Elucidation of DNA-Binding and Dimerization Activities of the Master Virulence Transcription Factor, VirF, From *Shigella flexneri* to Aid in Virulence-Targeted Inhibitor Design for Treatment of Shigellosis

by

Garrett T. Dow

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Medicinal Chemistry)
in the University of Michigan
2022

Doctoral Committee:

Professor George A. Garcia, Chair
Professor Amanda L. Garner
Professor Jonathan Z. Sexton
Professor David H. Sherman

Garrett T. Dow

gtdow@med.umich.edu

ORCID iD: 0000-0002-4683-9560

Dedication

To my wonderful wife and role model, Celina. Thank you for sticking by my side through everything and never letting me quit, even when I wanted to.

Acknowledgements

All this work would not have been done without the incredible support from my advisor, Dr. George Garcia. From the very beginning, he supported my research endeavors which allowed me to follow through with my proposed experiments while also providing valuable input into ways of improving my experimental designs. I learned a great deal from Dr. Garcia and will continue to use his scientific mantras throughout my career (mostly because they have been engrained in my head). Beyond the academic and research support, Dr. Garcia was always there for me through all my personal matters and knew how to tell some great dad jokes which often brought a smile to my face. Thank you for being such a great mentor. I learned so much about becoming a well-rounded scientist as well as how to manage a healthy work-life balance. I would also like to thank the members of my committee for all the experimental suggestions, advice, and support they provided throughout my graduate career: Dr. David Sherman, Dr. Amanda Garner, and Dr. Jonny Sexton.

To my lab members, past and present, thank you for providing me with a welcoming environment to discuss science, bounce ideas off each other, and to vent about graduate school. Specifically, I would like to thank Nicholas Ragazzone, who also had the "pleasure" of working on the *Shigella* project with me in the Garcia lab. This project presented many difficulties, and it was helpful having Nick around for troubleshooting and discussing the NFL and music when we needed a break from our research. I would also like to thank former lab members, Drs. Maxwell

Stefan and Anthony Emanuele, for providing me with the analytic tools to succeed in the Garcia lab and for providing a great foundation for me to build upon with my own research, respectively.

Lastly, I would like to thank my friends and family for their constant support throughout my time in graduate school. Firstly, I'd like to thank my parents, Angie and Ted, for their unwavering support and sending me care packages with lots of food and giftcards to restaurants. Thank you for letting a lifelong "germaphobe" pursue his passions of studying bacteria and antibacterial drug discovery. Most of all, I would like to thank my amazing wife, Celina. While going through medical school herself, Celina managed to provide me with all the tough love and support I needed to succeed. I could not have done any of this without you.

The research presented herein was supported by the University of Michigan College of Pharmacy and the Pharmaceutical Sciences Training Program (NIH GM007767).

Table of Contents

List of Tables

List of Figures

List of Equations

Abstract

The uncontrollable rise in resistance to current antimicrobial agents is a worldwide concern. Beyond traditional approaches to develop novel antibiotics, targeting bacterial virulence is considered a promising approach. It is thought that by targeting virulence, there will be a weaker selective pressure on the bacteria to develop resistance while also displaying little effect on the normal microbiota in the environment and within a host. Specifically, antimicrobial resistance in the bacterial genus, *Shigella,* is a significant problem requiring novel therapies to lessen its rise. These organisms, particularly *Shigella flexneri*, are responsible for the diarrheal disease Shigellosis which leads to approximately 200,000 deaths globally every year. A promising virulence target is the transcription factor, VirF. VirF is an AraC-family protein that contains an N-terminal dimerization domain and a C-terminal DNA-binding domain (DBD) harboring two helix-turn-helix (HTH) motifs. This protein activates transcription of two major virulence genes, *virB* and *icsA*, which allows the pathogen to invade and spread within colonic epithelial cells. Previous drug discovery campaigns identified VirF inhibitors which exhibited anti-virulence effects *in vivo* but failed to improve upon their efficacies for testing in animal models.

This thesis presents results focused on elucidation of both the VirF DNA-binding and dimerization domains to aid future virulence-targeted drug discovery. First, we developed homology models for the VirF DBD using structures of VirF-homologs from *E. coli*, GadX and MarA. We conducted alanine-scanning mutagenesis on seven residues within VirF which were based on an alignment with MarA residues that contributed base-specific interactions with its cognate promoter, *marRAB*. We elucidated VirF DNA-binding activity for its three cognate DNA

promoters (*pvirB*, *picsA*, and *prnaG*) using wildtype and seven DBD mutants in *in vitro* DNA-binding assays. Upon testing with *pvirB*, mutations to the N-terminal HTH exhibited significant reductions in DNA-binding, while the effects of mutations in the C-terminal HTH varied. When expanded to *picsA* and *prnaG*, WT MalE-VirF bound to these promoters through multiple binding shifts with the DBD mutants displaying similar binding trends compared to *pvirB*. Specifically, the VirF•*picsA* interaction was more sensitive than the other promoters where all mutations, except I189A, caused reductions in DNA-binding activity. To study VirF dimerization, we employed a LexA monohybrid β-galactosidase reporter assay which confirmed WT VirF dimerizes and identified alanine-mutations to Y132, L137, and L147 significantly affected dimerization. These mutations also significantly affected protein stability, but we successfully purified Y132A MalE-VirF which was capable of binding to the *virB* and *rnaG* promoters, albeit with weaker affinity or reduced protein:DNA ratio, respectively. I-TASSER generated full-length VirF structures which predicted the location of these residues and the potential contributions of Y132 for dimerization activity as well as L137 and L147 to protein stability.

To discover VirF-targeted inhibitors, compound screening was performed. Our previously optimized VirF-driven β-galactosidase reporter assay was used to screen a set of *S. flexneri* virulence inhibitors identified at GlaxoSmithKline. However, none of the hit compounds exhibited activity against VirF, so it is likely they are interacting with virulence targets outside of VirF. In addition, the LexA reporter assay was used to screen lead compounds, which previously exhibited inhibition of VirF transcriptional activation, against VirF dimerization, but none displayed dimerization inhibition. Despite these unsuccessful screens, the in-depth functional analysis of the VirF DNA-binding and dimerization domains will be crucial for the success of future VirF-targeted inhibitor discovery and design.

Chapter 1 Introduction

1.1 Antibiotic Resistant Infection and Novel Treatments to Combat this Pandemic

It is no surprise that infection continues to be a significant threat to humankind. While medicine has advanced significantly in treating infection and reducing corresponding mortality rates, more resistant bacterial, viral, and other microbial infections continue to spawn and cause widespread damage. As we experienced in 2020 with the COVID-19 pandemic, this problem will never go away unless epidemiologists and pharmaceutical scientists accurately predict and develop novel treatments to prevent or lessen the extent of the next pandemic. Unfortunately, there is an ongoing "silent" pandemic due to the rise of infection with antibiotic-resistant bacteria. Resistance naturally occurs in the environment but the overuse and misuse of these drugs by humans, particularly in agriculture, has only accelerated the rise and spread of these resistant organisms. In 2014, it was estimated that approximately 10 million people will die every year due to antimicrobial resistant infection (AMR) (1). Unfortunately, current estimates indicate that there were ~5 million deaths due to AMR in 2019 alone, of which 1.27 million were attributed to bacterial infection(2). These are substantial indications for a far more future mortality crisis than previously anticipated, let alone the financial strain this will incur globally. Based on estimates garnered from 2014, Thorpe et al. estimated that there was an additional annual cost of ~$1,400 to treat an AMR infection per patient resulting in a $2.2 billion total cost in the United States per year(3). The Centers for Disease Control (CDC) analyzed an AMR cost analysis study by Roberts et al. to which they extrapolated that the United States alone could incur upwards of $55 billion in costs annually due to AMR infection(4, 5). These are staggering projections and provide an

essential, worldwide need to discover novel therapeutics for treatment of AMR infection without exacerbating this ongoing crisis.

There are many approaches that scientists are exploring for the treatment of AMR bacterial infection. First and foremost is the discovery and development of novel antibiotics or improving upon existing ones. This can be accomplished through natural products drug discovery, high-throughput screens, small molecule syntheses, and target-based approaches, etc. There is a vast list of ongoing research studies in academic, industrial, and governmental labs aimed at accomplishing this goal. A roadmap to this drug discovery pipeline has been outlined in Miethke et al., importantly, the authors also highlight significant threats to this process including lack of funding and high societal costs(6). The most recent antibiotic to be FDA approved (Cefiderocol) was in 2019(7). This is due in part to the lack of research and development of new antibiotics in the pharmaceutical sector. There is a lack of "financial incentive" for antibiotic drug development compared to more profitable drugs used for the treatment of cancer or metabolic diseases, for example(8).

Another route to treat AMR bacterial infection is with vaccines. This route is garnering significant interest given the increased prevalence and success of the novel mRNA-based COVID-19 vaccines. Vaccines have been proven to be incredibly effective for the treatment of bacterial infection (e.g., pertussis and meningitis vaccines). Like antibiotic drug discovery, many studies are focused on developing vaccines for the treatment of numerous bacterial infections. However, bacteria have many potential targets and vaccines require proper formulations and adjuvants to elicit a strong and effective immune response(9, 10). Additionally, vaccines exhibit varied efficacies and durations of protection in malnourished, immunodeficient subjects found in

developing areas(11, 12). Given that infectious diseases are highly prevalent in less developed nations, these factors likely reduce the efficacy of vaccines in the short term.

Intriguingly, a promising route to treat AMR infection is through targeting bacterial virulence. Given virulence pathways are often not essential for bacterial cell viability, it is thought that by targeting these pathways, there will be no selective pressure on the bacteria to develop resistance to the drug(13). In addition, it is expected that virulence-targeting agents should have no effect on the host microbiome as well as non-pathogenic organisms in the environment, which often serve as reservoirs for developing and transferring antimicrobial resistance to pathogenic organisms(13, 14). However, this area of research is still speculative as there are only five such drugs approved by the FDA. All five drugs are antibodies that target bacterial toxins: BabyBIG (15) and BAT (NCT00360737) for treatment of *Clostridium botulinum*, Raxibacumab (16) and Obiltoxaximab (17, 18) for treatment of *Bacillus anthracis*, and Bezlotuxumab (19–21) for treatment of *Clostridium difficile*. There are still no small molecule FDA-approved, anti-virulence therapeutics; however, a few are presently in pre-clinical trials(22, 23).

1.2 Shigella flexneri

A pathogen that the CDC has labeled a "Serious Threat" is the microorganism, *Shigella*(24). Shigellosis, the main form of bacillary dysentery caused by the infection with *Shigella* spp., leads to ~450,000 cases, of which, 77,000 are drug-resistant in the U.S. per year(24). Globally, infection with *Shigella* spp. results in nearly 270 million cases and over 200,000 deaths per year primarily affecting children under the age of five(25). *Shigella* infection is prevalent in lower socioeconomic areas where there is lesser access to proper sanitation and adequate medications to treat the infection. This leads to higher mortality rates particularly in Africa, India, and Southeast Asia (**Figure 1.1**).

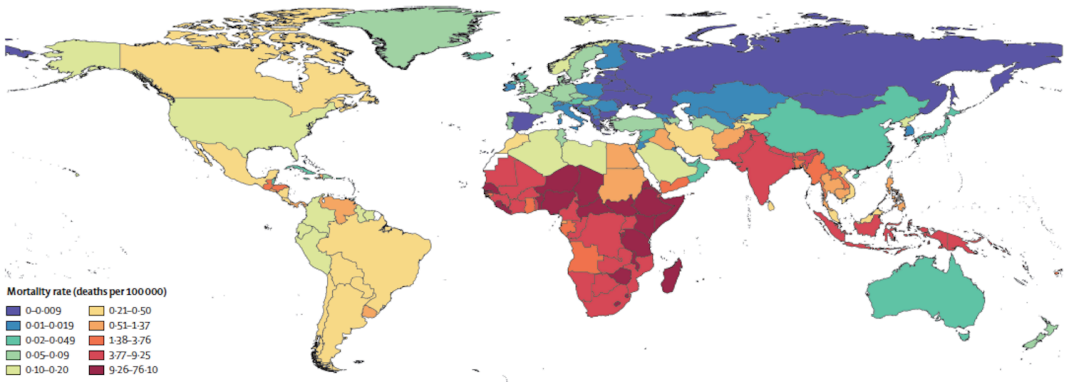


Figure 1.1: Diagram of mortality rates per nation due to *Shigella* infection in 2016. All age groups are included. This figure was adapted from Khalil et al., 2018.

While usually self-limiting, *Shigella* infection can be life-threatening. Treatments often include oral rehydration and prescription antibiotics. Currently, azithromycin or ciprofloxacin are the most effective drugs administered to clear the infection. However, resistance to both antibiotics has reached significantly high levels with 17% of isolated strains exhibiting resistance to either ciprofloxacin or azithromycin in the United States alone(24). Worldwide, resistance is significantly rising for currently prescribed antibiotics, with complete ciprofloxacin resistance noted as a serious concern by the World Health Organization (WHO)(26). While many mutations have increased *Shigella* resistance to antibiotics, mutations to *parE* (topoisomerase IV) and *mphA* (phosphotransferase) have significantly reduced their susceptibility to ciprofloxacin and azithromycin, respectively(27–30). Because of the rise in resistance globally, *Shigella* has been classified by the WHO as a "priority pathogen" in critical need for the development of novel therapeutics to treat this infection(31).

Shigella are gram-negative, facultative intracellular bacteria that specifically infect primates and humans. *Shigella* spp. are closely related to *E. coli*, specifically enteroinvasive *E. coli* (EIEC), and evolved their pathogenic phenotype and lifecycle following acquisition of the large virulence plasmid, pINV (**Figure 1.2**) (32–34). Infection with *Shigella,* known as shigellosis, is an acute intestinal infection which causes the destruction of the colonic epithelium resulting in

symptoms such as bloody or mucus-associated diarrhea, fever, and severe dehydration(35). One of the most globally prevalent and virulent species in the *Shigella* genus, *S. flexneri* (serotype 2a specifically), is an incredibly infectious pathogen requiring as little as 10 cells to establish an infection(36). Once ingested, *S. flexneri* bypass the acidity in the stomach by increasing expression of a homolog to an *E. coli* acid resistance gene, *rpoS*, contributing to its low infectious dose(37). Inside the colon, *S. flexneri* enter the basolateral membrane by one of three mechanisms: paracellular movement between epithelial cells, movement from specialized microfold cells (M-cells) into macrophages (**Figure 1.3 (0)**), or movement between cytokine recruited granulocytes(35, 38, 39). It has also been shown that *Shigella* can infect the lumenal face of epithelial cells through colonic crypts but it is not thought to be the primary route of invasion (**Figure 1.3 (4)**)(40). Normally, macrophages phagocytose the bacteria, but *S. flexneri* are highly adaptive pathogens and can induce macrophage pyroptosis in order to escape into the extracellular, basolateral space (**Figure 1.3 (1)**)(38). Upon invasion of an epithelial cell with its type 3 secretion system (T3SS) (**Figure 1.3 (2)**), it spreads to other epithelial cells through polymerizing host cell actin with the virulence factor, IcsA(41, 42) (**Figure 1.3 (3)**).

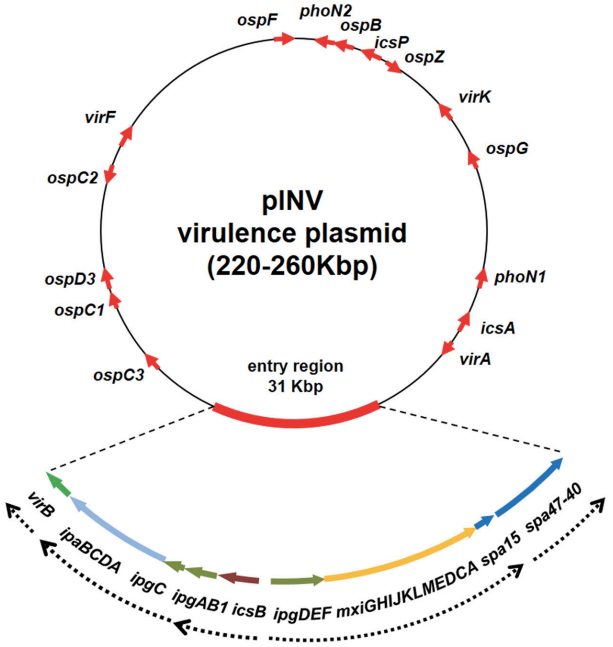


Figure 1.2: Diagram of the *Shigella flexneri* virulence plasmid. Virulence genes discussed in this introduction are *virF, icsA, virB, ipaB, ipaC,* and *ipaD*. This figure was adapted from Pasqua et al., 2017, *Frontiers in Microbiology*.

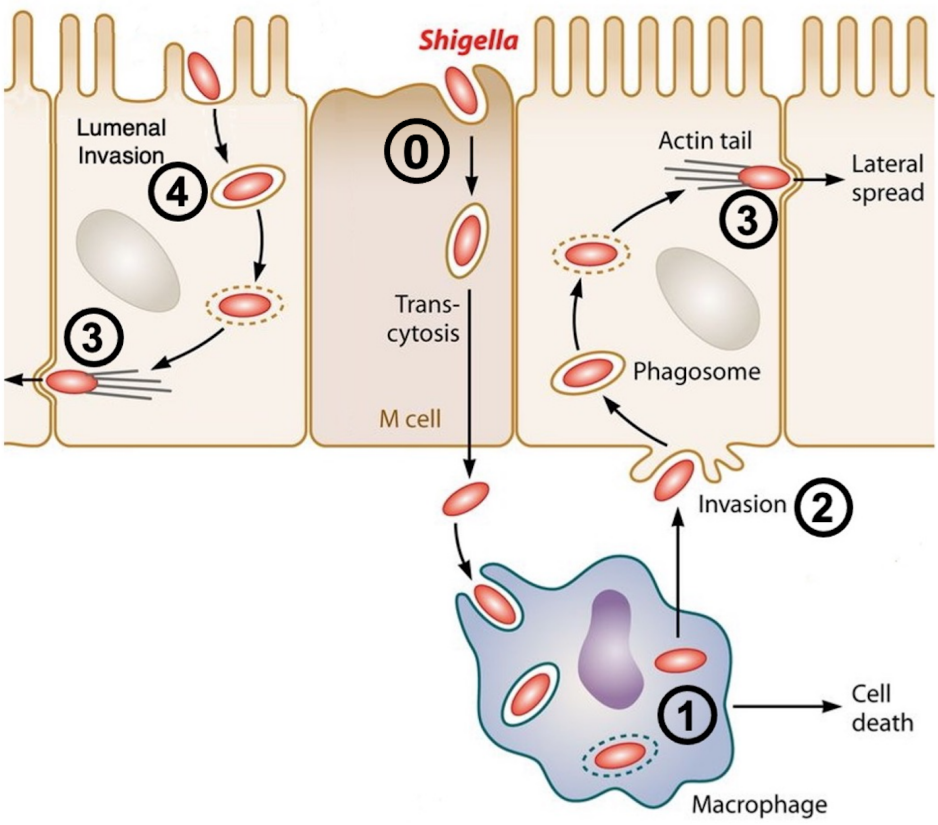


Figure 1.3: Diagram of *Shigella* invasion of colonic epithelial cells. The infection cycle is characterized by (**0**) transcytosis from M-cells to resident macrophages, (**1**) inducing macrophage pyroptosis and escape, (**2**) invasion of the basolateral membrane via T3SS, (**3**) and cell-to-cell spread via host actin polymerization facilitated by IcsA. Invasion may also occur through the lumenal side of the epithelial cell (**4**). The figure was adapted from Croxen et al., 2013, *Clinical Microbiology Reviews* and prepared by George A. Garcia, PhD.

The *Shigella flexneri* T3SS (**Figure 1.3 (1), 1.3 (2), and Figure 1.4**) allows the pathogen to invade the colonic epithelium, induce macrophage pyroptosis, and modulate the host immune response. Among the many proteins that serve as machinery or effector proteins in the T3SS, three Ipa (invasion plasmid antigen) proteins are most essential to the bacteria's virulence: IpaB, IpaC, and IpaD. IpaB is an important virulence factor used to both infect the colonic epithelium and induce macrophage pyroptosis. Located at the tip of the T3SS with IpaC and IpaD, IpaB causes lysis of the phagocytic vacuole allowing for bacterial access to the cytosol where IpaB can induce macrophage pyroptosis (inflamed cell death)(35, 38, 43–45). IpaB has also been shown to produce

pores in host cell membranes *in vitro* and binds directly to caspase-1 inside the macrophage which leads to secretion of IL-1β and IL-18(44, 46, 47). This promotes polymorphonuclear cells to the epithelial layer thus reducing its integrity and allowing for further passage of *S. flexneri* into the basolateral space. The translocon protein, IpaC, plays a crucial role by binding to IpaB at the tip of the T3SS allowing the protein complex to create pores in host cell membranes thus facilitating their invasion and effector secretion(45, 48, 49). IpaD is another important virulence factor at the T3SS tip that has been both implicated in caspase-1 independent macrophage apoptosis, translocon pore formation, and recruitment and secretion of IpaB and IpaC(45, 50–52).

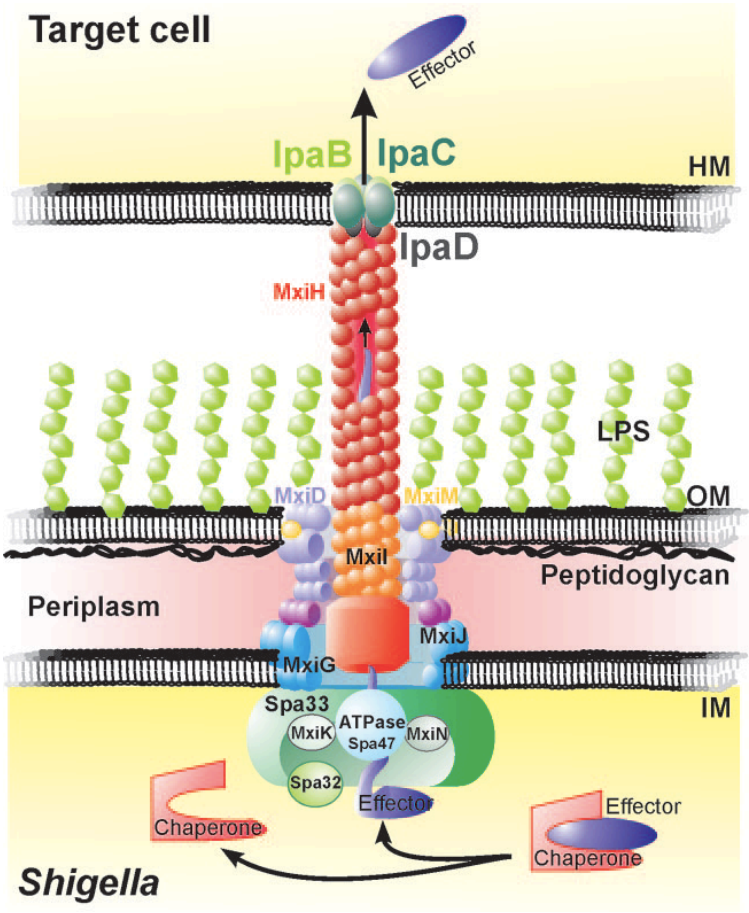


Figure 1.4: Diagram of *Shigella* T3SS used for invasion of colonic epithelial cells. The proteins detailed in the dissertation (IpaB, C, D) are located at the tip of the T3SS). Other machinery proteins (MxiD, MxiH, etc.), chaperones (IpgC), and ATPases (Spa47) are also included. This figure was adapted from Schroeder and Hilbi, *Clinical Microbiology Reviews,* 2008.

Another essential *S. flexneri* virulence factor is IcsA. *S. flexneri* are immotile organisms and require IcsA for cell-to-cell spread of the bacteria within the colonic epithelium (**Figure 1.3 (3)**). IcsA utilizes host factors such as N-WASP and Arp2/3 to facilitate host F-actin polymerization at one pole of the bacterial to propel it into adjacent cells(41, 42, 53–57). With

expression of a mutant form of IcsA, virulent *S. flexneri* did not cause colonic tissue damage to test animals due to their incapability of cell-to-cell spread(58).

1.3 VirF: The Master Transcriptional Activator of Virulence

Shigella flexneri virulence is primarily controlled by the transcription factor, VirF. VirF is an AraC-family protein that contains a highly conserved C-terminal DNA-binding domain (DBD) and an N-terminal dimerization domain (NTD). VirF belongs to a class of AraC-family proteins that respond to physical signals (i.e., temperature, pH, osmolarity) and are thought to dimerize, bind to DNA promoters, and recruit RNA polymerase in order to activate transcription of downstream virulence genes(59). The *virF* gene resides on the *Shigella* virulence plasmid, pINV (**Figure 1.2**), along with the genes involved in the T3SS and cell-to-cell spread, among many others(32, 42, 60, 61). This master transcriptional activator of virulence either directly (*icsA, virB*) or indirectly (*ipaB, ipaC, ipaD*, etc.) activates downstream virulence genes. Firstly, VirF activates transcription of *virB* (**Figure 1.5A**) allowing this secondary transcription factor to activate further downstream virulence factors such as the Ipa- and Mxi- proteins and others that form the T3SS or act as effector proteins(53). Structurally and functionally distinct from VirF and AraC-family proteins, VirB has been shown to oligomerize and bind to DNA promoters via one helix-turn-helix (HTH) motif to activate transcription of the downstream virulence genes(62). VirF also directly activates the gene responsible for cell-to-cell spread, *icsA* (**Figure 1.5B**) (60, 63). In addition to

being a transcriptional activator, VirF has been shown to be a transcriptional repressor of an antisense RNA known as RnaG, which attenuates transcription of *icsA* (**Figure 1.5C**)(63, 64).

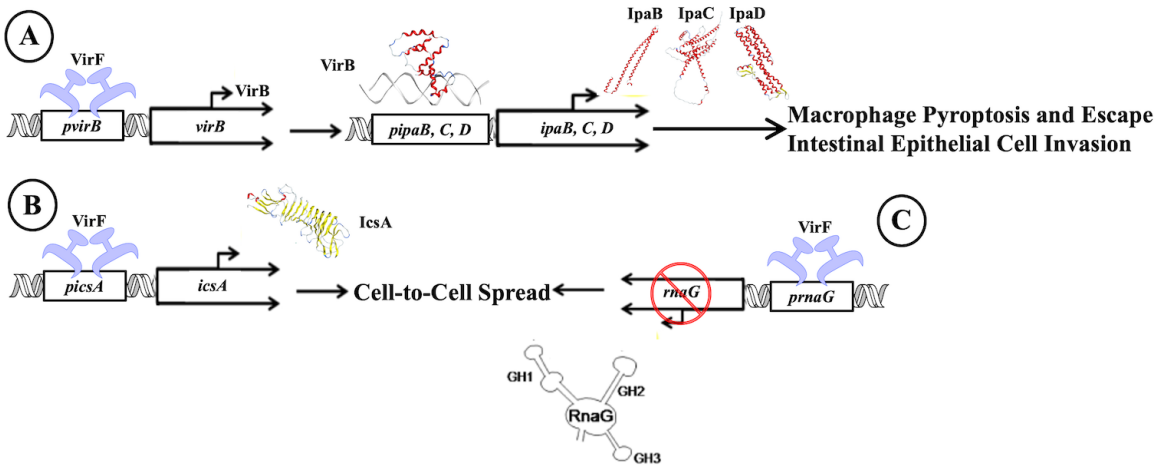


Figure 1.5: VirF controls the transcription of the *virB*, *icsA*, and *rnaG* genes. A) VirF interacts with the *virB* promoter to activate transcription of *virB*. VirB activates downstream *ipa* genes involved in macrophage pyroptosis and escape as well as intestinal epithelial cell invasion. Structures of VirB (PDB: 3W2A), IpaB (PDB: 5WKQ), IpaC (AlphaFold), IpaD (PDB: 2J0O) involved in these processes can be found according to their PDB identification numbers or via AlphaFold. B) VirF also directly activates transcription of *icsA* after binding to the *icsA* promoter but C) also represses transcription of *rnaG* within the same *icsA* regulatory region. Structures and a visual depiction of IcsA (PDB: 5KE1) and the antisense RNA, RnaG (figure adapted from Giangrossi et al., *Frontiers in Microbiology*, 2017) are included.

Although VirF controls necessary downstream virulence factors used by the pathogen to invade and spread, VirF is also regulated by a few bacterial proteins. Foremost, H-NS (heat stable nucleoid-structuring protein) is a chromosomal protein expressed in *S. flexneri* which acts as a silencer of virulence genes found on mobile genetic elements such as the *Shigella* virulence plasmid, pINV (**Figure 1.2**) (65). Virulence genes within pINV, including *virF*, are repressed at 30 ˚C but are derepressed at 37 ˚C(66). VirF also competes with H-NS at the *virB, icsA,* and *rnaG* promoters in order to activate or repress transcription(63, 67). VirF is also negatively regulated by itself. VirF exists in three forms of different molecular weights and one form, $VirF_{21}$, represses wildtype *virF* transcription by directly interacting with the *virF* promoter(68–70).

The *virF* gene is also positively affected by three proteins: FIS, IHF, and VirB. FIS (factor for inversion stimulation) is required for full expression of VirF at 37 ˚C and partially antagonizes

H-NS repressive activity at the *virF* promoter at 32 ˚C(66, 71). Another nucleoid-associated protein, IHF (integration host factor) positively regulates *virF* transcription and, in strains with defective IHF (*ihfa*-), significant reductions in Ipa and T3SS proteins were observed due to reduced VirF levels(72). Lastly, *virF* is positively regulated by one of the proteins it activates, VirB. Kane and Dorman determined that VirB mediates a positive feedback loop in which it positively regulates *virF* and *virB* transcription by antagonizing H-NS activity at both the *virF* and *virB* promoters(73).

1.4 Discovery and Development of VirF Inhibitors

VirF is a very promising anti-virulence target. In 1988, it was reported that when VirF was inactivated via Tn*5* insertion, there was a significant reduction in expression of IpaB, IpaC, IpaD and IcsA, which were all restored when VirF was re-introduced(41). Sansonetti *et al.* (74) also reported that *S. flexneri*, when expressing a mutant form of IcsA, did not present colonic tissue damage in test animals as it did not possess the ability for cell-to-cell spread. The importance of IcsA in *Shigella* was further demonstrated by the Theriot Lab, who made videos showing a dramatic drop in motility in a Δ*icsA* mutant(75). In addition, by targeting VirF with small molecules, we hypothesize that *S. flexneri* will have impaired abilities to invade the colonic epithelium and induce macrophage pyroptosis while also not pressuring the bacteria to develop resistance to the therapeutic. It has been shown that avirulent strains of *S. flexneri*, either through inactivating *virF* via insertion sequences or through removal of the virulence plasmid, the bacteria remain viable in cell culture(76, 77). Overall, these results strongly implicate VirF as a promising target for virulence-targeted therapy.

Previously, we screened ~140,000 compounds against VirF in a *Shigella*-based β-galactosidase reporter assay(78, 79). Five lead compounds (19615, 144092, 144143, 153578, 24904) were identified with IC_{50} values less than 100 μM in the reporter assay (**Figure 1.6**) (79). Of which, 19615 was confirmed to inhibit VirF DNA-binding activity for the *virB* promoter (*pvirB*) via electrophoretic mobility shift and fluorescence polarization assays(80). The remaining hits exhibited reductions in cell-to-cell spread (19615, 144092, 153578) and Caco2 invasion (144092) in plaque and invasion assays *in vivo*, respectively(79). However, a preliminary structure-activity-relationship with 19615 and a small series of commercially available analogs was performed but did not identify compounds with a higher potency than 19615(80). In Koppolu et al., a small molecule named SE-1 (**Figure 1.6**) was identified as a VirF inhibitor in a β-galactosidase reporter assays and also presented activity in *Shigella*-infected mouse fibroblast cells(81). Like 19615, SE-1 analogs did not exhibit higher potencies than SE-1(82). Overall, these screening campaigns successfully identified VirF inhibitors and showed that they exhibited *in vivo* activity but were unable to expand upon or improve these inhibitors to be progressed in the drug discovery pipeline (i.e., animal model testing).

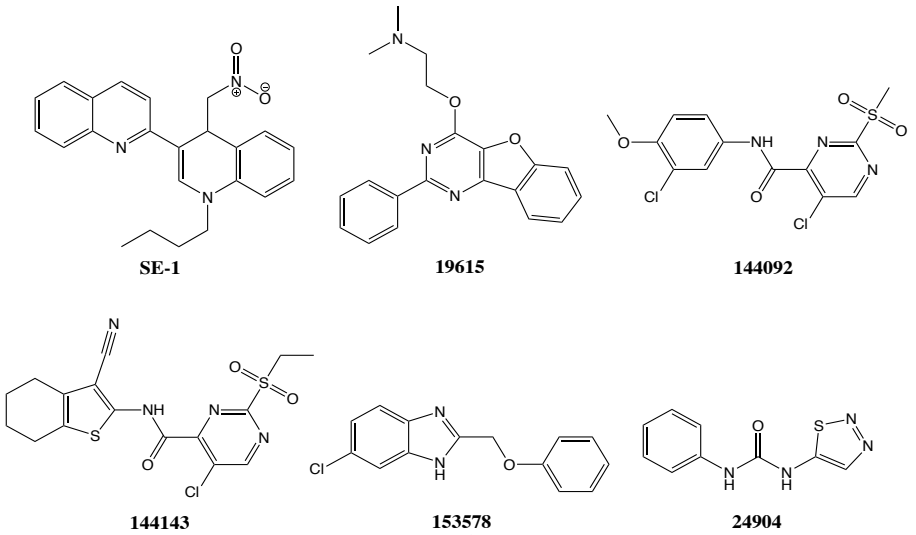


Figure 1.6: VirF inhibitors identified in Koppolu et al., 2013 (SE-1) and Emanuele et al., 2014 (19615, 144092, 144143, 153578, 24904).

Despite the identification of VirF inhibitors in Emanuele et al., the mechanism of inhibition for all these inhibitors was not identified(79). While 19615 exhibited activity against the MalE-VirF•*pvirB* DNA-binding interaction *in vitro*, it is possible the other compounds inhibited the *icsA* or *rnaG* promoter interactions. It is crucial to fully understand how VirF interacts with these other promoters *in vitro* (**Chapter 2**) to discover and develop VirF DNA-binding inhibitors, potentially with promoter-specific activity. In addition, VirF inhibitors could affect dimerization activity. However, no studies have directly probed VirF dimerization. This highlights the critical need to analyze VirF dimerization activity (**Chapter 3**) to screen for dimerization inhibitors. Dimerization inhibitors, like DNA-binding inhibitors, could exhibit successful *in vivo* activities and serve as potential virulence-targeted therapeutics.

More recently, we performed a 1.7 million compound, high-throughput, phenotypic screen in a *Shigella* intra-macrophage survival assay at GlaxoSmithKline in Tres Cantos, Spain(83). This phenotypic screen identified a set of compounds with potencies < 1 μM against *S. flexneri* virulence and provided further evidence that targeting *Shigella* virulence can yield efficacious drugs. However, the screen did not identify the molecular targets of the hits. Due to VirF's control of the *S. flexneri* virulence phenotype, it is possible these highly potent compounds could be interacting and inhibiting VirF. In **Chapter 4**, we employ our previously optimized VirF-driven β-galactosidase reporter assay to screen these hits against VirF.

1.5 Specific Aims

The long-term goal of this research is to develop novel, virulence-targeted therapeutics to be both effective treatments of shigellosis and to not further exacerbate the ongoing rise of antibiotic resistance. The overall objective of the research presented herein is to provide a functional analysis of the VirF DNA-binding and dimerization domains. Our central hypothesis

is that by studying and understanding the functions of VirF, discovery and development of lead compounds that target VirF will be more successful than previous studies. Ultimately, these VirF inhibitors are expected to display significant inhibitory effects on *S. flexneri* virulence (e.g., preventing macrophage pyroptosis, epithelial cell invasion, and cell-to-cell spread). This research and proposed future directions will provide proof of concept for future antibiotic research to control and prevent the global rise of antibiotic resistant bacterial spread and infection.

The following Specific Aims will be addressed:

AIM 1: Analysis of VirF DNA-binding Activity

- Elucidate VirF DNA-binding activity for the *virB, icsA,* and *rnaG* promoters using:
    - Homology modeling of the VirF DBD
    - Alanine-scanning mutagenesis
    - *In vitro* DNA-binding assays

AIM 2: Analysis of VirF Dimerization Activity

- Monitoring VirF dimerization using a LexA Monohybrid β-Galactosidase Reporter Assay
- Elucidation of dimerization activity by testing of VirF dimerization domain mutants:
    - Modeling of VirF dimerization domain via homology with AraC and ToxT
    - Alanine-scanning mutagenesis
    - Testing of putative dimerization mutants in LexA reporter assay
- Testing confirmed VirF dimerization mutants in *in vitro* DNA-binding assays to determine the effect of dimerization on DNA-binding

AIM 3: Screening of *Shigella* Virulence Inhibitors against VirF

Test GSK hits against VirF in a VirF-driven β-galactosidase reporter assay

1.6 References

1. O'Neill J. 2014. Review on Antimicrobial Resistance. Antimicrobial Resistance: Tackling a crisis for the health and wealth of nations. Rev Antimicrob Resist 1–16.

2. Murray CJ, Ikuta KS, Sharara F, Swetschinski L, Robles Aguilar G, Gray A, Han C, Bisignano C, Rao P, Wool E, Johnson SC, Browne AJ, Chipeta MG, Fell F, Hackett S, Haines-Woodhouse G, Kashef Hamadani BH, Kumaran EAP, McManigal B, Agarwal R, Akech S, Albertson S, Amuasi J, Andrews J, Aravkin A, Ashley E, Bailey F, Baker S, Basnyat B, Bekker A, Bender R, Bethou A, Bielicki J, Boonkasidecha S, Bukosia J, Carvalheiro C, Castañeda-Orjuela C, Chansamouth V, Chaurasia S, Chiurchiù S, Chowdhury F, Cook AJ, Cooper B, Cressey TR, Criollo-Mora E, Cunningham M, Darboe S, Day NPJ, De Luca M, Dokova K, Dramowski A, Dunachie SJ, Eckmanns T, Eibach D, Emami A, Feasey N, Fisher-Pearson N, Forrest K, Garrett D, Gastmeier P, Giref AZ, Greer RC, Gupta V, Haller S, Haselbeck A, Hay SI, Holm M, Hopkins S, Iregbu KC, Jacobs J, Jarovsky D, Javanmardi F, Khorana M, Kissoon N, Kobeissi E, Kostyanev T, Krapp F, Krumkamp R, Kumar A, Kyu HH, Lim C, Limmathurotsakul D, Loftus MJ, Lunn M, Ma J, Mturi N, Munera-Huertas T, Musicha P, Mussi-Pinhata MM, Nakamura T, Nanavati R, Nangia S, Newton P, Ngoun C, Novotney A, Nwakanma D, Obiero CW, Olivas-Martinez A, Olliaro P, Ooko E, Ortiz-Brizuela E, Peleg AY, Perrone C, Plakkal N, Ponce-de-Leon A, Raad M, Ramdin T, Riddell A, Roberts T, Robotham JV, Roca A, Rudd KE, Russell N, Schnall J, Scott JAG, Shivamallappa M, Sifuentes-Osornio J, Steenkeste N, Stewardson AJ, Stoeva T, Tasak N, Thaiprakong A, Thwaites G, Turner C, Turner P, van Doorn HR, Velaphi S, Vongpradith A, Vu H, Walsh T, Waner S, Wangrangsimakul T, Wozniak T, Zheng P, Sartorius B, Lopez AD, Stergachis A, Moore C, Dolecek C, Naghavi M. 2022. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. Lancet 399:629–655.

3. Thorpe KE, Joski P, Johnston KJ. 2018. Antibiotic-Resistant Infection Treatment Costs Have Doubled Since 2002, Now Exceeding $2 Billion Annually. Health Aff 37:662–669.

4. Roberts RR, Hota B, Ahmad L, Scott D, Foster SD, Abbasi F, Schabowski S, Kampe LM, Gr-Ciavarella G, Supino M, Naples J, Cordell R, Levy SB, Weinstein RA. 2009. Hospital and Societal Costs of Antimicrobial-Resistant Infections in a Chicago Teaching Hospital: Implications for Antibiotic Stewardship. Clin Infect Dis 49:1175–1184.

5. U.S. Department of Health and Human Services: Centers for Disease Control and Prevention. 2013. Antibiotic Resistance Threats in US, 2013 1–114.

6. Miethke M, Pieroni M, Weber T, Brönstrup M, Hammann P, Halby L, Arimondo PB, Glaser P, Aigle B, Bode HB, Moreira R, Li Y, Luzhetskyy A, Medema MH, Pernodet JL, Stadler M, Tormo JR, Genilloud O, Truman AW, Weissman KJ, Takano E, Sabatini S, Stegmann E, Brötz-Oesterhelt H, Wohlleben W, Seemann M, Empting M, Hirsch AKH, Loretz B, Lehr CM, Titz A, Herrmann J, Jaeger T, Alt S, Hesterkamp T, Winterhalter M, Schiefer A, Pfarr K, Hoerauf A, Graz H, Graz M, Lindvall M, Ramurthy S, Karlén A, van Dongen M, Petkovic H, Keller A, Peyrane F, Donadio S, Fraisse L, Piddock LJV, Gilbert IH, Moser HE, Müller R. 2021. Towards the sustainable discovery and development of new antibiotics.

Nat Rev Chem 5:726–749.

7. Andrei S, Droc G, Stefan G. 2019. FDA approved antibacterial drugs: 2018-2019. Discoveries 7:e102.

8. Plackett B. 2020. No money for new drugs. Nat Outlook 586:50–52.

9. Osterloh A. 2022. Vaccination against Bacterial Infections: Challenges, Progress, and New Approaches with a Focus on Intracellular Bacteria. Vaccines 10.

10. Poolman JT. 2020. Expanding the role of bacterial vaccines into life-course vaccination strategies and prevention of antimicrobial-resistant infections. npj Vaccines 5.

11. Prendergast AJ. 2015. Malnutrition and vaccination in developing countries. Philos Trans R Soc B Biol Sci 370:1–8.

12. Chen WH, Kozlovsky BF, Effros RB, Grubeck-Loebenstein B, Edelman R, Sztein MB. 2014. Vaccination in the elderly: an immunological perspective. Trends Immunol 30:351–359.

13. Clatworthy AE, Pierson E, Hung DT. 2007. Targeting virulence: A new paradigm for antimicrobial therapy. Nat Chem Biol 3:541–548.

14. Beceiro A, Tomás M, Bou G. 2013. Antimicrobial Resistance and Virulence: a Successful or Deleterious Association in the Bacterial World? Clin Microbiol Rev 26:185–230.

15. Jewell NP, Ph D, Hatheway CL. 2006. Human Botulism Ummine Globin for the Treatment of Infant Botulism. NEW Engl YOURNAL Med 354:462–471.

16. Migone T-S, Subramanian GM, Zhong J, Healey LM, Corey A, Devalaraja M, Lo L, Ullrich S, Zimmerman J, Chen A, Lewis M, Meister G, Gillum K, Sanford D, Mott J, Bolmer SD. 2009. Raxibacumab for the Treatment of Inhalational Anthrax. N Engl J Med 361:135–144.

17. Yamamoto BJ, Shadiack AM, Carpenter S, Sanford D, Henning LN, Gonzales N, O'Connor E, Casey LS, Serbina N V. 2016. Obiltoxaximab prevents disseminated Bacillus anthracis infection and improves survival during pre- and postexposure prophylaxis in animal models of inhalational anthrax. Antimicrob Agents Chemother 60:5796–5805.

18. Greig SL. 2016. Obiltoxaximab: First Global Approval. Drugs 76:823–830.

19. Rounds J, Strain J. 2017. Bezlotoxumab for Preventing Recurrent Clostridium difficile Infections. S D Med 70:422–423.

20. Babcock GJ, Broering TJ, Hernandez HJ, Mandell RB, Donahue K, Boatright N, Stack AM, Lowy I, Graziano R, Molrine D, Ambrosino DM, Thomas WD. 2006. Human monoclonal antibodies directed against toxins A and B prevent Clostridium difficile-induced mortality in hamsters. Infect Immun 74:6339–6347.

21. Lowy I, Molrine DC, Leav BA, Blair BM, Baxter R, Gerding DN, Nichol G, Thomas Jr. WD, Leney M, Sloan S, Hay CA, Ambrosino DM. 2010. Treatment with Monoclonal Antibodies against Clostridium difficile Toxins. N Engl J Med 362:197–205.

22. Dickey SW, Cheung GYC, Otto M. 2017. Different drugs for bad bugs: Antivirulence strategies in the age of antibiotic resistance. Nat Rev Drug Discov 16:457–471.

23. Ogawara H. 2021. Possible drugs for the treatment of bacterial infections in the future: anti-virulence drugs. J Antibiot (Tokyo) 74:24–41.

24. CDC. 2019. Antibiotic Resistance Threats in the United States, 2019. Centers Dis Control Prev 1–140.

25. Khalil IA, Troeger C, Blacker BF, Rao PC, Brown A, Atherly DE, Brewer TG, Engmann CM, Houpt ER, Kang G, Kotloff KL, Levine MM, Luby SP, MacLennan CA, Pan WK, Pavlinac PB, Platts-Mills JA, Qadri F, Riddle MS, Ryan ET, Shoultz DA, Steele AD, Walson JL, Sanders JW, Mokdad AH, Murray CJL, Hay SI, Reiner RC. 2018. Morbidity and mortality due to shigella and enterotoxigenic Escherichia coli diarrhoea: the Global Burden of Disease Study 1990–2016. Lancet Infect Dis 18:1229–1240.

26. Williams P, Berkley JA. 2018. Guidelines for the treatment of dysentery (shigellosis): a systematic review of the evidence. Paediatr Int Child Health 38:550–565.

27. Ranjbar R, Farahani A. 2019. Shigella: Antibiotic-Resistance Mechanisms And New Horizons For Treatment. Infect Drug Resist 12:3137–3167.

28. Xue C, Cai J, Kang H, Chen Y, Wang K, Qian H, Bao C, Li N, Guo Z, Zhang Z, Wang J, Ma P, Gu B. 2018. Two novel mutations in parE among Shigella flexneri isolated from Jiangsu Province of China, 2016. Ann Transl Med 6:306–306.

29. Bhattacharya D, Bhattacharya H, Thamizhmani R, Sayi DS, Reesu R, Anwesh M, Kartick C, Bharadwaj AP, Singhania M, Sugunan AP, Roy S. 2014. Shigellosis in Bay of Bengal Islands, India: Clinical and seasonal patterns, surveillance of antibiotic susceptibility patterns, and molecular characterization of multidrug-resistant Shigella strains isolated during a 6-year period from 2006 to 2011. Eur J Clin Microbiol Infect Dis 33:157–170.

