# Cross-talk between the core and pathogen-specific genomes in *Escherichia coli* via a bacterial anti-activator protein

Abigail Sue Mickey

Harrisburg, PA

B.S., The Pennsylvania State University, 2013

A Dissertation presented to the Graduate Faculty

of the University of Virginia in Candidacy for the Degree of

Doctor of Philosophy

Department of Microbiology, Immunology, and Cancer Biology

University of Virginia

July 2020

**Acknowledgments**

First, I would like to thank James Nataro for welcoming me into his lab. One of Jim's favorite quotes is by Albert Einstein, "Play is the highest form of research", and his has always allowed me to play and explore my ideas. Jim has been a patient mentor and allowed me the independence I needed to grow as a scientist. He has always been a strong advocate on my behalf, and for that I am especially grateful!

I would also like to acknowledge my lab family. Both Fernando Ruiz and Araceli Santiago have provided great advice and endless encouragement during this long journey. To the ladies that shared my lab space - Rachel Smith, Laura Gonyar, and Chantel Bailey – thank you for listening to my research woes and for simply being great friends.

Thank you to my committee members – Melissa Kendall, Alison Criss, Hervé Agaisse, Dan Engel, and Erik Hewlett. You have all challenged me to view my data with an unbiased, critical eye and present it with precise language. Your advice and comments have shaped me into a stronger scientist.

To my first research mentor at Penn State Harrisburg – Dr. Katherine Baker – your obvious passion for microbiology sparked my love for the subject. You helped that love to grow with your constant support and mentorship. Without you, I never would have sought to further my education.

A giant thanks to my family! To my parents – Frank and Donna – words cannot express my gratitude for your constant love and encouragement. The last

7 years have been filled with ups and a lot of downs, but through it all you have been there. I never could have accomplished so much without you both! To my sister and nephews – Catriona, Braeden, and Sawyer – thank you for loving me and supporting my journey.

And finally, I would like to thank my cats – Fluffy and Comet. Your fuzzy faces lift my spirits on my worst days. You have been my daily support system and a welcome source of laughter. I never could have done it without your silent love and encouragement!

## Table of contents

## List of abbreviations

AAF – Aggregative adherence fimbria

Aap – dispersin

Aar – AggR activated regulator

ANR – AraC negative regulator

CFU – colony forming units

CTD – C-terminal DNA binding domain

DAEC – diffusely adherent *E. coli*

DMEM – Dulbecco's modified Eagle's medium

EAEC – enteroaggregative *E. coli*

EAST1 – EAEC heat-stable toxin

EHEC – enterohemorrhagic *E. coli*

EPEC – enteropathogenic *E. coli*

ETEC – enterotoxigenic *E. coli*

Fis – factor for inversion stimulation

Fur – ferric-uptake regulator

Hcp – hemolysin coregulated protein

H-NS – histone-like nucleoid-structuring protein

HTH – helix-turn-helix

HU – heat unstable

IHF – integration host factor

Km - kanamycin

LB – Luria broth

LEE – locus of enterocyte effacement

MFI – mean fluorescent intensity

mRNA – messenger RNA

NTD – N-terminal domain

Pet – plasmid-encoded toxin

Pic – protein involved in intestinal colonization

RNAP – RNA polymerase

rRNA – ribosomal RNA

ShET1 – *Shigella* enterotoxin 1

SPR – surface plasmon resonance

sRNA – small RNA

T6SS – type VI secretion system

Tss – Type six subunit

UPEC – uropathogenic *E. coli*

tRNA – transfer RNA

WT – Wild type

## Abstract

Bacteria have evolved elegant and complex mechanisms to regulate gene expression and survive in diverse environmental conditions. Pathogenic bacteria are confronted by the compound challenges of both successful competition within their metabolic niche as well as the need to effectively orchestrate virulence gene expression at the right time and place. The regulatory complexities are still being deciphered for a large number of pathogens, and new, highly nuanced systems are being illuminated. This dissertation describes regulatory features of diarrheagenic *Escherichia coli* heretofore unknown.

Several Gram-negative pathogens colonize and compete in certain niches by acquiring new genes that are characterized by their high AT content (compared with the AT content of *E. coli*, which is ca. 50%). Although this aberrant AT speaks of exogenous ancestry, the maintenance of such loci permits concerted expression in part by virtue of proteins that bind such AT-rich segments. Diarrheagenic *E. coli* have thus used the histone-like structuring protein, H-NS, which binds and silences AT-rich genes. Exactly how H-NS is harnessed, and more broadly, how these pathogens both regulate cross-talk between acquired genes and the core *E. coli* genome as well as control the switch between nonpathogenic and pathogenic lifestyles, are subjects of intense and fascinating investigation.

Enteroaggregative *Escherichia coli* (EAEC), a common diarrheagenic *E. coli* pathotype, utilizes a member of the AraC family of transcriptional regulators, AggR, to regulate the expression of virulence genes which are necessary for host

colonization and virulence. In addition to controlling/promoting expression of virulence factors, AggR also regulates the expression of its own negative regulator, *aar*, which encodes a small protein that binds *in vitro* to the dimerization site of AggR and thus blocks AggR function. Recent transcriptomic profiling of EAEC strain 042 with and without functional *aar* revealed that core genes outside of the AggR regulon are also regulated by Aar (perhaps via AggR, or independently). One of the non-AggR-regulated genes encoded the global regulator H-NS. Moreover, published data from our research group suggest that Aar directly binds to H-NS and effects H-NS repression.

In this dissertation, I identify the functional significance of both Aar/AggR and Aar/H-NS binding events in the pathogenesis of EAEC. I propose a novel role for Aar that includes acting as an anti-activator of AggR and an anti-repressor of H-NS. My data suggest that during the initial steps of pathogenesis, Aar removes H-NS silencing at AggR-regulated AT-rich genes, thereby allowing unobstructed activation of these genes by AggR. Then, as the concentration of Aar increases, Aar binds to AggR and prevents further gene activation. In addition, I demonstrate that the expression of *aar* affects bacterial fitness *in vivo,* and I speculate that this occurs through modulation of *aggR* and *hns* expression. In summary, my work extends the characterization of a novel, but common, regulatory protein that regulates both virulence and core metabolic genes, and provides new insights into pathogenic gene regulation.

# Chapter One

## Introduction

## Enteroaggregative *Escherichia coli* infection and burden

First described in 1987 in a prospective study of pediatric diarrhea in Chile (1), enteroaggregative *Escherichia coli* (EAEC), a Gram-negative enteric pathogen, is now frequently associated with acute and persistent diarrhea in children, traveler's diarrhea, and diarrheal outbreaks (2–4). In some cases, EAEC has also been associated with urinary tract infections (5) and bacteremia leading to sepsis (6). Of increasing concern is the association of both clinical and subclinical EAEC infections with chronic intestinal inflammation which leads to childhood malnutrition and growth impairment (7–11). Surprisingly, subclinical infections are more significantly associated with growth deficit (8). The likelihood of EAEC presenting as a clinical infection characterized by watery diarrhea depends on host genetic susceptibility and EAEC genomic combination of virulence genes (8, 12). During diarrheagenic EAEC infection, pro-inflammatory cytokines IL-8, IL-6, IL-1β, and TNF-α, as well as TLR-5 expression are increased, but only IL-8 has been shown to be secreted (2). Even asymptomatic carriers have been shown to express increased levels of fecal IL-8, IL-1β lactoferrin, and leukocytes (2).

EAEC is transmitted by the fecal-oral route (13). Risk factors for EAEC include ingestion of contaminated food or water, travel to developing countries, poor hygiene, and host susceptibility (13, 14). Treatment of EAEC infections usually includes oral rehydration therapy as these infections are usually self-limiting and resistant to multiple antibiotics (13).

**EAEC virulence**

EAEC has been shown to colonize both the small and large intestinal mucosa (2) above the enterocyte brush border in a thick mucus layer (13). EAEC pathogenesis occurs in three stages: 1) adherence to the intestinal mucosa and formation of biofilm on the surface of enterocytes via aggregative adherence fimbria (AAF) and other adherence factors; 2) induction of inflammation and inflammatory cytokines; and 3) production of toxins resulting in mucosal toxicity and intestinal damage characterized by enlarged crypt openings, microvillus vesiculation, and increased epithelial cell extrusion (8, 13, 15, 16).

The prototypical EAEC strain 042 harbors the virulence plasmid pAA2 (IncFIC) encoding the virulence master regulator AggR and other virulence genes (17, 18). Toxins can be plasmid-encoded or chromosome-encoded. Toxins produced by EAEC are destructive to the tips and sides of intestinal villi and enterocytes (13). The plasmid-encoded toxins include: plasmid-encoded toxin (Pet), a cytopathic serine protease autotransporter that induces cell elongation, rounding, and exfoliation (19), and EAEC heat-stable enterotoxin (EAST1), encoded by *astA,* which is similar to the heat-stable toxin of enterotoxigenic *E. coli* (ETEC) (3, 12). The chromosome-encoded toxins include: protein involved in intestinal colonization (Pic), an autotransporter mucinase that promotes gastrointestinal colonization; and *Shigella* enterotoxin 1 (ShET1), whose mode of action is not yet understood (12, 13).

Host response to EAEC infection is dependent on the EAEC strain and the host innate immune response. Although several EAEC virulence genes have been

characterized, their clinical outcome during infection is not completely clear due to the genomic heterogeneity among isolated strains. The prevalence and association of virulence genes with diarrheal disease remains unclear since they vary by geographical location (7, 11). While EAEC strains harboring putative virulence genes are not always associated with disease, these virulence factors are associated with increased inflammatory markers and fecal cytokines (13). It has been shown that EAEC strain 042 adheres variably to human intestinal enteroids, demonstrating how unidentified inter-patient variations are contributing to EAEC disease (8, 20).

As there has been no single factor that appears to be consistent in symptomatic disease, a combination of plasmid-borne and chromosomal virulence genes under the control of AggR has been hypothesized (12). These factors are described below.

*AggR, the transcriptional activator of virulence genes*

AggR is a member of the AraC family of transcriptional activators, which regulate virulence gene expression in response to environmental changes (12). The AraC family is one of the largest groups of regulatory proteins in bacteria, with currently more than 10,000 homologs (21). EAEC and other mucosal colonizing pathogens tightly regulate virulence expression at the transcriptional level (22) via their respective AraC homologs (23). AraC family proteins have two structural domains: a conserved helix-turn-helix (HTH) C-terminal DNA binding domain

(CTD) and a variable N-terminal signaling domain (NTD) (21). Generally, the NTD is responsible for protein dimerization and inducer ligand binding (21).

Activation of the *aggR* promoter likely occurs through the “light switch” model best described in the activation of the *araBAD* promoter. When arabinose (the inducer) is absent, an AraC dimer binds to a distal half-site and a proximal half-site leading to looping of DNA and repression of the operon. However, when the inducer is present, it binds to the AraC dimer causing an allosteric change that allows the dimer to bind two adjacent proximal sites, leading to the opening of the DNA loop and transcriptional activation (21, 24, 25). Inducer molecules are environmental chemicals such as bicarbonate ions for RegA in *Citrobacter rodentium*, or bile fatty acids for ToxT in *Vibrio cholerae* (21). The actual inducer for the AggR regulon has not been determined, but the addition of glucose or maltose has been shown to induce a strong biofilm – a phenotypic readout for *aggR* expression (26).

Regulation of AggR can occur transcriptionally or post-transcriptionally. Transcriptionally, *aggR* is repressed by histone-like nucleoid-structuring protein, H-NS, binding to the promoter region (27). When EAEC is in an environment outside of the host, it conserves nutrients and energy by silencing virulence genes. This is mediated by the global regulator, H-NS (21). One way to counteract H-NS mediated silencing and allow virulence gene expression is via anti-repression (21). Anti-repression involves competition between a positive regulator, such as a member of the AraC family of transcriptional regulators (here AggR), and H-NS for DNA-binding at the promoter; the degree of regulation is dependent on the

presence or absence of the AraC family transcriptional regulator inducer and the abundance of the anti-repressor and repressor (21).

In addition to transcriptional repression, *aggR* can be transcriptionally activated by factor for inversion stimulation (Fis) and itself (26, 27). Fis an 11.2 kDa protein, which is strongly induced during logarithmic growth phase when there is a nutrient upshift, activates transcription of *aggR* by disrupting H-NS binding to DNA (28, 29). AggR acts to autoactivate its own expression by binding to a specific AggR binding sites within the DNA sequence for which there are two predicted: one upstream and one downstream of the transcriptional start site (27). However, studies suggest that only the upstream binding site is required for the ability of AggR to positively regulate itself (27).

Post-transcriptionally, *aggR* is repressed by a member of the AraC negative regulator (ANR) family, namely Aar (30). Aar is thought to disrupt AggR dimerization and binding to DNA by binding with high affinity to the central region of AggR, the dimerization domain (31). The mechanism of this is described in more detail later.

AggR was initially discovered for its role in the induction of aggregative adherence fimbriae (AAF) expression, but has since been further characterized to regulate approximately 44 plasmid-borne and chromosomal genes (12, 32, 33). Genes regulated by AggR include itself (27), the AAF, dispersin (*aap*), the *aat* secretion system, a type VI secretion system (*aaiA-Y*), and the AggR negative regulator (*aar*) (figure 1.1) (32).

*Aggregative adherence fimbria (AAF)*

EAEC is defined by its characteristic aggregative pattern of adherence to HEp-2 epithelial cells resembling stacked bricks (12). This adherence pattern is mediated by AAFs, which are semiflexible bundle-forming organelles (12). AAFs are related to the Dr family of adhesins, which are adhesins in diffusely adherent *E. coli* (DAEC) and uropathogenic *E. coli* (UPEC) that assemble via the usher/chaperone pathway (12, 34). As with the Dr family of adhesins, AAFs are encoded by four biogenesis genes, termed A, B, C, and D; comprising the major structural pilin subunit, the minor subunit, the usher, and the chaperone (respectively) (35, 36). Currently, there are five AAF variants, all of which are regulated by AggR (7): the major pillin of AAF/I is encoded by *aggA* (33); AAF/II is encoded by *aafA* (35); AAF/III is encoded by *agg3A* (13); AAF/IV is encoded by *agg4A* (previously *hdaA*) (37); and AAF/V is encoded by *agg5A* (38). Each fimbrial variant is only present in a minority of strains and generally only one variant is present in a strain (12, 39). To date, only 3% of characterized EAEC strains harbor and transcribe genes encoding both *agg3A* and *agg5A* pilin subunits which share the accessory genes encoded by *agg3DCB* (39).

The minor subunits (B) of AAF have a conserved structure, but the major subunits (A) of AAF have large structural differences that affect the pilus architecture (36). The diameter of individual fimbriae is 2-3 nm for AAF/I (40), 5 nm for AAF/II (40), and 3-5 nm for AAF/III; the latter are usually observed as individual filaments rather than the characteristic AAF bundles (41).

Although there are structural differences among the AAF variants, all variants (excluding AAF/V (42)) recognize a common receptor, fibronectin (36). AAF attachment to fibronectin is primarily based on electrostatic interactions by clusters of basic residues at the junction between pilus subunits (36). In addition to fibronectin, other extracellular matrix proteins such as laminin and cytokeratin 8 have been shown to bind the major subunits of AAFs (43). Bacterial interaction with extracellular matrix proteins, which are generally localized to the basement membrane, may occur during inflammation and the opening of tight junctions (44). AAF expression can disrupt tight junctions, inducing aberrant localization of claudin-1 and occludin, potentially permitting the penetration of bacterial toxins or the bacteria to the intestinal submucosa (45). Disruption of the tight junctions can also exacerbate diarrheal disease due to the loss of host ions and proteins into the lumen (45). MUC1, a transmembrane mucin expressed throughout the human gastrointestinal tract, has been identified as a possible luminal host cell receptor for EAEC (46). However, not all EAEC strains adhere via AAFs; in some atypical (strain does not encode *aggR*) EAEC strains, adherence is via a type IV fimbriae (12).

Not only are AAFs required for adherence to the small and large intestinal mucosa, but they are also required for biofilm formation (26). It has been reported that of all the diarrheagenic pathotypes, EAEC is the most likely to be resistant to antibiotics, presumably due to its ability to form biofilms (15).

EAEC strains carrying AAF/II have been found to be the most virulent in nature based on both genome studies on strains isolated from people and

experimental studies comparing virulence of strains encoding different AAF variants (2, 7, 8). This virulence is potentially due to more diverse virulence markers and induction of a more pronounced inflammatory response (2, 7, 8). In EAEC strain 042, genes required for AAF/II expression are encoded on the 65-MDa virulence plasmid pAA2 in a 23 kb cluster containing two regions (35). One cluster contains the chaperone (*aafD*), major subunit (*aafA*), and transcriptional activator (*aggR*); the other contains a nonfunctional chaperone (*aafD'*), usher (*aafC*), and minor subunit (*aafB*) (35). Interestingly, only *aafDA* appear to be AggR-regulated (35). During assembly, the major subunits (encoded by *aafA*) use donor strand complementation to assemble into linear polymers and with a single minor subunit (*aafB*) inserted at the polymer tip (36).

*Dispersin, an anti-aggregation protein*

Encoded by *aap* and regulated by AggR, dispersin is a 10-kDa secreted protein whose predicted function is to disperse EAEC across the intestinal mucosa (12, 47). EAEC strains harboring *aap* mutations aggregate more intensely, resulting in fewer individual bacteria and larger aggregates (47). Dispersin is a positively charged protein that is proposed to act by neutralizing the bacterium's negatively charged lipopolysaccharide surface, thereby releasing the positively charged AAF to extend from the bacterial surface (12, 47). Dispersin has been shown to be highly immunogenic in a volunteer challenge, making it a potential vaccine candidate (48). However, the dispersin-encoding gene (*aap*) has been identified in other pathogenic and nonpathogenic *E. coli* suggesting that it may not

be sufficient for pathogenicity, thus reducing enthusiasm for it as a vaccine candidate (49).

*Type VI secretion system, T6SS*

Bacteria can compete for an ecological niche by utilizing a type VI secretion system (T6SS) to translocate antibacterial effectors from predator bacteria to prey cells (host or other bacteria) (50). This translocation is achieved by the assembly of a macromolecular complex that spans the cell envelope and is composed of a core with 13 Type six subunits (Tss) and an extracellular sheath wrapped around a tube of hemolysin coregulated protein (Hcp) (51, 52). It is predicted that the sheath extends and contracts to propel the Hcp tube toward the prey cell (52). A second protein, valine-glycine repeat protein G (VgrG), caps the Hcp tube and acts as the puncturing device (52). ClpV, an AAA+-family of ATPase, recycles the Tss core components during contraction and is required for T6SS function (52). Expression of the T6SS can be controlled by environmental signals, bacterial signals, and quorum sensing (52).

EAEC strain 042 encodes two T6SS (Sci-I and Sci-II); however, only Sci-II is under AggR control (53). Sci-I is required for biofilm formation (52, 54) and is regulated at the transcriptional level by iron availability through the ferric-uptake regulator (Fur) repression and DNA-adenine methyltransferase (Dam)-catalyzed methylation (52, 55). In the absence of iron, Fur dissociates from its DNA binding sites and allows RNA polymerase to bind and initiate transcription. The loss of Fur also permits methylation at the site inhibiting the rebinding of Fur (52, 55).

Currently the physiological consequences of Sci-I are not known; however, since the anaerobic environment of the intestinal lumen favors the ferrous [Fe(II)] form of iron, it is predicted that Sci-I remains repressed by Fur *in vivo* (52). This prediction is supported by studies demonstrating that Sci-I has no contribution to virulence in animal models (52, 53).

Sci-II is upregulated in biofilm or aggregated conditions (53) and provides a growth advantage to the pathogen when present in a culture with nonpathogenic *E. coli* by killing the nonpathogenic *E. coli* (56). Sci-II is a cluster of 25 contiguous genes (*aaiA-aaiY*) located on the 117 kb pathogenicity island termed the PAI-1 that is inserted at the *pheU* transfer RNA (tRNA) locus on the chromosome (32, 34, 53). Genes within Sci-II include: *aaiP* encoding a homolog of ClpV; *aaiG* encoding a homolog to Vgr; *aaiC* encoding a putative effector protein associated with malnutrition in an epidemiological study (11, 53); and *aaiQ, aaiR,* and *aaiWXY* share sequence homology to mobile elements. It has been predicted that due to truncations in N- and/or C-terminus of *aaiQ, aaiR,* and *aaiX*, that these three ORFs are pseudogenes (53). However, genes spanning *aaiA-aaiP* are sufficient for AaiC secretion (53).

There is a specificity during assembly of the two T6SS tails, as the Hcp in Sci-I only interacts with sheath components from Sci-I and not from Sci-II (57). This was also true for disassembly of the contracted sheath via ClpV, where ClpV from Sci-I could not compensate for the ClpV from Sci-II and vice versa (57).

*AggR activated regulator, Aar*

Aar belongs to a large family of proteins that downregulate AraC homologs in at least 26 distinct pathogenic Gram-negative species termed AraC negative regulators (ANR) (31). The predicted protein structure of members of the ANR family includes three highly conserved α-helices, with at least α-helix 2 and 3 required for activity and oligomerization (31). Members of the ANR family can complement the function of other ANRs whose AraC homologs are phylogenetically close (i.e. Aar can downregulate Rns and CfaD from enterotoxigenic *E. coli*) (31).

In EAEC, Aar is thought to act as an anti-activator by binding directly to the dimerization domain of AggR (protein-protein, figure 1.3D), thus preventing AggR dimerization and function. Binding assays, such as surface plasmon resonance and bacterial two-hybrid assays with full length and truncated AggR, demonstrated Aar binding directly to AggR at the dimerization domain (AggR residues 69-181) (31). Bacterial two-hybrid assays and electrophoretic mobility shift assays suggested that Aar binding to AggR prevents AggR homodimerization, AggR binding to DNA, and functional activity (figure 1.2 blue cycle) (31). There is no evidence that Aar repressed AggR via binding to DNA or indirectly through another regulatory protein (31).

In addition to downregulating AggR in EAEC strain 042, Aar can regulate genes in an AggR-independent way, both positively and negatively (58). In fact, 88% of Aar regulated genes are AggR-independent (58). These genes encoded hypothetical proteins, proteins involved in metabolic function, transporter proteins, regulatory proteins, putative virulence factors, membrane proteins, phage proteins,

and proteins involved in diverse other functions (58). Of great interest was Aar's upregulation of nucleoid-associated proteins: H-NS, a putative histone-like DNA-binding protein, and StpA (58). To date, only Aar's relationship with H-NS has been extensively characterized. Aar was found to bind directly to H-NS in a bacterial two-hybrid assay and it was predicted that the second α-helix interacts with the oligomerization domain of H-NS (58). Aar binding to H-NS alters H-NS binding to DNA in distinct ways at different promoters, and therefore the effect of Aar on the H-NS regulon is heterogenous (figure 1.2 gold cycle) (58).

In an *in vivo* mouse model, *aar* was highly expressed in the intestinal lumen during the early stages of infection and expression was dependent on AggR. In addition, *aar* is required for maximal *hns* expression (58). In epidemiological studies, *aar* is more commonly associated with protection from diarrhea (7), and is a marker of less severe disease (8). In addition, the presence of *aar* is associated with nourishment (11).

**Gene regulation in bacteria**

Bacteria conserve energy by only expressing genes necessary for survival in the appropriate context (59). Gene regulation can occur at a transcriptional level or post-transcriptional level; however, transcriptional regulation is the most common and most efficient means of control (59, 60). Transcriptional regulation can occur at the level of DNA accessibility and binding (transcription initiation) or at the messenger RNA (mRNA) level (posttranscription initiation) (59, 60).

Most transcriptional regulation occurs during transcription initiation by DNA-binding transcriptional regulators with a HTH motif (59–61). Transcriptional regulators can be positive, negative, or both. During positive regulation, an activator is required for transcription initiation by RNA polymerase (RNAP). The activator may 1) increase the strength of RNAP binding to the promoter, 2) bring recognition sites together via bending or rotating the promoter, 3) enable RNAP to open the DNA strands, or 4) allow RNAP to escape the promoter and continue to the first gene (59, 60). A classic example of a positive regulator is AraC in *Escherichia coli* that induces the L-*ara* operon, which is responsible for utilization of L-arabinose (24, 25, 62). For negative regulation, a repressor binds to a DNA sequence and blocks transcription initiation by RNAP by 1) binding to the RNAP binding site, 2) bending the promoter and blocking RNAP binding, or 3) preventing RNAP from leaving the promoter (59, 60). A classic negative regulator is LacI in *E. coli* that binds to the operator regions of the promoter, bends the DNA, and prevents RNAP binding; thus, repressing the *lac* operon (responsible for utilization of lactose) in the absence of lactose (59, 63). Some regulators can act as both a positive and negative regulator; the function is determined by where the regulator binds to on the promoter (59).

During attenuation of transcription, transcription is initiated at the promoter region but if the gene products are not needed, transcription ends or slows before RNAP reaches the first structural gene (59, 60). A classic example is the *trp* operon in *E. coli* and *Bacillus subtilis* which encodes enzymes responsible for synthesizing the amino acid L-tryptophan. In this system, termination is controlled by speed of

translation, which depends on the presence or absence of tryptophan: in the presence of tryptophan, the ribosome translates through the *trp* codons resulting in an RNA terminator hairpin; in the absence of tryptophan, the ribosome pauses at the *trp* codons resulting in the formation of an anti-terminator loop (59, 60, 64).

Leader RNA, which are mRNA sequences upstream of the initiation codon for the first gene and are usually not translated, can also have a regulatory role via secondary structural changes that block access of the ribosome to the translation initiation region. This secondary structure can be changed or "melted" with the increase in temperature that breaks mRNA base pairing (59). The gene for the heat shock sigma, $\sigma^{32}$, in *E. coli* is an example of this (59, 65). This method is also used by pathogenic bacteria to turn on virulence genes that allow them to propagate in the host (i.e. the expression of *lcrF*, the gene encoding low calcium response gene F in *Yersinia pestis*) (59, 66).

Posttranscriptional regulation can also be achieved through riboswitch regulation, where the structure of mRNA can be altered through the binding of a small effector molecule (amino acids, vitamins, or nucleic acid bases) to the RNA (59). The transcription of genes encoding the aminoacyl-tRNA synthetase in *B. subtilis* is the first example of this (59). Transcription of these genes is upregulated through anti-termination when the bacterium is deprived of a specific amino acid, causing that amino acid's tRNA to be unaminoacyated, leading to more efficient attachment of the amino acid to its cognate tRNA (59, 67).

Small RNA molecules (sRNA) can act at both initiation and extension. At the level of transcription initiation, base pair binding of sRNA molecules to a strand

of DNA near the promoter can block ribosome binding (60). At the level of posttranscription initiation, sRNA molecules can bind to mRNA and change the stability of mRNA by triggering degradation (60).

Gene regulation can also occur post-translation by feedback inhibition. During feedback inhibition, the end product binds to the first enzyme and inhibits its activity. An example of this is the tryptophan pathway in *E. coli,* where tryptophan binds to anthranilate synthetase, the first enzyme of the tryptophan biosynthetic pathway, blocking the synthesis of tryptophan (59).

*Anti-activation*

Aar and other members of the ANR family act as anti-activators and block the function of their respective AraC homolog activators. Anti-activators function via protein-DNA or protein-protein binding. In protein-DNA binding, the anti-activator will bind to the DNA sequence upstream of the target promoter and block either binding by the activator (figure 1.3A) or interactions between the activator and RNAP (figure 1.3B). Protein-DNA binding anti-activation is seen in the opportunistic pathogen *Burkholderia cenocepacia,* where an AraC-type transcriptional homolog, CepS, is blocked from activating its target genes by the DNA bound LuxR-type transcriptional factor, CepR2 (figure 1.3B) (68). Only in response to an increase in a quorum sensing pheromone is CepR2 released from DNA (68).

In protein-protein binding, the anti-activator blocks the activating signal by either binding to the activator in an uninduced state, only freeing the activator in

the presence of an inducing signal (generally, the inducing signal causes the anti-activator to change binding partners) (figure 1.3C); or by binding to the activator only after an inducing signal or threshold is reached and blocking the binding of the activator to the DNA (figure 1.3D). Examples of the first include the type III secretion system in *Pseudomonas aeruginosa* and the bacterial phage shock protein stress response in both *E. coli* and the intestinal pathogen *Yersinia enterocolitica.* In these examples, activators ExsA and PspF (respectively) are only freed from their anti-activators (ExsD and PspA, respectively) when the system is induced; freeing a secondary protein for the anti-activator to bind instead of the activator (69, 70). Another example is the type III secretion system in *Shigella* spp. where activator MxiE is only freed from the anti-activator, OspD1, when the system is induced and concentrations of OspD1 are reduced (71). Then there are cases where the anti-activator only binds the activator after an inducing signal or threshold is reached; such is the case during quorum sensing in the plant pathogen *Agrobacterium tumefaciens* and the opportunistic pathogen *Pseudomonas aeruginosa* (72–76). Virulence factors coordinated by quorum sensing are activated by DNA binding by activators TraR (*A. tumefaciens*) or LasR (*P. aeruginosa*), and the threshold and response effects are controlled by the anti-activators TraM or QslA binding to the activators, thereby inducing an allosteric conformational change that disrupts the activators' binding to the DNA (72–76). In attaching-and-effacing intestinal pathogens such as *Citrobacter rodentium,* enteropathogenic *E. coli* (EPEC)*,* and enterohemorrhagic *E. coli* (EHEC), Ler is a positive regulator for most of the *LEE* (the locus of enterocyte effacement)

operons, including *grlRA* (encoding *Ler* activator GrlA, and anti-activator GrlR). To activate *Ler*, GrlA binds to promoter DNA fragments via its HTH motif, but in the presence of GrlR, GrlR binds the HTH motif of GrlA and displaces the DNA promoter fragment, thus blocking *Ler* activation (77, 78). These examples highlight the important role of anti-activators in pathogenesis and adaptation to the host environment.

*Nucleoid-associated proteins*

Gram-negative bacteria encode at least 12 types of nucleoid-associated proteins, each with their own DNA-binding preferences and expression patterns (79). These proteins are categorized as a family due to shared features such as binding to nucleic acids, low molecular masses, and basic charge (79). The most well-known include H-NS, Fis, heat unstable (HU), and integration host factor (IHF) (28). Structurally, Fis and H-NS are distinct, whereas the amino acid sequence of HU and IHF are homologous (28). The major histone-like proteins are not all expressed simultaneously. In contrast to the mostly constant levels of H-NS throughout growth, Fis is highly expressed during early exponential growth and IHF is expressed most abundantly during stationary phase (28). The expression of HU subunits varies with growth, but declines during stationary phase (28). HU and H-NS are two of the most abundant nucleoid associated proteins (80).

*Fis/IHF/HU* - Fis is an important regulator of virulence genes in bacterial pathogens such as *Shigella flexneri*, *Salmonella enterica* serovar Typhimurium, EPEC, and EAEC (28). Fis positively influences promoters involved in expressing components of the cellular protein expression machinery such as translation elongation factors, ribosomal proteins, tRNA, and ribosomal RNA (rRNA) (28). Fis- and H-NS-binding sites partially overlap (81), which allows Fis to disrupt H-NS repression at promoters such as *aggR* and *hns* (28). Fis can also act as transcriptional repressor of its own promoter and genes encoding DNA gyrase (*gyrA* and *gyrB*) via promoter occlusion or by preventing isomerization of a closed to an open transcription initiation complex (28).

IHF, consisting of an αβ structure, binds to unspecific sequences and bends DNA (82). However, IHF can also interact with specific, conserved nucleotide sequences to induce a sharp, 180˚ U-shaped bend that provides transcriptional control and aids in recombination and transposition (82, 83). In *Salmonella*, IHF has a positive regulatory role in the expression of *Salmonella* Pathogenicity Island-1 (SPI-1) genes and has been implicated in coordinating regulation of virulence gene expression with adaptation to stationary phase (84).

HU, a homolog of IHF, exists as a heterodimer of homologous subunits α and β (82). Dimerized HU binds and bends DNA that contains structural aberrations such as single-stranded lesions (28). HU and H-NS have different structural roles, where H-NS compacts DNA and HU does not. In fact, HU opens circular DNA molecules like the plasmid pUC19 and can relieve the repression of

H-NS (28). In *Salmonella*, HU affects the expression of genes in both SPI-1 and SPI-2 and has a role in bacterial invasion of epithelial cells, as well as survival in macrophages (85).

DNA bending usually promotes the formation of H-NS binding to DNA; however, DNA could also be bent in less favorable conformations for H-NS binding by other DNA benders Fis or IHF (86).

*H-NS* - H-NS, a 137 amino acid protein that is 15.4kDa (87), is involved in DNA organization, compaction, and transcriptional silencing, affecting up to 5% of *E. coli* genes (28, 88–90). The total amount of H-NS per cell is 14,000 monomers, and is higher in stationary phase than exponential growth due to the increase in DNA concentration (81). In prokaryotes, genetic material and associated molecules are located in a macromolecular complex called a nucleoid (88). DNA condensation occurs via H-NS non-specific interactions with DNA (89) in a similar manner as eukaryotic histones; however, H-NS does not have any sequence homology with histones (81). Organization and silencing require DNA binding and protein oligomerization (89, 91). Although H-NS is classified as a DNA-binding protein due to its strong binding to double-stranded DNA, H-NS has an affinity for all types of nucleic acid (81).

H-NS is composed of two domains connected by a flexible linker: the C-terminal DNA-binding domain composed of an antiparallel β-sheet and 2 α-helices (92), and the N-terminal oligomerization domain composed of 4 α-helices (81, 93, 94). The first 64 residues in the N-terminal domain are necessary for dimerization

(the basic building block of H-NS oligomerization) and residues 65-89 are partially responsible for the formation of higher-order structures (93). Higher-order structures, which are concentration-dependent (93), result from the oligomerization of dimers (28) and are required for DNA-binding activity (81).

H-NS monomers self-associate in a head-to-head (α-helix 1 and 2 of the N-terminal domain) and tail-to-tail (α-helix 3 and 4 of the N-terminal domain) manner to form a superhelix with protruding DNA binding sites containing a conserved Q/RGR amino acid motif that act as a hook to dock with the DNA minor groove (94, 95). This H-NS filament then acts to zipper adjacent DNA together (28). H-NS can be found in three isoforms suggesting H-NS may undergo post-translational modification, and all of these isoforms strongly bind to DNA (81).

DNA binding by H-NS has no consensus sequences but instead has a consensus structure – curved DNA (81). DNA bending promotes the binding of H-NS to DNA, but H-NS itself can also bend DNA (81, 91). H-NS recognizes the minor groove of AT-rich DNA with the affinity varying depending on the sequence: the highest affinity binding occurring at the “TpA step” which is a thymine followed immediately by an adenine (96). H-NS bending of DNA and recognition of curved DNA is dependent on the ability of H-NS to form oligomeric structures between the DNA-bound H-NS proteins (89, 91).

