BINDING PROPERTIES OF HFQ TO RNA AND GENOMIC DNA AND THE FUNCTIONAL IMPLICATIONS

A Thesis
Presented to
The Academic Faculty

by

Taylor Blanton Updegrove

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Biology

Georgia Institute of Technology
August 2011

BINDING PROPERTIES OF HFQ TO RNA AND GENOMIC DNA

AND THE FUNCTIONAL IMPLICATIONS

Approved by:

Dr. Roger M. Wartell, Advisor
School of Biology
Georgia Institute of Technology

Dr. Ingeborg Schmidt-Krey
School of Biology
Georgia Institute of Technology

Dr. Thomas DiChristina
School of Biology
Georgia Institute of Technology

Dr. Nicholas Hud
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Brian Hammer
School of Biology
Georgia Institute of Technology

Date Approved: April 27, 2011

To my family and friends,
for their constant love and support.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my mother and father for their unconditional love, support, and encouragement. Their dedication and commitment to my upbringing can never truly be matched. Even during my adolescence, which was marked with many acts of disobedience, my parents were always at my side to pull me out of whatever ditch that I might have fallen into. Over the years they spent an ungodly amount of money paying for my upbringing and always spent the time to instill character and strength into my personality. If not for their continuous love and support I truly would not be the person that I am today.

Many thanks also go to my brother who was always at my side to throw a punch or two when I needed it during my childhood. My longtime High School friends Samuel Escamilla and Scott Johnson should also be credited with enriching my personality and beliefs with good qualities. Both also provided me with innumerous days and nights of memorable times that will never be forgotten. My good friends Lee Katz and Byron Lee should also be recognized for carrying me under their wing during my beginning acclimation to student life here at Georgia Tech. Finally, the undergraduates that I would also like to recognize are Charles Terry and Ashley Brouillette. It never ceases to amaze me how cleaver and hard working those two are, and how much dedication and commitment they brought to the lab.

Lastly, I would like to thank my two academic advisors: Dr. Donald Lightfoot and Dr. Roger Wartell. Dr. Lightfoot guided me through my undergraduate career and Master of Science degree. He shaped my mind into thinking like a scientist and always encouraged me to put forth my best effort at everything I do. His incessant encouragement led me along the path to a Ph.D. and the knowledge he imparted on me gave me the tools to pursue a Ph.D. For that, I will always

be indebted to him. Dr. Wartell deserves appreciation for his commitment to me in his lab, and his continual guidance during my tenure here at Georgia Tech.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

A	Adenine
Bp	base pair
C	Cytosine
CD	circular dichoism
DEPC	Diethylpyrocarbonate
dH_2O	deionized water
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide 5'-triphosphate
EDTA	ethylenediaminetetraacetic acid
G	Guanine
H_2O_2	hydrogen peroxide
HPLC	high performance liquid chromatography
M	Molar
Min	Minute
mg	Milligram
ml	Milliliter
nm	Nanometer
nM	Nanomolar
OD_{260}	optical density at 260 nm
ODNs	Oligodeoxyribonucleotids

PCR	polymerase chain reaction
RBS	ribosomal binding site
RNA	ribonucleic acid
RNase E	ribonuclease E
RNase H	ribonuclease H
T	Thymine
TBE	tris-borate EDTA electrophoresis buffer
TE	tris-EDTA buffer
T_m	melting temperature
U	Uracil
UV	Ultraviolet
°C	degree Celsius
ΔG	Gibbs free energy
ΔH	Enthalpy
ΔS	Entropy
μg	Microgram
μM	Micromolar
μl	Microliter
ε	extinction coefficient

SUMMARY

Hfq, also known as Host Factor I, is an 11.2 Kilo-Dalton heat stable protein that is a required host factor for bacteriophage Qβ RNA replication in *E. coli*. Early studies have shown Hfq protein to be a global regulator of *E. coli* metabolism, which can be seen in the pleiotropic phenotypes of Hfq knockout mutants; *E. coli* Hfq mutants fail to respond to various stress insults, thus leaving the bacterium vulnerable. The broad impact of this protein appears to stem from its role in regulating the stability and/or translation of mRNA from a number of regulatory genes in an array of bacterial species. In *E. coli*, Hfq has been shown to work in concert with such known riboregulators (sRNAs) as RprA, RyhB, MicA, SgrS, DsrA, OxyS, and Spot42 RNA to up or down regulate targeted mRNAs by stimulating the proper pairing of the sRNA with its target mRNA. Hfq has also been shown to be involved in facilitating polyadenylation and degradation of mRNA through the recruitment of poly(A) polymerase and RNase E, respectively, and in the stimulation of CCA addition to the 3'end of tRNAs by enhancing the enzymatic activity of tRNA nucletidotranferase enzyme. The molecular mechanisms of the protein's broad RNA selectivity and diverse functions are not completely understood. Hfq recognizes and binds RNAs that have A-rich or U-rich sequences on at least two distinct surfaces and RNA secondary structure could be an important element in Hfq recognition and function.

The *rpoS* mRNA encodes a stress response sigma factor in *E. coli* that is required for stationary phase growth and survival to stress insults. *rpoS* mRNA is one of a growing number of mRNAs found to be regulated by sRNAs and Hfq. Translation initiation of *rpoS* mRNA is enhanced by two sRNAs, DsrA and RprA, which pair to the same site near the *rpoS* start codon in the presence of the Hfq protein. The interaction of *E. coli* Hfq with RprA and two portions of

the *rpoS* mRNA leader region was examined to explore Hfq's effect on promoting RprA-*rpoS* RNA binding *in vitro*. One *rpoS* RNA, *rpoS*-L, contained the entire 565-nucleotide untranslated leader region, while the other, *rpoS*-S, contained the 199-nucleotide sequence surrounding the start codon. An RNase H assay indicated both *rpoS* RNAs have similar secondary structures in the translation initiation region. Hfq formed two complexes with RprA in a gel mobility assay with binding parameters similar to values previously determined for DsrA. Unlike DsrA, Hfq binding to RprA was inhibited by poly(A) and influenced by Hfq mutations on both the distal and proximal surfaces. Hfq increased the level of RprA binding to both *rpoS* RNAs but showed a much larger enhancement when *rpoS*-L was examined. The lower affinity of RprA for *rpoS*-L versus *rpoS*-S in the absence of Hfq suggests that Hfq overcomes an inhibitory structure within *rpoS*-L in stimulating RprA binding. Similar results were obtained with DsrA. The results indicate that the full upstream leader sequence of *rpoS* mRNA influences Hfq-facilitated annealing of RprA and DsrA and is likely to be involved in its regulation.

The sRNAs DsrA and RprA enhance translation of *rpoS* mRNA by pairing to a site on this mRNA and disrupting an intramolecular stem-loop structure containing the ribosome binding site (RBS). The sRNA OxyS represses *rpoS* mRNA translation by an unknown mechanism. The binding of eleven mutant Hfqs to DsrA, RprA, OxyS, and two segments of the *rpoS* mRNA untranslated leader region was examined to explore RNA binding surfaces on Hfq. Mutant Hfqs were also tested for their ability to stimulate DsrA-*rpoS* RNA binding. Nine of the mutant Hfqs had single amino acid mutations located on the proximal, distal, or outer-circumference surface of the Hfq hexamer structure. Two mutant Hfqs had truncated C-terminal ends. Proximal surface mutations decreased Hfq binding to the three sRNAs and the *rpoS* RNA segment containing the RBS. Distal surface mutations lowered Hfq affinity to the *rpoS* RNA

region containing the $(ARN)_4$ sequence. Strong binding of Hfq to both the RBS and $(ARN)_4$ segments of *rpoS* mRNA was needed for maximum Hfq enhancement of DsrA•*rpoS* RNA annealing. The two truncated Hfqs and the Hfqs with circumference surface mutations behaved similar to wild type Hfq with regard to binding the sRNAs and both *rpoS* RNA segments, and in stimulating DsrA•*rpoS* RNA formation; this suggest the C-terminal tail and the circumference residues examined are not critical for RNA binding and *rpoS* regulation. The binding of OxyS to *rpoS* RNA in the presence and absence of Hfq was examined. Under both conditions very little OxyS•*rpoS* RNA complex was observed, suggesting against a mechanism of riboregulation involveing Hfq enhancing the annealing of OxyS to *rpoS* mRNA.

Hfq is involved in many aspects of posttranscriptional gene expression. Tight binding of Hfq to polyadenylate sequences at the 3' end of mRNAs influences exonucleolytic degradation, while Hfq binding to sRNAs and their targeted mRNAs facilitate their hybridization which in turn effects translation. Hfq binding to the sRNA DsrA and to an A-rich tract in the 5' leader region of the *rpoS* mRNA have been shown to be important for DsrA enhanced translation initiation of this mRNA. The complexes of Hfq-A_{18} and Hfq-DsrA provide models for understanding how Hfq interacts with these two RNA sequence/structure motifs. Different methods have reported different values for the stoichiometry of Hfq-A_{18} and Hfq-DsrA. In this work, mass spectrometry and analytical ultracentrifugation were utilized to provide direct evidence that the strong binding mode of the Hfq hexamer (Hfq_6) for A_{18} and domain II of DsrA (($DsrA_{DII}$), a 38-nt portion of DsrA that competes with full length DsrA for Hfq binding), involve 1:1 complexes. This stoichiometry was also supported by fluorescence anisotropy and a competition gel mobility shift experiment using wild type and truncated Hfq. More limited studies of Hfq binding to DsrA as well as to the sRNAs RprA, OxyS, and an 18-nt segment of

OxyS that binds Hfq were also consistent with 1:1 stoichiometry. Mass spectrometry of a sample containing Hfq_6, A_{18}, and $DsrA_{DII}$ exhibit intensity corresponding to a ternary 1:1:1 complex; however, the small intensity of this peak, and fluorescence anisotropy experiments did not provide evidence that this ternary complex is stable in solution.

Hfq has been studied extensively for its function as a modulator of gene expression at the posttranscriptional level. While most Hfq studies have focused on the protein's interaction with sRNAs and mRNAs, Hfq binding to DNA has been observed but is less explored. During the isolation of Hfq from *Escherichia coli*, we found genomic DNA fragments associated with the protein after multiple steps of purification. Sequences of 41 amplified segments from the DNA fragments associated with Hfq were determined. A large fraction of the DNA segments were predicted to have significant helical axis curvature and were from genes associated with membrane proteins, characteristics unexpected for nonspecific binding. Analysis by analytical ultracentrifugation indicated that A_{18} binding to Hfq disrupts Hfq-DNA interactions. The latter observation suggests Hfq binding to DNA involves its distal surface. This was supported by a gel mobility shift assay that showed single amino acid mutations on the distal surface of Hfq inhibited Hfq binding to duplex DNA, while six of seven mutations on the proximal surface and outer circumference of the hexamer did not prevent Hfq binding. Two mutated Hfq which have portions of their C-terminal domain removed also failed to bind to DNA. The apparent K_d of wild type Hfq binding to several duplex DNA fragments isolated from Hfq preps was estimated from a gel mobility shift assay to be ~400 nM, about ten fold less affinity than for Hfq target RNAs. Hfq was shown to display a wide range of affinities to single stranded DNA fragments of the same length but different sequence, indicating Hfq to bind single stranded DNA with sequence specificity.

CHAPTER 1

Introduction and literature review of Hfq

HISTORY

The history of Hfq dates back to the late 1960's when it was found that Hfq was a required host factor in *E. coli* for the replication of the Qβ RNA bacteriophage. Franze de Fernandez et. al. showed Hfq to be necessary for the initiation of Qβ RNA plus-strand synthesis *in vitro* (Franze de Fernandez et al. 1968; Franze de Fernandez et al. 1972). At the time the exact cellular role of Hfq was unknown; however, Hfq was thought to carry out functions other than enabling intracellular bacteriophage replication and ultimately cell death. For the most part of the next 20 years the focus of Hfq research has been mainly on its binding properties to RNA molecules such as Qβ RNA and poly(A) sequences of a given length and geometry (de Haseth and Uhlenbeck 1980b; de Haseth and Uhlenbeck 1980a). It was within this time period that Hfq was recognized as a high affinity RNA binding protein with a preference for adenylate rich sequences, and that the protein organizes itself into a multimeric form that is consistent with the molecular weight of a homohexamer in solution.

In 1994 the importance of Hfq in the cell was recognized when scientists were able to pinpoint the *hfq* gene on the *E. coli* genome and upon disruption of it, *hfq* mutants were shown to display prominent pleotropic phenotypes (Tsui et al. 1994). Hfq mutants were shown to have a decrease in growth rate, altered cell morphology (mutant cells were more elongated than wild type cells), altered protein synthesis, increase sensitivity to UV-light, and changes in sensitivity to osmolarity and oxidation. Shortly thereafter, similar phenotypes were also observed in *hfq* mutants of other bacterial species. It became evident that a majority of the phenotypes observed

from the *hfq* knockout mutants in *E. coli* matched the phenotypes that were caused by the disruption of the *rpoS* gene encoding the stationary phase sigma factor σ^s – a global regulator for stress response and stationary phase growth. Indeed, a couple of years after the first reported disruption of the *hfq* gene, it was realized that Hfq was necessary for RpoS expression at the mRNA translation level (Brown and Elliott 1996). However, not all of the phenotypes observed in the *hfq* null mutant could be attributed to defects in the expression of RpoS (Muffler et al. 1997).

In the late 1990's and early 2000's, after the initial discovery of function of trans-encoded sRNAs, it was realized that Hfq is involved in the expression of multiple genes at the posttranscriptional level by not only binding to and modulating the half-lives of mRNAs and sRNAs, but by enhancing the interaction of sRNAs to specific mRNAs and eliciting sRNA riboregulation (Storz et al. 2004; Valentin-Hansen et al. 2004; Brennan and Link 2007). Hfq was shown to be conserved in multiple bacteria from distinct phylogenetic groups, and very closely related, at least structurally, to the Sm and Sm-like (LSm) proteins found in eukaryotes and archaea, respectively (Sun et al. 2002; Valentin-Hansen et al. 2004). To date roughly half of the sRNAs discovered in *E. coli* require Hfq for function, and multiple sRNAs that are continually being discovered in other gram negative and gram positive bacteria also require Hfq for function (Brennan and Link 2007; Pichon and Felden 2007). Hfq was recently reported to be the major protein hub of the *E. coli* gene regulatory network (Butland et al. 2005), and is regarded today as a global regulator of bacterial gene expression.