30. Gaudreau C, Barkati S, Leduc JM, Pilon PA, Favreau J, Bekal S. 2014. Shigella spp. with Reduced Azithromycin Susceptibility, Quebec, Canada, 2012–2013. Emerg Infect Dis 20:854–856.

31. WHO. 2017. Global Antimicrobial Resistance Surveillance System (GLASS) Report. Geneva, World Health OrganizationWho. https://apps.who.int/iris/bitstream/handle/10665/279656/9789241515061-eng.pdf?ua=1.

32. Sansonnetti PJ, Kopecko DJ, Formal SB. 1982. Involvement of a plasmid in the invasive ability of Shigella flexneri. Infect Immun 35:852–860.

33. Hale TL, Sansonetti PJ, Schad PA, Austin S, Formal SB. 1983. Characterization of

Virulence Plasmids and Plasmid Associated Outer Membrane Proteins in Shigella flexneri, Shigella sonnei, and Escherichia coli. Infect Immun 40:340–350.

34. Pasqua M, Michelacci V, Di Martino ML, Tozzoli R, Grossi M, Colonna B, Morabito S, Prosseda G. 2017. The Intriguing Evolutionary Journey of Enteroinvasive E. coli (EIEC) toward Pathogenicity. Front Microbiol 8:1–12.

35. Schroeder GN, Hilbi H. 2008. Molecular Pathogenesis of Shigella spp.: Controlling Host Cell Signaling, Invasion, and death by Type III Secretion. Clin Microbiol Rev 21:134–156.

36. DuPont H, Levine M, Hornick R, Formal SB. 1989. Inoculum Size in Shigellosis and Implications For Expected Mode of Transmission. J Infect Dis 159:1126–1128.

37. Small P, Blankenhorn D, Welty D, Zinser E, Slonczewski JL. 1994. Acid and base resistance in Escherichia coli and Shigella flexneri: Role of rpoS and growth pH. J Bacteriol 176:1729–1737.

38. Jennison A V., Verma NK. 2004. Shigella flexneri infection: pathogenesis and vaccine development. FEMS Microbiol Rev 28:43–58.

39. Croxen MA, Law RJ, Scholz R, Keeney KM, Wlodarska M, Finlay BB. 2013. Recent Advances in Understanding Enteric Pathogenic Escherichia coli. Clin Microbiol Rev 26:822–880.

40. Arena ET, Campbell-Valois FX, Tinevez JY, Nigro G, Sachse M, Moya-Nilges M, Nothelfer K, Marteyn B, Shorte SL, Sansonetti PJ. 2015. Bioimage analysis of Shigella infection reveals targeting of colonic crypts. Proc Natl Acad Sci U S A 112:E3282–E3290.

41. Sakai T, Sasakawa C, Yoshikawa M. 1988. Expression of four virulence antigens of Shigella flexneri is positively regulated at the transcriptional level by the 30 kilo Dalton virF protein. Mol Microbiol 2:589–597.

42. Bernardini ML, Mounier J, D'Hauteville H, Coquis-Rondon M, Sansonetti PJ. 1989. Identification of icsA, a plasmid locus of Shigella flexneri that governs bacterial intra- and intercellular spread through interaction with F-actin. Proc Natl Acad Sci U S A 86:3867–71.

43. High N, Mounier J, Prévost MC, Sansonetti PJ. 1992. IpaB of Shigella flexneri causes entry into epithelial cells and escape from the phagocytic vacuole. EMBO J 11:1991–9.

44. Chen Y, Smith MR, Thirumalai K, Zychlinsky A. 1996. A bacterial invasin induces macrophage apoptosis by binding directly to ICE. EMBO J 15:3853–3860.

45. Blocker A, Gounon P, Larquet E, Niebuhr K, Cabiaux V, Parsot C, Sansonetti P. 1999. The tripartite type III secreton of Shigella flexneri inserts IpaB and IpaC into host membranes. J Cell Biol 147:683–693.

46. SenerovicL., Tsunoda SP, GoosmannC., Brinkmann V, Zychlinsky A, Meissner F, Kolbe

M. 2012. Spontaneous formation of IpaB ion channels in host cell membranes reveals how Shigella induces pyroptosis in macrophages. Cell Death Dis 3:e384-10.

47. Guichon A, Hersh D, Smith MR, Zychlinsky A. 2001. Structure-function analysis of the Shigella virulence factor IpaB. J Bacteriol 183:1269–1276.

48. Menard R, Sansonetti PJ, Parsot C. 1993. Nonpolar mutagenesis of the ipa genes defines IpaB, IpaC, and IpaD as effectors of Shigella flexneri entry into epithelial cells. J Bacteriol 175:5899–5906.

49. Harrington AT, Hearn PD, Picking WL, Barker JR, Wessel A, Picking WD. 2003. Structural Characterization of the N Terminus of IpaC from Shigella flexneri. Infect Immun 71:1255–1264.

50. Veenendaal AKJ, Hodgkinson JL, Schwarzer L, Stabat D, Zenk SF, Blocker AJ. 2007. The type III secretion system needle tip complex mediates host cell sensing and translocon insertion. Mol Microbiol 63:1719–1730.

51. Espina M, Olive AJ, Kenjale R, Moore DS, Ausar SF, Kaminski RW, Oaks E V., Middaugh CR, Picking WD, Picking WL. 2006. IpaD localizes to the tip of the type III secretion system needle of Shigella flexneri. Infect Immun 74:4391–4400.

52. Picking WL, Nishioka H, Hearn PD, Baxter MA, Harrington AT, Blocker A, Picking WD. 2005. IpaD of Shigella flexneri is independently required for regulation of Ipa protein secretion and efficient insertion of IpaB and IpaC into host membranes. Infect Immun 73:1432–1440.

53. Adler B, Sasakawa C, Tobe T, Makino S, Komatsu K, Yoshikawa M. 1989. A dual transcriptional activation system for the 230 kb plasmid genes coding for virulence-associated antigens of Shigella flexneri. Mol Microbiol 3:627–635.

54. Goldberg MB, Barzu O, Parsot C, Sansonetti PJ. 1993. Unipolar localization and ATPase activity of IcsA, a Shigella flexneri protein involved in intracellular movement. J Bacteriol 175:2189–2196.

55. Goldberg MB, Theriot JA. 1995. Shigella flexneri surface protein IcsA is sufficient to direct actin- based motility. Proc Natl Acad Sci U S A 92:6572–6576.

56. Rohatgi R., Ma L., Miki H., Lopez M., Kirchhausen T., Takenawa T., Kirschner M.W. 1999. The Interaction between N-WASP and the Arp2/3 Complex Links Cdc42-Dependent Signals to Actin Assembly. Cell 97:221–231.

57. Suzuki T, Miki H, Takenawa T, Sasakawa C. 1998. Neural Wiskott-Aldrich syndrome protein is implicated in the actin-based motility of Shigella flexneri. EMBO J 17:2767–2776.

58. May KL, Morona R. 2008. Mutagenesis of the Shigella flexneri autotransporter IcsA reveals novel functional regions involved in IcsA biogenesis and recruitment of host neural

Wiscott-Aldrich syndrome protein. J Bacteriol 190:4666–4676.

59. Martin RG, Rosner JL. 2001. The AraC transcriptional activators. Curr Opin Microbiol 4:132–137.

60. Sakai T, Sasakawa C, Yoshikawa M. 1988. Expression of four virulence antigens of Shigella flexneri is positively regulated at the transcriptional level by the 30 kilo Dalton virF protein. Mol Microbiol 2:589–597.

61. Prosseda G, Fradiani P, Lorenzo M Di, Falconi M, Micheli G, Casalino M, Nicoletti M, Colonna B. 1998. A role for H-NS in the regulation of the virF gene of Shigella and enteroinvasive Escherichia coli. Res Microbiol 149:15–25.

62. McKenna S, Beloin C, Dorman CJ. 2003. In vitro DNA-binding properties of VirB, the Shigella flexneri virulence regulatory protein. FEBS Lett 545:183–187.

63. Tran CN, Giangrossi M, Prosseda G, Brandi A, Di Martino ML, Colonna B, Falconi M. 2011. A multifactor regulatory circuit involving H-NS, VirF and an antisense RNA modulates transcription of the virulence gene icsA of Shigella flexneri. Nucleic Acids Res 39:8122–8134.

64. Giangrossi M, Prosseda G, Tran CN, Brandi A, Colonna B, Falconi M. 2010. A novel antisense RNA regulates at transcriptional level the virulence gene icsA of Shigella flexneri. Nucleic Acids Res 38:3362–3375.

65. Dorman CJ. 2004. H-NS: a universal regulator for a dynamic genome. Nat Rev Microbiol 2:391–400.

66. Falconi M, Colonna B, Prosseda G, Micheli G, Gualerzi CO. 1998. Thermoregulation of Shigella and Escherichia coli EIEC pathogenicity. A temperature-dependent structural transition of DNA modulates accessibility of virF promoter to transcriptional repressor H-NS. EMBO J 17:7033–7043.

67. Tobe T, Yoshikawa M, Mizuno T, Sasakawa C. 1993. Transcriptional Control of the Invasion Regulatory Gene virB of Shigella flexneri: Activation by VirF and Repression by H-NS. J Bacteriol 175:6142–6149.

68. Di Martino ML, Romilly C, Wagner EGH, Colonna B, Prosseda G. 2016. One Gene and Two Proteins: a Leaderless mRNA Supports the Translation of a Shorter Form of the Shigella VirF Regulator. MBio 7:1–10.

69. Skovajsova E, Colonna B, Prosseda G, Sellin ME, Martino ML Di. 2022. The VirF21:VirF30 protein ratio is affected by temperature and impacts Shigella flexneri host cell invasion. FEMS Microbiol Lett 369:1–9.

70. Sakai T, Sasakawa C, Makino S, Yoshikawa M. 1986. DNA Sequence and Product Analysis of the virF Locus Responsible for Congo Red Binding and Cell Invasion in Shigella flexneri 2a. Infect Immun 54:395–402.

71. Falconi M, Prosseda G, Giangrossi M, Beghetto E, Colonna B. 2001. Involvement of FIS in the H-NS-mediated regulation of virF gene of Shigella and enteroinvasive Escherichia coli. Mol Microbiol 42:439–452.

72. Porter ME, Dorman CJ. 1997. Positive Regulation of Shigella flexneri Virulence Genes by Integration Host Factor. J Bacteriol 179:6537–6550.

73. Kane KA, Dorman CJ. 2012. VirB-Mediated Positive Feedback Control of the Virulence Gene Regulatory Cascade of Shigella flexneri. J Bacteriol 194:5264–5273.

74. Sansonetti PJ, Arondel J, Fontaine A, d'Hauteville H, Bernardini ML. 1991. OmpB (osmo-regulation) and icsA (cell-to-cell spread) mutants of Shigella flexneri: vaccine candidates and probes to study the pathogenesis of shigellosis. Vaccine 9:416–422.

75. Theriot J, Portnoy D. 2022. Theriot Lab at the University of Washington: Publications and Movies.

76. Maurelli AT, Blackmon B, Curtiss Iii R. 1984. Loss of Pigmentation in Shigella flexneri 2a Is Correlated with Loss of Virulence and Virulence-Associated PlasmidINFECTION AND IMMUNITY.

77. Mills JA, Venkatesan MM, Baron LS, Buysse JM. 1992. Spontaneous Insertion of an ISJ-Like Element into the virF Gene Is Responsible for Avirulence in Opaque Colonial Variants of Shigella flexneri 2a. Infect Immun 60:175–182.

78. Hurt JK, McQuade TJ, Anthony E, Larsen MJ, Garcia GA. 2010. High-Throughput Screening of the Virulence Regulator VirF: A Novel Antibacterial Target for Shigellosis. J Biomol Screen 15:379–387.

79. Emanuele AA, Adams NE, Chen YC, Maurelli AT, Garcia GA. 2014. Potential novel antibiotics from HTS targeting the virulence-regulating transcription factor, VirF, from Shigella flexneri. J Antibiot (Tokyo) 67:379–386.

80. Emanuele AA, Garcia GA. 2015. Mechanism of Action and Initial, In Vitro SAR of an Inhibitor of the Shigella flexneri Virulence Regulator VirF. PLoS One 10:1–18.

81. Koppolu V, Osaka I, Skredenske JM, Kettle B, Hefty PS, Li J, Egan SM. 2013. Small-Molecule Inhibitor of the Shigella flexneri Master Virulence Regulator VirF. Infect Immun 81:4220–4231.

82. Jain P, Li J, Porubsky P, Neuenswander B, Egan SM, Aubé J, Rogers S. 2014. 3-Substituted Biquinolinium Inhibitors of AraC Family Transcriptional Activator VirF from S. flexneri Obtained Through In Situ Chemical Ionization of 3,4-Disubstituted Dihydroquinolines. RSC Adv 4:39809–39816.

83. Miljkovic M, Lozano S, Castellote I, de Cózar C, Villegas-Moreno AI, Gamallo P, Martinez DJ-A, Fernández-Álvaro E, Ballell L, Garcia GA. 2021. Novel inhibitors that target bacterial virulence identified via HTS against intramacrophage survival of Shigella flexneri.

Chapter 2 Analysis of VirF DNA-Binding Activity

2.1 Abstract

AraC-family proteins are found in many bacteria and regulate transcription of many genes involved in metabolism, stress, or virulence-related pathways. These proteins contain a conserved C-terminal DNA-binding domain (DBD) harboring two helix-turn-helix (HTH) motifs involved in binding to DNA promoters to either activate or repress gene transcription. VirF from *Shigella flexneri* is an AraC-homolog that activates transcription of downstream virulence genes needed for the bacteria to be infect and invade the human colon. Previous studies of the VirF DNA-binding activity were performed to identify DNA-binding inhibitors, but follow-up studies were unable to develop more potent inhibitors for pre-clinical testing. Further analyses of VirF DNA-binding activity through structural and *in vitro* studies are essential for overall understanding of this protein's function and continued inhibitor development.

2.2 Introduction

Widespread in bacteria, AraC family proteins contain a highly conserved DNA-binding domain (DBD) harboring two canonical helix-turn-helix (HTH) motifs(1, 2). AraC-family proteins are often highly insoluble and difficult to work with *in vitro*(3). VirF, a DNA-binding transcriptional activator found in *Shigella flexneri*, is an AraC-family protein that exhibits these difficult characteristics. Tran et al. isolated full-length VirF for testing (4) but in insufficient yields for extensive *in vitro* assays or X-ray crystallography. In contrast, much success with expression, purification, and solubility has been with MalE-VirF fusion proteins(5–7). The increased yields

and solubility have made it possible to test DNA-binding activities of VirF at higher concentrations than previous studies. However, to date there is still no three-dimensional structure available for VirF. Only a few AraC homologs have been crystallized including AraC (8, 9), GadX, MarA (10), and Rns (11) from *E. coli* as well ToxT (12, 13) from *Vibrio cholerae*. This has made it immensely difficult to study VirF's DNA-binding activity. Currently, there are two structures of AraC-family proteins that are bound to their cognate DNA-promoters: MarA with *marRAB*(10), and Rob with *micF*(14). These structures represent two distinct binding methods in which MarA interacts with DNA via both HTH motifs while Rob binds with only its N-terminal HTH motif. Porter and Dorman determined that mutations in both HTH motifs in VirF can affect its ability to activate β-galactosidase activity so it is likely that VirF utilizes its entire DBD to interact with its cognate promoters(15).

VirF is known to interact with three DNA promoters: *virB*, *icsA*, and *rnaG*. The most extensively studied VirF DNA-binding interaction is with the *virB* promoter. In the earliest studies of VirF, it was shown that VirF increased expression of IpaB, IpaC, and IpaD through activation of an "intermediate protein" known as VirB(16). Following this study, Tobe et al. confirmed that a VirF fusion protein, MalE-VirF, protected a stretch of approximately 100 nucleotides in a DNase I footprinting assay(5). This also was the first study where VirF was purified with a large molecular weight tag, MalE (maltose binding protein from *E. coli*), and successfully used in *in vitro* DNA-binding and transcription assays(5–7). Since these early results, this protein•DNA interaction has been utilized for screening of inhibitors *in vivo* and confirmation of DNA-binding *in vitro*. In the Garcia lab, we screened ~140,000 compounds in a *Shigella*-based β-galactosidase reporter assay assessing compound inhibition of MalE-VirF activation of a *pvirB- lacZ* fusion(17, 18). In Emanuel and Garcia, follow up studies identified one of the five hits, 19615 (**Figure 2.1**),

exhibited inhibition of MalE-VirF DNA-binding activity for a fluorescently-labeled *virB* promoter (*pvirB*) DNA fragment using an electrophoretic mobility shift assay (EMSA) (6). Koppolu et al. employed a similar reporter assay to screen for VirF inhibitors and identified one compound, SE-1 (**Figure 2.1)**, which also directly inhibited MalE-VirF DNA-binding in an EMSA using a labeled *pvirB*(7).

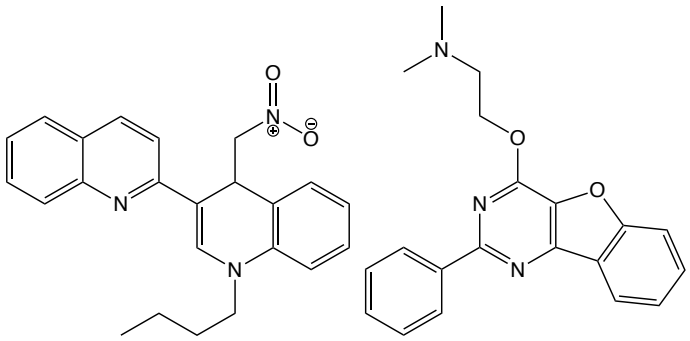


Figure 2.1: Lead VirF DNA-binding inhibitors identified in Emanuele and Garcia, 2015 and Koppolu et al., 2013.

VirF has also been shown to directly activate *icsA* and repress *rnaG* transcription(4). VirF regulates transcription of these genes in a complex DNA regulatory region via binding to both the *icsA* and *rnaG* promoter regions to activate transcription of *icsA* and repress *rnaG*(4). VirF also competes with the transcriptional silencer, H-NS, within this regulatory region(4, 19). Adding to the complexity of regulating *icsA* transcription, RnaG is a cis-encoded anti-sense RNA which downregulates *icsA* transcription(4, 20). VirF represses RnaG transcription thus promoting cell-to-cell spread. Unlike the *virB* promoter, this regulatory region is very interesting as it implicates VirF can act as both an activator (*virB* and *icsA*) as well as a repressor (*rnaG*). Tran et al. employed footprinting analyses which identified DNase I protected sites within the *icsA* regulatory region indicating the locations of VirF binding sites for the *icsA* and *rnaG* promoters (**Figure 2.2**) (4). Utilizing an i*n vitro* transcription assay, *icsA* transcription was activated only in the presence of VirF. Increases in VirF concentration led to significant reductions in RnaG production(4). Further

studies did not elaborate on VirF DNA-binding activity for these two DNA promoters but instead discovered that VirF has RNA-binding activity for the RnaG and the *icsA* mRNA(21).

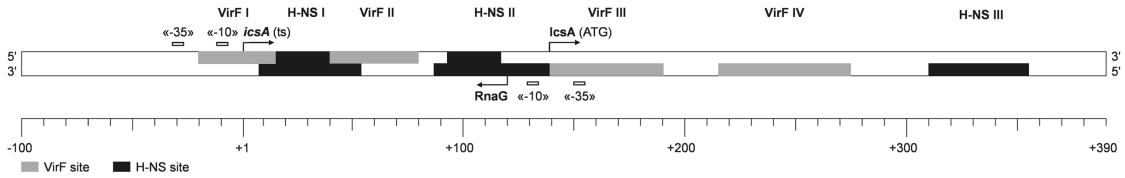


Figure 2.2: The complex *icsA* regulatory region. IcsA transcription is positively controlled by VirF but negatively controlled by H-NS and RnaG. RnaG transcription is downregulated by both VirF and H-NS. The figure is adapted from Tran et al., *Nucleic Acids Research*, 2011.

Overall, these previous studies identify validated approaches for studying VirF DNA-binding activity. Herein, we expand studies of the VirF DNA-binding activity to its two other cognate promoters, *picsA* and *prnaG*. Homology modeling was performed with two homologs, MarA and GadX, to predict structures for the VirF DBD which informed site-directed mutagenesis on the DBD to identify trends and differential binding activities with the DNA-binding mutants.

2.3 Results

2.3.1 Homology Modeling and Alignments of the VirF DNA-Binding Domain (DBD) with MarA and GadX[§]

VirF DBD homology models were generated by Nicholas Ragazzone using both the GadX and MarA as structural templates (**Figure 2.3)**. The VirF models were superimposed and overlayed with the MarA and GadX templates **(Figure 2.3A and 2.3B)** and exhibited RMSD values of 3.57 and 0.61, respectively. When the two homology models were overlayed (**Figure 2.3C**), it presented an RMSD value of 3.57. As depicted in Rhee et al., MarA binds to *marRAB* using both helix-3 and helix-6 which reside within the major grooves of the promoter(10). This characteristic is also visible for the MarA-based homology model (**Figure 2.3A**). The GadX-based model overlayed closer to the GadX template than the MarA-based model for its template (**Figure**

2.3A and 2.3B). For instance, helix-2 in the MarA-based model (**Figure 2.3A**) is shown as a free loop but when overlayed with MarA or the GadX-based model, it appears this region is helical. This is likely due to the VirF DBD having a lower sequence identity with MarA than for GadX. When the two models were superimposed and overlayed, helix-6 shared a similar position in space among the two homology models (**Figure 2.3C**) but helix-3 showed variability in its placement. However, the percent identity and RMSD values did not vary much between the two regions.

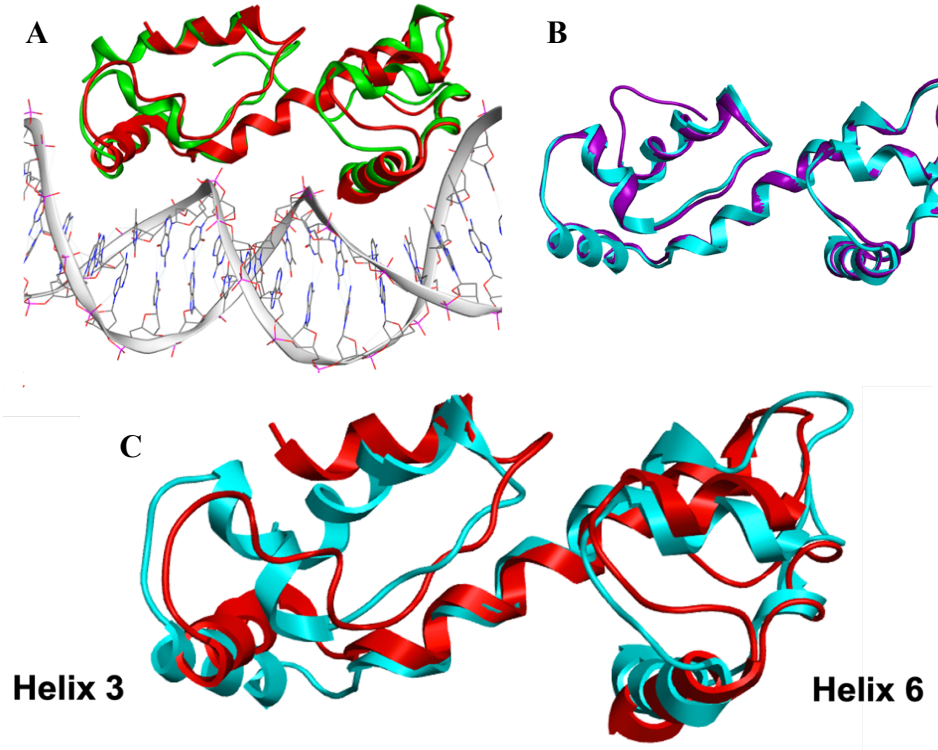


Figure 2.3: VirF DBD homology models and overlays with MarA and GadX.
(**A**): MarA-based homology model of the VirF DBD (Red) overlayed with the MarA•*marRAB* crystal structure (PDB ID: 1BL0) (**B**): The VirF DBD homology model (Cyan) created using GadX (PDB ID: 3MKL) as a structural template overlayed with the GadX structure (Purple). (**C**): Overlay of MarA- and GadX-based VirF DBD homology models, MarA-based (Red) and GadX-based (Cyan), overlayed to show shift in confirmation.

The MarA crystal structure was examined further to identify amino acids that make critical contacts with its cognate DNA promoter, *marRAB*. Based on this examination, three amino acid side chains in helix-3 and four in helix-6 interact with DNA (**Figure 2.4**)(10). Using EXPASY SIM-Alignment, the VirF DBD primary sequence was aligned with MarA which allowed us to identify VirF DBD amino acids that correspond to these seven MarA residues (**Figure 2.4**)(22).

The MarA-based model indicated that all seven predicted amino acids resided within helices-3 and -6 which further validates this homology model.

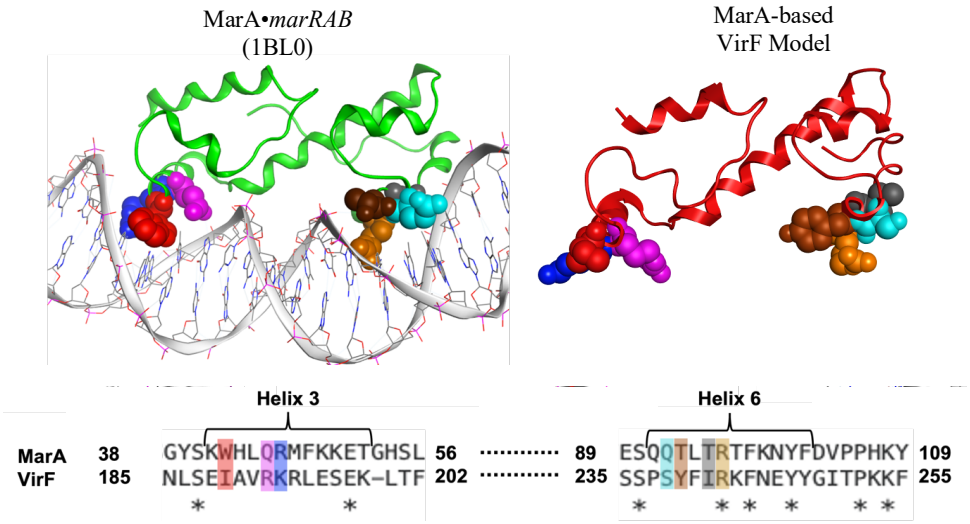


Figure 2.4: Alignment of VirF with MarA to predict crucial DNA-binding residues. Spatial orientations of the seven residues contributing to the MarA•*marRAB* (green) DNA-binding interaction and the corresponding, aligned residues in VirF (red). An alignment of MarA and the VirF DBD is included below, and the residues are color coded based on this alignment (* represents complete similarity between the aligned residues).

[§]The experiments and results within this sub-section were performed and analyzed by Nicholas Ragazzone. More details can be found in Ragazzone et al., 2022(23).

2.3.2 Purification of WT MalE-VirF and DBD mutants

Prior to purification of these proteins, *malE-virF*, within the expression plasmid pBAD202-MALvirF, was subject to alanine-scanning mutagenesis at the seven residues identified in the alignment above (**Figure 2.4**). The residues that were mutated include I189, R192, K193, S238, Y239, I241, and R242. Following mutagenesis and sequencing analysis of these mutants, VirF was heterologously expressed in TOP10 *E. coli*. VirF is expressed and purified as a fusion protein, MalE-VirF, to increase protein expression, yield, and solubility(5, 6). In Emanuele and Garcia(6), MalE-VirF expression occurred homologously in avirulent BS103 *Shigella flexneri* since

expression and purification in *E. coli* KS1000 yielded more truncated MalE rather than the desired fusion protein. Unfortunately, attempts were unsuccessful to express and purify MalE-VirF using this homologous expression method. To circumvent this, MalE-VirF was heterologously expressed in TOP10 *E. coli*. This cell line, like BS103 *S. flexneri*, contains the *araBAD* operon which allows for arabinose-inducible expression of MalE-VirF. Expression of WT MalE-VirF and the corresponding DNA-binding mutants was induced with 0.2% w/v arabinose.

After harvesting the TOP10 *E. coli* containing the pBAD202-MALvirF WT or mutant expression plasmids, the cells were subject to sonication and centrifugation to obtain the resultant lysate. Unlike in Emanuele and Garcia, protein purification (amylose resin gravity column) was performed using an AKTA FPLC. The lysate was loaded onto a 5 mL MBPTrap HP column (Cytiva) and MalE-VirF was eluted from the column using an elution buffer containing 10 mM maltose (**Figure 2.5**). Following elution, the eluted protein was concentrated and loaded onto a size exclusion column (Superdex 200 GL

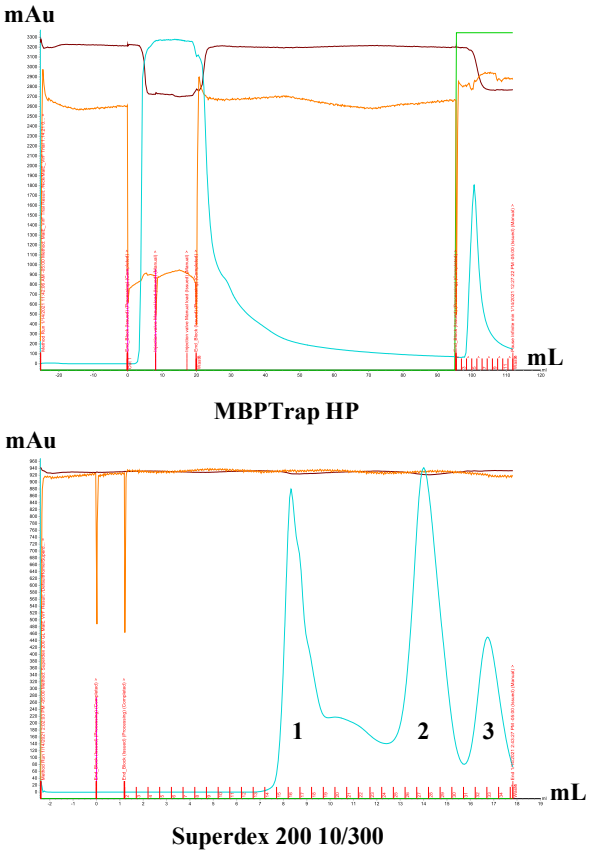


Figure 2.5: Chromatogram for MBPTrap HP and size exclusion chromatography (Superdex 200 10/300) for the WT purification. The Y-axes (mAu at 280 nm). The peak for the MBPTrap HP column was concentrated and loaded onto the Superdex column to separate MalE-VirF from impurities and truncations. Peaks are labeled according to what is normally seen in SDS-PAGE from WT MalE-VirF purifications (1=larger protein impurities and "aggregated MalE-VirF" species, 2=MalE-VirF fractions selected for dialysis and testing, 3=MalE or truncated MalE-VirF).

10/300) where protein impurities, aggregated MalE-VirF, and truncated MalE was separated from pure MalE-VirF (**Figure 2.5**). Elution peak 2 from the size exclusion chromatography purification resulted in pure and active WT MalE-VirF. Although the majority of protein impurities are removed from the MalE-VirF via size exclusion chromatography, significantly low levels remain in the purified sample SDS-PAGE of the MBP Elution fractions, and all three peaks from obtained from the Superdex 200 column purification was performed and showed that WT MalE-VirF (molecular weight ~73 kDa) was successfully purified from the impurities and truncated MalE (molecular weight ~43 kDa) using the Superdex 200 column (**Figure 2.6**). A Bradford assay confirmed the protein concentration to be 57.5 μM when compared to a set of bovine serum albumin (BSA) standards.

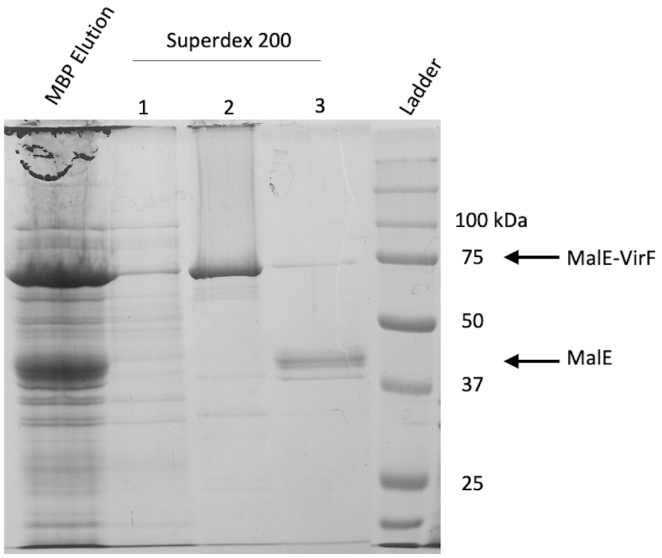


Figure 2.6: SDS-PAGE for WT MalE-VirF purification using an MBPTrap HP (MBP Elution) and Superdex 200 GL 10/300 gel filtration columns (Superdex 200 1, 2, 3 correspond to **Figure 2.5**). Locations of eluted proteins from Superdex column are noted by arrows.

After successful purifications of WT MalE-VirF, this method was used to express and purify the seven DBD mutants. Chromatograms and SDS-PAGE for these purifications can be found in **Figures S2.1 and S2.2**. The I189A, S238A, and I241A MalE-VirF proteins exhibited successful first purifications and yields sufficient for *in vitro* testing (**Table 2.1**). However,

R192A, K193A, Y239A, and R242A mutants required multiple purification attempts with varying amounts of culture needed to purify enough usable protein for testing (**Table 2.1**). R192A proved very difficult to purify even with increasing the amount of culture grown to 2 L (yield of 0.4 mg/L culture). Due to this low yield and concentration, the R192A protein stock was concentrated to 21 μM for a higher testable concentration. Overall, all mutants were successfully purified for use in DNA-binding assays to test their activities.

Table 2.1: Concentration and yield of WT MalE-VirF and DBD mutant purifications. Purifications performed from more than 1 L (i.e., 2 L) of cell culture are indicated with an asterisk.

Protein	WT	I189A	R192A	K193A	S238A	Y239A	I241A	R242A
Conc. (μM)	57.5	47.9	6.7	22	21.2	47.6	43.8	20.0
mg/L culture	2.1*	2.8	0.4*	0.6*	1.2	0.5*	2.6	1.2

2.3.3 WT MalE-VirF and DBD mutants binding to pvirB

To test how WT MalE-VirF and the DBD alanine-mutants bound to DNA, an electrophoretic mobility shift assay was used (EMSA; **Figure 2.7**). The I189A, Y239A, I241A, and R242A mutants exhibited titratable binding shifts like WT MalE-VirF. Conversely, R192A and K193A were unable to bind to *pvirB* at the tested concentrations while S238A MalE-VirF presented a visible reduction in DNA-binding activity. The active DBD mutants bound to *pvirB* with similar affinities to WT according to their experimental K_D values and overlapping 95% confidence intervals (**Figure 2.8**). While S238A bound with similar affinity to WT, the saturation (B_{Max}) of the DNA probe was significantly lower than the other mutants and WT. In addition, at higher tested concentrations for WT and the mutants, a significant amount of *pvirB* probe can be seen in the wells (**Figure 2.7**). These data points are not included in the fits provided in **Figure 2.8.** Several factors could be contributing to this observation, but it is likely that these proteins aggregate at high concentrations (> ~10 μM). Precipitate does form in the wells when the reactions are loaded at these higher concentrations. This probe-trapping could be due to MalE-VirF on the surface of the aggregates binding specifically to *pvirB* since reactions containing denature (via

boiling) MalE-VirF at the same concentrations do not trap the probe in the well (data not shown). Overall, the probe-trapping due to protein aggregation confounded the data analysis and binding fits of the EMSA results.

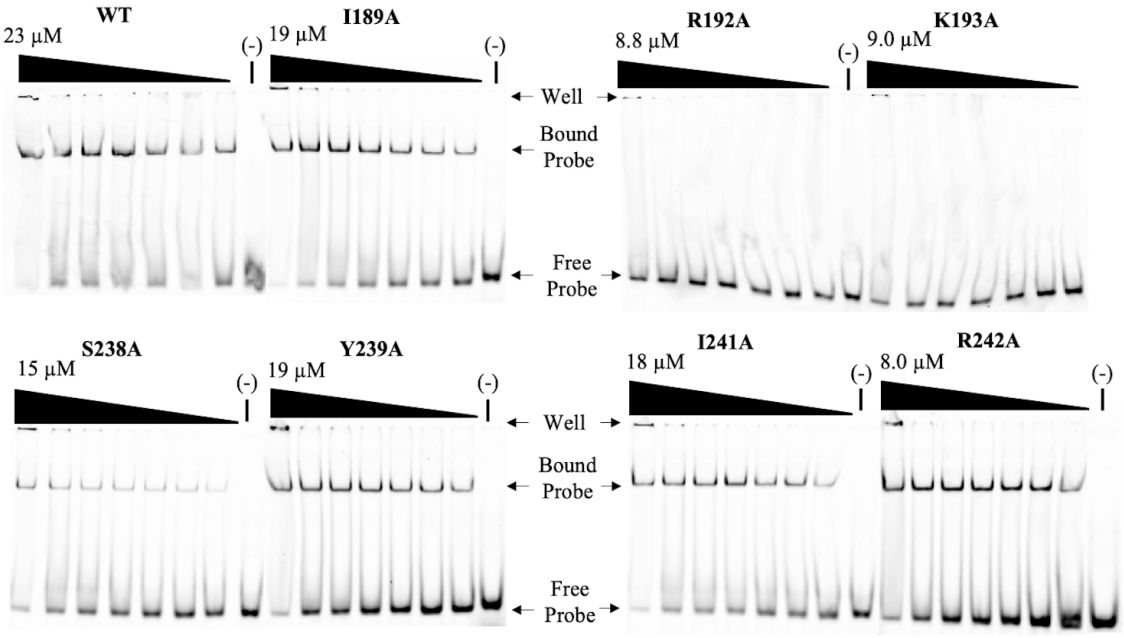
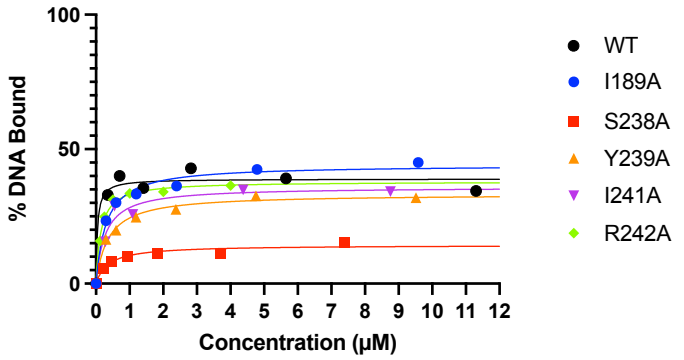


Figure 2.7: EMSAs testing the binding of WT MalE-VirF and DBD mutants for *pvirB*. Each protein was tested in two-fold serial dilutions from the starting concentration shown on the left of each gel. For gels testing the titrations of S238A, I241A, and R242A against *pvirB*, the negative control (-) was cropped and placed to the right of the titrations for presentation consistency. One negative control was used for R192A and K193A titrations as they were tested together on one gel.

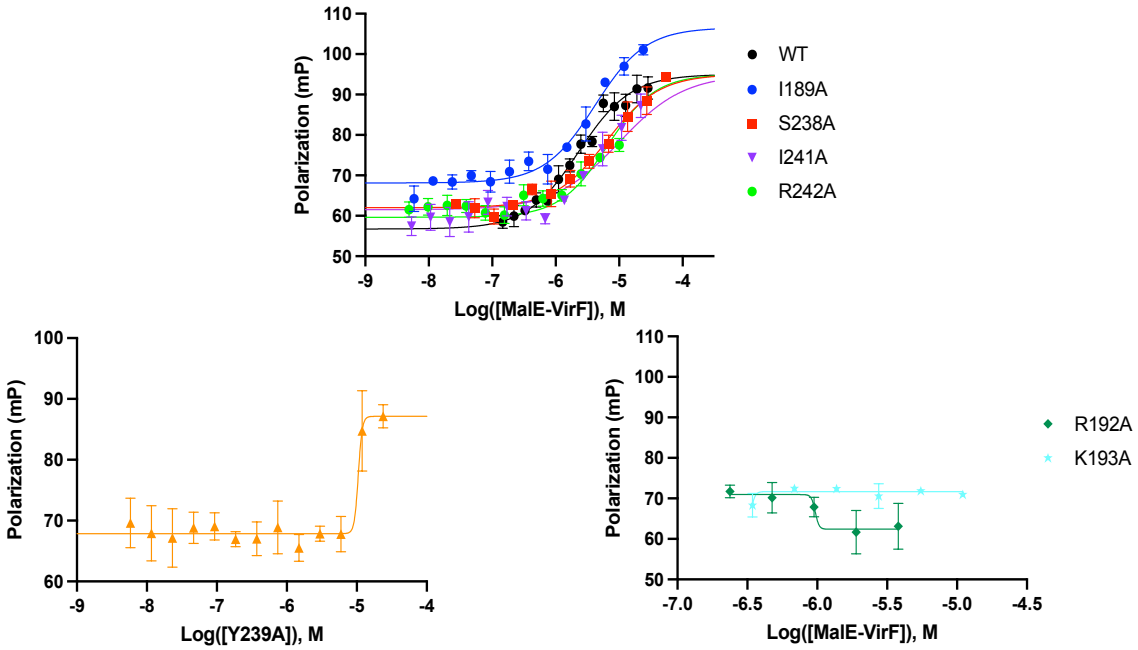


Protein	WT	I189A	S238A	Y239A	I241A	R242A
K_D, nM	43	300	370	360	270	140
(95%CI)	(0 - 180)	(180 - 460)	(160 - 760)	(250 - 510)	(79 - 620)	(100 - 200)
B_{Max}, %	40	44	14	33	36	38
(95%CI)	(34 - 44)	(40 - 48)	(12 - 17)	(31 - 36)	(29 - 43)	(35 - 41)

Figure 2.8: Non-linear regression analysis of WT MalE-VirF and DBD mutants I189A, S238A, Y239A, I241A, and R242A, binding to *pvirB* in the EMSA. Binding affinities (K_D) and probe saturations (B_{Max}) are presented in the table provided below. For each affinity and saturation value, 95 % confidence intervals (95 % CI) are presented in parentheses. The data points are single determinations from image density analyses of the EMSA gels.

To better quantitate the binding affinities of WT MalE-VirF and the DNA-binding mutants, fluorescence polarization (FP) was used (**Figure 2.9**). After the data was fitted to a non-linear regression (sigmoidal 4PL fit) with Prism, WT MalE-VirF exhibited a K_D of 2.3 µM for *pvirB*. I189A (4.2 µM) presented an insignificant change in affinity for *pvirB* compared to WT since these values contained overlapping 95% confidence intervals. However, I241A (6.5 µM) and R242A (10.1 µM), which showed qualitatively equivalent binding in the EMSA (**Figures 2.7 and 2.8**), exhibited 3 to 4-fold reduced binding relative to WT, respectively. The S238A mutant presented a 3-fold weaker affinity for *pvirB* (6.3 µM) compared to WT. This reduced affinity is consistent with the EMSA of S238A MalE-VirF where there was visually weaker binding and probe saturation compared to WT and the other active DNA-binding mutants. Additionally, the Hill slopes for WT and all mutants in the FP assay ranged from 0.7 to 1.5 (**Figure 2.9**). This indicates that the proteins are not likely aggregating (Hill slope >>1) and are likely binding with a 1:1 ratio

of protein to DNA. Currently, it is not known if MalE-VirF must dimerize to bind to DNA promoters, so it is unknown if this 1:1 binding ratio is with a monomer or preformed dimer. Unfortunately, Y239A did not present optimal binding to *pvirB* in the FP unlike in the EMSA (**Figure 2.9**). Y239A MalE-VirF bound to *pvirB* with an affinity of 10.5 µM but presented a very high Hill slope of 16 indicating significant aggregation was occurring in the FP assay. But when we tested for Y239A MalE-VirF aggregation in a turbidity assay, no aggregation was observed.



Protein	WT	I189A	S238A	Y239A	I241A	R242A
K_D, µM	2.3	4.2	6.3	10.5	6.5	10.1
(95%CI)	(1.7 - 2.9)	(2.7 - 18)	(5.1 – 7.9)	(ND)	(5.2 – 8.3)	(7.7 - 15)
Hill Slope	1.1	1.1	1.1	16	1.2	0.8
(95%CI)	(0.9 - 1.3)	(0.6 - 1.8)	(0.9 - 1.3)	(ND)	(0.9 - 1.5)	(0.6 - 1.1)
Span, mP	38	38	33	19	35	34

Figure 2.9: Non-linear regression analysis of WT MalE-VirF and DBD mutants binding to a fluorescein-labeled *virB* promoter in the FP assay. Sigmoidal curves were generated to calculate the binding affinities (K_D), Hill slopes, and spans (mP differences between the maximums and minimums of the fits) of all mutants and WT, except R192A and K193A, are presented in the table below. For each affinity and Hill slope, 95 % confidence intervals (95 % CI) are presented in parentheses (ND = not determined). Every protein concentration data point is tested in triplicate and error bars are included indicating the standard deviation.

When R192A and K193A MalE-VirF were both tested in the FP assay, no binding was observed (**Figure 2.9**). Given the loss of observed DNA-binding for the R192A and K193A mutants in both DNA-binding assays, it was possible that these proteins were inactive due to misfolding. To determine if they were misfolded, circular dichroism (CD) spectra were obtained and compared to WT. S238A MalE-VirF was also included to observe CD spectra for a weaker binding mutant. All spectra were like that of WT suggesting that the proteins were properly folded (**Figure 2.10**). It is entirely possible that there is local misfolding adjacent to the mutation sites, but we did not detect significant, global misfolding of these proteins.

