Mechanistic Studies of the Roles of the Transcriptional Activator ExsA and Anti-activator Protein ExsD in the Regulation of the Type Three Secretion system in Pseudomonas aeruginosa

Manisha Shrestha

Dissertation submitted to the faculty of Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

In

Biological Sciences

Florian D. Schubot (Chair)

Biswarup Mukhopadhyay

David L. Popham

Ann M. Stevens

April 25th, 2018

Blacksburg, VA, USA

Keywords: Pseudomonas aeruginosa, opportunistic pathogen, type three secretion system, transcription regulator, AraC proteins

Copyright 2018, Manisha Shrestha
Mechanistic Studies of the Roles of the Transcriptional Activator ExsA and Anti-activator Protein ExsD in the Regulation of the Type Three Secretion System in *Pseudomonas aeruginosa*  

Manisha Shrestha  

**ABSTRACT**

*Pseudomonas aeruginosa* is a ubiquitous opportunistic pathogen that is a substantial threat, particularly in hospital settings, causing severe infections in immunocompromised patients that may lead to death. *Pseudomonas aeruginosa* harbors a multitude of virulence factors that enable this pathogen to establish both acute and chronic infections in humans. A key determinant of acute infections is a hollow molecular needle structure used for injecting toxins into a host cell, called the type three secretion system (T3SS). The secretion machinery itself is highly complex and, together with the specific secreted factors, requires expression of more than 30 genes. Due to the high energy cost of its synthesis to the organism this system is highly regulated to finely time gene expression to coincide with host contact. ExsA, a member of the AraC-type transcription factor family, is the main transcriptional activator of all the genes necessary for expression of the T3SS. Members of the AraC family are characterized by the presence of two helix-turn-helix (HTH) motifs, which bind to the promoter DNA and activate transcription. ExsA uses its HTH containing C-terminal domain (CTD) to regulate gene expression from 10 different promoters. The N-terminal domain (NTD) of ExsA mediates dimerization and regulation of ExsA-activity. While most AraC-type activators are regulated by a small molecule ligands, ExsA is regulated by another protein, ExsD. As part of a four-protein signaling cascade, ExsD interacts directly with ExsA to prevent transcription of T3SS-associated genes under non-inducing conditions prior to host cell contact. The entire regulatory cascade
includes of two additional proteins, ExsC and ExsE. ExsA, ExsC, ExsD, and ExsE follow a partner-switching mechanism to link expression of the secretion system with host cell contact. Our laboratory is working to understand this unique signaling mechanism by determining the molecular basis for the regulation of this important virulence factor. Previous studies in the laboratory have solved the structures of ExsE, ExsC and ExsD, and shed light on how these proteins interact and compete for overlapping binding sites. However, it is still unclear as to how the ExsA and ExsD interact and thus how regulation is mediated at the molecular level.

In the presented study, we sought to map the molecular interface between ExsA and ExsD. First, the crystal structure of ExsA-NTD is presented wherein the dimerization interface of the protein was identified. Two of the well-studied AraC-type proteins, AraC and ToxT crystal structures have been solved by others in the presence of their respective ligands. Residues that were involved in ligand binding in AraC and ToxT were aligned with the residues in ExsA and analyzed for interaction with ExsD. However, this canonical binding pocket appeared to be not involved in the interaction between ExsA and ExsD. Structure directed site-specific mutagenesis was carried out to construct many different variants of ExsD and ExsA. Thus constructed variants were purified and analyzed in a functional assay. Using this approach, we were able to identify regions on ExsD and ExsA that are crucial for the interaction and for the regulation of ExsA-dependent transcription. It turns out that backbone interactions between the amino-terminal residues of ExsD and the beta-barrel region of the ExsA-NTD are pivotal. This result explains how ExsA and ExsC compete for ExsD binding, since both target the same regions on ExsD.
Mechanistic Studies of the Roles of the Transcriptional Activator ExsA and Anti-activator Protein ExsD in the Regulation of the Type Three Secretion system in *Pseudomonas aeruginosa*

**Manisha Shrestha**

**GENERAL AUDIENCE ABSTRACT**

*Pseudomonas aeruginosa* is an opportunistic pathogen that is notorious for causing severe infections in immunocompromised individuals. Acute *Pseudomonas aeruginosa* infections are characterized by immediate adverse effects. An initial acute infection may become chronic, leading to long-term morbidity and mortality in affected individuals. During the initial stages of infection *P. aeruginosa* uses the type three secretion system, a syringe-like structure, to puncture the host cell and inject potent toxins. The activation of the genes required for forming this structure is tightly controlled by an activator protein, ExsA. When *P.aeruginosa* is not invading a host, ExsA is inhibited by another protein called ExsD, to prevent the needless production of the secretion apparatus. The presented work explores the mechanism of how ExsD achieves this inhibition of ExsA. This information is of potential biomedical interest because a clear understanding of the molecular basis for the interaction could inform the development of a small-molecule mimic of ExsD to be used in therapy. In Chapter 2 we report the structure of the domain of ExsA that is known to bind ExsD. Also, in this chapter and more so in Chapter 3, we performed a detailed analysis of potential interacting regions and ultimately succeeded in identifying key interacting regions in both ExsA and ExsD.
To my parents, **Mrs. Rukmini Shrestha** and **Mr. Padam Lal Shrestha**, whose vision has been so much bigger than the opportunities they have had.
ACKNOWLEDGEMENTS

First of all, I would like to thank the Virginia Tech Graduate School, the Inter-departmental Microbiology Program and the Department of Biological Sciences for providing me with the support and letting me join this amazing community where I gained training to complete my Ph.D. work.

I would like to express my deep appreciation to my academic advisor, Dr. Florian D. Schubot for allowing me to work on the projects I did. Thanks to him, I have developed a keen interest in one the most important biomolecules, proteins. I plan to keep working with these amazing molecules. I would like to thank him for continuously giving me advice on the things that I lacked confidence in. I would also like to thank him for showing patience with me when I was struggling. I would compare his way of advising me as someone flying a kite, he would let me go on my own with the flow but whenever I would begin to drift, he would pull the strings which help me ascend with a fixed path. I thank you for your time.

I would also like to thank my committee members, Dr. Biswarup Mukhopadhyay, Dr. David Popham and Dr. Ann Stevens for helping me polish and steer my projects to the required direction. Their continuous support during the meetings and beyond has been monumental for bringing my project to this point.

I would also like to thank our former lab member, Robert C. Bernhards, Ph.D. My work built on his dissertation studies. He had a major role in my training and I thank him for his patience and thoroughness with which he trained me during the initial few months in this lab. I would also like to thank two former lab members, Xing Jing, Ph.D. and Yi Xiao, M.S.; their contribution to my training was also very important. I would like to thank my current fellow lab member, Jordan Mancl. We were two of the same six people cohort who joined the inter-departmental microbiology program. He has been a great resource for the past 5+ years I have been in this lab. His comments, attitude, and work ethic have helped me be the person who I am today. Thank you for everything Jordan. 😊 I wish you all the best for your bright future. I also thank Sarah Nikraftar M.S., Dr. Brianna Petruzzi, Zhixiong Sun, and Dr. Oscar Galagarza who were people in the same cohort as Jordan and I, and in whom I found friendship during the initial days of my graduate career. I would like to thank Sarah for being such an amazing friend.

I would like to thank other members of the Schubot lab, especially undergraduate students with whom I was able to work and enjoy while teaching them basic concepts of molecular biology. Yichen Fu, Natalya Ortolano, and Aaron Okwei were outstanding students. I would also like to thank Kylie Ryan, a fairly new graduate student who joined our lab in January 2016. I wish you good luck for your future endeavors.
I thank Dr. Birgit Scharf for all her help and guidance during my whole graduate career. She is an amazing scientist and one of the most wonderful human being I have ever come across. Because of her, I came to meet a group of most amazing people who joined her lab, and whom I feel very proud to call my friends. Dr. Ben Webb, and Dr. Hardik Zatakia, previous Scharf lab members, I thank you for all your help during my prelims and all the other times whenever I needed. Thank you, Katie Broadway, Timofey Araprov, Karl Compton, Floricel Gonzalez, Rafael Castneda, and Richard Sobe for your company, conversations, and guidance. I know all of you will do amazing things in future. I hope our paths cross in future. I would also like to thank Dr. Hualan Liu, Dr. Revathy Ramachandran, and Dr. Casey Bernhards for being outstanding senior graduate students and role models. Their attitude and personality helped me develop a very positive outlook towards graduate school. I would like to thank Dr. An Duy Duong for being my friend and from whom I sought advice. I would also like to thank Annette Fluri, Dr. Sherry Hilderth and Jody Jervis for being great neighbors when we inhabited in LSI. I cherish all our conversations and value your friendships. I thank Bidisha and Andreas for sharing the same headspace with me at several occasions. I wish all the graduate students in the Microbiology group; Cameron, Keane, Holly, Angie, Gary and Ian all the best for their future. I would also like to thank the members of the teaching lab, Katrina Lasley, Kevin Loah, (past members: Katie Rodgers, and Carla Tyler) for being a great resource and making my teaching experience so wonderful in past 5 years. I would also like to thank BGSA for helping foster a community for graduate students.

I would also like to thank the Nepalese community at Virginia Tech and Blacksburg, for helping me create home away from home. I would like to specially thank, Bishal Paudel, soon to be an M.D., for being the greatest source of my mental, emotional, and intellectual support for the four years when I knew him as an undergrad here at Tech. I would like to thank three of my very good friends, Dr. Anisha Shakya, Dr. Rajdeep Bomjan and Dr. Puja Sharma. I believe in the quote, “you are the company you keep” by Howard Schultz. All of the interactions with them has given the greatest source of happiness when all else seemed to be failing.

Last but not the least, I would like to thank my family, my parents, Maa and Baa, siblings, Barsha, Rupam and Kripam for being the constant sources of love, light, and wisdom in my life. Without them I do not think I would be here. I also thank my fiancé, Netain Shrestha and his family for showing me so much love in the past few months which has been the key ingredient needed for the final push.

I would like to thank every single person I interacted with since August 2012 when I was here, you have helped shape me into who I am today.
ATTRIBUTIONS

I disclose that all the experimental work reported herein, except for the statements below was performed by me. In addition, several colleagues contributed in writing, data collection and analysis in Chapter Two and Three is briefly reported here.

**Florian D. Schubot, Ph. D.,** Associate Professor in the Department of Biological Sciences at Virginia Tech, was the research advisor for all the projects presented in this dissertation. He conceived the experimental design, contributed reagents, materials and tools. He assisted in data analysis and revising all the chapters.

**Robert C. Bernhards, Ph.D.,** a previous graduate student in this laboratory helped in designing, executing, data analysis and writing of Chapter Three. The work presented here has been reported in his dissertation, submitted to the Graduate School, Virginia Tech in Spring 2013. His specific contribution has been mentioned on the title page of Chapter Three.

**Yi Xiao, M.S.,** a previous graduate student in this laboratory helped in designing, executing, data analysis in project presented in Chapter Two. His work was reported in this Masters’ Thesis submitted to the Graduate School, Virginia Tech in Spring 2013. Chapter Two has been published in the PloS One journal in August, 2015.
TABLE OF CONTENTS

LIST OF TABLES .......................................................................................................................... xi
LIST OF FIGURES ......................................................................................................................... xii
LIST OF ABBREVIATIONS ............................................................................................................ xiv

CHAPTER ONE: Literature Review ............................................................................................. 1-27
  Pseudomonas aeruginosa, the prototypical opportunistic pathogen ............................................... 2
  Virulence factors and infections caused by Pseudomonas aeruginosa ...................................... 2
  Type three secretion system (T3SS) in P. aeruginosa ................................................................ 4
  Effectors and their action in the host .......................................................................................... 6
  Post-translational regulation of Type three secretion system in P. aeruginosa ....................... 6
  Role of ExsA, an AraC/XylS-type regulator in transcriptional regulation of the T3SS ....... 7
  P. aeruginosa T3SS regulatory proteins and their interactions ................................................. 12
  Additional regulators of T3SS gene expression ........................................................................ 15
  ExsA-ExsD interaction: Possible drug intervention point ......................................................... 16
  Project objectives .................................................................................................................... 18
  Rationale and Approach ........................................................................................................... 19
  Approach ................................................................................................................................ 20
  References ............................................................................................................................... 22

CHAPTER TWO: Structural analysis of the regulatory domain of ExsA, a key transcriptional regulator of the type three secretion system in Pseudomonas aeruginosa ........................................ 28-56
  Abstract .................................................................................................................................... 29
  Introduction ............................................................................................................................... 30
  Materials and Methods ............................................................................................................ 33
    Preparation of ExsD and ExsA variants .................................................................................. 33
    Recombinant protein expression and purification ................................................................. 34
    Preparation and Crystallization of ExsA-NTD ...................................................................... 36
    X-ray diffraction data collection, structure determination and refinement ......................... 37
    ExsA-dependent in vitro transcription assays ....................................................................... 37
  Results ....................................................................................................................................... 38
    Crystal Structure of the regulatory domain of ExsA ............................................................. 39
    Helices α-2 and α-3 together form the core of the ExsA dimer interface ......................... 40
    The cavity within the ExsA-NTD beta-barrel is not involved in ExsD binding ............... 42
  Discussion ................................................................................................................................ 43
  References ............................................................................................................................... 45
  Figures ....................................................................................................................................... 52

CHAPTER THREE: Mechanistic Studies of the roles of transcriptional activator ExsA and anti-activator protein, ExsD in regulating the type three secretion system in Pseudomonas aeruginosa ......................................................... 57-92
  Abstract .................................................................................................................................... 58
  Introduction ............................................................................................................................... 59
  Methods .................................................................................................................................... 62
    Construction of ExsA and ExsD expression plasmids ......................................................... 62
    Construction of different variants of ExsD and ExsA ......................................................... 62
    Purification of ExsA and ExsD ............................................................................................. 64
    Limited proteolysis to obtain ExsDA20 .............................................................................. 66
  In vitro transcription assay ....................................................................................................... 67
    Bacterial adenylate cyclase two-hybrid assay ..................................................................... 68
  Results and Discussion ............................................................................................................ 70
The coiled-coil region of ExsD is not important for ExsA binding .................. 70
The amino-terminal 20 residues of ExsD are important for inhibiting ExsA-dependent transcription ........................................................................................................................................ 71
Conserved surface exposed residues on ExsA are not crucial for ExsD binding ........................................................................................................................................ 72
Residues near the amino-terminus of ExsA play an important role in the regulation mechanism in vitro ........................................................................................................................................ 74
Secondary structure elements play an important role in transcriptional regulation of T3SS via ExsD ........................................................................................................................................ 76
The Y24P and V26P mutations in ExsA do not fully disrupt the ExsD-ExsA interface ........................................................................................................................................ 78
Conclusion ........................................................................................................................................ 79
References ........................................................................................................................................ 81
Figures ........................................................................................................................................ 84

CHAPTER FOUR: Overall Conclusions ................................................................. 93-101
References ........................................................................................................................................ 98
Figures ........................................................................................................................................ 100

Appendix A- Supplemental data: Chapter Two .................................................................................. 102
Appendix B- Supplemental data: Chapter Three .................................................................................. 105
Appendix C- Additional variants of ExsA ........................................................................................... 116
LIST OF TABLES

Chapter Two
Table 1.1. Diffraction data and crystal structure refinement statistics............................ 55

Appendix B
Table B1. List of primers used to construct ExsD<sub>1-38</sub>-MBP variant.............................. 105
Table B2. List of primers used to construct variants of ExsA ............................................. 105
LIST OF FIGURES

Chapter One
- Figure 1.1 Cartoon showing exotoxin, structural, and regulatory proteins involved in the T3SS of *P. aeruginosa* ................................................................. 5
- Figure 1.2 Different operons and promoters regulated by ExsA ........................................... 8
- Figure 1.3 Binding site ExsA-dimer on promoters ................................................................. 10
- Figure 1.4 ExsACDE regulatory model ................................................................................. 12
- Figure 1.5 Different regulatory pathways controlling T3SS in *P. aeruginosa* ............... 15

Chapter Two
- Figure 2.1. Crystal structure of the ExsA-NTD .................................................................. 51
- Figure 2.2. Mapping of the ExsA dimer interface ............................................................... 52
- Figure 2.3. Impact of mutations in the conserved cavity of ExsA on ExsD binding .......... 53
- Figure 2.4. Cartoon model of a full-length dimeric ExsA-DNA complex ....................... 54

Chapter Three
- Figure 3.1. Effect of ExsD variants on ExsA dependent transcription *in vitro*............... 84
- Figure 3.2. The structurally ordered region encompassing amino acids 40 to 278 of ExsD provides more than bulk to the transcriptional regulation of ExsA. ......................... 85
- Figure 3.3. Conserved surface exposed residues of ExsA are not critical for ExsD binding ............................................................................................................... 86-87
- Figure 3.4. Side chains of individual residues near the amino-terminus of ExsA are not critical for the regulatory mechanism but modification of the amino-terminus can impact this regulation ........................................................................................................ 88-89
- Figure 3.5. Secondary structure elements play an important role in transcriptional regulation of T3SS via ExsD ................................................................................. 90-91
- Figure 3.6. Results for bacterial two-hybrid assay for the interaction between ExsA variants and ExsD .......................................................... 92

Chapter Four
- Fig.4.1 Hypothetical model of a possible ExsA-ExsD interface in full-length ExsA........... 100
- Fig. 4.2 Hypothetical model for how ExsD “locks down” ExsA to inhibit transcription.... 101

Appendix A
- A1. SDS-PAGE of ExsA-NTD, titration of ExsA N-terminal domain in ExsA-dependent transcription, crystal and sample X-ray diffraction image of ExsA-NTD ............................ 102
- A2. Schematic representation of the non-bonding contacts in the A/A’ and B/B” dimers of ExsA-NTD ........................................................................................................ 103
- A3. SDS-PAGE slices of the purified protein samples used in the various experiments .... 104

Appendix B
- Fig. B1. Commassie stained gels of purified variants of ExsA and ExsD .......................... 108
- Fig. B2. Effect of ExsDΔC-C on ExsA-dependent transcription at 30 and 37 °C............. 109
- Fig. B3. Effect of ExsDΔ20 on ExsA dependent transcription at 37 °C .......................... 110
- Fig. B4. ExsDΔ20 has no effect on ExsA-independent transcription ............................... 111
- Fig. B5. Effect of ExsDΔ1-46 on ExsA dependent transcription at 37 °C ..................... 112
- Fig. B6 ExsDΔ1-46 has no effect on ExsA-independent transcription ............................ 113
- Fig B7. Sequence identifier and alignment of ExsD from different bacteria ................. 114
Appendix C

Fig. C1. Are the main-chain amide groups of β-1 or β-2 of ExsA important for ExsD binding?.................................................................116

Fig. C2. Additional variants of ExsA had no effect by ExsD regulation, *in vitro*.....117
LIST OF ABBREVIATIONS

T3SS Type three secretion system
DTT Dithiothreitol
TCEP Tris(2-carboxyethyl)phosphate
BSA Bovine serum albumin
MDR Multidrug resistance
LB Luria-Bertani
RNAP RNA polymerase
dIdC deoxyinosinic-deoxycytidylic
EDTA Ethylenediaminetetraacetic acid
IPTG Isopropyl β-D-1-thiogalactopyranoside
TEV Tobacco etch virus
MBP Maltose-binding protein
NTA Nitrilotriacetic acid
SDS-PAGE Sodium dodecyl sulfate–polyacrylamide gel electrophoresis
ATP Adenosine triphosphate
GTP Guanosine triphosphate
CTP Cytosine triphosphate
UTP Uridine triphosphate
PCR Polymerase chain reaction
cAMP cyclic-adenosine monophosphate
**Pseudomonas aeruginosa**, the prototypical opportunistic pathogen

*Pseudomonas aeruginosa* is a Gram-negative ubiquitous opportunistic bacterium [1]. The *P. aeruginosa* (PAO1) genome is 6.3 Mbp in size with approximately 5500 predicted open reading frames [2]. *P. aeruginosa* expresses nearly 500 regulatory proteins that function to adapt the bacterium’s metabolism for survival under the specific nutrient-limited conditions, which characterize its many habitats such as aquatic systems, sewage, soil, plants, and animals [3]. *P. aeruginosa* is a facultative anaerobe and may use nitrate as a terminal electron acceptor [4]. Due to its metabolic flexibility it is a prototypical opportunistic pathogen. Colonization of healthy individuals is generally not harmful but infection of immunocompromised hosts can be life-threatening. *P. aeruginosa* may be found in the human gastrointestinal tract as a part of normal flora, however colonization increases upon hospitalization and *P. aeruginosa* contributes to 4.8-16% of nosocomial infections, the percentage of colonization can be as high as 50% in the immunocompromised hosts [5-7]. It can cause infections in different parts of human body including the eyes, ears, skin surface, burns, wounds, and lungs. Patients who use prosthetic medical devices like catheters and ventilators are also prone to *Pseudomonas* infections [1].

