Roles of ComW, HPF and YhaM in Competence Regulation in *Streptococcus pneumoniae*

by

Junqin Bai
M.S., Tianjin University, 2010

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Biological Sciences
in the Graudate College of the
University of Illinois at Chicago, 2018
Chicago, Illinois

Defense Committee:
Donald A. Morrison, Advisor
Peter G. Okkema, Chair
Yury Polikanov
Nancy E. Freitag, Microbiology & Immunology
Michael J. Federle, Medicinal Chemistry & Pharmacognosy

This thesis is dedicated to my parents for their endless love and encouragement

ACKNOWLEDGMENTS

I am very grateful to my advisor, Dr. Donald A. Morrison, for his unfailing support and encouragement over the past several years. Dr. Morrison has inspired me to see the beauty of the bacterial world and the art of science. He has taught me to be critical, to design experiments and analyze data, and to draw meaningful conclusions. I truly appreciate how patiently Dr. Morrison has explained matters of importance to me. His eagerness in carrying out research and his enthusiasm in mentoring students have inspired me to become a dedicated mentor in the future. He has always been positive and encouraging during my five years of training in his laboratory, and he has helped me to grow as a person and professional. This has been an important chapter of my life that I will always remember.

I would also like to thank the other members of my committee, Dr. Peter Okkema, Dr. Yury Polikanov, Dr. Nancy Freitag, and Dr. Michael Federle, for their gracious support, encouragement, and mentoring. My thanks to Dr. Okkema for his helpful suggestions and encouragement when I was preparing for our committee meetings. Thank you to Dr. Freitag for her intelligent advice and caring support. I thank Dr. Polikanov for his willingness to share his knowledge of the "sleeping giant" and his expertise with me. I want to thank Dr. Federle for allowing me to join his lab meetings for two years and for the opportunity to learn from the positive thinking meeting. I am honored to have had all these wonderful scientists on my committee. I could always approach them, share my ideas, and ask for their suggestions for implementing my experimental designs. I would not have earned my Ph.D. degree without their help.

In addition, I would like to thank my lab members Nicole Inniss, Duoyi Hu, and Med Jimson Jimenez for their wonderful help and for always being willing to spend time with me and

make sure that my experiments were going smoothly. Special thanks to Duoyi for his training in using Prism and Endnote and his kind-hearted encouragement and support.

I also thank previous lab members Dr. Yanina Tovpeko and undergraduate students for their wonderful help. It was great sharing a lab with all of them during the last five years. Furthermore, I would like to extend my gratitude to Dr. Laura Cook and Dr. Jennifer Chang from Dr. Federle's lab for their critical reading of my dissertation and insightful ideas. And I thank the undergraduate laboratory aides who helped me to prepare solutions. Without your help, I would not have finished my experiments.

I sincerely thank my writing advisor, Mr. Jon Mann, for his editing of my dissertation, sentence by sentence. His encouragement and willingness to support me really touched my heart. Without his help, I would not have finished my writing in a timely manner. I would also like to thank Amel Alqadah, Med Jimenez, and Bethanie Statler for their proofreading efforts and support.

It is also important that I thank my friends in the Biological Sciences department for continuously supporting me, their friendship and caring during my preparation for defense. I appreciate that throughout the ups and downs of my graduate life, I could always talk to them and lean on them for support and encouragement.

I thank the members of my church family for the prayers and love that sustained me through difficult days. And I thank God. "In all your ways acknowledge Him, and He will make your paths straight" (Proverbs 3:6).

Above all, I thank my mother and father, who encouraged me throughout my pursuit of my Ph.D. in Biological Sciences in the United States. I am also thankful to my younger brother, who has always supported me without hesitation.

Junqin (Heidi) Bai

CONTRIBUTION OF AUTHORS

My thesis Chapter III represents a published manuscript (Journal of Bacteriology. 2016. vol.198 no.17 2370-2378) for which I was the second author. I generated Figure 5 and together with Tovpeko for Figure 6, and played a critical role in the writing of the manuscript along with the first author Yanina Tovpeko and my research mentor, Dr. Donald Morrison. My work was critical to the conclusions of this manuscript because the direct interaction between ComW and SigX identified by yeast two-hybrid was fundamental for our understanding of the mechanism of ComW in competence regulation in *Streptococcus pneumoniae*.

TABLE OF CONTENTS

CHAPTER **PAGE**

TABLE OF CONTENTS (continued)

CHAPTER **PAGE**

TABLE OF CONTENTS (continued)

CHAPTER **PAGE**

LIST OF TABLES

TABLE **PAGE**

LIST OF FIGURES

FIGURE PAGE

LIST OF FIGURES (continued)

FIGURE PAGE

LIST OF ABBREVATIONS

α-gal	α-galactosidase
A	adenine
aa	amino acid (s)
Ab	aureobasidin A
ATP	adenosine triphosphate
bp	base pair(s)
BSA	bovine serum albumin
cAMP	cyclic adenosine monophoshpate
CAT	casein hydrolysate yeast extract medium
CFU	colony forming unit
CSP	competence stimulating factor
Cin-box	competence induced promoters
C-terminal	carboxy terminal
DNA	deoxyribonucleic acid
DNaseI	deoxyribonuclease I
DR	direct repeat
EDTA	ethylene diamine tetraacetic acid
Erm	erythromycin
HK	histidine kinase
kBP	kilo-base pairs
Kan	kanamycin

LIST OF ABBREVIATIONS (continued)

L	Leucine
Nov	novobiocin
N-terminal	amino terminal
OD	optical density
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PMSF	phenylmethyl sulfonyl fluoride
PNP-α-gal	p-nitrophenyl-α-D-galactopyranoside
RNA	ribonucleic acid
RNAP	RNA polymerase
RR	response regulator
SD	minimal synthetically defined medium for yeast
SDS	sodium dodecyl sulfate
Spc	spectinomycin
T	tryptophan
Tet	tetracycline
TE	Tris-EDTA
TGED	Tris-glycerol-EDTA-DTT
Tris	Tris (hydroxymethyl) aminomethane
YPD	yeast extract peptone with dextrose
X-α-gal	5-bromo-4-chloro-3-indolyl-D-galactopyranoside

SUMMARY

Competence for genetic transformation in *Streptococcus pneumoniae* is a transient state during which the bacteria can take up DNA from the environment and incorporate it into their genome. Transformation in *S. pneumoniae* depends on two early competence proteins, σ^X and ComW, which have similar temporal induction patterns that begin at 5 minutes, peak at 20 minutes, and cease by 40 minutes after competence stimulating peptides (CSP) exposure. ComW is crucial for σ^X activity and transformation.

To determine the exact mechanism of ComW during competence, yeast two-hybrid (Y2H) assays were performed to assess whether a direct interaction occurred between σ^X and ComW. I found a direct interaction between the two proteins by using the yeast two-hybrid system, including four tightly controlled reporter genes that are Gal4 transcription factor-dependent. A weak interaction between ComW and σ^X was supported by the fact that two of the reporter genes were activated. Lack of interaction between ComW and σ^A was observed in the Y2H assay. This finding, together with the Δ*comW* bypass mutations in σ^A, led to the hypothesis that by directly binding to σ^X, ComW acts to spark competition between σ^A and σ^X in favor of σ^X access to core RNAP.

In bacteria, the mechanisms of competence regulation have been studied at the transcriptional, translational, and post-translational levels. In 1979, Morrison and Baker observed that competence depended on the expression of a set of proteins. However, the molecular mechanism regulating protein expression during competence was unclear. Recently, the proteins HPF and YhaM have been identified by transcriptome as belonging to highly conserved σ^X-dependent regulons in all Streptococci. To investigate the roles of HPF and YhaM in the expression of competence-specific proteins during transformation, the deletion mutants Δ*yhaM,* Δ*Hpf,* and Δ*Hpf*Δ*yhaM* were examined for details of protein synthesis using pulse-chase labeling with S35-

methionine combined with one- and two-dimensional gel electrophoresis and competence assays in parallel. I found that HPF and YhaM affected the protein expression pattern slightly but not globally.

In bacteria, σ factor activity could be regulated by anti-σ factors. In the case of σ^X in *S. pneumoniae*, no anti-σ^X factor had been identified. To determine whether ComW acts as an anti-anti-σ in regulating σ^X activity during development of competence, a Δ*comW* suppressor mutation screen was performed using a Mariner T7 transposon mutation library. The MarinerT7 mutation library failed to identify a suppressor mutation, likely for one of two reasons. It may have produced either insertions on essential genes or a high density of mutations under our experimental conditions. In comparison, the spontaneous mutagenesis approach would have provided a sufficient but low density of amino acid substitutions for mutation mapping purposes.

I. INTRODUCTION

1. Transformation and its regulation in bacteria

Transformation is the process in which bacteria take up free-floating DNA released from dead or lysed cells and incorporate it into their genome. This process, transduction, and conjugation are the three primary mechanisms of horizontal gene transfer (HGT) in bacteria. Both transduction and conjugation depend on mobile genetic elements such as bacteriophages and plasmids. Transduction involves DNA transfer from one bacterium to another by viruses such as bacteriophages, and conjugation is the transfer of DNA between bacterial cells through direct contact and is primarily used for plasmid transfer. HGT is considered a major driving force in increasing bacteria fitness as well as in population dynamics by virtue of its incorporation of newly uptake DNA into a genome via either homologous recombination or insertion. Transformation is considered a powerful mechanism of HGT in natural bacterial populations. Consequently, the rising antibiotic resistance among bacteria resulting from HGT calls for increased efforts to study the mechanisms of HGT.

Natural bacterial transformation was first observed in *S. pneumoniae* in 1928 by British bacteriologist Frederick Griffith. In his experiment, known as the Griffith experiment, he reported that bacteria could obtain substances through a process called transformation (Griffith, 1928). In a series of experiments with pneumococci, an S strain (avirulent) must have taken up a substance from the heat-killed R strain (virulent), which caused an S strain-injected mouse to develop pneumonia and die. The substance that caused the bacterial transformation was revealed to be DNA by Oswald Avery in an era when scientists believed that proteins had been transferred in the Griffith experiment (Avery, 1944).

Most naturally transformable bacteria have transient competence, such as the competence lasting less than 40 minutes in *S. pneumoniae* (Alloing, Martin, Granadel, & Claverys, 1998; Ween, Gaustad, & Havarstein, 1999). *Neisseria gonorrhoeae* is exceptional in having constitutive competence (Jyssum, 1966). Natural competence can occur during the transition from log phase to stationary phase, such as in *B. subtilis* and *Vibrio* sp.(Frischer, Thurmond, & Paul, 1993). In *Haemophilus influenza,* competence begins when cells are transferred to a medium that does not allow growth (Smith, Danner, & Deich, 1981). With the rapid spread of large-scale sequencing of microbial genomes, transcriptomes, and proteomics, the number of naturally transformable species is rising. A total of 82 species have been shown to be naturally transformable (Johnston, Martin, Fichant, Polard, & Claverys, 2014).

The benefits of natural transformation and DNA uptake remain controversial. Uptake of exogenous DNA can provide nutrients, genome repair, or genome evolution. Exogenous DNA used as food (such as nucleotides, nitrogenous bases, and phosphates) can be beneficial for recipient bacteria immediately after uptake (Finkel & Kolter, 2001). Additionally, exogenous DNA supports genomic repair, which was evidenced by observations that competence was induced and may have been involved in DNA repair after UV exposure in *Legionella pneumophila* (Charpentier, Kay, Schneider, & Shuman, 2011). However, it was shown that uptake of homologous DNA did not improve *H. influenzae*'s ability to survive DNA damage (Redfield, 1993), and competent cells were less able to repair DNA lesions induced by UV compared to normal cells (Ahmed, Setlow, Grist, & Setlow, 1993). Yet nutrient limitation alone but not DNA damage induced competence in both *H. influenzae* and *B. subtilis* (Redfield, 1993). Nonetheless, the usefulness of exogenous DNA for evolution is strongly supported by growing findings in comparative genomics. For example, highly frequent HGT in *V. cholera* and inter- and intra-

species in Streptococci (Veening & Blokesch, 2017) showed that integrated DNA could effectively change the pathogenic and ecological character of recipient bacteria (Finkel & Kolter, 2001).

Transformation is regulated though quorum sensing (QS), a concentration-dependent manner of coordinating collective behaviors in a population (see Figure 1). A QS system includes intracellular inducer synthesis, secretion and accumulation outside of cells, detection by cognate transmembrane receptors, and activation of signaling cascades regulating gene expression (Ng, 2009). Gram-positive and Gram-negative bacteria both use QS systems. Gram-positive bacteria typically use small peptides and a two-component signal transduction system (TCSTS), which includes a membrane-bound kinase receptor and the cytoplasmic response regulator that direct gene expression (Monnet, Juillard, & Gardan, 2016). Gram-negative bacteria commonly use small molecules that can diffuse freely through membranes, such as acyl-homoserine lactone derivatives or other small molecules synthesized from S-adenosylmethionine (SAM) (Papenfort & Bassler, 2016). QS is crucial not only for many bacteria in nature but also for collective behaviors in higher organisms, such as the QS-like pattern used in nesting sites by honey bees and ants and hair follicle regeneration in animals (Papenfort & Bassler, 2016). QS has broad response regulation ranging from expression of virulence factors such as the PapR and PlcR, and NprX and NprR in *B. subtilis* to formation of biofilms such as the SHP and Rgg circuit in *Streptococcus pyogenes* and competence development governed by XIP/ComR in *Streptococcus mutans* (Declerck et al., 2007; Ng & Bassler, 2009). To precisely orchestrate gene expression in transformation regulation, bacteria have developed distinct mechanisms in diverse species.

1.1 Transformation regulation in Gram-negative bacteria

Gram-negative bacteria use extracellular signaling molecules called autoinducers (AI) to trigger QS to temporally control input and output dynamics in cells. AIs can be recognized by and

bound to specific receptors that locate in either the inner membrane or cytoplasm of cells, triggering dozens to hundreds of gene expressions. Two major receptors are widely used: one is LuxR-type receptors, a cytoplasmic transcription factor that detects AIs. This transcription factor contains two functional domains: a carboxy-terminal ligand binding domain and an amino-terminal ligand binding domain. The N-terminal can be bound and stabilized by AIs, resulting in functional transcription factor recognition of "lux boxes" upstream of target genes; another one is two-component membrane-bound histidine kinases. This receptor recognition of AIs is regulated by QS, which establishes a positive feedback loop to promote synchronous gene expression in the population (Khusainov et al., 2017). Transformation regulation in *Neiserria gonorrhoeae* and *H. influenzae* are illustrated briefly as two examples.

N. gonorrhoeae is a naturally transformable species that only live inside human beings and competence for genetic transformation last all phases of growth (Biswas, 1977). It is unique that competence does not depend on any soluble competence factor. During transformation, only those DNA that contain a genus-specific sequence (GCCGTCTGAA) can be transformed (Elkins, Thomas, Seifert, & Sparling, 1991). Its donor DNA can from autolysis or type IV secretion (T4SS). The processing of DNA uptake shared similar uptake apparatus as other bacteria such as *S. pneumoinae* which is described later in details (Hamilton & Dillard, 2006).

In the case of *H. influenzae,* competence is induced by energy starvation and depletion of nucleotide acids (Sinha, Mell, & Redfield, 2012), and regulated by cyclic adenosine monophosphate (cAMP), cAMP regulatory protein (CRP), and Sxy (Redfield et al., 2005). Like *N. gonorrhoeae*, *H. influenzae* only takes up genus-specific DNA (Smith & Venter, 1995). Under stress conditions, cAMP-dependent CRP, as a transcription factor, recognizes competence regulatory element (CRE) upstream of competence specific genes, activating expression of those

genes. One of those gene, *sxy*, its protein product acts as activator further activate gene expression (Sinha et al., 2012). In addition, the secondary structure of *sxy* mRNA is regulated by depletion of nucleotide acids, resulting in active form of *sxy* mRNA.

1.2 Transformation regulation in Gram-positive bacteria

In Gram-positive bacteria, transformation mainly relies on QS pathways involving a two-component signal transduction system (TCSTS) or cytoplasmic effectors. After the discovery of QS-mediated bioluminescence in *Vibrio fisheri,* QS by hormone-like substances was discovered in *S. pneumoniae* by Alexander Tomasz (Tomasz, 1965). A TCSTS serves as a basic stimulating response to decode extracellular signals to organisms in order to coordinate their behaviors. This system typically consists of a membrane-bound histidine kinase that senses extracellular signals and a corresponding response regulator that manipulates intracellular activity at the molecular level. QS allows bacteria to carry out group behaviors such as genetic transformation and biofilm formation. Quorum peptides are synthesized and secreted into extracellular space either by a general secretion system or by dedicated ABC transporters (Jimenez & Federle, 2014). Then peptides are brought back to the cell through oligopeptide permease or through membrane-bound two-component receptors, such as ComDE in some Streptococci groups (Ng & Bassler, 2009; Pestova, Havarstein, & Morrison, 1996). Transformation regulation in *S. pneumoniae* and *S. mutans* will be discussed in detail as follows.

The most important regulatory mechanism in bacteria is the use of different σ factors for expression of DNA uptake and incorporation apparatus (Hamoen, 2003). However, a striking exception is the competence regulation *B. subtilis,* that σ factor plays a minor role in the

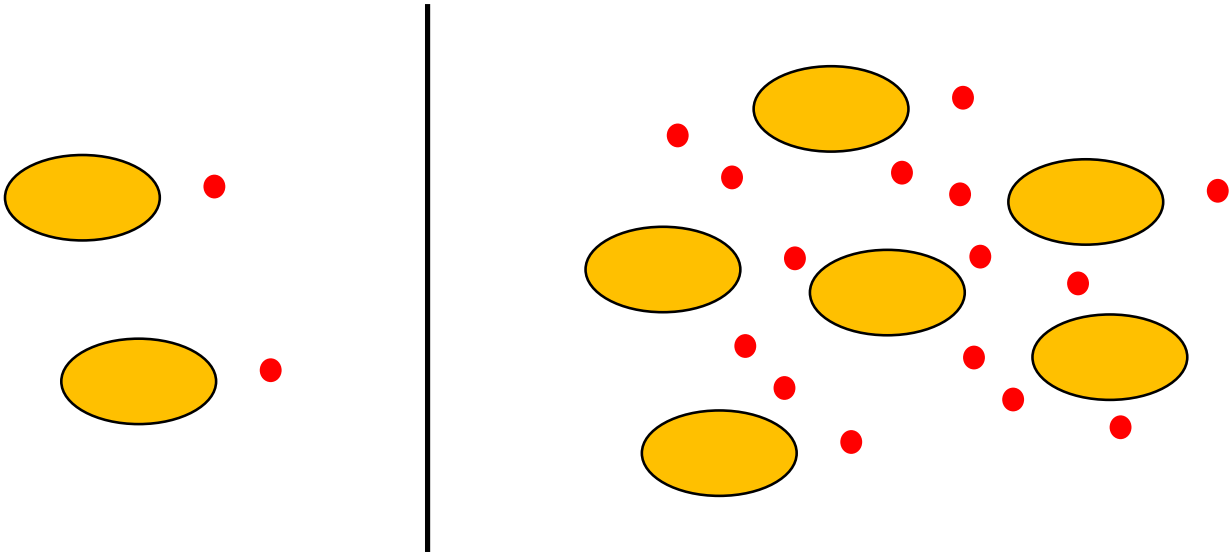


Figure 1. Overview of quorum sensing mechanism in bacteria. Extracellular molecules are transferred from cells either by molecular diffusion or transporters. In low cell density, bacteria secret. When population density increases, more molecules are secreted and accumulated to a critical point, intracellular signaling pathway is activated. This activates a positive-feedback loop to induce gene expression responding to environmental stimuli.

development of competence. The σ^H, encoded by *spo0H*, is known to be required for competence, but its major role in competence development seems to regulate the quorum-sensing pheromone PhrC (Soloman, 1996). ComK, the competence-specific transcription factor (Hahn, Luttinger, & Dubnau, 1996; Vansinderen et al., 1995), its transcription can be repressed by CodY (Serror, 1996), AbrB (Hahn, 1994, Vansinderen, 1995), and Rok (Hoa & Dubnau, 2002) during exponential phase (Hoa, Tortosa, Albano, & Dubnau, 2002; Yuksel, Power, Ribbe, Volkmann, & Maier, 2016). AbrB acts as a transcriptional repressor that downregulates *comK* transcription (Hahn, Roggiani, & Dubnau, 1995). Moreover, *comK* transcription can be activated by phosphorylated Spo0A via repression of *abrB* transcription, thus releasing ComK to be activated (Rahmer, Heravi, & Altenbuchner, 2015). In addition, another protein ComS is reported to protect ComK from degradation of MecA/ClpCP (Turgay & Dubnau, 1998) (see Figure 2). Recent findings indicate that a novel protein Kre also affects the stability of *comK* mRNA (Gamba, Jonker, & Hamoen, 2015). The ComK can also be regulated at stability level. ComK is rapidly degraded during the escape from competence by a protease complex that consists of MecA and ClpP/ClpC. MecA is an adaptor targeting degradation of ComK by ClpPC (Dahl & Rapoport, 1991). Another protein ComS is produced in response to quorum sensing system. It has a stronger affinity to MecA in competing with ComK for binding to MecA. So ComS's binding of ComK protects it from degradation by MecA and ClpPC. This is evidenced by the fact that ComK's degradation rate can be decreased in a strain carrying a multicopy *comS* plasmid and in a strain with an over-expression of ComS (Stiegelmeyer & Giddings, 2013; Turgay et al., 1998) (see Figure 2). Despite the various mechanisms of transformation regulation, the processes are to ensure bacteria to adapt the drastic changes of environmental conditions.

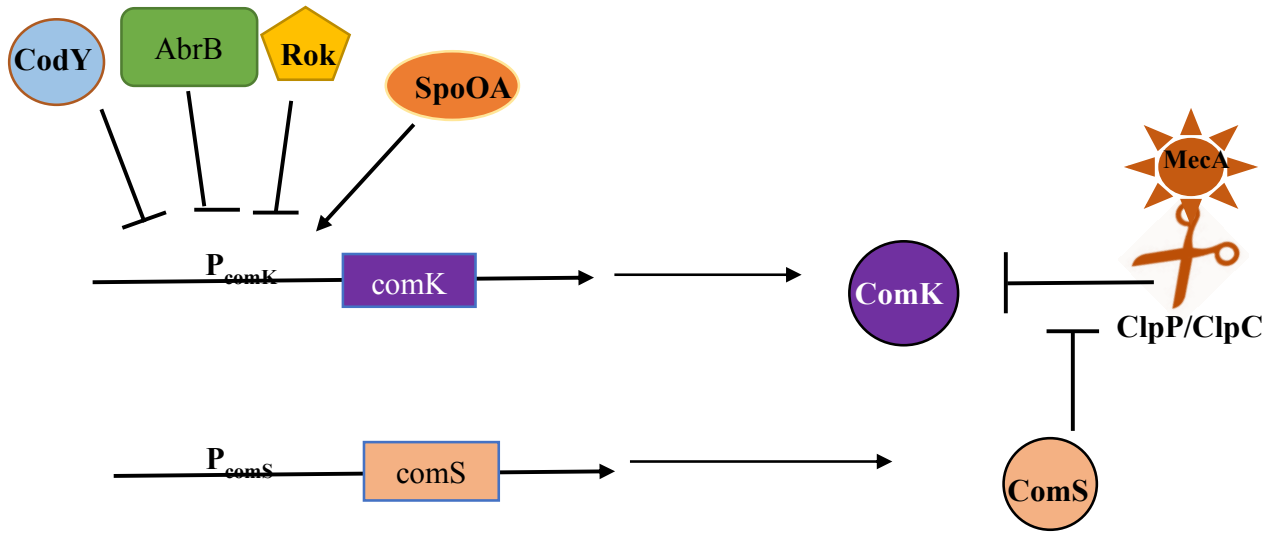


Figure 2. ComK regulation in *B. subtilis*. ComK is regulated at the transcriptional and stability levels. Transcription of ComK can be repressed by CodY, AbrB and/or Rok. At stability level, ComK protein can be degraded by MecA/ClpP/ClpC complex. ComS can bind to MecA, resulting in protection of ComK from degradation by ClpP/ClpC.

1.3 Transformation regulation in Streptococci

Streptococci are facultative anaerobes, Gram-positive, alpha-hemolytic, and oval shaped in either diploid or long chain form. Extensive phylogenetic data have revealed that all transformable streptococcal species share conserved uptake and processing proteins but differ in their inducing circuit and regulatory mechanisms. Competence is a specific physiological state during which bacteria can take up exogenous DNA from their environment and integrate it into their chromosomes. In Streptococci, competence for transformation involves internalization of DNA, recombination, and fratricide (S. Peterson, Cline, Tettelin, Sharov, & Morrison, 2000; S. N. Peterson et al., 2004; Tettelin et al., 2001). QS systems that regulate transformation in the Streptococci fall into two groups: ComRS and ComDE (Johnston, Martin, et al. 2014, Khan, 2016).

The mutans, pyogenic, salivarius, and bovis groups of streptococci use ComRS system (Fontaine et al., 2013; Mashburn-Warren, Morrison, & Federle, 2010; Morrison, Guedon, & Renault, 2013), that use a peptide encoded by *comS* known as XIP (σ^X-inducing peptide) to trigger the intracellular QS circuit mainly comprised of expression of ComRS and σ^X. ComR is the intracellular regulator that is activated by XIP binding, resulting in the activated expression of ComS and σ^X. The expression of σ^X activates competence-specific gene expression (Fontaine, Wahl, Flechard, Mignolet, & Hols, 2015).

The mitis and anginosus groups of streptococci use ComCDE system (Havarstein, Gaustad, Nes, & Morrison, 1996; Johnston et al., 2014). CSP pheromone encoded by *comC* and the TCSTS system consisted of ComDE regulate the gene expression on entry into competence. ComDE is activated by CSP binding and result in activation of σ^X expression, which is the master regulator of competence-specific gene expression (Lee & Morrison, 1999; S. N. Peterson et al., 2004).

Both ComRS and ComCDE systems form a positive feedback loop to increase σ^X expression thus activate expressions of σ^X-dependent genes. Two sets of genes are classified during competence: early gene expression is limited to bacteriocin production, immunity, and early gene expression, and late gene expression is associated with DNA uptake and recombination in the process of competence. In ComCDE system, pheromone not only stimulates expression of ComDE but also create a positive feedback loop to simultaneous increase expression of pheromone. In comparison, ComRS only induces expression of pheromone.

1.3.1 Transformation regulation in *Streptococcal pneumoniae*

As a member of the Mitis group, competence in *S. pneumoniae* is the most studied species, is transient, involving the specific expression of approximately 147 genes that are turned on and off abruptly, with regulation provided by a QS mechanism (Khan et al., 2016).

At a critical concentration, CSP acts through a TCSTS, the membrane-bound histidine kinase ComD, and the response regulator ComE to stimulate the development of competence (Havarstein et al., 1996). ComD and ComE, which are expressed during competence, belong to the TCSTS, whose activation is involved in autophosphorylation of an HK, ComD, and its cognate RR, ComE (Pestova et al., 1996).

ComE~P recognizes a ComE binding site (CEbs) upstream of several genes and activates expression of these genes, which are termed early genes. Early genes are ComE~P dependent, and their mRNA reaches a maximum 7-8 minutes after CSP exposure, while late competence genes are ComX dependent, and their mRNA peaks at 10-13 minutes (Lee & Morrison, 1999; Peterson et al., 2004). Among these early genes, the most important are *comAB, comDE, comX,* and *comW* (Lee, Dougherty, Madeo, & Morrison, 1999; Luo & Morrison, 2003; S. N. Peterson et al., 2004).

ComAB cleaves at a double-glycine motif of a CSP precursor (a 43 aa peptide) and the resulting mature 17 aa peptide is called CSP (Havarstein, Coomaraswamy, & Morrison, 1995). Two early competence proteins, ComX and ComW, play critical roles during competence development. Both proteins are transient, and their expression is regulated by the same QS system and TCSTS. The early protein σ^X is encoded by two identical genes, *comX1* and *comX2*, either of which can support normal competence development (Lee & Morrison, 1999; Luo, Li, & Morrison, 2003). Most operons in *S. pneumoniae* are controlled by σ^A, the housekeeping σ factor, but competence-specific genes are controlled by σ^X, the only alternative σ factor (σ) which can bind to RNAP core enzyme and recognize the non-canonical promoter sequence termed the "combox" (TACGAATA), which is upstream of late competence genes (S. Peterson et al., 2000; S. N. Peterson et al., 2004; Piotrowski, Luo, & Morrison, 2009) (see Figure 3). σ^X regulon includes genes conserved in streptococcal groups, and expression of genes involved in exogenous DNA uptake, recombination, fratricide, and immunity (Claverys & Havarstein, 2002; Luo & Morrison, 2003).

1.3.2 Transformation regulation by ComRS in *Streptococcus mutans*

As a member of Mutans group, *S. mutans* is a human oral colonizer and capable of natural genetic transformation during exponential growth (Kunal, 2012, Khan, 2016). In *S. mutans*, regulatory circuits for coordination of entry into competence is controlled by two systems ComCDE (Martin, Quentin, Fichant, & Claverys, 2006) and ComRS (Mashburn-Warren et al., 2010) in a manner of medium-dependent. Although the mechanisms for such behavior remain unclear, it is possible that free peptides in rich medium block XIP imports (Khan et al., 2016), or peptides eliminate the activity of XIP (Cook & Federle, 2014).

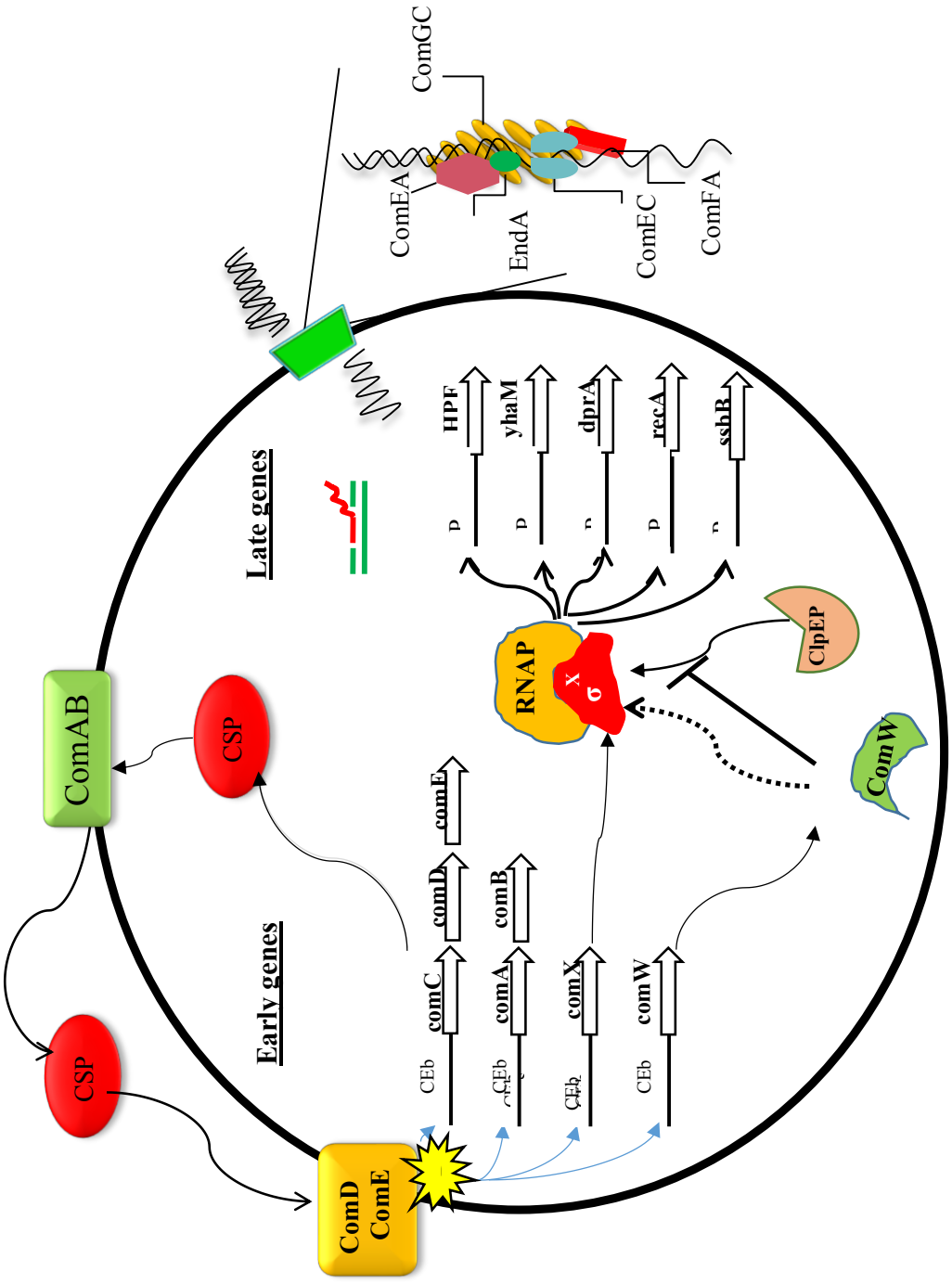


Figure 3. Transformation regulation in *S. penumoniae.* The expression of competence proteins ComAB and ComCDE are under the control of ComDE. Two proteins, σ^X and ComW play a role in expression of late competence genes. σ^X works as an alternative σ in transcribing competence specific late genes. ComW acts to protect σ^X protein from ClpEP proteolysis.

In peptide-free media such as CDM, development of competence is regulated by ComRS (Mashburn-Warren et al., 2010). ComS encodes a double-glycine signaling peptide, which is processed into XIP and exported via an unknown mechanism. At a critical concentration, XIP can be imported into the cytoplasm via oligopeptide permease. Then XIP binds ComR, forming a complex that elevates expression of ComS and σ^X, and the latter binds to RNAP, forming a whole enzyme that drives transcription of genes to respond to DNA uptake and transformasome machinery as well as murein hydrolase (S. Peterson et al., 2000). XIP activation of ComRS, result in a positive feedback loop to increase the production of ComS and XIP. Bacteriocin and immunity proteins are upregulated by ComED in response to elevated expression of σ^X.

In peptide-rich media such as THB, development of competence is regulated by ComCDE (Martin et al., 2006), that ComC encodes the precursor of CSP and ComDE is the histidine kinase senor and response regulator. ComDE is orthologous to BlpBH in *S. pneumoniae* that serve as TCSTS in the regulation of bacteriocin production. Thus, extracellular CSP induces the expression of competence protein as well as bacteriocin and immunity proteins. The precursor of CSP is processed and exported by ComAB transporter. CSP triggers competence via phosphorylated ComE, and triggers expression of bacteriocins and immunity proteins. The bacteriocin is considered to create pores that allow internalization of the competence pheromone XIP. XIP binding to ComR activates expression of ComS and σ^X. ComR creates a positive feedback loop for competence development, while σ^X activates expression involved in DNA integration, uptake, and recombination (see Figure 4).

1.3.3. Regulation of σ^X and ComW in *Streptococcus pneumoniae*

One critical early protein σ^X, encoded by *comX*, is the only known alternative σ factor in *S. pneumoniae* (Lee et al., 1999). Orthologs of *ComX (SigX)* have been found in all Streptococcal

genomes and transcribe late competence genes (Johnston and Martin, 2014). In fact, *comX* homolog seems to be present in all genomes of lactobacillales species (Martin et al., 2006). Expression of *comX* is regulated by two distinct QS systems: one is ComDE and CSP, and the other is ComRS and BIP (Havarstein, 2010).

During competence, σ^X forms the link between early and late genes. Pneumococci have two copies of *comX*, and either copy is sufficient for competence induction (Lee and Morrison, 1999; Luo and Morrison, 2003). Expression of σ^X accumulates to a high-concentration intracellular level via a positive feedback loop of ComDE and CSP. RNAP bound by σ^X forms holoenzyme, recognizing the cin-box of promoters of competence-specific genes that are required for DNA uptake and recombination (Campbell et al., 2001; Luo & Morrison, 2003; S. N. Peterson et al., 2004). As a transient protein directly under ComE regulation, σ^X first appears 5 minutes after addition of CSP, reaches a maximum protein level at 15 minutes, and disappears at 20 minutes (Luo and Morrison, 2003). The transient appearance of σ^X *in vivo* suggests that σ^X may be under tight control by proteases. ClpEP is the main protease complex to degrade the σ^X protein (Luo & Morrison, 2003; Piotrowski et al., 2009). The activity of σ^X in transcribing late genes has been studied extensively (Luo & Morrison, 2003; Sung & Morrison, 2005; Weng & Morrison, 2013) (see Figure 5).

Research has shown σ^X activity to be dependent on ComW. In the absence of ComW, σ^X activity decreases 90% during late gene transcription (Piotrowski et al., 2009). In a protease-deficient strain, the σ^X protein becomes stable and remains detectable until 40 minutes after CSP

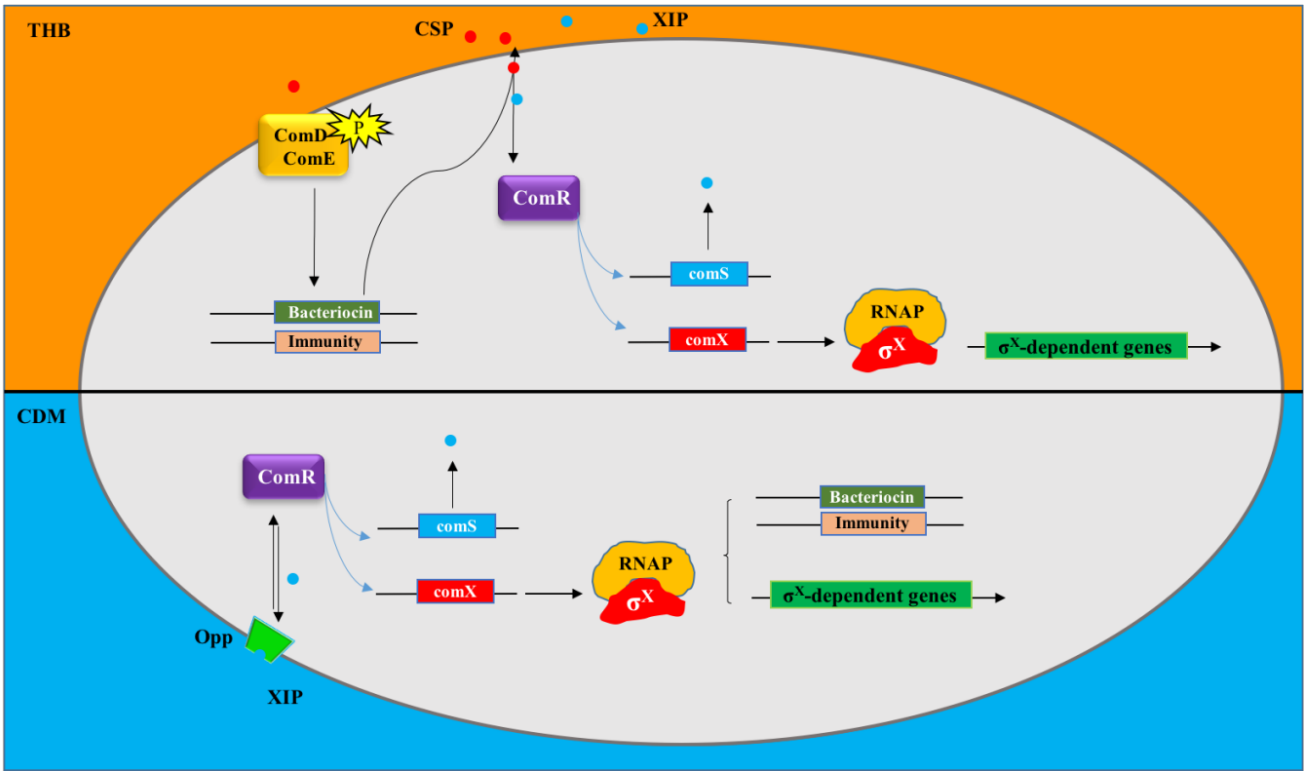


Figure 4. Transformation regulation in *S. mutans*. ComR can be activated by XIP. XIP is processed from its precursor ComS and recruited into the cytoplasm via Opp. The activated ComR-XIP can recognize a specific sequence in the promoter region of several genes and enhance their expression, including XIP precursor comS and comX. σ^X turns on expressions late genes that are involved in competence development.

induction. However, less than 0.1% transformation efficiency was detected in the *clpP-comW* double mutant, even though the σ^X protein accumulated at a high level (Sung & Morrison, 2005).

1.3.4. Regulation of ComW in *Streptococcus pneumoniae*

ComW is an early protein with a temporal induction pattern similar to that of σ^X (S. N. Peterson et al., 2004) and is required for full σ^X activity during transformation (Luo, Li, & Morrison, 2004). Ectopic expression of ComW and σ^X was found sufficient to restore transformation to 80% of a wild type in the *comW* deletion mutant (Luo et al., 2004; Sung & Morrison, 2005). Transformation efficiency decreases 10^4-fold in a *comW* deletion mutant (Luo, 2004). In a *comW* mutant, the amount of σ^X protein is 10% of a wild type when measured 15 minutes after CSP induction (Sung and Morrison, 2005). The σ^X protein is stabilized in the *clpEP* mutant, but the number of transformants remains as low as the number of the *comW* mutant, indicating that ComW is required for σ^X activity in late gene expression. ComW is conserved in 11 species in the mitis and anginous groups (*S. pneumoniae, S. mitis, S. pseudopneumoniae, S. oralis, S. tigurinus, S. dentisani, S. infantis, S. anginosus, S. oligofermentans, S. cristatus,* and *S. sinesis*). Among these species, the similarity to pneumococcal *comW* ranges from 86% in *S. pseudopneumoniae* to 40% in *S. tigurinusa* (Tovpeko, Bai, & Morrison, 2016). The molecular mechanism of ComW in the regulation of competence in *S. pneumoniae* is still unclear (see Figure 5). The sequence similarity of *comW* in all available species may offer clues to ComW's function in other species.

2. Sigma factors and transcription regulation in bacteria

2.1 The primary sigma factor and alternative sigma factors

Bacterial RNA polymerase (RNAP) is a large, multi-subunit enzyme containing five subunits (α, α, ω, β and β') called a core enzyme (Gross et al., 1998; Ishihama, 2000). A sigma (σ)

factor is a dissociable subunit of RNAP. When the σ factor binds to RNAP, forming a holoenzyme, this complex can start transcription at specific promoter sites by forming an open transcriptional complex and interacting with transcription activators to stimulate mRNA synthesis (Ishihama, 2000; Osterberg, del Peso-Santos, & Shingler, 2011).

It is common for bacteria to use multiple σ factors together with the RNAP core enzyme to regulate differential transcription of specific sets of genes (Osterberg et al., 2011). There are two distinct σ families: σ^{70} and σ^{54} (M. Lonetto, Gribskov, & Gross, 1992). The σ^{70} family members consist of up to four structurally conserved regions: 1, 2, 3, and 4 (Malhotra, Severinova, & Darst, 1996). Only regions 2 and 4 are well conserved in all members of the σ^{70} family, including the major RNAP and promoter binding determinants. Region 1 is found only in a subpopulation of the σ^{70} family, ensuring σ^{70} binding to DNA only when it is complexed with RNAP. Region 3 interacts with DNA upstream of certain promoters lacking the -35 consensus sequence. The σ^{54} family members lack sequence homology to the σ^{70} family and use a distinct pathway to form an open transcription complex (Campbell et al., 2002; Gruber & Gross, 2003).

The σ^{70} family members are classified into four major structurally and phylogenetically distinct groups. Group 1 consists of house-keeping σ factors, and groups 2 through 4 consist of alternative σ factors. The primary σ factor transcribes housekeeping genes, while the alternative σ factors control various adaptive responses to environmental changes. These alternative σ factors are often used in bacteria. Group 2 σ factors are usually associate with responses to stress conditions, including high temperature, nutrient limits, and other stresses related to the stationary

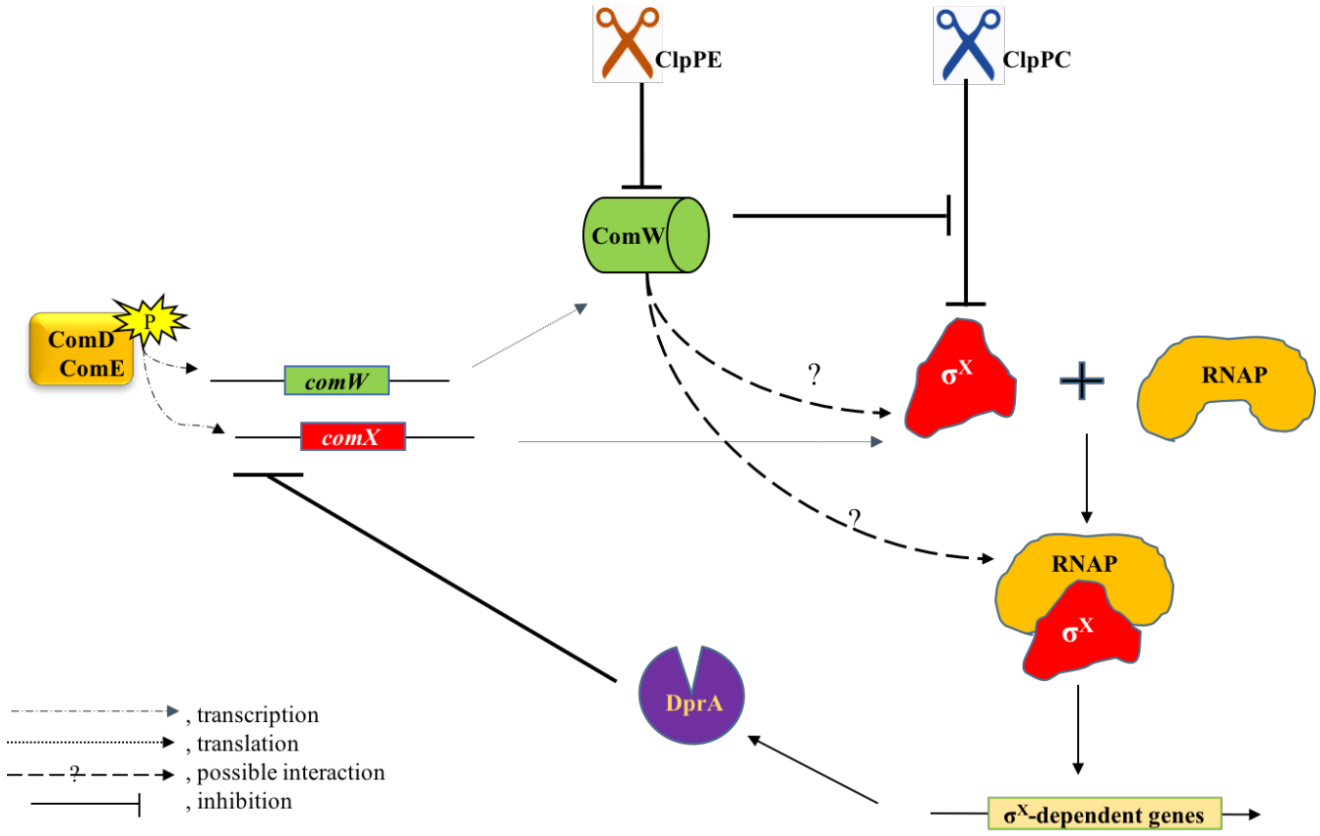


Figure 5. Regulation of σ^X and ComW in *S. pneumoniae*. *comX* and *comW* encode the most two important early proteins required for DNA uptake and recombination. Transcription of both genes are activated by phosphorylated ComE, and tightly controlled by a late competence protein, DprA. ComW plays a role in protecting σ^X from degradation of ClpPC, while itself can be degraded by ClpPE. ComW is required for σ^X activity, but the mechanism is unknown, which is represented by question marks. Green cylinder, ComW protein; green box, comW gene; yellow rectangles, ComDE; yellow star, phosphate; purple circles, DprA; orange circle, RNAP; pale yellow, σ^X-dependent genes.

phase. Group 3 σ factors mainly regulate flagellum biosynthesis, general stress, heat shock response, and sporulation. Group 4 σ factors are known as the extracytoplasmic function (ECF) group, that transcribes genes whose products are involved in extracytoplasmic functions, and each member are regulated by extracytoplasmic conditions, and these account for responses to signals outside cells or in the cell membrane (M. A. Lonetto & Buttner, 1994; Morrison et al., 2013; Paget, 2015). Bacteria can respond to environmental changes by switching their global gene transcription, which depends on specific σ factor binding to RNAP to direct the responses. In *E. coli*, RNAP (460 kDa) was first isolated in 1960 by Samuel and Jerard. RNAP in *E. coli* reaches approximately 2,000 molecules during the logarithmic phase, and two-thirds of the 4,000 genes in the genome are actively transcribed. Competition between σ factors for a limited amount of RNAP plays a central role in global transcriptional changes and cellular protein content.

2.2 Sigma factor switching

Bacteria use various mechanisms to direct RNAP transcription of specific genes in response to growth signals from the environment, and σ factor switching provides a mechanism for stress-specific gene expression driven by choice of a σ factor from its repertoire (Puri et al., 2014). To cope with environmental challenges, bacteria have developed different molecular strategies to regulate gene expression at the transcriptional level. The model organisms *Escherichia coli*, *Pseudomonas aeruginosa*, and *Methylobacterium extorguens* are used below to illustrate σ factor switching in bacteria.

2.2.1 Sigma factor switching in *Escherichia coli*

In *E. coli*, there is one primary sigma factor (housekeeping σ^{70}) and six alternative σ factors (envelope stress σ^{E}, flagellar regulation σ^{F}, stress response and stationary phase σ^{S}, nitrogen

assimilation σ^N, heat shock σ^H, and iron mebabolism σ^{FeCl}) (Cho & Palsson, 2014). The σ^S factor was identified as the master regulator of the transcription of genes under general stress and in the stationary phase in *E. coli*. The molecular mechanisms of σ^S activation in *E. coli* have been well characterized. The cellular level of σ^S can reach as high as approximately one-third the amount of σ^{70} (Paget, 2015). There is likely an aggressive competition in σ factor switching that facilitates σ^S recruitment to RNAP due to the lower affinity of σ^S for RNAP relative to σ^{70}. A σ factor binding protein, Crl, directly activates σ^S-dependent gene transcription as a σ^S-holoenzyme assembly promoter by directly interacting with σ^S (Banta et al., 2014; Pratt & Silhavy, 1998).

Another protein, the putative anti-σ^{70} factor Rsd in *Escherichia coli,* can bind to σ^{70} and sequester its activity, leading to enhanced σ^S capacity in competition over σ^{70} for access to core polymerase (RNAP) (Campbell et al., 2003). In the stationary phase, Rsd binds to region 4 of σ^{70} to block association with RNAP, leading to preferential binding of σ^S. Overexpression of Rsd causes reduced σ^{70}-dependent gene expressions, while the Rsd null mutant increases σ^{70}-dependent gene expressions (Piper, Mitchell, Lee, & Busby, 2009) (see Figure 6). Similarly, 6S RNA binds to σ^{70} to decrease its ability to transcribe genes, allowing σ^S-activated transcription.

2.2.2 Sigma factor switching in *Pseudomonas aeruginosa*

The Gram-negative pathogen *P. aeruginosa* have one primary σ factor (RpoD) and ten alternative σ factors (Fecl, Fecl2, PvdS, RpoS, AlgU, FpvI, FliA, RpoH, SigX, and RpoN) (Schulz et al., 2015). AlgU, an ECF σ factor, control transcription of genes specifically for alginate biosynthesis and cellular homeostasis (Hershberger, Ye, Parsek, Xie, & Chakrabarty, 1995). In the absence of stimuli, AlgU is tethered by the anti-σ factor MucA, which is in turn bound by MucB. In the presence of stress, the transmembrane protease AlgW is activated and cleaves MucA,

releasing MucA from MucB binding; followed by further MucP cleavage, AlgU is released into the cytoplasm. AlgU binds to RNAP by replacing other σ factors, initiating transcription of genes specifically for alginate synthesis (see Figure 7). A *mucA* mutant with a premature stop codon results in a severely truncated MucA, and protease Tsp acts to cleave AlgU instead of AlgW (Schulz et al., 2015; Schurr, Yu, MartinezSalazar, Boucher, & Deretic, 1996).

2.2.3 Sigma factor switching in *Methylobacterium extorquens*

M. extorquens is the most extensively studied Gram-negative facultative methylotroph bacterium, that can use methanol as carbon source. In *M. extorquens*, an ECF σ factor (σ^{ECF}) plays a critical role in adaption to environmental stress. The activity of σ^{ECF} is often negatively regulated by cognate anti-σ factors, preventing their access to the core RNAP. Specifically, σ^{ECF} is sequestered by the anti-σ factor NepR, blocking transcription under unstressed conditions. PhyR, conserved in α-proteobacteria, is a key regulator of general stress response in the methylotrophic plant colonizer *M. extorquens*. PhyR usually consists of a σ^{ECF}-like domain and a response regulator receiver domain (Gourion & Vorholt, 2008). Its σ^{ECF} domain is occluded by its receiver domain in the unphosphorylated form. PhyR mimics a σ factor sequestering NepR when its receiver domain is phosphorylated, releasing σ^{ECF} and allowing stress-related gene expressions (see Figure 8). A PhyR deletion mutant has very detrimental growth phenotypes under stress conditions (Francez-Charlot et al., 2009).

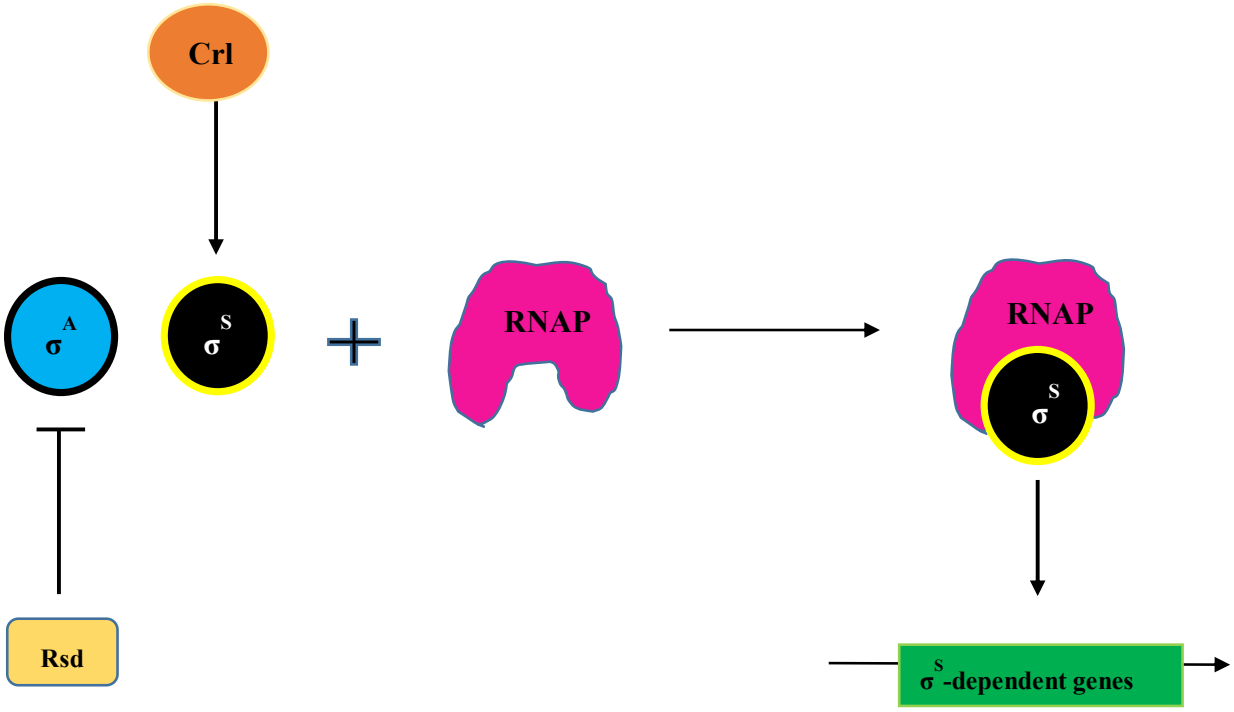


Figure 6. Sigma factor switching in *Escherichia coli*. σ^S and σ^A compete in binding to RNAP. Crl helps σ^S binding to RNAP by interacting with σ^S. Rsd works as an anti-σ^{70} in preventing its binding to RNAP.

