Xenogeneic Silencer MvaT Differs in its DNA Binding Mechanism from Family Members H-NS and Lsr2

by

Grace Tong

A thesis submitted in conformity with the requirements
for the degree of Master of Science

Department of Molecular Genetics
University of Toronto

Xenogeneic Silencer MvaT Differs in its DNA Binding Mechanism from Family Members H-NS and Lsr2

Grace Tong

Master of Science

Department of Molecular Genetics
University of Toronto

2015

Abstract

Xenogeneic silencers (H-NS, Lsr2 and MvaT) facilitate bacterial evolution through the cryptic maintenance of horizontally acquired genes, binding and transcriptionally silencing DNA that is more AT-rich than the resident genome. H-NS and Lsr2 employ an AT-hook motif to recognize and bind AT-rich DNA targets. MvaT lacks this motif. Structural data suggested a KGGN motif may act as an analogous AT-hook. *In vitro* electrophoretic mobility shift assays found that double mutant MvaT K97A/N100A had similar DNA binding affinity and discrimination for target (AT-rich DNA) compared to wild type. Competitive gel shifts also found that while MvaT bound GC-rich and AT-rich DNA with comparable affinities, MvaT greatly prefers AT-rich DNA. We present a model where, within the oligomer, individual binding domains bind tightly to AT-rich DNA and loosely to GC-rich DNA. Together these data suggest that MvaT binds DNA through a novel mechanism dissimilar to xenogeneic silencing family members H-NS and Lsr2.

Acknowledgments

I would firstly like to acknowledge and thank my supervisor Dr. William Navarre for his support and guidance. He has been a wonderful mentor and advisor- keen for my success and invested in the education process. Thank you for being an integral part of my growth as a person and scientist.

I would also like to thank my committee members Dr. Alex Ensminger, Dr. Timothy Hughes and Dr. Joshua Milstein for their guidance and advice. Thank you for asking the questions that needed to be asked and helping me through my graduate studies.

Thank you to all the Navarre lab members who have offered valuable advice and made bench work and lab life enjoyable. Thank you to Sabrina Ali, Steve Hersch, Andrea Leung, Jeremy Soo and Betty Zou. Thank you also to Vivian Cheung for her friendship.

To my friends and family, who have been an endless supply of encouragement through the tough times: thank you. Thank you to my fiancé David Chan, who has faithfully been with me through it all. Thank you to my church community, to Abby Cabunoc, Liz Van Dyck, Hannah Hong and Melinda Kinney who have prayed with and for me. And lastly thank you to my Mom and Dad and my brothers for their loving support.

Finally, I would like to praise and thank my God and Saviour. He has carried me through my Masters and His love and grace sustain me.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

bp: base pairs

ChIP: chromatin immunoprecipitation

DAPI: 4', 6-diamidino-2-phenylindole

DNA: deoxyribonucleic acid

DTT: dithiothreitol

EDTA: ethylenediaminetetraacetic acid

EMSA: electrophoretic mobility shift assay

ESX-1: ESAT 6 secretion system I

Fis: factor for inversion stimulation

FPLC: fast protein liquid chromatography

HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

Hfr: high frequency recombination

HMG-I(Y): high mobility group protein

H-NST: H-NS truncated

HSQC: heteronuclear single quantum coherence

IPTG: isopropyl β-D-1-thiogalactopyranoside

LB: Luria-Bertani

LEE: locus of enterocyte effacement

Ler: locus of enterocyte effacement (LEE) encoded regulator

Ni-NTA: nickel nitrilotriacetic acid

NMR: nuclear magnetic resonance

PCR: polymerase chain reaction

PDIM: phthiocerol dimycocerosate

PGL: phenolic glycolipid

PHYRE: protein homology/analogy recognition engine

Psi-BLAST: position specific iterated basic local alignment search tool

RNA: Ribonucleic acid

SPI: *Salmonella* pathogenicity island

SSC: *Staphylococcal* chromosomal cassette

List of Appendices

1 Introduction

1.1 Horizontal Gene Transfer

Free-living bacteria encounter many different environmental challenges and evolve in ways to combat and survive against a multitude of different stresses. Compared with other mechanisms of evolution (mutations, insertions, deletions, duplications), horizontal gene transfer allows bacteria to evolve in leaps and bounds through the acquisition of new genes with novel functions (Ochman et al., 2000; Jain et al., 2003). As such, horizontal gene transfer is recognized as a primary driving force behind bacterial evolution (Ochman et al., 2000; Boto, 2010) and has played an important role in the spread of antibiotic resistance and the emergence of new pathogens. Examples include the spread of methicillin-resistant *Staphylococcus aureus* through the transfer of a SCC*mec* mobile element (*Staphylococcal* chromosomal cassette containing the *mecA* gene) and the acquisition of pathogenicity island LEE (locus of enterocyte effacement) in enteropathogenic and enterohaemorrhagic *E. coli* (Katayama et al., 2000; McDaniel et al., 1997).

Horizontal gene transfer is defined as the acquisition of DNA through means other than that of heredity or vertical descent. In bacteria, horizontal gene transfer can occur through three main mechanisms: i) transformation, ii) transduction and iii) conjugation. Transformation is the uptake of DNA directly from the surrounding environment. Certain bacteria such as *Neisseria gonorrhoeae, Haemophilus influenzae, Bacillus subtilis* and *Streptococcus pneumoniae* are naturally competent and are able to transport DNA across their bacterial membrane(s) in their natural environment (Dubnau, 1999). Transduction is the transfer of bacterial DNA from one bacterium to another mediated through bacteriophages, which inadvertently package bacterial DNA as it exits the bacterial host and introduce this DNA to the next infected bacterium. Two mechanisms of transduction have been observed: i) generalized transduction, which is the

packaging and transfer of random bacterial DNA, and ii) specialized transduction, the packaging and transfer of bacterial DNA directly adjacent to integration sites. And finally conjugation is the transfer of DNA from one bacterium to another via direct cell-cell contact, in the form of a sex pili in Gram negative bacteria and secreted adhesins in some Gram positive bacteria. Shared DNA is usually in the form of self-transmissible plasmids but conjugation can even allow the transfer of chromosomal DNA through plasmids that have integrated into the chromosome (Hfr strains).

Upon acquisition, new DNA can subsequently be maintained as an episome in the recipient cell or integrated into the genome through mechanisms such as homologous recombination (if the species are closely related and the transferred sequence has sufficient homology), site specific recombination by bacteriophage or transposon integrases, or "illegitimate" incorporation via double-stranded break repair, in order to be passed on to future generations (Ochman et al., 2000).

The advent and age of full genome sequencing has demonstrated the high degree at which horizontal gene transfer occurs and its impact on bacterial genomes (Ochman et al., 2000; Koonin et al., 2001; Boto, 2010, Abby et al., 2012). For example, though *Escherichia coli* and *Salmonella* Typhimurium are very closely related bacterial species, comparisons of the two genomes has revealed that as much as 25% of the *Salmonella* genome is dissimilar and was acquired by horizontal gene transfer after its divergence from the last common ancestor with *E. coli* (Porwollik et al., 2003). Likewise, it is estimated that 18% of the genome of a typical strain of *E. coli* has been horizontally acquired (Lawrence et al., 1998). In fact, one defining speciation event in *Salmonella* as it diverged from E. coli was the acquisition of a horizontally acquired ~40kb pathogenicity island named ***S****almonella* **P**athogenicity **I**sland I (SPI-1) which encodes for a type 3 secretion system that allows the pathogenic *Salmonella* to inject effector proteins into

the host to rearrange actin polymerization and facilitate bacterial entry into host cells (Haraga et al., 2008).

1.2 Detection of Horizontally Acquired Genes

Horizontal gene acquisitions are pre-dominantly defined using two approaches: i) phylogenetic comparisons: where incongruent gene trees between closely related species are deemed the result of horizontal gene transfer as the most parsimonious explanation. This method requires having the genomes of fully sequenced bacteria in order to compare and construct phylogenetic trees. It should be noted that gene loss can also explain phylogenetic incongruencies and that classification of a gene as being horizontally acquired occurs when it is more likely than the alternate explanation of gene loss in all other related species (the most parsimonious explanation) (Ochman et al., 2000; Koonin et al., 2001; Boto, 2010; Zhaxybayeva et al., 2011). ii) compositional analyses: where the GC-content, codon usage bias and di- and tri-nucleotide frequencies of genes are examined and compared against the host genome (Ochman et al. 2000; Koonin et al., 2001; Boto, 2010; Zhaxybayeva et al. 2011). This latter method relies upon the observation that within a genome, the GC content and codon usage variation is small (Sueoka, 1962). The introduction of foreign DNA therefore is identifiable because they bear the mark of its source and differ from the host genome (Lawrence et al., 1997). Compositional analyses however are only able to identify recently acquired genes as genes slowly ameliorate over time (a result of mutational pressures acting on both the native and foreign genes) to look more like that of the native genome (Lawrence et al., 1997; Marri et al., 2008). Regardless, these two methods (phylogenetic comparisons and compositional analyses) do indeed uncover the majority of horizontally acquired genes (Daubin et al., 2003) and are robustly used for their identification.

Based on compositional analyses of genomes, one empirically noted observation is that many foreign acquired genes have a tendency to be more AT-rich in their composition compared to the surrounding host genome (Lawrence et al., 1997; Lawrence et al., 1998; Syvanen, 1994; Groisman et al., 1993; Navarre et al., 2006; Navarre et al., 2007). It is not known why this bias exists but Navarre et al. (2007) postulate that this bias allows bacteria to differentiate between self and non-self DNA. The acquisition of predominantly AT-rich foreign sequences is intriguing and yields two explanations to explain this bias: i) bacteria tend to acquire foreign DNA from sources that are more AT-rich than themselves and/or ii) bacteria have evolved a mechanism to facilitate the transfer and maintenance of AT-rich foreign sequences. While many foreign elements in genomes do indeed originate from AT-rich sources such as phages (in the form of morons) (Daubin et al. 2003), there is also compelling evidence for the latter explanation- that is, bacteria have evolved proteins that facilitate the maintenance of AT-rich foreign DNA (Navarre et al., 2006; Lucchini et al., 2006; Grainger et al., 2006; Oshima et al., 2006).

1.3 Xenogeneic Silencing Proteins

The introduction of novel and foreign genes into a genome presents several challenges for the host cell. First, these new DNA sequences must be integrated into the host genome in such a way as to not disrupt essential genes or existing regulatory networks. Secondly, expression of these new genes must not lead to a decrease in cellular fitness. And lastly, the expression of these genes must be properly regulated to avoid inappropriate expression (Dorman, 2009). Because of the factors listed above, in the absence of any facilitative system, the acquisition and integration of novel genes into the genome would most likely lead to deleterious repercussions on fitness. Engaging in horizontal gene transfer would therefore be unwise. To counteract this, bacteria have evolved sentinel proteins that silence foreign DNA,

providing a means by which foreign DNA uptake can occur without any consequences on fitness (Lucchini et al., 2006; Navarre et al., 2006; Navarre et al., 2007; Dorman, 2007). New genes are initially repressed and ultimately spread cryptically in the population by vertical gene transfer. Eventually a bacterial cell in the population may evolve mechanisms to properly regulate the newly acquired gene, leading to a new beneficial phenotype (Navarre et al., 2007; Ali et al, 2012). Genes that do not have the potential to improve fitness are eventually lost from the population. Xenogeneic silencing proteins (proteins that silence foreign genes) therefore act as a buffering system to allow for the maintenance of potentially beneficial genes acquired through horizontal gene transfer without suffering the fitness consequences that may be incurred through their inappropriate expression.

There are three protein families found in the bacterial domain that function as xenogeneic silencers: H-NS is the best studied and is found amongst several clades of Gram negative proteobacteria, Lsr2 which is found amongst the Gram positive actinobacteria (Gordon et al., 2008) and MvaT which is found in the Gram negative Pseudomonads (Tendeng et al., 2003). These xenogeneic silencers share little structural and sequence identity (~20% amino acid sequence identity) but were identified and grouped together due to their ability to functionally complement each other's phenotypes and their shared functional roles in their bacterial hosts (Gordon et al., 2008; Tendeng et al., 2003; Ali et al, 2012). Figure 1 depicts the distribution of xenogeneic silencing proteins identified thus far in the bacterial domain.

All three xenogeneic silencing proteins silence foreign DNA by forming higher order oligomers that bind to and repress transcription of AT-rich DNA sequences (Ali et al., 2012; Gordon et al., 2008; Castang et al., 2010). Transcriptional repression of AT-rich DNA by these xenogeneic silencing proteins over evolutionary time, has led to the accumulation and maintenance of AT-rich foreign genes specifically (as opposed to GC-neutral or GC-rich genes). This offers an explanation for the AT-bias of foreign genes where empirical observations found

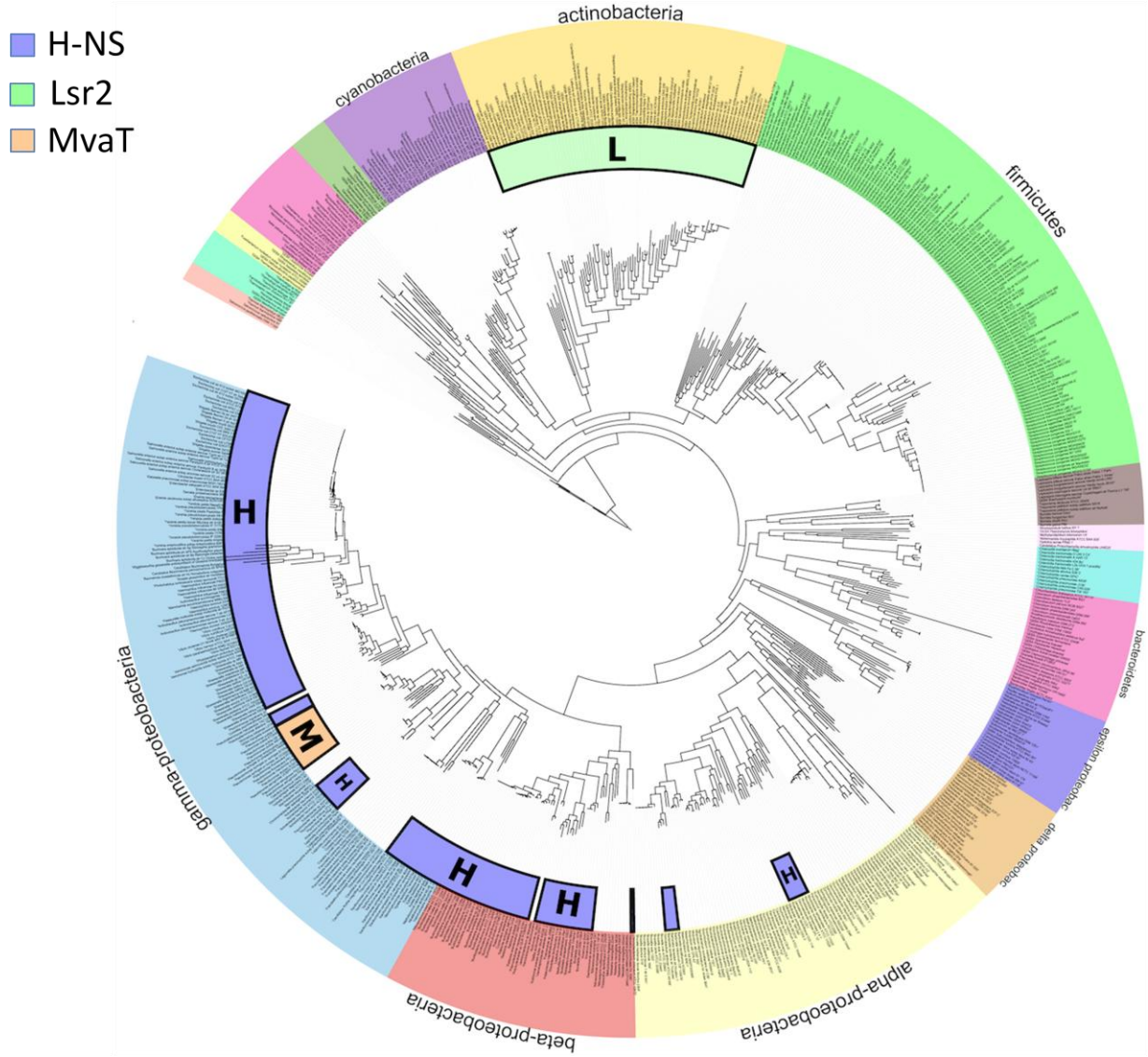


Figure 1: Xenogeneic silencing proteins within the bacterial domain. Phylogenetic tree of the bacterial domain, highlighting species where xenogeneic silencing proteins (to date) have been found. Figure made by Gavin Wilson and William Navarre.

foreign genes to be more AT-rich in composition compared to the host genome (Lawrence et al., 1997; Lawrence et al., 1998; Navarre et al., 2007).

