genes that are important for type IV pili-based social motility, surface spreading and normal development [15, 16, 30, 40, 41, 51, 52].

Here, we performed a preliminary characterization of the remainder of the confirmed Nla28 targets via mutational analysis. We found that all of the loci are important for development and two are important for surface spreading and presumably motility (**Table 4.5 and 4.6**). The most common molecular/cellular functional category assigned to these loci, based on the annotated genome sequence of wild-type strain DK1622 [16], is defense mechanisms (**Table 4.8 and Fig. 4.6A**). This includes genes that are predicted to encode heavy metal efflux transporter components. It also includes CRISPR-associated proteins, which have been characterized in a variety of bacterial species and are known to

be involved in adaptive immunity against foreign genetic elements [53-57]. As for the remaining Nla28 target loci characterized here, two have been linked to potential metabolic functions. Namely, one locus is predicted to be involved in lipid metabolism and the second in carbohydrate metabolism (**Table 4.8 and Fig. 4.6A)**. The other two loci characterized here have not been assigned a particular function. However, the three Nla28 targets characterized in a recent study [6, 21, 22] are worth mentioning here, since they are likely involved in production of secondary metabolites and hence could have a defense function, predation function or could be involved in signaling.

Table 4.8 Functional categories of confirmed targets of Nla28

Locus	Number of genes	NCBI gene definition	NCBI Cluster of Orthologous Genes (COG) Category:
actB	4	sigma-54-dependent Fis family transcriptional regulator	Signal transduction mechanisms
nla6	2	NtrC family HydH-HydG (metal tolerance)	Signal transduction mechanisms
nla28	2	NtrC family AtoS-AtoC (cPHB biosynthesis)	Signal transduction mechanisms
mrpB (MXAN5124)	1	sigma-54 dependent DNA-binding response regulator	Signal transduction mechanisms
pilA (MXAN5783)	6	Bacterial motility proteins	Cell motility/ Extracellular Structures
MXAN881	1	Hypothetical lipoprotein	Poorly Characterized

MXAN989	2	cation efflux system protein CusB	Inorganic ion transport and metabolism
MXAN2511	10	hypothetical protein	Poorly Characterized
MXAN5040	2	aldehyde dehydrogenase family protein	Lipid transport and metabolism
MXAN6732	1	class II aldolase/adducin domain protein	Carbohydrate transport and metabolism/ Amino acid transport and metabolism
MXAN7147	2	efflux transporter, RND family, MFP subunit	Defense mechanisms
MXAN7280	5	putative CRISPR-associated protein Crm3	Defense mechanisms

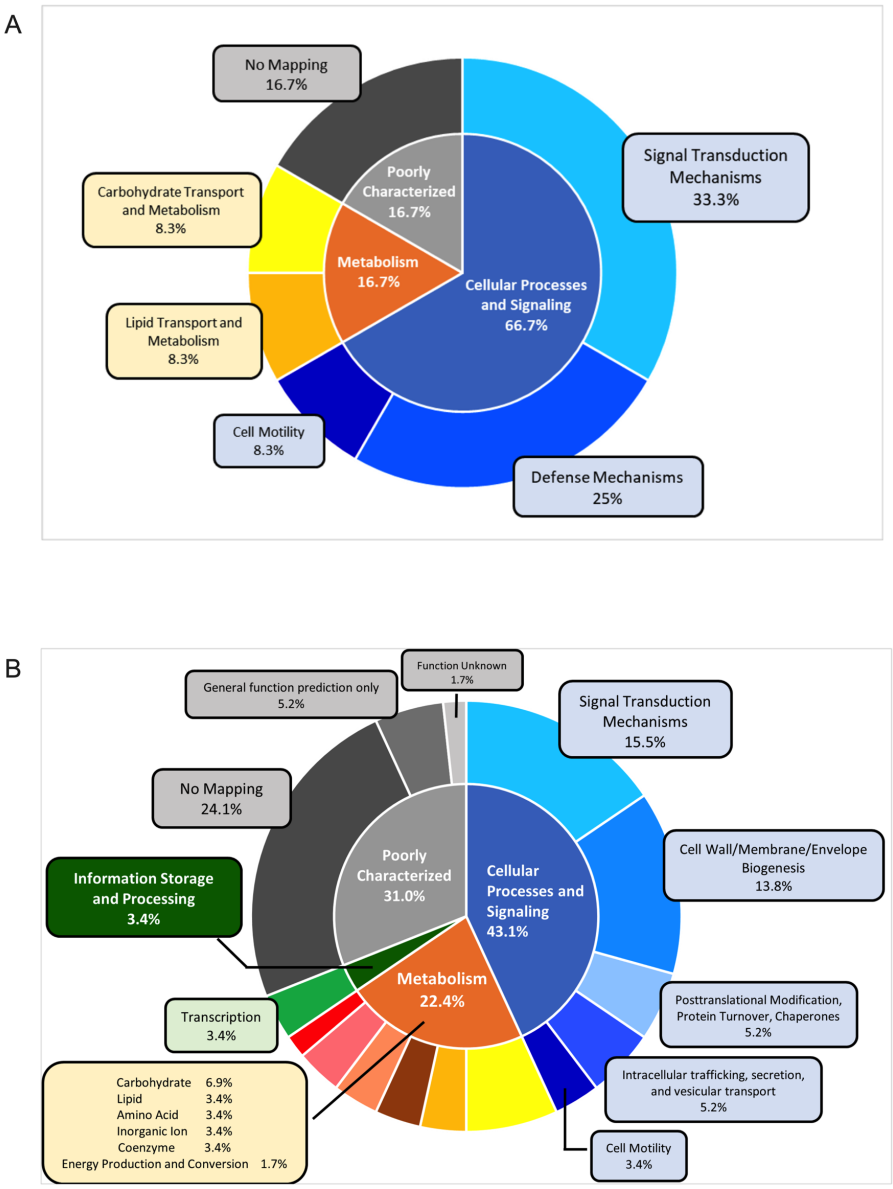


Figure 4.6 Functional category and subcategory distributions of Nla28 targets based on the Clusters of Orthologous Groups of proteins (COGs) database

(A) The two-layer pie chart shows COG functional categories, subcategories and their relative abundances of confirmed Nla28 targets in this study. (B) The two-layer pie chart shows COG functional categories, subcategories and their relative abundances of putative Nla28 targets in this study. The inner layer of pie charts represents categories; the outer layer represents subcategories.

It has been suggested that Nla28 is part of a general stress response induced by nutrient depletion, since Nla28 modulates transcription of many early, starvation-responsive developmental genes and genes that are highly expressed during the transition into stationary phase [6, 21, 22]. That some confirmed targets have defense functions makes sense if Nla28 is viewed as a regulator of starvation-induced stress. The metabolic functions of other Nla28 targets also make sense in this context, as one would expect nutrient depletion to be accompanied by changes in cellular metabolism. Additional information that supports the view of Nla28 as a regulator of *M. xanthus*' stress response comes from physiological studies in other Gram-negative bacteria. In particular, orthologues of MXAN881, MXAN989, MXAN5040, MXAN7147 and MXAN7280 have all been linked to stress responses in other Gram-negative bacteria [58-62] (**Table 4.8 and Fig. 4.6A**).

4.6.5 Regulatory, metabolic and cell envelope biogenesis are common functions among putative Nla28 targets

As noted above, 58 putative σ^{54} promoters and 102 genes were tagged as potential targets of Nla28 using the consensus Nla28 half binding site, bringing the total number of candidates for direct Nla28 regulation to 70 promoters and 140 genes. Based on the annotated genome sequence of wild-type strain DK1622 [16], we were able to place the putative Nla28 targets into molecular/cellular functional categories (**Table 4.9 and Fig. 4.6B**). The three most highly represented categories are metabolic functions, regulatory/signal transduction and cell envelope/cell wall biogenesis. Given that developing cells

are experiencing starvation, changes in cellular metabolism might be expected. Similarly, developing cells might be expected to express genes involved in cell wall or cell membrane structural changes, as developing cells are experiencing starvation-induced stress. Other notable categories of putative Nla28 targets include motility genes, post-translational modifications/protein turnover and protein secretion. Finally, many of the putative Nla28 targets have predicted orthologues linked to stress responses in other Gram-negative bacteria. For example, potential orthologues of MXAN162, MXAN934 MXAN2907 have been linked to envelope stress in Gram-negative species [63-66] (**Table 4.9 and Fig. 4.6B**).

Table 4.9 Functional categories of potential targets of Nla28

Locus	Number of genes	NCBI definition	NCBI Cluster of Orthologous Genes (COG) Category:
MXAN162	1	hypothetical protein	Poorly Characterized
MXAN179	3	metallo-beta-lactamase family protein	Poorly Characterized
MXAN255	1	peptidase homolog, M20 family	Amino acid transport and metabolism
MXAN419	2	hypothetical protein	Signal transduction mechanisms
MXAN496	2	glyoxalase family protein	General function prediction only
MXAN542	2	F5/8 type C domain protein	Poorly Characterized
MXAN562	1	phosphate-selective porin O and P	Poorly Characterized
MXAN854	1	hypothetical protein.	Poorly Characterized
MXAN859	1	putative lipoprotein	Poorly Characterized
MXAN909	2	hypothetical protein	Poorly Characterized
MXAN934	1	protease DO family protein	Posttranslational modification, protein turnover, chaperones
MXAN1043	2	glycosyl transferase	Cell wall/membrane/envelope biogenesis

MXAN1124	1	ABC transporter, ATP-binding protein	Cell wall/membrane/envelope biogenesis
MXAN1404	1	hypothetical protein TIGR00661	Poorly Characterized
MXAN1412	1	putative serine/threonine protein phosphatase	Signal transduction mechanisms
MXAN1433	1	M23 peptidase domain protein	Cell wall/membrane/envelope biogenesis
MXAN1752	2	isoquinoline 1-oxidoreductase, beta subunit	Energy production and conversion
MXAN1783	1	hypothetical protein	Poorly Characterized
MXAN1968	1	hypothetical protein	Poorly Characterized
MXAN2159 (*nla12*)	2	sigma-54 dependent transcriptional regulator, Fis family	Signal transduction mechanisms
MXAN2359	2	glycosyl transferase, group 2 family protein	Cell wall/membrane/envelope biogenesis
MXAN2362	1	glycosyl transferase, group 1 family protein	Cell wall/membrane/envelope biogenesis
MXAN2363	1	glutaminyl-tRNA synthetase	Translation, ribosomal structure and biogenesis
MXAN2447	7	FliP/MopC/SpaP family protein	Intracellular trafficking, secretion, and vesicular transport
MXAN2515	5	general secretion pathway protein C	Intracellular trafficking, secretion, and vesicular transport
MXAN2735	1	methyl-accepting chemotaxis protein	Cell motility
MXAN2752	2	hypothetical protein	Poorly Characterized
MXAN2777	1	tonB domain protein	Cell wall/membrane/envelope biogenesis
MXAN2907	1	CDP-diacylglycerol--serine O-phosphatidyltransferase	Lipid transport and metabolism
MXAN2947	1	isochorismatase family protein	Coenzyme transport and metabolism
MXAN2952	5	hypothetical protein	Poorly Characterized
MXAN3026	2	O-antigen polymerase family protein	Cell wall/membrane/envelope biogenesis
MXAN3149	2	peptidase, M6 (immune inhibitor A) family	Posttranslational modification, protein turnover, chaperones
MXAN3167	1	hypothetical protein	Poorly Characterized

MXAN3735	2	response regulator	Transcription/Signal transduction mechanisms
MXAN3778	1	DnaK family protein	Posttranslational modification, protein turnover, chaperones
MXAN3824	1	general secretion pathway protein G	Cell motility/ Extracellular structures/ Intracellular trafficking, secretion, and vesicular transport
MXAN3962	2	pyridine nucleotide-disulphide oxidoreductase	Amino acid transport and metabolism
MXAN3975	1	hypothetical protein	Poorly Characterized
MXAN3979	2	hypothetical protein	Poorly Characterized
MXAN4007	1	hypothetical protein	Poorly Characterized
MXAN4155	1	hypothetical protein	Poorly Characterized
MXAN4198	4	putative outer membrane macrolide efflux protein	Cell wall/membrane/envelope biogenesis
MXAN4236	1	CBS domain protein	Signal transduction mechanisms
MXAN4687	1	YHS domain protein	Inorganic ion transport and metabolism
MXAN4784	2	inorganic anion transporter, sulfate permease (SulP) family	Inorganic ion transport and metabolism
MXAN4877	1	aspartate kinase	Amino acid transport and metabolism
MXAN5543	6	ATPase, P-type (transporting), HAD superfamily, subfamily IC	Inorganic ion transport and metabolism
MXAN5615	1	hypothetical protein	Poorly Characterized
MXAN5715	1	response regulator/putative sensor histidine kinase	Signal transduction mechanisms
MXAN6209 (*sigC*)	1	RNA polymerase sigma-C factor	Transcription
MXAN6223	5	sensor histidine kinase	Signal transduction mechanisms
MXAN6297	1	ribulose-phosphate 3-epimerase	Carbohydrate transport and metabolism
MXAN6437	1	putative lipoprotein	Poorly Characterized
MXAN6788	1	conserved domain protein	Poorly Characterized
MXAN7157	1	enoyl-CoA hydratase/isomerase family protein	Lipid transport and metabolism

MXAN7225	2	putative sugar ABC transporter, ATP-binding protein	Carbohydrate transport and metabolism
MXAN7435	1	hydrolase, alpha/beta fold family	Coenzyme transport and metabolism

4.6.6 Conclusions

Nla28 is an EBP that begins modulating gene expression soon after *M. xanthus* cells encounter nutrient-poor conditions. In this study, we identified the direct targets of Nla28 to better understand Nla28's function and to help define the early gene regulatory pathways involved in *M. xanthus*' starvation response. Our data suggest that Nla28 might recognize its target promoters using tandem, imperfect repeats of an 8-bp sequence and that most of the Nla28 target promoters are intragenic. Seventy promoters and 140 genes were classified as potential targets of Nla28, suggesting that the Nla28 regulon might be relatively large. Some of the functions assigned to Nla28 target genes are regulatory, metabolism, defense-related functions, and cell envelope biogenesis. Many of these functions make sense in the context of Nla28's role as a general regulator of stress-associated genes and, based on the work presented here, we now know that several these genes are important for production of stress-resistant spores following starvation.

4.7 References

1. Mittal, S. and L. Kroos, *Combinatorial regulation by a novel arrangement of FruA and MrpC2 transcription factors during Myxococcus xanthus development.* Journal of bacteriology, 2009. **191**(8): p. 2753-2763.
2. Lee, J.-s., et al., *Combinatorial regulation of fmgD by MrpC2 and FruA during Myxococcus xanthus development.* Journal of bacteriology, 2011. **193**(7): p. 1681-1689.
3. Popham, D.L., et al., *Function of a bacterial activator protein that binds to transcriptional enhancers.* Science, 1989. **243**(4891): p. 629-635.
4. Sasse-Dwight, S. and J.D. Gralia, *Role of eukaryotic-type functional domains found in the prokaryotic enhancer receptor factor σ54.* Cell, 1990. **62**(5): p. 945-954.
5. Wedel, A. and S. Kustu, *The bacterial enhancer-binding protein NTRC is a molecular machine: ATP hydrolysis is coupled to transcriptional activation.* Genes & development, 1995. **9**(16): p. 2042-2052.
6. Giglio, K.M., et al., *A cascade of coregulating enhancer binding proteins initiates and propagates a multicellular developmental program.* Proceedings of the National Academy of Sciences, 2011. **108**(32): p. E431-E439.
7. O'Connor, K.A. and D.R. Zusman, *Development in Myxococcus xanthus involves differentiation into two cell types, peripheral rods and spores.* Journal of Bacteriology, 1991. **173**(11): p. 3318-3333.

8. O'Connor, K.A. and D.R. Zusman, *Behavior of peripheral rods and their role in the life cycle of Myxococcus xanthus.* Journal of bacteriology, 1991. **173**(11): p. 3342-3355.

9. O'Connor, K.A. and D.R. Zusman, *Patterns of cellular interactions during fruiting-body formation in Myxococcus xanthus.* Journal of bacteriology, 1989. **171**(11): p. 6013-6024.

10. Singer, M. and D. Kaiser, *Ectopic production of guanosine penta-and tetraphosphate can initiate early developmental gene expression in Myxococcus xanthus.* Genes & development, 1995. **9**(13): p. 1633-1644.

11. Harris, B.Z., D. Kaiser, and M. Singer, *The guanosine nucleotide (p) ppGpp initiates development and A-factor production inMyxococcus xanthus.* Genes & development, 1998. **12**(7): p. 1022-1035.

12. Kuspa, A., L. Plamann, and D. Kaiser, *A-signalling and the cell density requirement for Myxococcus xanthus development.* Journal of Bacteriology, 1992. **174**(22): p. 7360-7369.

13. Kuspa, A., L. Plamann, and D. Kaiser, *Identification of heat-stable A-factor from Myxococcus xanthus.* Journal of Bacteriology, 1992. **174**(10): p. 3319-3326.

14. Plamann, L., A. Kuspa, and D. Kaiser, *Proteins that rescue A-signal-defective mutants of Myxococcus xanthus.* Journal of bacteriology, 1992. **174**(10): p. 3311-3318.

15. Caberoy, N.B., et al., *Global mutational analysis of NtrC-like activators in Myxococcus xanthus: identifying activator mutants defective for motility and*

fruiting body development. Journal of Bacteriology, 2003. **185**(20): p. 6083-6094.