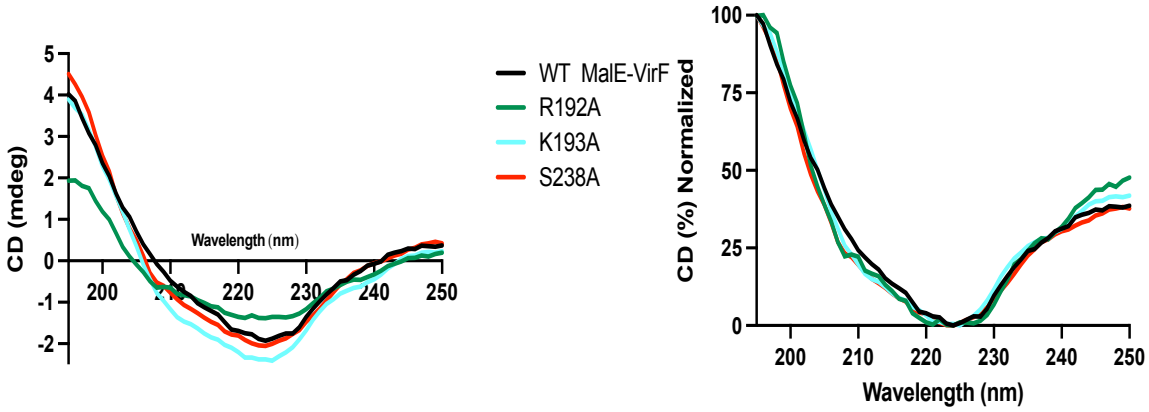


Figure 2.10: CD spectra of WT, R192A, K193A, and S238A MalE-VirF. Spectra for inactive mutants R192A and K193A as well as S238A (reduced activity) mutants were compared to WT MalE-VirF. Normalized data is presented in the curves to the right. Protein samples were exchanged into 10 mM NaH_2PO_4 (pH 7.5) and tested at 1 µM. Data were corrected by subtracting the contributions of the sodium phosphate buffer.

Based on footprinting analysis, VirF has been shown to have two binding sites on *pvirB*(5, 7). Additionally, VirF belongs to a subclass of AraC family proteins that are thought to dimerize(2). To better elucidate the MalE-VirF•*pvirB* binding interaction, the two VirF binding sites on *pvirB* were independently scrambled (sequence of promoters can be found in **Figure 2.11**) and DNA-binding was tested with WT MalE-VirF in the EMSA (**Figure 2.12**). WT MalE-VirF bound nearly identically to both the wild-type (*pvirB*) and binding site 2 scrambled probes (BS2 Scram). However, when binding site 1 was scrambled (BS1 Scram), a significant reduction in

binding was observed. At concentrations higher than 6 µM, binding was present albeit at a weaker density than the other promoters. In addition, a second binding shift was not visualized which could have indicated either two binding events, or dimerization, when MalE-VirF interacted with any of the *pvirB* promoters. The results from this EMSA indicate that binding site 1 is a higher affinity site for MalE-VirF whereas binding to site 2 may be significantly facilitated by MalE-VirF dimerization.

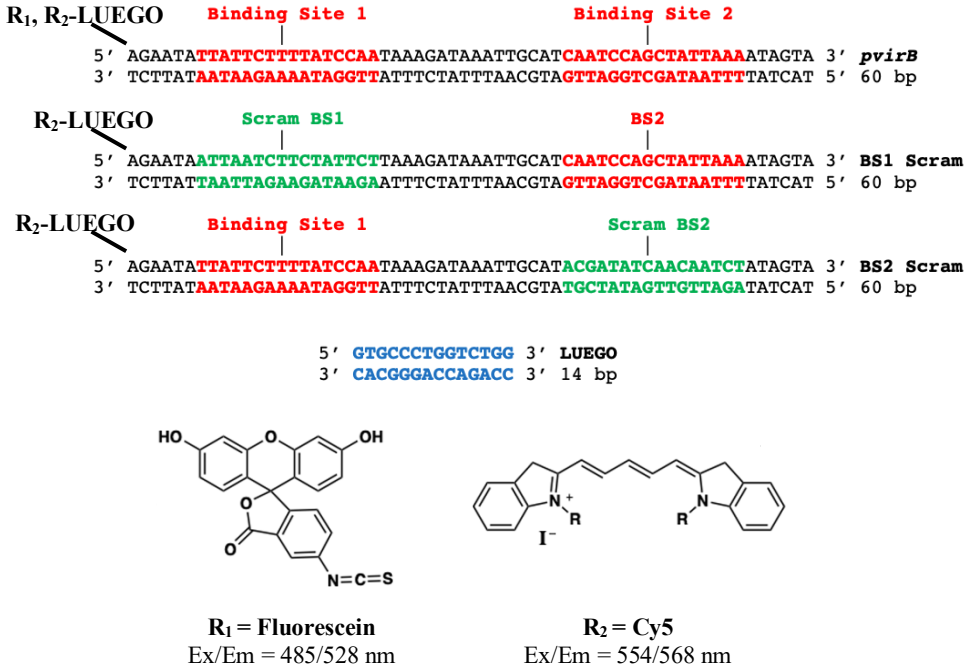


Figure 2.11: Wildtype and scrambled *pvirB* probes used in the EMSA and FP. DNA sequences for *pvirB* including VirF binding sites labeled in red (green for scrambled sites). Cy5 and fluorescein fluorophores used in the assays were placed at the 5' end of the top strands of all probes (EMSA probes contain a 5'-Cy5 label on the LUEGO sequence which resides on the 5' end of the promoter sequence, i.e., R_2-LUEGO). LUEGO stands for labeled universal electrophoretic gel shift oligonucleotide.

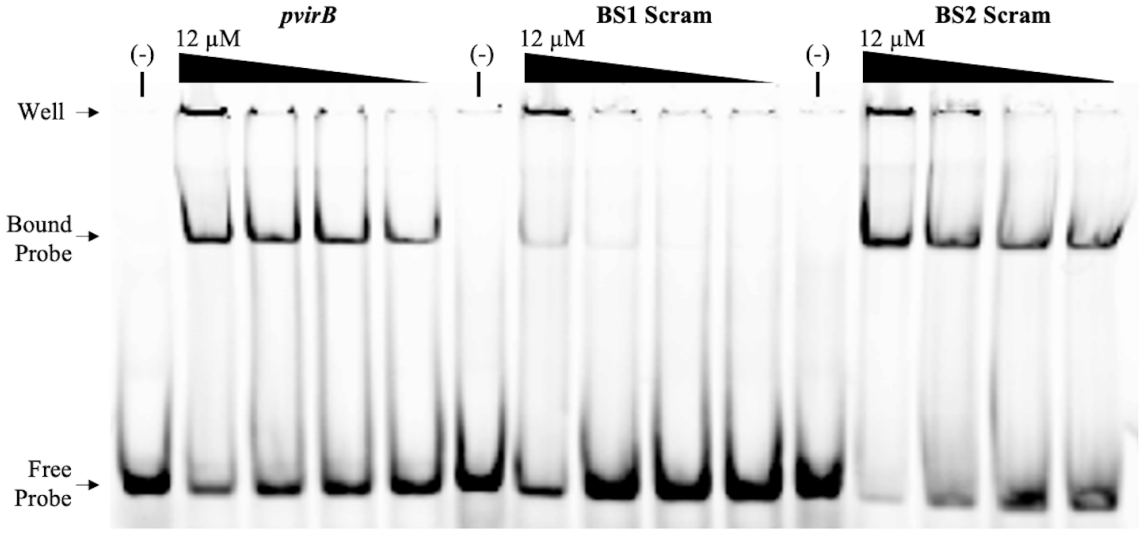


Figure 2.12: EMSA of WT MalE-VirF binding to wild-type (*pvirB*), binding site 1 scrambled (BS1 Scram), and binding site 2 scrambled (BS2 Scram) *virB* promoters. MalE-VirF was tested in two-fold serial dilutions from the starting concentration shown on the left of each gel

2.3.4 WT MalE-VirF and DBD mutants binding to picsA and prnaG

Published *in vitro* DNA-binding studies have focused on the MalE-VirF•*pvirB* interaction for the discovery of DNA-binding inhibitors or determination of binding affinities(6, 7). DNA-binding studies have not been performed for either the VirF•*picsA* or VirF•*prnaG* interactions. Using DNA footprinting analysis, Tran et al. identified DNase I protected sites within the *icsA* regulatory region suggesting the locations of the VirF binding sites within the *icsA* and *rnaG* promoters (4). The *icsA* (*picsA*) and *rnaG* (*prnaG*) promoters used in the EMSA and FP were designed based on this study. Synthetic overlapping primers covering the entire promoter sequences (**Figure 2.13**) were designed to be annealed and amplified in a PCR using Taq DNA polymerase (**Tables 2.4**). The resultant A-tailed PCR products were TOPO cloned into pCR™2.1-TOPO® vectors. Upon verification of the promoters' presence in these plasmids, PCR (**Tables 2.5)** was performed with Cy5-labeled DNA oligonucleotide primers to obtain the DNA probes used in the EMSA. The DNA promoter probes can be found in **Figure 2.13**.

Figure 2.13: EMSA *picsA* and *prnaG* probes. DNA sequences for *picsA and prnaG* including VirF binding sites labeled in red. EMSA probes contain a 5'-Cy5 label which reside on the top strand's 5' end of the promoter sequence.

WT MalE-VirF was titrated against both *picsA* and *prnaG* probes in an EMSA (**Figure 2.14**). Unlike in the MalE-VirF•*pvirB* binding interaction (**Figure 2.6**), multiple shifts were visible for both promoters. Three distinct shifts occurred when WT MalE-VirF was titrated against *picsA*. At MalE-VirF concentrations less than 4 µM, only one shift appears (S1). At higher concentrations, second and third shifts (S2 and S3) appear. Additionally, from 8 to 16 µM, S1 loses density which gives rise to higher probe densities in the well and at S2 and S3. This indicates that WT MalE-VirF interacts with *picsA* through multiple, consecutive binding events (Free → S1 → S2 → S3), or potentially through a dimerization event. Similarly, two shifts are observed when WT interacts with *prnaG*. S1 occurs at concentrations less than 2 µM, with the occurrence of S2 at concentrations greater than 2 µM. We also see the disappearance of S1 at 16 µM. Unlike in *picsA*, S3 was not visible with promoter at the tested concentrations. However, binding to these promoters was visibly weaker than compared to with *pvirB* (**Figure 2.7**).

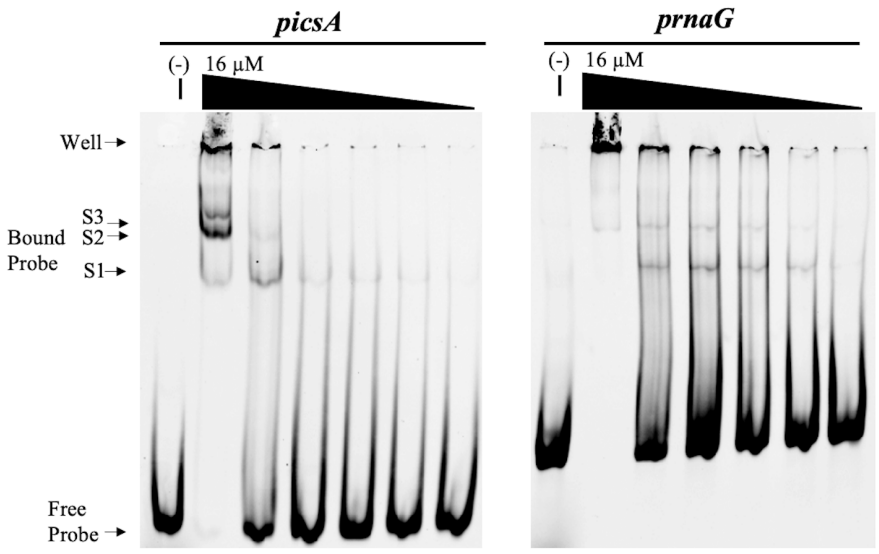


Figure 2.14: EMSA of WT MalE-VirF binding to *picsA* and *prnaG*. MalE-VirF was tested in two-fold serial dilutions from the starting concentration shown on the left of each gel. Binding shifts are indicated by the order in which they occur on the gel (S1 = first shift, S2 = second shift, S3 = third shift).

Probe saturation (%Bound) for each shift was calculated for each promoter. However, affinities for each binding event were unable to be determined due to the complexity of the binding schemes with these promoters. Similar to the WT MalE-VirF titration against *pvirB*, we see a significant amount of both probes stuck in the well at concentrations greater than 8 µM which confounded the data analysis. A kinetic simulation was performed for the *prnaG* EMSA (**Figure 2.15**) as described in Methods. The simulated curves for each species matched the experimental data fairly well and yielded values for the constants as follows: $k_1 = 0.05\ \mu M^{-1}$, $k_2 = 1.0\ \mu M^{-1}$, and $k_3 = 0.7\ \mu M^{-1}$. It is prudent to avoid over-interpreting these simulated values. However, they are consistent with a sequential binding scheme where the initial, 1:1 binding event is driven solely by the interaction of the MalE-VirF DNA binding domain with the DNA promoter and subsequent binding events have increased affinity due to the additional DNA-binding and/or dimerization interactions.

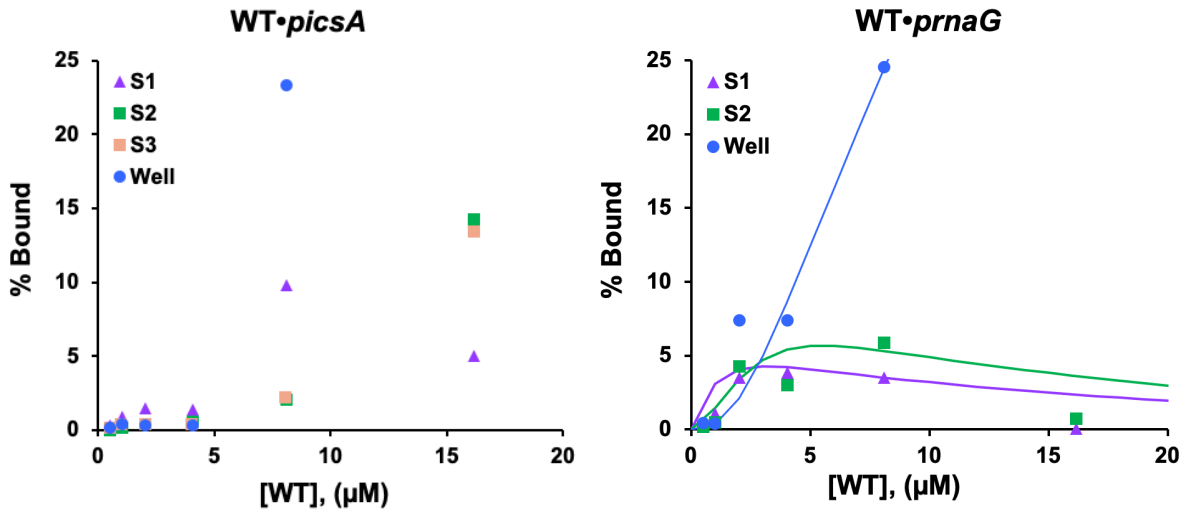


Figure 2.15: Plots of the WT•*picsA* and WT•*prnaG* DNA-binding interactions as studied in the EMSA (**Figure 2.14**). Plot of %Bound for *picsA* (S1, S2, S3, Well) and for *prnaG* at each binding shift (S1, S2, Well) are included. For *prnaG*, the curves are a simulation of consecutive binding kinetics. The plots for *picsA* could not be simulated.

In an attempt to identify these affinities, FP assays were used with promoters containing a 5' fluorescein rather than a 5' Cy5 fluorophore on its top strand. No binding curves were observed after analyzing the data associated with the FP assay. This is likely due to increased promoter size relative to MalE-VirF (70.8 kDa for *picsA*, 94.3 kDa for *prnaG*, and 73 kDa for WT MalE-VirF). Previously, we observed binding curves in FP with *pvirB* (**Figure 2.9**) since this promoter probe (37.7 kDa) was half the molecular weight of MalE-VirF. Because of the insignificant difference in molecular weight between the probes and MalE-VirF, this significantly reduced the assay sensitivity. FP assays measure polarization of light due to changes in molecular weight (i.e., binding of MalE-VirF to DNA increases polarization), so the molecular weight of the *icsA* and *rnaG* promoters did not change as significantly upon MalE-VirF binding compared to the *virB* promoter thus leading to a lower sensitivity. Furthermore, the FP assay would likely not be able to successfully observe the separate binding events as exhibited in the EMSA.

Similar to the *pvirB* binding studies, the binding sites within the *icsA* and *rnaG* promoters were independently scrambled and EMSAs testing WT MalE-VirF for these modified promoters

were performed. The sequences for these promoters can be found in **Figure 2.16B**. Compared to the wildtype *icsA* promoter, only two binding events (S1 and S2) were visible when either binding site 1 (BS1 Scram) or 2 (BS2 Scram) were scrambled. S1 shifts were not observed until > 11 μM WT MalE-VirF for BS1 Scram while > 6 μM for the BS2 Scram promoter (**Figure 2.16A**). While S3 was present with the wildtype promoter, neither scrambled promoter exhibited that shift at the same concentration (23 μM). This provides evidence that WT MalE-VirF may require both sites to form a higher oligomeric complex with the promoter but scrambling the sites does not significantly reduce binding activity. This was evident with WT MalE-VirF and *pvirB* BS1 Scram (**Figure 2.12**). Overall, it appears that BS1 is higher affinity site than BS2 within *picsA*. Conversely, WT exhibited a different binding scheme with the scrambled *prnaG* probes. When binding site 4 was scrambled (BS4 Scram), only one binding shift (S1) appeared which was distinct from the two shifts (S1 and S2) for both wildtype *prnaG* and BS3 Scram. Based on these observations, it is likely that binding site 4 is required for the development of a higher oligomeric complex between MalE-VirF and *prnaG*. Binding site 3 of *prnaG* may not be essential for this complex but this may only be true for the limited concentrations tested.

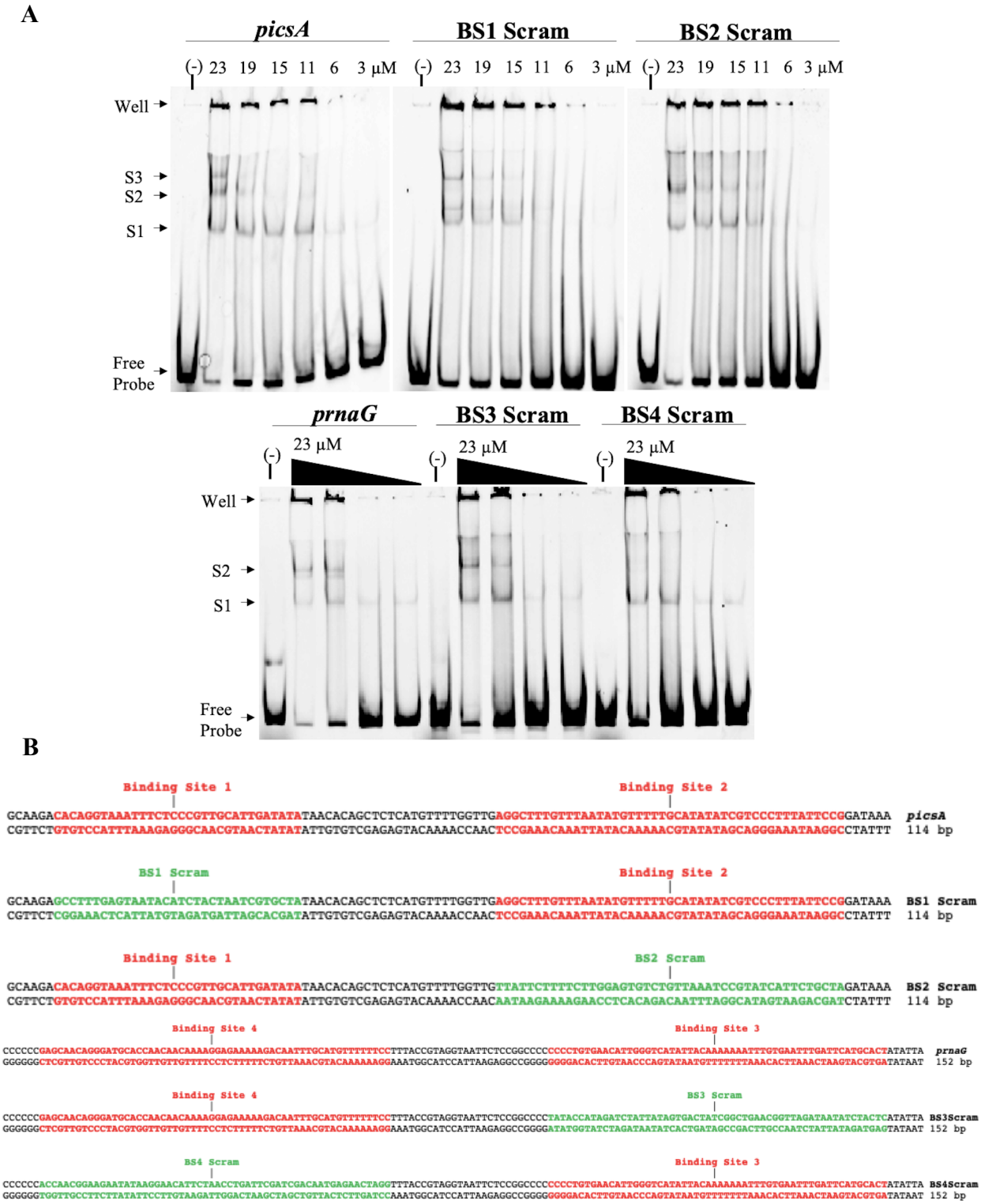
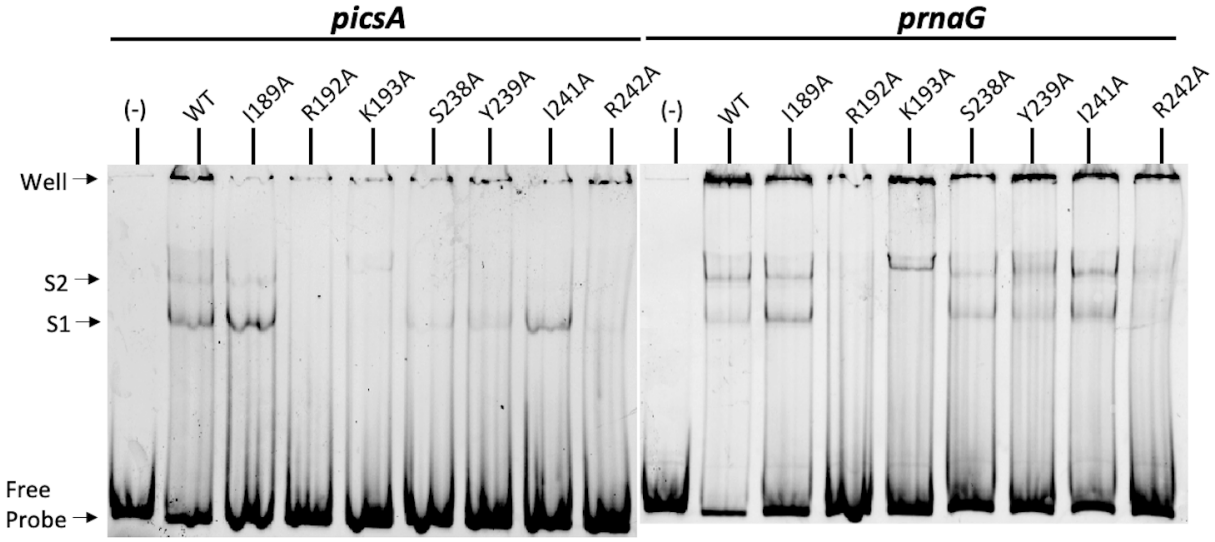


Figure 2.16: EMSAs of WT MalE-VirF for scrambled *picsA* and *prnaG* probes. A) EMSA testing the DNA-binding activity of WT MalE-VirF for wild-type and scrambled *picsA* and *prnaG* DNA probes. MalE-VirF was tested over a range of concentrations (23-2.8 µM) against 40 nM DNA promoter probe. A more comprehensive titration was included when testing with *picsA*. B) The sequences for the scrambled DNA promoters are included.

Following binding assays with WT, we tested the seven DNA-binding alanine mutants against *picsA* and *prnaG* to determine their differential activities for these promoters (**Figure 2.17;** titrations for active mutants against both promoters can be found in **Figure 2.18**). Since obtaining binding affinities for the binding interactions proved difficult, proteins were tested at or near 10 μM with both promoters so binding activities could be compared for the mutants and WT MalE-VirF. I189A was the only mutant that showed an insignificant change in binding for both promoters relative to WT. When the total %Bound values for both binding shifts (% (S1 + S2)) were calculated, I189A (15% for *picsA* and 14% for *prnaG*) presented slightly weaker binding for both promoters compared to WT (22% for *picsA* and 17% for *prnaG*) specifically for S2 when tested against *picsA* (%S2; 2.8% for I189A and 6.7% for WT). The S238A, Y239A, and I241A mutants all showed reductions in *picsA*-binding where S2 shifts were not observed. When these mutants were tested against *prnaG*, all presented S2's with lower % (S1 + S2) than WT. R242A exhibited significantly reduced binding for both *picsA* and *prnaG* (2.7% and 5.9% % (S1 + S2) for *picsA* and *prnaG*, respectively). R242A exhibited a barely visible S1 when tested against *picsA* and showed significantly less dense S1 and S2 shifts than WT in the *prnaG* EMSA. This mirrors the significant decrease in affinity of R242A MalE-VirF for *pvirB* in the FP (**Figure 2.7**).

R192A and K193A exhibited differential binding for these promoters compared to *pvirB*. Like with *pvirB*, R192A was unable to bind *picsA* and *prnaG* at 9 μM (highest testable concentration). Interestingly, K193A bound to both larger promoters but at a shift slightly above S2 (termed "S2" in **Figure 2.17).** When titrated against both promoters, only one shift appears on (**Figure 2.18**). This is significantly different from the other promoters where consecutive binding occurred. K193A might be capable of binding to these promoters only as a dimer or a higher oligomeric complex. K193A elutes at approximately the same volume and fractions as WT upon

purification with the Superdex 200 10/300 column (**Figure S2.1**) and it was shown that WT MalE-VirF exists as a monomer in solution (6), so it is unlikely that K193A MalE-VirF exists as a dimer following purification.



picsA								
Protein	**WT**	**I189A**	**R192A**	**K193A**	**S238A**	**Y239A**	**I241A**	**R242A**
µM	10	10	9	8.8	10	10	10	8
%S2	6.7	2.8	0	1.4*	1.6	1.9	1.1	1.4
%S1	15	12	0	0	1.6	2.0	5.3	1.3
% (S1 + S2)	22	15	0	1.4	3.2	3.9	6.4	2.7
prnaG								
%S2	7.5	5.8	0	7.8*	3.9	6.7	6.8	3.6
%S1	9.4	8.3	0	0	4.6	4.4	8.1	2.3
% (S1 + S2)	17	14	0	7.8	8.5	11	15	5.9

Figure 2.17: EMSAs of WT MalE-VirF and DBD mutants binding to *picsA* and *prnaG*. All proteins were tested at 10 µM or at the highest testable concentration ([R192A] = 9 µM, [K193A] = 8.8 µM, [R242A] = 8 µM) against 40 nM promoter. Tables of the probe saturation (% bound values for S1, S2, and (S1 + S2)) for each protein were calculated using ImageJ and are presented below. (*S2 for K193A does not align with the S2s for WT and the other mutants but is labeled thus for consistency)

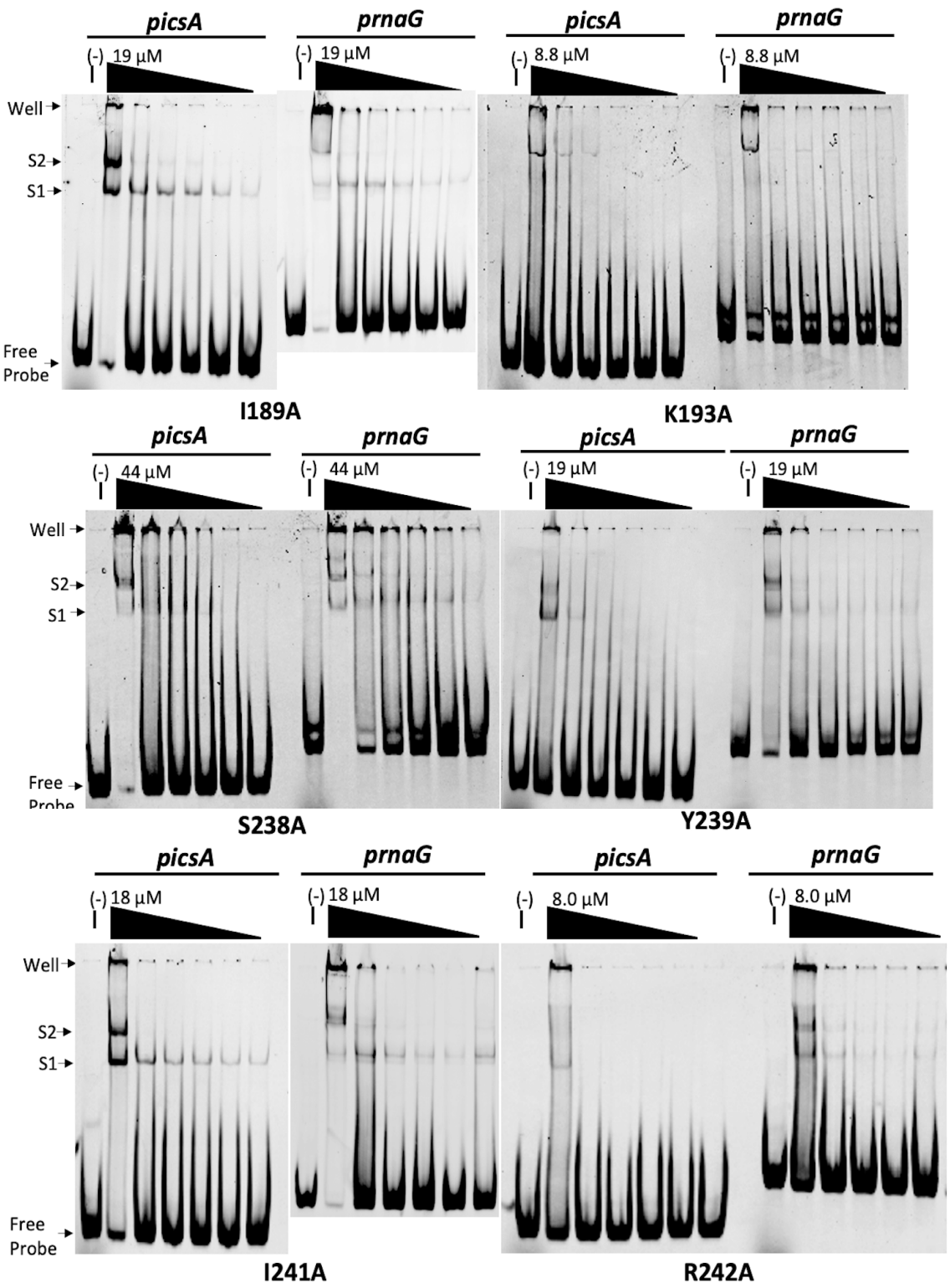


Figure 2.18: EMSAs of all active DBD mutants for *picsA* and *prnaG*. EMSAs testing binding of I189A and I241A for *picsA* and *prnaG* were performed on separate gels but were cropped and pasted to scale to match the other gels.

2.4 Discussion and Conclusions

2.4.1 Homology Modeling and Alignment with MarA and GadX

Since no VirF structure is available but some homologs have been crystallized, homology modeling was performed. Homology modeling is an effective structural prediction method and is often used when methods such as X-ray crystallography or NMR solution studies are unsuccessful. Two homologous structures from *E. coli*, GadX (PDB 3MKL) and MarA (PDB 1BL0) were selected as templates for homology modeling of VirF in free and DNA-bound conformations. The GadX DBD is more similar to VirF (29% sequence identity and 0.76 GMQE score) than to MarA (19% sequence identity and 0.70 GMQE score). In its crystal structure, GadX is free in solution, unbound to DNA. In contrast, MarA is bound to its cognate promoter, *marRAB*(10). Upon superimposing the two VirF homology models, the predicted structures displayed slightly different orientations of the HTH motifs, particularly helix-3 (**Figure 2.3C**). This could represent DNA-bound (MarA-based; **Figure 2.3A**) and unbound (GadX-based; **Figure 2.3B**) conformations for the VirF DBD. When superimposed with the MarA crystal structure, the MarA-based VirF homology model resided within the major grooves of *marRAB* allowing us to envision what the VirF•*pvirB* interaction would look like (**Figure 2.3A**). Although this might be an artifact of the homology model and superimposition with MarA, we cannot definitively state that VirF adopts this conformation when DNA-bound (i.e., interacts with its cognate promoters via both HTH unlike *E. coli* Rob). While the function of both HTH motifs has been somewhat debated (1, 24, 25), conformational variations for these motifs could be required for VirF, and other AraC-family proteins, to bind to their cognate promoters.

2.4.2 Elaboration of the MalE-VirF•pvirB DNA-binding Interaction

The structure of MarA served two purposes in our studies with VirF. First, MarA provided a structural template from which a VirF homology model could be generated. Secondly, the MarA was used to analyze crucial interactions of the MarA DBD with *marRAB* so that this analysis could be translated to VirF. Upon examination of the MarA crystal structure and an alignment with the VirF DBD (**Figure 2.4)**, seven residues were selected for alanine-scanning mutagenesis and subsequent *in vitro* testing. The seven mutations independently placed within the VirF DBD were I189A, R192A, K193A, S238A, Y239A, I241A, and R242A. Initial testing of WT MalE-VirF and these mutants was performed with the *virB* promoter. In the FP assay, WT MalE-VirF exhibited a micromolar affinity of 2.3 μM for *pvirB* (**Figure 2.9**). Non-conservative mutations to residues capable of hydrogen bonding and/or electrostatic interactions with *pvirB* (i.e., R192, K193, and R242) showed reductions or loss of DNA-binding activity. While R192 and K193 MalE-VirF alanine-mutants led to a complete loss of binding activity, R242A in MalE-VirF exhibited a > 4-fold change in binding relative to WT MalE-VirF but retained activity in both the EMSA and FP (**Figure 2.7 and 2.9**).

Few studies have characterized the VirF DNA-binding activity, specifically which amino acids contribute to promoter binding. Porter and Dorman performed mutagenesis on *virF* to test WT and mutant VirF activation of β-galactosidase activity via a *mxiC-lacZ* fusion(15). The results from this study translate to our studies since, in virulent *S. flexneri*, VirF activates transcription of *virB* (via binding to the *virB* promoter) which in turn positively regulates *mxiC* transcription and expression. This study identified mutations to the VirF DBD that significantly affected downstream *mxiC-lacZ* expression and activity but follow-up on these mutants was not performed to identify their direct effect on DNA-binding. Three mutations correlated between Porter and

Dorman and this study: K193A, Y239F, and I241N. All three mutations lead to a loss β-galactosidase assay *in vivo*(15). Consistent with Porter and Dorman, the K193A mutation led to a loss of DNA-binding activity when MalE-VirF was tested with *pvirB* in the EMSA and FP (**Figures 2.7 and 2.9**). Although the testable concentrations were limited to the low yield of the K193A MalE-VirF protein, there was still a significant change in DNA-binding activity compared to WT and the other active DNA-binding mutants. Overall, this is not surprising since losing lysine, a positively charged amino acid, could have deleterious effects on binding simply via loss of electrostatic interactions between the terminal charged amine of lysine and the *pvirB* phosphate backbone. In addition, disruptions of high energy hydrogen-bonds between the promoter bases and the lysine side chain could lead to a significant loss in binding. As expected, a loss of DNA-binding was observed with R192A MalE-VirF and was assumed to be for similar reasons as for K193A.

Conversely, DNA-binding activity for Y239A and I241A was observed in both the FP and EMSA (**Figures 2.7 and 2.9**). The corresponding VirF mutations in Porter and Dorman, Y239F and I241N, exhibited a loss of *in vivo* activity(15). Porter and Dorman observed expression of Y239F VirF in a western blot so it is possible that *in vivo* aggregation or misfolding of Y239F VirF could be occurring thus leading to a reduction in β-galactosidase activity. Our Y239A MalE-VirF displayed binding similar to WT MalE-VirF in the EMSA, but the FP assay revealed that aggregation (Hill slope = 16) is likely occurring *in vitro*. These findings may provide evidence that while mutations to Y239 lead to aggregation, it does not affect DNA-binding but significantly affect how VirF activates gene transcription. Despite the significant aggregation, we suspect that Y239 participates via hydrophobic or Van der Waals (VDW) interactions with *pvirB* or potentially through a weak/suboptimal hydrogen bond. Again, Porter and Dorman saw a loss of *in vivo*

activity with their I241N VirF(15). The I241N mutation introduced a non-conservative change which significantly altered both the polarity and hydrophobicity at that residue. I241N VirF expression was observed in a western blot, so it is likely that a non-conservative change at this position led to a decrease in β-galactosidase activity. However, our corresponding I241A MalE-VirF mutant presented a modest reduction in binding compared to WT MalE-VirF in the FP assay. I241A MalE-VirF incorporates a hydrophobically conservative change at that residue but significantly reduces the steric bulk. The modest change in DNA-binding activity of I241A MalE-VirF compared to WT suggests that I241 interacts with *pvirB* mostly through weak hydrophobic or VDW interactions. These interactions involve less energy compared to hydrogen bonding or electrostatic interactions which may account for this modest change in DNA-binding activity.

I189A, S238A, and R242A MalE-VirF had not been previously studied. Like I241A, I189A introduced a hydrophobically conservative change to the protein and we did not observe a significant difference in DNA-binding activity. This indicates that I189 interacts with *pvirB* most likely through weak hydrophobic or VDW interactions. S238A MalE-VirF displayed a decrease in binding affinity for *pvirB* in the FP assay and a qualitative reduction in probe saturation in the EMSA (**Figures 2.7 and 2.9**). The loss of the serine hydroxyl led to a ~3-fold reduction of DNA-binding activity relative to WT. This indicates that S238 participates via weak hydrophobic or VDW interactions despite containing a hydroxyl which could contribute to a hydrogen-bonding interaction with the bases within *pvirB*. Losing a high energy interaction like a hydrogen-bonding interaction would likely lead to a >10-fold loss in DNA-binding affinity, so it is unlikely that S238 contributes a hydrogen-bonding interaction with *pvirB*, but we cannot rule out a suboptimal hydrogen-bond.

Lastly, R242A MalE-VirF bound to *pvirB* ~4-fold weaker than WT. Initially, it was hypothesized that the loss of arginine would elicit a significant (> 10-fold) reduction in DNA-binding activity via loss of assumed electrostatic interactions like R192A and K193A. On the contrary, the alanine substitution resulted in a modest loss of affinity. Therefore, it seems unlikely that the R242 side chain contributes to the *pvirB* binding interaction via high energy interactions such as electrostatics or hydrogen-bonding. However, it cannot be ruled out that R242 interacts with *pvirB* through suboptimal/weak hydrogen bond or hydrophobic and VDW interactions. These proposed interactions based on the experimental data can be found in **Figure 2.19**. Overall, if alanine mutations to any of these residues, except I189A, arose in *S. flexneri*, the observed decreases in DNA-binding activity could have a significant effect on virulence. It has been shown that VirF needs to reach a threshold of activity (~40% relative to *Shigella* WT grown at 37 °C) to activate the virulence pathway(26), so these mutations could be sufficient to be detrimental to the bacteria.

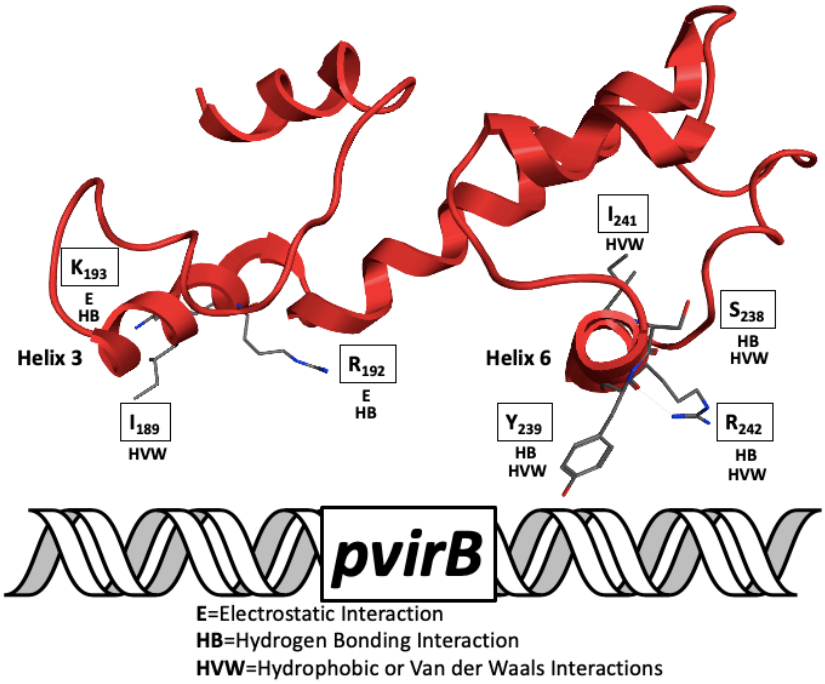


Figure 2.19: Diagram of studied MalE-VirF amino acids and their potential binding interactions with *pvirB* based on the experimental data. The diagram uses the MarA-based homology model (**Figure 2.3A**).

2.4.3 Comparing VirF Binding Schemes with the virB, icsA, and *rnaG Promoters*

VirF is known to bind three distinct promoters which regulate transcription of two downstream genes (*virB* and *icsA*) and an antisense RNA, RnaG(4, 5). VirF binding activity was expanded to these other two promoters: *picsA* and *prnaG*. Significant differences arose between MalE-VirF•*pvirB* and the •*picsA* and •*prnaG* binding interactions. In an EMSA, three consecutive binding shifts (S1, S2, and S3) appeared when MalE-VirF bound to *picsA* in the EMSA, but only two binding shifts (S1 and S2) were observed with *prnaG* (**Figure 2.14**). While the EMSA provided valuable qualitative evidence for how WT MalE-VirF interacts with these promoters, affinity calculations were unsuccessful (**Figure 2.15**). When comparing the three DNA promoters, *picsA* (114 bp) and *prnaG* (152 bp) are much longer than *pvirB* (60 bp) and contain much larger VirF binding sites (**Figures 2.11 and 2.13;** Comparison of all three promoters, side by side, can be found in **Figure S2.3**). Identified from footprinting analyses, the DNase I-protected VirF binding sites (**Figure 2.20A)** on these promoters varied from 100 bp (*pvirB*), 32-45 bp (*picsA*), or 56-58 bp (*prnaG*)(4, 5). Within the 100 bp protected region of the *virB* promoter, Koppolu et al. identified two 16 bp VirF binding sites(7). Based on the MarA•*marRAB* crystal structure (PDB 1BL0), MarA occupies a region of approximately 15 base pairs on the *marRAB* promoter(10). As we have performed homology models and alignments with MarA, it is highly probable that VirF binds to a sequence of similar length within its promoters as a monomer. In addition, due to the size of the VirF binding sites within *picsA* and *prnaG*, more than one binding event within those protected sites cannot be ruled out. Overall, the consecutive binding in the EMSA and probe comparisons provide suggestive evidence that VirF binds to these promoters via multiple binding events possibly facilitated by dimerization.

When testing the scrambled *picsA* and *prnaG* promoters in the EMSA (**Figure 2.16**), inconsistent binding schemes were observed. For both scrambled *picsA* probes, only two shifts were observed at 23 μM MalE-VirF compared to the three shifts observed with the wildtype promoter. This could indicate that VirF forms a higher oligomeric complex with *picsA* only in the presence of both VirF binding sites and at higher concentrations. Conversely, only one binding shift was seen when VirF binding site 4 (BS4 Scram) was scrambled in *prnaG* whereas two shifts were present with the BS3 Scram promoter like wildtype *prnaG*. This is inconsistent with the analysis presented in Tran et al. in which they identified VirF binding site 3 as the high affinity site(4). While it is possible that both binding sites in *prnaG* are needed for the formation of a higher oligomeric complex, it also appears likely that the binding sites were misidentified based on the footprinting assay (**Figure 2.20A**). Specifically for *prnaG*, VirF binding to site 3 (III; VirF•*icsA* Regulon) is clear in the left panel but when another primer was used for the extension (right panel), the protected site is less defined. This may be due to inconsistent loading in the gel between the different lanes or other mechanical errors. In addition, the other protected sites are not clear within the *icsA* regulatory region. In contrast, the protected site on the *virB* promoter (**Figure 2.20A**) is far clearer and indicates that MalE-VirF protects a stretch of ~100 nucleotides within the promoter region(5). Although the DNA-binding assays presented herein confirm VirF binds to both *picsA* and *prnaG* identified by Tran et al., the inconsistencies observed in our EMSAs testing the scrambled promoters and the unclear binding sites identified in Tran et al. make it difficult to draw too many conclusions from this data.