1.4 H-NS: Global Regulator and Xenogeneic Silencing Protein

Jacquet et al. first discovered H-NS in 1971 as a contaminant in the process of purifying *E. coli* RNA polymerase, and later in 1977, Varshavsky et al. rediscovered H-NS in a search for histone-like proteins in *E. coli*. The pleiotropic consequences that occur when *hns* is deleted/mutated in enteric bacteria suggest a global and important role for H-NS. In *E. coli, hns* mutant phenotypes include a very slow growth phenotype (Barth et al., 1995), mucoidy appearance on solid media (Harrison et al., 1994), loss in motility (Bertin et al., 1994), increased hemolytic activity (Gomez-Gomez, 1996), usage of β-glucosides as a carbon source (Defez et al., 1981) and sensitivity to serine in minimal media (Lejeune et al., 1989). Furthermore, deletion of *hns* in *Salmonella* is poorly tolerated unless other mutations in genes like *rpoS, phoP* or SPI-2 are also present (Navarre et al., 2006; Lucchini et al., 2006). The reason for these pleiotropic effects lies in its role as a global regulator of virulence and housekeeping genes. H-NS recognizes and binds to AT-rich DNA in a sequence-independent manner and represses transcription (Grainger et al., 2006; Oshima et al., 2006; Navarre et al. 2006; Lucchini et al., 2006). As such, H-NS acts as a xenogeneic silencing protein, buffering the potentially deleterious consequences of horizontal gene transfer by recognizing foreign genes through their GC content and preventing their spurious expression (Navarre et al., 2007).

The H-NS protein is organized into three structural domains: an N-terminal dimerization domain, a central oligomerization domain (Esposito et al., 2002; Arold et al., 2010) and a C-terminal DNA-binding domain (Shindo et al., 1999). This modular domain structure of H-NS allows it to bind DNA through its C-terminal domain and self-associate into higher-order structure through its N-terminal and central domains. The ability of H-NS to form these higher-

order oligomers is essential for its silencing function (Spurio et al., 1997; Ueguchi et al., 1997). DNase footprinting studies demonstrated that H-NS bound to large stretches of DNA and investigation of the *proV* promoter showed that H-NS binds at high affinity nucleation sites and oligomerizes in a cooperative manner to bind to adjacent lower-affinity sites (Bouffartigues et al., 2007). In this way, H-NS binds to DNA and oligomerizes cooperatively along AT-rich tracts of DNA to form a nucleoprotein filament that represses transcription. Atomic force microscopy and magnetic tweezer experiments have also demonstrated two different modes of oligomeric binding for H-NS. Atomic force microscopy observed that H-NS is able to bind and bridge different strands of DNA (Dame et al., 2000) while magnetic tweezers demonstrated H-NS' ability to stiffen and lengthen otherwise flexible DNA in a cooperative manner (Amit et al., 2003). These two modes of binding exist in high and low magnesium conditions respectively (Liu et al., 2010) and it is still unclear which mode of binding (or both) is physiologically relevant.

In addition to its role as a xenogeneic silencer, H-NS also plays a role as a nucleoid-associated protein that modulates DNA topology and organization within the cell (Spassky et al., 1984; Wang et al., 2011). H-NS also acts as a repressor of housekeeping genes such as the *bgl* and *proU* operons responsible for β-glucoside metabolism and transport of osmotic protectants respectively (Schnetz, 1995; Dole et al., 2004; Higgins, et al., 1988; Lucht, et al., 1994).

1.5 Transcriptional Repression and Derepression

H-NS-mediated transcriptional repression of genes occurs through multiple mechanisms. i) Occlusion of RNA polymerase: H-NS binding to the promoter region may sterically block RNA polymerase from binding and initiating transcription (Lucchini et al., 2006), ii) Trapping of RNA polymerase: H-NS may bind and form a loop, trapping RNA polymerase and

preventing elongation (Schroder et al., 2000; Shin, et al., 2005), and iii) Constraining supercoils: H-NS being a nucleoid-associated protein is able to alter the superhelicity of DNA and binding may alter DNA topology, making it unfavourable for transcription (Blot et al., 2006).

Just as transcriptional repression can occur in several ways, derepression of H-NS-silenced genes also occurs through a variety of different mechanisms (Stoebel et al., 2008). Site-specific transcription factors are able to displace xenogeneic silencers and activate transcription of the underlying gene. SsrA and SsrB for instance act as master regulators of SPI-2 in *Salmonella,* a virulence program that enables the bacterium to survive in the intracellular environment. SsrB is able to antagonize H-NS repression at several SPI-2 loci and displaces H-NS to activate transcription of genes such as *sifA* and *sifB* (Walthers et al., 2007; Walthers et al., 2011). SlyA is another example of an H-NS antagonist that activates transcription by displacing H-NS from promoters and allowing access for RNA polymerase, as is the case for the *hlyE* promoter (Westermark et al., 2000; Lithgow et al., 2007). SlyA however also has been shown to activate transcription at the *pagC* and *utgL* promoters not by displacing H-NS but rather through remodeling the H-NS-DNA complex to allow RNA polymerase to bind (Stoebel et al., 2008). VirB in *Shigella* also uses this mode of transcriptional derepression (Turner et al., 2007). Other mechanisms include limiting the spread of the H-NS nucleoprotein filament, observed with LeuO, an H-NS antagonist (Chen et al., 2005) and competing with H-NS for binding sites, demonstrated by Fis (factor for inversion stimulation), another nucleoid-associated protein (Falconi et al., 1996; Falconi et al., 2001). H-NS-like proteins (Ler and H-NST) have dominant negative effects by interfering with H-NS oligomerization. These molecules imitate the N-terminal dimerization domain of H-NS and upon binding to H-NS, sequester it, preventing it from engaging in transcriptional repression (Bustamante et al., 2001; Williamson et al., 2005). Other factors such as supercoiling, osmolarity and temperature have also been shown to alter H-NS-mediated silencing at various promoters (Schnetz et al., 1996; Amit et al, 2003).

1.6 Lsr2: an H-NS-like Molecule in Gram positive *Mycobacterium*

A screen for *Mycobacterium leprae* proteins that interacted with the human immune system uncovered an immunodominant T-cell antigen Lsr2 (Laal et al., 1991). Its role in the *Mycobacterium* host began to be uncovered in 2006 when Chen et al. screened a library of transposon insertion mutants of *Mycobacterium smegmatis* for mutants with altered colony morphology and uncovered the gene *lsr2* as one which had the greatest changes in colony morphology when disrupted. Further work to characterize this protein found that though it only shares ~20% amino acid sequence identity with H-NS, it shared in many H-NS-like properties such as a modular domain structure (Gordon et al. 2010), ability to dimerize and oligomerize, non-specifically bind to AT-rich DNA and bridge DNA (Chen et al. 2008). Through complementation studies, Lsr2 was able to functionally complement many Δ*hns* phenotypes in *E. coli* including mucoidal appearance, utilization of β-glucosides as a carbon source, hemolytic activity, and loss of motility. Likewise, H-NS was able to restore the altered colony morphology in a mutant *lsr2 M. smegmatis* background (Gordon et al. 2008). The authors also demonstrated the ability of Lsr2 to bind to known H-NS-regulated targets and not to non-targets, and be similarly displaced by H-NS antagonist SlyA (Gordon et al., 2008). ChIP-chip in both *M. tuberculosis* and *M. smegmatis* found that Lsr2 pulled down with specifically AT-rich sequences (47% GC or less in a host genome with a GC content of ~66%), and bound a significant portion of horizontally acquired genes (~58%) (Gordon et al. 2010). Lsr2 also bound and repressed the expression of many virulence genes such as the ESX-1 pathogenicity island, which is essential for virulence in *M. tuberculosis*, cell wall lipids PDIMs (phthiocerol dimycocerosates) and PGLs (phenolic glycolipids) and the *iniBAC* operon which is involved in multi-drug efflux (Gordon et al. 2010; Colangeli et al., 2007). Lsr2 proteins have also been identified in related actinomyces such as *Streptomyces*, *Nocardia* and *Rhodococcus* (Gordon et al., 2008).

Altogether, these studies provide strong evidence that Lsr2 belongs to the H-NS family of proteins as repressors of AT-rich foreign DNA.

1.7 MvaT: an H-NS-like Molecule in *Pseudomonas*

Rosenthal et al. (1988) originally identified MvaT as being involved in mevalonate metabolism in *Pseudomonas mevalonii*. Its role as an H-NS-like xenogeneic silencer was assigned in 2003, when Tendeng et al. screened a library of *Pseudomonas* genes for the ability to complement the *E. coli hns* phenotype of serine susceptibility in minimal media. They identified MvaT and further demonstrated that MvaT was able to complement other Δ*hns* phenotypes such as loss of motility and utilization of β-glucosides as a carbon source (Tendeng et al., 2003). Subsequent studies verified that MvaT indeed functioned like H-NS, sharing its modular domain structure (Tendeng et al. 2003; Castang et al., 2010), ability to bind AT-rich DNA and repress transcription (Castang et al., 2008), self-associate into dimers and oligomers (Castang et al., 2010) and form nucleoprotein filaments on DNA (Winardhi et al., 2012). In *P. aeruginosa*, MvaT binds approximately two-thirds of the horizontally acquired regions of genome plasticity (Castang et al., 2008) and represses and regulates the transcription of many virulence factors in *Pseudomonas* including genes involved in pyocyanin (toxin) production, *lecA* involved in quorum sensing (Diggle et al. 2002; Castang et al., 2008), and *cupA1* involved in biofilm formation (Vallet et al., 2004; Valet-Gely et al., 2005). Altogether, there is strong evidence that H-NS, Lsr2 and MvaT belong to the same functional family of proteins: binding and transcriptionally repressing AT-rich foreign DNA to facilitate bacterial evolution.

1.8 H-NS and Lsr2: AT-hook is Critical for DNA Binding

The ability for xenogeneic silencers H-NS, Lsr2 and MvaT to discriminate between AT-rich DNA targets and GC-rich or GC-neutral non-target DNA lies in their C-terminal DNA binding domain. Gordon et al. (2011) have used NMR spectroscopy to solve the DNA binding domains of H-NS and Lsr2. The authors demonstrated that despite the fact that H-NS and Lsr2 do not share much sequence homology (~20% amino acid sequence identity) or structural similarity (H-NS DNA binding domain is composed of two anti-parallel β-sheets, one α-helix and one 3_{10} helix while Lsr2 DNA binding domain consists of two α-helices linked by a long loop) (Gordon et al., 2010; Gordon et al., 2011), they both bind to AT-rich DNA via a small positively charged motif- QGR in H-NS and RGR in Lsr2 (Gordon et al., 2010; Gordon et al., 2011). Figure 2 illustrates a docking model demonstrating that this motif inserts into the narrow minor groove of AT-rich DNA such that the glutamine and arginine residues point out in opposite directions. Xenogeneic silencers non-specifically bind AT-rich DNA through the recognition of structural features of the minor groove of AT-rich. A-tract DNA (defined as 4 consecutive ApA, TpT or ApT steps) is more narrow in width due to negative propeller twisting, making it more electronegative (Rohs et al., 2009). A:T base pairs also lack an exocyclic 6-NH_2 group, allowing for deeper penetration of DNA-binding proteins (Ali et al., 2012; Rohs et al., 2009). These features enable the QGR/RGR motif to recognize and bind specifically to AT-rich DNA. This QGR/RGR motif is reminiscent of the AT-hook (PRGRP motif) found in eukaryotic non-histone chromatin protein HMG-I(Y) (Huth et al., 1997; Aravind et al., 1998). As such, we have coined the QGR and RGR motifs in H-NS and Lsr2 respectively as AT-hook motifs (Gordon et al., 2011). Mutagenesis of the glutamine and arginine residues in H-NS and Lsr2 to alanines completely abolished DNA binding, (assessed by a lack of change in the 2D ^{1}H-^{15}N HSQC spectra), reinforcing the essentiality of these residues for DNA binding (Gordon et al.,

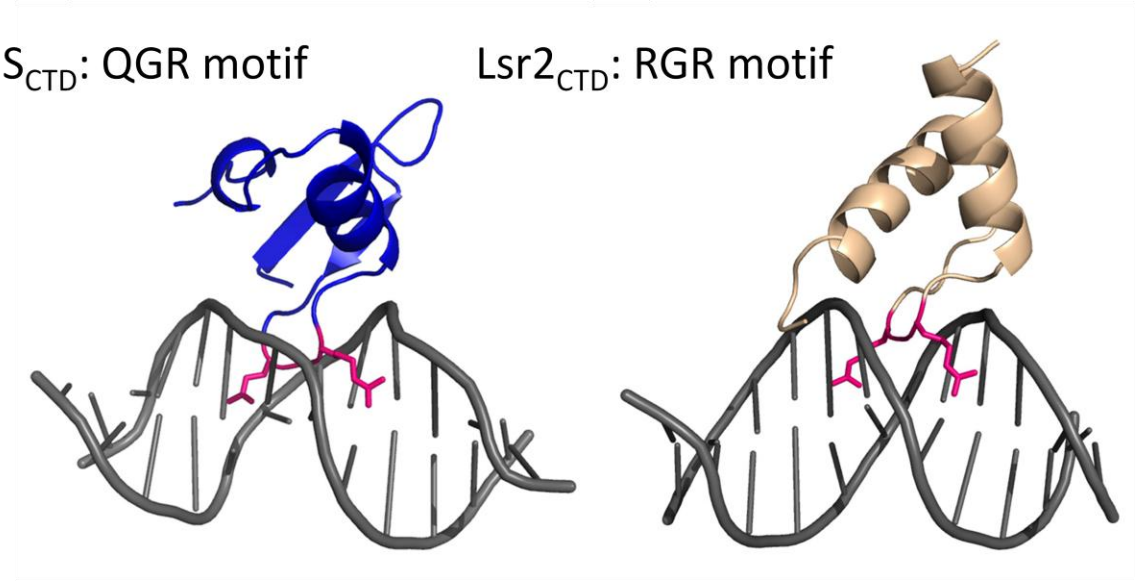


Figure 2: C-terminal DNA binding domains of H-NS and Lsr2. Despite sharing little sequence and structural identity, H-NS and Lsr2 both bind to AT-rich DNA via a small positively charged QGR and RGR motif respectively. These residues insert themselves into the minor groove of AT-rich DNA, pointing outward, reminiscent of the AT-hook found in eukaryotic non-histone chromatin protein HMG-I(Y). Figure from Ali et al, 2012 and work done by Gordon et al., 2010, Gordon et al., 2011.

2011). MvaT does not possess this QGR/RGR motif in its primary sequence and its mechanism of binding DNA has not yet been determined.

1.9 MvaT Prefers Wider and More Flexible AT-rich DNA

Our lab (in collaboration with Dr. Timothy Hughes) has investigated the sequence preferences of all three xenogeneic silencers *in vitro* using protein binding microarrays (PBM). Purified protein was applied to microarrays containing double-stranded 60-mer oligonucleotides (25nt constant primer sequence and 35nt variable sequence) such that all possible non-palindromic 8-mers were represented 32 times (Gordon et al., 2011). H-NS, Lsr2, and MvaT all displayed a clear preference for AT-rich DNA in this assay. The type of AT-rich DNA that each xenogeneic silencer bound however was different. For Lsr2 and H-NS, the highest scoring oligomers were A-tract DNA (defined as 4 consecutive ApA, TpT or ApT steps) interrupted by one or two adjacent TpA steps. The presence and preference for sparse TpA steps is presumably to allow the AT-hook motif to penetrate deeply into the minor groove as TpA steps widen it (Gordon et al., 2011; Ali et al., 2012). In contrast, multiple TpA steps (TATATATA) composed the highest scoring oligomers bound by MvaT (unpublished data). This indicates that MvaT binds a wider minor groove than H-NS and Lsr2 and prefers more flexible DNA (as TpA steps also increase the flexibility of DNA).

1.10 Thesis Rationale

Xenogeneic silencing proteins H-NS, Lsr2 and MvaT have global roles in their bacterial hosts as repressors of AT-rich DNA sequences. It is important to understand how these xenogeneic silencers recognize their target sequences, as it sheds light on the evolutionary impact they have had on the genome. These xenogeneic silencers also regulate many genes involved in virulence and pathogenicity and it is important to study how they function in order

to better understand the environmental cues and intracellular signaling which affect H-NS, Lsr2 and MvaT and contribute to pathogenicity. As well, this work has practical applications. Understanding the DNA binding mechanism of these xenogeneic silencers may allow us to potentially improve protein expression systems used for both research and drug production purposes where xenogeneic silencers may decrease or prevent yields of recombinant protein expression in a non-native system.

Previous work (Gordon et al., 2011) has characterized two of the three xenogeneic silencing proteins (known to date) and demonstrated the use of a Q/R GR AT-hook motif that H-NS and Lsr2 employ to recognize and bind AT-rich DNA. The DNA binding mechanism of the third xenogeneic silencing member, MvaT, has not been characterized or well-studied. MvaT lacks an obvious AT-hook motif and protein binding microarray studies hint at a possibly different mode of DNA binding. We were therefore interested in studying the DNA binding mechanism of MvaT, to shed light on the third and remaining xenogeneic silencing protein.