STRUCTURE AND HOMOLOGY

Hfq sequence analysis and conservation

Amino acid sequence analysis of Hfq against the non-redundant sequence data base at the NCBI web site has shown that the N-terminal portion of Hfq (consisting of residues 7-66) is highly conserved among a number of bacteria, and shares similarity with the Sm and LSm proteins found in eukaryotes and archaea (Sun et al. 2002; Valentin-Hansen et al. 2004). Nearly half of the completed or partially completed genomes of bacteria show the presence of a putative Hfq gene, and it is believed that as more genomes are sequenced, and the search parameters for data base mining become more refined, additional Hfq homologues will be discovered (Valentin-Hansen et al. 2004). The presence of the Sm1 motif sequence in *E. coli* Hfq, and Hfq's ability to bind RNA and form an oligomeric ring structure in solution (see below), support the notion that Hfq belongs to the eukaryotic and archaeal family of Sm and LSm proteins. However, the Hfq sequence lacks the highly conserved Sm2 motif sequence found in virtually all Sm and LSm proteins. Interestingly, at least one archaeal member - *Methanococcus jannaschii* - possesses an Hfq gene that can support riboregulation and functionally complement an *hfq* mutation in *E. coli* (Nielsen et al. 2007).

Hfq structure and oligomeric form

Studies involving sedimentation analysis, gel electrophoresis, transmission election microscopy, and MALDI-TOF mass spectrometry demonstrated Hfq to form homohexamers in solution (Updegrove et al. 2011; Franze de Fernandez et al. 1972; Carmichael et al. 1975; Moller et al. 2002; Zhang et al. 2002). The studies involving gel electrophoresis and mass

spectrometry also showed that Hfq can stably exist as multimers of less than six subunits (Updegrove et al. 2011; Carmichael et al. 1975). The first Hfq crystal structure (from *Staphylococcus aureus*), and homology modeling of the N-terminal domain of *E. coli* Hfq with a known LSm protein structure, showed Hfq monomer subunit to consist of a bent five-stranded antiparallel β-sheet capped by an N-terminal α-helix that collectively display the topology β5α1β1β2β3β4 (Figure 1.1) (Schumacher et al. 2002; Sun et al. 2002). The conserved Sm1 motif sequence encompasses the first three β strands, whereas the conserved Sm2 motif sequence (different than the Sm2 motif of Sm proteins) is composed of β strands 4 and 5. The cyclic hexamer is formed primarily by hydrophobic interactions between residues from β4 and β5 from opposing subunits. The subunit structures of Hfq and Sm proteins are very similar and display root mean squared deviations ranging from 0.85 $Å^2$ to 1.3 $Å^2$ (when comparing archaeal and human Sm proteins to Hfq) (Brennan and Link 2007).

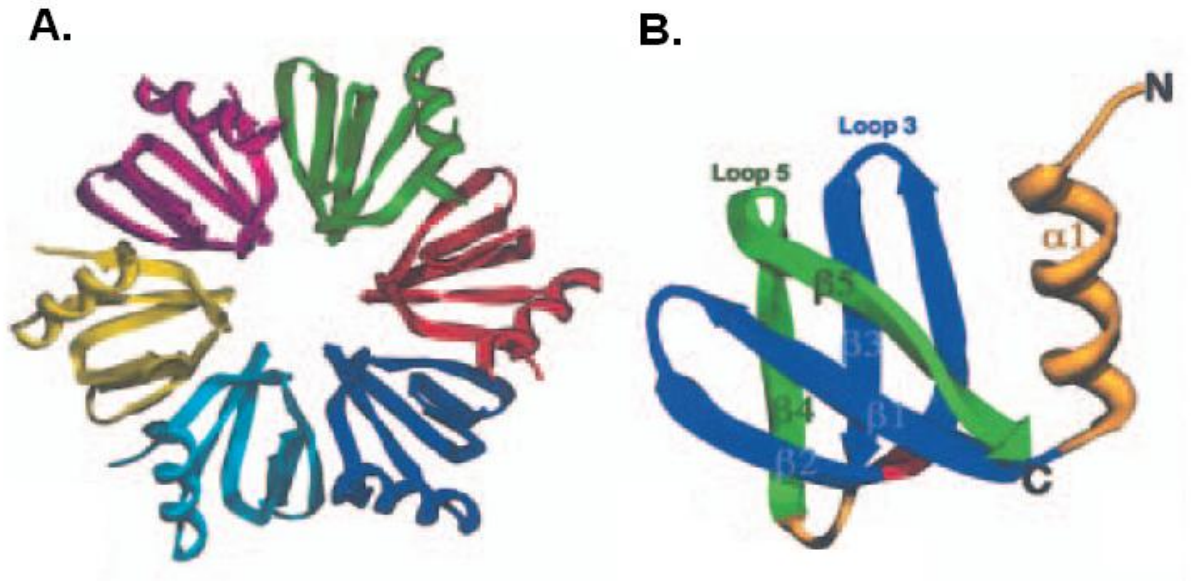


Figure 1.1: (**A.**) Structure of the *S. aureus* Hfq hexamer with each subunit colored differently. (**B.**) Ribbon diagram of an Hfq subunit. The Sm1 motif is colored blue and the Sm2 motif is green. Regions outside the two motifs, i.e. the N-terminal α-helix and the variable region are colored yellow. Hfq residue Gly-34, the sole conserved residue among Hfq and the Sm proteins, is blocked in red. This figure was adapted and modified from Valentin-Hansen et al. 2004.

Since the initial *S. aureus* crystal structure was determined, 13 unique structures of Hfq-proteins from seven different organisms have been deposited in the Protein Data Bank [*S. aureus* (amino acids 1-77) 1KQ1, 1KQ2; *E. coli* (amino acids 1-72) 1HK9, 3GIB (amino acids 2-69); *Pseudomonas aeruginosa* (amino acids 1-82) 1U1T, 1U1S, 3M4G, 3INZ; *M. jannaschii* (amino acids 1-71) 2QTX; *Synechocystis sp*. (amino acids 1-70) 3HFO; *Anabaena sp*. (amino acids 1-72) 3HFN; *Bacillus subtilis* (amino acids 1-78) 3HSB, and 3HSA] (Beich-Frandsen et al. 2011). Such structures show the Hfq toroid to have outer and central pore diameters of ~70 Å and 10 ± 2 Å, respectively, and a thickness of ~ 25 Å. The root mean square comparison between subunits of different Hfq proteins range between 0.37 $Å^2$ and 0.52 $Å^2$ (Brennan and Link 2007).

The length of the C-terminal end of each subunit varies between bacterial species, and is extremely dependant on bacterial phylogenetic group; the Hfq proteins of γ- and β-proteobacteria have extended C-termini, whereas some Gram-positive bacteria including *S. aureus* have Hfq proteins with short C-terminal extensions (Sun et al. 2002). Extremely long C-terminal ends (greater than 100 residues) have been reported from *Moraxella catarrhalis* and *Acinetobacter baylyi* (Attia et al. 2008; Schilling and Gerischer 2009). However, the exact function(s) of the extended C-terminal ends is currently a contentious issue, with recent studies presenting conflicting results in the possible role the C-terminal end plays in riboregulation and RNA binding (Beich-Frandsen et al. 2011; Olsen et al. 2011 ; Vecerek et al. 2008). What is generally agreed on is the lack of ordered structure of the extended C-terminal end, which can be seen through sequence based structure predication, along with multiple biophysical, molecular and structural studies (Beich-Frandsen et al. 2011). The regions on the C-termini that are predicted to be disordered and flexible may be critical for the stable interaction of Hfq with other proteins

(including Hfq itself) and/ or required for the Hfq RNA chaperone like activities. Further studies however would need to confirm this.

FUNCTION AND IMPORTANCE

Role of Hfq in trans-encoded sRNA riboregulation

Perhaps the most widely known function of Hfq in the cell is its role in posttranscriptional gene regulation through the concerted action of trans-encoded sRNAs. Because trans-encoded sRNAs arise from genes located at loci different from those encoding their RNA targets, sRNA-target complementarity is generally incomplete, and regulation is achieved by forming small, imperfect duplexes. This relaxation in binding, along with the fact that sRNAs are much longer than the typical sRNA-mRNA interaction site (~100 nt versus ~15 nt, respectively), allows for greater versatility in single sRNAs targeting multiple mRNAs, and reciprocally, single mRNAs targeted by multiple sRNAs (Valentin-Hansen et al. 2004). The presence of Hfq has been shown to enhance the interaction of sRNAs to mRNAs, although by a mechanism(s) not fully understood. Hfq has been postulated to act as a RNA chaperone that modulates sRNA and/ or mRNA structure, and alteration in the structure of one or both RNA partners allows more efficient binding. Evidence for this comes from the observation of Hfq altering the structure of sRNAs such as OxyS and DsrA (Zhang et al. 2002; Lease and Woodson 2004), and the mRNAs *rpoS* and *ompA* (Moll et al. 2003a; Lease and Woodson 2004). Alternatively (or additionally) Hfq my simply bring both RNAs in close proximity of each other, and the increase in local concentration of both RNAs would drive the equilibrium to favor RNA-RNA binding. The observation that single Hfq hexamers can simultaneously bind two RNA molecules would support this idea (Hwang et al. 2011; Updegrove et al. 2011). Regardless of the exact mechanism, the presence of Hfq was shown to enhance the formation of the following *E. coli* sRNA-mRNA pairs: DsrA-*rpoS* (Soper and Woodson 2008; Updegrove et al. 2008),

RprA-*rpoS* (Updegrove et al. 2008), OxyS-*fhlA* (Zhang et al. 2002), RyhB-*sodB* (Kawamoto et al. 2006), PtsG-*sgrS* (Kawamoto et al. 2006; Maki et al. 2008), and Spot 42-*galK* (Moller et al. 2002). In all tested cases the enhancement in stability correlated with an increase in kinetic association rate of the two RNAs

Binding of sRNA to mRNA target usually occurs around the mRNA region harboring the ribosome binding site (RBS) and AUG start codon, and such binding directly influences the translational status of the mRNA. Depending on the riboregulator and the mRNA target, sRNA-mRNA binding could result in translation repression or activation. RyhB-*sodB* and PtsG-*sgrS* are often cited examples where the sRNA mediates repression of the mRNA (Aiba 2007). In both cases, binding of the sRNA to the mRNA is sufficient for translational repression (Morita et al. 2006), probably by blocking ribosome binding. sRNA binding can also target both RNAs for degradation by RNase E (Morita et al. 2005). In this case Hfq is thought to act as a bridge that couples the sRNA-mRNA complex to the RNase E enzyme, since Hfq can directly interact with both RNase E (Ikeda et al. 2011) and the sRNA-mRNA complex (Maki et al. 2008), and can be found in a complex with all three molecules (Morita et al. 2005). The rapid degradation of both RNAs is a consequence of pairing, and suggests the sRNA act stoichiometrically rather than catalytically, which would allow for rapid termination of regulation. However, not all sRNAs target itself and/ or the mRNA for degradation upon binding, thus the sRNA can act catalytically as in the case of the *ybfM* mRNA that is down regulated by MicM sRNA. MicM binding to *ybfM* targets the mRNA for degradation while MicM is recycled and can partake in the pairing and destruction of multiple *ybfM* molecules (Overgaard et al. 2009).

Examples of activation of mRNA translation by sRNA binding are much less frequently observed than repression of translation. In most cases activation of translation proceeds through a

mechanism whereby sRNA binding removes an inhibitory RNA secondary structure that prevents efficient translation. An example of this would be the enhanced translation of *rpoS* mRNA upon binding of DsrA or RprA sRNAs in *E. coli*. Under normal growth conditions the ribosomal binding site of the *rpoS* mRNA is trapped in a secondary structure that represses translation. However, when *E. coli* are exposed to cold temperature, the induction of DsrA occurs and the binding of DsrA to the cis-inhibitory stem opens up the ribosome binding site for interaction with the ribosome and consequently causes more efficient translation (Brown and Elliott 1996; Sledjeski et al. 1996; Majdalani et al. 1998). After osmotic shock or stationary phase growth, RprA is induced and acts similarly as DsrA does to enhance *rpoS* translation (Majdalani et al. 2002).

Hfq modulates mRNA and sRNA half-life

Very early on in the studies of Hfq it became apparent that Hfq binds with high affinity to RNA targets at discrete sites. Both sRNAs and mRNAs that are dependent on Hfq for pairing seem to bind Hfq independent of each other and sometimes, but not always, the binding alters the half-life of the RNA *in vivo*. A particularly instructive example of this is the growth rate dependent regulation of the *ompA* mRNA. This mRNA encodes an outer membrane protein that is regulated during envelope stress. The half-life of the mRNA is determined by its 5' untranslated region (5' UTR), which contains a stabilizing stem-loop structure as well as RNase E recognition sites (the major endonuclease involved in mRNA decay as well as processing of tRNA and rRNA precursors (Kaberdin and Blasi 2006)). The efficiency of cleavage at the 5' UTR determines the *ompA* mRNA turnover rate. Hfq was identified as a factor present in slow-growing cells that bind to the 5' UTR and is essential for controlling ompA mRNA half-life. The

half-life of the mRNA in *hfq* mutant cells is much longer, and the growth rate dependence on mRNA stability is lost (Vytvytska et al. 1998). It was shown that binding of Hfq to the 5'UTR prevents binding of the 30S ribosomal subunit to the RBS, which normally protects the 5' UTR from endonucleolytic cleavage (Vytvytska et al. 2000; Moll et al. 2003b). The loss of nuclease protection by efficient ribosome binding and translocation seems to be a common theme for many Hfq binding mRNAs. For instance, the binding of Hfq to the mRNAs *mutS* (Tsui et al. 1997), *miaA* and *hfq* (Vecerek et al. 2005) reduce the cellular half-lives of these RNAs in part by preventing ribosome binding and increasing the susceptibility to nuclease degradation (Storz et al. 2004; Majdalani et al. 2005).