16. Goldman, B., et al., *Evolution of sensory complexity recorded in a myxobacterial genome.* Proceedings of the National Academy of Sciences, 2006. **103**(41): p. 15200-15205.
17. Diodati, M.E., et al., *Nla18, a key regulatory protein required for normal growth and development of Myxococcus xanthus.* Journal of bacteriology, 2006. **188**(5): p. 1733-1743.
18. Ossa, F., et al., *The Myxococcus xanthus Nla4 protein is important for expression of stringent response-associated genes, ppGpp accumulation, and fruiting body development.* Journal of bacteriology, 2007. **189**(23): p. 8474-8483.
19. Giglio, K.M., et al., *The enhancer binding protein Nla6 regulates developmental genes that are important for Myxococcus xanthus sporulation.* Journal of bacteriology, 2015. **197**(7): p. 1276-1287.
20. Sarwar, Z. and A.G. Garza, *The Nla28S/Nla28 two-component signal transduction system regulates sporulation in Myxococcus xanthus.* Journal of bacteriology, 2012. **194**(17): p. 4698-4708.
21. Sarwar, Z. and A.G. Garza, *Two-component signal transduction systems that regulate the temporal and spatial expression of Myxococcus xanthus sporulation genes.* Journal of bacteriology, 2016. **198**(3): p. 377-385.
22. Ma, M., R.D. Welch, and A.G. Garza, *The σ 54 system directly regulates bacterial natural product genes.* Scientific reports, 2021. **11**(1): p. 1-11.

23. Bush, M. and R. Dixon, *The role of bacterial enhancer binding proteins as specialized activators of σ54-dependent transcription.* Microbiology and Molecular Biology Reviews, 2012. **76**(3): p. 497-529.

24. Gao, F., et al., *Bacterial Enhancer Binding Proteins—AAA+ Proteins in Transcription Activation.* Biomolecules, 2020. **10**(3): p. 351.

25. Giglio, K.M., J. Eisenstatt, and A.G. Garza, *Identification of enhancer binding proteins important for Myxococcus xanthus development.* Journal of bacteriology, 2010. **192**(1): p. 360-364.

26. Studier, F.W. and B.A. Moffatt, *Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes.* Journal of molecular biology, 1986. **189**(1): p. 113-130.

27. Kaiser, D., *Social gliding is correlated with the presence of pili in Myxococcus xanthus.* Proceedings of the National Academy of Sciences, 1979. **76**(11): p. 5952-5956.

28. Hodgkin, J. and D. Kaiser, *Genetics of gliding motility in Myxococcus xanthus (Myxobacterales): two gene systems control movement.* Molecular and General Genetics MGG, 1979. **171**(2): p. 177-191.

29. Gronewold, T.M. and D. Kaiser, *The act operon controls the level and time of C-signal production for Myxococcus xanthus development.* Molecular microbiology, 2001. **40**(3): p. 744-756.

30. Wu, S.S. and D. Kaiser, *Markerless deletions of pil genes in Myxococcus xanthus generated by counterselection with the Bacillus subtilis sacB gene.* Journal of Bacteriology, 1996. **178**(19): p. 5817-5821.

31. Sun, H. and W. Shi, *Genetic studies of mrp, a locus essential for cellular aggregation and sporulation of Myxococcus xanthus.* Journal of bacteriology, 2001. **183**(16): p. 4786-4795.

32. Fisseha, M., et al., *Characterization of the regulatory region of a cell interaction-dependent gene in Myxococcus xanthus.* Journal of Bacteriology, 1996. **178**(9): p. 2539-2550.

33. Plamann, L., et al., *Evidence that asgB encodes a DNA-binding protein essential for growth and development of Myxococcus xanthus.* Journal of bacteriology, 1994. **176**(7): p. 2013-2020.

34. Wu, S.S. and D. Kaiser, *Regulation of expression of the pilA gene in Myxococcus xanthus.* Journal of bacteriology, 1997. **179**(24): p. 7748-7758.

35. Gronewold, T.M. and D. Kaiser, *Mutations of the act promoter in Myxococcus xanthus.* Journal of bacteriology, 2007. **189**(5): p. 1836-1844.

36. Studholme, D.J., M. Buck, and T. Nixon, *Identification of potential σN-dependent promoters in bacterial genomes.* Microbiology, 2000. **146**(12): p. 3021-3023.

37. Sarwar, Z. and A.G. Garza, *The Nla6S protein of Myxococcus xanthus is the prototype for a new family of bacterial histidine kinases.* FEMS microbiology letters, 2012. **335**(2): p. 86-94.

38. Kadam, S.V., et al., *Novel transcriptome patterns accompany evolutionary restoration of defective social development in the bacterium Myxococcus xanthus.* Molecular biology and evolution, 2008. **25**(7): p. 1274-1281.

39. Gorski, L. and D. Kaiser, *Targeted mutagenesis of ς54 activator proteins in Myxococcus xanthus.* Journal of Bacteriology, 1998. **180**(22): p. 5896-5905.

40. Wu, S.S. and D. Kaiser, *Genetic and functional evidence that type IV pili are required for social gliding motility in Myxococcus xanthus.* Molecular microbiology, 1995. **18**(3): p. 547-558.

41. Wu, S.S., *The role of Type IV pili in social gliding motility of Myxococcus xanthus*. 1998: Stanford University.

42. Hodgkin, J. and D. Kaiser, *Genetics of gliding motility in Myxococcus xanthus (Myxobacterales): genes controlling movement of single cells.* Molecular and General Genetics MGG, 1979. **171**(2): p. 167-176.

43. Shi, W. and D.R. Zusman, *The two motility systems of Myxococcus xanthus show different selective advantages on various surfaces.* Proceedings of the National Academy of Sciences, 1993. **90**(8): p. 3378-3382.

44. Sharon, E., et al., *Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters.* Nature biotechnology, 2012. **30**(6): p. 521-530.

45. Ezer, D., N.R. Zabet, and B. Adryan, *Physical constraints determine the logic of bacterial promoter architectures.* Nucleic acids research, 2014. **42**(7): p. 4196-4207.

46. Ezer, D., N.R. Zabet, and B. Adryan, *Homotypic clusters of transcription factor binding sites: A model system for understanding the physical mechanics of gene expression.* Computational and structural biotechnology journal, 2014. **10**(17): p. 63-69.

47. Mejía-Almonte, C., et al., *Redefining fundamental concepts of transcription initiation in bacteria.* Nature Reviews Genetics, 2020. **21**(11): p. 699-714.

48. Gronewold, T.M. and D. Kaiser, *act operon control of developmental gene expression in Myxococcus xanthus*. 2002, Am Soc Microbiol.

49. Sun, H. and W. Shi, *Analyses of mrp genes during Myxococcus xanthus development.* Journal of Bacteriology, 2001. **183**(23): p. 6733-6739.

50. McLaughlin, P.T., et al., *MrpC, a CRP/Fnr homolog, functions as a negative autoregulator during the Myxococcus xanthus multicellular developmental program.* Molecular microbiology, 2018. **109**(2): p. 245-261.

51. Wall, D. and D. Kaiser, *Type IV pili and cell motility.* Molecular microbiology, 1999. **32**(1): p. 01-10.

52. Bretl, D.J., et al., *Type IV-pili dependent motility is co-regulated by PilSR and PilS2R2 two-component systems via distinct pathways in Myxococcus xanthus.* Molecular microbiology, 2016. **102**(1): p. 37-53.

53. Jinek, M., et al., *A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity.* science, 2012. **337**(6096): p. 816-821.

54. Heler, R., L.A. Marraffini, and D. Bikard, *Adapting to new threats: the generation of memory by CRISPR-Cas immune systems.* Molecular microbiology, 2014. **93**(1): p. 1-9.

55. Marraffini, L.A., *CRISPR-Cas immunity in prokaryotes.* Nature, 2015. **526**(7571): p. 55-61.

56. Mojica, F.J. and F. Rodriguez-Valera, *The discovery of CRISPR in archaea and bacteria.* The FEBS journal, 2016. **283**(17): p. 3162-3169.

57. Jiang, F. and J.A. Doudna, *CRISPR–Cas9 structures and mechanisms.* Annual review of biophysics, 2017. **46**: p. 505-529.

58. Shaibe, E., E. Metzer, and Y.S. Halpern, *Control of utilization of L-arginine, L-ornithine, agmatine, and putrescine as nitrogen sources in Escherichia coli K-12.* Journal of bacteriology, 1985. **163**(3): p. 938-942.

59. Ma, D., et al., *Genes acrA and acrB encode a stress-induced efflux system of Escherichia coli.* Molecular microbiology, 1995. **16**(1): p. 45-55.

60. Bay, D.C., et al., *Secondary multidrug efflux pump mutants alter Escherichia coli biofilm growth in the presence of cationic antimicrobial compounds.* Research in microbiology, 2017. **168**(3): p. 208-221.

61. May, K.L., et al., *A stress response monitoring lipoprotein trafficking to the outer membrane.* Mbio, 2019. **10**(3): p. e00618-19.

62. McMahon, S.A., et al., *Structure and mechanism of a Type III CRISPR defence DNA nuclease activated by cyclic oligoadenylate.* Nature communications, 2020. **11**(1): p. 1-11.

63. Kim, K.I., et al., *Selective degradation of unfolded proteins by the self-compartmentalizing HtrA protease, a periplasmic heat shock protein in Escherichia coli.* Journal of molecular biology, 1999. **294**(5): p. 1363-1374.

64. Skorko-Glonek, J., et al., *The Escherichia coli heat shock protease HtrA participates in defense against oxidative stress.* Molecular and General Genetics MGG, 1999. **262**(2): p. 342-350.

65. Rowley, G., et al., *Pushing the envelope: extracytoplasmic stress responses in bacterial pathogens.* Nature Reviews Microbiology, 2006. **4**(5): p. 383-394.

66. Choi, U. and C.-R. Lee, *Distinct roles of outer membrane porins in antibiotic resistance and membrane integrity in Escherichia coli.* Frontiers in microbiology, 2019. **10**: p. 953.

Chapter 5

Nla28-Mediated σ^{54}-Dependent Regulation of Natural Product Genes in *Myxoccocus xanthus*

5.1 Author Contributions

Muqing Ma designed and conducted all experiments; collected, analyzed, and interpreted the data with Anthony Garza. Muqing Ma and Anthony Garza wrote the manuscript. Anthony Garza and Roy Welch conceived and supervised the research.

5.2 Abstract

Bacterial-derived polyketide and non-ribosomal peptide natural products are crucial sources of therapeutics and yet little is known about the conditions that favor activation of natural product genes or the regulatory machinery controlling their transcription. Recent findings suggest that the σ^{54} system, which includes σ^{54}-loaded RNA polymerase and transcriptional activators called enhancer binding proteins (EBPs), might be a common regulator of natural product genes. Here, we explored this idea by analyzing a selected group of putative σ^{54} promoters identified in *Myxococcus xanthus* natural product gene clusters. We show that mutations in putative σ^{54}-RNA polymerase binding regions and in putative Nla28 EBP binding sites dramatically reduce in vivo promoter activities in growing and developing cells. We also show in vivo promoter activities are reduced in a *nla28* mutant, that Nla28 binds to wild-type fragments of these promoters in vitro, and that in vitro binding is lost when the Nla28 binding sites are mutated. Together, our results indicate that *M. xanthus* uses σ^{54} promoters for transcription of at least some of its natural product genes. Interestingly, the vast majority of experimentally confirmed and putative σ^{54} promoters in *M. xanthus* natural product loci are located within genes and not in intergenic sequences.

5.3 Introduction

The σ^{54} regulatory system modulates transcription of a wide variety of bacterial genes. One crucial component of this regulatory system is the σ^{54} protein, which directs RNA polymerase to conserved DNA sequences located in the -12 and -24-bp regions of σ^{54} promoter elements [1, 2]. Enhancer binding proteins (EBPs), which are transcriptional activators, are also crucial for the normal function of the σ^{54} regulatory system. Namely, EBPs are ATPases that use the energy from ATP hydrolysis to help σ^{54} -RNA polymerase form an open promoter complex and initiate transcription [3-5]. Bacteria typically have one gene for σ^{54}, but often have multiple genes for EBPs; each EBP works with σ^{54} to regulate a subset of σ^{54} promoters, which the EBP identifies via specific tandem repeat sequences or enhancer elements [6, 7]. Interestingly, the tandem repeat binding sites of EBP dimers are typically located 80- to 150-bp upstream of the -24 and -12 regions of σ^{54} promoters; hence, it seems likely that many σ^{54} promoters have intrinsically curved DNA sequences or binding sequences for DNA bending proteins, as EBP dimers directly contact σ^{54} -RNA polymerase [8-10].

EBPs generally contain three domains: an N-terminal signaling domain, a central ATPase domain that is responsible for ATP hydrolysis and transcriptional activation, and a C-terminal DNA-binding domain (DBD) that recognizes a specific DNA sequence [7]. Typically, the N-terminal signaling domain modulates the ATPase activity of the EBP in response to an intracellular or extracellular signal. In some cases, the N-terminal domain binds directly to a signaling

molecule. However, the N-terminal domain of most EBPs is modified (eg., by phosphorylation) by a signal transduction partner such as a histidine protein kinase that detects the signal [11]. Because σ^{54}-RNA polymerase requires the energy from EBP-catalyzed ATP hydrolysis to initiate transcription and the EBP's ATPase activity is controlled by signal input, the σ^{54} system is able to tightly control transcription of its target genes.

Historically, σ^{54} was viewed as specialized regulatory system that was mainly dedicated to transcription of genes involved in nitrogen assimilation or nitrogen fixation [12, 13]. In recent years however, it has become clear that the σ^{54} system is important for transcription of many types of bacterial genes. For example, the σ^{54} system in *Bacillus cereus*, *Pseudomonas aeruginosa* and *Escherichia coli* modulates transcription of genes involved in amino acid metabolism, the response to reactive nitrogen species and the phage shock response, respectively [14-16]. In *Caulobacter crescentus*, *Pseudomonas aeruginosa*, *Pseudomonas putida* and *Vibrio cholerae* the σ^{54} system regulates genes that are important for flagellar biosynthesis and motility [17-21], and in *P. aeruginosa* the σ^{54} system is implicated in transcription of genes involved in quorum sensing, biofilm formation and virulence [22-24].

The σ^{54} system in the soil bacterium *Myxococcus xanthus* has been studied extensively and has some rare properties. Namely, *M. xanthus* is one of the rare bacterial species in which the σ^{54} system has been linked to growth in nutrient rich conditions. For example, inactivation of the *nla4* or *nla18* EBP gene severely impairs *M. xanthus* growth in nutrient rich media [25-27]. Presumably,

the relatively slow growth of the *nla4* mutant and *nla18* mutant is due at least in part to changes in the basal levels of the intracellular starvation signal (p)ppGpp and the resulting changes in gene expression.

M. xanthus also has an unusually large repertoire of 53 EBP genes [28]. All of the EBPs were characterized a number of years ago and many of the EBPs were implicated in motility [25, 29, 30] and in starvation-induced biofilm formation [25, 30-36], which yields spore-filled aerial structures called fruiting bodies. Six of the EBPs that begin functioning in the early to middle stages of biofilm development, which is also known as fruiting body development, form a regulatory cascade [30]. This EBP cascade is reminiscent of the sigma factor cascade that controls the sequential stages of spore development in *Bacillus subtilis* [37], as pairs of EBPs functioning at one stage of development directly activate transcription of an EBP gene important for the next developmental stage.

Nla28 is one of the early-functioning developmental EBPs that participates in the transcriptional cascade. A putative tandem repeat promoter binding site for Nla28 dimers was identified and analyzed using bioinformatics and experimentation [30, 38, 39]. The consensus Nla28 binding site [CT(C/G)CG(C/G)AG consensus half site], which was generated from these studies, was subsequently used to search the *M. xanthus* genome sequence for Nla28 target promoters/genes located outside the EBP cascade [28, 39]. A number of these putative Nla28 target promoters were identified in natural product gene clusters. This was an intriguing finding, as Volz *et al* (2012) previously showed that two *M. xanthus* EBPs (HsfA and MXAN4899) are capable

of binding to fragments of natural product gene promoters [40]. Furthermore, it was previously suggested that the σ^{54} system might be a key regulator of polyketide (PK) and non-ribosomal peptide (NRP) natural product genes in *M. xanthus*, and in bacteria in general, based on bioinformatics [38].

Here, we explored the idea that the σ^{54} system modulates expression of *M. xanthus* natural product genes. We focused on three potential σ^{54} promoter targets of Nla28 that are located in *M. xanthus* natural product gene clusters [38]. The activities of the three natural product promoters increase during growth, with peak activities occurring during the transition into stationary phase. The activities of the promoters also increase during starvation-induced development. These findings are consistent previous studies linking expression of bacterial natural product genes to changes in nutritional status [41-45]. Our in vivo and in vitro mutational analyses indicate that three natural product promoters are indeed σ^{54} promoter elements and direct targets of Nla28, as predicted via bioinformatics. An examination of the characterized and uncharacterized σ^{54} promoters in *M. xanthus* natural product gene clusters yielded a striking result; the vast majority are predicted to be intragenic not intergenic promoters. These results are consistent with previous experimental findings, which placed many characterized *M. xanthus* σ^{54} promoters in intragenic regions [30, 39].

5.4 Materials and Methods

5.4.1 Bacterial strains, plasmids and media

Bacterial strains and plasmids used in the study are listed in Table 5.1. *M. xanthus* strains were grown at 32℃ in CTTYE broth [1% Casitone, 0.2% yeast extract, 10 mM Tris (pH 8.0), 1 mM KH_2PO_4 (pH 7.6), 8 mM $MgSO_4$] or on CTTYE plates containing 1.5% agar. Fifty μg/ml of kanamycin or 10 μg/ml of tetracycline were added to CTTYE broth and CTTYE agar plates as needed. CTT soft agar (CTTSA), which is used to plate electroporated *M. xanthus* cells, contains 1% Casitone, 10 mM Tris (pH 8.0), 1.0 mM KH_2PO_4 (pH 7.6), 8.0 mM $MgSO_4$, and 0.7% agar. Submerged culture development of *M. xanthus* strains were carried out in 24-well polystyrene plates containing 100 μl of MC7 buffer [10 mM morpholinepropanesulfonic acid (MOPS; pH 7.0), 1 mM $CaCl_2$]. Unless otherwise stated, *E. coli* strains were grown in Luria-Bertani (LB) broth [0.5% yeast extract, 1% tryptone, 1% NaCl) or on LB plates containing 1.5% agar. LB broth and LB plates were supplemented with 100 μg/ml of ampicillin, 50 μg/ml of kanamycin or 10 μg/ml of tetracycline as needed. For Nla28-DBD expression, *E. coli* strains were grown in rich LB broth [0.5% yeast extract, 1% tryptone, 0.5% NaCl, 0.2% glucose] supplemented with 100 μg/ml of ampicillin.