**Virulence factors and infections caused by Pseudomonas aeruginosa**

*P. aeruginosa* was first discovered in wounds because of the blue-green coloration it gave to bandages [8]. This phenotype is due to the production of a redox active pigment, called pyocyanin, which causes oxidative stress necessary for virulence in animals [9]. *P. aeruginosa* is associated with both acute and chronic infection especially in immunocompromised individuals. Disruption of epidermal barriers, for instance cuts and burns, provides a primary path by which
this organism enters a human body leading to skin and tissue infections such as folliculitis, cellulitis, necrotizing fasciitis, green nail syndrome, and otitis externa [10]. Contact lens users are susceptible to *P. aeruginosa*-related keratitis that may lead to permanent corneal damage [11], while disruption of mucous membranes may lead to ventilator-associated pulmonary infections and catheter-associated urinary tract infections [12].

*P. aeruginosa* is particularly notorious for causing chronic lung infections in people with cystic fibrosis (CF). These patients have impaired mucosal membrane, due to a loss of function of the cystic fibrosis transmembrane regulator (CFTR) gene [13]. A non-functional CFTR gene results in loss of CFTR channels in the lungs, sweat glands and pancreas. In the lungs the channels are responsible for chloride and bicarbonate ions movement needed for maintaining pH and anti-microbial activity of mucosal surface liquid [13]. This is also important to maintain the normal composition of mucus, which is necessary for clearance of the lung ducts. Due to defects in the normal clearing of the pathogen, 80% of the adults with CF are colonized by *P. aeruginosa*; the patients are also subject to recurrent pneumonia leading to high mortality [14].

Virulence factors may be categorized depending on the type of disease symptoms they cause. Establishment of an acute infection involves LPS, adhesins, proteases, and the secretion of cytotoxins, which may lead to severe tissue damage, septic shock, systemic infection and death [15]. Proteases are exported by the type II secretion system (T2SS). Other structures like type IV pili (Tfp) and the type III secretion system (T3SS) can inject toxins into the host cells [16]. Chronic infections are less invasive and less cytotoxic compared to acute infections as the factors involved are selected for through the course of the chronic infection [15]. However, the chronic disease may persist for many years leading to long term morbidity and later mortality. Chronic infections are characterized by biofilm formation, increased antibiotic resistance, and
long term protection from host defenses [17]. Chronic infections are particularly important in
cystic fibrosis (CF) patients, as a large percentage of this patient group is continuously infected
by *P. aeruginosa* throughout their lifetime. Mortality in these patients is mainly due to persistent
lung infections leading to severe lung damage [18].

**Type Three secretion System (T3SS) in *P. aeruginosa*:**

The T3SS is believed to have evolved from a flagellar apparatus, which was modified
over time by horizontal gene transfer events from various organisms [19]. T3SSs have also been
termied transmembrane protein nanomachines, that inject effector proteins into the host
cytoplasm to interfere with critical signaling pathways and cause cellular remodeling by
changing either the permeability or morphology of the host cell [20-22]. Due to its prevalence
and prominent role during infection of many important Gram-negative pathogens the T3SS is
one of the most extensively studied secretion systems. The *P. aeruginosa* T3SS belongs to the
Ysc family, which also includes the T3SSs of *Yersinia* spp, *Aeromonas salmonicida, *Photorhabdus luminescens, Vibrio parahaemolyticus, Bordetella* spp, and *Desulfovibrio vulgaris*
[23]. There are five known operons encoding 36 genes encoding structural and regulatory
components of the T3SS in the chromosome of *P. aeruginosa*. Six additional genes encoding
effectors and their associated chaperones are located elsewhere in the chromosome [25]. The
syringe-like T3SS apparatus is a highly conserved structure. The apparatus is composed of inner
and outer membrane spanning rings formed by PscD/PscJ, and the secretin, PscC. PscN is an
ATPase attached to the basal part of the cytoplasmic apparatus. Originally, the ATPase had been
hypothesized to power the secretion process itself but later the power was shown to have been
derived from the proton motive force [26, 27]. The role of the T3SS ATPase has been best
studied in the analogous system in *Salmonella enterica* [26]. The effector proteins traverse the
20Å wide inner channel in unfolded form [26, 28, 29]. In 1996, through genetic screen, ExoS, a toxin secreted by T3SS was first identified in \textit{P. aeruginosa} [24]. \textit{S. enterica} InvC was shown to separate chaperones from their cognate T3SS substrates and help unfold the effector protein to prepare for secretion in an ATP-dependent manner [30]. In \textit{P. aeruginosa}, the subunits of PscF polymerize to form the needle, which is anchored to the basal body by the PscI inner rod protein, and protrudes out from the bacterial cell surface. The needle has an outer diameter of approximately 6-8 nm and is about 60-120 nm in length. The length of the needle is regulated by a molecular ruler protein called PscP [31]. PopB and PopD are referred to as the translocon (Figure 1.1).

These proteins form the translocon pore in the host cell membrane and are universally conserved among T3SS-bearing bacteria [32, 33]. Prior to attachment to the host cell, PcrV interacts with the translocon to form the needle tip [34]. Mutational analysis of PcrV has suggested that the regulation and secretion functions of PcrV are distinct processes as some alleles of PcrV could not regulate the secretion but were still cytotoxic [35, 36]. The translocon assembly precedes induction of effector expression and

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Cartoon showing the secreted exotoxin, as well as structural, and regulatory proteins involved in the T3SS of \textit{P. aeruginosa}. The needle complex spans from the inner-membrane of the bacterium to the host cell cytoplasmic membrane.}
\end{figure}
secretion [37] [38]. PcrH serves as the chaperone to the translocon proteins and facilitates timing and delivery of PopB and PopD to the basal body of the T3SS [39].

**Effectors and their action in the host**

The T3SS of *P. aeruginosa* translocates exoenzymes S, T, U and Y into the host cell cytoplasm where they serve as effectors causing a variety of effects that collectively serve to delay the innate immune response of the host (7). ExoS has ADP-ribosylating activity, and ExoT and ExoU are involved in stimulating GTPases activity to mediate the rearrangement of the host cell cytoskeleton [40]. ExoY, a homologue of *Bordetella pertussis* toxin and edema factor of anthrax toxin, is an adenylate cyclase, which causes morphological changes to the Chinese hamster ovary caused likely due to the difference in intracellular level of cAMP or varying distribution of the toxin inside the cell [41].

**Post-translational regulation of type three secretion system in P. aeruginosa**

Forty three different genes clustered in six operons are responsible for the formation, translocation, secretion, and regulation of the T3SS in *P. aeruginosa* [25, 42, 43]. T3SS expression and assembly are energy intensive processes and are thus highly regulated at all levels: prior to and post transcription as well as post-translationally [24, 43, 44]. Environmental conditions play a major role in this regulation. Contact with a host cell, decreases in calcium levels, and the presence of serum in the environment are considered the most important cues for induction of the system [42]. The first report of such regulation was obtained when Exoenzyme S was shown to be secreted only under calcium-depleted conditions [24]. The needle tip protein, PcrV maintains the secretion apparatus in an inactive state before host cell contact [36].
However, deletion of \textit{pcrV} abolishes translocation of proteins, which may be attributed either to loss of PcrV causing a permanent sealing of the needle tip or to a role of the protein in facilitating important conformational changes in the needle following host cell contact \cite{37}. Secretion is induced when \textit{P. aeruginosa} is grown in media containing low Ca\textsuperscript{2+} levels, but the relationship to \textit{in vivo} infections are not clear. Translocon proteins, however, are secreted in the presence of Ca\textsuperscript{2+} suggesting these proteins are secreted prior to host cell contact \cite{37}. There is a model which suggests that PopB and PopD are involved in relaying the signal provided by the host cell, while PcrV directly or indirectly blocks the secretion until an inducing signal is received \cite{45}. Other cytoplasmic proteins, called PcrG and PopN, due to their homology with \textit{Yersinia spp.} proteins, have been identified as negative regulators for secretions \cite{46}. PcrD and PopN are also shown to have some role in regulation in the \textit{Salmonella enterica} serovar Typhymurium T3SS in which loss of PopN, Pcr1 or the PcrD-Pcr1 interaction results in increased effector secretion \cite{47}.

**Role of ExsA, an AraC/XylS-type regulator in transcriptional regulation of the T3SS**

Genetic studies demonstrated an important regulatory role of the \textit{exsA} gene in control of ExoS synthesis \cite{42}. Further work demonstrated that \textit{exoS} is actually trans-regulated by four genes named \textit{exsC, exsE, exsA and exsD}, respectively \cite{48-50}. We now know that transcription of not only \textit{exoS}, but gene expression from all ten T3SS-related promoters is controlled by the single transcriptional activator, ExsA. A two-domain protein, the carboxy-terminal part of ExsA displays high sequence homology with the DNA-binding domains of regulatory proteins belonging to the AraC/XylS family \cite{51} (Figure. 1.2). Alignment of all 10 ExsA-dependent promoters has allowed for identification of the conserved promoter elements with a consensus
site of AaAAAAnwmMyGrCynnnmTGayAk [52]. (>90% conservation is represented by a nucleotide, conservation between 90 to 80% is represented by a lower case alphabet of the nucleotide. W = A or T, M = A or C, Y = C or T, R = A or G, K = G or T, and n = any nucleotide [52]). DNA- footprinting experiments suggested two monomers of ExsA bind to the promoter region extending from the -35 to -70 positions relative to the transcriptional start site [52]. ExsA-dependent promoters lie upstream of pscN, popN, pcrG, exsC, exsD, spcS, exoS, exoT, exoY, and exoU, and each promoter contains at least one ExsA binding site (TNAAAANA), which is centered ~15 base pairs upstream of the -35 RNA polymerase binding sites.

AraC/XylS proteins are characterized by a unique 100 amino acid conserved sequence, which folds into two DNA-binding helix-turn-helix motifs [53]. The AraC/XylS family includes more than 10,000 members, many of which have not yet been characterized [54]. Those that have been characterized function as transcription activators except CelD from E. coli, which is a

Figure 1.2 Different operons and promoters regulated by ExsA. ExsA is the master regulator of all the genes involved in the formation of the T3SS in P. aeruginosa.
repressor [55], while YbtA from *Yersinia* sp may act as both an activator and a repressor in response to an effector molecule [56]. Broadly, the AraC/XylS family regulates transcription of three main functions; carbon metabolism, stress response, and pathogenesis [53]. These transcription regulators have been widely found in prokaryotes only. They usually have two domains: an N-terminal domain that is weakly conserved at least at the sequence level and the aforementioned AraC domain at the C-terminal region [57]. Typically, these transcription factors are 250-300 amino acids in length, with a few exceptions. Ada from *Mycobacterium*, and YbbB from *B. subtilis* contain 500 amino acids, while MarA and SoxS of *E.coli* consist only of a single AraC domain. Exceptions include Rob and YbbB from *E.coli* where the AraC domain is positioned at the N-terminus and in the middle, respectively [53]. Of the two Helix-turn-Helix (HTH) motifs, the sequence of the second is more conserved across the entire family [53]. Usually AraC proteins associate with their consensus DNA-binding site as dimers, however monomers and higher order oligomers have also been reported [57]. Consequently, promoters for these proteins have more consensus binding sites for regulators [53]. However, sometimes an ExsA-dependent promoter displays only a single obvious consensus sequence, but DNase footprinting and EMSA studies show binding of at least two protein monomers (7). Promoter affinities vary not only between proteins, but also between promoter sites if a single regulator associates with more than one promoter [53]. The position of the binding site ranges from almost overlapping the -35 region to being located about 100 nucleotides upstream of the transcription start site, suggesting that the different regulators use different mechanisms to regulate transcription and interact with RNA polymerase (RNAP) [53]. If more than one consensus site is present those might be arranged as either direct repeats or as inverted repeats, indicating that the binding orientation is variable [53]. Such varying orientations have also been observed among
the various ExsA-dependent promoters. Some have only one obvious ExsA-binding site, others have two head-to-head sites, while still others force a head-to-tail orientation of the two bound ExsA molecules [52]. The consequences of these different orientations on ExsA-RNAP interactions are not clear. However, ExsA is a 278 amino acid protein with two functional domains. The N-terminal 178 residues constitute the regulatory and dimerization domain, while residues 179-278 are composed of a flexible linker region and the DNA binding domain [58, 59]. It is likely this linker region that enables the consistent formation of NTD-mediated ExsA dimers, while still permitting a flexible orientation of the AraC-domains on the DNA.

At concentrations of ~60 nM ExsA is monomeric in solution and binds specifically to its promoters with high affinity as reflected in apparent K_D values of 5 nM [52]. One ExsA binding site is highly conserved on all of its 10 promoters. Guanine and cytosine at positions -47 and -45 are present in each, as well as a conserved adenine in the -51 position, and a TGNNA sequence centered around position -36 [52]. Two ExsA molecules were found to bind to P_exoT and P_exsC promoters as shown by EMSA studies. DNase footprinting experiments also revealed a protected DNA region of around ~45bps, much larger than would be expected for binding of a single ExsA molecule [60]. A model for ExsA recruitment on its promoter has been proposed whereby a monomer of ExsA first binds near the -35 region and interactions at the NTD facilitates the binding of second monomer to a second upstream binding site [58, 59] (Figure 1.3). This was verified by King et. al. who found that promoter positions -45, -47, and -36 serve as the binding

<table>
<thead>
<tr>
<th></th>
<th>-50</th>
<th>-35</th>
</tr>
</thead>
<tbody>
<tr>
<td>-70</td>
<td>Binding site 2</td>
<td>Binding site 1</td>
</tr>
</tbody>
</table>

**Figure 1.3 Binding site of an ExsA-dimer on promoters.** Two monomers of ExsA bind between the -35 to -70 region of transcriptional start site. Binding site 1 consists of TGnA consensus sequence, while site-2 consists of an adenine-rich site.
site for the first molecule of ExsA prior to cooperative binding of a second molecule upstream [59]. Importantly, EMSA studies produced a lower molecular weight shift with DNA that were mutated at the adenine-rich site, and both binding sites proved necessary for P_exoT activation [52]. Brutinel observed that the ExsA-P_exsC promoter complex had lower mobility in a gel than corresponding complexes with P_exoT and P_exsD, suggesting that binding with this specific promoter may cause changes in DNA topology. Circular permutation-type EMSA assays subsequently demonstrated that ExsA bends the P_exsC promoter up to 78°; a much more dramatic effect than the comparatively modest effect of ExsA on P_exsD and P_exoT which were bent 31° and 20°, respectively [52]. This finding raised the question of whether the more extreme DNA bending of P_exsC is significant in the context of the regulatory mechanisms of the different operons.

ExsA is required for transcription of all T3SS genes [58]. Transcription activators help in transcription either by recruiting σ70-RNAP holoenzyme to the promoter region or by helping to form an open transcription initiation complex. The -35 site is spaced 21-22 bases from the -10 site in ExsA-dependent promoters. Normally, this spacing is about 17 bp. ExsA-mediated DNA bending was speculated to help overcome this abnormal spacing. However, spacing alone is not the only factor involved in the regulation because the same group demonstrated that T3SS genes were not transcribed in the absence of ExsA even when the promoters were redesigned to restore the canonical 17 bp spacing [61, 62]. Subsequent potassium permanganate modification studies and lag plots demonstrated ExsA to be directly involved in the recruitment of σ70 to RNAP, and to aid in the formation of the open complex [52, 61]. A truncated ExsA protein containing the CTD alone (ExsA-CTD_His6) has very low stability but can activate transcription in vivo and in vitro; however, the transcript level in vitro is reduced [58]. The importance of -35 region for
ExsA binding places this activator in the class II activator group in which the DNA-binding site is near the -35 promoter region. Similar to other class II activators, ExsA interacts with the 4.2 region of $\sigma^{70}$ shown by substitution in this region resulted in decreased ExsA-dependent transcription [63].

**P. aeruginosa T3SS regulatory proteins and their interactions**

ExsA is responsible for regulating the *exsCEBA* operon, which includes its own gene. Three of the proteins expressed from this operon, ExsE, ExsC, and ExsA and a fourth protein, ExsD, form a four-protein signaling cascade that ties expression of the T3SS to host-cell contact [64]. To understand the regulation of T3SS in more depth, McCaw *et. al.* took a closer look at all the proteins involved in the regulation of transcription. Hornef *et. al.* had previously shown that the T3SS exo-enzymes did not accumulate in the cytoplasm of secretion apparatus mutant strains.

*Figure 1.4 ExsACDE regulatory model.* This mechanism involves four regulatory proteins, ExsA, ExsC, ExsD, and ExsE. When there is no host cell contact with the bacterial-syringe, ExsA is inhibited by ExsD, its anti-activator. This prevents binding of ExsA to its promoters thus preventing transcription of T3SS genes. However, upon host cell contact, ExsE, is released from the needle complex, which frees the chaperone protein ExsC. ExsC now sequesters ExsD, freeing ExsA in the process. ExsA now binds to T3SS promoters and recruits RNA Polymerase.
thus providing the first evidence of a link between the up-regulation of T3SS genes and active secretion [65]. A comparison of the T3SS machinery of *P. aeruginosa* with that of *Yersinia enterocolitica* revealed many conserved factors except for ExsD, which is only present in *P. aeruginosa*. It was determined that loss of function exsD mutants expressed ExsA-related exoenzymes ExoS, ExoT, ExoU, and ExoY constitutively in a calcium-independent manner. These effects were also seen in strains lacking the secretion apparatus indicating that ExsD has a negative regulatory effect on the T3SS machinery regardless of the extracellular environment [44].

A closer look at the sequence of ExsC revealed characteristics of a T3SS chaperone [66]. T3SS chaperones generally perform two functions: **One**, they stabilize effector proteins in a secretion-competent state. **Two**, they serve as a three-dimensional targeting signals for directing their partner proteins to the T3SS apparatus. ExsE is a small highly charged secreted protein that negatively regulates T3SS expression. Two-hybrid experiments provided the first evidence for the strong interaction between ExsE and ExsC. This study also showed that ExsC interacts with ExsD [67]. The association of ExsC and ExsD is nineteen fold weaker than that between ExsC and ExsE [66, 68]. Similarly, the association of ExsD with ExsA is weaker than those in the ExsC-ExsD complex [69]. The model that ultimately emerged from these studies stated that, when the cells are in non-inducing state, ExsC binds to ExsE, while anti-activator ExsD is bound to ExsA and thereby represses transcription of T3SS genes. Upon induction, secretion of ExsE enables ExsC to sequester ExsD, which in turn frees ExsA to recruit RNA polymerase to the T3SS-related promoters [70] (**Figure 1.4**).