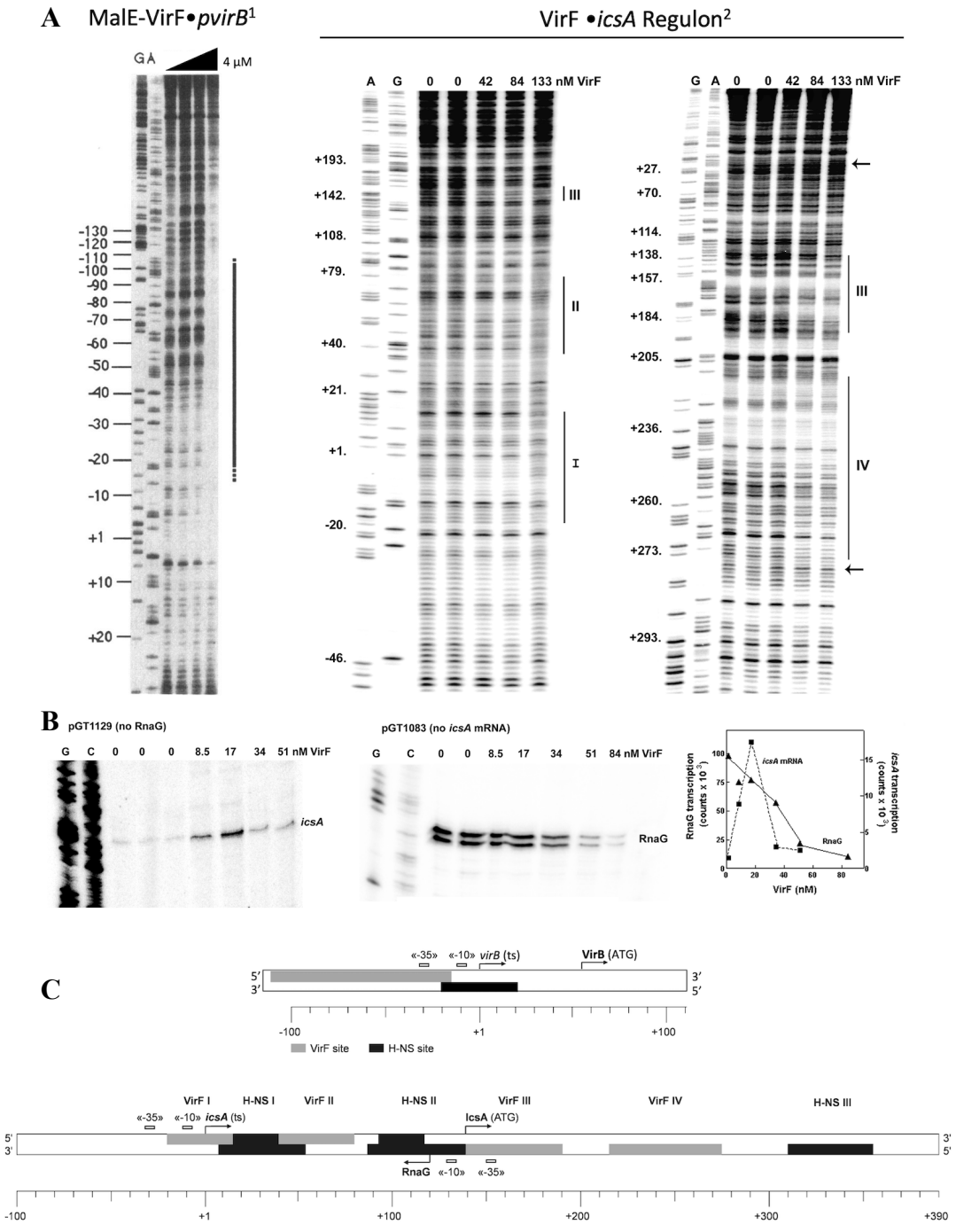


Figure 2.20: Comparison of VirF binding sites on the *virB*, *icsA*, and *rnaG* promoters from previous reports. A) Footprinting assays of VirF interacting with the *virB* promoter ([1]Tobe et al., *Journal of Bacteriology*, 1993) and with the *icsA* regulon ([2]Tran et al., *Nucleic Acids Research*, 2011). B) Effects of VirF on *in vitro* transcription of RnaG and *icsA* mRNA from Tran et al. C) Diagrams of the *virB* promoter (Tobe et al.), and the *icsA* regulon (Tran et al.) with VirF and H-NS binding sites. All figures were adapted from the respective reports.

VirF is a very interesting AraC-family transcription factor since it is both capable of promoting transcription (*virB* and *icsA*) while also repressing it (*rnaG*). VirF binds to the *icsA* regulatory region which contains both *icsA* and *rnaG* promoters and genes but manages to increase *icsA* transcription and significantly repress *rnaG* transcription (**Figure 2.20B**)(4). However, at higher VirF concentrations, there is a reduction in *icsA* transcription most likely due to VirF occupying all binding sites, specifically sites 3 and 4 in the *rnaG* promoter region, thus preventing transcription by forming a "transcriptionally inactive DNA-protein complex"(4). On the other hand, VirF significantly reduces *rnaG* transcription with higher concentrations tested. However, there are two bands present in the *in vitro* transcription assay monitoring RnaG (**Figure 2.20B**). While it is undiscussed in Tran et al., it is likely that RnaG adopts multiple forms or two transcripts of RnaG occur in this transcription assay(4).

The orientation in which VirF interacts with these promoters relative to the -35 and -10 RNA polymerase signals for each DNA promoter can explain how it acts as either an activator or a repressor. Generally, AraC-family proteins can activate either Class I or II promoters(2, 27). Class I promoters are DNA sequences in which the protein activator binds upstream of the -35 signal. For Class II promoters, protein activators overlap the -35 and/or -10 signal. Based on these criteria and the data provided in Tobe et al. (**Figure 2.20C**)(5), the *virB* promoter can be classified as a Class II promoter. Since the second VirF binding site within *pvirB* (based on those identified in Koppolu et al., 2013) was non-essential for binding (**Figure 2.12**), it also seems likely that the promoter could be classified as Class I since VirF binds upstream of the -35 region to recruit RNA polymerase(7). In addition, the MalE tag on VirF used in the footprinting assays reported in Tobe et al. may contribute significantly to the ~100 nucleotide protected region(5). On the other hand, both the *icsA* and *rnaG* promoters can be classified as Class II promoters (**Figure 2.20C**)(4). WT

VirF binding overlaps the -35 region of the *rnaG* promoter but overlaps the -10 region of the *icsA* promoter. In the report by Tran et al., it is unclear how VirF activates *icsA* transcription but it likely modulates the *icsA* promoter in a positive way to allow for more efficient binding and recruitment of RNA polymerase (i.e., via direct contact with VirF) to transcribe the *icsA* mRNA(4). For the *rnaG* promoter, VirF must be occluding RNA polymerase binding as it overlaps the -35 region thus reducing its transcription. This seems likely given *rnaG* transcription occurs from a stronger promoter without requiring VirF(20). Hence the binding of VirF reduces the strong activity of that promoter upon binding.

The *icsA* regulon is far more intricately regulated than the *virB* promoter(4, 5, 20). RnaG transcription arises from a strong promoter and represses *icsA* transcription by two mechanisms either through transcriptional interference via decreasing *icsA* promoter activity or by termination of the *icsA* transcript prematurely (i.e., transcriptional attenuation)(20). H-NS also negatively regulates this regulatory region at 30 ˚C (binding sites of H-NS to the promoter regions can be found in **Figure 2.20C**)(4). In addition, VirF has RNA-binding activity and can bind and sequester RnaG and the *icsA* mRNA thus preventing the transcription-downregulating RNA•RNA interaction to further promote *icsA* transcription(21). Similarly, the *virB* promoter is also negatively regulated by H-NS at 30 ˚C and, upon an increase in temperature to 37 ˚C, VirF outcompetes H-NS at that site to recruit RNA polymerase and promote transcription(5). However, an antisense RNA does not negatively regulate this promoter, nor is there a known RNA-binding interaction that further attenuates *virB* transcription.

2.4.4 Elaboration of the MalE-VirF•picsA and •prnaG DNA-binding Interactions

To further elaborate on the MalE-VirF DNA-binding activity, our binding mutants were tested with *picsA* and *prnaG* (**Figure 2.17 and 2.18**). I189A, R192A, and R242A mutants

presented similar binding trends to the *pvirB* studies (**Table 2.2**). Due to the insignificant change in DNA-binding for all three promoters, I189 likely contributes to all DNA-binding interactions via hydrophobic or VDW interactions. As previously observed with *pvirB*, R192A MalE-VirF did not bind to either *picsA* or *prnaG* indicating that this residue is crucial to its DNA-binding activities, regardless of the DNA promoter. The arginine likely contributes strong electrostatic or hydrogen-bonding interactions with the phosphate backbone or DNA base pairs of all promoters, respectively. Lastly, R242A MalE-VirF displayed a moderate reduction in binding for both promoters. Arginine-242 might confer weak hydrophobic/VDW or suboptimal hydrogen-bonding interactions for these promoters like our assessment with *pvirB*. We do not suspect it contributes electrostatic interactions for this promoter since binding is still attainable at our tested concentrations and knocking out a high energy interaction would likely confer a loss of DNA-binding activity (i.e., R192A).

Three mutants, S238A, Y239A, and I241A expressed differential binding when tested against *picsA* and *prnaG* (**Table 2.2**). S238A resulted in reduced binding to both promoters, however to a greater degree with the *icsA* promoter. This provides further evidence that the serine hydroxyl could be contributing via a weak or suboptimal hydrogen-bonding interaction with the DNA, irrespective of the promoter. S238 may also be involved in weak hydrophobic and VDW interactions which could also lead to a reduction in binding activity upon mutation. For Y239A, it bound nearly identically to WT when tested with *pvirB* (**Figure 2.7**), but in contrast, the mutation resulted in a larger reduction in binding to *picsA* relative to WT MalE-VirF. However, Y239A displayed a minimal reduction in binding to *prnaG*. The tyrosine hydroxyl could be more necessary for binding *picsA* while insignificant in the *pvirB* and *prnaG* DNA-binding interaction. The aromatic moiety of Y239 could also be participating in base-stacking with the DNA bases or

through weak hydrophobic/VDW interactions. Another mutant that exhibited differential binding for the three promoters was I241A. Since this mutation introduced a hydrophobically conservative change, it was expected to have an insignificant effect on DNA-binding, but it introduced a moderate reduction in *picsA* binding compared to *prnaG*. I241 may contribute a hydrophobic effect when the protein binds to *picsA* thus increasing the entropy and subsequently binding activity of the protein for this promoter. This hydrophobic effect might not be crucial for the other promoters. Most importantly, these mutations have more substantial effects on binding activity for *picsA* than for *prnaG*.

Unexpectedly, K193A MalE-VirF presented a distinctly different binding scheme (**Figure 2.17 and 2.18**). K193A exhibited only one binding shift ("S2") when titrated against both promoters while no binding was observed with *pvirB* (**Figure 2.7**). This mutation may potentially affect the oligomeric state of MalE-VirF hence why a single shift above S1 and S2 is observed. Our size exclusion chromatography purifications of K193A MalE-VirF suggest that it resides as a monomer *in vitro* when compared to WT (**Figure S2.1**). The accumulation of these observations suggests that the K193A mutation significantly affects the binding equilibria for the S1 and S2 consecutive binding events compared to WT and the other active DNA-binding mutants. The off rates may be higher for those early binding events and thus are unobservable in an EMSA. Alternatively, it is also possible that K193 is required for binding to all promoters at lower oligomeric states (monomer or dimer) but conformational changes when MalE-VirF forms higher oligomeric complexes, occupies larger portions of *picsA* and *prnaG in vitro* which may render the K193•DNA interaction non-essential. However, structural and/or biophysical studies are needed to determine if conformational changes upon protein dimerization affect VirF DNA-binding activity.

The elaboration of DNA-binding to *picsA* and *prnaG* provided interesting results regarding differential binding and potential effects on specificity. Firstly, mutations to the first helix-turn-helix motif (i.e., R192A and K193A) significantly reduced DNA-binding but K193A still exhibited binding for *picsA* and *prnaG* suggesting it does not affect binding when MalE-VirF is in a higher oligomeric complex with DNA. This N-terminal HTH motif seems crucial for DNA-binding activity and affinity for DNA promoters since it provides high energy interactions. Conversely, mutations to the second HTH motif (i.e., S238A, Y239A, and I241A) conferred differential activities for *picsA* and *prnaG*. Mutations to this C-terminal HTH motif more significantly reduces the DNA-binding activity of MalE-VirF for *picsA* but not for *prnaG*. Since all three promoters vary in lengths, apparent affinities, and length of individual VirF binding sites, it is difficult to determine why the trends between *pvirB* and *prnaG* match but differ for *picsA* within the second helix-turn-helix (**Table 2.2**). Potentially, the MalE-VirF•*picsA* DNA-binding interaction may be more intricate, where all interactions are needed to occur to bind at full affinity. This may not be as crucial in the MalE-VirF•*pvirB* and •*prnaG* interactions since mutations to the second HTH resulted in lesser changes to DNA-binding. These interactions could encompass lower energy complexes.

Overall, this data provides evidence that the two helix-turn-helix motifs in VirF provide different functions as the first motif confers DNA-binding affinity while the second may provide promoter specificity. However, the spatial orientation of which HTH participates in high affinity or promoter specific interactions across AraC-transcription factors is expected to be transcription factor specific. It has been shown that mutations to the C-terminal HTH motif of ToxT from *V. cholerae* display greater reductions in transcriptional activation than the N-terminal HTH motif for three promoters in a β-galactosidase reporter assay(28). Purified ToxT proteins harboring

mutations to either motif did not exhibit significant changes to DNA-binding activity compared to WT ToxT suggesting these mutations are likely affecting transcription elsewhere(28). Additionally, mutations to the N-terminal HTH motif in MarA were more disruptive than those of the C-terminal HTH, with the exception of R96A(29). As previously stated, Rob in *E. coli* interacts with *micF* through its N-terminal HTH motif solely (14) but this could be an artifact of the crystallization conditions or a lack of other specific factors or conditions that the protein requires to bind to its DNA promoter with both motifs. While these examples highlight the importance of one motif over the other, further studies are needed to validate this hypothesis for VirF.

Table 2.2: Amino Acids within the VirF DNA-binding domain that contribute to binding to *pvirB*, *picsA*, and *prnaG*. The effects of each alanine-mutation on DNA-binding activity compared to WT MalE-VirF are included.

Mutation	**Change in Binding**		
	pvirB	*picsA*	*prnaG*
I189A	Insignificant	Insignificant	Insignificant
R192A	Loss of binding	Loss of binding	Loss of binding
K193A	Loss of binding	Loss of S1 appearance of "S2"	Loss of S1 appearance of "S2"
S238A	Slight Reduction	Moderate Reduction	Slight Reduction
Y239A	Slight Reduction	Moderate Reduction	Slight Reduction
I241A	Slight Reduction	Moderate Reduction	Insignificant
R242A	Moderate Reduction	Moderate Reduction	Moderate Reduction

2.4.5 Role of the Dimerization Domain in VirF DNA-binding Activity

While this chapter has focused on the DNA-binding domain of VirF, this protein also contains an N-terminal dimerization domain which may influence its DNA-binding activity. It has been shown that a truncated form of VirF which lacks a portion of its dimerization domain, known as$VirF_{21}$, exhibited binding to the *virF* promoter at a different position than WT VirF, to negatively

autoregulate the expression of full-length VirF(30). This result suggested that the N-terminal dimerization domain may influence the DNA-binding domain's affinity for its promoters. Although VirF is thought to dimerize, this domain is severely understudied in VirF. Because of this, dimerization's effect on promoter binding or activation of gene transcription has not been elucidated. Porter and Dorman reported that when co-expressed with WT VirF and tested in a β-galactosidase assay, only two VirF mutations (I180N and Y224Ochre) presented dominant negative effects, suggesting that the protein dimerizes(15). There have not been any reported studies that directly test the dimerization activity of VirF and its effects on DNA-binding. Fortunately, this domain in AraC from *E. coli*, ToxT from *V. cholerae*, and HilD from *S. enterica* serovar Typhimurium, have been studied(8, 31, 32). Alterations to this domain through effector binding (AraC) or site-directed mutagenesis (ToxT), affected these homologs' abilities to bind to their cognate promoters (33, 34). As seen in **Figure 2.14**, there are multiple shifts arising when WT MalE-VirF interacts with its larger cognate promoters, *picsA* and *prnaG*. These observtions highly suggest that, beyond multiple DNA-binding events, dimerization could be occurring for these promoters. While multiple shifts were not present in EMSAs testing binding of WT MalE-VirF to wildtype and scrambled *pvirB* probes (**Figure 2.7 and 2.12**), it cannot be ruled out that MalE-VirF does not interact with this promoter as a dimer. VirF dimerization and its effects on DNA-binding activity are explored in **Chapter 3**.

2.4.6 Temporal Activation of VirF-Driven Promoters

Based on the binding data presented within this chapter, it can be speculated that the DNA-binding activity of VirF for its three cognate promoters reflects the timeline of *S. flexneri* infection of the human intestinal epithelium. In the early stages of the infection (**Figure 2.21 (1)**), *S. flexneri* undergo transcytosis from M cells into resident macrophages within the human intestinal

epithelium(35). To be able to invade intestinal cells, the bacteria must induce macrophage pyroptosis (inflamed cell death) via its type 3 secretion system (T3SS) but more specifically, IpaB. Located at the tip of the T3SS along with IpaD and IpaC, IpaB has been shown to induce macrophage pyroptosis through direct interactions and activation of caspase 1(36–38). All of these proteins are indirectly regulated by VirF (via direct activation of *virB*)(5, 16). Considering these proteins are required at the earliest stages of the infection, MalE-VirF's higher affinity for this promoter compared to *picsA* and *prnaG* makes sense, temporally. Before the bacteria create their replicative niche within epithelial cells, there may be a limited amount of VirF available to activate transcription of these necessary virulence factors. But there is an amplification effect since VirF activates transcription and subsequent expression of the downstream transcription factor, VirB. Because of the VirB amplification effect and high affinity of VirF for *pvirB*, a lower concentration of VirF (relative to that for *picsA* and *prnaG*) is required to promote activity of the T3SS and induce macrophage pyroptosis.

As there is reduced binding for *picsA* and *prnaG* (**Figure 2.14**) at low concentrations of MalE-VirF compared to *pvirB* (**Figure 2.7**), the former MalE-VirF•DNA interactions seem to be non-essential at low concentrations. In the later stages of infection (**Figure 2.21 (2)**), where there is increased intracellular replication and VirF production, VirF can reach levels significant enough to activate transcription of *icsA* and to repress *rnaG*. Production of IcsA and intracellular cell-to-cell spread through facilitating host cell actin polymerization is likely a significant energy-consuming process, so stringent regulation of these two promoters would seem to be necessary. *S. flexneri* may regulate these promoters in other substantial ways: multiple DNA-binding or dimerization events. As discussed previously, MalE-VirF interacts with these larger promoters, where multiple shifts are present, in a more complex fashion than with *pvirB*. The concentrations

of VirF needed for these multiple shifts could indicate that, to activate or repress these promoters, higher oligomeric complexes need to be formed. This is not as likely for the MalE-VirF•*pvirB* due to presence of solely one shift (**Figure 2.7**). However, probe-trapping in the wells of the EMSA could be indicative of higher order, promoter-activating complex for all three promoters or it is an active, soluble DNA-binding aggregate as discussed previously. Overall, the binding data for all three promoters suggest that VirF interacts with *pvirB* first to promote macrophage pyroptosis and intestinal epithelial cell invasion where subsequent intracellular replication and increases in [VirF] lead to activation of *icsA* transcription, *rnaG* repression, and cell-to-cell spread.

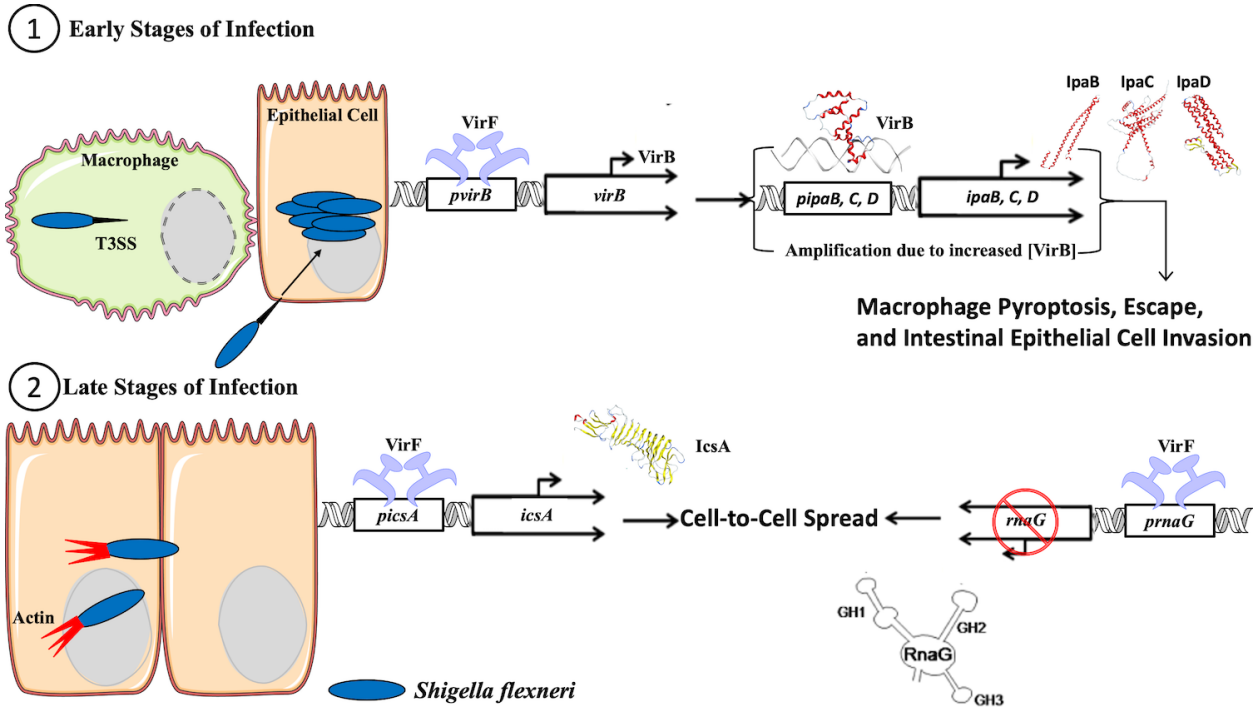


Figure 2.21: Temporal activation of VirF-controlled promoters during *S. flexneri* infection. **(1)** Early Stages of Infection. T3SS (black triangle), IpaB (PDB: 5WKQ), IpaC (AlphaFold), IpaD (PDB: 2J0O), and other effector proteins involved in macrophage pyroptosis, escape, and intestinal epithelial cell invasion (prior to replication) are controlled by the transcriptional activator, VirF. VirF activates the secondary transcription factor, VirB (PDB: 3W2A), which activates transcription the Ipa proteins and other proteins involved in the T3SS. **(2)** Late Stages of Infection. To spread cell-to-cell, VirF activates transcription of the IcsA (PDB: 5KE1) protein which is involved in facilitating host actin polymerization (red triangles). VirF also represses transcription of the antisense RNA, RnaG (figure adapted from Giangrossi et al., *Frontiers in Microbiology*, 2017), which is involved in attenuating icsA transcription, thus promoting cell-to-cell spread.

2.4.7 Conclusions

VirF, an AraC-family protein, is the master transcription factor of virulence for the intracellular pathogen, *Shigella flexneri.* These proteins are often unstable and insoluble *in vitro* which makes structural studies difficult. Using two homologs from *E. coli*, MarA and GadX, we developed homology models for the VirF DNA-binding domain (DBD). These two homology models provided insight into potential conformation changes the proteins could undergo in DNA-bound versus free states. An alignment between MarA and the VirF DBD identified seven corresponding residues (based upon the MarA•*marRAB* structure) that we mutated to alanine and tested in *in vitro* DNA-binding assays. The differential DNA-binding activities of these mutants that we observed validated the VirF homology models. Mutations to the N-terminal HTH motif (e.g., R192 and K193) exhibited the most significant effects on DNA-binding for *pvirB* while mutations to the C-terminal HTH motif varied. When DNA-binding activity was expanded to the *icsA* and *rnaG* promoters, multiple shifts were observed indicating potential dimerization activity. The same, corresponding mutants exhibited differential binding to these two promoters, specifically K193A, S238A, Y239A, and I241A. Of importance, K193A MalE-VirF solely bound to *picsA* and *prnaG* as a potential, higher oligomeric complex.

2.5 Materials and Methods

2.5.1 Materials

All standard buffer components were purchased from Millipore Sigma or Thermo Fisher. Specific reagents or biological products not purchased from these are noted in parentheses. DNA oligonucleotides were purchased from Integrated DNA Technologies. Equipment utilized for

these experiments was purchased from varying companies which are indicated in parentheses throughout.

2.5.2 Homology Modeling and Sequence Alignments

Performed by Nicholas Ragazzone, two protein templates which showed the highest sequence identity, homology, and GMQE (Global Model Quality Estimate) scores with the VirF C-terminal DNA-binding domain (DBD) were selected from the Protein Data Bank (PDB) using SWISS-MODEL: GadX (PDB: 3MKL) and MarA (PDB: 1BL0)(39–43). First, the VirF primary sequence for the DNA-binding domain (Uniprot Entry Number: P0A2T1; amino acids 144-262) was entered into the modeling and simulation platform, Molecular Operating Environment (MOE)(44). In the MOE homology model window, the VirF DBD sequence was entered as the modeling sequence and either the GadX or MarA template was selected as the primary structure template. The geometries for the models with the lowest root-mean-square-deviations (RMSDs) for each template were selected as the representative homology models: GadX-based or MarA-based. The VirF DBD, MarA, and GadX sequences were aligned using MOE. The alignments also provided percent-identity matches. MarA and VirF DBD were also aligned using EXPASY SIM-Alignment Tool(22).

2.5.3 Alanine-Scanning Mutagenesis of pBAD202-MALvirF

Site-directed mutagenesis of the *malE-virF* gene was performed using oligonucleotides presented in **Table 2.3**. The alanine mutation was incorporated using the GCG alanine codon for each MalE-VirF mutant (I189, R192, K193, S238, Y239, I241, R242). MalE-VirF numbering refers to the primary sequence of VirF from *Shigella flexneri* (Uniprot Entry Number: P0A2T1).

Table 2.3: DNA oligonucleotide primers designed and used for alanine-scanning mutagenesis of the MalE-VirF (pBAD202-MALvirF) expression plasmid. Alanine codons incorporated into the expression plasmids are bolded and annealing temperatures used in the Two-Step PCR are included.

Oligonucleotide	Annealing Temp. (˚C)	Sequence (5' → 3')
I189A Forward	56	CTTGAATTTATCAGAA**GCG**GCTGTTAGAAAACG
I189A Reverse	56	CAATCGTTTTCTAACAGC**CGC**TTCTGATAAATTCAAG
R192A Forward	56	GAAATAGCTGTT**GCG**AAACGATTGGAGAGTG
R192A Reverse	56	CACTCTCCAATCGTTT**CGC**AACAGCTATTTC
K193A Forward	58	CAGAAATAGCTGTTAGA**GCG**CGATTGGAGAGTG
K193A Reverse	58	CACTCTCCAATCG**CGC**TCTAACAGCTATTTCTG
S238A Forward	55	GGAATATCAAGCCCA**GCG**TATTTTATAAGGAAATTT
S238A Reverse	55	AAATTTCCTTATAAAATA**CGC**TGGGCTTGATATTCC
Y239A Forward	56	GGAATATCAAGCCCATCT**GCG**TTTATAAGGAAATTT
Y239A Reverse	56	AAATTTCCTTATAAA**CGC**AGATGGGCTTGATATTCC
I241A Forward	54	GCCCATCTTATTTT**GCG**AGGAAATTTAATGAAT
I241A Reverse	54	ATTCATTAAATTTCCT**CGC**AAAATAAGATGGGC
R242A Forward	54	GCCCATCTTATTTTATA**GCG**AAATTTAATGAATATTATGGT
R242A Reverse	54	ACCATAATATTCATTAAATTT**CGC**TATAAAATAAGATGGGC

Two-step polymerase chain reaction (PCR) mutagenesis was used on pBAD202-MALvirF with PFU Turbo DNA Polymerase (Agilent) in a MiniAmp™ Plus Thermal Cycler (Thermo Fisher). Each 50 µL reaction contained ~250 ng pBAD202-MALvirF, 1 µL dNTPs (NEB; 200 µM), 2.5 µL of forward or reverse DNA oligonucleotide primer (0.5 µM), 5 µL 10x PFU reaction buffer, 1 µL PFU Turbo DNA Polymerase (2.5 units). Final concentrations are represented in parentheses. The PCR cycles are presented in **Table 2.4**. Following the 1st PCR cycle, 25 µL of the reaction containing the forward primer was mixed with 25 µL of the reaction containing the reverse primer along with 1 µL of PFU Turbo DNA polymerase and the reaction proceeded as described in **Table 2.4**. Following the PCR cycles, the reactions were incubated with 2 µL DpnI restriction enzyme (NEB) at 37 ˚C prior to running on a 2% agarose gel with 1x TAE buffer (40 mM Tris base, 20 mM glacial acetic acid, 1 mM ethylenediaminetetraacetic acid (EDTA), pH 8.0). Bands migrating to the corresponding length of pBAD202-MALvirF (6418 nucleotides) were extracted and purified according to QIAGEN. Chemically competent TOP10 *E. coli* cells were transformed with the mutated pBAD202-MALvirF using the heat shock method The heat shock

method included incubating 2-4 µL of purified plasmid with 50 µL chemically competent TOP10 *E. coli* cells for 10 min on ice followed by 30 seconds at 42 ˚C in a water bath. Approximated 250 µL of 2xTY media (16 g bactotryptone, 10 g yeast extract, and 5 g NaCl per 1 L MilliQ purified water) was added, then the cell and plasmid mixture was allowed to shake at 190 rpm for 1 hour at 37 ˚C. Colonies were selected for overnight growth in 2xTY media supplemented with kanamycin (50 µg/mL) at 37 ˚C. On the following day, the cultures were subjected to the QIAGEN miniprep protocol. The mutants were confirmed by DNA sequence analysis via Eurofins Genomics.

Table 2.4: Two-Step PCR cycles used in alanine-scanning mutagenesis of pBAD202-MALvirF. Temperature, time at the corresponding temperature, and number of individual cycles are included below.

1st PCR Cycle		
Temperature (˚C)	**Time (sec or min)**	**Cycles**
95	30 sec	1
95	30 sec	4
Annealing Temperature (**Table 2.1**)	1 min	4
68	14 min	4
4	Held until next PCR cycle	1
2nd PCR Cycle		
95	30 sec	1
95	1 min	18
Annealing Temperature (**Table 2.1**)	1 min	18
68	14 min	18
68	10 min	1
4	Hold Overnight	1

2.5.4 Expression and Purifications of MalE-VirF proteins

WT and mutant MalE-VirF proteins were expressed similarly to that previously described in Emanuele and Garcia, 2015(6). Chemically competent TOP10 *E. coli* cells were transformed with the MalE-VirF expression plasmid, pBAD202-MALvirF, using the heat shock method (described in **2.5.3**). Following successful transformations, cells harboring the expression plasmid

were grown overnight in 10 mL 2xTY media supplemented with kanamycin (50 µg/mL) while shaking at 37 ˚C. Following overnight incubation, the overnight culture was added to 1 L 2xTY supplemented with kanamycin and allowed to grow at 37 ˚C until the cells reached OD_{600} ~0.5. At the appropriate OD_{600}, 10 mL of sterile filtered 20% w/v L-(+)-arabinose was added to induce MalE-VirF protein expression, and the cultures were allowed to shake and grow at 37 ˚C for an additional 5 hours. The cultures were centrifuged at 6,000xg for 15 min at 4 ˚C.

Following centrifugation, the pelleted cells were resuspended in MalE-VirF binding buffer (20 mM Tris-HCl, 1 mM EDTA, 1 mM dithiothreitol (DTT), 500 mM NaCl, pH 7.4) and supplemented with a "c0mplete, Mini Protease Inhibitor Cocktail" tablet (Roche) and 0.1 mM PMSF. Following resuspension, all subsequent steps were on ice or kept at 4 ˚C. The resuspended cells were lysed via sonication (6 cycles, 30 seconds pulse, 4 minutes rest, 60% of max pulse setting) using an ultrasonic XL2020 sonicator (Misonix). Cellular debris was separated via centrifugation (45 min, 25000xg, 4 ˚C). Following centrifugation, the supernatant was loaded onto a 5 mL MBPTrap HP column (Cytiva) using an AKTA FPLC (GE Healthcare). The column was then washed with 15 column volumes (CV) MalE-VirF binding buffer. Protein was eluted from the column using MalE-VirF elution buffer (20 mM Tris-HCl, 1 mM EDTA, 1 mM DTT, 500 mM NaCl, 5% glycerol, 10 mM maltose, pH 7.4). The eluted protein was collected in 1.5 mL fractions. Fractions with the highest resulting UV-280nm absorbance were collected and concentrated to approximately 600 µL using Amicon Ultra-15 MWCO 10 kDa centrifugal filter units and then filtered to remove any precipitate. The filtered protein was loaded onto a Superdex 200 GL10/300 gel filtration column to separate MalE-VirF from truncated MalE and other protein impurities eluted from the MBPTrap HP column. Protein was eluted from the column using MalE-VirF binding buffer and 0.5 mL fractions were collected and tested via SDS-PAGE to determine where

MalE-VirF eluted. The corresponding MalE-VirF fractions were dialyzed with a Slide-A-Lyzer dialysis cassette (10K MWCO; Thermo Fisher) at 4 ˚C overnight in MalE-VirF dialysis buffer (20 mM Tris-HCl, 1 mM EDTA, 5 mM DTT, 200 mM NaCl, 40% glycerol, pH 7.4). The concentration of each protein was tested with a Bradford assay (BioRad) using bovine serum albumin (BSA) standards. Protein stocks were flash frozen in liquid nitrogen and then stored at -80 ˚C.

2.5.5 Preparation of pvirB, picsA, and prnaG probes

A Cy5-labeled *virB* (*pvirB*) oligonucleotide promoter probe was annealed as previously described using the oligonucleotide primers found in **Table 2.5** (6). For the *pvirB* EMSA probe, 1 µL of the *pvirB* top strand was mixed with 1µL of the *pvirB* EMSA bottom strand and 1 µL of labeled universal electrophoretic gel shift oligonucleotide and 27 µL TE/NaCl (10 mM Tris-HCl, 0.1 mM EDTA, 50 mM NaCl, pH 8.0). All oligonucleotides were stored as 100 µM stocks in TE/NaCl buffer. The final concentration of the annealed promoter was 3.3 µM assuming complete annealing occurred. In a MiniAmp^{TM} Plus Thermal Cycler (Thermo Fisher), the mixture was heated to 94 ˚C for 2 min followed by 70 ˚C for an additional 2 min. The reaction was allowed to cool to room temperature prior to use in the EMSA.

Table 2.5: DNA oligonucleotides used for preparation of EMSA and FP promoter probes. Underlined sequences represent the individual binding sites on the *virB*, *icsA*, *rnaG* promoters, and their corresponding scrambled sites. The italicized sequence represents the LUEGO (labeled universal gel shift oligonucleotide) annealing site.

Oligonucleotide	Sequence (5' → 3')
LUEGO	Cy5-GTGCCCTGGTCTGG
pvirB Top Strand FP and EMSA	AGAATATTATTCTTTTATCCAATAAAGATAAATTGCATCAATCCAGC TATTAAAATAGTA (5'-6FAM used in FP)
pvirB Bottom Strand EMSA	TACTATTTTAATAGCTGGATTGATGCAATTTATCTTTATTGGATAAAA GAATAATATTCT*CCAGACCAGGGCAC*
pvirB Bottom Strand FP	TACTATTTTAATAGCTGGATTGATGCAATTTATCTTTATTGGATAAAA GAATAATATTCT
pvirB Top Strand Scram1 EMSA	AGAATAATTAATCTTCTATTCTTAAAGATAAATTGCATCAATCCAGC TATTAAAATAGTA
pvirB Bottom Strand Scram1 EMSA	TACTATTTTAATAGCTGGATTGATGCAATTTATCTTTAAGAATAGAA GATTAATTATTCT*CCAGACCAGGGCAC*
pvirB Top Strand Scram2 EMSA	AGAATATTATTCTTTTATCCAATAAAGATAAATTGCATACGATATCA ACAATCTATAGTA
pvirB Bottom Strand Scram2 EMSA	TACTATAGATTGTTGATATCGTATGCAATTTATCTTTATTGGATAAAA GAATAATATTCT*CCAGACCAGGGCAC*
picsA Taq Forward	GCAAGACACAGGTAAATTTCTCCCGTTGCATTGATATATAACACAGC TCTCATGTTTTGGTTG
picsA Taq Reverse	TTTATCCGGAATAAAGGGACGATATATGCAAAAACATATTAAACAA AGCCTCAACCAAAACATGAGAGCTGTGTTA
picsA BS1 Scram Taq Forward	GCAAGAGCCTTTGAGTAATACATCTACTAATCGTGCTATAACACAGC TCTCATGTTTTGGTTG
picsA BS2 Scram Taq Reverse	TTTATCTAGCAGAATGATACGGATTTAACAGACACTCCAAGAAAAG AATAACAACCAAAACATGAGAGCTGTGTTA
picsA EMSA PCR Forward	Cy5-GCAAGACACAGGTAAATTTCTCCCG
picsA EMSA PCR Reverse	TTTATCCGGAATAAAGGGACGATATATGC
picsA BS1 Scram EMSA PCR Forward	Cy5-GCAAGAGCCTTTGAGTAATACATCTACTAA
picsA BS2 Scram EMSA PCR Reverse	TTTATCTAGCAGAATGATACGGATTTAACAG
prnaG Taq Forward	CCCCCCGAGCAACAGGGATGCACCAACAACAAAAGGAGAAAAAGA CAATTTGCATGTTTTTTCCTTTACCGTAGGTAATTCTCCGGCCCC
prnaG Taq Reverse	TAATATAGTGCATGAATCAAATTCACAAATTTTTTTGTAATATGACC CAATGTTCACAGGGGGGGGCCGGAGAATTACCTACGGTAAA
prnaG BS4 Scram Taq Forward	CCCCCCACCAACGGAAGAATATAAGGAACATTCTAACCTGATTCGA TCGACAATGAGAACTAGGTTTACCGTAGGTAATTCTCCGGCCCC
prnaG BS3 Scram Taq Reverse	TAATATGAGTAGATATTATCTAACCGTTCAGCCGATAGTCACTATAA TAGATCTATGGTATAGGGGCCGGAGAATTACCTACGGTAAA
prnaG EMSA PCR Forward	Cy5-CCCCCCGAGCAACAGGGATGCA
prnaG EMSA PCR Reverse	TAATATAGTGCATGAATCAAATTCACAAAT
prnaG BS4 Scram EMSA PCR Forward	Cy5-CCCCCCACCAACGGAAGAATATAAGGAAC
prnaG BS3 Scram EMSA PCR Reverse	TAATATGAGTAGATATTATCTAACCGTTCAGC

To prepare the *icsA* (*picsA*) and *rnaG* (*prnaG*) promoters, two overlapping DNA primers were PCR amplified with Taq DNA polymerase (NEB) and the oligo nucleotides presented in **Table 2.5** according to manufacturer instructions (**Table 2.6**). Reactions included 1 μM forward

primer (*picsA* or *prnaG* Taq Forward; **Table 2.5**), 1 µM reverse primer (*picsA* or *prnaG* Taq Reverse; **Table 2.5**), 1x Taq DNA polymerase buffer, 200 µM dNTPs, and 1-unit Taq DNA polymerase (similar reaction volumes can be found in section **2.5.3**). The scrambled promoters were PCR amplified similarly with the following oligonucleotide primers pairs: *picsA* BS1 Scram Taq Forward/*picsA* Taq Reverse for *picsA* BS1 Scram, *picsA* Taq Forward/*picsA* BS2 Scram Taq Reverse (*picsA* BS2 Scram), *prnaG* Taq Forward/*prnaG* BS3 Scram Taq Reverse (*prnaG* BS3 Scram), and *prnaG* BS4 Scram Taq Forward/*prnaG* Taq Reverse for *prnaG* BS4 Scram. Following PCR and gel extraction from a 2% agarose gel via QIAGEN, A-tailed PCR product was TOPO cloned into the pCR™2.1-TOPO® vector (Invitrogen; TOPO™ TA Cloning™ Kit, with pCR2.1™-TOPO™) according to manufacturer protocols. The general protocol includes mixing 4 µL purified A-tailed PCR product with 1 µL salt solution (provided in kit), and 1 µl TOPO vector (Invitrogen) which is allowed to incubate at room temperature for > 5 min before transformation of TOP10 *E. coli*. The insertion of the *picsA* or *prnaG* promoters was verified via sequencing analysis.

One step PCR was performed using Cy5-labeled oligonucleotide primers (**Table 2.5**), the previously prepared TOPO plasmids, and Phusion DNA polymerase (NEB) according to manufacturer instructions. Reactions included 1 µM forward primer (*picsA* or *prnaG* EMSA PCR Forward; **Table 2.5**), 1 µM reverse primer (*picsA* or *prnaG* EMSA PCR Reverse; **Table 2.5**), ~250 ng template plasmid (pCR2.1-TOPO-*picsA* or -*prnaG*), 1x High Fidelity buffer, 200 µM dNTPs, and 1-unit Phusion polymerase (similar reaction volumes can be found in section **2.5.3**). PCR amplified promoters were loaded onto a 2% agarose gel in 1x TAE buffer (40 mM Tris Base, 20 mM glacial acetic acid, 1 mM EDTA, pH 8) and compared to a 100 bp DNA ladder (NEB). The Cy5-labeled scrambled promoters were PCR amplified similarly with the following

oligonucleotide primers pairs: *picsA* BS1 Scram EMSA PCR Forward / *picsA* EMSA PCR Reverse for *picsA* BS1 Scram, *picsA* EMSA PCR Forward / *picsA* BS2 Scram EMSA PCR Reverse (*picsA* BS2 Scram), *prnaG* EMSA PCR Forward / *prnaG* BS3 Scram EMSA PCR Reverse (*prnaG* BS3 Scram), and *prnaG* BS4 Scram EMSA PCR Forward / *prnaG* EMSA PCR Reverse for *prnaG* BS4 Scram. The Cy5-labeled scrambled promoters were PCR amplified similarly with the following oligonucleotide primers pairs: *picsA* BS1 Scram Taq Forward/*picsA* Taq Reverse for *picsA* BS1 Scram, *picsA* Taq Forward/*picsA* BS2 Scram Taq Reverse (*picsA* BS2 Scram), *prnaG* Taq Forward/*prnaG* BS3 Scram Taq Reverse (*prnaG* BS3 Scram), and *prnaG* BS4 Scram Taq Forward/*prnaG* Taq Reverse for *prnaG* BS4 Scram. Promoters were gel extracted using a gel extraction kit (QIAGEN) and the concentration of the purified product was determined using a BioTek Synergy H1 plate reader.

Table 2.6: One-Step PCR cycles used in production of A-tailed and Cy5-labeled *picsA* and *prnaG* promoters. Temperature, time at the corresponding temperature, and number of individual cycles are included below.

A-Tailed Promoters (Taq DNA Polymerase)		
Temperature (˚C)	**Time (sec or min)**	**Cycles**
95	30 sec	1
95	30 sec	30
62	1 min	30
68	30 sec	30
68	5 min	1
4	Hold Overnight	1
Cy5-labeled Promoters (Phusion DNA Polymerase)		
95	30 sec	1
95	30 sec	30
53 (*picsA*), 50 (*prnaG*)	50 sec	30
72	30 sec	30
72	10 min	1
4	Hold Overnight	1

2.5.6 Electrophoretic Mobility Shift Assay (EMSA)

EMSAs were performed similarly to previously described in Emanuele and Garcia, 2015(6). A 6% native polyacrylamide gel was prepared using 1.2 mL of 30% acrylamide/bis-acrylamide (29:1 ratio) solution, 3 mL 0.5x TBE buffer (0.25x final concentration; 22 mM Tris Base, 22 mM boric acid, 0.5 mM EDTA, pH 9.5), 12 µL tetramethylethylenediamine (TEMED), and 100 µL 10 % ammonium persulfate (APS). The gel solidified at room temperature for ~1 hour and then the empty gel was electrophoresed for 1 hour at 150 V in 0.25x TBE (pH 9.5) at 4 ˚C.

Reactions were prepared with 6 µL WT MalE-VirF or mutant protein, 6 µL *pvirB* (83 nM),1.5 µL Milli-Q H_2O, 1 µL salmon sperm DNA (0.7 mg/mL), and 0.5 µL BSA (0.07 mg/mL). Final concentrations are listed in parentheses and the final buffer conditions were 12 mM Tris-HCl, 100 mM NaCl, 16% glycerol, 2 mM DTT, and 0.44 mM EDTA, pH 7.4. To perform titrations of all MalE-VirF proteins against *pvirB*, two-fold dilutions were prepared using MalE-VirF dialysis buffer. MalE-VirF WT and mutant protein concentrations varied by purification yield. Negative control reactions were prepared with 6 native gel loading dye (300 mM Tris-HCl, 50% glycerol, 0.05% bromophenol blue, pH 7) instead of MalE-VirF proteins. All reactions were incubated at 37 ˚C for 15 min in a water bath before adding 6 µL of each reaction to the corresponding wells on the gel. The gel was electrophoresed at 150 V for an additional 1.5 hours in the dark at 4 ˚C. Gel visualization was performed using a Molecular Dynamics Typhoon 9200 Molecular Imager by excitation (Ex) at 607 nm and reading the 710 nm emission (Em). Quantitative data was obtained by measuring the density of the bands on the gel using ImageJ software(45). Prism 9 software(46) (one-site specific binding equation) was used to generate binding affinities (K_D) and B_{Max} (%).

2.5.7 Fluorescence Polarization (FP) Assay

The FP assays were adapted and optimized from Emanuele and Garcia, 2015(6). The assays were conducted in low-volume round-bottom 384-well plates (Corning). The fluorophore-labeled *pvirB* oligonucleotide probe was annealed as previously described using the oligonucleotide primers found in **Table 2.5** (6). The top strand of the *pvirB* promoters (**Table 2.5**) contained a 5'-fluorescein and was annealed to the corresponding, unlabeled bottom strand to prepare stocks with concentrations of 5 μM. The DNA probe was diluted with probe buffer (50 mM Tris-HCl, 1 mM EDTA, 5 mM DTT, and 200 mM NaCl, pH 7.4) to a final working concentration of 20 nM and supplemented with 1.4 mg/mL salmon sperm DNA (Invitrogen) and 0.14 mg/mL BSA. Next, 14-step two-fold serial dilutions of WT or DNA-binding mutants of MalE-VirF were prepared with MalE-VirF dialysis buffer then 10 μL of each dilution were added to the appropriate wells in triplicate for each concentration tested. Following loading of protein into the wells, 10 μL of the 20 nM DNA probe solution were added to each well. Final concentrations of DNA probe, salmon sperm DNA, and BSA in each reaction are 10 nM, 0.7 mg/mL, and 0.07 mg/mL, respectively. Negative control reactions were also tested in triplicate, containing 10 μL MalE-VirF dialysis buffer and 10 μL of the 20 nM DNA probe solutions to determine the baseline FP for *pvirB*. Additionally, for each protein concentration tested, a blank reaction, containing the corresponding tested protein concentration and the DNA probe solution lacking the labeled probe, was used to subtract fluorescent contributions of the tested MalE-VirF protein, salmon sperm DNA, BSA, and other buffer components from the test reactions. The plate was incubated at 37 °C for 2 hours in a humidified incubator before the raw fluorescence was measured using a Biotek Synergy H1 plate reader after excitation at 485 nm (fluorescein Ex/Em = 485/528). Polarization values were calculated from the corrected parallel ($F_{||}$) and perpendicular

($F_{\perp}$) fluorescence values using **Equation 2.1**. The G factor, often used in polarization calculations, was not included in our calculations as we determined it to be negligible under these conditions (data not shown).

Equation 2.1 $FP = (F_{\parallel} - F_{\perp}) / (F_{\parallel} + F_{\perp}) * 1000$

The plots were fit by non-linear regression to the following sigmoidal four-parameter equation using Prism 9(47) (GraphPad Software; **Equation 2.2**),

Equation 2.2 $mP = min + (max - min) / (1 + 10^{((logEC_{50}-X) * Hill\ Slope)})$

where max and min were the maximum and minimum plateaus of the mP (milli-polarization) and X is the log of sample concentration. When unconstrained, the values of max, min, EC_{50}, and Hill slope are fit by the regression plot. For all plots except the I189A mutant, the mP max was constrained to the observed mP max for WT MalE-VirF (e.g., 95 mP).