Table 5.1 Bacterial strains and plasmids

Bacterial	Relevant characteristics	Source or reference
Strains		
M. xanthus		
DK1622	Wild type	[46]
MM201	Δ*nla28* derivative of DK1622	This study
E. coli		
Top 10	Cloning host;	Invitrogen
BL21(DE3)	Nla28-DBD protein expression host;	New England Biolabs
Plasmids		
pCR 2.1 TOPO	Cloning vector; Km^r	Invitrogen
pREG1727	Vector containing promoterless *lacZ*; Km^r	[47]
pSWU22	Vector used for *nla28* inactivation; Tet^r	[25]
pMAL-c5x	Maltose-binding protein (MBP) fusion vector for Nla28-DBD protein expression; Amp^r	New England Biolabs
pMM101	WT MXAN1286 promoter in pREG1727; Km^r	This study
pMM101a	-12 region mutant MXAN1286 promoter in pREG1727;	This study
pMM101b	-24 region mutant MXAN1286 promoter in pREG1727;	This study
pMM101c	-12/24 spacer mutant MXAN1286 promoter in	This study
pMM101d	Half Nla28 EBP binding site mutant MXAN1286 promoter in pREG1727; Km^r	This study
pMM102	WT MXAN1579 promoter in pREG1727; Km^r	This study
pMM102a	-12 region mutated MXAN1579 promoter in pREG1727;	This study
pMM102b	-24 region mutated MXAN1579 promoter in pREG1727;	This study
pMM102c	-12/24 spacer mutated MXAN1579 promoter in	This study
pMM102d	Half Nla28 EBP binding site mutated MXAN1579 promoter in pREG1727; Km^r	This study
pMM103	WT MXAN1603 promoter in pREG1727; Km^r	This study
pMM103a	-12 region mutated MXAN1603 promoter in pREG1727;	This study
pMM103b	-24 region mutated MXAN1603 promoter in pREG1727;	This study
pMM103c	-12/24 spacer mutated MXAN1603 promoter in	This study
pMM104	WT MXAN3778 promoter in pREG1727; Km^r	This study
pMM104a	-12 region mutated MXAN3778 promoter in pREG1727;	This study
pMM104b	-24 region mutated MXAN3778 promoter in pREG1727;	This study
pMM104c	-12/24 spacer mutated MXAN3778 promoter in	This study
pMM104d	Half Nla28 EBP binding site mutated MXAN3778 promoter in pREG1727; Km^r	This study

pMM301	600-bp internal fragment of *nla28* gene in pSWU22 used for *nla28* inactivation; Tet^r	This study
pMM302	Fragment for Nla28-DBD expression in pMAL-c5x; Amp^r	This study

5.4.2 M. xanthus growth and development

M. xanthus strains were grown by inoculating cells into flasks containing CTTYE broth and incubating the cultures at 32℃ with vigorous swirling. Development was induced as previously described [25]. Briefly, *M. xanthus* cells were grown in CTTYE broth until the cultures reached a density of approximately 5 x 10^8 cells/ml, the cells were pelleted, the supernatant was removed, and the cells were resuspended in MC7 buffer to a density of 5 x 10^9 cells/ml. Forty μl aliquots of the cell suspensions were placed into polystyrene plate wells containing 100 μl of MC7 buffer and the polystyrene plates were transferred to a 32°C incubator for 24 hours.

5.4.3 Standard DNA procedures

Chromosomal DNA from wild-type *M. xanthus* strain DK1622[46] was extracted using ZYMO Research gDNA extraction kit. Oligonucleotides used in PCR reactions were synthesized by Integrated DNA Technologies (IDT) and are listed in Table 5.2. Plasmid DNA was extracted using the Promega Nucleic acid purification kit. Amplified and digested DNA fragments were purified using the Gel Extraction Minipreps kit of Bio Basic. For all kits, the manufacturer's protocols were used. The compositions of all plasmids and promoter fragments were confirmed by DNA sequencing (Genewiz).

Table 5.2 Oligonucleotides used in the current study

Application of oligonucleotides	Name	Sequence
Amplification of WT σ^{54} promoter fragments		
MXAN1286	MM101F	5’ gaaagcttaacgccgctcgcaagg 3’
	MM101R	5’ atcctagggaggtcggcgatgtg 3’
MXAN1579	MM102F	5’ gcaagcttggggttggcgtagaagat 3’
	MM102R	5’ atcctagggaggaacaacccccg 3’
MXAN1603	MM103F	5’ ataagcttgacgggctcgtgggg 3’
	MM103R	5’ cacctagggcctcccgcatcaa 3’
MXAN3778	MM104F	5’ aagctttcaccaccgggatgccgg 3’
	MM104R	5’ cctagaagggcagcttcacgacgagc 3’
2-bp mutation in -12 region of σ^{54} promoter fragments		
MXAN1286	MM101aF	5' cgcaggtagagctgaaccttccgcgccaac 3'
	MM101aR	5' gttggcgcggaaggcttagctctacctgcg 3'
MXAN1579	MM102aF	5' ccacccgcgcaaaaggattgcgccccatccc 3'
	MM102aR	5' gggatggggcgcaatccttttgcgcgggtgg 3'
MXAN1603	MM103aF	5' ccgcgcgccgagaaaccacccaggcc 3'
	MM103aR	5' ggcctgggtggtttctcggcgcgcgg 3'
MXAN3778	MM104aF	5'gacgacccggctcactcaaaatctccgtctcaacatc 3'
	MM104aR	5' gatgttgagacggagattttgagtgagccgggtcgtc 3'
2-bp mutation in -24 region of σ^{54} promoter fragments		
MXAN1286	MM101bF	5' gccttccgcgaaaacgcgcgattccggtgga 3'
	MM101bR	5' tccaccggaatcgcgcgttttcgcggaaggc 3'
MXAN1579	MM102bF	5' cgcaaggattgcgaaccatccccgcccg 3'
	MM102bR	5' cgggcggggatggttcgcaatccttgcg 3'
MXAN1603	MM103bF	5' accacccaggaaacttcgggcgacacccaga 3'
	MM103bR	5' tctgggtgtcgcccgaagtttcctgggtggt 3'
MXAN3778	MM104bF	5' cgcaatctccgtcaaaacatccccctcctgtcg 3'
	MM104bR	5' cgacaggaggggatgttttgacggagattgcg 3'
1-bp deletion in the spacer region of σ^{54} promoter fragments		
MXAN1286	MM101bF	5' gcgttggcgcggaaagcttaacgccg 3'
	MM101bR	5' cggcgttaagctttccgcgccaacgc 3'
MXAN1579	MM102bF	5' gggatggggcgcaataagcttggggttg 3'
	MM102bR	5' caaccccaagcttattgcgccccatccc 3'

MXAN1603	MM103bF	5' gaagtggcctgggaagcttgacgggc 3'
	MM103bR	5' gcccgtcaagcttcccaggccacttc 3'
MXAN3778	MM104bF	5' gggatgttgagacgggattgcgagtgagcc 3'
	MM104bR	5' ggctcactcgcaatcccgtctcaacatccc 3'
Amplification of 5'-labelled Cy5 WT		
σ^{54} promoter fragment for EMSAs		
Cy5-P_{1286}	MM109F	5' Cy5-gctcgtgtcgcgtccctacatcca 3'
	MM109R	5' acagggacagcgtccacacgacc 3'
Cy5-P_{1579}	MM110F	5' Cy5-gcgtcttcttgtccgtaccgcgga 3'
	MM110R	5' atacacccgaaagggaaggaacgccg 3'
Cy5-P_{3778}	MM110F	5' Cy5-aagttggtctcgaagcgcacgtcgc 3'
	MM110R	5' cagcaggcgctcgaggtcctcaa 3'
600-bp internal fragment of *nla28*	MM301F	5' tggacagcatccagcagggcgcct 3'
in pMM301	MM301R	5' ccacccgcgtcttgaggtccttgtt 3'
Fragment for Nla28-DBD	MM302F	5' ctggcgctcaacgtgacgggcgt 3'
expression	MM302R	5' acctgcaggctacgactcggcctccg 3'

5.4.4 Site-directed mutations

Site-directed mutations in putative σ^{54} promoter elements were generated using the Quick Lightning Mutagenesis Kit from Agilent Technologies and the manufacturer's protocol. Briefly, promoter fragments containing the putative σ^{54}-RNA polymerase binding site in the -12 and -24 regions and the upstream Nla28 tandem repeat binding site were cloned into pCR 2.1 TOPO vector (Invitrogen). Mutations in the -12 region, the -24 region, the spacer between the -12 and -24 regions were generated using primers carrying the appropriate nucleotide changes (**Table 5.2**), plasmids containing the promoter fragments and PfuUltra DNA polymerase. Parental plasmid DNA was removed by digesting with DpnI and transformed into *E. coli* for conversion into duplex form. Plasmid-borne

promoter mutations were verified by DNA sequence analysis. Promoter fragments carrying Nla28 binding site mutations (P_{EM1286}, P_{EM1579} and P_{EM3778}) were synthesized by IDT; the first putative Nla28 half binding site in each promoter was changed to AAAAAAAA. The mutant promoter fragments were then subcloned into the promoterless *lacZ* expression vector pREG1727 [47], introduced into *M. xanthus* strains and analyzed as described below.

5.4.5 In vivo analysis of wild-type and mutant promoters

Wild-type and mutant MXAN1286, MXAN1579, MXAN1603 and MXAN3778 promoter fragments were cloned into the promoterless *lacZ* expression vector pREG1727 to create *lacZ* transcriptional fusions [47]. The plasmids were introduced into strain DK1622 or a derivative of strain DK1622 carrying an insertion in the *nla28* gene, and cells carrying a plasmid integrated at the Mx8 phage attachment site in the chromosome were identified via PCR. The in vivo activities of wild-type and mutant promoters were determined by measuring the specific activities of β-galactosidase in cells developing in submerged cultures for 1, 2, 6, 12 or 24 hours, or growing in CTTYE broth for various amounts of time[48, 49]. Mean wild-type and mutant promoter activities were compared using a two-way analysis of variance (ANOVA) and Tukey's multiple comparisons *post hoc* tests, considering two independent variables in each group. The significance level was set at $p < 0.05$ or lower. GraphPad Prism software v8.4 (GraphPad Software, La Jolla, CA, USA) was used for all analyses.

5.4.6 Expression and purification of Nla28-DBD

A fragment of the *nla28* gene corresponding to the Nla28 DNA binding domain (Nla28-DBD) [30] was PCR amplified using gene-specific primers (**Table 5.2**), and then cloned into the pMAL-c5x vector. The resulting plasmid, which creates an N-terminal Maltose Binding Protein (MBP) fusion to Nla28-DBD, was introduced into *E. coli* strain BL21 (DE3) using electroporation. Cells containing the Nla28-DBD expression plasmids were grown in rich LB broth to a density of 2 x 10^8 cells/ml. Protein expression was induced by the addition of 0.3 mM IPTG to the culture and the subsequent incubation of the culture for 12 hours at 15 ℃. Cells were pelleted via centrifugation and resuspended in 25 ml column buffer (20 mM Tris-HCl, 200 mM NaCl, 1mM EDTA, 5 U/ml DNase I) per liter of culture. The resuspended cells were lysed by a combination of freeze-thawing and sonication and pelleted by centrifugation. The crude extract (supernatant) containing Nla28-DBD was diluted by adding 125 ml of cold column buffer to every 25 ml aliquot of crude extract. For purification of Nla28-DBD from the diluted crude extract, 15 ml of amylose resin was placed in a 2.5 x 10 cm column and the amylose resin was washed with 250 ml cold column buffer. Diluted crude extract containing Nla28-DBD was loaded onto amylose columns at a flow rate of 5 ml/min and washed with 600 ml cold column buffer at a flow rate of 10 ml/min. Nla28-DBD was eluted using100 ml cold column buffer containing 10 mM maltose; the flow rate was 5 ml/min and 20 fractions containing 5 ml were collected. The presence of eluted Nla28-DBD was detected by UV absorbance at 280 nm. Nla28-DBD- containing fractions were pooled and incubated with 1 mg

of Factor Xa at 4 ℃ overnight to cleave the MBP tag. Nla28-DBD was separated from MBP and concentrated to about 1 mg/ml using Amicon Ultra centrifugal filter units (EMD Millipore). SDS-PAGE and Bradford assays were used to determinate the purity and concentration of Nla28-DBD.

5.4.7 Electrophoretic mobility shift assays (EMSAs)

Purified Nla28-DBD was expected to bind to wild-type MXAN1286, MXAN1579 and MXAN3778 promoter fragments carrying a putative Nla28 tandem repeat binding site. Using the 5'Cy5-labelled oligonucleotides shown in Table 5.2, Cy5-labelled MXAN1286, MXAN1579 and MXAN3778 promoter fragments (Cy5-P_{1286}, Cy5-P_{1579} and Cy5-P_{3778}) were generated via PCR; each promoter fragment contained a putative wild-type Nla28 binding site. Three mutant derivatives of these 5' Cy5-labelled promoter fragments (Cy5-$P_{mut1286}$, Cy5-$P_{mut1579}$ and Cy5-$P_{mut3778}$) were synthesized; the first putative Nla28 half binding site in each promoter was changed to AAAAAAAA. All PCR-generated and synthesized promoter fragments were gel-purified and used in subsequent EMSAs. In EMSA reactions, 2 μM purified Nla28-DBD was incubated with 1 ng of 5' Cy5-labelled wild-type promoter fragment (Cy5-P_{1286}, Cy5-P_{1579} and Cy5-P_{3778}) or 5' Cy5-labelled mutant promoter fragment (Cy5-$P_{mut1286}$, Cy5-$P_{mut1579}$ and Cy5-$P_{mut3778}$) in EMSA buffer (25 mM Tris acetate, 8.0 mM magnesium acetate, 10 mM KCl, 1 mM DTT, pH 8.0) for 30 mins at 30 °C. The samples were then analyzed using PAGE under non-denaturing conditions and imaged using a Bio-Rad imager.

5.5 Results

5.5.1 Identifying putative σ^{54} promoter elements in M. xanthus natural product gene clusters

In a previous study [38], the algorithm developed by Studholme *et al.*[50] was used to examine whether the σ^{54} system might be a common regulator of bacterial natural product genes. Namely, 180 annotated PK and NRP gene clusters from 58 bacterial species were analyzed for sequences that closely match the σ^{54} promoter consensus in the -12 region and in the -24 region (ie., the regions of σ^{54}-RNA polymerase binding). The results, which uncovered 124 clusters with at least one σ^{54} promoter based on consensus matching, supported the idea that the σ^{54} system might be a general regulator of bacterial natural product genes.

The goal of the work presented here was to examine whether a major producer of bacterial natural products (*M. xanthus*) uses the σ^{54} system for transcription of PK and NRP gene clusters, as predicted in the bioinformatics analysis of Stevens *et al.* [38]. We focused on the putative σ^{54} promoters upstream of MXAN1286, MXAN1579, MXAN1603 and MXAN3778, as these genes were linked to natural product biosynthetic genes/clusters. Namely, the MXAN1286 gene is embedded in a putative NRP gene cluster and flanked by genes that are likely to be involved in NRP biosynthesis. MXAN1579 and MXAN1603 are predicted to be part of another *M. xanthus* NRP gene cluster. MXAN1603 is a putative non-ribosomal peptide synthetase gene. MXAN1579 is embedded in the cluster and flanked by genes such as MXAN1603 that are likely

to be involved in NRP biosynthesis. MXAN3778 is adjacent to a putative hybrid non-ribosomal peptide synthetase/type I polyketide synthase gene and possibly part of the same operon [38].

The MXAN1286, MXAN1579, MXAN1603 and MXAN3778 genes were also identified as potential targets of the EBP Nla28 [30, 39]; 8-bp repeat sequences, which are close matches to the consensus Nla28 half binding site, were identified upstream of the putative -12 and -24 regions (**Fig. 5.1**). It is notable that six of the eight putative Nla28 half binding sites and three of the four putative σ^{54}-RNA polymerase binding sites are located within protein coding sequences. Indeed, residence in an intragenic region is common among the putative PK and NRP σ^{54} promoters identified in the *M. xanthus* genome (**Fig. 5.2**, **Fig. 5.3, and Table 5.3**), and among the σ^{54} promoters known to be regulated by the EBP Nla28[30, 39]. It is also noteworthy that many of the putative σ^{54} promoter elements are located within operons; they might serve as internal promoters (**Fig. 5.2**, **Fig. 5.3, and Table 5.3**).