In stark contrast to mRNAs, the binding of Hfq to sRNAs seem to extend sRNA cellular half-life. The presence of Hfq stabilizes sRNAs such as DsrA (Madhugiri et al. 2010; Sledjeski et al. 2001), RprA (Madhugiri et al. 2010), RyhB (Moll et al. 2003a), PtsG (Masse et al. 2003), Spot42 (Moller et al. 2002). It has been established that Hfq binding protects RyhB and DsrA RNAs from cleavage by RNase E (Massè et al. 2003; Moll et al. 2003a). This mechanism of action can probably be extended to many of the riboregulators because Hfq binding sites on RNA may coincide with recognition sites for RNase E (Aiba 2007). From protection assays it has been shown that Hfq tends to bind AU rich single stranded regions abutted by one or more hair-pins (Moller et al. 2002; Schumacher et al. 2002; Zhang et al. 2002; Brescia et al. 2003), while RNase E recognition sites are also internal AU single stranded regions (Rauhut and Klug 1999; Mackie 1992). The enhanced half-life of some of the sRNAs by Hfq may also be explained by the Hfq dependent enhancement of sRNA binding mRNA target (Valentin-Hansen et al. 2004); sRNA in duplex form is more protected from single stranded specific nucleases. It is worth noting that not all sRNAs that bind Hfq and require Hfq for function are necessarily

stabilized by the presence of Hfq. The half-life of the Hfq binding sRNA OxyS, which requires Hfq for repression of translation of the *rpoS* mRNA and *fhlA* mRNA, is the same in *hfq* null mutants as in wild type cells (Zhang et al. 2002).

Role of Hfq in poly(A) metabolism and tRNA maturation

Beyond its role as a RNA chaperone, Hfq also functions in regulating mRNA half-life by binding with nanomolar to subnanomolar affinities to polyadenylate [poly(A)] sequence that have been added to the 3' ends of mRNAs. In *E. coli*, it has been estimated that > 90% of the total mRNA generated during exponential growth phase is polyadenylated posttranscriptionally (Mohanty and Kushner 2006). Unlike cytosolic eukaryotic mRNA, the addition of adenosine nucleotides to bacterial mRNAs enhances their degradation (Steege 2000). The size of the poly(A) tail of bacterial mRNA depends on the competition between poly(A) polymerase I (PAP I) and exonucleases like polynucleotide phosphorylase (PNPase), which attack the 3' end of mRNA (Vasil'eva Iu and Garber 2002). Hfq substantially enhances the poly(A) synthesis by PAP I *in vitro* by switching PAP I from a distributive to a processive catalytic mode, allowing for the more extensive adenylation of mRNAs. This can also be seen *in vivo* (Hajnsdorf and Regnier 2000). Moreover, Hfq protects poly(A) tails from the degradation by the exoribonucleases RNase II, PNPase, and RNase E (Folichon et al. 2003; Mohanty et al. 2004; Folichon et al. 2005), and Hfq has been shown to associate with PAP I and PNPase in pull down assays (Mohanty et al. 2004).

A more recent function that has been ascribed to Hfq involves its ability to stimulate the catalytic activity of the tRNA nucleotidyltransferase enzyme which synthesizes the 3'-terminal sequence C-C-A to bacterial tRNAs (Scheibe et al. 2007). This enzyme has a very close

evolutionary relationship with PAP I, and like the association of Hfq with PAP I substrate poly(A), Hfq was shown to specifically bind tRNA transcripts (Scheibe et al. 2007; Lee and Feig 2008), which seems to be a prerequisite for the observed effect on CCA-addition. Limited studies also demonstrate that Hfq displays ATPase like activity, where the binding of ATP to the distal side of Hfq can hydrolyze ATP to ADP, albeit relatively weakly when comparing to other known ATPases (Sukhodolets and Garges 2003; Arluison et al. 2007b). Further studies are needed to demonstrate Hfq's ATPase activity *in vivo* and the biological significance. Worth noting are *in vitro* experiments that demonstrate Hfq's chaperone like activity and enhancement of sRNA-mRNA pairing occur in the absence of ATP; thus, Hfq's involvement in its many RNA transactions probably do not require ATP hydrolysis (Moller et al. 2002; Zhang et al. 2002; Arluison et al. 2007a; Soper and Woodson 2008; Updegrove et al. 2008).

Role of Hfq in the bacterial nucleoid

Using immunostaining and fluorescence microscopy, it was shown that most (80-90%) of the cell's Hfq is in the cytoplasmic fraction and is associated with ribosomes, whereas 10-20% of Hfq are contained in the nucleoid region (Kajitani et al. 1994; Azam et al. 2000). Electron microscopy also confirmed the presence of Hfq in the nucleoid and cytoplasm, but also around the cell periphery in close association with the cell membrane (Diestra et al. 2009). The presence of Hfq in the nucleoid suggests Hfq to associate with DNA and possibly serve some function in association with DNA. Indeed, Hfq was found to be one of three most prevalent nucleoid proteins in exponentially growing cells (Ali Azam et al. 1999), and was shown to bind, albeit non-specifically, with supercoiled as well as linear DNA (Takada et al. 1997), thus suggesting Hfq to act as a structural protein for nucleoid formation. However, a more specialized role of

Hfq in the nucleoid could be in regulating gene transcription. A recent study demonstrated that the presence of Hfq had an impact on the transcription rate of certain genes (Le Derout et al. 2010). Further studies are needed to show if the direct binding of Hfq to specific DNA sites is the cause of the transcription modulation, or if Hfq is indirectly affecting transcription. It is worth mentioning that *in vitro* binding studies suggested Hfq has a higher affinity for curved DNA segments than linear DNA of the same length (Azam and Ishihama 1999). Based on the observation that promoters of genes often display significant helical curvature, one may speculate that Hfq binding to these regions may serve a role in transcribing the corresponding mRNA.

Role of Hfq in bacterial pathogens

Hfq has been shown to be required for the fitness and virulence of an increasing number of bacterial pathogens. Pathogens lacking Hfq are often sensitive to host defense mechanisms and their infections are highly attenuated in animal models. Defects in overall virulence of *hfq* mutants have been observed in such Gram negative pathogens as *Brucella abortus* (Robertson and Roop 1999) and *B. cepacia* (Sousa et al.), pathogenic *E. coli* (Kulesus et al. 2008; Shakhnovich et al. 2009), *Francisella tularensis* (Kadzhaev et al. 2009; Meibom et al. 2009), *Legionella pneumophila* (McNealy et al. 2005), *Neisseria meningitidis* (Fantappie et al. 2009; Pannekoek et al. 2009) and *N. gonorrhoeae* (Dietrich et al. 2009), *Pseudomonas aeruginosa* (Sonnleitner et al. 2003), *Salmonella typhimurium* (Sittka et al. 2007), *Vibrio cholerae* (Ding et al. 2004), *Y. pestis* (Geng et al. 2009), and more; and in the Gram positive pathogens *Listeria monocytogenes* (Christiansen et al. 2004) and *Staphylococcus aureus* (Liu et al. 2010). Virulence phenotypes of *hfq* mutants are more prevalent and dramatic in Gram-negative

pathogens, which can be seen in pathogen survival in rat and mice models, and increased sensitivity of pathogens to: starvation, oxidative stress, acid stress, heat stress, high osmolarity, iron limitation, heat exposure, detergent exposure, and antimicropeptide/antibiotic (Chao and Vogel 2010). In some cases the sensitivity to environmental stress or timing of virulence gene expression can be traced to transduction pathways involving one or more sRNAs and Hfq, such as the Qrr sRNAs system in *V. cholerae* (Hammer and Bassler 2007*)*. It is likely that as more pathogens are characterized, the role of Hfq in virulence will expand and possibly shed light on the pathogenic mechanism in these bacteria.

OBJECTIVES

One of the most widely studied system of riboregulation involving Hfq and sRNAs is in the regulation of the stationary phase sigma factor RpoS in *E. coli*. This system is one of the few examples in bacteria where sRNAs in conjunction with Hfq activate translation of a mRNA. Although much is known about this system of regulation, there are still many questions that need to be addressed. For instance, RprA is an sRNA activator of RpoS and very little is known about its binding properties to Hfq and how it compares with the more studied sRNA DsrA. Little is also known about what role the leader region of the *rpoS* mRNA plays in the regulation of RpoS through Hfq and sRNAs. In Chapter 2 we attempted to answer some of these questions. We established a framework for studying the binding of RprA to Hfq, *rpoS* to Hfq, and RprA binding to *rpoS* in the presence and absence of Hfq. We performed similar experiments with DsrA, a better studied activator of *rpoS*, to compare with RprA and shed light on the common aspects of RpoS regulation that involves these two sRNAs and Hfq.

The objective of Chapter 3 was to further explore RNA binding surfaces on Hfq that are specific for sRNAs and mRNAs. We looked at the binding affinities of Hfq mutants with amino acid changes on all four surface regions of Hfq to the sRNAs DsrA, RprA, OxyS, and to several size variants of the *rpoS* transcript. We also looked for correlations between the affinities of mutant Hfq for sRNAs and to different regions of the *rpoS* leader sequence, with changes in the ability of the Hfq to enhance DsrA binding to *rpoS*. Furthermore, since the mechanism of *rpoS* regulation by the sRNA OxyS is not known, we set out to determine if OxyS is able to bind *rpoS* and if Hfq enhances this interaction akin to what is seen with DsrA and RprA binding *rpoS*.

In Chapter 4 our objective was to determine the stoichiometry of Hfq and RNA in the high affinity complex. Previous studies using different experimental techniques have reported disparate results on the stoichiometry of Hfq bound to several RNA targets. Knowledge of the stoichiometry is very important when constructing models that would explain how Hfq might be stimulating the binding of sRNA regulators to mRNA targets. In this chapter we employed a variety of techniques not previously utilized to explore the binding of Hfq to DsrA, RprA, OxyS, and oligo A_{18}. Several of these methods allowed for the most direct possible way of determining the molecular weight of protein-RNA complexes and the deduction of RNA and protein stoichiometry. Furthermore, these methods were applied at RNA and protein concentrations thought to be present in the cell at physiological conditions, thus the measured stoichiometry likely reflects that which occurs in nature.

While the characterization of Hfq function at the RNA level has been analyzed extensively, studies of possible Hfq functions at the DNA level have been less explored. In Chapter 5 the objective was to characterize genomic DNA fragments that are present in Hfq preps after the protein purification procedure. Binding studies of Hfq to these DNA fragments evaluated the affinity Hfq has for DNA, and the sequences of DNA fragments found associated with Hfq was determined. The surface region on Hfq responsible for DNA binding was examined by conducing binding studies of the DNA to Hfq mutants with amino acid changes on all major surface domains of Hfq, and through binding competition experiments between preformed Hfq-DNA complexes and RNAs that bind Hfq on known surfaces. Lastly, we analyzed the binding of Hfq to single stranded DNA of a defined length, but variable sequence composition, to explore possible sequence specificity of Hfq for DNA.

REFERENCES

Aiba, H. 2007. Mechanism of RNA silencing by Hfq-binding small RNAs. *Curr Opin Microbiol* **10**(2): 134-139.

Ali Azam, T., Iwata, A., Nishimura, A., Ueda, S., and Ishihama, A. 1999. Growth phase-dependent variation in protein composition of the Escherichia coli nucleoid. *J Bacteriol* **181**(20): 6361-6370.

Arluison, V., Hohng, S., Roy, R., Pellegrini, O., Regnier, P., and Ha, T. 2007a. Spectroscopic observation of RNA chaperone activities of Hfq in post-transcriptional regulation by a small non-coding RNA. *Nucleic Acids Res* **35**(3): 999-1006.

Arluison, V., Mutyam, S.K., Mura, C., Marco, S., and Sukhodolets, M.V. 2007b. Sm-like protein Hfq: location of the ATP-binding site and the effect of ATP on Hfq-- RNA complexes. *Protein Sci* **16**(9): 1830-1841.

Attia, A.S., Sedillo, J.L., Wang, W., Liu, W., Brautigam, C.A., Winkler, W., and Hansen, E.J. 2008. Moraxella catarrhalis expresses an unusual Hfq protein. *Infect Immun* **76**(6): 2520-2530.

Azam, T.A., Hiraga, S., and Ishihama, A. 2000. Two types of localization of the DNA-binding proteins within the Escherichia coli nucleoid. *Genes Cells* **5**(8): 613-626.

Azam, T.A. and Ishihama, A. 1999. Twelve species of the nucleoid-associated protein from Escherichia coli. Sequence recognition specificity and DNA binding affinity. *J Biol Chem* **274**(46): 33105-33113.

Beich-Frandsen, M., Vecerek, B., Konarev, P.V., Sjoblom, B., Kloiber, K., Hammerle, H., Rajkowitsch, L., Miles, A.J., Kontaxis, G., Wallace, B.A. et al. 2011. Structural insights into the dynamics and function of the C-terminus of the E. coli RNA chaperone Hfq. *Nucleic Acids Res*.

Brennan, R.G. and Link, T.M. 2007. Hfq structure, function and ligand binding. *Curr Opin Microbiol* **10**(2): 125-133.

Brescia, C.C., Mikulecky, P.J., Feig, A.L., and Sledjeski, D.D. 2003. Identification of the Hfq-binding site on DsrA RNA: Hfq binds without altering DsrA secondary structure. *Rna* **9**(1): 33-43.

Brown, L. and Elliott, T. 1996. Efficient translation of the RpoS sigma factor in Salmonella typhimurium requires host factor I, an RNA-binding protein encoded by the hfq gene. *J Bacteriol* **178**(13): 3763-3770.

Butland, G., Peregrin-Alvarez, J.M., Li, J., Yang, W., Yang, X., Canadien, V., Starostine, A., Richards, D., Beattie, B., Krogan, N. et al. 2005. Interaction network containing conserved and essential protein complexes in Escherichia coli. *Nature* **433**(7025): 531-537.

Carmichael, G.G., Weber, K., Niveleau, A., and Wahba, A.J. 1975. The host factor required for RNA phage Qbeta RNA replication in vitro. Intracellular location, quantitation, and purification by polyadenylate-cellulose chromatography. *J Biol Chem* **250**(10): 3607-3612.

Chao, Y. and Vogel, J. 2010. The role of Hfq in bacterial pathogens. *Curr Opin Microbiol* **13**(1): 24-33.