Initially, ExsD was only shown to interact with ExsA, when the two proteins are co-expressed, thus leading to the hypothesis that folding intermediates of the two proteins interact,
which then fold completely upon release of ExsD from the ExsD-ExsA complex [71]. ExsD, a 276 amino acid protein of 31.4 KDa, forms a homo-trimer; however, the high binding affinity of ExsD for ExsC is able to break the trimer to form the observed 2:2 complexes between the two proteins [68, 71]. The interactions of ExsD with ExsA are necessarily weaker; therefore, Bernhards et al. speculated that ExsD trimerization may prevent interactions with ExsA when the two proteins are produced independently. It was shown using a monomeric variant of ExsD that this is indeed the case. In the same study, the authors demonstrated that the kinetic barrier posed by ExsD self-association may also be overcome by an increase in temperature. In vitro transcription assays have shown that independently purified wild-type ExsD and ExsA can interact at 37°C but not at 30°C. The monomeric variant of ExsD on the other hand repressed ExsA-dependent transcription in vitro even at 30°C [71].

A perhaps even more surprising feature of ExsD was revealed by structural studies. The structure uncovered a structural similarity between ExsD and the DNA binding repressor protein KorB. KorB represses the transcription through a combination of DNA looping and interactions with another protein [68]. Structural homology between ExsD and KorB may suggest that ExsD represses ExsA function in similar manner through direct interactions with not only ExsA but also DNA. Differential scanning fluorimetry studies have shown that ExsD binds to DNA [68]. EMSA studies with trimeric ExsD revealed no binding interactions; however, the monomeric variant of ExsD did give a significant shift [71].

Through structural and biochemical studies we have gained a good understanding of the interactions involving ExsC and ExsE, and also those of ExsD and ExsC [67]. However, many open questions remain with respect to the molecular basis for transcription activation by ExsA and the role(s) of ExsD in the regulation of T3SS-associated gene expression. The diverse
experimental results on mechanism of ExsD repression studies put forth a lot of questions and possibilities on how it represses ExsA. Therefore, more kinetic and mechanistic based studies are required in order to understand their interactions. The ExsACDE cascade is conserved in *Vibrio* species and similar mechanism for tying host cell contact to transcription of the T3SS through an AraC-type transcriptional activator have also been found in *Salmonella* and *Shigella flexneri* [72-75].

**Additional regulators of T3SS gene expression**

There are various other intrinsic and extrinsic factors that regulate the T3SS besides the ExsACDE regulatory cascade. There are a variety of environmental factors that have been identified [76]. The T3SS expression is known to be indirectly activated by RsmA, shown by a defective effector and translocation proteins in *rsmA* mutant [77]. RsmA activity is dependent upon two non-coding regulatory RNAs, RsmY and RsmZ. These small RNAs are activated by a GacA-GacS two-component system [78]. This system is known to regulate the switch between acute and chronic infections depending upon the presence of RsmA. The cAMP-signaling

![Figure 1.5 Different regulatory pathways controlling T3SS expression in *P. aeruginosa.*](image-url)
cascade and two additional pathways regulate the switch between acute and chronic infections [76]. cAMP regulates Vfr (virulence factor regulator), a global virulence gene expression regulator in *P. aeruginosa*. The mechanism of how Vfr affects the expression of the T3SS has yet to be worked out, however microarray studies showed that Vfr and cAMP stimulate expression of T3SS [79, 80]. Repression of the T3SS also occurs indirectly by two mini proteins, PtrB and PtrC, which are stress response proteins upregulated by the PrtR/PrtN pathway and MexT, respectively [81]. Repression also occurs indirectly by the RocA/R/S system and AlgR. RocA1 is a repressor of RocR, a component of a three-component system. Compared to wild type PA14, the type three secretion system was increased in the *rocR/S* mutant [82]. However, the mechanism underlying this phenotype is unknown. AlgR is considered one of the major factors responsible for antibiotic resistance in *P. aeruginosa* cells [80] (Figure 1.5).

Besides Vfr and RsmA, the small RNA-binding protein RsmF positively regulates T3SS production [83]. RsmF, an RsmA homolog in *P. aeruginosa* regulates T6SS, the biofilm formation, in addition to T3SS [84]. MucA is an anti-sigma factor related AlgU. The AlgU regulon activates AlgZR, which in turn decreases Vfr expression while increasing RsmYZ levels, both result in a decrease in T3SS gene expression [85].

**ExsA-ExsD interaction: Possible drug intervention point**

*P. aeruginosa* is a dangerous pathogen as it is resistant to multiple antibiotics. Control of the infection caused by this organism in hospital settings is therefore of prime importance. Due to multidrug resistance (MDR), alternative therapeutic agents are needed to combat infections caused by this organism. Targeting the T3SS might be one feasible option, as removal of the T3SS makes the organism less virulent [86]. Mortality related to *P. aeruginosa* infection as a
result of septic shock due to T3SS effectors has been reported [87]. Several studies targeting various factors of the T3SS has been conducted to date including PscF and ExoU [88-90]. PcrV can serve as potential antigen, which can be used to produce a vaccine and has been shown to protect against \textit{P. aeruginosa} in a murine pneumonic plague challenge model [91]. But this vaccine is not favored by hospitals to treat cystic fibrosis patients due to difficulty to assess colonization with the bacteria versus infection leading to ventilator-associated pneumonia [92, 93]. Studies targeting the development of inhibitors to effector proteins are also underway but this approach is not considered universally applicable as different isolates carry different types of effector genes [94]. Direct inhibition of ExsA would constitute an attractive and most efficient option for controlling infection. Therefore, if we understand the molecular basis for the regulation of ExsA by ExsD, we may be able to derive small-molecule ligands that mimic ExsD function.
PROJECT OBJECTIVES

Understanding the molecular basis for the ExsD-ExsA interaction will be critical for explaining how ExsD inhibits the T3SS gene expression. We do know that ExsD interferes with ExsA dimerization, a step that stabilizes ExsA-promoter interactions. While we have a broad understanding of how ExsA binds to the promoter site, a number of questions remain unanswered, particularly with respect to the dynamics of the domain interactions between the regulatory and the DNA binding domains of ExsA. This study could eventually facilitate future efforts to develop drugs that can mimic ExsD and thus control T3SS-related infections.

Therefore, there are two objectives of this research:

1. Solve the structure of ExsA-NTD.
   X-ray crystallography will be used to solve the structure of protein crystals formed by ExsA-NTD to obtain insights into ExsA dimerization and a roadmap for the mutational analysis described under objective 2.

2. Map the binding interface of ExsA and ExsD to understand the mechanism of regulation of T3SS. We used two different approaches to carry out this study.

   a. Use a combination of structure-guided site-directed mutagenesis and in vitro transcription experiments, to systematically map the ExsA-ExsD interface.

   b. Confirm that results showing disruption of regulation of ExsA by ExsD obtained from the in vitro transcription experiments mean a loss of ExsD-ExsA binding, using a cell-based bacterial adenylate cyclase two-hybrid (BACTH) system assay.
Rationale & Approach

**Rationale:** Because ExsD was shown to interfere with ExsA dimerization, we first hypothesized that the two helices involved in ExsA dimerization might be the target of ExsD binding as well. Alternatively, we also considered the possibility that ExsA might also use a region that is involved in small molecule binding in other AraC-type proteins to bind to ExsD. Our lab has solved the crystal structure of the ExsA NTD (encompassing amino acids 10-165). When we aligned this structure with other AraC proteins it showed high homology with the ToxT. The full length structure of ToxT with a 16-carbon fatty acid bound in its beta-barrel binding pocket has been solved and the NTD of ToxT also shows significant structural homology with AraC proteins [95]. Close comparison of the binding pockets of AraC and ExsA revealed that they share a few conserved residues. In AraC these residues are directly or indirectly involved in H-bond formation with arabinose, its natural ligand [96]. Since ExsA is not known to have any small molecule ligand, we hypothesized that these residues are responsible for binding of ExsD, its anti-activator.

Because of the unusual nature of the ExsD-ExsA interaction, we also used the crystal structure of the ExsA-NTD to identify representative and evenly distributed surface residues on ExsA that were also mutated and tested in our assays.

In order to map the ExsD binding site on ExsA the series of generated ExsA variants was tested in our *in vitro* transcription assay for their ability to initiate transcription in the absence and presence of ExsD. Variants that were not affected by ExsD but are still able to activate transcription were examined *in vivo* using a bacterial adenylate cyclase based reporter assay to directly probe the impact of the introduced mutations on ExsD-ExsA binding.
Approach:

A. Structure-guided site-directed mutagenesis coupled with in vitro transcription.

The loss of an effective interaction between ExsD and ExsA may be tracked indirectly using an in vitro transcription assay. Here, the level of inhibition by ExsD is gauged by comparing the amount of mRNA produced by RNAP from an ExsA-dependent promoter in the presence and absence 10 μM ExsD. Therefore, this is a functional assay. In the given experimental set-up ExsD decreases the amount of transcript produced from wild-type or native ExsA (wtExsA) by about 50%. If a mutation in exsA reduces the binding affinity of the resulting ExsA variant to ExsD, the presence of ExsD should not reduce the transcript levels.

B. Bacterial Adenylate Cyclase Two-Hybrid (BACTH) System

We wanted to corroborate the results obtained from our in vitro transcription assay, which constitutes an indirect means of testing the ExsA-ExsD interaction, in a cell-based assay for which we used the BACTH system. This approach also enables us to address the important mechanistic question as to whether a specific mutation truly interfered with ExsA-ExsD binding or primarily abolished the ability of ExsD to inhibit ExsA-dependent transcription. A prior study used bacterial two-hybrid systems to examine ExsA-ExsD interactions (6). BACTH determines interactions between two proteins in E. coli. This system manipulates the ability of adenylate cyclase to synthesize cAMP. The adenylate cyclase consists of fragments T18 and T25, which must interact with each other directly in order to synthesize cAMP. These two fragments can be fused with the target proteins and can be reconstituted in E. coli cya. cya catalyzes the formation of cAMP from ATP in E. coli [97]. If the target proteins physically interact with each other, in our case ExsA and ExsD, then cAMP is produced. Thus produced cAMP binds to the catabolite
activator protein CAP and activates alternative carbon source utilizing genes in *E.coli*. This activation can be visualized on differential media like MacConkey agar which is selective for Gram-negative bacteria and differentiates lactose fermenting versus non-fermenting bacteria. The interaction of wtExsA and wtExsD serves as a positive control.
REFERENCES


8. Vat-de-Griice F: CLASSICS IN INFECTIOUS DISEASES. REVIEWS OF INFECTIOUS DISEASES 1984, 6:3.


77. Mulcahy H, O'Callaghan J, O'Grady EP, Adams C, O'Gara F: The posttranscriptional regulator RsmA plays a role in the interaction between *Pseudomonas aeruginosa* and human airway epithelial cells by positively regulating the type III secretion system. *Infection and immunity* 2006, 74(5):3012-3015.


CHAPTER TWO

Structural analysis of the regulatory domain of ExsA, a key transcriptional regulator of the type three secretion system in Pseudomonas aeruginosa

Manisha Shrestha¹, Yi Xiao¹, Howard Robinson², Florian D. Schubot¹*

¹ Department of Biological Sciences, Virginia Polytechnic Institute & State University, Washington Street, Blacksburg, VA 24060, United States of America, ² Biology Department, Brookhaven National Laboratory, Upton, NY 11973–5000, United States of America

Corresponding Author: * fschubot@vt.edu

PLOS ONE | DOI:10.1371/journal.pone.0136533 August 28, 2015

Author Contributions
Conceived and designed the experiments: MS YX FDS.
Performed the experiments: MS YX HR FDS.
Analyzed the data: MS YX HR FDS.
Contributed reagents/materials/analysis tools: MS YX HR FDS.
Wrote the paper: MS YX FDS.

Key words
transcription activator, type three secretion system, Pseudomonas aeruginosa, ExsA, Crystal structure, regulator domain
Abstract

*Pseudomonas aeruginosa* employs a type three secretion system to facilitate infections in mammalian hosts. The operons encoding genes of structural components of the secretion machinery and associated virulence factors are all under the control of the AraC-type transcriptional activator protein, ExsA. ExsA belongs to a unique subfamily of AraC-proteins that is regulated through protein-protein contacts rather than small molecule ligands. Prior to infection, ExsA is inhibited through a direct interaction with the anti-activator ExsD. To activate ExsA upon host cell contact this interaction is disrupted by the anti-antiactivator protein ExsC. Here we report the crystal structure of the regulatory domain of ExsA, which is known to mediate ExsA dimerization as well as ExsD binding. The crystal structure suggests two models for the ExsA dimer. Both models confirmed the previously shown involvement of helix α-3 in ExsA dimerization but one also suggest a role for helix α-2. These structural data are supported by the observation that a mutation in α-2 greatly diminished the ability of ExsA to activate transcription in vitro. Additional in vitro transcription studies revealed that a conserved pocket, used by AraC and the related ToxT protein for the binding of small molecule regulators, although present in ExsA is not involved in binding of ExsD.
Introduction

Gram-negative bacteria survive under a broad range of environmental conditions. Numerous species entertain mutualistic relationships or infects plant and animal hosts using a diverse array of virulence mechanisms. Perhaps the most prominent among these virulence mechanisms is the type three secretion system (T3SS). The T3SS consists of a needle apparatus, a varying array of exported toxins (or effectors), and, to ensure precisely timed expression, a specialized set of regulatory proteins [1–3]. The needle complex, broadly conserved across bacterial species, is used to transport toxins directly from the bacterial cytosol into the host cell cytoplasm. The types of the exported toxins differ among bacterial species as they appear tailored for specific hosts or niches within a host [4, 5]. While the molecular targets may vary, the secreted virulence factors generally fall into three functional categories: factors that act to subvert the host immune system [5–12], those that induce apoptosis [5, 11, 13–20], or, in case of intracellular bacteria, those that mediate engulfment by the host cell [4, 21–25]. Expression of T3SS-associated genes is usually timed to coincide with host infection. Host sensing is accomplished through a variety of mechanisms such as a shift in nutrient conditions, changes in temperature and physical contact with a host cell [3, 26–30]. A number of well-known mammalian pathogens are among the Gram-negative bacteria that employ T3SSs to facilitate infection. These include Yersinia pestis, Salmonella enterica, Chlamydia species, Vibrio species, and Pseudomonas aeruginosa [1, 2, 5, 31]. P.aeruginosa causes opportunistic acute and chronic infections in wide range of animal and plant hosts [10, 32–34]. The T3SS of P.aeruginosa, transporting at least four potent cytotoxins into host cells, is essential for the establishment of acute infections [7, 14, 18–20, 35–40]; and there is some evidence to suggest that the T3SS is also important during the early stages of chronic infections [41–45]. Underlying the remarkable adaptability of P.aeruginosa to exploit a broad range of hosts are intricate regulatory networks formed by the biggest set of regulatory
proteins among all known bacterial species. Presumably to preserve energy and avoid premature
detection by the host organism expression of the T3SS is also coordinated by a complex network
of signaling pathways [7, 29, 46–48]. One of these pathways, the ExsA-ExsC-ExsD-ExsE
cascade, provides a direct link between bacterial host-cell contact and an upregulation of T3SS-
related gene expression [49, 50]. Following a non-canonical mechanism, signaling is mediated
by the formation and dissociation of three mutually exclusive protein-protein complexes [49–51].
Under non-inducing conditions the transcriptional activator ExsA is sequestered by the anti-
activator protein ExsD [52]. Under these conditions the type three secretion chaperone ExsC and
the 81 amino acid ExsE also form a tight complex [49–51]. Host-cell contact triggers opening
of the basally expressed secretion apparatuses leading to ExsE export [49, 50]. The liberated ExsC
binds ExsD to cause the release of ExsA [53]. The transcription factor ExsA in turn recruit RNA
polymerase to the transcription initiation sites of the eleven promoters that control the expression
of the T3SS genes [49, 50, 52, 54–57]. Over the years the details of the mechanism of
transcription activation have been worked out and we have gained a fairly clear picture for the
structural basis of the ExsC-ExsE and ExsC-ExsD interactions [56, 58]. In addition, the features
of ExsA-dependent promoters have also been elucidated [59]. These promoters contain
consensus regions similar to -35 and -10 sites of constitutive σ70-dependent promoters.
However, while canonical -35 and -10 sites are usually spaced by 17 nucleotides the separation is
21 to 22 bases pairs for ExsA-dependent promoters, a spacing that was shown to be critical for
ExsA-dependent transcription [59]. Each promoter contains two ExsA binding sites. One site
overlaps the -35 element, the site also involved in Sigma-70 binding, while the second site
encompasses the conserved adenine-rich region centered near the -51 position [60–63]. The
ExsA protein consists of a ~100 amino acid carboxy-terminal domain (ExsA-CTD) and a ~170
amino acid amino-terminal domain (ExsA-NTD). The two domains are connected by a flexible linker. ExsA-CTD contains two helix-turn-helix motifs required for binding to T3SS associated promoters [64]. The ExsA-NTD mediates homo-dimerization but is also target for ExsD binding [54, 55]. Most recently, dimerization of the amino-terminal domain of ExsA was shown to be essential for not only stabilizing ExsA-DNA interactions but also for facilitating a structural change in ExsA that permits sequential binding of two ExsA molecules to the promoter [65]. ExsD inhibits ExsA function by interfering with both dimerization and promoter binding of ExsA [54, 55]. ExsA constitutes a particularly attractive target for the development of novel therapeutics because it belongs to the AraC family, which is comprised entirely of bacterial and fungal proteins but not represented in higher eukaryotes [66]. From a standpoint of drug design perhaps the most interesting targets in this signaling cascade are the ExsA-DNA and the ExsA-ExsD interfaces. Previous mutagenesis, functional, and interface mapping studies have provided a reasonably clear view of the ExsA-DNA interactions [60–62]. However, the interface of the ExsA-ExsD complex has so far resisted detailed analysis. AraC-type proteins such as ExsA are usually regulated through small-molecule ligands [66–70]. In the cases of ToxT and AraC these ligands that bind in a conserved pocket within the beta-barrel structure of the regulatory domain [66–70]. Here for the first time we report the structure for the regulatory domain of ExsA that belongs to a group of AraC-type transcriptional factors, which is regulated not by a small molecule ligand but by another protein [22, 23, 25, 71, 72]. Additional mutational analysis coupled with functional assays demonstrate that the conserved cavity in ExsA, although present, is not required for binding of ExsD in vitro, suggesting that this subfamily of AraC proteins is regulated through interactions that involve distinct molecular interfaces.
Material and Methods

Preparation of ExsD and ExsA variants

The genes coding for the ExsA variants genes were generated by site-directed mutagenesis using Quik-Change (Stratagene) kit and the manufacturer’s suggested protocol. The following primers were used:

ExsAW77A 5’-3’ F: GGAAAGGACAGCCGAATACTCGCGATTCCATTATCTGCC
AGTTTCTACAAGGC; R: GCCTTGTAGAAAACCTGGCAGATAATGGAATCGCGAGT
ATTCGGCTGTCCTTTCC
ExsAR25A 5’-3’ F: GTCATTGGAACATTCCAACTTTCGAATACGCAGTAAACAAG
GAAGAGGGCGTATAT; R: ATATACGCCCTCTTCTTGTATTTACGCCTAGATT
GTTGGAATGTTCCTAATGAC
ExsAL117R 5’-3’ F: CTGCCACGCCTCGCGGTCCGGTG; R: CAACCGGCCAGCC
AGGCGTGGCAG
ExsAG124R 5’-3’ F: GCCGGTTGCGTCAAGAGGTTGAAGGAATTGC; R: GCAATTC
CTTCAACCTCTTTGACGCAACCAGGC
ExsAN27A 5’-3’ F: GTCATTGGAACATTCCACCTCTCTTCAATAAGGTAGCCAAG
AAGAGGGCGTATAT; R: ATATACGCCCTCTTCTTGGCTACCCCTGATTCAAGTT
GGAATGTTCCTAATGAC

The ExsAW77A gene served as template to create the exsAR25A-W77A double mutant using primers 5’-3’: F: GCCGGTTGCGTCAAGAGGTTGAAGGAATTGC; R: GCAATTCCT
TCAACCTCTTTGACGCAACCAGGC
**Recombinant protein expression and purification**