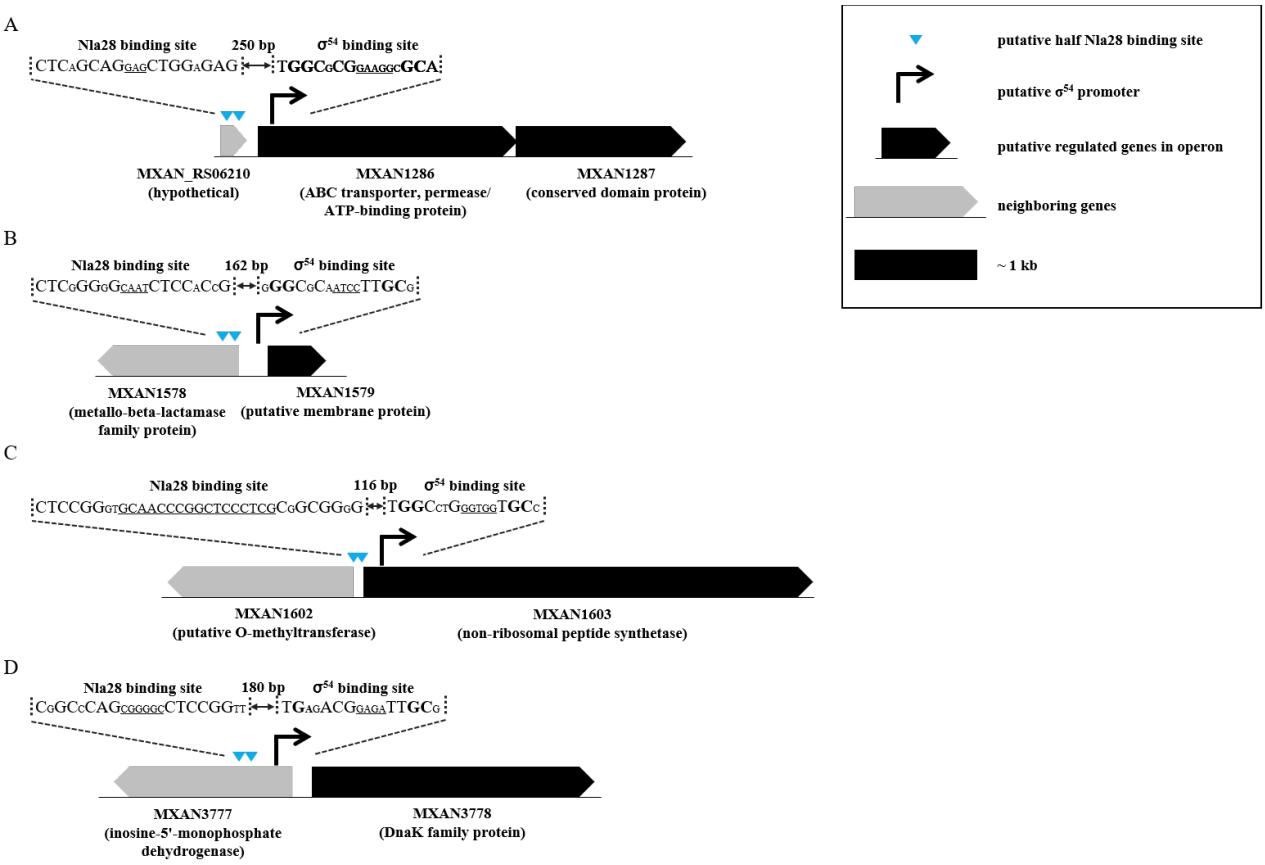


Figure 5.1 The promoter regions of the MXAN1286, MXAN1579, MXAN1603 and MXAN3778 natural product loci

Nucleotides the match those in the consensus Nla28 binding site or the consensus σ^{54} RNA polymerase binding site are relatively large. The conserved GC dinucleotide in -12 region and the conserved GG dinucleotide in -24 region of the putative σ^{54} RNA polymerase binding sites are in bold. The underlined nucleotides represent the spacers between the two half Nla28 binding sites or the spacers between -12 and -24 promoter regions.

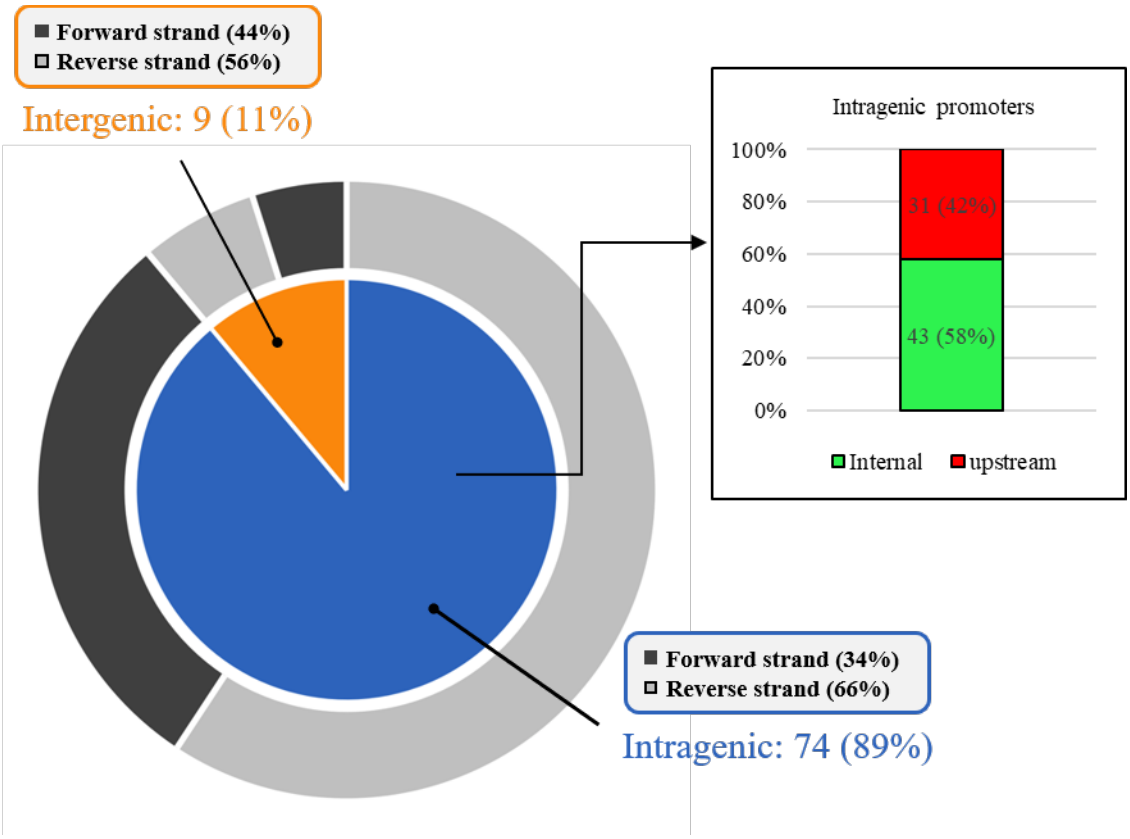
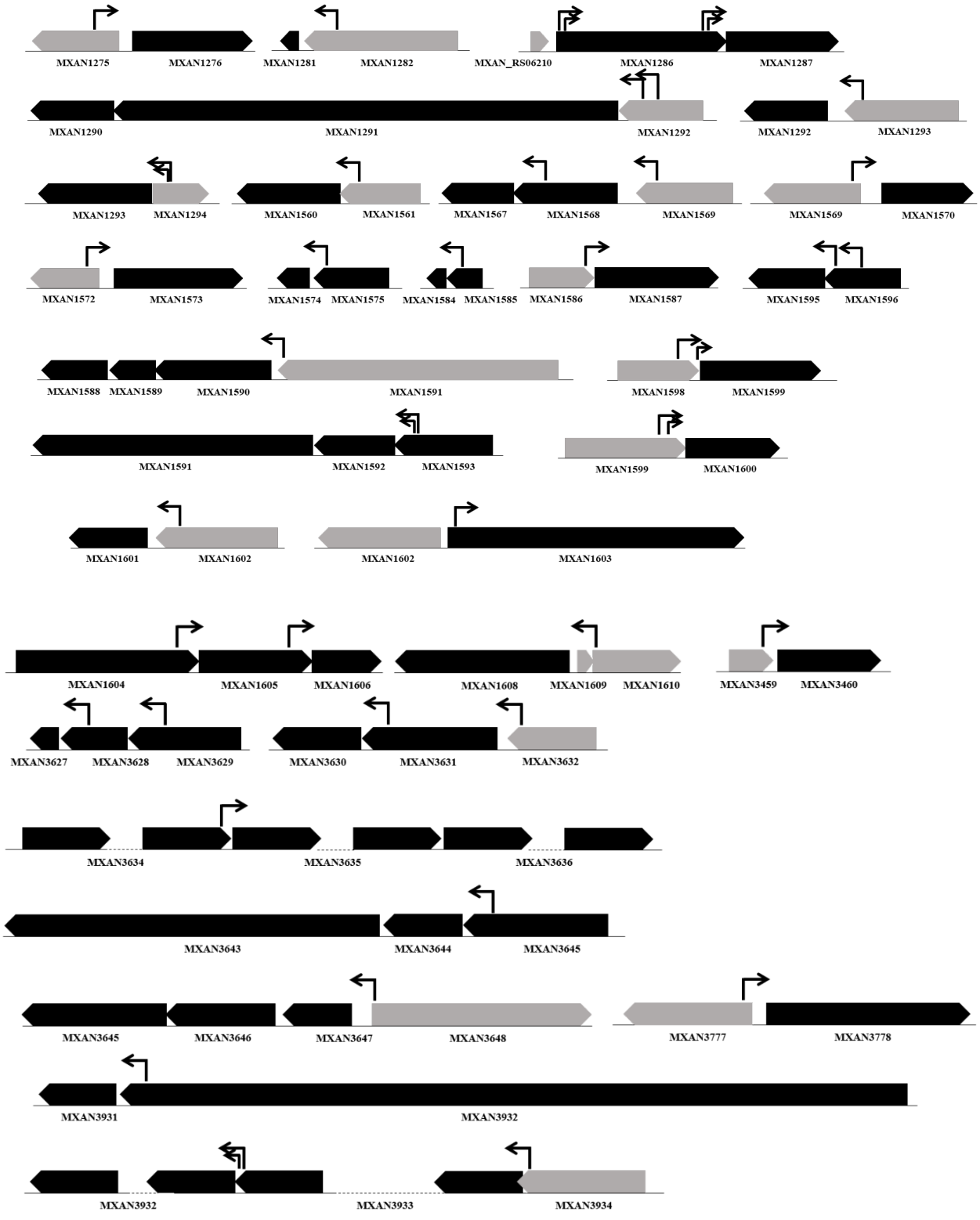
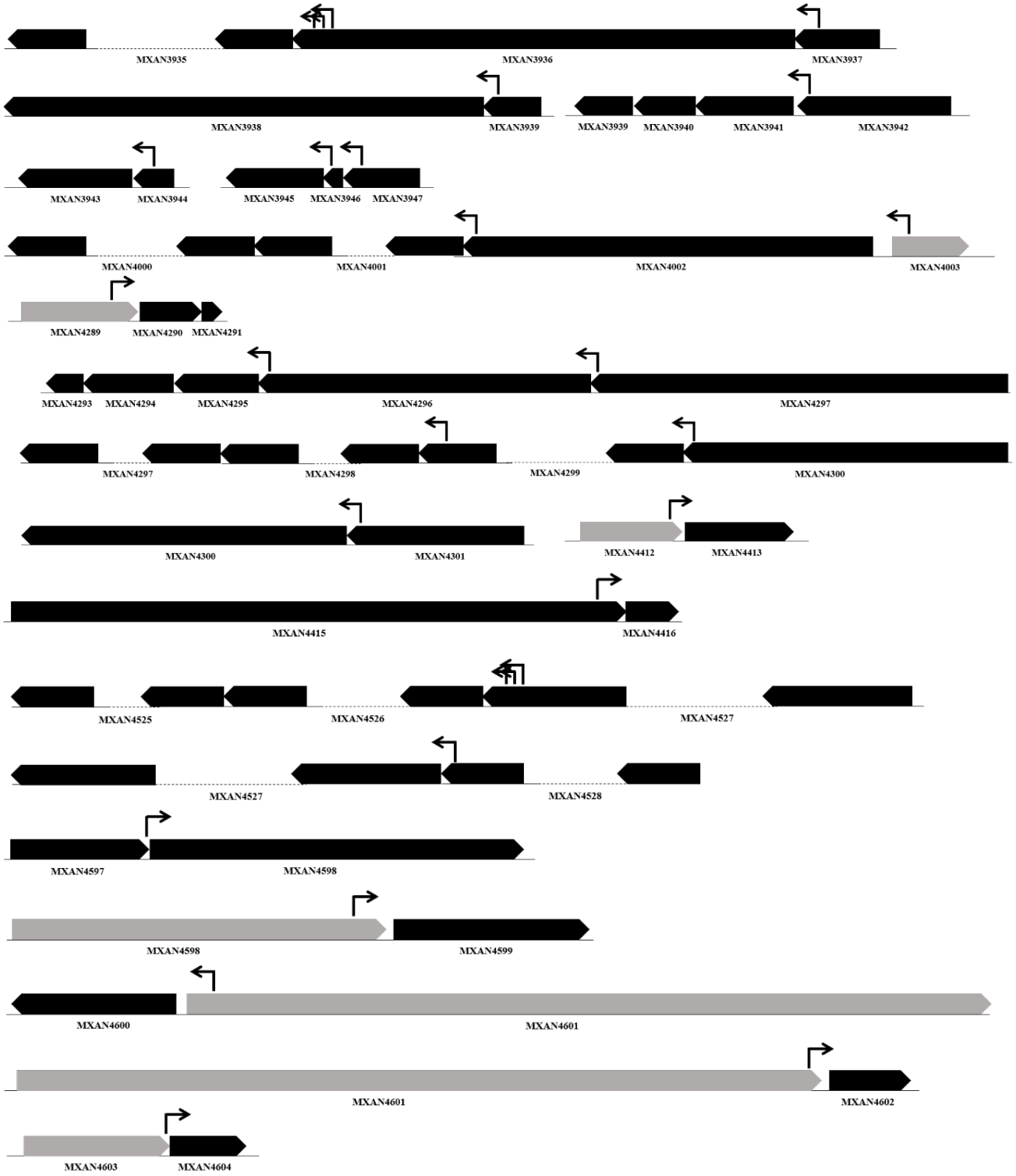


Figure 5.2 Locations of putative PK/NRP σ^{54} promoters identified in the *M. xanthus* genome

Of the 83 putative PK/NRP σ^{54} promoters identified in *M. xanthus* genome sequence, 74 (89%) are located in protein coding sequences (intragenic promoters) and 9 (11%) are located in non-coding sequences (intergenic promoters). Of the 74 intragenic promoters, 43 are located within a protein coding sequence in an operon or within the protein coding sequence of a single gene (internal promoters), and 31 are located in the protein coding sequence of an upstream gene (upstream promoters).

A


MXAN1275
MXAN1276
MXAN1281
MXAN1282
MXAN_RS06210
MXAN1286
MXAN1287
MXAN1290
MXAN1291
MXAN1292
MXAN1292
MXAN1293
MXAN1293
MXAN1294
MXAN1560
MXAN1561
MXAN1567
MXAN1568
MXAN1569
MXAN1569
MXAN1570
MXAN1572
MXAN1573
MXAN1574
MXAN1575
MXAN1584
MXAN1585
MXAN1586
MXAN1587
MXAN1595
MXAN1596
MXAN1588
MXAN1589
MXAN1590
MXAN1591
MXAN1598
MXAN1599
MXAN1591
MXAN1592
MXAN1593
MXAN1599
MXAN1600
MXAN1601
MXAN1602
MXAN1602
MXAN1603
MXAN1604
MXAN1605
MXAN1606
MXAN1608
MXAN1609
MXAN1610
MXAN3459
MXAN3460
MXAN3627
MXAN3628
MXAN3629
MXAN3630
MXAN3631
MXAN3632
MXAN3634
MXAN3635
MXAN3636
MXAN3643
MXAN3644
MXAN3645
MXAN3645
MXAN3646
MXAN3647
MXAN3648
MXAN3777
MXAN3778
MXAN3931
MXAN3932
MXAN3932
MXAN3933
MXAN3934


MXAN3935
MXAN3936
MXAN3937
MXAN3938
MXAN3939
MXAN3939
MXAN3940
MXAN3941
MXAN3942
MXAN3943
MXAN3944
MXAN3945
MXAN3946
MXAN3947
MXAN4000
MXAN4001
MXAN4002
MXAN4003
MXAN4289
MXAN4290
MXAN4291
MXAN4293
MXAN4294
MXAN4295
MXAN4296
MXAN4297
MXAN4297
MXAN4298
MXAN4299
MXAN4300
MXAN4300
MXAN4301
MXAN4412
MXAN4413
MXAN4415
MXAN4416
MXAN4525
MXAN4526
MXAN4527
MXAN4527
MXAN4528
MXAN4597
MXAN4598
MXAN4598
MXAN4599
MXAN4600
MXAN4601
MXAN4601
MXAN4602
MXAN4603
MXAN4604

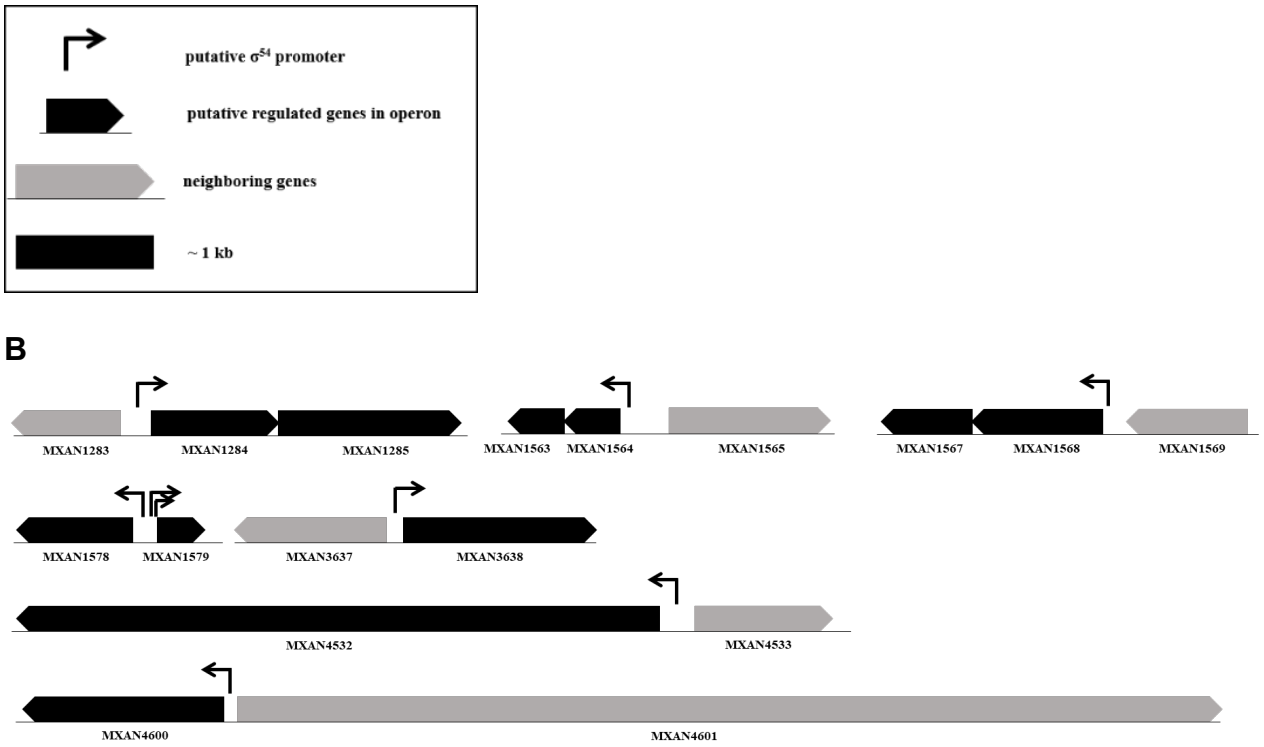


Figure 5.3 Locations of putative σ^{54} promoters in *M. xanthus* PK and NRP gene clusters

Of the 83 putative σ^{54} promoters identified in *M. xanthus* PK and NRP gene clusters[38], 74 are predicted to be intragenic (A) and 9 are predicted to be intergenic (B).