In 2013, our collaborator Dr. Bin Xia and his lab at Peking University, solved the NMR spectroscopy structure of the DNA-binding (or C-terminal) domain of MvaT in complex with an AT-rich fragment of DNA (Figure 3). Structural data suggested that a KGGN motif, that is well-conserved amongst *Pseudomonas,* and other arginine and lysine residues in proximity to this KGGN motif were predicted to make contact with DNA. A docking model demonstrates that the KGGN motif inserts into the minor groove of DNA, much like the QGR and RGR motif. The lysine and asparagine residues of this motif also point outward in the same way that had previously been observed with the glutamine and arginine residues of H-NS and Lsr2's AT-hook motif. We therefore hypothesized that based on the similarity between the KGGN motif and the QGR/RGR AT-hook motif that in MvaT, an elongated motif may carry out the function of that

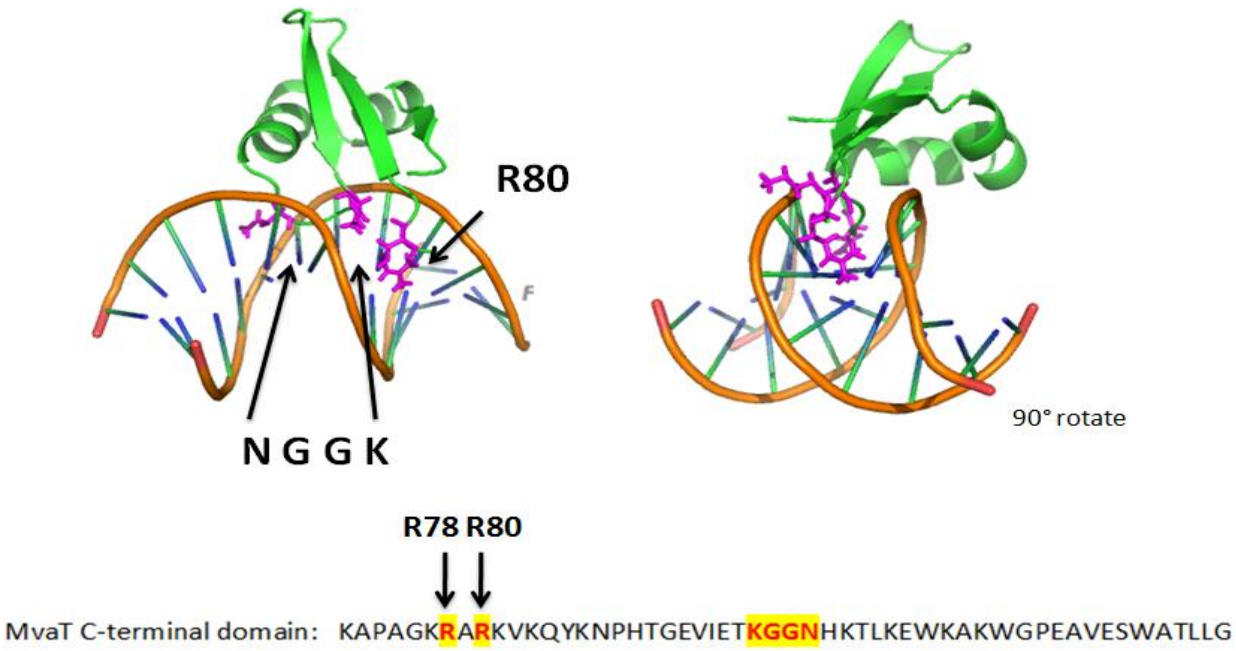


Figure 3: NMR structure of MvaT DNA binding domain. Docking model of MvaT C-terminal DNA binding domain (green) in complex with DNA demonstrate that R80, K97 and N100 residues (magenta) insert into the minor groove of AT-rich DNA. R78 is also predicted to make contact with DNA. Structure courtesy of Pengfei Ding, Shujuan Jin, Bin Xia, Peking University.

of the AT-hook in H-NS and Lsr2. This elongated motif would also explain MvaT's preference for DNA with wider minor grooves (PBM data).

To test the hypothesis that the KGGN motif is an elongated AT-hook motif, I have employed a similar technique to that used by Gordon et al. (2011), who demonstrated the significance and importance of the glutamine and arginine residues in H-NS and Lsr2 by changing them to alanine and evaluating DNA binding by NMR. I have conducted site-directed mutagenesis on Lys-97 and Asn-100 of the KGGN motif, changing these residues to alanine, and evaluated DNA binding *in vitro* by electrophoretic mobility shift assays (EMSAs). In this study, I address the question of how mutations to the lysine and asparagine residues of the KGGN motif affect DNA binding and recognition of target AT-rich DNA and ask whether or not this novel KGGN motif functions as an analogous AT-hook in *Pseudomonas*.

2 Materials and Methods

2.1 Construction of MvaT and H-NS Protein Expression constructs

The MvaT region from *Pseudomonas aeruginosa* was PCR-amplified using primers WNp524 and WNp525 (Table 1) and inserted into a pHSG576 low copy plasmid (Takeshita et al., 1987) using BamHI and HindIII restriction enzyme sites to make plasmid pWN597. Using pWN597 as a template, the MvaT gene alone was PCR-amplified using primers MvaT Full F and MvaT R and inserted into a pET21b plasmid. These primers however also amplified a portion of the multiple cloning site of pWN597 and so primers pET21b RBS F and MvaT RBS R were used to perform around the world PCR on the incorrect pET21b: MvaT vector to omit the extraneous region and ligated together to generate the final construct pGT009, which is a pET21b expression vector with C-terminally 6-His tagged MvaT under the control of the T7 promoter. $MvaT_{77\text{-}124}$ was cloned in an identical manner using primers MvaT 77 SacI pET21b and MvaT R to amplify from template pWN597 and primers pET21b RBS F and MvaT77 RBS R were used to exclude the extra multiple cloning site region. These constructs were generated by William Navarre, Emily Beckett, Andrea Leung and myself.

Alanine mutations were introduced into pGT009 (pET21b-MvaT) to generate MvaT K97A/N100A by site-directed mutagenesis. Complementary primers (GT031-GT033, GT037-GT039) flanking the mutation on each side were designed and used in a 2-step Phusion High-Fidelity DNA polymerase (Thermo Scientific) PCR reaction with pGT009 (pET21b-MvaT) and pET21b-$MvaT_{77\text{-}124}$ plasmids. Reactions were treated with DpnI restriction enzyme at 37°C for 1 hour to remove parental vector and cleaned up (Bio Basic EZ-10 spin column PCR product

Table 1: Oligonucleotide primers used in this study

Name	Description	Sequence	Reference
GT031	MvaT K97A	CCGGCGAAGTCATCGAGACCGCGG GCGGCAACCACAAGACTTT	Grace Tong
GT032	MvaT N100A	TCATCGAGACCAAGGGCGGCGCCC ACAAGACTTTGAAAGAGTG	Grace Tong
GT033	MvaT K97A/N100A	TCATCGAGACCGCGGGCGGCGCCC ACAAGACTTTGAAAGAGTG	Grace Tong
GT037	MvaT K97A	AAAGTCTTGTGGTTGCCGCCCGCGG TCTGCATGACTTCGCCGG	Grace Tong
GT038	MvaT N100A	CACTCTTTCAAAGTCTTGTGGGCGC CGCCCTTGGTCTCGATGA	Grace Tong
GT039	MvaT K97A/N100A	CACTCTTTCAAAGTCTTGTGGGCGC CGCCCGCGGTCTCGATA	Grace Tong
GT044	cupA1 340bp	GCGAAGCCGTGGTTCGAGTTGTT	Grace Tong
GT045	cupA1 340bp	ATCCCGGCCTCTCTTGCTTGTCTT	Grace Tong
GT049	PA3900 204 bp	CCGCAGGTGGCTGAACA	Grace Tong
GT050	PA3900 204 bp	CGAATGCGGTGCGTTGATGG	Grace Tong
GT068	hilA 289 bp	GGCATGATAATAGTGTATTCTCTT	Grace Tong
GT077	hilA 289 bp	CTCTCTCTGCACCAGGATA	Grace Tong
GT067	hilA 24 bp	AAGAGAATACACTATTATCATGCC	Grace Tong
GT071	hilA 101 bp	GTACTAACAGCAGAATTACTG	Grace Tong
SSA1	H-NS into pSSA2	AAAACATATGAGCGAAGCACTTAA AATTCTGAACAACATCC	Sabrina Ali
SSA2	H-NS into pSSA2	AAAACTCGAGTTCCTTGATCAGGA AATCTTCCAGTTGCTTACC	Sabrina Ali
WNp524	MvaT into pHSG576 (pWN597)	TTTAAGCTTGGATCCGGCGAAGGTC TTTCGACAG	William Navarre
WNp525	MvaT into pHSG576 (pWN597)	TTTAAGCTTGGATCCGTTTGTCCCA TGAAGAATACGG	William Navarre
MvaT Full F	MvaT into pET21b (incorrect)	GACGCCACCATGGGGCATATGGTG CACAGGATGTCCCTGATCAACGAA TATCGCGCCACG	Emily Beckett
MvaT R	MvaT into pET21b (incorrect)	GTTGGGGTAACTGGCCGCGGATTG GAAAGGTTAGCCGAGCAGGGTGGC CCAGCTCTCGAC	Emily Beckett
MvaT 77 SacI F pET21b	MvaT 77-124 into pET21b (incorrect)	AGCAAATGGGTCGGGATCCGAATT CGAGCTATGAAGCGCGCGCGCAAG GTCAAGCAGTAC	Emily Beckett
pET21b RBS F	Repair to generate pGT009	ATGTATATCTCCTTCTTAAAGTTAA ACAAAATTA	Emily Beckett
MvaT RBS R	Repair to generate pGT009	ATGTCCCTGATCAACGAATATCGCG CCA	Emily Beckett
MvaT77 RBS R	Repair to generate pET21b-MvaT 77-124	ATGAAGCGCGCGCGCAAGGTCAA	Emily Beckett

purification kit). DNA was transformed into *E. coli* DH5α and positive clones were selected for on LB-amp plates and verified by sequencing (TCAG DNA sequencing facility). H-NS from *Salmonella enterica* subspecies enterica Serovar Typhimurium strain LT2 was PCR-amplified using primers SSA1 and SSA2 and inserted into a pET21b vector between NdeI and XhoI restriction enzyme sites. This pSSA2 plasmid is a pET21b expression vector with C-terminally 6-His tagged H-NS under the control of the T7 promoter (Ali et al., 2011). This construct was generated by Sabrina Ali. The H-NS Q112A/R114A construct was generated by Blair Gordon and kindly provided by the Liu lab (Gordon et al., 2011).

2.2 Protein Purification

Protein expression constructs were transformed into an *E. coli* BL21 DE3 strain containing the plysS plasmid (a plasmid encoding T7 lysozyme, a natural inhibitor of T7 RNA polymerase), to prevent basal level of expression prior to induction. Cultures were incubated at 37°C with shaking until mid-log phase (OD_{600}= 0.5-0.6). Protein expression was then induced by the addition of 200μM IPTG, following which, the incubator was cooled to 18°C and cultures continued to incubate overnight. Cells from 4L of cultures were pelleted by centrifugation in a Sorvall Legend RT+ centrifuge swinging bucket rotor for 30 minutes at 3300rpm at 4°C and chilled at -80°C for 30 minutes. Pellets were resuspended in 50mL cold lysis buffer (10mM Tris pH 8.0, 500mM NaCl, 5mM imidazole, 2.5mM β-mercaptoethanol) and lysed by sonication (3 minutes in 30 second intervals on ice with a Branson sonifier 450 sonicator). Lysates were cleared by centrifugation at 8900rpm for 45 minutes at 4°C (Sorvall Legend RT+ fixed angle rotor) and the insoluble debris was discarded. Ni-NTA agarose resin (Qiagen) was washed and equilibrated in cell lysis buffer and incubated with the cleared cell lysates for 1 hour at 4 °C on a rocking platform. The cell lysate/Ni^{2+} resin mixture was applied to a gravity flow column and washed twice with wash buffer (10mM Tris pH 8.0, 300mM NaCl,

30mM imidazole) before elution with 500mM imidazole elution buffer (10mM Tris pH 8.0, 300mM NaCl, 500mM imidazole). Eluates were dialyzed against Buffer A (20mM Tris pH 8.0, 100mM NaCl, 1mM EDTA, 5% glycerol) at 4°C over 24 hours (changed twice into fresh buffer). Protein samples were further purified (to remove co-eluting contaminants as visualized by Coomassie staining on a 16% tricine gel) over a 5mL HiTrap Heparin HP column (GE Healthcare Life Sciences) using the ÄKTA FPLC system. The Heparin column was equilibrated with 4 column volumes of Buffer A prior to injection of dialyzed protein sample. A linear NaCl gradient from 0.1M to 1M NaCl was applied at a flow rate of 1mL/min to elute protein. Fractions containing the protein of interest (as visualized by Coomassie staining on a 16% tricine gel) were pooled and dialyzed again against Buffer A at 4°C. Protein samples were concentrated using Amicon Ultra-15 centrifugal filter 3kDa molecular weight cut-off (EMD Millipore). Protein concentration was measured using a spectrophotometer and calculated using the Beer-Lambert law ($A= \varepsilon cl$) where A is the absorbance at 280nm in a black quartz cuvette, ε is the predicted extinction coefficient (ExPASy) ($\varepsilon_{MvaT}= 20970$), c is the concentration in mols/L and l is the cuvette pathlength in centimetres (1cm). An equal volume of glycerol was added and protein was stored at -80°C. Visualization by SDS-PAGE confirmed the purify of protein. All proteins purified in this study were performed by myself with the exception of H-NS which was purified by Jeremy Soo according to the outlined procedure above ($\varepsilon_{H\text{-}NS}= 9970$).

2.3 Electrophoretic Mobility Shift Assays:

2.3.1 DNA Probes:

A 289bp fragment of the *Salmonella hilA* gene was PCR amplified from *Salmonella enterica* subspecies enterica Serovar Typhimurium strain SL1244 using primers GT077 and GT068. Shorter fragments (24bp and 101bp) of this original 289bp *hilA* DNA were used for gel shifts of truncated MvaT. The 24bp fragment was constructed by annealing complementary

oligonucleotide primers GT067 and GT068 (equimolar concentrations of each primer were heated in a 94°C water bath for 5 minutes and cooled to room temperature for 1 hour). Primers GT071 and GT068 were used to PCR amplify the 101bp fragment from *Salmonella* Typhimurium SL1344. The *cupA1* 340bp fragment was PCR amplified from *Pseudomonas aeruginosa* PAO1 using primers GT044 and GT045. A 204bp fragment of the *PA3900* gene was PCR amplified from *Pseudomonas aeruginosa* PAO1 using primers GT049 and GT050. The full sequence of each of these DNA fragments is found in the Appendix.

2.3.2 5' Radiolabelling of DNA Probes:

DNA probes were 5'end-radiolabelled with γ-^{32}P ATP using T4 polynucleotide kinase (New England Biolabs). 400nM DNA fragment, 4μL γ-^{32}P ATP (3000Ci/mmol, 10mCi/mL, Perkin Elmer), 1x polynucleotide kinase buffer and T4 polynucleotide kinase enzyme (New England Biolabs) were incubated in a total of 40μL at 37°C for 30 minutes. The reaction was stopped by the addition of 1μL of 0.5M EDTA and excess radioisotopes were removed using a G-25 spin column (GE Healthcare Life Sciences). Spin column resin was resuspended by vortexing the column upside down for 30 seconds. The cap was loosened one-quarter turn and the bottom closure was snapped off. The column was then placed into a 1.5mL microcentrifuge tube and spun at 2800rpm (735xg) in a tabletop centrifuge. The column was then placed into a fresh 1.5mL microcentrifuge tube and the radiolabelling reaction was applied. DNA was eluted by spinning at 2800 rpm (735xg) for 2 minutes in a tabletop centrifuge. 760μL water was added to a final volume of 800μL to dilute the DNA to a working stock of 20nM. DNA was aliquoted and stored at -20°C. 1μL of this stock will yield a final concentration of 1nM in a 20μL EMSA binding reaction.

2.3.3 Electrophoretic Mobility Shift Assay:

1nM radiolabelled DNA was incubated with varying concentrations of protein in binding buffer (15mM HEPES pH 7.9, 40mM KCl, 1mM EDTA, 0.5mM DTT, 5% glycerol). Addition of varying amounts of protein altered the concentrations of solutes from tube to tube. Buffer conditions were therefore normalized such that each reaction contained 8mM Tris pH 8.0, 15mM HEPES pH 7.9, 40mM NaCl, 40mM KCl, 1.4mM EDTA, 0.5mM DTT, 9.95% glycerol. Binding reactions were incubated at room temperature for 30 minutes. 2.5µL of 10x DNA loading dye (10mM Tris-HCl pH 7.5, 10mM EDTA, 65% sucrose, 0.3% bromophenol blue) was added to each reaction and samples were loaded onto a 6% native polyacrylamide gel that had been pre-run for 1 hour at 100 volts at 4°C. Samples were run at 70 volts for 165 minutes at 4°C, dried in a Gel Dryer (Labnet International) for 1 hour at 80°C and exposed overnight on a storage phosphor screen (GE Healthcare Life Sciences/Molecular Dynamics). Gels were visualized the following morning using a Typhoon 9400 imager with an image resolution of 50µm.