H-NS modulates many genes that are regulated by environmental signals such as temperature, osmolarity, growth phase, pH, or anaerobiosis (81). H-NS can stimulate gene expression such as genes involved in motility and chemotaxis, but only a few of these genes have been identified and the mechanism by which

this occurs is unknown (81, 87). The H-NS to DNA ratio is fairly constant *in vivo* but increases during cold shock (81). During the first 3-4 hours of cold shock, *hns* is prominently expressed via the binding and induction of the major cold-shock protein CspA on the *hns* promoter (81). This increase in H-NS was concomitant with a temporal inhibition of growth (81). Just the overproduction of H-NS appears to result in inhibition of growth, as growth declined at 37˚C when H-NS was overproduced from an inducible promoter (81). Several H-NS repressed promoters are induced by high osmolarity (i.e. the *proU* promoter). It is likely that de-repression occurs due to changes in the oligomerization state of H-NS which is sensitive to increased salt concentrations (81).

H-NS represses several genes that can be induced by a positive regulator such as the AraC family of transcriptional regulators (AggR in EAEC) (81). In some cases, the absence of H-NS removes the need of a positive regulator for gene expression (i.e. expression of *cfaAB*, encoding CFA/I fimbrial components, by the activator CfaD in ETEC; expression of *ctx*, encoding cholera toxin, by activators ToxT and ToxR in *Vibrio cholerae*) (97, 98). Positive transcriptional regulators and their H-NS repressed target genes (virulence factors and certain pili) can also be thermoregulated, expressed at 37˚C but not ambient temperatures (81). H-NS has been implicated in thermoregulation as the inactivation of *hns* leads to the expression of these thermoregulated H-NS repressed target genes at low temperatures (81).

H-NS also regulates its own expression. H-NS binds upstream of its own promoter, thus resulting in transcriptional auto-repression (99). This autoregulation

is important for maintaining a constant H-NS to DNA ratio during growth (81). Fis can stimulate *hns* expression directly by counteracting H-NS repression (81). H-NS expression can be negatively regulated post-transcriptionally by an anti-sense RNA (DsrA) and an RNA-binding protein (Hfq) (79).

There are three proposed mechanisms by which H-NS regulates transcription: 1) before transcription by blocking RNAP from binding to the promoter region (the H-NS nucleation motif and the RNAP promoter -10 element have potential overlapping sequences (95)) as is the case for the *proVWX* operon in *E. coli* (figure 1.4A) (100), or after transcription initiation 2) by either blocking the RNAP elongation process promoting Rho-dependent termination (i.e. *bgl* operon in *E. coli*) (figure 1.4B) (101) or 3) by forming a loop that interferes with promoter clearance (i.e. *rrnB* promoter) (figure 1.4C) (28, 86). H-NS can also indirectly act to repress gene expression via *rpoS*, the RNA polymerase sigma factor, $\sigma^S$ (81). H-NS affects *rpoS* expression at the post-transcriptional level by either influencing the efficiency of *rpoS* mRNA translation by acting as an inefficient RNA chaperone or the stability of RpoS by affecting the synthesis of factors involved in protein degradation (81). The expression of *rpoS* increases in high-osmolarity medium, at low temperature, upon entry into stationary phase, and at the onset of carbon/phosphate starvation (81).

Under non-intestinal environmental conditions, H-NS represses gene expression; however, this repression is relieved in response to specific stimuli in the intestinal environment (81). Counter-silencing can occur via the disruption of H-NS complexes by multimerization antagonists (i.e. H-NS like proteins such as

StpA, Hha or a truncated H-NS molecule, H-NST) (figure 1.5A), competition for DNA-binding by high-affinity sequence-specific DNA-binding proteins (i.e. anti-repressors such as AggR) (figure 1.5B), activation of certain promoters by the alternative σ-factor RpoS (figure 1.5C), or by changes in promoter geometry due to protein binding or environmental changes which disrupt H-NS complex formation (i.e. temperature-induced alterations at *virF* promoter) (figure 1.5D) (102). These mechanisms result in either the stabilization of RNAP binding or by allowing RNAP binding and procession (96). These mechanisms may not be mutually exclusive (102) and may not require the removal of H-NS from promoter regions (96). In pathogenic *E. coli*, Ler antagonizes H-NS by competing for binding to DNA (96, 103). VirB in *Shigella flexneri* is thought to counter-silence H-NS by binding to DNA, oligomerizing, and bending the DNA thus disrupting the H-NS-DNA filament (96). LeuO, a LysR-type transcriptional regulator in *Salmonella* and *E. coli,* is thought to create a barrier to H-NS polymerization by binding to DNA as a tetramer between the H-NS binding sites and promoter (96). Similar to VirB, LeuO is able to bend DNA and wrap it around its oligomerized forms (96).

H-NS-like proteins include StpA, Hha, and YdgT and they generally interact with H-NS and modify its DNA binding or oligomerization properties (95). StpA shares 57% sequence identity with H-NS and has overlapping biochemical activities – binding to curved DNA, transcriptional repression, and constraints of supercoils (81, 104). H-NS and StpA can form heteromeric structures (28), or StpA can form homomeric complexes via the N-terminal domains (105). Although H-NS

can substitute for StpA (104), StpA can only substitute for H-NS at some genes (81).

Hha has no clear DNA-binding domain but has a high degree of homology to the N-terminal domain of H-NS (105). In *Enterobacteriaceae*, the Hha family of proteins (including paralog YdgT in *Salmonella*) regulates virulence gene expression in response to environmental cues by forming complexes with H-NS-like proteins (105). By binding to the oligomerization domain of H-NS, Hha modifies H-NS oligomerization which influences the DNA binding activity of H-NS, altering H-NS target genes (95, 105). H-NS interacts with member of the Hha family to specifically target horizontally acquired genes but not core genes (105, 106). Hha has been shown to regulate genes in an H-NS/StpA-independent manner, with a significant number of these genes being involved in pathogenicity, cell motility, and iron uptake (105).

Horizontal gene transfer drives bacterial evolution by enabling the bacteria to evolve in significant bounds by acquiring new traits in a single genetic event instead of gradually accumulating beneficial mutations over time (96). However, unregulated expression of these new genes can impose a significant fitness cost and competitive disadvantage (96). Horizontally acquired genes have higher AT-content than ancestral genes, thus allowing bacteria to discriminate between “self” and “non-self” sequences, so that foreign gene expression can be silenced (96). H-NS is hypothesized to buffer the fitness consequences associated with the acquisition of foreign genes (107). This has been demonstrated in multiple pathogenic backgrounds. For example, H-NS has been shown to be central to the

evolution of *Salmonella* and the SPI-1 type III secretion system by buffering the negative fitness cost (107). In an H-NS mutant, *Salmonella* has severe growth defects that are primarily due to the misregulation of SPI-1, given many generations the bacteria compensate for this fitness cost by acquiring missense mutations in the H-NS paralog StpA which alter StpA's DNA binding and oligomerization properties to resemble H-NS. In addition, deletions within SPI-1 are acquired (107). In *E. coli* O157:H7, H-NS enables the utilization of a large variety of carbon and nitrogen sources for maximal carbon acquisition and energy generation, and regulates the tolerance of bile salts allowing strains to survive in changing nutrients and conditions (108).

In some instances, bacteria evolve to encode a second *hns* either on conjugative plasmids or chromosomally. Historically, plasmid-encoded and chromosome-encoded H-NS proteins have been considered to function equivalently, with plasmid-encoded H-NS providing additional H-NS to compensate for the depletion of the primary chromosomal H-NS caused by silencing newly acquired AT-rich genes (106). Several conjugative plasmids (IncH1 group common in *Salmonella*) can encode plasmid forms of H-NS and Hha. Plasmid R27, detected in several *S.* Typhi outbreaks, encodes a copy of *hns* and *hha.* The N- and C- terminal domains between H-NS on R27 ($H\text{-}NS_{R27}$) and the primary H-NS on the chromosome had high sequence conservation; however, the linker domains were significantly different (106). $H\text{-}NS_{R27}$ selectively silences horizontally transferred genes but does not affect core genes. In order for the

chromosomal H-NS to regulate the horizontally transferred plasmid genes, it requires Hha (106). This plasmid encoded *hns* allows for a selective discrimination between core and horizontally transferred DNA. These properties were also seen in a second chromosomally encoded *hns* in EAEC strain 042 which mainly targeted horizontally transferred genes, that were also targeted by Hha (109). A role for these gene duplications in virulence is supported by the fact that they are present in a high percentage of pathogenic *E. coli* (110).

Alterations to the level of *hns* can have detrimental effects on the fitness of the bacteria and bacterial pathogenesis. The loss of H-NS leads to pleiotropic phenotypes due to the involvement of H-NS in the regulation of seemingly unlinked genes (90). In UPEC, H-NS regulates all of the major classical virulence factors (87). A high infectious dose of UPEC lacking *hns* resulted in an increased number of deaths in models for urinary tract infection and sepsis, likely due to increased toxicity (87). However, despite increased expression of virulence factors, a low infectious dose was attenuated due to growth deficits (87). Over-production of H-NS has been associated with cell death due to the immediate inhibition of RNA and protein synthesis, and morphologically cells over expressing *hns* exhibited very dense, compact, spherical nucleoids (111).

**Project rationale**

RNA-seq data and binding assays suggested that Aar can affect the expression of *aggR* and *hns,* and that it is able to bind to both proteins (31, 58).

The hypothesis for this project is that Aar binding to AggR and H-NS has functional significance in bacterial physiology and pathogenesis. We investigated if there were effects on downstream genes when *aar* and *aggR* were expressed together or when *aar* and *hns* were expressed together. In addition, we investigated the consequences of interactions between Aar and AggR or H-NS on pathogenesis and bacterial fitness. Furthering our understanding of how Aar is affecting bacterial gene expression will help us to understand its role in pathogenesis and may reveal potential interventions against the pathogen.

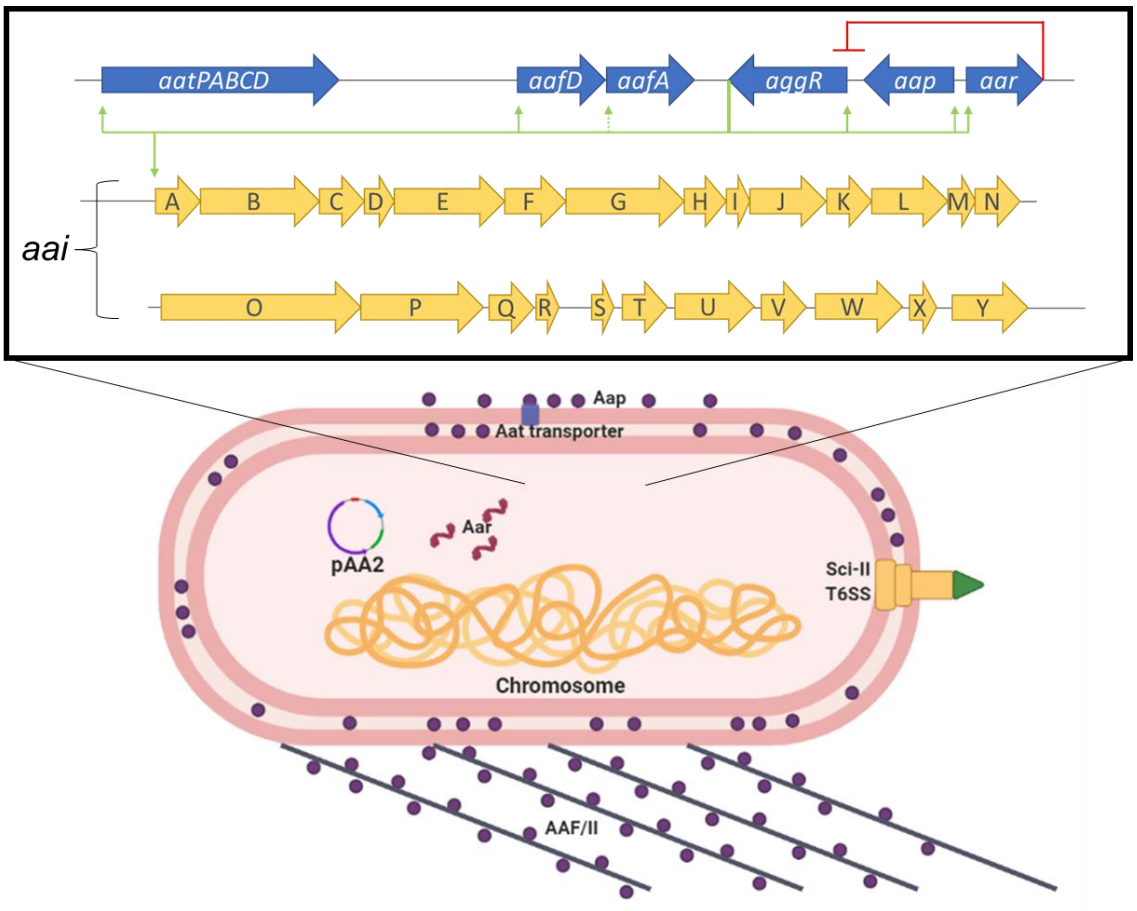



**Figure 1.1. Abbreviated AggR regulon in EAEC strain 042.** AggR is autoactivated and regulates approximately 44 chromosomal (*aai* operon encoding Type VI secretion system, Sci-II) and plasmid-borne (dispersin, *aap*; AAF/II genes, *aafDA; aat* transporter; negative regulator, *aar*) virulence genes.

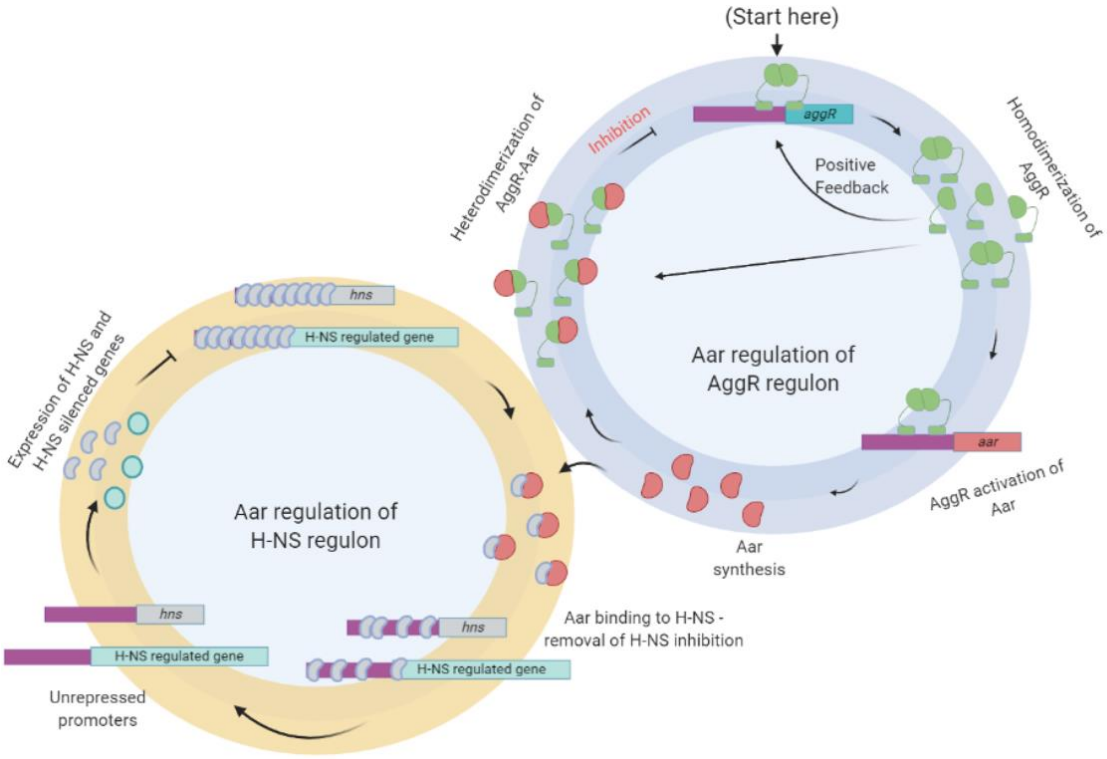



**Figure 1.2. Mechanism of Aar regulation of AggR and H-NS.** Aar acts as an anti-activator of the AggR regulon by binding directly to the dimerization domain of AggR, preventing AggR homodimerization, and inhibiting AggR activation of itself and other AggR-regulated genes (blue cycle). Aar can also regulate the global regulator H-NS by binding to the oligomerization domain of H-NS, which alters H-NS-DNA binding, resulting in the upregulation of H-NS and H-NS regulated genes (gold cycle).

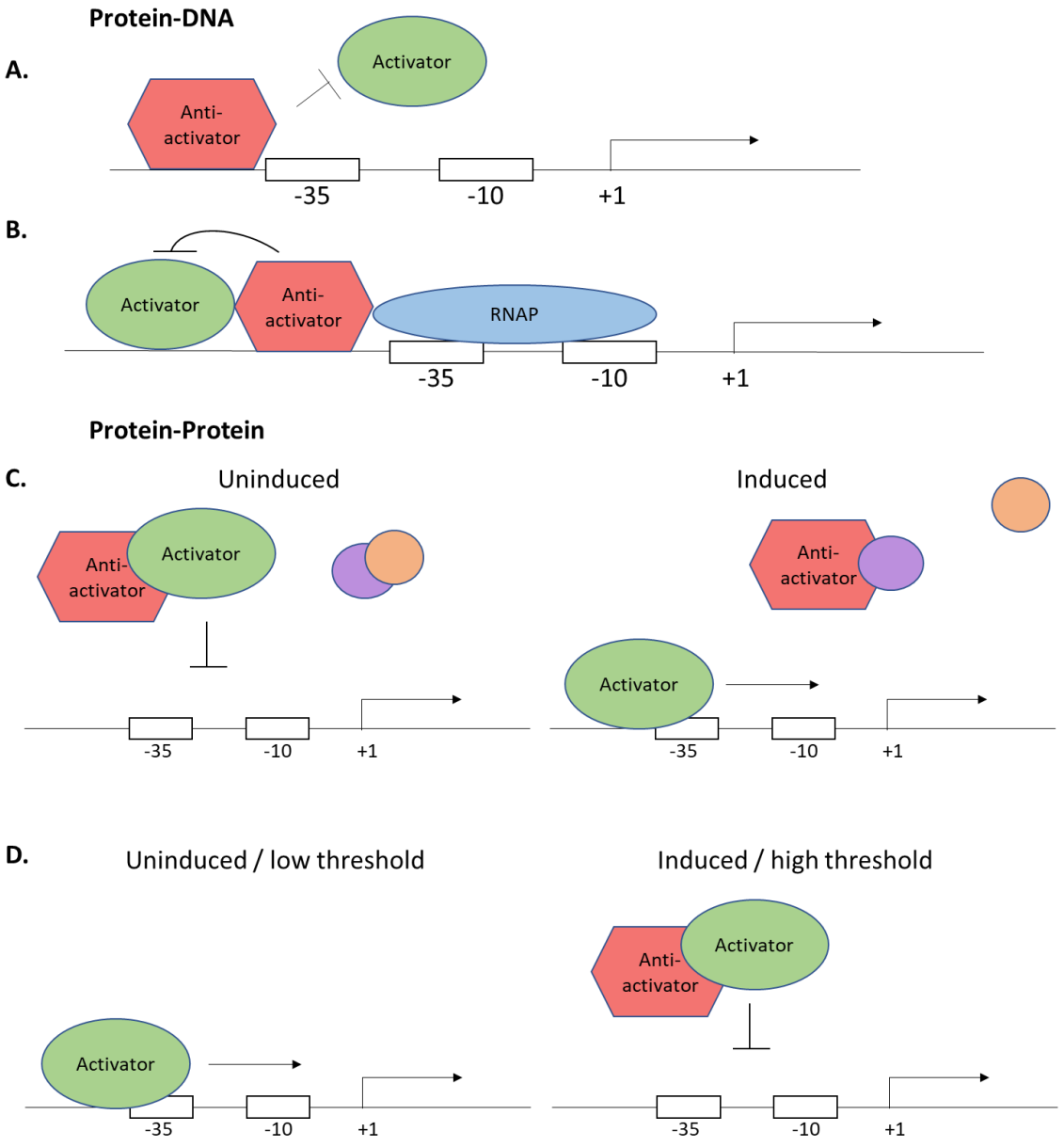



**Figure 1.3. Mechanisms of anti-activation.** Anti-activators (red hexagon) can function via protein-DNA binding by binding to the DNA sequence upstream of the target promoter thus blocking either (A) the binding by the activator (green oval) or (B) the interactions between the activator and RNA polymerase (blue oval). Anti-activators can also act via protein-protein binding where the anti-activator is either

(C) bound to the activator in an uninduced state and when induced change binding partners (purple circle) to free the activator or (D) binds to the activator only after an inducing signal or threshold is reached thus blocking the binding of the activator to the DNA. See text for references.

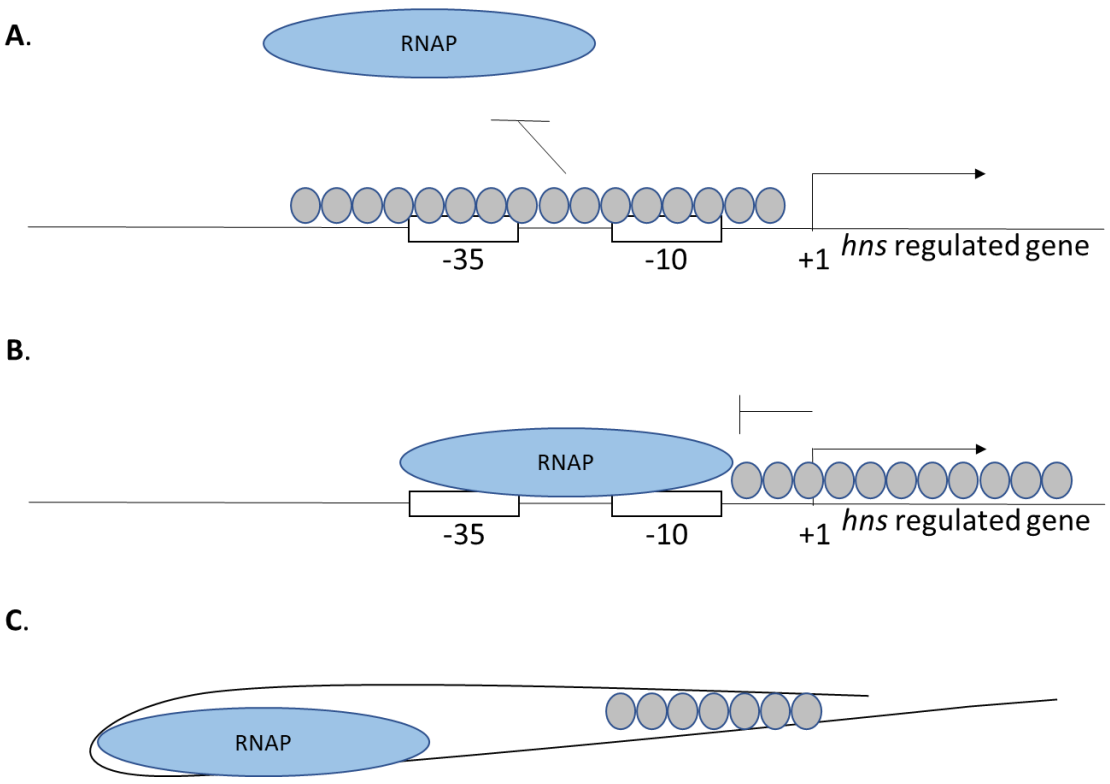



**Figure 1.4. Mechanisms of H-NS silencing.** H-NS (grey ovals) can regulate gene expression (A) by blocking RNAP (blue oval) from binding to the promoter region, (B) by blocking the RNAP elongation process, or (C) by forming a loop that interferes with promoter clearance. See text for references.

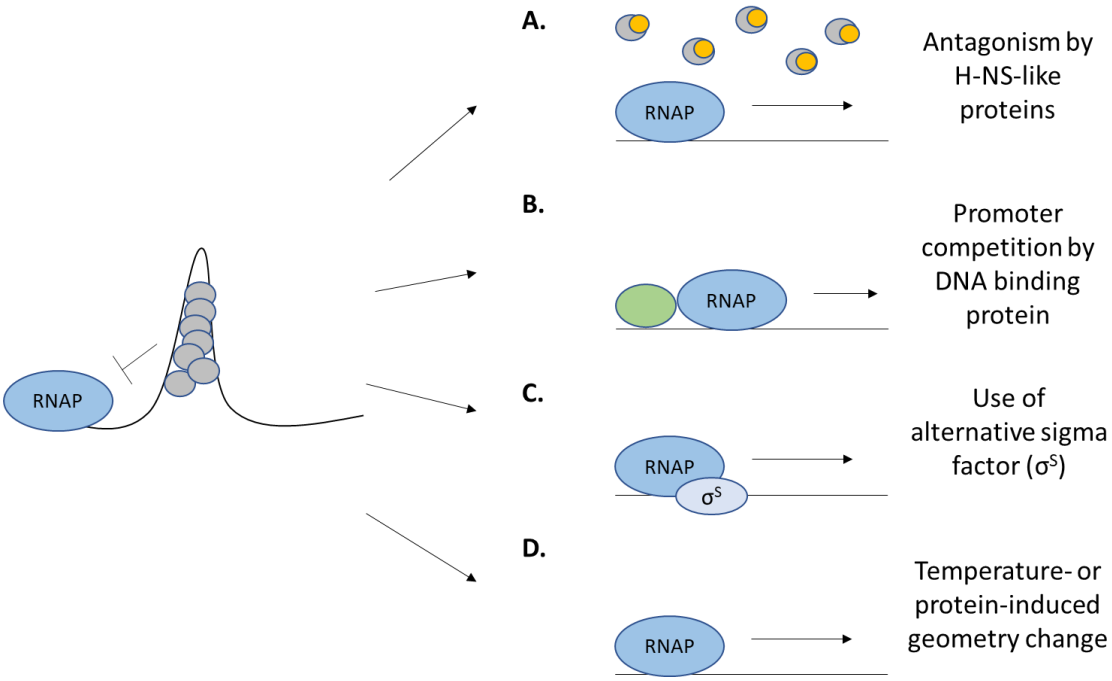



**Figure 1.5. Mechanisms of H-NS counter-silencing.** Counter-silencing can occur via (A) the disruption of H-NS complexes by multimerization antagonists, (B) competition for DNA-binding by high-affinity sequence-specific DNA-binding proteins, (C) activation of certain promoters by the alternative σ-factor RpoS, and/or (D) by changes in promoter geometry due to protein binding or environmental changes. See text for references.

# Chapter 2

## Dual function of Aar, a member of the new AraC negative regulator family, in *Escherichia coli* gene expression

Adapted from “Dual function of Aar, a member of the new AraC negative regulator family, in *Escherichia coli* gene expression”

Abigail S. Mickey and James P. Nataro

Infection and Immunity. 2020 May; 88(6).

**Abstract**

Enteroaggregative *Escherichia coli* (EAEC) is an *E. coli* pathotype associated with diarrhea and growth faltering. EAEC virulence gene expression is controlled by the auto-activated AraC family transcriptional regulator, AggR. AggR activates transcription of a large number of virulence genes, including Aar, which in turn acts as a negative regulator of AggR itself. Aar has also been shown to affect expression of *E. coli* housekeeping genes, including H-NS, a global regulator that acts at multiple promoters and silences AT rich genes (such as those in the AggR-regulon). Although Aar has been shown to bind both AggR and H-NS *in vitro*, functional significance of these interactions has not been shown *in vivo*. In order to dissect this regulatory network, we removed the complex inter-dependence of *aggR* and *aar* by placing the genes under the control of titratable promoters. We measured phenotypic and genotypic changes on downstream genes in EAEC strain 042 and *E. coli* K12 strain DH5α, which lacks the AggR regulon. In EAEC, we found that low expression of *aar* increases *aafA* fimbrial gene expression via H-NS; however, when *aar* is more highly expressed, it acts as a negative regulator via AggR. In DH5α, *aar* affected expression of *E. coli* genes in some cases via H-NS, and in some cases independent of H-NS. Our data supports the model that Aar interacts in concert with AggR, H-NS, and possibly other regulators, and that these interactions are likely to be functionally significant *in vivo*.

## Introduction

Enteroaggregative *Escherichia coli* (EAEC) is a common cause of traveler's diarrhea in industrial and developing countries and has been linked to growth failure in children (3, 4, 10, 112). Host colonization of EAEC is attributed to the presence of virulence genes which are controlled by AggR, a member of the AraC family of bacterial transcriptional regulators (27, 32, 33). A small protein named Aar (AggR Activated Regulator), whose expression is activated by AggR, has been described as a negative regulator of AggR (30, 32). Further characterization of Aar found that it belongs to a large family of proteins termed AraC Negative Regulators (ANR). The ANR family is found in hundreds of Gram-negative pathogens, and phylogenetically close homologs are able to complement function in ANR mutants (30).

In addition to regulating AggR expression, Aar has also been found to regulate genes encoding proteins outside of the AggR regulon, such as H-NS (58). H-NS is a global regulatory protein which usually acts as a repressor at a wide variety of promoters and genes that are AT rich and therefore intrinsically curved (79, 113). H-NS, AggR, and Aar have a complex dynamic. H-NS transcriptionally silences AraC transcriptional regulators; however, AraC transcriptional regulators may act as anti-repressors that counteract H-NS silencing in selected environments (21, 62). It has been hypothesized that regulation of AggR and H-NS by Aar is via Aar binding directly to either AggR or H-NS. Aar has been shown to bind both AggR and H-NS via surface plasmon resonance, the bacterial two-

hybrid system, pull down assays, and electrophoretic mobility shift assays (31, 58); however, functional significance of these interactions has not been elucidated.

We have previously postulated that Aar could be acting on AggR through direct formation of Aar/AggR complexes and/or through the formation of Aar/H-NS complexes, which could act to lift H-NS silencing of the regulon (58). The benefit to the bacterium of regulating virulence genes by two different interactions effected by one protein is unclear. In this study, we sought to better understand the mechanism by which Aar downregulates AggR-regulated genes and the functional significance (if any) of the hypothesized Aar/AggR and Aar/H-NS binding events.

## Materials and Methods

**Bacterial strains, plasmids, and growth conditions.** Bacterial strains and plasmids used in this study can be found in table 2.1. Bacteria were grown in Luria Broth (LB) and Dulbecco's modified Eagle's medium with 0.4% glucose (DMEM high glucose) (Gibco, Grand Island, NY) as previously described (32). When indicated, media were supplemented with carbenicillin (100μg/ml) and/or kanamycin (50μg/ml). For phenotypic titration studies, 0.01mM or 1mM IPTG and 0.01%, 0.05%, or 0.1% rhamnose was added as indicated below. For transcriptional studies, 5μM or 7.5μM IPTG and 0.00025%, 0.01%, or 0.1% rhamnose was added as indicated below. Inducer concentrations were selected

after a range of concentrations were screened to determine which had detectable effects on the expression of *aggR* or *aar*.

Mutagenesis of *aggR* and *hns* in 042Δ*aar* was accomplished by using lambda red technology (114). The loci (41,080-41,877 and 1,376,831-1,377,244; GenBank FN554767.1) in 042Δ*aar* were replaced with the kanamycin (km) resistance marker as previously reported (114). 042Δ*aar*Δ*aggR* and 042Δ*aar*Δ*aggR*Δ*hns* strains were identified by PCR using specific primers for *aggR*, *hns*, and a km resistance marker (Table 2.2). Deletion strains were cured of the km resistance using pCP20 as previously reported (114). Repair of *hns* in 042Δ*aar*Δ*aggR*Δ*hns* was done by using the lambda red recombination protocol to recombine a PCR product of *hns* with large flanking regions from 042 with the 042 Δ*aar*Δ*aggR*Δ*hns* km resistant strain and testing for recombination via km sensitivity. Primers for recombination of *hns* and screening for the repair are shown in Table 2.2.

Mutagenesis of *hns* in DH5α was accomplished by using lambda red technology and mutants were PCR screened and cured of km resistance as stated above.

**Titratable expression of *aggR* and *aar.*** For the independent expression of *aggR* and *aar,* plasmids pPlacZ-*aggR,* pPrham-*aar,* and pPlacZ-*aggR-D* were generated in this study (Table 2.1). Briefly, 1,086bp fragment containing the *lacZ* promoter region, the entire *aggR* gene, an HA-tag, a termination sequence, and flanked by restriction enzyme sites were synthesized by Genewiz Inc. by fragmentGENE

synthesis. The synthesized fragment was inserted into *Bam*HI and *Pst*I sites in pACYC177; the resulting plasmid was designated pPlacZ-*aggR*. pPlacz-*aggR*-*D* was similar to pPlacZ-*aggR* but only contains the dimerization site of *aggR,* comprising amino acid residues 69-181. pPrham-*aar* was generated similarly, but harboring a 548bp fragment containing the rhamnose promoter region, the entire *aar* gene, a 6Xhistidine tag, a termination sequence, and flanked by restriction enzyme sites. The synthesized fragment was inserted into the *Bam*HI and *Hind*III sites in pBR322.

**Biofilm production.** The biofilm assay previously described by Sheikh et al (2001) (26) was modified. Briefly, bacterial strains were grown in LB overnight at 37°C shaking. Overnight cultures of WT 042 and 042Δ*aafA* were diluted 1:20 in DMEM high glucose, and titration constructs were diluted 1:20 in LB with or without IPTG and rhamnose and inoculated into a 24 well polystyrene plate (Sigma-Aldrich). Bacteria were incubated for 3h at 37°C. After incubation plates were washed two times with PBS and fixed with 75% ethanol. The fixed biofilms were dried and stained with 0.5% Crystal Violet (Sigma). Biofilms were washed 4 times with PBS after staining and solubilized in 95% ethanol. The absorbance was determined at 570nm. Biofilms for 042Δ*aar*Δ*aggR*Δ*hns* were incubated for 5h at 37°C due to impaired growth.

**RNA extraction and qRT-PCR.** For quantitative reverse transcriptase PCR (qRT-PCR), EAEC strain 042, 042Δ*aar*Δ*aggR* and 042Δ*aar*Δ*aggR*Δ*hns* titration strains were grown aerobically in LB overnight at 37°C with shaking and then diluted 1:100

in DMEM high glucose or LB supplemented with IPTG and rhamnose concentrations as indicated and grown at 37°C. RNA from three biological replicates of each condition was extracted after 3h or 5h for 042Δ*aar*Δ*aggR*Δ*hns* titration strain. RNA was extracted using RNAprotect Bacteria Reagent (Qiagen) followed by a RNeasy Mini Kit (Qiagen). Primers used were previously published for EAEC (30, 58). qRT-PCR was performed using a one-step reaction in an ABI 7500-FAST sequence detection system (Applied Biosystems). All data were normalized to the levels of *rpoA* and analyzed using the comparative cycle threshold ($C_T$) method (115). The relative quantification method was used to determine the expression levels of target genes. Statistical significance was determined by ANOVA with post-hoc Tukey, and a *P* value of ≤0.05 was considered significant.