Table 5.3 Putative intragenic and intergenic σ^{54} promoters in *M. xanthus* PK/NRP gene clusters

Intragenic			
Regulated genes/operons	**Functions**	**Putative promoter −24/−12 sequence[a]**	**Position[b]**
MXAN1276	Glutamate-cysteine ligase family 2 protein	GT**GG**CGCGGCCCTT**GA**A	-476
MXAN1281	hypothetical protein	CT**GG**AACGCGTGCA**GC**A	-712
MXAN1286-1287	ABC transporter/ATP-binding protein;	TT**GG**ACCGGCTCGT**GC**G	+15
	hypothetical protein	TT**GG**CGCGGAAGGC**GC**A	+111
MXAN1287	hypothetical protein	AT**TG**CCCGCGCCCT**GC**T	-308
		GA**GG**CCCAACTGTT**GC**G	-233
MXAN1290-1291	Phosphotransferase; non-ribosomal peptide	CT**GG**CACGTGACGT**GT**T	-285
	synthetase	GT**GG**TGCGGCGGAT**GC**A	-488
MXAN1292	hypothetical protein	CT**GG**CGCGCGAGGA**AC**T	-521
MXAN1293	protoporphyrinogen oxidase	CC**GG**TACACACGTT**GG**T	-285
		GA**GA**CACCGGCGTT**GT**T	-324
MXAN1560	class I aminotransferase	GT**GG**CACGTCGCTG**GT**G	-284
MXAN1567	urea amidolyase-like protein	CT**GG**CGCGGTTGCT**GC**G	-407
MXAN1567-1568	hydrolase family protein;	CT**GC**CACGCGCCTG**GC**G	-443
MXAN1570	class V aminotransferase	AT**GA**CTCGCAGCTT**GC**G	-293
MXAN1573	AMP-binding domain protein	CT**GG**CCTCCACGGT**GG**T	-409
MXAN1574	TfoX domain-containing protein	CT**GG**CGCGCGCCGG**GG**T	-212
MXAN1584	hypothetical protein	GT**GG**CGCGGCCGCT**GC**T	-182
MXAN1587	hypothetical protein	GT**GG**GACATCCGGT**GC**G	-146
MXAN1588-1590	Aminotransferase; putative para-aminobenzoate	GA**GG**CACGCCGCAC**GC**T	-122
MXAN1591-1592	hypothetical proteins	CT**GC**TCTGCACGGA**GC**T	-217
		CT**GG**AAGTGGACCT**GC**T	-229
MXAN1595	hypothetical protein	AT**GG**CACGGGGGAC**GA**A	-115
		GA**GG**ATGACCGGCT**GC**T	-413
MXAN1599	putative methyltransferase	GT**GG**CCAGCCGGGC**GC**G	-281
		GT**GG**TGCCGCTGCT**GC**T	-44
MXAN1600	class I aminotransferase	GA**GG**CCCGCTCCGA**GC**T	-402
		CT**GG**CTCGGGCGTT**GC**G	-270
MXAN1601	fatty acid desaturase family protein	GT**GG**CCCGGGCACT**GC**G	-384
MXAN1603	putative non-ribosomal peptide synthetase	GT**GG**CCTGGGTGGT**GC**C	+84
MXAN1605	putative permease	CT**GG**AAGGGTCGCT**GC**T	-265
MXAN1606	hypothetical protein	GT**GG**CCGCCTTCAT**GC**T	-260
MXAN1608	hypothetical protein	TA**GG**AACGCACGAC**CC**T	-286
MXAN3460	Gfo/Idh/MocA family oxidoreductase	CT**GG**TCCACATCAT**GC**T	-182
MXAN3627	hypothetical protein	AT**GG**CGGAAGAGCT**GC**T	-305

MXAN3627-3628	hypothetical protein; non-ribosomal peptide	CT**GC**CACGCCAGCA**GC**T	-383
MXAN3630	polyketide synthase type I	CT**GG**CGGACATCTG**GC**A	-315
MXAN3630-3631	polyketide synthase type I	AA**GG**AACAGCTGTT**GC**G	-290
MXAN3635-3636	non-ribosomal peptide synthase/polyketide synthase	AA**GG**CGTGCGCATT**GC**T	-193
MXAN3643-3644	non-ribosomal peptide synthase; isochorismatase	GT**GG**AAAACCATCT**GC**T	-297
MXAN3645-3647	2,3-dihydroxybenzoate-AMP;	GT**GG**CACGCGAAAT**CC**C	-254
MXAN3778	DnaK family protein	TT**GA**GACGGAGATT**GC**G	-284
MXAN3931	hypothetical protein	GT**GG**GCCGGTTCCG**GC**T	-289
MXAN3932	polyketide synthase	CA**GA**CACGCCCATC**GC**T	-58
		TT**GA**AAAACCGGTT**GC**T	-75
MXAN3933	mixed type I polyketide synthase-peptide	GT**GG**TACTTCGCTT**AC**T	-222
MXAN3935	non-ribosomal peptide synthase/polyketide synthase Ta1	CT**GG**CACGAACCCT**GG**G	-205
		TT**GG**CGCTGAGCGT**GC**C	-301
		TT**GG**CTCGCGTCTG**GC**T	-413
MXAN3936	polyketide synthase	TT**GG**TCCGCCGGGC**GG**T	-318
MXAN3938	polyketide synthase	CT**GT**CGCTCGTGAC**GC**T	-140
MXAN3939-3941	enoyl-CoA hydratase; polyketide beta-ketoacyl:acyl	CT**GG**CTCTGCAACT**GC**G	-229
MXAN3943	cytochrome P450 family protein	GT**GG**CGCATGCCTT**TC**T	-436
MXAN3945	polyketide TA biosynthesis protein TaF	CT**GG**ACAGCCTGCG**GC**T	-82
MXAN3945-3946	polyketide TA biosynthesis protein TaF; putative	CT**GT**CGCGCGTGGT**GC**T	-205
MXAN4000-4001	non-ribosomal peptide synthase/polyketide	CT**CG**CCCGCAGGTT**GC**A	-161
MXAN4002	nonribosomal peptide synthetase	AG**GG**CACGGCCCGT**GA**T	-451
MXAN4290-4291	putative thioesterase; hypothetical protein	CT**GG**AGCGCGTGCT**GC**T	-354
MXAN4293-4295	hypothetical protein; transporting ATPase;	AT**GT**CGCCGTGCTT**GC**A	-118
MXAN4296	non-ribosomal peptide synthetase	CT**GG**GAGTGGACCT**GC**T	-134
MXAN4297-4298	polyketide synthase type I	AT**GC**GTCGGCAGTT**GC**T	-364
MXAN4299	non-ribosomal peptide synthase/polyketide	AT**GG**CGCTCGAGTT**GC**G	-164
MXAN4300	polyketide synthase type I	CT**GG**CGAAGCGGCT**GC**T	-167
MXAN4413	hypothetical protein	CT**GG**ACGACCGGTT**GG**T	-186
MXAN4416	cephalosporin hydroxylase family protein	CT**GG**GCCACGGATT**GC**T	-332
MXAN4525-4526	non-ribosomal peptide synthase; polyketide synthase type I	AT**GG**CGCTGGAGCT**GC**G	-243
		AT**GG**CGCTGCTGTT**GG**A	-359
		GT**GG**CAGGGCAGGT**GC**G	-452
MXAN4527	polyketide synthase	CT**GG**CTGAACAACT**GC**A	-156
MXAN4598	non-ribosomal peptide synthase	AT**GG**AACGGCGCAT**CC**T	-41
MXAN4599	M28 family peptidase	CT**GG**GCTACGGGGT**GC**T	-534
MXAN4600	radical SAM domain-containing protein	GT**GG**CCTGGGCGTC**GG**T	-493
MXAN4602	hypothetical protein	CT**GG**CCCGTCTGGA**GC**T	-241
MXAN4604	hypothetical protein	AT**GT**CGCTCGACTT**GC**T	-62

Intergenic Regulated genes/operons	Functions	Putative promoter[a] −24/−12 sequence	Position[b]
MXAN1284-1285	2-isopropylmalate synthase/homocitrate synthase family protein; tryptophan halogenase	AG**GG**CAAGGCATTT**CC**A	-121
MXAN1563-1564	alkyl hydroperoxide reductase C	CT**GG**CACGGTGACT**GC**T	-69
MXAN1567-1568	hydrolase family protein; LamB/YcsF	CT**GG**CACGCCAGCG**TC**T	-44
MXAN1578	metallo-beta-lactamase family protein	AC**GG**CGCAGCGCTT**GC**T	-45
MXAN1579	hypothetical protein	GG**GG**CGCAATCCTT**GC**G	-104
		GG**GG**CGCAGGTCTT**GC**G	-75
MXAN3638	M19 family peptidase	CT**GG**TACTTCGAGT**GC**A	-32
MXAN4532	non-ribosomal peptide synthase	GT**GG**CACAAGCTGC**GC**T	-173
MXAN4600	radical SAM domain-containing protein	GC**GG**TAAAGTCTTT**GC**T	-41

[a]-12 and -24 regions of putative σ^{54} promoters were identified in a previous study [37].

[b]Distance between the −12 region of the putative σ^{54} promoter and the predicted initiation codon.

5.5.2 Mutations in the putative -12 region, -24 region or spacer region impair the in vivo activities of natural product promoters

σ^{54} promoters typically have a GC dinucleotide in the −12 region and a GG dinucleotide in the −24 region [1, 2]. These dinucleotides and the 4-bp spacer between the -12 and -24 regions are often referred to as the hallmarks of σ^{54} promoters. Indeed, the putative σ^{54} promoters in the MXAN1286, MXAN1579 and MXAN1603 natural product loci appear to have these hallmarks (**Fig. 5.1**). As for the σ^{54} promoter identified in the MXAN3778 locus, one hallmark variation is apparent. Namely, the -24 region has GA instead of a GG dinucleotide. Despite this variation in -24 region dinucleotide, MXAN3778 was classified as a potential σ^{54} promoter, as a 1-bp change in either the GC or GG dinucleotide has been identified in a number of characterized σ^{54} promoters, including the σ^{54} promoters in the *M. xanthus asgE*, *spi* and *nla6* loci [30, 51-53].

To confirm that the natural product loci have bona fide σ^{54} promoter elements, we analyzed the putative σ^{54} promoter hallmarks via mutational analysis. In particular, a 446-bp DNA fragment of the MXAN1286 promoter region, a 402-bp fragment of the MXAN1579 promoter region, a 518 bp fragment of the MXAN1603 promoter region and a 600-bp fragment of the MXAN3778 promoter region were used to generate mutations in the putative -12, -24 and the spacer regions. We should note that the MXAN1286 promoter fragment contains both of the putative σ^{54} promoters located in the 5' end of the MXAN1286 gene and the MXAN1579 promoter fragment contains the two putative intergenic σ^{54} promoters that are oriented in the appropriate direction with respect to MXAN1579 transcription (**Table 5.3 and Fig. 5.3**). However, we focused the mutational analysis on the putative σ^{54} promoter that is the farthest downstream in the MXAN1286 coding sequence (+111 bp; see **Table 5.3**) and the putative σ^{54} promoter that is farthest upstream of the predicted transcriptional start site of MXAN1579 gene (-104 bp; see **Table 5.3**) after considering a number of factors, including the presence of σ^{54} promoter hallmarks, the overall match to the σ^{54} promoter consensus sequence [TGGCACG-4N-TTGC(T/A)] and the distance from putative Nla28 binding sites [1, 2]. The mutations that we generated are the following: the GC dinucleotide in the −12 region was replaced with a TT, the GG (GA in MXAN3778) dinucleotide in the −24 region was replaced with a TT, or 1 bp in the spacer between the −12 and −24 regions was deleted. Subsequently, wild-type and mutant promoter fragments were fused to the promoterless *lacZ* gene in plasmid pREG1727 [47] and the *lacZ* transcriptional fusion plasmids

were introduced into wild-type *M. xanthus* strain DK1622 (the plasmids integrated at the Mx8 phage attachment site in the chromosome).

Wild-type and mutant promoter activities during growth in CTTYE broth and fruiting body development in MC7 starvation buffer were inferred from the levels of *lacZ* expression. As shown in Figure 5.4A-C, the in vivo activities of the MXAN1286, MXAN1579 and MXAN3778 promoters increased about 4.6-fold, 1.8-fold and 2-fold, respectively, during growth in CTTYE broth. It is notable that peak levels of promoter activity occurred at the highest cell densities, which correspond to the transition into stationary phase and presumably nutrient depletion. This of course agrees with the data shown in Figure 5.4D-F, which revealed a 1.8- to 2.5-fold increase in the vivo activities of the three promoters during development in MC7 starvation buffer. Mutations in the -12 region, the -24 region and spacer dramatically reduced (about 3.2- to 11.1-fold) the activities of these promoters at all cell densities during growth and time points in development. Thus, mutations in the putative sites of σ^{54}-RNA polymerase binding substantially impacted the activities of the MXAN1286, MXAN1579 and MXAN3778 promoters in growing and developing cells, supporting the prediction that the three natural product loci use σ^{54} promoter elements for transcription.

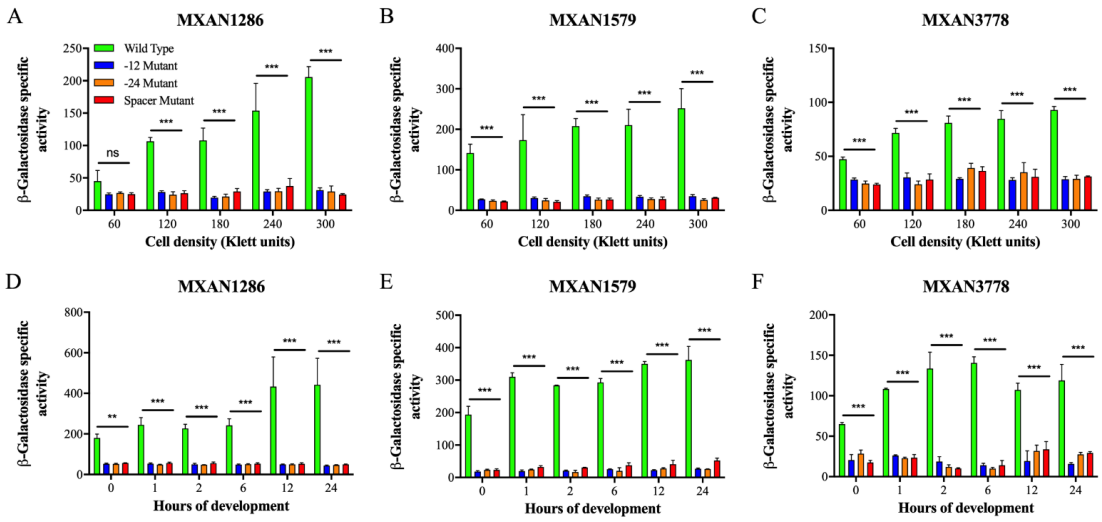


Figure 5.4 In vivo activities of wild-type MXAN1286, MXAN1579 and MXAN3778 promoters and derivatives of the promoters carrying a mutation in the putative -12 region, -24 region or spacer region

Wild-type and mutant fragments of the MXAN1286, MXAN1579 and MXAN3778 promoters were cloned into a *lacZ* expression vector and transferred to the wild-type *M. xanthus* strain DK1622. At various cell densities during growth (A-C) and time points during development (D-F), β-galactosidase-specific activities (defined as nanomoles of ONP produced per minute per milligram of protein) in cells carrying a wild-type or a mutant promoter fragment were determined. (N=3 at each density or time point; error bars are standard deviations of the means; ***p < 0.001; **p < 0.01; *p < 0.05 for in vivo activities of mutant promoters versus wild-type promoters).

In contrast to the other promoters, the MXAN1603 promoter only showed a slight increase (about 1.3-fold) in activity during growth in CTTYE (**Fig. 5.5A**) and during development in MC7 starvation buffer (**Fig. 5.5B**). Furthermore, mutations in the -12 region, the -24 region and the spacer caused a modest, but statistically significant decrease (P < 0.001) in promoter activity at all cell densities during growth and after 1h of development (**Fig. 5.5**). Thus, mutations

in the putative σ^{54}-RNA polymerase binding site in the MXAN1603 promoter region only had a modest impact on growth-related and developmental activities. Our interpretation of this result is that the fragment of the MXAN1603 promoter region contains an unidentified promoter element that is responsible for the majority of the observed development and growth-related activities and that the putative σ^{54}-type promoter element only makes a minor contribution to the observed activities.