All ExsA variants and ExsD were overexpressed in *E. coli* from a vector constructed by Gateway recombinational cloning (Invitrogen, Carlsbad, CA, USA). A tobacco etch virus (TEV) protease recognition site and the appropriate att recombination sites (attB1 and attB2) were added to the exsA and exsD genes during PCR, and the amplicons were subsequently recombined into pDONR201 (Invitrogen). The nucleotide sequences of the ORFs were verified, then recombined into the destination vector pDEST-HisMBP [82] to create the expression vectors pFS-HMBPExsD and pFS-HMBPExsA. These vectors were designed to produce either ExsA or ExsD as a fusion to the C-terminus of an N-terminally His6-tagged *E. coli* maltose-binding protein (MBP). Single colonies of *E. coli* BL21(DE3) CodonPlus RIL cells (Stratagene, La Jolla, CA, USA) containing either expression plasmid were used to inoculate 125 mL of Luria broth (LB) supplemented with 2 g/L dextrose, 100 μg/mL ampicillin, and 30 μg/mL chloramphenicol. The cultures were grown to saturation with shaking (225 rpm) overnight at 37°C and then diluted 66-fold into 6 L of fresh medium. pFS-HMBPExsA containing cultures were grown to an OD600 of 0.6–0.8 prior to induction with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG), while pFS-HMBPExsD containing cultures were induced at an OD600 of 0.6 using the same concentration of IPTG. ExsA expression was induced overnight at 18°C, whereas ExsD variants were expressed at 28°C for four hours. All cells were harvested by centrifugation at 5,000 x g for 15 minutes. For purification cell pastes were resuspended in 200 mL of 500 mM NaCl, 25 mM imidazole, 50 mM Tris-HCl (pH 7.4), 2 mM dithiothreitol (DTT) (buffer A), along with three tablets of Complete EDTA-free Protease Inhibitor Cocktail (Roche Applied Science, Indianapolis, IN, USA). Cells were lysed via sonication and centrifuged at 40,000 x g for 40 min. The supernatants were filtered through 0.45-μm polyethersulfone membranes and applied to a 30
mL Ni-NTA Superflow affinity column (Qiagen, Valencia, CA, USA) equilibrated with buffer A. For each run, the column was washed with five column volumes of buffer A, and proteins were eluted with a linear gradient from 25 to 250 mM imidazole (pH 7.4). The His6-MBP-ExsD protein was digested with 5 mg His-tagged TEV (S219V) Protease [83] while being dialyzed overnight in 150 mM NaCl, 50 mM Tris-HCl (pH 7.4), 25 mM imidazole (pH 7.4), and 1 mM DTT. The sample was then passed through a second Ni-NTA column to remove both the His6-MBP tag and the tagged TEV protease, using the same buffers as employed for running the initial Ni-NTA column. The ExsD sample was obtained from the flow through. This sample was diluted with 50 mM Tris-HCl (pH 7.4) and 2 mM DTT in order to lower the NaCl concentration to 50 mM. The ExsD sample was loaded onto a HiTrap Q HP column (GE Healthcare, Waukesha, WI, USA) that had been equilibrated with 50 mM NaCl, 25 mM Tris-HCl (pH 7.4), and 2 mM DTT. Elution was achieved by applying a linear gradient of NaCl from 50 mM to 1 M. The final purification step involved loading the sample onto a HighLoad 26/60 Superdex 200 prep grade column (GE Healthcare) preequilibrated with a running buffer containing 150 mM NaCl, 25 mM Tris-HCl (pH 7.4), and 2 mM Tris (2-Carboxyethyl) phosphine (TCEP). Purification of the ExsA protein followed a different protocol. Following the initial Ni-NTA affinity purification step, the His6-MBP-ExsA fusion protein was dialyzed against a buffer of 50 mM NaCl, 25 mM Tris-HCl (pH 7.4), and 2 mM DTT and loaded onto a HiTrap Q HP column (GE Healthcare) that had been equilibrated with the same buffer. The His6-MBP- ExsA fusion protein was eluted using a linear NaCl gradient from 0.05 M to 1 M. The sample was dialyzed against 2 L of 45 mM NaCl, 25 mM Tris-HCl (pH 7.15), and 2 mM DTT (buffer B) overnight. The sample was then loaded onto a HiTrap Heparin HP column (GE Healthcare) equilibrated in buffer B and eluted with a 0.05 M to 1 M gradient of NaCl. The NaCl concentration in the His-
MBP-ExsA sample was adjusted to 0.5 M, and the fusion protein was digested with 3 mg of His6-TEV(S219V) protease at 4°C overnight. Next, ExsA was run through a second Ni-NTA Superflow affinity column, this time collecting ExsA in the flow through. Finally, the sample was run through a HighLoad 26/60 Superdex 200 prep grade column (GE Healthcare) using 500 mM NaCl, 25 mM Tris-HCl (pH 7.4), and 2 mM TCEP (ExsA storage buffer). The sample was concentrated to 1 mg/mL, flash-frozen using liquid nitrogen, and stored at -80°C. A sample SDS-gel of the purified proteins is provided with the supplemental materials (S3 Fig).

**Preparation and Crystallization of ExsA-NTD**

ExsA full-length protein was digested by thermolysin (2 mg/mL) at 30°C for 1 hour, leaving only the N-terminal domain (verified by mass spectrometry). The product was applied onto a Superdex-26/60 gel-filtration column (GE Healthcare) (buffer: 500mM NaCl, 25mM Tris-HCl pH = 7.4, 2mM TCEP). The purified ExsA N-terminal domain protein was concentrated to 1.5mg/mL. Selenomethionine (SeMet) ExsA N-terminal domain protein was expressed in *Escherichia coli* BL21 (DE3) cells in minimal medium containing selenomethionine and was purified using the same protocol. As only properly folded protein can yield crystals, we wanted to show the purified ExsA-NTD protein is properly folded and functional. A titration of ExsA-NTD protein in the ExsA-dependent in vitro transcription assay is done to show that the ExsA-NTD protein can inhibit ExsA dependent transcription in a dose-dependent manner, which indicates that the purified ExsA-NTD protein competes for the dimerization site of full length ExsA, resulting in a reduced transcription activity as the DNA binding motif is located in the C-terminal domain of ExsA. The result indicates the ExsA-NTD protein is properly folded and functional (S1 Fig). Crystallization of the ExsA-NTD was performed using the hanging-drop
vapor-diffusion method at 25°C (S1 Fig). Crystals of ExsA-NTD were obtained using a reservoir solution containing 1.6M MgSO4, 0.1M MES pH 6.5 and 0.1M EGTA. The crystallization droplet consisted of 3μl protein solution (1.5 mg/mL ExsA N-terminal domain protein, 500mMNaCl, 25mM Tris-HCl pH = 7.4, 2mM TCEP) and 1μl reservoir solution. Rod-shaped crystals of ExsA-NTD appeared after several days. SeMet ExsA-NTD was crystallized under identical conditions. The crystals were flash-cooled in liquid nitrogen after soaking in a cryoprotection solution containing 90% reservoir solution and 10% (v/v) glycerol.

**X-ray diffraction data collection, structure determination and refinement**

X-ray data were collected at beamline X29A (National Synchrotron Light Source, Brookhaven National Laboratory) using an ADSC Q315 CCD detector. The X-ray diffraction data were processed using the XDS program package. Initial phases for SeMet-ExsA-NTD were obtained by PHASER using the single anomalous dispersion [84]. Model building was performed using COOT [85], and the PHENIX program suite was used for structure solution and refinement [86]. During the refinement data were cut off at 2.5 Å using the correlation of the observed data set with the refined model, CC1/2 as defined by Karplus and Diederichs [87] as selection criterion. Data collection and refinement statistics are summarized in Table 1. The schematic representation of the non-bonding contacts in the A/A’ and B/B” dimers in S2 Fig was generated using PDBsum [88]. The refined model was deposited in the protein data bank under the accession code 4ZUA.

**ExsA-dependent in vitro transcription assays.**

A previously published report indicated that E.coli RNA polymerase is can substitute for its P.aeruginosa counterpart in ExsA-dependent in vitro transcription assays [59]. Preliminary tests confirmed this observation. Therefore, in the interest of time all ExsA variants were analyzed
using commercially available E.coli RNAP holoenzyme (Epicentre Biotech). The linear DNA template used in each assay encompassed positions -207 to 94 of the PexsD promoter, relative to the transcription start site; and from this template, RNA polymerase synthesizes an 82 base mRNA transcript. The template was produced by PCR using forward primer 5′-CATCAGTTGCTGCTCAACAGCG-3′ and reverse primer 5′-CACCGCTTTCTCGGGAGTACTGC-3′. The PCR product was run on a 2% agarose gel and purified using the Wizard SV Gel and PCR Clean-up System (Promega, Madison, WI, USA). Each 30 μL transcription assay reaction contained 4.4 fmole of promoter template, 50.4 μM bovine serum albumin (to eliminate non-specific protein-protein interactions), 10 U purified RNA polymerase holoenzyme, 1 U Riboguard RNase Inhibitor (Epicentre Biotechnologies), 15 ng/μL poly(deoxyinosinic- deoxycytidylic) acid (to prevent non-specific transcription initiation), 133 mM NaCl, 32 mM Tris-HCl (pH 7.4), 10 mM MgCl2, 25 μMEDTA, 0.9 mM TCEP, 0.2 mM DTT, 50μM BSA and 15.5% glycerol. The time-course experiments contained 64 nM ExsA and either no ExsD or 50 μM ExsD. Samples were mixed and allowed to equilibrate at room temperature for five min. Samples were then pre-incubated for 10 min at 37°C. Next, 3 μL NTPs (stock concentrations of 200 μM ATP, CTP, GTP and 40 μMUTP) mixed with 0.2 μL (0.2 μCi) of 3.3 mM P32-alpha UTP was added to each sample to start the reaction, and samples were incubated at either 30 or 37°C, depending on the experiment. After the reactions were stopped by adding 12 μL 1X stop solution (3M ammonium acetate, 50 mM EDTA, 0.11 mg/mL glycogen), 170 μL 100% cold ethanol was added, and the samples were incubated at -20°C for one hr. Following centrifugation at 12,000 x g for 15 min, the supernatant was discarded and pellets were resuspended in 12 μL 1X TBE (Tris/Borate/EDTA)-urea sample buffer and heated at 70°C for five min. After a brief centrifugation, the samples were loaded onto a 10% TBE-urea gel and run
at 200 mV for 60 min. Gels were exposed to a storage phosphor screen (GE Healthcare) for 16 hrs. The phosphor screen was scanned using a Typhoon Trio Variable Mode Imager (GE Healthcare), and gel bands were quantified using Image Quant TL v2005 (Amersham Biosciences, Piscataway, NJ, USA). Each experiment was performed in duplicate.

**Results**

**Crystal structure of the regulatory domain of ExsA.**

Because full-length ExsA resisted crystallization the protein was subjected to limited proteolysis with thermolysin. Remarkably, the digest produced large amounts of a single fragment that could be readily purified via gel filtration chromatography. LC-MS analysis revealed that this proteolytic product encompasses the entire ExsA-NTD (amino acids 2–178). Initial crystallization conditions were identified from commercial sparse matrix screening kits. Careful optimization yielded crystals that diffracted X-rays to about 2.5 Å resolution. Seleno-methionine substituted sample was produced to resolve the phase problem [73], crystallized, and subjected to diffraction analysis at the absorption peak wavelength for Selenium. Single anomalous scattering phasing of these data with PHENIX [74] produced a readily traceable electron density map. Two molecules with very similar overall structures form the asymmetric unit of the crystal. The r.m.s.d. for the superposition of backbone atoms is about 2Å for the two chains. In chain A both termini encompassing residues 2 to 10 and 166–178, respectively produced no interpretable electron density. In chain B the termini were also flexible producing no density for residue ranges 2 to 9 and 167 to 178. A cartoon drawing of chain A is presented in Fig 1A. Despite low sequence similarity the overall fold of ExsA-NTD closely resembles those of the regulatory domains of AraC and ToxT. A structural alignment of ExsA-NTD with AraC-NTD produced an
r.m.s.d. of 3.1Å for the backbone superposition with 16% sequence identity. The overlay of ExsA-NTD with ToxT-NTD gave an r.m.s.d. of 3.2Å for the backbone superposition but only 6% of the structurally aligned residues are identical.

**Helices α-2 and α-3 together form the core of the ExsA dimer interface**

Overall, the fold of this domain is characterized by the formation of a beta barrel structure, which, in case of AraC and ToxT, houses the ligand binding pocket. The beta structure is flanked by two parallel carboxy-terminal helices α-2 and α-3 (Fig 1A). In AraC residues within these helices are critical for homo-dimerization by forming an antiparallel four-helix bundle wherein the two α-2 helices and the two α-3 helices are paired with each other [75]. More recently a mutational analysis of the corresponding region in ExsA demonstrated that α-3 residues C139, L140, K141, E143, L148, and F149 are important for optimal ExsA function [65]. The two molecules in the asymmetric unit are not interacting with each other through α-3. However, a dimer consistent with the experimental findings may be obtained from the ExsA-NTD crystal structure by generating crystallographic symmetry mates for the two independent molecules (Fig 1B). Our structural data suggest that L140, K141, E143, and L148, previously identified as critical for ExsA function [65], are all directly involved in ExsA dimerization (Fig 2A). The two resulting symmetric dimers consisting of either two chain A molecules (designated as A and A’ in Fig 1B) or two chain B molecules (designated as B and B’ in Fig 1B) are not structurally identical but similar. Both provide a structural support for the previously published experimental findings that L140, K141, E143, and L148, which are all positioned on helix α-3 are directly involved in ExsA dimerization [65]. C139 and F149, on the other hand, are not directly positioned at the observed interface suggesting that mutating these residues affects ExsA dimerization indirectly (Fig 2A). In Fig 1C chain A and chain B have been superposed and to
observe the relative positioning of the respective symmetry mates. Compared to A’ in the A/A’ dimer the B” molecule is rotated by roughly 23° around its α-3 helix in the B/B’ dimer. The result of this rotation is a smaller number of interactions in the B/B” dimer. In the A/A’ dimer helices α-2 and α-3 form the core of the interface, whereas no contacts are observed between residues from helix α-2 in the B/B” dimer. A schematic representation of the non-bonding contacts in the two dimers is provided in S2 Fig. It is possible that packing pressures led to the differences between the two dimer interfaces observed in the crystal. Therefore, we set out to determine whether or not contacts mediated by α-2 as observed in the A/A’ dimer are biologically important. To examine this possibility we generated two additional ExsA variants based on our structural data. Because the structural model shows a large dimerization interface, we opted to introduce changes, G124 to arginine and L117 to arginine that should not only reduce the number of intermolecular contacts but also should actively disrupt dimerization (Fig 2A). Dimerization of ExsA is mediated by the ExsA-NTD. While this has not been explicitly shown for ExsA, the AraC-domains of other family members have been shown to facilitate DNA binding and recruitment of the RNA polymerase holoenzyme to their cognate promoters [76–78]. Therefore, a loss of transcription activation as the result of a point mutation in ExsA-NTD would provide indirect support for the hypothesized location of the dimer interface. Both, ExsA-G124R and ExsA-L117R expressed stably and could be readily purified (S2 Fig). In order to ensure that the activity measurements occurred within the sensitive range in vitro transcription assays were performed for three different concentrations of each of the ExsA variants. Although, ExsA-G124R appeared to trend toward inducing lower levels of transcription these differences were not statistically significant given the observed experimental error (P-values ranged from 0.3 to 0.19 for the three different concentrations (Fig 2B). The ExsA-L117R variant, on the other hand,
was at least ten-fold attenuated compared to wild-type ExsA. Statistical analysis gave P values of less than 0.02 for the experiments conducted with 64 nM and 32 nM of ExsA-L117R. At 16 nM ExsA-L117R the lower transcript levels produced a larger experimental error so that the difference between the reactions conducted in the presence of wtExsA and those conducted with the ExsA-L117R variant was no longer statistically significant (P = 0.09). While the large experimental error does not allow us to conclude that G124 is positioned at the dimerization interface the data collectively support the hypothesis that, in addition to α-3, helix α-2 also has a functional role in ExsA dimerization as predicted by the structure of the A/A’ dimer.

The cavity within the ExsA-NTD beta-barrel is not involved in ExsD binding

As pointed out above, ExsA belongs to a subgroup of AraC-type proteins, which are regulated through protein-protein contacts. Beyond knowing that the interactions are facilitated by the regulatory domain, we know very little about the location and properties of the interface. The structures of the ligand-bound forms of ToxT and AraC-NTD revealed binding pockets in virtually the same location within the core of the beta barrel structure [70, 75, 79]. Therefore, we sought to determine if the structurally equivalent region in ExsA also participates in ExsD binding. To test this hypothesis residues, R25, N27, and W77 were respectively replaced by alanines to generate three single substitution variants. In addition an ExsA-R25A-W77A double mutant was also expressed, purified, and tested. The selection of the to-be-modified residues was based on two criteria: together these amino acids line a small cavity at the entrance of the beta barrel core (Fig 3A). Moreover, R25 as well as W77 are not only structurally conserved in AraC but also interact with the arabinose ligand of that protein. All four variants of ExsA expressed stably and were shown to induce ExsA-dependent transcription in vitro. None of the four variants showed a statistically significant reduction in its sensitivity to ExsD (Fig 3B). In fact,
the ExsA-R25A variant showed a statistically significant (P = 0.008) higher sensitivity to ExsD compared to wild-type. However, the effect was less pronounced with the ExsA-R25A-W77A doublemutant (P = 0.09). Based on these results, we speculate that the mutations in the pocket have a slight effect on ExsA stability causing it to undergo the conformational changes induced by ExsD more readily.

Discussion

AraC-type proteins constitute a large family of transcription factors primarily found in bacteria and fungi [66]. Different family members have been shown to regulate stress response, metabolism, and virulence [66]. While prominent members of this group such as MarA are composed of only an AraC-type DNA binding domain and regulated at the transcriptional level [80], the canonical AraC-type transcription factor contains an additional amino-terminal regulatory domain [66]. Generally, binding of small molecule ligands within the regulatory domain modulates the function of the cognate transcription factor. The mechanism of regulation appears to differ among members of this protein family. In AraC, binding of arabinose triggers a rearrangement of an aminoterminal extension, which in turn permits the AraC dimer to bind to two adjacent binding sites on the promoter and activate transcription [81]. In ToxT, binding of a fatty acid molecule within a structurally conserved pocket causes a more compact interaction between N-and C domains to inhibit ToxT function [70]. ExsA belongs to an intriguing subgroup of AraC-like factors that is regulated not through interactions with a small-molecule ligand but by another protein[52]. In P.aeruginosa binding of the anti-activator ExsD interferes with both DNA-binding and homodimerization of ExsA [54, 55]. The latter process has recently been shown to be important for facilitating a conformational change in ExsA that permits consecutive
binding of two ExsA molecules at adjacent sites on the promoter [65]. To date little is known about the structural basis of the ExsA-ExsD interactions or any of the other protein-protein complexes involving AraC-type transcriptional regulators. Despite a low level of sequence conservation the crystal structure of the ExsA-NTD closely resembles those of AraC and ToxT. In conjunction with previous work [65] our structural and experimental data suggest that the dimerization interface of ExsA is also conserved between ExsA and AraC. Yet, the ligand binding pocket used by AraC and ToxT appears to play no role in ExsD binding. The structure of the ToxT-fatty acid complex represents the only known structure of a full length AraC protein with the canonical domain architecture. Yet, this structure and biochemical studies of AraC highlight the mechanistic importance of the intramolecular interactions between NTD and CTD. For example, given the results of the present work and a recently published study, we now have an excellent idea about the location of the interface of the ExsA homodimer. Using the structures of full-length ToxT (pdb code: 3gbg) and the MarA-DNA complex (pdb code: 1bl0) [80] as templates we generated a model for a dimer of a full-length ExsA-DNA complex. In this model DNA binding domains are arranged at a 120° angle (Fig 4). Provided ToxT uses the same dimerization interface as AraC and ExsA this certainly explains why ToxT is inhibited by the fatty acid [70], however, it also raises the question what the “active” conformation of these proteins looks like. Is it a simple loosening of the domain interactions that is causing activation? If ExsA assumes similar conformations, does ExsD binding also encourage the formation of a rigid NTD-CTD interface in ExsA to inhibit transcription? At this point it has become quite clear that the regulatory mechanism of AraC is by no means representative for all AraC-type proteins. Therefore, perhaps contrary to general perception, much remains to be learned about this huge family of microbial transcription factors.
REFERENCES


**Fig 2.1. Crystal structure of the ExsA-NTD.**

A) Model of a monomer encompassing amino acids 2–166 which produced clearly defined electron density. Blue to red rainbow coloring traces the backbone from the N to the C-terminus. Secondary structure elements are numbered. (B) Packing contacts in the crystal suggest the possible structure of the biological dimer. Chains A and B constitute the asymmetric unit of the crystal. Application of two crystallographic two-fold axes produces two additional pairs of chains labeled with a prime and a double-prime, respectively. Contacts between either chains A and A’ or between chains B and B” are proposed to mediate ExsA dimerization *in vivo.* (C) Shown in gray are the overlaid backbones traces of chains A and B. Also displayed are the symmetry-related molecules A’ and B” to highlight similarities and differences between the two possible quaternary structures. The B” molecule is rotated by approximately 23° around helix α-3. The rotation is visualized by marking the angle between the P20 residues of A’ and B” in the figure.
Fig 2.2. Mapping of the ExsA dimer interface.