2.3.4 Competitive Gel Shifts:

Binding reactions occurred as described above with the following changes/additions: binding reactions were incubated for 15 minutes at room temperature before the addition of excess cold (unlabelled) DNA. The samples were further incubated another 15 minutes at room temperature. DNA loading dye was added and samples were electrophoresed and visualized as above.

2.3.5 Salt-induced Dissociation:

Binding reactions were performed as described above with the following changes: binding reactions were incubated for 15 minutes at room temperature before the addition of varying salt (NaCl) concentrations. The samples were further incubated another 15 minutes at

room temperature. DNA loading dye was added and samples were electrophoresed and visualized as above.

2.4 Nuclease Protection Assays:

2.4.1 DNase I Protection Assay:

Binding reactions (DNA with protein) occurred at room temperature for 15 minutes in DNase I buffer (10mM Tris pH 7.5, 2.5mM $MgCl_2$, 0.1mM $CaCl_2$). DNase I (Thermo Scientific) was added and the reaction was incubated at room temperature for 10 minutes. DNase I was heat inactivated at 70°C in the presence of 5mM EDTA for 10 minutes. The sample was spun down briefly and cooled to room temperature for 20 minutes. DNA loading dye was added and samples were electrophoresed and visualized as described above.

2.4.2 Exonuclease III Protection Assay:

Binding reactions (DNA with protein) occurred at room temperature for 15 minutes in NEBuffer 1 (10mM Bis-Tris-propane HCl pH 7.0, 10mM MgCl2, 1mM DTT, New England Biolabs). Exonuclease III (New England Biolabs) was added and the reaction was incubated at room temperature for 30 minutes. Exonuclease III was heat inactivated at 70° for 20 minutes. The sample was spun down briefly and cooled to room temperature for 20 minutes. DNA loading dye was added and samples were electrophoresed and visualized as described above.

3 Results

3.1 Generation of MvaT K97A/N100A Double Mutant

To determine whether or not the KGGN motif in MvaT functioned in a similar manner to the QGR AT-hook motif in H-NS, I purified wild type MvaT and a double mutant MvaT K97A/N100A for EMSA analysis. I cloned MvaT from *Pseudomonas aeruginosa* PAO1 into a pET vector and purified the protein using an *E. coli* BL21 DE3 expression system. I mutated residues Lys-97 and Asn-100 to alanine using site-directed mutagenesis to generate the MvaT K9A/N100A double mutant and likewise purified the MvaT double mutant.

3.2 DNA Fragments for EMSAs

I amplified three different DNA regions (*hilA*, *cupA1*, *PA3900*) by PCR from *Salmonella* Typhimurium and *Pseudomonas aeruginosa* to generate DNA fragments for gel shift assays. I selected these regions based on their ability to bind (or not bind) H-NS or MvaT in ChIP-chip studies (Navarre et al., 2006 and Castang et al., 2008), their known repression (or lack thereof) by H-NS or MvaT, and their GC content.

For positive controls I employed DNA fragments generated from the *hilA* and *cupA1* genes. H-NS has been shown to bind the promoter of and regulate the transcription of *hilA* in *Salmonella* Typhimurium (Navarre et al., 2006, Olekhnovich, et al., 2006). The *hilA* 289bp fragment used for gel shift assays has a GC content of 34% (*Salmonella* Typhimurium LT2 genome GC content is 52%). In *P. aeruginosa*, MvaT binds and regulates the AT-rich *cupA1* promoter (Castang et al., 2008, Vallet et al., 2004). For a negative control I employed a fragment of the *Pseudomonas PA3900* gene that lies within a ~180kb GC-rich region of the genome that does not contain any MvaT binding sites (Castang et al. 2008). The 340bp *cupA1* promoter fragment is relatively AT-rich (54% GC) and the 204bp fragment of *PA3900* is

relatively GC-rich (74% GC) compared to the *P. aeruginosa* genome average (*P. aeruginosa* PAO1 global genome % GC is 66%).

3.3 H-NS AT-hook Motif is Essential for DNA Binding and Discrimination of AT-rich Targets

Previously, Gordon et al (2011) determined that H-NS and Lsr2 bind to DNA through a small positively charged motif: QGR in H-NS and RGR in Lsr2. I have been able to confirm the role of these residues in DNA binding by performing electrophoretic mobility shift assays with both wild type H-NS and H-NS Q112A/R114A. Figure 4 demonstrates that H-NS is able to recognize and bind an AT-rich fragment of *hilA* DNA with greater affinity than non-target GC-rich *PA3900* DNA fragment (>6x affinity). H-NS binding to AT-rich DNA resulted in discrete shifted bands whereas binding of GC-rich DNA resulted in more diffuse smears until a higher concentration of protein (600nM) was present. The H-NS double AT-hook mutant (Q112A/R114A) was unable to bind DNA below protein concentrations of 600nM and the protein DNA complexes did not form discrete bands indicating that the mutations drastically reduced the affinity of H-NS for DNA. Notably, H-NS mutant protein shifted the GC-rich DNA fragment at the same protein concentration as the AT-rich fragment. This demonstrates the importance of the AT-hook motif in both DNA binding and distinguishing between target AT-rich DNA and non-target GC-rich DNA.

3.4 MvaT Binds GC-rich DNA with a Higher than Expected Affinity

To compare H-NS with MvaT, I conducted gel shifts with MvaT binding to either AT-rich or GC-rich DNA. Both H-NS and MvaT proteins bound AT-rich DNA with similar

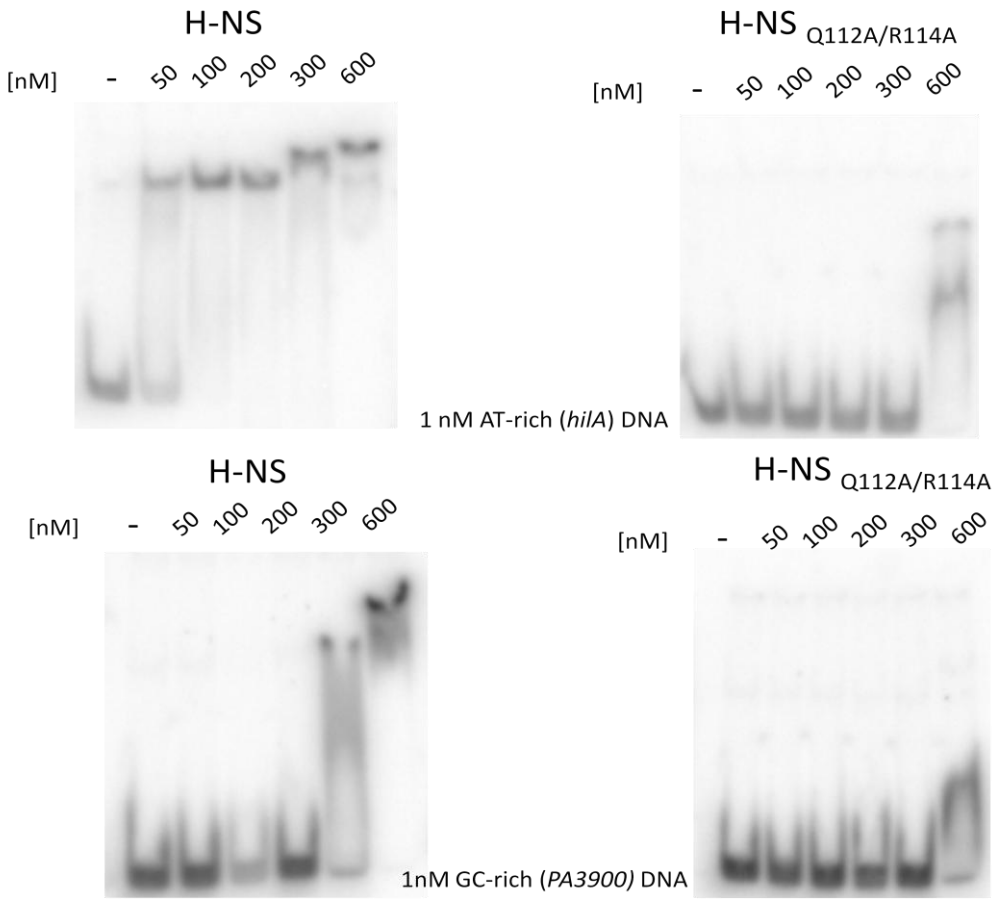


Figure 4: H-NS selectively binds AT-rich DNA via its AT-hook motif. H-NS has a higher affinity for AT-rich (*hilA)* DNA with binding and shifting occurring at 50nM (compared to 300nM with GC-rich (*PA3900*) DNA). Alanine mutations to the AT-hook motif of H-NS drastically alter its ability to bind DNA, demonstrating the essentiality of these residues for DNA binding. 1nM radiolabelled DNA was incubated with H-NS or H-NS AT-hook mutant (Q112A/R114A) for 30 minutes at room temperature. Complexes were resolved on a 6% native polyacrylamide gel. AT-rich DNA: 289bp fragment of *hilA* promoter (34% GC); GC-rich DNA: 204bp fragment of *PA3900* (74% GC).

affinities, (Figure 4 and 5) with shifting observed at 50nM protein. Binding of GC-rich DNA however was dissimilar between H-NS and MvaT. Compared to H-NS, MvaT demonstrates a higher affinity for GC-rich DNA. Whereas shifting of GC-rich DNA does not appear until 300nM H-NS, binding and shifting occurs at concentrations as low as 50nM MvaT (Figure 4 and 5). Between AT-rich and GC-rich DNA fragments, H-NS demonstrates a >6x affinity for AT-rich DNA over GC-rich DNA, while MvaT demonstrates an approximately 2.5x greater binding affinity, which is a much smaller difference in differential affinity. This was unexpected given that MvaT *in vivo* ChIP-chip data (Castang et al., 2008) and unpublished protein binding microarray data demonstrated that MvaT had a clear preference for AT-rich sequences.

3.5 MvaT Lys-97/Asn-100 Residues are not Essential for DNA Binding

To investigate the role of Lys-97 and Asn-100 in DNA binding, I conducted electrophoretic mobility shift assays with the double alanine mutant MvaT K97A/N100A. Figure 6 demonstrates that binding and shifting of both AT-rich (*cupA1*) and GC-rich (*PA3900*) DNA is very comparable to that of wild type MvaT. There is an approximately 2 fold difference in affinity between wild type MvaT and the Lys97/Asn100 double mutant, which is in stark contrast to the H-NS Q112A/R114A mutant, which shows much greater loss in DNA binding. These results demonstrate that *in vitro,* MvaT Lys-97 and Asn-100 residues are not essential for DNA binding and that these residues differ in their role from that of the QGR AT-hook motif in H-NS.

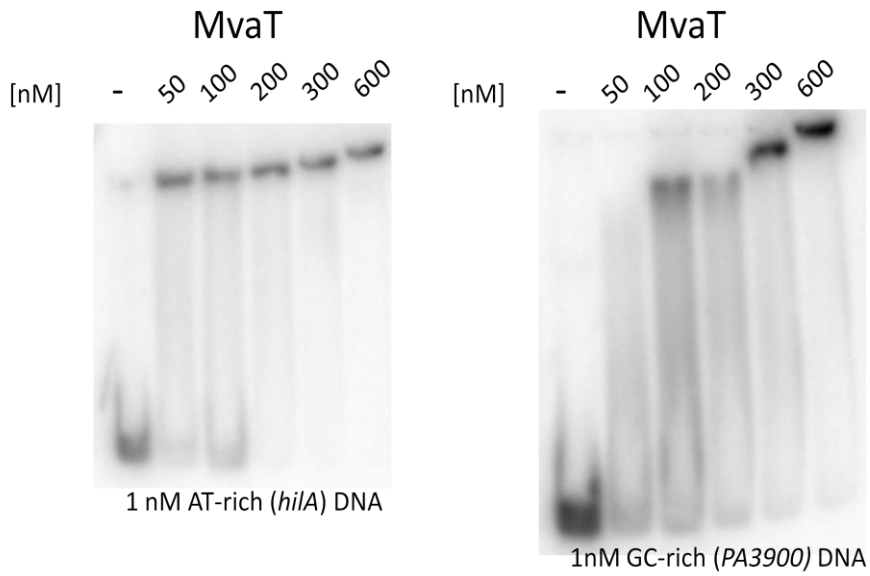


Figure 5: MvaT binds GC-rich DNA with higher affinity than H-NS. MvaT and H-NS both bind an AT-rich DNA fragment (*hilA*) with similar affinities. Whereas H-NS begins shifting GC-rich (*PA3900)*DNA at a concentration of 300nM, MvaT is able to bind and shift GC-rich (*PA3900*) DNA at concentrations as low as 50nM and a fully shifted complex is observed at 300nM indicating that MvaT has a higher affinity for GC-rich (*PA3900*) DNA. 1nM radiolabelled DNA was incubated with MvaT for 30 minutes at room temperature. Complexes were resolved on a 6% native polyacrylamide gel. AT-rich DNA: 289bp fragment of *hilA* promoter (34% GC); GC-rich DNA: 204bp fragment of *PA3900* (74% GC).

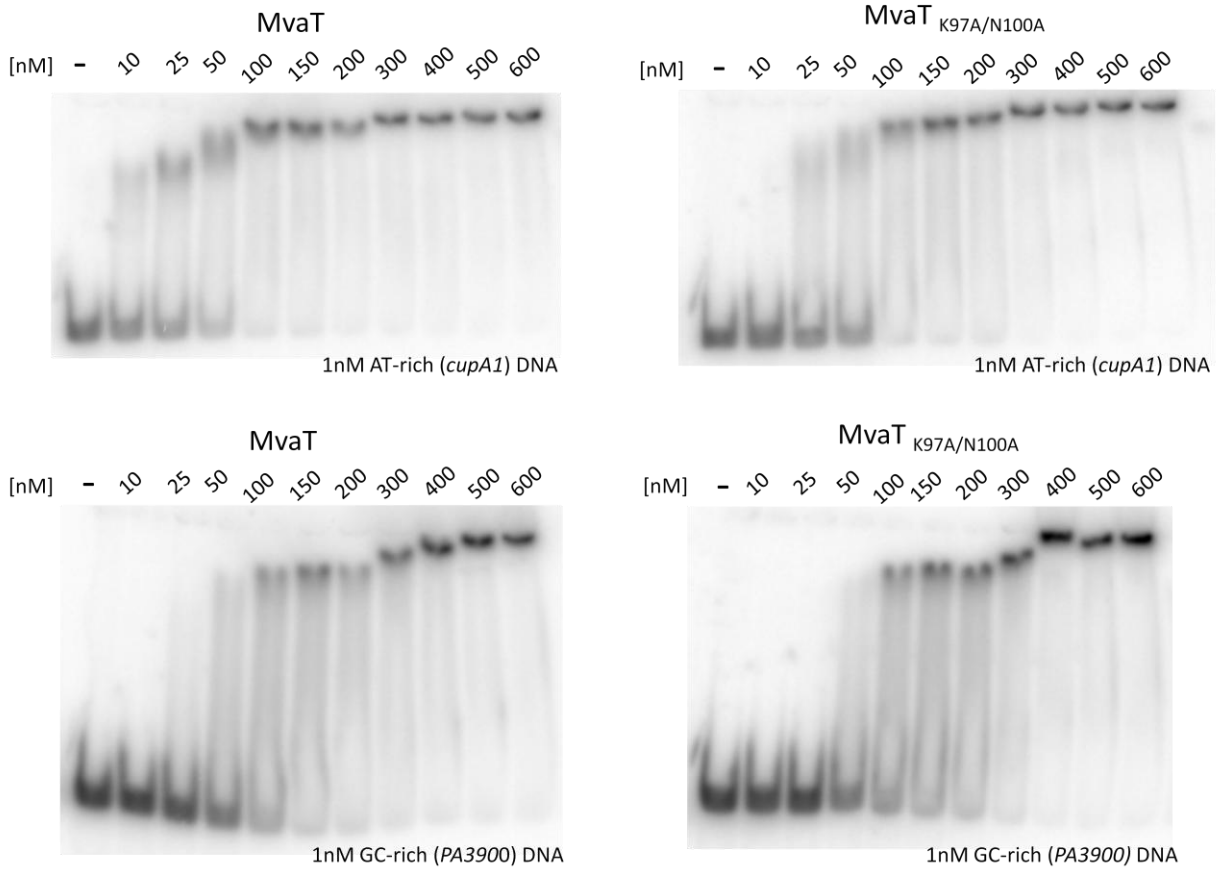


Figure 6: Lys-97 and Asn-100 residues are not essential for DNA binding by MvaT. Mutation of Lys-97 and Asn-100 does not abolish DNA binding and results in an approximately 2-fold loss in affinity for both AT-rich (*cupA1*) and GC-rich (*PA3900*) DNA compared to wild type MvaT. 1nM of radiolabelled DNA was incubated with protein for 30 minutes at room temperature before resolution on a 6% native polyacrylamide gel. AT-rich DNA: 340bp fragment of *cupA1* promoter (54% GC); GC-rich DNA: 204bp fragment of *PA3900* (74% GC).