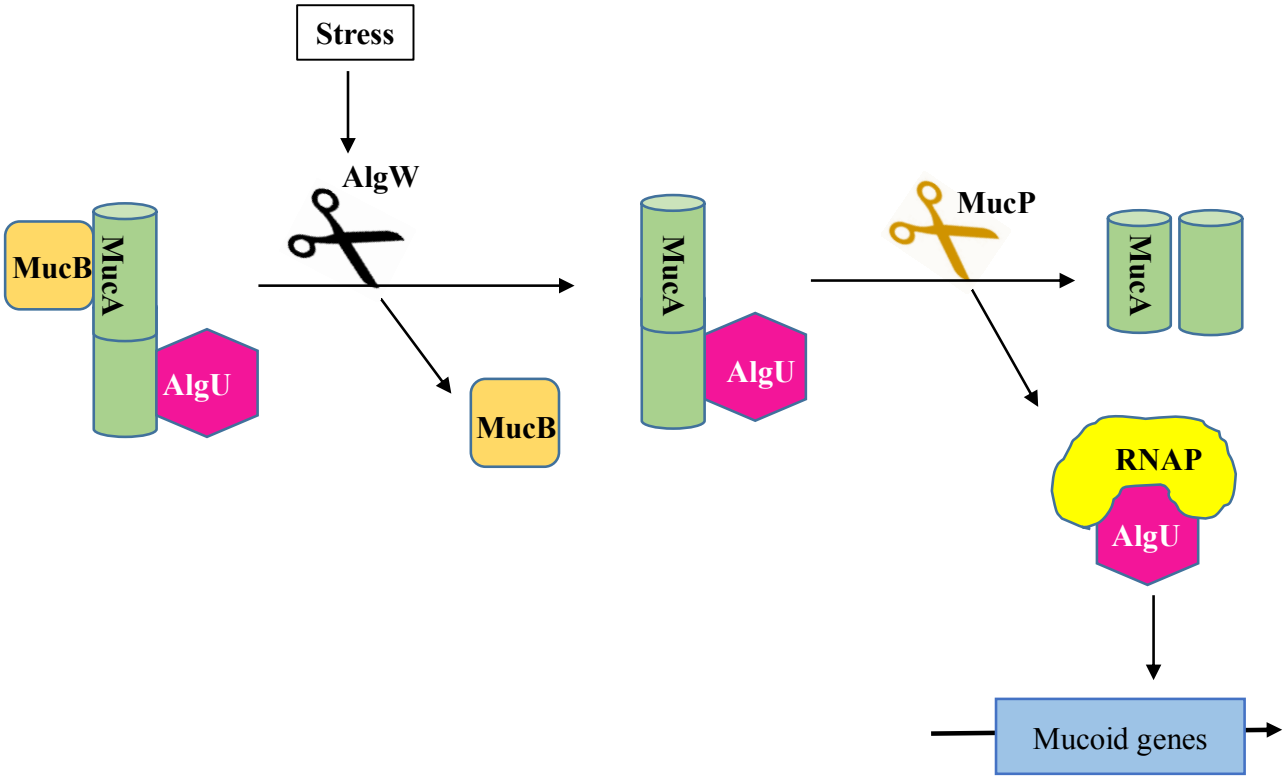


Figure 7. Sigma factor switching in *Pseudomonas aeruginosa*. Under normal growth conditions, MucA is bound by MucB and sequesters AlgU, inactivating AlgU. Under the stress conditions, AlgW cleaves the complex, and release MucB from MucA. Further cleavage by MucP results in release of AlgU from MucA. AlgU is released into cytoplasmic and active.

2.3 Sigma factor regulation

2.3.1 Sigma factor regulation at the transcriptional level

It is well-established that *rpoS* in *E. coli,* encoding σ^S factor, is a master regulator solely involved in response to environmental stress (HenggeAronis, 1996). On one hand, *rpoS* transcription can be negatively regulated by a transcription factor ArcA (Mika & Hengge, 2005), and a toxin-antitoxin (TA) system MqsR/MqsA (Wang et al., 2011). On the other hand, it can be positively regulated by CRP in response to nutritional stress. Suppression of *rpoS* transcription by direct binding of phosphorylated ArcA to *rpoS* promoter is believed to occur though inhibition of transcription activation by CRP, who shared the same binding sites to the promoter (Landini & Lacour, 2014; Mika & Hengge, 2005).

2.3.2 Sigma factor regulation at the translational level

The σ factor can also be regulated during translation. In *E. coli*, the alternative σ factor σ^H can be regulated at the translational level. Heat shock proteins (HSP) are transiently upregulated by the σ^H-RNAP holoenzyme. The σ^H factor is encoded by *rpoH*, and the 5' region of *rpoH* mRNA forms a temperature-sensitive secondary structure that sequesters the shine-dalgarno (SD) sequence. At high temperature, the secondary structure melts, thereby exposing free SD sites for initiation of *rpoH* translation (Morita & Yura, 1999).

2.3.3 Sigma factor regulation at the stability level

Regulated degradation of σ factors plays a major role in determining the quantity of σ factors in cells. In *E. coli*, the adaptor protein RssB directly stimulates σ^S degradation by proteasome ClpXP. A larger σ^S-RssB-ClpXP complex forms when ClpP accumulates to a specific amount. The complex degrades σ^S and releases RssB from ClpXP. Guanosine pentaphosphate

(ppGpp) stabilizes the σ^S protein by activating expression of the anti-adaptor proteins IraD and IraP, which bind to RssB and inhibit its activity.

2.3.4 Sigma factor regulation at the activity level

The activity of σ factors can be promoted by pro-σ factors. In *E. coli*, Crl is a small protein (15KDa) showed to positively regulate σ^S activity by direct binding to σ^S region 2, which is primarily responsible for -10 promoter, in formation of σ^S-holoenzyme (Banta et al., 2013). Crl is highly conserved in γ-proteobacteria, and Crl homologs from diverse species can interact with σ^S to promote transcription from σ^S-dependent promoters (Banta et al., 2014).

On the other hand, σ factors can be sequestered by anti-σ factors. For example, *B. subtilis* is known to encode seven σ^{ECF} factors, one of which is σ^W, which protects a bacterium from antimicrobial compounds (Cao et al., 2002; Helmann, 2016). Without antimicrobials, σ^W is sequestered by the anti-σ factor RsiW. Antimicrobials trigger the extracellular domain of protease PrsW and RasP to cleave RsiW, thus liberating σ^W to bind to RNAP and initiating σ^W-controlled gene transcription. Additionally, the activity of another alternative σ factor, σ^B, can be sequestered by anti-σ factor RsbW. However, under stress, RsbW can be antagonized by an anti-anti-σ factor, which dephosphorylates RsbV and releases free σ^B (see Figure 9).

3. Transcription initiation and its regulation

3.1 Transcription initiation

Transcription in bacteria relies on promoter recognition and transcription initiation. RNAP-bound σ factors recognize specific promoters with other initiation factors, unwinding the double-stranded DNA to form the transcriptional open complex. Subsequently, the transcription cycle proceeds to elongation, termination and σ factor recycle (Browning and Busby, 2016). Bacteria

positively or negatively regulate transcription by using transcriptional factors and small molecule effectors that affect promoter activity (Gruber and Gross, 2003, Browning and Busby, 2004). This thesis focuses on the regulation of transcription initiation by promoter recognition, RNAP binding, and σ factor binding factors. Regulation of transcription initiation by σ factor binding regulators is similar with the regulation of σ switching.

3.2 Transcription initiation regulated by promoter- and RNAP-binding factors

Transcription initiation can be activated or repressed by promoter-binding factors. Several mechanisms of transcription activation have been characterized. Transcription activation can be enhanced by transcription factors binding directly to the promoter sequence, including activators binding upstream of the promoter and contacting the carboxy-terminal domain of the alpha-subunit. That either activates binding to the promoter adjacent to the -35 element and direct interaction with σ subdomains or induces a conformational change in the promoter for recruitment of RNAP. For instance, the cI protein of bacteriophage (λcI) activates transcription by directly binding to a promoter and σ^{70} in forming a complex (Nickels, Dove, Murakami, Darst, & Hochschild, 2002).

Transcription repression occurs when a repressor blocks promoter sites that overlap with the sites recognized by RNAP, preventing the binding of RNAP to those sites; or the repressor can loop the promoter region and prevent recognition of the promoter by RNAP; alternatively, the repressor can repress activators' binding to the promoter, thereby disabling activation for transcription initiation. For example, GalR mediates repression of the galactose operon in *E. coli* by looping the promoter region. Another example is CytR, an anti-activator, which represses transcription by binding to CytR/CRP-regulated promoters with the help of activator cAMP

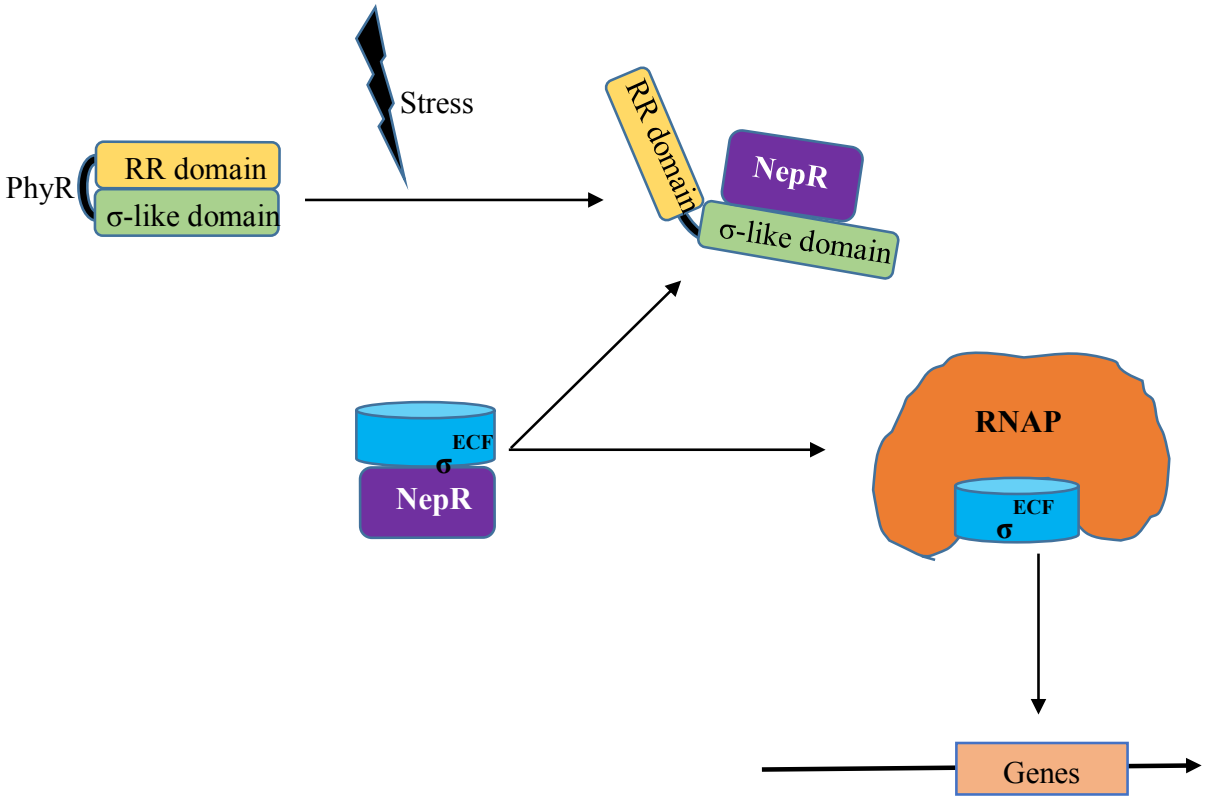


Figure 8. Sigma factor switching in *Methylobacterium extorquens*. Under normal conditions, PhyR is inactive by interaction between response regulator domain (RR) and σ–like domain. σ^{ECF} is sequestered by anti-σ factor NepR. Under stress conditions, RR domain of PhyR is phosphorylated resulting two domains are separate. The σ–like domain can bind to NepR, and release σ^{ECF}. Activated σECF gets access to RNAP to activate stress response gene expression.

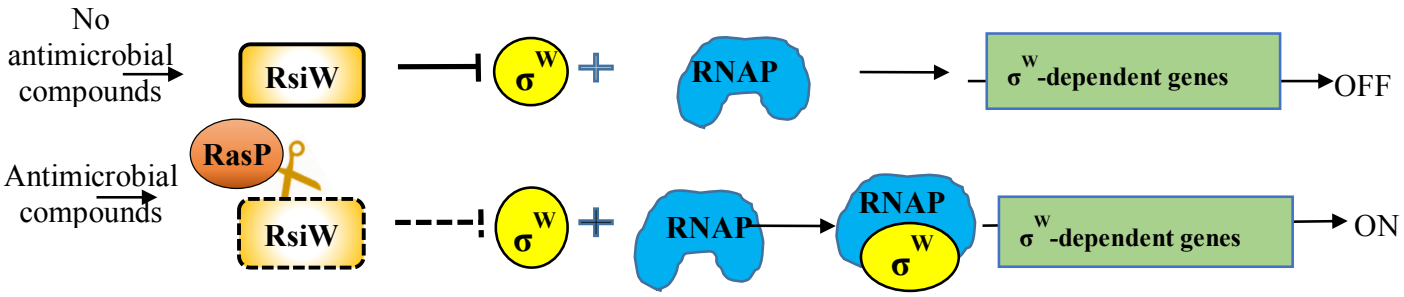


Figure 9. The σ^W regulation at the activity level in *B. subtilis*. Anti-sigma factor RsiW sequesters σ^W under no antimicrobial compound condition. Addition of antimicrobial compounds triggers protease RasP and PrsW digestion of RsiW, resulting in the release of σ^W. σ^W binding to RNAP activates σ^W–dependent gene expression.

receptor protein (CRP), subsequently preventing binding of RNAP (Browning & Busby, 2016; ValentinHansen & Pedersen, 1996). In *S. aureus*, there is a set of promoters, α/CTD/upstream dependent, where transcription initiation can be inhibited by a phage protein Gp67. Gp67 inhibits transcription by binding to σ^A to interfere the promoter recognition of σ^A-holoenzyme. Only promoters where activity depends on α/CTD/upstream can be inhibited by Gp67 (Osmundson & Darst, 2012).

Transcription initiation can also be regulated by RNAP binding factors that bind directly to RNAP instead of the promoter, affecting polymerase kinetics such as open complex formation. For instance, in *E. coli*, RNAP-binding transcription factor DksA regulates transcription by binding to the secondary channel of RNAP. DksA together with ppGpp inhibits specific promoters of rRNAs and most tRNAs (Haugen & Gourse, 2008).

4. Translation initiation and its regulation by ribosome-associated factors in bacteria

Protein synthesis includes four major phases: initiation, elongation, termination, and recycling. Various mechanisms regulate translation in bacteria. Regulation of translation initiation by ribosome-associated factors is the primary focus here. The bacterial 70S ribosome, which is composed of one small 30S and one large 50S subunit, provides a platform upon which aminoacyl-tRNAs can form a polypeptide chain. The 30S subunit ensures that the mRNA is positioned correctly and that the tRNAs are paired with the correct codons. The 50S subunit is the peptidyl-transferase center (PTC) for peptide bond formation. In addition, the 70S ribosome contains A-, P-, and E- sites for tRNA binding, with peptidyl-tRNA access in the P-site, an aa-tRNA in the A-site, and the E-site for uncharged tRNA exit.

Under nutrient deprivation, the stringent enzyme RelA interacts with ribosomes containing uncharged or deacylated tRNA at the A-site, with the interaction facilitated by synthesized ppGpp, a broad regulator in many bacteria. Rapid synthesis of ppGpp, which reaches millimolar concentrations in cells within minutes, inhibits transcription of components of translational apparatus such as rRNAs and tRNAs (Starosta & Wilson, 2014). Other factors have been identified that stop translation, such as RelE, HPF, and RMF, or that re-program translation, such as MazF and BipA. During the stationary phase, the RelE toxin cleaves mRNA at the A-site, causing truncated 3' end of mRNAs lacking a stop codon. The anti-toxin RelB inactivates RelE. Starosta has proposed the use of BipA to regulate translation of a specific subset of mRNAs under stress conditions, with factors such as tmRNA, ArfA and ArfB extensively immerging to rescue stalled ribosomes (Starosta et al., 2014). Understanding of the translation- regulating mechanism will provide insight into how bacteria respond to stress and will help to identify new targets for novel antimicrobial medications.

4.1 Translation regulation during transformation in *Streptococcus pneumoniae*

During competence regulation in *S. pneumoniae*, a global shift in protein synthesis was first observed by Morrison (1979). The global protein synthesis pattern was pulse-labeled with S-methionine followed by one-dimensional (1D) fluorogram analysis, revealing two prominent features of protein synthesis responding to competence. One was a striking decrease in the synthesis of many proteins, and the other was a robust increase of about 11 principal bands. The upshift pattern can be explained by the expression of σ^X and a set of new gene transcriptions (Lee et al., 1999). Transcriptome data identify many competence-induced genes and a cascade of transcriptional controls of the QS pathway. However, the mechanism for the downshift pattern of the housekeeping mRNA remains unclear (see Figure 10).

4.2 Core regulons - ribosome-associated factors during transformation in Streptococci

Recent transcriptome data comparing and combining all published data have identified three major regulons and a core set of 27 panstreptococcal competence-activated genes within the σ^X regulon shared by all streptococci (see Figure 11). Of the 27 genes identified, 18 are required for internalization of DNA and recombination (Khan et al., 2016; S. N. Peterson et al., 2004). The 9 additional genes are invariably over-expressed at competence and play an underappreciated role in competence regulation. Streptococcal species were grouped based on sequencing results from sources using 16S RNA sequencing (Kawamura & Ezaki, 1995) and whole genome sequencing (Richards et al., 2014). ComDE-dependent regulation exists in the anginosus and mitis groups, whereas ComRS-dependent regulation occurs in the salivarius, bovis, pyogenic, and mutans groups.

4.3 Translation regulation by ribosome-associated factors in other bacteria

During logarithmic growth of bacteria, most ribosomes are involved in translation and formed 70S complexes on mRNA. After translation, the two subunits are separated and recycled for the next binding process to mRNA. Studies suggested that ribosome hibernation can prevent ribosomes from being degraded by ribonucleases (Gonzalez, 2017; Yoshida & Wada, 2014).

Bacteria use different strategies to survive under harsh conditions. Two mechanisms of ribosome hibernation are known: one mediated by RMF and a short HPF in part of γ-proteobacteria, and the other involved by only long HPF in most bacteria. RMF is only found widely in the γ-proteobacteria class. Ribosome hibernation factor (HPF)-controlled ribosome inactivation is critical for survival of bacteria during a prolonged stationary phase. Multiple sequence alignment and phylogenetic analyses indicate that most bacteria have at least one HPF homolog. Bacteria use different mechanisms to regulate ribosome contents and activity. In *E. coli*, the number of

ribosomes increases proportionally with the rate of growth according to cellular energy needs, but the number of ribosomes in *L. lactis* increases less when it shifts from a nutrient-limited medium into a fresh medium, indicating regulations of ribosome activity is preferred over production of ribosome contents (Scott & Hwa, 2010). Cells control wasteful protein synthesis by controlling the activity of existing ribosomes (Gonzalez, 2017). Almost all bacteria possess proteins that facilitate 100S ribosome formation. In Gram-negative bacteria, 100S ribosomes form by means of RMF and HPF producing 90S dimer that is then stabilized into the mature 100S form. In Gram-positive bacteria, ribosome hibernation is mediated solely by a long HPF homolog.

RMF- and short HPF-mediated 100S ribosome formation occur mostly in the stationary phase, whereas long HPF-mediated 100S ribosome formation exists in whole growth phases. The 100S ribosome is translationally inactive but sets apart into two translationally active 70S ribosomes after shift from a nutrient-limited medium to a fresh medium. The physiological role of ribosome dimerization has remained enigmatic. Lack of HPF in human pathogens *L. monocytogenes, P. aeruginosa, B. anthracius,* and Clostridium species showed loss of competitive advantage under stress conditions and defect in infection in mouse model (Kline & Portnoy, 2015). Explore researches in ribosome hibernation, detailing cases of species have been shown that ribosome hibernation mediated by HPF and/or RMF could contribute to bacterial pathogenesis, and the 100S ribosome may be a new target for treatment of chronic infections (Beckert et al., 2017; Khusainov et al., 2017).

4.3.1 Translation regulation by stress-specific ribosomes in *Escherichia coli*

Bacteria use diverse protein factors that interact with ribosome to cope with different types of stress. *E. coli* has more than 33 toxin-antitoxin (TA) systems, that involved in programed cell-death, bacterial persistence, and temporial cell arrest under stress conditions (Zhang & Inouye,

2005). MazEF TA system is widely distributed in bacteria (Pandey & Gerdes, 2005). The name "MazE" originates from the Hebrew term "ma-ze," meaning "What is it?" (Metzger et al., 1988). In *E. coli*, MazEF is the most intensively studied toxin and anti-toxin system which cleaves cellular mRNA, and the benefits of MazEF regulation is controversial: some suggested a role of bactericidal (Hazan, Sat, & Engelberg-Kulka, 2004), and other suggested bacteriostatic (Pedersen & Gerdes, 2002). MazE encodes an antitoxin that sequesters the toxin MazF by forming a heterohexamer, where the C-terminal of MazE extends around a MazF dimer (Simanshu & Patel, 2013). During the stationary phase, the active MazF cleaves mRNA at ACA sequences, removing the SD sequence both *in vitro* and *in vivo* (Zhang et al., 2005). At the same time, MazF cleaves 16 rRNA of the 30S subunit at the ACA sequence, removing the anti-SD sequence. In summary, MazF reprograms translation of a specific set of leaderless mRNAs by specific ribosomes (Starosta et al., 2014) (see Figure 12).

4.3.2 Translation regulation by HPFsa in *Staphylococcus aureus*

Suppression of translation during stress conditions is a strategy of bacterial adaptation to the environment. In *S. aureus*, a 22.2 kDa stationary-phase protein HPFsa may suppress translation initiation by dimerizing the 70S ribosomes into an inactive 100S complex. Unlike *E. coli*, which contains a short form of HPF that inactivates ribosomes in concert with Ribosome Modulation Factor (RMF) and YfiA, *S. aureus* has a long form of HPF that functions equivalently to RMF and HPF. HPF exists in all growth phases of *S. aureus*, but the loss of HPF causes massive ribosome degradation at the onset of the stationary phase and hyperactive translation in the logarithmic stage. In one study, HPF null mutant drastically increased ribosome occupancy at the 5′ end of specific mRNAs under nutrient-limited conditions (Basu & Yap, 2016; M. Ueta & A. Wada, 2010). In addition, N-terminal domain (NTD) of HPFsa binds to the 30S subunit of ribosomes and C-

terminal domain directly interacts with another ribosome to form 100S ribosomes. Binding sites of the NTD and 30S subunit are similar to HPF and YfiA in *E. coli* (Khusainov et al., 2017). Ribosome dimer formation is a strategy of responding to stress conditions shared among bacteria. Ribosome hibernation shuts down protein translation and provides protection against antibiotics. HPFsa binding to the 30S subunit may occlude several antibiotic binding sites that are active translational sites. By robustly blocking ribosomal activity during translation of mRNAs, *S. aureus* can preserve its ribosomal integrity and survive under harsh conditions.

4.3.3 Formation of 100S ribosomes by long HPF in *Listeria monocytogenes*

During ribosome hibernation, *L. monocytogenes* HPF is required to dimerize pairs of 70S ribosomes into inactive 100S ribosomes. One study found that deletion of HPF in *L. monocytogenes* led to decreased virulence in a mouse model with infection and increased sensitivity to antibiotic exposure (McKay & Portnoy, 2015). The HPF deletion mutant was sensitive specifically to the ribosome-acting class of aminoglycoside antibiotics but not to other classes of bacterial antibiotics. HPF is needed for optimal fitness and pathogenesis, and HPF-deficient bacteria was impaired under certain conditions (Kline et al., 2015).

4.3.4 HPF binding-induced stabilization of dimerized 100S in *Escherichia.coli*

Ribosome hibernation in *E. coli* has been studied extensively in the past two decades. During stationary phase, 30S subunits can be bound by RMF, and its bindings sites overlap with the mRNA channel of the 30S subunit Shine-Dalgarno sequence, preventing interaction between SD-anti-SD helix formation between mRNA and 16S ribosomal RNA. RMF alone can promote dimerization of the 70S ribosome to the 90S form (Masami Ueta & Akira Wada, 2010).

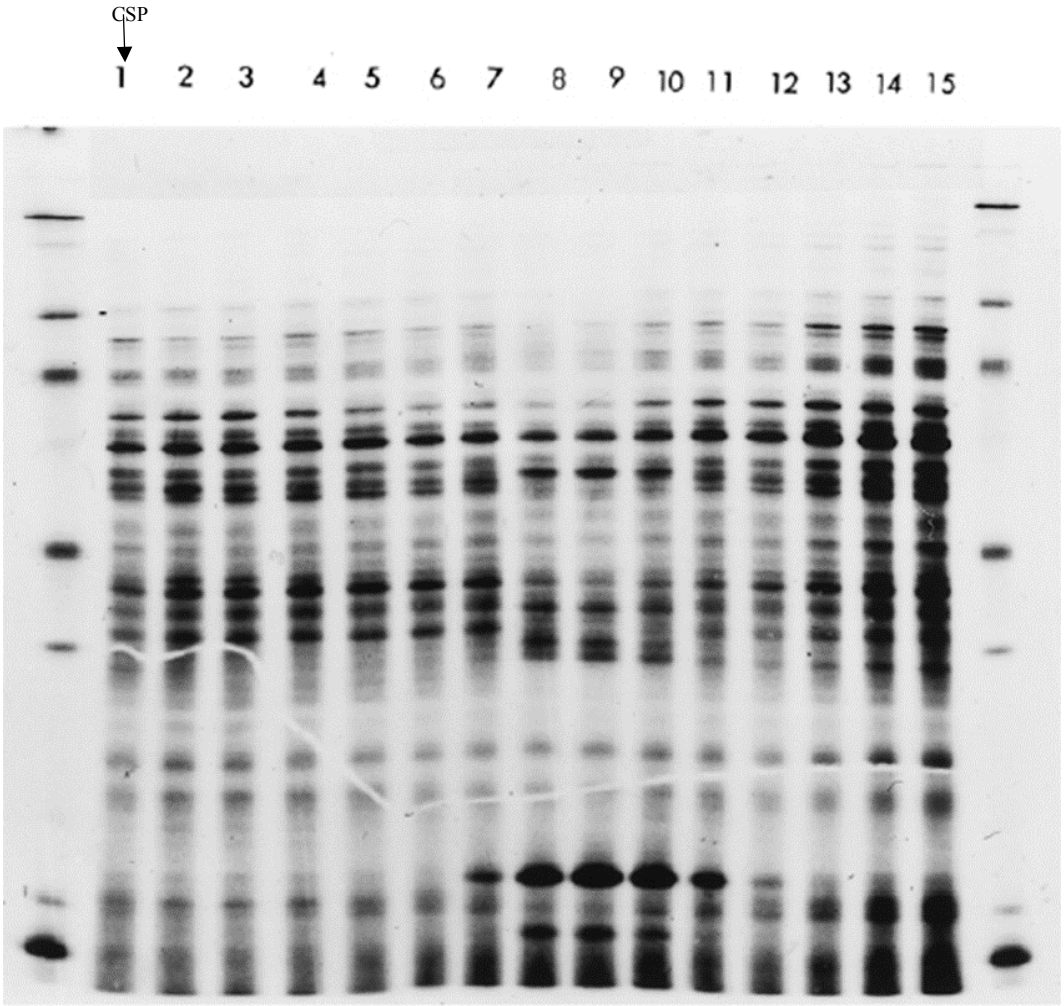


Figure 10. Competence depends on a small set of proteins in *S. pneumoniae*. The 1D gel is cited from Morrison and Lee (2000). CSP was added at time T_1, and continusouly exposure until time T_{15}. A pneumococcal culture was pulse-labeled with S35-methionine for successive 3-minute periods during a competence cycle. A set of proteins showed increased intensity during T_7 to T_{11}, while a set of proteins showed decreased intensity. Competence at more than 25% was corresponding to T_9 to T_{14}.

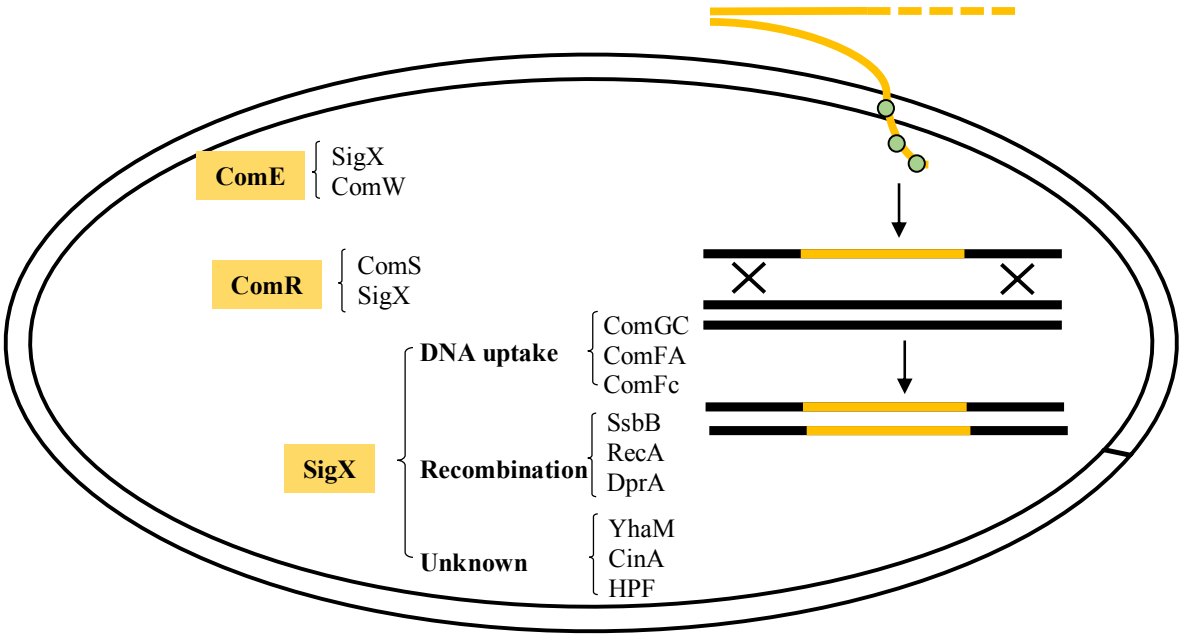


Figure 11. Core regulons in streptococci. A core set of competence genes in diverse streptococcal groups. Three set of regulons include ComE, ComR and σ^X. ComE loci are associated with bacteriocin production and immunity. ComR regulons provide the link between early genes and late genes. The σ^X regulons include proteins required for DNA uptake and subsequent recombination. There are nine genes have unknown roles in competence.

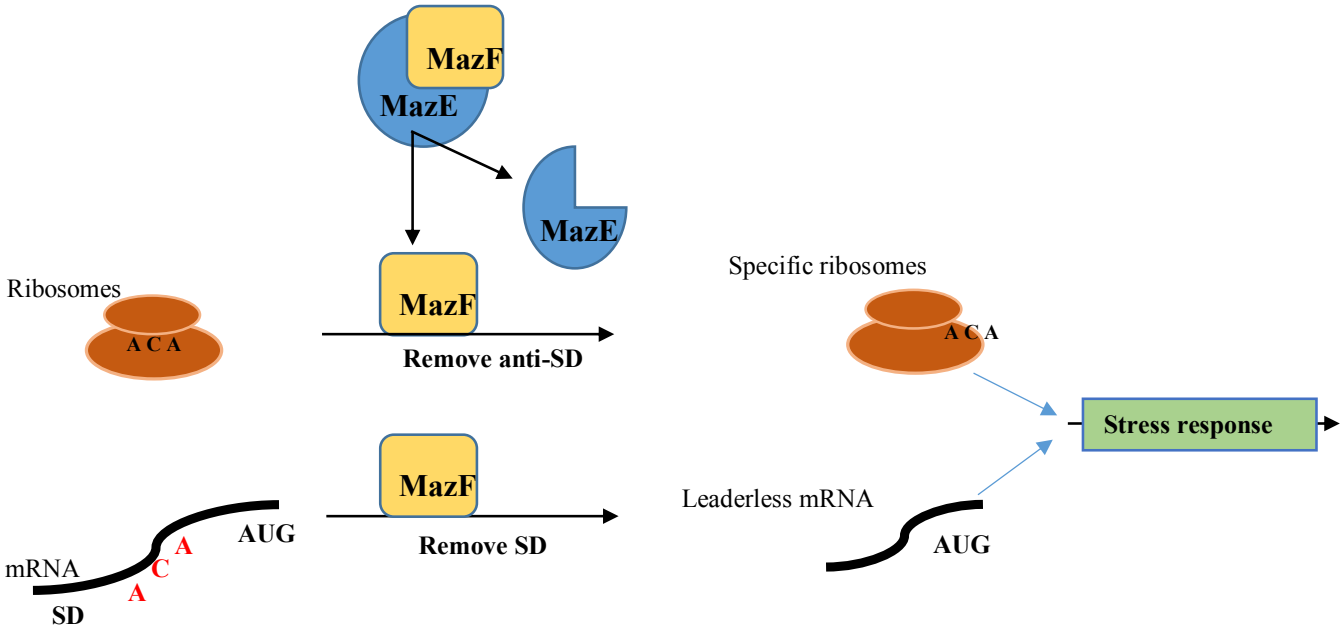


Figure 12. Translation regulation mediated by MazEF in *E. coli*. Under stress conditions, toxin MazF is released from its anti-toxin MazE, and become active. MazF is an endoribonuclease that cleaves at ACA sites at 16r RNA of 30S ribosomal subunits, and 5' end of a specific sets of mRNAs. Cleavage of 5' end of mRNA results in leaderless mRNA that can be recognized by the specific set of ribosomes cleaved and generated by MazF. MazEF mediates the stress-induced gene expression.

HPF and YfiA (a paralog of HPF) showed nearly identical binding sites to 30S subunits with those of mRNA, tRNA, and initiation factors, blocking translation initiation. HPF and YfiA binding induce a conformational change of the 30S head domain, causing 100S dimer formation (Polikanov & Steitz, 2012). The ribosome hibernation as an interconversion system between active and inactive forms of ribosomes is an important survival strategy for *E. coli*.

4.3.5 Sufficiency of YfiA for ribosome dimerization in *Lactococcus lactis*

An additional stationary-phase protein YfiA, with 40% sequence similarity with long HPF is required to dimerize ribosomes under nutrient limited conditions. It was shown that YfiA could dimerize *L. lactis* ribosomes without the presence of other protein factors (Puri et al., 2014). In addition, YfiA substitution for HPF_{ec} and RMF_{ec} *in vitro* can dimerize the *E. coli* ribosome. In examining this heterologous function, Puri and colleagues proved that YfiA alone is sufficient for ribosome dimerization. Moreover, the YfiA deletion mutant was similar to the wild type in that it survived starvation, but the mutant had a prolonged lag phase when cells were transferred to a fresh medium. *L. lactis* did not adjust its ribosome content to meet the needs of protein synthesis, suggesting an excess capacity of ribosomes at low growth rates (Puri et al., 2014). By reversibly inactivating ribosomes, cells do not waste resources on unnecessary protein synthesis but rather are able to recover protein activity as soon as nutrients again become available.

4.3.6 Translation regulation by HPF in *Bacillus subtilis*

Long HPF has been shown to be sufficient for 100S formation in various bacteria under nutrient deprivation and other stress conditions. HPF is required for ribosome dimerization for the efficient regrowth after transfer stationary-phase cells to fresh medium (Akanuma et al., 2016). Recently Beckert and colleagues found long-form HPF in *B. subtilis* dimerizes 100S ribosomes in a divergent mechanism compared to *E. coli*. NTD binds to sites of 30S subunits overlapped with

mRNA and tRNAs, while CTD unexpectedly binds to another CTD from the second 70S ribosome, thus forming 100S ribosomes. It is believed that NTD binding to the 30S subunit interferes with binding of mRNA and tRNA, blocking their access to the 70S ribosome and preventing translation initiation. Yet the physiological significance and formation of ribosome dimers remain unclear.

4.3.7 Translation regulation by YhaM-mediated RNAs in *Bacillus subtilis*

Controlling mRNA stability is an important means of gene expression in bacteria. In *B. subtilis*, the major 3′ to 5′ exoribonuclease, PNPase, is present in addition to a second 3′ to 5′ exoribonuclease, RNase R. In addition, Oussenko identified another 3′ to 5′ exoribonuclease that was Mn+ dependent, the YhaM protein. This protein was found to contain a C-terminal HD domain in metal-dependent phosphohydrolases and an OB-fold domain in many oligosaccharides- and oligonucleotide-binding proteins (Oussenko & Bechhofer, 2002).

Generally, rRNA operons in bacteria are transcribed as 30S RNA precursors that are matured at the transcriptional level by RNase III, yielding 16S, 23S, and 5S rRNA precursors that are further matured by other ribonucleases. For 23S rRNA, maturation occurs by cleavage by nuclease Mini-III at both sides of a double-stranded RNA in the RNase III mutant. In the absence of Mini-III, 23S rRNA maturation can be catalyzed by 3' to 5' exoribonuclease, principally RNase PH and YhaM, together with the 5' to 3' exoribonuclease RNase J1 (Redko & Condon, 2010).

5. Significance of Competence Study

Competence for genetic transformation was first observed by Frederick Griffith in 1928 during transfer of thermostable material between different strains of *S. pneumoniae*, and the thermostable material was identified as DNA (Avery et al., 1944). Almost a hundred years later, the regulation of competence has been studied extensively. However, many competence regulatory

mechanisms remain unknown. Researchers have argued that the main benefits of natural competence and DNA uptake apply to DNA employed for food, evolution, and genome repair and maintenance, and newly incorporated DNA can be used as a nucleotide resource or for genetic evolution by homologous recombination (Charpentier et al., 2011; Finkel & Kolter, 2001; Ochman, Lawrence, & Groisman, 2000; Redfield, 1993). Transformation of genes between species provides genome plasticity beyond the species' boundaries (Daubin & Szoellosi, 2016). Almost all transformable bacteria share common mechanisms of DNA uptake and processing, but their mechanisms of competence regulation vary (Johnston et al., 2014). Study of competence regulation in *S. pneumoniae* could provide valuable clues for an extensive competence regulation repertoire in all naturally transformable species. Moreover, as acquiring virulence or antibiotic resistance genes poses a threat of contagion, greater understanding of competence regulation would support the development of vital vaccines.

6. Specific Aims

Transformation in *S. pneumoniae* has been studied extensively for several decades, and two early competence proteins, σ^X, and ComW, have been of interest to our laboratory for almost two decades. σ^X is the master regulator of competence and the only alternative σ factor in *S. pneumoniae*. ComW is crucial for σ^X activity and transformation. Extensive research has been performed to decipher the regulatory mechanism of ComW in competence. The absence of ComW causes σ^X activity during transcription of late genes to decrease by 90% and transformation efficiency to decrease by 99.9% (Bartilson et al., 2001; Luo et al., 2004). ComW has been hypothesized to indirectly support σ^X activity by interacting with an anti-anti-σ factor, by outcompeting σ^A in σ factor switching, or by directly facilitating σ^X access to RNAP (Tovpeko et

al., 2016; Weng et al., 2013). In this three-phase study, two phases were performed to test hypotheses regarding ComW regulation in competence and to develop a model of competence regulation that predicts control of the activity of the alternative σ factor by means of a small protein. The third phase of this study was conducted to investigate the role of HPF and YhaM in competence regulation.

The first study phase began with suppressor screening to identify an anti-σ factor with which ComW interacts during transformation. A MarinerT7 transposon mutagenesis screening was used to provide a mutant library containing random mutations that might avoid the requirement for ComW during transformation. To identify the anti-σ factor, several compounds of transformation assays were used to screen mutants that were transformable in the Δ*comW* background. Transformants were tested, and linkage of the Δ*comW* suppressor was examined. It was hoped that finding this kind of mutation would provide a clue to the role of ComW in regulating σ^X activity.

The second phase of this study was intended to determine whether ComW directly interacted with σ^X, which would indicate the role of ComW in competence regulation. Yeast two-hybrid was used to determine whether ComW interacted directly with σ^X or σ^A. σ^X and ComW have a similar expression pattern that appears at 5 minutes, peaks at 20 minutes, and disappears by 40 minutes after CSP exposure. A direct interaction between the two transient proteins is possible, and the preliminary result of an earlier study (Weng, 2009) showed a direct interaction between the two. Vector swap in interacting pairs of proteins as well as constructions of fusion expression with truncated forms was used to confirm the interactions identified by yeast two-hybrid assay.

The third phase of this study was performed to investigate the role of HPF and YhaM in competence regulation. HPF and YhaM have been identified as belonging to core regulons that are

highly conserved in all Streptococci; however, their role in competence regulation remains unknown. Sequence similarities to proteins with known mechanisms in other bacteria suggested that the activities of the two proteins might be related to ribosome activity and protein expression re-programming during competence. In addition to the sequence similarities, Morrison and Baker (1979) observed that competence depended on expression of a set of proteins. To determine a connection between HPF and YhaM and the protein expression pattern change during competence, 1-dimensional and 2-dimensional electrophoresis gels were employed in addition to pulse-labeling of samples.

This study aimed to investigate the roles of ComW, HPF, and YhaM during competence regulation. Understanding the role of ComW in competence regulation will pave the way for a *de novo* mechanism in which σ factor activity can be regulated by an activator in Gram-positive bacteria. Furthermore, investigation of the role of HPF and YhaM may shed light on mechanisms of re-programming of protein expressions under specific conditions in Streptococci.

II. MATERIALS AND METHODS

2.1 Strains, plasmids and culture media

S. pneumoniae was grown in complete CAT medium, a standard peptide-rich medium, with 1.5% agar to make CAT agar plates. CAT medium was prepared as follows: 5 g yeast extract, 5 g tryptone, 10 g enzymatic casein hydrolysate (ICN Nutritional Biochemicals), and 5 g NaCl in 1 L H_2O; autoclaved for 40 minutes at 121°C; and was then added to 0.016 M K_2HPO_4 and 0.2% glucose.

S. pneumoniae was grown in peptide-free medium, also called chemical defined medium (CDM). CDM contains per liter of H_2O: NaCl, 8.8 g; glucose 2.0 g; BSA (bovine serum albumin), 0.8 g; $MgSO_4$, 0.7 g; $MnSO_4$, 5mg; $CaCl_2$, 33 mg; $NaHCO_3$, 2.5 g; choline, 10 mg; iron ($Fe(NO_3)_3 \cdot 9H_2O$, 1mg; $FeSO_4 \cdot 7H_2O$, 5 mg); phosphate (K_2HPO_4, 0.2 g; KH_2PO_4, 1 g; $NaH_2PO_4H_2O$, 3.2 g; $Na_2HPO_4.7H_20$, 13.9 g); bases (adenine, 20 mg; guanine hydrochloride, 20 mg; uracil, 20 mg); vitamins (riboflavin,0.14 mg; thiamine hydrochloride, 0.3 mg; pyridoxine, 0.3 mg; uracil, 50 mg; nicotinic acid, 0.3 mg; biotin, 0.006 mg); amino acids (L-arginine, 200 mg; L-asparagine, 10 mg; adenine, 50 mg; L-glutamine, 30 mg; L-histidine, 150 mg; L-tyrosine, 50 mg; L-isoleucine, 6.6 mg; L-lysine, 420 mg; L-phenylalanine, 50 mg; L-proline, 50 mg; L-serine, 120 mg; L-threonine,195 mg; L-tryptophan, 50 mg; L-valine, 6 mg).

Escherichia coli was grown in Luria Broth (LB) medium, prepared from 1% tryptone, 0.5% yeast extract, and 5% NaCl, with 1.5 % agar to make LB agar plates.

Saccharomyces cerevisiae was grown in yeast extract-peptone-dextrose (YPD), or synthetic selective media (SD), and stored with glycerol at a final concentration of 15%. YPD was prepared from 20 g of dextrose, 20 g of peptone, and 10 g of yeast extract in 1 L of H_2O. In case of SD medium was prepared using: yeast nitrogen base without amino acids, 6.7g; 2% glucose;

and any of the amino acids if necessary: histidine, 20 mg; leucine, 120 mg; lysine, 60 mg; arginine, 20 mg; tryptophan, 20 mg; tyrosine, 20 mg; threonine, 40 mg; phenylalanine, 50 mg; methionine, 20 mg; uracil, 20 mg; and adenine, 20 mg in 1 Liter of H_2O; supplemented with 1.5% agar to make YPD or SD agar plates.

The strains used in this study are listed in Table I. CP2137, a Δ*comA* derivative of pneumococcal Rx1 strain was used as the WT standard. The primers are listed in Table II. Plasmids used in this study are listed in Table III. Antibiotics concentration used in this study are listed in Table IV.

2.2 DNA preparation

Genomic DNA of pneumococcus was extracted by chloroform and ethanol precipitation as described previously. A small volume of frozen stock cells (OD_{550} 0.2) was inoculated with 10 mL CAT, 0.016 M phosphate, and 0.2% glucose. Cells were grown until OD_{550} 0.2 at 37°C. Cultures were then chilled on ice and then centrifuged at 10,000 g for 15 minutes at 4 °C. The supernatant was discarded, and cells were pelleted. Pelleted cells were then resuspended in 1/10 volume of buffer I (0.05 M Tris-HCl pH 7.6; 10 mM EDTA, 0.1 M NaCl). The centrifugation was repeated, and cell pellets were resuspended in buffer II (buffer I, 0.5% Triton X-100, 0.2% SDS) and incubated at 37°C for 10 minutes before relocating to 65°C for 20 minutes. The lysate was incubated at -20°C overnight, and following day, the same volume of chloroform was added. The solution was then mixed vigorously through shaking and followed by centrifugation at 10,000 g for 30 minutes at -4°C. Then the supernatant was collected, and 1/10 volume of 3M sodium acetate with RNaseA was added to the resuspension mix. Then 2.5 folds of ice-cold ethanol were added to precipitate DNA. This solution was centrifuged under the same previously mentioned conditions and was left to let the DNA pellet to air dry.

Table I. Bacterial strains used in this study

Strain	Description	Source
S. pneumoniae		
CP1344	CP1250, but Δ*clpC*::NPctet^R	P. Luo
CP1759	CP1250, but SP18::PcSpc^R	CP1250×aPL06
CP1894	CPM7 X CP1389 (*dprA*::kan^R, *ssbB::lacZ,* Cm^R, kan^R)	A. Piotrowski
CP1896	CP 1250, aga::*comX*::*comW*::PcKan, Kan^R	A. Piotrowski
CP2055	CP1334, but Δ(spr1630spr1631)::kan^R::Cheshire; Sm^R Em^R	Weng, 2009
CP2137	*hex malM511 str-1 bgl-1*; Δ*comA*, Δ*cps*, Sm^r Cm^r LacZ Δ*ssbB*::pEVP3::*ssbB*; *SsbB*+	Weng, 2010
CP2158	CP2137, but Δ*clpC*::P$_{aga}$*::comX::comW::*kan^R	This study
CP2159	CP2137, but *clpC*$^+$; *aga::comX::comW::kan^R;* Δ*comW::*spc^R	This study
CP2165	*P$_{aga}$::comX::comW*, Δ*clpC* ::tet^R	This study
CP2166	CP2165, Δ*comW*::spc^R	This study
CP2167	CP2137, but Δ*comW*::*spc^R*	This study
CP2168	CP2137, P$_{aga}$*::comW::kan^R*, Δ*comW::P$_{fsck}$::clpC::spc^R*	This study
CP2173	CP2137,*P$_{aga}$::comX::comW*, Δ*clpC*::tet^R; Δ*comW*::P$_{fcsk}$::*clpC*::*spc^R*	This study
CP2180	CP2137, Δ*Hpf* (*yfiA*)::Pctet^R	This study
CP2181	CP2137, Δ*yhaM* (*cbf1*)::Pckan^R	This study
CP2182	CP2137, Δ*Hpf* (*yfiA*)::Pctet^R; Δ*yhaM* (*cbf1*)::Pckan^R	This study
CP2184	CP2137, rif^R, *rpoB* point mutation, serine$_{408}$ to proline$_{408}$	This study
CP2190	CP2137, Δ*metE*::Pckan^R	This study
CP2191	CP2137, Δ*metEF*::Pckan^R	This study
E. coli		
DH5α	F– Φ80*lac*Z, ΔM15, Δ(*lac*ZYA-*arg*F), *end*A1, *rec*A1, *hsd*R17 (rK–, mK+), *pho*A, *sup*E44 λ– *thi*-1, *rel*A1, *gyr*A96	Invitrogen
S. cerevisiae		
GlodY2H	MATa, *leu2-,3 trp1-901, 112, Δgal4, Δgal80, Mel1*	Clontech
Y187	MATα, *trp1-901, leu2-3, 112, Δgal4, Δgal80, Mel1*	Clontech

The dried DNA was resuspended in 5 mL TE buffer (10 mM Tris-HCl, pH 7.6, 1 mM EDTA).

PCR amplification was performed in a 50 µL reaction solution: 1 µL Phire HotStart II polymerase, 100 ng template DNA, 10mM dNTP, 25 µM primers, and dH_2O. The PCR conditions are as follows: 95°C for 3 minutes, then denaturing at 95°C for 30 seconds followed by a variable annealing time, then elongation at 72°C for 5 seconds and were repeated for a total of 30 cycles before a final extension at 72°C for 10 minutes. The annealing time can vary based on primers used in the experiment.

Colony PCR was performed by using one freshly grown single clone which was resuspended in 10µL dH_2O as template DNA. PCR amplification was performed in a 50µL reaction solution: 1 µL Phire HotStart II polymerase, 1µL template DNA, 10mM dNTP, and 25 µM primers. The PCR reactions were: 98°C for 10 minutes; 30 cycles of 98°C for 30 seconds, an annealing time of 5 seconds, 72°C for 1 minute/k bp product; 72°C for 10 minutes. The annealing time can vary based on primers used in the experiment.

2.3 Transformation

Transformation of pneumococcus was done by growing 10 mL culture in CAT with glucose and phosphate to OD_{550} 0.05-0.2 depending experimental designs. Then induced to competence with 0.5 mM of $CaCl_2$, 0.04% of BSA and 250 ng/mL of CSP (NeoBioSci, Cambridge, MA) (Pozzi et al., 1996). CSP is a synthetic heptadecapeptide with a sequence composition of EMRLSKFFRDFILQRKK and stored at −20°C. Then 10-100 ng donor DNA was added into the induced culture and incubated at 37°C for 60 minutes. After this the culture was diluted with 1.5 mL CAT and 10 µg/mL DNase I, and then the diluted culture was plated in CAT agar with selective antibiotics as follows: 1.5 mL diluted culture was mixed with 1.5 mL warm CAT agar (55°C) and poured over 3 mL of pre-casted CAT agar in a small petri dish. After the agar solidified, the third

layer of CAT agar was added. Lastly, CAT agar with antibiotics was added, and plates were incubated at 37°C. Transformation efficiency was determined by the ratio of numbers of antibiotics resistant colonies over counts of viable colonies. The average of three parallel counts was used to represent transformation efficiency.

Transformation with raffinose induction was performed using 10 mL culture grown in CAT with glucose and phosphate to OD_{550} 0.05-0.2 depending on the experimental designs at 37°C. The culture was then transferred to 30°C. Raffinose was added to the culture at a final concentration of 0.1%. The culture was left to transform with the 10-100 ng donor DNA in solution at 30°C for 60 minutes. Afterwards, the cells were diluted and plated as described above.

Transformation with fucose induction was performed with 1% fucose at 37°C. The following process was followed as described in the raffinose induction.

2.4 Competence kinetics

Competence kinetics was performed by using freshly grown culture in CAT with glucose and phosphate at OD_{550nm} 0.05 before splitting the volume of the solution in half. One half was induced by the addition of CSP and the other half without CSP. The time of CSP addition was recorded as the 0-minute mark, and small amounts of samples (150 µL-1 mL) were taken out every 5 minutes and exposed to an antibiotics resistance donor DNA for 3 minutes at 37°C.Then the culture was incubated with DNase I (5 mg/mL in 0.15 M NaCl stock) at a final concentration of 10 µg/mL for 60 minutes. Cultures were chilled at room temperature, and then diluted and plated. Antibiotics resistance transformants were counted after 16-20 hours of incubation at 37°C. Competence kinetics was plotted and recorded as the number of transformants over the time post induction.

Table II. Primers used in this study

Primer	Sequence 5'-3'	Location
DAM1089	TCAACTGTAGTGGGTTGAAG	*transposaseB*
DAM1090	GAAACCGATACCGTTTACGA-CTAGTACCAGCCTCAACAAG	*transposaseB*
DAM1091	CTTGTTGAGGCTGGTACTAG-TCGTAAACGGTATCGGTTTC	*cheshire*
DAM1092	CGTAGCATGTTTGCTCGAA-GGTCAGTAAATTGGACATAG	*cheshire*
DAM1093	CTATGTCCAATTTACTGACC-TTCGAGCAAACATGCTACG	*clpC*
DAM1094	CACTAGTTCTAGTACCAGCC-GCTAGCTGCCTAGTTTGC	*clpC*
DAM1095	GCAAACTAGGCAGCTAGC-GGCTGGTACTAGAACTAGTG	*purA*
DAM1096	TACATGGCAGGACTCTCAC	*purA*
DAM1099	TGAACGCTCTCCTGAGTAG	*Pfcsk*
DAM1100	TAGACATGGCAATCAACAAG	*Pfcsk*
DAM1101	CTTCCATAGCCTATCTAGTG	*purA*
DAM1102	TCCTGATCCAAACATGTAAG	*purA*
FP-r1	AAGGCGGAAGAAGGAGACAGAC	*rpoB*
RP-r1	CCACAGTTTCAACGGCAGTATC	*rpoB*
FP-r2	TCGACCGTACTCGTAAGATTCC	*rpoB*
RP-r2	ACCTCTGTAATAACCGCTTCCC	*rpoB*
JB17	GTACGGATCCAGGGGAAAATTATGATTAAAGAATTGTAT	*comX*
JB18	GCATCTCGAGAAAGAGATAATAATCATCTAGCCAGAGAC	*comX*
JB19	GTACGGATCCCTGGCTAGATGATTATTATCTCTTTC	*comX*
JB20	GCATCTCGAGCTAATGGGTACGGATAGTAAACTC	*comX*
JB21	GCATCTCGAGCCCTCACTTATACGATGACTGATCTCAC	*comX*
JB22	GTACGGATCCTGAAACACTAAGAGATTATAGAAAC	*comX*
JB37	GTCAGGATCC-TTATGTTACAAAAAATTTATGAGCAGATG	*comW*
JB38	GCATGTCGACTACTAAAATTACCTCAACAAGAAATAAAC	*comW*
JB41	GTCAGGATCCTTATGTTACAAAAAATTTATGAGCAGATG	*comW*
JB42	GCATCTCGAGTACTAAAATTACCTCAACAAGAAATAAAC	*comW*
JB45	GTACGGATCCAGGGGAAAATTATGATTAAAGAATTGTAT	*comX*
JB46	GCATGTCGACAAAGAGATAATAATCATCTAGCCAGAGAC	*comX*
JB47	GTACGGATCCCTGGCTAGATGATTATTATCTCTTTC	*comX*
JB50	GTACGGATCCTGAAACACTAAGAGATTATAGAAAC	*comX*
JB51	GTACGGATCCCAGCCCTAGAAGAATTGGAACG	*comX*
JB52	GCATCTCGAGTCAATTTGCTCTTCTGTATAAG	*rpoD*
JB53	GCATCTCGAGTACTGGTTCTTGGGCAATCT	*rpoD*
JB54	GTACGGATCC-AAGAACCAGTATCTCTTGAAAC	*rpoD*
JB70	GTACGGATCCAATGAATATTACCTTCATTTATGGG	*comX*
JB72	GCATGTCGAC-TCAATTTGCTCTTCTGTATAAG	*rpoD*
JB73	GCATGTCGAC-TACTGGTTCTTGGGCAATCT	*rpoD*
JB91	AACAGCGACTTCGACCAATGAG	*yfiA*
JB92	TAAGCAGTCGACCATATGAGTACCTTCTTTCTAAAC	*yfiA*
JB93	TAAGCAGCGGCCGCAATGTGATTTATCGTCGTGAG	*yfiA*

JB94	AGAGCCCAAAGCATCCATGTC	*yfiA*
JB95	TCTCCAAGCCTTGCAGGAATC	*cbf1*
JB96	TAAGCAGTCGACCATCAATCTCGCCACTATCATC	*cbf1*
JB97	TAAGCAGCGGCCGCCGTTCCTTCTATAAACCAGATTTAG	*cbf1*
JB98	TTCTTGGGCATTGTCCGCAAAGACC	*cbf1*
JB99	TAAGCAGTCGAC-CCGGGCCCAAAATTTGTTTGAT	*tetR*
JB100	TAAGCAGCGGCCGC-CCAAAGTTGATCCCTTAACG	*tetR*
JB112	TAAGCAGTCGAC-GGGATCCGTTTGATTTTTAATG	*kanR*
JB113	TAAGCAGCGGCCGC-ACAGTTGCGGATGTACTTCAG	*kanR*
JB139	TAAGCAGTCGAC-GGATCCCCCGTTTGATTTTTAATG	*spcR*
JB140	TAAGCAGCGGCCGC-CAAAAAAGCGGTTAGCTCCTTC	*spcR*
JB163	TGATTCAAGCGCCGACAGGG	*novR*
JB164	TGTGTCAGGCGCTCACTAAC	*novR*
JB165	AGCAGCCATAGTTGCAGTAG	*metE*
JB166	TAAGCAGCGGCCGC-ATTAGGGTTTGATATAGTTTG	*metE*
JB169	TAAGCAGTCGAC-AAACCATTCTTCTCAGGTGAGGG	*metEF*
JB170	CAGCAATTGGTCCGTCAAATGG	*metEF*

2.5 SDS-PAGE

SDS-PAGE was performed as described previously using the Bio-Rad Mini- and Maxi-protein gel apparatus. Each gel has 15-wells (0.75mm or 1.5 mm) and is discontinuous gel composed of 10%-15% separation gel and 4.5% stacking gel. SDS-PAGE was usually casted the day before use, and with details discussed later. Protein samples were prepared by mixing with Laemmili buffer (100 mM Tris-HCl (pH 6.8), 20% glycerol, 0.2% bromophenol blue, 4% SDS, and 200 mM dithiothreitol) and heated at 95ºC for 10 minutes. The gel was run at 100V until the dye reached the borderline of the two gels and then ran at 200V until the bottom of the separation gel. Protein samples were stained with Coomassie brilliant blue following standard procedures as discussed in later sections.