Christiansen, J.K., Larsen, M.H., Ingmer, H., Sogaard-Andersen, L., and Kallipolitis, B.H. 2004. The RNA-binding protein Hfq of Listeria monocytogenes: role in stress tolerance and virulence. *J Bacteriol* **186**(11): 3355-3362.

de Haseth, P.L. and Uhlenbeck, O.C. 1980a. Interaction of Escherichia coli host factor protein with oligoriboadenylates. *Biochemistry* **19**(26): 6138-6146.

-. 1980b. Interaction of Escherichia coli host factor protein with Q beta ribonucleic acid. *Biochemistry* **19**(26): 6146-6151.

Diestra, E., Cayrol, B., Arluison, V., and Risco, C. 2009. Cellular electron microscopy imaging reveals the localization of the Hfq protein close to the bacterial membrane. *PLoS One* **4**(12): e8301.

Dietrich, M., Munke, R., Gottschald, M., Ziska, E., Boettcher, J.P., Mollenkopf, H., and Friedrich, A. 2009. The effect of hfq on global gene expression and virulence in Neisseria gonorrhoeae. *FEBS J* **276**(19): 5507-5520.

Ding, Y., Davis, B.M., and Waldor, M.K. 2004. Hfq is essential for Vibrio cholerae virulence and downregulates sigma expression. *Mol Microbiol* **53**(1): 345-354.

Fantappie, L., Metruccio, M.M., Seib, K.L., Oriente, F., Cartocci, E., Ferlicca, F., Giuliani, M.M., Scarlato, V., and Delany, I. 2009. The RNA chaperone Hfq is involved in stress response and virulence in Neisseria meningitidis and is a pleiotropic regulator of protein expression. *Infect Immun* **77**(5): 1842-1853.

Folichon, M., Allemand, F., Regnier, P., and Hajnsdorf, E. 2005. Stimulation of poly(A) synthesis by Escherichia coli poly(A)polymerase I is correlated with Hfq binding to poly(A) tails. *FEBS J* **272**(2): 454-463.

Folichon, M., Arluison, V., Pellegrini, O., Huntzinger, E., Regnier, P., and Hajnsdorf, E. 2003. The poly(A) binding protein Hfq protects RNA from RNase E and exoribonucleolytic degradation. *Nucleic Acids Res* **31**(24): 7302-7310.

Franze de Fernandez, M.T., Eoyang, L., and August, J.T. 1968. Factor fraction required for the synthesis of bacteriophage Qbeta-RNA. *Nature* **219**(154): 588-590.

Franze de Fernandez, M.T., Hayward, W.S., and August, J.T. 1972. Bacterial proteins required for replication of phage Q ribonucleic acid. Pruification and properties of host factor I, a ribonucleic acid-binding protein. *J Biol Chem* **247**(3): 824-831.

Geng, J., Song, Y., Yang, L., Feng, Y., Qiu, Y., Li, G., Guo, J., Bi, Y., Qu, Y., Wang, W. et al. 2009. Involvement of the post-transcriptional regulator Hfq in Yersinia pestis virulence. *PLoS One* **4**(7): e6213.

Hajnsdorf, E. and Regnier, P. 2000. Host factor Hfq of Escherichia coli stimulates elongation of poly(A) tails by poly(A) polymerase I. *Proc Natl Acad Sci U S A* **97**(4): 1501-1505.

Hammer, B.K. and Bassler, B.L. 2007. Regulatory small RNAs circumvent the conventional quorum sensing pathway in pandemic Vibrio cholerae. *Proc Natl Acad Sci U S A* **104**(27): 11145-11149.

Hwang, W., Arluison, V., and Hohng, S. 2011. Dynamic competition of DsrA and rpoS fragments for the proximal binding site of Hfq as a means for efficient annealing. *Nucleic Acids Res*.

Ikeda, Y., Yagi, M., Morita, T., and Aiba, H. 2011. Hfq binding at RhlB-recognition region of RNase E is crucial for the rapid degradation of target mRNAs mediated by sRNAs in Escherichia coli. *Mol Microbiol* **79**(2): 419-432.

Kaberdin, V.R. and Blasi, U. 2006. Translation initiation and the fate of bacterial mRNAs. *FEMS Microbiol Rev* **30**(6): 967-979.

Kadzhaev, K., Zingmark, C., Golovliov, I., Bolanowski, M., Shen, H., Conlan, W., and Sjostedt, A. 2009. Identification of genes contributing to the virulence of Francisella tularensis SCHU S4 in a mouse intradermal infection model. *PLoS One* **4**(5): e5463.

Kajitani, M., Kato, A., Wada, A., Inokuchi, Y., and Ishihama, A. 1994. Regulation of the Escherichia coli hfq gene encoding the host factor for phage Q beta. *J Bacteriol* **176**(2): 531-534.

Kawamoto, H., Koide, Y., Morita, T., and Aiba, H. 2006. Base-pairing requirement for RNA silencing by a bacterial small RNA and acceleration of duplex formation by Hfq. *Mol Microbiol* **61**(4): 1013-1022.

Kulesus, R.R., Diaz-Perez, K., Slechta, E.S., Eto, D.S., and Mulvey, M.A. 2008. Impact of the RNA chaperone Hfq on the fitness and virulence potential of uropathogenic Escherichia coli. *Infect Immun* **76**(7): 3019-3026.

Le Derout, J., Boni, I.V., Regnier, P., and Hajnsdorf, E. 2010. Hfq affects mRNA levels independently of degradation. *BMC Mol Biol* **11**: 17.

Lease, R.A. and Woodson, S.A. 2004. Cycling of the Sm-like protein Hfq on the DsrA small regulatory RNA. *J Mol Biol* **344**(5): 1211-1223.

Lee, T. and Feig, A.L. 2008. The RNA binding protein Hfq interacts specifically with tRNAs. *RNA* **14**(3): 514-523.

Liu, Y., Wu, N., Dong, J., Gao, Y., Zhang, X., Mu, C., Shao, N., and Yang, G. 2010. Hfq is a global regulator that controls the pathogenicity of Staphylococcus aureus. *PLoS One* **5**(9).

Mackie G.A. 1992. Secondary structure of the mRNA for ribosomal protein S20. Implications for cleavage by ribonuclease E. *J Biol Chem* **267**(2): 1054-1061.

Madhugiri, R., Basineni, S.R., and Klug, G. 2010. Turn-over of the small non-coding RNA RprA in E. coli is influenced by osmolarity. *Mol Genet Genomics* **284**(4): 307-318.

Majdalani, N., Cunning, C., Sledjeski, D., Elliott, T., and Gottesman, S. 1998. DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. *Proc Natl Acad Sci U S A* **95**(21): 12462-12467.

Majdalani, N., Hernandez, D., and Gottesman, S. 2002. Regulation and mode of action of the second small RNA activator of RpoS translation, RprA. *Mol Microbiol* **46**(3): 813-826.

Majdalani, N., Vanderpool, C.K., and Gottesman, S. 2005. Bacterial small RNA regulators. *Crit Rev Biochem Mol Biol* **40**(2): 93-113.

Maki, K., Uno, K., Morita, T., and Aiba, H. 2008. RNA, but not protein partners, is directly responsible for translational silencing by a bacterial Hfq-binding small RNA. *Proc Natl Acad Sci U S A* **105**(30): 10332-10337.

Massè, E., Escorcia, F.E., and Gottesman, S. 2003. Coupled degradation of a small regulatory RNA and its mRNA targets in Escherichia coli. *Genes Dev* **17**(19): 2374-2383.

McNealy, T.L., Forsbach-Birk, V., Shi, C., and Marre, R. 2005. The Hfq homolog in Legionella pneumophila demonstrates regulation by LetA and RpoS and interacts with the global regulator CsrA. *J Bacteriol* **187**(4): 1527-1532.

Meibom, K.L., Forslund, A.L., Kuoppa, K., Alkhuder, K., Dubail, I., Dupuis, M., Forsberg, A., and Charbit, A. 2009. Hfq, a novel pleiotropic regulator of virulence-associated genes in Francisella tularensis. *Infect Immun* **77**(5): 1866-1880.

Mohanty, B.K. and Kushner, S.R. 2006. The majority of Escherichia coli mRNAs undergo post-transcriptional modification in exponentially growing cells. *Nucleic Acids Res* **34**(19): 5695-5704.

Mohanty, B.K., Maples, V.F., and Kushner, S.R. 2004. The Sm-like protein Hfq regulates polyadenylation dependent mRNA decay in Escherichia coli. *Mol Microbiol* **54**(4): 905-920.

Moll, I., Afonyushkin, T., Vytvytska, O., Kaberdin, V.R., and Blasi, U. 2003a. Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs. *RNA* **9**(11): 1308-1314.

Moll, I., Leitsch, D., Steinhauser, T., and Blasi, U. 2003b. RNA chaperone activity of the Sm-like Hfq protein. *EMBO Rep* **4**(3): 284-289.

Moller, T., Franch, T., Hojrup, P., Keene, D.R., Bachinger, H.P., Brennan, R.G., and Valentin-Hansen, P. 2002. Hfq: a bacterial Sm-like protein that mediates RNA-RNA interaction. *Mol Cell* **9**(1): 23-30.

Morita, T., Maki, K., and Aiba, H. 2005. RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes Dev* **19**(18): 2176-2186.

Morita, T., Mochizuki, Y., and Aiba, H. 2006. Translational repression is sufficient for gene silencing by bacterial small noncoding RNAs in the absence of mRNA destruction. *Proc Natl Acad Sci U S A* **103**(13): 4858-4863.

Muffler, A., Traulsen, D.D., Fischer, D., Lange, R., and Hengge-Aronis, R. 1997. The RNA-binding protein HF-I plays a global regulatory role which is largely, but not exclusively, due to its role in expression of the sigmaS subunit of RNA polymerase in Escherichia coli. *J Bacteriol* **179**(1): 297-300.

Nielsen, J.S., Boggild, A., Andersen, C.B., Nielsen, G., Boysen, A., Brodersen, D.E., and Valentin-Hansen, P. 2007. An Hfq-like protein in archaea: crystal structure and functional characterization of the Sm protein from Methanococcus jannaschii. *RNA* **13**(12): 2213-2223.

Olsen, A.S., Moller-Jensen, J., Brennan, R.G., and Valentin-Hansen, P. 2010. C-terminally truncated derivatives of Escherichia coli Hfq are proficient in riboregulation. *J Mol Biol* **404**(2): 173-182.

Overgaard, M., Johansen, J., Moller-Jensen, J., and Valentin-Hansen, P. 2009. Switching off small RNA regulation with trap-mRNA. *Mol Microbiol* **73**(5): 790-800.

Pannekoek, Y., Huis in 't Veld, R., Hopman, C.T., Langerak, A.A., Speijer, D., and van der Ende, A. 2009. Molecular characterization and identification of proteins regulated by Hfq in Neisseria meningitidis. *FEMS Microbiol Lett* **294**(2): 216-224.

Pichon, C. and Felden, B. 2007. Proteins that interact with bacterial small RNA regulators. *FEMS Microbiol Rev* **31**(5): 614-625.

Rauhut, R. and Klug, G. 1999. mRNA degradation in bacteria. *FEMS Microbiol Rev* **23**(3): 353-370.

Robertson, G.T. and Roop, R.M., Jr. 1999. The Brucella abortus host factor I (HF-I) protein contributes to stress resistance during stationary phase and is a major determinant of virulence in mice. *Mol Microbiol* **34**(4): 690-700.

Scheibe, M., Bonin, S., Hajnsdorf, E., Betat, H., and Morl, M. 2007. Hfq stimulates the activity of the CCA-adding enzyme. *BMC Mol Biol* **8**: 92.

Schilling, D. and Gerischer, U. 2009. The Acinetobacter baylyi Hfq gene encodes a large protein with an unusual C terminus. *J Bacteriol* **191**(17): 5553-5562.

Schumacher, M.A., Pearson, R.F., Moller, T., Valentin-Hansen, P., and Brennan, R.G. 2002. Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: a bacterial Sm-like protein. *Embo J* **21**(13): 3546-3556.

Shakhnovich, E.A., Davis, B.M., and Waldor, M.K. 2009. Hfq negatively regulates type III secretion in EHEC and several other pathogens. *Mol Microbiol* **74**(2): 347-363.

Sittka, A., Pfeiffer, V., Tedin, K., and Vogel, J. 2007. The RNA chaperone Hfq is essential for the virulence of Salmonella typhimurium. *Mol Microbiol* **63**(1): 193-217.

Sledjeski, D.D., Gupta, A., and Gottesman, S. 1996. The small RNA, DsrA, is essential for the low temperature expression of RpoS during exponential growth in Escherichia coli. *Embo J* **15**(15): 3993-4000.

Sledjeski, D.D., Whitman, C., and Zhang, A. 2001. Hfq is necessary for regulation by the untranslated RNA DsrA. *J Bacteriol* **183**(6): 1997-2005.

Sonnleitner, E., Hagens, S., Rosenau, F., Wilhelm, S., Habel, A., Jager, K.E., and Blasi, U. 2003. Reduced virulence of a hfq mutant of Pseudomonas aeruginosa O1. *Microb Pathog* **35**(5): 217-228.

Soper, T.J. and Woodson, S.A. 2008. The rpoS mRNA leader recruits Hfq to facilitate annealing with DsrA sRNA. *RNA* **14**(9): 1907-1917.

Sousa, S.A., Ramos, C.G., Moreira, L.M., and Leitao, J.H. 2010. The hfq gene is required for stress resistance and full virulence of Burkholderia cepacia to the nematode Caenorhabditis elegans. *Microbiology* **156**(Pt 3): 896-908.

Steege, D.A. 2000. Emerging features of mRNA decay in bacteria. *RNA* **6**(8): 1079-1090.

Storz, G., Opdyke, J.A., and Zhang, A. 2004. Controlling mRNA stability and translation with small, noncoding RNAs. *Curr Opin Microbiol* **7**(2): 140-144.

Sukhodolets, M.V. and Garges, S. 2003. Interaction of Escherichia coli RNA polymerase with the ribosomal protein S1 and the Sm-like ATPase Hfq. *Biochemistry* **42**(26): 8022-8034.