For qRT-PCR on *aar* and *aggR* in *E. coli* K12 strain DH5α was transformed with titratable *aar, aggR*, and their corresponding empty vector plasmids, and was grown aerobically in LB overnight at 37°C with shaking. A 1:100 dilution was made in LB with 0.1mM IPTG and 0.1% rhamnose and grown shaking at 37°C for 3 hours. RNA was extracted and qRT-PCR performed as above. Primers used were previously published for EAEC (30, 58).

## Results

**Independent expression of *aggR* and *aar* affects biofilm formation.** Although we have observed that Aar binds both AggR and H-NS (31, 58), the mutual

interdependence of these genes obfuscates the functional implications of these putative protein-protein interactions. Specifically, 1) H-NS has been shown to bind to AT-rich structural genes (79, 113), which include both AggR and Aar; 2) AggR is the activator of Aar gene expression (32); 3) Aar has been shown to repress AggR expression (30); and 4) transcriptomic data suggested that Aar may activate expression of the H-NS-encoding gene (58). Therefore, in order to better dissect the roles and contributions of these interdependent regulators in the control of gene expression in EAEC, we assembled systems in which expression of the genes could be controlled independently. Accordingly, we first constructed a derivative of EAEC strain 042 that harbored mutations in *aar* and *aggR*, and then introduced plasmids that carried the structural genes of *aar* and *aggR* under independently controllable promoters.

Plasmid pP*rham-aar* (designated here p*aar*) features the *aar* gene under the control of the rhamnose promoter; the plasmid is built on a pBR322 backbone (pMB1 replicon) and confers resistance to ampicillin. In preliminary experiments, we demonstrated that there were growth differences between LB and LB with rhamnose after 4h, likely due to rhamnose catabolism, so all experiments were performed at 3h post-induction unless stated otherwise (Figure 2.1). Similarly, plasmid pP*lacz-aggR* (designated here p*aggR*) features *aggR* under the *lacZ* promoter, and is built on a pACYC177 backbone (p15A replicon) that confers resistance to kanamycin.

In EAEC, AggR production induces the expression of *aafA,* leading to the formation of a bacterial biofilm (26, 32); thus biofilm formation is a ready phenotypic screen for *aggR* expression. Strain 042Δ*aar*Δ*aggR* did not produce an observable biofilm on a polystyrene substratum after 3 h of incubation at 37°C (data not shown). To assess the effect of p*aggR* expression in strain 042Δ*aar*Δ*aggR,* we subjected the strain to increasing concentrations of IPTG. As predicted, we observed a concentration-dependent increase in biofilm formation (figure 2.2A).

Expression of *aar* in 042Δ*aar*Δ*aggR* via introduction of plasmid p*aar* was affected by increasing concentrations of rhamnose; such a construct did not display expression of *aggR*. As predicted from our previous observations that *aafA* and resultant biofilm formation requires AggR (26, 32), we observed no change in biofilm formation in this construct under conditions of increasing rhamnose concentrations (figure 2.2B).

To confirm that changes seen in biofilm formation correlated with changes in the expression of *aggR, aar,* and *aafA*, qRT-PCR was performed for these gene transcripts. We discovered that inducer concentrations lower than that which produced observable biofilms were found to maximize mRNA transcript production by qRT-PCR; induction curves for qRT-PCR demonstrated that lower concentrations of the inducers were necessary to detect differences at the RNA level (Figure 2.3A and 2.3B). The combination of *aar* and *aggR* expression induced by the lower concentrations of rhamnose and IPTG (respectively), lead to measurable changes in *aafA* expression (Figure 2.3C). Gene expression of *aggR,*

*aar,* and *aafA* confirmed that biofilm formation parallels *aafA* gene expression (Figure 2.3A-C).

**Aar has a paradoxical effect on *aafA* expression**. We have previously observed that Aar serves as a negative regulator of the AggR regulon, and that the two proteins bind to each other *in vitro* (30, 31). As predicted from this model, we observed a concentration-dependent decrease in biofilm formation with increasing expression of *aar* (increasing concentrations of rhamnose) under conditions of constant *aggR* expression (figure 2.4A, colored bars with same fill pattern). qRT-PCR measurements of *aafA* transcription support that this decrease in biofilm formation was associated with a decrease in *aafA* expression (figure 2.4B, colored bars with same fill pattern) under constant *aggR* expression and increasing *aar* expression (figure 2.5A and B, colored bars with same fill pattern).

Unexpectedly given our model, we observed a paradoxical effect: in the presence of *aggR*, low levels of *aar* expression lead first to increased biofilm (figure 2.4A, dark grey bars to blue bars), followed by the expected concentration-dependent decrease in biofilm formation at higher *aar* concentrations. This effect was not due to the effects of the inducers themselves (figure 2.6A and B).

**The paradoxical effect of Aar requires H-NS**. RNAseq transcriptomic data suggested that Aar has an effect on expression of the histone-like protein H-NS and binding assays suggested that the two proteins physically interact (58). Like many members of the AraC family of transcriptional activators, AggR and the genes that it regulates (i.e. *aafA*) are repressed by H-NS (21, 62). Members of the

AraC family of transcriptional regulators are thought to counteract H-NS-induced silencing in select environments (21, 62). The paradoxical effect of Aar expression on AggR-dependent *aafA* expression suggested the action of another regulator, and we hypothesized that this regulator was H-NS.

To test this hypothesis, we constructed 042Δ*aar*Δ*aggR*Δ*hns*, transformed the strain with p*aar* and p*aggR*, and measured biofilm production with varying IPTG and rhamnose concentrations. 042Δ*aar*Δ*aggR*Δ*hns* had a growth defect compared to 042Δ*aar*Δ*aggR*; growth of 042Δ*aar*Δ*aggR*Δ*hns* at 5h produced an $OD_{600}$ similar to that observed in 042Δ*aar*Δ*aggR* at 3h (figure 2.7). Rhamnose catabolism had no effect on the growth phase of 042Δ*aar*Δ*aggR*Δ*hns* at 5h (figure 2.7). Similar to 042Δ*aar*Δ*aggR*, a low-level increase in *aar* led to a decrease in biofilm growth (figure 2.8A, colored bars with same fill pattern). However, in the absence of *hns,* the increase of biofilm formation at low levels of *aar* was no longer observed (figure 2.8A, dark grey bars to blue bars). The effect of low levels of Aar on biofilm production was rescued when *hns* was restored (figure 2.8B). qRT-PCR analysis confirmed that in 042Δ*aar*Δ*aggR*Δ*hns*, under conditions of constant *aggR* expression (figure 2.9A) but increasing *aar* expression (figure 2.9B), the direct concentration-dependent expression of *aafA* by *aggR* is maintained, as is the negative effect by *aar* (figure 2.8C); however, the paradoxical effect of *aar* on *aafA* is lost in the absence of H-NS (figure 2.8C). As seen in the biofilm assay (figure 2.8B), qRT-PCR supported that the restoration of *hns* rescued the paradoxical effect (figure 2.8D) under conditions of constant *aggR* expression (figure 2.9C) but

increasing *aar* expression (figure 2.9D). The paradoxical effect was only seen in the presence of AggR production; i.e. *aar* expression by itself did not affect *aafA* expression regardless of the presence of *hns.* Taken together, these data suggest a tripartite model of AggR/Aar/H-NS interaction, consistent with our *in vitro* observation that Aar binds to both AggR and H-NS.

***aggR* diminishes the *aar*-induced upregulation of non-AggR regulated genes in *E. coli* K12 via *aar*.** The results of our AggR and Aar controllable gene expression studies reveal concentration-dependent effects of the two regulators consistent with the previously published model of AggR/Aar protein-protein binding (31); i.e. that *aafA* expression may depend on the concentration of AggR unbound to Aar. If the mechanism is in fact due to protein-protein interaction, then binding of AggR to Aar might also reduce the activity of the latter protein. We sought to utilize a simplified system with which to probe hypothetical interference of Aar activity by AggR. Such an effect would add another regulatory dimension to the tripartite protein-protein interaction system.

We have previously observed in strain 042 that Aar activates housekeeping genes that are AggR-independent (58); if true, we would hypothesize that expression of such genes would similarly be affected in a K12 background. Interrogating this effect in K12 would eliminate the effect of AggR on other genes of the AggR regulon, which does not exist in K12.

We transformed *E. coli* DH5α separately with p*aar* and p*aggR* or their corresponding empty vector controls (pBR322 and pACYC177, respectively). To

interrogate a possible inhibitory effect of AggR on Aar, we chose to study two chromosomal *E. coli* core genes previously shown to be affected by *aar* in strain 042 (*orf1228* and *orf2223*) (58). As predicted, we observed that in K12 strain DH5α, when *aar* was expressed in p*aar*, the expression levels of *orf1228* and *orf2223* were increased (figure 2.10A). The expression of *aggR* alone in DH5α had no effect on the expression of either target gene in the absence of *aar* (figure 2.10A).

The expression of *aggR* simultaneously with *aar* caused only a small reduction to the observed increase in gene expression (figure 2.10A). Given that AggR is a DNA-binding protein, we repeated this experiment using an *aggR* construct comprising only the AggR dimerization domain (amino acids 69-181), therefore lacking the DNA binding helix-turn-helix C-terminal region (pP*lacZ-aggR-D* is referred to as p*aggR-D*) (31); we previously reported that the AggR dimerization domain binds Aar in the bacterial two-hybrid system (31). As seen with the full length *aggR*, the expression of the *aggR-D* in DH5α had no effect on the expression of *orf1228* and *orf2223* (figure 2.10B); when *aggR-D* and *aar* were both expressed in DH5α, the expression levels of the two queried genes were both significantly decreased compared to the level of expression when *aar* was expressed alone (figure 2.10B), suggesting that the expression of *aggR-D* in this system affects the expression of *orf1228* and *orf2223* via *aar,* consistent with protein-protein interaction of the two proteins. The expression levels of *aar* and

*aggR* were similar whether expressed alone or together (figure 2.11A); the same was true for *aar* and *aggR-D* (figure 2.11B).

***aar* upregulates gene expression in *E. coli* K12 via *hns.*** Several of the Aar-controlled genes affected in 042 that are AggR-independent are thought to be under H-NS control, based on previous literature (58, 90, 116, 117). We hypothesize that the effect of Aar on these genes is via the proposed model of Aar/H-NS protein-protein binding.

We used DH5α and DH5αΔ*hns* transformed with p*aar* to probe the expression of *ompX*, a gene known to be under the regulation of H-NS (90). The expression of *ompX* was increased in the absence of *hns*, confirming that *ompX* expression in DH5α is regulated by *hns* (figure 2.12A). In DH5α, we found that the expression of *ompX* was significantly increased when *aar* was induced (figure 2.12A). However, when *aar* was induced in DH5αΔ*hns*, there was no change in the expression of *ompX* (figure 2.12A), suggesting that *hns* is required for the effect of *aar* on *ompX.*

***hns* diminishes the *aar* induced upregulation of non-H-NS regulated genes in *E. coli* K12 via *aar.*** We reasoned that if Aar bound H-NS in the bacterium, the interactions of the two genes may be mutually interfering and the activity of Aar in the bacterium may be diminished by expression of *hns,* similarly to what was demonstrated with AggR (figure 2.10A-B). We therefore sought to use the DH5α system to probe hypothetical interference of Aar activity by H-NS.

We transformed *E. coli* DH5α∆*hns* separately with p*aar* and our previously published pKNTHNS (designated here p*hns*) (58) or their corresponding empty vector controls, pBR322 and pKNT25, respectively. We determined that *orf1228* was not affected by H-NS expression (figure 2.13); therefore, we chose to probe the expression of this gene to interrogate a possible inhibitory effect of H-NS on Aar activity. We performed the experiment in DH5α∆*hns* to remove any effects that the native *hns* could have on the system. When *aar* was expressed in DH5αΔ*hns*, the expression of *orf1228* was increased (figure 2.12B). The expression of *hns* alone had no effect on the expression of *orf1228* in the absence of *aar* (figure 2.12B). As suspected, when *hns* and *aar* are expressed together, the effect of *aar* on *orf1228* is significantly reduced (figure 2.12B). This suggests that the expression of *hns* is affecting the expression of *orf1228* via *aar,* consistent with the proposed model of a direct interaction of the two proteins.

**Discussion**

Although binding of Aar to AggR and H-NS has been previously demonstrated in artificial systems, we have not yet provided evidence that either of these binding phenomena have a functional role in the bacterium. In this work, we constructed a series of experimental systems to probe potential inter-relationships among Aar, AggR, and H-NS, regulators that are expected to have mutual inter-dependence. By using a system in which we remove the transcriptional inter-dependence of *aar* and *aggR*, our data suggest functional

roles for Aar binding to both AggR and H-NS in EAEC 042. In addition, employing *E. coli* DH5α, we observe evidence for Aar function in the absence of AggR, both via H-NS and potentially other regulators of the core *E. coli* genome.

As predicted, increasing *aar* expression lead to a concentration-dependent decrease in *aggR*-regulated *aafA* expression and biofilm formation. This relationship was inversely reciprocal: expression of *aafA* was increased by increasing *aggR* expression, but decreased by *aar; aafA* expression seemed to correlate best with excess *aggR* abundance over the level of *aar* expression. Previously published data failed to reveal evidence that Aar binds to DNA (31, 58), and because we removed the transcriptional dependence of *aggR* and *aar* on one another, our data best supports the model that Aar binds and sequesters AggR, and that *aafA* expression levels would be determined by free AggR protein concentrations.

In titration experiments, we were surprised to observe a phenotypic effect on biofilm formation with low expression of *aar* that was contrary to our hypothesis regarding how Aar would affect the expression of AggR-regulated genes. At the lowest levels of *aar* expression, we observed a paradoxical increase in *aafA* expression, independent of *aggR* expression. This effect was abrogated in an *hns* mutant. These data suggest a potential role for dual binding of both AggR and H-NS by Aar. It is tempting to speculate that Aar has a higher affinity for H-NS than for AggR, given that *aar* has a positive effect on *aafA* through *hns* first and then a negative effect via *aggR*. Such a nuanced effect could permit early expression of

*aafA in vivo* before the time required for cycles of *aggR* transcription and translation and subsequent binding to the *aafA* promoter. Given that there is evidence for EAEC infection of both the duodenal (48) and the colonic mucosae (19), this dual regulation could provide distinct pathogenic timing.

It has previously been shown that Aar acts upon H-NS-regulated promoters differentially (58). Due to the low expression of *aar* that is necessary to observe changes in *aafA* through *hns,* our data suggest that Aar may remove H-NS from the *aafA* gene (possibly the structural gene itself) thereby permitting AggR to upregulate expression. This affinity for removing H-NS may extend to other AggR-regulated promoters. It is possible that Aar could add specificity to the removal of H-NS at AggR-regulated genes over other H-NS regulated genes, thus allowing for a timed de-repression of those specific genes. The underlying mechanism of how Aar is leading to differential expression of various genes is unclear.

H-NS is a global regulator of *E. coli* gene expression (79, 113), and the putative binding of Aar to H-NS suggested that Aar may have global effects on EAEC gene expression beyond the AggR regulon. For this to prevail *in vivo*, one would expect effects of Aar on gene expression in *E. coli* K12, which is devoid of the AggR regulon; we not only observed such effects in a K12 system, but our data suggest still more global complexity accompanying *aar* expression. Our observations in a K12 system rule out the need for a pathogen-specific intermediary protein.

Although we posit that Aar acts via protein-protein interaction, demonstration of protein-protein binding is not definitive evidence that this phenomenon occurs *in vivo*. The use of multiple assays suggested AggR/Aar binding: surface plasmon resonance, bacterial two-hybrid system, and EMSA (31). The *in vivo* data presented here confirm inter-relationships among these regulatory proteins in ways that would be difficult to ascribe to alternate mechanisms. Importantly, the expression of *aggR* alone had no effect on *orf1228* and *orf2223*; therefore, the decrease in expression of these two genes in the presence of both *aar* and *aggR* suggests that AggR may be binding free Aar and preventing Aar from activating the genes. By demonstrating an *aar* effect on gene expression through *aggR* (in EAEC 042) and an *aggR* effect on gene expression through *aar* (in DH5α), our data supports the previously published model that the effect of *aar* and *aggR* is through protein-protein binding.

Targeting the *E. coli* gene *ompX,* previously reported to be under H-NS control (90), we confirmed that expression of *aar* induced expression of *ompX* in an *hns*-dependent manner. Surprisingly, however, our data suggest that the Aar effect on *orf1228* expression persists even in an *hns* mutant, suggesting that Aar may act in concert with still another regulator beyond AggR and H-NS. *E. coli* possesses additional histone-like proteins which may be responsible for the effect of *aar* on *orf1228* expression, and these are the targets of ongoing research in our laboratory. As we predicted, expression of the AggR dimerization domain (which does not occur naturally in K12 and does not bind DNA but has been shown to

bind Aar *in vitro*) demonstrated an Aar-inhibitory effect in a K12 background. These data strongly support the hypothetical model wherein AggR and Aar bind directly, thereby inhibiting activity of both proteins. The effects we observed show interdependence of AggR, Aar, and H-NS; however, the data do not prove that the mechanism is direct protein-protein interactions.

Based on our data and previous studies (30, 31, 58), we propose a model to illustrate a dual function of Aar in EAEC virulence gene expression (Figure 2.14). In abiotic environments, H-NS binds AT rich genes and silences their expression (i.e. *aggR* and *aggR* regulated genes). When the bacteria reach the host, temperature change and inducer molecules induce *aggR* expression. AggR upregulates *aar* expression, and at early stages of induction when *aar* expression is low but detectable, Aar binds and relieves H-NS silencing from AggR-regulated genes. This results in immediate upregulation of previously silenced genes by AggR. As the expression of *aar* increases, Aar begins to bind AggR in addition to H-NS, preventing AggR dimerization and therefore reducing activation of *aggR*-regulated genes.

The data presented in this paper support the model that Aar is binding to both AggR and H-NS and that both interactions have functional significance. In EAEC strain 042, Aar has a dual function in virulence gene expression: first, when present at low concentrations, Aar removes the inhibitory effect of H-NS on fimbrial gene expression; and then when the concentration of Aar increases, Aar acts as a negative regulator, turning off AggR-activated virulence genes. Our data suggest

that not only is Aar an anti-activator, but it can also act as an anti-repressor. The role of Aar on genes of the core *E. coli* genome is more difficult to decipher but could play a role in the switch from the non-pathogenic to the pathogenic lifestyle. Further research will address the concerted action of this complex regulatory circuitry.

**Table 2.1. Strains and plasmids used**

| Name | Characteristics | Source |
|---|---|---|
| | **Strains** | |
| **EAEC 042** | Enteroaggregative E. coli strain 042 | (48) |
| **042Δ*aar*Δ*aggR*** | 042 derivative carrying *aar* and *aggR* deletions | This study |
| **042Δ*aar*Δ*aggR*Δ*hns*** | 042 derivative carrying *aar*, *aggR* and *hns* deletions | This study |
| **042Δ*aar*Δ*aggR hns* repair** | 042 derivative carrying *aar* and *aggR* deletions with repaired *hns* | This study |
| **042Δ*aafA*** | 042 derivative carrying *aafA* deletion | (118) |
| **E. coli DH5α** | K12 strain | Lab collection |
| **DH5αΔ*hns*** | DH5α derivative carrying *hns* deletion | This Study |
| | | |
| | **Plasmids** | |
| **pBR322** | Cloning and expression vector ($Amp^R$, $Tet^R$) | Lab collection |
| **pACYC177** | Cloning and expression vector ($Amp^R$, $Km^R$) | Lab collection |
| **pKNT25** | BACTH vector for fusions to the C-terminus of fragment T25 | Lab collection |
| **pPrham-aar** | pBR322 derivative encoding Aar under the rhamnose promoter | This study |
| **pPlacZ-aggR** | pACYC177 derivative encoding AggR under the lacZ promoter | This study |

| | | |
|---|---|---|
| **pPlacZ-aggR-D** | pACYC177 derivative encoding AggR dimerization domain 69-181 under the lacZ promoter | This study |
| **pKNTHNS** | Plasmid encoding orf 1292 fused to the T25 fragment of CyaA | (58) |

**Table 2.2. Primers used**

| Name | DNA sequences (5->3) |
|---|---|
| **Primers employed for deletion of *aggR* (Region 41,080-41,877; GenBank FN554767.1)** | |
| **LRAggRFd** | TTTTGCCGTTACGCACCACTCCGTCAGTAGCTGAACAGG AGGGACAGCTGATAGAAACAGAAGCCACTGGAGCACCTC AAAAACACCATCATACACTAAATCAGTAAGTTGGCAGCAT CACCAACTTCAGCCATCTCAATATGTTTATAGCAATCTCA AATAATGATATGAAACATGTTTGTGTAGGCTGGAGCTGCT TC |
| **LRAggRrev** | AAGAATACGATAAATAATTTCCTATTGTAATTATAAGCGTA AAAATCATATCCCACATGACGATGTGGAAATTAACAAACG TATTTTATATGAGTTAAAAATATATCTTTTTATTGATAAGAG TTAGGTCATTCTAACGCAGATTGCCTGATAAAGACATTTT TTTCATGTGAGAATGATATGGGAATTAGCCATGGTCC |
| **Primers employed for deletion of *hns* (Region 1,376,831-1,377,244; GenBank FN554767.1)** | |
| **hns lambda forward** | CGGCGCAAATAGGGCTATATGCCGCGTCTTTTCTGGCTA ATTTTATGAAAAGATATTTATTGGCGGCACAAAATAAAGAA CAATTTTGAATTCCTTACATTCCTGGCTATTGCACAACTGA ATTTAAGGCTCTATTATTACCTCAACAAACCACCCCAATAT AAGTTTGAGATTACTACAGTGTAGGCTGGAGCTGCTTC |
| **hns lambda reverse** | AAGTAACATCCGTATCGGTGTTATCCACGAAACGGCGTT GAGTAATCGACGCCGTTTTTTTATAGCTTATTCTTATTAAA TTGTCTTAAACCGGACAATAAAAAATCCCGCCGATGGCG GGATTTTTAAGCAAGTGCAATCTACAAAAGATTATTGCTT GATCAGGAAATCGTCGAGGGAATGGGAATTAGCCATGGT CC |
| **Primers employed for screening of 042Δ*aar*Δ*aggR* and 042Δ*aar*Δ*aggR*Δ*hns* and DH5αΔ*hns*** | |
| **AggRFDLR** | TTCAGCCATCTCAATATGTTTATAGCA |
| **AggRrevLR** | TGGACTGTTGCGATCGTGAAGCC |
| **AggRFdLR1** | TACCGGGTTGAGAAGCGGTGTAA |

| | |
|---|---|
| **KanrevLR1** | TTGTCCAGATAGCCCAGTAGCTG |
| **Hns sense** | ATGAGCGAAGCACTTAAAATTCTGAACAACATCCGTACTC TTCGTGCGC |
| **Hns rev** | TTATTGCTTGATCAGGAAATCGTCGAGGGATTTACC |
| **H-NS (upstream) sense** | ATCCTTCTGAGCTATCATTACAACTGCC |
| **Primers employed for recombining *hns* in 042Δ*aar*Δ*aggR*Δ*hns* and screening of 042Δ*aar*Δ*aggR hns* repair** | |
| **HnsFd** | CCCTTACGAAGCCTTGCATAATCCTTCTGAG |
| **HnsRv** | GGTGAAAGCGTACCGATGGTTGGC |

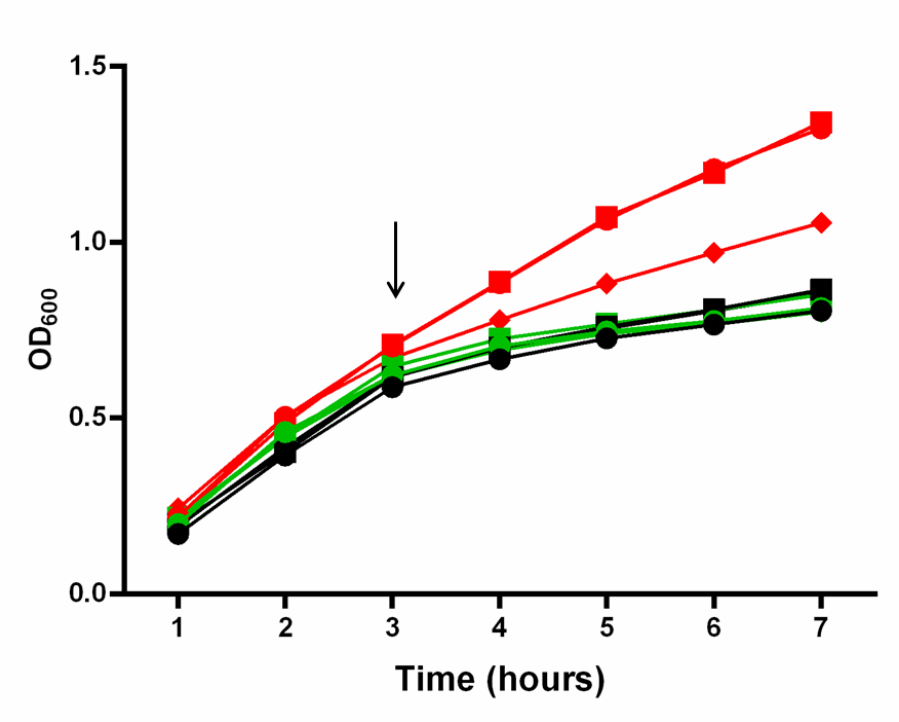



**Figure 2.1. Growth curves in presence and absence of inducer molecules.** Growth curves were measured for WT 042 (circles), 042Δ*aar*Δ*aggR* (squares), and 042Δ*aar*Δ*aggR*(p*aar*)(p*aggR*) (diamonds) in LB (black), LB+1mM IPTG (green), or LB+1% rhamnose (red). Growth curve data are representative of at least three independent experiments.

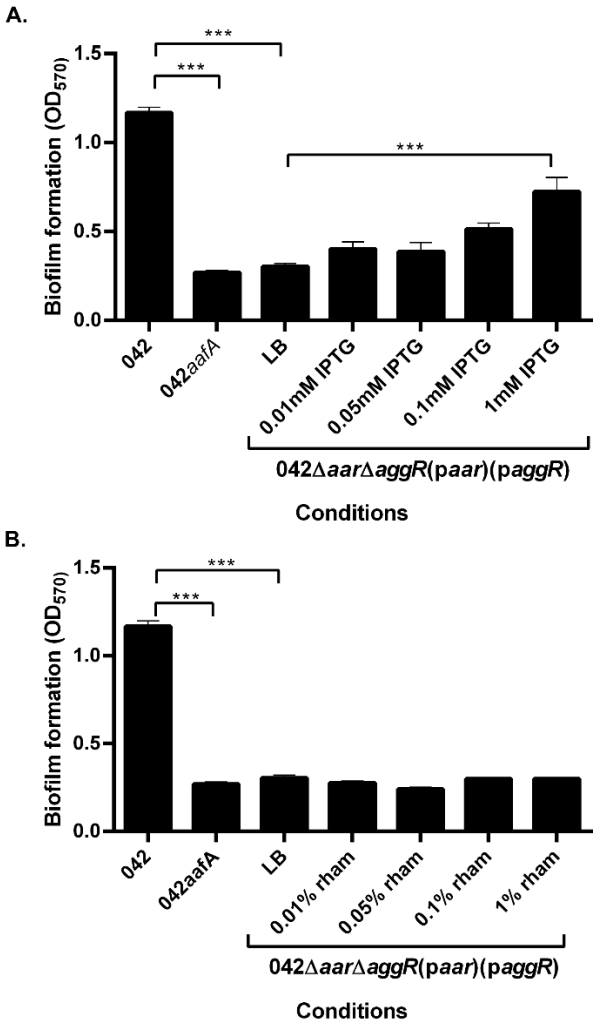



**Figure 2.2. Biofilm formation in the presence and absence of inducer molecules.** (A) Biofilm formation was measured using crystal violet staining after 3h in 042 and 042Δ*aafA* in DMEM high glucose and in 042Δ*aar*Δ*aggR*(p*aar*)(p*aggR*) in LB with varying concentrations of IPTG. (B) Biofilm formation was measured using crystal violet staining after 3h in 042 and 042Δ*aafA* in DMEM high glucose and in 042Δ*aar*Δ*aggR*(p*aar*)(p*aggR*) in LB with varying concentrations of rhamnose. Biofilm data are representative of at least

three independent experiments. Asterisks indicate significant differences by ANOVA (*, P<0.05; **, P<0.005; ***, P<0.0005).

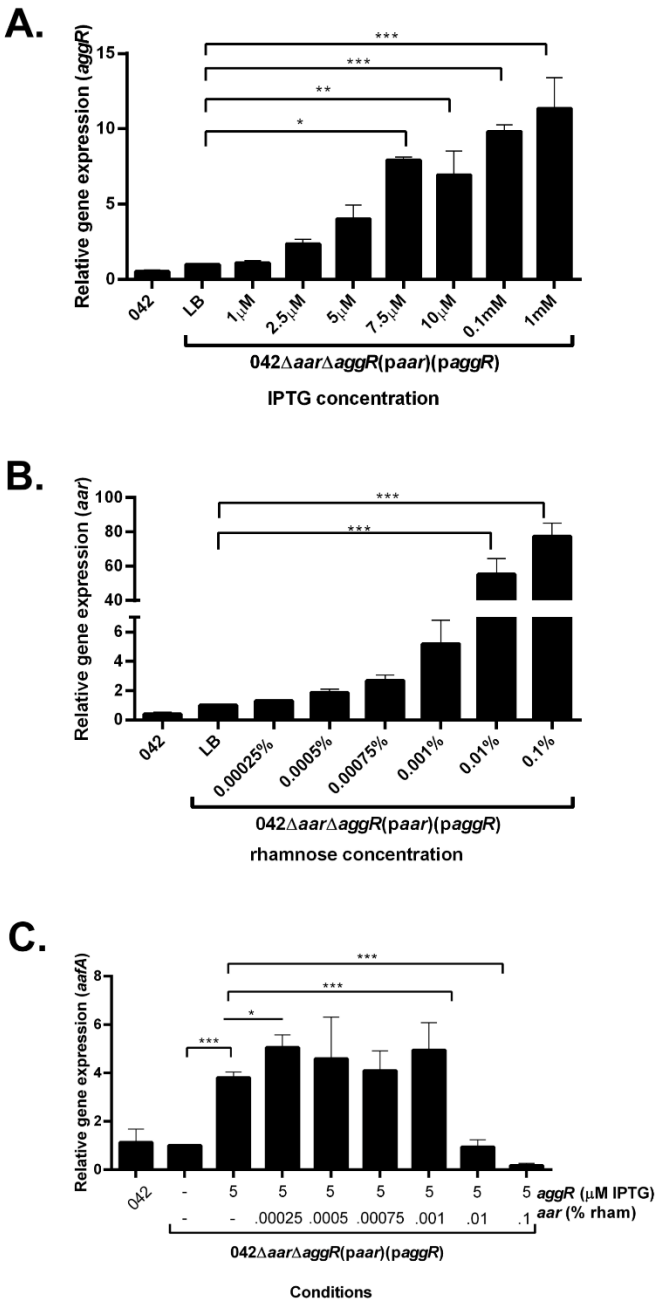



**Figure 2.3. Gene expression after induction of *aggR* by IPTG or *aar* by rhamnose.** (A) 3hr *aggR* expression measured using qRT-PCR on 042 in DMEM high glucose and 042Δ*aar*Δ*aggR*(p*aar*)(p*aggR*) in LB with varying concentrations of IPTG. (B) 3hr *aar* expression measured using qRT-PCR on 042 in DMEM high

glucose and 042Δ*aar*Δ*aggR*(p*aar*)(p*aggR*) in LB with varying concentrations of rhamnose. (C) 3hr *aafA* expression measured using qRT-PCR on 042 in DMEM high glucose and 042Δ*aar*Δ*aggR*(p*aar*)(p*aggR*) in LB with 0.005mM IPTG and a range of concentrations of rhamnose. qRT-PCR data are representative of at least three independent experiments. Asterisks indicate significant differences by ANOVA (*, P<0.05; **, P<0.005; ***, P<0.0005). The difference between *aggR* (5μM) and *aggR* (5μM) with *aar* (0.00025%) was found to be significant with a two-tailed paired t-test.

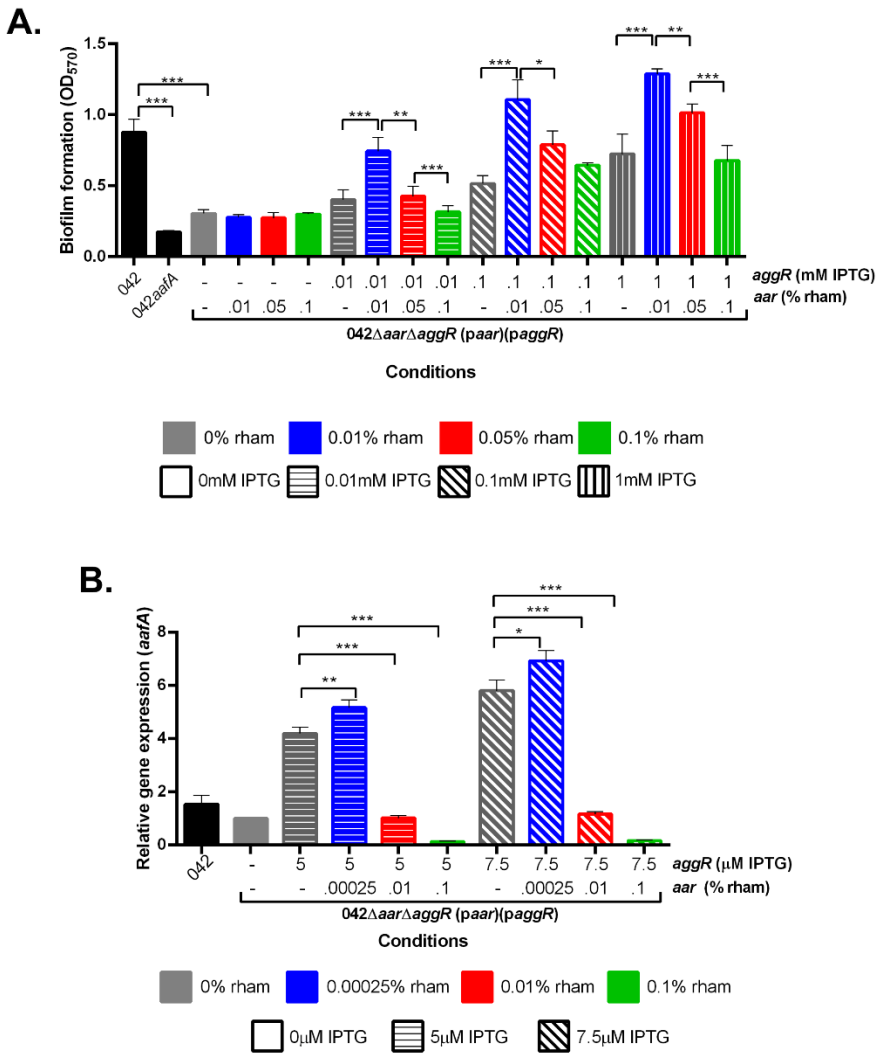



**Figure 2.4. Biofilm formation and gene expression of *aafA* in 042*ΔaarΔaggR* titrated with *aar* and *aggR*.** (A) Biofilm growth at 3 hours post induction with increasing concentration of IPTG and rhamnose. *aggR* expression was induced with 0.01mM IPTG (horizontal fill pattern), 0.1mM IPTG (diagonal fill pattern), or 1mM IPTG (vertical fill pattern). *aar* expression was induced with 0.01% rham (blue), 0.05% rham (red), or 0.1% rham (green). (B) qRT-PCR analysis of *aafA*

using titratable *aar* and *aggR*. *aggR* expression was induced with either 5µM IPTG (horizontal fill pattern) or 7.5µM IPTG (diagonal fill pattern). *aar* expression was induced with 0.00025% rham (blue), 0.01% rham (red), or 0.1% rham (green). Biofilm data and qRT-PCR data are representative of at least three independent experiments. Asterisks indicate significant differences by ANOVA (*, $P<0.05$; **, $P<0.005$; ***, $P<0.0005$).