(A) The shown A/A’ ExsA-NTD interface suggests involvement of helix α-2 in ExsA dimerization. Previously identified interface residues are indicated in the same color as the protein backbone. G124 and L117 are colored violet and yellow in the respective molecules. (B) Shown is a sample gel of measurements testing the impact of the L117R and G124R mutations on the ability of ExsA to activate transcription in vitro. Three concentrations of each protein were tested to ensure that the experiments were conducted in a sensitive range. (C) Graphical representation of the in vitro transcription assays from triplicate experiments. Going from left to right: wtExsA, ExsAG124R, and ExsAL117R.
Fig 2.3. Impact of mutations in the conserved cavity of ExsA on ExsD binding.

Three residues lining the cavity within the beta-sandwich structure of ExsA were mutated with alanine to determine if these residues are involved in ExsD binding. (A) Cartoon depiction of a full-length model of an ExsA-DNA complex. This model was generated by overlaying the structures of ExsA-NTD and a homology model of ExsA-CTD (based on the MarA-DNA crystal structure) onto the structure of ToxT. The mutated residues are depicted as ball-and-stick. (B) Results of in vitro transcription assays measuring the impact of the indicated mutations on ExsA-ExsD interactions. Plotted in the chart is the percent change in obtained transcript level when 10 μM ExsD is added to the reaction. A sample gel showing transcript bands is presented above the chart. Experiments were conducted in duplicate.
Fig 2.4. Cartoon model of a full-length dimeric ExsA-DNA complex.

This model was generated by first overlaying the structure of the ExsA-NTD A/A’ dimer and a homology model of ExsA-CTD (based on the MarA-DNA crystal structure) onto the structure of ToxT. Subsequently, crystallographic two-fold axis was applied to create a model of the full-length protein with a dimer interface corresponding to A/A’ dimer observed in the crystal.
Table 1.1. Diffraction data and crystal structure refinement statistics.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength (Å)</td>
<td>1.075</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>55.51–2.5 (2.589–2.5)</td>
</tr>
<tr>
<td>Space group</td>
<td>P 4_3 2_1</td>
</tr>
<tr>
<td>Unit cell (Å)</td>
<td>a = b = 69.9, c = 191.8</td>
</tr>
<tr>
<td>Total reflections</td>
<td>1,261,113 (1,19,391)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>17,161 (1,581)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>73.4 (75.1)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.36 (94.05)</td>
</tr>
<tr>
<td>Mean l/σ(l)</td>
<td>91.07 (4.53)</td>
</tr>
<tr>
<td>Wilson B-factor (Å²)</td>
<td>57.80</td>
</tr>
<tr>
<td>Rmerge</td>
<td>0.26 (4.10)</td>
</tr>
<tr>
<td>CC1/2</td>
<td>0.96 (0.495)</td>
</tr>
<tr>
<td>CC</td>
<td>0.79 (0.814)</td>
</tr>
<tr>
<td>Rwork</td>
<td>0.239 (0.393)</td>
</tr>
<tr>
<td>Rfree</td>
<td>0.269 (0.403)</td>
</tr>
<tr>
<td>Number of non-hydrogen atoms</td>
<td>2,401</td>
</tr>
<tr>
<td>Macromolecules</td>
<td>2,290</td>
</tr>
<tr>
<td>Water</td>
<td>11</td>
</tr>
<tr>
<td>Protein residues</td>
<td>306</td>
</tr>
<tr>
<td>RMS (bonds) (Å)</td>
<td>0.01</td>
</tr>
<tr>
<td>RMS (angles) (°)</td>
<td>1.38</td>
</tr>
<tr>
<td>Ramachandran favored (%)</td>
<td>53</td>
</tr>
<tr>
<td>Ramachandran outliers (%)</td>
<td>0.67</td>
</tr>
<tr>
<td>Clash score</td>
<td>6.66</td>
</tr>
<tr>
<td>Average B-factor (Å²)</td>
<td>77.7</td>
</tr>
<tr>
<td>Macromolecules (Å²)</td>
<td>77.8</td>
</tr>
<tr>
<td>Solvent (Å²)</td>
<td>61.4</td>
</tr>
</tbody>
</table>

*a*The values in parentheses relate to the highest resolution shell from 2.589–2.5 Å.

*b*Rmerge = Σ[I–(1/I)]/ΣI, where I is the observed intensity, and I is the average intensity obtained from multiple observations of symmetry-related reflections after rejections.

*c*CC1/2 = Pearson correlation coefficient between random half-datasets

*d*CC = [(2CC1/2)/(1+CC1/2)]0.5

*e*Rwork = Σ|Fobs|–|Fc|/Σ|Fobs|, where Fobs and Fc are the observed and calculated structure factors, respectively.

f*Rfree defined in Ref. [72].

doi:10.1371/journal.pone.0138533.t001
CHAPTER THREE

Mechanistic Studies of the Roles of Transcriptional Activator ExsA and Anti-activator protein, ExsD in Regulating the Type Three Secretion System in Pseudomonas aeruginosa

Manisha Shrestha, Robert C. Bernhards, Florian D. Schubot

Author Contributions
RCB and FDS: Conceived and designed experiments of ExsD variants (ExsDΔ20, ExsDΔCC, ExsD1-46). This data has been reported in RCB dissertation.
RCB: conducted experiments for ExsD variants (ExsDΔ20, ExsDΔCC, ExsD1-46). This data has been reported in RCB dissertation.
MS and FDS: Conceived and designed experiments for all ExsA variants and ExsD1-38-MBP construct
MS: Constructed, purified and performed experiments for all ExsA variants and ExsD1-38-MBP construct
MS, RCB, and FDS: Performed data analysis and wrote the manuscript

Keywords: ExsA, type three secretion system, Pseudomonas aeruginosa, site-directed mutagenesis, in vitro transcription assay, BACTH
Abstract

*Pseudomonas aeruginosa* is an opportunistic pathogen associated with both chronic and acute infections. The type three secretion system (T3SS), present in almost all pathogenic Gram-negative bacteria, is an important virulence mechanism during acute infection. The T3SS functions as a molecular syringe; puncturing cell membranes during infection and releasing cytotoxic proteins directly into the host cell cytoplasm. In *Pseudomonas aeruginosa* expression of T3SS-related genes is regulated by a signaling cascade involving the proteins ExsA, ExsC, ExsD, and ExsE. ExsA, a member of the AraC/XylS family of transcription factors, is responsible for activating transcription of all T3SS-associated genes. The function of ExsA is inhibited through direct binding of the anti-activator protein ExsD. The ExsA-ExsD interaction is intriguing because AraC-type regulators are commonly regulated by small molecule ligands and not other proteins. To illuminate the molecular basis of the regulation, we set out to identify the regions and residues of the two proteins that are involved in this novel interaction. Using a combination of structure-guided site-directed mutagenesis and *in vitro* transcription experiments, we systematically searched for the ExsA-ExsD interface. We had previously demonstrated that the interaction between ExsA-ExsD does not involve the putative ligand binding pocket that other AraC-type proteins use to bind their cognate regulator. We have now determined that the amino-terminal 20 residues of ExsD are critical for the interaction. On the ExsA side we discovered critical interactions between ExsD and the backbone of the central sheet structure of ExsA. Follow-up, bacterial-two-hybrid experiments suggest the presence of a larger protein-protein interface. The involvement of the amino-terminus of ExsD in β-strand-β-strand interactions with ExsA also nicely explains how ExsC relieves the ExsD-mediated inhibition of T3SS gene expression because the same region of ExsD also forms β-strand-β-strand interactions with ExsC upon host cell contact.
Introduction

*Pseudomonas aeruginosa* is a ubiquitous Gram-negative bacterium that causes opportunistic infections in many organisms including humans [1, 2]. *P. aeruginosa* can infect the blood, skin, lungs, bones, eyes, and heart of the human body [2]. Infections may follow two distinct paths: acute or chronic [3]. Chronic infections are characterized by the formation of persistent biofilms and are particularly prevalent among patients with cystic fibrosis [4]. Acute *P.aeruginosa* infections progress much more rapidly and are difficult to treat due to the inherent high levels of resistance of the organism to antibiotics [5-7]. While acute infection is mediated by numerous virulence factors, the type three secretion system (T3SS), which is conserved in many pathogenic Gram-negative bacteria, has emerged as a central virulence determinant [5]. This system consists of a multi-protein molecular syringe structure, which spans all the way from the inner membrane of the bacterial cell to the host cell membrane. This syringe pierces host cell membranes and translocates effector proteins into the host cytoplasm [8]. The effector proteins are pathogen-specific and cause a range of adverse effects in the host [9-11]. The T3SS of *Pseudomonas aeruginosa* transports four different well-characterized toxins, ExoS, ExoT, ExoU and ExoY [11]. Their respective effects include disruption of the host cell actin cytoskeleton, inhibition of host cell cytokinesis, phospholipase-induced cell death and inhibition of phagocytosis [6, 9]. Transcription of these effector molecules along with the expression of all of the T3SS structural and regulatory components is finely controlled through a number of regulatory mechanisms. A cascade involving four different proteins ExsA, ExsC, ExsD and ExsE, has been shown to mechanistically link host-cell contact with upregulation of T3SS-related promoters [12-18]. ExsA is the master transcriptional activator for all of these promoters [15]. Prior to infection, ExsA is inhibited by ExsD through the formation of a 1:1 complex that
prevents ExsA homo-dimerization and thus interferes with ExsA-DNA interactions as well [12, 14]. At this point the T3SS chaperone ExsC is sequestered by the 81 amino acid protein ExsE. Upon host-cell contact, ExsE is secreted from the bacterial cell into the host. This secretion of ExsE causes the stimulation of T3SS gene expression as the liberated ExsC protein now sequesters ExsD, to permit ExsA to bind to its cognate promoters and recruit RNAP [19]. Over the years significant structural and biochemical insights into the interactions of ExsE with ExsC and ExsD with ExsC have been gained. The crystal structure of the ExsC-ExsE complex revealed a complex that closely resembles other T3SS chaperone effector complexes. The structure not only explained the previously observed 2:1 stoichiometry but also offered insights into formation of the distinct ExsC-ExsD complex, which has a lower affinity and assumes an unusual 2:2 stoichiometry [13, 14, 16]. However, the structural and biochemical characterization of the ExsD-ExsA reaction interface has so far been elusive. ExsA is a class II type transcriptional activator belonging to the prominent AraC/XylS family of proteins [15]. Members of this family of activators are involved in regulating a variety of processes, including metabolism and virulence [20-29].

The characteristic AraC domain consists of two helix-turn-helix motifs that mediate binding to the cognate promoter elements of the transcription factors [24]. The canonical AraC-type proteins feature an additional amino-terminal dimerization domain that also mediates regulatory mechanisms through interactions with cellular signaling molecules [26, 29-32]. ExsA, a 278 amino acid protein, represents a prototypical AraC protein possessing two functional domains: The carboxy terminal AraC domain interacts with promoter elements and with the 4.2 region of σ70, whereas the NTD facilitates dimerization and binds to the anti-activator ExsD [19]. Invariably, two ExsA monomers bind near the -35 elements of the ten different promoters with
dimerization occurring upon binding to their two recognition sites on the DNA [15, 19]. The proximal DNA binding site, termed site 1 consists of conserved GnC and TGnnA sequences required for binding of the two HTH motifs [15, 18, 33]. Site 2 is not as well conserved, but plays a pivotal role in determining promoter strength by enhancing ExsA-binding affinity and affecting the degree of promoter bending which in turn improves recruitment of RNAP [15, 34]. Prior studies have shown ExsD to sequester ExsA by forming a 1:1 complex until host cell contact has occurred [14]. This direct interaction was first demonstrated using a bacterial-two hybrid assay [12], before two additional studies revealed that ExsD interferes with ExsA function by disrupting the dimerization and thus preventing efficient promoter binding of ExsA [12, 34]. However, exactly how ExsD achieve this has remained unclear. Prior work has also shown that ExsD binds to the NTD of ExsA but this interaction does not involve the canonical ligand binding site of AraC-type transcription factors [12, 33].

Because the ExsA-ExsD interaction plays a central role in regulating P. aeruginosa virulence, development of ExsD mimicking drugs could offer a novel avenue for therapeutic invention during P. aeruginosa infection. Toward this end, we sought to improve our understanding as to how ExsD interferes with ExsA dimerization. Building on previous work, we took a structure-guided mutagenesis approach in conjunction with in vitro transcription measurements in an attempt to map the ExsA-ExsD interface at a molecular level. Subsequent bacterial-two-hybrid studies were used to differentiate between those mutations that disrupted the interaction of ExsD and ExsA entirely and those variants that still permit an interaction but no longer permit transcription inhibition. Through this work we discovered why ExsC-ExsD and ExsD-ExsA interactions are mutually exclusive.
Methods

Construction of ExsA and ExsD expression plasmids

Expression and purification of wild-type (wt) ExsA and wtExsD followed previously published protocols [33, 35]. Briefly, both proteins were respectively overexpressed as His6-MBP-ExsA and His6-MBP-ExsD fusions from the pFS-HMBPExsA and pFS-HMBPExsD vector constructed by Gateway recombination cloning (Invitrogen, Carlsbad, CA, USA). ExsA was overexpressed in E. coli BL21 from a vector constructed by Gateway recombination cloning (Invitrogen, Carlsbad, CA, USA). A tobacco etch virus (TEV) protease recognition site and the appropriate att recombination sites (attB1 and attB2) were added to the exsA gene during PCR, and the amplicons were subsequently recombined into pDONR201 (Invitrogen). The nucleotide sequences of the genes were verified, and then recombined into the destination vector pDEST-His6MBP [36] to create the expression vector pFS-HMBPExsA. This vector was designed to produce ExsA as a fusion to the C-terminus of an N-terminally His6-tagged E. coli maltose-binding protein (MBP). The ExsD expression vector was constructed in similar manner.

Construction of different variants of ExsD and ExsA

The ExsDM59R variant was generated by site-directed mutagenesis using Quik-Change (Stratagene) and the manufacturer’s suggested protocol. The following primers were used:

5´–CTGCAGCGGCGTCCGCGTGGGCTGGGCTGGAGC–3´,
5´–GGCGCGGCCAGCCGCCAGCAGGAAGAC–3´.

The ExsDΔC-C variant was created by two sequential rounds of PCR. The first set of primers was 5´–GTGGAGAACTGTACTTCCAGGGGTATTGGAGCAGGAAGAC–3´ and

5´–GTGGAGAACCTGTACTTCCAGGGGTATTGGAGCAGGAAGAC–3´. These primers were used to amplify codons for 1-137 residues of ExsD.
The second set was

5´–CGGGTCAACCTCGAGGAGGATCGGCACTGGCG–3´

5´–GGGGACAACTTTGTACAAGAAAGTTGCTCATACTGGCAGAGCTGA–3´, these were used to amplify codons for ExsD residues 203-276. Codons to add four glycines were introduced in between the two fragments of DNA. The two products obtained from the first PCR were then used in the second PCR as templates to amplify the ExsDΔC-C with primers forward: 5´-GTGGAGAACCTGTACTTCCAGGGTATGGAGCAGGAAGAC-3´, reverse: 5´-GGGGACAACCTTTGTACAAGAAAGTTGCTCATACTGGCAGAGCTGA-3´. The ExsDΔC-C gene was then introduced into the pDONR201 plasmid and later introduced into the pDEST-His-MBP expression plasmid via Gateway recombinational cloning. This expression plasmid was subsequently transformed into E. coli BL21RIL cells for expression.

The expression plasmid for the ExsD1-38-MBP construct was created using overlap PCR. The primers used in the two rounds of PCR are listed in Table B1. The final construct encoded a His$_6$-ExsD(1-38)-MBP fusion protein containing a TEV-cleavage site, after the hexahistidine tag, a two glycine spacer, ExsD residues 1-38, a single alanine spacer, and the entire reading frame cloned into the MBP pKL66 vector [37].

The 1-46 residue ExsD peptide used in the experiments was purchased from New England Peptide, Inc. (Garner, MA).

All the ExsA variants were constructed using QuikChange (Agilent) using complementary primers with substitution of the bases which was used to amplify pDONR201-wtExsA. Using the Gateway Cloning method, LR reaction the gene was cloned in destination vector where the target protein was cloned in-frame of the C-terminus of the maltose-binding protein. The list of primers used can be found in Appendix B of this dissertation.
Purification of ExsA and ExsD proteins

These protocols have been described previously [33, 35]. A single colony of *E. coli* BL21(DE3) CodonPlus RIL cells (Stratagene, La Jolla, CA, USA) containing the pFS-HMBPExsA expression plasmid was used to inoculate 125 mL of Luria broth (LB) supplemented with 2 g/L dextrose, 100 μg/mL ampicillin, and 30 μg/mL chloramphenicol. The cultures were grown with shaking (225 rpm) to saturation overnight at 37 °C and then diluted 66-fold into 6 L of fresh medium. They were grown to an OD$_{600}$ of 0.8 then induced with IPTG at a final concentration of 1 mM. The induction temperature was 18 °C, and the cultures were shaken overnight. Cells were harvested by centrifugation at 5,000 x g for 15 min. The cell pastes were resuspended in 200 mL of 500 mM NaCl, 25 mM imidazole, 50 mM Tris-HCl (pH 7.4), 2 mM DTT (buffer A), along with three tablets of Complete, EDTA-free Protease Inhibitor Cocktail (Roche Applied Science, Indianapolis, IN, USA). The cells were lysed via sonication and centrifuged at 40,000 x g for 25 min. The supernatants were filtered through 0.45-μm polyethersulfone membranes and applied to a 30 mL Ni-NTA Superflow affinity column (Qiagen, Valencia, CA, USA) equilibrated with buffer A. During the run, the column was washed with five column volumes of buffer A, and proteins were eluted with a linear gradient from 25 to 250 mM imidazole (pH 7.4). The His$_6$-MBP-ExsA fusion protein sample was then dialyzed against a buffer of 50 mM NaCl, 25 mM Tris-HCl (pH 7.4), and 2 mM DTT and loaded onto a HiTrap Q HP column (GE Healthcare) that had been equilibrated with the same buffer. The His$_6$-MBP-ExsA fusion protein was eluted using a linear NaCl gradient from 0.05 M to 1 M. The sample was dialyzed against 2 L of 45 mM NaCl, 25 mM Tris-HCl (pH 7.15), and 2 mM DTT (buffer B) overnight. The sample was then loaded onto a HiTrap Heparin HP column (GE Healthcare) equilibrated in buffer B and eluted with a 0.05 M to 1 M gradient of NaCl. The NaCl
concentration in the His$_6$-MBP-ExsA sample was adjusted to 0.5 M, and the fusion protein was digested with 3 mg of His-tagged TEV(S219V) protease [38] at 4°C overnight. The TEV protease cleaves at the TEV recognition site located between MBP and ExsA. After digestion, ExsA was run through a second Ni-NTA Superflow affinity column to remove both the His$_6$-MBP tag and the protease, using the same buffers as the first Ni-NTA column. The ExsA sample was collected in the flow through. Thus collected ExsA was dialyzed with 500 mM NaCl, 25 mM Tris-HCl (pH 7.4), and 2 mM TCEP [tris(2-carboxyethyl)phosphine] (ExsA storage buffer). The samples were flash-frozen using liquid nitrogen, and stored at -80 °C. Expression on purification of all variants followed essentially the same protocols with the exception that protein expression was induced at 14°C.