3.6 Oligomerization of MvaT is Required for Stable DNA Binding

MvaT forms extended filaments when binding DNA, giving this molecule highly cooperative behaviour. Specifically, it appears that MvaT-DNA complexes result from a multiple low-affinity interactions that are coupled to make an apparently higher affinity interaction (i.e. a “Velcro”-like mechanism). This cooperative behaviour makes it difficult to accurately quantify the dissociation constant (K_D) of individual MvaT molecules by EMSA. MvaT- DNA complexes form partially shifted smears, not discrete bands, and there are different types of fully shifted complexes that form at high protein concentrations (Figure 5, 6).

To determine the affinity of MvaT for DNA without interference from its cooperative binding behaviour, I performed gel shift assays with truncated MvaT protein that contained only the DNA-binding domain (MvaT residues 77-124) (Figure 7). Gel shifts with this truncated construct however did not result in discrete shifted complexes as was the case with full-length protein. The resulting smear is indicative that the DNA-binding domain alone is unable to form a stably shifted complex with DNA, which likely results from a weak interaction and high on-off rates. Mutating Asn-100 to alanine in this truncated construct did not further impair complex formation.

To improve complex stability, I increased the gel acrylamide percentage (from 6, up to 10%) to enhance the 'caging' effect. Polyacrylamide matrices stabilize complexes by enclosing the protein and DNA such that even if dissociation occurred, the components cannot diffuse far away and proximity will lend itself to re-association. Increasing the polyacrylamide concentration therefore creates smaller pores and increases the caging effect to prevent diffusion. In addition to increasing the polyacrylamide concentration, I also increased the electrical voltage from 70 volts to 200 volts to minimize and prevent gradual complex dissociation over the lengthy gel run time (165 minute run time reduced to 40 minutes).

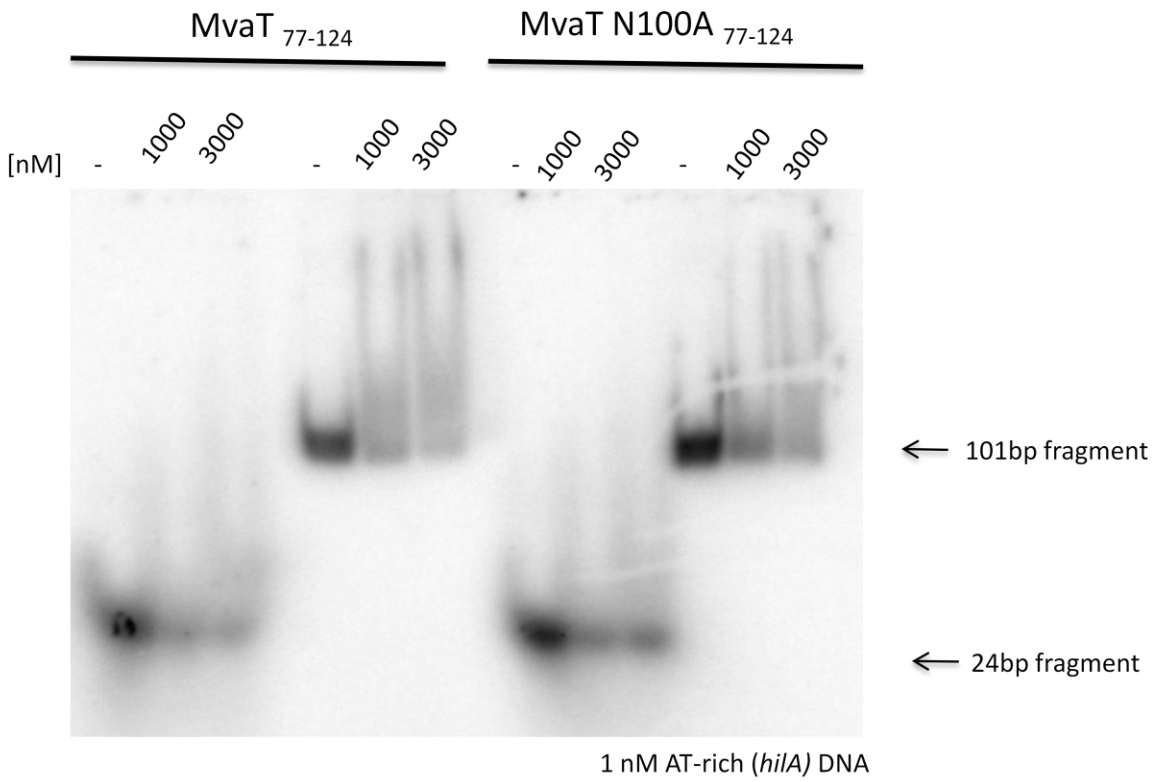
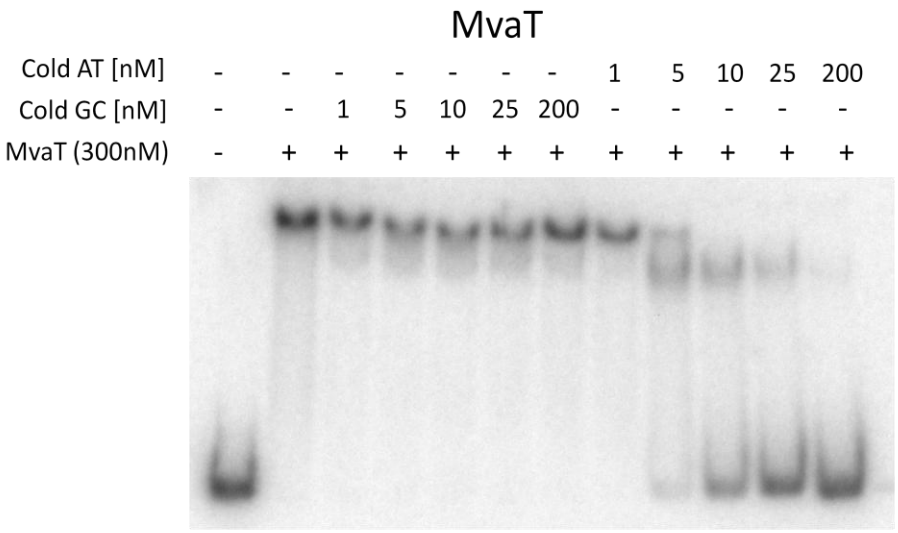


Figure 7: MvaT DNA-binding domain (residues 77-124) does not stably form complexes with AT-rich DNA. Truncated MvaT containing only the C-terminal DNA-binding domain was purified to eliminate the cooperative oligomeric binding of MvaT to DNA. 1nM radiolabelled DNA was incubated with $MvaT_{77\text{-}124}$ for 30 minutes at room temperature and resolved on a 10% native polyacrylamide gel at 200 volts for 40 minutes. Two different sized fragments of AT-rich DNA were used: 24bp and 101bp fragments of *hilA* promoter (34% GC).

Changes in either one of these parameters failed to improve complex stability. The dimerization and oligomerization properties of MvaT are therefore responsible and necessary for the formation of a stable nucleoprotein filament structure on DNA, supporting a model that MvaT-DNA complexes form from the coupling of multiple weak binding interactions. However at this point we cannot exclude the possibility that additional DNA binding contact surfaces lie within the oligomerization domain.

3.7 MvaT strongly prefers AT-rich DNA in competition assays

In vivo, MvaT preferentially and solely binds to AT-rich DNA sequences, a result our lab has corroborated with protein binding microarrays (Castang et al., 2008; unpublished data). *In vitro,* however, MvaT was able to bind to GC-rich DNA and did so with a higher than expected affinity (Figure 5). To reconcile these findings, I sought to determine if MvaT displayed a preference for its *in vivo* target sequence when presented with both AT-rich and GC-rich DNA through competitive gel shift assays. Pre-formed radiolabelled DNA-protein complexes were challenged by the addition of excess cold, unlabelled DNA. Dissociation of the radiolabelled complex by competing cold DNA is visualized as a release or shift down of the radiolabelled DNA back to its free form. Figure 8 demonstrates that MvaT shows a remarkable affinity and specificity for AT-rich DNA. Once bound, even an excess of 200x GC-rich DNA is unable to perturb the bound-complex. Furthermore the MvaT complex pre-formed on radiolabeled AT-rich DNA was not easily competed by cold AT-rich DNA, requiring greater than 25-fold excess competitor for complete dissociation. When the pre-bound complex was challenged with cold otherwise identical AT-rich DNA, two complexes can be seen. These likely represent specific and non-specifically bound DNA complexes. The upper band, with lower mobility, is a fully saturated MvaT-DNA complex and the lower band represents a complex where non-specifically bound MvaT dissociated away and only MvaT bound to high affinity sites remained. In



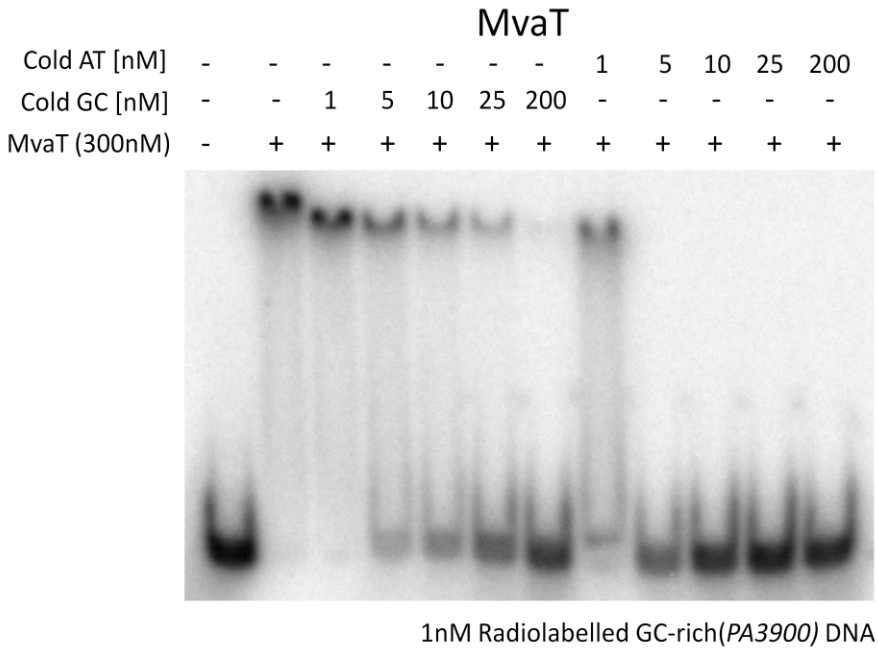


Figure 8: MvaT strongly prefers AT-rich DNA over GC-rich DNA. The presence of even 200x excess cold GC-rich (*PA3900*) DNA was unable to dissociate the pre-bound MvaT-AT-rich (*cupA1*) DNA complex. Conversely, GC-rich (*PA3900*) bound DNA complexes were quickly dissociated in the presence of the preferred AT-rich (*cupA1*) DNA sequence. Pre-bound complexes of MvaT (300nM) with 1nM radiolabelled AT-rich DNA (incubated for 15 minutes at room temperature) were challenged with the addition of excess unlabelled DNA. Samples were incubated for an additional 15 minutes at room temperature and resolved on a 6% native polyacrylamide gel. AT-rich DNA: 340bp fragment of *cupA1* promoter (54% GC); GC-rich DNA: 204bp fragment of *PA3900* (74% GC).

contrast to the strong association of MvaT to AT-rich DNA, MvaT bound to GC-rich DNA is very easily dissociated in the presence of the preferred AT-rich target. The addition of as little as 5x excess of AT-rich DNA results in complete dissociation of the original MvaT-GC-rich DNA complex. Altogether, this data suggests that although MvaT is able to bind to GC-rich DNA (Figure 5) it does not form stable complexes and that localized regions of the nucleoprotein filament may briefly dissociate from the DNA. On AT-rich DNA MvaT oligomers appear to form a more stable complex and therefore cannot be removed by excess competitor DNA.

3.8 MvaT-AT-rich DNA Complexes are More Stable

Salt-induced dissociation of MvaT-DNA complexes confirmed that the MvaT-AT-rich DNA complex is more stable (Figure 9). Pre-formed complexes of MvaT with either AT-rich or GC-rich DNA were challenged with increasing concentrations of salt. With AT-rich DNA, complex dissociation began at salt concentrations of 500mM compared to GC-rich DNA where complex dissociation began at 100-200mM NaCl. Together with Figure 8, these results demonstrate that MvaT has a strong affinity and preference for AT-rich DNA and forms much stabler complexes with target AT-rich DNA *in vitro*.

3.9 MvaT Lys-97/Asn-100 Residues Do Not Alter Strong Preference for AT-rich DNA

The AT-hook in H-NS has two roles: i) binding DNA and ii) distinguishing between target AT-rich DNA and non-target GC-rich DNA. Figure 6 demonstrates that Lys-97 and Asn-100 are not essential for the former. I therefore tested whether or not these residues play a role in the latter, helping MvaT differentiate between AT-rich and GC-rich DNA. Using competition assays as described above, I challenged bound complexes of MvaT K97A/N100A with excess cold DNA.

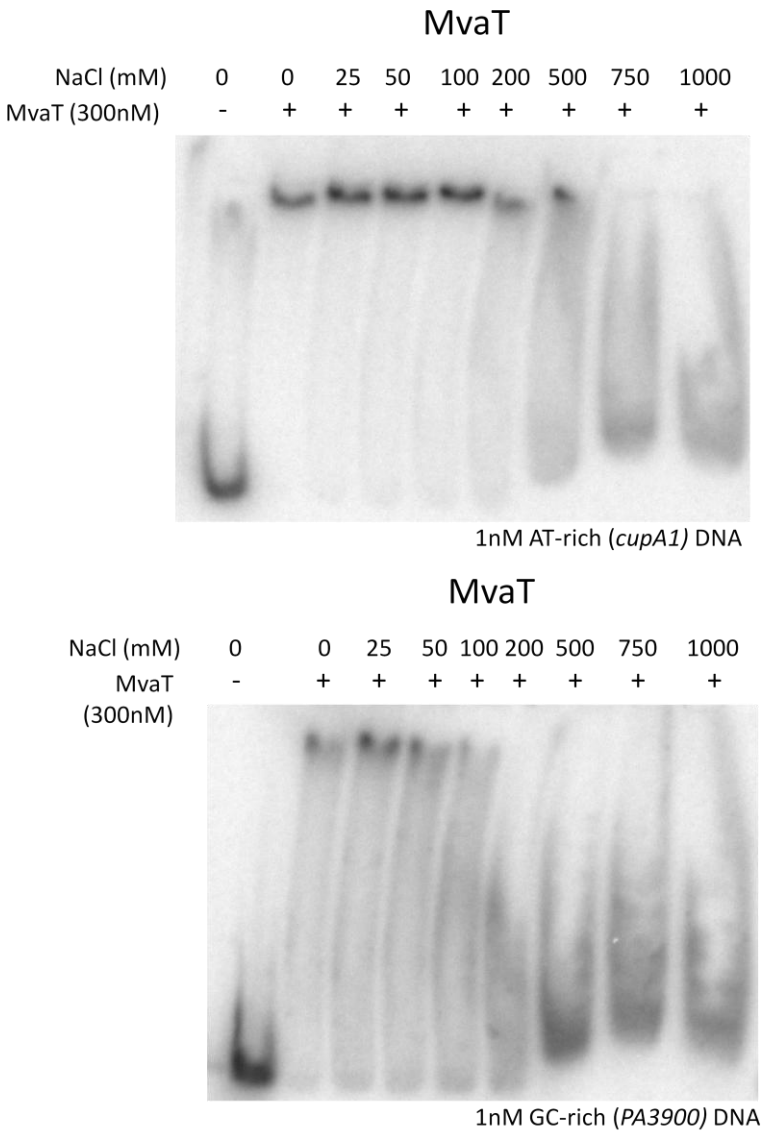


Figure 9: MvaT forms more stable complex with AT-rich DNA. Salt-induced dissociation of MvaT complexes show that MvaT-AT-rich (*cupA1*) DNA complexes are more resistant to dissociation. Pre-formed MvaT-DNA complexes were challenged by varying concentrations of salt. 1nM radiolabelled DNA was incubated with 300nM MvaT for 15 minutes at room temperature. Salt solutions were added and incubated for a further 15 minutes at room temperature before being resolved on a 6% native polyacrylamide gel. High salt concentrations affected DNA migration and resulted in warped lanes. AT-rich DNA: 340bp fragment of *cupA1* promoter (54% GC); GC-rich DNA: 204bp fragment of *PA3900* (74% GC).

Comparing Figure 8 with Figure 10, it is clear that alanine mutations to Lys-97 and Asn-100 do not alter the behaviour of MvaT complexes on AT-rich DNA. As such, lysine-97 and asparagine-100 of MvaT are not essential for the discrimination of AT-rich target DNA from GC-rich non-target DNA *in vitro*. Along with Figure 6, Figure 10 establishes that the KGGN motif in MvaT is not synonymous with the AT-hook motif found in H-NS and Lsr2. Alanine mutations to these residues did not significantly alter MvaT DNA binding and recognition of preferred AT-rich DNA *in vitro*. Therefore, despite strong structural similarity between the KGGN motif in MvaT and the QGR and RGR motifs in H-NS and Lsr2 respectively, based on the above data, the KGGN motif does not represent an analogous AT-hook motif and the mechanism that MvaT employs to bind DNA and recognize AT-rich targets is still unknown.