Sun, X., Zhulin, I., and Wartell, R.M. 2002. Predicted structure and phyletic distribution of the RNA-binding protein Hfq. *Nucleic Acids Res* **30**(17): 3662-3671.

Takada, A., Wachi, M., Kaidow, A., Takamura, M., and Nagai, K. 1997. DNA binding properties of the hfq gene product of Escherichia coli. *Biochem Biophys Res Commun* **236**(3): 576-579.

Tsui, H.C., Feng, G., and Winkler, M.E. 1997. Negative regulation of mutS and mutH repair gene expression by the Hfq and RpoS global regulators of Escherichia coli K-12. *J Bacteriol* **179**(23): 7476-7487.

Tsui, H.C., Leung, H.C., and Winkler, M.E. 1994. Characterization of broadly pleiotropic phenotypes caused by an hfq insertion mutation in Escherichia coli K-12. *Mol Microbiol* **13**(1): 35-49.

Updegrove, T., Wilf, N., Sun, X., and Wartell, R.M. 2008. Effect of Hfq on RprA-rpoS mRNA pairing: Hfq-RNA binding and the influence of the 5' rpoS mRNA leader region. *Biochemistry* **47**(43): 11184-11195.

Updegrove, T.B., Correia, J.J., Chen, Y., Terry, C., and Wartell, R.M. The stoichiometry of the Escherichia coli Hfq protein bound to RNA. *RNA* **17**(3): 489-500.

Valentin-Hansen, P., Eriksen, M., and Udesen, C. 2004. The bacterial Sm-like protein Hfq: a key player in RNA transactions. *Mol Microbiol* **51**(6): 1525-1533.

Vasil'eva Iu, M. and Garber, M.B. 2002. [The regulatory role of the Hfq protein in bacterial cells]. *Mol Biol (Mosk)* **36**(6): 970-977.

Vecerek, B., Moll, I., and Blasi, U. 2005. Translational autocontrol of the Escherichia coli hfq RNA chaperone gene. *RNA* **11**(6): 976-984.

Vecerek, B., Rajkowitsch, L., Sonnleitner, E., Schroeder, R., and Blasi, U. 2008. The C-terminal domain of Escherichia coli Hfq is required for regulation. *Nucleic Acids Res* **36**(1): 133-143.

Vytvytska, O., Jakobsen, J.S., Balcunaite, G., Andersen, J.S., Baccarini, M., and von Gabain, A. 1998. Host factor I, Hfq, binds to Escherichia coli ompA mRNA in a growth rate-dependent fashion and regulates its stability. *Proc Natl Acad Sci U S A* **95**(24): 14118-14123.

Vytvytska, O., Moll, I., Kaberdin, V.R., von Gabain, A., and Blasi, U. 2000. Hfq (HF1) stimulates ompA mRNA decay by interfering with ribosome binding. *Genes Dev* **14**(9): 1109-1118.

Zhang, A., Wassarman, K.M., Ortega, J., Steven, A.C., and Storz, G. 2002. The Sm-like Hfq protein increases OxyS RNA interaction with target mRNAs. *Mol Cell* **9**(1): 11-22.

CHAPTER 2

Effect of Hfq on RprA-*rpoS* mRNA pairing: Hfq-RNA binding and the influence of the 5' *rpoS* mRNA leader region

INTRODUCTION

Short non-coding RNAs (sRNA), approximately 100 nt long, have been shown to be involved in post-transcriptional regulation of mRNA expression in *Escherichia coli* and other bacteria (Gottesman et al. 2001; Gottesman 2004; Storz et al. 2004; Majdalani et al. 2005). Characterization of a number of sRNAs indicate that many are induced under stress conditions and act in trans by base pairing to target sites on specific mRNAs, inhibiting or enhancing translation initiation. A characteristic common to this class of sRNA is their strong binding affinity to Hfq (Wassarman et al. 2001), a protein with RNA chaperone activity (Moll et al. 2003). Initially identified as a host factor for the replication of the RNA phage Qβ (Franze de Fernandez et al. 1968), the importance of Hfq to cell metabolism in *E. coli* was demonstrated by the widespread pleiotropic effects caused by null mutants of the *hfq* gene (Tsui et al. 1994). Phylogenetic and structure analyses showed that Hfq is well conserved in many other bacterial species, and is closely related to the Sm family of RNA binding proteins in archaea and eucaryotes (Arluison et al. 2002; Moller et al. 2002a; Sun et al. 2002; Zhang et al. 2002). During the past several years *E. coli* Hfq was demonstrated to be important to the regulation of mRNA expression by a number of sRNAs including DsrA (Sledjeski et al. 2001), OxyS (Zhang et al. 1998; Zhang et al. 2002), RprA (Majdalani et al. 2002), Spot42 (Moller et al. 2002a; Moller et al. 2002b), RyhB (Geissmann and Touati 2004) and SgrS (Kawamoto et al. 2006). Hfq homologues

in other bacterial species appear to play a similar role in sRNA regulation (Lenz et al. 2004; Antal et al. 2005; Bossi and Figueroa-Bossi 2007).

rpoS was among the first genes shown to be regulated by sRNAs (Brown and Elliott 1996; Sledjeski et al. 1996; Altuvia et al. 1997). This gene encodes the stationary phase sigma factor required to transcribe a group of genes expressed in stationary phase or during growth under stress conditions. *rpoS* is regulated by three sRNAs, 109 nt OxyS (Zhang et al. 1998; Zhang et al. 2002), 87 nt DsrA, and 105 nt RprA. OxyS decreases *rpoS* mRNA translation while DsrA and RprA enhance translation in the presence of Hfq (Sledjeski et al. 1996; Zhang et al. 1998; Sledjeski et al. 2001; Majdalani et al. 2002). *In vivo* studies on the effect of mutations to DsrA and RprA and their *rpoS* mRNA target site indicated that both sRNAs hybridize to one strand of a predicted duplex segment in *rpoS* mRNA just upstream of the start codon (Cunning et al. 1998; Majdalani et al. 1998; Majdalani et al. 2002). This interaction renders accessible the ribosome binding site (RBS) on the other strand. *In vitro* studies provided direct evidence for the DsrA-*rpoS* mRNA interaction (Lease and Woodson 2004) and showed that Hfq independently binds to DsrA and a site near the RBS of the *rpoS* mRNA (Lease and Woodson 2004; Mikulecky et al. 2004). Hfq was also shown to enhance DsrA binding to a 140 nt *rpoS* RNA (Lease and Woodson 2004), but by an amount (1.8-fold) thought to be small relative to the influence of Hfq on DsrA regulation of *rpoS* mRNA translation *in vivo* (Majdalani et al. 2005).

Hfq has been shown to enhance the binding of a sRNA to its mRNA target for several other sRNA-mRNA pairs in addition to DsrA and *rpoS* mRNA. They include Spot42 and *galK* mRNA (Moller et al. 2002a), OxyS and *fhlA* mRNA (Zhang et al. 2002), RyhB and *sodB* mRNA (Geissmann and Touati 2004), and SgrS and *ptsG* mRNA (Kawamoto et al. 2006). How Hfq enhances sRNA-mRNA pairing is not fully understood. For DsrA and *rpoS* mRNA as well as

some of the above RNA pairs, *in vitro* experiments indicate that Hfq is not required to maintain stable sRNA-mRNA complexes once they form (Zhang et al. 2002; Lease and Woodson 2004; Afonyushkin et al. 2005; Kawamoto et al. 2006). This suggests that Hfq's RNA chaperone activity may alter the conformation of a sRNA or its mRNA target creating a metastable conformation that enables intermolecular hybridization (Brescia et al. 2003). Enhanced hybridization may also result from an Hfq complex sequestering the two RNAs simultaneously thereby increasing their local concentration and allowing thermal fluctuations or transient binding-release by Hfq to drive hybridization.

Evidence that Hfq binding can induce conformational change to a mRNA or sRNA has been obtained from *in vitro* nuclease footprinting studies with *sodB* mRNA (Brescia et al. 2003; Geissmann and Touati 2004), *ompA* mRNA (Udekwu et al. 2005), and OxyS sRNA (Zhang et al. 2002). This assay did not, however, indicate a significant alteration to the secondary structure of *rpoS* RNA (Lease and Woodson 2004) . FRET studies showed that Hfq binding to DsrA alters the distance between this RNA's 5' and 3' ends (Vecerek et al. 2008), although it did not produce a significant change to the DsrA CD spectrum (Brescia et al. 2003). The above studies indicate that Hfq binding can alter DsrA conformation, but major distortion in secondary structure is not evident. The role of an Hfq-induced conformational change to RNA in the annealing of a sRNA to its mRNA site remains uncertain.

Three-dimensional crystal structures have been obtained of the *Staphylococcus aureus* Hfq (Schumacher et al. 2002), a truncated version of the *E. coli* Hfq (residues 4-72) (Sauter et al. 2003), and *Pseudomonas aerogenosa* Hfq (Nikulin et al. 2005). The structures are very similar, each forming a hexameric toroid with an outer diameter of ~70 Å and a central cavity ~ 10 Å wide (Figure 2.7). The crystal structure of *S. aureus* Hfq with the heptanucleotide AU_5G shows

this RNA binding in a circular contour adjacent to the central cavity on Hfq's proximal surface (Schumacher et al. 2002). Mutagenesis studies indicate that residues along the corresponding contour of *E. coli* Hfq (Mikulecky et al. 2004) as well as residues F39 and R16 on the proximal surface (Sun and Wartell 2006) influence Hfq binding to DsrA. Mutations to some residues on the distal surface of *E. coli* Hfq did not have a significant effect on its binding to DsrA or *rpoS* RNA (Mikulecky et al. 2004), but they strongly influenced Hfq binding to poly(rA) (Mikulecky et al. 2004; Sun and Wartell 2006). A recent investigation showed that residues beyond position 65 on the C-terminal end of *E. coli* Hfq are needed for sRNA regulation *in vivo* and influence Hfq binding to *rpoS* mRNA and DsrA *in vitro* (Vecerek et al. 2008).

In the current work we examined the binding of *E. coli* Hfq to RprA and two lengths of the *rpoS* mRNA leader region, and the impact of Hfq on the interaction of RrpA with the *rpoS* RNAs. RprA is similar to DsrA in that it binds to the same target region on *rpoS* mRNA, enhances translation, and requires Hfq. Yet RprA differs in sequence and length from DsrA. Comparison of the interactions of RprA, Hfq and *rpoS* RNAs with previous results on DsrA, Hfq and *rpoS* mRNA can be expected to reveal common features and differences that may shed light on how Hfq enhances the pairing of these sRNAs with their common target site.

Our studies show that Hfq forms two complexes with RprA in a gel shift assay with binding parameters similar to those determined for Hfq and DsrA (Lease and Woodson 2004; Mikulecky et al. 2004). Unlike DsrA, Hfq binding to RprA was inhibited by poly(A) and influenced by Hfq mutations on the distal and proximal surfaces. RprA bound to the two *rpoS* RNAs examined but with different affinities. The 654 nt *rpoS* RNA containing the entire upstream leader sequence showed much weaker binding to RprA than the 210 nt *rpoS* RNA although both *rpoS* RNAs have the RprA target sequence. Hfq induced a much larger

enhancement of RprA binding to the 654 nt *rpoS* RNA compared to the 210 nt *rpoS* RNA. This is consistent with a previous *in vivo* result (Cunning et al. 1998) indicating that the full upstream leader sequence of *rpoS* mRNA plays an important role in Hfq activity.

MATERIALS AND METHODS

Plasmid construction and transcription of *rpoS*, RprA and DsrA RNAs

DNA plasmid templates for *rpoS* RNA transcription were constructed from PCR amplification of segments of the *rpoS* gene in purified *E. coli* K12 DNA. Two primer pairs were used to amplify DNA fragments containing 646 bp and 199 bp of the *rpoS* 5'UTR. Subsequent rounds of PCR amplification used standard procedures positioned an EcoRI restriction site and phage T7 promoter sequence upstream of the *rpoS* DNAs and a HindIII restriction site downstream. PCR products of this second round were purified by agarose gel electrophoresis and ligated into the polylinker region of pUC19. The shorter *rpoS* DNA insert created the plasmid designated pUC19*rpoS*-S and the plasmid with the longer *rpoS* DNA insert was designated pUC19*rpoS*-L. Plasmids were transformed into DH5α cells and colonies selected on ampicillin-LB plates. Isolated plasmid DNA was sequenced to verify the inserted DNA sequences.

The run off transcript from HindIII digested pUC19*rpoS*-L includes the start point of *rpoS* mRNA transcription, 565 nt behind the *rpoS* start codon (Lange et al. 1995) and 75 nt in front of this AUG. This 654 nt transcript (*rpoS*-L) also included six nucleotides (GGGAGA) at the 5' end from the T7 promoter sequence and five nucleotides from the HindIII site at the 3'end. The transcribed product from pUC19*rpoS*-S (*rpoS*-S) has the same 3' end as *rpoS*-L but has only 127 nt of the *rpoS* sequence upstream of the start codon. *rpoS*-S is 210 nt and includes six nucleotides of the T7 promoter sequence at its 5' end and the HindIII 3' end sequence (Figure 2.2). A 157 nt segment of the *rpoS* mRNA leader region was produced for a limited number of studies by transcribing pUC19*rpoS*-S digested with the restriction enzyme DraI. This transcript

ends 32 nt downstream of the start codon. RprA was produced from a plasmid constructed in a similar manner to the other plasmids. The run off transcript, 112 nt, added two G's at the 5' end from the T7 promoter and five nucleotides due to the HindIII site at the 3' end to the RprA sequence. Secondary structures predicted using RNA Structure 4.5 (Mathews et al. 2004) do not suggest that the additional nucleotides alter folding of the RprA sequence. DsrA was produced from a pUC19 based plasmid using a similar approach. Plasmid was digested by DraI for run off transcription. The RNA contained the two G's from the T7 promoter at the 5'end and the stretch of 6 A's at the 3' end of the DsrA sequence.