The His$_6$-MBP-ExsD protein was digested with 5 mg His-tagged TEV (S219V) protease [38], while being dialyzed overnight in 150 mM NaCl, 50 mM Tris-HCl (pH 7.4), 11.6 mM imidazole (pH 7.4), and 1 mM DTT. The sample was then passed through a second Ni-NTA column to remove both the His$_6$-MBP tag and the protease, using the same buffers as the first Ni-NTA column. The protein sample was collected in the flow through. The sample was diluted with 50 mM Tris-HCl (pH 7.4) and 2 mM DTT in order to lower the NaCl concentration to 50 mM. The ExsD sample was loaded onto a HiTrap Q HP column (GE Healthcare, Waukesha, WI, USA) that had been equilibrated with 50 mM NaCl, 25 mM Tris-HCl (pH 7.4), and 2 mM DTT, and elution was achieved by applying a linear gradient of NaCl from 50 mM to 1 M. Finally, gel filtration was performed using 150 mM NaCl, 25 mM Tris-HCl (pH 7.4), and 2 mM TCEP (ExsD storage buffer). ExsDM59R was purified in the same manner as wild-type ExsD, except that no anion exchange was performed. All purification steps were performed at 4°C. Protein samples were flash-frozen using liquid nitrogen and stored at -80°C.
E. coli BL21 (DE3) cells were transformed with the expression plasmid for ExsDΔC-C and induced with 1mM IPTG at an OD<sub>600</sub> of 0.6. The cells were incubated for 12 h at 18°C, harvested by centrifugation and the pellet was stored at -80°C. Thawed cells were lysed by sonication in binding buffer (500 mM NaCl, 25 mM imidazole pH 7.4, 50 mM Tris-HCl pH 7.4, 2 mM DTT), containing 50 mM of PMSF. The cell lysate was clarified by centrifugation at 15000 x g for 1 hour at 4°C and applied onto a HisTrap Ni-NTA column (GE Healthcare) (wash buffer: 150 mM NaCl, 25 mM imidazole pH 7.4, 50 mM Tris-HCl pH 7.4, 2 mM DTT. elution buffer: 150 mM NaCl, 250 mM imidazole pH 7.4, 50 mM Tris-HCl pH7.4, 2 mM DTT.). TEV protease was added into the elute to cleave the histidine affinity tag as well as the maltose binding protein tag, and the product ExsD coiled coil deleted protein was further purified using a Hi-Trap Ni-NTA column (wash buffer: 150 mM NaCl, 25 mM imidazole pH 7.4, 50 mM Tris-HCl pH 7.4, 2 mM DTT. elution buffer: 150 mM NaCl, 250 mM imidazole pH 7.4, 50 mM Tris-HCl pH 7.4, 2 mM DTT).

**Limited proteolysis of ExsD to generate ExsDΔ20**

ExsDΔ20 was obtained by performing limited proteolysis on wild type ExsD using thermolysin. The digest contained 500 μL of 2.5 mg/mL ExsD in 150 mM NaCl, 25 mM Tris-HCl pH 7.4 and 2 mM TCEP as well as 500 μL 2× thermolysin buffer (20 mM Tris-HCl pH 8.0, 4 mM CaCl2, 400 mM NaCl, and 10% glycerol). To start the reaction, 10 μL of 0.25 mg/mL thermolysin in 1 × thermolysin buffer was added and the sample was incubated at 37°C for one hour. The reaction was stopped by adding 10 μL of 0.5 M EDTA pH 8.0. In order to remove precipitated protein, the sample was centrifuged at 10,500 x g at 4 °C for 10 min then filtered using a spin-column at
10,000 x g at 4 °C for 3 min. The sample was loaded onto a HighLoad 26/60 Superdex 200 prep grade column (GE Healthcare). The column was washed with 150 mM NaCl, 25 mM Tris-HCl pH 7.4, and 2 mM TCEP pH. 8.0. Sample fractions were run on SDS-PAGE. Fractions containing ExsDΔ20 were pooled then the protein sample was aliquoted, flash-frozen, and stored at -80 °C.

**In vitro transcription assay**

The P\textsubscript{exsD} promoter template encompassed positions -207 to 94 of the P\textsubscript{exsD} promoter and it codes for an 82 base mRNA transcript. The P\textsubscript{exsD} template was produced by PCR using forward primer 5´-CATCAGTTGCTGCTCAACAGCG-3´ and reverse primer 5´-CACCGCTTCTCGGGAGTACTGC-3´. PCR products were run on 2% agarose gels and purified using the Wizard SV Gel and PCR Clean-up System (Promega, Madison, WI, USA). Each 30 μL transcription assay reaction contained 4.4 fM of promoter template, 2U RNA polymerase from *E.coli*, 1 U RiboGuard RNase Inhibitor (Epicentre Biotechnologies), 15 ng/μL poly(deoxyinosinic- deoxycytidylic) acid (to prevent non-specific transcription initiation), 133 mM NaCl, 32 mM Tris-HCl (pH 7.4), 10 mM MgCl\textsubscript{2}, 25 μM EDTA, 0.9 mM TCEP, 0.2 mM DTT, 50μM BSA and 15.5% glycerol. Each experiment contained 128 nM ExsA. Samples were mixed and allowed to equilibrate at room temperature for five minutes. To start the reaction, 3 μL NTPs (stock concentrations of 200 μM ATP, CTP, GTP and 40 μM UTP) mixed with 0.2 μL (0.2 μCi) of 3.3 mM P\textsuperscript{32}-alpha UTP was added to each sample, and samples were incubated at 37 °C for 10 minutes. All the assays were performed at this temperature unless specified. The reactions were stopped by adding 12 μL 1X stop solution (3M ammonium acetate, 50 mM EDTA, 0.11 mg/mL glycogen). Then 160 μL 100% cold ethanol was added, and the samples were incubated at -20°C for one hr. Following centrifugation at 12,000 x g for 15 min, the
supernatant was discarded and pellets were resuspended in 12 μL 1X TBE (Tris/Borate/EDTA)-urea sample buffer and heated at 70 °C for five minutes. After a brief centrifugation, the samples were loaded onto a 10% TBE-urea gel and run at 200 mV for 60 min. Gels were exposed to a storage phosphor screen (GE Healthcare) for 16 hr. The phosphor screen was scanned using a Typhoon Trio Variable Mode Imager (GE Healthcare), and gel bands were quantified using the Image Quant TL v2005 software package (Amersham Biosciences, Piscataway, NJ, USA). Each experiment was performed at least in triplicate.

**Bacterial adenylate cyclase two-hybrid assay**

BACTH determines direct interactions between two proteins in *E. coli*. This system manipulates the ability of adenylate cyclase to synthesize cAMP. The adenylate cyclase consists of fragments T18 and T25, which do not interact on their own but must be brought together to synthesize cAMP [39, 40]. These two fragments are encoded for on separate plasmids where they can be fused with the genes of target proteins and co-expressed in *cyA E. coli* strain BTH101. If the target proteins physically interact with each other, then two fragments conjointly synthesize cAMP. Thus produced cAMP binds to catabolite activator protein CAP, which activates alternative carbon source utilizing genes in *E.coli*. This activation can be visualized easily on differential media like MacConkey agar as the colonies with interacting proteins turn pink. The interaction of wtExsA and wtExsD served as a positive control. The T18 fragment is supplied in a pUT18 vector, which is a high copy number vector and the T25 fragment is supplied in the low copy number pKNT25 vector. Genes of both partner proteins were fused to the 5’-ends of the fragment genes because the N-termini of both ExsD and ExsA were shown to play key roles in their mutual interaction. All ExsA variants that were attenuated in their response to ExsD *in vitro*, and wtExsA, were tested for binding to wtExsD. All of the exsA
mutant genes were added to both the pKNT25 and the pUT18 vector in-frame with the fragment 25 and fragment 18, respectively. *exsD* was ligated in-frame with fragment 18 and co-transformed in *E. coli cya* BTH101 strain. Successfully ligated constructs were re-constituted in chemically competent BTH101 with the respective partner plasmid of ExsD. The strain was also co-transformed with positive control vectors (pUT18C-zip and pNT25-zip) to determine the efficiency of competence and serve as additional positive control for the assay on MacConkey plates. The empty T18 and T25 vectors were co-transformed into BTH101 to serve as negative controls for the assay. After growing the cells overnight in LB under ampicillin (100μg/ml) and kanamycin (50μg/ml) selection at 37°C, the bacteria were plated on MacConkey agar plates and incubated at 30°C for 1–3 days. Bacteria co-transformed with plasmids expressing interacting proteins produced a Cya+ phenotype. On MacConkey agar the associated increased expression of β-galactosidase is manifested by the development of a pink color. The lack of an interaction, on the other hand, results in colorless colonies on the indicator plates.
Results and Discussion

The coiled-coil region of ExsD is not important for ExsA binding

Prior studies established several important facts that informed our experimental analysis. ExsD prevents ExsA dimerization but does not bind to the canonical ligand binding pocket. The targeted ExsA dimer interface consists of a four-helix bundle formed by the helices α-2 and α-3 from each of the two molecules [33]. Therefore, we hypothesized that ExsD, rather than acting through interactions with the commonly used binding pockets, interferes directly with ExsA homodimerization by using its extensive coiled-coil region to form a competing four-helix bundle in the ExsD-ExsA complex. In order to test this hypothesis, the ExsDΔC-C variant was constructed wherein residues 138-202 of the original sequence were replaced by four glycine residues. Following overexpression and purification, the ExsDΔC-C variant was examined using our established in vitro transcription assay (Fig. B1). We predicted if the coiled-coil region was indeed important for ExsA binding as hypothesized, then we would see a loss of inhibition of ExsA-dependent transcription by ExsDΔC-C in the assay. Multiple concentrations of ExsDΔC-C were tested at 37°C (Fig. B2). To our surprise, rather than being attenuated, ExsDΔC-C turned out to be a more potent inhibitor than wild-type ExsD (IC\textsubscript{50} = 0.97 ± 0.4 μM) (Fig. 3.1), which had an IC\textsubscript{50} of about 7 μM [35]. Our results clearly demonstrate that the coiled-coil region of ExsD is not important for ExsA binding or inhibition. We were also able to rationalize the observed higher potency of the ExsDΔC-C variant. We and other have demonstrated that ExsD forms homotrimers in solution, with the coiled-coil region forming a substantial part of the trimerization interface. While deletion of the coiled-coil region may not have fully disrupted the trimer, the change likely reduced the stability of the trimer. This is significant because we have previously demonstrated that the monomeric ExsDM59R variant is a more potent inhibitor of
ExsA-dependent transcription than wtExsD [35]. Under our assay conditions ExsDM59R gave an IC₅₀ value of 1.3 ± 0.1 μM, which is remarkably similar to the value we measured for ExsDΔC-C and therefore consistent with our explanation that ExsDΔC-C is a better inhibitor than wtExsD because the former protein is in fact monomeric under our assay conditions.

**The amino-terminal 20 residues of ExsD are important for inhibiting ExsA-dependent transcription**

In previous work, we uncovered that the amino-terminus of ExsD contains a single binding motif that is crucial for the interaction with the T3SS chaperone ExsC [13]. ExsE, the fourth protein in the cascade, contains two copies of this motif explaining its higher affinity for the ExsC dimer compared to ExsD. Because the ExsC-ExsD and ExsD-ExsA interaction are also mutually exclusive, we hypothesized that the amino-terminus of ExsD may also be involved in binding of ExsA. In order to test this hypothesis we took advantage of our previous discovery that limited proteolysis of full-length ExsD with thermolysin produces a highly resistant ExsD fragment that lacks 20 amino acids at the N-terminus. This strategy was originally used to aid crystallization and structure determination of ExsDΔ20 [35]. Therefore, we know that the proteolysis does not affect the tertiary structure of ExsD. ExsDΔ20 was tested in the *in vitro* transcription assay at 37°C. While the precise determination of an IC₅₀ value was not possible due to the limits on ExsDΔ20 solubility, we can estimate it to be around IC₅₀ ~40 μM and thus about 6-7 times higher than for wtExsD, indicating that the ExsD N-terminus is indeed important not only for ExsD-ExsC interaction but also for binding of ExsD to ExsA (Fig. 3.1 and Fig. B3). To confirm that the observed effect of ExsDΔ20 is specific to inhibit ExsA-dependent transcription, we performed an *in vitro* transcription assay time course using a DNA template containing the ExsA-independent RNA-1 promoter and observed no effect (Fig. B4).
Next we sought to determine if the ExsD N-terminus is not only important for ExsA binding but sufficient for ExsA regulation. To test this, a synthetically made polypeptide composed of the first 46 amino acids of the ExsD N-terminus (ExsD_{1-46}) was purchased and examined in the *in vitro* transcription assay at 37°C. ExsD_{1-46} does indeed inhibit ExsA-dependent transcription, albeit weakly (Fig. 3.1 and Fig. B5). Once again, we confirmed that the observed effect of ExsD_{1-46} is specific to ExsA-dependent transcription by performing an *in vitro* transcription assay time course experiment using a DNA template containing the RNA-1 promoter (Fig. B6). ExsD_{1-46} showed no significant inhibition, indicating that ExsD_{1-46} action is dependent on ExsA. Since removal of the coiled-coiled region made the ExsD variant more potent as stated above, we questioned whether the remaining 245 amino acids at the C-terminus of ExsD was just present to provide bulk needed to sterically interfere with ExsA dimerization and could be non-specific. In order to test this, we generated a construct contains the first 38 residues of ExsD fused to the amino-terminus of MBP. However, this construct showed no inhibition of ExsA-dependent transcription even at 20 μM (Fig. 3.2). Collectively, these results suggest that the amino terminus of ExsD is important for ExsA regulation but other regions of ExsD also play a specific role, suggesting the presence of an extensive ExsD-ExsA interface.

**Conserved surface exposed residues on ExsA are not crucial for ExsD binding**

We have previously shown that mutations in putative binding pocket of ExsA, which is conserved in AraC and ToxT for binding their respective ligands do not affect ExsD binding [33]. ExsD also does not appear to directly compete with ExsA homodimerization [33]. Therefore, we took a broader approach for identifying residues of ExsA that are critical for ExsD
binding. To limit the scope of the number of variants we needed to generate, we applied two selection criteria. The candidate residues should be positioned on the surface of ExsA-NTD and the residues should be conserved among ExsA homologs. There are thousands of AraC-like proteins but ExsD homologs are rare. Therefore, we sought to compare only AraC-type proteins that also had an ExsD protein ligand to identify additional residues that we could target via site-directed mutagenesis. We first identified ExsD homologs and subsequently screened for the presence of an ExsA homolog in the same operon even though we are cognizant of the fact the exsD and exsA are not located on the same operon in P. aeruginosa. Seven different exsA-exsD pairs were identified using this approach (Fig. B7). The ExsA sequences were aligned and amino acid conservation was determined. Highly conserved residues with large predicted surface area in the structure of ExsA-NTD from P. aeruginosa (more than 37% of the residue exposed) were chosen for further investigation. Eleven evenly distributed residues were selected for mutagenesis to generate variants carrying single alanine substitutions: E37, T48, Q90, R91, L95, E98, R101, L129, L140, E144, and F151. All variants stably expressed and could be readily purified using our generic protocols for purifying ExsA (Fig. B1). However, preliminary tests revealed that only eight of the ExsA variants were able to stably activate in vitro transcription from an ExsA-dependent template. None of the eight functional variants was attenuated in its ability to interact with ExsD as addition of ExsD inhibited ExsA-dependent transcription at levels comparable to those seen for wtExsA (Fig. 3.3).

We reasoned that the interactions between the two proteins- ExsA (~31kDa) and ExsD (~31kDa) may encompass a large interface therefore a single alanine mutation may not perturb this interaction sufficiently to show a significant effect in the in vitro transcription assay. We therefore decided to introduce disruptive substitutions of the surface residues that would actively
interfere with ExsD-ExsA binding. Seven residues were selected for the substitution: R25F, N27F, T48R, E98F, L140R, E144F, and F151R. Like all the previous variants, all seven variants were initially expressed as His-MBP fusion proteins and readily purified, however, upon TEV-protease digestion, ExsAE98F proved unstable and was eliminated from further experiments. Of the remaining six, ExsAL140R and ExsAF151R had lost the ability to activate transcription. In corroboration to this finding the F151L substitution also results in decrease of ExsA dimerization according to a LexA two-hybrid assay [Personal communication, Yahr Lab]. The other ExsA variants retained their ability to activate transcription similar to that of wtExsA with ExsAR25F appearing to work even somewhat better than the wild-type protein (Fig 3.3). However, once again all four variants proved sensitive to the addition of ExsD in the in vitro transcription assay to a comparable degree as wtExsA, suggesting that the four modified residues are not critical for ExsD binding.

Residues near the amino-terminus of ExsA play an important role in the regulation mechanism in vitro

As we were probing the entire surface of ExsA-NTD residues critical for the interaction with ExsD, we continued to examine other regions that are known to be critical for function in either ExsA or other AraC type proteins. In this context, we noted that the flexible N-terminus of AraC had been shown to play a pivotal role in the regulation of the protein [41]. The flexible region is shorter in ExsA compared to AraC but the first ten amino acids are disordered in the crystal structure of the ExsA-NTD, which is indicative of a dynamic segment of the protein that could serve a similar function in ExsA as it does in AraC [33]. To test this hypothesis we initially replaced the first seven codons of ExsA with five glycines to create a Gly$_5$$\Delta$7ExsA variant. The five glycine were inserted to ensure successful cleavage of the His$_6$-MBP-Gly$_5$$\Delta$7ExsA fusion protein with TEV protease after the initial purification step. The Gly$_5$$\Delta$7ExsA protein promoted
ExsA-dependent transcription at levels comparable to those observed for wtExsA. When ExsD was included in the assay, we were excited to find that the anti-activator protein had no measurable effect on transcription activation by Gly5Δ7ExsA (Fig. 3.4), suggesting that we had finally identified ExsA residues pivotal for ExsD binding. To corroborate the direct role of the first seven amino acids of ExsA in the interaction with ExsD, we subsequently created a Δ7ExsA variant, where the five glycine residues have been removed. We had initially anticipated problems for this construct with the removal of the His6-MBP tag but observed efficient cleavage by TEV protease. The purified Δ7ExsA protein readily induced transcription from the ExsA-dependent promoter. However, unexpectedly, unlike the Gly5Δ7ExsA variant, the function of Δ7ExsA was inhibited by ExsD, suggesting that the previously observed loss of inhibition of the Gly5Δ7ExsA variant had been caused by an indirect effect (Fig. 3.4).