3.10 Attempts to Probe Differential Local Off-rate Model using Nuclease Protection Assays

EMSA assays with MvaT and AT-rich and GC-rich DNA demonstrated that MvaT was able to bind non-target GC-rich DNA with a comparable affinity to that of target AT-rich DNA (only a ~2.5 fold reduction) (Figure 5). However, competitive assays demonstrated a very strong preference for AT-rich DNA (Figure 8). To reconcile these two findings, we propose a model discussed in section 4.2. Briefly, we propose that MvaT binds weakly to GC-rich DNA but is able to form stable oligomers on GC-rich DNA as a result of cooperative binding (multiple weak binding events result in a stable complex). MvaT however, prefers and binds much tighter to AT-rich DNA and is not easily dissociated once bound. We propose that whilst bound to DNA, MvaT units are connected to one another via their N-terminal and central oligomerization domains and form an oligomer, but individual binding domains can freely dissociate and associate with DNA in accordance to their differential affinities and off-rates for AT-rich and

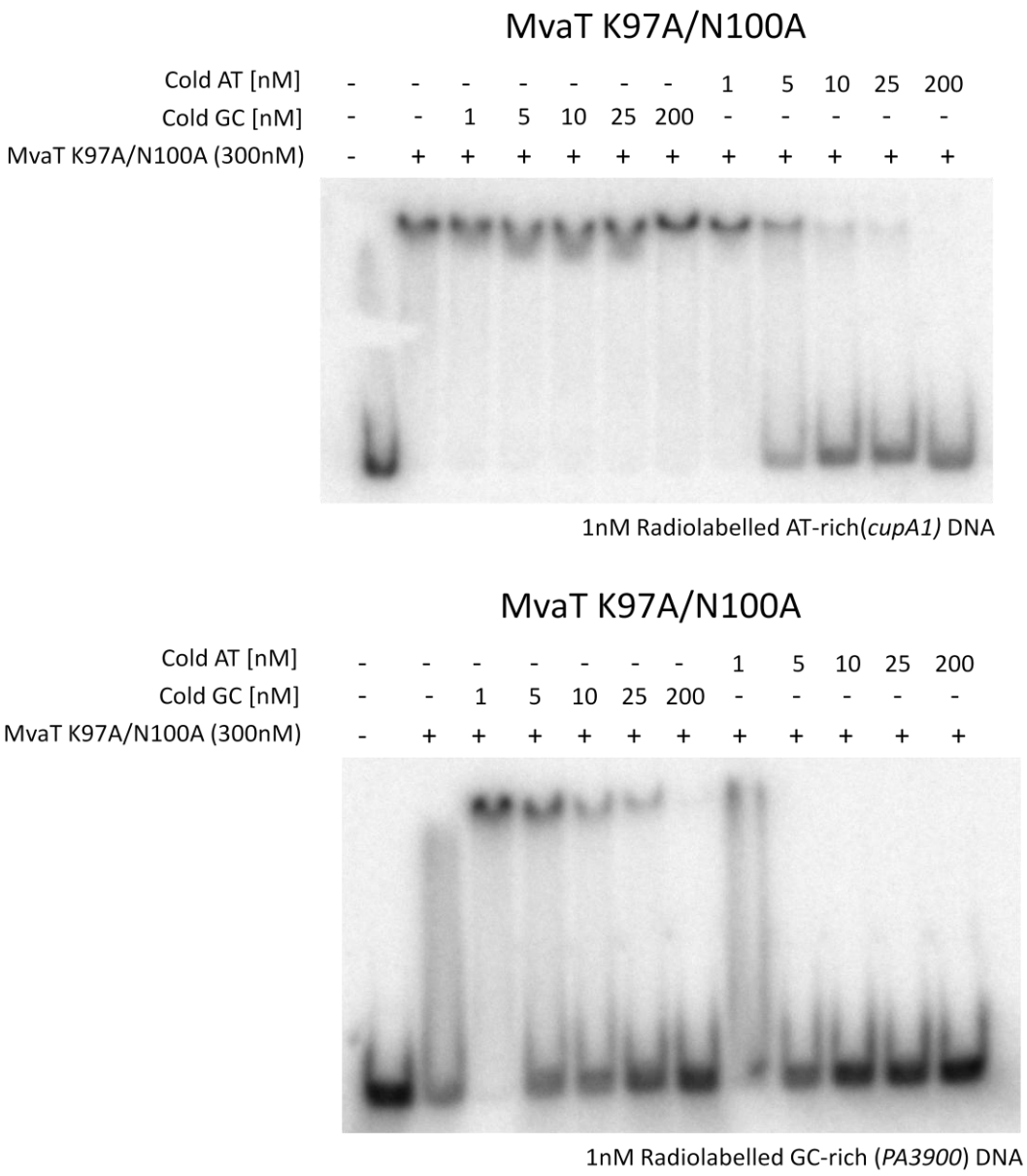


Figure 10: Alanine mutations to Lys-97 and Asn-100 do not alter preference for AT-rich DNA. MvaT K97A/N100A does not display an altered preference for AT-rich (*cupA1*) DNA compared to wild type MvaT (Figure 8). Pre-bound complexes of MvaT K97A/N100A (300nM) with 1nM radiolabelled AT-rich DNA (incubated for 15 minutes at room temperature) were challenged with the addition of excess unlabelled DNA. Samples were incubated for an additional 15 minutes at room temperature and resolved on a 6% native polyacrylamide gel. AT-rich DNA: 340 bp fragment of *cupA1* promoter (54% GC); GC-rich DNA: 204bp fragment of *PA3900* (74% GC).

GC-rich DNA. In this way, MvaT oligomers bound to DNA form breathable complexes and we propose that MvaT bound to GC-rich DNA forms a more breathable and flexible oligomer compared to that formed on AT-rich DNA, which is a tighter association. To test this model, we reasoned that the 'breathability' of MvaT bound to GC-rich DNA might leave DNA vulnerable to nuclease digestion while AT-rich DNA would be better protected against nucleases in the presence of MvaT. Two nucleases were selected: DNase I and Exonuclease III for nuclease protection assays. DNase I is an endonuclease and the traditional enzyme used for protection assays (DNA footprinting). Caveats with using DNase I however include its higher affinity for cleaving AT-rich DNA due to enhanced binding in the narrow minor groove of AT-rich DNA (Suck, 1994; Lazarovici et al., 2013), as well as the fact that my DNA probe is 5' labelled and a single nick downstream of this label would result in a loss of visible probe. Exonuclease III is a 3' to 5' exonuclease and was chosen as an alternative nuclease to preserve the 5' radiolabelled phosphate. Figure 11 depicts the results of this nuclease protection assay. MvaT did not offer significant protection from DNase I and ExoIII degradation (there is a reduction in band intensity in lanes containing MvaT and nuclease for both AT-rich and GC-rich DNA compared to lanes with MvaT alone), and a differential protection of AT-rich versus GC-rich DNA was not seen at multiple concentrations of enzyme (Figure 11-data only shown for one concentration). Altogether, nuclease protection assays were unable to shed light on the differential localized off rates of MvaT from AT-rich and GC-rich DNA.

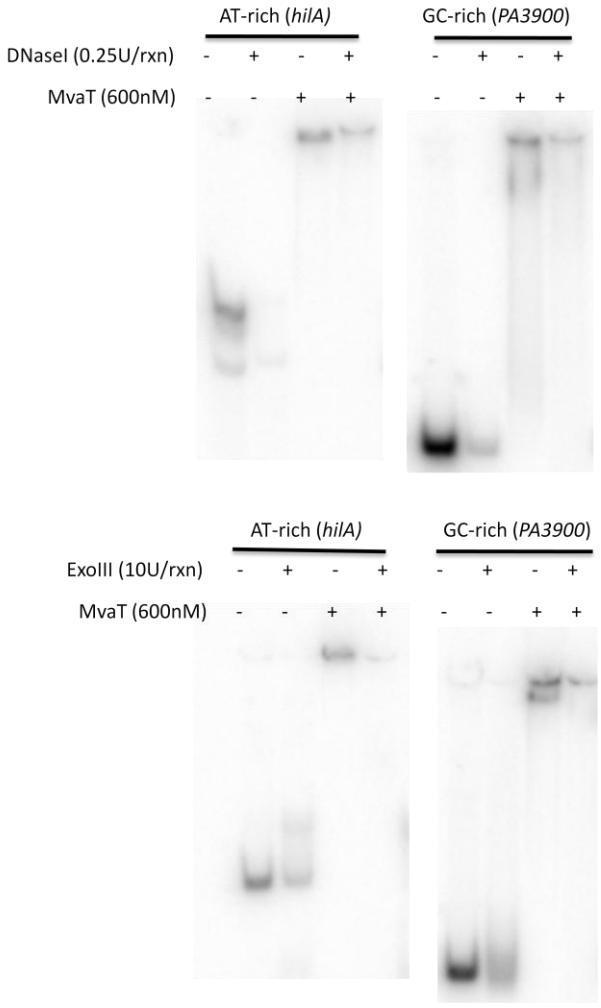


Figure 11: MvaT does not differentially protect AT-rich better than GC-rich DNA from DNA nucleases. MvaT-bound DNA complexes did not offer significant protection from nucleases. AT-rich (*hilA*) DNA was not better protected than GC-rich (*PA3900*) DNA. 1nM radiolabelled DNA was incubated with 600nM MvaT for 15 minutes at room temperature. DNase I and Exonuclease III was added and incubated at room temperature for 10 minutes and 30 minutes respectively. Enzymes were heat inactivated at 70°C for 20 minutes (DNase I was also heat inactivated in the presence of 5mM EDTA), cooled to room temperature for 20 minutes and resolved on a 1nM radiolabelled DNA was incubated with 600nM MvaT for 15 minutes at room temperature. DNase I and Exonuclease III was added and incubated at room temperature for 10 minutes and 30 minutes respectively. Enzymes were heat inactivated at 70°C for 20 minutes (DNase I was also heat inactivated in the presence of 5mM EDTA), cooled to room temperature for 20 minutes and resolved on a native polyacrylamide gel. AT-rich DNA: 289bp fragment of *hilA* promoter (34% GC); GC-rich DNA: 204bp fragment of *PA3900* (74% GC). native polyacrylamide gel. AT-rich DNA: 289bp fragment of *hilA* promoter (34% GC); GC-rich DNA: 204bp fragment of *PA3900* (74% GC).

4 Discussion

4.1 MvaT KGGN Motif is Not an AT-hook

H-NS, Lsr2 and MvaT belong to a family of xenogeneic silencing proteins. They function in their host bacteria as transcriptional repressors of AT-rich DNA, a characteristic feature of many horizontally acquired genes. In doing so, they act as genome sentinels, allowing bacteria to participate in horizontal gene transfer without suffering fitness consequences resulting from inappropriate expression of these foreign genes (Navarre et al., 2007). The ability for these proteins to discriminate and bind AT-rich DNA lies in their C-terminal DNA binding domain. The structure of the DNA binding domain of H-NS and Lsr2 has been previously characterized (Figure 2) and a small positively charged motif, QGR/RGR (coined the AT-hook motif) is responsible for recognizing the narrow and more electronegative minor groove of AT-rich DNA (Gordon et al., 2010; Gordon et al., 2011). Mutation of these positively charged residues to alanine completely abolished DNA binding, highlighting their importance (Gordon et al., 2011). The structure and characterization of the DNA binding domain of MvaT has not been reported in the literature.

In 2013, our collaborator Dr. Bin Xia at Peking University provided us with a structure of the DNA binding domain of MvaT (Figure 3). Based on the NMR structure, certain residues that made contact with DNA and were highly conserved between all MvaT-like molecules were of interest to us. Specifically, K97 and N100 inserted into the minor groove of DNA such that they pointed away from each other, much in the same way that the glutamine and arginine residues did in H-NS and Lsr2. We were therefore interested in determining whether or not this conserved KGGN motif functioned as an elongated AT-hook motif in MvaT. Alanine mutations to these residues yielded a double mutant MvaT K97A/N100A and its DNA binding profile was

determined using EMSAs. My results indicate that despite structural data, the KGGN motif does not act analogous to the AT-hook motif. Alanine mutations did not abolish DNA binding (albeit it did slightly reduce binding affinity) and did not alter MvaT's preference for AT-rich DNA *in vitro* (Figure 6, 10). Surprisingly, in the absence of competitive DNA, MvaT was able to bind to GC-rich DNA with comparable affinity (Figure 5). This is in contrast to H-NS, which has a clearer preference for AT-rich DNA (Figure 4). Together these results demonstrate that the KGGN motif in MvaT is not analogous to the AT-hook motif found in H-NS and Lsr2. All the work done to characterize the DNA binding domain of MvaT thus far has been *in vitro.* And while evidence in this study suggests that MvaT employs a unique and still uncharacterized mechanism to bind AT-rich genes, to definitively conclude that the KGGN motif does not play a role in DNA binding, *in vivo* experiments with MvaT K97A/N100A need to be conducted. A simple assay would be to determine *in vivo* binding of the double alanine mutant in *Pseudomonas aeruginosa* through ChIP-chip and compare the results against that of wild type MvaT as performed by Castang et al., 2008.

Though the KGGN motif is not essential for DNA binding or specificity, there are slight quantitative and qualitative differences between wild type MvaT and MvaT K97A/N100A proteins binding to DNA. Figure 6 demonstrates that the double mutant does have a slightly reduced affinity for DNA (approximately 2 fold). Binding to AT-rich DNA is also slightly different between wild type and the double mutant (Figure 8 and 10 top panels right side). Upon competition with cold AT-rich DNA, MvaT is released from radiolabelled AT-rich DNA. With wild type MvaT, we see the presence of a second, higher mobility complex when greater than 5nM (5 times) excess cold DNA is added. The formation of a smaller second band indicates that the AT-rich DNA fragment contained binding sites with high affinity and lower affinity. Upon competition, MvaT bound to sites of lower affinity dissociated, however MvaT bound to high affinity sites remained. The reduction of MvaT molecules bound to the radiolabelled DNA

results in the formation of this second smaller, higher mobility complex. This second band is not seen with MvaT K97A/N100A indicating that the mutations do affect DNA binding in some way, however this alteration is not enough to cause drastic defects as observed with the H-NS AT-hook double mutant (Figure 4).

4.2 MvaT Differential Local Off-Rates from AT-rich and GC-rich DNA: a Proposed Model

Based on our current data, we propose a model in which, unlike H-NS, MvaT is able to form stable nucleoprotein filaments on both AT-rich and GC-rich DNA. However the oligomeric structure formed on each is different. We propose that within the oligomeric nucleofilament structure that MvaT forms, there is localized association and dissociation from DNA such that individual units of MvaT are able to bind and release from DNA as dictated by entropy. Localized refers to individual units/molecules of MvaT rather than the entire MvaT oligomer formed on DNA. This localized dissociation from DNA does not disturb the overall bound oligomeric state and MvaT molecules are still attached to adjacent molecules via their oligomerization domains (as this is the thermodynamically stable state). In this way, the MvaT oligomer is a flexible, breathable structure on DNA, where MvaT units within the oligomer are able to bind to and release from the oligomer-coated DNA strand without altering the overarching bound state of the MvaT oligomer on DNA. Our data points to a model where oligomers formed on AT-rich DNA are much tighter, ‘locking’ the protein onto the DNA. MvaT units therefore have a relatively slow localized off-rate from AT-rich DNA. MvaT oligomers on GC-rich DNA, in contrast, are not as rigidly locked on and form a more 'breathable' complex. MvaT units experience a faster localized off-rate from GC-rich DNA. In the scenario of a competition assay, MvaT units bound to AT-rich DNA would not experience fast localized off-rates from DNA and predominantly remain bound to the AT-rich DNA. When however, the

MvaT units do locally dissociate from DNA, and come into contact with competing GC-rich DNA, the association to GC-rich DNA is not long lived (fast localized off-rate with GC-rich DNA) and the MvaT unit releases and once again binds the original AT-rich DNA fragment. When the inverse conditions are tested, MvaT units bound to radiolabelled GC-rich DNA experience a fast localized off -rate and bind to competing AT-rich DNA. These MvaT units once bound to AT-rich DNA stay associated and bring the AT-rich DNA fragment in close proximity to adjacent MvaT molecules bound to the original GC-rich DNA. Because of the proximity of the competing AT-rich DNA and the fast localized off-rate of MvaT from GC-rich DNA, adjacent MvaT units will dissociate from the GC-rich fragment and bind the AT-rich DNA. Once bound to AT-rich DNA, MvaT does not easily dissociate and in this way, competing AT-rich DNA acts as a sink, rapidly drawing away MvaT molecules from previously bound GC-rich DNA. This model reconciles the fact that MvaT is able to form stable shifted complexes with GC-rich DNA (Figure 5) at lower than expected protein concentrations, but demonstrate a stronger affinity for AT-rich DNA in competition assays (Figure 8). This model is illustrated in Figure 12.

4.3 Other Methods to Test Differential Local Off-Rate Model

Nuclease protection assays were attempted here to test this model of differential off-rates between AT-rich and GC-rich DNA, but were unable to demonstrate a preferential protection of AT-rich DNA (Figure 11). 5' labelling of my DNA fragment with γ-^{32}P meant that enzyme cleavage a single nucleotide downstream of my 5' label would result in the loss of visible probe. As such, Exonuclease III was chosen as an alternative nuclease to circumvent this issue. A better approach however, is to label my DNA probes with α-^{32}P, such that every nucleotide is visible. Repeating the DNase I nuclease protection assay with this α-labelled DNA may reveal more about the state of oligomeric MvaT bound to AT-rich and GC-rich DNA.