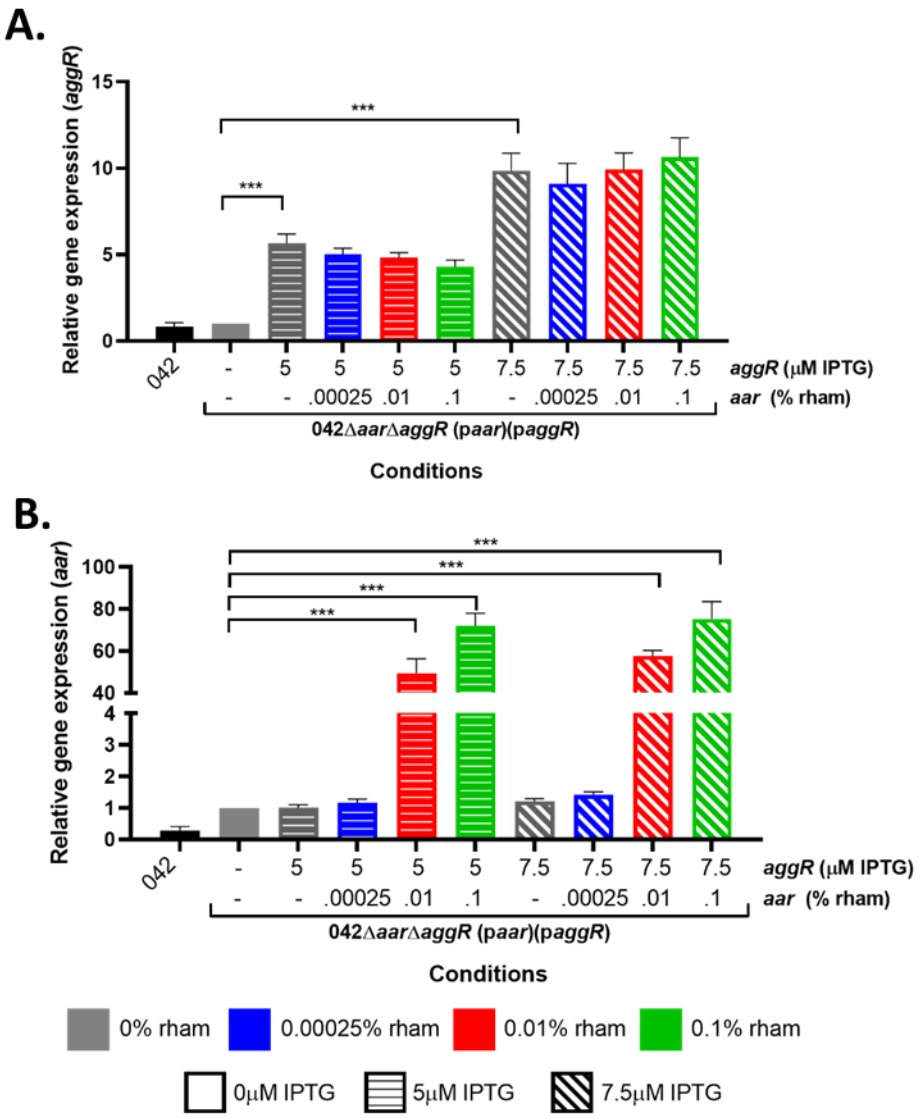



**Figure 2.5. Titration expression of *aar* and *aggR* in 042*ΔaarΔaggR*.** (A) qRT-PCR analysis of *aggR* and (B) *aar* using titratable *aar* and *aggR. aggR* expression was induced with either 5μM IPTG (horizontal fill pattern) or 7.5μM IPTG (diagonal fill pattern). *aar* expression was induced with 0.00025% rham (blue), 0.01% rham (red), or 0.1% rham (green). Biofilm data and qRT-PCR data are representative of

at least three independent experiments. Asterisks indicate significant differences by ANOVA (*, P<0.05; **, P<0.005; ***, P<0.0005).

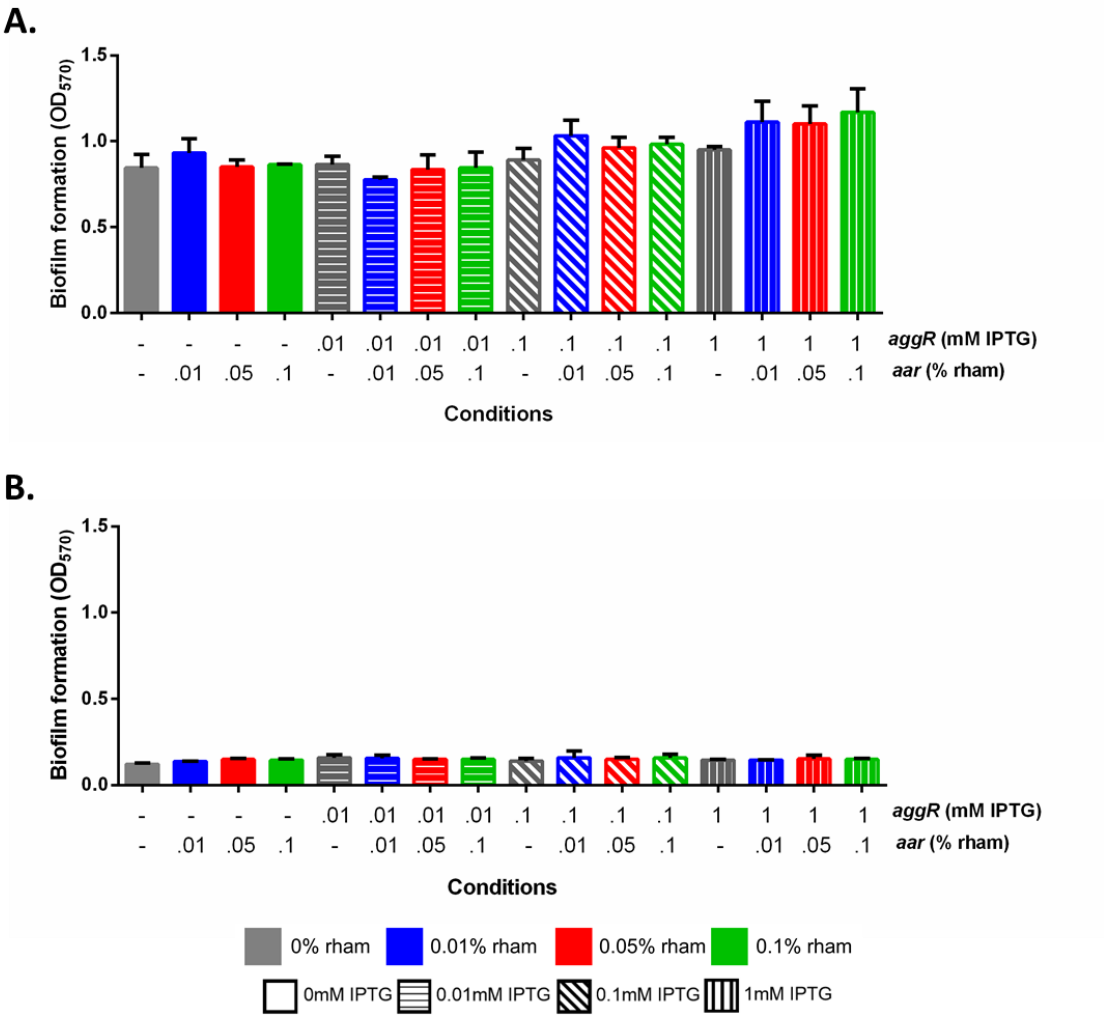



**Figure 2.6. Inducer effects in titration constructs.** (A) Biofilm growth of 042*ΔaarΔaggR*(p*aggR*) at 3 hours post induction with increasing concentration of IPTG and rhamnose. (B) Biofilm growth of 042*ΔaarΔaggR*(p*aar*) at 3 hours post induction with increasing concentration of IPTG and rhamnose. *aggR* expression was induced with 0.01mM IPTG (horizontal fill pattern), 0.1mM IPTG (diagonal fill pattern), or 1mM IPTG (vertical fill pattern). *aar* expression was induced with 0.01% rham (blue), 0.05% rham (red), or 0.1% rham (green). Biofilm data are representative of at least three independent experiments.

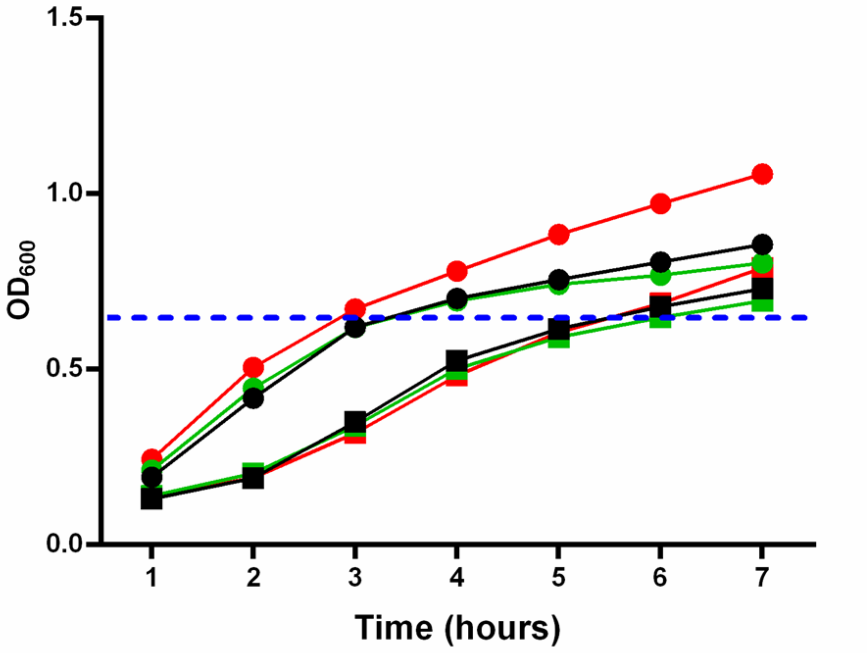



**Figure 2.7. Growth curve of 042Δ*aar*Δ*aggR* and 042Δ*aar*Δ*aggR*Δ*hns* in different conditions.** Growth curves were measured for 042Δ*aar*Δ*aggR*(p*aar*)(p*aggR*) (circles) and 042Δ*aar*Δ*aggR*Δ*hns*(p*aar*)(p*aggR*) (squares) in LB (black), LB+1mM IPTG (green), or LB+1% rhamnose (red). Growth curve data are representative of at least three independent experiments.

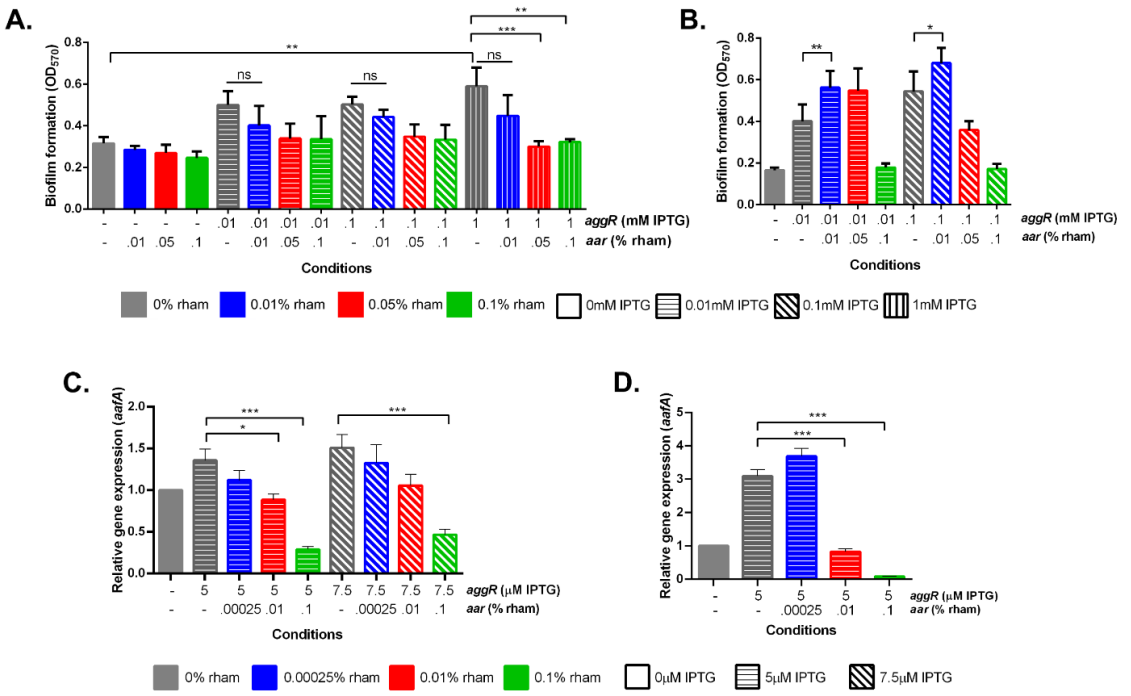



**Figure 2.8. Biofilm formation and expression of *aafA* in 042Δ*aar*Δ*aggR*Δ*hns* and *hns* repair titrated with *aar* and *aggR*.** (A) Biofilm growth in 042Δ*aar*Δ*aggR*Δ*hns* at 5 hours post induction with increasing concentration of IPTG and rhamnose. (B) Biofilm growth in *hns* repair at 3 hours post induction with increasing concentration of IPTG and rhamnose. *aggR* expression was induced with 0.01mM IPTG (horizontal fill pattern), 0.1mM IPTG (diagonal fill pattern), or 1mM IPTG (vertical fill pattern). *aar* expression was induced with 0.01% rham (blue), 0.05% rham (red), or 0.1% rham (green). (C) qRT-PCR analysis of *aafA* using titratable *aar* and *aggR* in 042Δ*aar*Δ*aggR*Δ*hns* after 5h. (D) qRT-PCR analysis of *aafA* using titratable *aar* and *aggR* in the *hns* repaired 042Δ*aar*Δ*aggR* after 3h. *aggR* expression was induced with either 5µM IPTG (horizontal fill pattern) or 7.5µM IPTG (diagonal fill pattern). *aar* expression was induced with

0.00025% rham (blue), 0.01% rham (red), or 0.1% rham (green). Biofilm data and qRT-PCR data are representative of at least three independent experiments. Asterisks indicate significant differences by ANOVA (*, $P<0.05$; **, $P<0.005$; ***, $P<0.0005$).

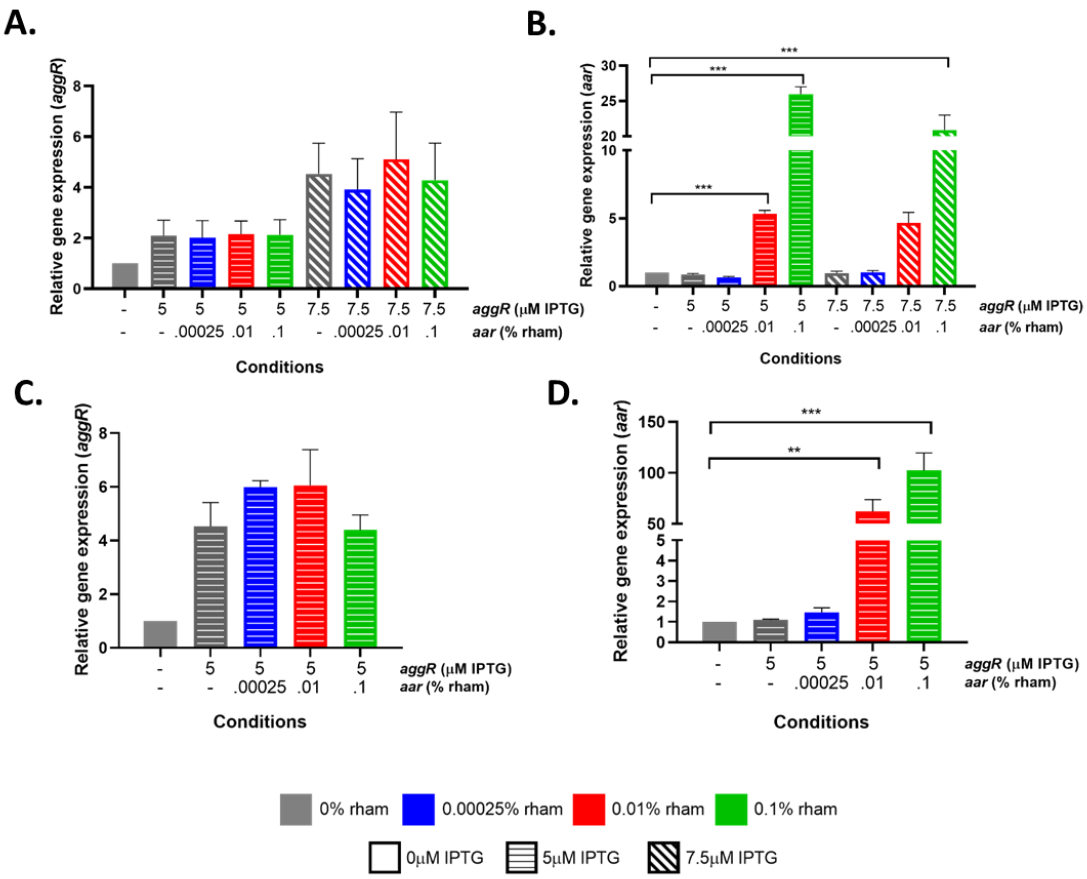



**Figure 2.9. Titration expression of *aar* and *aggR* in 042Δ*aar*Δ*aggR*Δ*hns* and *hns* repair.** (A) qRT-PCR analysis of *aggR* and (B) *aar* using titratable *aar* and *aggR* in 042Δ*aar*Δ*aggR*Δ*hns* after 5h. (C) qRT-PCR analysis of *aggR* and (D) *aar* using titratable *aar* and *aggR* in the *hns* repaired 042Δ*aar*Δ*aggR* after 3h. *aggR* expression was induced with either 5μM IPTG (horizontal fill pattern) or 7.5μM IPTG (diagonal fill pattern). *aar* expression was induced with 0.00025% rham (blue), 0.01% rham (red), or 0.1% rham (green). qRT-PCR data are representative of at least three independent experiments. Asterisks indicate significant differences by ANOVA (*, P<0.05; **, P<0.005; ***, P<0.0005).

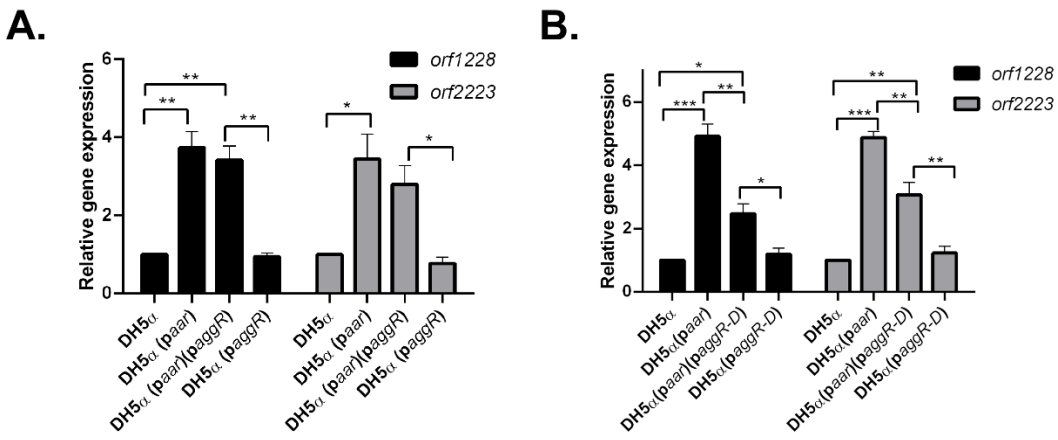



**Figure 2.10. The effect of *aar* and *aggR* on gene expression in DH5α transformed with p*aar* and/or p*aggR*/p*aggR-D*.** (A) DH5α was transformed with p*aar* and p*aggR* expressing full length *aggR* or their corresponding empty vectors pBR322 and pACYC177, respectively. Transcriptional levels of *E. coli* chromosomal genes *orf1228* and *orf2223* were analyzed by qRT-PCR. (B) DH5α was transformed with p*aar* and p*aggR-D* expressing the AggR dimerization domain or their corresponding empty vectors. Transcriptional levels of *E. coli* chromosomal genes *orf1228* and *orf2223* were analyzed by qRT-PCR. RT-PCR data are representative of at least three independent experiments. Asterisks indicate significant differences by ANOVA (*, P<0.05; **, P<0.005; ***, P<0.0005).

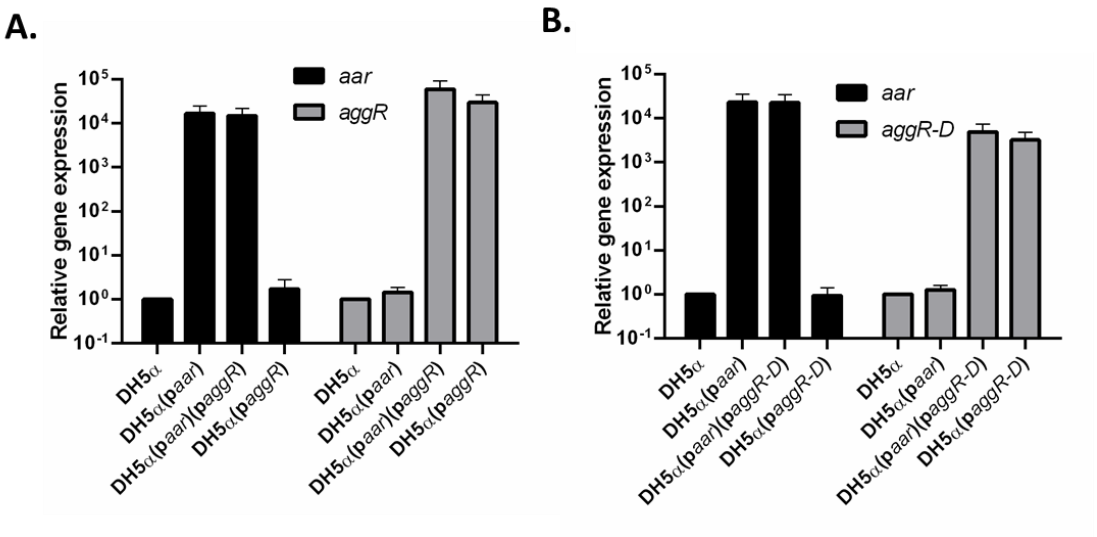



**Figure 2.11. Expression levels of *aar*, *aggR*, and *aggR-D* in DH5α transformed with p*aar* and/or p*aggR*/p*aggR-D*.** (A) DH5α was transformed with p*aar* and p*aggR* expressing full length *aggR* or their corresponding empty vectors pBR322 and pACYC177, respectively. Transcriptional levels of *aar* and *aggR* were analyzed by qRT-PCR. (B) DH5α was transformed with p*aar* and p*aggR-D* expressing the AggR dimerization domain or their corresponding empty vectors. Transcriptional levels of *aar* and *aggR-D* were analyzed by qRT-PCR. RT-PCR data are representative of at least three independent experiments.

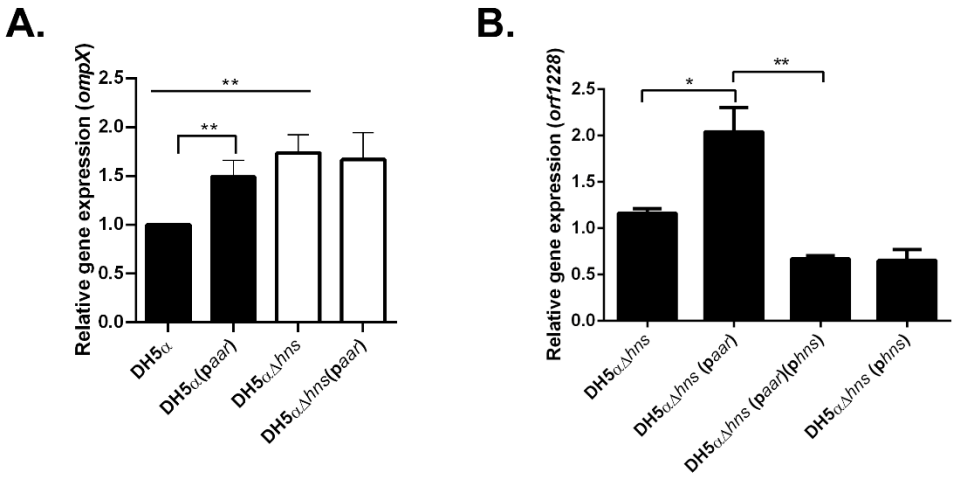



**Figure 2.12. The effect of *aar* on gene expression in the presence or absence of *hns* in DH5α.** (A) DH5α and DH5αΔ*hns* were transformed with p*aar* or its corresponding empty vector pBR322. Transcriptional levels of *ompX* were analyzed by qRT-PCR. (B) DH5αΔ*hns* was transformed with p*aar* and p*hns* or their corresponding empty vectors pBR322 and pKNT25 respectively. Transcriptional levels of *orf1228* was analyzed by qRT-PCR. RT-PCR data are representative of at least three independent experiments. Asterisks indicate significant differences by ANOVA (*, P<0.05; **, P<0.005; ***, P<0.0005).

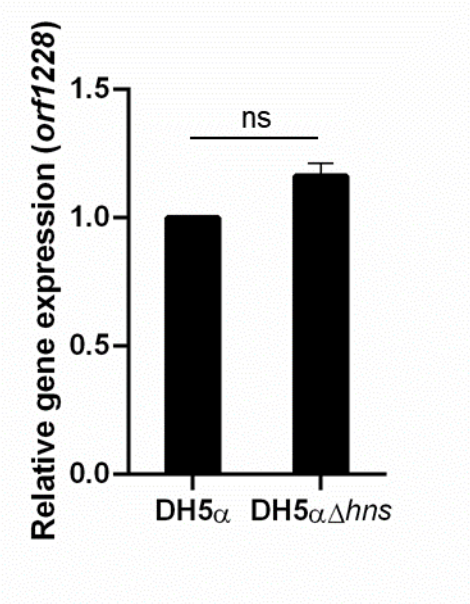



**Figure 2.13. Expression of *orf1228* in DH5α and DH5αΔ*hns*.** Transcriptional levels of *orf1228* in DH5α and DH5αΔ*hns* were analyzed by qRT-PCR. RT-PCR data are representative of at least three independent experiments.

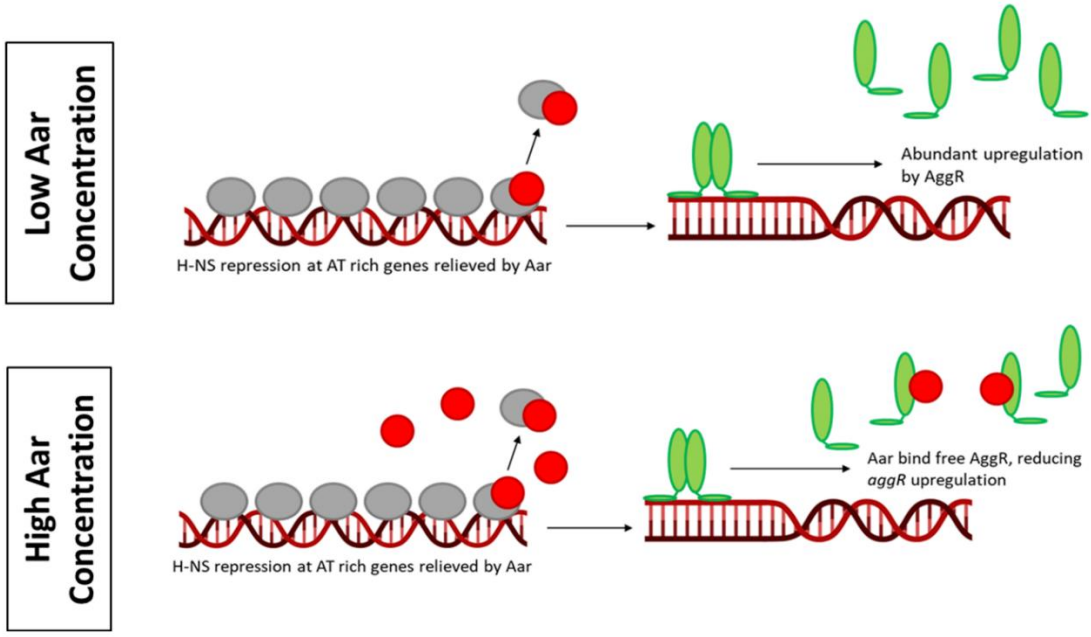



**Figure 2.14. Proposed mechanism of AggR-Aar-Hns interaction *in vivo*.** When the concentration of Aar (red circles) is low, Aar removes H-NS (grey ovals) repression at AT-rich genes). This allows AggR (green ovals) to abundantly upregulate gene expression. When the concentration of Aar is high, Aar removes H-NS repression but also binds to AggR. AggR is still able to upregulate gene expression but not as abundantly.

# Chapter Three

## The presence of Aar, a negative regulator in EAEC, affects relative colonization in pathogenic and nonpathogenic *E. coli*

**Abstract**

Enteroaggregative *Escherichia coli* (EAEC) is an *E. coli* pathotype that is associated with growth impairment in children. Multiple putative virulence genes in EAEC are regulated by the transcriptional regulator AggR including at least 44 genes and its own negative regulator, Aar. The presence of *aar* in EAEC strains has been associated with statistical protection from severe disease outcomes. Recent *in vitro* experiments have found that Aar has a role independent from AggR and can regulate *E. coli* core genes. These core genes include global regulators such as H-NS and Hha, which can affect bacterial fitness and survival. We utilized an *in vivo* competition assay to determine how Aar is affecting colonization and bacterial fitness. Expression of *aar* decreased competitiveness in EAEC strain 042 partially through the repression of *aggR*. A specific gene in the AggR regulon that could be responsible for colonization in mice was not apparent. We studied the effect of *aar* on core genes in K12 strain DH5α to factor out effects of AggR and its regulon. In DH5α, we demonstrated that *aar* reduces *in vivo* competitiveness, and that this is likely due to *aar* modulating *hns* expression. Our data from *in vivo* competition assays suggest that Aar is affecting colonization and bacterial fitness in EAEC through both the AggR regulon and core genes.

## Introduction

There is growing global concern about enteroaggregative *Escherichia coli* (EAEC) infection and its association with growth impairment in children (8). Epidemiological studies are frequently used to try to elucidate what factors are associated with pathogenesis and how these factors are causing disease. EAEC colonization can result in either diarrhea, which is dependent upon host susceptibility and the bacterial complement of virulence genes (8, 11, 119), or subclinical infection, which is nonetheless associated with reduced height for children at 2 years of age (8). In both types of infections, EAEC is associated with malnutrition (8).

Colonization is mediated by the expression of virulence genes that are controlled by AggR, a member of the AraC family of transcriptional regulators. AggR regulates approximately 44 chromosomal and plasmid genes including the aggregative adherence fimbriae (AAF) (33), dispersin (*aap*), components of a type VI secretion system (*aaiA*-*aaiY*) (27), and the negative regulator of AggR (*aar*) (30, 32). Aar belongs to a large family of proteins termed AraC Negative Regulators (ANR) which are found in hundreds of Gram-negative pathogens (30).

Epidemiological studies have described a protective role for *aar* in EAEC pathogenesis (11). Not only is *aar* associated with non-diarrheagenic strains (7, 119), but strains lacking *aar* are associated with malnutrition (11). This protection could be due to the role Aar has on downregulating AggR-regulated virulence

genes. In fact, AAF/II- (one of the five AAF variants) encoding genes are frequently associated with diarrhea and malnutrition (2, 7, 11).

These epidemiological data are supported by findings in an *in vivo* mouse model which looked at the virulence of the *aar* homolog *orf02851* in the mouse pathogen *Citrobacter rodentium.* It was observed that strains lacking *orf02851* showed sustained high levels of bacteria in feces and high levels of colonization, compared to wildtype strains (30). In addition, mice given *C. rodentium* lacking *orf02851* experienced more weight loss and transmission electron microscopy of colonic tissue revealed more tissue damage compared to mice given the wildtype strain (30).

Recent characterization of Aar revealed that Aar has a role outside of regulating *aggR* expression: *aar* affects *E. coli* core genes such as *hns* in an AggR-independent way (58). H-NS is a global regulator that often acts to silence intrinsically curved, AT-rich genes and promoters (79, 113) including AraC transcriptional regulators like AggR (21, 62). In *E. coli,* H-NS has been suggested to play an essential role in the adaptability of bacterial cells by regulating genes involved in accomodating environmental changes such as temperature, pH, and osmolarity (87, 90, 92). Aar has been shown to affect *hns* expression *in vivo* and is required for maximal *hns* expression (58).

The role of Aar in EAEC pathogenesis and fitness has been of great interest to our lab since its discovery. In this chapter, we sought to understand how *aar* is

affecting colonization and whether this effect is through the influence of Aar on the AggR regulon and/or on the genes in the core genome, such as *hns.*

## Method and Materials

**Bacterial strains, plasmids, and growth conditions.** Bacterial strains and plasmids used in this study can be found in table 3.1. Bacteria were grown in Luria Broth (LB) and Dulbecco's modified Eagle's medium with 0.4% glucose (DMEM high glucose) (Gibco, Grand Island, NY) as previously described (32). When indicated, media was supplemented with streptomycin (100µg/ml), carbenicillin (100µg/ml), kanamycin (50µg/ml), and/or rifampicin (100µg/ml).

Streptomycin-resistant derivatives of *E. coli* K12 strain DH5α and DH5α∆*hns* were generated by selection on Luria Broth (LB) agar plates with 100µg/ml streptomycin. Rifampicin-resistant derivatives of *E. coli* K12 strain DH5α was generated by selection on Luria Broth (LB) agar plates with 100µg/ml rifampicin.

The cloning vector pACYC177 was made ampicillin sensitive by digesting the plasmid with restriction endonuclease BamHI and PstI, then ligating the 3kb product.