RNAs were synthesized using the MEGAscript T7 kit (Ambion) according to manufacturer's protocol. RNAs were purified by ammonium acetate precipitation after digestion of the template with DNase (Epicentre). RNAs were characterized by native and denaturing gel electrophoresis, and their concentrations determined by UV absorbance and RiboGreen fluorescence assay (Invitrogen Inc). ^{32}P labeling of the RNAs was carried out using standard protocols: RNA was dephosphorylated with shrimp alkaline phosphatase, radioactively labeled at the 5' end with [γ-^{32}P]ATP and T4 polynucleotide kinase, and purified by gel electrophoresis followed by elution from gel slices in 0.5 M ammonium acetate and ethanol precipitation. RNAs were dissolved in DEPC treated water or 2 mM sodium citrate and 0.1 mM EDTA and stored at -70 ^{o}C.

Purification and characterization of wild type (wt) and mutant Hfq

The Impact-CN intein system (New England Biolabs) was used to purify Hfq proteins as previously described (Sun and Wartell 2006). The *E. coli hfq* gene was amplified by PCR using *E. coli* chromosomal DNA as the template. PCR products were digested with SapI and cloned

into a SapI-SmaI digested pTYB11 plasmid. Protein purification was carried out according to the recommendation of the manufacturer using strain ER2566. Cell lysis was carried out using a french press. The cell lysate was centrifuged and the supernatant loaded onto a chitin column. The column was extensively washed with the lysis/wash buffer of 20 mM Tris (pH 8.3) and 1 M NaCl prior to incubation of the column with this buffer plus 40 mM dithiothreotol. The eluted protein was concentrated and buffer-exchanged to 0.5 M NaCl and 20 mM Tris at pH 8.3 using centrifugation filtration units.

Mutant Hfq proteins were produced as described previously from plasmids containing mutant *hfq* genes generated using the QuikChange Mutagenesis Kit from Stratagene Inc (Sun and Wartell 2006). Plasmid constructs were verified by DNA sequencing. The same procedure used to purify wt Hfq was used to purify the mutant proteins. All proteins showed the 11 kD monomer band by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) with purity estimated to be ~95% from Commassie blue staining. Characterization was also carried out using analytical sedimentation velocity centrifugation in 0.5 M NaCl and 20 mM Tris (pH 8.3), circular dichroism, and UV absorbance spectra. The absorbance ratio of A_{274}/A_{250} indicated less than 5% of contaminating nucleic acid.

RNase H degradation assay

DNA oligonucleotide probes were purchased commercially, dissolved in TE and their concentrations evaluated by UV absorbance. The sequences of the five probes in the 5' to 3' direction were: (1) GCTCCTAC, (2) CGATTTAT, (3) GCAAATAAC, (4) GACGGAAC, and (5) CGCAGCGG. A master mixture (60 μL) containing RNA and E. coli RNase H1 (USB) was preincubated for 15 min at 25 °C in RNase H buffer with 0.2 unit of prime RNase inhibitor. The

RNase H buffer contained 20 mM KCl, 10 mM $MgCl_2$, 20 mM Tris (pH 7.5), 0.1 mM EDTA, 0.1 mM DTT. 7 μL from this mixture was added to each reaction tube with 3 μL of a given DNA probe in RNase H buffer. The final DNA and RNA concentrations were 15 μM and 1.5 μM respectively in each tube. Each reaction was incubated at 37 °C for 15 min. 30% glycerol loading buffer was mixed directly to the samples which were heated for one minute at 80 ^{o}C, placed in ice, and then run immediately into an 8% polyacrylamide gel for *rpoS*-S RNA and 5% gel for *rpoS*-L RNA. For *rpoS*-L, nucleic acid bands were visualized by staining the gel with SYBR Gold (Invitrogen). For *rpoS*-S, the master mixture contained ^{32}P labeled RNA as a marker. Gels were scanned and analyzed using a Fujifilm Image Reader FLA-3000 in fluorescence or IP mode. The percentage of RNase H induced degradation of RNA was evaluated by comparing the band intensities of the full length RNAs with and without DNA probes.

Electrophoretic gel mobility shift assay

Binding reactions between wt Hfq and mutant Hfq and ^{32}P labeled *rpoS*-S or RprA RNAs were prepared in 15 μL volumes. 5 μL of 12 nM RNA (4 nM final concentration) were mixed with 7.5 μL of an Hfq solution to give the appropriate Hfq concentration and 2.5 μL of loading buffer (0.25% bromophenol blue, 30% glycerol). The final reaction solvent consisted of 20 mM Tris–HCl (pH 8.0), 100 mM NH_4Cl, 50 mM NaCl, 50 mM KCl and 5% glycerol. The RNA and Hfq solutions were prepared in the above solvent except for the glycerol. Reactions were incubated for five minutes at 37 °C and then for ten minutes at room temperature (25 ^{o}C). For experiments in which poly(A) or A_{18} were added, an additional 10 minute incubation at 25 ^{o}C followed the additions prior to loading samples.

Samples were analyzed by electrophoresis on native 5% polyacrylamide gels [29:1 (w/w) acrylamide /bisacrylamide] gels with 3% glycerol in 0.5X TBE. Gels were run at 80 – 100 V at room temperature. The fraction of ^{32}P-labeled RNA that was free, or in Hfq•RNA complexes was determined from the counts in each band relative to the total counts in each lane. Hfq binding to *rpoS*-L RNA was carried out in a similar manner except that SYBR Gold stain was employed to image and analyze Hfq binding. Labeling of *rpoS*-L with ^{32}P was also carried out, but degradation of this longer RNA during purification compromised its use.

Association of *rpoS*-S or *rpoS*-L RNA with ^{32}P-labeled RprA in the absence or presence of Hfq was evaluated using the same reaction volume and buffer described above. Varying concentrations of unlabeled *rpoS*-S or *rpoS*-L RNA were added to 4 nM RprA and incubated for 60 minutes at 25 ^{o}C prior to electrophoresis at 4 ^{o}C on a native 5% polyacrylamide gel. Gels were run at 100 - 115 V for 60 to 90 minutes and analyzed using a Fujifilm Image Reader FLA-3000.

Evaluation of binding constants

Equilibrium binding parameters describing the interaction of Hfq with RprA were determined from a least squares fit of the gel shift data to a model that assumed Hfq_6 may cooperatively bind to two different independent sites on RprA. The fraction of ^{32}P-labeled RprA that was free or in the observed complexes C_1 and C_2 (f_o, f_1, and f_2 respectively) was calculated from the counts in each band relative to the total counts in the lane. The experimental fractions are related to the model by the equations

$$f_1 = ([H]/K_1)^{n_1} / Q_2 \qquad (1a)$$

$$f_2 = ([H]/K_1)^{n_1} ([H]/K_2)^{n_2} / Q_2 \qquad (1b)$$

$$f_o = 1\text{-}f_2 - f_1 \qquad (1c)$$

with $$Q_2 = 1 + ([H]/K_1)^{n_1} + ([H]/K_1)^{n_1} ([H]/K_2)^{n_2} \qquad (2)$$

[H] is the free Hfq_6 concentration, K_1 and K_2 are the equilibrium dissociation constants for Hfq binding to the first and second site respectively, and n_i is the Hill cooperativity coefficient for binding site i. [H] was determined from the equation

$$[H] = [H_T] - n_1 f_1 [R_T] - (n_1 + n_2) f_2 [R_T] \qquad (3)$$

$[H_T]$ and $[R_T]$ are the total concentrations of Hfq_6 and RprA respectively. This model is the same as that employed by Lease and Woodson (Lease and Woodson 2004) except that it considers the possibility of different n values for the two sites and accounts for the Hfq concentration in terms of moles hexamer rather than monomer. Experimental values of f_1, f_2, $[H_T]$ and R_T and an assumed (n_1,n_2) pair were employed in eq. (3) to calculate [H] values which were then employed in eq. (1) and (2) to solve for K_1 and K_2 by a non-linear least squares fit to the data (Sigma Plot, SPSS Inc). The process was iterated varying (n_1, n_2) in 0.1 increments between $1< n_i < 3$. Solutions were assessed for convergence and the best fit determined from the maximum correlation coefficient R^2.

Hfq binding to *rpoS*-S was analyzed using a similar approach with a model that assumed Hfq_6 may cooperatively bind to five independent sites. The Hill coefficient for all sites was assumed to be the same ($n_i = n$). The fractions of *rpoS*-S in the five complexes were related to the binding parameters by equations similar to eq. (1);

$$f_j = \prod_{i=1}^{j} ([H]/K_i)^{n_i} / Q_5 \ , \quad \text{for} \quad j = 1 \text{ to } 5 \qquad (4a)$$

$$Q_5 = 1 + \sum_{j=1}^{5} \prod_{i=1}^{j} ([H]/K_i)^{n_i} \qquad (4b)$$

Unlike the gel shift results with *rpoS*-S, Hfq binding to *rpoS*-L RNA did not display complexes as discrete bands. A broad band was observed that decreased in mobility with increasing Hfq (Figure 2.4B). In order to quantify the affinity of Hfq_6 to this RNA a simplified analysis was employed that considers the RNA to be in two states, free or bound;

$$nH + R \leftrightarrow H_n \bullet R \qquad (5)$$

R represents *rpoS*-L RNA and n represents an average stoichiometry of the complexes. The dissociation constant K of eq. (5) can be related to the fraction of unbound or free RNA, $f = [R]/[R_T]$, by the following equation (Fried and Crothers 1984);

$$\log[(1-f)/f] = \log K + n \log[H] \qquad (6)$$

K is related to the geometric mean of the 'n' intrinsic binding constants that lead to $H_n \bullet R$, i.e., $K = (K_d)^n$ where $K_d = (K_1 K_2 \ldots K_n)^{1/n}$. The intercept of the plot of $n^{-1} \log[(1-f)/f]$ vs log[H] yields K_d, a value that reflects the mean of the binding constants for the n binding steps leading to $H_n \bullet R$.

Comparison of the binding affinity of various mutant and wild-type Hfq's for RprA and *rpoS*-S also employed the above simplified analysis. This approach was used since the weak affinity of some mutant Hfq's produced broad sometimes overlapping complex bands making them difficult to quantify. This compromised the ability to compare all mutant Hfq's using the more detailed analysis described by eq (1) to (4).

RESULTS

Hfq binding to RprA

Immunoprecipitation experiments have shown that Hfq binds to RprA (Wassarman et al. 2001); however, the affinity of Hfq for this sRNA has not been previously investigated. Figure 2.1A shows a gel mobility shift experiment of 4 nM RprA with varying amounts of Hfq. Two Hfq-RprA complexes are observed (C_1 and C_2). Their relative mobility implies that more Hfq is bound to the second complex. Association and dissociation of Hfq with RprA was rapid. Experiments showed all of the RprA shifted to the C_1 band within ~30 seconds after adding 40 nM Hfq_6 to 4 nM RprA, and preformed C_1 complex dissociated completely to free RprA within ~30 seconds after adding saturating amounts of poly-rU to trap free Hfq (data not shown). This rapid association and dissociation behavior is similar to what was observed for DsrA (Lease and Woodson 2004).

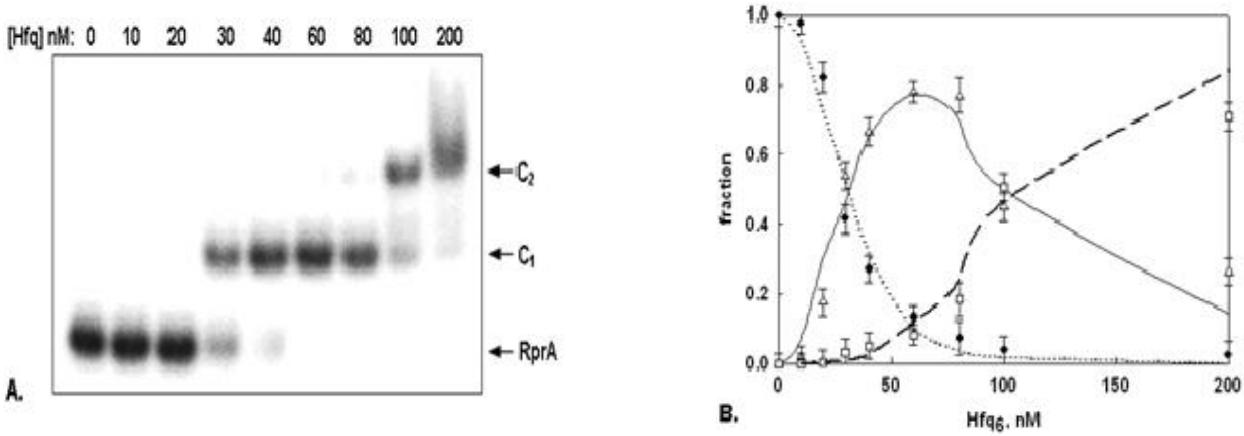


Figure 2.1: (**A)** Hfq binding to ^{32}P-labeled RprA assessed by a gel mobility shift assay. The Hfq_6 concentration varied from 0 to 200 nM. Hfq-RprA complexes are designated C_1 and C_2. (**B)** Fraction of total RprA concentration (f) as a function of Hfq_6 concentration in free RprA, (♦), complex C_1 (Δ), and complex C_2 (□) bands. Error bars based on three experiments. Lines are least-squares fit of binding model described in eqs (1) and (2) using parameters in Table 2 (f_o •••••• ; f_1 —— ; f_2 - - - - -) .

Equilibrium dissociation constants for the formation of complexes C_1 and C_2 were evaluated using the Hill cooperative binding model (eqs. (1) to (3)). This model assumes that one or more Hfq_6 may bind RprA to form complex C_1 and additional Hfq_6 bind C_1 to form complex C_2. The least squares fit yielded equilibrium dissociation constants of $K_1 = 24 \pm 3$ nM and $K_2 = 96 \pm 9$ nM with Hill coefficients ranging from $2.1 < (n_1, n_2) \leq 2.8$ (Figure 2.1B). The correlation coefficient R^2 was 0.945 ± 0.02 in this range. Values of n_1 and n_2 outside this range reduced the quality of the fit. If we fixed n_1 and n_2 to 1, we obtained an R^2 of 0.77. Thus approximately two Hfq_6 bound to each site gave a markedly better fit to the data than one Hfq_6 per site. Mixing increasing amounts of Hfq_6 to 100 nM RprA, a concentration above the K_d of complex C_1, saturated the C_1 complex at a 2:1 molar ratio of Hfq_6 to RprA (Figure S2.1) consistent with the model fitting analysis.

rpoS mRNA leader sequence and characteristics of the RBS secondary structure

Hfq has been previously shown to bind to a portion of the *rpoS* mRNA leader sequence that surrounds the AUG start codon (Lease and Woodson 2004; Mikulecky et al. 2004). The *rpoS* RNAs were ~140 nt long extending from -134 to +3 or -128 to +12 relative to the start codon, and predicted to fold up in the manner displayed in Figure 2.2. Since RprA and DsrA bind to a sequence within this region (Majdalani et al. 1998; Majdalani et al. 2002), and DsrA binding makes the ribosome binding site (RBS) accessible (Lease and Woodson 2004), this ~140 nt region is essential for Hfq-sRNA stimulated translation. If this *rpoS* RNA region is all that is required for the functional interactions of RprA and Hfq, upstream or downstream sequences are not expected to have a strong influence. However, *in vivo* studies indicate that *rpoS* mRNA

sequence more than 220 nt upstream of the start codon is required for regulation by Hfq (Cunning et al. 1998).