2.6 Bradford Assay

Protein concentration was determined by Bio-Rad protein Assay (Bradford, 1976; Luo & Morrison, 2003). BSA was used as a standard to set up the formula. Protein concentration ranged from 0.2 to 0.9 mg/mL was measured, otherwise, protein samples were diluted. 0.9 mL of dye was added to 0.1 mL samples and was incubated for 5 minutes before measurement of absorbance at OD_{595}.

Table III. Pladmids used in this study

Plasmid	Annotation	Description
pAD*	pAD*	pGADT, HA-Tag, Amp^R, -*Leu2*
pBD*	pBD*	pGBKT, c-Myc epitope tag, Kan^R, -*Trp1*
pAD-*SigX*$_{160}$	pMJB01	pGADT-*SigX*$_{160}$, HA-Tag, Amp^R, -*Leu2*
pAD-*SigX*$_{N100}$	pMJB02	pGADT-*SigX*$_{N100}$, HA-Tag, Amp^R, -*Leu2*
pAD-*SigX*$_{C60}$	pMJB03	pGADT-*SigX*$_{C60}$, HA-Tag, Amp^R, -*Leu2*
pBD-*comW*	pMJB04	pGBKT-*comW*, c-Myc epitope tag, Kan^R, -*Trp1*
pBD-*SigX*$_{160}$	pMJB05	pGBKT-*SigX*$_{160}$, c-Myc epitope tag, Kan^R, -*Trp1*
pBD-*SigX*$_{N100}$	pMJB06	pGBKT-*SigX*$_{N100}$, c-Myc epitope tag, Kan^R, -*Trp1*
pBD-*SigX*$_{C60}$	pMJB07	pGBKT-*SigX*$_{C60}$, c-Myc epitope tag, Kan^R, -*Trp1*
pAD-*comW*	pMJB08	pGADT-*comW*, HA-Tag, Amp^R, -*Leu2*
pAD-*sigA*$_{369}$	pMJB09	pGADT-*sigA*$_{369}$, HA-Tag, Amp^R, -*Leu2*
pAD-*sigA*$_{N286}$	pMJB10	pGADT-*sigA*$_{N286}$, HA-Tag, Amp^R, -*Leu2*
pAD-*sigA*$_{C83}$	pMJB11	pGADT-*sigA*$_{C83}$, HA-Tag, Amp^R, -*Leu2*
pBD-*sigA*$_{369}$	pMJB12	pGBKT-*sigA*$_{369}$, c-Myc epitope tag, Kan^R, -*Trp1*
pBD-*sigA*$_{N286}$	pMJB13	pGBKT-*sigA*$_{N286}$, c-Myc epitope tag, Kan^R, -*Trp1*
pBD-*sigA*$_{C83}$	pMJB14	pGBKT-*sigA*$_{C83}$, c-Myc epitope tag, Kan^R, -*Trp1*
pAD-*sigX*$_{R2-4}$	pMJB15	pGADT-*sigX*$_{R2-4}$, HA-Tag, Amp^R, -*Leu2*
pAD-*sigX*$_{R2}$	pMJB16	pGADT-*sigX*$_{R2}$, HA-Tag, Amp^R, -*Leu2*
pAD-T*	pAD-T*	pGADT-antigen T, HA-Tag, Amp^R, -*Leu2*
pBD-53*	pBD-53*	pGBKT-murine p53, HA-Tag, kan^R, -*Trp1*
pBD-lam*	pBD-lam*	pGBKT-lamin, HA-Tag, kan^R, -*Trp1*

* indicate plasmids from Clontech, Mountain View, CA

Table IV. Antibiotic concentration list

Antibiotic	Abbreviation	Mechanism	Stock mg/mL	Final μg/mL
Erythromycin	Erm^R	Binds to 50S ribosomal subunits, inhibiting translation	20	0.25
Kanamycin	Kan^R	Binds to 30S ribosomal subunits, inhibiting translation	50	200
Novobiocin	Nov^R	*gyrB* point mutation, targeting DNA gyrase	100	2.5
Spectinomycin	Spc^R	Inhibits translocation of tRNA, inhibiting translation	100	100
Tetracycline	Tet^R	Prevents codon-anticodon interactions, inhibiting translation	15	0.25
Rifampicin	Rif^R	*rpoB* point mutation, binds to RNAP, inhibiting RNA synthesis	50	10

III Interaction of SigX and ComW in Competence Development in *Streptococcus pneumoniae*

Yanina Tovpeko, Junqin Bai and Donald A. Morrison

(Paper: Journal of Bacteriology. 2016. vol.198 no.17 2370-2378)

3.1 Abstract

In *Streptococcus pneumoniae,* competence is regulated by two sets of competence responsive genes after exposure to CSP (Competence Stimulating Peptides): early and late genes, defined by maximal expression between 7.5 to 10 minutes and 12.5 to 15 minutes post CSP exposure. The expression of early genes under ComE-P (phosphorylated ComE) is required for development of competence, while late genes under σ^X (ComX) regulation are required for DNA uptake and recombination. Two early competence proteins σ^X and ComW play critical roles during competence development. The expression of both proteins is transient and regulated by the same pheromone circuit and TCSTS. σ^A is the primary sigma factor while σ^X is the only alternative σ factor identified. σ^X directs transcription of competence specific late genes by binding to RNA polymerase (RNAP). Full σ^X activity in transcribing late competence genes requires ComW, a protein that is conserved in eleven species of the mitis and anginous groups. Its absence causes a 90% decrease of late gene expression and a 10^4-fold decrease in the amounts of transformants. But its exact role in regulating competence is unknown. ComW may interact with σ^X directly to facilitate the access of σ^X to RNAP. In addition, transcription of competence specific genes can be activated by ComW, serving as a transcription factor that binds directly to DNA. Alternatively, ComW may indirectly promote σ^X for access to RNAP by antagonizing σ^A.

To gain insight into the molecular role of ComW, yeast two-hybrid assay was used to investigate whether ComW binds to σ^X directly or antagonizes σ^A. Gal4 transcription Activation

Domain (AD) fusion constructs, full-length $SigX_{160}$ and truncated $SigX_{160}$ were generated. Gal4 Binding Domain (BD) fusion constructs. BD-ComW and AD-σ^A were also generated. Potential interactions between the two fusion proteins will bring the AD and BD close enough to turn on reporter genes in yeast. Diploids containing plasmids with fusions of $SigX_{160}$ and ComW did turn on expression of several reporter genes. Truncated σ^X showed no interaction with ComW. This suggests that a complete $SigX_{160}$ is required for interacting with ComW. σ^A didn't interact with ComW. Interestingly, Y. Tovpeko identified σ^A mutations that bypass the ComW requirement for competence. The Y2H results indicated that ComW interacts directly with σ^X but not σ^A. Taken together with σ^A bypass mutations, we were led to the hypothesis that ComW acts to promote σ^X assembly into RNAP rather than antagonizing σ^A. In the future, *in vitro* transcription can be used to determine a direct promotion of ComW in σ^X activity. Using a pro-σ factor to regulate access of an alternative σ factor to RNAP would be a novel mechanism of gene transcription in gram-positive bacteria.

3.2 Introduction

Bacteria use various mechanisms to direct RNAP transcription of specific genes in response to growth signals from the environment. Many mechanisms of transcription activation have been characterized. For example, anti-σ^{70} factor Rsd in *Escherichia coli* can bind to σ^{70} and sequester its activity, leading to enhanced σ^S capacity in competition over σ^{70} for access to core polymerase (RNAP) (Campbell et al., 2003). Another protein, Crl, activates σ^S-dependent gene transcription directly by promoting σ^S-holoenzyme formation (Banta et al., 2014; Pratt & Silhavy, 1998). Transcription activation can be enhanced by transcription factors binding directly to a promoter sequence. The cI protein of bacteriophage (λcI) activates transcription by direct binding to the promoter and σ^{70}, resulting in formation of a complex (Nickels et al., 2002).

Competence in *S. pneumoniae* is regulated by a QS mechanism and TCSTS. σ^X is the only σ factor known to initiate competence-specific gene expression (Luo et al., 2004). A small (10KDa) protein, ComW, is required for σ^X stability and activity. A Δ*comW* mutant exhibits an approximately 10^4-fold decrease in transformants compared to a wild type strain (Bartilson et al., 2001; Luo et al., 2004). Previous studies in our laboratory have shown that even though large amounts of σ^X were accumulated in a Δ*clpP*Δ*comW* double mutant strain, the level of competence was significantly reduced (Sung & Morrison, 2005). However, competence was restored by 80% by re-introducing *comW* to the chromosomal locus (Luo et al., 2004). The expression of three σ^X-dependent genes (*ssbB, celB,* and *cglA*) was also significantly reduced in a Δ*comW* mutant—a reduction of at least 90% compared to that of the same late genes in a wild type. In fact, competence was consistently decreased by 99.9% in these strains (Luo et al., 2004). Both the N- and C-termini of ComW have been proven important for σ^X protein accumulation, but the N-terminus is more important for σ^X activity. Specifically, competence was only 0.1% of a wild type in the strain expressing a N-terminus His-tagged-ComW, a level as low as in a Δ*comW* mutant. However, the level of competence of the strain expressing a C-terminus His-tagged ComW was 76% of a wild type (Sung & Morrison, 2005).

ComW and σ^X proteins have shown similar expression patterns during competence development: appearing at 5 minutes, peaking at 20 minutes, and disappearing by 40 minutes (Sung & Morrison, 2005). This raises the question of whether ComW and σ^X interact directly. Yeast two-hybrid (Y2H) and affinity purification assays have been performed to ask whether there was a direct interaction between ComW and σ^X. However, the two assays yielded inconsistent results. The Y2H assay showed a direct interaction between σ^X and ComW, while the affinity purification assay showed no such interaction (Piotrowski, 2009). In the Y2H assay, a strong

interaction between ComW and σ^X was reported in a strong selective medium (Weng, 2011). However, this Y2H result was not validated by additional methods. In the affinity purification assay, His-σ^X and His-ComW were affinity-purified from competent cells 15 minutes after CSP induction and were detected by the anti-His antibody. The co-purification assay with His-σ^X and ComW-His showed that the two proteins were not co-purified (Piotrowski et al., 2009). However, the His-tag may have blocked any potential interaction between termini of the two proteins. Whether there is a direct interaction between ComW and σ factors remains unclear.

Recently, a Δ*comW* suppressor screen revealed that nonsynonymous amino acid substitutions of σ^A partially restored competence in a Δ*comW* deletion mutant (Tovpeko & Morrison, 2014), which suggests that σ^A mutation avoided the requirement for ComW during competence. Those mutations were mainly located in regions 2 and 4 of σ^A, which might have affected the affinity of σ^A to RNAP. These findings lead to the hypothesis that ComW promotes σ^X access to RNAP in competition with σ^A and that this competition can be enhanced the weaker affinity of σ^A in gaining access to RNAP. To test this hypothesis, an improvedY2H assay was applied to determine whether these proteins would interact directly.

3.3 Methods

3.3.1 Yeast two-hybrid assay

Recombinant plasmids were transformed into competent yeast of *MATα* or *MATa* mating type by the lithium acetate (LiAc)-mediated method (Gietz & Woods, 2002), with the transformed haploid cells selected on SD-L or SD-T plates at 30°C for 72 hours. Plasmids from single clones were extracted and verified by sequencing plasmid junctions and stored at -80°C with 15% glycerol. Several successfully transformed bait and prey colonies were inoculated into YPDA

medium and grown overnight. By cross-mating between a and α mating types of yeast haploids, diploid strains were obtained on SD-LT. Five diploid transformants were picked up and saved from each cross-mating. Small identical volumes of five individual frozen strains were mixed by resuspension with SD and diluted to inoculate about 10^5 cells onto patches of SD-LT agar with 5-bromo-4-chloro-3-indolyl-D-galactopyranoside (X-α-gal). Diploids on SD-LT-X-α-gal were incubated at 30°C for 5 days before determining interacting partners by the production of blue product.

The Golden yeast two-hybrid system has four independent reporter genes (*aur1-C, ade2, his3,* and *mel1*) to provide a more stringent selection condition. The four reporter genes contain three complete heterologous three completely heterologous Gal4-responsive promoter elements: G1, G2, and M1. Yeast colonies that express *mel1* encoding a-galactosidase turn blue in the presence of the chromogenic substrate X-α-gal. *Aur1-C* encodes the inositol phosphoryl ceramide synthase, which confers aureobasidin A resistance (Ab^R). The other two auxotrophic selection markers, *ade2* and *his3,* are used to reduce false positives during the screening. Diploids that grow on selective media indicate the presence of a direct binary interaction. A strong interaction between a prey and a bait will activate all four reporter genes, while a weak interaction can activate only one or two of these genes. Autoactivation leading to false positive can happen if bait protein has transcription activation activity or interacts nonspecifically with the Gal4 activaiton domain.

3.3.2 Quantitative α-galactosidase assay

To quantify the strength between interacting proteins, α-galactosidase activity was measured in each diploid. Because α-galactosidase is a secreted enzyme and it can be assessed directly from the supernatant of the culture after centrifugation (Aho & Uitto, 1997; Goddard & Davey, 2005). For each interaction pair, five single colonies were picked up from the selection

plates, mixed in 3 mL of a defined medium (SD) and grew overnight. When the OD_{595} reached 1, cells were pelleted at 13,000 rpm for 1 minute, and the supernatant containing released α-galactosidase was collected and transferred to a transparent bottom 96-well microplate. Three volumes (3V) of assay buffer (0.33 M sodium acetate pH 4.5; 33 mM p-nitrophenyl-α-D-galactopyranoside) were incubated with one volume (1V) of the supernatant at 29°C, and the reactions were stopped at different time points by the addition of one volume (1V) 2M Na_2CO_3. The α-galactosidase in the supernatant turned the colorless chromogenic substrate p-nitrophenyl-α-D-galactopyranoside into p-nitrophenol, which displayed a strong OD reading at 405 nm. This reaction can be represented as follows:

$$\text{PNP-α-gal} + H_2O \xrightarrow{\text{α-galactosidase}} \text{p-nitrophenol} + \text{D-galactose}$$

The OD_{405} value was proportional to the moles of p-nitrophenol liberated. The OD_{405} value was recorded at different time points using a scanner and then used to calculate the concentration of α-galactosidase in Miller units. A blank containing no enzyme was used to calibrate the spectrophotometer for OD_{405} readings. For α-galactosidase kinetics, the ratio between OD_{405} and OD_{595} was plotted over time and was used to represent the hydrolysis rate of α-galactosidase. Following measurement of the color developments during hydrolysis of PNP-α-gal, signals were recorded as an indication of the *mel1* expression activated by interacting pairs of proteins.

One unit of α-galactosidase activity was defined as 1 μmol of PNP-α-gal converted into p-nitrophenol and D-galactose in assay buffer per minute at pH 4.5 at 29°C. In the following formula, the OD_{595} of the culture is used to normalize the OD_{405} to the amounts of cells.

$$\text{Miller Units/mL α-galactosidase} = OD_{405} \times V_f / OD_{595} \times 18.5 \times 60(\text{min}) \times V_s$$

V_f: final volume of the assay (200 μL)

V_S: volume of supernatant added (16 µL)

18.5: the light path (1 cm) x 18.5 mL/µmol for the formula.

3.4 Results

3.4.1. Cloning variant σ^X and ComW into shuttle plasmids

In the Y2H system, two domains of the split Gal4 transcription factor: the activating domain (Gal4-AD) and binding domain (Gal4-BD), are unable to interact with each other and activate gene expression alone. BD binds to the upstream sequence, and AD recruits RNAP and activates transcription of a reporter gene. However, if the two domains can be reconstituted, a functional transcription factor will activate expressions of reporter genes by direct binding to their promoters (Fields & Song, 1989; Gross et al., 1998; Keegan & Ptashne, 1986). Sequences encoding these two domains are in two different vector plasmids, the GAL4-pGBKT7 (pBD) and GAL4-pGADT7 (pAD) vectors. A bait gene is cloned into the GAL4-BD with *TRP1* marker, while a prey gene is cloned into the GAL4-AD with *LEU2* marker. *TRP1* and *LEU2* encode genes to synthesize tryptophan and leucine respectively. They will complement the mutations of *trp1* and *leu2* in the host strains.

To determine whether ComW functions to promote σ^X access to RNA polymerase by binding σ^X directly as well as σ^X domains important for the interaction, a well-designed Y2H assay was used. Shuttle plasmids pAD and pBD, along with other plasmids used as controls were obtained from Clontech. The recombinant plasmids used for the Y2H assay were constructed by restriction enzyme digestion and ligation of insert DNAs into shuttle vectors (see Figure 13). The regions of the *S. pneumoniae* R6 chromosome encoding *sigX* (*comX*) were amplified by PCR: full-length *sigX* ($sigX_{160}$), N-terminal *sigX* ($sigX_{N100}$), and C-terminal *sigX* ($sigX_{C60}$). Then pAD-

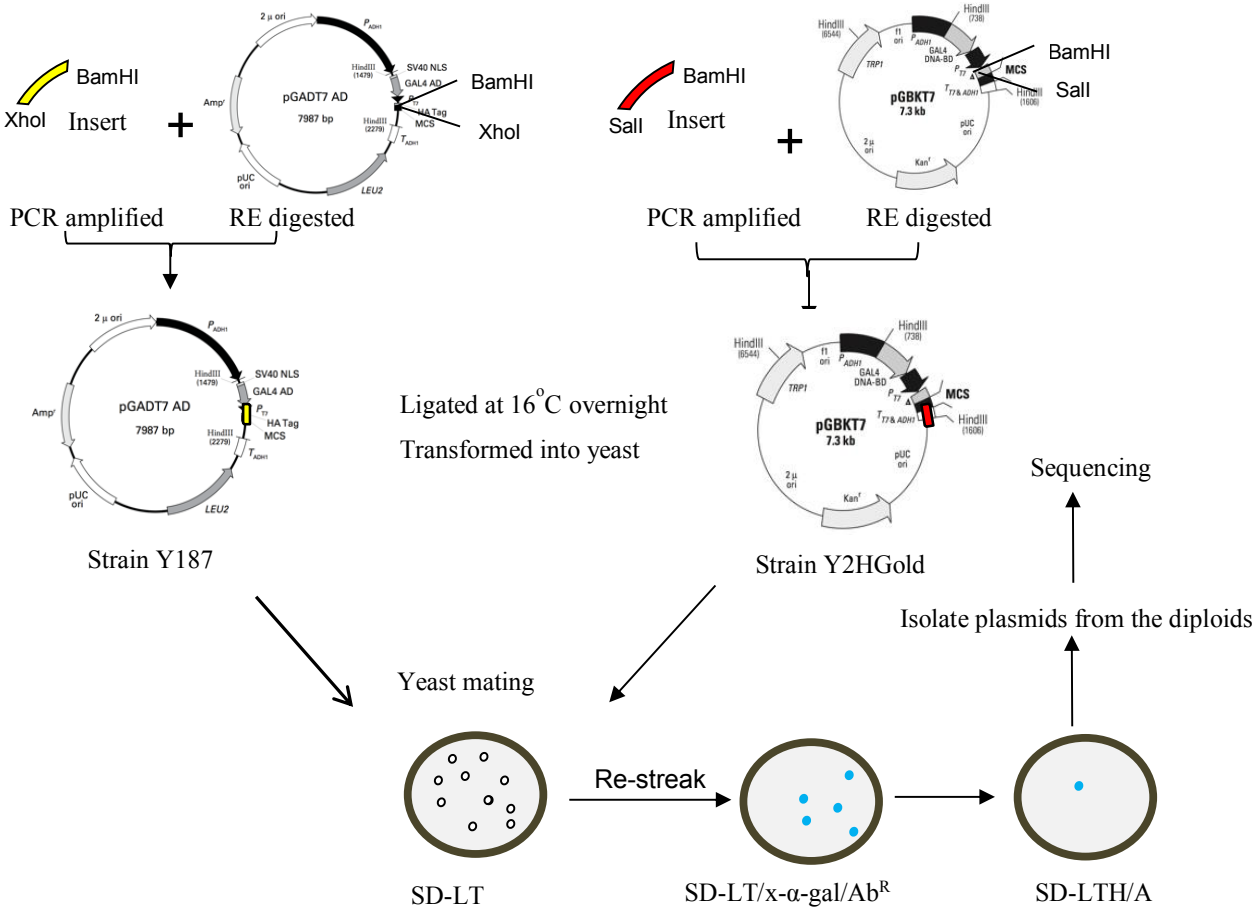


Figure 13. Experimental design of Y2H assay. Three versions of *comX* were cloned into shuttle plasmids pAD and *comW* was cloned into pBD individually in *E. coli*. Then the recombinant plasmids were transformed into yeast haploid a and α cells. Recombined plasmids pAD with prey were transformed into yeast strain Y187 (MATα), and recombined pBD with bait were transformed into yeast strain Y2HGold (MATa). After yeast mating, diploids were selected from SD-LT plates: AD-*sigX*$_{160}$ + BD-*comW,* AD-*sigX*$_{N100}$ + BD-*comW*, AD-*sigX*$_{C60}$ + BD-*comW,* ϕBD + AD-*sigX*$_{160}$, BD-*comW* + ϕAD, PC and NC. Colonies grown on SD-LT were streaked on SD-LT/x-α-gal/Ab^{R} plates to detect interacting prey and bait. Further interaction between prey and bait was determined on more stringent SD-LTH/A plates.

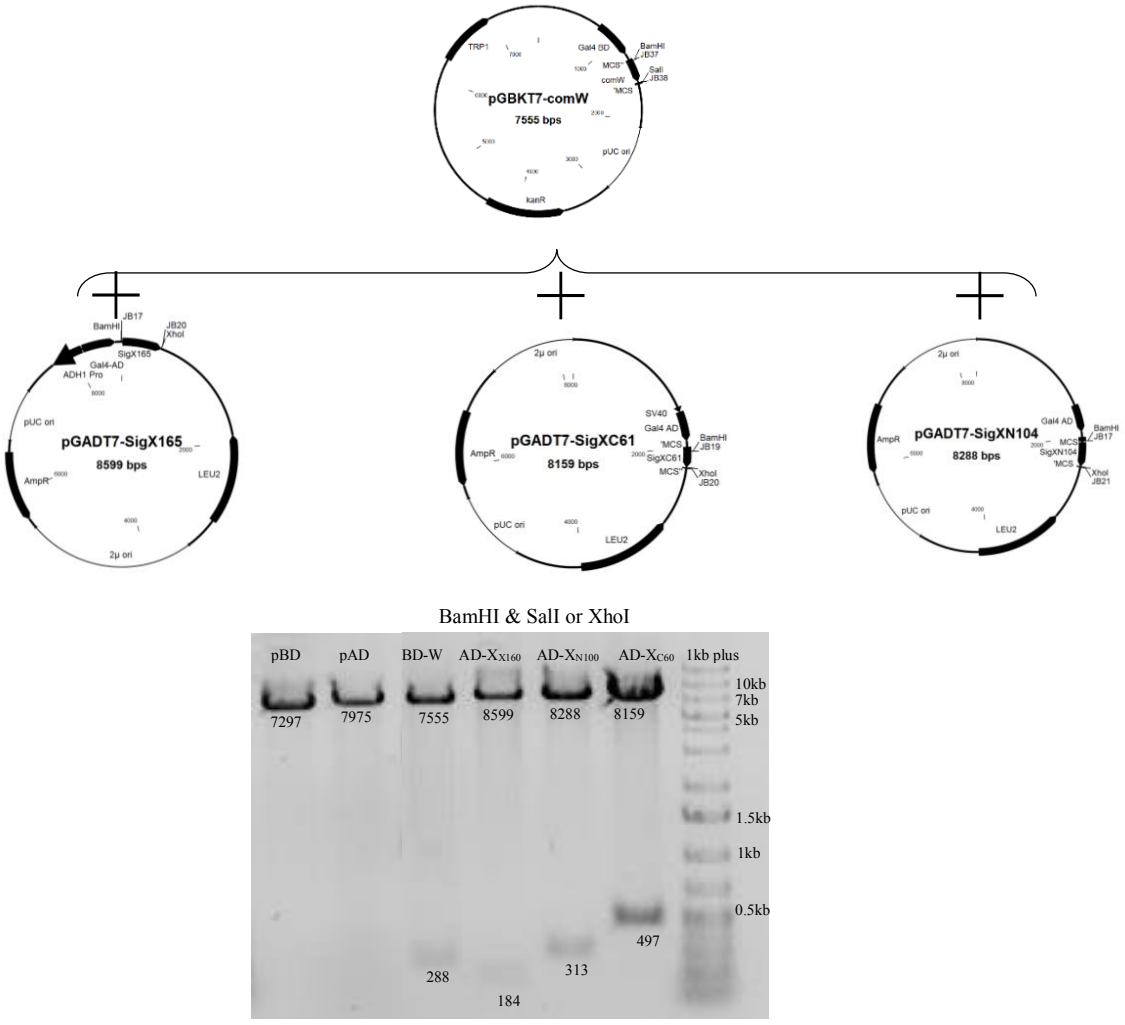


Figure 14. Cloning *sigX* and *comW* into shuttle plasmids pAD and pBD for Y2H assays. To construct recombinant plasmids, *comX* (*sigX*) was inserted into plasmid pAD at BamHI and XhoI sites, and comW was inserted into pBD at BamHI and SalI sites. Ligation between vectors and inserts were carried out. Resulting recombinant plasmids pAD-$SigX_{160}$, pAD-$SigX_{N100}$, pAD-$SigX_{C60}$ and pBD-*comW* were confirmed by digestion with BamHI and either SalI or XhoI. Linearized pAD and pBD are 7975 bp and 7297 bp, respectively. About 100 ng plasmid DNA was digested in each reaction and loaded on 1% agarose electrophoresis gel. $sigX_{160}$, 497 bp;, $sigX_{N100}$, 313 bp; $sigX_{C60}$, 184 bp, and *comW*, 288 bp.

SigX$_{160}$ (MJB01), pAD-*SigX*$_{N100}$ (pMJB02), and pAD-*SigX*$_{C60}$ (pMJB03) were created by ligating the three versions of the *sigX* amplicon into pAD after digestion of molecules by BamHI and XhoI (see Figure 14). Similarly, pBD-*comW* (pMJB04) was created by ligating *comW* into pBD after digestion of both molecules by BamHI and SalI. The primers used in this study are listed in Table II. Recombinant plasmids were first transformed into *E. coli* with selection on ampicillin or kanamycin plates. DNA from selected single colonies was purified and correct in-frame cloning was confirmed by sequencing the junction area of DNA using Sanger sequencing.

Another set of recombinant plasmids was constructed for vector swap Y2H assays. *SigX* and *comW* were cloned into pBD and pAD respectively using PCR followed by restriction enzyme digestion and ligation at BamHI and either XhoI or SalI sites. Recombinant plasmids were confirmed by sequencing the multiple cloning sites (MCS) and were named BD-*sigX*$_{N100}$ (pMJB05), BD–*sigX*$_{C60}$ (pMJB06), BD-*sigX*$_{160}$ (pMJB07), and AD-*comW* (pMJB08).

Truncated σ^X proteins were also examined in this assay. No structure is available for pneumococcal σ factors. However, by sequence comparisons to σ^A in *E. coli*, σ^X was divided into four regions: region 1 (1-17aa), region 2 (18-98aa), region 3 (99-112aa), and region 4 (113-159aa). Recombinant plasmids AD-*sigX*$_{R2\text{-}4}$ (pMJB15) and AD-*sigX*$_{R2}$ (pMJB16) were constructed to test the truncated σ^X in interaction with ComW.

The recombinant plasmids transformed into yeast haploid cells were selected on SD-L or SD-T plates and mated. By cross-mating different mating types of the four haploids, yeast diploid strains were obtained: pAD-*sigX*$_{160}$ + pBD-*comW* (pMJB01 & pMJB04), pAD-*sigX*$_{N100}$ + pBD-*comW* (pMJB02 & pMJB04), and pAD-*sigX*$_{C60}$ + pBD-*comW* (pMJB03 & pMJB04). For the domain swap, diploids were obtained: pAD-*comW* + pBD-*sigX*$_{N100}$ (pMJB08 & pMJB05), pAD-*comW* + pBD–*sigX*$_{C60}$ (pMJB08 & pMJB06), and pAD-*comW* + pBD-*sigX*$_{160}$ (pMJB08 &

pMJB07). For the truncated σ^X domains, diploids were obtained: pBD-*comW* + pAD-*sigX*$_{R2-4}$ (pMJB05 & pMJB15), pAD-*comW* + pAD–*sigX*$_{R2}$ (pMJB05 & pMJB16). Diploids with recombinant plasmids were verified by extracting pAD-prey and pBD-bait plasmids from outgrown yeast diploids, followed by restriction enzyme digestions, and Sanger sequencing of the inserts.

3.4.2 Interaction between full-length of σ^X and ComW

To determine fusion protein interactions in Y2H, diploids were selected on SD-LT were screened for activation of reporter genes by streaking onto four other selection plates: SD-LT/X-α-gal/Ab^R, triple (SD-LTA/X-α-gal/Ab^R or SD-LTH/X-α-gal/Ab^R), and quadruple (SD-LTHA/X-α-gal/Ab^R). When the fusion proteins interacted, a functional Gal4 transcription factor was reconstituted and followed by activation of expression of the Gal4-responsive gene *mel1*, which was monitored by means of formation of blue colonies on X-α-gal plates. Triple and quadruple selection media were used to select cells having a strong interaction. Empty vectors without inserts were used as controls to eliminate any prey or bait autoactivation.

The activation of two reporter genes *mel1* and *aur1-C* in diploids containing AD-*SigX*$_{160}$ and BD-*comW* was indicated by blue colonies on plates with X-α-gal and aureobasidin (50 ng/mL). This suggested a weak interaction between SigX$_{160}$ and ComW. Because the interaction was weak, expression of only two reporter genes, *mel1* and *aur1-C,* was activated. No activation of *his3* gene was supported by no growth on SD-LTH plates (see Figure 15). In addition, three versions of *sigX* and *comW* were cloned as fusion proteins with vector swap. In the Y2H assay with vector swap, the expression of recombinant proteins between BD-*SigX*$_{160}$ and AD-*comW* activated *mel1*, *aur-C*, and *his3* genes. The results confirmed the weak interaction between σ^X and ComW (see Figure 16).

In summary, ComW and $SigX_{160}$ interacted in both pairings pBD-*$SigX_{160}$* + pAD-*comW* and pBD-*com*W + pAD-*SigX160* indicated by blue colonies (see Figure 16). Diploids containing the following pairs of fusion proteins formed white colonies on X-α-gal indicator plates: ϕBD + pAD-*comW*, ϕBD + pAD-*$SigX_{160}$*, ϕAD + pBD-*$SigX_{160}$*, ϕBD + ϕAD, and pBD-*comW* + ϕAD. Control partners used in the assay were pBD-53 + pAD-T as positive controls (PC) and pBD-lam and pAD as negative controls (NC) (see Figure 15). Diploids containing P53 + T antigen have activated *mel* expression, resulting in colonies with a strong blue color on X-α-gal plates, while diploids containing lamin + T antigen could not activate *mel* expression indicated by white colonies. Interestingly, $SigX_{N100}$ and $SigX_{C60}$ could also weakly interact with ComW in this experiment, which was indicated by the growth of blue colonies on SD-LT/X-α-gal medium (see Figure 15).

Diploids containing recombinant plasmids BD-*comW* (pMJB04) and either AD-*$sigX_{R2}$* (pMJB16) or AD-*$sigX_{R2\text{-}4}$* (pMJB15) were not able to activate expression of any reporter gene, suggesting that no interaction exists between truncated σ^X and ComW. These results indicated that both the N- and C- termini of σ^X are required for this interaction to occur.

In addition, an independent Y2H assay conducted by Weng (2011) was performed. The plasmids used in his Y2H assay were reconstructed and bait and prey proteins were cloned into pGBDUC2 and pACT2, respectively, following his methods. The recombinant plasmid DNA was transformed into yeast haploids NSY752 (MAT*α*) and NSY468 (MAT*a*), respectively. The two transformed haploid cells with different mating types were selected in SD-L or SD-U media and mated to generate diploids in SD-LU agar plates. The diploids were tested on selective media: SD-LU, SD-LUH, SD-LUH+1mM 3AT, and SD-LUH+3mM 3AT. The four test plates were then incubated, and growth was recorded.

In the Y2H assay, *sigX*$_{160}$ and *comW* were respectively fused with the DNA binding domain and activation domain of the GAL4 transcription factor. Then the two fusion plasmids, pACT2-*sigX*$_{160}$ and pGBDUC2-*comW*, were transformed into yeast mating haploid cells. To determine protein interactions, diploids containing interacting fusion proteins were selected in media lacking a histidine supplement. A strong interaction between SigX$_{160}$ and ComW was detected in the medium SD-LUH+3AT.

3.4.3 Prey and bait interaction-induced differentiatial activation of reporter genes

Four reporter genes were used to detect possible interaction between prey and bait proteins. Because each of the four genes contained a unique Gal4-responsive promoter, it was possible to decrease potential false positive signals caused by activation of nonspecific promoters that occurs occasionally in Y2H screening. Growth and color indication on selection media were used to reveal differential activation of *Aur-C, Mel1, Ade2,* and *His3* reporter genes.

Activation of *Mel1* and *His3* expression by fusion expressions of *comW* & *SigX*$_{160}$, *comW* & *SigX*$_{N100}$, *comW* & *SigX*$_{C60,}$ and *SigX*$_{160}$ & *SigX*$_{160}$ was indicated by growth on SD-LTH and blue colonies on SD-LT-X-α-gal media. The growth and color on those media varied from clone to clone: after 7 days of incubation, some clones grew well but were pale blue (*comW* & *SigX*$_{N100}$ and *SigX*$_{160}$ & *SigX*$_{160}$), whereas another displayed slow growth but with deeper blue color (*SigX*$_{160}$ & *comW*). The presence of growth revealed an interaction between the pairs of proteins after 5 days of incubation at 30^{o}C, while the absence of growth on SD-LTH and pink colonies on SD-LT/X-α-gal indicated lack of interaction. Such differential activation of reporter genes made it difficult to evaluate the strength of each interaction.

3.4.4 Determination of interaction strength between σ^X and ComW

To determine the interaction strength between σ^X and ComW, the α-galactosidase activity was measured. The α-galactosidase activity, which is encoded by the *melI* reporter gene, has been used in the Y2H assay to determine whether an interaction exists between prey and bait (Wagemans & Lavigne, 2015). Colonies turn blue when an interaction occurs between prey and bait on X-α-gal plates. However, as a sole indicator, the color change is not very accurate for determining an interaction. Various degrees of blue can cause biased judgments about the result, especially because the color change is sensitive to incubation time. To further clarify the result, a quantitative liquid α-gal assay is more sensitive and can be used to determine the strength of the interactions identified by plate screening. The stronger the interaction between prey and bait, the more α-galactosidase produced.

The α-galactosidase is a secreted enzyme, and its activity from the supernatant of diploids cells grown in SD-LT was measured in the presence of PNP-α-gal. The biggest slope of the linear portion of the α-galactosidase activity over time directly represents the most robust kinetics of α-galactosidase contained in each sample.

The α-galactosidase activity was evaluated with the murine tumor suppressor P53 protein and SV40 T antigen, and lamin and T antigen. For each interaction protein pair, several colonies were picked up from SD-LT/X-α-gal/Ab^R selection plates and grew in the SD-LT medium.

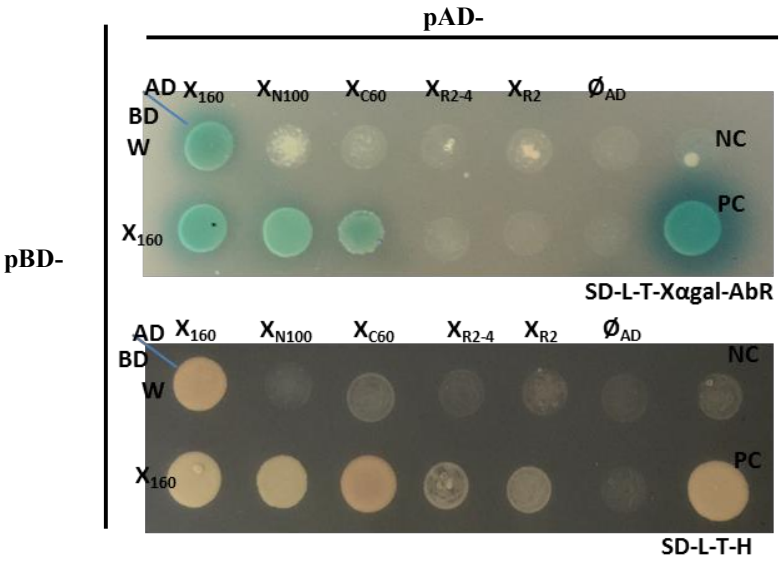


Figure 15. The interaction between SigX$_{160}$ and ComW evaluated by the Y2H assay. Diploids containing fusion proteins were plated on SD-LT-X-α-gal-Ab^{R} (Top) and SD-LTH (Bottom) plates. Diploids that activate α-galactosidase expression allow appearance as blue colonies on plates with X-α-gal. Plates were incubated at 30^{o}C for 5 days. Proteins fused to pAD are indicated on the top; proteins fused to pBD are indicated on the left. øAD, empty pAD vector. Growth on SD-LTH plates suggests activated his reporter gene. Yeast diploids with murine p53 protein and SV40 large T antigen showed strong blue color. Yeast diploids with lamin and SV40 large T antigen did not grow on both selection plates.

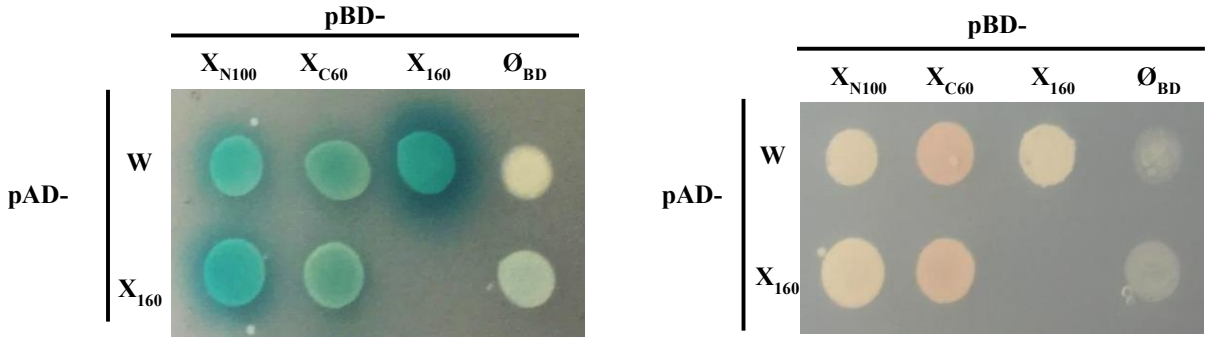


Figure 16. Y2H assays with vector swap to confirm the interactions. Diploids with divergent SigX and ComW were evaluated by the Y2H assay. (Left) Diploids selected from SD-LT were incubated in two selection media: SD-LT-X-α-gal-Ab^R and SD-LTH at 30°C for five days to allow color development and detectable growth. The interaction between $SigX_{160}$ and ComW was confirmed by strong blue spots on SD-LT-X-α-gal-Ab^R plates, and $SigX_{N100}$ and $SigX_{160}$ showed faint blue spots. (Right) Visible growth of colonies with all interacting protein pairs on SD-LTH plates.

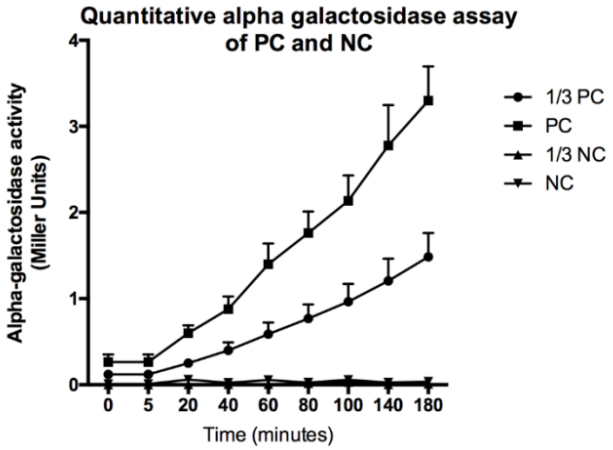
Cultures reached one absorbance unit (1A) was collected to perform the PNP-α-gal assay. Positive control samples (P53 and T antigen) were prepared when one absorbance unit as well as 1/3 absorbance unit of yeast cells, because we expect a strong α-galactosidase activity stimulated by the interaction between P53 and T antigen (see Figure 17). The relative α-galactosidase activity of testing proteins in PNP-α-gal assays was determined by comparisons to the α-galactosidase activity of 1/3 unit of P53 and T antigen. Based on the results of PNP-α-gal assay, the strength of interactions can be ranked as follows: p53 + T antigen>> BD-*SigX*$_{160}$ + AD-*W* > BD-*W* + AD-*SigX*$_{160}$ > lamin + T antigen (see Figure 17). Thus, the liquid α-galactosidase assay with PNP-α-gal was able to detect protein interactions and allowed a quantitative assessment of interaction strength.

3.4.5 Lack of interaction between σ^A and ComW

To determine whether an interaction between σ^A and ComW, Y2H was performed. *RpoD*, encoding σ^A, was cloned into pAD using PCR followed by restriction enzyme digestion and ligation at BamHI and XhoI sites (see Figure 18). The constructed plasmids were first transformed into *E. coli* and purified with a Zymo Miniprep kit. Sanger sequencing confirmed the construction of plasmids pAD-*sigA*$_{369}$ (pMJB09), pAD-*sigA*$_{N286}$ (pMJB10), pAD-*sigA*$_{C83}$ (pMJB11), and pBD-*comW* (pMJB04). The first three plasmids were then transformed into the yeast haploid strain Y2HGold (*MATa*) and the fourth into Y187 (*MATα*). The transformed haploid cells were selected in an SD-L or SD-T medium.

Similarly, for the domain swap assay, three recombinant plasmids were constructed: pBD-*sigA*$_{369}$ (pMJB12), pBD-*sigA*$_{N286}$ (pMJB13), and pBD-*sigA*$_{C83}$ (pMJB14). After cross-mating, three diploid strains were obtained: pAD-*comW* + pBD-*sigA*$_{369}$ (pMJB12 & pMJB08), pAD-

A.



B.

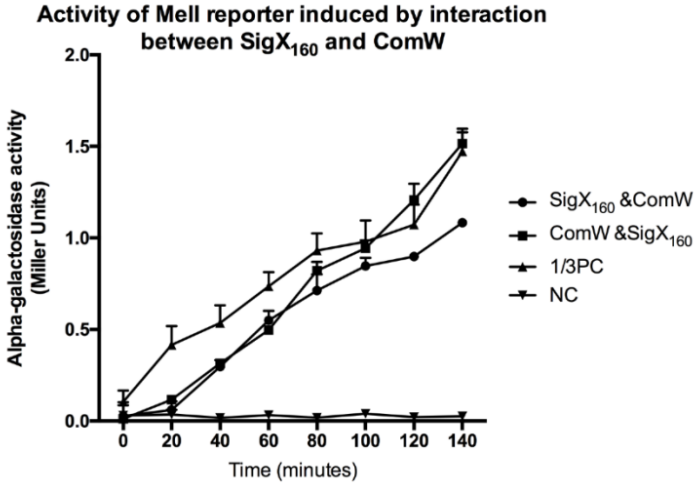


Figure 17. Quantitative α-galactosidase measurements of various interacting protein pairs. A. Evaluation of α-galactosidase activity using PNP-α-gal with PC and NC. The interaction between murine p53 protein and SV40 large T antigen had strong α-galactosidase activity, while no detectable activity in lamin and SV40 large T antigen. 1/3 of both NC and PC were measured to test the sensitivity of the assay. B. Interaction strength between protein pairs AD-$SigX_{160}$ and BD-ComW, and BD-$SigX_{160}$ and AD-ComW were measured with samples collected every twenty minutes after CSP addition. Error bars represent the standard deviation of three independent replicates. Plates were incubated for 140 minutes to allow for color development.

comW + pBD-*sigA*$_{N286}$ (pMJB13 & pMJB08), and pAD-*comW* + pBD-*sigA*$_{C83}$ (pMJB14 & pMJB08).

By cross-mating four different a and α haploids, yeast diploids were obtained: AD-*sigA*$_{369}$ + BD-*comW* (pMJB09 & pMJB04), AD-*sigA*$_{N286}$ + BD-*comW* (pMJB10 & pMJB04), and AD-*sigA*$_{C83}$ + BD-*comW* (pMJB11 & pMJB04). For the domain swap, diploids were obtained: pAD-*comW* + pBD-*sigA*$_{369}$ (pMJB08 & pMJB12), pAD-*comW* + pBD-*sigA*$_{N286}$ (pMJB08 & pMJB13), and pAD-*comW* + pBD-*sigA*$_{C83}$ (pMJB08& pMJB14). Diploids selected on SD-LT were screened for activation of reporter genes by streaking onto four other selection plates: SD-LT/X-α-gal/Ab^{R}), triple (SD-LTA/X-α-gal/Ab^{R} or SD-LTH/X-α-gal/Ab^{R}), and quadruple (SD-LTHA/X-α-gal/Ab^{R}).

The σ^{A} and ComW fusion expression did not activate expression of any of the reporter genes, which suggested that no direct interaction occurred between σ^{A} and ComW. In addition, a domain swap assay showed no activation of reporter genes by the two proteins, confirming that σ^{A} did not interact with ComW directly (see Figure 19).

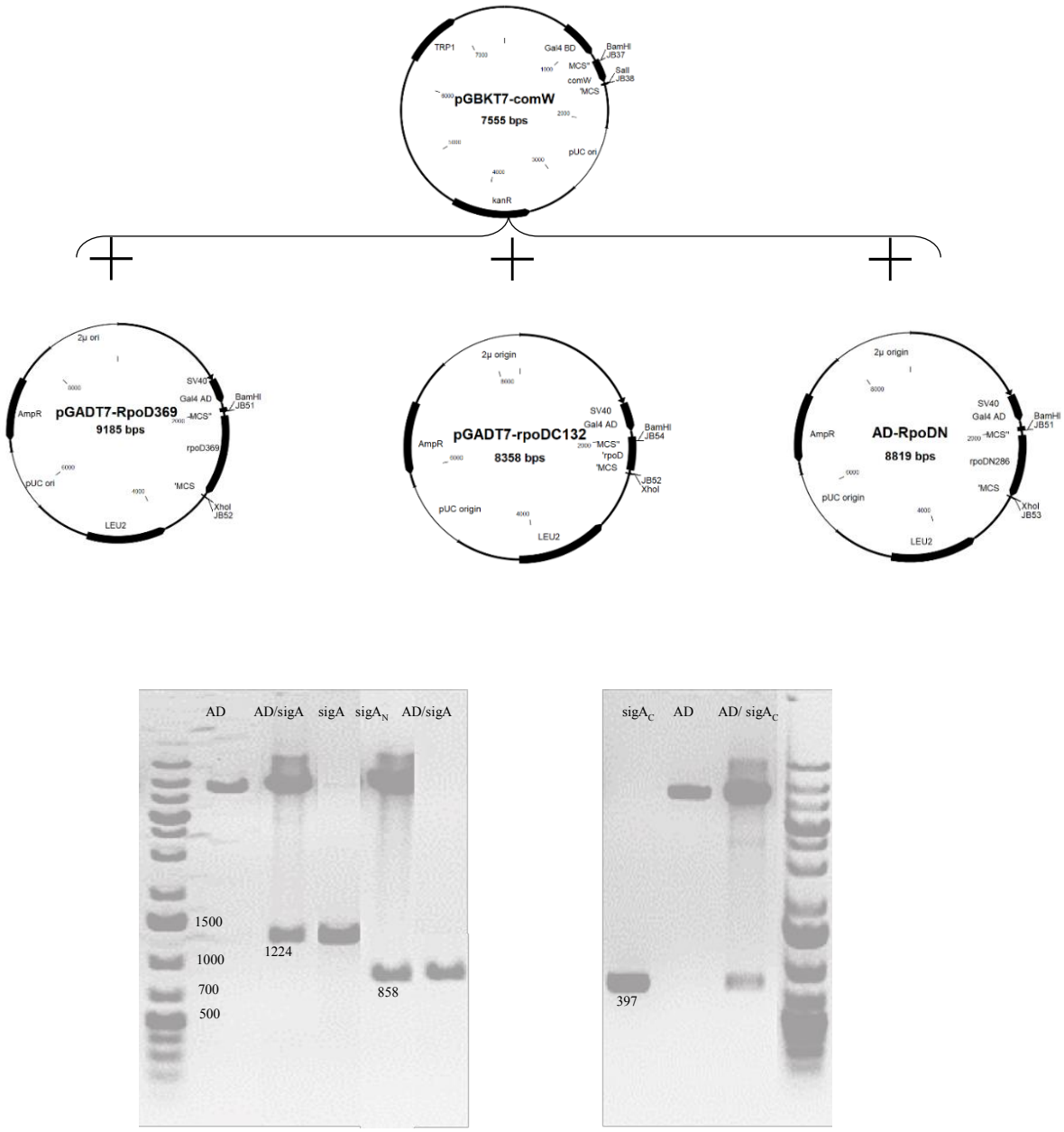


Figure 18. Cloning *sigA* and *comW* into shuttle vectors for Y2H assays. *sigA* (rpoD) was cloned into plasmid pAD at BamHI and XhoI sites, and *comW* was inserted into pBD at BamHI and SalI sites. Recombinant plasmids pAD-*sigA*$_{369}$, pAD-*sigA*$_{N286}$, pAD-*sigX*$_{C132}$ and pBD-comW were confirmed by digestion products: *sigA*$_{369}$, 1224 bp; *sigA*$_{N286}$, 858 bp; *sigA*$_{C132}$, 397 bp. Linearized pAD and pBD are 7975 bp and 7297 bp, respectively. About 100 ng plasmid DNA was digested per reaction and loaded on 1% agarose gel for electrophoresis assay.

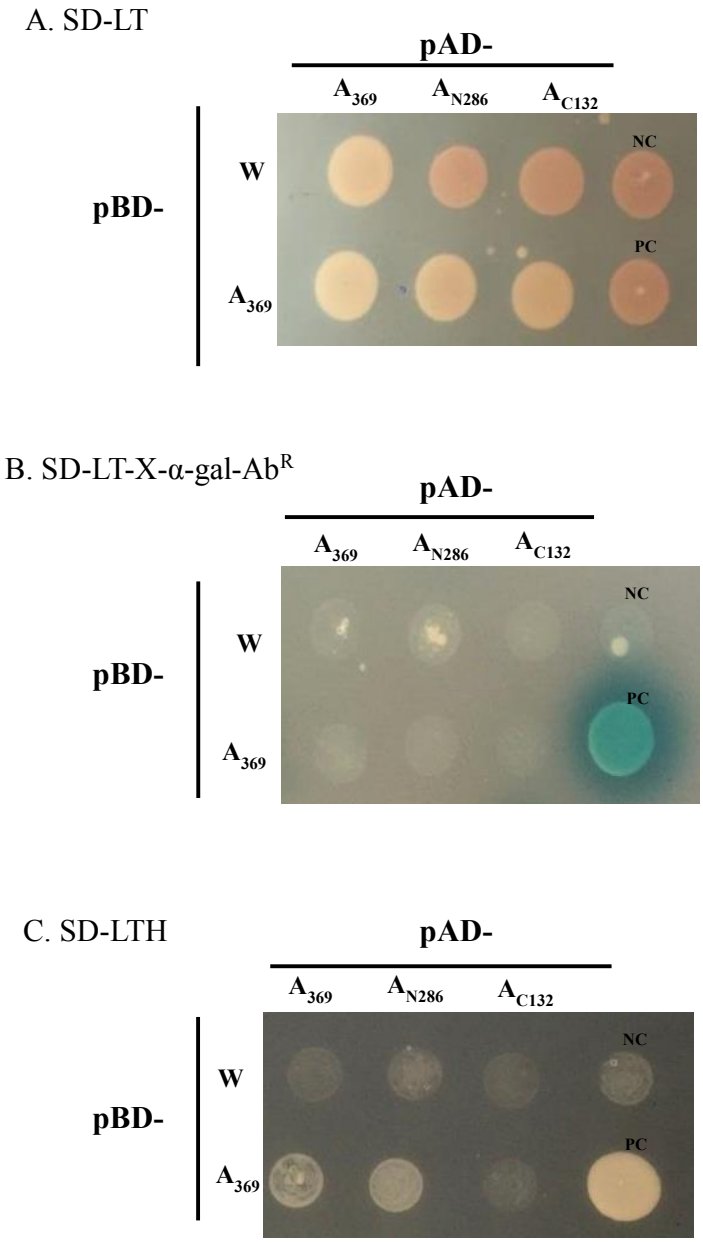


Figure 19. Lack of interaction between SigA and ComW evaluated by the Y2H assay. Diploids containing fusion proteins were plated on SD-LT (Top), SD-LT-X-α-gal-Ab^R (Middle) and SD-LTH (Bottom) plates. All diploids could grow on SD-LT plates. Interacting proteins that activate α-galactosidase allowing appearance as blue colonies on plates with X-α-gal. Proteins fused to pAD are indicated on the top; proteins fused to pBD are indicated on the left. øAD, empty pAD vector. Interacting proteins that activate his reporter gene to allow growth on SD-LTH plates. Positive control: a strong interaction between murine p53 protein and SV40 large T antigen. Negative control: lack of interaction between lamin and SV40 large T antigen.

3.4.6 Strong σ^X self-interaction

Yeast diploid strains with pAD-*SigX*$_{160}$ and pBD-*SigX*$_{160}$ were obtained and tested for activation of reporter genes. Diploids selected on SD-LT were plated onto four other selection plates: SD-LT/X-α-gal/Ab^R), triple (SD-LTA/X-α-gal/Ab^R or SD-LTH/X-α-gal/Ab^R), and quadruple (SD-LTHA/X-α-gal/Ab^R). The pAD-*SigX*$_{160}$ and pBD-*SigX*$_{160}$ fusion expression activated all four reporter genes. The σ^X self-interaction led to colony growth and blue color on X-α-gal plates and colony growth on histidine and adenine auxotroph media, suggesting that strong self-interaction of the σ^X protein occurred (see Figure 20).

Quantitative measurement of interaction strength between pAD-*SigX*$_{160}$ + pBD-*SigX*$_{160}$ was determined based on α-galactosidase activity in the presence of PNP-α-gal. In Figure 21, α-galactosidase activity in the supernatant was measured at endpoint after 24 hours of incubation at 29^{o}C. Protein pairs with pAD-*SigX*$_{160}$ + pBD-*W* showed weaker interaction with less than 1.8 Miller units, and pAD-*W*+ pBD-*SigX*$_{160}$ showed relatively stronger interaction with about 2.5 Miller units. A strong self-interaction was confirmed by almost 4.0 Miller units by pAD-*SigX*$_{160}$ + pBD-*SigX*$_{160}$. A positive control containing p53 and T antigen had 8 Miller units while negative control containing lamin and T antigen had as low as 0 readouts. This endpoint α-galactosidase assay was able to differentiate interaction strength between protein pairs.

In summary, a strong SigX self-interaction suggests SigX could form dimer after its expression, and ComW is required for breaking the intermolecular interaction of SigX.

3.5 Discussion

RNAP recognizes and transcribes specific sets of genes under the direction of σ factors. Bacteria use one predominant σ factor to transcribe housekeeping genes required for metabolism and normal growth, while alternative σ factors coordinate the gene expression required for

response to environmental changes. The σ factor switching in many bacteria has been studied extensively. However, σ factor switching between the alternative σ factor σ^X and primary σ factor σ^A in *S. pneumoniae* remains unclear. Piotrowski (2009) and Luo (2003) have estimated the number of molecules per cell of RNAP (426.9 kDa), SigX (19.9 kDa) and ComW (8.8 kDa) per cell as: 2000, 3000, and 500 copies, respectively. Based on these values, the molar ratio of SigX to RNAP is 1.72, but during protein co-purification, this ratio is 0.67. This variation suggests that σ factor switching might occur during competence development.

Several σ factor switching mechanisms may provide a clue to the molecular role of ComW in regulating σ^X activity. There have been extensive studies of σ factor at the transcriptional and translational levels. In addition, σ factor can be regulated at the activity level, either positively by a pro-activator or negatively by an anti-σ factor during σ factor switching for binding to RNAP.

In the case of *E. coli*, σ^S is regulated directly by Crl as well as indirectly by Rsd (Banta et al., 2014; Sharma & Chatterji, 2010). On the one hand, Crl interacts directly with σ^S to mediate σ^S-holoenzyme formation. On the other hand, binding of the anti-σ factor Rsd to σ^{70} prevents σ^{70}/RNAP assembly, resulting to σ^S/RNAP assembly (Campbell et al., 2003; Yuan et al., 2008). In addition, σ factor activity can be regulated by promoter recognition and binding affinity. Another activator protein, SoxS in *E. coli*, recognizes and contacts Sox-box sequences upstream of genes that specifically respond to oxidative conditions. Specifically, SoxS facilitates the holoenzyme to target promoters of those genes through direct binding to the Sox-box sequences before holoenzyme binding to the DNA.

To determine the molecular mechanism of ComW in regulating σ^X, yeast two-hybrid was used. In this study, the Y2H results are the first indication that ComW interacts directly with σ^X but not with σ^A. In the future, the interaction between σ^X and ComW can be validated by additional

methods such as surface plasmon resonance (SPR). To perform the SPR assay, His6-tagged-σ^X will be overexpressed in *E. coli* and immobilized onto sensor chips serving as the ligand. Commercially synthesized ComW could then be passed over the chip surfaces to determine ComW and σ^X binding affinity.