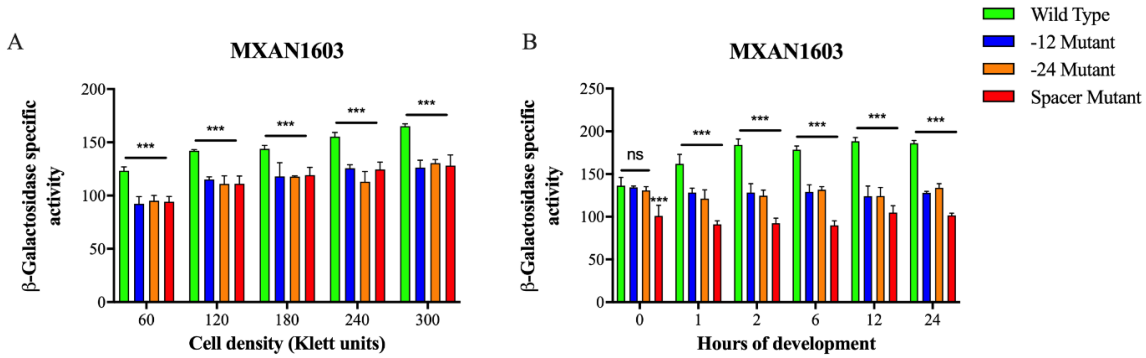


Figure 5.5 In vivo activities of the wild-type MXAN1603 promoter and derivatives of the promoter carrying a mutation in the putative -12 region, -24 region or spacer region

Wild-type and mutant fragments of the MXAN1603 promoter were cloned into a *lacZ* expression vector and transferred to the wild-type *M. xanthus* strain DK1622. At various cell densities during growth (A) and time points during development (B), β-galactosidase-specific activities in cells carrying a wild-type or a mutant promoter fragment were determined. (N=3 per group; Error bars: Mean ± SD; ***p < 0.001; **p < 0.01; *p < 0.05 for in vivo activities of mutant promoters versus wild-type promoters).

5.5.3 The in vivo activities of natural product promoters are impacted by inactivation of the nla28 gene

EBPs are essential for transcription at σ^{54} promoters, as EBP-mediated ATP hydrolysis opens the σ^{54}-RNA polymerase promoter complex so that transcription can initiate[3-5]. Since the σ^{54} promoters in the MXAN1286, MXAN1579, and MXAN3778 loci were identified as potential targets of Nla28, we determined whether the activities of wild-type promoter fragments are reduced as predicted in a mutant containing an inactivated *nla28* gene[25] (Note that the MXAN1603 σ^{54} promoter was not analyzed further because it is unlikely to be the primary promoter used during growth or development). MXAN1286, MXAN1579, and MXAN3778 promoter activities in wild-type and *nla28* mutant cells grown in CTTYE broth are shown in Figure 5.6A-C. As predicted, inactivation of *nla28* abolished the growth phase regulation of all three promoters and caused about 3.1- to 4-fold reduction in peak promoter activities at the highest cell densities. Inactivation of *nla28* also abolished the developmental activities of the promoters, as the promoters did not show the typical increases in activities when *nla28* cells were placed in MC7 starvation buffer (**Figure 5.6D-F**). Indeed, the peak developmental promoter activities in *nla28* mutant cells were reduced about 3.4- to 5.0-fold relative to the corresponding peak activities in wild-type cells. These findings indicate that Nla28 is crucial for the observed growth-related and developmental activities of the MXAN1286, MXAN1579, and MXAN3778 natural product promoters.

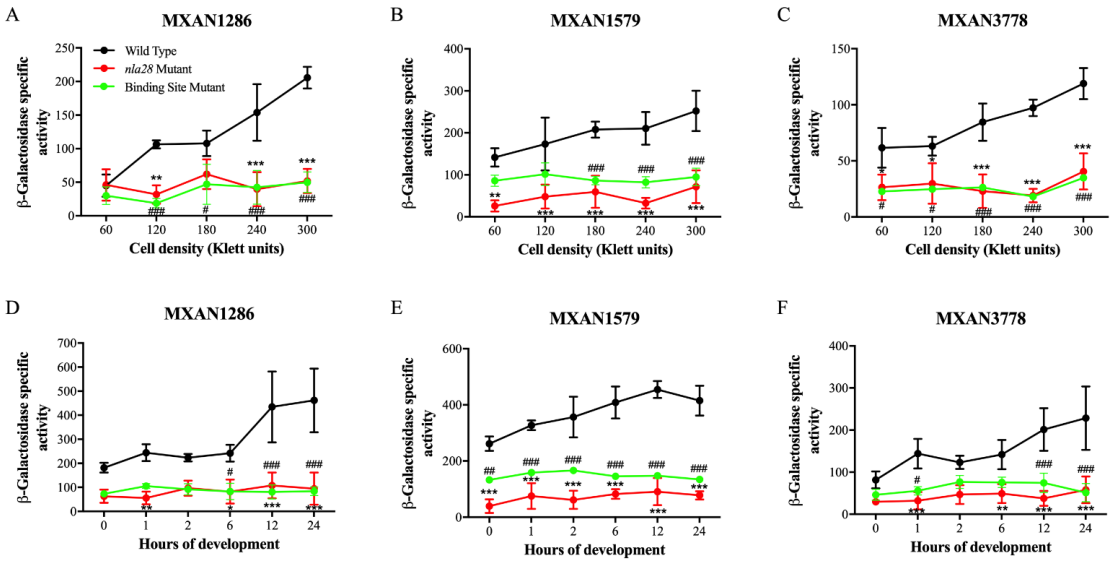


Figure 5.6 In vivo activities of MXAN1286, MXAN1579 and MXAN3778 promoters in nla28 mutant cells and when Nla28 binding sites are mutated

Wild-type fragments of the MXAN1286, MXAN1579 and MXAN3778 promoters were cloned into a *lacZ* expression vector and transferred to the wild-type *M. xanthus* strain DK1622 (Wild Type) or to a derivative of strain DK1622 with an inactivated *nla28* gene (*nla28* mutant). In addition, fragments of the MXAN1286, MXAN1579 and MXAN3778 promoters containing one mutated Nla28 half binding site were cloned into a *lacZ* expression vector and transferred to strain DK1622 (Binding site mutant). At various cell densities during growth (A-C) and time points during development (D-F), β-galactosidase-specific activities in cells carrying a wild-type or a mutant promoter fragment were determined. (N=3 at each density or time point; error bars are standard deviations of the means; ***p < 0.001; **p < 0.01; *p < 0.05 for in vivo activities of wild-type promoters in *nla28* mutant versus wild-type cells; $^{###}$p < 0.001; $^{##}$p < 0.01; $^{#}$p < 0.05 for in vivo activities of mutant promoters versus wild-type promoters in wild-type cells).

5.5.4 Mutations in putative Nla28 half sites impact the in vivo activities of natural product promoters

As noted above, we identified 8-bp repeat sequences, which are close matches to the consensus Nla28 half binding site [CT(C/G)CG(C/G)AG], in the σ^{54} promoters under study here (see **Fig. 5.1**). To examine whether the σ^{54} promoters are directly regulated by Nla28 and to further confirm that the promoters are members of the σ^{54} family, mutations were generated in the putative Nla28 binding sites in the MXAN1286, MXAN1579, and MXAN3778 promoter fragments noted above. Namely, the distal (relative to the -12 and -24 regions) Nal28 half binding site in each promoter fragment was converted to all A nucleotides. Wild-type and mutant promoters were introduced into wild-type strain DK1622 and promoter activities during growth in CTTYE broth and development in in MC7 starvation buffer were determined (**Fig. 5.6**). The data revealed that Nla28 binding site mutations abolish the growth phase regulation of all three promoters (**Fig. 5.6A-C**). Indeed, the peak mutant promoter activities, which were observed at the highest cell density, were reduced about 2.6- to 5-fold compared to that of the corresponding wild-type promoter. Similarly, wild-type MXAN1286, MXAN1579, and MXAN3778 promoters showed increased activities during development and the Nla28 binding site mutations abolished this developmental regulation (**Fig. 5.6D-F**). Furthermore, the peak developmental activities of the mutant promoters were reduced from about 2.8- to 4.8-fold. These findings are consistent with the idea that Nla28 directly regulates the MXAN1286, MXAN1579, and MXAN3778 σ^{54} promoters, that the 8-bp repeats

that we identified are Nla28 binding sites and that Nla28 is crucial for growth-related and developmental promoter activities.

5.5.5 Purified Nla28-DBD binds to natural product promoter fragments carrying a wild-type Nla28 binding site in vitro, but not to fragments carrying a mutated Nla28 binding site

Electrophoretic mobility shift assays (EMSAs) were used to confirm that the MXAN1286, MXAN1579, and MXAN3778 natural product promoters are targets of the Nla28 EBP. In particular, we used EMSAs to determine whether the purified DNA binding domain of Nla28 (Nla28-DBD) is capable of binding a fragment of the MXAN1286 promoter, MXAN1579 promoter, and MXAN3778 promoter. Each promoter fragment, which corresponded to DNA upstream of -12 and -24 regions, contained a putative binding site for a Nla28 dimer. As shown in Figure 5.7, Nla28-DBD is capable of binding to a MXAN1286, MXAN1579 and MXAN3778 promoter fragment that has a wild-type Nla28 binding site. However, when the distal Nla28 half binding site in each promoter fragment was converted to all A nucleotides, no Nla28-DBD binding was detected (**Fig. 5.7**). These findings provide further support that the Nla28 EBP directly regulates the σ^{54} promoter elements of the MXAN1286, MXAN1579 and MXAN3778 natural product loci and that the tandem repeats that we identified in the σ^{54} promoter elements are Nla28 binding sites.

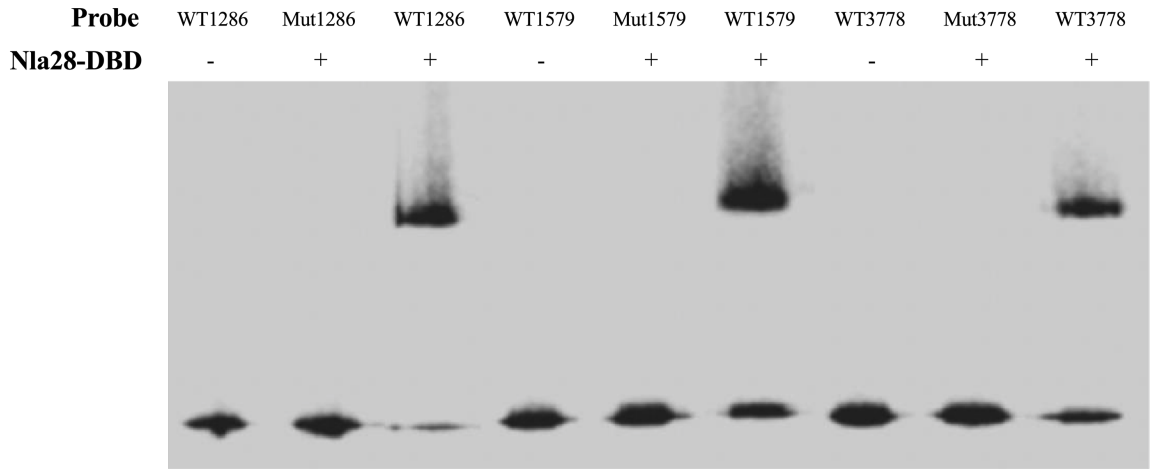


Figure 5.7 EMSAs performed with Nla28-DBD and a MXAN1286, MXAN1579 or MXAN3778 promoter fragment carrying a wild-type or mutated Nla28 binding site

Binding reactions were performed with (+) or without (-) 2 μM of purified Nla28-DBD and a Cy5 end-labeled promoter fragment containing a wild-type (WT) or mutated (Mut) Nla28 binding site.

5.6 Discussion

For decades, bacterial-derived PK and NRP natural products have been a crucial source of therapeutic agents such as antibiotics and yet little information about the regulation of these genes has been uncovered. In a notable study in 2012, Volz *et al* showed that two *M. xanthus* EBPs (HsfA and MXAN4899) are capable of binding to fragments of natural product gene promoters [40]. With this information and the preliminary data from Nla28 studies in mind, Stevens *et al*. asked if the σ^{54} system might be a common regulator natural product genes [38]. Namely, a bioinformatics analysis was used to search for putative σ^{54} promoters in the sequences of 180 PK or NRP gene clusters from 58 bacterial species. The

results, which revealed that about 70% of natural product gene clusters have at least one putative σ^{54} promoter, suggested that the σ^{54} system might indeed be a common regulator of natural product genes.

One of the goals of this study was to analyze the bioinformatics data experimentally and the putative natural product promoter targets of Nla28 seemed particularly well suited for such a study, given our knowledge of Nla28-mediated regulation. Furthermore, *M. xanthus* is an excellent system to study natural product gene regulation, as this bacterium is a major producer of PKs and NRPs and over 80 putative σ^{54} promoters were identified in the PK and NRP gene clusters of strain DK1622 [38] (**Table 5.3 and Fig. 5.3**).

Our in vivo and in vitro data indicate that three of the *M. xanthus* natural product promoter regions characterized here (MXAN1579, MXAN3778 and MXAN1286) contain bona fide σ^{54} promoter elements and that the σ^{54} promoters are targets of the EBP Nla28. Our results also indicate that the σ^{54} promoters are crucial for both growth-related and developmental activities. We propose that the fourth natural product promoter region that we characterized, MXAN1603, also contains a σ^{54} promoter. However, it seems likely that the MXAN1603 promoter region contains an unidentified promoter element that is responsible for the majority of the observed developmental and growth-related activities and the σ^{54}-type promoter only makes a minor contribution to these activities.

Previous work indicated that the Nla28 EBP is a response regulator that forms a two component signal transduction system with the membrane-bound sensor histidine kinase Nla28S [54, 55]. Nla28 begins modulating gene

expression in the early stages of starvation-induced fruiting body development [25, 30, 39], which led to the suggestion that the Nla28/Nla28S signal transduction system might be a general regulator of starvation-induced or stress-responsive genes [55]. The findings presented here are consistent with this idea. In particular, the MXAN1286, MXAN1579, and MXAN3778 σ^{54} promoters are induced in the early-middle stages of fruiting body development and the developmental activities of these promoters are dependent on Nla28 (**Fig. 5.6D-F**). The activities of the three promoters also increase during growth, with peaks corresponding to the transition into stationary phase and presumably nutrient depletion. As shown in Figure 5.6A-C, this growth-phase regulation is abolished when Nla28 is absent, further linking the Nla28/Nla28S signal transduction system to the starvation response.

The above results point to nutrient depletion/starvation as the potential trigger for Nla28/Nla28S-mediated activation of the MXAN1286, MXAN1579 and MXAN3778 promoters. We suggest that the stringent response and accumulation of (p)ppGpp might also be a key inducing factor, as accumulation of this starvation signal is known to initiate *M. xanthus* fruiting body development [56, 57]. Furthermore, it is reasonable to speculate that (p)ppGpp accumulates when *M. xanthus* transitions from exponential growth to stationary phase, since this seems to be the case in other bacterial species [58]. In addition, bacterial natural product gene expression was previously linked to (p)ppGpp accumulation [59] and stringent response-associated genes such as *relA* and *spoT* have been linked to synthesis of bacterial natural products [41, 42, 60].

It is interesting that three of the four natural product σ^{54} promoters that we characterized are located within the coding sequences of genes (intragenic) and not in intergenic regions (**Fig. 5.1**). In the MXAN1286 and MXAN1603 loci, the σ^{54} promoters are located in the coding sequence of the first gene of an operon and in the 5' end of a single gene, respectively. In the other case (MXAN3778), the σ^{54} promoter is located in the coding sequence of an upstream gene (**Fig. 5.1**). With the exception of MXAN1603, the Nla28 binding sites of the natural product promoters are also intragenic, located in the coding sequence of an upstream gene (**Fig. 5.1**). These findings are counter to the commonly held view that bacterial promoter elements are typically located in intergenic regions [61-63], but are supported by additional pieces of bioinformatic and experimental data. First, the vast majority of the putative natural product σ^{54} promoters shown in Figure 5.3 and listed in Table 5.3 are intragenic. Secondly, the majority of the characterized σ^{54} promoter targets of the *M. xanthus* EBP Nla6 are intragenic, located in the coding sequence of an upstream gene [53]. Thirdly, the vast majority of the developmental σ^{54} promoter targets of Nla28, which we recently characterized, are located in the coding sequence of an upstream gene or appear to be internal promoters in operons [30, 39]. Lastly, genomic-wide studies in *Escherichia coli* and *Salmonella enterica* serovar *Typhimurium* uncovered an abundance of potential intragenic σ^{54} promoters and a number of these promoters were subsequently confirmed experimentally [64, 65]. Together, these results suggest that σ^{54} promoter elements might indeed be commonly located in intragenic regions.

Many of the natural product σ^{54} promoters shown in Figure 5.3 and listed in Table 5.3 appear to be internal promoters; promoters located within a gene in an operon instead of upstream of the first gene in the operon. There are examples of internal promoters oriented in the antisense direction relative to the genes in an operon [65-68]; however, this does not seem to be the case for the above internal σ^{54} promoters. Instead, each promoter is predicted to yield mRNA corresponding to a subset of an operon's genes. As noted in a recent review, one potential advantage of using internal promoters would be the ability to express only the genes needed in a particular environment [69].