**Mouse colonization and competitive infections**. The streptomycin-treated mouse model was adapted from previous reports (120). Briefly, 4-week-old male C57BL/6 mice (Jackson Laboratories) were provided with drinking water ad libitum

containing 5g/liter streptomycin 48 h prior to inoculation and for the duration of the experiment. Inoculation strains were grown overnight in 3ml LB at 37˚C statically, diluted 1:100 in 20ml warmed DMEM high glucose with phenol red and incubated, shaking, to early logarithmic phase (approximately 3 hours) at 37˚C. Bacteria were pelleted by centrifugation at 4,000rpm for 15 minutes and resuspended in 2m DMEM high glucose with phenol red to a final concentration of 1 x $10^7$ CFU/ml. For competitive infections, a 1:1 mixed suspension of the two competing strains were prepared at a final total concentration of 2 x $10^7$ CFU/ml. Prior to infection with bacterial strains, mice were given 200µl half-saturated sodium bicarbonate solution orogastrically to neutralize gastric acid. Approximately 15 min later, 100µl of the inoculum was administered orogastrically. Bacterial inocula were quantified by plate counts. Fresh fecal pellets were collected every day for 7 days postinfection. Feces were weighed, diluted, and homogenized in sterile PBS; serial dilutions were plated on LB agar with antibiotics. Antibiotic concentrations used in these media were 100µg/ml streptomycin for single strain experiments; 100µg/ml streptomycin and 100µg/ml carbenicillin or 100µg/ml streptomycin and 50µg/ml kanamycin for competition experiments. The limit of detection was approximately $10^3$ CFU/g feces.

For enumeration of bacteria in intestinal sections and contents, a 2- to 3-cm section of the proximal duodenum, mid-jejunum, distal ileum, mid- colon, and partial cecum were excised. Lumenal contents were expelled, weighed, and serial dilutions were made in PBS for plating onto LB agar; bacterial counts were

expressed as CFU/gram intestinal content. Intestinal segments were rinsed in PBS to remove fecal material, homogenized in 1mL PBS with 1.0mm diameter zirconia/silica beads (BioSpec) for 30s using Mini-BeadBeater (BioSpec, Bartlesville, OK), and diluted in PBS for plating onto LB agar. Bacterial counts were determined as described above.

Competitive ratio was calculated as $\log(S_{1F}/S_{2F})/(S_{1i}/S_{2i})$, where $S_{1F}$ represents strain 1 final CFU, $S_{2F}$ represents strain 2 final CFU, $S_{1i}$ represents strain 1 initial CFU, and $S_{2i}$ represents strain 2 initial CFU.

***in vitro* competition assays.** Competing strains were grown overnight in 3ml LB at 37˚C shaking, diluted 1:100 in 10ml LB and incubated, shaking, to mid-late logarithmic phase (approximately 3 hours) at 37˚C. The optical density at 600nm of each culture was measured and normalized to approximately 0.4. A mixed suspension of 1:1 of the two competing strains was diluted 1:100 into 2ml LB or DMEM high glucose and incubated at 37˚C statically. Bacteria inocula were quantified by plate count. Every two hours for the first 8h, 100µl was sampled and diluted in PBS for plating onto LB agar. After 24h, 1:100 dilution was made into 1ml LB or DMEM high glucose and incubated at 37˚C statically for another 24h. In addition, the 24h culture was diluted in PBS for plating onto LB agar. This process of dilution into fresh media and enumeration after 24h was continued for 72h post infection. Competitive ratio was determined as described above.

**RNA extraction and qRT-PCR.** For quantitative reverse transcriptase PCR (qRT-PCR), DH5α(pBAD30) and DH5α(pAar) grown aerobically in LB overnight at 37°C

with shaking and then diluted 1:100 in LB and grown at 37°C. RNA from three biological replicates of each condition was extracted after 3h. RNA was extracted using RNAprotect Bacteria Reagent (Qiagen) followed by a RNeasy Mini Kit (Qiagen). Primers used were previously published for EAEC (Santiago 2014 and 2017). qRT-PCR was performed using a one-step reaction in an ABI 7500-FAST sequence detection system (Applied Biosystems). All data were normalized to the levels of rpoA and analyzed using the comparative cycle threshold (CT) method (22). The relative quantification method was used to determine the expression levels of target genes. Statistical significance was determined by ANOVA with post-hoc Tukey, and a P value of ≤0.05 was considered significant.

## Results

**Expression of *aar* had no effect on colonization in mice given a single strain of EAEC.** We sought to determine what effect *aar* has on *in vivo* colonization. We hypothesized that since *aar* mutants lack the negative regulator of AggR (which induces expression of virulence genes) and are characterized as hyper-fimbriated, strains lacking *aar* would colonize more abundantly. We also hypothesized that the presence or absence of *aar* would affect the location of colonization. To test these hypotheses, we infected streptomycin treated, 4-week-old C57BL/6 mice from Jackson Labs with either $10^8$ colony forming units (CFUs) EAEC WT 042 or with $10^8$ CFUs 042Δ*aar.* Fecal shedding was followed for 7 days post infection. No differences in fecal CFUs of the challenged strains were measured between the

two groups (figure 3.1A). We collected tissue samples from the duodenum, jejunum, ileum, cecum, and colon after 3- and 7-days post infection to determine colonization CFUs. Consistent with fecal shedding, there were no differences in tissue colonization between the two groups in any intestinal section at 3 (figure 3.1B) or 7 days (figure 3.1C).

**Expression of *aar* decreases competitiveness *in vivo.*** The pre-administration of streptomycin to mice clears a metabolic niche by suppressing the commensal microbiota; thus, minor (but biologically significant) effects on colonization fitness may be obscured in this model. Therefore, although we did not detect differences in colonization between 042 and 042Δ*aar* when given individually, we hypothesized that when given together, there may be differences (as previously published (120)), i.e., we hypothesized that strains lacking *aar* would outcompete strains with *aar*. In order to distinguish between competing strains, differential antibiotic resistance was employed. EAEC wildtype strain 042 was transformed with plasmid pBR322 to render the strain carbenicillin resistant; 042Δ*aar* was already kanamycin resistant from the lambda-red allelic deletion process. Streptomycin-treated, 4-week-old C57BL/6 mice were orally gavaged with $10^8$ CFU of equal amounts of 042pBR322 and 042Δ*aar*. Feces from the mice were collected daily for 7 days and tissue colonization in the ileum, cecum, and colon was measured on days 3 and 7 post-infection. Fecal and tissue samples were diluted and plated on either LB with streptomycin/carbenicillin (identifies 042) or streptomycin/kanamycin (identifies 042Δ*aar*). It was apparent after just 1 day that

042Δ*aar* outcompeted 042, as determined by fecal shedding; this trend remained throughout the 7 days of the experiment (Figure 3.2A). The competitive ratio from tissue colonization and intestinal contents supported the observation that 042Δ*aar* outcompeted 042 at 3 days (figure 3.2B) and 7 days (figure 3.2C) in all locations tested.

We sought to confirm that the loss of *aar* was responsible for the competitive advantage, so we compared 042Δ*aar* with 042Δ*aar* complemented with a plasmid containing *aar* (here designated pAar) (30). Plasmid pAar expresses the *aar* gene under the control of the truncated *aar* promoter; the plasmid is built on a pBAD30 backbone (p15A replicon) and confers resistance to ampicillin (30). The expression of *aar* reversed the competitive phenotype of 042Δ*aar* in fecal shedding (figure 3.2D) and in tissue colonization and intestinal contents at 3 days (figure 3.2E) and 7 days (figure 3.2F). These data suggested that the presence of pBR322 in the 042 strain was not accounting for its relative lack of fitness.

It is possible that the increase in competitiveness of 042Δ*aar* compared to 042 may have been due to metabolic differences. To account for this possibility, we performed *in vitro* competition assays using Luria broth (LB; a nutrient-rich medium) and DMEM with high glucose (a minimal medium). We checked the competitive ratio every two hours for the first 8 hours and then again after 1, 2, and 3 days. During the first 8 hours, 042Δ*aar* outcompeted the wildtype 042 in both LB (figure 3.3A) and DMEM with high glucose (figure 3.3C). There were no competitive differences after 1, 2, and 3 days in either the rich medium or the

minimal medium (figure 3.3B and 3.3D). Thus, competitive differences seen *in vivo* are not recapitulated *in vitro*.

**Expression of *aggR* increases EAEC competitiveness *in vivo*.** Our hypotheses tested above were all based on the known negative effect of *aar* on *aggR* and AggR-regulated genes. Thus, we predicted that the loss of *aggR* would (similar to *aar* expression) decrease competitiveness when compared to WT 042. Mice were orally gavaged with a 1:1 ratio of 042(pBR322) and 042Δ*aggR,* and fecal shedding was followed for 7 days. As predicted, the loss of *aggR* decreased the yield of the mutant strain compared to WT 042 (Figure 3.4A).

We next examined if the increased competitiveness of 042 lacking *aar* was due to *aggR*; we therefore, compared colonization of WT 042 to an 042 strain with both *aar* and *aggR* deleted. The additional loss of *aggR* in 042Δ*aar* did indeed reverse the colonization defect manifested by 042Δ*aar* (figure 3.4B). Interestingly, 042Δ*aar*Δ*aggR* was not competitively different from WT 042, suggesting that *aggR* is at least partially responsible for the increased competitiveness observed when *aar* is absent.

AggR activates a large number of genes, and it is therefore possible that the loss of *aggR* is affecting the metabolism of the bacterium simply by increasing energy expenditure. We therefore performed *in vitro* competition assays to interrogate this possibility. In rich medium (LB), the strain lacking *aggR* out-competed the wildtype strain during the initial 8 hours of infection (figure 3.5A) and after 2 days (figure 3.5B). In a minimal medium, the wildtype 042 and 042Δ*aggR*

grew similarly to one another during the initial 8 hours as determined by quantification of CFUs (figure 3.5C). However, over 2-3 days, the loss of *aggR* increased competitiveness (figure 3.5D). It was not surprising that when both *aar* and *aggR* (two energy expensive regulators) were deleted, the double mutant out-competed the wildtype, starting during the initial hours of infection in both LB (figure 3.5E) and DMEM with high glucose (figure 3.5G). This pattern was continued into days 2-3 in LB (figure 3.5F) and DMEM with high glucose (figure 3.5H). This suggests (but does not prove) that the loss of competitiveness in strains lacking *aggR in vivo* is not due to metabolic effects.

**The specific gene regulated by AggR that is responsible for 042Δ*aggR* decreased competitiveness *in vivo* is unclear.** We sought to determine which of the AggR-regulated genes could be responsible for the decrease in *in vivo* competitiveness in the *aggR* mutant. Genes that would likely be responsible for competitiveness include genes encoding colonization factors; colonization factors comprise proteins that promote adhesion to tissue, invasion of the mucosa, or evasion of the host immune system. In EAEC strain 042, genes encoding the aggregative adherence fimbriae (*aafA*), dispersin (*aap*), a type VI secretion system (*aaiC* and *aaiP*), and hypothetical capsule proteins (*orf3* and *orf4*) were targeted.

Contrary to our hypothesis, single deletions to *aafA*, *aap, aaiC, aaiP, orf3,* and *orf4* increased competitiveness compared to the WT 042(pBR322) (figures 3.6A-F, respectively). Thus, we did not identify a specific AggR-dependent gene that was responsible for the loss of competitiveness seen in the *aggR* mutant.

**Expression of *aar* in DH5α decreases competitiveness.** Our findings above suggest that the increased competitiveness seen in the *aar* mutant strain of 042 is only partially due to the effect of *aar* on *aggR*. We have previously reported that Aar regulates genes outside of the AggR regulon (58), and that these genes, which are part of the *E. coli* core genome, can be affected by expression of *aar* in K12 (121). The use of K12 strain DH5α allowed us to explore how *aar* is affecting competitiveness and fitness via core genes without the potential interference of AggR and its regulon.

We next tested if the expression of *aar* had a similar effect on competitiveness in DH5α as it did in EAEC strain 042. We again employed different antibiotic resistance phenotypes to distinguish between the two different bacterial strains. In this experiment DH5α was transformed with either the empty plasmid pACYC177 (conferring kanamycin resistance; here designated pACYC) or with pAar (conferring carbenicillin resistance). Mice were orally gavaged with a 1:1 ratio of DH5α(pACYC) and DH5α(pAar). As predicted, the expression of *aar* decreased the competitiveness of DH5α compared to WT DH5α in fecal shedding for all 7 days (figure 3.7A), and in tissue colonization and intestinal contents at 3 days (figure 3.7B) and 7 days post infection (figure 3.7C).

The addition of a plasmid-encoded gene can have a detrimental effect on metabolic efficiency (122). Using the *in vitro* competition assay previously described, we determined that the addition of pAar did not affect fitness in LB during the initial 8hr (figure 3.8A) and throughout 3 days (figure 3.8B). In DMEM

high glucose, the strain harboring pAar initially outcompeted the wildtype during the first 2 hours, but then the wildtype started to outcompete the strain with pAar at 8 hr (figure 3.8C). Day 1 post infection in DMEM high glucose resulted in DH5α(pAar) outcompeting DH5α(pACYC), but during day 2 and 3 both strains were recovered equally (figure 3.8D).

**Expression of *aar* in DH5α increases the expression of housekeeping regulators.** We next sought to better understand how the expression of *aar* in DH5α decreases bacterial competitiveness compared to the wildtype. RNA sequencing in EAEC strain 042, revealed that *aar* expression affects the expression of housekeeping regulators such as *hns* and *hha* (58), and we have previously shown that the expression of *aar* in DH5α affects chromosomal genes (121). If the expression of *aar* in DH5α affects the expression of housekeeping regulators, like it does in EAEC strain 042, then these genes may be responsible for the effect *aar* has on competitiveness in DH5α. We transformed DH5α with either pAar or its corresponding empty vector pBAD30. As predicted, the expression levels of *hns* and *hha* were increased when *aar* was induced (figure 3.9).

**Changes in the expression of *hns* in DH5α lead to decreased competitiveness.** The qRT-PCR data above suggest that *aar* increases the expression of *hns* in DH5α. We hypothesized that any change to the natural level of *hns* would decrease *in vivo* competitiveness. Competitiveness was measured for DH5αΔ*hns* and DH5α(pHns) against wildtype DH5α. Mice were orally gavaged

with a 1:1 ratio of DH5α that was transformed with pBR322 to give the strain carbenicillin resistance and DH5α∆*hns* that is kanamycin resistant or with DH5α that was transformed with pACYC to confer kanamycin resistance, and DH5α transformed with pHns. Fecal shedding was followed for 7 days and colonization was measured at 3- and 7-days post infection. As predicted, wildtype DH5α outcompeted the *hns* mutant in fecal shedding for all 7 days sampled (figure 3.10A) and in colonization after 3- and 7-days (figure 3.10B and C, respectively). The DH5α with increased *hns* was also outcompeted by the wildtype DH5α in fecal shedding (figure 3.10D) and in colonization after 3- and 7- days (figure 3.10E and F, respectively). Thus, alterations to the native level of *hns* expression reduce bacterial competitiveness.

**The expression of *aar* has no effect on competitiveness in DH5α when *hns* is absent.** To elucidate if the decrease in competitiveness in DH5α when *aar* is expressed is due to the effect of *aar* on *hns*, we orally gavaged mice with a 1:1 ratio of DH5α∆*hns* (kanamycin resistant) and DH5α∆*hns*(pAar) (kanamycin sensitive and carbenicillin resistant). We measured CFUs in fecal shedding for 7 days and in tissue colonization after 3- and 7- days post infection. The strain expressing *aar* had a slight competitive advantage for the first 2 days of infection but starting at day 3 there was no difference competitively between the two strains in fecal shedding (figure 3.11A). Tissue colonization and intestinal contents reflected fecal shedding with no differences in competitiveness measured after 3

and 7 days in any intestinal compartment sampled (figure 3.11B and C, respectively).

**Plasmids are stable *in vivo* for the duration of the study.** Due to the use of multiple plasmids for antibiotic resistance or gene expression, it was important to determine how stable the plasmids were *in vivo* when not under the pressure of antibiotic selection. Mice were orally gavaged with $10^8$ CFUs of DH5α transformed with either pBR322, pACYC, pAar, or pHns. Feces were collected every other day for 7 days to enumerate CFUs of bacteria that grew on LB with streptomycin (streptomycin resistance is on the bacterial chromosome) or LB with streptomycin and carbenicillin or kanamycin (carbenicillin and kanamycin resistance are located on the plasmids). There were no differences in CFU counts between the streptomycin and streptomycin with carbenicillin or kanamycin for DH5α transformed with pBR322 (figure 3.12A), pACYC (figure 3.12B), pAar (figure 3.12C), or pHns (figure 3.12D) for all days sampled.

The use of empty vectors to provide selective antibiotic resistance could potentially affect competitiveness. We first examined if DH5α transformed with pBAD30 (the corresponding empty vector for both pAar and pHns) was competitively different from DH5α transformed with pACYC (plasmid used to provide selective kanamycin resistance). Mice were orally gavaged with a 1:1 ratio of the two strains and CFUs in fecal shedding were measured daily for 7 days. There were no differences in the competitive ratio of DH5α transformed with pACYC or pBAD30 for all 7 days sampled (figure 3.13A).

It was also important to ensure that the addition of the empty plasmids we used in order to add antibiotic resistance did not affect competitiveness compared to DH5α without the addition of a plasmid. Rifampicin-resistant colonies of DH5α were selected after growth on LB agar plates with 100μg/ml rifampicin. Mice were orally gavaged with a 1:1 ratio of rifampicin resistant DH5α and DH5α transformed with either pBR322 or pACYC and CFUs in fecal shedding were measured daily for 7 days. DH5α outcompeted DH5α with pACYC and pBR322 for the first 3 days of infection; however, after 3 days there were no differences in competitiveness (figure 3.13B and C, respectively).

**Discussion**

There are obvious drawbacks when using mouse models of infection to study a human pathogen, namely the proven host specificity of enteric pathogens (123). In addition, mice and humans have different intestinal lengths and intestinal transit times that may factor into timing and colonization (124). These differences in host anatomy likely resulted in the lack of differences in the abundance or location of colonization when mice were given a single strain of 042Δ*aar* compared to wildtype 042 despite *aar* downregulating the AggR-regulon, which includes AAF (30). The lack of differences could also be due to both strains' ability to colonize the same niche that is opened by the treatment of streptomycin and the lack of competition for that niche. However, determining if and how *aar* is affecting

colonization and bacterial competitiveness in mice provides some initial insight into how EAEC may be affecting pathogenesis in humans.

Here we use *in vivo* and *in vitro* competition assays to better understand how *aar* is affecting bacterial fitness and to illuminate which *aar*-affected genes are responsible for the phenotype. Our data suggest that expression of *aar* decreases the competitiveness of both pathogenic EAEC strain 042 and nonpathogenic K12 strain DH5α. In 042, *aar* is affecting competitiveness partially through downregulating the expression of *aggR*. In DH5α, which lacks the AggR regulon, the decreased competitiveness caused by *aar* is due (at least in part) to the effect on *hns*. This suggests that Aar serves as a molecular liaison, connecting virulence and core fitness genes. Presumably, the effect in K12 affects the metabolic fitness of the bacterium, although an effect on adherence is not ruled out.

As stated above, the lack of differences seen in mice given a single strain may be due to lack of competition for the niche, so both strains were given to the mice in a 1:1 ratio to measure their competitiveness. Fecal shedding demonstrated that 042Δ*aar* was shed at a greater ratio than that of 042. Early on during the course of infection, this could have been due to 042Δ*aar* being washed out of the intestines, with 042 colonizing better; however, since the trend continued for 7 days, this did not support that hypothesis. Tissue colonization was measured to also confirm that fecal shedding was reflective of colonization.

*In vitro* competition between 042Δ*aar* and 042 suggested that during early growth, 042Δ*aar* was outcompeting wildtype 042. This may suggest an early

metabolic advantage to the mutant; however, the competitive ratio was not significantly different from zero. The loss of *aar* also affects the expression of the AggR regulon, which includes AAF and *aap* (30). The increase in *aap* may cause the bacteria to be less aggregated, which could affect CFU counts. We have previously seen in our lab that in the *aap* mutant, the bacteria cluster more tightly together, and this results in lower CFU counts compared to the actual number of bacteria detected by qPCR.

We had predicted that the main cause of 042Δ*aar* outcompeting the WT is due to *aar* downregulating the AggR regulon (30); therefore, it was not surprising that the *aggR* mutant was outcompeted by WT 042. The *in vitro* competition assay suggested that 042Δ*aggR* has a metabolic advantage over the WT strain; these combined data suggest that the detrimental effect in 042Δ*aggR* on *in vivo* competition is due to the loss of some other, presumably colonization-specific, factor. We sought to identify this AggR-dependent colonizing factor by competing WT 042 against 042 mutants in the major fimbrial gene (*aafA*), dispersin (*aap*), the type VI secretion effector (*aaiC*), the type VI secretion ATPase (*aaiP*) (53), or genes predicted to be involved in the capsule (*orf3* or *orf4*) (32). All of the mutants outcompeted the WT *in vivo*. It is likely that the six genes tested produce proteins that are host species specific and do not contribute to colonization in mice (123); therefore, the loss of these virulence factors likely gave the strains a metabolic advantage compared to the WT. It is possible that the AggR-dependent mouse colonizing factor may be one of the remaining >35-AggR regulated genes not

explored here (although nearly all of these genes are part of molecular complexes operating in concert with the products of the genes tested). We could use signature-tagged mutagenesis to investigate all AggR regulated genes with each gene mutation tagged with a different DNA sequence (125), as this would allow an unbiased, high-throughput analysis. It is quite possible that, similar to epidemiological studies in humans (7, 8, 11, 119), it is the combination of genes that is responsible for effect of AggR.

The additional loss of *aggR* in 042Δ*aar* resulted in a similar competitiveness compared with WT 042. This suggested that the increased competitiveness in 042Δ*aar* is only partially due to the unregulated increase in *aggR* expression. If the increased competitiveness seen in the *aar* mutant was solely due to *aar*'s effect on *aggR*, then the competitive ratio of the double mutant compared to WT would have a similar pattern compared with the *aggR* mutant against WT.

Since AggR is only partially responsible for the increased competitiveness seen in the *aar* mutant, we sought to determine what other factor or factors could be contributing. As previously stated, outside of AggR, *aar* effects *E. coli* core genes (58), and so to determine if decreased competitiveness when *aar* is expressed is due to *aar*'s effect on core genes, we chose to perform experiments in K12 strain DH5α, which lacks the AggR regulon. As predicted, the expression of *aar* in DH5α decreased competitiveness when compared to WT DH5α. The *in vitro* competition assay demonstrated that this was not due to a metabolic effect. Although we have demonstrated that core genes that are affected by *aar* in EAEC

strain 042 are similarly altered in DH5α (121), we had not previously measured the effect of *aar* on global regulators in DH5α. In support of previous work, global regulators *hns* and *hha* were both increased when *aar* was expressed in DH5α. Hha is involved in regulating virulence genes in response to environmental cues, usually by forming complexes with H-NS (105). H-NS has a more global effect and can regulate approximately 5% of *E. coli* genes (90); therefore we chose to focus on how *aar* is affecting *hns.* In our hands, any change, whether an increase or decrease, compared with the natural levels of *hns* resulted in decreased competition compared to the WT, demonstrating the importance of *hns* on bacterial fitness. It was exciting to find that *aar* had no effect on the competition of DH5α when the strain lacked *hns*. This suggested that, in DH5α, the expression of *aar* is affecting competitiveness by altering the global regulator, *hns*. These data are consonant with previous observations from our laboratory that Aar directly binds H-NS *in vitro* (58).

All of our competitive pairs included the presence of at least one plasmid to confer antibiotic resistance. When inoculated into mice, the strains are no longer under antibiotic pressure to maintain the plasmid, and because we are selecting for the plasmid during enumeration, the loss of the plasmid could result in an underestimation of that specific strain. Thus, it was important to ensure that the plasmids were stable *in vivo* during the course of the experiment. Our data suggested that the plasmids were stable *in vivo* and that the increased competitiveness in mutant strains is not due to wildtype strains losing plasmids. In

addition, we determined that the presence of the empty plasmid vectors had no effect on competition; thus, changes to competitiveness with the addition of pAar or pHns were not due to the presence of the plasmid.

Based on our results, we propose a model to explain how Aar is affecting colonization (figure 3.14). In pathogenic *E. coli,* Aar downregulates virulence genes by downregulating *aggR*. This decrease in virulence genes hinders the bacterium's ability to colonize maximally. However, the effect on virulence genes is not the only way Aar is decreasing colonization potential; Aar is also increasing the expression of core genes, specifically *hns.* We speculate that the level of *hns* expression within a bacterium is adjusted for maximal fitness of the bacterium; any variations to this natural level will result in decreased fitness. Therefore, Aar is affecting colonization via two routes: decreasing expression of virulence genes and increasing expression of core genes.

The data presented here demonstrate the complex nature of Aar in EAEC and raises the important question - why does *aar* exist if its expression negatively affects colonization? We are tempted to speculate that because the ANR family is only found in pathogenic *E. coli,* the purpose of *aar* may be to temper the bacterium's ability to cause harm to the host. It is currently beyond our capabilities to directly test this hypothesis since there is no whole animal model that fully recapitulates EAEC pathogenesis; however, we can use epidemiological studies (11) and studies done in mice using a mouse pathogen (*C. rodentium*) with an *aar* homolog (30) to elucidate effects of Aar. Evolutionarily, the benefit of Aar

tempering the harmful effects of the bacterium could be to keep hosts healthy so that they remain ambulatory and therefore lead to further transmission of the bacteria to new hosts.

**Table 3.1. Strains and plasmids used**

| Name | Characteristics | Source |
|---|---|---|
| | **Strains** | |
| **EAEC 042** | Enteroaggregative *E. coli* strain 042 | (48) |
| **042Δ*aar*** | 042 derivative carrying *aar* deletion | (30) |
| **042Δ*aggR*** | 042 derivative carrying *aggR* deletion | (26, 47) |
| **042Δ*aar*Δ*aggR*** | 042 derivative carrying *aar* and *aggR* deletions | (121) |
| **042Δ*aafA*** | 042 derivative carrying *aafA* deletion | (118) |
| **042Δ*aap*** | 042 derivative carrying *aap* deletion | (47) |
| **042Δ*aaiC*** | 042 derivative carrying *aaiC* deletion | (53) |
| **042Δ*aaiP*** | 042 derivative carrying *aaiP* deletion | (53) |
| **042Δ*orf3*** | 042 derivative carrying *orf3* deletion | (32) |
| **042Δ*orf4*** | 042 derivative carrying *orf4* deletion | (32) |
| **DH5α (strep$^R$)** | K12 strain DH5α derivative that is resistant to streptomycin | This study |
| **DH5αΔ*hns* (strep$^R$)** | DH5α derivative carrying *hns* deletion and is resistant to streptomycin | This study |
| | | |
| | **Plasmids** | |
| **pBR322** | Cloning and expression vector (Amp$^R$, Tet$^R$) | Lab Collection |
| **pACYC177 (Amp$^s$)** | Cloning and expression vector (Km$^R$) | This study |
| **pAar** | $P_{aar}$-*aar* from 042 strain cloned into pBAD30 (Amp$^r$) | (30) |
| **pHns** | *hns* from 042 strain cloned into pBAD30 (Amp$^r$) | |
| **pBAD30** | Cloning expression vector (Amp$^R$) | Lab Collection |

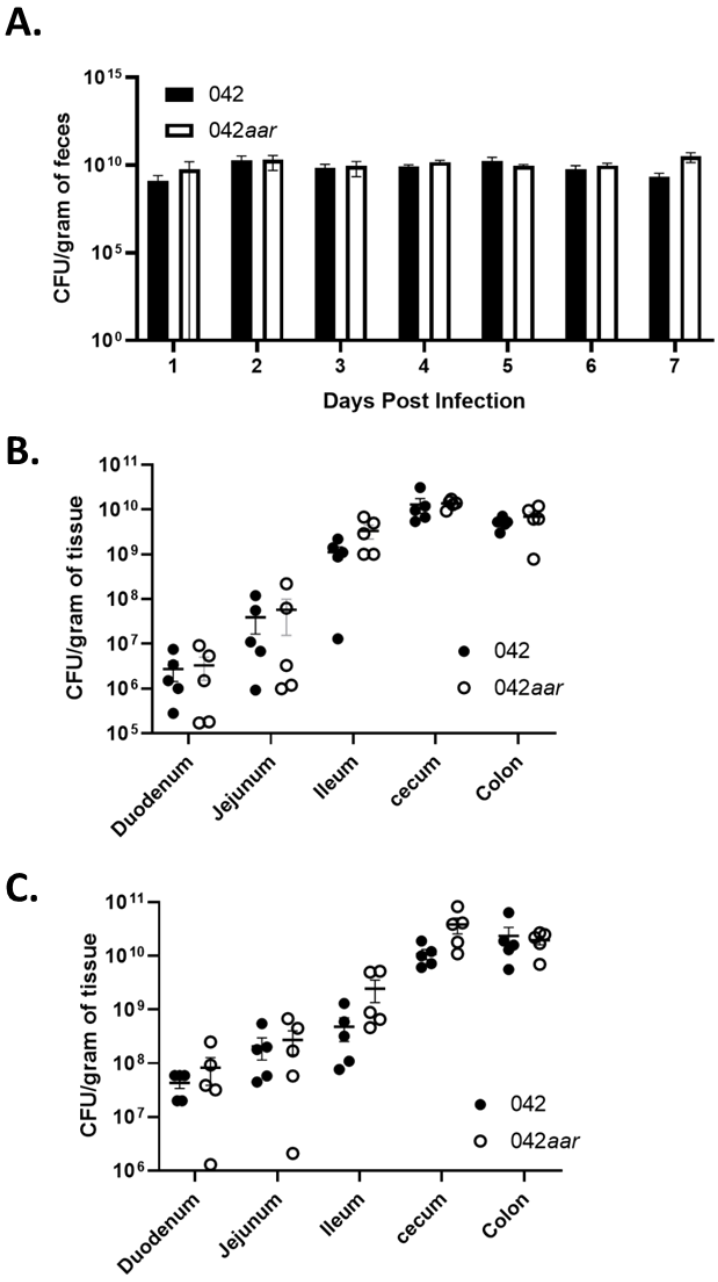



**Figure 3.1. Colonization and fecal shedding of mice infected with WT 042 or 042Δ*aar*.** (A) Fecal shedding of 042 and 042Δ*aar* measured for 7 days post infection. (B) Tissue colonization of 042 and 042Δ*aar* after 3 days post infection and (C) 7 days post infection. CFU data are representative of at least three

independent experiments. Differences were not statistically significant via Mann-Whitney test.

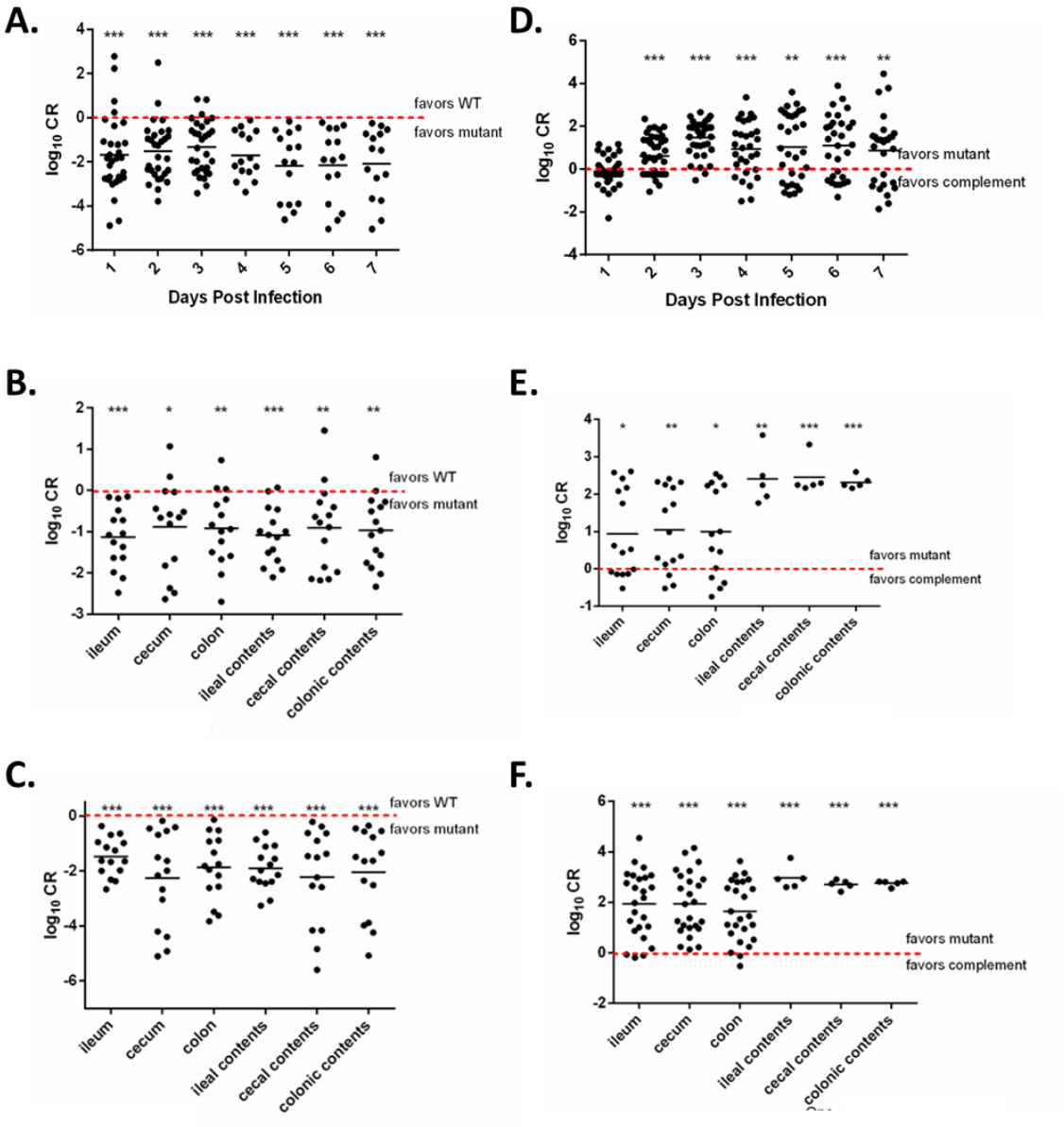



**Figure 3.2. Loss of *aar* leads to increased competitiveness *in vivo*.** (A) Competitive ratio of 042 and 042Δ*aar* from fecal shedding was determined every day for 7 days post infection. (B) Competitive ratio of 042 and 042Δ*aar* for tissue colonization after 3 days and (C) 7 days. (D) The competitive ratio of 042Δ*aar* and 042Δ*aar*(pAar) from fecal shedding. (E) The competitive ratio of 042Δ*aar* and

042Δ*aar*(pAar) for tissue colonization after 3 days and (F) 7 days. The competitive ratios combine at least three independent experiments. Asterisks indicate significant differences by ANOVA (*, P<0.05; **, P<0.005; ***, P<0.0005).