2.5.8 Circular Dichroism (CD)

Buffer exchange was performed on WT and selected MalE-VirF and MarA mutant proteins to place the samples in 10 mM NaH_2PO_4 (pH 7.5) using Amicon Ultra-0.5 centrifugal filter units (3 MWCO; MilliporeSigma). The samples were centrifuged (13,000xg, 4 ˚C, 10 min), flow-through was discarded, and then diluted again with 10 mM NaH_2PO_4 (pH 7.5). This process was repeated for 5-10 rounds. The concentration of each sample was determined via Bradford assay (BioRad) compared to a series of BSA standards. For CD testing, samples were loaded into a 1 mm path-length quartz cuvette and CD spectra were collected from 195-250 nm using a JASCO J-810 spectrometer. Spectra of a sample containing only buffer was used to correct the raw data for each protein. JASCO Spectra Manager was used to visualize the spectra and export the raw data to be plotted using Prism 9 software(48).

2.5.9 Kinetic Simulation of WT MalE-VirF Binding to prnaG

The EMSA binding shift data for WT MalE-VirF binding to the *prnaG* promoter, plotted in **Figure 2.14**, were compared to a kinetic simulation of three consecutive reactions as shown below.

$$\text{For A} \xrightarrow[P]{k_1} \text{B} \xrightarrow[P]{k_2} \text{C} \xrightarrow[P]{k_3} \text{D}$$

Equation 2.3

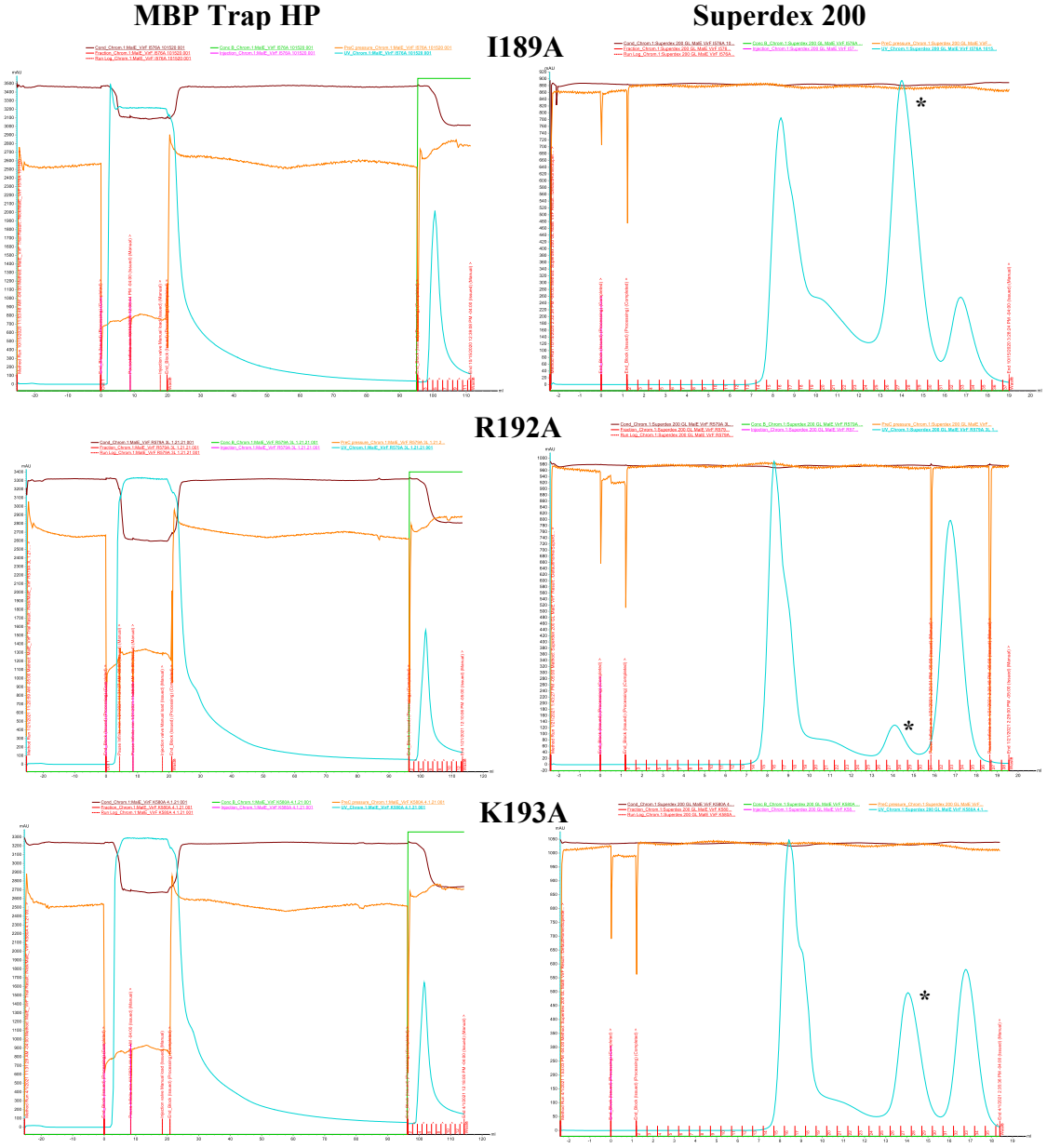
$$B = [A_0]\,[k_1]\,[\exp(-k_1x)/(k_2-k_1) + \exp(-k_2x)/(k_1-k_2)]$$

$$C = [A_0]\,[k_1k_2]\,[\exp(-k_1x)/((k_2-k_1)(k_3-k_1)) + (\exp(-k_2x)/((k_1-k_2)(k_3-k_2)) + (\exp(-k_3x)/((k_1-k_3)(k_2-k_3))]$$

$$D = [A_0]\,[k_2k_3*(1-\exp(-k_1x))/((k_2-k_1)(k_3-k_1)) + k_1k_3*(1-\exp(-k_2x))/((k_1-k_2)(k_3-k_2)) + k_1k_2*(1-\exp(-k_3x))/((k_1-k_3)(k_2-k_3))]$$

In this simulation, A = unbound DNA, A_0 = unbound DNA at x = 0, B = S1, C = S2, D = Well, P = WT MalE-VirF, and x = [WT MalE-VirF] (**Figure 2.14**). An Excel spreadsheet (adapted from "Biochemistry Online")(49) was created wherein A_0, k_1, k_2, and k_3 can be varied and the resultant curves plotted on the same graph with the experimental EMSA data. This model was derived for simulating consecutive chemical reactions where x = time (Adapted from Chemistry Stack Exchange online)(50). In that case, the k values are interpreted as unimolecular rate constants. In this adaptation for equilibrium binding, x = concentration and the k values have units of 1/concentration, suggesting that they may approximate equilibrium constants for the corresponding binding steps.

2.6 Supplemental Figures



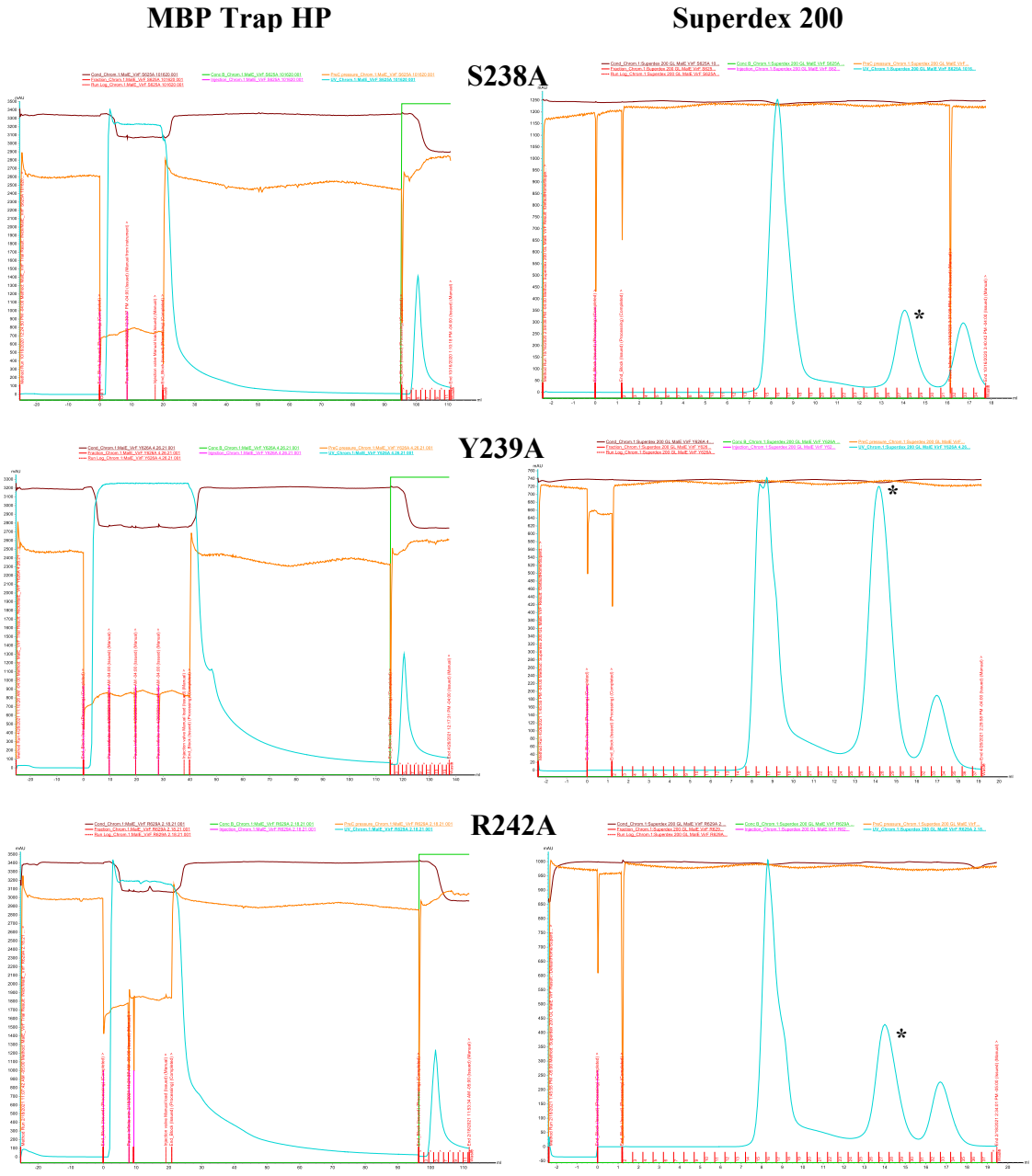
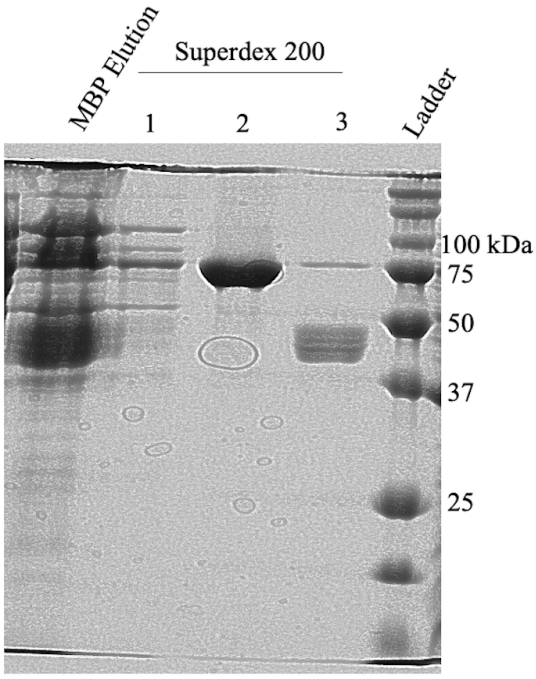
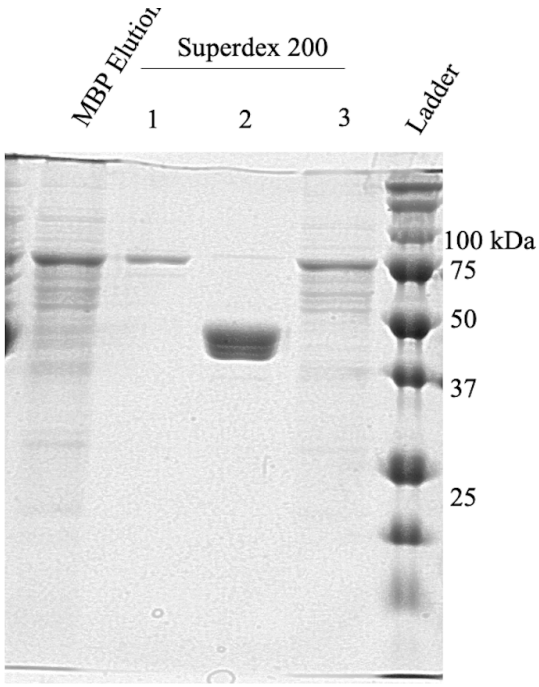


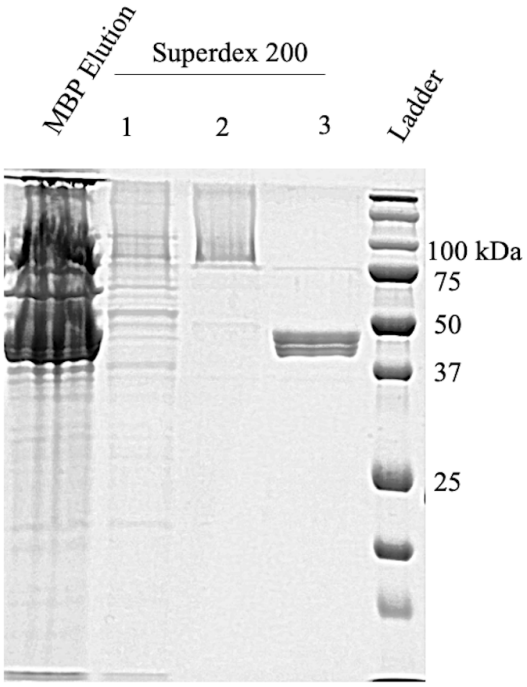
Figure S2.1: Chromatograms for MBPTrap HP and size exclusion chromatography (Superdex 200 10/300) for the MalE-VirF DNA-binding domain mutant purifications. The Y-axes (mAu at 280 nm) are not to scale between the mutant purifications. The peak for the MBPTrap HP column was concentrated and loaded onto the Superdex column to separate MalE-VirF from impurities and truncations (* peak containing pure MalE-VirF isolated for dialysis and *in vitro* testing). Note: The file for the I241A chromatograms was corrupt and is therefore not in this figure.



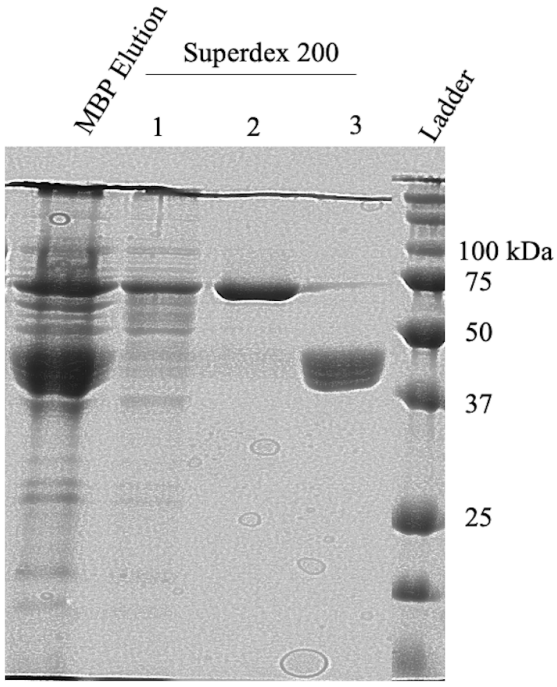
I189A



R192A



K193A



S238A

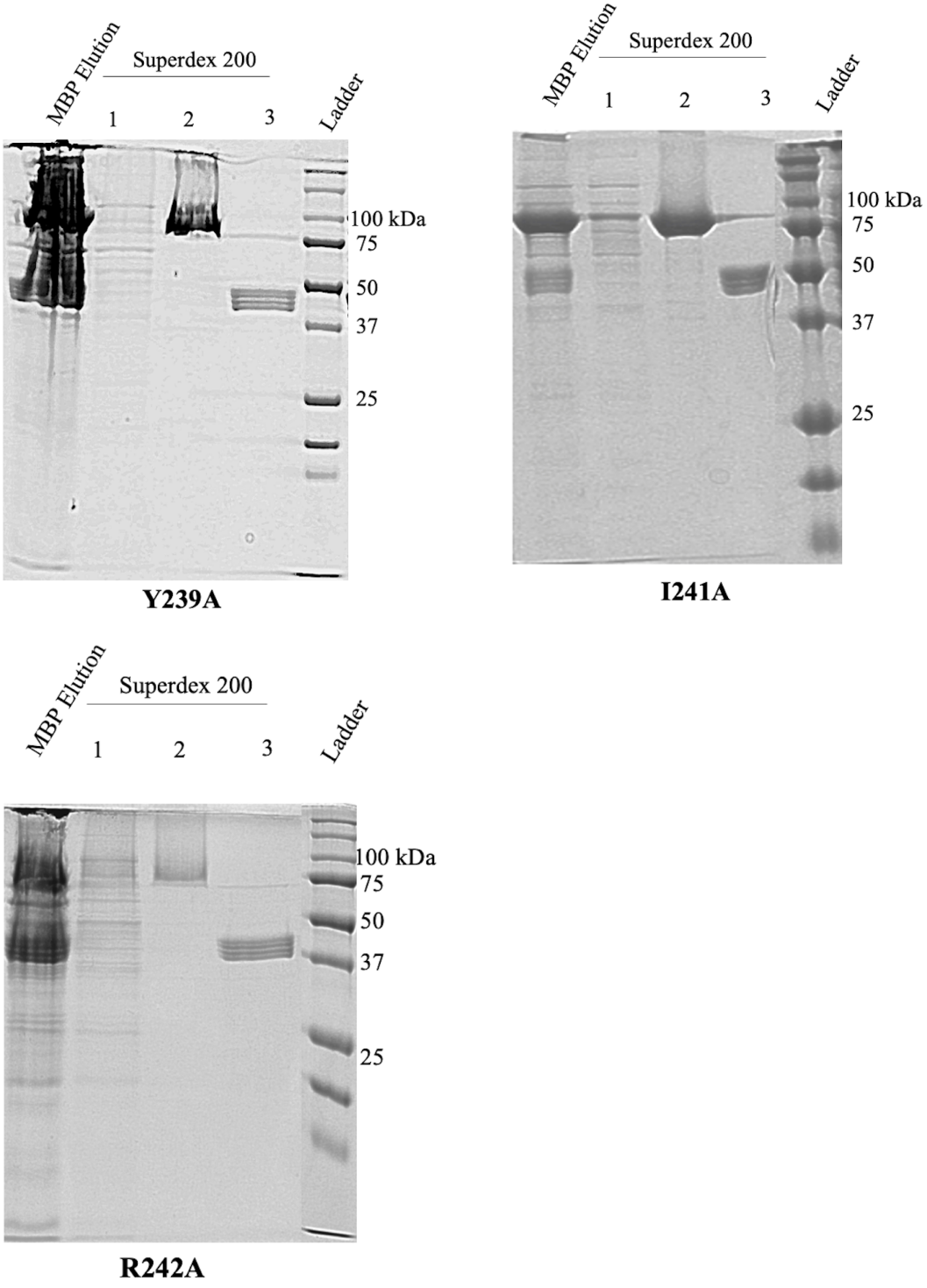


Figure S2.2: SDS-PAGE for I189A, R192A, K193A, S238A, Y239A, I241A, and R242A MalE-VirF purifications using an MBPTrap HP (MBP Elution) and Superdex 200 GL 10/300 gel filtration columns (Superdex 200 1, 2, 3 correspond to **Figure 2.5**). Protein ladders for the S238A, I241A, and R242A gels are cropped and placed to the right of the gels for consistency.

Figure S2.3: EMSA probes for the DNA promoters tested in this report. DNA sequences for *pvirB*, *picsA,* and *prnaG* sequences including protein binding sites labeled in red. For *pvirB*, both binding sites are 16 bp long. For *picsA*, binding site 1 is 32 bp long and binding site 2 is 45 bp long. For *prnaG*, binding site 3 is 56 bp long and binding site 4 is 58 bp long. The fluorophore, Cy5, is attached to the 5' end of the top strand of each promoter probe. VirF binding sites for these promoters are based on previous DNase I footprinting analyses (Tran et al., *Nucleic Acids Research,* 2011 and Tobe et al., *Journal of Bacteriology*, 1993).

2.7 References

1. Gallegos M, Schleif R, Bairoch A, Hofmann K, Ramos JL. 1997. AraC/XylS family of transcriptional regulators. Microbiol Mol Biol Rev 61:393–410.

2. Martin RG, Rosner JL. 2001. The AraC transcriptional activators. Curr Opin Microbiol 4:132–137.

3. Schleif R. 2003. AraC protein: A love-hate relationship. BioEssays 25:274–282.

4. Tran CN, Giangrossi M, Prosseda G, Brandi A, Di Martino ML, Colonna B, Falconi M. 2011. A multifactor regulatory circuit involving H-NS, VirF and an antisense RNA modulates transcription of the virulence gene icsA of Shigella flexneri. Nucleic Acids Res 39:8122–8134.

5. Tobe T, Yoshikawa M, Mizuno T, Sasakawa C. 1993. Transcriptional Control of the Invasion Regulatory Gene virB of Shigella flexneri: Activation by VirF and Repression by H-NS. J Bacteriol 175:6142–6149.

6. Emanuele AA, Garcia GA. 2015. Mechanism of Action and Initial, In Vitro SAR of an Inhibitor of the Shigella flexneri Virulence Regulator VirF. PLoS One 10:1–18.

7. Koppolu V, Osaka I, Skredenske JM, Kettle B, Hefty PS, Li J, Egan SM. 2013. Small-Molecule Inhibitor of the Shigella flexneri Master Virulence Regulator VirF. Infect Immun 81:4220–4231.

8. Soisson SM, MacDougall-Shackleton B, Schleif R, Wolberger C. 1997. Structural Basis for Ligand-Regulated Oligomerization of AraC. Science (80-) 276:421–425.

9. Rodgers ME, Schleif R. 2009. Solution Structure of the DNA Binding Domain of AraC Protein. Proteins 77:202–208.

10. Rhee S, Martin RG, Rosner JL, Davies DR. 1998. A novel DNA-binding motif in MarA: The first structure for an AraC family transcriptional activator. Proc Natl Acad Sci U S A 95:10413–10418.

11. Midgett CR, Talbot KM, Day JL, Munson GP, Kull FJ. 2021. Structure of the master regulator Rns reveals an inhibitor of enterotoxigenic Escherichia coli virulence regulons. Sci Rep 11:1–13.

12. Li J, Wehmeyer G, Lovell S, Battaile KP, Egan SM. 2016. 1.65 Å resolution structure of the AraC-family transcriptional activator ToxT from Vibrio cholerae. Acta Crystallogr Sect Struct Biol Commun 72:726–731.

13. Lowden MJ, Skorupski K, Pellegrini M, Chiorazzo MG, Taylor RK, Kull FJ. 2010. Structure of Vibrio cholerae ToxT reveals a mechanism for fatty acid regulation of virulence genes. Proc Natl Acad Sci U S A 107:2860–2865.

14. Kwon HJ, Bennik MHJ, Demple B, Ellenberger T. 2000. Crystal structure of the Escherichia coli Rob transcription factor in complex with DNA. Nat Struct Biol 7:424–430.

15. Porter ME, Dorman CJ. 2002. In Vivo DNA-Binding and Oligomerization Properties of the Shigella flexneri AraC-Like Transcriptional Regulator VirF as Identified by Random and Site-Specific Mutagenesis. J Bacteriol 184:531–539.

16. Adler B, Sasakawa C, Tobe T, Makino S, Komatsu K, Yoshikawa M. 1989. A dual transcriptional activation system for the 230 kb plasmid genes coding for virulence-associated antigens of Shigella flexneri. Mol Microbiol 3:627–635.

17. Hurt JK, McQuade TJ, Anthony E, Larsen MJ, Garcia GA. 2010. High-Throughput Screening of the Virulence Regulator VirF: A Novel Antibacterial Target for Shigellosis. J Biomol Screen 15:379–387.

18. Emanuele AA, Adams NE, Chen YC, Maurelli AT, Garcia GA. 2014. Potential novel antibiotics from HTS targeting the virulence-regulating transcription factor, VirF, from Shigella flexneri. J Antibiot (Tokyo) 67:379–386.

19. Dorman CJ. 2004. H-NS: a universal regulator for a dynamic genome. Nat Rev Microbiol 2:391–400.

20. Giangrossi M, Prosseda G, Tran CN, Brandi A, Colonna B, Falconi M. 2010. A novel antisense RNA regulates at transcriptional level the virulence gene icsA of Shigella flexneri. Nucleic Acids Res 38:3362–3375.

21. Giangrossi M, Giuliodori AM, Tran CN, Amici A, Marchini C, Falconi M. 2017. VirF relieves the transcriptional attenuation of the virulence gene icsA of Shigella flexneri affecting the icsA mRNA-RnaG complex formation. Front Microbiol 8:1–12.

22. Huang X, Miller W. 1991. A Time-Efficient, Linear-Space Local Similarity Algorithm. Adv Appl Math 12:337–357.

23. Ragazzone NJ, Dow GT, Garcia A. 2022. Elucidation of Key Interactions between VirF and the virB Promoter in Shigella flexneri Using E. coli MarA- and GadX-Based Homology Models and In Vitro Analysis of the DNA-Binding Domains of VirF and MarA. J Bacteriol 204.

24. Bhende PM, Egan SM. 1999. Amino Acid-DNA Contacts by RhaS: an AraC Family Transcription Activator. J Bacteriol 181:5185–5192.

25. Dorman CJ, McKenna S, Beloin C. 2001. Regulation of virulence gene expression in Shigella flexneri, a facultative intracellular pathogen. Int J Med Microbiol 291:89–96.

26. Durand JMB, Dagberg B, Uhlin BE, Björk GR. 2000. Transfer RNA modification, temperature and DNA superhelicity have a common target in the regulatory network of the virulence of Shigella flexneri: The expression of the virF gene. Mol Microbiol 35:924–935.

27. Browning DF, Busby SJW. 2016. Local and global regulation of transcription initiation in bacteria. Nat Rev Microbiol 14:638–650.

28. Childers BM, Weber GG, Prouty MG, Castaneda MM, Peng F, Klose KE. 2007. Identification of Residues Critical for the Function of the Vibrio cholerae Virulence Regulator ToxT by Scanning Alanine Mutagenesis. J Mol Biol 367:1413–1430.

29. Gillette WK, Martin RG, Rosner JL. 2000. Probing the Escherichia coli transcriptional activator MarA using alanine-scanning mutagenesis: Residues important for DNA binding and activation. J Mol Biol 299:1245–1255.

30. Di Martino ML, Romilly C, Wagner EGH, Colonna B, Prosseda G. 2016. One Gene and Two Proteins: a Leaderless mRNA Supports the Translation of a Shorter Form of the Shigella VirF Regulator. MBio 7:1–10.

31. Prouty MG, Osorio CR, Klose KE. 2005. Characterization of functional domains of the Vibrio cholerae virulence regulator ToxT. Mol Microbiol 58:1143–1156.

32. Narm K-E, Kalafatis M, Slauch JM. 2020. HilD, HilC, and RtsA Form Homodimers and Heterodimers To Regulate Expression of the Salmonella Pathogenicity Island I Type III Secretion System. J Bacteriol 202:1–13.

33. Laronde-Leblanc N, Wolberger C. 2000. Characterization of the Oligomeric States of Wild Type and Mutant AraC. Biochemistry 39:11593–11601.

34. Childers BM, Cao X, Weber GG, Demeler B, Hart PJ, Klose KE. 2011. N-terminal Residues of the Vibrio cholerae Virulence Regulatory Protein ToxT Involved in Dimerization and Modulation by Fatty Acids. J Biol Chem 286:28644–28655.

35. Jennison A V., Verma NK. 2004. Shigella flexneri infection: pathogenesis and vaccine development. FEMS Microbiol Rev 28:43–58.

36. Zychlinsky A, Kenny B, Ménard R, Prévost M -C, Holland IB, Sansonetti PJ. 1994. IpaB mediates macrophage apoptosis induced by Shigella flexneri. Mol Microbiol 11:619–627.

37. Hilbi H, Moss JE, Hersh D, Chen Y, Arondel J, Banerjee S, Flavell RA, Yuan J, Sansonetti PJ, Zychlinsky A. 1998. Shigella-induced apoptosis is dependent on caspase-1 which binds to IpaB. J Biol Chem 273:32895–32900.

38. Guichon A, Hersh D, Smith MR, Zychlinsky A. 2001. Structure-function analysis of the Shigella virulence factor IpaB. J Bacteriol 183:1269–1276.

39. Guex N, Peitsch MC, Schwede T. 2009. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. Electrophoresis 30:162–173.

40. Bertoni M, Kiefer F, Biasini M, Bordoli L, Schwede T. 2017. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. Sci Rep

7:1–15.

41. Bienert S, Waterhouse A, de Beer TAP, Tauriello G, Studer G, Bordoli L, Schwede T. 2017. The SWISS-MODEL Repository-new features and functionality. Nucleic Acids Res 45:D313–D319.

42. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, De Beer TAP, Rempfer C, Bordoli L, Lepore R, Schwede T. 2018. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 46:W296–W303.

43. Benkert P, Biasini M, Schwede T. 2011. Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics 27:343–350.

44. ULC CCG. 2022. Molecular Operating Environment (MOE), 2020.09. Montreal, QC, Canada.

45. Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. Nat Methods 9:671–675.

46. Software GP. Non-linear Regression (One site -- Specific binding) was performed using GraphPad Prism version 9.0.0 for Mac OS X. San Diego, California USA.

47. Software GP. Nonlinear Regression (Sigmoidal, 4PL, X is log(concentration)) was performed using GraphPad Prism version 9.0.0 for Mac OS X. San Diego, California USA.

48. Software GP. Normalize was performed using GraphPad Prism version 9.0.0 for Mac OS X. San Diego, California USA.

49. Jakubowski H. 2016. Biochemistry Online: An Approach Based on Chemical Logic (Chapter 6B Kinetics of Simple and Enzyme-Catalyzed Reactions). Creat Commons Attrib 40 Int Licens. https://employees.csbsju.edu/hjakubowski/classes/ch331/bcintro/default.html.

50. de Azevedo R. 2017. Chemistry Stack Exchange (How to derive the rate laws for three consecutive reactions?). Stack Exch Inc.

Notes:

The majority of this this chapter has been published in "Elucidation of Key Interactions between VirF and the *virB* Promoter in *Shigella flexneri* Using E. coli MarA- and GadX-Based Homology Models and *In Vitro* Analysis of the DNA-Binding Domains of VirF and MarA."

Nicholas J. Ragazzone*, Garrett T. Dow*, George A. Garcia. *J. Bacteriol.,* **2022**, 204(*9*). * Co first Author.(23)

Sections discussing VirF DNA-binding to *picsA* and *prnaG* will be included in "Elucidation of the DNA-binding activity of VirF from *Shigella flexneri* for the *icsA* and *rnaG* promoters and characterization of the N-terminal domain to identify residues crucial for dimerization." Garrett T. Dow, Anna M. Young, George A. Garcia. **2022**, Manuscript in Preparation.

Chapter 3 Analysis of VirF Dimerization Activity

3.1 Abstract

AraC-family proteins are characterized by a highly conserved DNA-binding domain consisting of two helix-turn helix motifs, but some proteins contain an N-terminal dimerization domain that often contributes environmental sensing or ligand binding activities. Dimerization domains for other AraC-family homologs have been well-studied and have implicated this domain's importance in transcriptional activation, DNA-binding, and as a target for anti-virulence inhibitor design. The dimerization domain of VirF has been severely understudied. Further analysis of VirF dimerization is crucial to understanding if it dimerizes and how this activity contributes to its DNA-binding activity.

3.2 Introduction

In addition to the highly conserved C-terminal DNA-binding domain, most AraC-family proteins contain a larger, N-terminal domain(1, 2). These domains are found in homologs involved in sugar catabolism and virulence regulation(2). The N-terminal domains are often involved in protein homodimerization and have additional activities such as small molecule binding (AraC, RhaR) or environmental sensing (ToxT)(3–5). Dimerization activities for AraC(4, 6, 7), ExsA(8, 9), XylS(10), HilD(11, 12), ToxT(5, 13–16), and others have been studied. The crystal structure of the N-terminal domain of AraC (PDB ID: 2ARC) indicates that dimerization occurs via a coiled-coil interface (**Figure 3.1**)(4, 7). This coiled coil dimerization interface has been further analyzed

and identified an α-helix in this interface has been identified as important for dimerization(6, 7, 9, 10, 13–16).

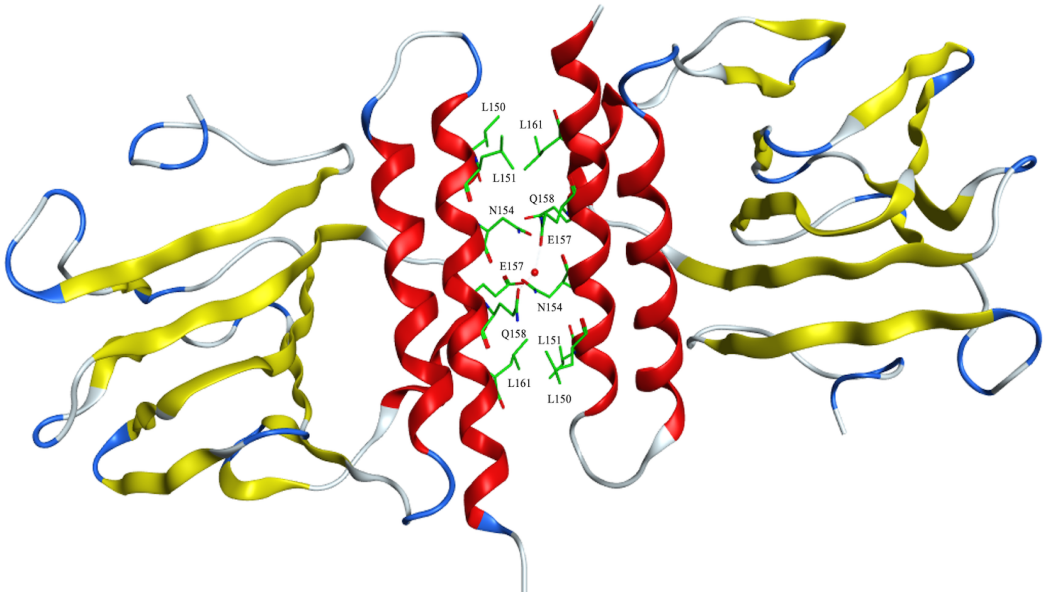


Figure 3.1: Dimerization interface of AraC based upon the crystal structure (PDB ID: 2ARC) from Bustos et al., 1993. The interaction consists of a coiled coil interface of two leucine triads (L150, L151, L161) and a central coordinating water contributing hydrogen bonding with N154 and Q158.

VirF is an AraC-family protein that contains both N-terminal and C-terminal domains with different activities. The C-terminal DNA-binding activity of VirF has been extensively studied via footprinting, *in vitro* DNA-binding assays, site-directed mutagenesis, and *in vivo* β-galactosidase reporter assays which have identified crucial DNA-binding interactions, DNA-binding inhibitors, and how the protein interacts with its three DNA promoters(17–24). To date, no direct results have been attributed to VirF's capability to dimerize. In 2002, a study by Porter and Dorman(19) reported that two site-directed mutants, I180N and Y224Ochre, contributed to a dominant negative effect when co-expressed and tested with WT VirF in a β-galactosidase reporter assays suggesting the protein dimerizes. Herein, we use a LexA monohybrid β-galactosidase reporter assay to study VirF dimerization directly and probe mutants of VirF's putative

dimerization α-helix that contributes to the dimerization activity and how these mutants affect DNA-binding activity for its cognate DNA promoters.

3.3 Results

3.3.1 Identification of WT VirF Dimerization in the LexA Monohybrid β-Galactosidase Reporter Assay

Based on the DNA-binding results obtained in **Chapter 2**, we hypothesized that VirF must have an active N-terminal dimerization domain. Two mutants studied in Porter and Dorman suggested that VirF actively dimerized *in vivo* (19) but unfortunately no studies have directly assayed VirF dimerization. To test our hypothesis, a previously optimized LexA monohybrid β-galactosidase assay was used. This intracellular assay utilizes the DNA-binding protein, LexA, from *E. coli* to bind to a DNA repressor sequence, *sulA,* which negatively controls the transcription and expression of β-galactosidase (25–27). LexA must dimerize to bind *sulA* and thus is unable to repress transcription of β-galactosidase without the dimerization domain. In this assay, a truncation mutant of LexA (LexA Δ1-87) harboring only the DNA-binding domain recovers its repressive DNA-binding activity when a dimerization-capable protein is fused to the C-terminus of LexA Δ1-87. We cloned and ligated *virF* into the test vector, pSR658 (LexA Δ1-87), which expresses the N-terminal DNA-binding domain of LexA. The assay utilizes an *E. coli* strain, SU101, which harbors the repressor sequence, *sulA*, which controls *lacZ* transcription. A diagram of the assay is presented in **Figure 3.2**. Upon LexA-VirF fusion protein (pSR658-VirF) expression in SU101 *E. coli*, there was a statistically significant ($p < 0.0001$) decrease in β-galactosidase activity compared to both the negative control, pSR658, and the positive control, pDD506 (**Figure 3.3**). Additionally, LexA-VirF and LexA-CAT (chloramphenicol acetyltransferase; pDD506)

samples displayed a statistically significant difference. This difference is likely due to differential expression of the two fusion proteins or because CAT has a stronger homodimerization affinity than WT VirF. Overall, these results indicate that VirF actively dimerizes, and this is the first reported result of this activity.

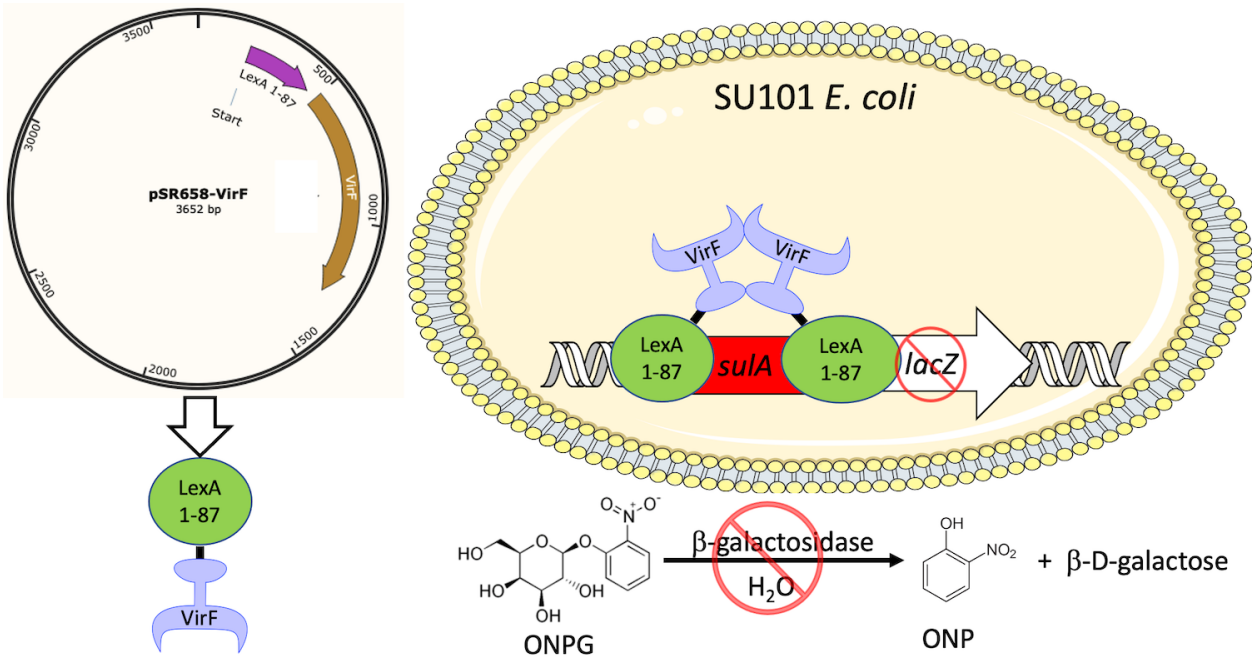


Figure 3.2: Diagram of VirF homodimerization as tested in the intracellular LexA monohybrid β-galactosidase reporter assay. LexA-VirF, expressed with IPTG from the pSR658-VirF plasmid, dimerizes and the LexA DNA-binding domain binds to *sulA* which represses *lacZ* transcription. This repression leads to a lack of ONP production resulting in a clear reaction. ONP produces a yellow color to the reactions and can be measured spectrophotometrically (OD_{420}).

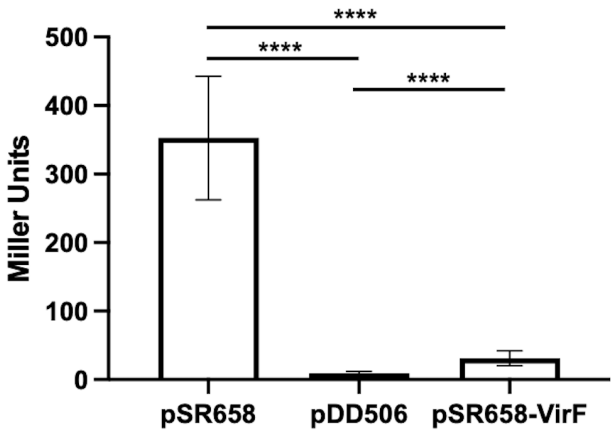


Figure 3.3: Testing WT VirF dimerization (via β-galactosidase suppression) in LexA monohybrid β-galactosidase reporter assay. Expression of LexA Δ1-87 (pSR658), LexA-VirF (pSR658-VirF), and LexA-CAT (pDD506) was induced with 1 mM IPTG in SU101 *E. coli*. All tests were performed with nine replicates (n=9). Student's T-tests were performed to compare all tests (**** $p<0.0001$).

3.3.2 Identification of Amino Acids within the VirF Dimerization α-Helix that contribute to Dimerization

An α-helix within the dimerization domain of AraC-family proteins has been implicated to be important for dimerization(6, 7, 9, 10, 12–15). First, we performed a sequence alignment (using Clustal Omega) of VirF with AraC from *E. coli* and ToxT from *V. cholerae* to identify the sequence of this putative dimerization α-helix (**Figure 3.4**)(28). Based on the alignment and the known dimerization helices of AraC and ToxT from their respective crystal structures(6, 15), we predicted a sequence of 21 amino acids of a potential dimerization helix within the VirF N-terminal domain. There was not much similarity between the three sequences despite significant similarities in the aligned, downstream DNA-binding domains suggesting this dimerization interaction is highly specific for individual AraC homologs.

```
AraC 142 QGEGRYSELLAINLL-EQLLLRRMEAINESLHPPMD 176
ToxT 134 FREINYSDEF--LKVFFSGFFSKVEKKYNSIFITDD 167
VirF 121 IKEMPFGKRKIYSLACLLSAVSDEEALYTSISIASS 156
           *  :..            .   *    *:    .
```

Figure 3.4: Sequence alignments of VirF with AraC and ToxT using Clustal Omega. Alignments include each proteins' C-terminal DNA-binding domain and the last 60 amino acids of the N-terminal domains. Primary sequences were obtained from Uniprot. (Uniprot Entry Numbers: AraC=P0A9E0, ToxT=P0C6D6, VirF=P0A2T1). Dimerization α-helices of ToxT (PDB ID: 3GBG) and AraC (PDB ID: 2ARC, 2ARA) are bolded and are based on their crystal structures. Underlined amino acids within ToxT (Prouty et al., 2005; Childers et al., 2011) and AraC (Soisson et al., 2007; LaRonde-LeBlanc and Wolberger, 2000) have been identified as important for protein dimerization. The predicted dimerization α-helix sequence of VirF is bolded red. Similarities of the residues are noted by * (fully similar), : (strongly similar), and . (weakly similar).

All non-alanines in this predicted region (18 of the 21 amino acids) were mutated to alanine via site-directed mutagenesis. Following mutagenesis, the ability of these LexA-VirF mutants to dimerize was tested in the β-galactosidase reporter assay (**Figure 3.5**). The data was normalized for every test based on its corresponding uninduced control (0 mM IPTG; 100% β-galactosidase activity; **Figure S3.1**). Most mutations within this region negatively affected β-galactosidase

suppression; however, three amino acids, Y132A, L137A, and L147A, resulted in > 50% β-galactosidase activity (100% activity indicates a maximal loss of dimerization activity as indicated by β-galactosidase suppression) when LexA-VirF mutant expression was induced at 0.1 mM IPTG. When mutant expression was induced with 1 mM IPTG, L137A again exhibited > 50% β-galactosidase activity. All three exhibited statistically significant differences from the corresponding WT reactions (0.1 mM IPTG mutant vs. 0.1 mM IPTG WT, and 1 mM IPTG mutant vs. 1 mM IPTG WT) with L137A presenting p values < 0.0001 for both inducer concentrations. To determine if these mutants would have an additive effect on disrupting β-galactosidase suppression, a triple mutant harboring all three mutations (Y132A/L137A/L147A) was generated. As expected, the triple mutant displayed a statistically significant loss in β-galactosidase suppression compared to WT and even exhibited a significant reduction compared to L137A when protein expression was induced with 1 mM IPTG ($p = 0.0033$). There was no significant difference between the corresponding 0.1 mM IPTG tests.