Nevertheless, we reasoned that, if the five glycines could interfere with ExsD signaling, the interacting residues must be in close proximity to the amino-terminus of ExsA. Therefore, we created eleven additional variants to systematically probe this area of transcription factor. The following individual mutations were introduced: Q2A, K5A, G8P, R9A, K10A, Q11A, I12A, S14A, C15A, H16A, and W17A. The G8P mutation was designed to probe how important the flexibility of the amino-terminus is for permitting ExsD binding. All eleven variants were successfully purified and all specifically induced transcript production from the ExsA-dependent promoter. However, once again, the transcription activation functions of these variants were also inhibited by ExsD to a degree comparable to wtExsA (Fig. 3.4). Overall, these findings indicate that conservation of individual residues of the ExsA N-terminus is not critical for ExsD binding (Fig. 3.4).
Secondary structure elements play an important role in transcriptional regulation of T3SS via ExsD

The indirect impact on ExsD binding caused by the added glycines in the Gly5Δ7ExsA variant drew our attention to the nearby beta barrel structure that houses the ligand binding pocket in canonical AraC family enzymes. We have previously shown that conserved pocket residues are not critical for ExsD binding but we decided to probe as to whether the beta-sheet structure itself was involved in the interaction by forming intermolecular β-strand to β-strand contacts with ExsD. This hypothesis is particularly attractive because we have previously shown that the amino-terminus of ExsD is not only important for ExsA binding but also interacts with T3SS chaperone ExsC via such β-strand to β-strand interactions [13]. In order to test the hypothesis that ExsD and ExsA are primarily interacting through backbone contacts we introduced four proline substitutions into the outmost beta strands of ExsA-NTD β-1 and β-2. We reasoned, because prolines’ amide groups are not available for hydrogen-bonding contacts their introduction should interfere with ExsD binding if our hypothesis was correct. Proline-substituted variants of residues Ser-14, His-16, Tyr-24 and Val-26, whose amide groups are predicted to be available for mediating hydrogen bonding interactions with incoming β-strands were constructed. All four variants expressed stably, were purified, and tested in our in vitro transcription assay. Excitingly, both the ExsAY24P and the ExsAV26P variants proved insensitive to inhibition by ExsD during the assay (Fig 3.5). In fact, the amount of produced transcript even increased slightly when ExsD was added to the ExsAY24P containing assay (Fig. 3.5). In contrast, the S14P and H16P substitutions in ExsA had no effect on ExsD binding (Fig 3.5), suggesting that strand β-2 but not β-1 mediates contacts with ExsD. As will be shown later in our bacterial-two-hybrid studies a single proline substitutions did not completely abrogate ExsA-ExsD binding but weakened the contacts sufficiently so as to interfere with the regulatory
process. Therefore, we hypothesize that the increase in transcript levels observed for the ExsAY24P variant in the presence of ExsD stems from a structural stabilization of this variant during the assay through its, albeit weakened, interaction with ExsD. We attempted to create an ExsAY24P-V26P double mutant in order to determine if we would observe an additive effect but that protein was not stable expressed, likely because the kinking caused by the two prolines interfered with the interactions between β2 and β7 within the beta barrel structure of ExsA.

To determine if, as hypothesized, backbone contacts are the primary determinants for ExsD-ExsA binding two additional ExsA variants were constructed. Here, residues 13-17, which form the β1 strand and β2 strand formed by amino acids 23-27 were respectively substituted with five alanines each. The idea for these changes was informed by previous work showing that stretches of alanines may assume different secondary structures depending on the flanking residues and overall structural context [42]. We speculated that the adjacent beta-barrel structure should induce strand formation in both instances despite the concurrent introduction of five substitutions. This uncertainty would have been a bigger concern if we had attempted to introduce loss-of-function mutations but our model predicted that both variants should still be fully functional and remain susceptible to the presence of ExsD during the in vitro transcription experiment. Both variants carrying the penta-alanine substitutions readily expressed and purified without issues. Both proteins activated transcription from the ExsA specific promoter and, consistent with our model, both were indeed inhibited by ExsD as least as strongly as wtExsA (Fig 3.5). In conjunction with our findings that the Y24P and V26P mutations alone did interfere with the interactions of ExsA with ExsD these finding provide compelling evidence that it is the backbone of the ExsA β2 that is pivotal for ExsD binding presumably through strand-strand interactions.
**The Y24P and V26P mutations in ExsA do not fully disrupt the ExsD-ExsA interface**

The *in vitro* transcription experiments represent an indirect means for probing the interactions between ExsA and ExsD. The assay offers the advantage that any misfolded ExsA variants are rapidly identified and discarded when they fail to induce transcription. In order to ask the question as to whether or not any mutations in ExsA disrupt ExsA-ExsD binding entirely, we used the bacterial-two-hybrid system (BACTH) as a complementary tool to directly measure their interaction. We were confident that the system would be well-suited for the analysis because bacterial-two-hybrid studies played an important role in the initial discovery of the ExsA-ExsD interaction [12]. The selected BACTH system manipulates the ability of adenylate cyclase to synthesize cAMP. The reconstituted adenylate cyclase consists of fragments T18 and T25, which must be brought in close proximity by the interacting proteins to catalyze cAMP synthesis. These two fragments are fused with the target proteins for co-expression in *E. coli*. If the target proteins physically interact, in our case ExsA and ExsD, cAMP is produced. In BTH101 *E. coli*, cAMP activates expression of a β-galactosidase from the *mal* promoter allowing for fermentation of maltose, which is visualized on MacConkey agar plates by the presence of pink colonies. Remarkably, all tested variants of ExsA including Gly5Δ7ExsA, ExsAY24P and ExsAV26P still showed an interaction with ExsD (Fig 3.6). It therefore appears that the strand-strand interactions between ExsD and ExsA are critical for proper positioning of ExsD on ExsA to allow interference with ExsA dimerization, but the entire interface likely extends to additional regions of both proteins. This finding is consistent with our prior work on ExsD variants, which demonstrated that the amino terminus of ExsD, which likely forms the beta-strand contacts with ExsA, is necessary but not sufficient for efficient inhibition of ExsA-dependent transcription.
Conclusion

ExsA is the central transcriptional activator for expressing the genes responsible for the T3SS. ExsA belongs to a select sub-group of AraC-type transcription factors that are regulated by other protein factors. Interestingly, all of these proteins are regulators of T3SSs and in all cases the regulation is tied to a signaling cascade that links activation of T3SS gene expression to host contact [43-46]. These contacts could be of some biomedical interest because T3SS are pivotal virulence determinants in these pathogens, suggesting that a small-molecule ligand mimicking the action of the cognate ligand might be effective in combating infection. This concept is particularly intriguing in light of the fact that AraC-type proteins are not conserved in mammals but only found in microorganisms. In *P. aeruginosa*, ExsA is regulated through a four-protein ExsA-ExsC-ExsD-ExsE cascade, wherein inhibition of ExsA by ExsD under non-inducing conditions is alleviated under inducing conditions when ExsE is secreted and the liberated T3SS chaperone sequesters ExsD. Previous studies have shown that, ExsD, the anti-activator protein directly binds with the NTD of ExsA and prevents transcription activation by interfering with ExsA dimerization and promoter binding [14, 15, 17]. Although the interactions of ExsD and ExsA and similar pairs in other organisms have been known for some time now the molecular basis for these interactions has remained a mystery. With hindsight it becomes quite clear why analysis of the ExsA-ExsD interface has proven so difficult. The interaction does not directly involve the dimerization interface, the DNA binding site or even the canonical ligand binding pocket of ExsA. Instead ExsA-ExsD binding likely involves numerous weaker interactions none of which alone is absolutely required. Moreover, the important backbone contacts between the ExsA beta-barrel and the amino-terminus of ExsD are insensitive to single amino acid changes unless a proline is introduced. Structural and biochemical analysis has
implicated the amino-terminus of ExsD as important contributor for its interaction with ExsC [13] and we have now shown that the same region is also pivotal for ExsA binding. Remarkably, the amino-terminus was actually absent in the crystal structure and to date we have been unable to identify any other pivotal residues among the remaining ~240 amino acids that actually form the stable tertiary structure of the ExsD protein. Even deletion of the extensive coiled-coil region of ExsD was well-tolerated. Nevertheless, the results here presented clearly suggest that other ExsD regions also contact ExsA.

Because of the likely extensive nature of the ExsA-ExsD interface, rather than more additional mutational analysis structural studies of an ExsD-ExsA complex may offer the most promising strategy. To date the complex has eluded all of our attempts for crystallization, but perhaps a trimmed construct encompassing only the ExsA-NTD and a minimal ExsD construct might be more amenable to crystallization. Such a structure would also clarify as to how an interaction primarily mediated by backbone contacts derives its specificity. In the case of binding of ExsD to ExsC, the beta-strand interaction are augmented by side-chain contacts of residues within a conserved sequence motif in ExsD [13]. It remains to be seen, how specificity is accomplished in the ExsD-ExsA system. Here interactions need to be necessarily weaker to permit ExsC to sequester ExsD after ExsE secretion. Therefore, specificity is likely accomplished through a series of rather transient contacts that are nevertheless pivotal for signal transduction. This is somewhat discouraging in the context of drug design as it will be difficult to mimic the effect of ExsD with a small-molecule ligand. Perhaps a better strategy would be to either directly target the DNA binding site of ExsA or its dimerization interface. In fact, there is some precedent for attempting the former [47].
REFERENCES


82


**Fig. 3.1. Effect of ExsD variants on ExsA dependent transcription in vitro.**

**a.** Shows the structure of the ExsD trimer, in yellow is the coiled-coil region highlighting its contribution to trimer formation.

**b.** Shows a model of an ExsD monomer and of the hypothetical model of the ExsDΔC-C variant.

**c.** Analysis of four different ExsD-derived polypeptides with respect to their ability to inhibit ExsA dependent transcription in vitro. Estimated IC$_{50}$ values for the four construct where: IC$_{50}$ (wtExsD) = 7.5 μM, IC$_{50}$ (ExsDΔ20) = 40 μM, IC$_{50}$ (ExsDΔC-C) = 1 μM, and IC$_{50}$ (ExsD(1-46)) = 50 μM.
Fig. 3.2. The structurally ordered region encompassing amino acids 39 to 278 of ExsD provides more than bulk to the transcriptional regulation of ExsA. A fusion construct containing the first 38 residues of ExsD fused to MBP was purified and tested in the assay. This variant did not yield any inhibition of ExsA-dependent transcription even at 20 μM. Higher band denotes ExsA-independent transcript which is 108 bases long (shown on its own in lane 1), while the lower band represents the 82 bases long ExsA-dependent transcript. Four independent reactions were quantified to plot the graph.
Fig 3.3. Conserved surface residues of ExsA are not critical for ExsD binding. a. Structure of the ExsA-NTD wherein the targeted surface residues are highlighted. b-d. Shows the representative autoradiograms of three independent \textit{in vitro} transcription experiments analyzing ExsA variants containing substitutions in conserved surface residues. Quantitative analysis of the autoradiograms was performed with ImageQTL. None of the eight functional variants of ExsA carrying mutation in conserved surface residues lost interaction with ExsD. Introduction of disruptive substitutions by either adding residues with opposite charge or significantly larger side chains produced four additional variants that could be stably purified and tested in the assay. All four of them were also sensitive to presence of ExsD in the assay.
Fig. 3.4. Side chains of individual residues near the amino-terminus of ExsA are not critical for the regulatory mechanism but modification of the amino-terminus can impact this regulation. In the structure of ExsA-NTD, the first 10 residues are disordered. In AraC, amino acids in the same region but longer stretch plays a role in the arabinose stabilization, thus regulation. a. The Gly5Δ7ExsA variant was insensitive to the presence of ExsD in the assay denoted. However, Δ7ExsA protein behaved like wtExsA. a., b. and c. Point substitutions in eleven more residues in this periphery of the ExsA amino-terminus showed no effect on ExsD regulation in our in vitro transcription assay.
a. | ExsA (128nM) | WT | WT | 13-17ala | 13-17ala | 23-27ala | 23-27ala |
| ExsD (10μM) | 0  | 10 | 0         | 10        | 0          | 10        |

![Graph showing % decrease in transcript production in presence of ExsD](image)

b. | ExsA (128nM) | WT | S14P | H16P | Y24P | V26P |
| ExsD (10μM) | 0  | 10   | 0    | 10   | 0    |

![Graph showing % change in transcript production in presence of ExsD](image)
Fig. 3.5. Secondary structure elements play an important role in transcriptional regulation of T3SS via ExsD. Because ExsD uses backbone contacts to bind to ExsC, we tested if the same was true for its interaction with ExsA. **a.** Residues in β-strands β-1 and β-2 of ExsA were replaced by alanines to determine if the side chains are important for ExsA binding. Both variants still responded to ExsD demonstrating that the side chains of the substituted residues are not critical. **b.** In order to probe if backbone contacts mediated by β-1 and β-2 are important for ExsD binding proline substitutions were made in both strands. Both the Y24P and the V26P substitution impacted ExsD dependent regulation. **c.** Shows a hypothetical model of the interaction of the amino-terminal residues of ExsD with the beta barrel structure of ExsA. Y24 and V26 of β-2 are highlighted.
Fig. 3.6. Results for bacterial two-hybrid assay for the interaction between ExsA variants and ExsD. When grown on MacConkey agar pink colonies indicate the two proteins are interacting with each other. + indicate the positive control for the assay where pKT25-ZIP and pUT18C-ZIP are reconstituted in the BTH101 cells. – indicate the negative control for the assay consisting of E. coli transformed with the empty pKNT25 and pUT18 empty vectors. All ExsA variants which showed loss of regulation by ExsD in vitro, still show binding with ExsD.
CHAPTER FOUR

OVERALL CONCLUSIONS
The primary objective of this work was to gain structural and functional insights into the unique regulatory mechanism governing the actions of ExsA, the central transcriptional activator of all T3SS-related genes in the human pathogen *P. aeruginosa*. This work addresses the major remaining gap in our understanding of the molecular interaction between the players in the ExsA-ExsC-ExsD-ExsE signaling cascade that enables the organism to tie T3SS expression to host cell contact [1-3]. Unlike the ubiquitous two-component and phosphorelay systems, this cascade operates not through phosphate transfer but via a partner-switching mechanism that is triggered when ExsE is secreted [3, 4]. The domain architecture of ExsA displays all the classical features of canonical AraC-type transcription factors. The effect of ExsD binding on ExsA-dimerization and DNA interactions also follows the expected pattern [1, 5]. Yet, ExsA does not utilize conserved reaction interfaces for this interaction [5]. Using the solved crystal structure of ExsA-NTD described in Chapter 2 as road map, I systematically analyzed surfaces and the canonical binding pocket to determine which regions are pivotal for ExsD binding. In Chapter 2, I describe my analysis of the conserved binding pocket and the probing of residues near the ExsA dimerization interface, while I expanded the scope of my study to include conserved surface regions of ExsA-NTD and its amino-terminus in Chapter 3. In Chapter 3 my work was combined with studies of a former graduate student, Robert Bernhards, Ph.D., who had previously discovered that the amino-terminal region of ExsD is pivotal for ExsA binding, which provided a critical clue for my work. Following the somewhat fortuitous discovery that modification of the ExsA amino-terminal residues indirectly affects ExsD binding, we hypothesized that it was the backbone of the ExsA beta barrel structure located near the ExsA N-terminus that is involved in the association with ExsD. This made sense because we knew from prior work that the amino-terminus of ExsD does form a beta strand to interact with T3SS
chaperone ExsC [6]. We also knew that ExsD-ExsC and ExsD-ExsA interactions are mutually exclusive [6, 7]. The Y24P and V26P mutations created in ExsA to test our new hypothesis indeed led to dramatic attenuation of the regulation by ExsD thus lending support to our model of a backbone-mediated binding between ExsD and ExsA. However, the results of the subsequently performed BACTH assays hint at the presence of a more substantial interface as both variants were still shown to interact with ExsD.

Previous work has shown that ExsD binding interferes with ExsA dimerization and DNA binding [1]. As those two processes are strongly correlated it may well be that ExsD simply interferes with dimerization to prevent ExsA from properly engaging with the promoter. However, it is still somewhat unclear how ExsD accomplishes this. ExsD does not appear to compete directly with dimerization. ExsD may simply use its bulk to sterically interfere with homodimerization. However, there is an alternative possibility. The regulatory mechanisms of AraC, and ToxT proteins have been studied the most extensively [8, 9]. In AraC, binding of arabinose causes a repositioning of the aforementioned flexible N-terminus, which has been proposed to interact in an intermolecular fashion with the DNA-binding domain of the other molecule. The net result is a conformationally locked AraC dimer incapable of activating transcription. Only binding of arabinose alters the position of the N-terminus to unlock the structure [10-12]. In ToxT, fatty acid binding actually forces a conformational change within each molecule of the ToxT dimer to also lock the NTD and CTD in a rigid and inactive conformation. Here not the amino-terminus but the flexible inter-domain linker plays an important in the regulatory process [8, 9, 12]. It is noteworthy that in both cases direct NTD-CTD interaction are pivotal for the regulatory mechanism. To date we have no experimental evidence to suggest that the NTD and CTD of ExsA form important specific contacts. Clearly,
unlike was the case for AraC, the amino-terminal seven residues of ExsA are not pivotal. The role of the linker region has yet to be examined more closely. The ExsA dimer has been previously shown to bind to its promoters in a head-to-tail arrangement [13], suggesting that the linker should be flexible to permit stable dimerization mediated by the NTDs, and at the same time allow binding of the CTDs to the promoters. Therefore, it is conceivable that ExsD binding prevents ExsA dimerization by locking the two domains of ExsA into a rigid conformation. Using the structure of full-length ToxT as template and our structure of ExsA-NTD we had previously generated a model of an ExsA dimer. Interestingly, if we now consider where the amino-terminus of ExsD binds to ExsA the beta strand inserts directly at the interface of the two domains (Fig 4.1). Therefore, it is entirely possible that, in line with what has been observed for ToxT and AraC as well, ExsD acts by targeting the contacts between the NTD and CTD of ExsA thus forming a locked conformation of ExsA (Fig 4.2).

At this point it is unclear if our findings may be extended to other systems. Despite their relative abundance there is not much known about how other AraC-type transcription factors are regulated. The little we do know appears to point toward a diversity of regulatory mechanisms. Even within the small subgroup of AraC proteins that ExsA belongs to it is completely unclear whether the regulation of ExsA is unique or if a unifying theme will eventually emerge. Feedback loops that link activation of a central AraC-type transcription factor to host cell contact have been identified in several other important pathogens such as Yersinia pestis [14, 15], Salmonella enterica [16], and Shigella flexneri [17], and Vibrio parahaemolyticus [18]. However, only Vibrio parahaemolyticus uses a homologous four-protein cascade to control T3SS gene expression [19]. Other pathogens use somewhat different mechanisms to achieve the coupling of gene expression and active secretion, there are striking parallels: In Yersinia pestis,
Shigella flexneri, Salmonella enterica T3SS chaperones LcrH, IpgC, and SicA interact with the transcription factor following secretion of their cognate effectors [20-22]. Here chaperone binding actually activates transcription presumably by aiding dimerization or alternatively by stabilizing a particular mode of dimerization, akin to the effect arabinose has on the geometry of the AraC dimerization interface. Alternatively, chaperone binding could once again act by unlocking rigid contacts between regulatory and DNA binding domains to facilitate T3SS gene expression. Clearly, more work will be needed to determine if there is a unifying theme by which these proteins are regulated, but current projections would point toward a great diversity of mechanism custom-tailored for the specific regulatory function of the protein.
REFERENCES


17. Parsot C, Ageron E, Penno C, Mavris M, Jamoussi K, d’Hauteville H, Sansonetti P, Demers B: A secreted anti-activator, OspD1, and its chaperone, Spa15, are involved in the control of
transcription by the type III secretion apparatus activity in *Shigella flexneri*. *Molecular Microbiology* 2005, **56**(6):1627-1635.


Fig. 4.1 Hypothetical model of a possible ExsA-ExsD interface in full-length ExsA Shown in teal is the ExsA-NTD structure, shown in marine blue is the model for the ExsA-CTD generated from MarA-DNA structure. Shown in magenta is the possible position of the amino-terminus of ExsD right at the interface between the NTD and CTD of ExsA and far away from the ExsA-dimer interface. Shown in green is another molecule of ExsA.
Fig. 4.2 Hypothetical model for how ExsD “locks down” ExsA to inhibit transcription. A. ExsA-NTD dimerization and DNA-binding conformation of ExsA-dimer inducing transcription activation by ExsA. B. Upon binding of ExsD at the ExsA-CTD and NTD interface the conformation changes so that prevents dimerization and DNA-binding of ExsA-monomers. Therefore, there is no induction of transcription activation.
**APPENDIX A**

**SUPPLEMENTAL DATA: CHAPTER TWO**

---

**A1.** (A) SDS-PAGE analysis of the purification of ExsA-NTD after digestion of ExsA with thermolysin via gel filtration chromatography. (B) Titration of ExsA N-terminal domain in ExsA dependent transcription. The in vitro transcription assay was performed as described under Materials and Methods. In addition to the usually added 64 nM of ExsA WT samples contained the indicated concentration of the ExsA-NTD protein. (C) Crystals and sample X-ray diffraction image of ExsA-NTD.
A2. Schematic representation of the non-bonding contacts in the A/A’ and B/B” dimers of ExsA-NTD.