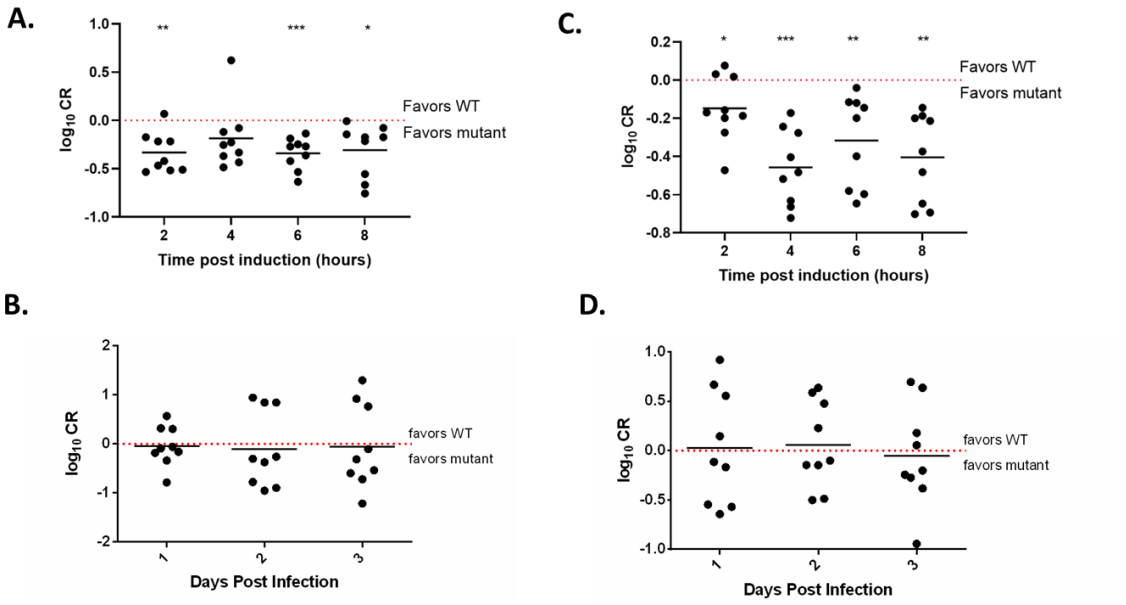



**Figure 3.3. *In vitro* competitive ratio of wildtype EAEC strain 042 and 042Δ*aar*.** (A) Competitive ratio was determined in LB for the first 8 hours post infection and (B) the first 3 days post infection. (C) Competitive ratio was determined in DMEM high glucose for the first 8 hours post infection and (D) the first 3 days post infection. The competitive ratios combine at least three independent experiments. Asterisks indicate significant differences by ANOVA (*, P<0.05; **, P<0.005; ***, P<0.0005).

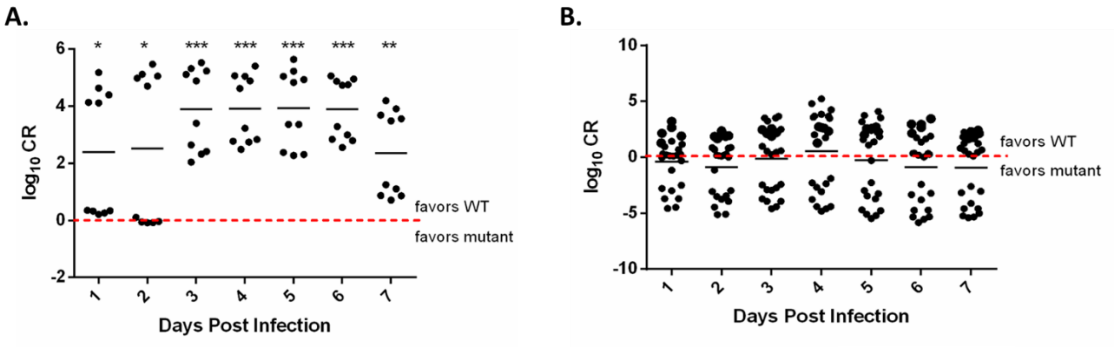



**Figure 3.4. The loss of *aggR* on competitive ratio.** (A) Competitive ratio of 042 and 042Δ*aggR* in fecal shedding. (B) Competitive ratio of 042 and 042ΔΔ*aar*Δ*aggR* in fecal shedding. The competitive ratios combine at least three independent experiments. Asterisks indicate significant differences by ANOVA (*, P<0.05; **, P<0.005; ***, P<0.0005).

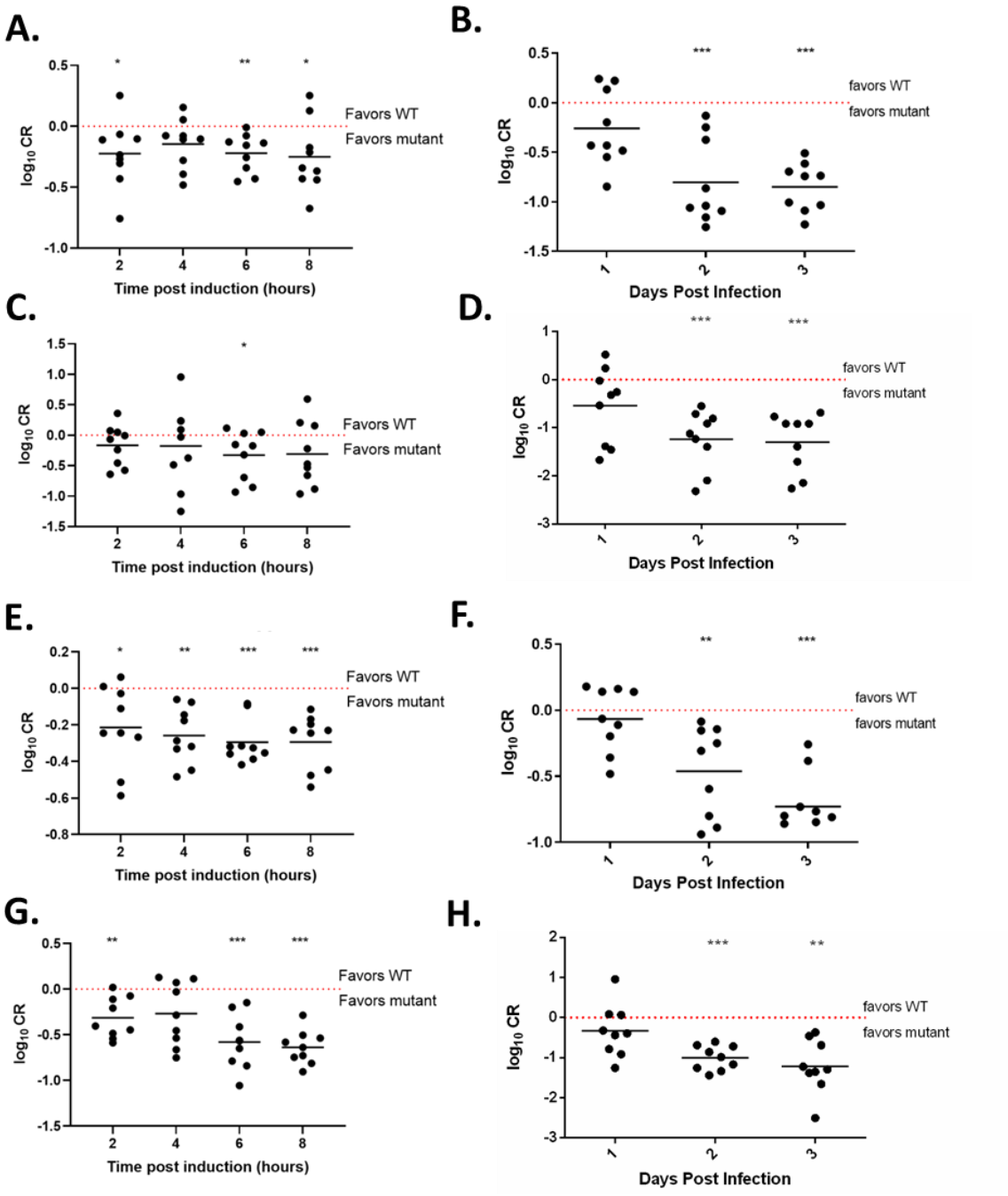



**Figure 3.5. *In vitro* competitive ratio of wildtype EAEC strain 042 and *aggR* mutants.** (A) Competitive ratio of 042 and 042Δ*aggR* was determined in LB for the first 8 hours post infection and (B) the first 3 days post infection. (C) Competitive

ratio of 042 and 042Δ*aggR* was determined in DMEM high glucose for the first 8 hours post infection and (D) the first 3 days post infection. (E) Competitive ratio of 042 and 042Δ*aar*Δ*aggR* was determined in LB for the first 8 hours post infection and (F) the first 3 days post infection. (G) Competitive ratio of 042 and 042Δ*aar*Δ*aggR* was determined in DMEM high glucose for the first 8 hours post infection and (H) the first 3 days post infection. The competitive ratios combine at least three independent experiments. Asterisks indicate significant differences by ANOVA (*, $P<0.05$; **, $P<0.005$; ***, $P<0.0005$).

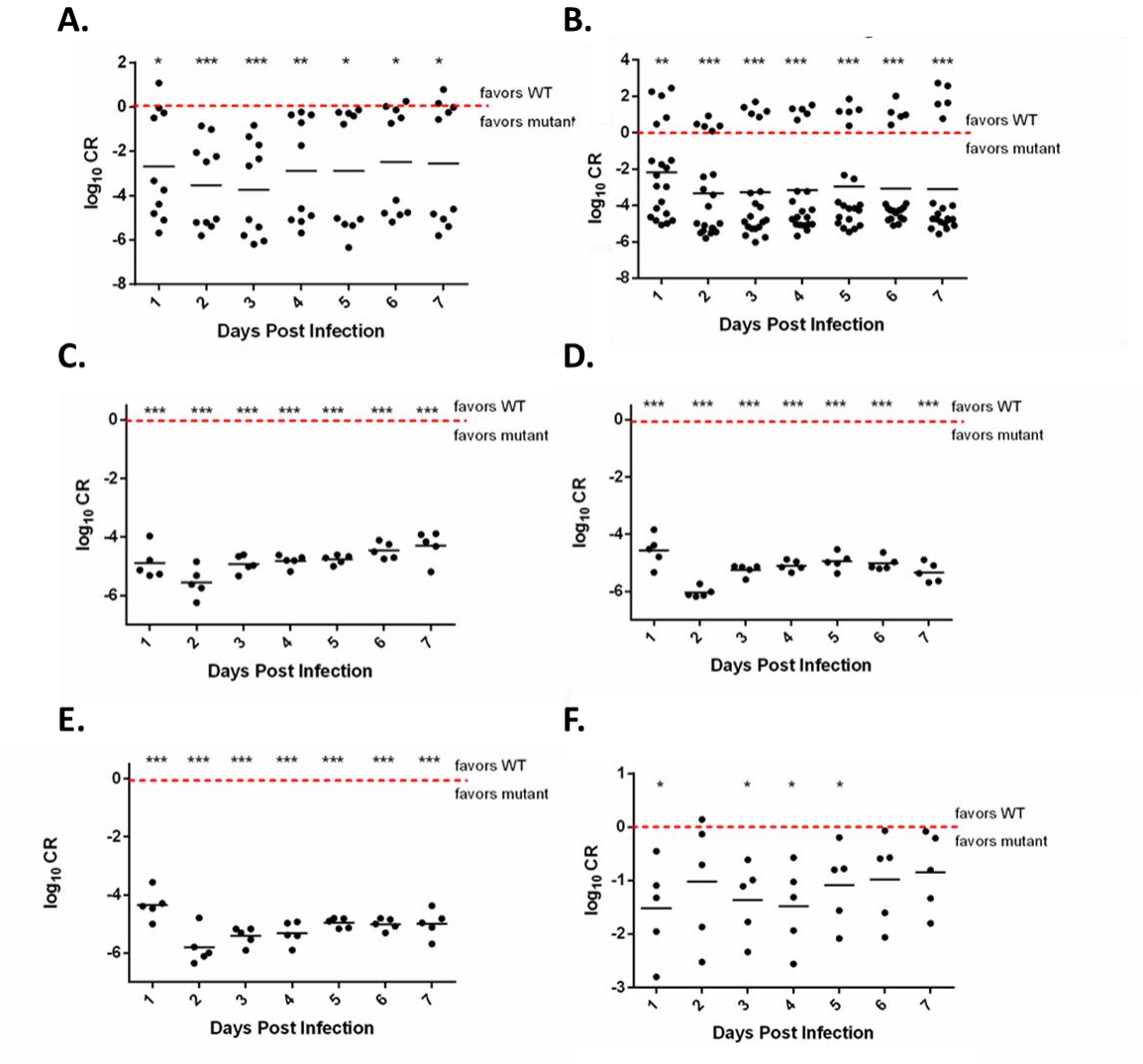



**Figure 3.6. Competitive ratio in fecal shedding between 042 and AggR-regulated gene mutations.** (A) Competitive ratio of 042 against 042Δ*aafA,* (B) 042Δ*aap,* (C) 042Δ*aaiC,* (D) 042Δ*aaiP,* (E) 042Δ*orf3,* and (F) 042Δ*orf4.* The competitive ratios are representative of at least one independent experiments. Asterisks indicate significant differences by ANOVA (*, P<0.05; **, P<0.005; ***, P<0.0005).

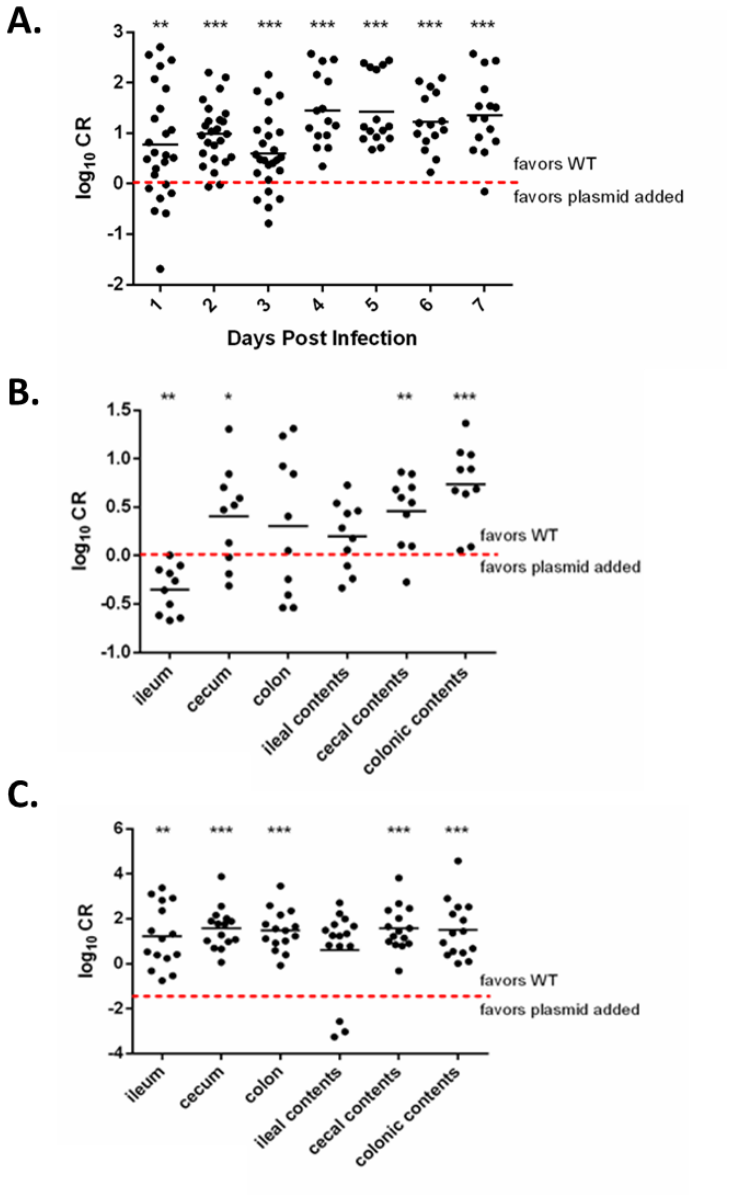



**Figure 3.7. Addition of *aar* to K12 strain DH5α leads to decreased competitiveness *in vivo*.** (A) Competitive ratio of DH5α and DH5α(pAar) from fecal shedding was determined every day for 7 days post infection. (B) Competitive ratio of DH5α and DH5α(pAar) for tissue colonization after 3 days and (C) 7 days.

The competitive ratios combine at least three independent experiments. Asterisks indicate significant differences by ANOVA (*, P<0.05; **, P<0.005; ***, P<0.0005).

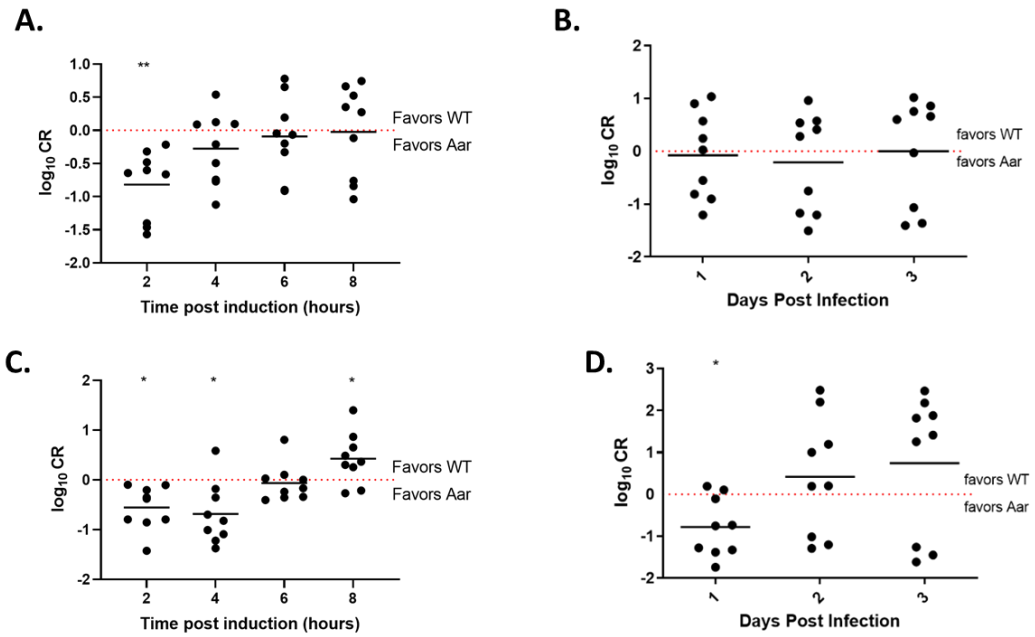



**Figure 3.8. *In vitro* competitive ratio of wildtype K12 strain DH5α and DH5α(pAar).** (A) Competitive ratio was determined in LB for the first 8 hours post infection and (B) the first 3 days post infection. (C) Competitive ratio was determined in DMEM high glucose for the first 8 hours post infection and (D) the first 3 days post infection. The competitive ratios combine at least three independent experiments. Asterisks indicate significant differences by ANOVA (*, P<0.05; **, P<0.005; ***, P<0.0005).

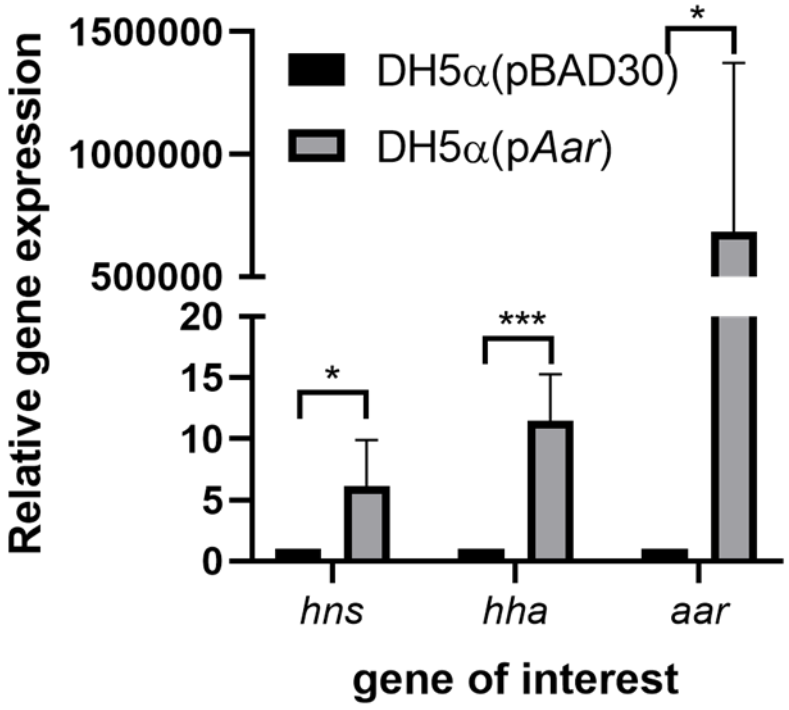



**Figure 3.9. Gene expression of housekeeping regulators is increased in DH5α(pAar) compared to DH5α.** DH5α was transformed with pAar or its corresponding empty vector pBAD30. Transcriptional levels of *hns, hha*, and *aar* were analyzed by qRT-PCR. qRT-PCR data are representative of at least three independent experiments. Asterisks indicate significant differences by multiple t-test (*, P<0.05; **, P<0.005; ***, P<0.0005).

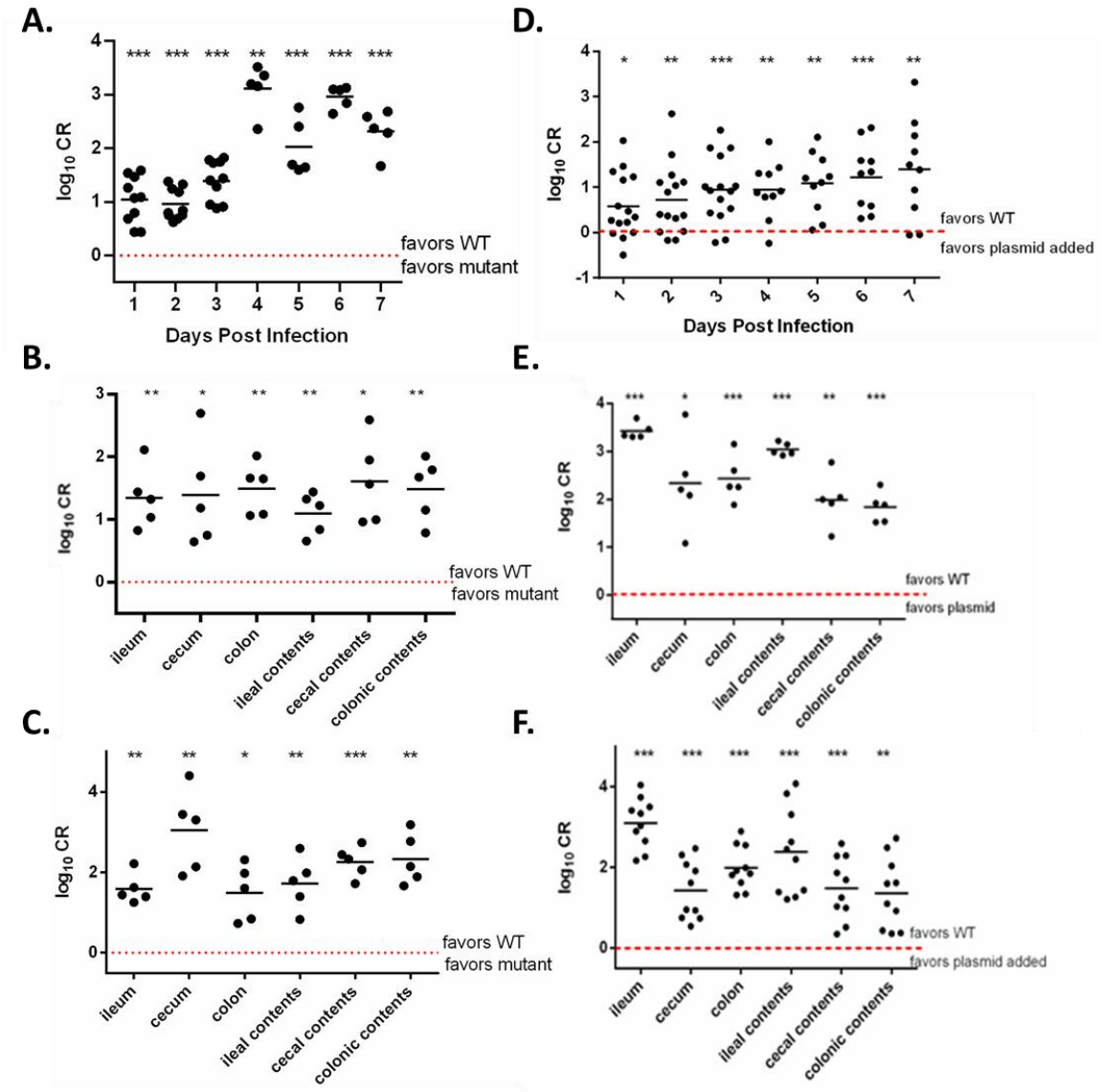



**Figure 3.10. Changes to the native expression of *hns* in DH5α reduce competitiveness.** (A) Competitive ratio of DH5α and DH5αΔ*hns* from fecal shedding was determined every day for 7 days post infection. (B) Competitive ratio of DH5α and DH5αΔ*hns* for tissue colonization after 3 days and (C) 7 days. (D) The competitive ratio of DH5α and DH5α(pHns) from fecal shedding. (E) The competitive ratio of DH5α and DH5α(pHns) for tissue colonization after 3 days and

(F) 7 days. The competitive ratios combine represent at least one independent experiments. Asterisks indicate significant differences by ANOVA (*, $P<0.05$; **, $P<0.005$; ***, $P<0.0005$).

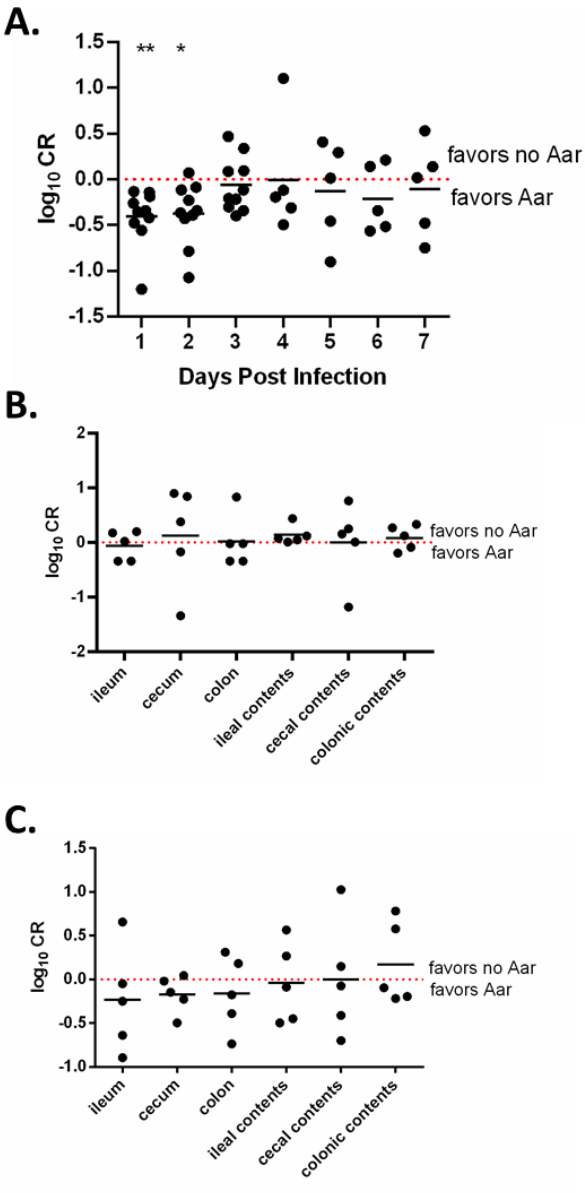



**Figure 3.11. *aar* has no effect on competitiveness in DH5α when *hns* is absent.** (A) Competitive ratio of DH5αΔ*hns* and DH5αΔ*hns*(pAar) from fecal shedding was determined every day for 7 days post infection. (B) Competitive ratio of DH5αΔ*hns* and DH5αΔ*hns*(pAar) for tissue colonization after 3 days and (C) 7

days. The competitive ratios are from one experiment. Asterisks indicate significant differences by ANOVA (*, $P<0.05$; **, $P<0.005$; ***, $P<0.0005$).

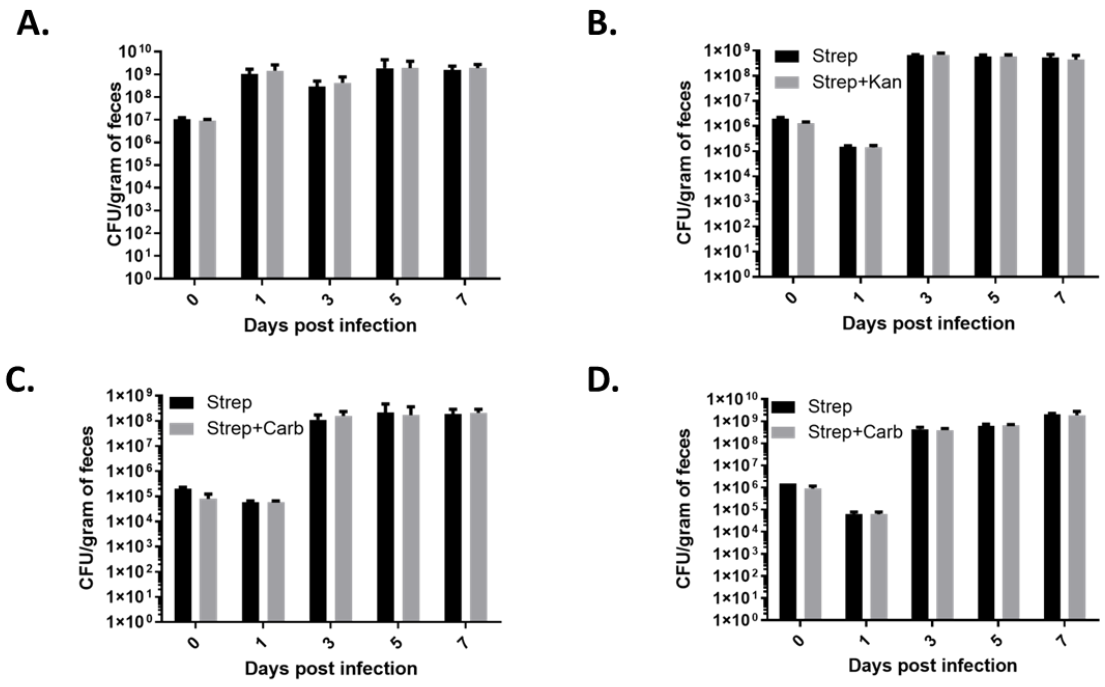



**Figure 3.12. *in vivo* plasmid stability.** Mice were orally gavaged with DH5α transformed with (A) pBR322, (B) pACYC177 carb$^{s}$, (C) pAar, or (D) pHns. Feces were collected every other day and CFUs were plated on LB with streptomycin (black bars) or LB with streptomycin/carbenicillin or kanamycin (grey bars) to determine plasmid stability. CFUs are from one experiment. There were no statistical differences.

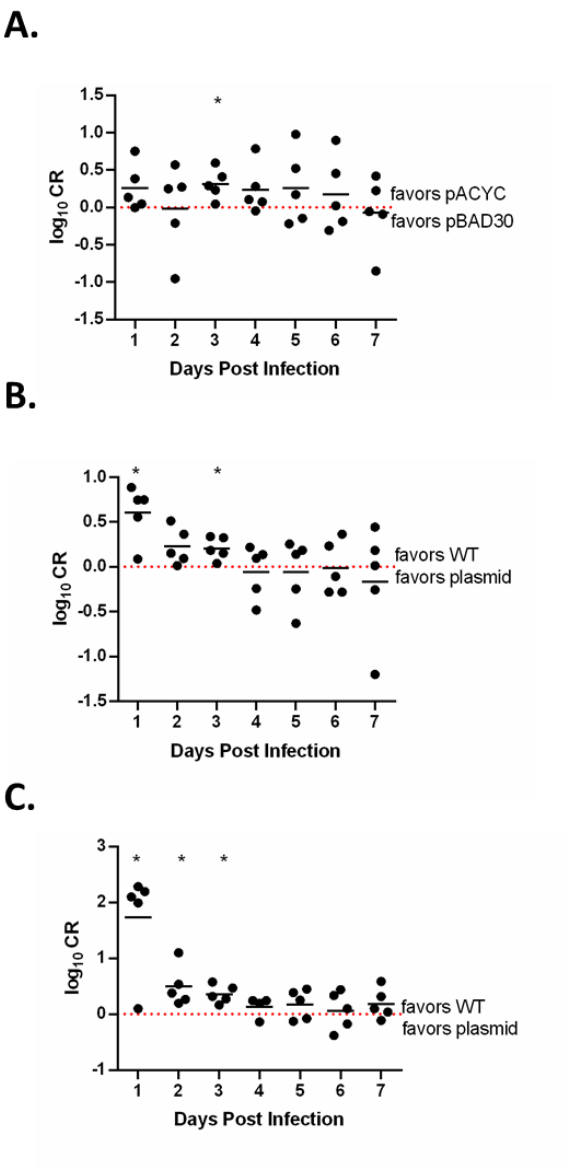



**Figure 3.13. The presence of empty vectors has no effect on the competitive ratio in DH5α.** (A) Competitive ratio of DH5α(pACYC) and DH5α(pBAD30) (B) DH5α and DH5α(pACYC) and (C) DH5α and DH5α(pBR322) from fecal shedding was determined every day for 7 days post infection. The competitive ratios represent a single experiment. Asterisks indicate significant differences by ANOVA (*, P<0.05; **, P<0.005; ***, P<0.0005).