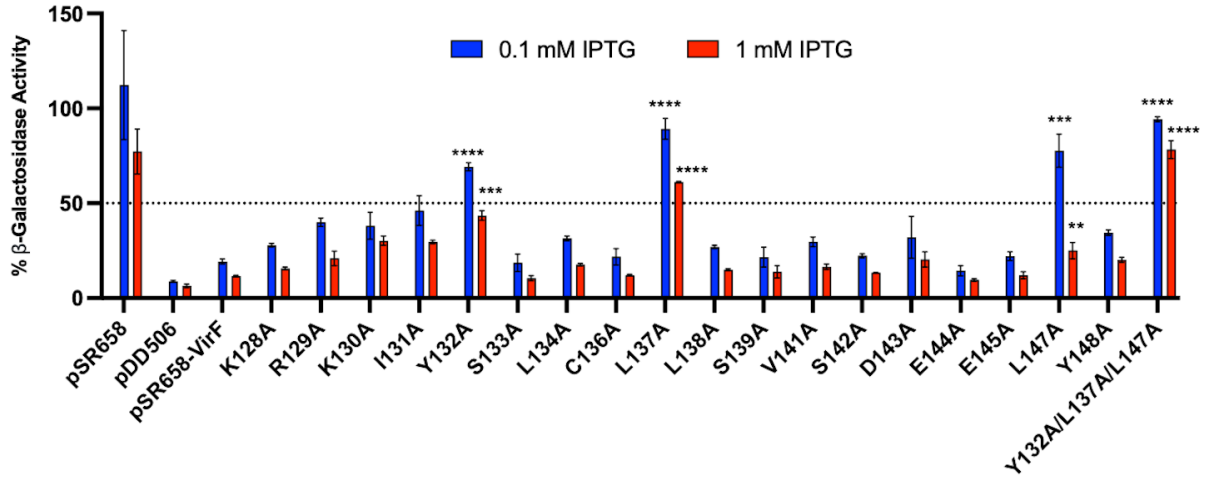


Figure 3.5: LexA monohybrid β-galactosidase reporter assay probing dimerization activities (via β-galactosidase suppression) of WT VirF and dimerization α-helix mutants. Proteins were expressed with 0.1 or 1 mM IPTG. All tests were performed in triplicate. Data was normalized to each test's uninduced control (0 mM IPTG; 100% β-galactosidase activity; **Figure S3.1**). Data plotted indicates the average of each triplicate and standard deviations are represented by error bars. Student's T-tests were used to compare either 0.1 mM or 1 mM IPTG tests of VirF mutants with the corresponding WT tests (pSR658-VirF) at 0.1 mM or 1 mM IPTG (* p=0.0332, ** p=0.0021, *** p=0.0002, **** p<0.0001).

Western blots were performed to compare expression levels of WT VirF with the Y132A, L137A, L147A, and triple mutants (**Figure 3.6**). It was possible that differential levels of expression of the mutants compared to WT VirF could be responsible for the assay results. Whole cell samples were obtained before the assay was performed and used in the western blots. Using the same samples, two western blots were performed and analyzed. In **Figure 3.6A**, Y132A and L147A displayed similar expression levels compared to WT. There was minimal expression for L137A, but a complete lack of expression seen with the triple mutant. These results were consistent with a second western blot (**Figure 3.6B**). Both L137A and triple mutant expressions were markedly lower than WT and the other mutants in both experiments, although the mutants exhibited lower relative expression in the second western blot. Additionally, in both blots, WT LexA-VirF expression is weaker than LexA-CAT (pDD506) suggesting that the significant differences in their dimerization activities is expected from differential expression levels. In summary, it is possible that mutations to L137 significantly effect expression thus present as a false positive (i.e., contributes to dimerization activity) in the LexA reporter assay as both a single and triple mutant. We suggest that VirF Y132A and L147A mutants have reduced dimerization activity via reduction of β-galactosidase suppression since they both exhibit similar expression levels compared to WT. However, the L137A and the triple mutant reduce β-galactosidase suppression likely through reduced expression relative to WT suggesting that these mutations, especially L137A, are not crucial to VirF dimerization activity.

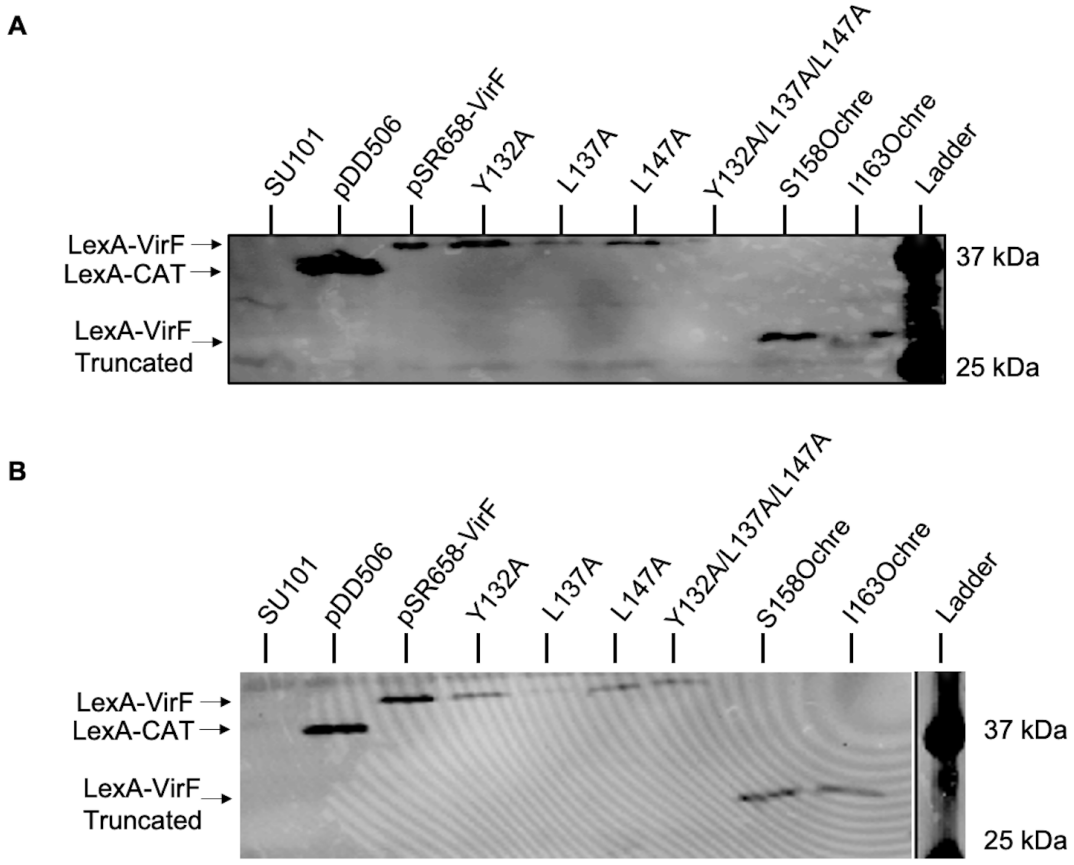


Figure 3.6: Western blots of WT VirF and dimerization mutants induced with 1 mM IPTG in the LexA monohybrid β-galactosidase reporter assay. Samples were obtained prior to the assay being performed. The primary antibody targeted the DNA-binding domain of LexA (rabbit anti-LexA antibody; Upstate ®) and the secondary antibody was a Cy5-labeled anti-rabbit antibody (Cytiva). Two gels were performed with the same samples (A and B) to control for gel loading errors or

3.3.3 Analysis of the Dimerization Activity of VirF N-Terminal Truncation Mutants

Due to VirF containing both N-terminal dimerization and C-terminal DNA-binding domains, we hypothesized that an N-terminal VirF truncation mutant, lacking the C-terminal DNA-binding domain, would be capable of dimerization. To test this hypothesis, we first prepared two truncation mutants at Y148 and I153 by incorporating an ochre stop codon (TAA) into the pSR658-VirF plasmid. These mutants were tested in the LexA reporter assay and compared to WT VirF (**Figure 3.7**). However, these constructs displayed significant reductions in β-

galactosidase suppression compared to WT VirF and at levels similar to the previously studied triple mutant (data not shown). It is probable that truncations at Y148 and I153 negatively affected the folding of the N-terminus, particularly within the dimerization α-helix. To correct for this, S158 and I163Ochre mutants were designed and obtained to include extra, flanking amino acids to allow for proper folding of the local secondary structure. Although they did not resemble WT VirF activity, both displayed < 50% β-galactosidase activity at both inducer concentrations (**Figure 3.7**). However, there was a significant decrease in β-galactosidase suppression when compared to WT. A western blot was also performed on samples obtained from the truncation mutants (**Figure 3.6**). Expression of S158Ochre appears equal to that of WT VirF (pSR658-VirF) but I163Ochre presented a less dense band on the western blot (**Figure 3.6A**) indicating it may be expressed to a lesser extent than WT VirF and S158Ochre VirF. Gel loading likely influenced the first western blot as the band was equivalently dense in the second western blot (**Figure 3.6B**). In summary, we suspect that VirF is capable of dimerizing without its C-terminal DNA-binding domain albeit with lower dimerization activity than full-length VirF. This is supported by both

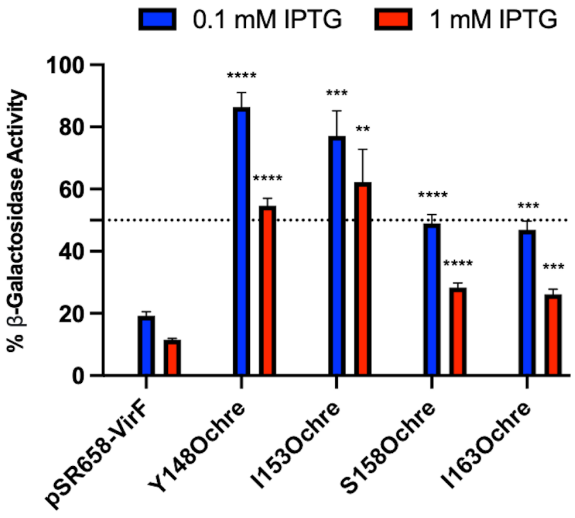


Figure 3.7: LexA monohybrid β-galactosidase reporter assay comparing dimerization activities (via β-galactosidase suppression) of WT VirF with truncation mutants. Proteins were expressed with 0.1 or 1 mM IPTG. All tests were performed in triplicate. Data was normalized to each test's uninduced control (0 mM IPTG; 100% β-galactosidase activity; **Figure S3.1**). Data plotted indicates the average of each triplicate and standard deviations are represented by error bars. Student's T-tests were used to compare either 0.1 mM or 1 mM IPTG tests of VirF mutants with the corresponding WT tests (pSR658-VirF) at 0.1 mM or 1 mM IPTG (* p=0.0332, ** p=0.0021, *** p=0.0002, **** p<0.0001).

S158Ochre and I163Ochre VirF truncation mutants displaying both similar expression levels as well as reductions in β-galactosidase suppression compared to WT VirF.

3.3.4 Purification of VirF Dimerization α-Helix Mutants

Alanine-scanning mutagenesis was performed on pBAD202-MALvirF to obtain the Y132A, L137A, L147A and triple mutants for purification and *in vitro* testing. After the mutants were verified via sequencing analysis, purifications were first attempted based on the protocols outlined in **Chapter 2 Methods: 2.5**. The MalE-VirF mutants were unable to be purified via this method since most of the protein resided in the earliest fractions from the Superdex 200 column purification (data not shown; left-most elution peak in the Superdex 200 chromatograms in **Figure S2.1**). Next, homologous protein expression in BS103 *S. flexneri* was performed as described in Emanuele and Garcia, 2015(18). As with the expression and purification from TOP10 *E. coli*, mutant purifications were unsuccessful following homologous expression (data not shown).

Due to the unsuccessful purifications, we hypothesized that mutant protein expression at 37 ˚C was causing significant aggregation or misfolding. To test for and alleviate this problem, protein expression was induced overnight at 16 ˚C in TOP10 *E. coli* harboring pBAD202-MALvirF mutant constructs instead of the previous expression conditions (37 ˚C for 5 hours). By reducing the temperature and increasing induction time, protein expression should be significantly slowed down to potentially allow for proper folding. Following overnight expressions, sonication, and centrifugation, the resulting lysate was incubated with amylose resin (NEB) overnight at 4 ˚C. Given that the 16 ˚C inductions of WT MalE-VirF expression resulted in low yield of protein (data not shown), we used this modified protocol to allow for as much protein binding to the resin as possible. On the next day, purifications were performed with a gravity column and a Superdex 200 GL 10/300 column (**Figure 3.8**). One mutant purification successfully produced a small

amount of pure Y132A MalE-VirF (elution peak 2 **Figure 3.8**; SDS-PAGE **Figure 3.9**). However, the yield (0.01 mg/L culture) was drastically lower than WT MalE-VirF (2.1 mg/L culture; see **Figure S2.1** to observe the Superdex 200 chromatogram for WT). To obtain a mutant that reduced VirF dimerization but allowed for a greater yield of protein, phenylalanine and serine were incorporated at Y132 instead of alanine. These mutations provided incremental changes to the residue which could positively affect protein folding and/or purification yield. Purifications were unsuccessful following both 37 and 16 ˚C inductions in TOP10 *E. coli* (data not shown). Interestingly, these mutants were tested in the LexA reporter assay, and it was observed that Y132S significantly reduced β-galactosidase suppression and Y132F slightly reduced β-galactosidase suppression relative to Y132A (**Figure 3.10**). In summary, incremental mutations to Y132 had no positive effect on protein purification but suggested that the aromatic moiety of Y132 is likely a major contributor to VirF dimerization.

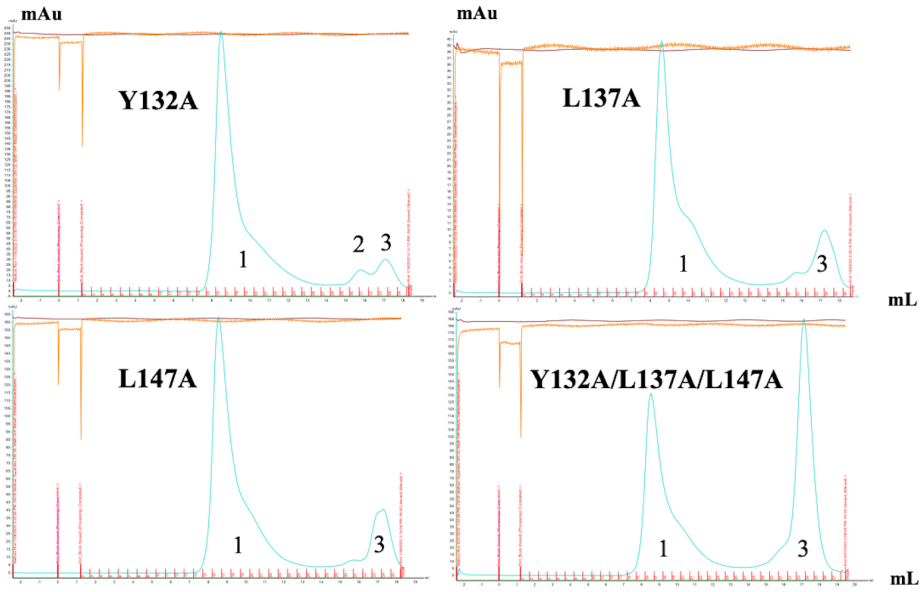


Figure 3.8: Chromatograms for MalE-VirF dimerization mutant purifications. Chromatograms are from size exclusion chromatography of Y132A, L137A, L147A, and triple mutant (Y132A/L137A/L147A). Size exclusion chromatography was performed using a Superdex 200 GL 10/300 column. The Y-axes (mAu at 280 nm) are not scaled proportionately to the WT purification. Peaks are labeled according to what is normally seen in SDS-PAGE from WT MalE-VirF purifications (1=impurities and "aggregated MalE-VirF", 2=pure MalE-VirF, 3=MalE).

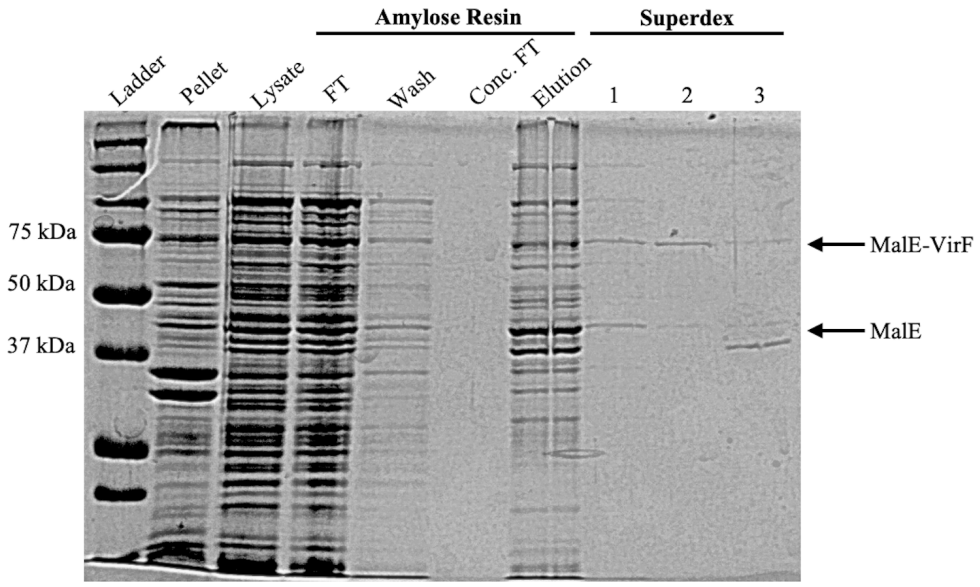


Figure 3.9: SDS-PAGE for the Y132A MalE-VirF purification. Locations of eluted proteins from Superdex column are noted by arrows (MalE or MalE-VirF). Abbreviations for lanes include "FT" flow-through and "Conc. FT" = Amicon Ultra-15 centrifugal filter unit concentrator flow-through.

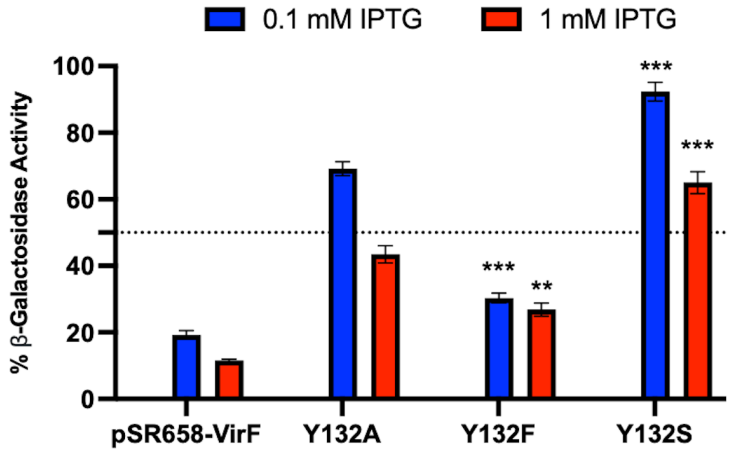


Figure 3.10: LexA monohybrid β-galactosidase reporter assay testing dimerization (via β-galactosidase suppression) of Y132F and Y132S VirF mutants compared to previously studied WT (pSR658-VirF) and Y132A VirF. All tests were performed in triplicate. Data was normalized to each test's uninduced control (0 mM IPTG; 100% β-galactosidase activity). Data plotted indicates the average of each triplicate and standard deviations are represented by error bars. Student's T-tests were used to compare either 0.1 mM or 1 mM IPTG tests of VirF mutants with the corresponding Y132A tests at 0.1 mM or 1 mM IPTG (* p=0.0332, ** p=0.0021, *** p=0.0002, **** p<0.0001). WT VirF data was included for reference.

As depicted in **Figure 3.8 and 3.9**, a significant amount of MalE-VirF mutant is found in chromatogram peak 1 after elution from the Superdex 200 column and a band can be visualized migrating to approximately 75 kDa (MalE-VirF molecular weight = 73 kDa) via SDS-PAGE. To further explore MalE-VirF dimerization mutant expressions, we performed a western blot to (**Figure 3.11**). Expression for all mutants was observed when protein expression was induced with 0.2% w/v arabinose for 5 hours at 37 ˚C (+) compared to uninduced samples (-). In addition, samples obtained from a Y132A/L137A/L147A triple mutant purification and confirmed MalE-VirF was eluted from both amylose resin gravity and Superdex 200 column purifications.

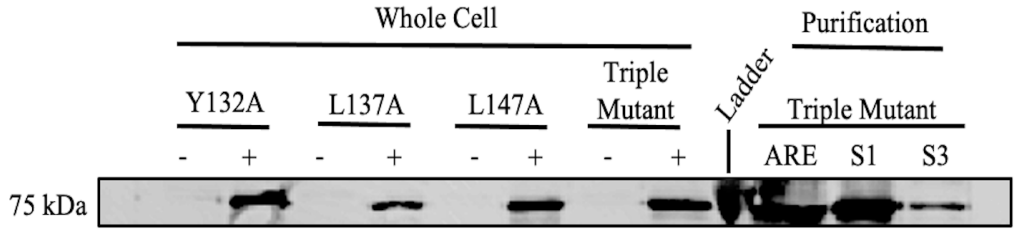


Figure 3.11: Western blot of MalE-VirF dimerization mutant expressions and purification of the VirF triple mutant (Y132A/L137A/L147A). Whole cell samples were obtained from TOP10 *E. coli* cultures before (-) and after (+) induction with 0.2% w/v arabinose at 37 ˚C for 5 hours. Samples from the purification were obtained from the amylose resin elution (ARE), Superdex 200 elution peak 1 (S1), and elution peak 3 (S3).

3.3.5 In vitro DNA-binding Activity of Y132A MalE-VirF

An EMSA was performed to examine how Y132A bound to *pvirB*, *picsA*, and *prnaG* (**Figure 3.12**). Due to the low yield from the protein purification, an FP assay testing binding for *pvirB* was not performed as it would exhaust the entire protein stock. At the highest testable concentration, 1.8 μM, a shift was observed when tested against *pvirB*. Based on the density of this shift compared to WT at the same concentration, Y132A binds less strongly to *pvirB* than WT. For *picsA*, there was no shift when Y132A was tested unlike WT where a shift was observed. As seen in **Chapter 2: Figure 2.13**, WT elicits binding shifts at concentrations around 1 μM when

interacting with *picsA* so we suspect this mutation affects its ability to bind to this promoter. It is possible that a shift would be visible at higher concentrations of Y132A. Surprisingly, Y132A bound to *prnaG* at 1.8 µM but at a lower migrated band. This shift was absent when [Y132A] was reduced to 0.5 µM. The occurrence of this lower shift indicates that Y132A can bind to *prnaG* at a lower ratio of protein:DNA suggesting a potential monomeric form of VirF binding to this promoter.

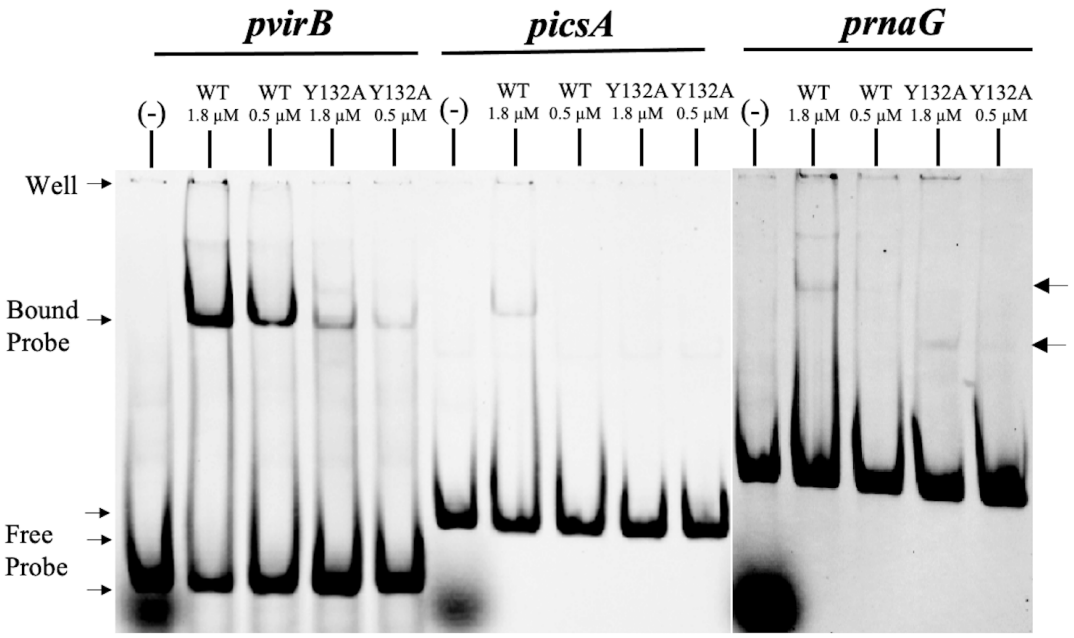


Figure 3.12: EMSA testing Y132A MalE-VirF binding to all three DNA promoters: *pvirB*, *picsA*, and *prnaG*. WT and Y132A were tested at 1.8 µM or 0.5 µM against 83 nM *pvirB*, 40 nM *picsA*, or 40 nM *prnaG*. All reactions were performed on one gel but the section of the gel testing *prnaG* binding was saturated to a greater extent to visualize the shifts.

3.3.6 Screening CCG Compounds against VirF Dimerization

Compounds that did not inhibit DNA-binding activity of MalE-VirF in the EMSA and FP(18), were examined in the LexA reporter assay (**Figure 3.13**). As these compounds showed inhibition *in vivo* with a VirF-driven β-galactosidase reporter assay(23, 24), one possible route for inhibition is by targeting protein dimerization. Compounds tested in the assay include 19615 (lead

DNA-binding inhibitor), 144092, 144143, and 24904. SU101 *E. coli* cultures harboring either pDD506 or pSR658-VirF were grown up with 10 μM of each inhibitor and 1 mM IPTG to induce protein expression. No compounds exhibited a significant loss of dimerization activity compared to the negative control (1 mM IPTG with no compound). If inhibition of dimerization activity was occurring, results would resemble the positive controls ("pDD506" and "pSR658-VirF" with 0 mM IPTG). In addition, the identified VirF inhibitors did not affect LexA-CAT dimerization activity. These reactions served a dual purpose as a counter screen to identify compounds which have activity against LexA DNA-binding activity. Seemingly, these compounds do not inhibit VirF dimerization activity at 10 μM.

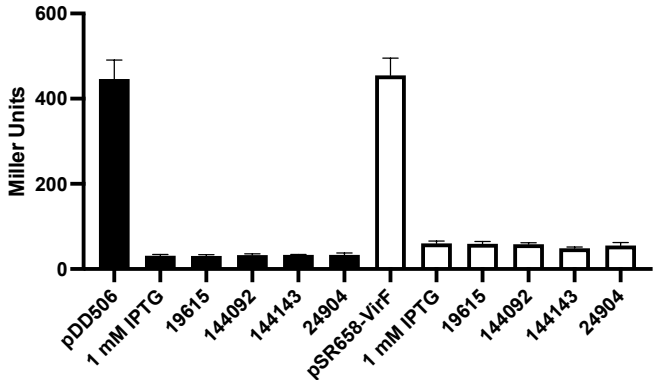


Figure 3.13: LexA monohybrid β-galactosidase reporter assay testing inhibition of VirF dimerization (via β-galactosidase suppression) with identified CCG inhibitors (19615, 144092, 144143, 24904). Controls include SU101 cells grown with pSR658-VirF or pDD506 with 0 or 1 mM IPTG for protein induction. Compounds were tested at 10 μM with 1 mM IPTG included. All tests were performed in triplicate. Data plotted indicates the average of each triplicate and standard deviations are represented by error bars.

3.3.7 Full-length VirF Modeling

In AraC-family proteins, there is a greater variation across the N-terminal domains than the C-terminal domains, so developing homology models for full length proteins based on homologous structures is difficult. To circumvent this, a structure prediction program, I-TASSER, was used to develop models of full-length VirF(29). The program generated multiple predictive structural models of VirF (**Figure 3.14**). Additionally, ToxT from *V. cholerae* (PDB ID: 3GBG) was identified as the closest structural similarity to VirF in the Protein Data Bank (PDB). The top generated VirF model was compared to the structure of ToxT and presented an RMSD of 1.19 Å

and a TM-score of 0.926. The TM-score is an assessment of topological structure similarity between different structures where 1 represents a perfect match and anything above 0.5 represents a good match.

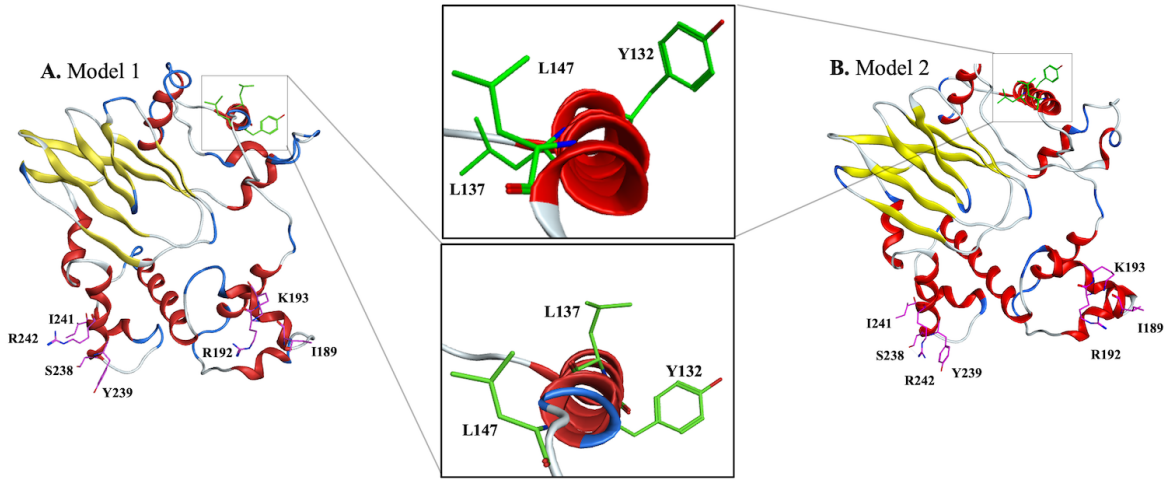


Figure 3.14: I-TASSER predicted structures of full-length VirF. The predicted structures were modeled using Molecular Operating Environment (MOE) The two structures presented the highest C-scores of -0.28 (**A.** Model 1) and -0.30 (**B.** Model 2). Boxed insets of the dimerization α-helix depict the spatial orientation of the three residues that contributed to dimerization in the LexA reporter assay.

Molecular Operating Environment (MOE) was used to analyze the predicted structures of VirF generated by I-TASSER (**Figure 3.14**)(30). The top two predicted structures of VirF displayed confidence scores (C-scores) of -0.28 (Model 1: **Figure 3.14A**) and -0.30 (Model 2: **Figure 3.14B**). C-scores range from -5 to 2 where the higher value infers higher confidence in the prediction. Firstly, both structures exhibit the two canonical helix-turn-helix motifs found in AraC-family proteins. Amino acids mutated and tested for their differential contributions to DNA-binding activity in **Chapter 2** were predicted by I-TASSER to be located within two α-helices as in the locations presented in **Figure 2.3.** The location of the residues and structural similarities to our VirF DNA-binding domain homology models validates these structure predictions as potential models for studying the full-length structure of VirF. When inspecting the dimerization domain in both models, the 21 amino acids we selected to mutate predominately resided on a single α-helix. For model 1 (**Figure 3.14A**), amino acids E144 through Y148 lie outside the predicted α-

helix and instead lie in a random coil. In Model 2 (**Figure 3.14B**), all 21 amino acids lie within a single α-helix. Of note, Y132A, L137A, and L147A all resided on the predicted dimerization α-helix in both models apart from L147 in model 1 (**Figure 3.14A**). There is little variation in the positioning of Y132 and L147 in both models. Y132A is solution-facing in both models while L147 can be found somewhere between solution-facing and protein-facing. However, considerable variation in spatial placement is observed with L137 between the two structures. L137 is solution-facing in model 1 (**Figure 3.14A**) but is protein-facing in model 2 (**Figure 3.14B**). Overall, these models exhibit high structural similarity to our homology models and another AraC homolog, ToxT, which makes these structures useful tools for studying VirF.

3.4 Discussion

3.4.1 Comparing VirF Dimerization Activity to other AraC-Family Homologs

AraC-family proteins, specifically those involved in controlling bacterial virulence, often contain N-terminal dimerization domains(1, 2). Characterization of AraC-family dimerization domains have been performed for AraC from *E. coli*(4, 6, 7), XylS from *Pseudomonas putida*(10), ExsA from *Pseudomonas aeruginosa*(8, 9), ToxT from *Vibrio cholerae*(5, 13–16) as well as HilD, HilC, RtsA from *Salmonella enterica* serovar Typhimurium(11, 12). While a few methods have been employed to study this domain, the LexA β-galactosidase reporter assay has most often been utilized(25–27). We hypothesized that VirF had a functional and active N-terminal dimerization domain. Employing the LexA reporter assay, we identified that full-length VirF has an active dimerization domain (**Figure 3.3**). To our knowledge, this is the first experimental report of VirF's ability to dimerize. Additionally, we showed that VirF is capable of homodimerization without the presence of its DNA-binding domain (I158Ochre and S163Ochre LexA-VirF mutants; **Figure**

3.7). Although there is still a significant difference in dimerization activity relative to WT, we suspect that the folding of local structure in this domain is crucial to dimerization. Dimerization activity would likely benefit from increased length to these truncations as already observed from Y148 to I163 truncations. N-terminal truncations of other AraC-family proteins have also exhibited dimerization in the LexA reporter assay. Active truncations include AraC(4, 6, 7), ToxT(5, 13, 14, 16), and ExsA (via heterodimerization of ExsA WT with an N-terminal Domain truncation)(9). These data suggest that the N-terminal and C-terminal domains can function independently of each other. Yet, attempts to demonstrate DNA binding by purified VirF DNA-binding domain constructs have been unsuccessful (data not shown; performed by Nicholas Ragazzone). While the VirF dimerization domain can function independently, the DNA-binding domain may require dimerization for proper folding and/or orientation of the DNA-binding domain to be active.

Studies have implicated a putative α-helix in the N-terminal domain that participates in the dimerization interface of AraC-family proteins(6, 7, 9, 12–16). Complete mutagenesis and follow-up studies of ToxT established D141 and F151 within this α-helix as crucial for dimerization(5, 13, 14). In 2019, models for the ToxT dimer interface pointed out two residues within this domain, G151 and K158, contributed to dimerization(16). ToxT G150L and K158E mutations significantly reduced dimerization in the LexA reporter assay through imparting steric clashes with valine (G151L) or introducing a charge repulsion (K158E) with the other subunit in the dimer interface(16). In XylS from *P. putida*, α-helix residues L193 and L194 were found crucial to XylS dimerization (10). More recently, Narm et al. reported that a proposed putative dimerization helix in HilD, HilC, and RtsA from *S. enterica* serovar Typhimurium influenced dimerization(12). In the LexA reporter assay, alanine mutations to L186 in HilD, and L171 and

F175 in HilC reduced dimerization activity relative to WT but also exhibited reductions in *hilA* transcription, a downstream transcription factor activated by both HilD, HilC, and RtsA(12). In AraC, a crystal structure depicted a leucine triad consisting of L150, L151, and L161 within the coiled-coil dimer interface(6). In addition, N154, E157, and Q158 all coordinate to a water within this interface(6). LaRonde-LeBlanc and Wolberger reported that a quadruple mutation within this α-helix, L150K/L151K/N154A/L161S, forced AraC into a monomer even in the presence of arabinose(7).

Alignments with ToxT and AraC allowed us to predict that a sequence of 21 amino acids belongs to this putative dimerization α-helix (**Figure 3.4**). Given the in-depth studies about this α-helix in other AraC homologs, it is not surprising that α-helix mutations in VirF significantly reduced dimerization activity. Three residues most significantly reduced dimerization activity relative to WT VirF: Y132, L137, and L147 (**Figure 3.5**). The contributions of leucines in AraC homolog dimerization have been well documented. These residues are often found in coiled-coil motifs, such as leucine zippers, and are involved in protein dimerization(31, 32). Y132 provides an unlikely contribution to VirF. Y132 may be contributing either hydrophobic or hydrogen bonding interactions to the dimer interface. If the VirF dimer interface contains an anti-parallel coiled-coil interaction like AraC (6, 7) or ToxT(16), then it is likely Y132 participates in hydrogen-bonding with S139 or S142 with the other α-helix subunit in the interface. In addition, tyrosines are uncharged at physiological pH and hence could also make hydrogen-bonding interactions with adjacent negatively charged sidechains of D143, E144, or E145 in VirF. Despite tyrosine's capability for hydrogen-bonding, Y132 more likely contributes hydrophobic or charge transfer interations with the other subunit in the dimer interface. This is supported by the observation that a Y132 mutation to serine knocked out dimerization while a mutation to phenylalanine presented

near WT levels of dimerization (**Figure 3.10**) maintaining the importance of the residue's aromatic moiety for VirF dimerization. This is not the first aromatic amino acid that has been identified as a contributor for dimerization: F151 in ToxT(13).

3.4.2 Mutations to the VirF Dimerization Domain Affect MalE-VirF Stability

Purifications of these mutants proved mostly unsuccessful. Differing expression and purification conditions were employed including different cell lines (TOP10 *E. coli* or BS103 *S. flexneri*), induction temperatures (16 ˚C or 37 ˚C), and overnight incubations with amylose resin. Purifications of L137A and L147A MalE-VirF dimerization mutants were unsuccessful, with the majority of protein eluting early on the chromatograms via Superdex 200 column (peak 1; **Figure 3.8**). When the Superdex 200 elution peak 1 fractions were concentrated and tested in the EMSA, there was no observable binding (data not shown) suggesting these fractions contain either misfolded or aggregated MalE-VirF protein. Mutations to L137 and L147, as well as to Y132, could potentially lead to effects on protein stability via increased protein aggregation or misfolding. Given that alanines are commonly found in α-helices, it was not expected that alanine mutations in this domain would destabilize the local secondary structure. Based on the predicted VirF models, L137 might be participating in crucial intramolecular interactions to help stabilize the α-helix or its orientation to the rest of VirF (**Figure 3.14A**). Furthermore, two western blots analyzing expression of L137A of LexA-VirF and MalE-VirF (**Figure 3.6 and 3.11**) display less intense bands with respect to the other mutants and/or WT. However, the L137A mutation does not completely knock out expression. Unlike L137A and L147A, the Y132A introduced a significant change to the protein electrostatically and hydrophobically. It is surprising that purification of only this mutant was successful. Both VirF models predict the location of Y132 to be solution-facing, so a mutation to this residue might not be completely destabilizing if this

wildtype residue does not contribute to other intramolecular protein interactions like L137 or L147. Even with incremental and modest mutations to Y132 via incorporations of phenylalanine and serine, purifications of these mutants were again unsuccessful. Overall, mutations to this domain significantly affect protein stability and imply the importance of the dimerization domain for full VirF activity.

3.4.3 Effect of Disrupting Dimerization on MalE-VirF DNA-Binding Activity

Fortunately, we successfully purified a small amount of Y132A MalE-VirF. Y132A exhibited DNA-binding activity for *pvirB* and *prnaG* but none for *picsA* (**Figure 3.12**). Although Y132A significantly reduced VirF dimerization relative to WT (**Figure 3.5**), there was an EMSA binding shift similar to that for WT for the *virB* promoter. With a visible reduction in the density of the shift relative to WT, the Y132A mutation may favor the protein being a monomer but not completely restrict it from dimerizing and binding to *pvirB*. Hence, the MalE-VirF•*pvirB* DNA-binding interaction likely occurs with a ratio of 2:1 protein to DNA. This is further supported by the binding of Y132A for the *rnaG* promoter. The occurrence of a lower binding shift more closely migrating with the free promoter in the EMSA, implicates that Y132A binds to *prnaG* as a monomer. Despite the binding observed for *prnaG*, none was seen with *picsA*. It cannot be ruled out that Y132A can bind to *picsA* as a monomer or dimer at higher concentrations than we were able to test. A speculative analysis of band migrations (**Figure 3.15**) supports these assumptions

based on the data despite the EMSA being a native gel lacking SDS to coat the proteins and that the protein•DNA complexes are not expected to run linearly with respect to size on this gel but rather logarithmically. The shifts observed in the EMSA suggest that WT MalE-VirF binds to all three promoters as a dimer initially. Later binding events likely consist of two VirF dimers binding to the *icsA* and *rnaG* promoters (4:1 ratio).

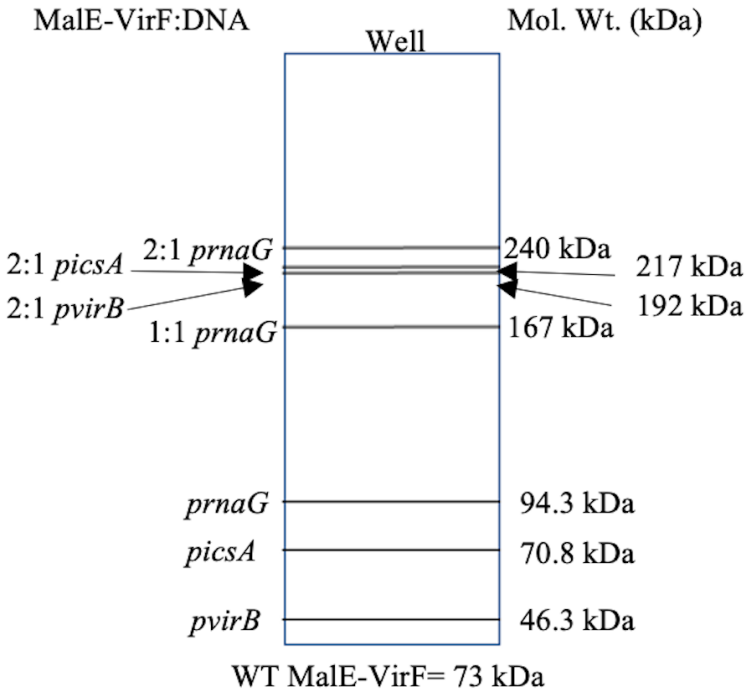


Figure 3.15: Speculative analysis of band migrations from the EMSA presented in **Figure 3.12**. Protein:DNA ratios are included for all three DNA promoters and the known molecular weights and locations of the three DNA promoters in the gel are indicated at the bottom of the gel.

Alterations to the dimerization domain have profound inhibitory effects on DNA-binding activity for AraC-family proteins. In 2014, Marsden et al. prepared a model of ExsA from *P. aeruginosa* based on the AraC dimer interface crystal structure (PDB 2ARC) to which they suggested L140 and L148 participated in ExsA "self-association"(9). It was shown that an ExsA L140A/L148A double mutant reduced binding of a second ExsA monomer to the *exsC* promoter suggesting that mutations to the ExsA dimerization domain affect the formation/stability of the wildtype $ExsA_2$•*exsC* promoter complex(9). Mutations to the ToxT dimerization domain, F151A and K158E, reduced DNA-binding activity for the *tcpA* promoter with complete loss of binding seen with K158E(13, 16). As depicted by a ToxT crystal structure bound to an unsaturated fatty

acid in the dimerization domain (PDB 6P7R), there is a perpendicular shift to the recognition helix of the second helix-turn-helix motif when compared to the MarA•*marRAB* crystal structure (PDB 1BL0)(16, 33). Cruite et al. assert that inhibition of the ToxT dimerization domain directly affects DNA-binding activity(16). These mutagenesis and crystallography studies provide evidence and rationale that dimerization is critical for DNA-binding activity. However, the mutagenesis studies do not completely match those seen with VirF since a reduction in dimerization introduced a monomeric binding event with *prnaG*. It is highly probably that VirF interacts with its DNA promoters distinctly from ToxT (potentially only as a dimer) or ExsA (two consecutive binding events with monomers). Physiologically, VirF likely requires full dimerization activity to bind to its DNA promoters to activate *icsA* and *virB* transcription. Further studies of how VirF dimerization affects DNA-binding need to be performed to solidify these claims. In addition, the role of this dimerization α-helix in transcription should be explored further. Due to the issues explained above with the potential aggregation or misfolding, we were unable to adapt our VirF transcription assay(23, 24) to test the contributions of these dimerization mutants in the assay (data not shown).

3.4.4 Screening Dimerization Inhibitors to Discover a Novel Anti-Virulence Therapeutic

Targeting dimerization activity of AraC-family promoters is expected to be highly specific. Unlike the DNA-binding domain, there is a low similarity found between different homologs in this family of proteins. For example, three residues show varying levels of similarity between AraC, ToxT, and VirF in our proposed alignment (**Figure 3.4**). Because of this low sequence similarity and the differing promoter binding and activation schemes by these proteins, it is likely the activity of this domain is highly organism-specific and therefore attractive for the discovery and development of specific anti-virulence inhibitors. A significant amount of work has already

been done to develop dimerization inhibitors for ToxT. Firstly, a small molecule, virstatin, was speculated and shown to inhibit ToxT dimerization and subsequent transcriptional activation of the *ctx* promoter(34, 35). Woodbrey et al. synthesized inhibitors based on the "folded" structure of cis-palmitoleate bound within the ToxT dimerization domain (PDB: 3GBG) and identified two compounds (3b and 5a) which exhibited greater than 10-fold better affinity for ToxT than virstatin(36). These inhibitors also had significant effects on ToxT transcriptional activation and *V. cholerae* autoagglutination activity(36). A follow-up study by the same group elaborated and improved upon these inhibitors and they identified a new lead that protected infant mice from *V. cholerae* infection(37).

Five compounds have shown reproducible inhibition of VirF in a VirF-driven β-galactosidase reporter assay(23, 24). Of which, 19615 was tested and confirmed as an inhibitor of the MalE-VirF•*pvirB* DNA-binding interaction but an initial structure-activity-relationship campaign was unsuccessful in finding a more potent analog(18). In addition, the other four inhibitors (144092, 144143, 153578, and 24904) were not identified as inhibiting DNA-binding hence could have activity against VirF dimerization. When tested at 10 μM in the LexA reporter assay (**Figure 3.13**), no compound exhibited activity against VirF dimerization when compared to the 0 and 1 mM IPTG controls. Unfortunately, we were unable to test 153578 in the assay due to insufficient sample. It cannot be ruled out that these compounds inhibit dimerization at much higher concentrations but since the VirF inhibitors exhibited IC_{50}'s ranging from 20-90 μM in the initial β-galactosidase reporter assay(23), some level of inhibition would be expected at 10 μM if that was the mechanism of inhibition. It is likely those compounds inhibit VirF transcriptional activation in some other capacity, likely through RNA polymerase recruitment. Despite these

negative results, this assay proved useful for screening these compounds and further optimization of it could be a valuable assay for a high-throughput screen.