Tovepko and Morrison (2014) have identified point mutations in σ^A predominantly in regions 2 and 4, which can bypass Δ*comW* bypass phenotype (Tovpeko & Morrison, 2014). Comparisons of the predicted structure of σ^A in *S. pneumoniae* with resolved structure of primary σ factors in *E. coli, T. aquaticus,* and *S. aureus*, indicate that region 4 of σ^A forms a "palm and wrist" structure, with the palm acting as a hydrophobic surface for cradling of the β subunit of RNAP and the wrist contacting core residues in the β and β' subunits. This finding suggests that σ^A containing point mutations in regions 2 and 4 have weakened σ^A binding affinity to RNAP. In consequence, σ^X binding affinity to RNAP has been enhanced simultaneously without the help of ComW. The results indicate that ComW facilitates σ^X binding to core RNAP in competition with σ^A during competence. Even though the predominant mutations were found in the regions critical for interaction with RNAP subunits, the exact mechanism of ComW has still not been revealed by direct proofs. Taken together, data suggest that ComW may tip the σ^A and σ^X competition in favor of σ^X-holoenzyme assembly by directly binding to σ^X, thus weakening σ^A affinity to RNAP eliminating the ComW requirement for σ^A and σ^X switching during competence development.

Considering what is known about the mechanism of σ factor regulation in bacteria, several non-exclusive explanations exist for ComW promotion of σ^X-holoenzyme assembly. First, σ factors must undergo conformation change after they are expressed but immediately before they bind to RNAP. ComW may aid assembly by disrupting the σ^X intramolecular interaction to facilitate σ^X interaction with core RNAP. Second, ComW may facilitate σ^X activity by promoting

σ^X-holoenzyme formation. To examine this hypothesis *in vitro*, σ^X-dependent gene transcription levels can be compared directly between ComW incubation with σ^X prior to σ^X-holoenzyme formation and ComW addition to pre-formed σ^X-holoenzyme. Then the transcription levels of late genes could be measured using gel electrophoresis. If ComW does promote σ^X-holoenzyme formation, the late gene transcription levels would be higher in the case where ComW is added prior to σ^X-holoenzyme formation. As a third explanation, ComW may interact with σ^X-holoenzyme and act as a σ^X activator to facilitate the formation of σ^X-holoenzyme. And fourth, ComW may act as a transcription factor to increase the promoter binding specificity of σ^X-holoenzyme, either by providing additional functional determinants to the promoter or by nullifying the effect of a repressor. Promoter binding assay could be used to determine whether ComW directly binds upstream of σ^X-dependent genes. Alternatively, ComW may act as a protease protector of σ^X-holoenzyme, as the σ^X protein decreased and disappeared much faster without the presence of ComW.

To achieve a clear understanding of the exact mechanism of ComW, future research should incorporate *in vitro* transcription assay and electrophoretic mobility shift assay (EMSA) to determine which of the above explanations is valid. Alternative σ factors always compete with the primary σ when attempting to bind to a limited number of free RNAP core enzymes. Genetic screening of mutagenized ComW or SigX can be used to identify important residues of ComW (SigX) required for its respective interaction with SigX (ComW). To mutagenize ComW (SigX) for its loss of function in interacting with SigX (ComW), error-prone PCR and followed by reverse yeast two-hybrid (Reverse Y2H) can be used. A novel error prone PCR with mutazyme DNA polymerase, exhibits increased error rates of 2×10^3 to 7×10^3 because of increased misinsertion and misextension frequencies, whereas wild type Taq exhibits error rates ranging from 0.1×10^4

to 2 x 10^4 per nucleotide. A commercial mutazyme, GeneMorph II, can be used to manipulate mutation rates by varying the number of amplification cycles performed and/or the starting amount of target DNA in the reaction (Cadwell & Joyce, 1994). Yeast toxic markers have been used for negative selection including *ura3* (Vidal, 1996). *ura3* can be used to allow both positive and negative growth selection. 5-Fluoroorotic acid (5-FOA) can be converted to the toxic form by URA3 expression. Strains expressing *ura3* coding gene will be phenotypically ura+, while loss of activation of transcription will be ura- and 5-FOA resistance (For^R) after selection (Vidal, 1999). ura3 transcriptional reporter fusion will be used to indicate SigX and ComW interaction defective mutant in Reverse Y2H. Diploids containing missense mutants in loss of interaction between SigX and ComW cannot grow on plates without uracil supplement, but can grow on 5-FOA selective media. Conversely, diploids containing mutants not affecting the interaction will grow on uracil but not 5-FOA selective media, while *comW* wild type allele will not grow on 5-FOA plates. So ComW interaction defective mutant can be rescued from 5-FOA selective plates.

Understanding the role of ComW will pave the way for a *de novo* mechanism by which σ factors switching can be regulated by an activator in Gram-positive bacteria.

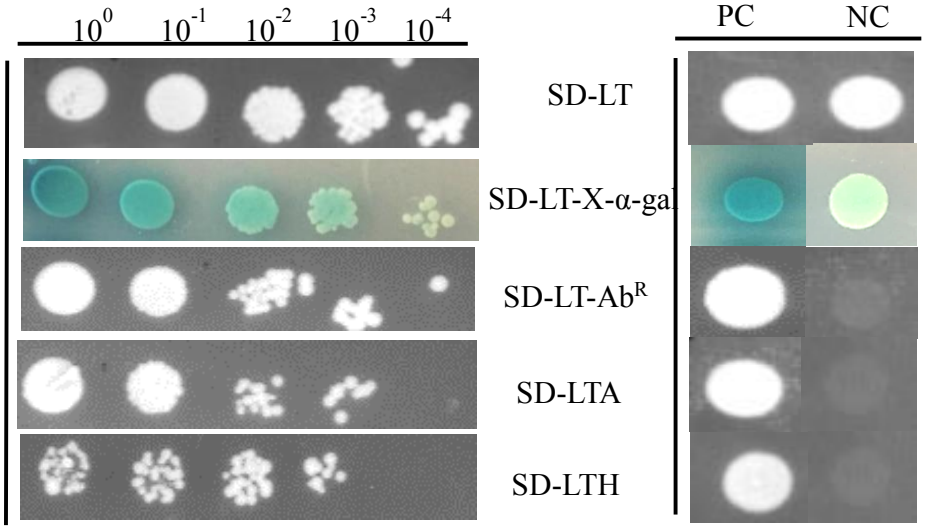


Figure 20. A strong self-interaction of $SigX_{160}$ evaluated by Y2H. Diploids with recombinant plasmids pAD-$SigX_{160}$ and pBD-$SigX_{160}$ were selected with all selective media SD-LT-X-α-gal-Ab^R, SD-LT-Ab^R, SD-LTA, and SD-LTH. Four successive dilutions of cells were done, and diluted cells were plated. Colonies with stronger interaction protein pairs grown on selective media: SD-LTH, SD-LTA, and SD-LT-Ab^R.

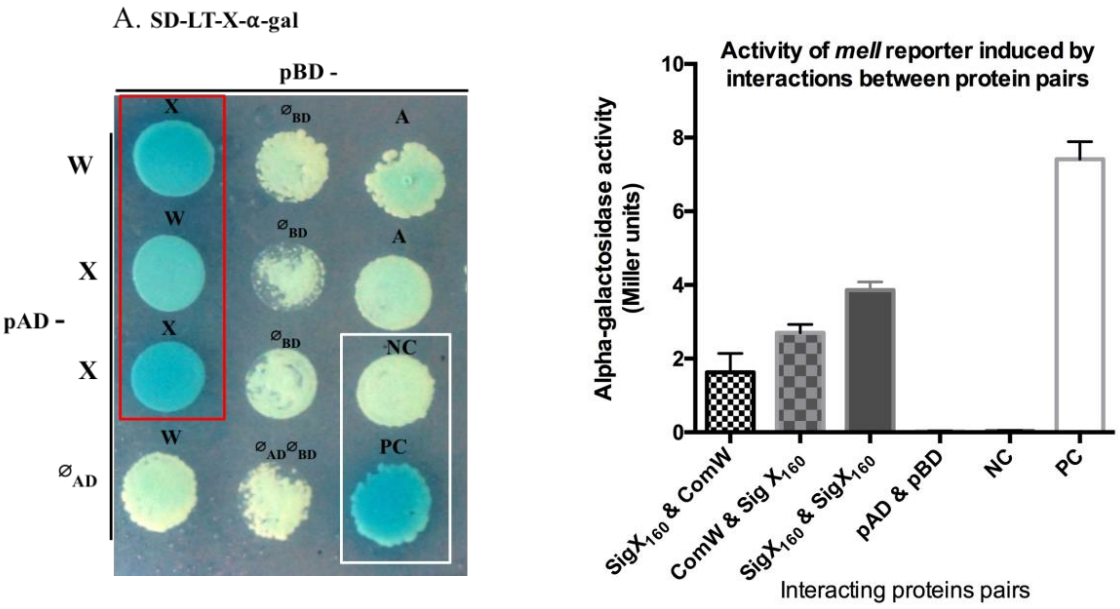


Figure 21. Protein interaction strength determined by α-galactosidase assay. (Left) Protein interactions evaluated by SD-LT-X-α-gal-Ab^R plates. (Right) Quantitative measurements of α-galactosidase activity. The interaction between SigX$_{160}$ and ComW was compared to a strong interaction between murine p53 protein and SV40 large T antigen in Positive Control and no interaction between lamin and SV40 large T antigen in negative Control. The interaction of SigX$_{160}$ itself had almost as strong as positive control indicated in a red bar. Internal control of empty vectors pAD and pBD showed no interaction as the negative control.

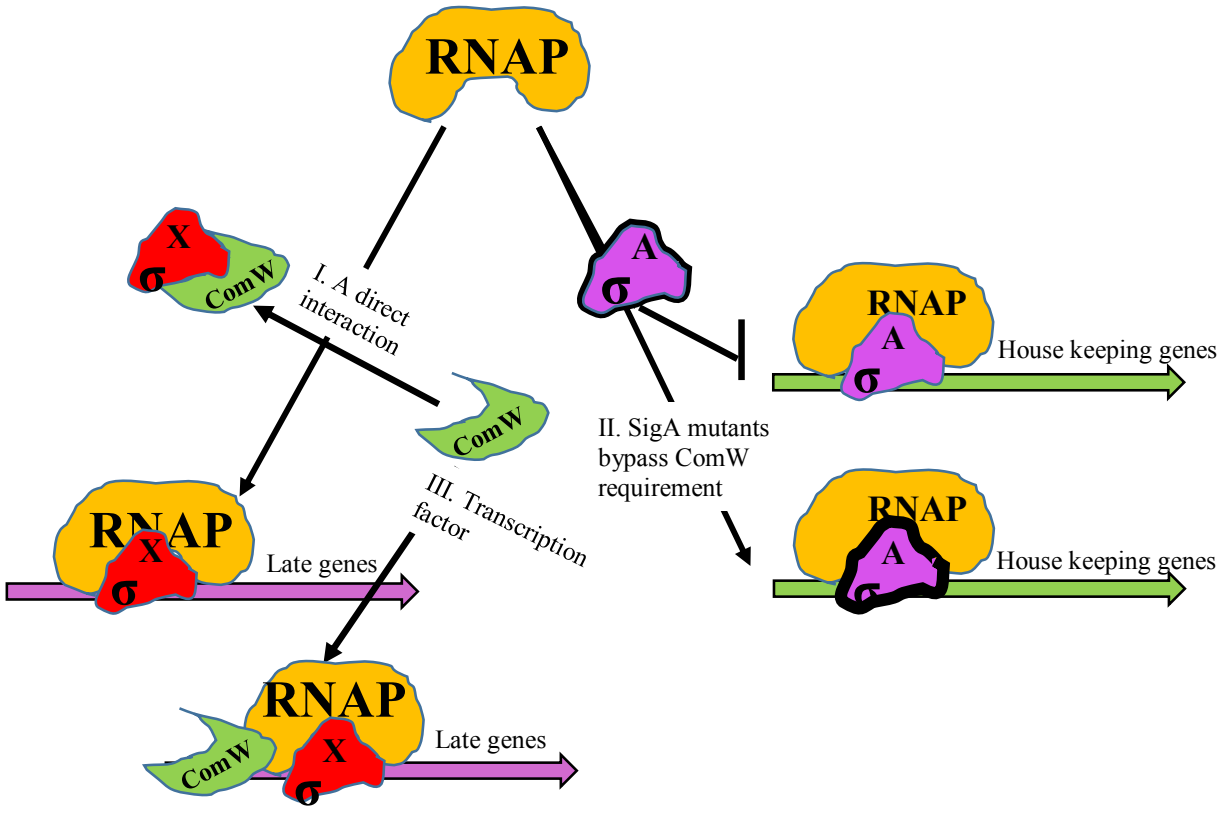


Figure 22. Hypothesized model of ComW during competence. I. A direct interaction between ComW and SigX, promoting SigX access to RNAP directly. II. SigA mutants bypass ComW requirement. III, ComW acts as a transcription factor by binding to SigX-dependent promoter to facilitate SigX access to RNAP.

IV. Role of HPF and YhaM in Competence Regulation in *S. pneumoniae*

4.1 Abstract

Competence in *S. pneumoniae* is tightly regulated, turning on and off briefly but robustly. About 147 competence responsive genes are induced after exposure to CSP. Recent transcriptome profiling combined and compared all available transcriptome data revealed a set of 27 "core SigX regulon" genes shared by all species in streptococci. Nine out of 27 of these genes were invariably overexpressed at peak competence and participated in unknown mechanisms for competence regulation (Khan et al., 2016). Interestingly, two of those nine genes may be a key to decipher phenomenon described in which three sets of mRNAs are repressed, and a set of protein expression decreased. However, the responsible regulatory mechanism is unknown.

Two of these nine core genes encode proteins with sequence similarities to proteins that target ribosome activity in other bacteria at sites critical for translation initiation. The first one, hpf_{spn} (*yfiA)*, encodes a protein homologous to a ribosome-modulating protein known for ribosome "hibernation." The other, *yhaM* (*cbf1),* is related to *yhaM* in *B. subtilis* which participates in the maturation of the 3' end of the 23S rRNA and mRNA degradation in *S. aureus* and *B. subtilis*. Significantly, MazEF mediates translation reprogramming by modifying mRNAs and generating stress-ribosomes in *E. coli*. HPF and YhaM may cooperate in a novel and unexpected mechanism related to mRNA, rRNA modification and/or ribosome reprogramming specific to competence, which may be key to the repressed mRNA and decreased protein expression.

To study the effects of *Hpf* and *yhaM* on transformation efficiency and protein expression, single and double deletion mutants were constructed. The transformation efficiency in the deletion mutants was expected to be low compared to a wild type strain strain, and the protein expression pattern in the deletion mutants was expected to swithed globally comparied with a wild type strain.

The two ribosomal-activity regulatory factors, *Hpf* and *yhaM,* may affect the protein synthesis pattern and contribute to the decreased transformation efficiency. Transformation efficiency decreased by 60% in the double mutant. Protein synthesis was pulse labeled with S^{35}-methionine and analyzed by 1-D and 2-D gels, revealing a global protein pattern shift in the wild type at peak competence. Some proteins appeared to have reduced production, especially one with approximate molecular weight of 55 kDa. In contrast, production of others increased significantly, especially with molecular weights of 14, 34 and 36 kDa. A critical comparison of 1-D and 2-D images between the double mutant and wild type showed the same protein expression pattern. HPF_{spn} and YhaM appear not to be responsible for a broad reprogramming of protein synthesis during competence. However, they may have more specific effects that account for the loss in transformation in the double mutant.

4.2 Introduction

In response to physical and chemical environmental stimuli, bacteria develop rapid mechanisms to regulate gene expression and protein activity. Protein synthesis that switches on and off during responses to stringent conditions is precisely regulated (Bosdriesz, Molenaar, Teusink, & Bruggeman, 2015; Keiler, 2015). In *S. pneumoniae*, competence can be induced by a pheromone CSP, and more than one hundred genes are reported to be upregulated immediately (Peterson et al., 2004). By whole genome-wide random mutagenesis, Peterson revealed 23 genes required for transformation in *S. pneumoniae* (2004). Recently, a genome-wide transcriptome study combined and compared all available transcriptome data revealed a set of 27-30 genes grouped into core σ^X regulon with orthologs shared by all species in Streptococci (Khan et al., 2016; (Khan et al., 2017). Of the 27 genes discovered in the study, 18 were confirmed to be required for the process of DNA uptake and recombination (Khan et al., 2016). Of the 18 genes,

12 are indispensable for DNA uptake, while the other 6 are important for subsequent recombination. Surprisingly, the remaining 9 genes, which are upregulated during competence, have unknown functions in competence.

Transformation in Pneumococci depends on the synthesis of a small set of proteins (Morrison & Baker, 1979). Radioactive labeling of synthesized proteins has been widely used to monitor protein synthesis pattern changes. Pneumococcal protein synthesis during competence development was first studied using radioactive S35-methionine pulse labeling in a largely chemically defined medium (CDM) with low methionine (Morrison, 1978; Morrison & Baker, 1979). In that study, protein synthesis was compared in induced and uninduced cultures using CSP in a series of successive 3-minute periods exposure to S35-methionine during competence development, followed by extraction and separation in a 1D gel. As summarized by Lee and Morrison in their 2000 review, a global protein synthesis pattern was revealed in a temporarily competent culture exhibiting two prominent features of protein synthesis: one set of proteins showed a striking increase in the synthesis, while the other set showed a robust decrease during competence.

As we know, 147 genes are upregulated during competence (Khan et al., 2016). It is typical to observe an upregulated protein synthesis pattern during CSP induction in response to mRNA newly synthesized during competence. Peterson et al.(2004) found three sets of genes with transiently repressed mRNA in a wild type during the response to CSP: the first set of three genes which were adjacent to a cluster of ribosomal protein genes of both 30S and 50s subunits, decreased dramatically but briefly to about half the precompetence levels at 12.5 minutes or 15 minutes, that was also observed by Rimini et al. (2000); the levels of the second set of genes, which encoded proteins involved in carbohydrate or amino acid metabolism, decreased continuously

from the beginning of CSP treatment but recovered to basal level after 15 minutes; and the levels of the third set of two genes encoding alcohol dehydrogenases decreased immediately upon CSP addition and lasted for 15 minutes. However, the mechanism for the downregulated mRNA remains unclear.

Collectively, the downregulated mRNA adjacent to ribosomal proteins may in line with the decrease of protein synthesis, that mRNA maximally decreased coincided with dramatically decreased proteins at peak competence, that is the maximum expression of σ^X. However, linkage of the repressed mRNA to σ^X expression in response to CSP remains not established.

The set of core σ^X-regulons provide a rich repertoire for an innovative study of competence regulation. This study aimed to address three questions. First, does the repressed mRNA pattern explain the decreased protein synthesis pattern? Second, how are these patterns related to competence regulation? And third, does any gene in the set of nine uncharacterized core competence genes play a role in the downregulated protein synthesis pattern? To investigate the hypothesis, the set of nine uncharacterized core genes were under consideration. Their protein products include four orthologs of well-characterized proteins and five proteins with unknown functions but with broad predicted roles. These roles include participating in DNA repair and protection and acting as ribosomal subunit interface protein and cmp binding factor (CBF).

Of particular interest is the sequence similarity of two genes, *Hpf and yhaM,* to proteins that target ribosome activity and rRNA processing in other bacteria at sites critical for translational initiation (see Figures 22 and 23). Our hypothesis was whether HPF and YhaM involved in a novel translational mechanism in regulating protein synthesis in response to CSP. The gene *Hpf* falls into the family known as ribosome hibernation factors (HPF). HPF family proteins have been classified into three types, long HPF, short HPF, and YfiA, a ribosome associated inhibitor A

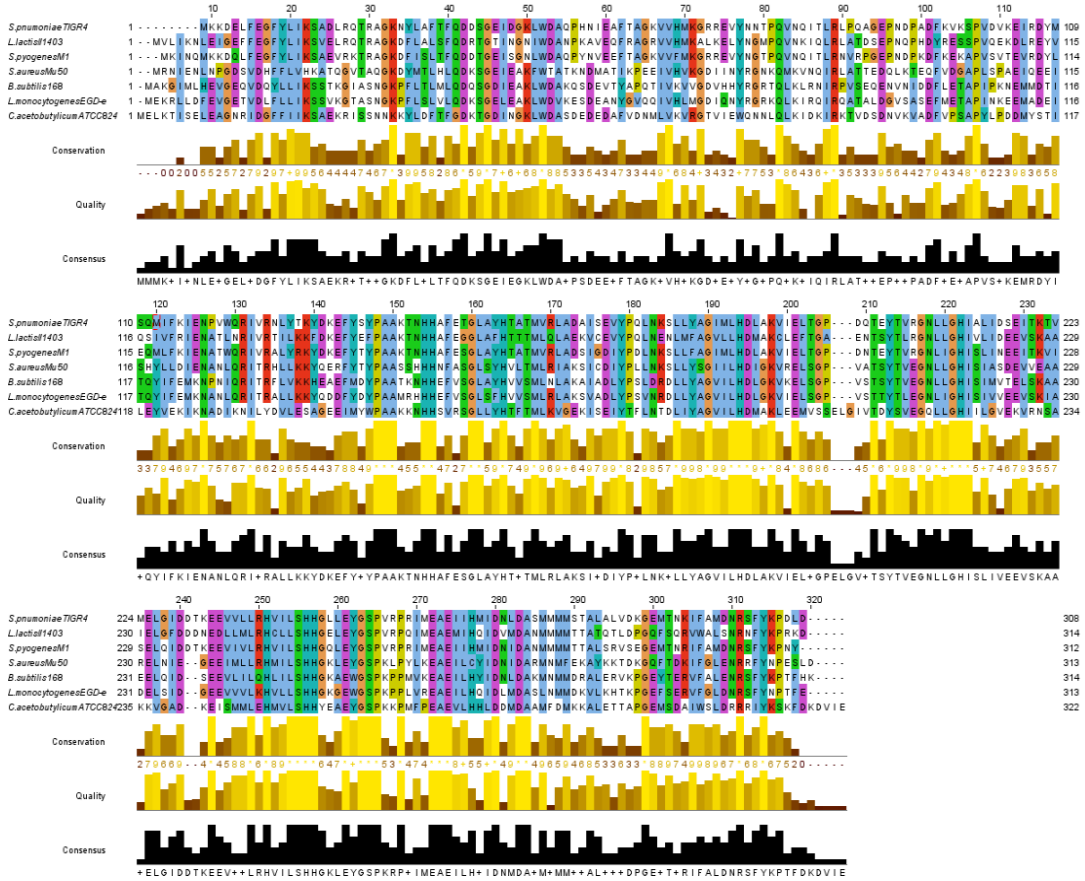


Figure 23. Alignments of *Hpf* (*yfiA*) homologs. Sequences from *Streptococcus. pneumoniae, Lactococcus. lactis, Streptococcus. pyogenes, Staphylococcus.aureus, Bacillus. subtilis, Listeria monocytogenes, and Clostridium acetobutylicum.* Colour schemes used according to physicochemical attributes of amino acids. According to ClustalX, colour codes demonstrate the physicochemical properties. The consensus row shows a '+' when more than two residues are abundant. The highlighted residue suggests that its frequency is higher than a threshold, and it can be conserved and important residue.

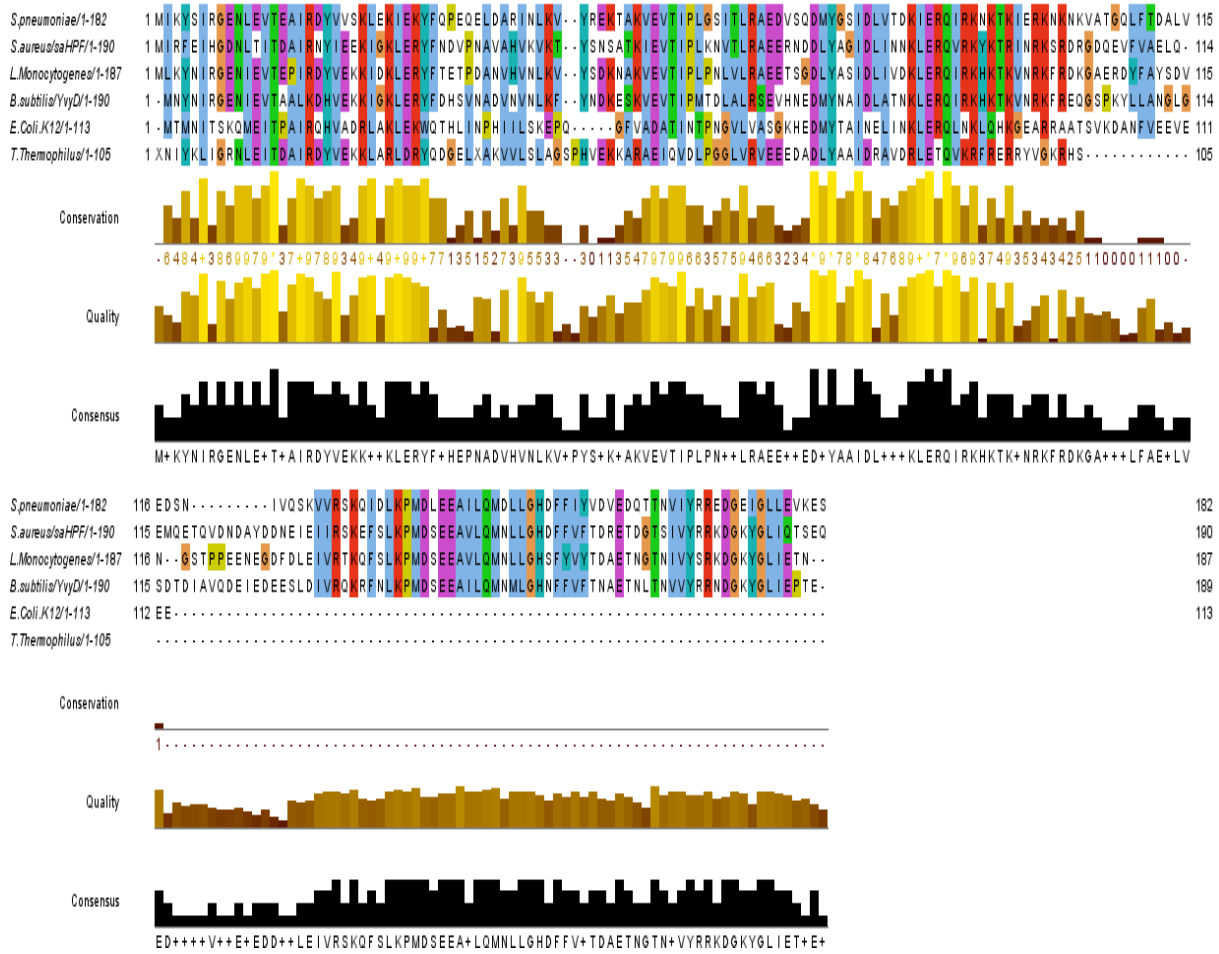


Figure 24. Alignments of *yhaM* (*cbf1*) homologs. Sequences from *Streptococcus. pneumoniae, Staphylococcus. aureus, Listeria monocytogenes, Bacillus. subtilis*, *E. coli,* and *Thermus Themophilus.* Colour schemes used according to physicochemical attributes of amino acids. According to ClustalX, colour codes demonstrate the physicochemical properties. The consensus row shows a '+' when more than two residues are abundant. The highlighted residue suggests that its frequency is higher than a threshold, and it can be conserved and important residue.

(Masami Ueta et al., 2010). In *E. coli,* a short HPF and a ribosome modulation factor (RMF, 6.5 kDa) mediate formation of 100S ribosomal dimers when cells enter stationary-phase and/or stress conditions (Polikanov et al., 2012). HPF and RMF binding to ribosomes prevent tRNA and mRNA from interacting with ribosomes, resulting in abolished translation.

It has been demonstrated that 100S ribosomes are only formed by a long HPF in bacteria except for gammaproteobacteria. Streptococcal HPF belongs to the long HPF class, and its expression in *S. mutans* increases by 3.2-fold in CSP induced competence in TSB (a peptide-rich medium) (Khan et al., 2016), and 12.5-fold in XIP induced competence in CDM (Khan et al., 2017). This HPF is part of the ComFAC operon, downstream of ComFA and ComFC, which encode two key components of the DNA uptake apparatus (Khan et al., 2016). The protein product of another gene, *yhaM*, downstream of *ccs50* (DNA recombination protein RmuC), is related to a protein exhibiting exoribonuclease activity in *B. subtilis.* In *S. mutans*, its expression increases by 5.9-fold in CSP induced competence in TSB (Khan et al., 2016), and 16.5-fold in XIP induced competence in CDM (Khan et al., 2017). In addition, Wenderska and co-workers confirmed the increased expression of HPF and YhaM by 4.9-fold and 2.1-fold, respectively in the XIP/CDM condition in their whole-genome transcriptome study (Wenderska et al., 2017). Notably, the expression of both genes increased during competence in all streptococcal groups; their presence may partially explain downregulated protein expression in the form of repression of protein expression by inactivation of ribosomes mediated by HPF in *E. coli* under stress conditions. Inactivation of ribosomes may result in suppression of normal growth protein synthesis and substitution of a narrowly focused expression regime.

Three of the best-characterized cases where global protein synthesis is regulated by ribosome modifications are as follows. In the first case, starvation-induced ribosome hibernation

by HPF and RMF in *E. coli* (Polikanov et al., 2012), a nutrient-limited induced long HPF_{Sa} to differentially repress translation in *S. aureus* (Basu & Yap, 2016), and energy-starving induced dimerization of ribosomes by YfiA in *L. lactis* (Puri et al., 2014). In the second case, stress-induced translation of specific proteins in the MazEF regulon. MazF can be activated by release from MazE sequestration, which is degraded by the proteolytic complex ClpAP. Activated MazF degrades most transcripts and removes 43 nts upstream of specific mRNAs, thus creating leaderless mRNA; meanwhile, it also removes 43 nts of 16S rRNA containing the anti-SD sequence. The newly generated ribosomes specifically recognize leaderless mRNAs and selectively initiate translation (Vesper et al., 2011). In the third case, stress-induced maturation of specific rRNA in *B. subtilis* by YhaM, which is characterized as an exoribonuclease together with exoribonuclease RNase J1 and RNase PH (Redko & Condon, 2010) and maturation of 23S rRNA in *S. aureus* with PNPase and PNaseR (Bonnin & Bouloc, 2015).

Considering the sequence similarity of HPF and YhaM in other bacteria, it was hypothesized that HPF and YhaM in *S. pneumoniae* regulate competence by controlling translation activity through either modification of either ribosomes by means of rRNA maturation or formation of ribosomal dimer or mRNA at the stability level. The modifying activity could affect ribosome reprogramming specific to competence in such a way as to sharpen the translation pattern. The working hypothesis for the role of HPF and YhaM was that the two genes generate the translational downshift pattern during competence induced by CSP in Baker CDM.

4.3 Methods

4.3.1 Construction of mutants

To construct Δ*Hpf* mutant, a DNA fragment containing tet^R was transformed to strain CP2137 to replace *Hpf*. *Hpf*-upstream (1348 bp) and downstream (1274 bp) fragments were amplified by two separate PCRs with primer pair JB91/JB92 and JB93/JB94 with CP2137 genome DNA as template. The tetracycline resistant gene (tet^R, 1526 bp) was amplified with primer JB99/JB100 with 6mt genomic DNA. Then all three fragments were ligated, and the ligation DNA was transformed to strain CP2173. Transformants selected on tetracycline plates were verified by Sanger sequencing the *Hpf* locus, and the correct clone was named as CP2180. Then a kanamycin gene (Kan^R, 957 bp) was amplified with primers JB112 and JB113 from a 6mt genomic DNA. *yhaM*-upstream (1086 bp) and downstream (1027 bp) regions were amplified by two separate PCRs with primer pair JB95/JB96 and JB97/JB98 with CP2137 genome DNA as template. Again, all three fragments were ligated, and the ligation DNA was transformed into CP2180 and CP2137. Transformants were selected on kanamycin plates, and one clone from each transformation assay was verified by Sanger sequencing and named as CP2182 and CP2181, respectively.

4.3.2 Pulse-chase labeling assay

When cells grown in met-cys-free medium reached OD_{550} 0.1, a portion of the culture was induced to competence by adding BSA, $CaCl_2$, and CSP as described. At specific time points, a 1mL competent culture was collected and chilled on ice. After all samples collected, following centrifugation at 10,000 G for 10 minutes at 4°C. Discard supernatant and resuspend pellets in 1mL cold PBS (2.56g Na_2HPO_4. $7H_2O$, 8g NaCl, 0.2g KCl, 0.2g KH_2PO_4) containing 1mg/mL cold methionine, and pelleting again as above, the PBS washed cells were then resuspended in pre-warmed (37°C) labeling medium.

S35-met stock (MP Biomedical) was thaw at RT and prepared as 10 uCi/uL working solution in pre-warmed (37^{o}C) medium. During pulse phase: 20uL of S35-met working solution was added into 1mL competent culture and incubated for 3 minutes at 37^{o}C; while during chase phase: 50 uL ice cold methionine (20 mg/mL) was added, and removed, chilling on ice. After all samples were collected during a cycle of competence induction, all samples were lysed in 300 uL lysis buffer (10 uL Tris-HCl, 1M, pH 8.0; 10 uL EDTA, 100 mM; 10 uL PMSF, 100 mM; 40 uL 10% Triton; 2 uL DNaseI, 5mg/ml; 25 uL RNase, 2 mg/mL; 0.1 mL 10x protease inhibitor; 50 uL $MgCl_2$,100 mM; 50 uL $CaCl_2$,100 mM per1mL), incubated at 37°C for 15 minutes, supplemented with 1.5 uL Sarkosyl, 10%; 15 uL; EDTA, 100 mM; incubated at 37°C for 3 minutes, and chilled on ice. Then released protein from lysis were precipitated by TCA or acetone. Alternatively, 4 volumes of ice-cold acetone were added to cell lysates to precipitate for overnight at -20 ^{o}C.

4.3.3 TCA precipitation

To precipitate protein from total cell lysates, 40% trichloroacetic acid (TCA, w/v) solution was prepared by dilution from 100% TCA stock. Cell lysates were mixed with 5 volume of ice cold TCA (40%) for precipitation on ice for 1 hour, then centrifuged at 10,000G for 10 minutes at 4^{o}C. Discard supernatant, washed with 1mL ice-cold acetone, make sure pellet resuspend completely. Repeat the step with acetone wash and let pellet dry in a hood for at least one hour. The dried pellets were resuspended in 22 uL 1x Laemmli sample buffer (0.0625 M Tris-HCl, pH.8, 2% SDS, 0.002% bromphenol blue, 10% glycerol, and 5% β-mercaptoethanol).

4.3.4 Acetone precipitation

Total protein from cell lysate could be precipitated by acetone. Apply 4 volumes of -20^{o}C-cold acetone to the cell lysate and allow precipitation overnight. Centrifuge at 10,000 g for 10 minutes at 4^{o}C. Discard supernatant, washed with 1mL -20^{o}C-cold acetone, make sure pellet

resuspend completely. Repeat the step with acetone wash and let pellet dry in hood for at least one hour. The dried pellets were resuspended in 1x Laemmli sample buffer as above.

4.3.5 Determination of radioactivity

Liquid scintillation counting (LSC) was used to measure S35-methionine incorporation. First, 5mL scintillation cocktail was mixed with 2uL protein samples dissolved in Laemmi sample buffer. S35-methionine is not a strong isotope, and the signal could be quenched by extra salt in H2O. Soaking protein samples in scintillation cocktail for 4 hours to overnight could be helpful to stabilize the signals. Second, the mixture was scanned by scintillation counter (C14 channel). Then, the amounts of S35-met in each fraction were calculated by scintillation readout using CPM (counts per minute) minus background counts. Percentage of S35-met incorporation into newly synthesized proteins was calculated by the ratio of CPM in the total cell lysate to CPM of the initial culture.

4.3.6 One-dimensional radiograph

The sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was used in a discontinuous running system using Bio-Rad protein mini gel system (0.75 mm). SDS-PAGE gel was prepared first by casting 10% resolving gel (10 mL solution: 3.4mL, 30% Acrylamide/Bis, 37.5:1; 100 uL10% SDS; 100 uL 10% APS; 5uL, TEMED; 2.5ml, 1.5M Tris-HCl, pH 8.8). Allow for resolving gel polymerized for 30-60 minutes. Secondly, 4.5 % stacking gel was prepared (4 mL solution: 0.6 mL, 30% Acrylamide/Bis, 37.5:1; 37.5 uL10% SDS; 37.5 uL 10% APS; 3.75 uL, TEMED; 0.94 mL, 0.5M Tris-HCl, pH 6.8;) with a comb was placed onto the top to form wells during polymerization. Allow for 30 minutes to be polymerized. An appropriate amount of protein samples dissolved in Laemmli buffer were loaded in pre-polymerized SDS-PAGE gel, and samples were separated under 100V constant voltage for 30 minutes, and 200V for 30 minutes. The

Coomassie Brilliant Blue G-250 staining procedure (staining for 30 minutes with staining buffer of 12.5% Coomassie blue (w/v), 22.5% methanol, and 7.5% acetic acid; and destaining for 1 hour with destaining buffer of 30% methanol and 10% acetic acid) was used to reveal protein expression pattern.

4.3.7 Two-dimensional radiograph

To performed 2D gel analysis, protein samples were prepared with SDS buffer, together with vortex and glass beads, and sonication as reported by Kendrick company. Separation of protein samples started with isoelectric focusing (IEF) and followed by SDS-PAGE. IEF separation was carried bout in 3.3 mm diameter glass tubes using 2% pH 4-8 mix servalytes. The internal standard was tropomyosin with MW 33,000 and pI 5.2. SDS-PAGE was carried out with 10% acrylamide slab gel (1mm). On each 2D-PAGE, there were 150 uL of samples containing 100 ug of protein. The samples were diluted in 1:1 SDS buffer: urea buffer.

4.3.8 Gel Drying with a vacuum heat dryer and vacuum pump

After Coomassie staining procedure, an SDS-PAGE gel was equilibrated with four series of buffer for 5 minutes each (for a big gel 20 minutes each): 10% ethanol, 20% ethanol, 30% ethanol, and 40% ethanol with 10% glycerol. Then discard all buffer, and place on the top of the gel a piece of Whatman paper, with the size slightly bigger than the gel, then cover the whole sandwich with saran wrap. Next, place the wrapped gel under a vacuum heat dryer, set Tm as 80^{o}C for 2-3 hours, then turn on a vacuum pump. Remove dried gel and placed under cassette exposure overnight.

4.3.9 Gel Drying with cellophane membranes

A SDS-PAGE gel can also be dried with cellophane membrane after Coomassie staining procedure. First, wet a cellophane membrane, and then place on the acrylic frame. Place the gel

on the membrane, then cover the gel with a second wet cellophane membrane. Place the acrylic frame and tight the frame with clips. Place the frame under the hood, and let it dry overnight or longer. Remove the dried gel and place under cassette exposure overnight.

4.4 Results

4.4.1 Baker CDM use in a pulse-chase analysis for protein synthesis

Morrison and Baker (1979) reported a dramatic global shift of protein synthesis accompanying competence development. Their 1D-PAGE protein gel revealed dozens of newly synthesized proteins, while many proteins were down-shifted (Morrison & Baker, 1979; Morrison & Lee, 2000). However, there is no comprehensive catalogue of actual protein synthesis shifts. To determine whether HPF and YhaM were implicated in this downshift phenomenon, the protein patterns were studied in a wild type and Δ*Hpf*Δ*YhaM* mutant by means of 1D and 2D gel electrophoresis.

To determine whether nascent proteins in competent cells can be efficiently labeled, CDMs were compared and combined of known CDMs. The first chemical semi-defined or completely defined medium to culture pneumococcal strains was developed (Adams & Roe, 1945; Rappaport & Guild, 1959). Based on previous recipes, Sicard formulated a synthetic medium allowing cultivation of lab strains that previous media could not support, which established the nutritional requirement for most pneumococcal strains (Sicard, 1964). Later the CDM most utilized for pneumococcal strains was described by Lacks and Van de Rijin (Lacks, 1970; Van de Rijin, 1980). Morrison formulated CDM components by mixing the buffers and vitamin mixture of Lacks with the amino acid mixture of Sicard, and adding a small amount of yeast extract to support growth (Morrison, 1977, 1978; Morrison & Baker, 1979). In the CDM formulated by Morrison and Baker,

here called Baker CDM, 98% of the components were defined and could support competence development with CSP induction. The amounts of total amino acids and especially methionine was limited in Baker CDM with the latter limited to only 1 μg/mL. This defined medium could be used to label proteins with S35-methionine (Morrison, 1977, 1978; Morrison & Baker, 1979).

In this study, Baker CDM was found to provide sufficient nutrition for pneumococcal growth (see Figure 25). In comparing the growth curves in Baker CDM and the peptide-rich medium CAT, a growing pneumococcal population doubled at a regular interval of 40 minutes per generation in both media. After 10 generations from original OD_{550} 0.01, the population reached a stationary phase.

In addition, Baker CDM was found to be suitable for the development of competence and incorporation of S35-methionine in wild type cells. Competence kinetics could be monitored by transformation a donor DNA into Strain CP2137, and it typically starts 5 minutes after CSP induction, peaks between at 20 minutes, and ceases at 40 minutes. During competence development initiated by CSP addition which was recorded as 0 minutes, the level of competence was measured at each of the specific time points indicated in Figure 24B. Transformation efficiency during incubation in the medium was determined by transforming a pure PCR DNA. Transformation efficiency was then determined by the percentage of antibiotic resistant colonies to the original cells per milliliter, in this case, novobiocin resistance (Nov^R) colonies per milliliter of cells. Maximal transformation efficiency was achieved by using a saturating amount of donor DNA, which was 1μg/mL purified PCR DNA. Unfortunately, transformation efficiency was only 16% at peak competence, raising the question of how to modify the medium to improve the level of transformation.

To determine S35-methionine incorporation in Baker CDM in parallel, S35-methionine was pulse incorporated beginning at 0 minutes and ending at 40 minutes after addition of CSP to the Baker CDM. A small portion of competence cells was exposed to S35-methionine for 3 minutes at 37^{o}C, followed by chilling that stopped protein synthesis. When competence reached a peak between 15 and 20 minutes, up to 4% the of S35-methionine was incorporated in a 3-minute pulse (see Figure 24). As a comparison, Morrison reported incorporation of 10% to 20% of the S35-methionine (Morrison, 1977, 1978). In summary, Baker CDM was needed to be modified to provide high transformation efficiency while remaining sufficient isotope incorporation.

4.4.2 Increased transformation efficiency with sufficient isotope incorporation

An efficient level of competence is required to study protein expression patterns. However, the competence achieved in Baker CDM was less than 20%, which suggests that only a small population became competent and that the amount of newly synthesized competent proteins may not have been sufficient to reveal a protein pattern shift from no competence to full competence. Consequently, Baker CDM and a complete CDM described in a previous study (Chang & Federle, 2011) were compared. The complete CDM used to induce competence achieved transformation efficiency of up to 90%. Two major differences observed in the comparison were (1) a much higher concentration of cysteine in the complete CDM than in Baker CDM and (2) different amounts of amino acids in the two (see Table V). In this study, a modified Baker CDM was prepared by modifying the complete CDM in the direction of Baker CDM; specifically, a new amino acid mix and bases mix were made to match the compositions in Baker CDM, and then the solution-making method described (Chang et al., 2011) was followed. Subsequently, the washed-cells approach suggested by Morrison (Morrison & Baker, 1979) was followed to resuspend cells in the modified Baker CDM containing some or all of the components of the Baker recipe to determine which

components could support competence development in order to reproduce the protein downshift pattern observed by Morrison and Baker. The washed-cells approach was used to eliminate components in the growth medium that might affect pulse-chase assay.

It was challenging to synchronize levels of competence and radioisotope incorporation. A rich medium may provide a high level of competence but reduce radioisotope incorporation, whereas a poor medium may boost the latter while compromising the former. Several parameters, including cell density, supplementation, and temperature, were examined in efforts to balance the radioisotope incorporation level while maintaining a high competence level.

A wild type culture collected from complete CDM was centrifuged, the pelleted cells were resuspended to a specific cell density, and the cells were washed with Baker CDM. The washed-cells approach was used to eliminate methionine and cysteine residue from 2% CAT that was added to support cell growth. To optimize levels of competence and radioisotope incorporation, different supplements were added to Baker CDM and compared. Because 1% CAT can mildly improve transformation efficiency of a wild type strain, 2% and 5% CAT were assessed to determine whether they could be used to further increase the competence level with minimal compromise of radioisotope incorporation; independently, a small addition of yeast extract was tested, as it was formerly found to support competence (Morrison, 1978). The amounts of methionine in 2% CAT and yeast extract (10.2 and 8 μg/mL, respectively) were rigorously compared because non-labeled methionine was competitively incorporated into newly synthesized proteins and excluded S35-methionine at a concentration more than 0.02 μg/mL in pulse-chase assays (see Figure 26). Cysteine in growth medium can also affect S35-methionine incorporation because cysteine allows endogenous synthesis of methionine (Sperandio et al., 2007). The amounts of cysteine in Baker CDM and complete CDM were 140 and 625 μg/mL, respectively, which

resulted in a substantial difference in S35-methionine incorporation, indicated by more than 2.5-fold decrease in incorporation efficiency (see Figure 26). Three other supplements—tryptone, casein, and NaCl/phosphate/glucose—were examined in competence and pulse-chase assays (see Figure 27). All three showed decreased levels of competence and isotope incorporation. In addition, no detectable competence was observed in the medium without supplementation.

To conclude, adding 2% CAT to the modified Baker CDM plus cell-wash process resulted in a competence level greater than 70%, while the radioisotope incorporation level was up to 1%, which was sufficient for detecting protein patterns in a signal sensitivity test. Moreover, five versions of CDM (I through IV and Baker CDM) differing in their amounts of amino acids were tested in the two assays; the results showed that Baker CDM had optimal levels of both competence and isotope incorporation (see Figure 28). CDM-I compromised the level of transformation to less than 50%, while CDM-II, -III, and -IV compromised both levels. The assays resulted in the identification of Baker CDM as the best medium permitting efficient incorporation of radioisotopes into newly synthesized proteins of competence cells as well as highly efficient uptake of donor DNA.

4.4.3 Lack of improved isotope incorporation in Baker CDM in *metEF* mutant

To eliminate the cysteine effect in a pulse-chase analysis, a Δ*metEF::tet*^{R} mutant was constructed. MetE (methionine synthase) together with MetF (methylenetetrahydrofolate reductase) and FolD (5-methyltetrahydrofolate) provides a methyl group to homocysteine to form methionine (Sperandio, 2007, Afzal, 2016). Deletion of MetEF would eliminate endogenous synthesis of methionine from the cysteine present in CDM. Using homologous recombination, a tetracycline-resistant (Tet^R) DNA was transformed to replace the *metEF* locus using PCR amplified DNA with primers JB165/166 and JB169/170. Cysteine can be processed into

homocysteine and further metabolized into methionine. Therefore, methionine can be synthesized endogenously even when little or no methionine is added to the medium. Unlabeled methionine can compete fiercely with S35-methionine during pulse-chase of native protein products.

Isotope incorporation efficiency was determined following a method described by Morrison and Baker (1979). Using this method, after CSP exposure, samples were collected at peak competence when the culture was exposed to S35-methionine in successive 3-minute intervals. The isotope incorporation efficiency results for wild type and Δ*metEF* were compared and were found to be similar. This finding was not surprising because smu1487 was previously reported to be involved in endogenous methionine synthesis (Sperandio, 2007).

4.4.4 Competence decrease in Δ*Hpf* and Δ*yhaM* mutants

Previous research revealed that synthesis of a set of proteins decreased during competence and returned to normal post-competence (Morrison & Baker, 1979; Morrison & Lee, 2000). To determine whether *Hpf* and *yhaM* played a role in this down-regulation of protein expressions during competence in logarithmic growth of pneumococcal cells, deletion mutants of *Hpf (yfiA)* and *yhaM (cbf1)* were constructed by first replacing *Hpf* with a tetracycline resistance (tet^R) gene, and then replacing *yhaM* with a kanamycin resistance (kan^R) gene.

Competence assay was performed with cultures of wild type and mutants (Δ*Hpf*, Δ*yhaM,* and Δ*Hpf*Δ*yhaM*) exposure to CSP at a specific cell density. A small aliquot of the culture was then collected at different time points and continuously exposed to a donor DNA for 3 minutes. Transformants were selected on antibiotic plates, and their transformation efficiencies were calculated as described in Methods. Compared to the wild type, the transformation efficiencies of mutant strains Δ*Hpf*, Δ*yhaM,* and Δ*Hpf*Δ*yhaM* were 54%, 37%, and 36% respectively at peak

competence (see Figure 29). These results showed that HPF and YhaM were required for full competence.

4.4.5 Similar protein expression patterns in Δ*Hpf*Δ*yhaM* and wild type on 1D gel

1D electrophoretic separation of denatured proteins is by far the most fundamental and important technique in proteomics studies (F. Gibson et al., 2008). To examine components of newly synthesized proteins during competence induced by extracellular CSP, samples were labeled by successively exposing to S35-methionine for 3 minutes during a cycle of competence appearance and disappearance in the modified Baker CDM. Competence in the wild type strain peaked 20 minutes after CSP exposure and disappeared after 40 minutes (see Figure 30). The efficiency of isotope incorporation was calculated following the standard method described in Methods. At the time points indicated in Figure 30, each sample containing 10^4 Cpm was loaded on the 1D-SDS gels.

A global protein pattern shift was observed in the wild type, the Δ*Hpf,* and Δ*Hpf*Δ*yhaM* mutants during competence, but no significant difference between the three. CP1894 (Δ*dprA*) was also examined as an internal control for an extended competence-specific protein expression pattern, as the *dprA* deletion mutant was reported to have prolonged a competence phenotype (Weng et al., 2013). In all four strains, some protein expressions decreased, especially two with approximate molecular weights of 55 and 65 kDa. In contrast, some expressions increased dramatically, especially those with molecular weights of 9, 14, 20, 34, 36, 45, 60, and 70 kDa (see

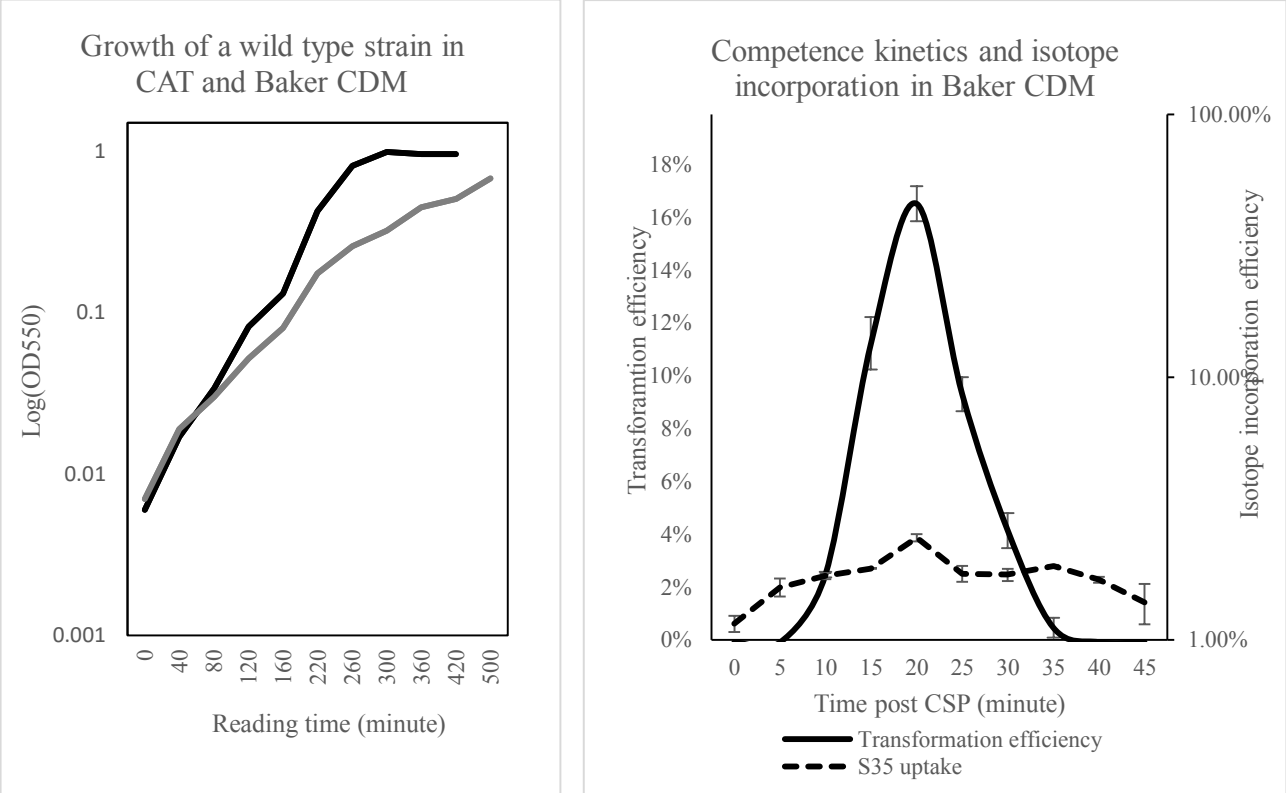


Figure 25. Bakers CDM can be used for pulse label of protein expression and competence. A (Left panel), growth patterns of a wild type strain was monitored by culture optical density at 550 nm from 0 to 500 minutes in CAT and Bakers CDM. Black: CAT; grey, Bakers CDM. B (Right panel), competence kinetics in a wild type strain in Baker CDM was tracked by the number of NovR transformants/mL after culture was transformed to a donor Nov^R DNA at various time points. At parallel, culture taken at indicated time points were exposed to S35-methionine for successive 3 minutes, and isotope incorporation efficiency was calculated by the formula described in methods, and all values were plotted over time.

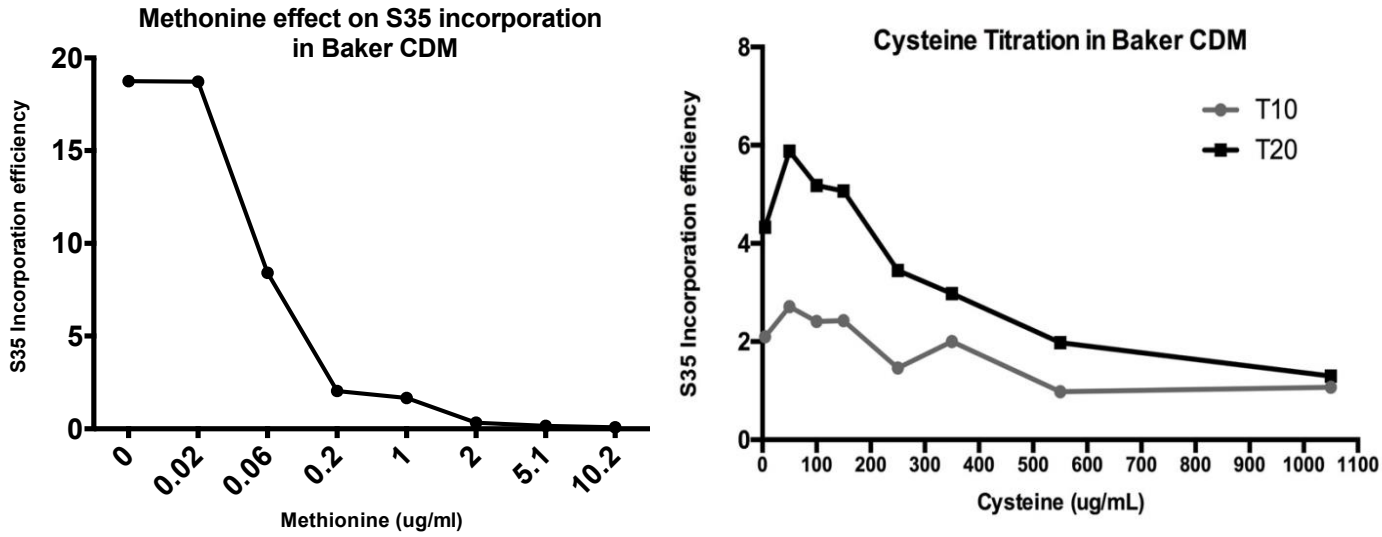


Figure 26. Effects of methionine and cysteine on efficiency of S35-methionine incorporation. A (Left panel), effects of non-labeled methionine on S35-methionine incorporation. B (Right panel), effects of cysteine on S35-methionine incorporation. Cysteine was added into Baker CDM with the indicated amount ranging from 0 to 1100 ug/mL. The incorporation efficiency was calculated following standard procedure.