A second and relatively large group of putative natural product σ^{54} promoters is apparent from scanning Figure 5.3 and Table 5.3. Namely, many of the putative σ^{54} promoters are located in the coding sequence of a gene that is upstream of the single gene or operon they are predicted to regulate, even though the single gene or operon has an upstream and adjacent intergenic region that is long enough to accommodate a σ^{54} promoter (σ^{54}-RNA polymerase binding site). The question that arises is, why would a σ^{54} promoter be located so far upstream of the single gene or operon whose transcription it aims to regulate? One possibility is that the natural product σ^{54} promoter is positioned at a distant location to produce a relatively long 5' untranslated region (UTR) in the mRNA and that the 5' UTR provides an additional layer of regulation for natural product gene expression. We should note that a substantial number of the internal σ^{54} promoters are predicted to yield relatively long 5' UTRs as well. Why would many natural product genes need multiple levels of regulation? Natural

product synthesis is often energy expensive, and many natural products are toxic at high levels; hence, tight and signal-responsive regulation of natural product gene expression might be important in many cases.

Of course, addressing questions about the role and mechanism of σ^{54}-mediated regulation of natural product genes and other types of bacterial genes will require researchers to go beyond predicting where σ^{54} promoters are located. Indeed, we argue that the work of confirming putative σ^{54} promoters and analyzing the mRNAs under σ^{54} control is crucial, as the results will help address a number of long-held assumptions about bacterial promoters and bacterial gene regulation in general.

5.7 References

1. Barrios, H., B. Valderrama, and E. Morett, *Compilation and analysis of sigma(54)-dependent promoter sequences.* Nucleic Acids Res, 1999. **27**(22): p. 4305-13.
2. Feklístov, A., et al., *Bacterial sigma factors: a historical, structural, and genomic perspective.* Annu Rev Microbiol, 2014. **68**: p. 357-76.
3. Popham, D.L., et al., *Function of a bacterial activator protein that binds to transcriptional enhancers.* Science, 1989. **243**(4891): p. 629-635.
4. Sasse-Dwight, S. and J.D. Gralia, *Role of eukaryotic-type functional domains found in the prokaryotic enhancer receptor factor σ54.* Cell, 1990. **62**(5): p. 945-954.
5. Wedel, A. and S. Kustu, *The bacterial enhancer-binding protein NTRC is a molecular machine: ATP hydrolysis is coupled to transcriptional activation.* Genes & development, 1995. **9**(16): p. 2042-2052.
6. Gao, F., et al., *Bacterial Enhancer Binding Proteins—AAA+ Proteins in Transcription Activation.* Biomolecules, 2020. **10**(3): p. 351.
7. Bush, M. and R. Dixon, *The role of bacterial enhancer binding proteins as specialized activators of σ54-dependent transcription.* Microbiology and Molecular Biology Reviews, 2012. **76**(3): p. 497-529.
8. Su, W., et al., *DNA-looping and enhancer activity: association between DNA-bound NtrC activator and RNA polymerase at the bacterial glnA promoter.* Proceedings of the National Academy of Sciences, 1990. **87**(14): p. 5504-5508.

9. Carmona, M. and B. Magasanik, *Activation of transcription at σ54-dependent promoters on linear templates requires intrinsic or induced bending of the DNA.* Journal of molecular biology, 1996. **261**(3): p. 348-356.

10. Rippe, K., et al., *Transcriptional activation via DNA-looping: visualization of intermediates in the activation pathway of E. coli RNA polymerase· σ54 holoenzyme by scanning force microscopy.* Journal of molecular biology, 1997. **270**(2): p. 125-138.

11. Shingler, V., *Signal sensory systems that impact σ54-dependent transcription.* FEMS microbiology reviews, 2011. **35**(3): p. 425-440.

12. Magasanik, B., *Regulation of transcription of the glnALG operon of Escherichia coli by protein phosphorylation.* Biochimie, 1989. **71**(9-10): p. 1005-1012.

13. Yamada, A., et al., *Evolutionary trend of phylogenetic diversity of nitrogen fixation genes in the gut community of wood-feeding termites.* Molecular ecology, 2007. **16**(18): p. 3768-3777.

14. Hayrapetyan, H., et al., *Bacillus cereus ATCC 14579 RpoN (Sigma 54) is a pleiotropic regulator of growth, carbohydrate metabolism, motility, biofilm formation and toxin production.* PLoS One, 2015. **10**(8): p. e0134872.

15. Arai, H., et al., *Transcriptional regulation of the flavohemoglobin gene for aerobic nitric oxide detoxification by the second nitric oxide-responsive regulator of Pseudomonas aeruginosa.* Journal of Bacteriology, 2005. **187**(12): p. 3960-3968.

16. Weiner, L., J. Brissette, and P. Model, *Stress-induced expression of the Escherichia coli phage shock protein operon is dependent on sigma 54 and modulated by positive and negative feedback mechanisms.* Genes & development, 1991. **5**(10): p. 1912-1923.

17. Anderson, D.K., et al., *Regulation of the Caulobacter crescentus rpoN gene and function of the purified σ 54 in flagellar gene transcription.* Molecular and General Genetics MGG, 1995. **246**(6): p. 697-706.

18. Arora, S.K., et al., *A transcriptional activator, FleQ, regulates mucin adhesion and flagellar gene expression in Pseudomonas aeruginosa in a cascade manner.* Journal of bacteriology, 1997. **179**(17): p. 5574-5581.

19. Jiménez-Fernández, A., et al., *Complex interplay between FleQ, cyclic diguanylate and multiple σ factors coordinately regulates flagellar motility and biofilm development in Pseudomonas putida.* PloS one, 2016. **11**(9): p. e0163142.

20. Correa, N.E., F. Peng, and K.E. Klose, *Roles of the regulatory proteins FlhF and FlhG in the Vibrio cholerae flagellar transcription hierarchy.* Journal of bacteriology, 2005. **187**(18): p. 6324-6332.

21. Martinez, R.M., et al., *Characterization of two outer membrane proteins, FlgO and FlgP, that influence Vibrio cholerae motility.* Journal of bacteriology, 2009. **191**(18): p. 5669-5679.

22. Shao, X., et al., *RpoN-dependent direct regulation of quorum sensing and the Type VI secretion system in Pseudomonas aeruginosa PAO1.* Journal of bacteriology, 2018. **200**(16).

23. Matsuyama, B.Y., et al., *Mechanistic insights into c-di-GMP–dependent control of the biofilm regulator FleQ from Pseudomonas aeruginosa.* Proceedings of the National Academy of Sciences, 2016. **113**(2): p. E209-E218.

24. Lloyd, M.G., et al., *Targeting the alternative sigma factor RpoN to combat virulence in Pseudomonas aeruginosa.* Scientific reports, 2017. **7**(1): p. 1-12.

25. Caberoy, N.B., et al., *Global mutational analysis of NtrC-like activators in Myxococcus xanthus: identifying activator mutants defective for motility and fruiting body development.* Journal of Bacteriology, 2003. **185**(20): p. 6083-6094.

26. Diodati, M.E., et al., *Nla18, a key regulatory protein required for normal growth and development of Myxococcus xanthus.* Journal of bacteriology, 2006. **188**(5): p. 1733-1743.

27. Ossa, F., et al., *The Myxococcus xanthus Nla4 protein is important for expression of stringent response-associated genes, ppGpp accumulation, and fruiting body development.* Journal of bacteriology, 2007. **189**(23): p. 8474-8483.

28. Goldman, B., et al., *Evolution of sensory complexity recorded in a myxobacterial genome.* Proceedings of the National Academy of Sciences, 2006. **103**(41): p. 15200-15205.

29. Wu, S.S. and D. Kaiser, *Genetic and functional evidence that type IV pili are required for social gliding motility in Myxococcus xanthus.* Molecular microbiology, 1995. **18**(3): p. 547-558.

30. Giglio, K.M., et al., *A cascade of coregulating enhancer binding proteins initiates and propagates a multicellular developmental program.* Proceedings of the National Academy of Sciences, 2011. **108**(32): p. E431-E439.

31. Kaufman, R.I. and B.T. Nixon, *Use of PCR to isolate genes encoding sigma54-dependent activators from diverse bacteria.* Journal of bacteriology, 1996. **178**(13): p. 3967-3970.

32. Gorski, L. and D. Kaiser, *Targeted mutagenesis of ς54 activator proteins in Myxococcus xanthus.* Journal of Bacteriology, 1998. **180**(22): p. 5896-5905.

33. Guo, D., Y. Wu, and H.B. Kaplan, *Identification and characterization of genes required for early Myxococcus xanthus developmental gene expression.* Journal of Bacteriology, 2000. **182**(16): p. 4564-4571.

34. Sun, H. and W. Shi, *Genetic studies of mrp, a locus essential for cellular aggregation and sporulation of Myxococcus xanthus.* Journal of bacteriology, 2001. **183**(16): p. 4786-4795.

35. Kirby, J.R. and D.R. Zusman, *Chemosensory regulation of developmental gene expression in Myxococcus xanthus.* Proceedings of the National Academy of Sciences, 2003. **100**(4): p. 2008-2013.

36. Jelsbak, L. and D. Kaiser, *Regulating pilin expression reveals a threshold for S motility in Myxococcus xanthus.* Journal of Bacteriology, 2005. **187**(6): p. 2105-2112.

37. Fimlaid, K.A. and A. Shen, *Diverse mechanisms regulate sporulation sigma factor activity in the Firmicutes.* Current opinion in microbiology, 2015. **24**: p. 88-95.

38. Stevens, D.C., et al., *Alternative sigma factor over-expression enables heterologous expression of a type II polyketide biosynthetic pathway in Escherichia coli.* PloS one, 2013. **8**(5): p. e64858.

39. Giglio, K.M. and A.G. Garza, *Fruiting Body Development in Myxococcus xanthus: A Multicellular Developmental Programme that Leads to Sporulation.* Bacterial Spores: Current Research and Applications, 2012: p. 19.

40. Volz, C., C. Kegler, and R. Müller, *Enhancer binding proteins act as hetero-oligomers and link secondary metabolite production to myxococcal development, motility, and predation.* Chemistry & biology, 2012. **19**(11): p. 1447-1459.

41. Knauber, T., et al., *Mutation in the rel gene of Sorangium cellulosum affects morphological and physiological differentiation.* Molecular microbiology, 2008. **69**(1): p. 254-266.

42. Washio, K., et al., *Identification and characterization of the genes responsible for the production of the cyclic lipopeptide arthrofactin by*

Pseudomonas sp. MIS38. Bioscience, biotechnology, and biochemistry, 2010. **74**(5): p. 992-999.

43. Alam, M.T., et al., *Metabolic modeling and analysis of the metabolic switch in Streptomyces coelicolor.* BMC genomics, 2010. **11**(1): p. 1-9.

44. Nieselt, K., et al., *The dynamic architecture of the metabolic switch in Streptomyces coelicolor.* BMC genomics, 2010. **11**(1): p. 1-9.

45. Shin, H., et al., *Production of antimicrobial substances by strains of myxobacteria Corallococcus and Myxococcus.* Microbiology and Biotechnology Letters, 2013. **41**(1): p. 44-51.

46. Kaiser, D., *Social gliding is correlated with the presence of pili in Myxococcus xanthus.* Proceedings of the National Academy of Sciences, 1979. **76**(11): p. 5952-5956.

47. Fisseha, M., et al., *Characterization of the regulatory region of a cell interaction-dependent gene in Myxococcus xanthus.* Journal of Bacteriology, 1996. **178**(9): p. 2539-2550.

48. Kroos, L., A. Kuspa, and D. Kaiser, *A global analysis of developmentally regulated genes in Myxococcus xanthus.* Developmental biology, 1986. **117**(1): p. 252-266.

49. Kaplan, H., A. Kuspa, and D. Kaiser, *Suppressors that permit A-signal-independent developmental gene expression in Myxococcus xanthus.* Journal of bacteriology, 1991. **173**(4): p. 1460-1470.

50. Studholme, D.J., M. Buck, and T. Nixon, *Identification of potential σN-dependent promoters in bacterial genomes.* Microbiology, 2000. **146**(12): p. 3021-3023.

51. Keseler, I.M. and D. Kaiser, *An early A-signal-dependent gene in Myxococcus xanthus has a sigma 54-like promoter.* Journal of Bacteriology, 1995. **177**(16): p. 4638-4644.

52. Garza, A.G., et al., *Control of asgE expression during growth and development of Myxococcus xanthus.* Journal of bacteriology, 2000. **182**(23): p. 6622-6629.

53. Giglio, K.M., et al., *The enhancer binding protein Nla6 regulates developmental genes that are important for Myxococcus xanthus sporulation.* Journal of bacteriology, 2015. **197**(7): p. 1276-1287.

54. Sarwar, Z. and A.G. Garza, *The Nla28S/Nla28 two-component signal transduction system regulates sporulation in Myxococcus xanthus.* Journal of bacteriology, 2012. **194**(17): p. 4698-4708.

55. Sarwar, Z. and A.G. Garza, *Two-component signal transduction systems that regulate the temporal and spatial expression of Myxococcus xanthus sporulation genes.* Journal of bacteriology, 2016. **198**(3): p. 377-385.

56. Singer, M. and D. Kaiser, *Ectopic production of guanosine penta-and tetraphosphate can initiate early developmental gene expression in Myxococcus xanthus.* Genes & development, 1995. **9**(13): p. 1633-1644.

57. Harris, B.Z., D. Kaiser, and M. Singer, *The guanosine nucleotide (p) ppGpp initiates development and A-factor production inMyxococcus xanthus.* Genes & development, 1998. **12**(7): p. 1022-1035.

58. Cashel, M., *The stringent response.* Escherichia coli and Salmonella typhimulium: cellular and molecular biology, 1996. **2**: p. 1458-1496.

59. Hesketh, A., J. Sun, and M. Bibb, *Induction of ppGpp synthesis in Streptomyces coelicolor A3 (2) grown under conditions of nutritional sufficiency elicits actII-ORF4 transcription and actinorhodin biosynthesis.* Molecular microbiology, 2001. **39**(1): p. 136-144.

60. Kang, S.G., et al., *Actinorhodin and undecylprodigiosin production in wild-type and relA mutant strains of Streptomyces coelicolor A3 (2) grown in continuous culture.* FEMS microbiology letters, 1998. **168**(2): p. 221-226.

61. Jacques, P.-É., et al., *Detection of prokaryotic promoters from the genomic distribution of hexanucleotide pairs.* BMC bioinformatics, 2006. **7**(1): p. 423.

62. Wang, H. and C.J. Benham, *Promoter prediction and annotation of microbial genomes based on DNA sequence and structural responses to superhelical stress.* BMC bioinformatics, 2006. **7**(1): p. 248.

63. Rangannan, V. and M. Bansal, *Identification and annotation of promoter regions in microbial genome sequences on the basis of DNA stability.* Journal of biosciences, 2007. **32**(1): p. 851-862.

64. Bonocora, R.P., et al., *Genome-scale mapping of Escherichia coli σ 54 reveals widespread, conserved intragenic binding.* PLoS Genet, 2015. **11**(10): p. e1005552.

65. Bono, A.C., et al., *Novel DNA binding and regulatory activities for σ54 (RpoN) in Salmonella enterica serovar Typhimurium 14028s.* Journal of bacteriology, 2017. **199**(12).

66. Dornenburg, J.E., et al., *Widespread antisense transcription in Escherichia coli.* MBio, 2010. **1**(1).

67. Hartkoorn, R.C., et al., *Genome-wide definition of the SigF regulon in Mycobacterium tuberculosis.* Journal of bacteriology, 2012. **194**(8): p. 2001-2009.

68. Fitzgerald, D.M., R.P. Bonocora, and J.T. Wade, *Comprehensive mapping of the Escherichia coli flagellar regulatory network.* PLoS Genet, 2014. **10**(10): p. e1004649.

69. Mejía-Almonte, C., et al., *Redefining fundamental concepts of transcription initiation in bacteria.* Nature Reviews Genetics, 2020. **21**(11): p. 699-714.

Chapter 6

Conclusions and Future Directions

6.1 Conclusions and Future Directions

In bacteria, starvation leads to profound changes in behavior and physiology. Some of these changes have economic and health implications, as the starvation response has been linked to the formation of biofilms, virulence, and antibiotic resistance. Bacterial gene regulation in response to environmental changes is dynamic, complex, and spread across multiple subpopulations. σ^{54} is an important and conserved regulatory factor in a majority of bacteria. Unlike other σ factors, σ^{54} regulates the transcription of numerous functional genes in an EBP-dependent manner after receiving external environmental signals. The role of σ^{54}-controlled regulatory pathways in these evolutionary dynamics in bacterial starvation response is being increasingly recognized. Although the view that NtrC-like activator-mediated σ^{54} regulation involves in early bacterial stress response is not new [1], systematic analysis of gene regulation pathways and transcriptional changes in this survival event has only begun in recent years. To better understand how starvation contributes to changes in bacterial physiology and resistance, in the first project we identified the starvation-triggered gene regulatory network associated with transcriptional activator Nla28, determined the mechanism by which starvation-responsive genes are activated by Nla28, and showed that several of the genes are important for the formation of a highly resistant cell type. Using a reasonable combination of sequence comparisons by bioinformatics tools, gene expression profiling and phylogenetic classification of encoded proteins, potential functional counterparts to many of the genes identified here were found in other bacterial species and some of these

counterparts have been implicated in stress resistance. Our findings not only confirmed that Nla28 activator is crucial for the regulation of many *M. xanthus* developmental genes in response to starvation, but also uncovered the changes in molecular or cellular functions of *M. xanthus* under starvation.