3.4.5 Predicted VirF Models Suggest Modes for Dimerization

The dimerization domain does not have a high similarity and identity among AraC homologs, unlike the C-terminal DNA-binding domain. Here, the VirF N-terminal domain shows a 36% similarity and 14% identity when compared to the N-terminal domains of ToxT while exhibiting 23% similarity and 9% identity with the AraC N-terminal domain(30). This makes it difficult to prepare homology models for the dimerization domain or even the full-length protein. Hence, we used I-TASSER to generate 3D models for full-length VirF (**Figure 3.14**). VirF model 2 (**Figure 3.14B**) placed the 21 predicted amino acids (**Figure 3.4**) on one α-helix. The Y132 residue is solution-facing in both models which may indicate its contribution to the VirF dimerization interface. There was variation in placement of the L137 and L147 residues, specifically L137. The orientation of L137 in these models could be providing evidence for different VirF conformations when it is capable of dimerizing (L137A in solution-facing position; **Figure 3.14A**) or when the protein is monomeric (L137A is protein-facing; **Figure 3.14B**). Although this is speculation based on two predicted models, further studies of the dimerization domain are needed to understand how conformational changes of the dimerization helix promote or prevent dimerization.

Recently, AlphaFold deposited a predicted structure of VirF into their database based on the primary sequence from Uniprot (P0A2T1) to which we used for our studies and modeling(38, 39). Their model (**Figure 3.16**) also harbors two helix-turn-helix motifs in which contain our seven studied amino acids. Yet, within the dimerization domain, the dimerization α-helix is split in the middle with a random coil leaving Y132 and L137 on one helix and L147 on the other helix,

but the spatial locations of these amino acids are consistent with the I-TASSER generated structures. Upon inspection of AlphaFold's structure, they note that this random coil and second helix containing L147 have lower confidence levels than the other helix containing Y132 and L137. The per-residue confidence score (pLDDT) for this region falls between 70 and 90 whereas the first helix lies above 90 (100 is high confidence). Even though this is not a large decrease in confidence, it may be indicative that their model is not as accurate as those generated by I-TASSER (**Figure 3.14**). Furthermore, I-TASSER identified regions within the VirF primary sequence which exhibited α-helical or β-strand characteristics and it determined our predicted 21 amino acids all exhibited a strong likelihood to be within an α-helix. We find it highly unlikely that this α-helix would be split but in two by a random coil, but further structural studies are needed to support that claim.

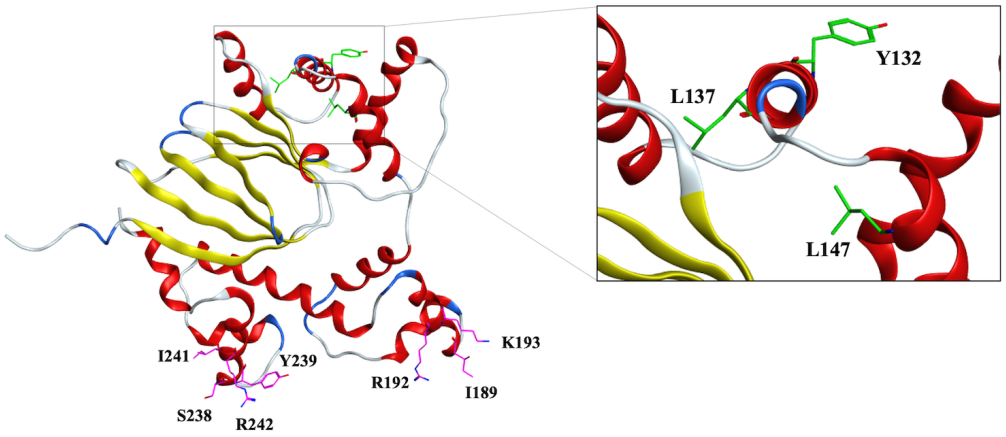


Figure 3.16: AlphaFold predicted structure of full-length VirF. The predicted structure was modeled using Molecular Operating Environment (MOE). Boxed insets of the dimerization α-helix depict the spatial orientation of the three residues that contributed to dimerization in the LexA reporter assay.

3.4.6 Conclusions

To our knowledge, this is the first study that directly tested and confirmed that VirF dimerizes. Using alignments and comparisons to crystallized, and well-studied homologs, AraC and ToxT, we selected 21 amino acids that we suggest lie in a putative dimerization α-helix

implicated in these homologs and others. In a LexA reporter assay, alanine-mutations to Y132, L137, and L147 displayed significant reductions in VirF dimerization activity compared to WT. Upon purification of these proteins for *in vitro* testing, only Y132A was successfully purified albeit with a significant reduction in yield compared to WT MalE-VirF. Mutations to this domain impart negative effects to MalE-VirF stability through affecting folding or inducing protein aggregation. Y132A exhibited a distinct binding scheme compared to WT for all three promoters especially *prnaG*. The reduced dimerization activity of Y132A MalE-VirF allowed the protein to bind as a monomer to *prnaG* whereas WT binds likely as a dimer at the same tested concentration. This data suggests that the protein likely binds to its DNA promoters as a dimer to activate transcription physiologically and that the Y132A mutation shifts equilibrium of MalE-VirF to a monomeric state but not completely restricting it from dimerizing (binds similarly to WT with *pvirB*). Full-length models of VirF predict the locations of these three residues and suggest their roles in dimerization (Y132), stability (L137 and L147), and protein conformation (L137).

3.5 Methods

3.5.1 Materials

Buffer components and chemicals were purchased from ThermoFisher. Specific reagents and biological products (restriction enzymes, antibodies, etc.) that were not purchased from ThermoFisher are indicated in parentheses. DNA oligonucleotides were purchased from Integrated DNA Technologies. Equipment utilized for these experiments was purchased from varying companies which are indicated in parentheses throughout.

3.5.2 LexA Monohybrid β-Galactosidase Reporter Assay

VirF homodimerization was studied using a LexA monohybrid β-galactosidase reporter assay as previously described(25, 40). The open reading frame of *virF* was PCR amplified (**Table 3.1**) from the MalE-VirF expression vector (pBAD202-MALvirF) using Phusion DNA Polymerase (NEB) with the DNA oligonucleotide primers found in **Table 3.2** according to manufacturer instructions. Reactions included 0.5 µM forward primer (VirF PCR SacI Forward), 0.5 µM reverse primer (VirF PCR PstI Reverse), ~250 ng template plasmid (pBAD202-MALvirF), 1x High Fidelity buffer, 200 µM dNTPs, and 1-unit Phusion polymerase (similar reaction volumes can be found in **Chapter 2 Methods**: **2.5**). Following PCR, *virF* and the vector, pSR658, were digested with SacI and PstI restriction enzymes (NEB) according to manufacturer instructions. Digestion reactions (50 µL total volume) were prepared with ~1 µg *virF* or pSR658, 5 µl 10x rCutsmart Buffer, 1 µL SacI-HF (20-units; NEB), and 1 µL PstI-HF (20-units; NEB) and were allowed to incubate at 37 ˚C for 15 min prior to 2% agarose gel electrophoresis and extraction (QIAGEN). After gel extraction, the digested products of *virF* and pSR658 were ligated with T4 DNA Ligase (NEB) according to manufacturer instructions. Ligation reactions were prepared with a 1:3 molar ratio of vector (pSR658) to insert (*virF*) included reaction volumes of 10 µL digested pSR658, 5 µL digested *virF*, 2 µL T4 DNA Ligase buffer, and 1 µL T4 DNA Ligase (400-units; NEB). The reaction was allowed to incubated overnight at 16 ˚C prior to transformation of TOP10 *E. coli* (**Table 3.3**) and plating on LB/Agar plates supplemented with tetracycline (15 µg/mL) (**Chapter 2 Methods: 2.5**). Resultant colonies were grown overnight in 2xTY media supplemented with tetracycline and were miniprepped (QIAGEN) the following day. Sequencing analysis was performed to verify *virF* ligation into pSR658 using the oligonucleotide primers in **Table 3.2**.

Table 3.1: PCR cycles used cloning and amplifying *virF* from pBAD202-MALvirF. Temperature, time at the corresponding temperature, and number of individual cycles are included below.

***virF* PCR (Phusion DNA Polymerase)**		
Temperature (˚C)	**Time (sec or min)**	**Cycles**
95	30 sec	1
95	30 sec	30
54	50 sec	30
72	110 sec	30
72	10 min	1
4	Hold Overnight	1

Table 3.2: DNA oligonucleotide primers designed and used for cloning and amplification of *virF* from the MalE-VirF (pBAD202-MALvirF) expression plasmid. Restriction cutsites are italicized and *virF* sequences are underlined. Sequencing primers are also included.

Oligonucleotide	Sequence (5' → 3')
VirF PCR SacI Forward	GGTGGT*GAGCTC*GGACATAAAAACAAAATAG
VirF PCR PstI Reverse	GGCGGC*CTGCAG*TTAAAATTTTTTATGATATAAG
LexA Sys Seq Forward	AAGGCGCTGGCACGCAAAGG
LexA-fusion Seq Reverse	CTGCGTTCTGATTTAATCTG

Table 3.3: *E. coli* cell lines used for the LexA Monohybrid β-Galactosidase Reporter Assay (SU101), cloning, and expression of MalE-VirF mutant proteins.

E. coli Strain	Genotype
TOP10	F- *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ *lac*X74 *rec*A1 *ara*D139 Δ(*araleu*)7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G
SU101	*lexA71:Tn5(Def)sulA211Δ(lacIPOZYA)169 /F'lacIqlacZΔM15::Tn9*

SU101 *E. coli* (**Table 3.3**) were transformed with either pSR658 (negative control; LexA Δ1-87), pSR658-VirF (test; LexA Δ1-87 in frame with WT or mutant VirF), or pDD506 (positive control; LexA-CAT). Transformed SU101 *E. coli* were grown overnight at 37 ˚C in LB media (10 g bactotryptone, 5 g yeast extract, 5 g NaCl, 1 mL 1 M NaOH per 1 L MilliQ water) supplemented with tetracycline (15 µg/mL), kanamycin (50 µg/mL), and chloramphenicol (30 µg/mL). On the following day, 10 mL of LB were inoculated with 0.2 mL overnight culture and supplemented with the appropriate antibiotics and either 0, 0.1, or 1 mM IPTG to induce expression of the LexA controls or LexA-VirF fusion proteins. The cells were grown at 37 ˚C for approximately 5 hours

until they reached OD_{600}~0.5. Final optical densities were recorded. At the appropriate density, 0.5 mL culture was added to 0.5 mL of 1x Z buffer (60 mM $Na_2HPO_4 \cdot 7H_2O$, 10 mM KCl, 1 mM $MgSO_4 \cdot 7H_2O$, pH 7.0) supplemented with 50 mM β-mercaptoethanol and 10 µL of 0.1% sodium dodecyl sulfate (SDS). Using a pipet, one drop of chloroform was added to each reaction and then vortexed vigorously for 10 seconds. Following ~5 minutes at room temperature, 0.2 mL of ortho-nitrophenyl β-galactoside (ONPG; 4 mg/mL in 1x Z buffer without β-mercaptoethanol) was added to each reaction and vortexed briefly. The reactions incubated for 10 minutes at room temperature and were terminated with 0.5 mL of 1 M Na_2CO_3. The terminated reactions were centrifuged at 13,000xg for 2 minutes to remove cellular debris. The absorbance of the product, o-nitrophenyl, was measured at OD_{420}. Miller units were calculated using **Equation 3.1**. All controls and test reactions were performed in triplicate.

Equation 3.1: Miller Units = $(OD_{420} / (T * V * OD_{600})) * 1000$

T = Time of reaction (10 min), V = Volume of culture added (0.5 mL)

3.5.3 Sequence Alignment and Site-Directed Mutagenesis of the Dimerization Domain

VirF, ToxT, and AraC primary sequences were obtained from Uniprot(38). The sequences were aligned and evaluated using Clustal Omega (28) to determine the approximate sequence for the VirF dimerization α-helix. Site-directed mutagenesis of the VirF dimerization α-helix was performed using DNA oligonucleotide primers found in **Table 3.4**. Either an alanine (GCG), serine (AGC), phenylalanine (TTT), or ochre stop (TAA) codon was incorporated into *virF* within the test plasmids, pSR658-VirF or pBAD202-MALvirF. MalE-VirF numbering refers to the primary sequence of VirF from *Shigella flexneri*. Two step PCR was used on either pBAD202-MALvirF (for expression of MalE-VirF dimerization mutants) or pSR658-VirF (for expressing LexA-VirF dimerization mutants) with PFU Turbo DNA Polymerase (Agilent) in a MiniAmp™

Table 3.4: DNA oligonucleotide primers designed and used for site-directed mutagenesis of pSR658-VirF and pBAD202-MALvirF. Alanine, phenylalanine, serine, or stop codons incorporated into the expression plasmids are bolded. Corresponding annealing temperatures used in the two-step PCR mutagenesis are included.

Oligonucleotide	Annealing Temperature (˚C)	Sequence (5' → 3')
K128A Mut Forward	56	GAGATGCCTTTCGGC**GCG**AGAAAGATCTATAG
K128A Mut Reverse	56	CTATAGATCTTTCT**CGC**GCCGAAAGGCATCTC
R129A Mut Forward	53	GCCTTTCGGCAAA**GCG**AAGATCTATAG
R129A Mut Reverse	53	CTATAGATCTT**CGC**TTTGCCGAAAGGC
K130A Mut Forward	55	CCTTTCGGCAAAAGA**GCG**ATCTATAGTTTAGC
K130A Mut Reverse	55	GCTAAACTATAGAT**CGC**TCTTTTGCCGAAAGG
I131A Mut Forward	53	CGGCAAAAGAAAG**GCG**TATAGTTTAGC
I131A Mut Reverse	53	GCTAAACTATA**CGC**CTTTCTTTTGCCG
Y132A Mut Forward	55	CAAAAGAAAGATC**GCG**AGTTTAGCTTGCC
Y132A Mut Reverse	55	GGCAAGCTAAACT**CGC**GATCTTTCTTTTG
Y132F Mut Forward	50	CAAAAGAAAGATC**TTT**AGTTTAGCTTGCC
Y132F Mut Reverse	50	GGCAAGCTAAACT**AAA**GATCTTTCTTTTG
Y132S Mut Forward	53	CAAAAGAAAGATC**AGC**AGTTTAGCTTGCC
Y132S Mut Reverse	53	GGCAAGCTAAACT**GCT**GATCTTTCTTTTG
S133A Mut Forward	54	AAAAGAAAGATCTAT**GCG**TTAGCTTGCCTTTTA
S133A Mut Reverse	54	TAAAAGGCAAGCTAA**CGC**ATAGATCTTTCTTTT
L134A Mut Forward	53	AAAGATCTATAGT**GCG**GCTTGCCTTTTA
L134A Mut Reverse	53	TAAAAGGCAAGC**CGC**ACTATAGATCTTT
C136A Mut Forward	54	GATCTATAGTTTAGCT**GCG**CTTTTATCAGCTG
C136A Mut Reverse	54	CAGCTGATAAAAG**CGC**AGCTAAACTATAGATC
L137A Mut Forward	55	TAGTTTAGCTTGC**GCG**TTATCAGCTGTTT
L137A Mut Reverse	55	AAACAGCTGATAA**CGC**GCAAGCTAAACTA
L138A Mut Forward	58	GTTTAGCTTGCCTT**GCG**TCAGCTGTTTCTG
L138A Mut Reverse	58	CAGAAACAGCTGA**CGC**AAGGCAAGCTAAAC
S139A Mut Forward	57	GCTTGCCTTTTA**GCG**GCTGTTTCTGATG
S139A Mut Reverse	57	CATCAGAAACAGC**CGC**TAAAAGGCAAGC
V141A Mut Forward	57	GCCTTTTATCAGCT**GCG**TCTGATGAGGAAG
V141A Mut Reverse	57	CTTCCTCATCAGA**CGC**AGCTGATAAAAGGC
S142A Mut Forward	57	TTTTATCAGCTGTT**GCG**GATGAGGAAGCTTT
S142A Mut Reverse	57	AAAGCTTCCTCATC**CGC**AACAGCTGATAAAA
D143A Mut Forward	56	ATCAGCTGTTTCT**GCG**GAGGAAGCTTTATA
D143A Mut Reverse	56	TATAAAGCTTCCTC**CGC**AGAAACAGCTGAT
E144A Mut Forward	54	CAGCTGTTTCTGAT**GCG**GAAGCTTTATATAC
E144A Mut Reverse	54	GTATATAAAGCTTC**CGC**ATCAGAAACAGCTG
E145A Mut Forward	54	CTGTTTCTGATGAG**GCG**GCTTTATATACTTC
E145A Mut Reverse	54	GAAGTATATAAAGC**CGC**CTCATCAGAAACAG
L147A Mut Forward	55	CTGATGAGGAAGCT**GCG**TATACTTCGATATC
L147A Mut Reverse	55	GATATCGAAGTATA**CGC**AGCTTCCTCATCAG
Y148A Mut Forward	55	GATGAGGAAGCTTTA**GCG**ACTTCGATATCGATAG
Y148A Mut Reverse	55	CTATCGATATCGAAGT**CGC**TAAAGCTTCCTCATC
S158Ochre Mut Forward	50	GCTTCTTCTCTT**TAA**TTTTCTGATCAGATAAG
S158Ochre Mut Reverse	50	CTTATCTGATCAGAAAA**TTA**AAGAGAAGAAGC
I163Ochre Mut Forward	50	GTTTTTCTGATCAG**TAA**AGGAAGATTGTTG
I163Ochre Mut Reverse	50	CAACAATCTTCCT**TTA**CTGATCAGAAAAAC

Plus Thermal Cycler (ThermoFisher) according to manufacturer instructions (**Table 3.5**).

Protocols and reaction volumes for the site-directed mutagenesis can be found in **Chapter 2 Methods: 2.5**. All plasmids and mutants were confirmed by sequencing analysis.

Table 3.5: Two-step PCR cycles used for site-directed mutagenesis of the pSR658-VirF and pBAD202-MALvirF vectors. Temperature, time at the corresponding temperature, and number of individual cycles are included below.

1st PCR Cycle		
Temperature (˚C)	**Time (sec or min)**	**Cycles**
95	30 sec	1
95	30 sec	4
Annealing Temperature (**Table 3.4**)	1 min	4
68	8 min (pSR658-VirF) or 14 min (pBAD202-MALvirF)	4
4	Held until next PCR cycle	1
2nd PCR Cycle		
95	30 sec	1
95	1 min	18
Annealing Temperature (**Table 3.4**)	1 min	18
68	8 min (pSR658-VirF) or 14 min (pBAD202-MALvirF)	18
68	10 min	1
4	Hold Overnight	1

3.5.4 Purification of VirF Dimerization Mutants

WT MalE-VirF was expressed and purified as identically described in **Chapter 2 Methods: 2.5**. For the dimerization mutants, TOP10 *E. coli* cells (**Table 3.3**) harboring the corresponding pBAD202-MALvirF mutant plasmids were grown, and protein was expressed as previously described in **Chapter 2**. Following sonication and centrifugation, the resulting lysate was incubated with 4 mL of amylose resin (NEB) while shaking gently at 4 ˚C overnight. Following overnight incubation, the lysate-resin mixture was loaded onto a gravity column and washed with 50 mL binding buffer (20 mM Tris-HCl, 500 mM NaCl, 1 mM EDTA, 1 mM DTT, pH 7.4). To elute the protein, 50 mL of elution buffer (20 mM Tris-HCl, 500 mM NaCl, 1 mM EDTA, 1 mM DTT, 10 mM maltose, 5% glycerol, pH 7.4) was gently applied to the column as to

not disturb the resin bed. The eluent was concentrated to approximately 600 µL using Amicon Ultra-15 MWCO 10 kDa centrifugal filter units and then filtered to remove any precipitate. A Superdex 200 GL 10/300 was used to separate the protein based on size as previously described in **Chapter 2**. Purified MalE-VirF dimerization mutants were dialyzed overnight at 4 ˚C in dialysis buffer (20 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA, 5 mM DTT, 40% glycerol, pH 7.4) with a Slide-A-Lyzer 10 kDa MWCO dialysis cassette (ThermoFisher). Protein purifications were confirmed via SDS-PAGE and protein concentrations were determined using a Bradford assay (Bio-Rad) compared to bovine serum albumin (BSA) standards.

3.5.5 Western Blot

Samples of all VirF samples (obtained from LexA monohybrid assay or pBAD202-MALvirF expressions) were loaded onto a 10% polyacrylamide gel (containing a 6% stacking gel) and separated based on size via SDS-PAGE. Proteins were transferred to a polyvinylidene difluoride (PVDF) membrane using a Mini Trans-Blot Cell (Bio-Rad) with 1x transfer buffer (25 mM Tris Base, 200 mM glycine, 20% methanol) for 2 hours at 300 milli-amps in a cold room (4˚C). After the transfer, the membrane was washed with water prior to adding 20 mL of blocking buffer (50 mM Tris-HCl, 150 mM NaCl, 0.1% Tween 20, 3% BSA, 0.02% sodium azide, pH 7.5). The blocking solution and membrane were allowed to shake for 30 min at room temperature. Primary antibody (anti-LexA DNA-binding domain antibody from Upstate ® or Maltose Binding Protein Rabbit anti-Tag, Polyclonal, Invitrogen) was prepared with a 1:1000 w/v ratio in blocking solution. After the blocking step, the buffer was discarded, and 20 mL of primary antibody was poured onto the membrane and allowed to gently shake overnight at 4˚C. On the following day, the primary antibody solution was discarded, and the membrane was rinsed five times with wash buffer (50 mM Tris-HCl, 150 mM NaCl, pH 7.5). Following the washes, 20 mL of secondary

antibody solution was applied and allowed to incubate in the dark while shaking gently for 1 hour at room temperature. The secondary antibody (ECL Plex™ G-A-R IgG, Cy®5 from Cytiva) was diluted 1:1250 w/v in wash buffer. Following incubation, the membrane was rinsed five times with wash buffer followed by rinses with MilliQ water. The membrane was imaged using a Typhoon 9200 Molecular Imager (Molecular Dynamics) with excitation at 607 nm and reading the emission at 710 nm.

3.5.6 Electrophoretic Mobility Shift Assay

EMSAs were performed identically to the protocol described in **Chapter 2 Methods: 2.5.**

3.5.7 Full-length VirF Structure Predictions and Modeling

The full-length primary sequence of VirF was acquired from UniProt (38) and was inputted into I-TASSER (29) to generate predictive structural model of the protein. Molecular Operating Environment (MOE) (30) was used to visualize and manipulate the generated models.

3.5.8 Screening Compounds in the LexA Monohybrid β-Galactosidase Reporter Assay

To screen for compound inhibition of VirF dimerization activity, compounds were included during grow-ups of the cultures and tested in the LexA reporter assay (protocol is previously described). Compounds tested were identified as VirF inhibitors in Emanuele et al., 2014 and Emanuele and Garcia, 2015(18, 23). Cultures of SU101 *E. coli* harboring the pSR658-VirF or pDD506 plasmids were grown with 10 μM of VirF inhibitor (19615, 144092, 144143, or 24904), appropriate antibiotics, and 1 mM IPTG to induce protein expression. Negative controls of cultures containing 0 mM IPTG were also included. The reactions involving the pDD506 plasmid served as a counter screen to determine if inhibition occurred via blocking LexA DNA-

binding activity. All reactions were performed in triplicate and Miller Units were obtained for all tests (**Equation 3.1**).

3.6 Supplemental Figures

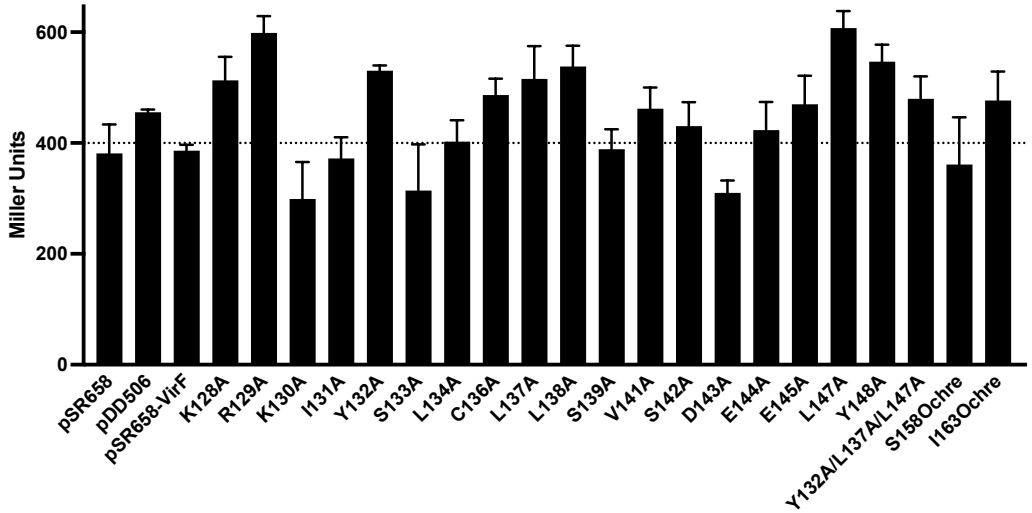


Figure S3.1: Raw data for uninduced samples, prior to normalization, of WT VirF and the dimerization mutants in the LexA monohybrid β-galactosidase reporter assay. Uninduced samples (0 mM IPTG) indicate 100% β-galactosidase activity. All tests were performed in triplicate. Data plotted indicates the average of each triplicate and standard deviations are represented by error bars.

3.7 References

1. Martin RG, Rosner JL. 2001. The AraC transcriptional activators. Curr Opin Microbiol 4:132–137.

2. Gallegos M, Schleif R, Bairoch A, Hofmann K, Ramos JL. 1997. AraC/XylS family of transcriptional regulators. Microbiol Mol Biol Rev 61:393–410.

3. Tobin JF, Schleif RF. 1990. Transcription from the rha Operon psr Promoter. J Mol Biol 211:1–4.

4. Bustos SA, Schleif RF. 1993. Functional domains of the AraC protein. Proc Natl Acad Sci U S A 90:5638–5642.

5. Prouty MG, Osorio CR, Klose KE. 2005. Characterization of functional domains of the Vibrio cholerae virulence regulator ToxT. Mol Microbiol 58:1143–1156.

6. Soisson SM, MacDougall-Shackleton B, Schleif R, Wolberger C. 1997. Structural Basis for Ligand-Regulated Oligomerization of AraC. Science (80-) 276:421–425.

7. Laronde-Leblanc N, Wolberger C. 2000. Characterization of the Oligomeric States of Wild Type and Mutant AraC. Biochemistry 39:11593–11601.

8. Brutinel ED, Vakulskas CA, Yahr TL. 2010. ExsD Inhibits Expression of the Pseudomonas aeruginosa Type III Secretion System by Disrupting ExsA Self-Association and DNA Binding Activity. J Bacteriol 192:1479–1486.

9. Marsden AE, Schubot FD, Yahr TL. 2014. Self-association is required for occupation of adjacent binding sites in Pseudomonas aeruginosa type III secretion system promoters. J Bacteriol 196:3546–3555.

10. Ruíz R, Marqués S, Ramos JL. 2003. Leucines 193 and 194 at the N-terminal domain of the XylS protein, the positive transcriptional regulator of the TOL meta-cleavage pathway, are involved in dimerization. J Bacteriol 185:3036–3041.

11. Paredes-Amaya CC, Valdés-García G, Juárez-González VR, Rudiño-Piñera E, Bustamante VH. 2018. The Hcp-like protein HilE inhibits homodimerization and DNA binding of the virulence-associated transcriptional regulator HilD in Salmonella. J Biol Chem 293:6578–6592.

12. Narm K-E, Kalafatis M, Slauch JM. 2020. HilD, HilC, and RtsA Form Homodimers and Heterodimers To Regulate Expression of the Salmonella Pathogenicity Island I Type III Secretion System. J Bacteriol 202:1–13.

13. Childers BM, Cao X, Weber GG, Demeler B, Hart PJ, Klose KE. 2011. N-terminal Residues of the Vibrio cholerae Virulence Regulatory Protein ToxT Involved in Dimerization and Modulation by Fatty Acids. J Biol Chem 286:28644–28655.

14. Childers BM, Weber GG, Prouty MG, Castaneda MM, Peng F, Klose KE. 2007. Identification of Residues Critical for the Function of the Vibrio cholerae Virulence Regulator ToxT by Scanning Alanine Mutagenesis. J Mol Biol 367:1413–1430.

15. Lowden MJ, Skorupski K, Pellegrini M, Chiorazzo MG, Taylor RK, Kull FJ. 2010. Structure of Vibrio cholerae ToxT reveals a mechanism for fatty acid regulation of virulence genes. Proc Natl Acad Sci U S A 107:2860–2865.

16. Cruite JT, Kovacikova G, Clark KA, Woodbrey AK, Skorupski K, Kull FJ. 2019. Structural basis for virulence regulation in Vibrio cholerae by unsaturated fatty acid components of bile. Commun Biol 2:1–9.

17. Tran CN, Giangrossi M, Prosseda G, Brandi A, Di Martino ML, Colonna B, Falconi M. 2011. A multifactor regulatory circuit involving H-NS, VirF and an antisense RNA modulates transcription of the virulence gene icsA of Shigella flexneri. Nucleic Acids Res 39:8122–8134.

18. Emanuele AA, Garcia GA. 2015. Mechanism of Action and Initial, In Vitro SAR of an Inhibitor of the Shigella flexneri Virulence Regulator VirF. PLoS One 10:1–18.

19. Porter ME, Dorman CJ. 2002. In Vivo DNA-Binding and Oligomerization Properties of the Shigella flexneri AraC-Like Transcriptional Regulator VirF as Identified by Random and Site-Specific Mutagenesis. J Bacteriol 184:531–539.

20. Tobe T, Yoshikawa M, Mizuno T, Sasakawa C. 1993. Transcriptional Control of the Invasion Regulatory Gene virB of Shigella flexneri: Activation by VirF and Repression by H-NS. J Bacteriol 175:6142–6149.

21. Ragazzone NJ, Dow GT, Garcia A. 2022. Elucidation of Key Interactions between VirF and the virB Promoter in Shigella flexneri Using E. coli MarA- and GadX-Based Homology Models and In Vitro Analysis of the DNA-Binding Domains of VirF and MarA. J Bacteriol Accepted-A.

22. Koppolu V, Osaka I, Skredenske JM, Kettle B, Hefty PS, Li J, Egan SM. 2013. Small-Molecule Inhibitor of the Shigella flexneri Master Virulence Regulator VirF. Infect Immun 81:4220–4231.

23. Emanuele AA, Adams NE, Chen YC, Maurelli AT, Garcia GA. 2014. Potential novel antibiotics from HTS targeting the virulence-regulating transcription factor, VirF, from Shigella flexneri. J Antibiot (Tokyo) 67:379–386.

24. Hurt JK, McQuade TJ, Anthony E, Larsen MJ, Garcia GA. 2010. High-Throughput Screening of the Virulence Regulator VirF: A Novel Antibacterial Target for Shigellosis. J Biomol Screen 15:379–387.

25. Daines DA, Granger-Schnarr M, Dimitrova M, Silver RP. 2002. Use of LexA-based system to identify protein-protein interactions in vivo. Methods Enzymol 358:153–161.

26. Daines DA, Silver RP. 2000. Evidence for multimerization of neu proteins involved in polysialic acid synthesis in Escherichia coli K1 using improved LexA-based vectors. J Bacteriol 182:5267–5270.

27. Dmitrova M, Younès-Cauet G, Oertel-Buchheit P, Porte D, Schnarr M, Granger-Schnarr M. 1998. A new LexA-based genetic system for monitoring and analyzing protein heterodimerization in Escherichia coli. Mol Gen Genet 257:205–212.

28. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7.

29. Yang J, Zhang Y. 2015. I-TASSER server: New development for protein structure and function predictions. Nucleic Acids Res 43:W174–W181.

30. ULC CCG. 2022. Molecular Operating Environment (MOE), 2020.09. Montreal, QC, Canada.

31. Crick FHC. 1953. The packing of α-helices: simple coiled-coils. Acta Crystallogr 6:689–697.

32. Walshaw J, Woolfson DN. 2003. Extended knobs-into-holes packing in classical and complex coiled-coil assemblies. J Struct Biol 144:349–361.

33. Gillette WK, Martin RG, Rosner JL. 2000. Probing the Escherichia coli transcriptional activator MarA using alanine-scanning mutagenesis: Residues important for DNA binding and activation. J Mol Biol 299:1245–1255.

34. Hung DT, Shakhnovich EA, Pierson E, Mekalanos JJ. 2005. Small-molecule inhibitor of Vibrio cholerae virulence and intestinal colonization. Science (80-) 310:670–674.

35. Shakhnovich EA, Hung DT, Pierson E, Lee K, Mekalanos JJ. 2007. Virstatin inhibits dimerization of the transcriptional activator ToxT. Proc Natl Acad Sci U S A 104:2372–2377.

36. Woodbrey AK, Onyango EO, Pellegrini M, Kovacikova G, Taylor RK, Gribble GW, Jon Kull F. 2017. A new class of inhibitors of the AraC family virulence regulator Vibrio cholerae ToxT. Nat Sci Reports 7:1–11.

37. Woodbrey AK, Onyango EO, Kovacikova G, Kull FJ, Gribble GW. 2018. A Modified ToxT Inhibitor Reduces Vibrio cholerae Virulence in Vivo. Biochemistry 57:5609–5615.

38. Bateman A, Martin MJ, Orchard S, Magrane M, Agivetova R, Ahmad S, Alpi E, Bowler-Barnett EH, Britto R, Bursteinas B, Bye-A-Jee H, Coetzee R, Cukura A, Silva A Da, Denny P, Dogan T, Ebenezer TG, Fan J, Castro LG, Garmiri P, Georghiou G, Gonzales L, Hatton-Ellis E, Hussein A, Ignatchenko A, Insana G, Ishtiaq R, Jokinen P, Joshi V, Jyothi D, Lock A, Lopez R, Luciani A, Luo J, Lussi Y, MacDougall A, Madeira F, Mahmoudy M, Menchi M, Mishra A, Moulang K, Nightingale A, Oliveira CS, Pundir S, Qi G, Raj S, Rice D, Lopez

MR, Saidi R, Sampson J, Sawford T, Speretta E, Turner E, Tyagi N, Vasudev P, Volynkin V, Warner K, Watkins X, Zaru R, Zellner H, Bridge A, Poux S, Redaschi N, Aimo L, Argoud-Puy G, Auchincloss A, Axelsen K, Bansal P, Baratin D, Blatter MC, Bolleman J, Boutet E, Breuza L, Casals-Casas C, de Castro E, Echioukh KC, Coudert E, Cuche B, Doche M, Dornevil D, Estreicher A, Famiglietti ML, Feuermann M, Gasteiger E, Gehant S, Gerritsen V, Gos A, Gruaz-Gumowski N, Hinz U, Hulo C, Hyka-Nouspikel N, Jungo F, Keller G, Kerhornou A, Lara V, Le Mercier P, Lieberherr D, Lombardot T, Martin X, Masson P, Morgat A, Neto TB, Paesano S, Pedruzzi I, Pilbout S, Pourcel L, Pozzato M, Pruess M, Rivoire C, Sigrist C, Sonesson K, Stutz A, Sundaram S, Tognolli M, Verbregue L, Wu CH, Arighi CN, Arminski L, Chen C, Chen Y, Garavelli JS, Huang H, Laiho K, McGarvey P, Natale DA, Ross K, Vinayaka CR, Wang Q, Wang Y, Yeh LS, Zhang J. 2021. UniProt: The universal protein knowledgebase in 2021. Nucleic Acids Res 49:D480–D489.

39. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596:583–589.

40. Miller JH. 1972. Assay of β-Galactosidase, p. 352–355. *In* Experiments in Molecular Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor.

Note:

The majority of this this chapter will be published in "Elucidation of the DNA-binding activity of VirF from *Shigella flexneri* for the *icsA* and *rnaG* promoters and characterization of the N-terminal domain to identify residues crucial for dimerization." Garrett T. Dow, Anna M. Young, George A. Garcia. **2022**, Manuscript in Preparation.

Chapter 4 Screening of *Shigella* Virulence Inhibitors against VirF

4.1 Abstract

Antibacterial resistance continues to rise due to incidences of negligence by doctors overprescribing these medicines, patients not properly completing their antibiotic regimen, and significant overuse of these drugs in livestock and for other agricultural purposes. Beyond expanded efforts to develop novel antibiotics or improve existing ones, there are a few routes to combat this crisis. Novel therapeutic strategies include enhanced public health measures, vaccine development, and targeting virulence. Through targeting bacterial virulence, scientists expect there to be a reduced or lack of selective pressure for bacteria to develop resistance to the prescribed medication. Compound library screening is a promising approach to identify small molecules that inhibit bacterial virulence as an alternative to traditional antimicrobial drug design. Following library screening, mechanism of inhibition studies can be a valuable tool to further characterize and develop these virulence inhibitors.

4.2 Introduction

Virulence-targeted therapy is a promising approach to treating bacterial infection. It is currently thought that by targeting virulence pathways, which are often non-essential for cell viability, there will be no selective pressure on the bacteria to develop resistance(1). With the ongoing “unstoppable” rise in antimicrobial resistant strains, developing drugs that may contribute little or no selective pressure on bacterial development of resistance is very enticing. To date, there are only five virulence-targeted drugs approved by the Food and Drug Administration (FDA),

all of which are antibodies that target bacterial toxins(2–8). However, no small molecules are currently FDA approved anti-virulence medications but a few small molecule agents are presently in pre-clinical trials(9, 10).

Despite the lack of approved small molecule anti-virulence drugs, there has been extensive, and promising, research done on identifying virulence inhibitors of *Shigella flexneri.* Two groups have studied the effects of Type 3 secretion system (T3SS)-directed inhibitors via targeting the Spa47 ATPase (11) and IpaD at the tip of the T3SS(12). Case and coworkers studied the effects of three previously identified *Yersinia pestis* T3SS ATPase inhibitors (13) on *S. flexneri* Spa47 ATPase via docking, kinetics analyses with purified oligomeric Spa47, *Shigella* growth rates, and T3SS of effector IpaC(11). These studies showed that these T3SS inhibitors have cross-species reactivity and even exhibited IC_{50}'s around 50 μM for Spa47 *in vitro*(11). Dey et al. identified four small molecule scaffolds that bind to T3SS tip protein, IpaD, from a surface plasmon resonance (SPR) screen(12). Saturation transfer difference (STD), 2D ^{1}H-^{15}N TROSY, and ^{1}H-^{13}C HSQC NMR experiments were utilized to further characterize these binding interactions(12). Since the purpose of this screen was to identify scaffolds to build upon, no *in vivo* data was performed with these compounds.

The master transcriptional activator of virulence in *S. flexneri*, VirF, is considered an attractive target for virulence-directed therapeutics. Located on a large, horizontally acquired virulence plasmid, VirF activates transcription and subsequent expression of necessary downstream virulence factors such as VirB and IcsA. It has been shown that when VirF was inactivated via Tn*5* insertion, significant reductions in expression were seen for four major downstream virulence factors/antigens (IpaB, IpaC, IpaD, and IcsA) but were restored to near WT levels when VirF was re-introduced(14). Affecting VirF activity through altering the length of

one of cognate promoters (*virB* promoter)(15), or via incorporation of the entire virulence plasmid into the bacterial chromosome(16), VirF was unable to activate transcription of *virB* resulting in a non-invasive phenotype. Sansonetti *et al.* also presented data that *S. flexneri*, when expressing a mutant form of IcsA, did not lead to colonic tissue damage in test animals as it did not possess the ability for cell-to-cell spread(17). Ultimately, it was shown that when the virulence plasmid was removed, *S. flexneri* lost pathogenicity(18).

Perhaps more extensively studied, inhibition of VirF with small molecules has been performed by a couple of groups. Koppolu and coworkers identified a small molecule compound, SE-1 (**Figure 4.1**), which showed considerable activity against VirF and appeared to reduce downstream virulence gene transcription (including *virB*, *icsA*, and *ipaB*) and *Shigella* invasion of L-929 mouse fibroblast cells(19). When SE-1 analogues were tested, it was discovered that their core scaffold decomposed into a reactive quinolinium core and the researchers were unable to develop a more potent inhibitor than SE-1(20). In the Garcia Lab, two large screens of 42,000 and 100,000 compound libraries in a *Shigella*-based β-galactosidase reporter assay were performed to identify potential inhibitors of VirF(21, 22). Five lead compounds were identified from these screens: 19615, 144092, 144143, 153578, and 24904 (**Figure 4.1**) (22). These hits presented IC_{50} values less than 100 μM with 19615, 144092 and 153578 presenting reductions in cell-to-cell spread in a plaque assay while 144092 also showed a reduction in Caco2 cell invasion(22). Mechanism of inhibition studies confirmed that 19615 inhibited VirF DNA-binding activity *in vitro*(23). However, an initial SAR campaign based on commercially available analogues was performed but resulted in no compounds that exhibited higher potency than 19615(23). Overall, these studies exemplify validated approaches to identify VirF inhibitors while also confirming these compounds have anti-virulence activity *in vivo*.

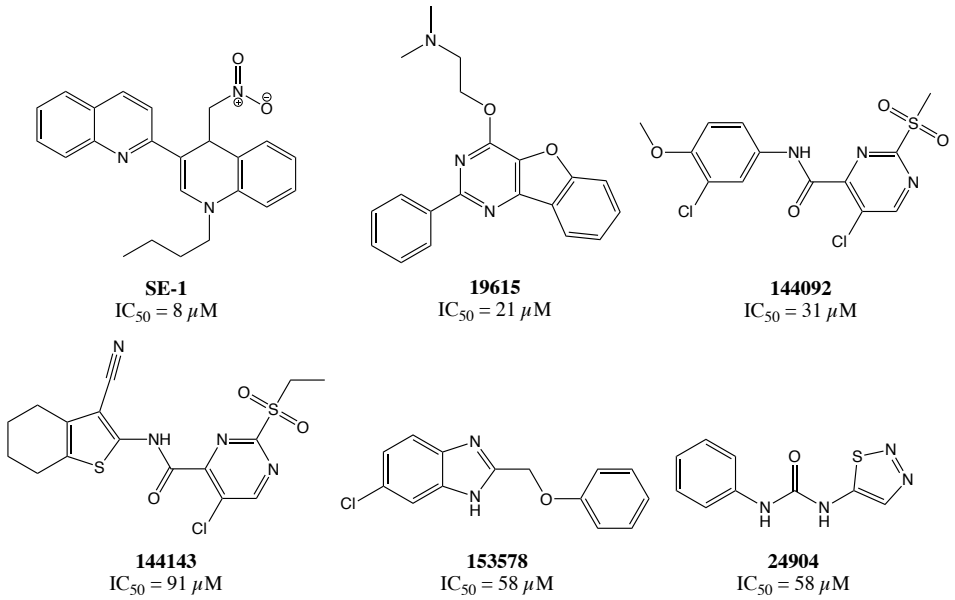


Figure 4.1: VirF inhibitors and corresponding IC_{50} values from β-galactosidase reporter assays identified in Koppolu et al., 2013 (SE-1) and Emanuele et al., 2014 (19615, 144092, 144143, 153578, 24904) (19,22).

Recently, in a collaboration with GlaxoSmithKline (GSK; funded by the Tres Cantos Open Lab Foundation, Tres Cantos, Spain), we performed a 1.7 million compound, high-throughput screen in a *Shigella* intra-macrophage survival assay(24). This large-scale phenotypic screen sought to identify high potency virulence inhibitors of *S. flexneri* with the goal of progression through pre-clinical and clinical trials as potential therapeutics to treat shigellosis. While the phenotypic screen generated a small set of hits with high potency (≤ 1 μM) against *S. flexneri* virulence, the molecular targets for these compounds are unknown. Herein, the primary objective was to further characterize these GSK hits. Employing our previously optimized VirF-driven β-galactosidase reporter assay, we sought to identify if any of these hits presented activity against our primary molecular target, VirF.

4.3 Results and Discussion

An initial 1.7 million compound screen was performed at GlaxoSmithKline in Tres Cantos, Spain by Dr. Marija Miljkovic to test compound inhibition of *Shigella* intra-macrophage survival

(24). The primary assay measured nano-luciferase activity of 2457T *S. flexneri* (harboring a nano-luciferase expression plasmid) following an 18-hour incubation period. *S. flexneri* are known to induce macrophage pyroptosis and escape into the extracellular space (25), so compounds were selected that inhibited *S. flexneri* survival in THP-1 macrophages. Proceeding confirmational screening and compound filtering, 3976 hits were advanced to dose response (DR) testing (**Figure 4.2 A**). DR assays for intra-macrophage survival and for antibacterial activity, chemical clustering, and computation analyses resulted in 84 hits (80 with IC_{50}'s <1.1 μM) with an additional 9 hits selected from an initial 58K primary screen (screened prior to the 1.7M compound library) (**Figure 4.2 B**). Three primary chemotypes were observed from these hits including arylnitriles, diarylsulfonamides, and phenyl-oxazolo-tetrazoles. Following DR and computation analyses, secondary screening was performed to remove compounds cytotoxic to THP-1 macrophages, Caco2 epithelial cells, and HepG2 liver cells, 44 hit compounds were identified as very promising virulence inhibitors (**Figure 4.2 C**). These compounds exhibited little to no antibacterial activity against *S. flexneri*, no cytotoxicity to Caco2 or HepG2 cells, and little inhibition of nano-luciferase (IC_{50} nano-luciferase/intramacrophage survival ratio > 4) (24). These compounds are incredibly promising for further development as *S. flexneri* virulence inhibitors and in mechanism of inhibition studies.