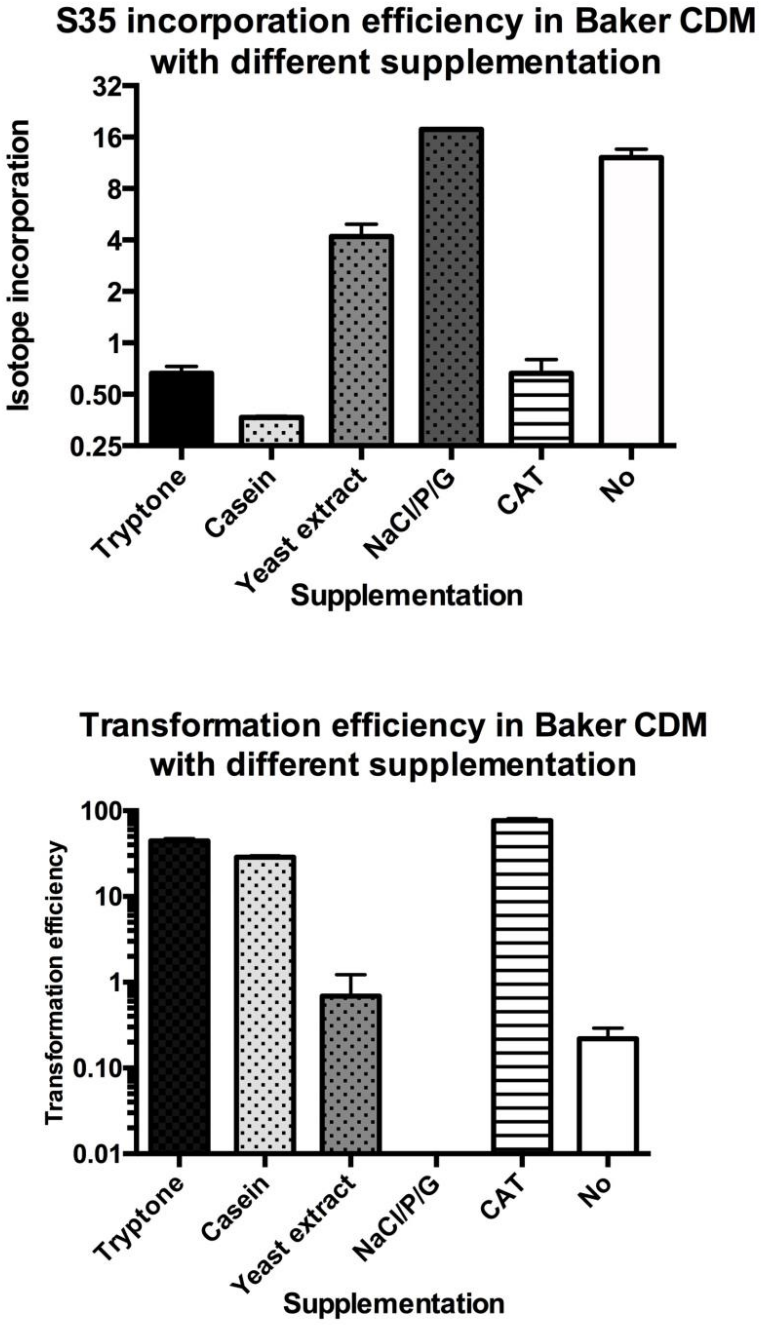


Figure 27. Effects of supplements on the efficiency of transformation and S35-methionine uptake. Effect of tryptone, casein, yeast extract, NaCl/P/G, CAT, and no addition was examined in pulse-chase labeling assay (Top) and transformation assay (Bottom). 2% (w/v) of each was added.

Figure 31). Transformation efficiency determined as CFU transformed by a rifampicin-resistant PCR DNA was 40% in the Δ*Hpf*Δ*yhaM* mutant compared to the wild type (see Figure 30).

Moreover, to objectively evaluate the protein pattern, more than five independent 1D gels were analyzed by comparing intensities of bands consistently showing differences (see Figure 32). Intensities of samples collected at three-time points—T0', T20', and T40'—were normalized and analyzed in the wild type and the Δ*Hpf*Δ*yhaM* mutant. The intensities of protein bands with 14, 34, and 36 kDa significantly increased at peak competence. In the wild type and Δ*Hpf*Δ*yhaM* strains, a protein band with 55 kDa significantly decreased in intensity at peak competence. A critical comparison of 1D images of the mutant and wild type showed slightly different protein expression patterns locally but not globally.

4.4.6 Similar protein expression patterns in Δ*Hpf*Δ*yhaM* and wild type on 2D gel

Two-dimensional gel electrophoresis is an effective strategy for determining a global protein shift. For each sample, about 100 µg/mL of protein was precipitated from acetone. A total of 350 mg of protein with 3 mcuries of S35-methionine was submitted to the Kendrick Company for 2D analysis. The newly synthesized proteins were visualized as radioactive spots in a pattern of several hundred spots. Spots on the 2D gel representing up- and down-regulated proteins during competence were labeled, and integrated signal intensities were analyzed using Image J. A wild type population was divided into two: one treated with CSP and the other untreated. Followed by pulse-chase labeling of competence-specific proteins, hot spots on the two 2D gel images were compared and analyzed (see Figure 33). Significantly different protein spots identified on the 2D gels of the CSP treated and the untreated cultures indicated differentiated expression of at least 29

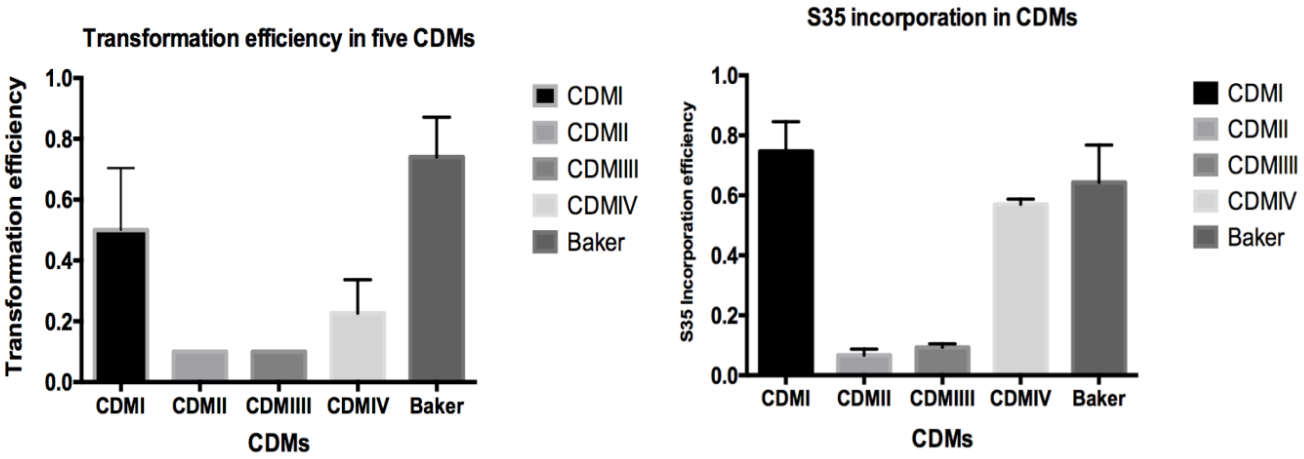


Figure 28. Comparisons of CDM in the transformation and pulse-labeling assays. A (Left). Transformation efficiency was monitored by transforming a NovR donor DNA in a wild type strain. Transformation efficiency in five media: CDM-I, CDM-II, CDM-III, CDM-IV and Baker CDM was compared. Error bar represented at least three independent assays. B (Right). In the pulse-labeling assay, the efficiency of S35-methionine incorporated into cells in five media was compared.

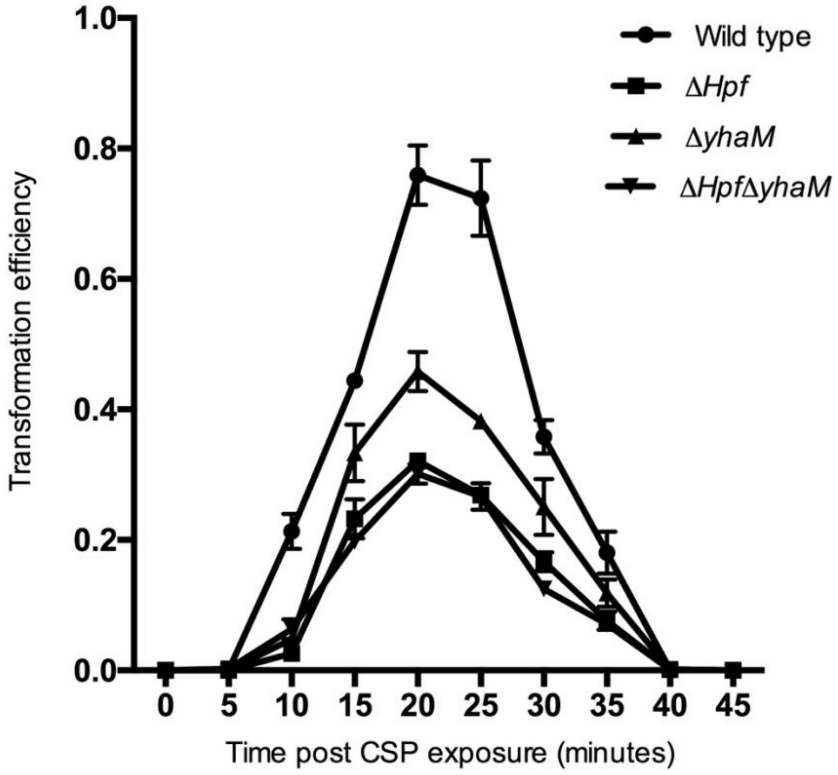


Figure 29. Competence is impaired by Δ*Hpf*Δ*yhaM* in the modified Baker CDM. Each of the four strains of wild type, Δ*Hpf*, Δ*yhaM*, and Δ*Hpf*Δ*yhaM* were transformed by exposing to a novR PCR DNA for 3 minutes. The transformation efficiency was determined by the number of transformants divided by the number of cell in a parallel culture. The competence kinetics assay was performed in modified Baker CDM, and the transformation efficiency of each sample collected at the 5-minute interval was calculated and plotted over times.

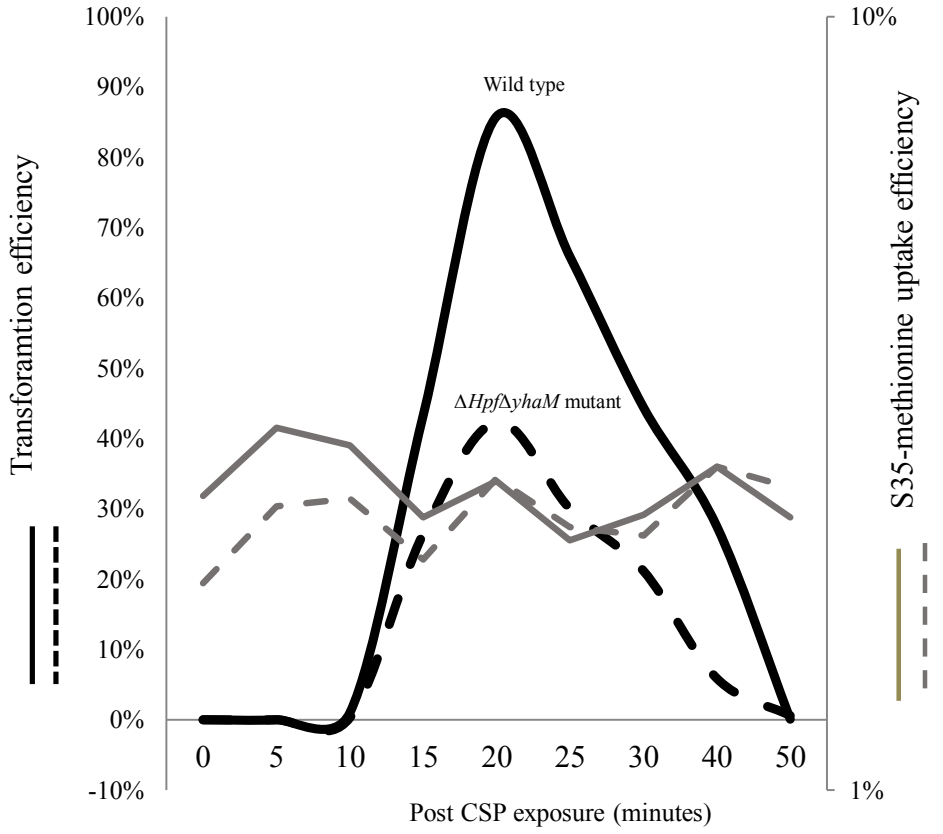


Figure 30. Combined assays of competence and pulse-chase labeling proteins with S35-methionine. The mutant Δ*Hpf*Δ*yhaM* showed a similar pattern of competence development but lower level with a wild type strain. S35-methionine incorporation efficiency was collected and calculated before loading protein samples on a PAGE gel. Competence kinetics are in black lines: solid, wild type; dashed, the mutant. The 2nd Y-axis represents S35-methionine incorporation efficiency. Time course of isotope incorporation is in grey lines: solid, wild type; dashed, the mutant.

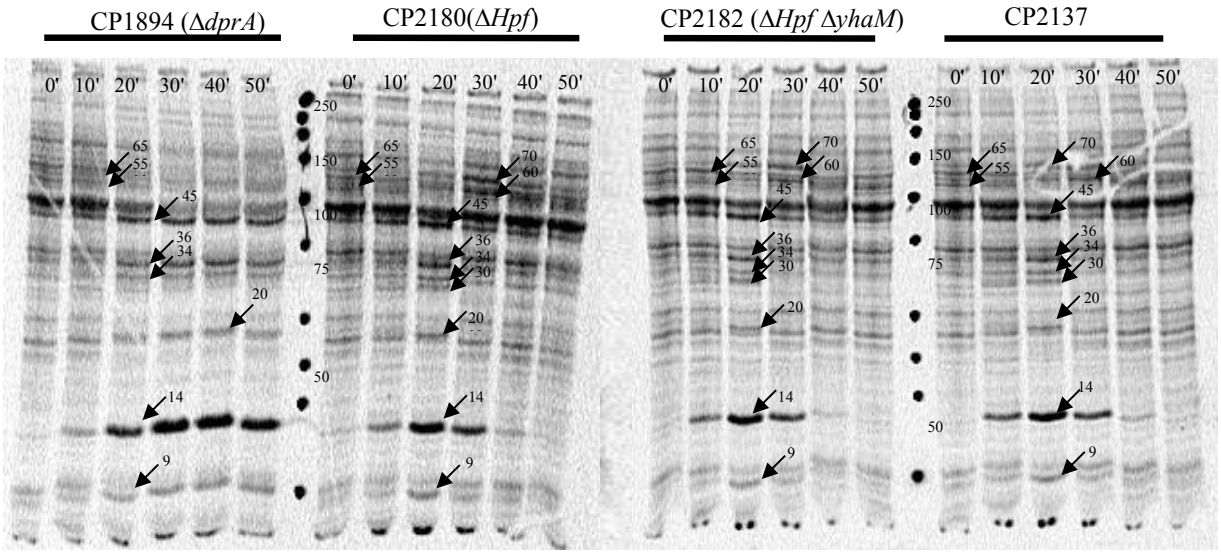


Figure 31. The protein expression patterns of Δ*Hpf*Δ*yhaM* and wild type strains on 1D gels
The S35-methionine pulse-labeled proteins from strains CP1894 (Δ*dprA*), CP2180 (Δ*Hpf*), CP2182 (Δ*Hpf*Δ*yhaM*), and CP2137 (wild type) were loaded on 1D-SDS gels. Critical comparisons of samples across a complete competence cycle from 0 to 50 minutes were done with all four strains. Arrows indicate the expression of proteins having a significant difference between peak competence and pre- and post-competence.

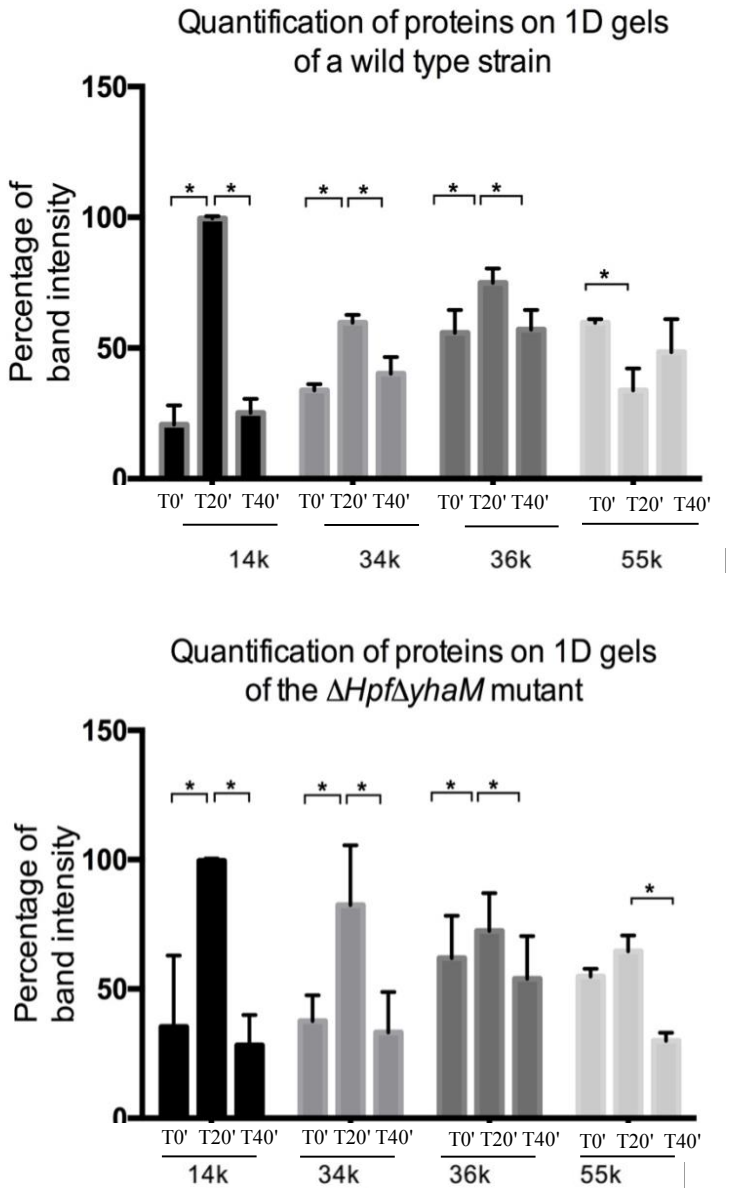


Figure 32. Quantification of proteins on 1D-PAGE. Proteins with a significant difference between peak competence (T20′) and pre- and post-competence were (T0′ and T40′) were analyzed. Amount of each protein was quantified by measuring the intensity of the band by image J. Four proteins with a significant difference during the cycle of competence were found in wild type and Δ*Hpf*Δ*yhaM* stain. Error bar represented at least five 1D gels of each culture and significance was with $p<0.05$.

proteins were differentiated: 15 proteins were downregulated especially 55 kDa, and 14 proteins were upregulated especially 14, 34, and 36 kDa.

Critical comparison of protein patterns on 2D gels of Δ*Hpf*Δ*yhaM* and wild type showed slight differences (see Figure 34). Most protein spots were similar, but at least six proteins appeared to be different. One spot with an estimated size of 34 kDa present on the wild type gel but missing from the Δ*Hpf*Δ*yhaM* gel was speculated to be YhaM (34 kDa). Two other spots on the Δ*Hpf*Δ*yhaM* gel showed decreased expression patterns compared to the wild type gel, suggesting the two proteins were HPF and/or YhaM-dependent. Three other spots on the Δ*Hpf*Δ*yhaM* gel showed increased expression patterns under Δ*Hpf*Δ*yhaM* background, suggesting that the three proteins may be have been repressed by HPF and/or YhaM. However, during competence, HPF and YhaM proteins did not affect global protein patterns during competence. Our results showed only a local effect of HPF and YhaM on patterns of six sepcific proteins, suggesting that HPF and YhaM may not be responsible for a broad reprogramming of protein synthesis. However, they may have had additional and more specific effects that accounted for the 60% transformation efficiency loss in the mutant.

4.5 Discussion

HPF is generally related to stress conditions such as HPFec and is starvation-specific in *E. coli*. In γ-proteobacteria, the 100S dimer is observed in only the stationary phases. However, HPFsa is present throughout all growth phases in *S. aureus*. Under nutrient-limited conditions,

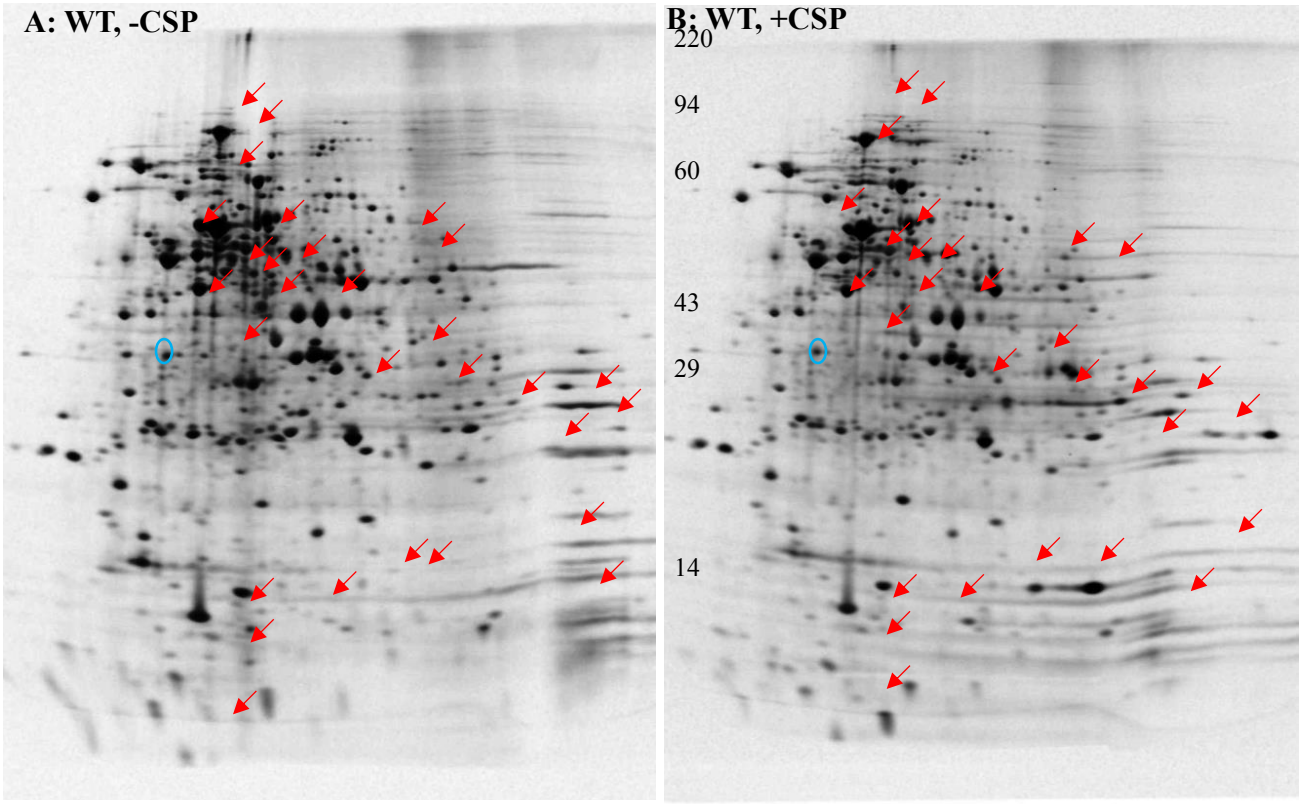


Figure 33. The protein expression patterns response to CSP induction on 2D gels. Protein expression patterns were compared in samples collected from the culture without and with exposure to CSP (A and B respectively) at 20 minutes after CSP addition. All red arrows indicate difference between the two gels. Blue circle represented an internal control tropomycin, with pI 5.2, and molecular weight 33 KDa.

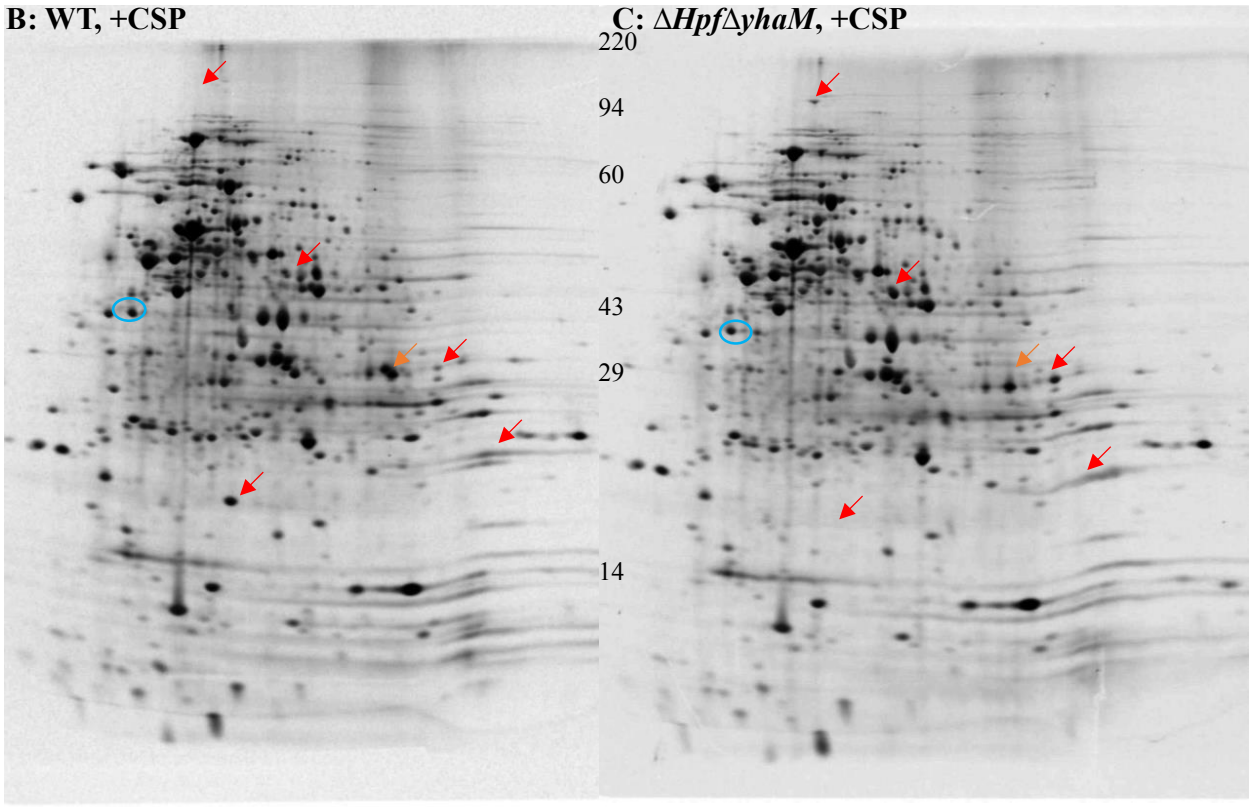


Figure 34. The protein patterns of wild type and Δ*Hpf*Δ*yhaM* strains. Samples were collected at 20 minutes after adding CSP that was at peak competence. Red and orange arrows were used to indicate a difference between wild type and Δ*Hpf*Δ*yhaM* strains. Blue circle represented an internal control tropomycin, with pI 5.2, and molecular weight 33 KDa. Orange arrow indicated speculated YhaM based on its size.

HPFsa was found to suppress translation initiation for stress adaptation, playing a role in ribosome integrity through dimerization of ribosomes (M. Ueta et al., 2010).

In this study, transformation efficiency was decreased to 36% in the Δ*Hpf*Δ*yhaM* mutant and 54% and 37% in the single mutants Δ*Hpf* and Δ*yhaM*, respectively. Expression of HPF is upregulated in response to CSP-induced competence in all streptococcal groups. Based on sequence homologs with identified proteins, their predicted functions in regulating competence are related to ribosome activity. No global protein pattern shift was observed, but a few bands were slightly different in the Δ*Hpf*Δ*yhaM* mutant. It seems that HPF and YhaM did not play a role in reprogramming global protein expression. Although HPF and YhaM may not contribute at the translational level, subtler roles may remain to be discovered. Competence is not considered to be a stress condition, and this may apply to the scenario if HPF and YhaM are stress-specific in *S. pneumoniae*.

In addition, the fact that we did not observe a global protein pattern shift may be attributable to samples' not being harvested at precise time points or the fraction of competent cells might have been significantly less than 1 when 2D-PAGE was applied. Samples were incubated for 20 minutes after CSP exposure, the time expected for typical peak competence. However, the timings for development of competence may have varied day-to-day by a few minutes depending on the cell-wash process, the pre-warming process after cold treatment, and the growth conditions on individual days of the experimental period. In addition, isotope incorporation efficiency was not high even though modified Baker CDM was used. In attempting to achieve sufficient signals on 1D radiographs, overly saturated protein samples may have been used, resulting in a compromised sensitivity of intensity change of protein bands. If such issues can be resolved, a 2D analysis could be applied to study protein patterns. Specifically, spots on a 2D gel with consistent ratios below or

above a specific value could be cut from the gel using a spot picker robot. After protein extraction by tryptic digestion, the resulting peptides could be analyzed by mass spectrometry. Alternatively, to specifically determine the role of HPF in the expression of six proteins showing differentiated patterns in the mutant background, *in vitro* translation could be used. In this assay, template DNA from the six genes could be incubated with reconstituted recombinant components including 0.3mM amino acid, in the PURExpression system at 37^{o}C for 2 hours. The newly synthesized protein could be labeled with 1.2μM S35-methionine, and protein products could be detected by autoradiograph. To understand the general role of HPF in expression of proteins globally, protein profiling could then be applied. Actively translated mRNA could be tracked after addition of chloramphenicol and digestion with nuclease, and the results could be compared in the mutant and wild type backgrounds. Any observed differences in gene translation could be used to represent the effect of HPF on protein synthesis.

It is still unclear what mechanism regulates decreased mRNA as well as the down-regulated protein pattern revealed on 1D images. One set of the three repressed mRNAs showed minimal mRNA coinciding with the maximum amount of σ^{X} protein, and most of the genes belonged to a cluster of ribosomal protein genes. It is possible that HPF and/or YhaM played a role along with unknown factors in repression of mRNA in *S. pneumoniae*. Clues regarding competence regulation may be provided by the additional seven genes in the core regulons: SP2044 (*ackA*, bifunctional acetaldehyde-coenzymeA/alcoholdehydrogenase), SP1941 (*cinA,* competence damage-inducible protein A), SP0021 (*dut,* deoxyuridine 5'-triphosphatenucleotide hydrolase), SP0979 (*pepF,* oligopeptidase), SP0782 (*pilC,* membrane protein of pilus assembly), SP1088 (*radC,* DNA repair protein), and SP1981 (*rmuC,* DNA recombination protein). These are collectively referred to as the genome sequence of strain TIGR4 (Khan et al., 2016). Some of the genes include orthologs of

well-characterized proteins, and some have domains that fall into known broad functional categories. To determine their roles in competence regulation, single and multiple deletion mutations of all nine genes could be carried out. For each mutant with a global effect on protein synthesis, protein samples from induced and uninduced cultures generated in parallel could be compared and comprehensively evaluated by means of 2D electrophoresis. Moreover, ribosomal structure change, mRNA and rRNA structure, and enrichment change could be examined in each mutant. Determination of mechanisms of competence regulation in any of the nine genes would have broad application in Streptococci and other bacteria having a homolog.

Table V. Comparisions of amino acids in different CDMs

Amino Acid Components						
(µg/mL)	CDM-I	CDM-II	CDM-III	CDM-IV	Baker NM	Complete NM
DL-Alanine	56.9			56.9		100.0
L-Arginine	200.0	298.5	3284.0	200.0	200.0	100.0
L-Asparagine	10.0	11.2	40.0	10.0	10.0	100.0
L-Aspartic acid	35.5			35.5		100.0
L-Cysteine					140.0	625.0
L-Cystine						50.0
L-Glutamic acid	141.3			141.3		100.0
L-Glutamine	30.0	30.2	7.0	30.0	30.0	200.0
Glycine	16.7	16.7		16.7		100.0
L-Histidine	150.0	193.1	1438.0	150.0	150.0	100.0
L-Isoleucine	6.6	231.5	7495.0	231.5	6.6	100.0
L-Leucine	228.2			228.2	6.0	100.0
L-Lysine	420.0	614.0	6466.0	420.0	420.0	100.0
L-Methionine						
L-Phenylalanine	50.0	157.6	3586.0	157.6	50.0	100.0
L-Proline	50.0	82.7	1090.0	50.0	50.0	100.0
Hydroxy-L-Proline						100.0
L-Serine	120.0	209.9	2996.0	120.0	120.0	100.0
L-Threonine	195.0	262.7	2256.0	195.0	195.0	200.0
L-Tryptophan	50.0	77.2	905.0	50.0	50.0	100.0
L-Tyrosine	50.0	94.3	1477.0	50.0	50.0	100.0
L-Valine	6.0	125.1	3970.0	125.1	6.0	100.0

V. Construction of a pneumococcal host strain to identify Δ*comW* suppressor mutations

5.1 Abstract

ComW is required for full activity of σ^X; it appears to protect σ^X from proteolysis and was hypothesized to act as an anti-anti-σ^X. The aim of this project was to construct a facultative host strain for selecting novel Δ*comW* suppressor mutations. By introducing a raffinose promoter locus P_{aga}*::comX::comW* into a wild type strain, the strain CP2158 could be induced into competence by 0.2% raffinose with high transformation efficiency (10^6 CFU). However, leaky expression of the raffinose promoter under non-permissive conditions was observed. The leaky expression can be suppressed by expression of *clpC* controlled by fucose inducible promoter, resulting in strain CP2173 (*P_{aga}::comX::comW, ΔclpC*::tet^R; *ΔcomW::P_{fcsk}::clpC*::spc^R) with low transformation efficiency (10^2 CFU) under non-permissive condition. Therefore, CP2173 can be used as a highly transformable recipient for a backcross of putative suppressor mutations from their original isolation background. Subsequently, transformants can be evaluated for suppression of the Δ*comW* phenotype under non-permissive conditions.

To seek a possible Δ*comW* suppressor mutant, we used a marinerT7 transposon mutant library (T7::spc^R) to screen for mutations that restore competence to the Δ*comW::kan^R* strain. We performed two successive library screening steps to enrich suppressor mutations from the library, followed by two independent transforming cycles for selecting potential individual suppressor strains. Suppressor strains backcrossed and recovered from spc^Rkan^R plates were tested for competence by transforming a pure PCR donor DNA. Progeny from each backcross were examined for restoration of a competence phenotype. Interestingly, most mutations were linked to *ΔcomW::kan^R*, suggesting that mariner T7 transposon insertions may have disrupted essential genes and produced predominate null mutations. Surprisingly, sequencing result revealed that all

competence restored mutants having Δ*comW::kan*^{R} inserted ectopically into genome, the mechanism of which was poorly understood. An alternative screening approach should be used for introducing single base substitutions to allow recovery of subtler bypass mutations instead of such stringent screening.

5.2 Introduction

Streptococcus pneumoniae is naturally competent, with its genomes encoding all proteins necessary for DNA uptake and recombination. During transformation, several genes are responsible for competence, two of which are *comX* and *comW*. Competence is induced by a mechanism called QS that allows cells within a population to communicate. The competence regulation pathway consists of ComAB and ComCDE, which are responsible for importing and sensing the pheromone secreted by neighboring cells, which is called CSP (Havarstein et al., 1995). These proteins also stimulate internal changes such as transcription of gene *comX,* which encodes an alternative σ factor. Genetic transformation is stimulated by σ^X, thus allowing transcription of late genes during competence. The σ^X is the first link between QS and induction of competence-specific genes (Lee & Morrison, 1999). While σ^X is essential for the competence response, it alone is not sufficient for full competence (Luo & Morrison, 2004). Thus, there must be another competence-specific gene expressed that plays a significant role in transformation. The gene *comW* is a CSP-dependent regulator that is necessary for accumulation of the σ^X and thus for development of competence. The mechanism by which ComW promotes cell competence is still under investigation (Luo et al., 2004; Sung & Morrison, 2005), but one hypothesis is that ComW protects σ^X from proteolysis by interacting with an anti-σ factor. As in *B. subtilis*, σ^F reaches a peak level about 2-fold higher than that of σ^A after sporulation, but the affinity of the core for σ^F is 25-fold

lower than that for $\sigma^{A,}$ which suggests that an anti-σ^{A} exists (Lord, 1999). In *S. pneumoniae*, the molar concentration of SigX was almost 2-fold that of RNAP, but affinity purification showed a molar concentration of SigX that was 0.67-fold that of RNAP, so it is possible that an anti-σ factor exists.

To test this hypothesis, a transposon mutant library was generated using suppressor mutation screen in the *ΔcomW* background, where competence is severely deficient. It is well known that suppressor mutations are those that alleviate or revert the phenotypic effects of an existing mutation selectively generated under experimental conditions. Screening suppressions are used to identify point mutations that affect a biological process of interest. Mutants with a restored competence phenotype were collected and enriched several times to identify a possible anti-σ factor. Mutants were expected to have insertions that disrupt the putative anti-σ factor, allowing restoration of a competence phenotype to a *ΔcomW* strain.

In addition, to enrich the pool of antibiotic resistance markers during compound transformation assays, a rifampicin-resistance maker was generated by spontaneous mutations. Since 1967, rifampicin has been used in combination with other antibiotics to treat pneumococcal infections, tuberculosis, and many other diseases. However, rifampicin-resistance has been developed in many pathogens. Several mechanisms have been proposed to account for rifampicin resistance, including rifampicin uptake, modification, and metabolism. Bacterial resistance to rifampicin is due to point mutations in *rpoB*, which encodes the β subunit of RNA polymerase (RNAP) (Aubry-Damon, Soussy, & Courvalin, 1998; Jin & Gross, 1988). Rifampicin acts by binding to the β subunit and causing the premature termination of DNA transcription. Rifampicin-resistance in *S. pneumoniae* has been related to *rpoB*, and three domains have been identified in *rpoB*: I (amino acids 406 to 434), II (amino acids 560 to 572), and III (amino acids 573 to 600)

(Enright, Zawadski, Pickerill, & Dowson, 1998; Meier, Utz, Aebi, & Muhlemann, 2003; Padayachee & Klugman, 1999). The sequence assignment was based on the sequence of *rpoB* in *E. coli* related to rifampicin-resistance. Rifampicin acts by binding to the β subunit of RNAP, resulting in premature termination of DNA transcription (Campbell et al., 2001). Mutations conferring high levels of resistance are frequently located in a binding pocket of the β subunit of RNAP, and lower levels of resistance are away from this binding pocket.

To determine the role of ComW in regulation of competence, a *ΔcomW* suppressor screen was carried out by means of MarinerT7 mutagenesis. The competence restoration mutation in the *ΔcomW* background was hypothesized to be a possible anti-σ factor. To select this type of mutants, a facultative host strain was constructed to conditionally express SigX and ComW conditionaly. In addition, a highly efficient donor rifampicin resistant donor DNA was generated by spontaneous mutation.

5.3 Methods

5.3.1 DNA ligation by Gibson assembly

Gibson assembly can ligate up to 6 DNA fragments in one single isothermal reaction to create one seamlessly assembled DNA (D. G. Gibson et al., 2009). By designing primers with 20 extra bases at the end with homologous overlapping regions between fragments. 0.2-1pmols of each of the PCR amplified fragments were incubated simultaneously at 50^{o}C for 60 minutes. Then the assembly product was transformed into *E. coli*. Sanger sequencing was used to verify correct assembled DNA.

To generate a DNA containing Δ*comW*::P_{fcsk}::*clpC*:spc^{R}, four fragments were ligated with a Gibson assembly kit (NEB). Each fragment was PCR amplified first: 1003bp-*comW* upstream

(*transposase B*) amplified from CP1759 genomic DNA with primers DAM1089 and DAM1090; 692bp-T1T2-fucose promoter amplified from CP2055 genomic DNA with primers DAM1091 and DAM1092; 2613bp-*clpC* fragment amplified from CP1759 genomic DNA with primers DAM1093 and DAM1094, and 1515bp-*comW* downstream (*purA*) amplified from CP1759 genomic DNA as template by primers DAM1095 and DAM1096. Then 2000 ng of a 5 kb Gibson DNA in one reaction which was around 0.5 pmols was generated. The assembly DNA was examined on DNA electrophoresis gel and confirmed by Sanger sequencing of the homologous regions during ligation.

5.3.2 Construction libraries of MarinerT7 mutants

A genome-wide disruption library in *S. pneumoniae* with the MarinerT7 transposon was constructed using a mini-transposon, which contains a MmeI recognition site that can be recognized by MelI and cut downstream DNA 20 bp away. The mini-transposon was mixed with pneumococcal genomic DNA *in vitro* and then used to transform a bacterial population (Bijlsma et al., 2007; Burghout et al., 2007; Plasterk, Izsvak, & Ivics, 1999). After cleavage by MmeI restriction nuclease, a DNA library containing 16 bp of pneumococcal genome DNA was generated. Then the library was ligated to an adapter of known sequence to serve as DNA barcode sequences, which enables next generation sequencing. A PCR amplification was used to get a 160 bp sequence with 16 bp of pneumococci genomic DNA flanked by sepcific sequences, based on which different DNA could be identified, and the 16 bp reads could be mapped to the genome.

5.3.3 Enrichment of libraries

The first library, WL1 (Tn::spc^R), containing possible *comW* suppressor mutants was created by transforming the MarinerT7 transposon library into a wild type strain. 10^5 CFU spectinomycin plates were collected as a pool through resuspension of all colonies in 1mL CAT

broth incubated in 5% CO_2. The whole library was grown overnight to produce a working pool stock, and then transformed with *ΔcomW::kan*^{R} DNA, and around 10^6 transformants were rescued, named as WL2 (Tn::spc^R, *ΔcomW::kan*^{R}). To obtain an enriched suppressor mutant library, the pool of kan^R spc^R transformants grew to OD_{550} 0.1 and was exposed to CSP and nov^R PCR DNA as described from the competence assay. All libraries were stored at -80°C with 15% glycerol. The recovered library pool is estimated to contain 20,000 *ΔcomW* suppressor mutations.

5.3.4 Estimation of spontaneous mutation frequency

Estimation of spontaneous mutation with low- and high-efficiency present in a mutant library was obtained by determining the number of trimethoprim-resistant mutants as described (Sung and Morrison, 2005). Trimethoprim-resistance was conferred by inactivation of proteins of thymidine synthetase and/or Ami transporters. Estimation of spontaneous mutation occurring with a low efficiency was determined by the number of rifampicin-resistant mutants as described later. A small amount of wild type cells (100 µL) at OD_{550} 0.2 was plated on CAT medium containing 100 to 200 µg/mL trimethoprim or 10 µg/mL rifampicin, and mutants were selected and counted to estimate the mutation frequency. The frequencies of trimethoprim-resistance and rifampicin-resistance mutations occurred were 10^{-4} and 10^{-6}, respectively.

5.3.5 Selection for mutants with various rifampicin-resistance

Pneumococcal cells were incubated until OD_{550} 0.2 (10^8 CFU) in CAT supplemented with glucose and phosphate. Spontaneous mutant strains were randomly selected by plating 10^8 wild type pneumococcal cells directly on CAT agar plates containing rifampicin ranging from low to high levels of rifampicin: 0.5 µg/mL, 4 µg/mL, and 10 µg/mL. 20 mg/mL rifampicin stock was made by dissolving rifampicin in methanol. Then 5 mL of pneumococcal cells were plated, and several single colonies were detected after 48 hours' incubation at 37°C. Several individual

colonies were picked up and outgrew in CAT to OD_{550} 0.2, then made into frozen stock with 15% glycerol at stored at -80°C for further use.

5.3.6 PCR amplification and sequencing of *rpoB* mutations

An 8.5kb fragment containing *rpoB* gene was amplified from genomic DNA with primers FP-r1 and RP-r1 and 5.9 kb fragment containing *rpoB* with primers FP-r2 and RP-r2. PCR reaction was carried out in 50 µL including 5 µL of each primer, 10 mM dNTP, 10 µL 5X Phire Buffer and 1 µL of and Phire Hotstart II polymerase, 100 ng of template DNA, and dH2O brings up to 50 uL. Reactions were performed in following condition: 98°C for 30 seconds, 30 cycles for 98°C for 10 seconds, 55°C for 30 seconds, and 72°C for 2 minutes, and a final extension at 72°C for 10min. PCR DNA was purified with Zymo PCR cleanup and purification kit (Zymo Research). The *rpoB* gene from different clones was identified by Sanger sequencing. Mutation frequency was determined by the number of rifampicin-resistant single colonies divided by the number of viable cells on sandwich plates (cells were embedded in CAT agar) described by Sung and Morrison (2015).

5.4 Results

5.4.1 Construction of Δ*comW* suppressor mutation library by MarinerT7 transposon

To objectively determine the role of ComW in the regulation of competence, a suppressor screening for rare mutations was carried out to rescue a Δ*comW* phenotype. Strain CP2137 (Δ*comA)* was transformed with DNA from the MarinerT7 library containing spectinomycin-resistance (Tn::spc^R) (obtained from P. Burghout). The transformation resulted in the random insertion of the DNA into each *S. pneumoniae* R6 genome (Bijlsma et al., 2007; Burghout et al., 2007) (see Figure 35). The DNA from MarinerT7 library contains the inverted repeat sequence specifically

recognized by endonuclease MmeI that cleaves downstream DNA 20 bp away from its recognition site, resulting 16 bp of pneumococcal genomic DNA included in MmeI digested DNA fragments (the recognization site located 4 bp before the end of the terminal repeat). Therefore, MmeI digested DNA that is rescued from the suppressor screening can be PCR amplified with a MmeI recognition primer and an adapter primer, resulting 120 bp product that can be sequenced in parallel (see Figure 36). The transposon library was expected to contain 20,000 unique transposon mutants per mL to give 10 times coverage of the genome, which means each gene (2,000 genes estimated in one genome) has ample (10-time) opportunity to acquire the antibiotic resistance. The library was transformed into a wild type strain at OD_{550} 0.2 (2 x10^8 CFU). The transformation rate obtained with the transposon DNA was 1%, resulting to 2 x10^6 CFU. Thus 100-time of the transposon library in the total of spc^R transformants. All 10^6 transformants were collected as a pool named WL1. Then the pool was grown to OD_{550} 0.1 (10^8 CFU) and transformed with a Δ*comW::kan*^{R} DNA, resulting in 10^6 transformants (*ΔcomA, ΔcomW*::kan^R, spc^R) collected as a pool named library WL2. The WL2 was grown to OD_{550} 0.1 (10^8 CFU) and then was induced to competence and transformed with a PCR DNA to maximally enrich potential *ΔcomW* suppressor mutants. 10^3 transformants were obtained on novobiocin selection plates which was consistent with the expected number that by consideration of 0.01% efficiency of a *ΔcomW* strain and 10% efficiency of a nov^R PCR donor DNA. At this point, the transformants were expected to be a mixture of cells lacking a competence phenotype and cells that were potentially restored to competence by virtue of a Δ*comW* suppressor (see Figure 37 and Table VI). The population of *ΔcomW* suppressor mutants was enriched by a round of transformation of the entire pool of WL2, with at least 50% of the progeny containing potential Δ*comW* suppressors. Twelve individual nov^R clones were collected and named WS1-12 (Δ*comA*; $spc^R kan^R nov^R$).

5.4.2 Identification of a Δ*comW* suppressor by determining competence linkage to Tn::spc^R

After several successive transformation assays were used to screen Δ*comW* suppressor mutations, 12 individual clones (WS1-12) were examined for competence by transforming an erm^R donor DNA. Colonies showing a transformation efficiency higher than a Δ*comW* mutant were defined as transformable clones. The results showed that 10 of the 12 WS clones were transformable (see Table VI). All the 10 transformable strains were grown out and their DNAs were used as donors to backcross a wild type strain.

In the backcross approach, 100 ng genomic DNA from each of the 10 strains were transformed into wild type strain CP2137. Resulting transformants were selected on spectinomycin and kanamycin plates to isolate transformants containing both Δ*comW::kan^R* and Tn::spc^R markers inherited from the transposon library. Two individual nov^R subclones from each of the 10 WS strains were picked up and subjected to a competence assay to study the linkage of competence to spc^{R}kan^R; the resulting transformants were named BWS1-20 (see Figure 38).

To assess transformability of those BWS strains, an erythromycin-resistant (erm^R) DNA was used as a donor DNA in transformation assays. Depending on the nature of the donor strain's specific suppressor, three possible patterns were expected from backcross progeny: (a) all spc^{R}kan^R progeny are competent because Tn::spc^R insertion was itself a suppressor; (b) no spc^{R}kan^R progeny are competent because the suppressor was not linked to either marker; and (c) spc^{R}kan^R progeny are of mixed phenotypes, some showing competence and some not because the suppressor was not spc^R but was linked to the kan^R or spc^R marker. In comparing the transformation efficiency of backcrossed progeny to that of a Δ*comW* mutant, 11 progeny were transformable, and 9 progeny were not (see Table VII). Thus, backcrosses results were of type c that some progeny showed competence while others did not, suggesting that

Table VI. Transformation efficiency of libraries and subclones

Culture[a]	Transformation efficiency[b]		Relative efficiency[c]
	CFU Nov^R	CFU Erm^R	
Library WL1	10^6	-	1%
Library WL2	10^3	-	0.001%
WS 1	10^6	10^6	10%
WS 2	10^6	10^6	10%
WS 3	10^6	10^6	10%
WS 4	10^6	10^6	10%
WS 5	10^6	10^6	10%
WS 6	10^6	10^6	10%
WS 7	10^6	10^6	10%
WS 8	10^6	10^6	10%
WS 9	10^6	10^6	10%
WS 10	10^6	10^6	10%
WS 11	10^6	10^3	0.01%
WS 12	10^6	10^3	0.01%
Wild type (CP2137)	10^6	-	10%
CP2166 (Δ*comW*)	10^3	-	0.01%

a. Pool or sub-clone in screen.
b. Transformation efficiency: Novobicin or erythromycin-resistant CFU/ml, determined in triplicates.
c. Relative efficiency: Transformation efficiency of a pool or sub-clone compared to a wild type strain.

the Δ*comW* suppressor was in fact not linked to Tn::*spc*^{R} but might be linked to the kan^R or spc^R markers.

To address this issue, a second backcross was performed by transforming strain CP2137 with genomic DNA of 8 WS strains but transformants were challenged with a single drug and in parallel with double drug selections. A rigorous comparison of competence level between progeny in the two parallel selection conditions was used to establish whether the competence phenotype was linked to the Δ*comW*::kan^R or Tn::spc^R. In this second backcross approach, 8 WS strains (WS 13-20) were backcrossed to a strain CP2137, but the resulting transformants were selected on kanamycin or kanamycin and spectinomycin plates. Kanamycin selection of transformants was used to determine the linkage of competence to Δ*comW*::kan^R. To do this, transformability of each backcross clone was studied by transforming a pure nov^R PCR donor DNA and selection of the transformants (spc^R kan^R nov^R) on novobiocin agar (see Figure 38). All eight WS strains were competent, and this was consistent with the results that transformants selected on the double antibiotic resistance plates. Next, five kan^R clones were picked up from each backcross and subjected to another round of transformation using erm^R DNA and transformants selected on erythromycin plates (see Table VIII). In parallel, seven spc^{R}kan^R clones were examined in the same way by two rounds of transformation with nov^R and erm^R DNA to determine the linkage of competence to spc^{R}kan^R (see Table IX). Unexpectedly, most mutations were linked to *ΔcomW::*kan^R. Sanger sequencing was used to verify correct integration of Δ*comW*::kan^R into *comW* locus during transformation. To conclude, two categories of mutations were in fact enriched: 75% of the mutations were linked to the *ΔcomW::*kan^R and 25% of the mutations were linked to *comX2* (see Table X). The exact mechanism for the misinsertions was not well-understood.

Table VII. Competence of 1st backcross clones

	Competence of 1st backcross clones			Competence of 1st backcross subclones	
Suppressor Isolate[a]	Transformation efficiency (nov^R)[b]	Competence[c]	Suppressor Isolate[a]	Transformation efficiency (erm^R)[b]	Competence[c]
BWS 1	1%	yes	BWS 1-1	< 0.01%	no
			1-2	< 0.01%	no
BWS 2	1%	yes	BWS 2-1	< 0.01%	no
			2-2	< 0.01%	no
BWS 3	1%	yes	BWS 3-1	1%	yes
			3-2	< 0.01%	no
BWS 4	1%	yes	BWS 4-1	1%	yes
			4-2	2/0	yes
BWS 5	1%	yes	BWS 5-1	1%	yes
			5-2	< 0.01%	no
BWS 6	1%	yes	BWS 6-1	1%	yes
			6-2	< 0.01%	no
BWS 7	1%	yes	BWS 7-1	< 0.01%	no
			7-2	< 0.01%	no
BWS 8	1%	yes	BWS 8-1	1%	yes
			8-2	1%	yes
BWS 9	1%	yes	BWS 9-1	1%	yes
			9-2	1%	yes
BWS 10	1%	yes	BWS 10-1	1%	yes
			10-2	1%	yes
BWS 11	<0.01%	no	-	-	-
			-	-	-
BWS 12	<0.01%	no	-	-	-
			-	-	-

a. Subclone isolated in the screen.
b. Transformation efficiency was determined by novobiocin-resistant CFU/ml compared to a wild type strain.
c. Competence was determined by level of transformation efficiency compared to 0.01% of the Δ*comW* mutant.

Table VIII. Competence of 2nd backcross clones with kanamycin selection

Competence of 2nd backcross clones[a] - **Kan^R**					
Suppressor Isolate[b]	Transformation Efficiency[c]	Competence[d]	Suppressor Isolate	Transformation Efficiency[c]	Competence[d]
BWS 13-1	1%	yes	BWS 17-1	1%	yes
13-2	1%	yes	17-2	1%	yes
13-3	1%	yes	17-3	< 0.01%	no
13-4	1%	yes	17-4	< 0.01%	no
13-5	< 0.01%	no	17-5	< 0.01%	no
BWS 14-1	< 0.01%	no	BWS 18-1	1%	yes
14-2	< 0.01%	no	18-2	1%	yes
14-3	< 0.01%	no	18-3	1%	yes
14-4	< 0.01%	no	18-4	< 0.01%	no
14-5	< 0.01%	no	18-5	< 0.01%	no
BWS 15-1	1%	yes	BWS 19-1	< 0.01%	no
15-2	< 0.01%	no	19-2	< 0.01%	no
15-3	< 0.01%	no	19-3	< 0.01%	no
15-4	< 0.01%	no	19-4	< 0.01%	no
15-5	< 0.01%	no	19-5	< 0.01%	no
BWS 16-1	1%	yes	BWS 20-1	< 0.01%	no
16-2	1%	yes	20-2	1%	yes
16-3	1%	yes	20-3	1%	yes
16-4	1%	yes	20-4	1%	yes
16-5	1%	yes	20-5	1%	yes

a. Backcrossed progeny were selected on kanamycin plates.
b. Subclone isolated in the screen.
c. Transformation efficiency was determined by novobiocin-resistant CFU/ml compared to a wild type strain.
d. Competence was determined by comparision to the level of transformation efficiency of 0.01% in the Δ*comW* mutant.

Table IX Competence of 2nd backcross clones with kanamycin and spectinomycin

Competence of 2nd backcross clones - **Kan^{R}Spc^R**					
Suppressor Isolate	Transformation Efficiency	Competence	Suppressor Isolate	Transformation Efficiency	Competence
BWS 13-1	1%	yes	BWS 17-1	1%	yes
13-2	1%	yes	17-2	1%	yes
13-3	1%	yes	17-3	1%	yes
13-4	1%	yes	17-4	< 0.01%	no
13-5	< 0.01%	no	17-5	< 0.01%	no
13-6	< 0.01%	no	17-6	< 0.01%	no
13-7	< 0.01%	no	17-7	< 0.01%	no
BWS 14-1	< 0.01%	no	BWS 18-1	1%	yes
14-2	< 0.01%	no	18-2	1%	yes
14-3	< 0.01%	no	18-3	1%	yes
14-4	< 0.01%	no	18-4	1%	yes
14-5	< 0.01%	no	18-5	1%	yes
14-6	< 0.01%	no	18-6	1%	yes
14-7	< 0.01%	no	18-7	1%	yes
BWS 15-1	1%	yes	BWS 19-1	< 0.01%	no
15-2	1%	yes	19-2	< 0.01%	no
15-3	1%	yes	19-3	< 0.01%	no
15-4	< 0.01%	no	19-4	< 0.01%	no
15-5	< 0.01%	no	19-5	< 0.01%	no
15-6	< 0.01%	no	19-6	< 0.01%	no
15-7	< 0.01%	no	19-7	< 0.01%	no
BWS 16-1	1%	yes	BWS 20-1	1%	yes
16-2	1%	yes	20-2	1%	yes
16-3	1%	yes	20-3	1%	yes
16-4	1%	yes	20-4	1%	yes
16-5	1%	yes	20-5	1%	yes
16-6	1%	yes	20-6	< 0.01%	no
16-7	1%	yes	20-7	-	-

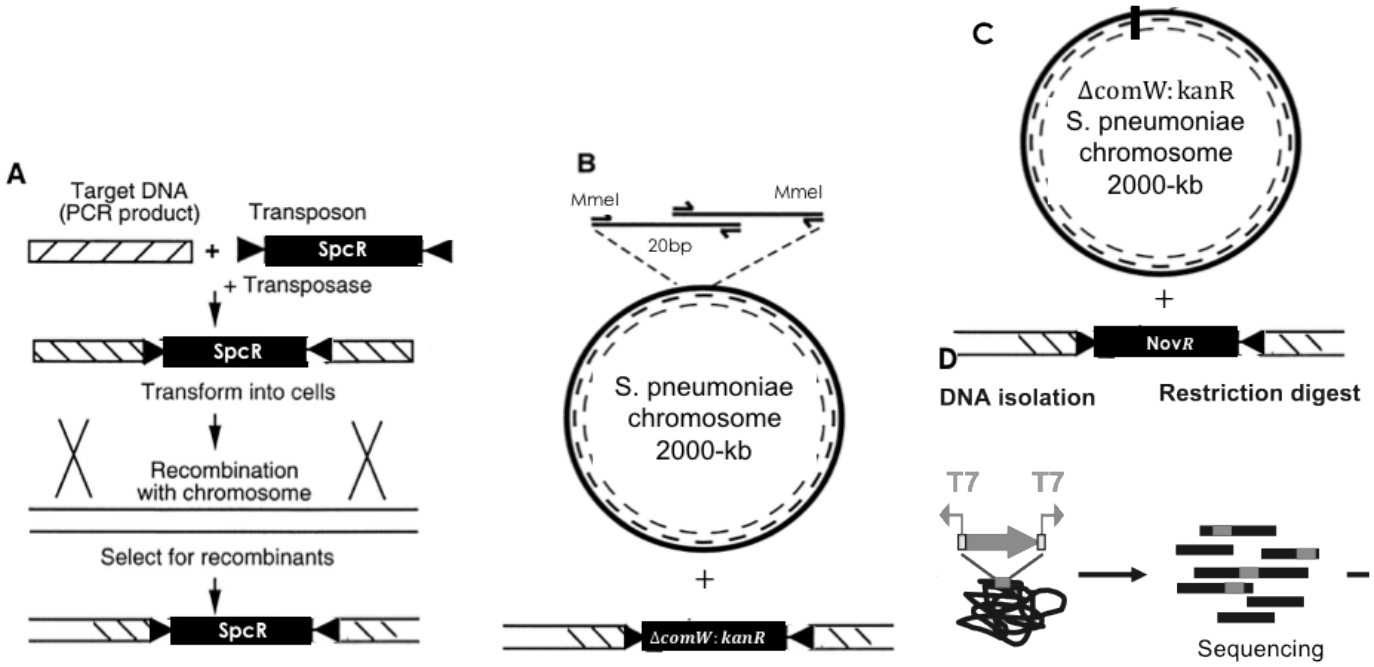


Figure 35. Schematic depiction of T7 transposon mutation library. A. A MarinerT7::spc^R transposon library was constructed by first transposing the transposon T7::spcR, which contains a MmeI restriction site with each inverted repeat into pneumococcal R6 DNA fragments. Then the mutant library was transformed to a wild type strain CP2137, which can randomly insert into genome after transformation. B. The library was transformed with a Δ*comW*::kan^R DNA to generate a Δ*comW* suppressor library. C. competence level of individual clones from the library was measured by novR transformants after transformation. D. strains with restored competence were analyzed by sequencing the 20 bp fragment near MelI sites.