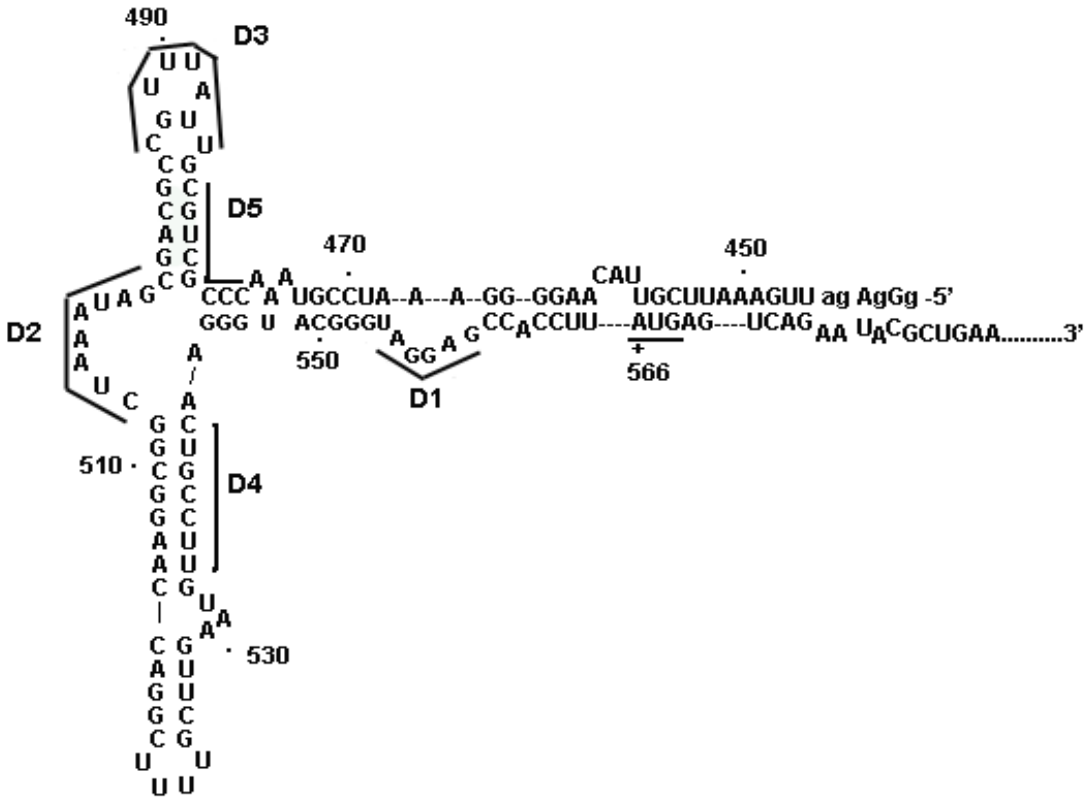
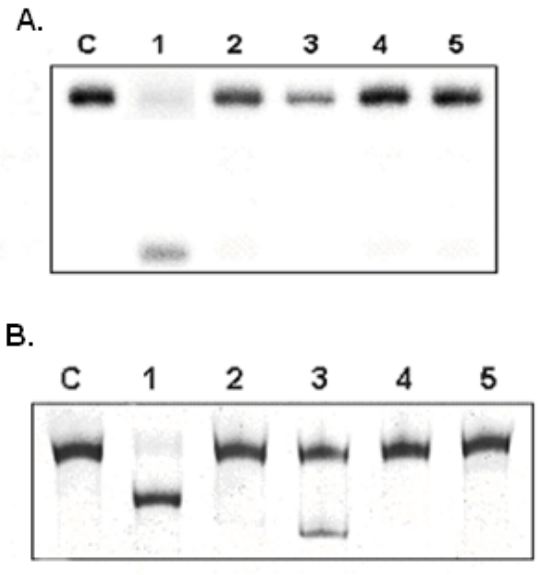


Figure 2.2: Predicted secondary structure of *rpoS*-S RNA except for 58 nt at 3' end. AUG start codon is underlined. Numbers designate base position from transcription start point. Segments D1 - D5 are target locations of DNA probes used in RNase H assay.

In order to explore the influence of the sequence upstream and downstream of the RBS on Hfq interaction with *rpoS* mRNA and RprA-*rpoS* RNA pairing *in vitro*, two transcripts of the *rpoS* mRNA leader region were synthesized. *rpoS*-L RNA contains the start point of transcription 565 nt upstream of the start codon and 75 nt downstream from this AUG (Materials and Methods). *rpoS*-S RNA includes 127 nt of the *rpoS* mRNA sequence upstream of the start

codon and has the same downstream sequence as *rpoS*-L. An RNase H assay was employed to examine if the secondary structure in the vicinity of the RBS was preserved in these two RNAs.

The predicted secondary structure of the sequence surrounding the RBS of the *rpoS* mRNA is shown in Figure 2.2 along with the location of five sites targeted by complementary DNA probes in the RNase H assay. RNase H degradation of *rpoS*-S and *rpoS*-L RNAs with the five DNA probes is shown in Figure 2.3. Probes 1 and 3 target sites predicted to be unpaired and both induce significant RNA degradation. Probe 1 induced approximately 90 ± 7 % degradation to both RNAs, while probe 3 produced 57± 10 % degradation of *rpoS*-S RNA and 30 ± 7 % degradation of *rpoS*-L RNA. DNA probes 4 and 5 target sites are predicted to be base paired, and as expected both probes produced very small amounts of degradation in both RNAs. Probe 2 targets a site predicted to form an unpaired loop, however little degradation was observed for both RNAs. The DNA-RNA hybrid expected to be formed by probe 2 has the lowest predicted T_m among the DNA-RNA hybrids. This factor and/or the RNA 3D-structure may account for the small amount of degradation produced. The results from this assay indicate that the secondary structure surrounding the RBS is similar for *rpoS*-S and *rpoS*-L RNAs.



Probe:	*rpoS*-S RNA % degradation (±7%):	*rpoS*-L RNA % degradation (±7%):
1	91	85
2	6	9
3	57	30
4	6	4
5	8	10

Figure 2.3: Results of RNase H assay of *rpoS*-S and *rpoS*-L using five DNA oligomer probes (see Materials and Methods). (**A)** Degradation of *rpoS*-S: C, control (no DNA); lanes 1 – 5, degradation with corresponding DNA probes 1 – 5, respectively. (**B)** Degradation of *rpoS*-L: C, control (no DNA); lanes 1 – 5, degradation with corresponding DNA probes 1 – 5, respectively.

Hfq binding to *rpoS*-S and *rpoS*-L RNAs

Electrophoretic gel mobility shift assays were carried out to determine the binding affinity of Hfq to *rpoS*-S and *rpoS*-L RNAs. When 4 nM *rpoS*-S RNA was titrated with increasing amounts of Hfq five discrete bands were observed (Figure 2.4A). The faster mobility bands decreased in intensity as the slower mobility bands appeared. Increasing the *rpoS*-S RNA concentration resulted in an increase of the faster mobility complexes at the expense of the slower complexes (data not shown). Both observations are consistent with the complexes (C_i, i = 1-5) representing one *rpoS*-S RNA with increasing numbers of bound Hfq_6 per RNA. The data was fit to the model described in Materials and Methods associated with eq. (4). The evaluated equilibrium constants for the five complexes (C_1 - C_5) were 50 nM, 66 nM, 89 nM, 92 nM and 97 nM (±10%), respectively, with n of 2.4. The correlation coefficient R^2 was 0.93. Hill

coefficients ranging from 2.2 to 2.6 produced similar results for the K_i values and R^2. If n was set equal to 1, R^2 was lower (0.875) and the least-squares solution for K_1 - K_5 was physically unrealistic. The binding constants were 91 nM, 63 nM, 134 nM, 119 nM, and 63 nM, respectively, no longer increasing in value in the order of appearance of the complexes. Thus, approximately two Hfq_6 per site produced a better fit to the *rpoS*-S RNA gel shift data than one Hfq_6 per site.

Hfq binding to *rpoS*-L RNA is shown in Figure 2.4B. As the ratio of Hfq_6 to *rpoS*-L RNA increased the complex(es) migrated as a broad band with decreasing mobility. Since discrete complex bands were not observed, the data was analyzed using a binding model that considered *rpoS*-L RNA as either free or bound (i.e., eqs. (5) and (6)). The apparent K_d was estimated to be 35 nM ± 10 nM from three experiments indicating a slightly greater affinity than the strongest binding site (50 nM) determined for *rpoS*-S RNA.

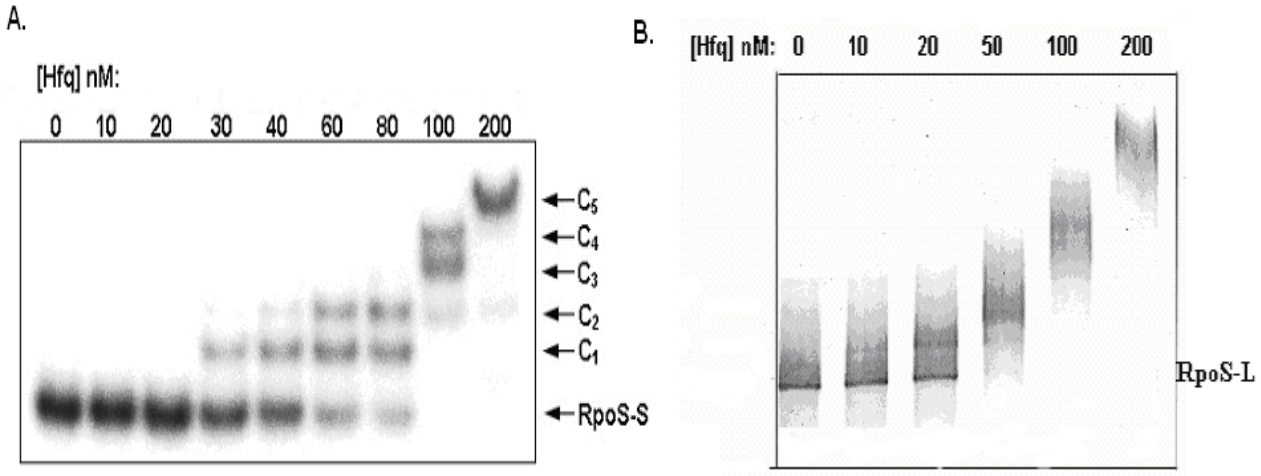


Figure 2.4: Hfq binding to *rpoS*-S (**A**) and *rpoS*-L **(B)** by gel mobility shift assays. The Hfq_6 concentration varied from 0 to 200 nM. Hfq-*rpoS*-S complexes designated C_1 - C_5. An 8% polyacrylamide gel was employed for *rpoS*-S and a 5% polyacrylamide gel employed for *rpoS*-L.

These results show that Hfq binds to the *rpoS*-S and *rpoS*-L RNAs with similar affinity. Multiple discrete complexes are formed with the 210 nt *rpoS*-S RNA, and the decreasing mobility of the Hfq- *rpoS*-L RNA band with increasing Hfq_6 suggests that multiple Hfq_6-RNA complexes also form for this RNA. The inability to resolve discrete bands may be attributed to this RNA's longer length and/or the overlapping mobility of different complexes. More than one discrete Hfq-RNA complex has been previously observed in gel shift experiments for several sRNAs or mRNA targets of Hfq (Zhang et al. 2002; Lease and Woodson 2004; Mikulecky et al. 2004; Mohanty et al. 2004). Unless this behavior is an artifact of the gel conditions, it indicates Hfq_6 either commonly binds to several different sites on its RNA target molecules, or cooperatively associates with RNA-bound Hfq_6, or both.

Table 2.1: Equilibrium binding parameters of Hfq-RNA and RprA-*rpoS* RNA complexes[a]

Components	Complexes	K_d values (nM)	n
Hfq and RprA	2	24, 96	2.1 - 2.8
Hfq and *rpoS*-S	5	50, 66, 89, 92, 97	2.2- 2.6
Hfq and *rpoS*-L		35 ± 10[b]	
RprA and *rpoS*-S		120 ± 20[b]	
RprA and *rpoS*-L		2500 ± 300[b]	

[a] Parameters of Hfq binding RprA and *rpoS*-S were obtained by least squares fit of gel shift data to cooperative binding models described by equations (1) to (4). 'Complexes' are the number of Hfq-RNA gel bands observed. 'K_d values' are the equilibrium dissociation constants for complexes in the order at which they appeared with increasing Hfq_6 concentration. ' n' is the Hill cooperative binding parameter. [b] Apparent K_d values of Hfq binding to *rpoS*-L, and RprA binding to *rpoS*-S or *rpoS*-L. Evaluated from the Hfq_6 concentration at which 50% free RNA is shifted to complexes. Values are averaged from three experiments. Reactions incubated at 25 ^{o}C.