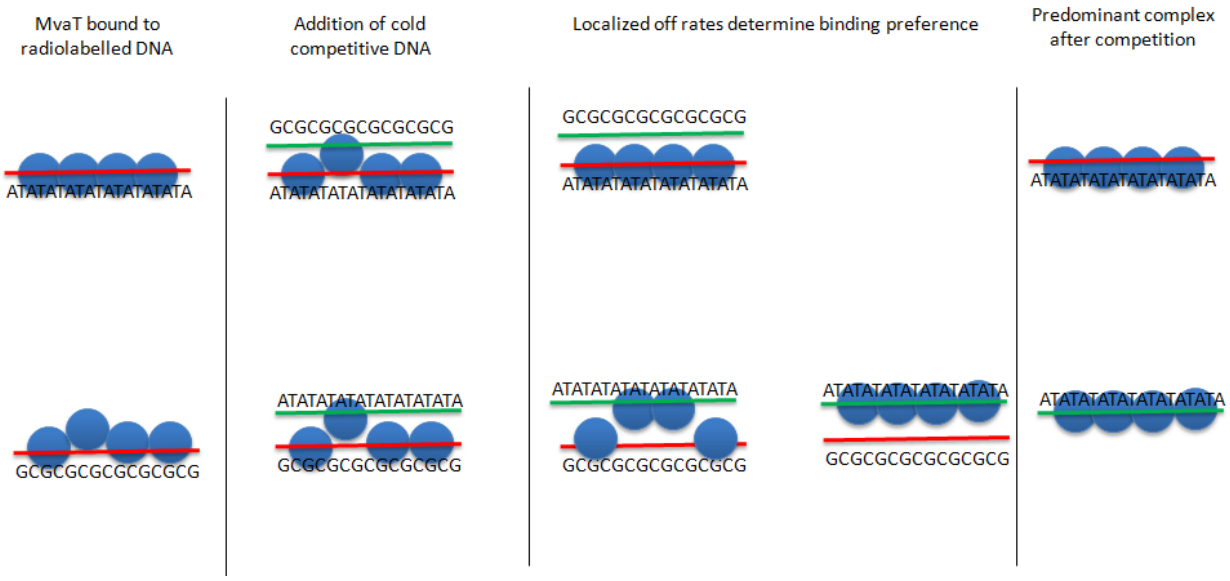


Figure 12: Model of MvaT's faster localized off-rate from GC-rich DNA. Radiolabelled DNA is depicted in red and competing cold DNA is depicted in green. MvaT oligomers bound to DNA are able to locally associate and dissociate from DNA while still maintaining the oligomeric state. Panel 1: MvaT bound to AT-rich DNA is more tightly bound and exhibits a slower localized off-rate. The MvaT oligomer bound to GC-rich DNA is more breathable and has a faster localized off-rate. Panel 2: The addition of cold competing DNA allows MvaT units that have locally dissociated from the radiolabelled DNA to bind to the cold DNA. Panel 3: The differential localized off-rate of MvaT for AT-rich and GC-rich DNA determines whether or not the MvaT units dissociate from cold GC-rich DNA and re-bind the original radiolabelled AT-rich DNA (top) or whether adjacent MvaT molecules also dissociate from the original radiolabelled GC-rich DNA and bind to cold AT-rich DNA. Panel 4 depicts the predominant MvaT-DNA complexes following addition of competitive DNA. In both scenarios, MvaT is bound preferentially to AT-rich DNA.

Another way to test the breathability of MvaT on AT-rich and GC-rich DNA would be through the use of DNA minor groove binding dyes that fluoresce upon binding to DNA. The increase in fluorescence as MvaT subunits dissociate and dye molecules gain access to DNA, could theoretically provide a direct read-out of the breathability of the MvaT oligomer. However, commercially available minor groove binding dyes such as Hoescht and DAPI bind with very high affinity and have been shown to displace known minor groove binding proteins such as histones and TATA-binding protein in eukaryotes (Loontiens et al., 1990; Trotta et al., 1998; Chiang et al., 1994).

Circular dichroism is another potential method used to probe the differential stability of MvaT on AT-rich and GC-rich DNA. This technique measures the differential absorbance of left and right circularly polarized light by optically active (chiral) molecules and can be used to measure the release of MvaT oligomers from DNA as a function of temperature. The differential stability of MvaT bound to AT-rich DNA and GC-rich DNA may also be explained by a structural difference in the MvaT oligomer that forms. If so, circular dichroism could theoretically also be used again to measure the susceptibility of MvaT-DNA complexes to guandium-induced protein denaturation.

4.4 MvaT may be Evolutionarily Related to H-NS

Given the shared functional properties between MvaT and the other two xenogeneic silencers, as well as the structural data that suggested that the KGGN motif inserted into the minor groove of DNA in a similar manner to that of the AT-hook in H-NS/Lsr2, the results we obtained which suggested that the KGGN motif did not function analogous to that of the AT-hook motif (QGR and RGR) were surprising. Phylogenetically, H-NS and MvaT are more closely related than H-NS is to Lsr2. Lsr2 is found among the distant Gram positive actinobacteria whereas H-NS and MvaT are both found within Gram negative proteobacteria.

On a phylogenetic tree, the Pseudomonads are evolutionary offshoots of bacterial species that contain H-NS (Figure 1). It is intriguing therefore that MvaT independently evolved within a narrow group of bacteria where H-NS is more widespread among related species. The reason for this is unclear. Structurally, MvaT is also more similar to H-NS than Lsr2 is to H-NS. The DNA binding domain of MvaT consists of three anti-parallel β-sheets, and two α-helices, which is quite similar to that of H-NS (two anti-parallel β-sheets, one α-helix and one 3_{10} helix) (Figure 13 A).

Sequence alignments indicate that MvaT can be linked to *E. coli* H-NS through H-NS-like molecules encoded by the plant pathogens *Xylella* and *Xanthomonas*. These H-NS-like molecules contain the canonical H-NS AT-hook motif but share similarity with MvaT in regions immediately upstream of the hook. Structural predictions using PHYRE2 (a protein structure prediction program) with the C-terminal amino acid sequence of MvaT yielded H-NS from *Xylella fastidiosa* (63% confidence, 29% identity) as the highest scoring match by a considerable margin. Reiterative Psi-BLASTs with full length MvaT also yields hits in H-NS from *Xanthomonas* and *Xylella* species. It is interesting therefore to note that though MvaT is more closely related to H-NS compared with Lsr2, MvaT does not share the same mechanism of DNA binding as H-NS. It is possible that Lsr2 and H-NS may have arisen independently and covergently evolved the AT-hook motif and MvaT diverged from H-NS to evolve a unique mechanism to target and bind DNA.

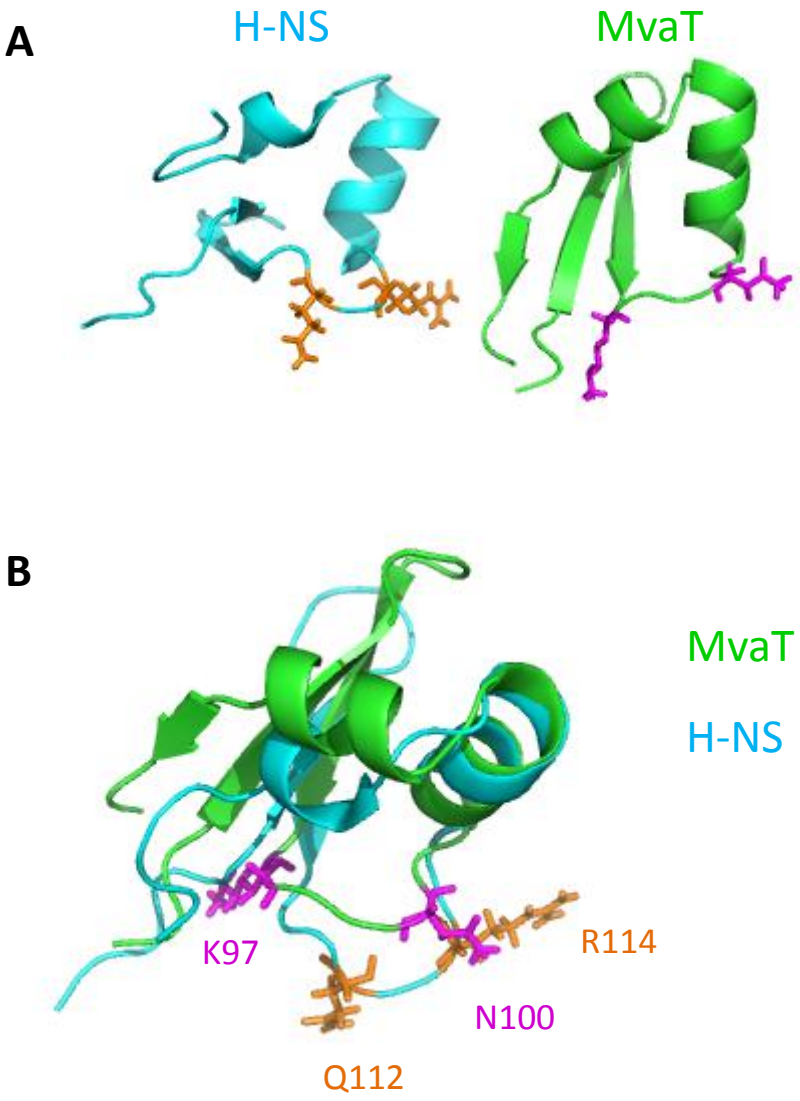


Figure 13: Comparisons of H-NS and MvaT DNA binding domains. A) Side-by-side comparisons of H-NS DNA binding domain (from *Salmonella* Typhimurium LT2 pdb 2l93) and MvaT (Dr. Bin Xia). B) Superimposition of H-NS and MvaT DNA binding domains using the 'super' command in PyMOL.

4.5 Significance

Previous work in the literature has characterized the DNA binding domain of two xenogeneic silencing proteins: H-NS and Lsr2, and found that they share an AT-hook motif that is essential for DNA binding and discrimination between target AT-rich DNA and non-target DNA. The structure and characterization of the DNA binding domain of the third and final xenogeneic silencer MvaT remained elusive until this study, which is the first to investigate the DNA binding domain of MvaT. Even though MvaT and H-NS appear related (evolutionarily and structurally), experiments shown in this thesis demonstrate that they do not share the same DNA binding mechanism (that is, MvaT lacks an apparent AT-hook motif). This is significant because given the identical roles that these proteins play in their host cell, this is the first study to demonstrate stark mechanistic differences among the family of xenogeneic silencers H-NS, Lsr2 and MvaT. Questions remain as to why MvaT and not H-NS is found in *Pseudomonas*, as well as the reason behind the divergent DNA binding mechanism. Further mutagenesis and characterization of MvaT residues in the DNA binding domain will need to be employed to delineate the mechanism by which this xenogeneic silencing protein binds to AT-rich DNA. Altogether, this study sheds light on the xenogeneic silencer MvaT and demonstrates that it does not employ an analogous AT-hook motif to bind DNA.

References

Abby, S.S., Tannier, E., Gouy, M., and Daubin, V. (2012). Lateral gene transfer as a support for the tree of life. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 4962-4967

Ali, S.S., Beckett, E., Bae, S.J., and Navarre, W.W. (2011). The 5.5 protein of phage T7 inhibits H-NS through interactions with the central oligomerization domain. *J. Bacteriol.* **193**, 4881-4892

Ali, S.S., Xia, B., Liu, J., and Navarre, W.W. (2012). Silencing of foreign DNA in bacteria. *Curr. Opin. Microbiol.* **15,** 175-181

Amit, R., Oppenheim, A.B., and Stavans, J. (2003). Increased bending rigidity of single DNA molecules by H-NS, a temperature and osmolarity sensor. *Biophys. J.* **84**, 2467-2473

Aravind, L., and Landsman, D. (1998). AT-hook motifs identified in a wide variety of DNA-binding protein. *Nucleic Acids Res.* **26**, 4413-4421

Arold, S.T., Leonard, P.G., Parkinson, G.N., and Ladbury, J.E. (2010). H-NS forms a superhelical protein scaffold for DNA condensation. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 15728-15732

Barth, M., Marschall, C., Muffler, A., Fischer, D., and Hengge-Aronis, R. (1995). Role for the histone-like protein H-NS in growth phase-dependent and osmotic regulation of sigma S and many sigma S-dependent genes in *Escherichia coli*. *J. Bacteriol.* **177**, 3455-3464

Bertin, P., Terao, E., Lee, E.H., Lejeune, P., Colson, C., Danchin, A., and Collatz, E. (1994). The H-NS protein is involved in the biogenesis of flagella in *Escherichia coli*. *J. Bacteriol.* **176**, 5537-5540

Blot, N., Mavathur, R., Geertz, M., Travers, A., and Muskhelishvili, G. (2006). Homeostatic regulation of supercoiling sensitivity coordinates transcription of the bacterial genome. *EMBO Rep.* **7**, 710-715

Boto, L. (2010). Horizontal gene transfer in evolution: facts and challenges. *Proc. Biol.Sci.* **277**, 819-827

Bouffartigues, E., Buckle, M., Badaut, C., Travers, A., and Rimsky, S. (2007). H-NS cooperative binding to high-affinity sites in a regulatory element results in transcriptional silencing. *Nat. Struct. Mol. Biol.* **14,** 441-448

Bustamante, V.H., Santana, F.J., Calva, E., and Puente, J.L. (2001). Transcriptional regulation of type III secretion genes in enteropathogenic *Escherichia coli*: Ler antagonizes H-NS-dependent repression. *Mol. Microbiol.* **39**, 664-678

Castang, S., and Dove, S.L. (2010). High-order oligomerization is required for the function of the H-NS family member MvaT in *Pseudomonas aeruginosa*. *Mol Microbiol.***78**, 916-931

Castang, S., McManus, H.R., Turner, K.H., and Dove, S.L. (2008). H-NS family members function coordinately in an opportunistic pathogen. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 18947-18952

Chen, C.C., and Wu, H.Y. (2005). LeuO protein delimits the transcriptionally active and repressive domains on the bacterial chromosome. *J. Biol. Chem.* **280**, 15111-15121

Chen, J.M., Ren, H., Shaw, J.E., Wang, Y.L., Li., M., Leung, A.S., Tran, V., Berbenetz, N.M., Kocíncová, D., Yip, C.M., Reyrat, J.M and Liu, J. (2008). Lsr2 of *Mycobacterium tuberculosis* is a DNA-bridging protein. *Nucleic. Acids Res.* **36**, 2123-2135

Chen, J.M., German, G.J., Alexander, D.C., Ren, H., Tan, T., and Liu, J. (2006). Roles of Lsr2 in colony morphology and biofilm formation of *Mycobacterium smegmatis*. *J. Bacteriol.* **188**, 633-641

Chiang, S.Y., Welch, J., Rauscher, F.J. 3rd., and Beerman, T.A. (1994). Effects of minor groove binding drugs on the interaction of TATA box binding protein and TFIIA with DNA. *Biochemistry* **33**, 7033-7040

Colangeli, R., Helb, D., Vilchèze, C., Hazbón, M.H., Lee, C.G., Safi, H., Sayers, B., Sardone, I., Jones, M.B., Fleischmann, R.D., Peterson, S.N., Jacobs, W.R. Jr., and Alland, D. (2007). Transcriptional regulation of multi-drug tolerance and antibiotic-induced responses by the histone-like protein Lsr2 in *M. tuberculosis*. *PLoS Pathog.* **3**, e87

Dame, R.T., Wyman, C., and Goosen, N. (2000). H-NS mediated compaction of DNA visualised by atomic force microscopy. *Nucleic Acids Res.* **28**, 3504-3510

Daubin, V., Lerat, E., and Perrière, G. (2003). The source of laterally transferred genes in bacterial genomes. *Genome Biol.* **4**, R57

Defez, R., and De Felice, M. (1981). Cryptic operon for beta-glucoside metabolism in *Escherichia coli* K12: genetic evidence for a regulatory protein. *Genetics* **97**, 11-25

Diggle, S.P., Winzer, K., Lazdunski, A., Williams, P., and Cámara, M. (2002). Advancing the quorum in *Pseudomonas aeruginosa*: MvaT and the regulation of N-acylhomoserine lactone production and virulence gene expression. *J. Bacteriol.* **184**, 2576-2586

Dole, S., Nagarajavel, V., and Schnetz, K. (2004). The histone-like nucleoid structuring protein H-NS represses the *Escherichia coli bgl* operon downstream of the promoter. *Mol. Microbiol.* **52**, 589-600

Dorman, C.J. (2007). H-NS, the genome sentinel. *Nat. Rev. Microbiol.* **5**, 157-161

Dorman, C.J. (2009). Regulatory integration of horizontally-transferred genes in bacteria. *Front Biosci. (Landmark Ed).* **14**, 4103-4112