The width of the dashed lines is proportional to the number of contacts between the residues. This schematic was generated using PDBsum [88]. The RasMol amino acid color scheme colors amino acids according to traditional amino acid properties. The outer parts of a protein that are polar are visible in bright colors and non-polar residues darker. This color scheme is similar to the Shapely scheme that is available under the RasMol website.
A3. SDS-PAGE slices of the purified protein samples used in the various experiments
### APPENDIX B

Supplemental data: Chapter Three

**Table B1. List of primers used to construct ExsD<sub>1</sub>-38-MBP variant**

<table>
<thead>
<tr>
<th>Primers Used</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>6HisTEVExsD</td>
<td>CCTGCCAGCGAACCAGCTTTTGCAGGAGTACTGCTTTATCGTTCTTCCCTGCTCCA TTCCCCCTTGGAGAGACAGTCTGAGGTGATGAGTGGTCTG</td>
</tr>
<tr>
<td>ExsD11-38MBP</td>
<td>GCG GTG TTC GCT GCG AGG CGG GTA TCC GTG GTG GGC TCG GAC GCC CGC TCG CGG GTG CGG GTG CGG GTG TAC GCA TCG GCA GAA GAA GGT AAA CTG GTA</td>
</tr>
<tr>
<td>5’ extension</td>
<td>CATATGGGGCAGCCATCAACCATCAACCATCA</td>
</tr>
<tr>
<td>3’ extension</td>
<td>ATCGCCGTTAATCCAGATTACCAGTCTTACC</td>
</tr>
</tbody>
</table>

**Table B2. List of primers used to construct variants of ExsA**

<table>
<thead>
<tr>
<th>Primers Used</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>E37A-fwd</td>
<td>GTATATGTGTTCTGCTGCGCGGCGCAACTGACCGTC</td>
</tr>
<tr>
<td>E37A-rev</td>
<td>GACGGGTCAATGTCGGCCCGCGAGCAGAATCATATAAC</td>
</tr>
<tr>
<td>T48A-fwd</td>
<td>GTCCAGGACATCGATTCGGCTTTTTGCCTTGCCGCTTG</td>
</tr>
<tr>
<td>T48A-rev</td>
<td>CAGGCGCCAGCCAATGACTGCTGCTCCTGGCAG</td>
</tr>
<tr>
<td>Q90A-fwd</td>
<td>GTTTTCTAAGGCTTTTCGCGACGCTTCCGGCCTGGTGTT</td>
</tr>
<tr>
<td>Q90A-rev</td>
<td>CAACAGCGCGCGGGAAGCGTGGCGAAGCCTTGGTAGAAAC</td>
</tr>
<tr>
<td>R91A-fwd</td>
<td>CTACAAGGGCTTCGTCAGGATTCGGCCTGCTG</td>
</tr>
<tr>
<td>R91A-rev</td>
<td>CACTCAACAGCGCGCGAATGCTTGAGCAGAAGCTTGGTAG</td>
</tr>
<tr>
<td>L95A-fwd</td>
<td>CGCGGCGCTTCGGCGCGGGAAGTGTGAG</td>
</tr>
<tr>
<td>L95A-rev</td>
<td>CTGCTATTGAGTCAGTCAAGGCTG</td>
</tr>
<tr>
<td>E98A-fwd</td>
<td>GAAAGACGCTCAGCTGACATCAACACAGCG</td>
</tr>
<tr>
<td>E98A-rev</td>
<td>CAATCAGGTCAGTCAAGGCTG</td>
</tr>
<tr>
<td>R101A-fwd</td>
<td>GTGGAGTGGACTGAGCGCGGCGAGCGCAGCCGCTG</td>
</tr>
<tr>
<td>R101A-rev</td>
<td>GCACGGGCGCTTCGTCAGGATTCGGCCTGCTG</td>
</tr>
<tr>
<td>L129A-fwd</td>
<td>GGCGGTTGAAGGAATTGGCTGTGATGGCAGATCCG</td>
</tr>
<tr>
<td>L129A-rev</td>
<td>CAGGATGCTGCTGGCAACGCGCAATTCCTTCCAACCC</td>
</tr>
<tr>
<td>E144A-fwd</td>
<td>CTGCTTTGAGAGTCAGGCGTTGCTGAGTGTGTC</td>
</tr>
<tr>
<td>E144A-rev</td>
<td>GAAAGACGCTCAGCTGACATCAACACAGCG</td>
</tr>
<tr>
<td>L140A-fwd</td>
<td>CATCCGCCGTAGCTCGCTCCTCGGCAAGAGATCGAGGAGTTGAGT</td>
</tr>
<tr>
<td>L140A-rev</td>
<td>CATCAGGCAACCCTCCGATCTCCGATCTGGCAGGCGAGCAGATCAGAGG</td>
</tr>
<tr>
<td>F151A-fwd</td>
<td>GAGACATCGCTTCTTCTTGCAGGAGTGGTGTTGAGGAGGCGGAAGGAGCATAGAAG</td>
</tr>
<tr>
<td>F151A-rev</td>
<td>CATCAGGCAACCCTCCGATCTCCGATCTGGCAGGCGAGCAGATCAGAGG</td>
</tr>
<tr>
<td>TevExsA2-7GlyFwd</td>
<td>GAGAACCCTGTAATTCCAGGCTGGGCTTCTTGGCTGAAAGTACAGGAGT</td>
</tr>
<tr>
<td>TevExsA2-7GlyRev</td>
<td>AGAGCTTTATCTGCTTTTCCGCTCCCTGACCACCCACCGCTGAAAGTACAGGAGT</td>
</tr>
<tr>
<td>ExsA N27F-rev</td>
<td>CATATAGCGCCTTCTTCTTGTAAATCCCTGTATTAGGAAAGT</td>
</tr>
<tr>
<td>ExsAT48R-fwd</td>
<td>CAGGACATCGCTTCTTCTTCTTGGCTGAAAGT</td>
</tr>
</tbody>
</table>

105
| ExsA T48R-rev | CAGCGGCAAGGCAAAACCTGGAAATCGATGTCCCTG |
| ExsA L95R-fwd | CGCTTCGGCGCGCGGTAGTTGAGTGAAGTC |
| ExsA L95R-rev | GACTTCACCTCAACCCGCGCCGCAAGCG |
| ExsA E98F-fwd | CGCGCGCCTGTTGAGTTTCGTCGAGCGTTGCAGCG |
| ExsA E98F-rev | CGTCGCAACGCTCGACGAAACTCAACAGCGCGCCG |
| ExsA L129R-fwd | GGGTTGAAGAGATTTGGCGTGTGCATGAGCATCCG |
| ExsA L129R-rev | CGGATGCTCATGCAACAGCAATTCCTTCAACCC |
| ExsA L140R-fwd | GATGCTCGCCTGCCGGAAGATCGAGGAGT |
| ExsA L140R-rev | CGTCGCAACGCTCGACGAAACTCAACAGCGCGCCG |
| ExsA E144F-fwd | GCCTGCCTGAAAGATCGAGTTCTTGCTGATGCTCTTCCG |
| ExsA E144F-rev | CGCGAAGAGCATCAGCAAGAACTCGATCTTCAGGCAGGC |
| ExsA F151R-fwd | TGATGCTCTTCCGCGCGCAGTCCGCGAGGGGC |
| ExsA F151R-rev | GCCCCTGCAGGGACTGCGCGCGGAAGAGCATCA |
| Δ7ExsA-fwd | GAGAACCTGTACTTCCAGGGCCGAAAGCAGATAACGTCT |
| Δ7ExsA-rev | AGACGTATATCTGCTTTGCGCCCTGGAAGATCACAGGTTCCT |
| Δ5ExsA-fwd | AACCTGTAATCCTGCTCTTGCAG |
| Δ5ExsA-rev | GCCAAGAGACTGGAAGTACAGGT |
| Δ6-10ExsA-fwd | CAAGGAGCCAAAATAACGTCT |
| Δ6-10ExsA-rev | AGACGTATATTTGGCTCTT |
| Q2A ExsA-fwd | GGAGAACCTGTACTTCCAGGGTGTCAGAGGAGCCAATAC |
| Q2A ExsA-rev | GATTGTCACCTGCTGGAAGATCGAGGTCTCC |
| K5AExsA-fwd | CTTCGGGCAAGATCGCGCTCGGTACCTTGGCCG |
| K5AExsA-rev | CCAGGGTCAAGGAGCGCATCTCTTGCGCCGAAAG |
| R9AExsA-fwd | CAAGACGTATATCTGCTCTGCGCCAAAGAGATTTGGC |
| R9AExsA-rev | GGAGCCAATACCTGCTCTTGCCGCAAGCAGATAACGTCTT |
| K10AExsA-fwd | TGACAAAGACGTATATCGGCTCGGCCAAAGAGATTTGGCT |
| K10AExsA-rev | GAGCAAGATACCTGCTCTTGCGCCAGGATACAGGT |
| Q11AExsA-fwd | CCAAAGAAGACGTATACCTGCTCTTGCGGCAAGGAGATTT |
| Q11AExsA-rev | CAAATCTCTTGCGCCAAAGAGCATAACGTCTT |
| ExsAd11F | GAGAACCTGTACTTCCAGGGCATAACGTCT |
| ExsAd11R | AGACGTATATCGGCTTGAAGATACAGGT |
| ExsAd20F | GAGAACCTGTACTTCCAGGGCACTTCTC |
| ExsAd20R | TTCGAAAGTGCCCTGGAAGTACAGGT |
| ExsAG8Pf | CGTTATCTGCTTCTTCCGCGGGAAGATTTGGGCTCTT |
| ExsAG8Pr | GGCCAAGGAGCCAATACTCTTCCCGCCAAAGCAGATTAACGT |

106
| ExsAI12Af | ATCTCCTTGCGCGGGAAGACAGCCAACGTCTTGCATTGGAACAT | |
| ExsAI12Ar | TGTTCCAATGACAAGACGTGCCTTGCTTTTCGGCGCCAAGAGAAT | |
| ExsAS13Af | GGGCAGGAAGACGATAACGCCTTGTCATTGGAACATTC | |
| ExsAS13Ar | GAATGTTCCAATGACAAGGCGTTATCTGCTTTTCGGCAAG | |
| ExsAC15Af | TGTTCCAATGACAAGGCGTTATCTGCTTTTCGGCCAAGAGA AT | |
| ExsAC15Ar | CCAATGACAAGACGTTATCTGCTTTTCGGCAAGAGAT | |
| ExsAH16Af | TGGCCGCGGGAAGACGATAACGGCTTGTCATTGGAACATTC | |
| ExsAH16Ar | GGAATGTTCCAATGACAAGGCGTTATCTGCTTTTCGGCAAGAGA AT | |
| ExsAW17Af | GGAATGTTCCAATGACGTTATCTGCTTTTCGGCAAGAGAT | |
| ExsAW17Ar | GGAATGTTCCAATGACAAGGCGTTATCTGCTTTTCGGCAAGAGA AT | |
| ExsA2_7alaf | GAGAACCTTGACTTCCAGGGTGAGCACGCGAGCTGTGGCCGAAAAG | |
| ExsA2_7alar | AGACGTATCTTGCCTGCGCCGACGAGCTGCGCAGTATCTGCTTTTCGGCAAGAGA AT | |
| ExsA13_17alaf | TCTTCTGGCGCGGAAAGACGATAGCGGCTGCTGCTGCGAACATTC | |
| ExsA13_17alar | GTATTGCCGCAATTTTCGCAAGCAGCAGCAGCCGCTATCTGCTTTTCGGCAAGAGA AT | |
| ExsA23_27alaf | CATTGGAACATCTCACTTTCCGCGCAAGCAGCAGCAGCCGCTATCTGCTTTTCGGCAAGAGA AT | |
| ExsA23_27alar | AACATATACGCCCTCTTCTTTGGCTGGCGGCTGGGCTGCGAACATTCCAACTTTCGAAGAGGCGTATATGTT | |
| ExsAT13P_f | CTTGGCGGGAAGACGATAACCGTCTTGCATTGGAACATTC | |
| ExsAT13P_r | GAATGTTCCAATGACAAGGCGTTATCTGCTTTTCGGCAAGAGA AT | |
| ExsAR25P_f | CCAATGACAAGACGTTATCTGCTTGCAAGGCAAAGAGGCGTA | |
| ExsAR25P_r | GTATTGCCGCAATTTTCGCAAGCAGCAGCAGCAGCCGCTATCTGCTTTTCGGCAAGAGA AT | |
| S14Pfwd | GGCGCAGCGACAAGCAGCCGCTTGTGGCGCCAAGGAAAGGCGGAAGGCGTATATGTT | |
| S14Prev | GAATGTTCCAATGGAACGAGCCTTGATGCTTGTGGCCTGGCCAAGGAAAGGCGGAAGGCGTATATGTT | |
| H16Pfwd | AAAGCAGATAACGTCTTGCTTTTGCAAACATTCCAACCTTTTCG | |
| H16Prev | CGAGAACGTGTTGGCAGCAGCAGAAGGTTGAATGTGCTTTTGCCAATGAGGCTGAGGCAAGGCGGAAGGCGTATATGTT | |
| Y24Pfwd | GTATTGCCGCAATTTTCGCGCAACCGGCGGCTGCGCGGCAAGGAAGGCG | |
| Y24Prev | CTCTTTCTGACCTTCCCTGTTGACCTATCCGGCTGGAAGGCG | |
| V26Pfwd | TCCAACTTCGAAATACCGGCGAACAAGAAGGAGGCGGAAGGCGTATATGTT | |
| V26Prev | TCCAACTTCGAAATACCGGCGAACAAGAAGGAGGCGGAAGGCGTATATGTT | |
| Y24Afwd | GTTTCCGCAAGAACGTGTTGGCAGCAGGCTGTTGGCGGAAGGCG | |
| Y24Arev | CTCTTTCTGACCTTCCCTGTTGACCTATCCGGCTGGAAGGCG | |
Fig. B1. Commassie stained gels of variants of ExsA and ExsD. A. Amino-terminal regions modified ExsA. Panel B. Alanine substituted variants of ExsA. C. Disruptive substitution variants of ExsA. D. Proline substituted variants of ExsA. E. Different variants of ExsD.
Fig. B2. Effect of ExsΔC-C on ExsA-dependent transcription at 30 and 37 °C. Autoradiograms and graphical representations for the *in vitro* transcription of an 82 base nucleotide transcript from an ExsA-dependent *P*exsD promoter template with 64 nM ExsA and a titration of ExsΔC-C at 30 and 37 °C. Wild type ExsD at 50 μM served as a control. Each experiment was performed in duplicate. The IC50 values were 4.3 ± 0.4 μM and 0.97 ± 0.4 μM at 30 and 37 °C, respectively.
Fig. B3. Effect of ExsΔ20 on ExsA dependent transcription at 37 °C. Autoradiogram for the in vitro transcription of an 82 base nucleotide transcript from an ExsA-dependent PexsD promoter template with 64 nM ExsA and a titration of ExsΔ20 at 37 °C. Wild type ExsD at 50 μM served as a control. This experiment was performed in triplicate.
Fig. B4. ExsDΔ20 has no effect on ExsA-independent transcription. Autoradiogram and graphical representation for the *in vitro* transcription of a 108 nucleotide transcript from an ExsA-independent RNA-1 promoter template with and without 50 μM ExsDΔ20 at 37 °C. This experiment was performed in duplicate.
Fig. B5. Effect of ExsD1-46 on ExsA dependent transcription at 37 °C. Autoradiogram for the *in vitro* transcription of an 82 base nucleotide transcript from an ExsA-dependent PexsD promoter template with 64 nM ExsA and a titration of ExsD1-46 at 37 °C. This experiment was performed in triplicate.
**Fig. B6. ExsD1-46 has no effect on ExsA-independent transcription.** Autoradiogram and graphical representation for the in vitro transcription of a 108 nucleotide transcript from an ExsA-independent RNA-1 promoter template with and without 50 μM ExsD1-46 at 37 °C. This experiment was performed in triplicate.
Fig. B7. Sequence identifier and alignment of ExsD from different bacteria

<table>
<thead>
<tr>
<th>Sequence identifier</th>
<th>Protein name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q9I321_PSEAE ExsD</td>
<td>Pseudomonas aeruginosa</td>
</tr>
<tr>
<td>Q699Q3_AERHY ExsD</td>
<td>Aeromonas hydrophila</td>
</tr>
<tr>
<td>A5Y876_9GAMM ExsD</td>
<td>Aeromonas veronii</td>
</tr>
<tr>
<td>A4SUG1_AERS4 ExsD</td>
<td>Aeromonas salmonicana</td>
</tr>
<tr>
<td>Q7N0V3_PHOLL Uncharacterized protein</td>
<td>Photorhabdus luminescens subsp. laumondii</td>
</tr>
<tr>
<td>Q87P20_VIBPA Uncharacterized protein</td>
<td>Vibrio parahaemolyticus serotype O3:K6</td>
</tr>
</tbody>
</table>

CLUSTAL O(1.2.4) multiple sequence alignment
APPENDIX C

ADDITIONAL VARIANTS OF ExsA

Initially, in order to directly probe the strands β1 and β2, we did not duly consider the directions of the amide-group when designing suitable variants and the created constructs: ExsAT13P and ExsAR25P. Thr-13 lies at the beginning of β1 and Arg-25 lies in the middle of the β2 strand of ExsA (Fig. C1). However, based on our crystal structure of ExsA-NTD their amide groups are engaged in intramolecular strand-strand interaction within the beta barrel of ExsA. Unsurprisingly, ExsAR25P proved unstable and was eliminated from the study. Along with these, another variant similar to the glycine substituted variant of ExsA, five alanines were used to replace the first seven amino acids of ExsA. This variant was called 1-7ala ExsA. The motivation behind constructing this variant was to observe whether addition of five alanines would have similar effect as that of adding five glycines.

Fig. C1. Are the main-chain amide groups of β-1 or β-2 of ExsA important for ExsD binding? A. Shown is the β-barrel of ExsA-NTD as cartoon-model. Amino acids of β1, β2, β7, and β8 are shown as sticks. Intra-molecular hydrogen bonding contacts are shown in yellow dashed lines. Side-chains of amino acids that were of substituted with proline (described in Chapter 3) are labeled. B. Shows the same model as shown in A. with selected residues now altered to proline, highlighting the resulting removal of possible H-bonding contacts by the main-chain amide group with any incoming β-sheet.

Both the new variants of ExsA, 1-7ala and T13P was able to induce ExsA-dependent transcription in similar manner to that of wtExsA, however, upon addition of ExsD to the
experiment reduced the amount of transcript produced. This indicates that, the selected residues are not important for regulation by ExsD.

<table>
<thead>
<tr>
<th>ExsA (128nM)</th>
<th>WT</th>
<th>1-7ala</th>
<th>T13P</th>
</tr>
</thead>
<tbody>
<tr>
<td>ExsD (10µM)</td>
<td>0</td>
<td>10</td>
<td>0</td>
</tr>
</tbody>
</table>

![Graph showing % change in transcript production in presence of ExsD](image)

**Fig. C2.** Additional variants of ExsA had no effect by ExsD regulation, *in vitro*. Results of *in vitro* transcription assay, showing addition of five alanines did not change the induction of transcription activation nor was there any change in the regulation by ExsD, *in vitro*. Similar results were seen for a proline substitution at the T13 residue, indicated by reduction in transcripts level similar to that of wtExsA by ExsD.