Most of the Nla28-targeted genes confirmed in this project encode regulatory or signal transduction factors including Nla28, ActB, Nla6 and MrpB EBPs. These EBPs modulates transcription of various types of developmental genes indicating that Nla28 might be a common regulator of external stimulus sensor and signal transducer when nutrient availability is changed. It also consists with the findings that some *M. xanthus* EBPs are auto-regulated or inter-regulated during development [2]. In addition, a couple of remaining Nla28-targeted genes confirmed in the project are characterized to function in defense mechanisms of *M. xanthus*, such as heavy metal transport and adaptive immunity against foreign genetic elements, while others are linked to lipid and carbohydrate metabolism. These gene regulations might work through decreasing the permeability and fluidity of cellular membrane to effectively protect *M. xanthus* cells from being affected by starvation. Our findings are also evidenced by the preliminary study of my second project that these Nla28-targeted genes are likely involved in natural product synthesis for defense and predation [3]. Also, the collection of 140 putative Nla28 targets in this project further links Nla28-mediated *M. xanthus* starvation response to metabolic functions, regulatory/signal transduction and cell envelope/cell wall biogenesis.

Overall, my first project provides a preliminary mapping of the genes that are engaged in modulating starvation response in *M. xanthus*.

In the past few years, the understanding of σ^{54}-dependent transcription has significantly progressed owing to the structural analysis of the σ^{54}-RNAP complex and the application of transcriptome sequencing technology [4, 5]. Concurrently, an increasing number of σ^{54}-dependent EBPs and target genes have been identified using the bioinformatic method by analyzing the conserved AAA+ domain and special binding sites on the promoter sequence, respectively [6-8]. In specific, regulons and binding sites of σ^{54} in *B. cereus* group and *Salmonella enterica* were identified by DNA microarray, ChIP-chip and computational predictions, which were also used to analyze EBP domains and gene organization and predict operons. Their results revealed several novel σ^{54}-dependent metabolic pathways in *Bacillus* and *Salmonella* bacteria. However, whether these EPSs have redundant regulatory functions and how they competitively interact with σ54 factors remain unknown. Furthermore, the upstream signals received by EBPs to activate σ54-dependent regulatory pathways remain elusive.

Our profiling and classification of Nla28-target genes associated with *M. xanthus* starvation response provides an important resource for myxobacterial researchers to delineate the regulatory and functional pathways responsible for survival of *M. xanthus* under nutrient stress. From this, it will be possible for researchers to draw a detailed Nla28 activator-mediated genetic regulatory maps outlining the transitions or complete changes in molecular or cellular functions

when *M. xanthus* confronts nutrient limitation. However, yet much remains unknown about the details of this regulatory network. In general, to dissect the complex regulatory network of σ^{54}, the following research-gaps should be addressed: (1) identification of candidate EBPs and characterization of their functions, (2) characterization of the interaction of EBPs with σ^{54} factor, (3) clarification of the redundant or unique regulatory functions of EBPs, (4) identification of the upstream signals of σ^{54} factor, and (5) identification of the conserved and specific regulatory pathways of σ^{54} factor in different bacteria [9]. As to Nla28-mediated σ^{54} regulation in *M. xanthus* specifically, which Nla28-targeted genes are regulated by Nla28 earlier right after starvation, while which ones are regulated relatively in later stages? Are these 140 putative Nla28 target genes directly regulated by Nla28 or indirectly regulated through other branches of pathways? Where does Nla28 recognize and bind in direct-regulated genes? Are there any positive or negative feedback loops in this regulation network?

First, to identify early and late Nla28 targets in this regulatory network, one approach is to determine the patterns and timelines of transcript abundance of these genes. A dynamic genome-wide transcriptome analysis by next-generation RNA sequencing will help with this identification. By analyzing the transcriptomic data of these genes in *M. xanthus* in a dynamic developmental cycle, we will know their accurate expression patterns and timelines in response to starvation. As for distinguishing Nla28 direct targets from indirect targets and identifying Nla28 binding sites, ChIP-sequencing can be applied to massively and parallelly analyze the interaction between Nla28 and all of the putative targets. Also, it will

determine the specific DNA sequence where Nla28 binds in each targeted gene. After obtaining more DNA sequences that Nla28 binds, we will be able to better define Nla28 binding sites by calculating nucleotide frequency. Finally, gene overexpression or knockdown will help identify positive or negative feedback loops in Nla28-mediated regulatory network.

My second project has provided preliminary evidence for the hypothesis that the σ^{54} system is a conserved regulator of PK/NRP natural product genes in bacteria using *M. xanthus* as a model. Since we have only validated three out of 11 *M. xanthus* natural product genes identified containing putative σ^{54} promoters. Also, our previous bioinformatics scanning on *M. xanthus* genome yielded 60 natural product operons (over 150 genes) with putative σ^{54} promoter elements. The future work we could do is to test whether the vast majority of these putative PK and NRP promoters are bona fide σ^{54} promoters [2, 10]. Once we demonstrate that σ^{54} is a common regulator of PK/NRP natural product genes in *M. xanthus*, we could evaluate the effectiveness of *M. xanthus* as a general heterologous host for PK and NRP natural product biosynthetic gene clusters. Unlike in other bacteria that σ^{54} is not required for growth under nutrient rich condition and is expressed at relatively low levels compared to σ^{70} factor [11-13], *M. xanthus* is a bacterium that requires σ^{54} for growth in rich media and expresses σ^{54} at relatively high levels [14]. Presumably, σ^{54} is relatively abundant in *M. xanthus* during growth so that it can effectively compete with other σ factors for RNA polymerase binding and activate transcription of genes with growth-related functions. If PK or NRP gene clusters commonly use σ^{54} promoters for

transcription of important genes, it is predicted that such promoters are more likely to be active in *M. xanthus* during vegetative growth than in bacteria that primarily employ σ^{54} in response to stress. For this reason, and the fact that *M. xanthus* produces a variety of building blocks for synthesis of natural products, we believe that *M. xanthus* is a good candidate for a general heterologous host to highly increase the yield of PKs and NRPs.

The results of recent studies support our speculation. In particular, when the oxytetracycline gene cluster of *Streptomyces rimosus* was placed in *M. xanthus* and the cells were grown in nutrient rich media, oxytetracycline (a type II PK) was produced at relatively high levels [15]. In contrast, when the oxytetracycline gene cluster was placed in an *E. coli* strain that was engineered to produce PK building blocks and the cells were grown in nutrient rich media, the σ^{54} promoter of the crucial oxyB gene was active and oxytetracycline was detected only after rpoN was over-expressed [16]. As confirmed by our previous study, rpoN over-expression was required for activation of two *M. xanthus* σ^{54} promoters in growing *E. coli* cells, suggesting that activation of foreign σ^{54} promoters is a general problem in *E. coli* and perhaps other bacteria under standard growth conditions. Importantly, the production levels of oxytetracycline in *M. xanthus* was 5-fold higher than in the rpoN over-expressing strain of *E. coli* and sufficient for biochemical analyses [15, 16]. In future, further evaluation on the effectiveness of *M. xanthus* as a general heterologous host for PK and NRP natural product biosynthetic gene clusters should be performed. Specifically, we can determine whether *M. xanthus* heterologously expresses the natural product

gene clusters for tyrocidine, which is an NRP and the main component of the clinical antibiotic tyrothricin, and the type I PK antibiotic erythromycin. In addition, we should also determine whether the tyrocidine and erythromycin production levels are sufficient for biochemical analyses; high production levels are crucial for the identification of new PK and NRP natural products. Tyrocidine and erythromycin can be selected because they originated from bacteria that are distantly related to *M. xanthus*, their chemical structures are known, and their corresponding biosynthetic gene clusters contain at least one σ^{54} promoter.

Non-model bacteria including *M. xanthus* are an untapped reservoir of regulatory elements including promoters and transcriptional factors (TF). As noted in this dissertation, the expression level of natural product genes heavily relies on the activity of their promoters and relevant regulations. Many manifests as construction of inducible promoters have been purposed as important components in heterologous expression systems to increase gene expression level. Chemically inducible promoters with TFs are the most applied type of inducible heterologous expression system due to the natural abundance of small molecule sensing TFs. However, most bacterial strains rely on only one or a small number of inducible systems due to the difficulty in expanding the availability of TF and promoters. Emerging industrial hosts like *Rhodococcus opacus* and *Halomonas* spp. have relied on one inducible promoter each, respectively [17, 18].

To overcome the limited number of inducible heterologous expression systems available within non-model strains, a common strategy is to appropriate

a handful of well-characterized *E. coli*-derived inducible systems (e.g. pBAD, pTET, pLac) that can function in non-model bacteria [19-21]. This strategy, most often simply adjusting expression levels of the TF to improve fold induction in non-model organisms, has various degrees of success. Reducing TetR expression by mutating the -10 box of the σ^{70} promoter led to an increase in dynamic range by 67-fold upon induction of anyhydrotetracycline-inducible system (pTET) in *R. opacus* [17]. Likewise, controlling LacI repressor concentration with a weak promoter increased the overall change in expression by 80-fold of IPTG-inducible *E. coli* (pTrc) promoter in *P. putida* [22]. While useful in the lab, these systems are often limited industrially due to the expense, varying half-lives of inducers or complications from serving a dual role as inducer and carbon source. Therefore, engineering inducible expression systems based on modifying promoter elements (i.e. RNAP bind site, TF binding site) as modular has become a new strategy for non-model bacteria [23]. For example, proper placement of the TF binding site on a strong native constitutive promoter or phage promoter blocks RNAP from binding to the promoter [18, 24]; Addition of two TF binding sites in pBAD promoter lowered basal expression and increased induction levels by 2-fold compared to one binding site in *Synechocystis* [19]. In addition, altering DNA sequence of promoter elements, as well as the TF itself, can further fine-tune the expression system's performance to respond with the appropriate sensitivity and signal output according to the growth conditions to increase metabolite productivity [25, 26]; Minor changes in the TF binding site, promoter, or 5'-UTR (including secondary structures) can have drastic effects on

translation efficiency [27]. Screening a library of over 10,000 members comprising *E. coli* and *Vibrio cholerae* σ^{54}-like promoters showed that short, CT-rich sequence motifs complementing the AG-rich Shine-Dalgarno sequence causes an inhibitory translation mechanism across σ^{54} promoters [28]. Based on the previous findings above and the fact that σ^{54} system possess relative higher activity during vegetative growth in *M. xanthus* than in *E. coli*, engineering inducible heterologous PK/NRP expression system via modifications on σ^{54} promoter elements including RNAP binding sites in -12/-24 regions, EBP binding sites, or even on EBPs' conformations to regulate interactions between EBPs and promoters, might be a feasible approach to construct a highly efficient natural product synthesis host in *M. xanthus* for industrial applications.

Overall, the existence of the alternative σ^{54} system in regulating starvation response and natural product synthesis indicates that evolution in bacteria might have developed various gene regulation pathways to accomplish certain biofunctions for survival, and probably many of them are yet to be discovered. The employment of σ^{54} system in response to starvation despite the dominance of the σ^{70} family is likely due to its near absolute dependence on a EBP activator. Presumably, similar to eukaryotic cells, bacteria adjust the extent to which the EBPs are utilized in response to severe conditions and during secondary metabolism, and such strict, almost hash, regulation must have conferred the advantage necessary for their survival. One of the major benefits for bacteria to engage σ^{54}-controlled site-specific transcript initiation for any gene or gene cluster is that it sets the stage for the evolution of gene-specific regulation,

involving activators. More evolutionary studies investigating the selection of gene regulation pathways and bacterial functions under certain environmental changes will help us better understand how distinct mechanisms evolved in response to different gene transcriptions, based on the specific sigma factor that is involved.

6.2 References

1. Ossa, F., et al., *The Myxococcus xanthus Nla4 protein is important for expression of stringent response-associated genes, ppGpp accumulation, and fruiting body development.* Journal of bacteriology, 2007. **189**(23): p. 8474-8483.
2. Giglio, K.M., et al., *A cascade of coregulating enhancer binding proteins initiates and propagates a multicellular developmental program.* Proceedings of the National Academy of Sciences, 2011. **108**(32): p. E431-E439.
3. Ma, M., R.D. Welch, and A.G. Garza, *The σ 54 system directly regulates bacterial natural product genes.* Scientific reports, 2021. **11**(1): p. 1-11.
4. Yang, Y., et al., *Structures of the RNA polymerase-σ54 reveal new and conserved regulatory strategies.* Science, 2015. **349**(6250): p. 882-885.
5. Sysoeva, T.A., et al., *Crystallization and preliminary X-ray analysis of the ATPase domain of the σ54-dependent transcription activator NtrC1 from Aquifex aeolicus bound to the ATP analog ADP–BeFx.* Acta Crystallographica Section F: Structural Biology and Crystallization Communications, 2013. **69**(12): p. 1384-1388.
6. Liu, B., et al., *Computational prediction of sigma-54 promoters in bacterial genomes by integrating motif finding and machine learning strategies.* IEEE/ACM transactions on computational biology and bioinformatics, 2018. **16**(4): p. 1211-1218.

7. Peng, Q., et al., *Identification of metabolism pathways directly regulated by sigma54 factor in Bacillus thuringiensis.* Frontiers in microbiology, 2015. **6**: p. 407.

8. Bono, A.C., et al., *Novel DNA binding and regulatory activities for σ54 (RpoN) in Salmonella enterica serovar Typhimurium 14028s.* Journal of bacteriology, 2017. **199**(12).

9. Yu, C., et al., *The Regulatory Functions of σ54 Factor in Phytopathogenic Bacteria.* International Journal of Molecular Sciences, 2021. **22**(23): p. 12692.

10. Giglio, K.M., et al., *The enhancer binding protein Nla6 regulates developmental genes that are important for Myxococcus xanthus sporulation.* Journal of bacteriology, 2015. **197**(7): p. 1276-1287.

11. Buck, M., et al., *The bacterial enhancer-dependent ς54 (ςN) transcription factor.* Journal of bacteriology, 2000. **182**(15): p. 4129-4136.

12. Jishage, M., et al., *Regulation of RNA polymerase sigma subunit synthesis in Escherichia coli: intracellular levels of four species of sigma subunit under various growth conditions.* Journal of bacteriology, 1996. **178**(18): p. 5447-5451.

13. Jurado, P., L.A. Fernández, and V. de Lorenzo, *Sigma 54 levels and physiological control of the Pseudomonas putida Pu promoter.* Journal of bacteriology, 2003. **185**(11): p. 3379-3383.

14. Keseler, I.M. and D. Kaiser, *σ54, a vital protein for Myxococcus xanthus.* Proceedings of the National Academy of Sciences, 1997. **94**(5): p. 1979-1984.

15. Stevens, D.C., et al., *Heterologous expression of the oxytetracycline biosynthetic pathway in Myxococcus xanthus.* Applied and Environmental Microbiology, 2010. **76**(8): p. 2681-2683.

16. Stevens, D.C., et al., *Alternative sigma factor over-expression enables heterologous expression of a type II polyketide biosynthetic pathway in Escherichia coli.* PloS one, 2013. **8**(5): p. e64858.

17. DeLorenzo, D.M., W.R. Henson, and T.S. Moon, *Development of chemical and metabolite sensors for Rhodococcus opacus PD630.* ACS synthetic biology, 2017. **6**(10): p. 1973-1978.

18. Zhao, H., et al., *Novel T7-like expression systems used for Halomonas.* Metabolic engineering, 2017. **39**: p. 128-140.

19. Immethun, C.M., et al., *Physical, chemical, and metabolic state sensors expand the synthetic biology toolbox for Synechocystis sp. PCC 6803.* Biotechnology and Bioengineering, 2017. **114**(7): p. 1561-1569.

20. Hanko, E.K., N.P. Minton, and N. Malys, *Characterisation of a 3-hydroxypropionic acid-inducible system from Pseudomonas putida for orthogonal gene expression control in Escherichia coli and Cupriavidus necator.* Scientific reports, 2017. **7**(1): p. 1-13.

21. Kelly, C.L., et al., *A rhamnose-inducible system for precise and temporal control of gene expression in cyanobacteria.* ACS synthetic biology, 2018. **7**(4): p. 1056-1066.

22. Cook, T.B., et al., *Genetic tools for reliable gene expression and recombineering in Pseudomonas putida.* Journal of Industrial Microbiology and Biotechnology, 2018. **45**(7): p. 517-527.

23. Ji, C.-H., H. Kim, and H.-S. Kang, *Synthetic inducible regulatory systems optimized for the modulation of secondary metabolite production in Streptomyces.* ACS Synthetic Biology, 2019. **8**(3): p. 577-586.

24. DeLorenzo, D.M. and T.S. Moon, *Construction of genetic logic gates based on the T7 RNA polymerase expression system in Rhodococcus opacus PD630.* ACS synthetic biology, 2019. **8**(8): p. 1921-1930.

25. Jha, R.K., et al., *A protocatechuate biosensor for Pseudomonas putida KT2440 via promoter and protein evolution.* Metabolic engineering communications, 2018. **6**: p. 33-38.

26. Hanko, E.K., N.P. Minton, and N. Malys, *A transcription factor-based biosensor for detection of itaconic acid.* ACS synthetic biology, 2018. **7**(5): p. 1436-1446.

27. Rohlhill, J., N.R. Sandoval, and E.T. Papoutsakis, *Sort-seq approach to engineering a formaldehyde-inducible promoter for dynamically regulated Escherichia coli growth on methanol.* ACS Synthetic Biology, 2017. **6**(8): p. 1584-1595.

28. Levy, L., et al., *A synthetic oligo library and sequencing approach reveals an insulation mechanism encoded within bacterial σ54 promoters.* Cell Reports, 2017. **21**(3): p. 845-858.

VITA

Name of Author: Muqing Ma

Place of Birth: Shenyang, Liaoning, China

Undergraduate & Graduate Schools Attended:

Jilin University, Changchun, Jilin, China

Degrees Awarded:

Master of Science in Biological Engineering, 2016, Jilin University

Bachelor of Science in Biological Engineering, 2013, Jilin University

Publications:

Ma M., Garza A. G., Lemon. D.J., Caro E. A., Ritchie L., Ryan C., Spearing V. M., Murphy K. A. & Welch R. D. Identifying the gene regulatory network of the starvation-induced transcriptional activator Nla28. *Submitted to Journal of bacteriology*, under revision.

Ma, M., Welch, R. D., & Garza, A. G. (2021). The σ^{54} system directly regulates bacterial natural product genes. *Scientific reports*, *11*(1), 1-11.