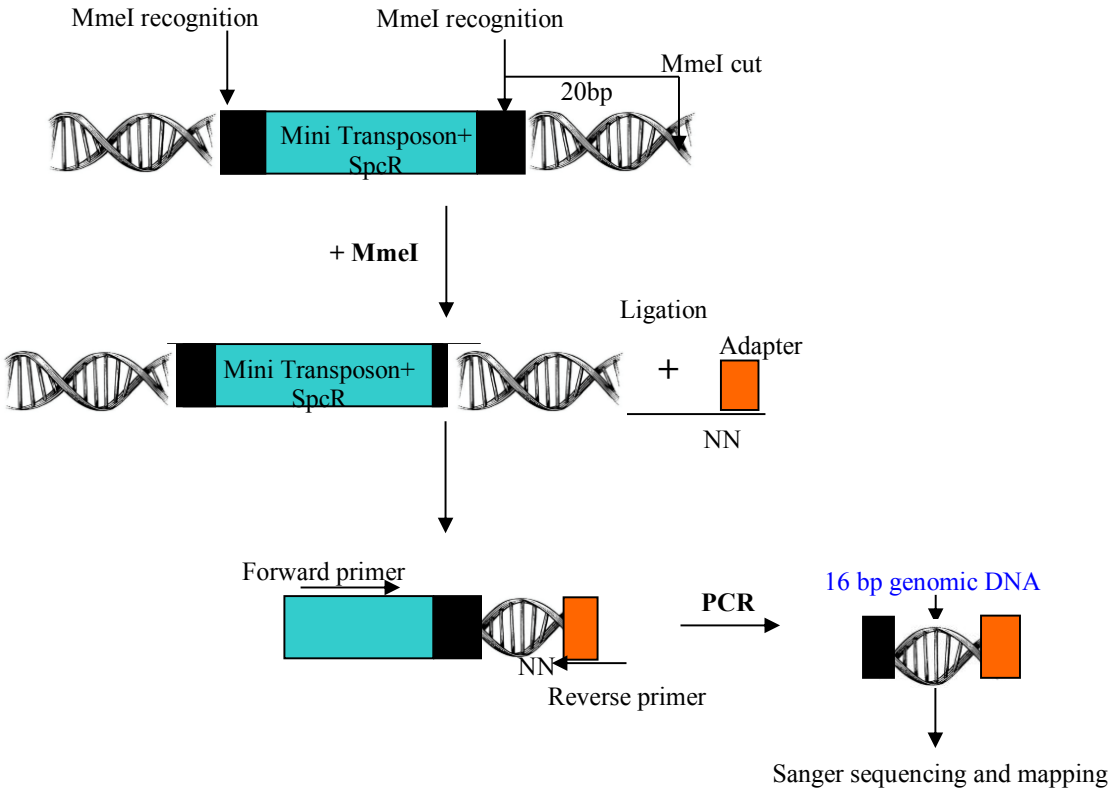


Figure 36. Structure of MarinerT7 transposon mutation DNA. DNA from suppressor strains was isolated and digested by MelI, which contains 16 bp of pneumococcal genomic DNA. The digested product was ligated to an adapter, which enables sequencing the DNA. A PCR amplification was performed to obtain a small DNA with 16 bp of pneumococcal DNA and adapter sequence. Then the 16 bp reads can be mapped to the genome.

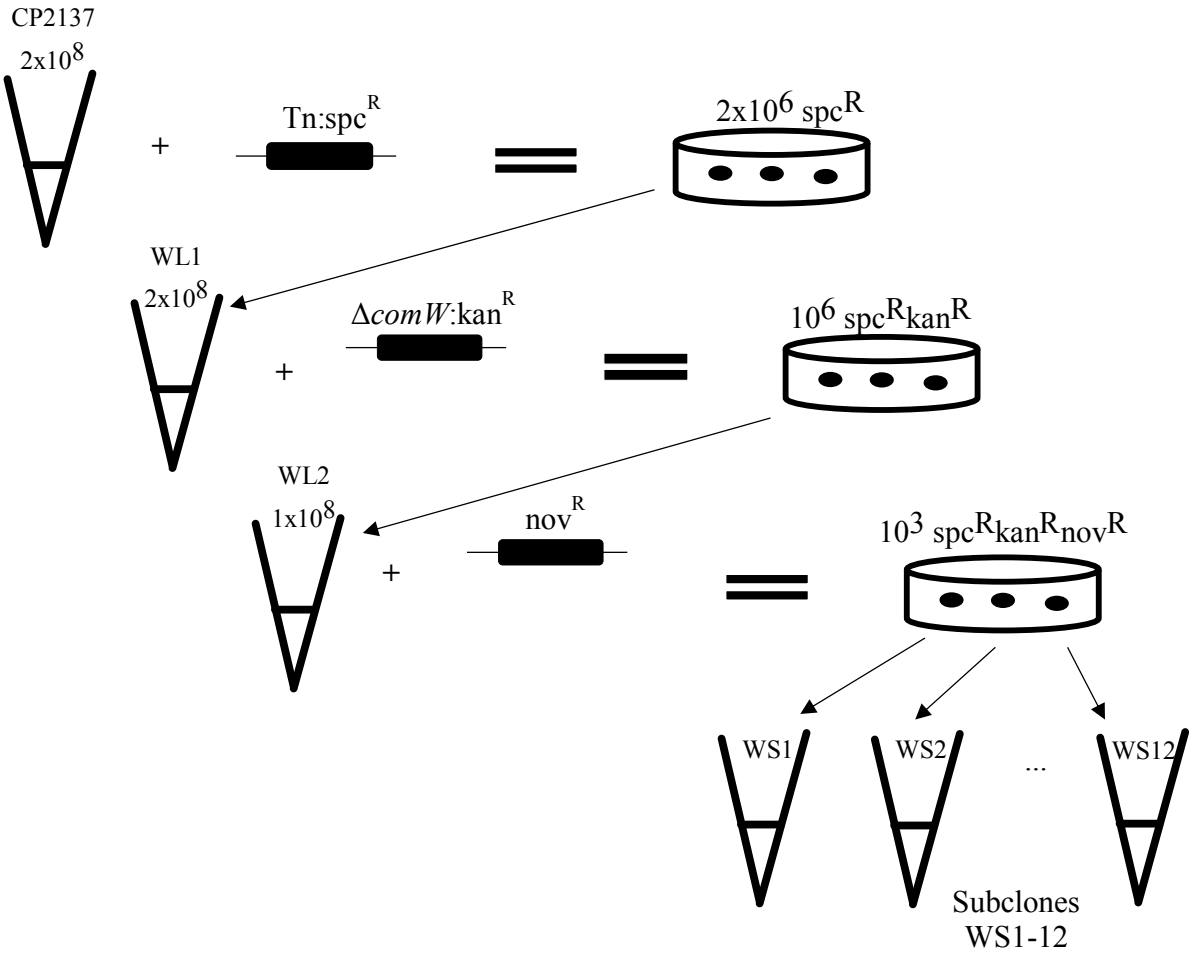


Figure 37. Strategy for the enrichment of Δ*comW* suppressor mutants. The pneumococcal Tn::spc^R mutant library was transformed into wild type strain CP2137 at OD_{550} 0.1. Then 10^6 transformants were obtained from spectinomycin plates as a pool and named as WL1. The Δ*comW* mutant was created by transforming Δ*comW*::kan^R into WL1, resulting 10^6 spc^{R}kan^R transformants were collected as a pool WL2. Followed by another transformation using nov^R as donor DNA to select clones containing possible Δ*comW* suppressors. Twelve individual nov^R colonies were picked up and named WS1-12 for further study.

In summary, the marinerT7 mutation library has failed to identify a mutant with a single amino acid substitution, suggesting that Δ*comW*-suppressing Tn::spc^R mutations were rare compared to a spontaneous mutation. The frequency of a spontaneous mutation was estimated as 10^{-6} by rifampicin-resistant single point mutation discussed later. The restoration of competence could be due to either a point mutation in a suppressor or Δ*comW*::kan^R itself creating the suppressor mutations.

5.4.3 Construction of facultative host strains to identify Δ*comW* suppressor mutants

A mutant with restored competence phenotype could be a possible anti-σ factor that bypass the requirement of ComW during competence. To verify such a putative anti-σ factor through the screen of *ΔcomW* suppressors, a facultative host strain was needed to be constructed.

To construct the facultative host strain, *comW* expression was designed to be inducible that its expression could be controlled. The endogenous *comW* was deleted but an ectopic copy of *comX* and *comW* were placed under a raffinose inducible promoter P_{aga}, because *S. pneumoniae* can use sugar as sources of energy, such as glucose, raffinose, and fucose (Rosenow & Trias, 1999; Russell & Ferretti, 1992). A donor DNA containing P_{aga}-*comX*-*comW*-*kan*^{R} was amplified from strain CP1896 and transformed into wild type strain CP2137. The transformants were selected on kanamycin plates, and several individual clones were tested for correct integration of P_{aga}-*comX*-*comW*-*kan*^{R} by Sanger sequencing of the *aga*-*rafE* locus. A correct subclone was retained and named CP2158 (see Figure 40).

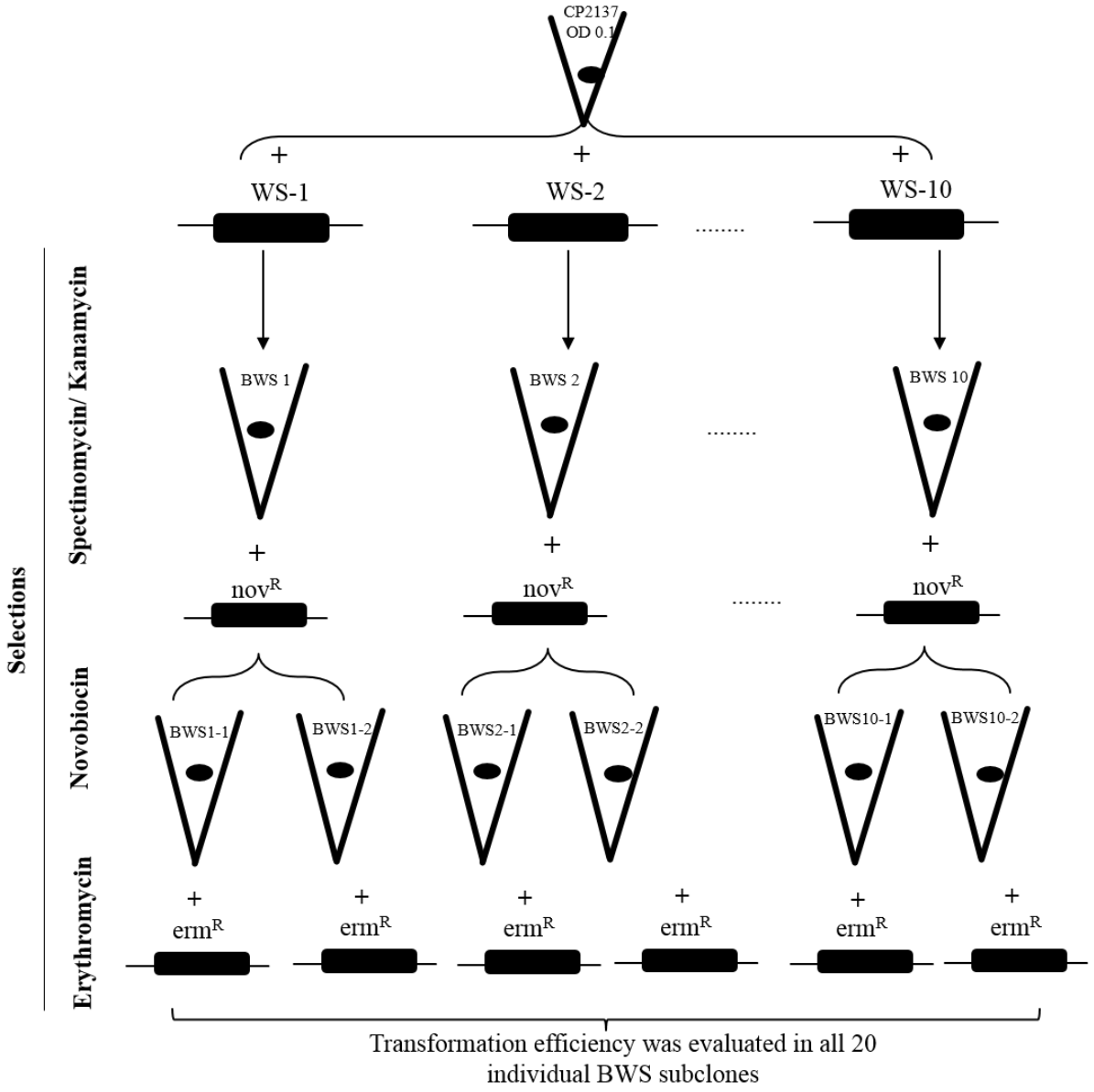


Figure 38. Backcross to determine linkage of competence to Tn::spc^R; ΔcomW::kan^R. Genome DNA from 12 strains WS1-12 was isolated and backcrossed to strain CP2137. The transformability of backcross progeny (BWS1-12) was examined for competence by transforming a novR DNA. Then 10 out of 12 WS strains were competent and collected as BWS1-10. Two individual clones of the resulting transformants from each BWS strains were selected on plates with the double drug selection of spectinomycin and kanamycin and named as BWS1-2. The transformation assay with ermR DNA as donor was performed to examine competence of progeny in all 20 individual BWS strains.

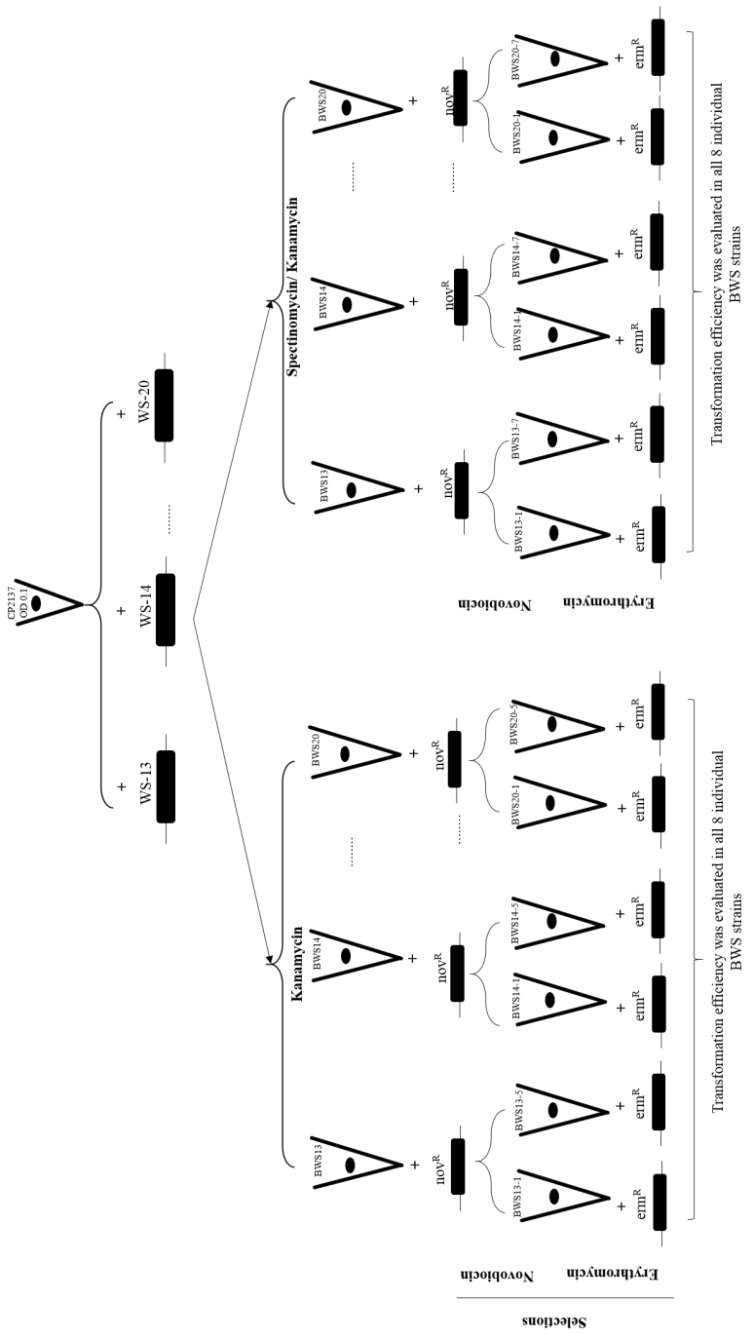


Figure 39. Backcross to determine linkage of competence to Δ*comW*::kan^{R}. Genomic DNA from 8 individual WS13-20 strains was isolated and backcrossed into a wild type strain, CP2137. Seven or five individual clones of the resulting transformants from each backcross were selected on plates with kanamycin only or double selection of spectinomycin and kanamycin and named as BSW1-5 or BSW1-7, respectively. The first round of transformation assay was performed by transforming nov^{R} DNA, following by the second round of transformation assay with erm^{R} DNA. Transformation efficiency of individual BWS strains was evaluated by ermR transformants.

Table X. Competence linkage to antibiotic-resistant genes

Suppressor Isolate	Competence of Backcross clones		Suppressor linkage to marker	
	$Spc^R Kan^R$ Com^+/Com^-	Kan^R Com^+/Com^-	$Spc^R Kan^R$	Kan^R
1st Backcross				
WS-1	0/2	-	no	-
WS-2	0/2	-	no	-
WS-3	1/1	-	yes	-
WS-4	2/0	-	yes	-
WS-5	1/1	-	yes	-
WS-6	1/1	-	yes	-
WS-7	0/1	-	no	-
WS-8	2/0	-	yes	-
WS-9	2/0	-	yes	-
WS-10	2/0	-	yes	-
WS-11	-	-	-	-
WS-12	-	-	-	-
2nd Backcross				
WS-13	4/1	4/3	yes	yes
WS-14	0/5	0/7	no	no
WS-15	1/4	3/4	yes	yes
WS-16	5/0	7/0	yes	yes
WS-17	2/3	3/4	yes	yes
WS-18	3/2	7/0	yes	yes
WS-19	0/5	0/7	no	no
WS-20	4/1	5/1	yes	yes

To test the activity of P_{aga}, 0.2% raffinose was used as suggested (Rosenow et al., 1999) to activate the promoter in the standard growth medium (CAT with 0.2% glucose and phosphate). Activation of P_{aga} was expected to turn on expression of ComW and σ^X proteins, which was represented by transformation efficiency of the host strain with raffinose induction. In this study, transformation efficiency was used to indicate the level of P_{aga} activity.

Transformation efficiency was found to be low (10^4 CFU) in CP2158 even under inducible conditions (Figure 41). As we know, ComW and σ^X proteins are stable in a protease *clpC* deletion mutant (Sung & Morrison, 2005). Thus, to increase the transformation efficiency of the host strain under a permissive condition, *clpC* was replaced by homologous recombination by transforming a CP1344 genomic DNA containing *ΔclpC::tet*^{R}. Transformants selected from tetracycline plates were confirmed by Sanger sequencing of *clpC* locus and named CP2165. The absence of ClpC was expected to stabilize the ectopically expressed σ^X and ComW. The transformation efficiency of CP2165 was determined by transforming a nov^R PCR donor DNA at 30°C (permissive condition), resulting 10^5 CFU (see Figure 41).

Moreover, to delete the original *comW* gene in CP2165, a *ΔcomW::spc*^{R} donor DNA was transformed, and transformants selected from spectinomycin plates were picked up and sequenced. Sequencing of the *comW* locus verified the correct integration of *ΔcomW::spc*^{R} to the recipient strain. The correct clone was named CP2166 to serve as the facultative host strain.

5.4.4 Evaluation of the facultative host strain under two conditions

The facultative *ΔcomW* host strain needed to provide two conditions: first a permissive condition allowing a library of *ΔcomW* suppressors to be transformed and then a non-permissive condition allowing a putative suppressor mutant to be tested.

Transformation assay was performed to evaluate the host strain under two conditions: the yield of transformation with raffinose induction at 30°C reached a level, as high as in a wild type with CSP induction at 37°C (10^5 CFU). In the non-permissive condition induced by CSP, CP2166 transformed higher (10^4 CFU) than a *ΔcomW* mutant (10^2 CFU). The unexpectedly high level of transformation of CP2166 under non-permissive condition would be a barrier to identify a suppressor mutant (see Figure 42).

To address this issue, temperature sensitivity tests were carried out since the temperature was found to play an important role in the efficiency of the raffinose promoter (Rosenow et al., 1999). Transformation efficiency of all three strains: the host strain, wild type, and *ΔcomW* mutant were examined under different temperatures (data not shown). The levels of transformation reduction above 37°C were observed in all three strains. Consequently, it was concluded that increasing the temperature in the non-permissive condition could not increase discrimination between the *ΔcomW* and WT phenotypes because both the facultative and control strains showed the same amount of decrease in competence from 37°C to 39°C. Increased temperatures did not increase the contrast in the transformation of the facultative strain between the permissive and non-permissive conditions.

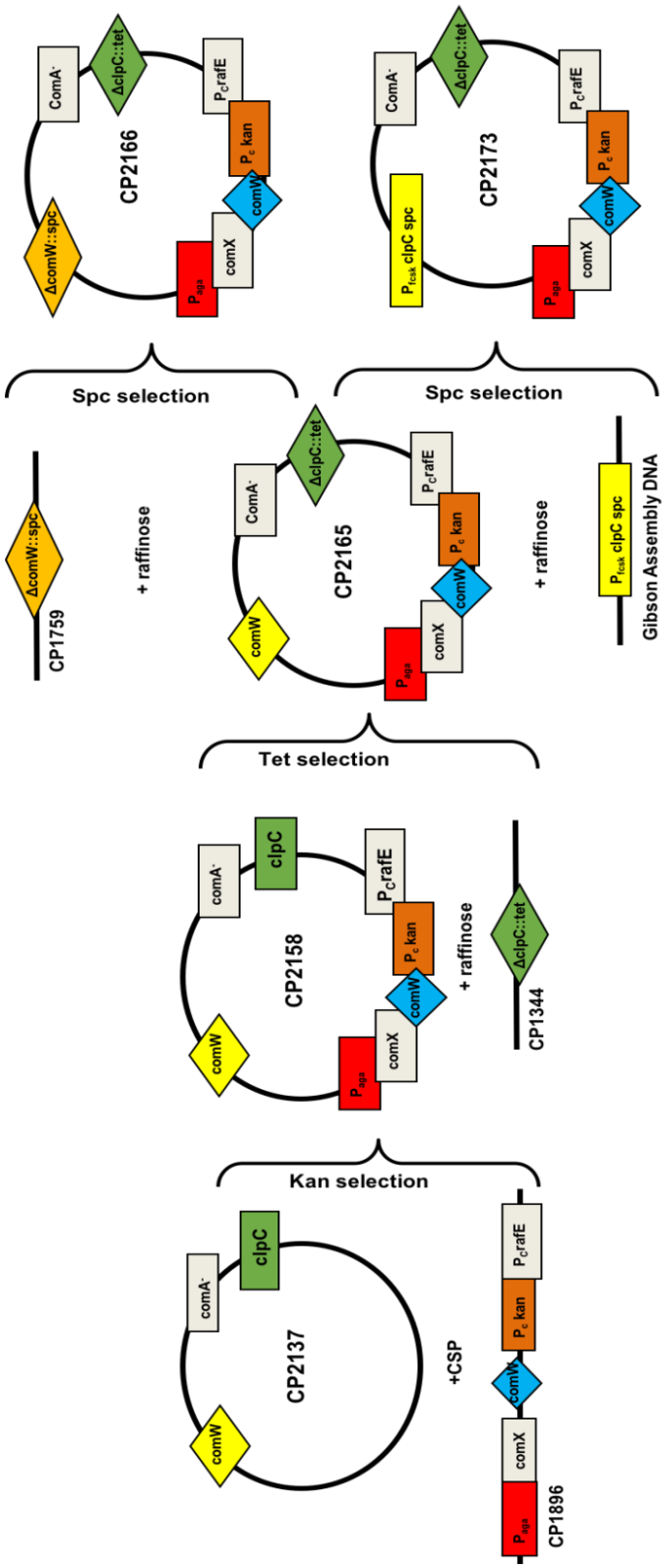


Figure 40. Schematic of facultative host strains construction. Creation of new strains stated with a wild type strain CP2137 and transformation were carried out by donor DNA indicated. All donor DNA were isolated from the strains indicated besides Gibson assembly DNA.

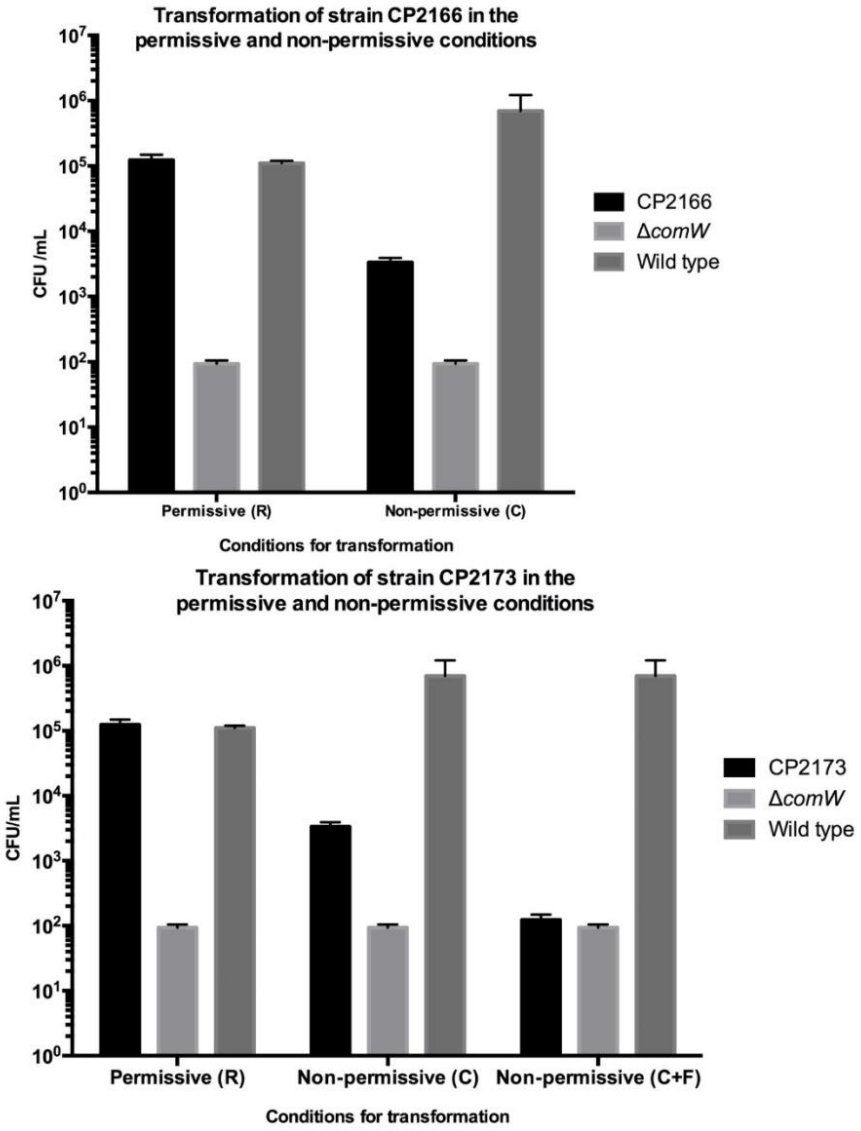


Figure 41. The transformation efficiency of host strains constructed in this study. (Top). The transformation level of strain CP2166 (Paga-*sigX-comW*, Δ*clpC*::tet^R, Δ*comW*::spc^R) was examined under permissive and non-permissive conditions. (Bottom). The transformation level of strain CP2137 (Paga-*sigX-comW*, Δ*clpC*::tet^R, Δ*comW*::Pfcsk:*clpC*::spc^R) was determined under the two conditions. C: CSP; R: raffinose; C+F: CSP+fucose.

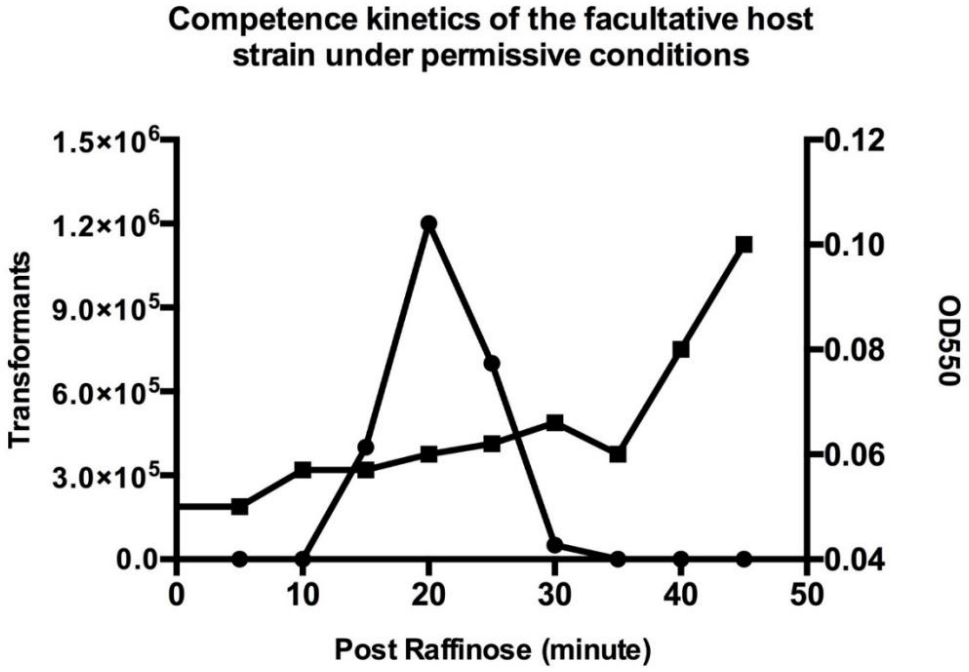


Figure 42. Competence kinetics of the host strain CP2173 with raffinose induction. The number of transformants/mL after a 3-minute exposure to 100 ng donor DNA was measured in strain CP2173 after raffinose induction. Samples were collected at 5-minute intervals. Meanwhile, OD550 was recorded in a 50-minute period.

5.4.5 Inability of sucrose to repress *aga* promoter under non-permissive condition

Raffinose is an activator of transcription of the aga locus, which encode α-galactosidase. The aga promoter has a catabolite repressor binding site (CRE) in its promoter region, which can be regulated by carbon catabolite repression (CCR). Sucrose has been shown to be the only catabolite that can repress P_{aga}, even though glucose is a more common catabolite repressor in prokaryotes (Rosenow, et al., 1999); however, sucrose was shown to be the repressor of aga promoter, as no α-galactosidase induction was detected in the medium.

To test whether sucrose could be used to repress leaky expression of P_{aga} under the non-permissive condition, 0.2% and 1% sucrose were examined in competence assays. The number of transformants counted under 0, 0.2%, and 1% conditions was similar (data not shown). This indicated that sucrose could not be used to repress P_{aga} under the non-permissive condition. This inability of sucrose to repress *aga* promoter may be attributable to incomplete CRE during strain construction when P_{aga} was introduced into pneumococcal cells or to my use of a medium different from the induced medium used by Rosenow (1999).

5.4.6 Decrease transformability by ectopic *clpC* expression under non-permissive condition

In the facultative host strain, the presence of 0.2% raffinose was expected to induce expression of ComW to allow DNA uptake at a high transformation efficiency during competence, whereas the presence of CSP was expected to turn off expression of *comW* to eliminate competence. However, under the non-permissive condition with CSP induction at 37°C, CP2166 had a transformation efficiency of approximately 1%, which was 100-fold higher than that of a *ΔcomW* mutant. The P_{aga}-*comX*-*comW* locus, which was inducible by raffinose, had a leaky expression of ComW under a uninduced condition. This may have contributed to the higher transformation efficiency under the non-permissive condition.

To overcome this problem, *clpc* was re-introduced to the host strain. By placing *clpc* under a fucose inducible promoter (P_{fcsk}), the host strain was expected to eliminate the leaky expression of σ^X and ComW from the P_{aga} promoter under when it was induced with CSP. To construct a donor DNA with *clpC* placing under P_{fcsk}, Gibson assembly was used to efficiently ligate four fragments in one reaction that contains P_{fcsk}, clpC, and spc^R, and *comW* flanking regions. By homologous recombination, the ligated DNA fragment was transformed into CP2165 to replace the original *comW* locus. The resulting transformants from spectinomycin plates were tested for correct integration of the Gibson assembly DNA into the *comW* locus of the genome. The correct transformant was retained and named CP2173. The new strain was *comW* deficient but had an ectopic *clpC* (see Figure 39).

The new host strain was expected to be inducible by fucose to activate expression of ClpC to eliminate transformation stimulated by leaky expression of ComW from the P_{aga} promoter under the non-permissive condition. Transforming the host strain with a 5mc genomic DNA under 1% fucose induction resulted in 0.01% efficiency. In addition, one prominent characteristic of competence in *S. pneumoniae* is a rapid on and off process, competence kinetics of the host strain was similar to a typical one that competence appears at 5 minutes, peaks at 20 minutes, and disappears within 40 minutes, lasting 40 minutes under experimental conditions (see Figure 41).

Strain CP2173 was tested under two conditions by transforming a genomic DNA (Nov^R) in transformation assays, the results showed that under permissive condition with raffinose induction CP2173 produced a wild type-like level of transformation efficiency, whereas under non-permissive condition with 1% fucose and CSP produced *ΔcomW* mutant-like level of transformation efficiency (see Figure 41).

The performance of the host strain is crucial in determining whether the suppressor mutation is inherited as a single gene locus or as multiple gene loci. Generally, cells uptake one region of genomic DNA containing about five genes, but less frequently they take up multiple noncontiguous regions on the genome. If the observed transformation efficiency of the backcrossed strain is very similar to that of the permissive condition, it would suggest that the suppressor mutation was located within a single gene locus. However, if the transformation efficiency is not as high as expected, it would suggest that the suppressor phenotype requires two (or more) mutations on two (or more) loci to be transformed into the facultative strain (see Figure 42).

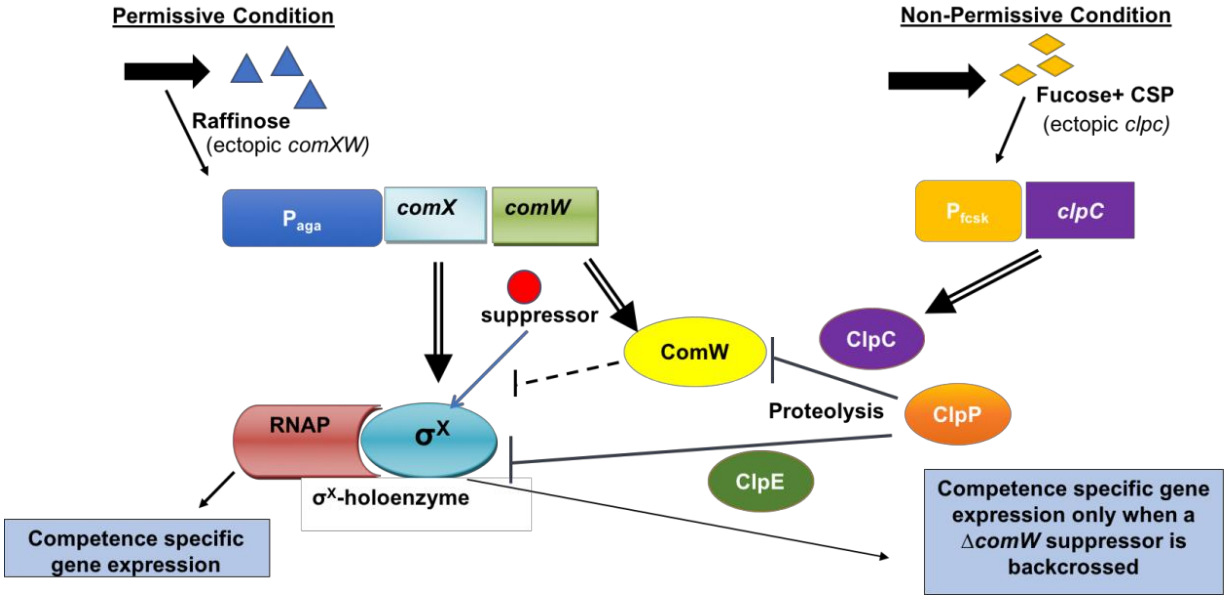


Figure 43. Model of the host strain under permissive and non-permissive conditions. Ectopic copy of *comX* and *comW* were cloned under Paga which were induced with raffinose. Under the permissive condition that ComW and σ^X proteins were present, ComW could bind to a suppressor that release the sequestration of and σ^X, resulting σ^X-dependent gene expression. Under non-permissive condition with ClpC presence, leaky expression of ComW was eliminated. Expression of σ^X-dependent genes occurred only when suppressor was mutated and could not bind to σ^X without the help of ComW.

5.4.7 Generation of rifampicin-resistance mutants

To increase the available antibiotic resistant donor DNA in compound transformation assays during library enrichments, a rifampicin resistance mutant was generated. As it is known that the limited variety of antibiotic resistant markers can hinder selection of transformants. In the case of rifampicin resistant donor DNA has been reported to be both feasible and sensitive.

Bacterial resistance to rifampicin is due to point mutations in *rpoB*, which encodes the β subunit of RNAP (Padayachee & Klugman, 1999). The crystal structure of RNAP in complex with rifampicin has been identified as having a binding pocket for rifampicin. The binding pocket is near the active center of the main channel in a region of the β subunit where transcription occurs (Campbell et al., 2001). Moreover, the rifampicin-resistant mutations isolated were predicted to affect the formation of the binding pocket and decrease its affinity for rifampicin. In the present study, spontaneous pneumococcal mutant isolates were selected on CAT agar plates containing rifampicin at 0.5, 4, or 10 μg/mL, which covered a range of resistances from low (minimal inhibitory concentration (MIC) > 0.5 and < 4 μg/mL) to high (MIC, > 4 μg/mL) (see Figure 44). About 10^8 wild type cells were plated, and 200 CFU were obtained on 0.5 μg/mL rifampicin, which indicated a spontaneous mutation frequency of 2 $x10^{-6}$. Also, 50 CFU were obtained on 4 μg/mL and 10 μg/mL rifampicin plates, indicating a frequency of $0.5x10^{-6}$ (see Table XI). The frequency of low-level resistance was four times higher than that of high-level resistance. Among the mutants generated from 10 μg/mL rifampicin plates, ten clones were picked up and were stored for further examination. Then genomic DNA was extracted from the ten independent single clones and the *rpoB* locus of each genomic DNA was amplified. An 8.5 kb *rpoB* fragment was amplified with primers FP-r1 and RP-r1 and subjected to Sanger sequencing to verify mutations on *rpoB*. Point mutations were determined by comparing sequenced results to a reference sequence in a previous

study (Padayachee & Klugman, 1999) (see Figure 45). The strain containing the rifampicin-resistance mutation was verified by sequencing the *rpoB* locus and named CP2184.

In previous studies, mutants with high-level rifampicin-resistance were mapped to point mutations in domains I and III, and mutants with low-level resistance were mapped to domain II of *rpoB* (Meier et al., 2003). Surprisingly, in the present study, an amino acid change from serine$_{408}$ to proline was found in domain I (amino acids 406 to 434), the first time that this point mutation has been observed. No point mutation was found in domain II or III of *rpoB*. The PCR fragment was transformed into competent pneumococcal cells and made rifampicin sensitive cells resistant. The significance of the mutations was confirmed by transformation.

5.4.8 Rifampicin-resistance PCR DNA as a highly efficient donor DNA

Typically, these competence kinetics tests lasted for 40 minutes with 5-minute intervals of exposure to DNA. For these tests, wild type strains were grown in CAT during competence induction. Samples were collected at three different OD_{550S}—0.05, 0.1, and 0.2—to induce competence. The 8.5 kb rif^R PCR DNA was used as a donor in 5-minute exposure to cells. For these ODs, cells developed into competence after 5 minutes, reached peak competence at 20 minutes and disappeared at 40 minutes. With OD_{550} 0.05, competence reached almost 80%. With OD_{550} 0.1, competence was reduced by 10%, and with OD_{550} 0.2, it was reduced by about 20%. These findings were consistent with previous laboratory research in which lower OD produced robust competence (see Figure 46).

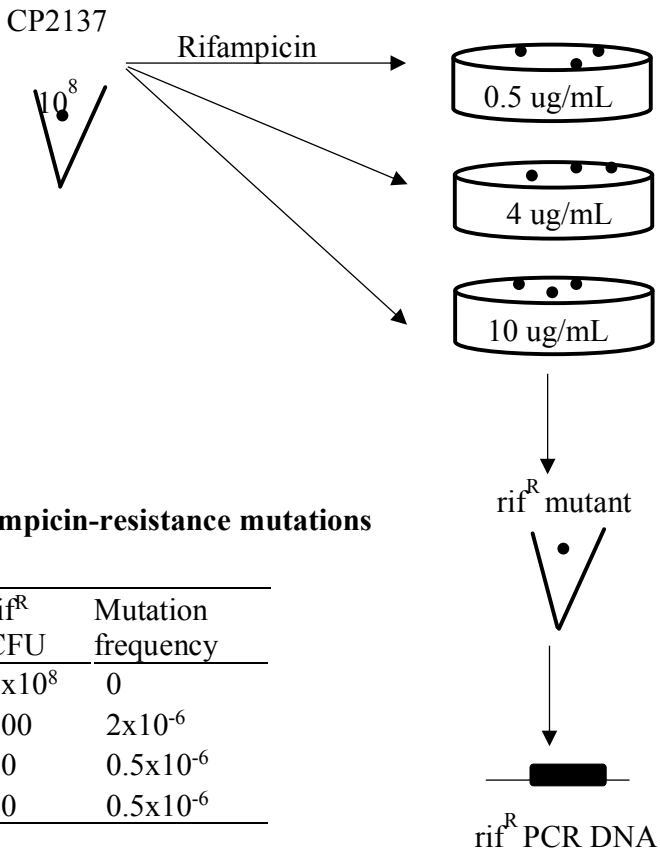


Table XI. The frequency of rifampicin-resistance mutations

Rifampicin (ug/mL)	rif^R CFU	Mutation frequency
0	$1x10^8$	0
0.5	200	$2x10^{-6}$
4	50	$0.5x10^{-6}$
10	50	$0.5x10^{-6}$

Figure 44. Generation of rifampicin-resistant mutants. Spontaneous pneumococcal mutant isolates were selected on CAT agar plates containing from low to high level of rifampin at 0.5, 4 or 10 μg/mL. Clones grown on rifampicin plates were counted, and results were shown in the table XI. Single clones were picked up from 10 μg/mL plates, outgrown to isolate genomic DNA. Then Sanger sequencing was used to verify mutations on *rpoB*.

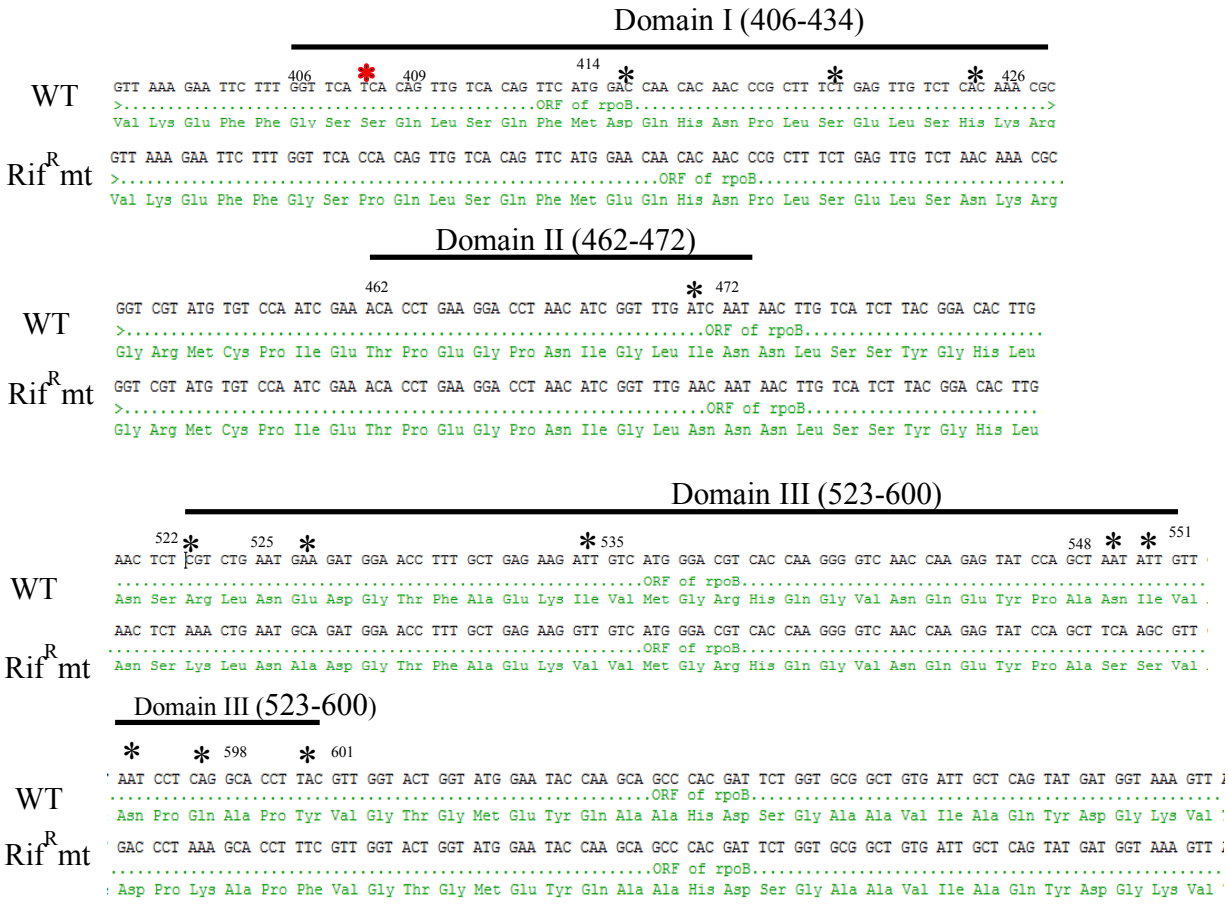


Figure 45. Nucleic acid substitutions in rifampin-resistant strains containing mutations. Mutations have been reported previously as indicated by asterisks. The sequence assignment is based on (Padayachee & Klugman, 1999). In *S. pneumoniae*, one single point mutation was located on Serine 408, and proline substitution conferred the high-level of rifampicin resistance.

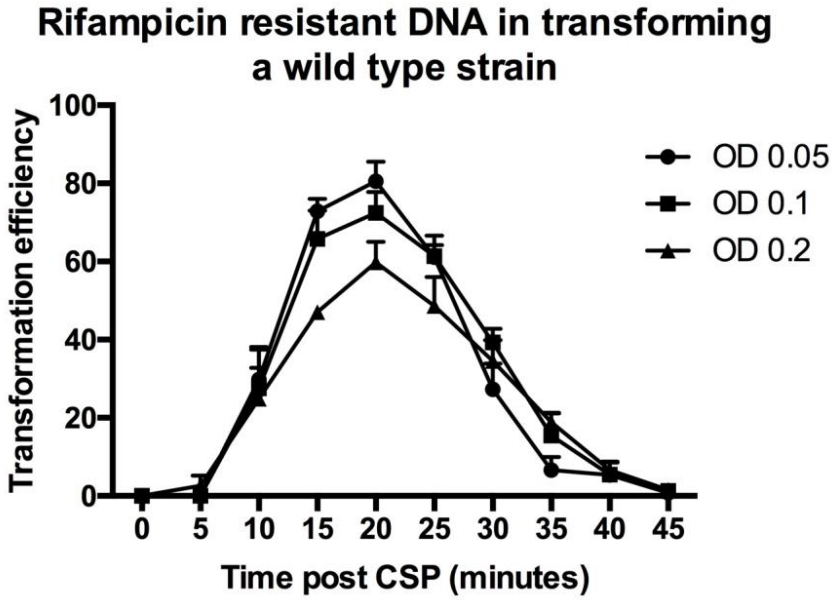


Figure 46. *rpoB* rif^{R} mutation can be used as a high efficient donor DNA in transformation. Rifampicin resistant mutant was selected on high-level of rifampicin plates. *rpoB* gene was amplified and used as a highly efficient donor DNA in transformation assay under OD550 0.05, 0.1, and 0.2. Competence kinetics showed similar curve as wild type transformed routinely with a genomic or PCR DNA. The highest level of competence was achieved at OD 0.05.

5.5 Discussion

MarinerT7 transposon mutagenesis did not identify a *ΔcomW* suppressor mutation. Restoration of competence in the *ΔcomW* background was independent of transposon insertion Tn::spc^R, possibly because transposon insertions have interrupted essential genes. In this study, 75% of competent progeny were linked to *ΔcomW::*kan^R and 25% of that to *comX2*. It was puzzling to determine competence restoration by *ΔcomW::*kan^R, but the exact mechanism was unknown. Sequencing results showed that insertion occurred undesirably and caused the original chromosomal *comW* was not deleted by donor DNA *ΔcomW::*kan^R during the transformation event. This could be the possibility that some genes were preferred to be mutagenized by a plasmid-based transposon instead of the Mariner system and resulted non-reproducible outcomes (Bijlsma et al.,2007). To conclude, the restoration of competence in the mutants was provided by the original *comW* rather than an interrupted *ΔcomW* suppressor. Failure to identify such a suppressor mutant in *ΔcomW* is taken as presumptive evidence that the MarinerT7 transposon library has interrupted essential genes. As protein activity could be regulated at the stability level through protease, it is possible that a point mutation that constitutively activated to increase the affinity of a suppressor to its specific protease may have relieved the effects of *comW* deletion. Alternatively, a less stringent mutagenesis strategy might be useful to identify a suppressor mutant, in which the selection condition allows it to survive with point mutations.

This was supported by the finding using spontaneous mutation screening in combination with whole genome sequencing (Tovpeko & Morrison, 2014). They found that the spontaneous mutagenesis allowed recovery of subtler suppressor mutations. Most single amino acid changes in *ΔcomW* suppressor mutants were mapped on *rpoD* (σ^A). Mutations on the housekeeping σ factor

may explain why a MarinerT7 transposon library was not successful in identifying a *ΔcomW* suppressor mutation.

In addition, the facultative host strain has proven to be a tightly controlled strain with properties that optimal transformability under permissive conditions, and deficient transformability under non-permissive conditions. Numerous competence kinetics tests were performed to determine the conditions under which the host strain became the most competent. Moreover, a highly efficient rif^R donor DNA was generated to enrich options of antibiotic resistant markers in compound transformation assays.

VI. GENERAL CONCLUSION AND DISCUSSION

It has been 13 years since ComW (SP0018) was identified as a positive regulator during competence in *S. pneumoniae*. ComW is 10 kDa in size, its expression parallels to that of σ^X during competence, and it is crucial for σ^X stability and activity. In the Δ*comW* mutant, σ^X activity in transcription of late genes is reduced by 90%, resulting in a 99.9% decrease in competence. Furthermore, Andrew and Morrison (2009) found that both N- and C-termini of ComW are required for σ^X stability and that the N-terminus is more important for σ^X activity.

What we know about ComW triggered our interest in understanding its mechanism in regulating competence and several attempts have been made. The first question was whether ComW affects transcription of σ^X-dependent genes *in vitro*. In Luo and Morrison's (2003) study, *in vitro* transcription assays indicated that ComW might not be required for σ^X-dependent late genes. However, there were limitations in this study that may compound the results. Transcription of late genes was not determined relative to the amount of σ^X protein present; consequently, it was not clear whether transcription levels changed with the holoenzyme purified from competent cells or with the reconstituted holoenzyme formed by σ^X and RNAP. Moreover, the assays were performed with over-expression of σ^X from *E. coli*, which may have unknown factors playing the role of ComW that led to σ^X activation without the assistance of ComW. The transcription levels were not precisely measured to reveal the differences with and without ComW during σ^X-mediated transcription. The second question was whether ComW and σ^X interact directly. Y2H and affinity purification assays have been performed in different studies to determine whether a direct interaction occurs between ComW and σ^X (Piotrowski et al., 2009; Weng et al., 2013). However, the assays yielded inconsistent results. The Y2H assay indicated a direct interaction between σ^X and ComW, while the affinity purification assay indicated no such interaction. The affinity

purification assay is limited in that a weak interaction between σ^X and ComW may not be detected.

Based on previous studies of ComW and σ^X during CSP-induced competence, Luo and Morrison (2003) hypothesized that ComW activates σ^X by post-translational modification, such as phosphorylation or methylation, or that it plays a role in antagonizing σ^A from core RNAP. In the latter case, ComW would not be required for transcription in *in vitro* assays (Luo & Morrison, 2003).

Given this information, **the first goal** of this study was to determine whether a direct interaction occurred between σ^X and ComW. The Y2H assay was repeated with an improved experimental design to unambiguously determine whether a direct interaction exists between σ^X and ComW. This independent Y2H assay confirmed a direct interaction between σ^X and ComW by using different shuttle vectors and a set of four tightly controlled reporter genes. ComW showed a weak interaction with σ^X, which was indicated by the fact that only two of the reporter genes were activated. This finding was confirmed by shuttle vector swap and examination of truncated σ^X. Interestingly, ComW did not interact with σ^A in the Y2H assay. There was a recent finding of Δ*comW* bypass mutations in σ^A, especially in regions 2 and 4, which interfere with the interaction with core RNAP (Tovpeko & Morrison, 2014). In summary, our current working model of ComW is as follows: ComW may spark competition between σ^A and σ^X in favor of σ^X access to core RNAP by directly binding to σ^X, or the σ^A mutations in regions 2 and 4 weaken its binding affinity to core RNAP, eliminating the requirement for ComW during competence initiation.

Our Y2H results could be confirmed by surface plasmon resonance (SPR) and mass spectrometry (MS). Mapping the residues that are important for a physical interaction of σ^X with ComW will validate the interaction, as well as decipher the significance of the interaction. Such kinds of residues of σ^X could be identified by means of combined assays of error-prone PCR and

library screening. This screening could include a search for σ^X mutations (σ^X*) disrupting σ^X-ComW interactions, but still maintaining the ability of σ^X to transcribe late genes with core RNAP. The σ^X is highly conserved in many Streptococcus species, so characterizing conserved residues of σ^X regarding its interaction with ComW may help us to understand competence regulation in other *Streptococcus* species.

Deciphering the mechanism by which ComW facilitates σ^X activity will offer a model for an alternative σ factor regulation, as many streptococcal and lactococcal species possess homologous competence regulatory operons of *S. pneumoniae*. Investigating the regulatory mechanism of competence could have a broad impact on our understanding of the high rate of horizontal gene transfer among Streptococci.