Dubnau, D. (1999). DNA uptake in bacteria. *Annu. Rev. Microbiol.* **53**, 217-244

Esposito, D., Petrovic, A., Harris, R., Ono, S., Eccleston, J.F., Mbabaali, A., Haq, I., Higgins, C.F., Hinton, J.C., Driscoll, P.C., and Ladbury, J.E. (2002). H-NS oligomerization domain structure reveals the mechanism for high order self-association of the intact protein. *J. Mol. Biol.* **324**, 841-850

Falconi, M., Brandi, A., La Teana, A., Gualerzi, C.O., and Pon, C.L. (1996). Antagonistic involvement of FIS and H-NS proteins in the transcriptional control of *hns* expression. *Mol. Microbiol.* **19**, 965-975

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

Gomez-Gomez, J.M., Blazquez, J., Baquero, F., and Martinez, J.L. (1996). *Hns* mutant unveils the presence of a latent haemolytic activity in *Escherichia coli* K-12. *Mol. Microbiol.* **19**, 909-910

Gordon, B.R., Imperial, R., Wang, L., Navarre, W.W., and Liu, J. (2008). Lsr2 of *Mycobacterium* represents a novel class of H-NS-like proteins. *J. Bacteriol.* **190**, 7052-7059

Gordon, B.R., Li, Y., Cote, A., Weirauch, M.T., Ding, P., Hughes, T.R., Navarre, W.W., Xia, B., and Liu, J. (2011). Structural basis for recognition of AT-rich DNA by unrelated xenogeneic silencing proteins. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 10690-10695

Gordon, B.R., Li, Y., Wang, L., Sintsova, A., van Bakel, H., Tian, S., Navarre, W.W., Xia, B., and Liu, J. (2010). Lsr2 is a nucleoid-associated protein that targets AT-rich sequences and virulence genes in *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 5154-5159

Grainger, D.C., Hurd, D., Goldberg, M. D., and Busby, S.J. (2006). Association of nucleoid proteins with coding and non-coding segments of the *Escherichia coli* genome. *Nucleic Acids Res.* **34**, 4642-4652

Groisman, E.A., Sturmoski, M.A., Solomon, F.R., Lin, R., and Ochman, H. (1993). Molecular, functional, and evolutionary analysis of sequences specific to *Salmonella*. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 1033-1037

Haraga, A., Ohlson, M.B., and Miller, S.I. (2008). *Salmonellae* interplay with host cells. *Nat. Rev. Microbiol.***6**, 53-66

Harrison, J.A., Pickard, D., Higgins, C.F., Khan, A., Chatfield, S.N., Ali, T., Dorman, C.J., Hormaeche, C.E., and Dougan, G. (1994). Role of *hns* in the virulence phenotype of pathogenic s*almonellae*. *Mol. Microbiol.* **13**, 133-140

Higgins, C.F., Dorman, C.J., Stirling, D.A., Waddell, L., Booth, I.R., May, G., and Bremer, E. (1988). A physiological role for DNA supercoiling in the osmotic regulation of gene expression in *S.* typhimurium and *E. coli*. *Cell* **52**, 569-584

Huth, J.R., Bewley, C.A., Nissen, M.S., Evans, J.N, Reeves, R, Gronenborn, A.M., and Clore,

G.M. (1997). The solution structure of an HMG-I(Y)-DNA complex defines a new architectural minor groove binding motif. *Nat. Struct. Biol.* **4**, 657-665

Jacquet, M., Cukier-Kahn, R., Pla, J., and Gros, F. (1971). A thermostable protein factor acting on *in vitro* DNA transcription. *Biochem. Biophys. Res. Commun.* **45**, 1597-1607

Jain, R., Rivera, M.C., Moore, J.E., and Lake, J.A. (2003). Horizontal gene transfer accelerates genome innovation and evolution. *Mol. Biol. Evol.* **20**, 1598-1602

Katayama, Y., Ito, T., and Hiramatsu, K. (2000). A new class of genetic element, *Staphylococcus* cassette chromosome *mec*, encodes methicillin resistance in *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* **44**, 1549-1555

Koonin, E.V., Makarova, K.S., and Aravind, L. (2001). Horizontal gene transfer in prokaryotes: quantification and classification. *Annu. Rev. Microbiol.* **55**, 709-742

Laal, S., Sharma, Y.D., Prasad, H.K., Murtaza, A., Singh, S., Tangri, S., Misra, R.S., and Nath, I. (1991). Recombinant fusion protein identified by lepromatous sera mimics native *Mycobacterium leprae* in T-cell responses across the leprosy spectrum. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 1054-1058

Lawrence, J.G., and Ochman, H. (1997). Amelioration of bacterial genomes: rates of change and exchange. *J. Mol. Evol.* **44**, 383-397

Lawrence, J.G., and Ochman, H. (1998). Molecular archaeology of the *Escherichia coli* genome. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 9413-9417

Lazarovici, A., Zhou, T., Shafer, A., Dantas Machado, A.C., Riley, T.R., Sandstrom, R., Sabo, P.J., Lu, Y., Rohs, R., Stamatoyannopoulos, J.A., and Bussemaker, H.J. (2013) Probing DNA shape and methylation state on a genomic scale with DNase I. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 6376-6381

Lejeune, P., Bertin, P., Walcon, C., Willemot, K., Colson, C., and Danchin, A. (1989). A locus involved in kanamycin, chloramphenicol and L-serine resistance is located in the *bglY-galU* region of the *Escherichia coli* K12 chromosome. *Mol. Gen. Genet.* **218**, 361-363

Lithgow, J.K., Haider, F., Roberts, I.S., and Green, J. (2007). Alternate SlyA and H-NS nucleoprotein complexes control *hylE* expression in *Escherichia coli* K-12. *Mol. Microbiol.* **66**, 685-698

Liu, Y., Chen, H., Kenney, L. J., and Yan, J. (2010). A divalent switch drives H-NS/DNA-binding conformations between stiffening and bridging modes. *Genes Dev.* **24**, 339-344

Loontiens, F.G., Regenfuss, P., Zechel, A., Dumortier, L., and Clegg, R.M. (1990). Binding characteristics of Hoechst 33258 with calf thymus DNA, poly[d(A-T)], and d(CCGGAATTCCGG): multiple stoichiometries and determination of tight binding with a wide spectrum of site affinities. *Biochemistry* **29**, 9029-9039

Lucchini, S., Rowley, G., Goldberg, M.D., Hurd, D., Harrison, M., and Hinton, J.C. (2006). H-NS mediates the silencing of laterally acquired genes in bacteria. *PLoS Pathog.* **2**, e81

Lucht, J.M., Dersch, P., Kempf, B., and Bremer, E. (1994). Interactions of the nucleoid-associated DNA-binding protein H-NS with the regulatory region of the osmotically controlled *proU* operon of *Escherichia coli*. *J. Biol. Chem.* **269**, 6578-6586

Marri, P.R., and Golding, G.B. (2008). Gene amelioration demonstrated: the journey of nascent genes in bacteria. *Genome* **51**, 164-168

McDaniel, T.K., and Kaper, J.B. (1997). A cloned pathogenicity island from enteropathogenic *Escherichia coli* confers the attaching and effacing phenotype on *E. coli* K-12. *Mol. Microbiol.* **23**, 399-407

Navarre, W.W., McClelland, M., Libby, S.J., and Fang, F.C. (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

Navarre, W.W., Porwollik, S. Wang, Y., McClelland, M., Rosen, H., Libby, S.J., and Fang, F.C. (2006). Selective silencing of foreign DNA with low GC content by the H-NS protein in *Salmonella*. *Science* **313**, 236-238

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

Olekhnovich, I.N., and Kadner, R.J. (2006). Crucial roles of both flanking sequences in silencing of the *hilA* promoter in S*almonella enterica*. *J. Mol. Biol.* **357**, 373-386

Oshima, T., Ishikawa, S., Kurokawa, K., Aiba, H., and Ogasawara, N. (2006). *Escherichia coli* histone-like protein H-NS preferentially binds to horizontally acquired DNA in association with RNA polymerase. *DNA Res.* **13**, 141-153

Porwollik, S., and McClelland, M. (2003). Lateral gene transfer in *Salmonella*. *Microbes Infect.* **5**, 977-989

Rohs, R., West, S.M., Sosinsky, A., Liu, P., Mann, R.S., and Honig, B. (2009). The role of DNA shape in protein-DNA recognition. *Nature* **461**, 1248-1253

Rosenthal, R.S., and Rodwell, V.W. (1988). Purification and characterization of the heteromeric transcriptional activator MvaT of the *Pseudomonas mevalonii mvaAB* operon. *Protein Sci.* **7**, 178-184

Schnetz, K. (1995). Silencing of *Escherichia coli bgl* promoter by flanking sequence elements. *EMBO J.* **14**, 2545-2550

Schnetz, K., and Wang, J.C. (1996). Silencing of the *Escherichia coli bgl* promoter: effects of template supercoiling and cell extracts on promoter activity *in vitro*. *Nucleic Acids Res.* **24**, 2422-2428

Schröder, O., and 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

Shin, M., Song, M., Rhee, J.H., Hong, Y., Kim, Y.J., Seok, Y.J., Ha, K.S., Jung, S.H., and Choy, H.E. (2005). DNA looping-mediated repression by histone-like protein H-NS: specific requirement of Esigma70 as a cofactor for looping. *Genes Dev.* **19**, 2388-2398

Shindo, H., Ohnuki, A., Ginba, H., Katoh, E., Ueguchi, C., Mizuno, T., and Yamazaki, T. (1999). Identification of the DNA binding surface of H-NS protein from *Escherichia coli* by heteronuclear NMR spectroscopy. *FEBS Lett.* **455**, 63-69

Spassky, A., Rimsky, S., Garreau, H., and Buc, H. (1984). H1a, an *E. coli* DNA-binding protein which accumulates in stationary phase, strongly compacts DNA *in vitro*. *Nucleic Acids Res.* **12**, 5321-5340

Spurio, R., Falconi, M., Brandi, A., Pon, C.L., and Gualerzi, C.O. (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

Stoebel, D.M., Free, A., and Dorman, C.J. (2008). Anti-silencing: overcoming H-NS-mediated repression of transcription in Gram-negative enteric bacteria. *Microbiology* **154,** 2533-2545

Suck, D. (1994). DNA recognition by DNase I. *J. Mol. Recognit.**7**, 65-70*

Sueoka, N. (1962). On the genetic basis of variation and heterogeneity of DNA base composition. *Proc. Natl. Acad. Sci. U.S.A.* **48**, 582-592

Syvanen, M. (1994). Horizontal gene transfer: evidence and possible consequences. *Annu. Rev. Genet.* **28**, 237-261

Takeshita, S., Sato, M., Toba, M., Masahashi, W., and Hashimoto-Gotoh, T. (1987). High-copy-number and low-copy-number plasmid vectors for *lacZ* alpha-complementation and chloramphenicol-or kanamycin-resistance selection. *Gene* **61,** 63-74

Tendeng, C., Soutourina, O.A., Danchin, A., and Bertin, P.N. (2003). MvaT proteins in *Pseudomonas* spp.: a novel class of H-NS-like proteins. *Microbiology* **149**, 3047-3050

Trotta, E., and Paci, M. (1998). Solution structure of DAPI selectively bound in the minor groove of a DNA T.T mismatch-containing site: NMR and molecular dynamics studies. *Nucleic Acids Res.* **26**, 4706-4713

Turner, E.C., and Dorman, C.J. (2007). H-NS antagonism in *Shigella flexneri* by VirB, a virulence gene transcription regulator that is closely related to plasmid partition factors. *J. Bacteriol.* **189**, 3403-3413

Ueguchi, C., Seto, C., Suzuki, T., and Mizuno, T. (1997). Clarification of the dimerization domain and its functional significance for the *Escherichia coli* nucleoid protein H-NS. *J. Mol.*

Biol. **274**, 145-151

Valet-Gely, I., Donovan, K.E., Fang, R., Joung, J.K., and Dove, S.L. (2005). Repression of phase-variable *cup* gene expression by H-NS-like proteins in *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U.S.A.* **102,** 11082-11087

Vallet, I., Diggle, S.P., Stacey, R.E., Cámara, M., Ventre, I., Lory, S., Lazdunski, A., Williams, P., and Filloux, A. (2004). Biofilm formation in *Pseudomonas aeruginosa*: fimbrial *cup* gene clusters are controlled by the transcriptional regulator MvaT. *J. Bacteriol.* **186**, 2880-2890

Varshavsky, A.J., Nedospasov, S.A., Bakayev, V.V., Bakayeva, T.G., and Georgiev, G.P. (1977). Histone-like proteins in the purified *Escherichia coli* deoxyribonucleoprotein. *Nucleic Acids Res.* **4**, 2725-2745

Walthers, D., Carroll, R.K., Navarre, W.W., Libby, S.J., Fang, F.C., and Kenney, L. J. (2007). The response regulator SsrB activates expression of diverse *Salmonella* pathogenicity island 2 promoters and counters silencing by the nucleoid-associated protein H-NS. *Mol. Microbiol.* **65**, 477-493

Walthers, D., Li, Y., Liu, Y., Anand, G., Yan, J., and Kenney, L.J. (2011). *Salmonella enterica* response regulator SsrB relieves H-NS silencing by displacing H-NS bound in polymerization mode and directly activates transcription. *J. Biol. Chem.* **286**, 1895-1902

Wang, W., Li, G.W., Chen, C., Xie, X.S., and Zhuang, X. (2011). Chromosome organization by a nucleoid-associated protein in live bacteria. *Science* **333**, 1445-1449

Westermark, M., Oscarrson, J., Mizunoe, Y., Urbonaviciene, J., and Uhlin, B.E. (2000). Silencing and activation of ClyA cytotoxin in *Escherichia coli*. *J. Bacteriol.* **182**, 6347-6357

Williamson, H.S., and Free, A. (2005). A truncated H-NS-like protein from enteropathogenic *Escherichia coli* acts as an H-NS-antagonist. *Mol. Microbiol.* **55**, 808-827

Winardhi, R.S., Fu, W., Castang, S., Li, Y., Dove, S.L., and Yan, J. (2012). Higher order oligomerization is required for H-NS family member MvaT to form gene-silencing nucleoprotein filament. *Nucleic Acids Res.* **40**, 8942-8952

Zhaxybayeva, O., and Doolittle, W.F. (2011). Lateral gene transfer. *Curr. Biol.* **21**, R242-246

Appendix: Sequences of EMSA DNA fragments

hilA fragments from *Salmonella* Typhimurium SL1344
cupA1 fragment from *Pseudomonas aeruginosa* PAO1
PA390 fragment from *Pseudomonas aeruginosa* PAO1

hilA (24bp): primers GT067 and GT068 (anneal complementary oligonucleotide primers)

AAGAGAATACACTATTATCATGCC

hilA (101bp): primers GT071 and GT068 (PCR amplification from *S.* Typhimurium SL1344)

GTACTAACAGCAGAATTACTGAAACAGTAGATTCTATCCTAACGACTTGTATTAGTT
ATTATAACTTTTCACCCTGTAAGAGAATACACTATTATCATGCC

hilA (289bp): primers GT077 and GT068 (PCR amplification from *S.* Typhimurium SL1344)

CTCTCTCTGCACCAGGATATACGGCAGCGTCCATTCGATAATCACAGTTAGTTATAA
CAATATTATTACCAACATGTCAGTTATTTAAAGCACAGGCATAAGCTAAATAATCA
AATGTTAAAAACATATAAACCCGAGCCCGTAGAATATGACATTAAGCTCATAATAA
AAGCTCAACCTGACCGTTAGTACTAACAGCAGAATTACTGAAACAGTAGATTCTAT
CCTAACGACTTGTATTAGTTATTATAACTTTTCACCCTGTAAGAGAATACACTATTA
TCATGCC

cupA1 (340bp): primers GT044 and GT045 (PCR amplification from *P. aeruginosa* PAO1)

GCGAAGCCGTGGTTCGAGTTGTTGACTTGTCGCAAGACGGATTCTCCGAACAGCAT
CGGCGCAGCGCCCACTTGCAACGAGAGTGCGCACGCCGCCGTTGGCACTCGTCTGC
GCACCGGACTTCTGAGCGGCGGCGGAAAAAGCTTGCTGCGCAAGGAGAAGCGCCT
GCGGGTGAACATGGAAATTGCCCGGGAAAAGTCGACGGTAATTATTCGTTGAATAA
CAGGGAAGTGCATACTCCAACTTGGAAGTTAACTGCATGGCACCATCCATGCATTA
CACGGATGACGGTTTTTATTATAAGTCCAAACTGGCGAAGACAAGCAAGAGAGGCC
GGGAT

PA3900 (204bp): primers GT049 and GT050 (PCR amplification from *P. aeruginosa* PAO1)

CCGCAGGTGGCTGAACAGGCCGTGCGCTGGCTGGTCGAGCTGCAGGGCGGCGCCG
ACGACGAACGCCTGCGCCAGGCCTGGCAGCGCTGGCGCCAGGCGGCCCCGGAACA
CGAGCAGGCCTGGCGCCACATCGAGGCGGTCAACCAGCGCCTGGCCGGGATAGGC
ACGCCACTGGCCCTGGCCGCCATCAACGCACCGCATTCG