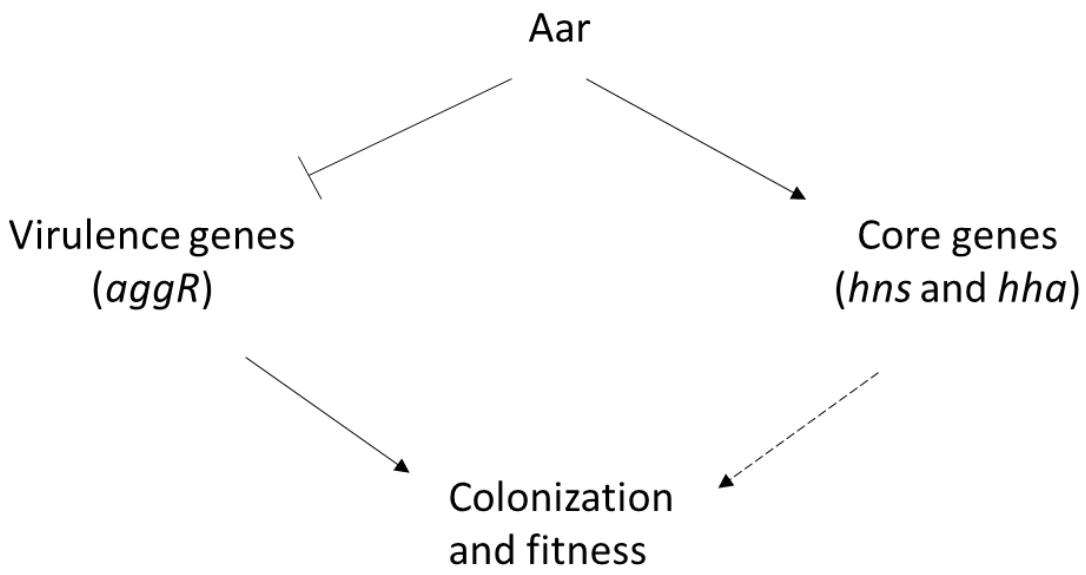



**Figure 3.14. The expression of *aar* decreases bacterial colonization and fitness via virulence and core genes.** Colonization and fitness are hindered by *aar* downregulating *aggR* which decreases virulence gene expression. In addition, colonization and fitness are decreased by *aar* increasing the expression of core genes such as *hns* and *hha.*

# Chapter Four

## Discussion, Future Directions, and Implications to the Field

## Summary

Previously, the binding of Aar to the virulence regulator AggR and the global regulator H-NS was observed *in vitro*, but functional significance of these binding events was unclear (31, 58). Our work has provided support for the hypotheses that both of these binding events in fact do occur *in vivo,* and moreover that each binding event has functional significance on a microscale (gene expression) (121) and on a macroscale (bacterial fitness). In addition, our work suggests that Aar may regulate gene expression via a novel mechanism that includes both anti-activation and counter-silencing.

Our data suggest that on a microscale, interaction of Aar with both AggR and H-NS affects virulence gene expression. The increased expression of *aafA* resulting from low expression of *aar* was contrary to our original hypothesis that Aar was acting solely as an anti-activator. By identifying an additional role of Aar in counter-silencing H-NS during initial pathogenesis, we now hypothesize a more complex mechanism for Aar in both virulence gene and core metabolic gene expression. Early in pathogenesis, we infer that by removing H-NS and its inhibitory effect, Aar upregulates the expression of fimbrial genes; then, as the concentration of Aar increases, Aar binds to AggR and prevents further upregulation of fimbrial genes, either by removing the activator or by preventing additional binding events (121). Drawing from all the data presented in this dissertation, we propose the following integrated mechanism for the contribution of Aar. First, upon bacterial entry to the intestine (figure 4.1A), *aggR* is induced by

a host signal (possibly glucose or maltose) (26) and begins a positive feedback loop to upregulate its own expression and the expression of genes within its regulon (including fimbrial genes and *aar*) (figure 4.1B) (27, 32). The initial low concentration of Aar binds to and removes H-NS from AT-rich genes (including genes in the AggR-regulon). Our data do not permit us to conclude whether this H-NS effect occurs at the affected promoters or at the level of the structural genes, as H-NS has been shown to bind to both sites (102). Removing H-NS from the relevant genes enables unchecked AggR activation (figure 4.1C). However, as the concentration of Aar increases, there is more free Aar available to bind to AggR, thereby inhibiting its activation (figure 4.1D).

In addition, we hypothesize that Aar may have a role in the timing of response by virulence genes to environmental changes. This hypothesis is supported not only by our Aar/AggR titration data, but also by mathematical modeling, which predicts that the presence of a negative regulator decreases a genetic circuit's response time to change. We were able to experimentally support the mathematical model by measuring the hourly abundance of AafA in WT 042 and 042Δ*aar* (see Appendix 1), where we observed earlier expression of AafA on the bacterial surface when *aar* was expressed, compared to when it was absent.

On a macroscale, we have demonstrated that the expression of *aar* affects bacterial fitness both by regulating the expression of virulence genes and by regulating core genes. Our work in the streptomycin-treated mice enabled us to investigate how changes in genes regulated by Aar affect the likelihood of bacterial

survival; however, because this is only a colonization model, we were unable to investigate EAEC pathogenesis in that model (126). As EAEC is a human pathogen, many of its virulence factors (i.e. fimbriae) are host-specific and therefore will not elicit the same response in mice as in humans (123). There is currently no mouse model for EAEC pathogenesis in which the AggR regulon plays its full role in recapitulating all aspects of human disease. The best animal model for studying EAEC pathogenesis is arguably the gnotobiotic piglet, which due to similarities to human gastrointestines, nutrition, metabolism, and immunology, demonstrate similar effects caused by EAEC on colonic pathology and host health; however, there are obvious drawbacks to this model such as cost and scalability (126).

Together, these data demonstrate that there is functional significance for each binding pair (Aar/AggR and Aar/H-NS) at the level of gene expression and more broadly at the level of bacterial fitness. These findings in EAEC expand our knowledge of how the bacterium is regulating virulence gene and house-keeping gene expression in a nuanced manner.

**Future Directions**

The work proposed provides important advances to our understanding of the contribution of the novel protein Aar to bacterial gene regulation. However, the work leaves some additional questions open for further experimentation. Our

proposed mechanism for how Aar is affecting AggR-regulated genes via AggR and H-NS was initially formed by the observation that using the bacterial two-hybrid system, Aar bound to both AggR and H-NS (31, 58). This suggested that were both binding events to occur in the live bacterium, they might occur at different protein concentrations, and therefore at different times during the pathogenic sequence. Given that the *in vivo* concentrations of the various proteins are unknown, it was impossible to predict a priori what these timing phenomena would be. The data presented above provide a plausible scenario. However, more needs to be done.

We would first propose to investigate in greater detail the strength of binding affinities between untagged or tagged (6XHis or HA) Aar/H-NS (figure 4.2A and B) and AggR/H-NS (figure 4.2C and D) using surface plasmon resonance (SPR) (127, 128). We hypothesize that H-NS will have a low or undetectable binding affinity for AggR, as there is no evidence that there is direct binding between the two proteins. We have previously determined the binding affinity of Aar/AggR using SPR (31). Next, we propose to investigate how those binding affinities influence the outcomes of protein-protein interactions between Aar, AggR, and H-NS via competition assays using SPR (127). We would immobilize either H-NS (figure 4.2E) or AggR (figure 4.2F) to CM5 sensor chips and expose them to Aar until a steady state equilibrium is reached. The third interacting partner (AggR or H-NS) would then be added to the sensor chip, and the dissociation rate of Aar from the chip bound ligand measured. The dissociation rate will be directly proportional to

the strength of Aar interacting with the third analyte (127). Based on our proposed model (figure 4.1), we hypothesize that Aar will have a higher binding affinity for H-NS than for AggR. Lastly, we propose to investigate if Aar is able to displace AggR and H-NS from DNA. Using SPR, we would immobilize either the *aggR* or *hns* promoter on sensor chips and add increasing concentrations of the promoter's corresponding protein to first determine the binding affinities of AggR (figure 4.2H) and H-NS (figure 4.2G) for their respective promoters (27, 99, 127, 128). We propose to follow these experiments by adding increasing concentrations of Aar to the immobilized promoters with respective protein at steady state equilibrium and measure dissociation rate of AggR (figure 4.2J) and H-NS (figure 4.2I) from DNA. We hypothesize that Aar can readily remove H-NS from DNA but not AggR.

We are still unsure how H-NS is acting to silence genes in the AggR regulon, i.e. whether it is binding to the structural genes themselves thereby impeding transcription. As H-NS has no sequence specificity and only binds to curved DNA associated with AT-rich DNA stretches (89, 91, 101, 129), it is possible that H-NS binds to both the promoter and the structural gene of *aafA* and other AggR-regulated genes which are AT-rich at both sites (table 4.1) (Note: these results may affect the DNA sequence we use for SPR experiments described above). We propose to identify how H-NS is silencing the AT-rich gene, *aafA*, by constructing a system in which we use the *lacZ* gene and promoter to screen for differences in β-galactosidase activity in the presence and absence of H-NS. The following constructs are proposed: $P_{lacZ}$ controlling the expression of *aafA* fused with *lacZ*

($P_{lacZ}$*aafA::lacZ*), $P_{aafA}$ controlling the expression of *lacZ* ($P_{aafA}$*lacZ*), and as a control $P_{lacZ}$ controlling the expression of *lacZ* ($P_{lacZ}$*lacZ*) (figure 4.3). If H-NS is silencing at the level of the promoter, then strains harboring $P_{aafA}$*lacZ* will have higher β-galactosidase activity when H-NS is not present compared to when H-NS is present. If H-NS is acting at the level of the structural gene, then strains harboring $P_{lacZ}$*aafA::lacZ* will have higher β-galactosidase activity when H-NS is not present. β-galactosidase activity should not be affected by the presence or absence of H-NS in strains harboring $P_{lacZ}$*lacZ*. Finally, H-NS could be (and likely is) affecting expression at both the promoter and structural gene level, so it is likely that both constructs harboring either component would be silenced in the presence of H-NS (figure 4.3).

We have only briefly investigated the residues on Aar that are specific for binding to AggR, but not for binding to H-NS (31, 58). Understanding where AggR and H-NS are binding to Aar will provide further insight into how Aar is regulating gene expression (i.e. are the binding sites overlapping and therefore preventing Aar/AggR/H-NS complexes, or are the binding sites different and therefore it is possible for the formation of Aar/AggR/H-NS complexes). Bacterial two-hybrid assays and complementation assays with derivatives of Aar suggest that α-helices 2 and 3 are necessary for Aar activity on AggR and for Aar oligomerization (31). We propose to utilize the previously synthesized Aar derivatives that were constructed to identify AggR binding sites to identify residues within Aar that are necessary for Aar/H-NS binding.

The mechanism that is least defined is if Aar is preferentially targeting H-NS binding to genes of the AggR regulon (horizontally transferred genes) and, if so, how. We would identify common genes targeted by both Aar and H-NS by transcriptomic analysis by RNA-seq. We hypothesize that, similar to Hha, Aar preferentially targets horizontally transferred genes, as ANRs are only present in pathogens which harbor pathogenicity islands/virulence genes that are AT-rich and horizontally acquired. For example, Hha and its paralogue YdgT, are small proteins that bind to H-NS, but not to DNA, to modulate H-NS silencing of horizontally acquired genes (130). In *Salmonella*, microarray analysis revealed that Hha and YdgT contribute to silencing the *Salmonella* pathogenicity islands (SPIs) and did not affect the silencing of typical ancestral genes that contain short AT-rich stretches (131). Thus, Hha and YdgT may be silencing large stretches of AT-rich DNA by facilitating the stable formation of extended nucleoprotein filaments (131). Only 5% of H-NS regulated genes are also regulated by Hha, suggesting that in addition to curved AT-rich stretches, some H-NS modulated genes must exhibit other structural motifs that are required for H-NS-Hha heteromeric complexes to specifically modulate (130). Although Aar counter-silences H-NS rather than strengthening H-NS silencing (as is the case with Hha), in EAEC, it is possible that Aar preferentially binds larger stretches of H-NS oligomers associated with horizontally transferred genes, or identifies a currently unknown structural motif, thus leading to our hypothesized Aar specificity for H-NS binding to AggR-regulated genes. The transcriptomic analysis by RNA-seq will

allow for a comparison of genes regulated by both Aar and H-NS, which we hypothesize will have large AT-rich stretches.

The work presented in this dissertation provides evidence that Aar is affecting gene expression and bacterial fitness through genes associated with virulence, and genes associated with core metabolic functions. We anticipate that the described future directions will provide new insights into the mechanism by which Aar is able to act as both an anti-activator and an anti-repressor on gene expression. Better understanding of this mechanism will broaden our understanding of how bacteria evolve to fine-tune the integration and expression of virulence genes with the expression of necessary survival genes.

**Implications to the field**

Aar modifies gene expression via a novel mechanism: previously, a regulator of gene expression that acts as an anti-activator on one regulon and an anti-repressor on another had not been identified. Anti-repression or counter-silencing of H-NS at virulence genes regulated by members of the AraC transcriptional regulators was thought to occur solely through the sequence-specific member of the AraC transcriptional regulators via competition for DNA-binding (102). Therefore, our work identifies a novel regulator that initially works in concert with the AraC transcriptional regulator AggR in EAEC strain 042 to counter-silence H-NS repression of AT-rich genes under AggR regulation. In

addition to aiding AggR in counter-silencing, our research supports the initial characterization of Aar acting as an anti-activator of AggR. Our work is unique in that we demonstrated that expression of *aar* in a basic *E. coli* K12 strain was able to affect core gene expression in a manner similar to that in pathogenic *E. coli*; thus, ruling out the need for a pathogen-specific intermediary protein.

We found that AggR is critical for optimal EAEC fitness in mice, which is evidence that, despite having evolved as a human-specific pathogen, AggR regulates a non-host-specific colonization factor. There are still numerous uncharacterized AggR-regulated proteins, one of which is a putative transcriptional regulator and others include hypothetical proteins (32). Although our *in vivo* mouse model did not show a role for any individual AggR-regulated virulence factor, we cannot rule out the possibility that one or more of the factors investigated are essential for EAEC persistence and pathogenicity in humans.

Additionally, we have shown a novel role for Aar in bacterial fitness at the level of the core genome. These results are not only important to our understanding of EAEC virulence but may also provide insight into how bacteria fine-tune the expression of horizontally transferred genes and “housekeeping” metabolic genes. Many Gram-negative pathogens encode a member of the ANR family (30, 31); therefore, ANRs are likely a common mechanism among pathogens to act as another level of regulation of horizontally transferred genes in coordination with H-NS like proteins.

In summary, our work supports the previously published hypothesis that Aar is regulating gene expression through interactions with both AggR and H-NS. In addition, our data provide the first evidence that Aar is affecting bacterial fitness via the repression of virulence factors and the induction of core genes. Further characterization of this novel dual mechanism could provide insight into other evolutionary adaptations that pathogens are acquiring to ensure their survival and propagation, and could lead to new therapeutic interventions.

**Table 4.1. AT content of AggR-regulated promoters and genes**

| Gene | Promoter AT % | Gene AT % |
|---|---|---|
| *aggR* | 57 | 71 |
| *aafA* | 69 | 61 |
| *aafD* | 64 | 64 |
| *aap* | 66 | 60 |
| *aar* | 61 | 65 |
| *aat* operon | 69 | 70 |
| *aai* operon | 62 | 74 |

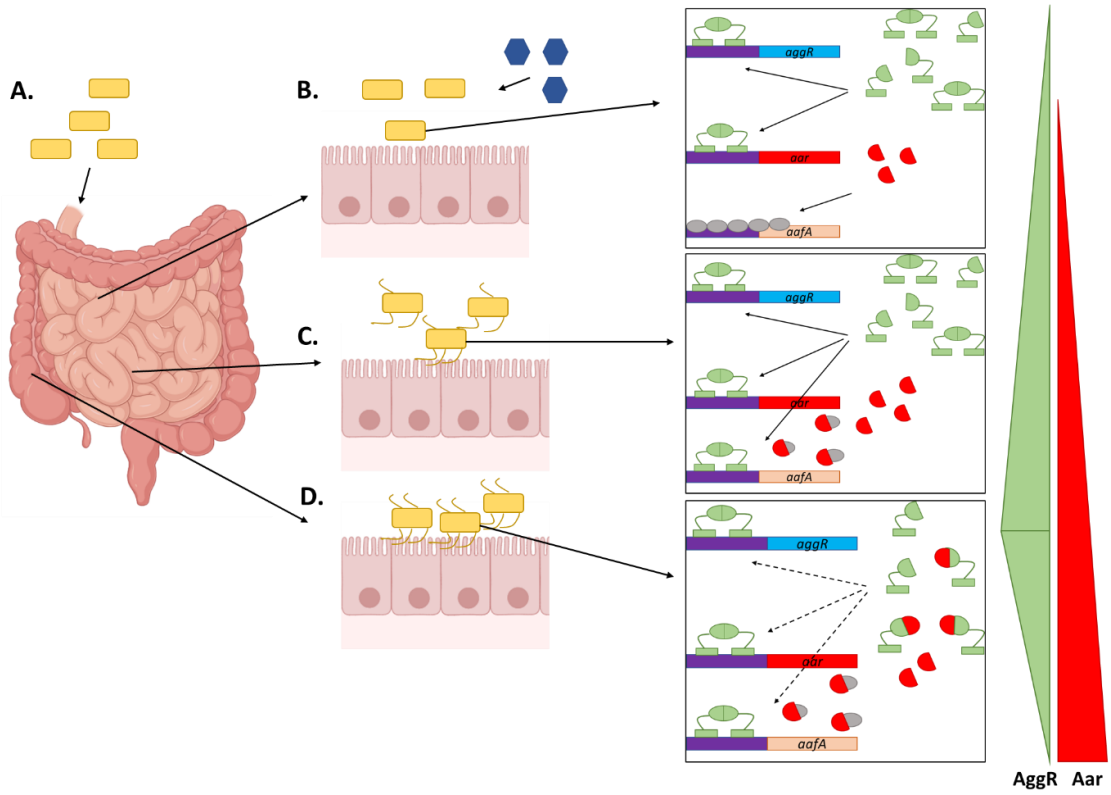



**Figure 4.1. Proposed mechanism of Aar regulation in EAEC.** A) Bacteria enter the small intestines and B) encounter an inducing signal (glucose/maltose – blue hexagon) that activates the expression of *aggR*. Production of AggR (green lamp figures) leads to the autoactivation of its own expression, as well as the expression of the negative regulator, *aar* (red half circles). As the concentration of Aar is initially low, there is only enough Aar to bind to H-NS (grey ovals) at AT-rich promoters (P*aafA*) and remove H-NS repression. C) Removal of H-NS allows for immediate binding of AggR at AggR-regulated promoters and a rapid increase in the expression of AT-rich genes (*aafA* resulting in increased fimbrial expression). During this time AggR is still autoactivating its own expression and the expression of *aar* (continued increase in the concentration of AggR and Aar). D) The continued

increase in the concentration of Aar results in Aar/AggR binding and inhibition of AggR autoactivation.

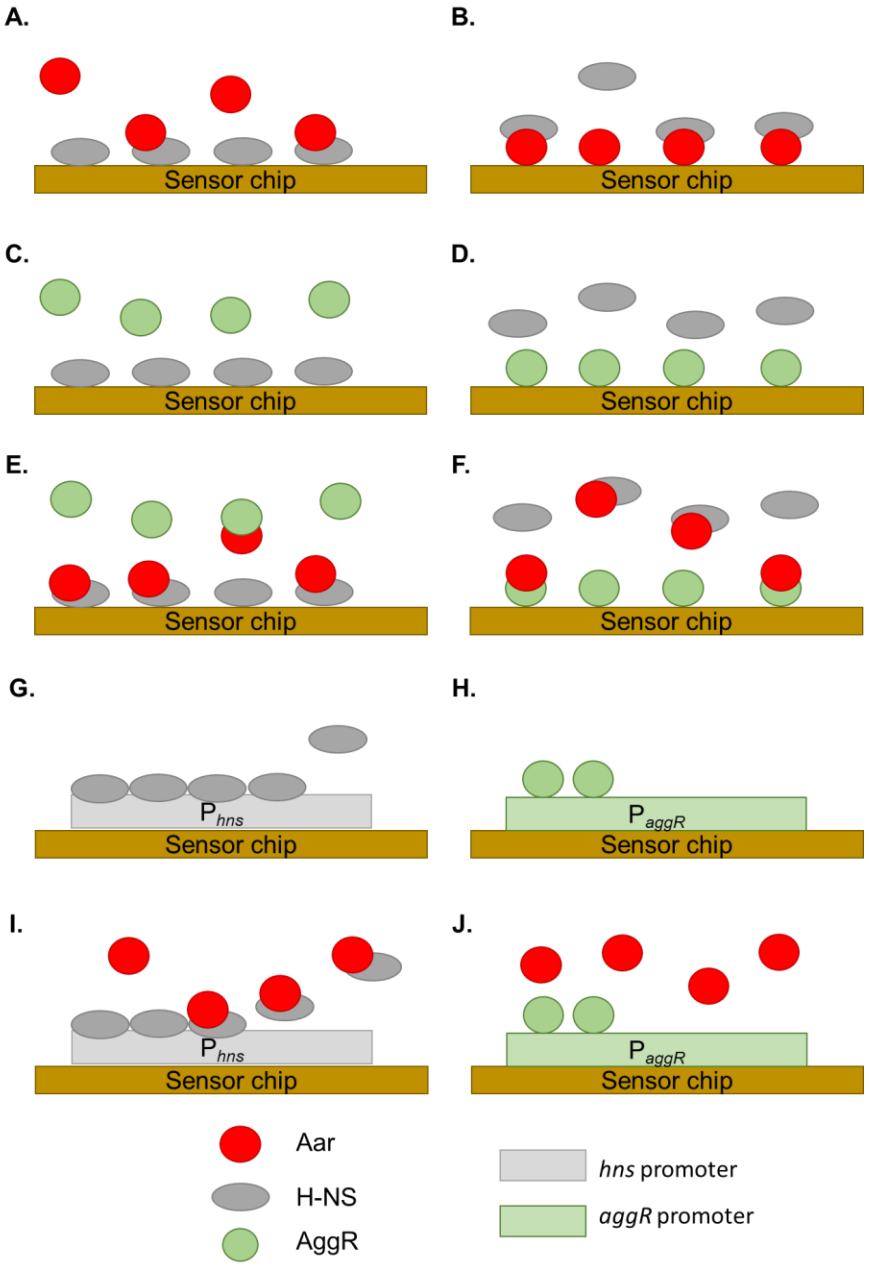
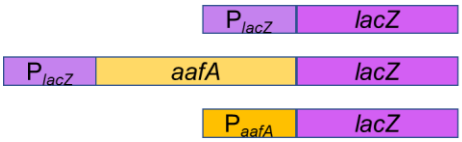



**Figure 4.2. SPR strategy to determine binding affinities and dissociation rates between proteins.** The binding affinity of Aar/H-NS will be measured using SPR when A) H-NS (ligand) (grey oval) is bound to the sensor chip while Aar (analyte) (red circle) flows over the bound H-NS and when B) Aar (ligand) is bound

to the sensor chip while H-NS (analyte) flows over the bound Aar. The binding affinity of AggR/H-NS will be measured using SPR when C) H-NS is bound to the sensor chip while AggR (analyte) (green circle) flows over the bound H-NS and when D) AggR is bound to the sensor chip while H-NS flows over the bound AggR. The dissociation rate of Aar bound to E) H-NS will be measured after the addition of AggR and to F) AggR after the addition of H-NS. The binding affinities of G) H-NS and H) AggR to their respective promoters bound to the sensor chip will be measured. The dissociation rate of I) H-NS and J) AggR from their respective promoters after the addition of Aar will be measured.




| β-galactosidase activity | | | |
|---|---|---|---|
| If H-NS silences region | | If H-NS does not silence region | |
| WT | Δ*hns* | WT | Δ*hns* |
| ++ | ++ | ++ | ++ |
| + | ++ | ++ | ++ |
| + | ++ | ++ | ++ |

**Figure 4.3. Potential outcomes for determining where H-NS is silencing AT-rich gene expression – promoter or coding region.** WT DH5α or DH5αΔ*hns* will be transformed with a plasmid encoding $P_{lacZ}$ controlling the expression of *lacZ* (positive control), $P_{lacZ}$ controlling the expression of *aafA* fused with *lacZ* (AT-rich gene construct), or $P_{aafA}$ controlling the expression of *lacZ* (AT-rich promoter construct). β-galactosidase activity will be measured in the WT and *hns* mutant strains for each construct to determine if H-NS is affecting the level of *lacZ* expression. If H-NS is silencing *lacZ* expression then WT β-galactosidase activity will be lower in the WT strain compared to the *hns* mutant.

# Appendix One

## Aar modulates the abundance and timing of AAF/II expression

## Abstract

Enteroaggregative *Escherichia coli* (EAEC) utilize a positive auto-activator, AggR, to activate the timely expression of virulence factors. In addition, AggR regulates the expression of its negative regulator, Aar. Mathematical modeling suggests contrasting behavior for auto-activating and auto-repressing promoters. We sought to characterize the behavior of the AggR-Aar system to test these hypothetical models. The abundance and timing of AAF/II expression was interrogated in WT 042 and 042Δ*aar* using flow cytometry. The mean fluorescent intensity and the percentage of bacteria expression AAF/II was calculated for each strain every hour for 12 hours. The mean fluorescent intensity demonstrated that WT 042 expressed AAF/II earlier during induction than 042Δ*aar*; however, this effect was lost at later time points when 042Δ*aar* expressed for AAF/II than WT 042. This supports the mathematical model that suggests that the presence of a negative auto-regulator allows for a quicker response time to environmental changes. Although, 042Δ*aar* generally expressed more fimbria per bacteria, the percentage of bacteria expressing AAF/II revealed that at later time points (>8hr) WT 042 had more bacteria expressing AAF/II (homogenous population) than 042Δ*aar.* Collectively, this data supports the role for a negative auto-regulator in bacteria based on a mathematical model.

## Introduction

Many Gram-negative pathogenic bacteria utilize a member of the AraC family of transcriptional regulators to respond to environmental changes and regulate the expression of virulence factors such as fimbria (21, 33, 113, 132–136). These AraC homologs are often partnered with a member from the AraC negative regulator (ANR) family to regulate their expression (30, 31). In enteroaggregative *Escherichia coli* (EAEC), an *E. coli* pathotype associated with childhood growth impairment (7, 8, 10, 11), the AraC homolog, AggR, regulates at least 44 virulence genes including genes encoding the aggregative adherence fimbria (*aafDA*) (32, 35). AggR and its regulon are downregulated by the ANR homolog, Aar (30).

The role of a positive and negative auto-regulator has been described with the use of mathematical modeling and experimentally using artificial genetic circuits (137–139). According to mathematical modeling, auto-regulators decrease the response time by producing bi-stability, ensuring that a subpopulation has a selective advantage to specific environmental changes and will survive (137, 140, 141). In contrast, the role of a negative autoregulator (i.e. Aar) is to decrease cell-to-cell fluctuations leading to a more homogenous population (139). Although the population is more homogenous, negative auto-regulators also enable a rapid response to environmental changes (137).

The combination of both a positive and negative auto-regulator has not been extensively studied. We hypothesize that Aar will decrease bacterial response time to an environmental change and lead to a more homogenous

population. Here we sought to better understand how this AggR-Aar pair in EAEC regulates virulence gene expression.

## Materials and Methods

**Bacterial strains, and growth conditions.** Bacterial strains used in this study were EAEC strain 042 (48) and 042Δ*aar* (30). Bacteria were grown in Luria Broth (LB) and Dulbecco's modified Eagle's medium with 0.4% glucose (DMEM high glucose) (Gibco, Grand Island, NY) as previously described (32).

**Flow cytometry for AafA.** For flow cytometry, a modified protocol was followed. Briefly, EAEC strain 042 and 042Δ*aar* were grown aerobically in LB overnight at 37˚C with shaking and then diluted 1:100 in 500ml DMEM high glucose and grown at 37˚C with shaking. Bacterial aliquots of the following volumes were collected at specified time points: 100ml at 1hr and 2hr; 50ml at 3hr; 30ml at 4hr; 15ml at 5hr, 6hr, and 7hr; and 10ml at 8hr, 9hr, 10hr, 11hr, 12hr, and 24hr. Bacterial aliquots were centrifuged at 4,000rpm for 20 minutes, pellets were washed in 500µl PBS and centrifuged at 14,000rpm for 5 minutes. Bacterial pellets were suspended in 1ml 1% paraformaldehyde (PFA).

For antibody staining, the following volumes of fixed bacteria were added to one of the 1mL wells in a 96-well plate (USA Scientific): 200µl of 1-5hr fixed bacteria with 800µl PBS; 100µl of 6-12hr fixed bacteria with 900µl PBS; and 50µl of 24hr fixed bacteria with 950µl PBS. Centrifuged at 3,250rpm for 10 minutes and

carefully vacuumed out supernatant. Resuspended bacteria in 500µl of 10% horse serum (Thermofisher) in PBS and incubated at 4˚C for 30 minutes with continuous rotation. Centrifuged at 3,250 rpm for 10 minutes and carefully vacuumed out supernatant. Resuspended in 500µl of 10% horse serum and 1:1000 AafA antibody and incubated overnight in the dark at 4˚C with continuous rotation. Centrifuged bacteria at 3,250rpm for 10 minutes and washed bacteria twice with 10% horse serum. Resuspended bacteria in 500µl 10% horse serum and 1:1000 FITC goat anti-rabbit Ig (BD Biosciences), wrapped in foil, and incubated at 4˚C with continuous rotation for 1.5hr. Centrifuged at 3,250rpm for 10 minutes and washed twice with 10% horse serum. Resuspended bacteria in 1ml PBS and ran flow.

## Results

**AafA is more abundantly expressed on 042Δ*aar* compared to WT 042 at later time points.** Initial characterization of 042Δ*aar* described it as hyper-fimbriated compared to WT 042 (30), therefore we hypothesized that Aar is modulating the expression (abundance) of virulence factors. However, we additionally hypothesized that the expression of *aar,* the negative auto-regulator, in WT 042 would result in earlier expression of virulence factors. We used flow cytometry to determine the amount of AafA, the major subunit of the surface exposed aggregative adherence fimbria (AAF), on individual bacteria. As predicted, the mean fluorescent intensity (MFI) was higher in 042Δ*aar* than in WT 042 but only

after 6hr post-induction (figure A1.1). For the first 4hr, the MFI was similar in both strains; however, again as predicted there was a higher MFI in WT 042 at 5hr compared to 042Δ*aar* (figure A1.1).

**WT 042 has more individual bacteria expressing AafA than 042Δ*aar.*** Mathematical modeling suggests that a positive auto-regulator leads to a bimodal population; a negative auto-regulator leads to a homogenous population (137, 139–141). We hypothesized that Aar is altering on/off states. We used flow cytometry to determine the percentage of bacterial cells expressing AafA. WT 042 was found to have a higher percentage of bacteria expressing AafA compared to 042Δ*aar* starting at the first hour (figure A1.2). In addition, the maximum percentage of bacteria expressing AafA was reached by 7 hours post-induction and this population distribution was maintained for 24 hours (figure A1.2).

## Discussion

AraC-ANR pairs have evolved to work together and tightly regulate virulence gene expression in response to an environmental change. Previous data suggested that ANR homologs regulate the abundance of virulence factors expressed (30); however, mathematical models suggest that the expression of a negative auto-regulator has a role in the timing of gene expression and population dynamics (137, 139). We were able to investigate how Aar is regulating gene expression by using flow cytometry to measure hourly AafA expression.

AafA expression was more abundant on individual bacterium when *aar* was not expressed; supporting the previously published hypothesis that Aar is regulating the abundance of virulence factors (30). Additionally, by collecting hourly samples of the bacteria after induction in DMEM high glucose, we were able to examine the kinetics of AafA expression which suggested that AafA expression peaked first in WT 042. This observation was in agreement with mathematical modeling which states that the expression of a negative auto-regulator allows for a quicker response to environmental changes (137).

The expression of Aar did affect the percentage of the bacterial population that expressed AafA, leading to a higher percentage of AafA positive bacteria and thus a more homogenous population. This supported the role for Aar predicted by mathematical modeling (139). Although WT 042 had a higher percentage of bacteria expressing AafA, the difference between WT 042 and 042Δ*aar* (approximately 10%) may not be biologically significant. Interestingly, this tempts us to speculate that when a positive auto-regulator and a negative auto-regulator are present in the same bacterium that the phenotype associated with the presence of a negative auto-regulator over-rides that of the positive auto-regulator.

Aar 1) modulates the expression (abundance) of virulence factors; 2) modulates the timing of virulence gene expression, allowing for a quicker response to environmental changes; and 3) modulates population dynamics. Individually, these provide an evolutionary advantage to the bacterium. By reducing the abundance of virulence factors, Aar is decreasing the likelihood of recognition by

the host immune system and ensures the continued health of the host for future propagation. In addition, bacterial survival depends on the bacterium's ability to respond to environmental changes; by allowing for a quicker response time, Aar is enhancing the bacterium's likelihood of survival.

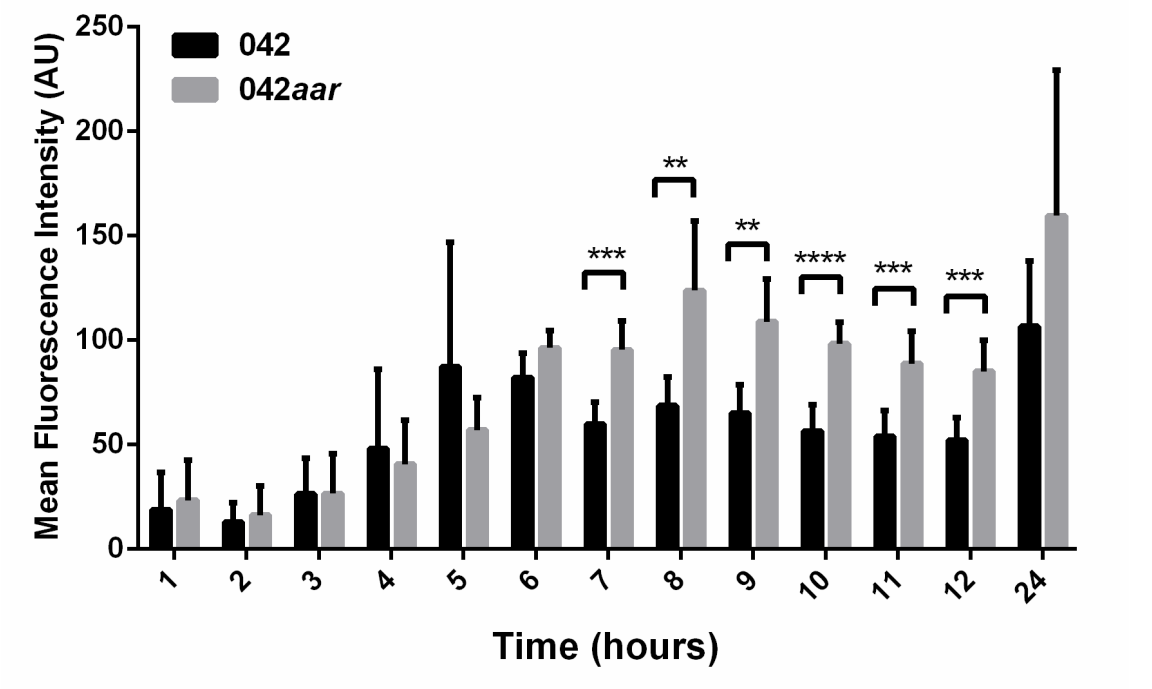



**Figure A1.1. Abundance of AAF/II on WT 042 and 042Δ*aar*.** WT 042 (black bars) and 042Δ*aar* (grey bars) were grown in DMEM high glucose at 37˚C, shaking and aliquots were collected every hour for the first 12 hours and then at 24 hours post induction. Bacteria were fixed and stained for AafA. Flow cytometry was used to determine the abundance of anti-FITC on individual bacterium. Flow data are representative of at least three independent experiments. Asterisks indicate significant differences by a two-tailed paired T-test (*, $P<0.05$; **, $P<0.005$; ***, $P<0.0005$).