The second goal of this study was to investigate a new mechanism for competence regulation by re-creating a striking protein pattern switch first discovered by Baker and Morrison (Morrison & Baker, 1979). The well-known mechanisms of competence regulation occur at the transcriptional, translational, and post-translational levels. At the transcriptional level, genes are selected to be transcribed in response to environmental changes such as extracellular CSP accumulation and competence induction. At the translational level, mRNA conformation change results in concealed or revealed binding sites (such as in the sequence for ribosome-, activator- or repressor-binding), which respectively blocks or facilitates ribosome binding for translation initiation. However, I hypothesized that previously unidentified factors affect ribosome structure or activity during growth of unstressed pneumococcal cells. This hypothesis was inspired by a genome-wide survey of gene expression during competence performed by Morrison in collaboration with the University of Oslo (2016). Several newly identified genes belonging to a set of genes called core σ^X regulons, which are shared by all core genomes of Streptococci species

including *S. pneumoniae*, may provide some clues to the protein pattern switch. These genes may reveal a new mechanism that reprograms ribosomes for translation of a special class of mRNA as well as point to several specific core regulons potentially responsible for the reprogramming of ribosomes.

The new mechanism by which ribosomes could be modified to globally regulate protein expression challenges conventional understanding of the regulation of gene expression, under which ribosomes have been regarded as unchanging complexes for protein synthesis. There are three well-characterized cases where global protein synthesis is regulated by ribosome modifications. First, HPF induces ribosome inactivation while protecting from degradation under nutrient-deficient conditions in *E. coli* (Polikanov et al., 2012), *S. aureus* (Basu & Yap, 2016), and *L. lactis* (Puri et al., 2014). Second, MazEF induces the translation of a specific set of proteins under stress induced in *E. coli* (Vesper et al., 2011). And third, YhaM as an exoribonuclease together with other nuclease induces rRNA maturation (Oussenko et al., 2002; Redko & Condon, 2010).

The mechanism of translation suppression of the housekeeping mRNAs remains unclear. Transcriptome data for Streptococci revealed a new set of nine genes belonging to the core σ^X regulons that remain unexplained, and two of the nine have sequence similarities to HPF and YhaM, which modify ribosomes at the activity level in other bacteria (Khan et al., 2016). In our study, three deletion mutants (Δ*Hpf, Δyha*M*, and Δ*Hpf*Δyha*M*) were examined for details of protein synthesis labeled using pulse-chase assays, combined with 1D and 2D gel electrophoresis, and competence assays in parallel. The results suggested that HPF and YhaM regulate the protein expression pattern slightly but not globally. We have successfully generated a protein pattern shift on 1D gel images by overcoming the dilemma of non-synchronized levels of competence and

isotope incorporation. Protein synthesis was classified into down- and up-shifted patterns in a wild type strain. Surprisingly, quantitative analyses of 1D gel images revealed only a few different bands between the double deletion mutant and the wild type strain, which was confirmed by a few different spots on 2D gel images.

The additional seven genes in the new set are also included in core regulons, playing unknown roles in competence regulation. They are SP2044 (*ackA*), SP0021(*dut*), SP0979 (*pepF*), SP1941 (*cinA*), SP0782 (*pilC*), SP1088 (*radC*), and SP1981 (*rmuC*), collectively referred to genome sequence of strain TIGR4 (Khan et al., 2016). Some of the genes include orthologs of well-characterized proteins, and some have domains that fall into known broad functional categories.

Sequence comparisons of HPF between *S. pneumoniae* and *S. aureus* showed 44.8% of similarity. Translational inhibition mediated by HPF_{sa} occurred in *in vitro* translation assay in *S. aureus* (Khan et al., 2016). A growing number of studies link HPF mediated ribosome activity with translational regulation in bacteria. For example, HPF is required for rRNA preservation and ribosome inactivation under starving conditions in *P. aeruginosa* (Akiyama et al., 2017). The rRNA preservation is important for ribosomal integrity since either ribosomal proteins or rRNA dissociated from complex becomes susceptible to proteolysis. Recently, crystal structures of ribosomes with HPF, RMF, and YfiA revealed that their binding sites overlap with the transcription initiation factors on the 30S subunit, resulting in translation inactivation in *E. coli* (Polikanov et al., 2012). However, all known research of ribosomes inactivation mediated by HPF were studied under stress conditions. Competence occurring in logarithmic growing culture is not considered as a stress condition. This may explain the result that HPF slightly repressed expression of a few genes but not globally during competence. The slight protein pattern change in targeting

a few proteins was observed on 1D and 2D gel images, indicating that HPF and YhaM may not provide a mechanism in modifying ribosomal activity globally. However, this cannot eliminate the possibility that HPF and YhaM play a role in repressing translation of house-keeping genes. Our results showed 60% decrease of competence in the double deletion mutant.

An alternative strategy to determine the role of HPF in regulating protein expression during competence regulation may be using ribosome profiling (ribosome-seq), which produces a "global snapshot". The assay can start with cell lysates harvested from the mutant Δ*Hpf* and wild type strain. Then ribosomes actively translating mRNAs *in vivo* can be captured and followed by nuclease digestion to get ribosome footprinting. Purified ribosomes with protected mRNA fragments should be recovered by specific centrifugation thus the released mRNA can be ligated to a linker for reverse transcription. After those steps, PCR and high-throughput sequencing can be used to map ribosome-associated mRNA to genome in parallel samples harvested from the mutant and wild type strains. If distinct translation patterns are observed between these two strains, HPF plays a role in gene translation. No difference observed suggests HPF may not play a role in gene translation during competence, and more subtle roles may remain to be discovered.

The third goal of this study was to identify a Δ*comW* suppressor mutant. MarinerT7 system is a powerful method to generate random mutant libraries without polar effects on downstream genes (Rubin et al., 1999). We took this advantage and introduced a MarinerT7 transposon mutation library into a pneumococcal wild type culture, following several transformation assays to screen mutants with restored competence. Surprisingly, our results showed that competence restoration to Δ*comW* background was not linked to the marinerT7 transposon library but to Δ*comW*. No Δ*comW* suppressor mutant was identified despite the use of a mutation library with a saturation level extrapolated as 10-fold (2 x10^4 members of the mutation

library) coverage of the genome. It is possible that the library does not represent all genes of the pneumococcal genome (Burghout et al., 2007).

MarinerT7 was shown to have a slightly higher saturation rate than a plasmid-based transposon pGh9T7::ISS1 (Bijlsma et al., 2007; Burghout et al., 2007). Comparisons of transposon libraries generated by both mutagenesis showed approximately 9% of the total number of genes had the preference to be mutagenized by one of the two transposons, and 4% was identified as "blind spots" suggested by missing readings in the sequencing results. In addition, PCR-based amplification of DNA adjacent to the insertion site could lead to non-reproducible results in the presence of more than 500 mutants in one single screening (Bijlsma et al., 2007). These could be the reasons why we did not find any Δ*comW* suppressor mutants in our screening. In our case, the genes that overcome *comW* deletion could not be targeted by the Mainer-based mutagenesis system. Another possibility is that transposition occurred on essential genes that were detrimental under selective conditions and could not be detected from the mutant pools. This hypothesis is supported by a recent finding (Tovpeko & Morrison, 2014). Tovpeko performed a whole genome screening (WGS) and identified single-base substitutions in σ^A especially in regions that interfere with its interaction with core RNAP bypassing the requirement of ComW during transformation. In summary, MarinerT7 transposon mutagenesis probably produced insertions on essential genes or high-density of mutations under our experimental conditions compared to WGS which provided a sufficient but low-density of base substitutions to allow for mapping. Regardless from all pitfalls described above, MarinerT7 can still be a useful tool to identify non-essential genetic elements under selection that can shed light on microbiota and host interactions and provide useful targets for new clinical interventions.

In conclusion, our data indicated that ComW functions to facilitate σ^X access to core RNAP

by directly binding to σ^X. The exact molecular mechanism of ComW can be explored by *in vitro* transcription. The σ^X-dependent gene transcript levels will be compared directly between incubation with σ^X prior to σ^X-holoenzyme formation, and ComW added to the pre-formed σ^X-holoenzyme. Then transcription levels of late genes will be measured via gel electrophoresis. We can conclude that ComW facilitates σ^X-holoenzyme formation in the case that late gene transcription levels are higher where ComW is added prior to σ^X-holoenzyme formation. Alternatively, ComW might act to increase the promoter binding specificity of σ^X-holoenzyme by providing additional functional determinants to the promoter, which is still consistent with our finding that a direct interaction occurs between ComW and σ^X.

APPENDIX A

S. pneumoniae **strains constructed in this study.**

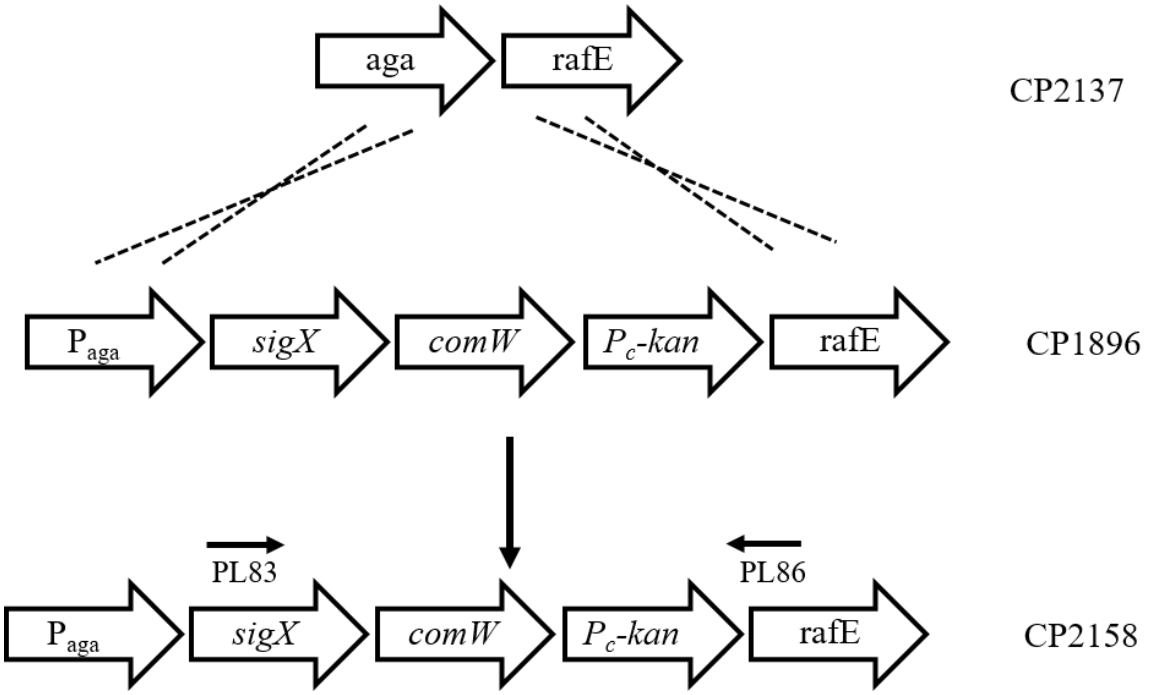


Figure 47: Construction of CP2158. To create a strain where *sigX* and *comW* are both under the control of the *aga* promoter, a genomic DNA CP1896 encompassing *aga* and *comX* under the aga promoter was transformed into CP2137. The structure of one clone (CP2158) was verified by sequencing the PCR product of primers PL83 and PL86 with primers AP19 and CKS36.

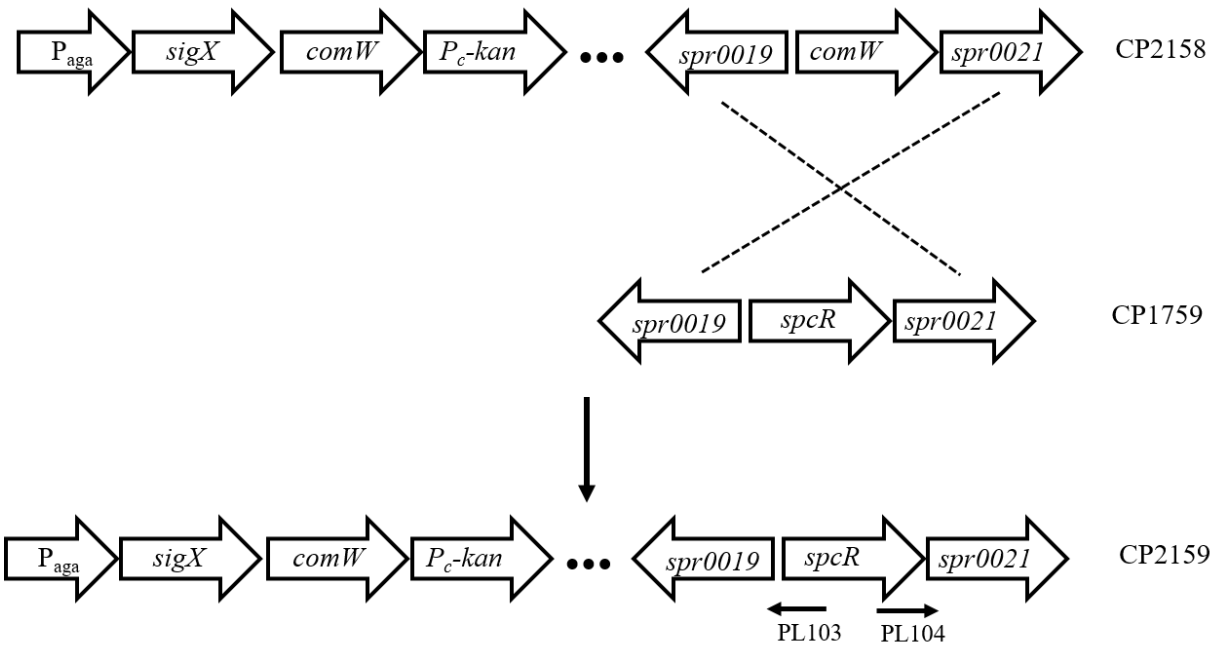


Figure 48. Construction of CP2159. To create a strain where *comW* deletion strain, a genomic DNA CP1759 encompassing *comW*::*spc*^{R} was transformed into CP2158. The structure of one clone (CP2159) was verified by sequencing the PCR product of primers PL106 and PL144 with primers PL103 and PL104.

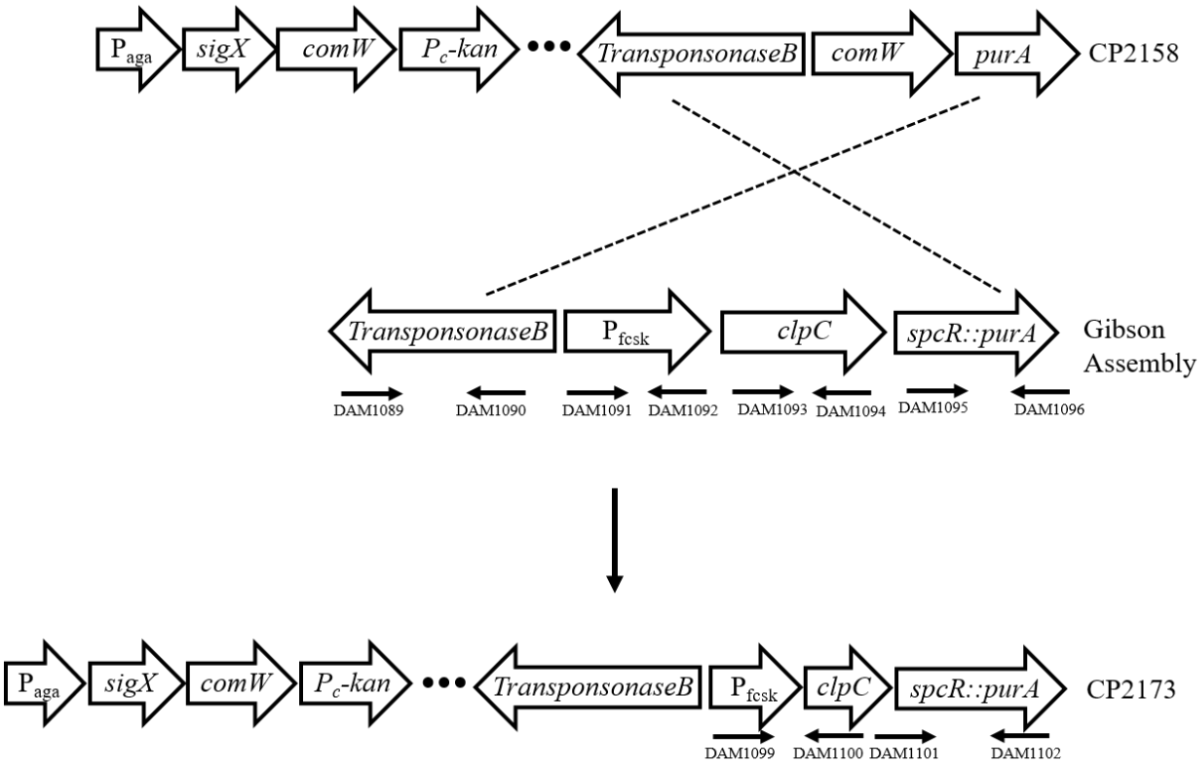


Figure 49. Construction of CP2173. To create a strain where *clpC* was under fucose inducible promoter, a Gibson assembly DNA encompassing *comW* upstream and downstream, *clpC* under fucose promoter was constructed. The *comW* upstream, downstream fragments, and *clpC* were amplified with template DNA CP1759 with primers DAM1089 and DAM1090, and DAM1095 and DAM1096, and DAM1093 and DAM1094. Fucose promoter was amplified with primers DAM1091 and DAM1092 with template CP2055. All fragments were ligated in one reaction by Gibson assembly and transformed into CP2158. The structure of one clone (CP2173) was verified by sequencing the PCR product of DAM1089 and DAM1096 with primers DAM1099 and DAM1100, and DAM1101 and DAM1102.

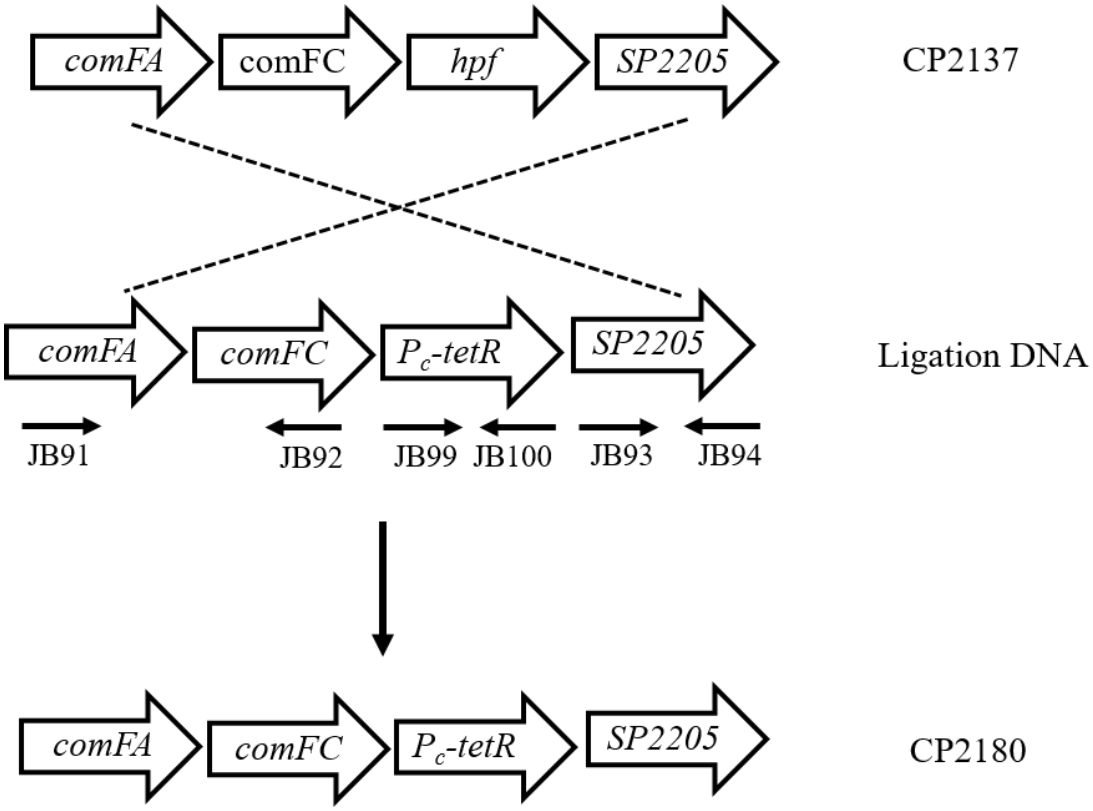


Figure 50. Construction of CP2180. To create a strain where *hpf* deletion mutant, *hpf* upstream *comFA* and *comFC*, and downstream SP2205 were amplified with template CP2137 with primers JB91 and JB92, and JB93 and JB94. The tetrycycline resistance gene was amplified with genomic DNA 6mt with primers JB99 and JB100. The upstream, downstream, and tetrycycline resistance gene were digested with SalI, NotI and SalI/NotI accordingly. Then the digested DNA was ligated and transformed into CP2173. The structure of one clone (CP2180) was verified by Sanger sequencing the *hpf* locus using primers JB117 and JB118, and JB119 and120.

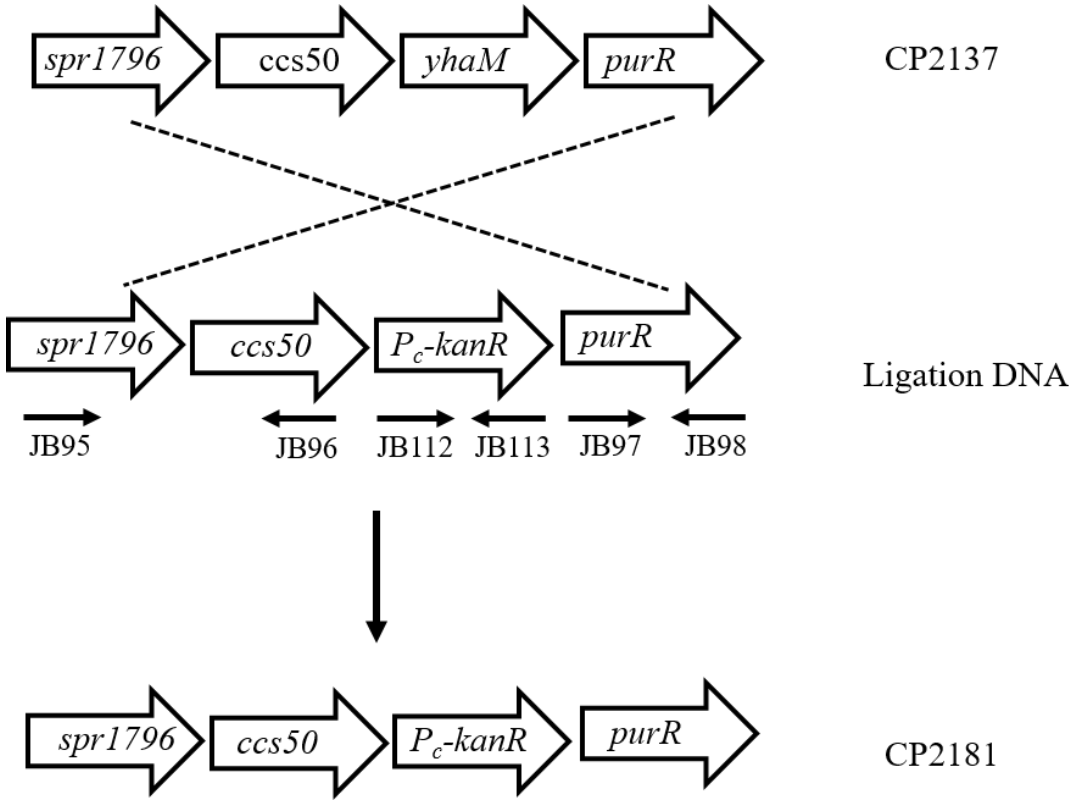


Figure 51. Construction of CP2181. To create a strain where *yhaM* deletion mutant, *yhaM* upstream *spr1796* and *ccs50*, and downstream purR were amplified with template CP2137 with primers JB95 and JB96, and JB97 and JB98. The kanamycin resistance gene was amplified with genomic DNA 6mt with primers JB112 and JB113. The upstream, downstream, and kanamycin resistance gene were digested with SalI, NotI and SalI/NotI accordingly. Then the digested DNA was ligated and transformed into CP2173. The structure of one clone (CP2181) was verified by Sanger sequencing the yhaM locus using primers JB121 and JB122, and JB123 and JB124.

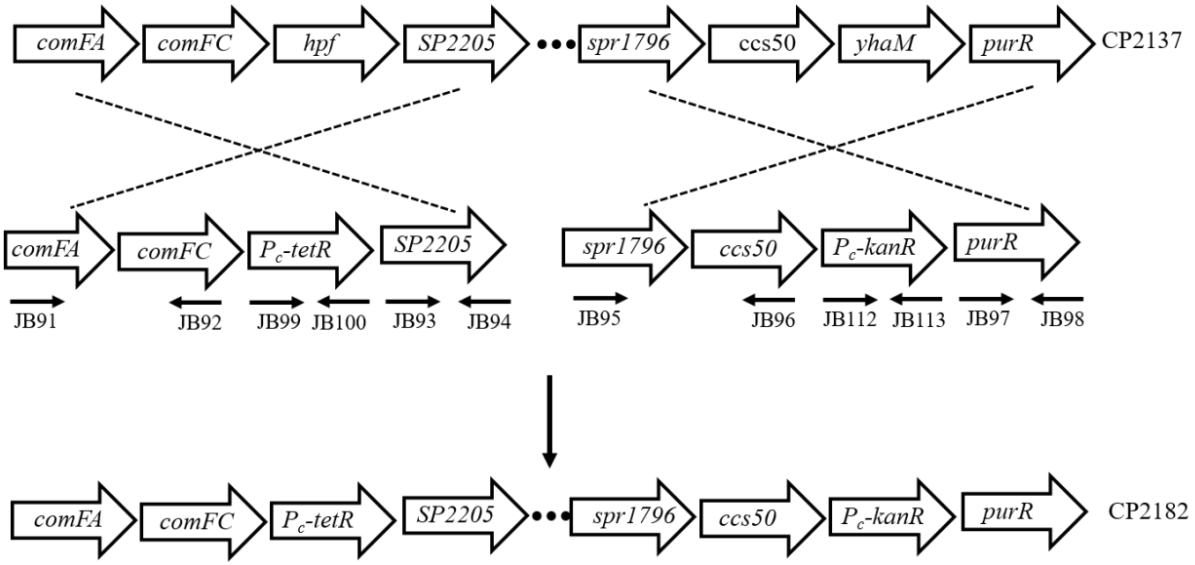


Figure 52. Construction of CP2182. To create a strain where *hpf* and *yhaM* double deletion mutant, *hpf* locus was replaced with tetrycycline resistance gene by transforming a genomic DNA CP2180 into CP2181. The structure of one clone (CP2182) was verified by Sanger sequencing the hpf locus using primers JB117 and JB118, and JB119 and120, and the yhaM locus using primers JB121 and JB122, and JB123 and JB124.

APPENDIX B

Plasmids constructed in this study.

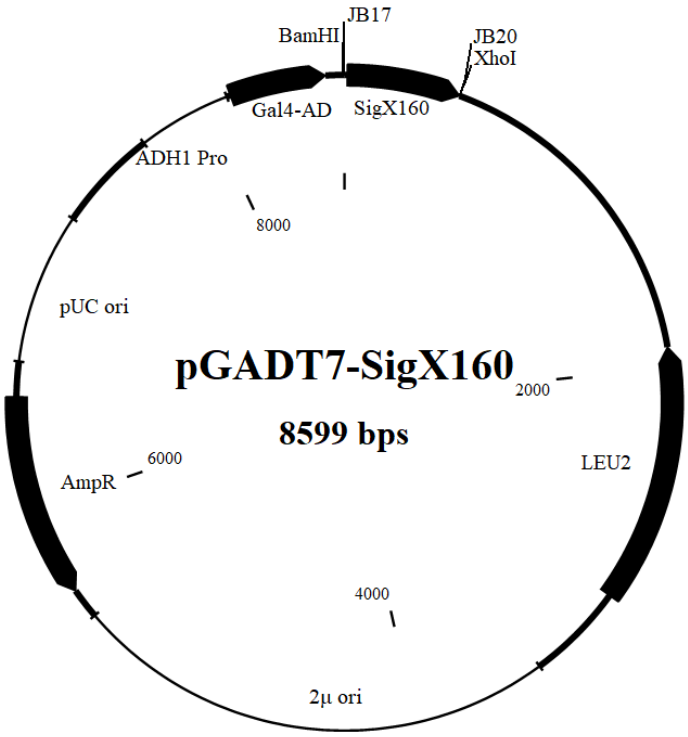


Figure 53. Construction of plasmid pMJB01 (pGADT7-*sigX*$_{160}$). To clone *sigX*$_{160}$ into pGADT7 vector, *sigX*$_{160}$ was amplified using primers JB17 and JB20 and digested by BamHI and XhoI. The PCR product was cloned into pGADT7 BamHI and XhoI sites by ligation. Then the ligated DNA was transformed into *E. coli* DH5α. The correct insertion was verified by sequencing the PCR product using primer T7 and JB66.

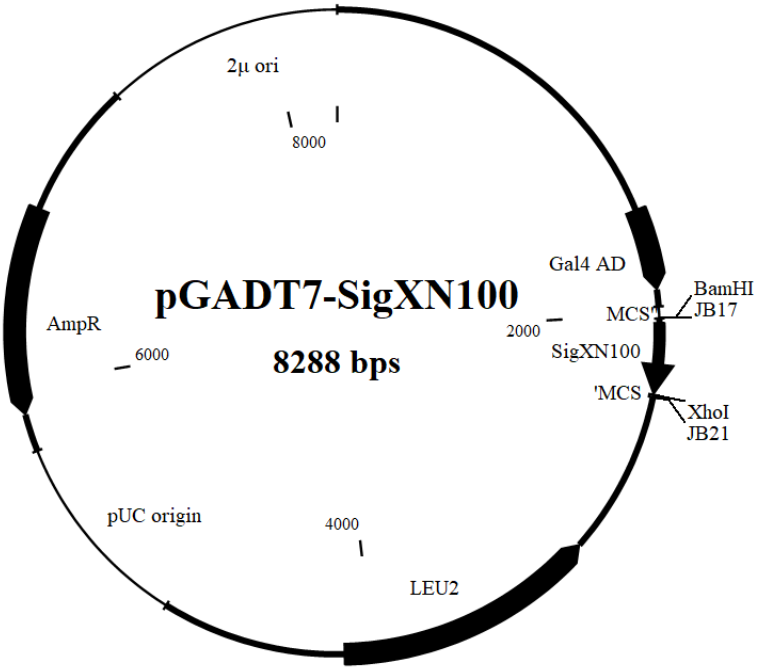


Figure 54. Construction of pMJB02 (pGADT7-*sigX*$_{N100}$). To clone *sigX*$_{N100}$ into pGADT7 vector, *sigX*$_{N100}$ was amplified using primers JB17 and JB21 and digested by BamHI and XhoI. The PCR product was cloned into pGADT7 BamHI and XhoI sites by ligation. Then the ligated DNA was transformed into *E. coli* DH5α. The correct insertion was verified by sequencing the PCR product using primer T7 and JB66.

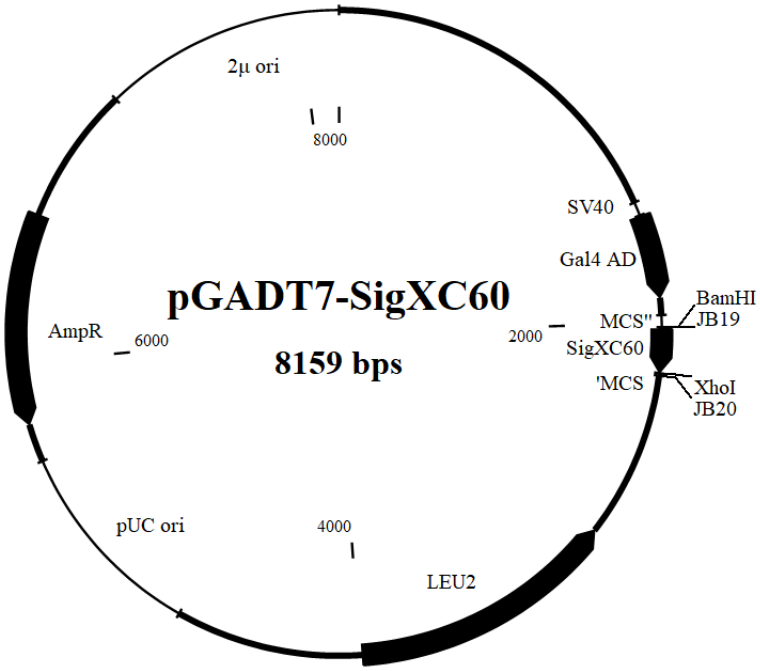


Figure 55. Construction of pMJB03 (pGADT7-*sigX$_{C60}$*). To clone *sigX$_{C60}$* into pGADT7 vector, *sigX$_{C60}$* was amplified using primers JB19 and JB20 and digested by BamHI and XhoI. The PCR product was cloned into pGADT7 BamHI and XhoI sites by ligation. Then the ligated DNA was transformed into *E. coli* DH5α. The correct insertion was verified by sequencing the PCR product using primer T7 and JB66.

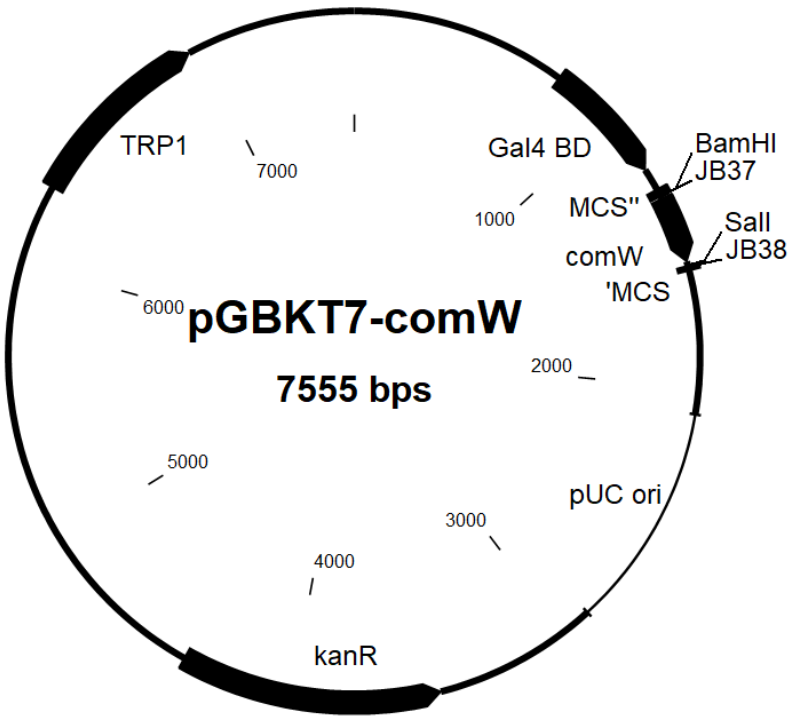


Figure 56. Construction of pMJB04 (pGBKT7-*comW*). To clone *comW* into pGBKT7 vector, *comW* was amplified using primers JB37 and JB38 and digested by BamHI and SalI. The PCR product was cloned into pGBKT7 BamHI and XhoI sites by ligation. Then the ligated DNA was transformed into *E. coli* DH5α. The correct insertion was verified by sequencing the PCR product using primer T7 and JB67.

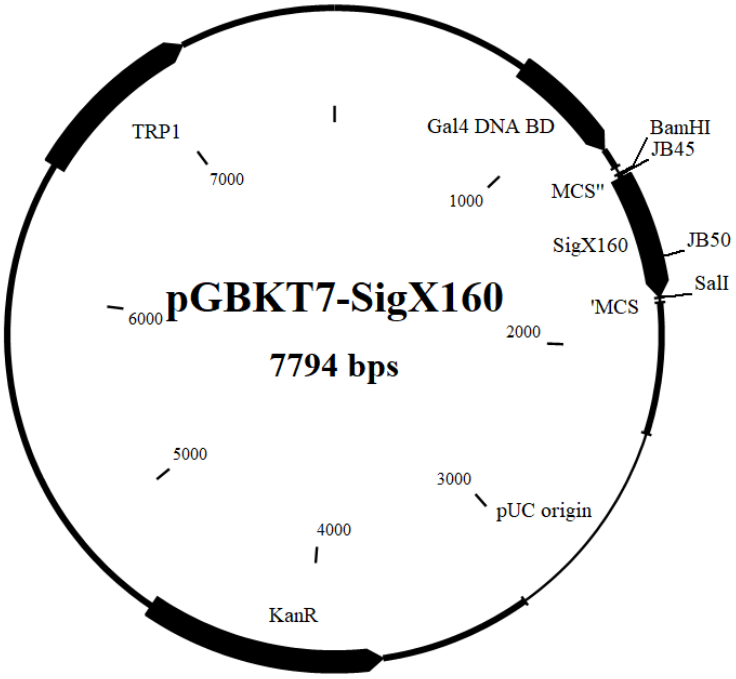


Figure 57. Construction of pMJB05 (pGBKT7-*sigX*$_{160}$). To clone *sigX*$_{160}$ into pGBKT7 vector, *sigX*$_{160}$ was amplified using primers JB45 and JB50 and digested by BamHI and SalI. The PCR product was cloned into pGBKT7 BamHI and XhoI sites by ligation. Then the ligated DNA was transformed into *E. coli* DH5α. The correct insertion was verified by sequencing the PCR product using primer T7 and JB67.

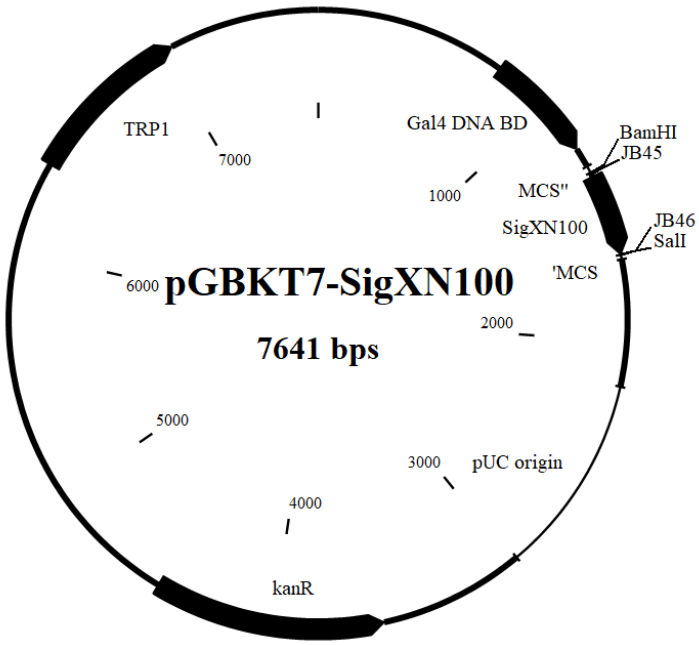


Figure 58. Construction of pMJB06 (pGBKT7-*sigX*$_{N100}$). To clone *sigX*$_{N100}$ into pGBKT7 vector, *sigX*$_{N100}$ was amplified using primers JB45 and JB46 and digested by BamHI and SalI. The PCR product was cloned into pGBKT7 BamHI and SalI sites by ligation. Then the ligated DNA was transformed into *E. coli* DH5α. The correct insertion was verified by sequencing the PCR product using primer T7 and JB67.

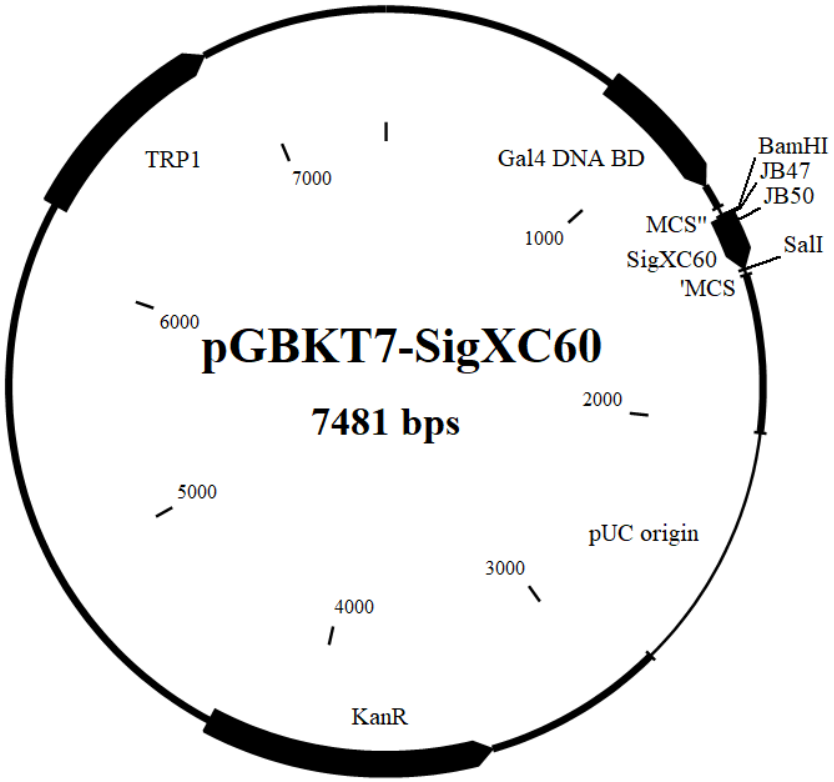


Figure 59. Construction of pMJB07 (pGBKT7-*sigX$_{C60}$*). To clone *sigX$_{C60}$* into pGBKT7 vector, *sigX$_{C60}$* was amplified using primers JB47 and JB50 and digested by BamHI and SalI. The PCR product was cloned into pGBKT7 BamHI and SalI sites by ligation. Then the ligated DNA was transformed into *E. coli* DH5α. The correct insertion was verified by sequencing the PCR product using primer T7 and JB67.

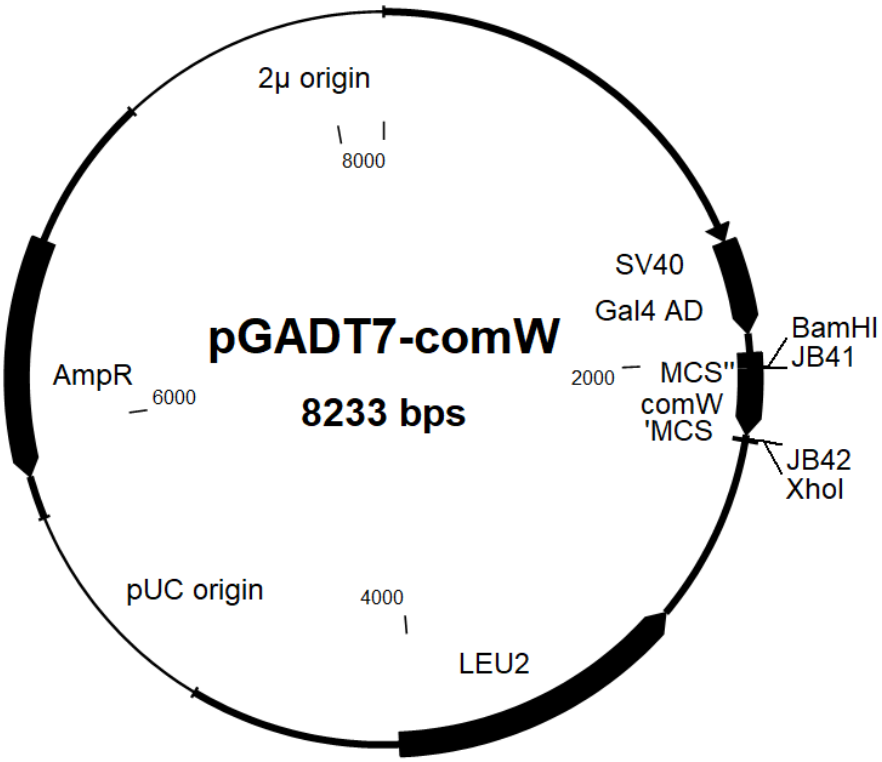


Figure 60. Construction of pMJB08 (pGADT7-*comW*). To clone *comW* into pGADT7 vector, *comW* was amplified using primers JB41 and JB42 and digested by BamHI and XhoI. The PCR product was cloned into pGADT7 BamHI and XhoI sites by ligation. Then the ligated DNA was transformed into *E. coli* DH5α. The correct insertion was verified by sequencing the PCR product using primer T7 and JB66.

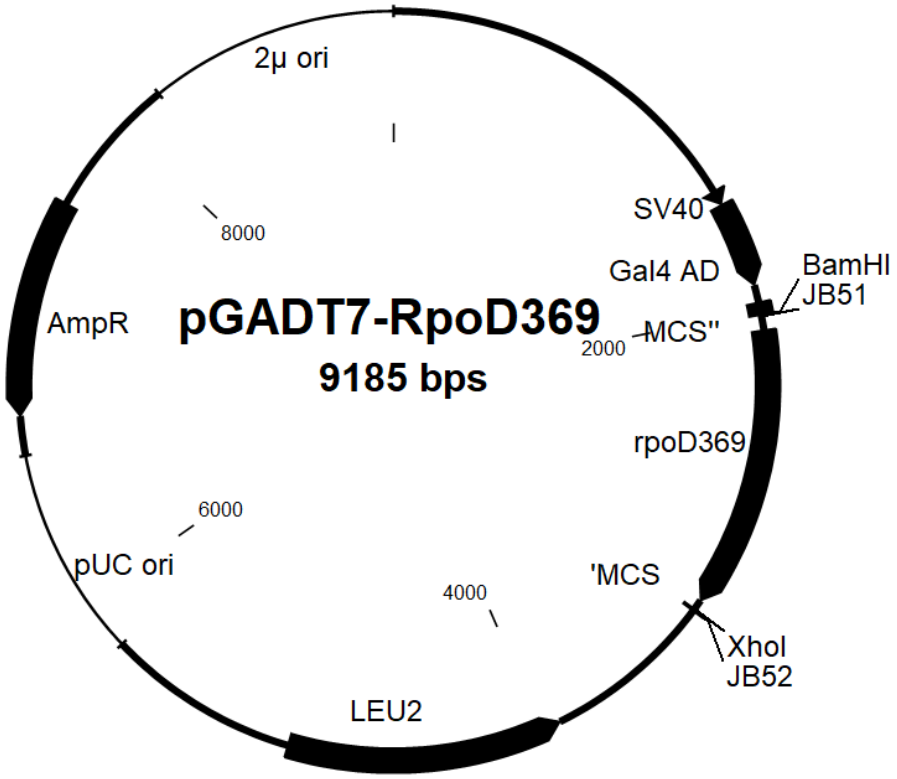


Figure 61. Construction of pMJB09 (pGADT7-*rpoD*$_{369}$). To clone *rpoD*$_{369}$ into pGADT7 vector, *rpoD*$_{369}$ was amplified using primers JB51 and JB52 and digested by BamHI and XhoI. The PCR product was cloned into pGADT7 BamHI and XhoI sites by ligation. Then the ligated DNA was transformed into *E. coli* DH5α. The correct insertion was verified by sequencing the PCR product using primer T7 and JB66.

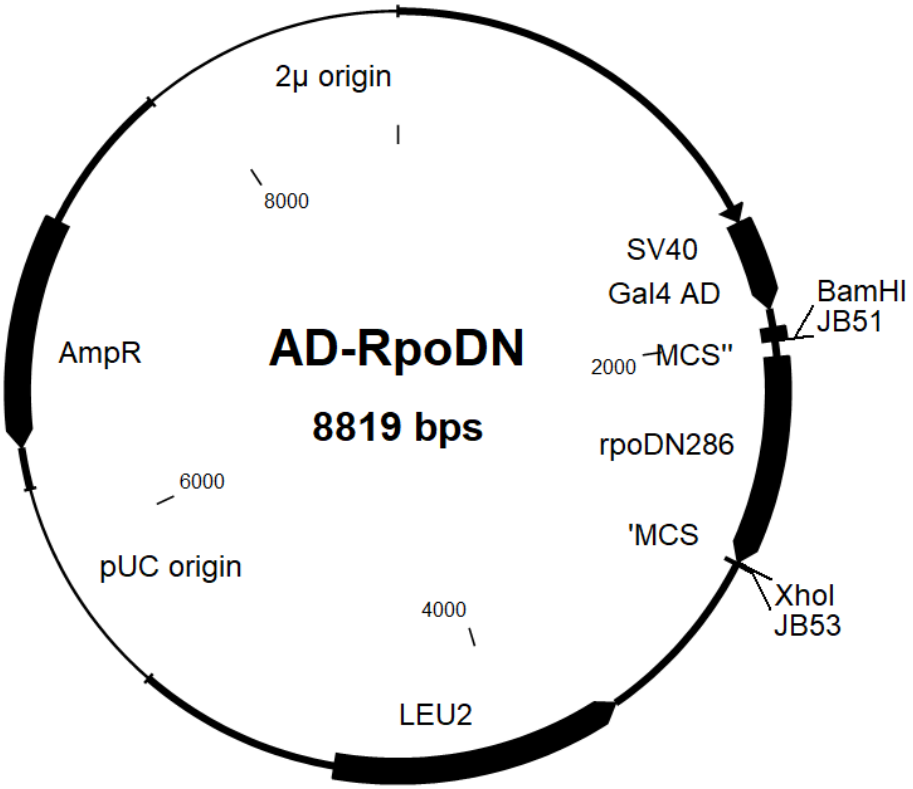


Figure 62. Construction of pMJB10 (pGADT7-*rpoD*$_{N286}$). To clone *rpoD*$_{N286}$ into pGADT7 vector, *rpoD*$_{N286}$ was amplified using primers JB51 and JB53 and digested by BamHI and XhoI. The PCR product was cloned into pGADT7 BamHI and XhoI sites by ligation. Then the ligated DNA was transformed into *E. coli* DH5α. The correct insertion was verified by sequencing the PCR product using primer T7 and JB66.

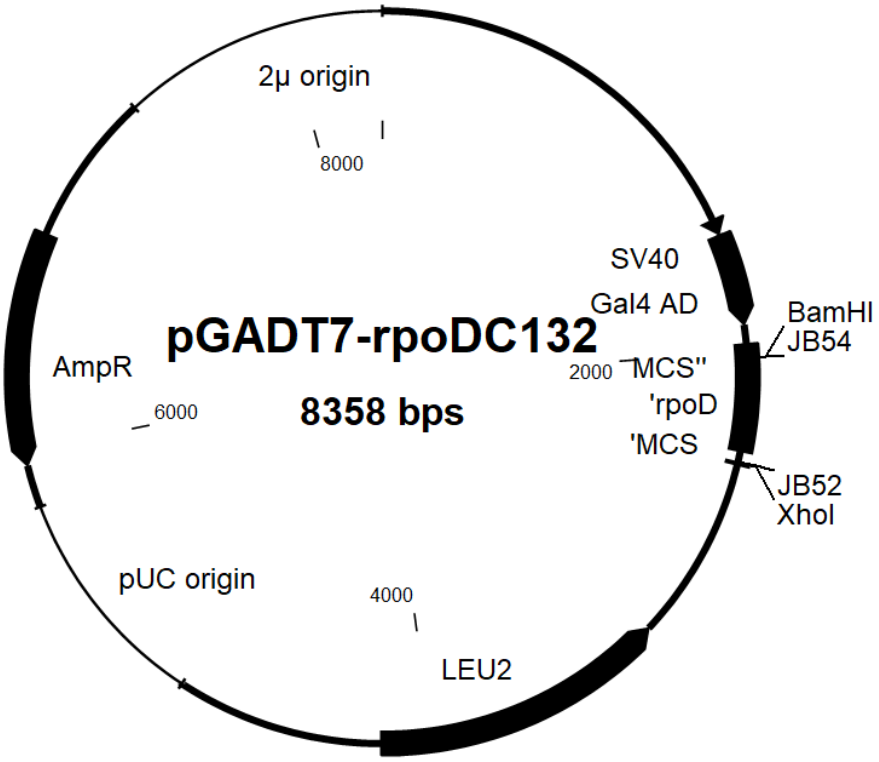


Figure 63. Construction of pMJB11 (pGADT7-*rpoD*$_{C132}$). To clone *rpoD*$_{C132}$ into pGADT7 vector, *rpoD*$_{C132}$ was amplified using primers JB52 and JB54 and digested by BamHI and XhoI. The PCR product was cloned into pGADT7 BamHI and XhoI sites by ligation. Then the ligated DNA was transformed into *E. coli* DH5α. The correct insertion was verified by sequencing the PCR product using primer T7 and JB66.

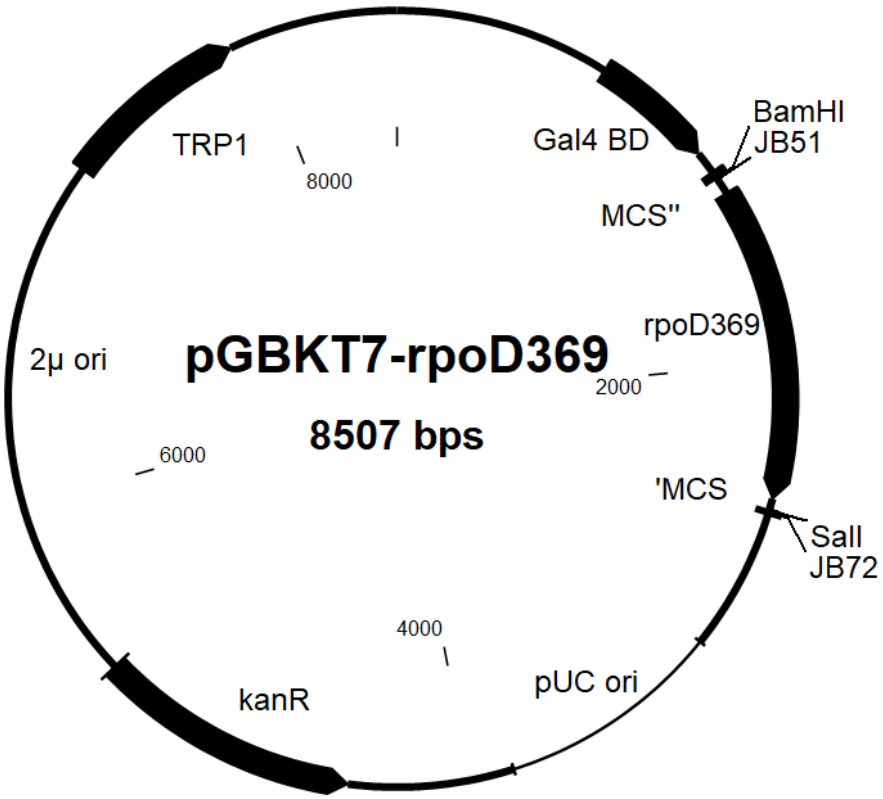


Figure 64. Construction of pMJB12 (pGBKT7-*rpoD*$_{369}$). To clone *rpoD*$_{369}$ into pGBKT7 vector, *rpoD*$_{369}$ was amplified using primers JB51and JB72 and digested by BamHI and SalI. The PCR product was cloned into pGBKT7 BamHI and SalI sites by ligation. Then the ligated DNA was transformed into *E. coli* DH5α. The correct insertion was verified by sequencing the PCR product using primer T7 and JB67.

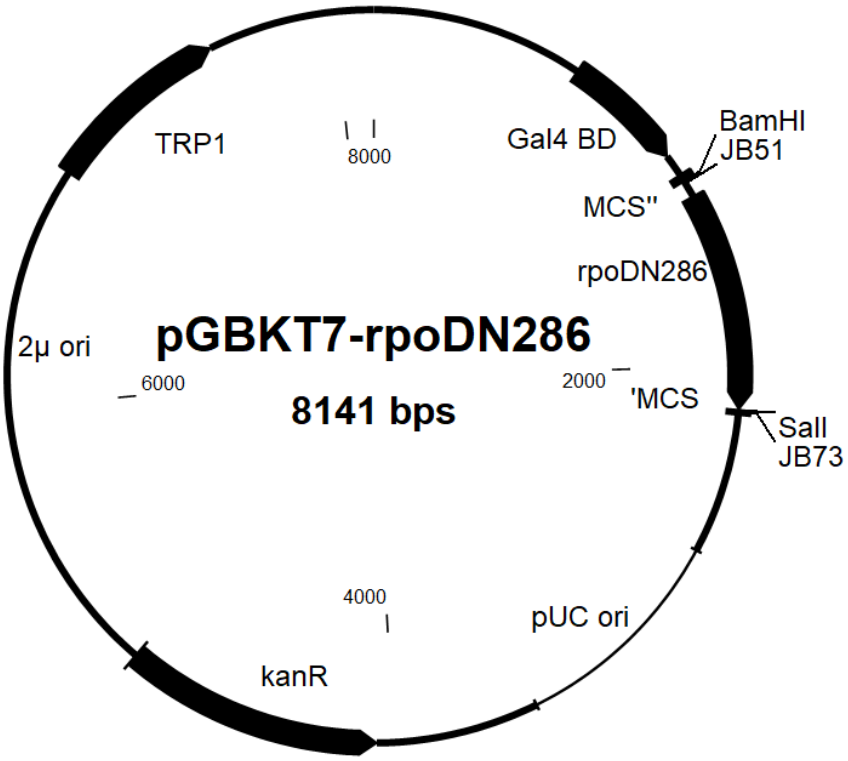


Figure 65. Construction of pMJB13 (pGBKT7-*rpoD*$_{N286}$). To clone *rpoD*$_{N286}$ into pGBKT7 vector, *rpoD*$_{N286}$ was amplified using primers JB51and JB73 and digested by BamHI and SalI. The PCR product was cloned into pGBKT7 BamHI and SalI sites by ligation. Then the ligated DNA was transformed into *E. coli* DH5α. The correct insertion was verified by sequencing the PCR product using primer T7 and JB67.