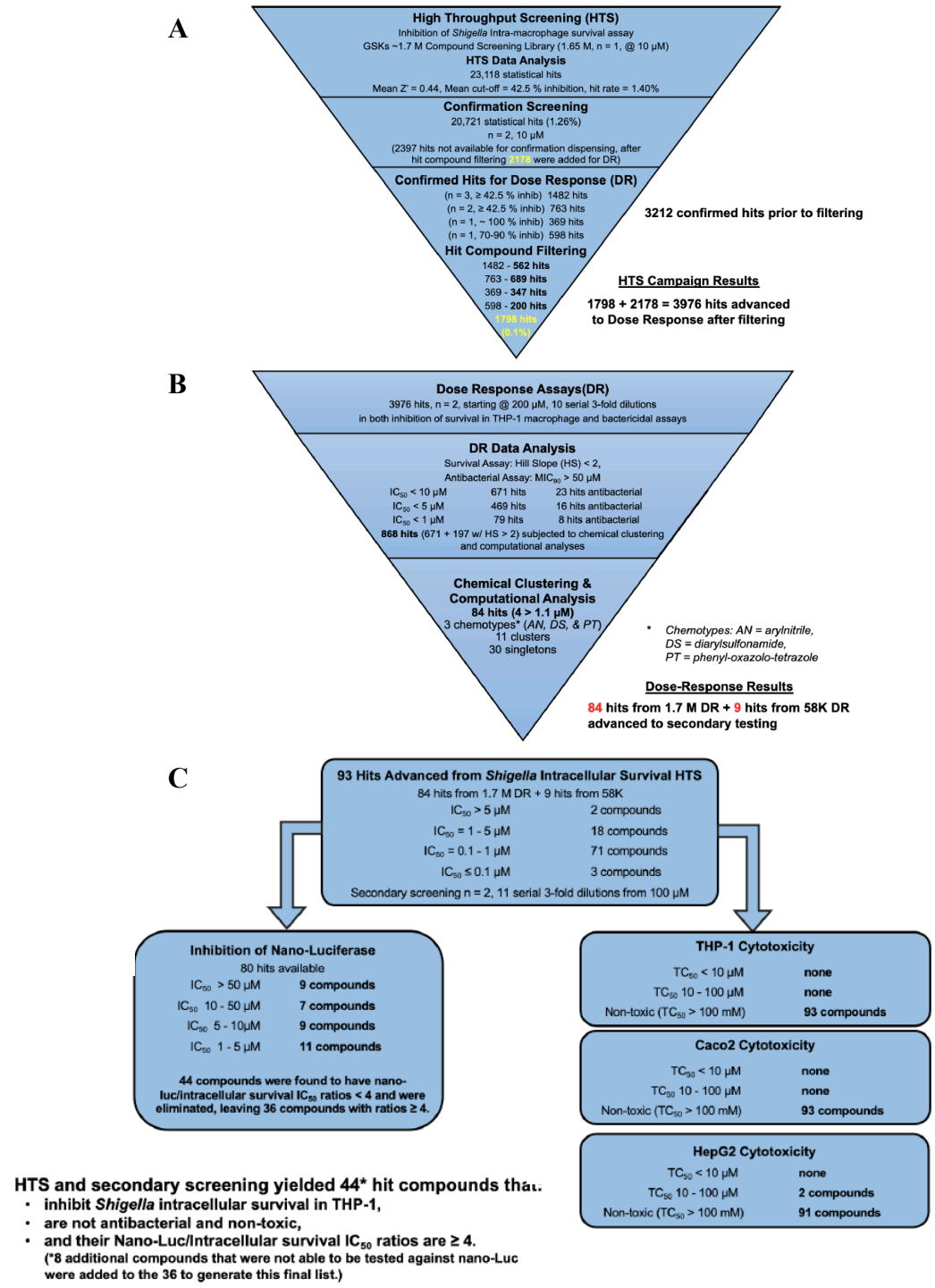


Figure 4.2: Screening cascade of the 1.7M compound library screen at GSK in Tres Cantos, Spain. (**A**): The primary and confirmation screening cascade identified 3976 potential hits. (**B**): Dose response (DR), chemical clustering, and computational analysis cascade indicates the selection of 93 hits that were moved on through the screening campaign to be tested in secondary screening. (**C**): Secondary screening cascade of the final 93 hit compounds was performed to identify compounds exhibiting nano-luciferase inhibition and cytotoxicity in THP-1 macrophages, Caco2 epithelial cells, or HepG2 liver cells. All figures were adapted from Miljkovic et al., 2021 (24).

A set of 33 compounds was received from GSK for follow-up testing (**Table 4.1**). Of these compounds, only 11 were from the 44 hit compounds (bolded in **Table 4.1**) identified in Miljkovic et al., 2021(24). Unfortunately, we were unable to receive the entire set of 44 due to a lack of dry compound (not dissolved in DMSO) availability for the remaining 33. In a collaboration with the Vahlteich Medicinal Chemistry Core (VMCC) at the University of Michigan, we requested the synthesis of three other compounds to test against VirF: GSK1109683A, GSK1131048A, and GSK1668619A (Labeled with an * in **Table 4.1**). These compounds were selected to be synthesized due to their significant characteristics (potency, chemical clustering, computational analyses, etc.) identified from the primary and secondary screens with two of which being a part of the set of 44 (Bolded and * in **Table 4.1**; GSK1109683A, GSK1131048A).

To determine if any of these hit compounds inhibited VirF activity in cells, we employed our previously optimized VirF-driven β-galactosidase reporter assay(21, 22). Initially, all compounds were tested at 10 μM. Only 10 compounds exhibited a positive value for percent inhibition. GR80356X resulted in a 98.4% percent inhibition of VirF activated transcription of β-galactosidase but this was seen as a false positive due to it being bactericidal. These results were also produced in the primary and secondary screening at GSK(24). Four other compounds exhibited > 10% inhibition in the assay: GSK1510765A, GSK510287A, GSK567121A, and GW395044X. Given that most of these compounds exhibited < 1 μM IC_{50} values in the intra-macrophage survival assays, the weak inhibition of VirF at 10 μM is inconsistent with VirF being the virulence target of these hit compounds.

Table 4.1: GSK hit compounds tested in the VirF-driven β-galactosidase reporter assay. Compounds are labeled with the GSK names and include the chemotype class in which they belong (singletons lie outside of the primary three classes). Percent inhibition (%Inhibition) at 10 μM (all compounds) and 50 μM (selected compounds) is placed to the right of the table. Compounds that are bolded are part of the 44 hits identified from the screen. Those labeled with an * were synthesized by our collaborators at the VMCC. Note: No structures are presented due to this work being unpublished.

GSK Hits	Chemotype	%Inhibition (10 μM)	%Inhibition (50 μM)
GR80356X	Singleton	98.4	-
GSK1102565A	Arylnitrile	-36.4	-
GSK1109683A*	Arylnitrile	-2.5	-
GSK1131048A*	Arylnitrile	7.2	-
GSK1235082A	Diarylsulfonamide	-6.5	-
GSK1243982A	Diarylsulfonamide	-19.5	-
GSK1344002A	Diarylsulfonamide	-13.7	-
GSK1510765A	Diarylsulfonamide	17.1	-11.3
GSK1510767A	Diarylsulfonamide	-20.5	-
GSK1516277A	Singleton	5.1	-6.3
GSK1648896A	Arylnitrile	-36.9	-31.2
GSK1668619A*	Diarylsulfonamide	4.8	-
GSK2276018A	NSR	8.3	-66.2
GSK2668123A	Singleton	-109.2	-
GSK347415A	Phenyl-oxazolo-tetrazole	-24.1	-
GSK510287A	Arylnitrile	10.5	-24
GSK510292A	Arylnitrile	-8.7	-24.5
GSK510418A	Arylnitrile	3.7	-30.7
GSK566838A	Singleton	-4	-
GSK567121A	Diarylsulfonamide	10.3	-18.8
GSK639094A	Diarylsulfonamide	-2.2	-
GSK732092A	Arylnitrile	-0.4	-
GSK733298A	Diarylsulfonamide	2.9	-
GSK737408A	Diarylsulfonamide	-14.3	-
GSK753049A	NSR	-23.7	-
GSK784540A*	Diarylsulfonamide	-19.3	-
GSK826577A	Singleton	0.9	-
GSK864774A	Singleton	8.5	-23.2
GSK877561A	Diarylsulfonamide	-8.2	-
GSK898826A	Diarylsulfonamide	-22	-
GSK945897A	Singleton	-8.6	-
GW395044X	Singleton	17.2	-29.2
SB-807542	NSR	-2.2	-
SB-807561	Diarylsulfonamide	-21.6	-
SB-807606	NSR	-5	-
SB-807639	NSR	-0.6	-

Due to concerns that these hits were insignificant in the β-galactosidase assay, compounds that exhibited positive values for percent inhibition at 10 μM were tested again at 50 μM (i.e.,

GSK1510765A, GSK1516277A, GSK2276018A, GSK510287A, GSK510418A, GSK567121, GSK864774A, and GW395044X). Additionally, we wanted to determine if there was a dose-response effect seen with compounds that exhibited negative percent inhibitions at 10 μM so GSK1648896A and GSK510292A were selected. As expected, none of the selected compounds exhibited a dose-response effect and were identified as not being VirF inhibitors.

The *S. flexneri* virulence cascade includes a complex network of transcription factors, effector proteins, structural proteins, and chaperones that regulate how the bacteria invade and spread within the human colon as well as moderate the host innate immune response. Although VirF is the primary transcription factor that regulates essential downstream proteins and transcription factors (e.g., VirF and IcsA) necessary for *S. flexneri* virulence, there are many more targets that could be inhibited with small molecules to elicit a lack of virulence. Among the many potential virulence factors in *S. flexneri*, four virulence factors, directly or indirectly activated by VirF, could be potential targets inhibited by the hit compounds identified at GSK: VirB, IpaB, IpaD, or IcsA. VirB is a secondary transcription factor which activates transcription of *ipaB*, *-C*, and *ipaD* leading to the expression of the T3SS resulting in its invasive phenotype(26). IpaB has also been shown to induce macrophage apoptosis so that the bacteria can escape from macrophages and proceed to invade the basolateral membrane of adjacent epithelial cells(27, 28). IpaD is another important virulence factor that is located at the tip of the T3SS apparatus and functions by controlling the recruitment and secretion of IpaB and IpaC and has even been implicated in caspase-1 independent apoptosis of macrophages(29, 30). Additionally, VirF also activates transcription of *icsA*, initiating cell-to-cell spread by facilitating the polymerization of host cell actin at one pole of the bacterium(14, 31). Importantly, the results of genetic knock-out or mutant studies of virulence pathway genes (e.g., *virF*, *virB*, *icsA, ipaB, ipaD)* in *S. flexneri* have provided

convincing evidence that inactivation of these pathways reduces or eliminates pathogenicity(16, 17, 29, 32, 33). Hence, through targeting any of these four additional virulence factors with small molecules, it is expected that *S. flexneri* will lose its virulence phenotype and pathogenicity, as seen in the *Shigella*-based intra-macrophage survival assay. The inability to escape from macrophages and to invade epithelial cells, allows for macrophages to properly phagocytose and eradicate the pathogen. Mechanism of inhibition studies testing inhibition of these four factors with the GSK-identified hits is essential for further development of these virulence inhibitors but is outside the scope of this dissertation work.

4.4 Conclusions

The selected GSK hits provided promising results as potential *S. flexneri* virulence inhibitors from the initial primary and secondary screens at GSK, but none significantly affected VirF activity in our VirF-driven β-galactosidase reporter assay. Due to the complexity of molecular targets that contribute to *S. flexneri* virulence, it is highly likely these compounds are targeting other necessary virulence factors with high potency rather than VirF.

4.5 Methods

4.5.1 Materials

All standard buffer components were purchased from Millipore Sigma or Thermo Fisher. Equipment utilized for these experiments was purchased from varying companies which are indicated in parentheses throughout this section.

4.5.2 VirF-Driven β-Galactosidase Reporter Assay

Prior to assay set-up, compounds were received from GSK and approximately 1 mg of each were diluted in appropriate volumes of dimethyl sulfoxide (DMSO) to prepare 20 mM stocks. The assay was performed as previously described(21, 22). First, to set up the assay, avirulent, BS103 *S. flexneri* cells harboring either pMAL-virF-lacZ (negative control or test) or pMAL-ΔvirF-lacZ (positive control) were grown overnight in 2xTY media (16 g bactotryptone, 10 g yeast extract, 5 g NaCl per liter MilliQ water) supplemented with 100 µg/mL carbenicillin. Following overnight incubation, the cultures were diluted to OD_{600}=0.012 using 2xTY media supplemented with carbenicillin. In a sterile tube, 2xTY media was supplemented with 100 µg/mL carbenicillin and 15 µM (or 75 µM) of the desired GSK compound. In a sterile 96-well UV-transparent plate, 50 µL of culture containing BS103 cells harboring either plasmid was pipetted into the wells. Then, 25 µL of culture containing the desired test compound was pipetted in the corresponding wells. Every compound was tested in triplicate against BS103 *S. flexneri* cells harboring the pMAL-virF-lacZ plasmid. Compound concentrations were tested either at 10 µM (15 µM stock in media prior to plate addition) or 50 µM (75 µM stock in media). Controls included a blank (no cells for background corrections due to media components), negative control (BS103 *S. flexneri* harboring pMAL-virF-lacZ plasmid without compound addition), and a positive control (BS103 *S. flexneri* harboring pMAL-ΔvirF-lacZ) which were all tested in triplicate. The plates were briefly spun down at 1000xg for 1 min using a tabletop centrifuge (Beckman Coulter Allegra Series) prior to incubation for 20-24 hours at 30 ˚C in a humidified incubator (5% CO_2). Following the incubation period, 75 ml of 1x Z buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, 0.1% Triton X-100, pH 7.0) containing 0.5 mg/mL chlorophenyl red-β-D-galactopyranoside (CPRG; Millipore Sigma) to the corresponding wells. The plates were incubated at room

temperature for 10 min then the OD_{570} was measured. Each control or compound test reaction was averaged prior to calculation of percent inhibition. To calculate percent inhibition for each compound test, **Equation 4.1** was used.

Equation 4.1 ((Positive Control – Test) / (Positive Control)) * 100

4.6 References

1. Clatworthy AE, Pierson E, Hung DT. 2007. Targeting virulence: A new paradigm for antimicrobial therapy. Nat Chem Biol 3:541–548.

2. Jewell NP, Ph D, Hatheway CL. 2006. Human Botulism Ummine Globin for the Treatment of Infant Botulism. NEW Engl YOURNAL Med 354:462–471.

3. Yamamoto BJ, Shadiack AM, Carpenter S, Sanford D, Henning LN, Gonzales N, O'Connor E, Casey LS, Serbina N V. 2016. Obiltoxaximab prevents disseminated Bacillus anthracis infection and improves survival during pre- and postexposure prophylaxis in animal models of inhalational anthrax. Antimicrob Agents Chemother 60:5796–5805.

4. Migone T-S, Subramanian GM, Zhong J, Healey LM, Corey A, Devalaraja M, Lo L, Ullrich S, Zimmerman J, Chen A, Lewis M, Meister G, Gillum K, Sanford D, Mott J, Bolmer SD. 2009. Raxibacumab for the Treatment of Inhalational Anthrax. N Engl J Med 361:135–144.

5. Greig SL. 2016. Obiltoxaximab: First Global Approval. Drugs 76:823–830.

6. Rounds J, Strain J. 2017. Bezlotoxumab for Preventing Recurrent Clostridium difficile Infections. S D Med 70:422–423.

7. Babcock GJ, Broering TJ, Hernandez HJ, Mandell RB, Donahue K, Boatright N, Stack AM, Lowy I, Graziano R, Molrine D, Ambrosino DM, Thomas WD. 2006. Human monoclonal antibodies directed against toxins A and B prevent Clostridium difficile-induced mortality in hamsters. Infect Immun 74:6339–6347.

8. Lowy I, Molrine DC, Leav BA, Blair BM, Baxter R, Gerding DN, Nichol G, Thomas Jr. WD, Leney M, Sloan S, Hay CA, Ambrosino DM. 2010. Treatment with Monoclonal Antibodies against Clostridium difficile Toxins. N Engl J Med 362:197–205.

9. Dickey SW, Cheung GYC, Otto M. 2017. Different drugs for bad bugs: Antivirulence strategies in the age of antibiotic resistance. Nat Rev Drug Discov 16:457–471.

10. Ogawara H. 2021. Possible drugs for the treatment of bacterial infections in the future: anti-virulence drugs. J Antibiot (Tokyo) 74:24–41.

11. Case HB, Mattock DS, Dickenson NE. 2018. Shutting Down Shigella Secretion: Characterizing Small Molecule Type Three Secretion System ATPase Inhibitors. Rev del Col Am Cardiol 57:6906–6916.

12. Dey S, Anbanandam A, Mumford BE, De Guzman RN. 2017. Characterization of Small-Molecule Scaffolds That Bind to the Shigella Type III Secretion System Protein IpaD. ChemMedChem 12:1534–1541.

13. Swietnicki W, Carmany D, Retford M, Guelta M, Dorsey R, Bozue J, Lee MS, Olson MA. 2011. Identification of small-molecule inhibitors of Yersinia pestis type III secretion system YscN atpase. PLoS One 6.

14. Sakai T, Sasakawa C, Yoshikawa M. 1988. Expression of four virulence antigens of Shigella flexneri is positively regulated at the transcriptional level by the 30 kilo Dalton virF protein. Mol Microbiol 2:589–597.

15. Tobe T, Yoshikawa M, Mizuno T, Sasakawa C. 1993. Transcriptional Control of the Invasion Regulatory Gene virB of Shigella flexneri: Activation by VirF and Repression by H-NS. J Bacteriol 175:6142–6149.

16. Colonna B, Casalino M, Fradiani PA, Zagaglia C, Naitza S, Leoni L, Prosseda G, Coppo A, Ghelardini P, Nicoletti M. 1995. H-NS regulation of virulence gene expression in enteroinvasive Escherichia coli harboring the virulence plasmid integrated into the host chromosome. J Bacteriol 177:4703–4712.

17. Sansonetti PJ, Arondel J, Fontaine A, d'Hauteville H, Bernardini ML. 1991. OmpB (osmo-regulation) and icsA (cell-to-cell spread) mutants of Shigella flexneri: vaccine candidates and probes to study the pathogenesis of shigellosis. Vaccine 9:416–422.

18. Maurelli AT, Blackmon B, Curtiss Iii R. 1984. Loss of Pigmentation in Shigella flexneri 2a Is Correlated with Loss of Virulence and Virulence-Associated PlasmidINFECTION AND IMMUNITY.

19. Koppolu V, Osaka I, Skredenske JM, Kettle B, Hefty PS, Li J, Egan SM. 2013. Small-Molecule Inhibitor of the Shigella flexneri Master Virulence Regulator VirF. Infect Immun 81:4220–4231.

20. Jain P, Li J, Porubsky P, Neuenswander B, Egan SM, Aubé J, Rogers S. 2014. 3-Substituted Biquinolinium Inhibitors of AraC Family Transcriptional Activator VirF from S. flexneri Obtained Through In Situ Chemical Ionization of 3,4-Disubstituted Dihydroquinolines. RSC Adv 4:39809–39816.

21. Hurt JK, McQuade TJ, Anthony E, Larsen MJ, Garcia GA. 2010. High-Throughput Screening of the Virulence Regulator VirF: A Novel Antibacterial Target for Shigellosis. J Biomol Screen 15:379–387.

22. Emanuele AA, Adams NE, Chen YC, Maurelli AT, Garcia GA. 2014. Potential novel antibiotics from HTS targeting the virulence-regulating transcription factor, VirF, from Shigella flexneri. J Antibiot (Tokyo) 67:379–386.

23. Emanuele AA, Garcia GA. 2015. Mechanism of Action and Initial, In Vitro SAR of an Inhibitor of the Shigella flexneri Virulence Regulator VirF. PLoS One 10:1–18.

24. Miljkovic M, Lozano S, Castellote I, de Cózar C, Villegas-Moreno AI, Gamallo P, Martinez DJ-A, Fernández-Álvaro E, Ballell L, Garcia GA. 2021. Novel inhibitors that target bacterial virulence identified via HTS against intramacrophage survival of Shigella flexneri. PLoS Pathog Submitted.

25. Jennison A V., Verma NK. 2004. Shigella flexneri infection: pathogenesis and vaccine development. FEMS Microbiol Rev 28:43–58.

26. Adler B, Sasakawa C, Tobe T, Makino S, Komatsu K, Yoshikawa M. 1989. A dual transcriptional activation system for the 230 kb plasmid genes coding for virulence-associated antigens of Shigella flexneri. Mol Microbiol 3:627–635.

27. High N, Mounier J, Prévost MC, Sansonetti PJ. 1992. IpaB of Shigella flexneri causes entry into epithelial cells and escape from the phagocytic vacuole. EMBO J 11:1991–9.

28. Zychlinsky A, Kenny B, Ménard R, Prévost M -C, Holland IB, Sansonetti PJ. 1994. IpaB mediates macrophage apoptosis induced by Shigella flexneri. Mol Microbiol 11:619–627.

29. Blocker A, Gounon P, Larquet E, Niebuhr K, Cabiaux V, Parsot C, Sansonetti P. 1999. The tripartite type III secreton of Shigella flexneri inserts IpaB and IpaC into host membranes. J Cell Biol 147:683–693.

30. Veenendaal AKJ, Hodgkinson JL, Schwarzer L, Stabat D, Zenk SF, Blocker AJ. 2007. The type III secretion system needle tip complex mediates host cell sensing and translocon insertion. Mol Microbiol 63:1719–1730.

31. Bernardini ML, Mounier J, D'Hauteville H, Coquis-Rondon M, Sansonetti PJ. 1989. Identification of icsA, a plasmid locus of Shigella flexneri that governs bacterial intra- and intercellular spread through interaction with F-actin. Proc Natl Acad Sci U S A 86:3867–71.

32. Picking WL, Nishioka H, Hearn PD, Baxter MA, Harrington AT, Blocker A, Picking WD. 2005. IpaD of Shigella flexneri is independently required for regulation of Ipa protein secretion and efficient insertion of IpaB and IpaC into host membranes. Infect Immun 73:1432–1440.

33. Makino S, Sasakawa C, Kamata K, Kurata T, Yoshikawa M. 1986. A genetic determinant required for continuous reinfection of adjacent cells on large plasmid in S. flexneri 2a. Cell 46:551–555.

Chapter 5 Conclusions and Future Directions

5.1 Conclusions

5.1.1 VirF DNA-Binding Activity

Using two VirF homologs from *E. coli*, MarA and GadX, we developed homology models for the *Shigella flexneri* VirF DNA-binding domain (DBD). Our two homology models provided insight into how the DNA-binding domain might undergo conformational changes from unbound to DNA-bound states. Upon aligning the VirF DBD with MarA and inspection of the MarA•*marRAB* crystal structure, we selected and mutated seven amino acids to alanine to be tested in *in vitro* DNA-binding assays. When binding was tested against *pvirB*, mutations to the N-terminal helix-turn-helix (HTH) motif (R192 and K193) significantly reduced or knocked out DNA-binding activity. However, mutations to the C-terminal HTH motif varied in their effects on DNA-binding. These trends were also observed when VirF DNA-binding activity was expanded to the *icsA* and *rnaG* promoters except for K193A. The K193A mutant exhibited binding to both *picsA* and *prnaG* at a higher migrating initial binding shift relative to WT, suggesting this mutant is capable of binding to these DNA promoters only as a higher oligomeric complex. Alanine-mutations to S238, Y239, and I241 exhibited differential effects on VirF DNA-binding across the three promoters. Overall, the *picsA* binding interaction was more sensitive than the other two promoter interactions where mutations to all residues except I189 caused moderate to significant losses in DNA-binding activity. A diagram of the contributions of these residues for VirF DNA-binding activity is included in **Figure 5.1**.

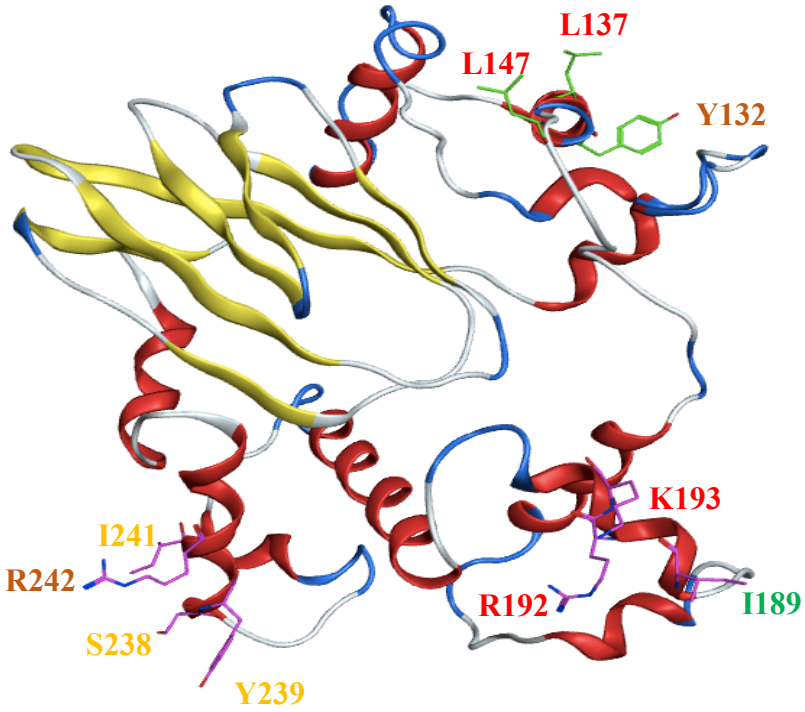


Figure 5.1: Summary of critical DNA-binding and dimerization contributing residues within VirF. The residue color indicates its relative contribution to DNA-binding or dimerization (Red=Significant, Orange=Moderate, Yellow=Slight, Green=Insignificant). The diagram uses VirF Model 1 from **Figure 3.14,** and the structure was modeled using Molecular Operating Environment (MOE).

Importantly, multiple binding events were observed when WT and MalE-VirF DNA-binding mutants were tested with the *icsA* and *rnaG* promoters, indicating these promoters may require higher oligomeric states (i.e., VirF actively dimerizes) to activate transcription of downstream virulence genes. In addition, the lower affinities exhibited for these promoters compared to *pvirB* may indicate how *S. flexneri* temporally activates these three promoters. Due to the lower affinities and multiple binding events observed for the *icsA* and *rnaG* promoters, it is likely that these promoters are activated in later infection stages when the bacteria are replicating and produce enough VirF to activate transcription of *icsA* thus promoting the energy-costly process of cell-to-cell spread. On the other hand, the *virB* promoter requires less VirF based on our calculated affinity (2.3 μM) which indicates the promoter is more active in the earlier stages of *Shigella* infection when the bacteria are not replicating as often and requires IpaB and the T3SS to

induce macrophage pyroptosis, escape, and invade epithelial cells. Overall, the results presented in **Chapter 2** provide a more complete functional postulate for how VirF interacts with its three cognate DNA promoters *in vitro* and may be useful in the future discovery and development of novel VirF DNA-binding inhibitors.

5.1.2 VirF Dimerization Activity

To our knowledge, this is the first report that provides experimental evidence that VirF dimerizes *in vivo*. Using alignments with ToxT and AraC, 18 non-alanine residues within our predicted dimerization α-helix were mutated to alanine and tested in a LexA reporter assay and compared to WT VirF. Three alanine mutations at Y132, L137, and L147 exhibited significant decreases in dimerization activity relative to WT VirF (**Figure 5.1**). In addition, VirF N-terminal truncation mutants were found to be capable of dimerizing, providing evidence that VirF dimerization is more complex than previously thought and that the two domains may be able to function independently. Based on western blots of these mutants from the assay, L137A and Y132A/L137A/L147A (triple mutant) LexA-VirF proteins displayed weaker expression levels compared to WT LexA-VirF indicating these mutations have a negative effect on protein expression and/or stability.

Dimerization mutant protein expression was performed under both homologous (BS103 *S. flexneri*) and heterologous (TOP10 *E. coli*) expression conditions but with little success. Despite the significant expression observed for all dimerization mutants in a western blot, we were unable to isolate pure protein to be used in *in vitro* DNA-binding assays. It is likely that mutations to this domain, specifically L137A, impart negative effects to MalE-VirF stability through affecting folding or inducing protein aggregation. Fortunately, one dimerization mutant, Y132A MalE-VirF, was successfully purified, albeit to a substantially low yield, and used in *in vitro* DNA-

binding assays. Y132A bound similarly to WT MalE-VirF with *pvirB*, exhibited no binding for *picsA*, but displayed different binding behavior than WT for *prnaG*. The reduced dimerization activity of Y132A MalE-VirF allowed the protein to bind as a monomer to *prnaG* whereas WT likely binds as a dimer at the same concentration. These results strongly suggest that, physiologically, VirF binds to its promoters as a dimer to activate transcription and that the Y132A mutation shifts equilibrium of MalE-VirF to a monomeric state but does not completely restrict it from dimerizing. This was observed when Y132A MalE-VirF bound *pvirB* similarly to WT.

I-TASSER-generated full-length models of VirF identified the locations of these three residues within our predicted dimerization α-helix and implicated their roles in dimerization. L137 and L147 likely contribute to VirF stability *in vivo* while Y132 is involved in the dimerization interface. Lastly, the orientation of the α-helix and L137 provides suggestive evidence that VirF may undergo conformational changes in the N-terminal domain upon dimerization. Overall, this is the first report that confirmed VirF dimerizes *in vivo*.

5.1.3 Screening GSK Hits Against VirF

A high-throughput screening campaign utilizing an *S. flexneri* intra-macrophage survival primary assay identified 93 hit compounds with IC_{50}'s < 1.1 μM. Of these 93 hits, we tested a set of 33 compounds obtained from GSK in our VirF-driven β-galactosidase reporter assay. At concentrations of 10 μM, there was no significant reductions in VirF-activated β-galactosidase activity. Hits with >10% inhibition at 10 μM were tested again at 50 μM but no dose-response was observed. Overall, none of the hit compounds exhibited activity against VirF that would justify any further studies. As the primary screen was phenotypic, the hits identified by GSK are likely interacting with *Shigella* molecular targets outside of VirF.

5.2 Future Directions

5.2.1 Expand VirF Nucleic Acid Binding Activity Studies to RNA

VirF promotes transcription of virulence genes via three different mechanisms. First, it interacts with DNA promoters to activate transcription of downstream virulence genes. This has been observed for both the *virB* and *icsA* promoters(1, 2). Secondly, VirF acts as a repressor by binding to the promoter region for *rnaG*(2). Finally, VirF has also been shown to act post-transcriptionally via binding to the RnaG RNA making it capable of preventing a "kissing complex" between RnaG and the *icsA* mRNA transcripts, thus promoting complete transcription (and subsequent translation) of the *icsA* mRNA (3). Without VirF present, a terminator hairpin within the *icsA* mRNA is formed, resulting in premature termination of the transcript(3).

The contents of this dissertation focused on the DNA-binding activity of MalE-VirF for its cognate DNA promoters but did not elaborate on VirF's RNA-binding activity. Using the experimental tools and mutants we employed in **Chapter 2**, the interaction of MalE-VirF with a subset of RNAs (e.g., *icsA* mRNA, RnaG, *virB* mRNA, etc.) could be elucidated. Since few VirF DBD mutants showed insignificant (I189A) or minimal effects (S238A, Y239A, I241) on DNA-binding, it is possible that these residues may contribute to VirF RNA-binding activity. In addition to the mutants studied herein, those proposed in Giangrossi et al.(3) (H212, S236, N245, T251, K253, and K254) can also be studied. Studying MalE-VirF RNA-binding activity will give a more complete understanding of how this protein interacts with nucleic acids and may elucidate how it acts both transcriptionally and post-transcriptionally to regulate virulence gene expression.

5.2.2 Analyze the Roles of DNA-Binding and Dimerization Mutants on Transcription

This study was primarily focused on the effect of mutations on *in vitro* DNA-binding activity but can be expanded to understand the effects of these mutations *in vivo*. β-galactosidase reporter assays would be incredibly useful in identifying how these proteins affect transcriptional activation with the *virB* and *icsA* promoters. I attempted to mutate the β-galactosidase reporter plasmid designed by Drs. Julie Hurt and Anthony Emanuele (4, 5) to do this, but was unsuccessful due to its large size. Because of this, it may be prudent to obtain the cell strain and plasmid used in Porter and Dorman to study VirF mutations on transcription(6). Using this system, a more comprehensive mutagenesis study on the DNA-binding domain outside of our selected mutations can be performed to understand how the two HTH motifs are involved in promoter activation and whether one motif exhibits high affinity interactions or provides promoter-specificity as postulated. This system would also be useful in elucidating how the VirF dimerization domain contributes to transcriptional activation. Such studies have been performed on ToxT(7) and ExsA(8) and determined that mutations to the N-terminal dimerization domain can elicit significant reductions in transcriptional activation of β-galactosidase. Like ToxT and ExsA, we hypothesize that mutations to the dimerization domain will incur a reduction in β-galactosidase activation since it is likely VirF is required to be a dimer to activate transcription.

5.2.3 High-Throughput Screen to Identify Dimerization Inhibitors

While broad spectrum antibiotics are incredibly effective, the downsides of these agents include disruption of the host microbiome and spread of resistance from non-pathogenic to pathogenic species(9–12). Conversely, since narrow-spectrum antibiotics, or in this case anti-virulence therapeutics, are pathogen-specific, they are expected to reduce the spread of resistance

between organisms. Due to the lack of similarity between AraC-family proteins in the N-terminal domain compared to the DBD, developing inhibitors that target dimerization is expected to be highly pathogen-specific, hence narrow-spectrum. Screening for VirF dimerization inhibitors can be accomplished with the LexA monohybrid β-galactosidase reporter assays(13–16). Since mutations to the VirF dimerization domain resulted in decreases in β-galactosidase activity without significant change in protein expression compared to WT (i.e., Y132A), we hypothesize that this system can be optimized to perform a high-throughput screen for inhibitors. Chemical compound libraries from the University of Michigan Center for Chemical Genomics (CCG), ChemDiv, ChemBridge, or others could be purchased and used for this screening campaign. However, proceeding with the libraries outside of the CCG would be better given we have already screened the CCG library against VirF (4, 5) and none of the hits inhibited VirF dimerization (**Figure 3.13**). Overall, the proposed screen can produce more chemical matter to expand upon in our search to develop a virulence-targeted inhibitor of *Shigella flexneri*.

5.2.4 Target Identification and Mechanism of Action for the GSK Hits

There are numerous virulence targets that could be inhibited by the GSK hits outside of VirF. As stated in **Chapter 1**, the T3SS and proteins such as IpaB, IpaD, and VirB are all essential for *Shigella* virulence. In addition, the T3SS ATPase, Spa47, could also be a potential molecular target (**Figure 5.2**). Firstly, GSK hits should be screened against VirB. Our VirF-driven β-galactosidase reporter assay could be modified to contain *virB* and use the *icsB* promoter to drive β-galactosidase expression for use in a *S. flexneri* 2a *virB* knockout cell strain(17). Other reporter strains that could be employed are the VirB-driven β-galactosidase reporter assay (*mxiC-lacZ* fusion) from Porter and Dorman (6) or the *icsB* and *icsP* promoter-controlled β-galactosidase reporter assays designed by the Wing Lab at UNLV(18). If the GSK hits are identified as VirB

inhibitors, mechanism of inhibition studies can proceed by testing the hits against VirB DNA-binding activity in EMSA and FP assays as described in Gao et al.(17). However, we would employ fluorescent-labeling to our *icsB* and *icsP* promoters due to safety concerns regarding radio-labeled nucleotides.

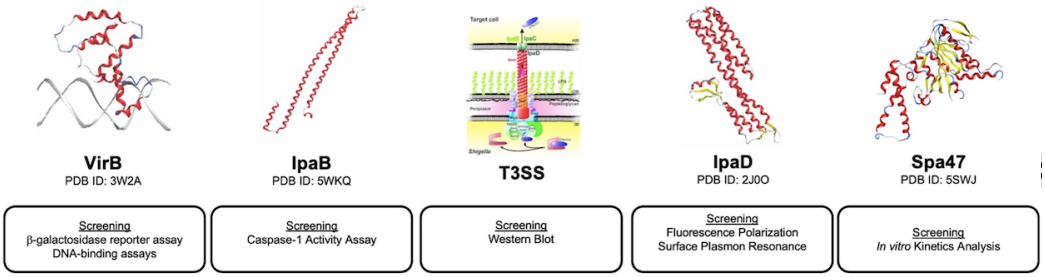


Figure 5.2: Diagram of proposed *S. flexneri* virulence factors that can be screened against with GSK hits. Protein structures were obtained from the Protein Data Bank and the T3SS figure was Schroeder and Hilbi, *Clinical Microbiology Reviews*, 2008. The figure was prepared in collaboration with Dr. George Garcia.

Another prominent virulence factor that could be targeted in the *Shigella* intra-macrophage survival assay at GSK is IpaB. Since it has been shown that IpaB is directly involved in macrophage pyroptosis, inhibition of this protein with any of the 93 GSK hits could have drastic effects on *Shigella* virulence within macrophages(19, 20). To determine if IpaB is a molecular target, a caspase-1 activity assay could be utilized and optimized based on Senerovic et al.(21). This assay studies IL-1β secretion (i.e., caspase-1 activity) in THP-1 macrophages infected with virulent 2457T *S. flexneri.* Compounds can be tested at 10 µM and compared to the negative (DMSO) and positive controls (caspse-1 irreversible inhibitor, Ac-YVAD-cmk; InvivoGen)(22). If the compounds show activity against cell-based caspse-1 activity, follow-up studies can be done against purified IpaB *in vitro*. IpaB has been shown to produce pores in host membranes *in vitro* resulting in macrophage pyroptosis(21, 23). Following purification of IpaB using its cognate chaperone IpgC as described in Birket et al.(23), IL-1β levels can be measured following priming of the THP-1 macrophage cells with LPS (*S. flexneri* 2a; Sigma) and co-incubation with purified IpaB.

The T3SS is comprised of ~25 proteins and its assembly and function is essential for *Shigella* virulence, suggesting it could also be a target in the primary screen performed at GSK(24, 25). Interestingly, the T3SS is highly conserved between enteropathogens so it is considered a very promising virulence target that is currently being explored for the development of anti-virulence therapeutics. In addition, the T3SS relies upon the ATPase, Spa47, to facilitate secretion of effector proteins which allow the bacteria to invade the colonic epithelium and subvert the host innate immune system(26). To test if the GSK hits inhibit the T3SS or Spa47, we can employ T3SS secretion and Spa47 *in vitro* kinetics assays as described in Case et al., 2018 and 2020(27, 28). Inhibition of T3SS translocator secretion will be monitored and quantified using western blotting with anti-IpaC antibodies following activation of the T3SS with Congo red and compared to untreated and cytoplasmic controls. For screening of Spa47 inhibitors, *in vitro* kinetics analyses of compounds on ATPase activity using a colorimetric malachite green assay kit (BioAssay Systems) can be performed(28).

Lastly, another target worth pursing is IpaD. The tip of the T3SS "nano-syringe" is constructed of a pentamer of IpaD that senses deoxycholate (DOC) in the environment, promoting the recruitment of IpaB and IpaC to form a translocon pore in the host cell membrane. This acts as a channel to allow the passage of bacterial effector proteins (29, 30). To probe IpaD as a potential target, we can utilize an FP assay adapted from Dickenson et al, 2013(31). By fluorescently labeling IpaB and co-incubating with IpaD, 1 mM DOC, and GSK inhibitor at 10 μM, we can measure the FP of the interaction. Any compound that prevents the increase in fluorescence polarization of IpaB following its interaction with IpaD will be viewed as a potential IpaD inhibitor. In addition, a direct IpaD binding assay can be performed using surface plasmon resonance (SPR) as described in Dey et al(32) for further analysis of the inhibitors.

Overall, target identification of the *Shigella* virulence inhibitors identified in the primary screen at GSK is essential to improving the efficacy of the potential therapeutic. This is a subset of the essential *S. flexneri* virulence factors but there are numerous more that could be explored.

5.3 Final Statements

The results presented within this dissertation expand on the DNA-binding and dimerization activities of VirF considerably. With this range of knowledge on VirF function, it will provide valuable information for the discovery and development of either VirF DNA-binding or dimerization inhibitors. The proposed future directions provide a strong basis for the continued efforts to study VirF as well as follow-up on the GSK hits. There is still a long way to go to develop a novel virulence-targeting *Shigella* inhibitor but the studies on VirF as well as other bacterial virulence factors provide a great pathway forward to lessen the rise of antibacterial resistance.

5.4 References

1. Tobe T, Yoshikawa M, Mizuno T, Sasakawa C. 1993. Transcriptional Control of the Invasion Regulatory Gene virB of Shigella flexneri: Activation by VirF and Repression by H-NS. J Bacteriol 175:6142–6149.

2. Tran CN, Giangrossi M, Prosseda G, Brandi A, Di Martino ML, Colonna B, Falconi M. 2011. A multifactor regulatory circuit involving H-NS, VirF and an antisense RNA modulates transcription of the virulence gene icsA of Shigella flexneri. Nucleic Acids Res 39:8122–8134.

3. Giangrossi M, Giuliodori AM, Tran CN, Amici A, Marchini C, Falconi M. 2017. VirF relieves the transcriptional attenuation of the virulence gene icsA of Shigella flexneri affecting the icsA mRNA-RnaG complex formation. Front Microbiol 8:1–12.

4. Emanuele AA, Adams NE, Chen YC, Maurelli AT, Garcia GA. 2014. Potential novel antibiotics from HTS targeting the virulence-regulating transcription factor, VirF, from Shigella flexneri. J Antibiot (Tokyo) 67:379–386.

5. Hurt JK, McQuade TJ, Anthony E, Larsen MJ, Garcia GA. 2010. High-Throughput Screening of the Virulence Regulator VirF: A Novel Antibacterial Target for Shigellosis. J Biomol Screen 15:379–387.

6. Porter ME, Dorman CJ. 2002. In Vivo DNA-Binding and Oligomerization Properties of the Shigella flexneri AraC-Like Transcriptional Regulator VirF as Identified by Random and Site-Specific Mutagenesis. J Bacteriol 184:531–539.

7. Childers BM, Weber GG, Prouty MG, Castaneda MM, Peng F, Klose KE. 2007. Identification of Residues Critical for the Function of the Vibrio cholerae Virulence Regulator ToxT by Scanning Alanine Mutagenesis. J Mol Biol 367:1413–1430.

8. Marsden AE, Schubot FD, Yahr TL. 2014. Self-association is required for occupation of adjacent binding sites in Pseudomonas aeruginosa type III secretion system promoters. J Bacteriol 196:3546–3555.

9. Melander RJ, Zurawski D V., Melander C. 2018. Narrow-spectrum antibacterial agents. Medchemcomm 9:12–21.

10. Jernberg C, Löfmark S, Edlund C, Jansson JK. 2007. Long-term ecological impacts of antibiotic administration on the human intestinal microbiota. ISME J 1:56–66.

11. Karam G, Chastre J, Wilcox MH, Vincent JL. 2016. Antibiotic strategies in the era of multidrug resistance. Crit Care 20:1–9.

12. Jernberg C, Löfmark S, Edlund C, Jansson JK. 2010. Long-term impacts of antibiotic exposure on the human intestinal microbiota. Microbiology 156:3216–3223.

13. Daines DA, Granger-Schnarr M, Dimitrova M, Silver RP. 2002. Use of LexA-based system

to identify protein-protein interactions in vivo. Methods Enzymol 358:153–161.

14. Daines DA, Silver RP. 2000. Evidence for multimerization of neu proteins involved in polysialic acid synthesis in Escherichia coli K1 using improved LexA-based vectors. J Bacteriol 182:5267–5270.

15. Dmitrova M, Younès-Cauet G, Oertel-Buchheit P, Porte D, Schnarr M, Granger-Schnarr M. 1998. A new LexA-based genetic system for monitoring and analyzing protein heterodimerization in Escherichia coli. Mol Gen Genet 257:205–212.

16. Miller JH. 1972. Assay of β-Galactosidase, p. 352–355. *In* Experiments in Molecular Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor.

17. Gao X, Zou T, Mu Z, Qin B, Yang J, Waltersperger S, Wang M, Cui S, Jin Q. 2013. Structural insights into VirB-DNA complexes reveal mechanism of transcriptional activation of virulence genes. Nucleic Acids Res 41:10529–10541.

18. Karney MMA, McKenna JA, Weatherspoon-Griffin N, Karabachev AD, Millar ME, Potocek EA, Wing HJ. 2019. Investigating the DNA-Binding Site for VirB, a Key Transcriptional Regulator of Shigella Virulence Genes, Using an In Vivo Binding Tool. Genes (Basel) 10:1–12.

19. Chen Y, Smith MR, Thirumalai K, Zychlinsky A. 1996. A bacterial invasin induces macrophage apoptosis by binding directly to ICE. EMBO J 15:3853–3860.

20. Guichon A, Hersh D, Smith MR, Zychlinsky A. 2001. Structure-function analysis of the Shigella virulence factor IpaB. J Bacteriol 183:1269–1276.

21. SenerovicL., Tsunoda SP, GoosmannC., Brinkmann V, Zychlinsky A, Meissner F, Kolbe M. 2012. Spontaneous formation of IpaB ion channels in host cell membranes reveals how Shigella induces pyroptosis in macrophages. Cell Death Dis 3:e384-10.

22. Garcia-Calvo M, Peterson EP, Leiting B, Ruel R, Nicholson DW, Thornberry NA. 1998. Inhibition of human caspases by peptide-based and macromolecular inhibitors. J Biol Chem 273:32608–32613.

23. Birket SE, Harrington AT, Espina M, Smith ND, Terry CM, Darboe N, Markham AP, Middaugh CR, Picking WL, Picking WD. 2007. Preparation and characterization of translocator/chaperone complexes and their component proteins from Shigella flexneri. Biochemistry 46:8128–8137.

24. Hu B, Morado DR, Margolin W, Rohde JR, Arizmendi O, Picking WL, Picking WD, Liu J. 2015. Visualization of the type III secretion sorting platform of Shigella flexneri. Proc Natl Acad Sci U S A 112:1047–1052.

25. Wagner S, Grin I, Malmsheimer S, Singh N, Torres-Vargas CE, Westerhausen S. 2018. Bacterial type III secretion systems: A complex device for the delivery of bacterial effector proteins into eukaryotic host cells. FEMS Microbiol Lett 365:1–13.

26. Burgess JL, Jones HB, Kumar P, Toth RT, Middaugh CR, Antony E, Dickenson NE. 2016. Spa47 is an oligomerization-activated type three secretion system (T3SS) ATPase from Shigella flexneri. Protein Sci 25:1037–1048.

27. Case HB, Mattock DS, Dickenson NE. 2018. Shutting Down Shigella Secretion: Characterizing Small Molecule Type Three Secretion System ATPase Inhibitors. Rev del Col Am Cardiol 57:6906–6916.

28. Case HB, Mattock DS, III BRM, Dickenson NE. 2020. Novel Noncompetitive Type Three Secretion System ATPase Inhibitors Shut Down Shigella Effector Secretion. Biochemistry 59:2667–2678.

29. Arizmendi O, Picking WD, Picking WL. 2016. Macrophage apoptosis triggered by IpaD from Shigella flexneri. Infect Immun 84:1857–1865.

30. Bernard AR, Jessop TC, Kumar P, Dickenson NE. 2017. Deoxycholate-Enhanced Shigella Virulence is Regulated by a Rare π-Helix in the Type Three Secretion System Tip Protein IpaD. Biochemistry 56:6503–6514.

31. Dickenson NE, Arizmendi O, Patil MK, Toth IV RT, Middaugh R, Picking WD, Picking WL. 2013. The N-terminus of IpaB provides a potential anchor to the Shigella type III secretion system tip complex protein IpaD. Biochemistry 52:8790–8799.

32. Dey S, Anbanandam A, Mumford BE, De Guzman RN. 2017. Characterization of Small-Molecule Scaffolds That Bind to the Shigella Type III Secretion System Protein IpaD. ChemMedChem 12:1534–1541.