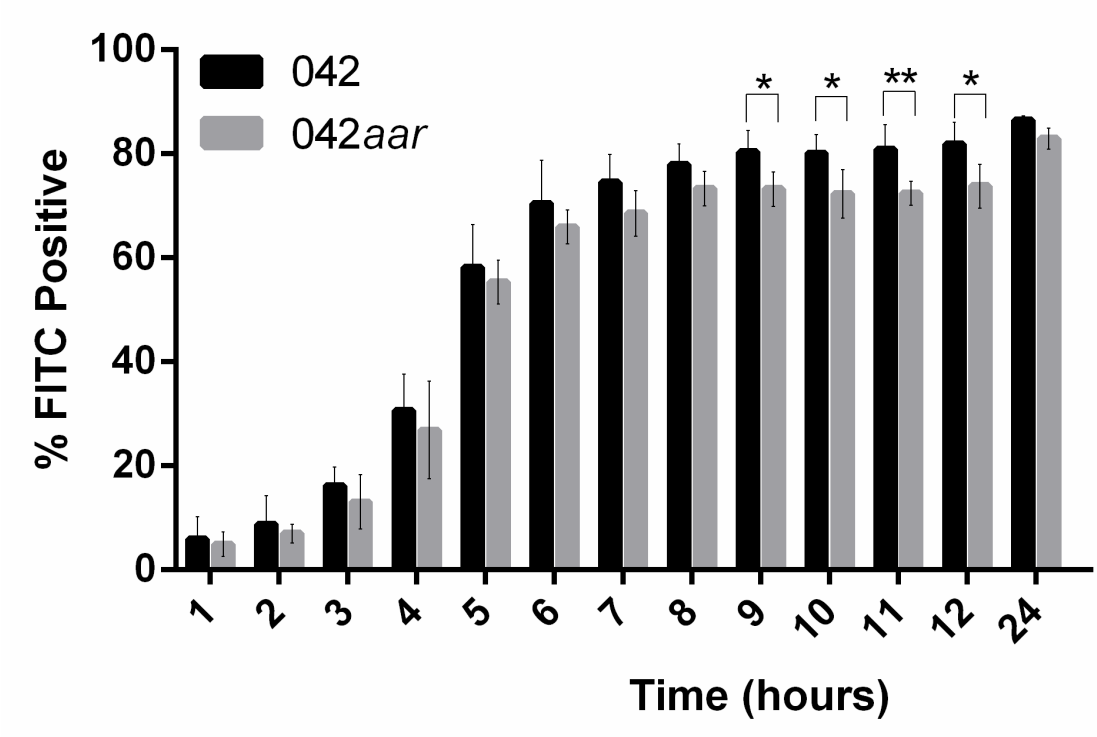



**Figure A1.2. Percentage of bacteria expressing fimbria in cultures of WT 042 and 042Δ*aar*.** WT 042 (black bars) and 042Δ*aar* (grey bars) were grown in DMEM high glucose at 37˚C, shaking and aliquots were collected every hour for the first 12 hours and then at 24 hours post induction. Bacteria were fixed and stained for AafA. Flow cytometry was used to determine the percentage of bacteria expressing anti-FITC. Flow data are representative of at least three independent experiments. Asterisks indicate significant differences by a two-tailed paired T-test (*, P<0.05; **, P<0.005; ***, P<0.0005).

## Citations


1. Nataro JP, Kaper JB, Robins-Browne R, Prado V, Vial P, Levine MM. 1987. Patterns of adherence of diarrheagenic Escherichia coli to HEp-2 cells. Pediatr Infect Dis J 6:829–831.

2. Gupta D, Sharma M, Sarkar S, Thapa BR, Chakraborti A. 2016. Virulence determinants in enteroaggregative Escherichia coli from North India and their interaction in in vitro organ culture system. FEMS Microbiol Lett 363.

3. Huang DB, Nataro JP, DuPont HL, Kamat PP, Mhatre AD, Okhuysen PC, Chiang T. 2006. Enteroaggregative Escherichia coli is a cause of acute diarrheal illness: a meta-analysis. Clin Infect Dis 43:556–563.

4. Cohen MB, Nataro JP, Bernstein DI, Hawkins J, Roberts N, Staat MA. 2005. Prevalence of diarrheagenic Escherichia coli in acute childhood enteritis: a prospective controlled study. J Pediatr 146:54–61.

5. Boll EJ, Struve C, Boisen N, Olesen B, Stahlhut SG, Krogfelt KA. 2013. Role of enteroaggregative Escherichia coli virulence factors in uropathogenesis. Infect Immun 81:1164–1171.

6. Herzog K, Engeler Dusel J, Hugentobler M, Beutin L, Sägesser G, Stephan R, Hächler H, Nüesch-Inderbinen M. 2014. Diarrheagenic enteroaggregative Escherichia coli causing urinary tract infection and bacteremia leading to sepsis. Infection 42:441–444.

7. Lima IFN, Boisen N, Quetz J da S, Havt A, de Carvalho EB, Soares AM, Lima NL, Mota RMS, Nataro JP, Guerrant RL, Lima AÂM. 2013. Prevalence of enteroaggregative Escherichia coli and its virulence-related genes in a case-control study among children from north-eastern Brazil. J Med Microbiol 62:683–693.

8. Lima AAM, Medeiros PHQS, Havt A. 2018. Enteroaggregative Escherichia coli subclinical and clinical infections. Curr Opin Infect Dis 31:433–439.

9. Steiner TS, Lima AA, Nataro JP, Guerrant RL. 1998. Enteroaggregative Escherichia coli produce intestinal inflammation and growth impairment and cause interleukin-8 release from intestinal epithelial cells. J Infect Dis 177:88–96.

10. Acosta GJ, Vigo NI, Durand D, Riveros M, Arango S, Zambruni M, Ochoa TJ. 2016. Diarrheagenic Escherichia coli: Prevalence and Pathotype Distribution in Children from Peruvian Rural Communities. Am J Trop Med Hyg 95:574–579.

11. Havt A, Lima IF, Medeiros PH, Clementino MA, Santos AK, Amaral MS, Veras HN, Prata MM, Lima NL, Di Moura A, Leite ÁM, Soares AM, Filho JQ, Houpt ER, Nataro JP, Guerrant RL, Lima AA. 2017. Prevalence and virulence gene profiling of enteroaggregative Escherichia coli in malnourished and nourished Brazilian children. Diagn Microbiol Infect Dis 89:98–105.

12. Nataro JP. 2005. Enteroaggregative Escherichia coli pathogenesis. Curr Opin Gastroenterol 21:4–8.

13. Huang DB, Mohanty A, DuPont HL, Okhuysen PC, Chiang T. 2006. A review of an emerging enteric pathogen: enteroaggregative Escherichia coli. J Med Microbiol 55:1303–1311.

14. Scavia G, Staffolani M, Fisichella S, Striano G, Colletta S, Ferri G, Escher M, Minelli F, Caprioli A. 2008. Enteroaggregative Escherichia coli associated with a foodborne outbreak of gastroenteritis. J Med Microbiol 57:1141–1146.

15. Okhuysen P, DuPont H. 2010. Enteroaggregative*Escherichia coli* (EAEC): A Cause of Acute and Persistent Diarrhea of Worldwide Importance. J Infect Dis 202:503–505.

16. Harrington SM, Strauman MC, Abe CM, Nataro JP. 2005. Aggregative adherence fimbriae contribute to the inflammatory response of epithelial cells infected with enteroaggregative Escherichia coli. Cell Microbiol 7:1565–1578.

17. Chaudhuri RR, Sebaihia M, Hobman JL, Webber MA, Leyton DL, Goldberg MD, Cunningham AF, Scott-Tucker A, Ferguson PR, Thomas CM, Frankel G, Tang CM, Dudley EG, Roberts IS, Rasko DA, Pallen MJ, Parkhill J, Nataro JP, Thomson NR, Henderson IR. 2010. Complete genome sequence and comparative metabolic profiling of the prototypical enteroaggregative Escherichia coli strain 042. PLoS One 5:e8801.

18. Hüttener M, Prieto A, Espelt J, Bernabeu M, Juárez A. 2018. Stringent Response and AggR-Dependent Virulence Regulation in the Enteroaggregative Escherichia coli Strain 042. Front Microbiol 9:717.

19. Henderson IR, Hicks S, Navarro-Garcia F, Elias WP, Philips AD, Nataro JP. 1999. Involvement of the enteroaggregative Escherichia coli plasmid-encoded toxin in causing human intestinal damage. Infect Immun 67:5338–5344.

20. Rajan A, Vela L, Zeng X-L, Yu X, Shroyer N, Blutt SE, Poole NM, Carlin LG, Nataro JP, Estes MK, Okhuysen PC, Maresso AW. 2018. Novel Segment- and Host-Specific Patterns of Enteroaggregative Escherichia coli Adherence to Human Intestinal Enteroids. MBio 9.

21. Yang J, Tauschek M, Robins-Browne RM. 2011. Control of bacterial virulence by AraC-like regulators that respond to chemical signals. Trends Microbiol 19:128–135.

22. Yang J, Hocking DM, Cheng C, Dogovski C, Perugini MA, Holien JK, Parker MW, Hartland EL, Tauschek M, Robins-Browne RM. 2013. Disarming bacterial virulence through chemical inhibition of the DNA binding domain of an AraC-like transcriptional activator protein. J Biol Chem 288:31115–31126.

23. Yang J, Hart E, Tauschek M, Price GD, Hartland EL, Strugnell RA, Robins-Browne RM. 2008. Bicarbonate-mediated transcriptional activation of divergent operons by the virulence regulatory protein, RegA, from Citrobacter rodentium. Mol Microbiol 68:314–327.

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

25. Schleif R. 2010. AraC protein, regulation of the l-arabinose operon in Escherichia coli, and the light switch mechanism of AraC action. FEMS Microbiol Rev 34:779–796.

26. Sheikh J, Hicks S, Dall'Agnol M, Phillips AD, Nataro JP. 2001. Roles for Fis and YafK in biofilm formation by enteroaggregative Escherichia coli. Mol Microbiol 41:983–997.

27. Morin N, Tirling C, Ivison SM, Kaur AP, Nataro JP, Steiner TS. 2010. Autoactivation of the AggR regulator of enteroaggregative*Escherichia coli in vitro* and*in vivo*. FEMS Immunol Med Microbiol 58:344–355.

28. Dorman CJ, Deighan P. 2003. Regulation of gene expression by histone-like proteins in bacteria. Curr Opin Genet Dev 13:179–184.

29. Ball CA, Osuna R, Ferguson KC, Johnson RC. 1992. Dramatic changes in Fis levels upon nutrient upshift in Escherichia coli. J Bacteriol 174:8043–8056.

30. Santiago AE, Ruiz-Perez F, Jo NY, Vijayakumar V, Gong MQ, Nataro JP. 2014. A large family of antivirulence regulators modulates the effects of

transcriptional activators in Gram-negative pathogenic bacteria. PLoS Pathog 10:e1004153.

31. Santiago AE, Yan MB, Tran M, Wright N, Luzader DH, Kendall MM, Ruiz-Perez F, Nataro JP. 2016. A large family of anti-activators accompanying XylS/AraC family regulatory proteins. Mol Microbiol 101:314–332.

32. Morin N, Santiago AE, Ernst RK, Guillot SJ, Nataro JP. 2013. Characterization of the AggR regulon in enteroaggregative Escherichia coli. Infect Immun 81:122–132.

33. Nataro JP, Yikang D, Yingkang D, Walker K. 1994. AggR, a transcriptional activator of aggregative adherence fimbria I expression in enteroaggregative Escherichia coli. J Bacteriol 176:4691–4699.

34. Harrington SM, Dudley EG, Nataro JP. 2006. Pathogenesis of enteroaggregative Escherichia coli infection. FEMS Microbiol Lett 254:12–18.

35. Elias WP, Czeczulin JR, Henderson IR, Trabulsi LR, Nataro JP. 1999. Organization of biogenesis genes for aggregative adherence fimbria II defines a virulence gene cluster in enteroaggregative Escherichia coli. J Bacteriol 181:1779–1785.

36. Berry AA, Yang Y, Pakharukova N, Garnett JA, Lee W-C, Cota E, Marchant J, Roy S, Tuittila M, Liu B, Inman KG, Ruiz-Perez F, Mandomando I, Nataro JP, Zavialov AV, Matthews S. 2014. Structural Insight into Host Recognition by Aggregative Adherence Fimbriae of Enteroaggregative Escherichia coli. PLoS Pathog 10:e1004404.

37. Boisen N, Struve C, Scheutz F, Krogfelt KA, Nataro JP. 2008. New adhesin of enteroaggregative Escherichia coli related to the Afa/Dr/AAF family. Infect Immun 76:3281–3292.

38. Jønsson R, Struve C, Boisen N, Mateiu RV, Santiago AE, Jenssen H, Nataro JP, Krogfelt KA. 2015. Novel aggregative adherence fimbria variant of enteroaggregative Escherichia coli. Infect Immun 83:1396–1405.

39. Jønsson R, Struve C, Boll EJ, Boisen N, Joensen KG, Sørensen CA, Jensen BH, Scheutz F, Jenssen H, Krogfelt KA. 2017. A Novel pAA Virulence Plasmid Encoding Toxins and Two Distinct Variants of the Fimbriae of Enteroaggregative Escherichia coli. Front Microbiol 8:263.

40. Czeczulin JR, Balepur S, Hicks S, Phillips A, Hall R, Kothary MH, Navarro-Garcia F, Nataro JP. 1997. Aggregative adherence fimbria II, a second

fimbrial antigen mediating aggregative adherence in enteroaggregative Escherichia coli. Infect Immun 65:4135–4145.

41. Bernier C, Gounon P, Le Bouguénec C. 2002. Identification of an aggregative adhesion fimbria (AAF) type III-encoding operon in enteroaggregative Escherichia coli as a sensitive probe for detecting the AAF-encoding operon family. Infect Immun 70:4302–4311.

42. Jønsson R, Liu B, Struve C, Yang Y, Jørgensen R, Xu Y, Jenssen H, Krogfelt KA, Matthews S. 2017. Structural and functional studies of Escherichia coli aggregative adherence fimbriae (AAF/V) reveal a deficiency in extracellular matrix binding. Biochim Biophys Acta Proteins Proteom 1865:304–311.

43. Farfan MJ, Inman KG, Nataro JP. 2008. The major pilin subunit of the AAF/II fimbriae from enteroaggregative Escherichia coli mediates binding to extracellular matrix proteins. Infect Immun 76:4378–4384.

44. Walia B, Castaneda FE, Wang L, Kolachala VL, Bajaj R, Roman J, Merlin D, Gewirtz AT, Sitaraman SV. 2004. Polarized fibronectin secretion induced by adenosine regulates bacterial-epithelial interaction in human intestinal epithelial cells. Biochem J 382:589–596.

45. Strauman MC, Harper JM, Harrington SM, Boll EJ, Nataro JP. 2010. Enteroaggregative Escherichia coli disrupts epithelial cell tight junctions. Infect Immun 78:4958–4964.

46. Boll EJ, Ayala-Lujan J, Szabady RL, Louissaint C, Smith RZ, Krogfelt KA, Nataro JP, Ruiz-Perez F, McCormick BA. 2017. Enteroaggregative Escherichia coli Adherence Fimbriae Drive Inflammatory Cell Recruitment via Interactions with Epithelial MUC1. MBio 8.

47. Sheikh J, Czeczulin JR, Harrington S, Hicks S, Henderson IR, Le Bouguénec C, Gounon P, Phillips A, Nataro JP. 2002. A novel dispersin protein in enteroaggregative Escherichia coli. J Clin Invest 110:1329–1337.

48. Nataro JP, Yikang D, Cookson S, Cravioto A, Savarino SJ, Guers LD, Levine MM, Tacket CO. 1995. Heterogeneity of Enteroaggregative “Escherichia coli” virulence demonstrated in volunteers. J Infect Dis 171:465–468.

49. Monteiro BT, Campos LC, Sircili MP, Franzolin MR, Bevilacqua LF, Nataro JP, Elias WP. 2009. The dispersin-encoding gene (aap) is not restricted to enteroaggregative Escherichia coli. Diagn Microbiol Infect Dis 65:81–84.

50. Schwarz S, Hood RD, Mougous JD. 2010. What is type VI secretion doing in all those bugs? Trends Microbiol 18:531–537.

51. Cascales E, Cambillau C. 2012. Structural biology of type VI secretion systems. Philos Trans R Soc Lond B, Biol Sci 367:1102–1111.

52. Silverman JM, Brunet YR, Cascales E, Mougous JD. 2012. Structure and regulation of the type VI secretion system. Annu Rev Microbiol 66:453–472.

53. Dudley EG, Thomson NR, Parkhill J, Morin NP, Nataro JP. 2006. Proteomic and microarray characterization of the AggR regulon identifies a pheU pathogenicity island in enteroaggregative Escherichia coli. Mol Microbiol 61:1267–1282.

54. Aschtgen M-S, Bernard CS, De Bentzmann S, Lloubès R, Cascales E. 2008. SciN is an outer membrane lipoprotein required for type VI secretion in enteroaggregative Escherichia coli. J Bacteriol 190:7523–7531.

55. Brunet YR, Bernard CS, Gavioli M, Lloubès R, Cascales E. 2011. An epigenetic switch involving overlapping fur and DNA methylation optimizes expression of a type VI secretion gene cluster. PLoS Genet 7:e1002205.

56. Brunet YR, Espinosa L, Harchouni S, Mignot T, Cascales E. 2013. Imaging type VI secretion-mediated bacterial killing. Cell Rep 3:36–41.

57. Douzi B, Brunet YR, Spinelli S, Lensi V, Legrand P, Blangy S, Kumar A, Journet L, Cascales E, Cambillau C. 2016. Structure and specificity of the Type VI secretion system ClpV-TssC interaction in enteroaggregative Escherichia coli. Sci Rep 6:34405.

58. Santiago AE, Yan MB, Hazen TH, Sauder B, Meza-Segura M, Rasko DA, Kendall MM, Ruiz-Perez F, Nataro JP. 2017. The AraC Negative Regulator family modulates the activity of histone-like proteins in pathogenic bacteria. PLoS Pathog 13:e1006545.

59. Snyder L, Champness W. 2007. Regulation of Gene Expression: Operons, p. 499–546. *In* Molecular genetics of bacteria, 3rd ed. ASM Press, Washington, DC.

60. Schaechter M, Ingraham JL, Neidhardt FC. 2006. Coordination and regulation, p. 221–247. *In* Microbe. ASM Press, Washington, DC.

61. Bervoets I, Charlier D. 2019. Diversity, versatility and complexity of bacterial gene regulation mechanisms: opportunities and drawbacks for applications in synthetic biology. FEMS Microbiol Rev 43:304–339.

62. Egan SM. 2002. Growing repertoire of AraC/XylS activators. J Bacteriol 184:5529–5532.

63. Lewis M. 2005. The lac repressor. C R Biol 328:521–548.

64. Yanofsky C. 2004. The different roles of tryptophan transfer RNA in regulating trp operon expression in E. coli versus B. subtilis. Trends Genet 20:367–374.

65. Yura T. 1996. Regulation and conservation of the heat-shock transcription factor sigma32. Genes Cells 1:277–284.

66. Schwiesow L, Lam H, Dersch P, Auerbuch V. 2015. Yersinia type III secretion system master regulator lcrf. J Bacteriol 198:604–614.

67. Yousef MR, Grundy FJ, Henkin TM. 2005. Structural transitions induced by the interaction between tRNA(Gly) and the Bacillus subtilis glyQS T box leader RNA. J Mol Biol 349:273–287.

68. Ryan GT, Wei Y, Winans SC. 2013. A LuxR-type repressor of Burkholderia cenocepacia inhibits transcription via antiactivation and is inactivated by its cognate acylhomoserine lactone. Mol Microbiol 87:94–111.

69. Thibault J, Faudry E, Ebel C, Attree I, Elsen S. 2009. Anti-activator ExsD forms a 1:1 complex with ExsA to inhibit transcription of type III secretion operons. J Biol Chem 284:15762–15770.

70. Yamaguchi S, Reid DA, Rothenberg E, Darwin AJ. 2013. Changes in Psp protein binding partners, localization and behaviour upon activation of the Yersinia enterocolitica phage shock protein response. Mol Microbiol 87:656–671.

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

72. Zheng Z, Fuqua C, Chen L. 2012. The quorum sensing transcriptional regulator TraR has separate binding sites for DNA and the anti-activator. Biochem Biophys Res Commun 418:396–401.

73. Chen G, Jeffrey PD, Fuqua C, Shi Y, Chen L. 2007. Structural basis for antiactivation in bacterial quorum sensing. Proc Natl Acad Sci USA 104:16474–16479.

74. Qin Y, Su S, Farrand SK. 2007. Molecular basis of transcriptional antiactivation. TraM disrupts the TraR-DNA complex through stepwise interactions. J Biol Chem 282:19979–19991.

75. Seet Q, Zhang L-H. 2011. Anti-activator QslA defines the quorum sensing threshold and response in Pseudomonas aeruginosa. Mol Microbiol 80:951–965.

76. Fan H, Dong Y, Wu D, Bowler MW, Zhang L, Song H. 2013. QsIA disrupts LasR dimerization in antiactivation of bacterial quorum sensing. Proc Natl Acad Sci USA 110:20765–20770.

77. Tauschek M, Yang J, Hocking D, Azzopardi K, Tan A, Hart E, Praszkier J, Robins-Browne RM. 2010. Transcriptional analysis of the grlRA virulence operon from Citrobacter rodentium. J Bacteriol 192:3722–3734.

78. Padavannil A, Jobichen C, Mills E, Velazquez-Campoy A, Li M, Leung KY, Mok YK, Rosenshine I, Sivaraman J. 2013. Structure of GrlR-GrlA complex that prevents GrlA activation of virulence genes. Nat Commun 4:2546.

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

80. Williams RM, Rimsky S. 1997. Molecular aspects of the E. coli nucleoid protein, H-NS: a central controller of gene regulatory networks. FEMS Microbiol Lett 156:175–185.

81. Atlung T, Ingmer H. 1997. H-NS: a modulator of environmentally regulated gene expression. Mol Microbiol 24:7–17.

82. Dilweg IW, Dame RT. 2018. Post-translational modification of nucleoid-associated proteins: an extra layer of functional modulation in bacteria? Biochem Soc Trans 46:1381–1392.

83. Dorman CJ. 2009. Chapter 2 Nucleoid-Associated Proteins and Bacterial Physiology, p. 47–64. *In* Elsevier.

84. Mangan MW, Lucchini S, Danino V, Cróinín TO, Hinton JCD, Dorman CJ. 2006. The integration host factor (IHF) integrates stationary-phase and virulence gene expression in Salmonella enterica serovar Typhimurium. Mol Microbiol 59:1831–1847.

85. Mangan MW, Lucchini S, Ó Cróinín T, Fitzgerald S, Hinton JCD, Dorman CJ. 2011. Nucleoid-associated protein HU controls three regulons that coordinate virulence, response to stress and general physiology in

Salmonella enterica serovar Typhimurium. Microbiology (Reading, Engl) 157:1075–1087.

86. Schröder O, Wagner R. 2000. The bacterial DNA-binding protein H-NS represses ribosomal RNA transcription by trapping RNA polymerase in the initiation complex. J Mol Biol 298:737–748.

87. Müller CM, Dobrindt U, Nagy G, Emödy L, Uhlin BE, Hacker J. 2006. Role of histone-like proteins H-NS and StpA in expression of virulence determinants of uropathogenic Escherichia coli. J Bacteriol 188:5428–5438.

88. Dorman CJ. 2013. Genome architecture and global gene regulation in bacteria: making progress towards a unified model? Nat Rev Microbiol 11:349–355.

89. Dame RT, Wyman C, Goosen N. 2001. Structural basis for preferential binding of H-NS to curved DNA. Biochimie 83:231–234.

90. Hommais F, Krin E, Laurent-Winter C, Soutourina O, Malpertuy A, Le Caer JP, Danchin A, Bertin P. 2001. Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS. Mol Microbiol 40:20–36.

91. Spurio R, Falconi M, Brandi A, Pon CL, Gualerzi CO. 1997. The oligomeric structure of nucleoid protein H-NS is necessary for recognition of intrinsically curved DNA and for DNA bending. EMBO J 16:1795–1805.

92. Bertin P, Hommais F, Krin E, Soutourina O, Tendeng C, Derzelle S, Danchin A. 2001. H-NS and H-NS-like proteins in Gram-negative bacteria and their multiple role in the regulation of bacterial metabolism. Biochimie 83:235–241.

93. Esposito D, Petrovic A, Harris R, Ono S, Eccleston JF, Mbabaali A, Haq I, Higgins CF, Hinton JCD, Driscoll PC, Ladbury JE. 2002. H-NS oligomerization domain structure reveals the mechanism for high order self-association of the intact protein. J Mol Biol 324:841–850.

94. Arold ST, Leonard PG, Parkinson GN, Ladbury JE. 2010. H-NS forms a superhelical protein scaffold for DNA condensation. Proc Natl Acad Sci USA 107:15728–15732.

95. Landick R, Wade JT, Grainger DC. 2015. H-NS and RNA polymerase: a love-hate relationship? Curr Opin Microbiol 24:53–59.

96. Will WR, Navarre WW, Fang FC. 2015. Integrated circuits: how transcriptional silencing and counter-silencing facilitate bacterial evolution. Curr Opin Microbiol 23:8–13.

97. Jordi BJ, Dagberg B, de Haan LA, Hamers AM, van der Zeijst BA, Gaastra W, Uhlin BE. 1992. The positive regulator CfaD overcomes the repression mediated by histone-like protein H-NS (H1) in the CFA/I fimbrial operon of Escherichia coli. EMBO J 11:2627–2632.

98. Nye MB, Pfau JD, Skorupski K, Taylor RK. 2000. Vibrio cholerae H-NS silences virulence gene expression at multiple steps in the ToxR regulatory cascade. J Bacteriol 182:4295–4303.

99. Falconi M, Higgins NP, Spurio R, Pon CL, Gualerzi CO. 1993. Expression of the gene encoding the major bacterial nucleotide protein H-NS is subject to transcriptional auto-repression. Mol Microbiol 10:273–282.

100. Ueguchi C, Mizuno T. 1993. The Escherichia coli nucleoid protein H-NS functions directly as a transcriptional repressor. EMBO J 12:1039–1046.

101. Rangarajan AA, Schnetz K. 2018. Interference of transcription across H-NS binding sites and repression by H-NS. Mol Microbiol 108:226–239.

102. Navarre WW, McClelland M, Libby SJ, Fang FC. 2007. Silencing of xenogeneic DNA by H-NS-facilitation of lateral gene transfer in bacteria by a defense system that recognizes foreign DNA. Genes Dev 21:1456–1471.

103. Winardhi RS, Gulvady R, Mellies JL, Yan J. 2014. Locus of enterocyte effacement-encoded regulator (Ler) of pathogenic Escherichia coli competes off histone-like nucleoid-structuring protein (H-NS) through noncooperative DNA binding. J Biol Chem 289:13739–13750.

104. Zhang A, Rimsky S, Reaban ME, Buc H, Belfort M. 1996. Escherichia coli protein analogs StpA and H-NS: regulatory loops, similar and disparate effects on nucleic acid dynamics. EMBO J 15:1340–1349.

105. Solórzano C, Srikumar S, Canals R, Juárez A, Paytubi S, Madrid C. 2015. Hha has a defined regulatory role that is not dependent upon H-NS or StpA. Front Microbiol 6:773.

106. Baños RC, Vivero A, Aznar S, García J, Pons M, Madrid C, Juárez A. 2009. Differential regulation of horizontally acquired and core genome genes by the bacterial modulator H-NS. PLoS Genet 5:e1000513.

107. Ali SS, Soo J, Rao C, Leung AS, Ngai DH-M, Ensminger AW, Navarre WW. 2014. Silencing by H-NS potentiated the evolution of Salmonella. PLoS Pathog 10:e1004500.

108. Erol I, Jeong K-C, Baumler DJ, Vykhodets B, Choi SH, Kaspar CW. 2006. H-NS controls metabolism and stress tolerance in Escherichia coli O157:H7 that influence mouse passage. BMC Microbiol 6:72.

109. Prieto A, Bernabeu M, Aznar S, Ruiz-Cruz S, Bravo A, Queiroz MH, Juárez A. 2018. Evolution of Bacterial Global Modulators: Role of a Novel H-NS Paralogue in the Enteroaggregative Escherichia coli Strain 042. mSystems 3.

110. Bernabeu M, Sánchez-Herrero JF, Huedo P, Prieto A, Hüttener M, Rozas J, Juárez A. 2019. Gene duplications in the E. coli genome: common themes among pathotypes. BMC Genomics 20:313.

111. Spurio R, Dürrenberger M, Falconi M, La Teana A, Pon CL, Gualerzi CO. 1992. Lethal overproduction of the Escherichia coli nucleoid protein H-NS: ultramicroscopic and molecular autopsy. Mol Gen Genet 231:201–211.

112. Mohamed JA, Huang DB, Jiang Z-D, DuPont HL, Nataro JP, Belkind-Gerson J, Okhuysen PC. 2007. Association of putative enteroaggregative Escherichia coli virulence genes and biofilm production in isolates from travelers to developing countries. J Clin Microbiol 45:121–126.

113. Porter ME, Mitchell P, Roe AJ, Free A, Smith DGE, Gally DL. 2004. Direct and indirect transcriptional activation of virulence genes by an AraC-like protein, PerA from enteropathogenic Escherichia coli. Mol Microbiol 54:1117–1133.

114. Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc Natl Acad Sci USA 97:6640–6645.

115. Anonymous. 1997. Applied Biosystems Prism 7700 sequence detection system: user bulletin no. 2. PerkinElmer Corp., Norwalk, CT.

116. Krin E, Danchin A, Soutourina O. 2010. Decrypting the H-NS-dependent regulatory cascade of acid stress resistance in Escherichia coli. BMC Microbiol 10:273.

117. Gutierrez C, Gordia S, Bonnassie S. 1995. Characterization of the osmotically inducible gene osmE of Escherichia coli K-12. Mol Microbiol 16:553–563.

118. Izquierdo M, Alvestegui A, Nataro JP, Ruiz-Perez F, Farfan MJ. 2014. Participation of integrin α5β1 in the fibronectin-mediated adherence of enteroaggregative Escherichia coli to intestinal cells. Biomed Res Int 2014:781246.

119. Boisen N, Scheutz F, Rasko DA, Redman JC, Persson S, Simon J, Kotloff KL, Levine MM, Sow S, Tamboura B, Toure A, Malle D, Panchalingam S, Krogfelt KA, Nataro JP. 2012. Genomic characterization of enteroaggregative Escherichia coli from children in Mali. J Infect Dis 205:431–444.

120. Harrington SM, Sheikh J, Henderson IR, Ruiz-Perez F, Cohen PS, Nataro JP. 2009. The Pic protease of enteroaggregative Escherichia coli promotes intestinal colonization and growth in the presence of mucin. Infect Immun 77:2465–2473.

121. Mickey AS, Nataro JP. 2020. Dual function of Aar, a member of the new AraC negative regulator family, in Escherichia coli gene expression. Infect Immun 88.

122. San Millan A, MacLean RC. 2017. Fitness costs of plasmids: a limit to plasmid transmission. Microbiol Spectr 5.

123. Krogfelt KA. 1991. Bacterial adhesion: genetics, biogenesis, and role in pathogenesis of fimbrial adhesins of Escherichia coli. Rev Infect Dis 13:721–735.

124. Schwarz R, Kaspar A, Seelig J, Künnecke B. 2002. Gastrointestinal transit times in mice and humans measured with 27Al and 19F nuclear magnetic resonance. Magn Reson Med 48:255–261.

125. Mazurkiewicz P, Tang CM, Boone C, Holden DW. 2006. Signature-tagged mutagenesis: barcoding mutants for genome-wide screens. Nat Rev Genet 7:929–939.

126. Philipson CW, Bassaganya-Riera J, Hontecillas R. 2013. Animal models of enteroaggregative Escherichia coli infection. Gut Microbes 4:281–291.

127. Drescher DG, Selvakumar D, Drescher MJ. 2018. Analysis of protein interactions by surface plasmon resonance. Adv Protein Chem Struct Biol 110:1–30.

128. Nguyen HH, Park J, Kang S, Kim M. 2015. Surface plasmon resonance: a versatile technique for biosensor applications. Sensors (Basel) 15:10481–10510.

129. Kahramanoglou C, Seshasayee ASN, Prieto AI, Ibberson D, Schmidt S, Zimmermann J, Benes V, Fraser GM, Luscombe NM. 2011. Direct and indirect effects of H-NS and Fis on global gene expression control in Escherichia coli. Nucleic Acids Res 39:2073–2091.

130. Madrid C, Balsalobre C, García J, Juárez A. 2007. The novel Hha/YmoA family of nucleoid-associated proteins: use of structural mimicry to modulate the activity of the H-NS family of proteins. Mol Microbiol 63:7–14.

131. Navarre WW. 2016. The impact of gene silencing on horizontal gene transfer and bacterial evolution. Adv Microb Physiol 69:157–186.

132. Hodson C, Yang J, Hocking DM, Azzopardi K, Chen Q, Holien JK, Parker MW, Tauschek M, Robins-Browne RM. 2017. Control of Virulence Gene Expression by the Master Regulator, CfaD, in the Prototypical Enterotoxigenic Escherichia coli Strain, H10407. Front Microbiol 8:1525.

133. Hart E, Yang J, Tauschek M, Kelly M, Wakefield MJ, Frankel G, Hartland EL, Robins-Browne RM. 2008. RegA, an AraC-like protein, is a global transcriptional regulator that controls virulence gene expression in Citrobacter rodentium. Infect Immun 76:5247–5256.

134. Munson GP. 2013. Virulence regulons of enterotoxigenic Escherichia coli. Immunol Res 57:229–236.

135. Weber GG, Klose KE. 2011. The complexity of ToxT-dependent transcription in Vibrio cholerae. Indian J Med Res 133:201–206.

136. Gonyar LA, Kendall MM. 2014. Ethanolamine and choline promote expression of putative and characterized fimbriae in enterohemorrhagic Escherichia coli O157:H7. Infect Immun 82:193–201.

137. Rosenfeld N, Elowitz MB, Alon U. 2002. Negative autoregulation speeds the response times of transcription networks. J Mol Biol 323:785–793.

138. Mitarai N, Jensen MH, Semsey S. 2015. Coupled positive and negative feedbacks produce diverse gene expression patterns in colonies. MBio 6.

139. Rosenfeld N, Young JW, Alon U, Swain PS, Elowitz MB. 2005. Gene regulation at the single-cell level. Science 307:1962–1965.

140. Eldar A, Elowitz MB. 2010. Functional roles for noise in genetic circuits. Nature 467:167–173.

141. To T-L, Maheshri N. 2010. Noise can induce bimodality in positive transcriptional feedback loops without bistability. Science 327:1142–1145.