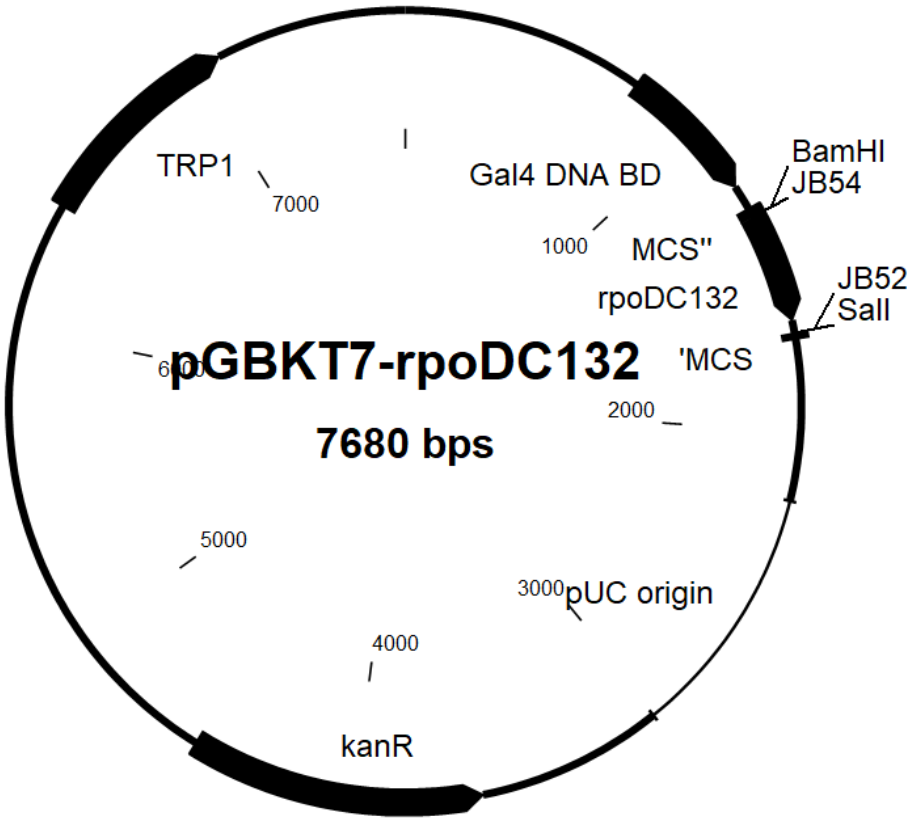


Figure 66. Construction of pMJB14 (pGBKT7-*rpoD*$_{C132}$). To clone *rpoD*$_{C132}$ into pGBKT7 vector, *rpoD*$_{C132}$ was amplified using primers JB52 and JB54 and digested by BamHI and SalI. The PCR product was cloned into pGBKT7 BamHI and SalI sites by ligation. Then the ligated DNA was transformed into *E. coli* DH5α. The correct insertion was verified by sequencing the PCR product using primer T7 and JB67.

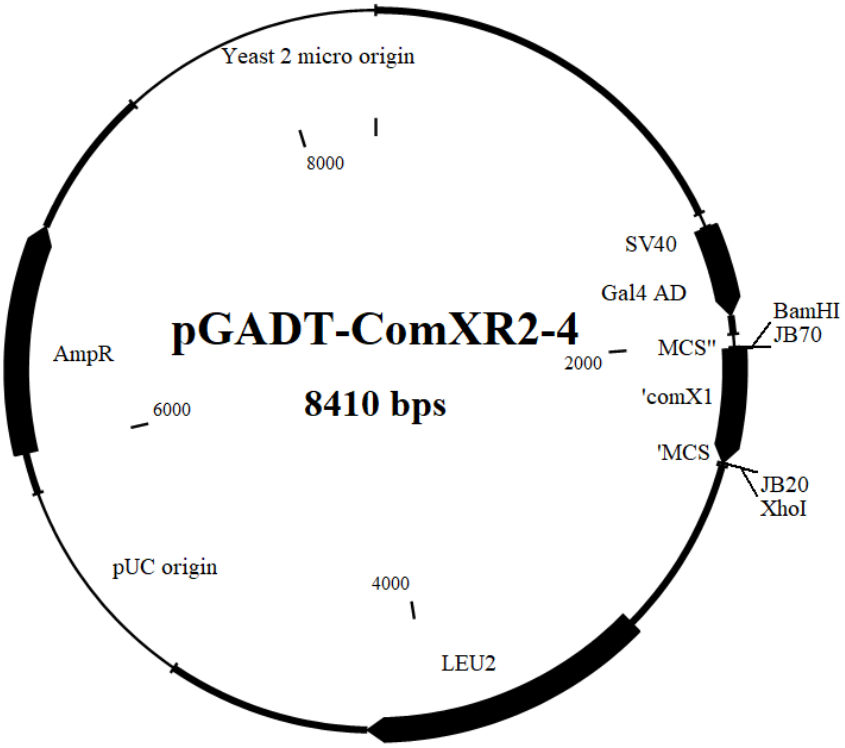


Figure 67. Construction of pMJB15 (pGADT7-*comX*$_{R2-4}$). To clone *comX*$_{R2-4}$ into pGADT7 vector, *comX*$_{R2-4}$ was amplified using primers JB20 and JB70 and digested by BamHI and XhoI. The PCR product was cloned into pGADT7 BamHI and XhoI sites by ligation. Then the ligated DNA was transformed into *E. coli* DH5α. The correct insertion was verified by sequencing the PCR product using primer T7 and JB66.

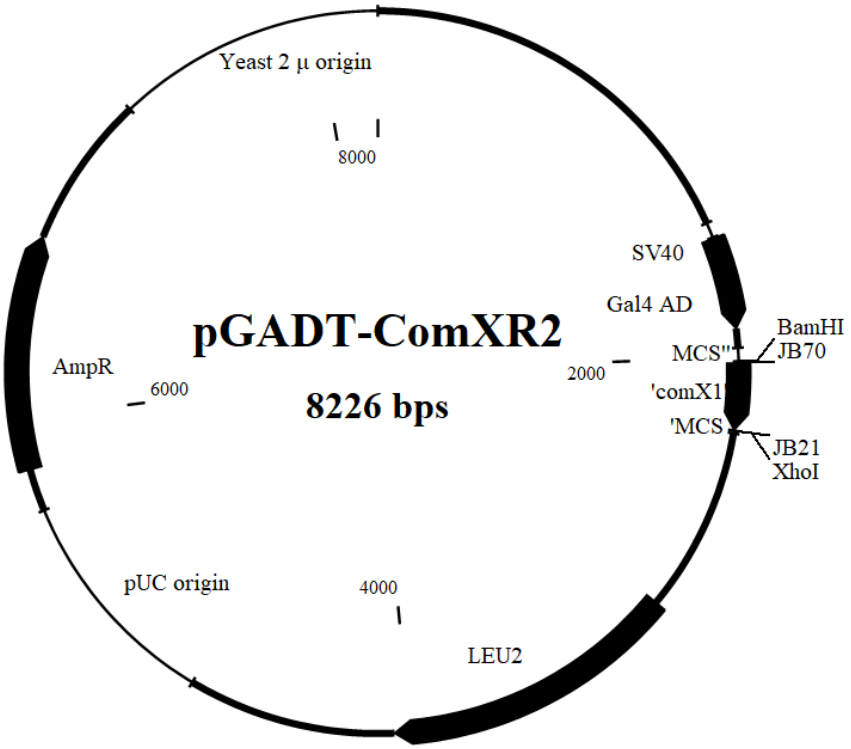


Figure 68. Construction of pMJB16 (pGADT7-*comX*$_{R2}$). To clone *comX*$_{R2}$ into pGADT7 vector, *comX*$_{R2}$ was amplified using primers JB21 and JB70 and digested by BamHI and XhoI. The PCR product was cloned into pGADT7 BamHI and XhoI sites by ligation. Then the ligated DNA was transformed into *E. coli* DH5α. The correct insertion was verified by sequencing the PCR product using primer T7 and JB66.

CITED LITERATURE

Adams, M. H., & Roe, A. S. (1945). A Partially Defined Medium for Cultivation of Pneumococcus. *Journal of Bacteriology, 49*(4), 401-409.

Ahmed, F. E., Setlow, R. B., Grist, E., & Setlow, N. (1993). DNA-damage, photorepair, and survival in fish and human-cells exposed to uv-radiation. *Environmental and Molecular Mutagenesis, 22*(1), 18-25. doi:10.1002/em.2850220105

Aho, S., Arffman, A., Pummi, T., & Uitto, J. (1997). A novel reporter gene MEL1 for the yeast two-hybrid system. *Analytical Biochemistry, 253*(2), 270-272. doi:10.1006/abio.1997.2394

Akanuma, G., Kazo, Y., Tagami, K., Hiraoka, H., Yano, K., Suzuki, S., Kawamura, F. (2016). Ribosome dimerization is essential for the efficient regrowth of Bacillus subtilis. *Microbiology-Sgm, 162*, 448-458. doi:10.1099/mic.0.000234

Akiyama, T., Williamson, K. S., Schaefer, R., Pratt, S., Chang, C. B., & Franklin, M. J. (2017). Resuscitation of Pseudomonas aeruginosa from dormancy requires hibernation promoting factor (PA4463) for ribosome preservation. *Proceedings of the National Academy of Sciences of the United States of America, 114*(12), 3204-3209. doi:10.1073/pnas.1700695114

Alloing, G., Martin, B., Granadel, C., & Claverys, J. P. (1998). Development of competence in Streptococcus pneumoniae: pheromone autoinduction and control of quorum sensing by the oligopeptide permease. *Molecular Microbiology, 29*(1), 75-83. doi:10.1046/j.1365-2958.1998.00904.x

Aubry-Damon, H., Soussy, C. J., & Courvalin, P. (1998). Characterization of mutations in the rpoB gene that confer rifampin resistance in Staphylococcus aureus. *Antimicrobial Agents and Chemotherapy, 42*(10), 2590-2594.

Banta, A. B., Chumanov, R. S., Yuan, A. H., Lin, H., Campbell, E. A., Burgess, R. R., & Gourse, R. L. (2013). Key features of sigma(S) required for specific recognition by Crl, a transcription factor promoting assembly of RNA polymerase holoenzyme. *Proceedings of the National Academy of Sciences of the United States of America, 110*(40), 15955-15960. doi:10.1073/pnas.1311642110

Banta, A. B., Cuff, M. E., Lin, H., Myers, A. R., Ross, W., Joachimiak, A., & Gourse, R. L. (2014). Structure of the RNA Polymerase Assembly Factor Crl and Identification of Its Interaction Surface with Sigma S. *Journal of Bacteriology, 196*(18), 3279-3288. doi:10.1128/jb.01910-14

Bartilson, M., Marra, A., Christine, J., Asundi, J. S., Schneider, W. P., & Hromockyj, A. E. (2001). Differential fluorescence induction reveals Streptococcus pneumoniae loci regulated by competence stimulatory peptide. *Molecular Microbiology, 39*(1), 126-135. doi:10.1046/j.1365-2958.2001.02218.x

Basu, A., & Yap, M. N. (2016). Ribosome hibernation factor promotes Staphylococcal survival and differentially represses translation. *Nucleic Acids Res, 44*(10), 4881-4893. doi:10.1093/nar/gkw180

Beckert, B., Abdelshahid, M., Schafer, H., Steinchen, W., Arenz, S., Berninghausen, O., Wilson, D. N. (2017). Structure of the Bacillus subtilis hibernating 100S ribosome reveals the basis for 70S dimerization. *Embo Journal, 36*(14), 2061-2072. doi:10.15252/embj.201696189

Bijlsma, J. J. E., Burghout, P., Moosterman, T. G., Bootsma, F. J., de Jong, A., Hermans, P. W. M., & Kuipers, O. P. (2007). Development of genomic array footprinting for identification

of conditionally essential genes in Streptococcus pneumoniae. *Applied and Environmental Microbiology, 73*(5), 1514-1524. doi:10.1128/aem.01900-06

Bonnin, R. A., & Bouloc, P. (2015). RNA Degradation in Staphylococcus aureus: Diversity of Ribonucleases and Their Impact. *International Journal of Genomics*. doi:10.1155/2015/395753

Bosdriesz, E., Molenaar, D., Teusink, B., & Bruggeman, F. J. (2015). How fast-growing bacteria robustly tune their ribosome concentration to approximate growth-rate maximization. *Febs Journal, 282*(10), 2029-2044. doi:10.1111/febs.13258

Bradford, M. M. (1976). rapid and sensitive method for quantitation of microgram quantities of protein utilizing principle of protein-dye binding. *analytical Biochemistry, 72*(1-2), 248-254. doi:10.1016/0003-2697(76)90527-3

Browning, D. F., & Busby, S. J. W. (2016). Local and global regulation of transcription initiation in bacteria. *Nature Reviews Microbiology, 14*(10), 638-650. doi:10.1038/nrmicro.2016.103

Burghout, P., Bootsma, H. J., Kloosterman, T. G., Bijlsma, J. J. E., de Jongh, C. E., Kuipers, O. P., & Hermans, P. W. M. (2007). Search for genes essential for pneumococcal transformation: The RadA DNA repair protein plays a role in genomic recombination of donor DNA. *Journal of Bacteriology, 189*(18), 6540-6550. doi:10.1128/jb.00573-07

Cadwell, R. C., & Joyce, G. F. (1994). Mutagenic PCR. *Pcr-Methods and Applications, 3*(6), S136-S140.

Campbell, E. A., Korzheva, N., Mustaev, A., Murakami, K., Nair, S., Goldfarb, A., & Darst, S. A. (2001). Structural mechanism for rifampicin inhibition of bacterial rna polymerase. *Cell, 104*(6), 901-912.

Campbell, E. A., Muzzin, O., Chlenov, M., Sun, J. L., Olson, C. A., Weinman, O., . . . Darst, S. A. (2002). Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Molecular Cell, 9*(3), 527-539. doi:10.1016/s1097-2765(02)00470-7

Campbell, E. A., Tupy, J. L., Gruber, T. M., Wang, S., Sharp, M. M., Gross, C. A., & Darst, S. A. (2003). Crystal structure of Escherichia coli sigmaE with the cytoplasmic domain of its anti-sigma RseA. *Mol Cell, 11*(4), 1067-1078.

Cao, M., Kobel, P. A., Morshedi, M. M., Wu, M. F. W., Paddon, C., & Helmann, J. D. (2002). Defining the Bacillus subtilis sigma(W) regulon: A comparative analysis of promoter consensus search, run-off transcription/macroarray analysis (ROMA), and transcriptional profiling approaches. *Journal of Molecular Biology, 316*(3), 443-457. doi:10.1006/jmbi.2001.5372

Chang, J. C., LaSarre, B., Jimenez, J. C., Aggarwal, C., & Federle, M. J. (2011). Two Group A Streptococcal Peptide Pheromones Act through Opposing Rgg Regulators to Control Biofilm Development. *PLoS Pathog, 7*(8). doi:10.1371/journal.ppat.1002190

Charpentier, X., Kay, E., Schneider, D., & Shuman, H. A. (2011). Antibiotics and UV Radiation Induce Competence for Natural Transformation in Legionella pneumophila. *Journal of Bacteriology, 193*(5), 1114-1121. doi:10.1128/jb.01146-10

Cho, B.-K., Kim, D., Knight, E. M., Zengler, K., & Palsson, B. O. (2014). Genome-scale reconstruction of the sigma factor network in Escherichia coli: topology and functional states. *Bmc Biology, 12*. doi:10.1186/1741-7007-12-4

Claverys, J. P., & Havarstein, L. S. (2002). Extracellular-peptide control of competence for genetic transformation in Streptococcus pneumoniae. *Frontiers in Bioscience, 7*, D1798-D1814. doi:10.2741/claverys

Cook, L. C., & Federle, M. J. (2014). Peptide pheromone signaling in Streptococcus and Enterococcus. *FEMS Microbiol Rev, 38*(3), 473-492. doi:10.1111/1574-6976.12046

Dahl, M. K., Msadek, T., Kunst, F., & Rapoport, G. (1991). Mutational analysis of the bacillus-subtilis degu regulator and its phosphorylation by the degs protein-kinase. *Journal of Bacteriology, 173*(8), 2539-2547.

Daubin, V., & Szoellosi, G. J. (2016). Horizontal Gene Transfer and the History of Life. *Cold Spring Harbor Perspectives in Biology, 8*(4). doi:10.1101/cshperspect.a018036

Declerck, N., Bouillaut, L., Chaix, D., Rugani, N., Slamti, L., Hoh, F., Arold, S. T. (2007). Structure of PlcR: Insights into virulence regulation and evolution of quorum sensing in Gram-positive bacteria. *Proceedings of the National Academy of Sciences of the United States of America, 104*(47), 18490-18495. doi:10.1073/pnas.0704501104

Elkins, C., Thomas, C. E., Seifert, H. S., & Sparling, P. F. (1991). Species-specific uptake of dna by gonococci is mediated by a 10-base-pair sequence. *Journal of Bacteriology, 173*(12), 3911-3913.

Enright, M., Zawadski, P., Pickerill, P., & Dowson, C. G. (1998). Molecular evolution of rifampicin resistance in Streptococcus pneumoniae. *Microbial Drug Resistance-Mechanisms Epidemiology and Disease, 4*(1), 65-70. doi:10.1089/mdr.1998.4.65

Fields, S., & Song, O. K. (1989). A novel genetic system to detect protein protein interactions. *Nature, 340*(6230), 245-246. doi:10.1038/340245a0

Finkel, S. E., & Kolter, R. (2001). DNA as a nutrient novel role for bacterial competence gene homologs. *Journal of Bacteriology, 183*(21), 6288-6293. doi:10.1128/jb.183.21.6288-6293.2001

Fontaine, L., Goffin, P., Dubout, H., Delplace, B., Baulard, A., Lecat-Guillet, N., Hols, P. (2013). Mechanism of competence activation by the ComRS signalling system in streptococci. *Molecular Microbiology, 87*(6), 1113-1132. doi:10.1111/mmi.12157

Fontaine, L., Wahl, A., Flechard, M., Mignolet, J., & Hols, P. (2015). Regulation of competence for natural transformation in streptococci. *Infect Genet Evol, 33*, 343-360. doi:10.1016/j.meegid.2014.09.010

Francez-Charlot, A., Frunzke, J., Reichen, C., Ebneter, J. Z., Gourion, B., & Vorholt, J. A. (2009). Sigma factor mimicry involved in regulation of general stress response. *Proceedings of the National Academy of Sciences of the United States of America, 106*(9), 3467-3472. doi:10.1073/pnas.0810291106

Frischer, M. E., Thurmond, J. M., & Paul, J. H. (1993). Factors affecting competence in a high-frequency of transformation marine vibrio. *journal of general microbiology, 139*, 753-761.

Gamba, P., Jonker, M. J., & Hamoen, L. W. (2015). A Novel Feedback Loop That Controls Bimodal Expression of Genetic Competence. *Plos Genetics, 11*(6). doi:10.1371/journal.pgen.1005047

Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., III, & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods, 6*(5), 343-U341. doi:10.1038/nmeth.1318

Gibson, F., Anderson, L., Babnigg, G., Baker, M., Berth, M., Binz, P. A., Jones, A. R. (2008). Guidelines for reporting the use of gel electrophoresis in proteomics. *Nature Biotechnology, 26*(8), 863-864. doi:10.1038/nbt0808-863

Gietz, R. D., & Woods, R. A. (2002). Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. In C. Guthrie & G. R. Fink (Eds.), *Guide to Yeast Genetics and Molecular and Cell Biology, Pt B* (Vol. 350, pp. 87-96).

Goddard, A., Ladds, G., & Davey, J. (2005). Development of a semi-quantitative plate-based alpha-galactosidase gene reporter for Schizosaccharomyces pombe and its use to isolate a constitutively active Mam2. *Yeast, 22*(1), 31-41. doi:10.1002/yea.1190

Gonzalez, R. L. (2017). Survivor: Ribosome Edition. *Embo Journal, 36*(14), 1996-1998. doi:10.15252/embj.201797518

Gourion, B., Francez-Charlot, A., & Vorholt, J. A. (2008). PhyR is involved in the general stress response of Methylobacterium extorquens AM1. *Journal of Bacteriology, 190*(3), 1027-1035. doi:10.1128/jb.01483-07

Gross, C. A., Chan, C., Dombroski, A., Gruber, T., Sharp, M., Tupy, J., & Young, B. (1998). The functional and regulatory roles of sigma factors in transcription. *Cold Spring Harbor Symposia on Quantitative Biology, 63*, 141-155. doi:10.1101/sqb.1998.63.141

Gruber, T. M., & Gross, C. A. (2003). Multiple sigma subunits and the partitioning of bacterial transcription space. *Annual Review of Microbiology, 57*, 441-466. doi:10.1146/annurev.micro.57.030502.090913

Hahn, J., Luttinger, A., & Dubnau, D. (1996). Regulatory inputs for the synthesis of ComK, the competence transcription factor of Bacillus subtilis. *Molecular Microbiology, 21*(4), 763-775. doi:10.1046/j.1365-2958.1996.371407.x

Hahn, J., Roggiani, M., & Dubnau, D. (1995). The major role of spooa in genetic competence is to down-regulate abrb, an essential competence gene. *Journal of bacteriology, 177*(12), 3601-3605.

Hamilton, H. L., & Dillard, J. P. (2006). Natural transformation of Neisseria gonorrhoeae: from DNA donation to homologous recombination. *Molecular Microbiology, 59*(2), 376-385. doi:10.1111/j.1365-2958.2005.04964.x

Haugen, S. P., Ross, W., & Gourse, R. L. (2008). Advances in bacterial promoter recognition and its control by factors that do not bind DNA. *Nature Reviews Microbiology, 6*(7), 507-519. doi:10.1038/nrmicro1912

Havarstein, L. S. (2010). Increasing competence in the genus Streptococcus. *Molecular Microbiology, 78*(3), 541-544. doi:10.1111/j.1365-2958.2010.07380.x

Havarstein, L. S., Coomaraswamy, G., & Morrison, D. A. (1995). An unmodified heptadecapeptide pheromone induces competence for genetic-transformation in streptococcus-pneumoniae. *Proceedings of the National Academy of Sciences of the United States of America, 92*(24), 11140-11144. doi:10.1073/pnas.92.24.11140

Havarstein, L. S., Gaustad, P., Nes, I. F., & Morrison, D. A. (1996). Identification of the streptococcal competence-pheromone receptor. *Molecular Microbiology, 21*(4), 863-869. doi:10.1046/j.1365-2958.1996.521416.x

Hazan, R., Sat, B., & Engelberg-Kulka, H. (2004). *Escherichia coli* mazEF-mediated cell death is triggered by various stressful conditions. *Journal of Bacteriology, 186*(11), 3663-3669. doi:10.1128/jb.186.11.3663-3669.2004

Helmann, J. D. (2016). Bacillus subtilis extracytoplasmic function (ECF) sigma factors and defense of the cell envelope. *Current Opinion in Microbiology, 30*, 122-132. doi:10.1016/j.mib.2016.02.002

HenggeAronis, R. (1996). Back to log phase: sigma(s) as a global regulator in the osmotic control of gene expression in Escherichia coli. *Molecular Microbiology, 21*(5), 887-893. doi:10.1046/j.1365-2958.1996.511405.x

Hershberger, C. D., Ye, R. W., Parsek, M. R., Xie, Z. D., & Chakrabarty, A. M. (1995). The algt (algu) gene of pseudomonas-aeruginosa, a key regulator involved in alginate biosynthesis,

encodes an alternative sigma-factor (sigma(e)). *Proceedings of the National Academy of Sciences of the United States of America, 92*(17), 7941-7945. doi:10.1073/pnas.92.17.7941

Hoa, T. T., Tortosa, P., Albano, M., & Dubnau, D. (2002). Rok (YkuW) regulates genetic competence in Bacillus subtilis by directly repressing comK. *Molecular Microbiology, 43*(1), 15-26. doi:10.1046/j.1365-2958.2002.02727.x

Ishihama, A. (2000). Functional modulation of Escherichia coli RNA polymerase. *Annual Review of Microbiology, 54*, 499-518. doi:10.1146/annurev.micro.54.1.499

Jimenez, J. C., & Federle, M. J. (2014). Quorum sensing in group A Streptococcus. *Front Cell Infect Microbiol, 4*, 127. doi:10.3389/fcimb.2014.00127

Jin, D. J., & Gross, C. A. (1988). Mapping and sequencing of mutations in the escherichia-coli rpob gene that lead to rifampicin resistance. *Journal of Molecular Biology, 202*(1), 45-58. doi:10.1016/0022-2836(88)90517-7

Johnston, C., Martin, B., Fichant, G., Polard, P., & Claverys, J. P. (2014). Bacterial transformation: distribution, shared mechanisms and divergent control. *Nat Rev Microbiol, 12*(3), 181-196. doi:10.1038/nrmicro3199

Jyssum, K. (1966). Genetic factors determining competence in transformation of Neisseria meningitidis. II. A reversible loss of competence. *Acta pathologica et microbiologica Scandinavica, 67*(4), 493-502.

Kawamura, Y., Hou, X. G., Sultana, F., Miura, H., & Ezaki, T. (1995). Determination of 16s ribosomal-rna sequences of streptococcus-mitis and streptococcus-gordonii and phylogenetic-relationships among members of the genus streptococcus. *international Journal of Systematic Bacteriology, 45*(2), 406-408.

Keegan, L., Gill, G., & Ptashne, M. (1986). Separation of dna-binding from the transcription-activating function of a eukaryotic regulatory protein. *Science, 231*(4739), 699-704. doi:10.1126/science.3080805

Keiler, K. C. (2015). Mechanisms of ribosome rescue in bacteria. *Nat Rev Microbiol, 13*(5), 285-297. doi:10.1038/nrmicro3438

Khan, R., Junges, R., Amdal, H. A., Chen, T., Morrison, D. A., & Petersen, F. C. (2017). A positive feedback loop mediated by Sigma X enhances expression of the streptococcal regulator ComR. *Scientific Reports, 7*. doi:10.1038/s41598-017-04768-5

Khan, R., Rukke, H. V., Hovik, H., Amdal, H. A., Chen, T., Morrison, D. A., & Petersen, F. C. (2016). Comprehensive Transcriptome Profiles of Streptococcus mutans UA159 Map Core Streptococcal Competence Genes. *mSystems, 1*(2). doi:10.1128/mSystems.00038-15

Khusainov, I., Vicens, Q., Ayupov, R., Usachev, K., Myasnikov, A., Simonetti, A., . . . Hashem, Y. (2017). Structures and dynamics of hibernating ribosomes from Staphylococcus aureus mediated by intermolecular interactions of HPF. *Embo Journal, 36*(14), 2073-2087. doi:10.15252/embj.201696105

Kline, B. C., McKay, S. L., Tang, W. W., & Portnoy, D. A. (2015). The Listeria monocytogenes hibernation-promoting factor is required for the formation of 100S ribosomes, optimal fitness, and pathogenesis. *J Bacteriol, 197*(3), 581-591. doi:10.1128/JB.02223-14

Landini, P., Egli, T., Wolf, J., & Lacour, S. (2014). sigmaS, a major player in the response to environmental stresses in Escherichia coli: role, regulation and mechanisms of promoter recognition. *Environmental Microbiology Reports, 6*(1), 1-13. doi:10.1111/1758-2229.12112

Lee, M. S., Dougherty, B. A., Madeo, A. C., & Morrison, D. A. (1999). Construction and analysis of a library for random insertional mutagenesis in Streptococcus pneumoniae: Use for

recovery of mutants defective in genetic transformation and for identification of essential genes. *Applied and Environmental Microbiology, 65*(5), 1883-1890.

Lee, M. S., & Morrison, D. A. (1999). Identification of a new regulator in Streptococcus pneumoniae linking quorum sensing to competence for genetic transformation. *Journal of Bacteriology, 181*(16), 5004-5016.

Lonetto, M., Gribskov, M., & Gross, C. A. (1992). The sigma-70 family - sequence conservation and evolutionary relationships. *Journal of Bacteriology, 174*(12), 3843-3849.

Lonetto, M. A., Brown, K. L., Rudd, K. E., & Buttner, M. J. (1994). Analysis of the streptomyces-coelicolor sige gene reveals the existence of a subfamily of eubacterial rna-polymerase sigma-factors involved in the regulation of extracytoplasmic functions. *Proceedings of the National Academy of Sciences of the United States of America, 91*(16), 7573-7577. doi:10.1073/pnas.91.16.7573

Luo, P., Li, H. Y., & Morrison, D. A. (2003). ComX is a unique link between multiple quorum sensing outputs and competence in Streptococcus pneumoniae. *Molecular Microbiology, 50*(2), 623-633. doi:10.1046/j.1365-2958.2003.03714.x

Luo, P., Li, H. Y., & Morrison, D. A. (2004). Identification of ComW as a new component in the regulation of genetic transformation in Streptococcus pneumoniae. *Molecular Microbiology, 54*(1), 172-183. doi:10.1111/j.1365-2958.2004.04254.x

Luo, P., & Morrison, D. A. (2003). Transient association of an alternative sigma factor, ComX, with RNA polymerase during the period of competence for genetic transformation in Streptococcus pneumoniae. *Journal of Bacteriology, 185*(1), 349-358. doi:10.1128/jb.185.1.349-358

Malhotra, A., Severinova, E., & Darst, S. A. (1996). Crystal structure of a sigma(70) subunit fragment from *E.coli* RNA polymerase. *Cell, 87*(1), 127-136. doi:10.1016/s0092-8674(00)81329-x

Martin, B., Quentin, Y., Fichant, G., & Claverys, J.-P. (2006). Independent evolution of competence regulatory cascades in streptococci? *Trends in Microbiology, 14*(8), 339-345. doi:10.1016/j.tim.2006.06.007

Mashburn-Warren, L., Morrison, D. A., & Federle, M. J. (2010). A novel double-tryptophan peptide pheromone controls competence in Streptococcus spp. via an Rgg regulator. *Molecular Microbiology, 78*(3), 589-606. doi:10.1111/j.1365-2958.2010.07361.x

McKay, S. L., & Portnoy, D. A. (2015). Ribosome Hibernation Facilitates Tolerance of Stationary-Phase Bacteria to Aminoglycosides. *Antimicrobial Agents and Chemotherapy, 59*(11), 6992-6999. doi:10.1128/aac.01532-15

Meier, P. S., Utz, S., Aebi, S., & Muhlemann, K. (2003). Low-level resistance to rifampin in Streptococcus pneumoniae. *Antimicrobial Agents and Chemotherapy, 47*(3), 863-868. doi:10.1128/aac.47.3.863-868.2003

Mika, F., & Hengge, R. (2005). A two-component phosphotransfer network involving ArcB, ArcA, and RssB coordinates synthesis and proteolysis of sigma(S) (RpoS) in E-coli. *Genes & Development, 19*(22), 2770-2781. doi:10.1101/gad.353705

Monnet, V., Juillard, V., & Gardan, R. (2016). Peptide conversations in Gram-positive bacteria. *Critical Reviews in Microbiology, 42*(3), 339-351. doi:10.3109/1040841x.2014.948804

Morita, M., Kanemori, M., Yanagi, H., & Yura, T. (1999). Heat-induced synthesis of sigma(32) in Escherichia coli: Structural and functional dissection of rpoH mRNA secondary structure. *Journal of Bacteriology, 181*(2), 401-410.

Morrison, D. A. (1977). Transformation in pneumococcus - existence and properties of a complex involving donor deoxyribonucleate single strands in eclipse. *Journal of Bacteriology, 132*(2), 576-583.

Morrison, D. A. (1978). Transformation in pneumococcus - protein-content of eclipse complex. *Journal of Bacteriology, 136*(2), 548-557.

Morrison, D. A., & Baker, M. F. (1979). Competence for genetic-transformation in pneumococcus depends on synthesis of a small set of proteins. *Nature, 282*(5735), 215-217. doi:10.1038/282215a0

Morrison, D. A., Guedon, E., & Renault, P. (2013). Competence for Natural Genetic Transformation in the Streptococcus bovis Group Streptococci S. infantarius and S. macedonicus. *Journal of Bacteriology, 195*(11), 2612-2620. doi:10.1128/jb.00230-13

Morrison, D. A., & Lee, M. S. (2000). Regulation of competence for genetic transformation in Streptococcus pneumoniae: a link between quorum sensing and DNA processing genes. *Research in Microbiology, 151*(6), 445-451. doi:10.1016/s0923-2508(00)00171-6

Ng, W.-L., & Bassler, B. L. (2009). Bacterial Quorum-Sensing Network Architectures. *Annual Review of Genetics, 43*, 197-222. doi:10.1146/annurev-genet-102108-134304

Nickels, B. E., Dove, S. L., Murakami, K. S., Darst, S. A., & Hochschild, A. (2002). Protein-protein and protein-DNA interactions of sigma(70) region 4 involved in transcription activation by lambda cI. *Journal of Molecular Biology, 324*(1), 17-34. doi:10.1016/s0022-2836(02)01043-4

Ochman, H., Lawrence, J. G., & Groisman, E. A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature, 405*(6784), 299-304. doi:10.1038/35012500

Osmundson, J., Montero-Diez, C., Westblade, L. F., Hochschild, A., & Darst, S. A. (2012). Promoter-Specific Transcription Inhibition in Staphylococcus aureus by a Phage Protein. *Cell, 151*(5), 1005-1016. doi:10.1016/j.cell.2012.10.034

Osterberg, S., del Peso-Santos, T., & Shingler, V. (2011). Regulation of Alternative Sigma Factor Use. In S. Gottesman & C. S. Harwood (Eds.), *Annual Review of Microbiology, Vol 65* (Vol. 65, pp. 37-55).

Oussenko, I. A., Sanchez, R., & Bechhofer, D. H. (2002). Bacillus subtilis YhaM, a Member of a New Family of 3'-to-5' Exonucleases in Gram-Positive Bacteria. *Journal of Bacteriology, 184*(22), 6250-6259. doi:10.1128/jb.184.22.6250-6259.2002

Padayachee, T., & Klugman, K. P. (1999). Molecular basis of rifampin resistance in Streptococcus pneumoniae. *Antimicrobial Agents and Chemotherapy, 43*(10), 2361-2365.

Paget, M. S. (2015). Bacterial Sigma Factors and Anti-Sigma Factors: Structure, Function and Distribution. *Biomolecules, 5*(3), 1245-1265. doi:10.3390/biom5031245

Pandey, D. P., & Gerdes, K. (2005). Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Research, 33*(3), 966-976. doi:10.1093/nar/gki201

Papenfort, K., & Bassler, B. L. (2016). Quorum sensing signal-response systems in Gram-negative bacteria. *Nat Rev Microbiol, 14*(9), 576-588. doi:10.1038/nrmicro.2016.89

Pedersen, K., Christensen, S. K., & Gerdes, K. (2002). Rapid induction and reversal of a bacteriostatic condition by controlled expression of toxins and antitoxins. *Molecular Microbiology, 45*(2), 501-510. doi:10.1046/j.1365-2958.2002.03027.x

Pestova, E. V., Havarstein, L. S., & Morrison, D. A. (1996). Regulation of competence for genetic transformation in Streptococcus pneumoniae by an auto-induced peptide pheromone and a

two-component regulatory system. *Molecular Microbiology, 21*(4), 853-862. doi:DOI 10.1046/j.1365-2958.1996.501417.x

Peterson, S., Cline, R. T., Tettelin, H., Sharov, V., & Morrison, D. A. (2000). Gene expression analysis of the Streptococcus pneumoniae competence regulons by use of DNA microarrays. *Journal of Bacteriology, 182*(21), 6192-6202. doi:Doi 10.1128/Jb.182.21.6192-6202.2000

Peterson, S. N., Sung, C. K., Cline, R., Desai, B. V., Snesrud, E. C., Luo, P., Morrison, D. A. (2004). Identification of competence pheromone responsive genes in Streptococcus pneumoniae by use of DNA microarrays. *Molecular Microbiology, 51*(4), 1051-1070. doi:10.1046/j.1365-2958.2003.03907.x

Piotrowski, A., Luo, P., & Morrison, D. A. (2009). Competence for Genetic Transformation in Streptococcus pneumoniae: Termination of Activity of the Alternative Sigma Factor ComX Is Independent of Proteolysis of ComX and ComW. *Journal of Bacteriology, 191*(10), 3359-3366. doi:10.1128/Jb.01750-08

Piper, S. E., Mitchell, J. E., Lee, D. J., & Busby, S. J. W. (2009). A global view of Escherichia coli Rsd protein and its interactions. *Molecular Biosystems, 5*(12), 1943-1947. doi:10.1039/b904955j

Plasterk, R. H. A., Izsvak, Z., & Ivics, Z. (1999). Resident aliens - the Tc1/mariner superfamily of transposable elements. *Trends in Genetics, 15*(8), 326-332. doi:10.1016/s0168-9525(99)01777-1

Polikanov, Y. S., Blaha, G. M., & Steitz, T. A. (2012). How hibernation factors RMF, HPF, and YfiA turn off protein synthesis. *Science, 336*(6083), 915-918. doi:10.1126/science.1218538

Pozzi, G., Masala, L., Iannelli, F., Manganelli, R., Havarstein, L. S., Piccoli, L., Morrison, D. A. (1996). Competence for genetic transformation in encapsulated strains of Streptococcus pneumoniae: Two allelic variants of the peptide pheromone. *Journal of Bacteriology, 178*(20), 6087-6090.

Pratt, L. A., & Silhavy, T. J. (1998). Crl stimulates RpoS activity during stationary phase. *Molecular Microbiology, 29*(5), 1225-1236. doi:10.1046/j.1365-2958.1998.01007.x

Puri, P., Eckhardt, T. H., Franken, L. E., Fusetti, F., Stuart, M. C., Boekema, E. J., Poolman, B. (2014). Lactococcus lactis YfiA is necessary and sufficient for ribosome dimerization. *Mol Microbiol, 91*(2), 394-407. doi:10.1111/mmi.12468

Rahmer, R., Heravi, K. M., & Altenbuchner, J. (2015). Construction of a Super-Competent Bacillus subtilis 168 Using the P-mt/A-comKS Inducible Cassette. *Front Microbiol, 6*. doi:10.3389/fmicb.2015.01431

Rappaport, H. P., & Guild, W. R. (1959). Defined medium for growth of two transformable strains of Diplococcus pneumoniae. *Journal of Bacteriology, 78*, 203-205.

Redfield, R. J. (1993). Genes for breakfast-the have-your-cake-and-eat-it-too of bacterial transformation. *Journal of Heredity, 84*(5), 400-404.

Redfield, R. J., Cameron, A. D. S., Qian, Q., Hinds, J., Ali, T. R., Kroll, J. S., & Langford, P. R. (2005). A novel CRP-dependent of competence genes in regulon controls expression Haemophilus influenzae. *Journal of Molecular Biology, 347*(4), 735-747. doi:10.1016/j.jmb.2005.01.012

Redko, Y., & Condon, C. (2010). Maturation of 23S rRNA in Bacillus subtilis in the absence of Mini-III. *J Bacteriol, 192*(1), 356-359. doi:10.1128/JB.01096-09

Richards, V. P., Palmer, S. R., Bitar, P. D. P., Qin, X., Weinstock, G. M., Highlander, S. K., Stanhope, M. J. (2014). Phylogenomics and the Dynamic Genome Evolution of the Genus Streptococcus. *Genome Biology and Evolution, 6*(4), 741-753. doi:10.1093/gbe/evu048

Rosenow, C., Maniar, M., & Trias, J. (1999). Regulation of the alpha-galactosidase activity in Streptococcus pneumoniae: Characterization of the raffinose utilization system. *Genome Research, 9*(12), 1189-1197. doi:10.1101/gr.9.12.1189

Rubin, E. J., Akerley, B. J., Novik, V. N., Lampe, D. J., Husson, R. N., & Mekalanos, J. J. (1999). In vivo transposition of mariner-based elements in enteric bacteria and mycobacteria. *Proceedings of the National Academy of Sciences of the United States of America, 96*(4), 1645-1650. doi:10.1073/pnas.96.4.1645

Russell, R. R. B., Aduseopoku, J., Sutcliffe, I. C., Tao, L., & Ferretti, J. J. (1992). A binding protein-dependent transport-system in streptococcus-mutans responsible for multiple sugar metabolism. *Journal of Biological Chemistry, 267*(7), 4631-4637.

Schulz, S., Eckweiler, D., Bielecka, A., Nicolai, T., Franke, R., Dotsch, A., . . . Haussler, S. (2015). Elucidation of Sigma Factor-Associated Networks in Pseudomonas aeruginosa Reveals a Modular Architecture with Limited and Function-Specific Crosstalk. *PLoS Pathog, 11*(3). doi:10.1371/journal.ppat.1004744

Schurr, M. J., Yu, H., MartinezSalazar, J. M., Boucher, J. C., & Deretic, V. (1996). Control of AlgU, a member of the sigma(E)-like family of stress sigma factors, by the negative regulators MucA and MucB and Pseudomonas aeruginosa conversion to mucoidy in cystic fibrosis. *Journal of Bacteriology, 178*(16), 4997-5004.

Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z., & Hwa, T. (2010). Interdependence of Cell Growth and Gene Expression: Origins and Consequences. *Science, 330*(6007), 1099-1102. doi:10.1126/science.1192588

Sharma, U. K., & Chatterji, D. (2010). Transcriptional switching in Escherichia coli during stress and starvation by modulation of Sigma 70 activity. *FEMS Microbiol Rev, 34*(5), 646-657. doi:10.1111/j.1574-6976.2010.00223.x

Sicard, A. M. (1964). A new synthetic medium for diplococcus pneumoniae, and its use for the study of reciprocal transformations at the amia locus. *Genetics, 50*, 31-44.

Simanshu, D. K., Yamaguchi, Y., Park, J.-H., Inouye, M., & Patel, D. J. (2013). Structural Basis of mRNA Recognition and Cleavage by Toxin MazF and Its Regulation by Antitoxin MazE in Bacillus subtilis. *Molecular Cell, 52*(3), 447-458. doi:10.1016/j.molcel.2013.09.006

Sinha, S., Mell, J. C., & Redfield, R. J. (2012). Seventeen Sxy-Dependent Cyclic AMP Receptor Protein Site-Regulated Genes Are Needed for Natural Transformation in Haemophilus influenzae. *Journal of Bacteriology, 194*(19), 5245-5254. doi:10.1128/jb.00671-12

Smith, H. O., Danner, D. B., & Deich, R. A. (1981). GENETIC-TRANSFORMATION. *Annual Review of Biochemistry, 50*, 41-68. doi:10.1146/annurev.bi.50.070181.000353

Smith, H. O., Tomb, J. F., Dougherty, B. A., Fleischmann, R. D., & Venter, J. C. (1995). Frequency and distribution of dna uptake signal sequences in the haemophilus-influenzae rd genome. *Science, 269*(5223), 538-540. doi:10.1126/science.7542802

Sperandio, B., Gautier, C., McGovern, S., Ehrlich, D. S., Renault, P., Martin-Verstraete, I., & Guedon, E. (2007). Control of methionine synthesis and uptake by MetR and homocysteine in Streptococcus mutans. *J Bacteriol, 189*(19), 7032-7044. doi:10.1128/JB.00703-07

Starosta, A. L., Lassak, J., Jung, K., & Wilson, D. N. (2014). The bacterial translation stress response. *FEMS Microbiol Rev, 38*(6), 1172-1201. doi:10.1111/1574-6976.12083

Stiegelmeyer, S. M., & Giddings, M. C. (2013). Agent-based modeling of competence phenotype switching in Bacillus subtilis. *Theoretical Biology and Medical Modelling, 10*. doi:10.1186/1742-4682-10-23

Sung, C. K., & Morrison, D. A. (2005). Two distinct functions of ComW in stabilization and activation of the alternative sigma factor ComX in Streptococcus pneumoniae. *Journal of Bacteriology, 187*(9), 3052-3061. doi:10.1128/jb.187.9.3052-3061.2005

Tettelin, H., Nelson, K. E., Paulsen, I. T., Eisen, J. A., Read, T. D., Peterson, S., Fraser, C. M. (2001). Complete genome sequence of a virulent isolate of Streptococcus pneumoniae. *Science, 293*(5529), 498-506. doi:DOI 10.1126/science.1061217

Tomasz, A. (1965). Control of the competent state in Pneumococcus by a hormone-like cell product: an example for a new type of regulatory mechanism in bacteria. *Nature, 208*(5006), 155-159. doi:10.1038/208155a0

Tovpeko, Y., Bai, J. Q., & Morrison, D. A. (2016). Competence for Genetic Transformation in Streptococcus pneumoniae: Mutations in sigma(A) Bypass the ComW Requirement for Late Gene Expression. *Journal of Bacteriology, 198*(17), 2370-2378. doi:10.1128/jb.00354-16

Tovpeko, Y., & Morrison, D. A. (2014). Competence for Genetic Transformation in Streptococcus pneumoniae: Mutations in sigma(A) Bypass the comW Requierement. *Journal of Bacteriology, 196*(21), 3724-3734. doi:10.1128/jb.01933-14

Turgay, K., Hahn, J., Burghoorn, J., & Dubnau, D. (1998). Competence in Bacillus subtilis is controlled by regulated proteolysis of a transcription factor. *Embo Journal, 17*(22), 6730-6738.

Ueta, M., Wada, C., & Wada, A. (2010). Formation of 100S ribosomes by HPF (hibernation promoting factor) homolog in gram positive bacteria. *Genes & Genetic Systems, 85*(6), 425-425.

Ueta, M., Wada, C., & Wada, A. (2010). Formation of 100S ribosomes in Staphylococcus aureus by the hibernation promoting factor homolog SaHPF. *Genes Cells, 15*(1), 43-58. doi:10.1111/j.1365-2443.2009.01364.x

ValentinHansen, P., SogaardAndersen, L., & Pedersen, H. (1996). A flexible partnership: The CytR anti-activator and the cAMP-CRP activator protein, comrades in transcription control. *Molecular Microbiology, 20*(3), 461-466. doi:10.1046/j.1365-2958.1996.5341056.x

Vansinderen, D., Luttinger, A., Kong, L. Y., Dubnau, D., Venema, G., & Hamoen, L. (1995). ComK encodes the competence transcription factor, the key regulatory protein for competence development in bacillus-subtilis. *Molecular Microbiology, 15*(3), 455-462.

Veening, J.-W., & Blokesch, M. (2017). Interbacterial predation as a strategy for DNA acquisition in naturally competent bacteria. *Nature Reviews Microbiology, 15*(10), 621-629. doi:10.1038/nrmicro.2017.66

Vesper, O., Amitai, S., Belitsky, M., Byrgazov, K., Kaberdina, A. C., Engelberg-Kulka, H., & Moll, I. (2011). Selective translation of leaderless mRNAs by specialized ribosomes generated by MazF in Escherichia coli. *Cell, 147*(1), 147-157. doi:10.1016/j.cell.2011.07.047

Wagemans, J., & Lavigne, R. (2015). Identification of protein-protein interactions by standard gal4p-based yeast two-hybrid screening. *Methods in molecular biology (Clifton, N.J.), 1278*, 409-431. doi:10.1007/978-1-4939-2425-7_27

Wang, X., Kim, Y., Hong, S. H., Ma, Q., Brown, B. L., Pu, M.,Wood, T. K. (2011). Antitoxin MqsA helps mediate the bacterial general stress response. *Nature Chemical Biology, 7*(6), 359-366. doi:10.1038/nchembio.560

Ween, O., Gaustad, P., & Havarstein, L. S. (1999). Identification of DNA binding sites for ComE, a key regulator of natural competence in Streptococcus pneumoniae. *Molecular Microbiology, 33*(4), 817-827. doi:10.1046/j.1365-2958.1999.01528.x

Wenderska, I. B., Latos, A., Pruitt, B., Palmer, S., Spatafora, G., Senadheera, D. B., & Cvitkovitch, D. G. (2017). Transcriptional Profiling of the Oral Pathogen Streptococcus mutans in Response to Competence Signaling Peptide XIP. *mSystems, 2*(1). doi:10.1128/mSystems.00102-16

Weng, L., Piotrowski, A., & Morrison, D. A. (2013). Exit from Competence for Genetic Transformation in Streptococcus pneumoniae Is Regulated at Multiple Levels. *Plos One, 8*(5). doi:10.1371/journal.pone.0064197

Yoshida, H., & Wada, A. (2014). The 100S ribosome: ribosomal hibernation induced by stress. *Wiley Interdiscip Rev RNA, 5*(5), 723-732. doi:10.1002/wrna.1242

Yuan, A. H., Gregory, B. D., Sharp, J. S., McCleary, K. D., Dove, S. L., & Hochschild, A. (2008). Rsd family proteins make simultaneous interactions with regions 2 and 4 of the primary sigma factor. *Mol Microbiol, 70*(5), 1136-1151. doi:10.1111/j.1365-2958.2008.06462.x

Yuksel, M., Power, J. J., Ribbe, J., Volkmann, T., & Maier, B. (2016). Fitness Trade-Offs in Competence Differentiation of *Bacillus subtilis*. *Front Microbiol, 7*, 888. doi:10.3389/fmicb.2016.00888

Zhang, Y. L., Zhang, J. J., Hara, H., Kato, I., & Inouye, M. (2005). Insights into the mRNA cleavage mechanism by MazF, an mRNA interferase. *Journal of Biological Chemistry, 280*(5), 3143-3150. doi:10.1074/jbc.M411811200

VITA

Junqin (Heidi) Bai, Ph.D.
Biological Sciences
900 S Ashland Ave, Room 4110, Chicago, IL. 60607 Email: haijingbai0402@gmail.com

EDUCATION

Doctor of Philosophy in Biological Sciences **Dec 2017**
Dept. of Biological Sciences, University of Illinois at Chicago (UIC), Illinois
Master of Science in Mathematics **Dec 2017**
Dept. of Mathematics, Statistics, and Computer Science, UIC, Illinois
Master of Science in Biochemistry and Molecular Biology **June 2010**
School of Chemical Engineering & Technology, Tianjin University, Tianjin, China
Bachelor of Science in Biology Technology **June 2007**
School of Life Science, Shanxi Normal University, Linfen, China
Bachelor of Arts in English **June 2007**
School of Foreign Linguistics, Shanxi Normal University, Linfen, China

RESEARCH EXPERIENCE

***Research Assistant*, UIC-Morrison Lab, Chicago, IL** **2012-2017**

- Characterize roles of ComW, HPF, and YhaM in competence regulation in *S. pneumoniae*
- Construct mutants and host strains by sequence alignment and genome editing
- Study global protein expression by pulse-chase labeling and 1D and 2D electrophoresis
- Analyze protein expression patterns using Photoshop, ImageJ, and Prism
- Determine binary direct interaction between SigX, SigA, and ComW by Y2H assay
- Screen Δ*comW* suppressors were by MarinerT7 transposon library and pGh9::ISS1
- Analyze sequencing results using Clone Manager and Benching

***Visiting Scholar*, University of Wisconsin-Madison-Peterson Lab, Madison, WI** **2010-2011**

- Identified the role of *Sox9 in* mediating dioxin toxicity in prostate development
- Studied Aryl Hydrocarbon receptor (AHR) target genes causing TCDD toxicity in prostate
- Evaluated β-catenin-responsive gene expression by real-time RT-PCR in the β-catenin dominant stable mutant
- Performed TCDD treatment by microinjection
- Conducted UGS transplantation by routine histology and animal inoculation and surgery
- Observed prostate budding by scanning electron microscopy

***Research Assistant*, School of Chemical Engineering and Technology, Tianjin University, China** **2007-2010**

- Applied mRNA stabilizer to construct the *lipA* overexpression strain in *Bacillus subtilis*

***Research Assistant*, School of Life Science, Shanxi Normal University, China** **2005-2007**

- Preserved yogurt flavor by preventing its over-acidification during storage

INSTRUCTIONAL AND ADDITIONAL EXPERIENCE

- **Teaching Assistant, Department of Biological Sciences, UIC, Chicago, IL 2011-present**
- **Mentorship Experience-Undergraduate Honors Capstone Research-UIC, IL 2013-2017**

Internship **2006-2007**

- Interned as a teacher in First Municipal Senior High School, Linfen, Shanxi, China,
- Served as a teacher in Li Yang Crazy English Conversation Association, Linfen, Shanxi, China

PUBLICATIONS

- Tovpeko, Y., **Bai, J.**, and Morrison, D. A. 2016. Competence for genetic transformation in *S. pneumoniae*: Mutations in sigma bypass the ComW requirement for late gene expression. *Journal of Bacteriology*, 198 (17). pp. 2370-2378.
- **Bai, J.**, and Morrison, D. A. 2018. Role of HPF and YhaM in competence regulation in *S. pneumoniae*. Manuscript in preparation for submission.

RESEARCH PRESENTATIONS

Oral Presentation

- **Bai, J.,** and Morrison, D. A. 2013. Facultative ComW host strain for evaluation of new ComW suppressor mutations in pneumococcus. 57th Annual Wind River Conference on Prokaryotic Biology, Estes Park, Colorado.

Poster Presentations

- **Bai, J.,** and Morrison, D. A. 2017. Role of HPF and YhaM in competence regulation in *Streptococcus pneumoniae*. 61st Annual Wind River Conference on Prokaryotic Biology, Estes Park, Colorado.
- **Bai, J.**, Tovpeko, Y., and Morrison, D. A. 2016. Uncovering a new mechanism of global translational regulation in *Streptococcus pneumoniae*. 60th Annual Wind River Conference on Prokaryotic Biology, Estes Park, Colorado.
- **Bai, J.,** and Morrison, D. A. Identification of an interaction between the competence regulators SigX and ComW of *Streptococcus pneumoniae* by yeast two-hybrid.
  59th Annual Wind River Conference on Prokaryotic Biology, Estes Park, Colorado.
  21st Annual Midwest Microbial Pathogenesis Conference, Chicago, Illinois.
  5th Molecular Genetics of Bacteria and Phages Meeting, Madison, Wisconsin.
- Inniss, N., Tovpeko, Y., **Bai, J.,** and Morrison, D. A. 2012. ComX dependent competence in *Streptococcus pneumoniae*: Characterization of ComW suppressors. 56th Annual Wind River Conference on Prokaryotic Biology, Las Vegas, Nevada.
- Lin, T. M., **Bai, J.,** and Peterson, R. E. 2011. Early stabilization of β-catenin in urogenital sinus (UGS) epithelia disrupts ventral prostate (VP) development. School of Pharmacy, University
- of Wisconsin, Madison, Wisconsin.

HONORS & AWARDS

- Arnold Ravin-Muriel Rogers National Science Foundation Travel Grants, Wind River, Colorado, 2012-2017
- Liberal Arts and Sciences Ph.D. Student Travel Award, UIC, Chicago, IL, 2012-2017
- Graduate College Student Presenter Award, UIC, Chicago, IL, 2012-2017

- Department of Biological Sciences Travel Award, UIC, Chicago, IL, 2012-2017
- Graduate Student Council Travel Award, UIC, Chicago, IL, 2012-2017
- Undergraduate Scholarship, Shanxi Normal University, Shanxi, China, 2004-2007

PROFESSIONAL STRENGTHS & SKILLS

- Molecular Genetics: PCR, genotyping, DNA sequencing, DNA and RNA isolation, plasmid construction, blotting**,** gel electrophoresis, qPCR, RT-PCR, genome editing, yeast two-hybrid
- Protein: Protein expression and purification, IP, SDS-PAGE, 2D-PAGE, 3D structure analysis
- Microbiology: Sterilization, aseptic techniques, strain isolation and identification
- Physiology: Dissection of mice organs, animal inoculations and surgery, microinjection
- Histology: tissue culture, immunohistochemistry, RNA *in situ* hybridization, SEM
- Scientific Writing
- Bioinformatics: Sequencing alignment, phylogenetic analysis
- Certificates: Statistical Analysis Software Base Certificate (SAS)
  Senior High School Teacher Certificate in China
- Computer Skills: Blackboard, Prism, SAS, R, Pymol, Benchling, Image J

UIC

Junqin Bai <jbai7@uic.edu>

Request for copyright permission

PermissionsFrance <permissionsfrance@elsevier.com> Mon, Nov 27, 2017 at 12:17 AM
To: Junqin Bai <jbai7@uic.edu>

ELSEVIER
MASSON

Issy-Les-Moulineaux, 27-Nov-2017

To the attention of Junqin Bai

-

Dear,

As per your request below, we hereby grant you permission to reprint the material detailed in your request at no charge **in your thesis** subject to the following conditions:

1. If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permissions are not obtained then that materials may not be included in your publication.

2. Any modification of the material is likely to harm the moral right of the authors and therefore should be first submitted and approved by the authors who are the sole owner of the moral right.

3. Suitable and visible acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

   "Reproduced from *Authors name. Article title. Journal title year; volume number(issue number):first page-last page.* Copyright © *year [if applicable: name of learned society,* published by*]* Elsevier Masson SAS. All rights reserved."

4. Your thesis may be submitted to your institution in either print or electronic form.

5. Reproduction of this material is confined to the purpose for which permission is hereby given.

6. This permission is granted for non-exclusive world **English** rights only. For other languages please reapply separately for each one required. Permission excludes use in an electronic form other than submission. Should you have a specific electronic project in mind please reapply for permission.

7. Should your thesis be published commercially, please reapply for permission.

Yours sincerely,
Regina Lavanya Remigius

Senior Copyrights Coordinator, GR - Copyrights

ELSEVIER | Global Book Production

From: Junqin Bai [mailto:jbai7@uic.edu]
Sent: Thursday, November 2, 2017 4:38 AM
To: PermissionsFrance <permissionsfrance@elsevier.com>
Subject: Request for copyright permission

***** External email: use caution *****