Binding of RprA to *rpoS*-S and *rpoS*-L RNAs and the influence of Hfq

In vivo studies indicate that RprA stimulation of *rpoS* translation is enhanced by Hfq (Majdalani et al. 2002), presumably by Hfq promoting the binding of RprA to its *rpoS* mRNA target site. In the current work, we examine the equilibrium binding of RprA to the core *rpoS* RNA (*rpoS*-S) and the full leader region of the *rpoS* mRNA (*rpoS*-L) at 25 ^{o}C, and the influence of Hfq on these interactions. Figure 2.5A shows a polyacrylamide gel of RprA binding to *rpoS*-S RNA in the absence of Hfq. The equilibrium dissociation constant for forming the RprA· *rpoS*-S complex was ~120 ± 20 nM based on three trials. When a 157 nt portion of *rpoS*-S RNA was employed, which was missing 53 nt at the 3' end, the K_d value was similar (data not shown). Figure 2.5B shows a similar gel shift experiment of RprA binding to *rpoS*-L RNA. The stability of the RprA•*rpoS*-L RNA complex was considerably weaker with a K_d ~2500 nM. The complete upstream leader sequence of *rpoS*-L RNA apparently inhibits RprA binding.

The influence of Hfq on the binding of RprA to *rpoS*-S and *rpoS*-L RNAs was examined next. Hfq_6 (40 nM) was mixed with 4 nM RprA, and varying concentrations of *rpoS*-S RNA were added (Figure 2.5C). Hfq_6 forms the strong binding complex with RprA (C_1) in the absence of *rpoS*-S RNA. As *rpoS*-S RNA was added, the free RprA band initially increased and then decreased as the RprA•*rpoS*-S RNA band formed. The latter band was assigned to RprA•*rpoS*-S RNA since its mobility relative to free RprA was the same as in Figure 2.5A. The initial increase in free RprA indicates that *rpoS*-S RNA can displace Hfq from the RprA-Hfq complex. Although Hfq has a greater affinity for RprA than *rpoS*-S RNA (Table 2.1), the larger concentration of *rpoS*-S RNA vs RprA overcomes this difference in affinity. The *rpoS*-S RNA concentration at which the intensity of the RprA•*rpoS*-S RNA band equaled the intensity of the free RprA band was 80 ± 20 nM. The stability of the RprA• *rpoS*-S RNA complex increased by ~1.5 fold in the

presence of 40 nM Hfq_6. Figure 2.5D shows the effect of 200 nM Hfq_6 on the formation of the RprA•*rpoS*-L RNA complex. The apparent binding constant of RprA for *rpoS*-L RNA was reduced to ~75 nM, a ~30-fold increase in stability of RprA for the full length *rpoS* mRNA leader region in the presence of Hfq_6.

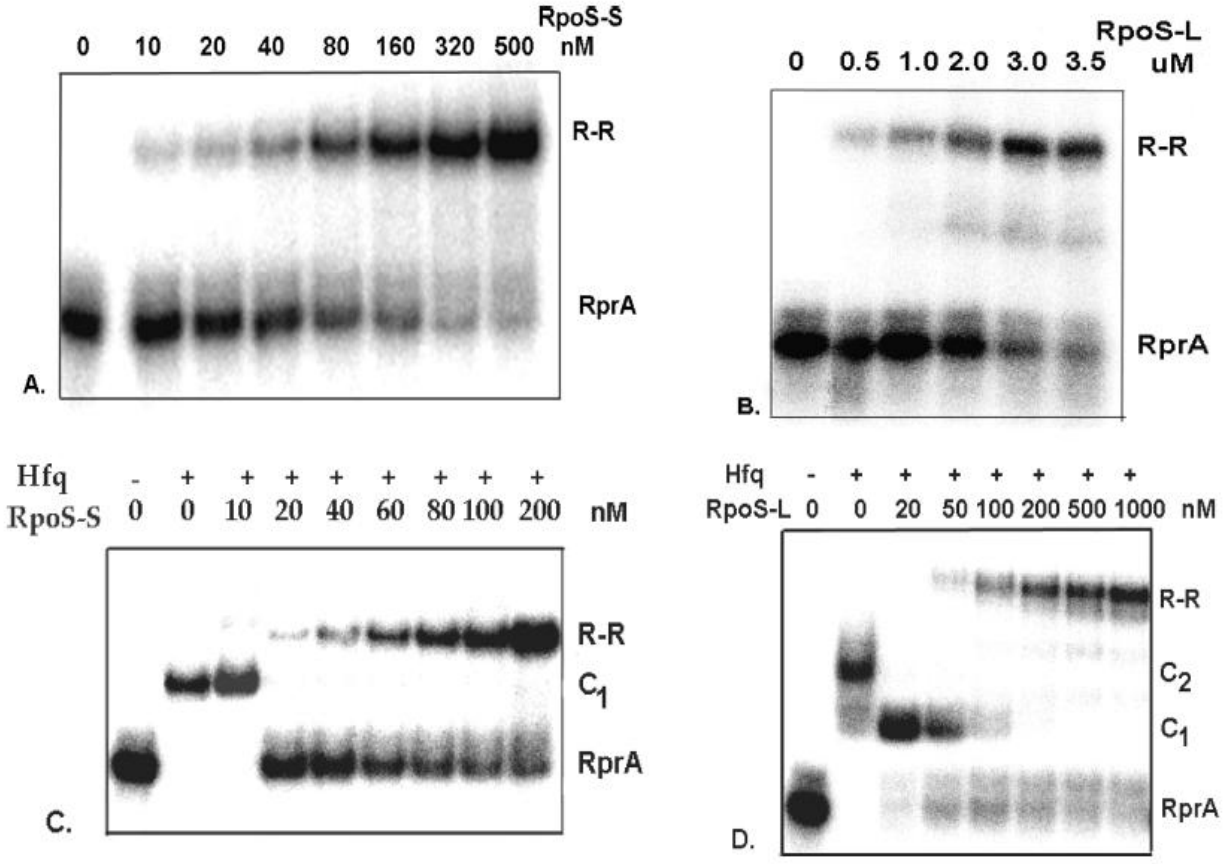


Figure 2.5: (**A)** Binding of *rpoS*-S to ^{32}P-labeled RprA. *rpoS*-S (0 – 500 nM) was added to 4 nM RprA and run in 5% polyacryamide gels (see Materials and Methods). R-R denotes the RprA-*rpoS*-S complex. (**B)** Binding of *rpoS*-L to ^{32}P-labeled RprA. *rpoS*-L (0 - 3.5 μM) was added to 4 nM RprA and run on 5% polyacrylamide gels. (**C)** Binding of *rpoS*-S to RprA as in panel (**A)** in the absence (-) or presence (+) of 50 nM Hfq_6. C_1 denotes the strong Hfq-RprA complex of Figure 2.1. (**D)** Binding of *rpoS*-L to RprA as in panel (**B)** in the absence (-) or presence (+) of 200 nM Hfq_6.

In addition to the much larger effect of Hfq on the binding of RprA to *RpoS*-L RNA compared with RprA binding *rpoS*-S RNA, several points are worth noting in the above experiments. RprA binding to both *rpoS* RNAs in the absence of Hfq required incubation times of approximately 40 minutes to ensure equilibrium. When Hfq was present the time required to reach equilibrium was reduced (Figure S2.2). Thus, Hfq enhancement of the stability of the RNA hybrids correlates with an increased association rate. A more detailed analysis of association-dissociation kinetics is being investigated. The concentrations of Hfq_6 that were employed in Figures 2.5C and 2.5D optimized formation of the RprA•*rpoS* RNA complex, i.e. gave the lowest apparent equilibrium dissociation constants. If more than 100 nM of Hfq_6 was incubated with the 4 nM RprA prior to adding *rpoS*-S RNA, formation of RprA•*rpoS*-S RNA decreased. A similar result was observed if more than 200 nM Hfq_6 was employed in the *rpoS*-L RNA titration of RprA (data not shown). This suggests that excess Hfq sequesters the added *rpoS* RNA or binds to secondary RprA sites inhibiting formation of the RprA•*rpoS* RNA complex.

A number of coupled reactions govern the distribution of RprA in the presence of *rpoS* RNA and Hfq. We have examined three of them as independent reactions: Hfq binding to each RNA, and RprA binding to *rpoS* RNA. A fourth reaction to consider is Hfq binding to RprA•*rpoS* RNA to form a ternary complex. Previous work showed that Hfq can form a ternary complex with DsrA and *rpoS* RNA (Lease and Woodson 2004). Figure 2.6 illustrates the influence of Hfq on preformed RprA•*rpoS*-S RNA complex.

Adding up to 40 nM Hfq_6 reduced the amount of RprA•*rpoS*-S RNA complex by ~15%. A small amount of the Hfq_6•RprA band was detected (~5%), but no ternary complex was observed. Hfq may also be binding to *rpoS*-S RNA which is at a substantially higher concentration than RprA and is not visible in the gel. The ability of Hfq to dissociate a truncated

model of DsrA•*rpoS* RNA and shift the equilibrium to Hfq bound to the individual RNAs was previously observed (Arluison et al. 2007). When 80 nM or more Hfq_6 was added, a small amount of ternary Hfq•RprA•*rpoS*-S complex was observed and the Hfq_6•RprA complex (C_1) increased in intensity.

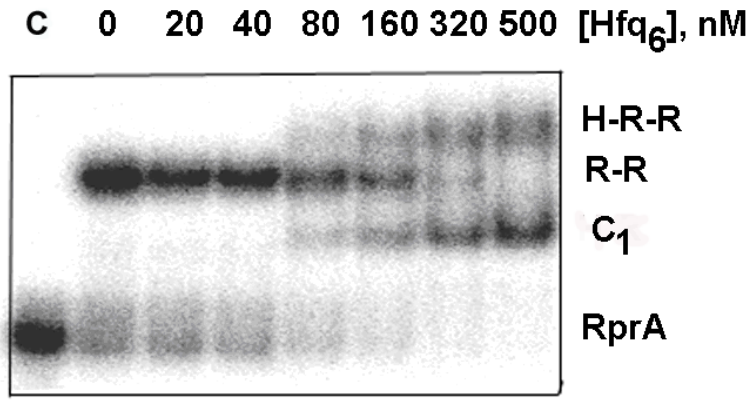


Figure 2.6: Effect of Hfq_6 on preformed *rpoS*-S-RprA complex. The first lane is control (C) with 4 nM ^{32}P-labeled RprA. Other lanes included 4 nM RprA preincubated with 200 nM *rpoS*-S to form RprA-*rpoS*-S complex (R-R). Hfq_6 (0 - 500 nM) was added and incubated for 10 minutes prior to loading in a 5% polyacrylamide gel. C_1 denotes a strong Hfq-RprA complex and H-R-R the Hfq-RprA-*rpoS*-S complex.

Although the affinity of Hfq for RprA• *rpoS* RNA could not be determined directly, the ternary complex appears to be less stable than Hfq_6•RprA or Hfq_6•*rpoS* RNA. Figure 2.6 shows that the ternary complex is only observed when the Hfq_6 concentration is high relative to RprA and *rpoS*-S RNA. The effect of Hfq_6 concentration on preformed RprA•*rpoS*-L complexes gave similar results (data not shown). It may be worth noting that Hfq_6 rapidly associates with RprA• *rpoS*-S RNA (within ~ 30 seconds), and that the ternary complex dissociates within 30 seconds to RprA• *rpoS*-S RNA and Hfq_6 when challenged by saturating amounts of poly(U) (data not shown). This rapid association and dissociation of Hfq with the RprA•*rpoS* RNA hybrid is similar to observations made with the DsrA• *rpoS* RNA complex (Lease and Woodson 2004).

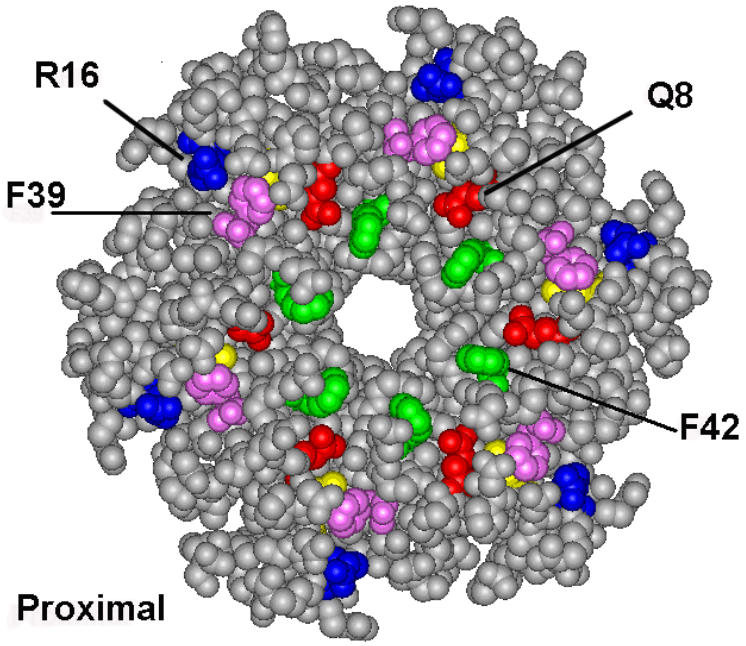


Figure 2.7: Space filling model of the proximal surface of the *E. coli* Hfq hexamer core showing the location of four residues mutated to alanine: F42A (green), Q8A (red), F39A (purple), and (R16A) blue.

Effect of Hfq mutations on RprA and *rpoS*-S binding

The effect of mutations to Hfq on its ability to bind RprA and *rpoS*-S was examined by the gel mobility shift assay. Five mutant proteins, designated Hfq-F42A, Hfq-F39A, Hfq-R16A, Hfq-Q8A, and Hfq-Y25A were employed. The first four proteins have mutations on the proximal face of Hfq while Hfq-Y25A has substitutions on the distal surface (Figure 2.7). Figure 2.8A shows representative gel-shift experiments of three mutant Hfq proteins with RprA, and Figure 2.8B displays the experimental data on the fraction of free RprA shifted (F_B) as a function of Hfq concentration. Each mutant Hfq protein formed two complexes with RprA but with varying affinities. Since the weak affinity of some of the mutant Hfq produced smeared bands making intensities difficult to evaluate, apparent K_d values were determined using the simplified binding

model described by eqs. (5) and (6) (Figure S2.3). Hfq-Q8A, Hfq-R16A, and Hfq-F42A gave apparent K_d values within 50% of that of wt Hfq. Mutants Hfq-Y25A and Hfq-F39A exhibited weaker affinities with apparent K_d values 3.5- and 5.0-fold higher that wt Hfq, respectively.

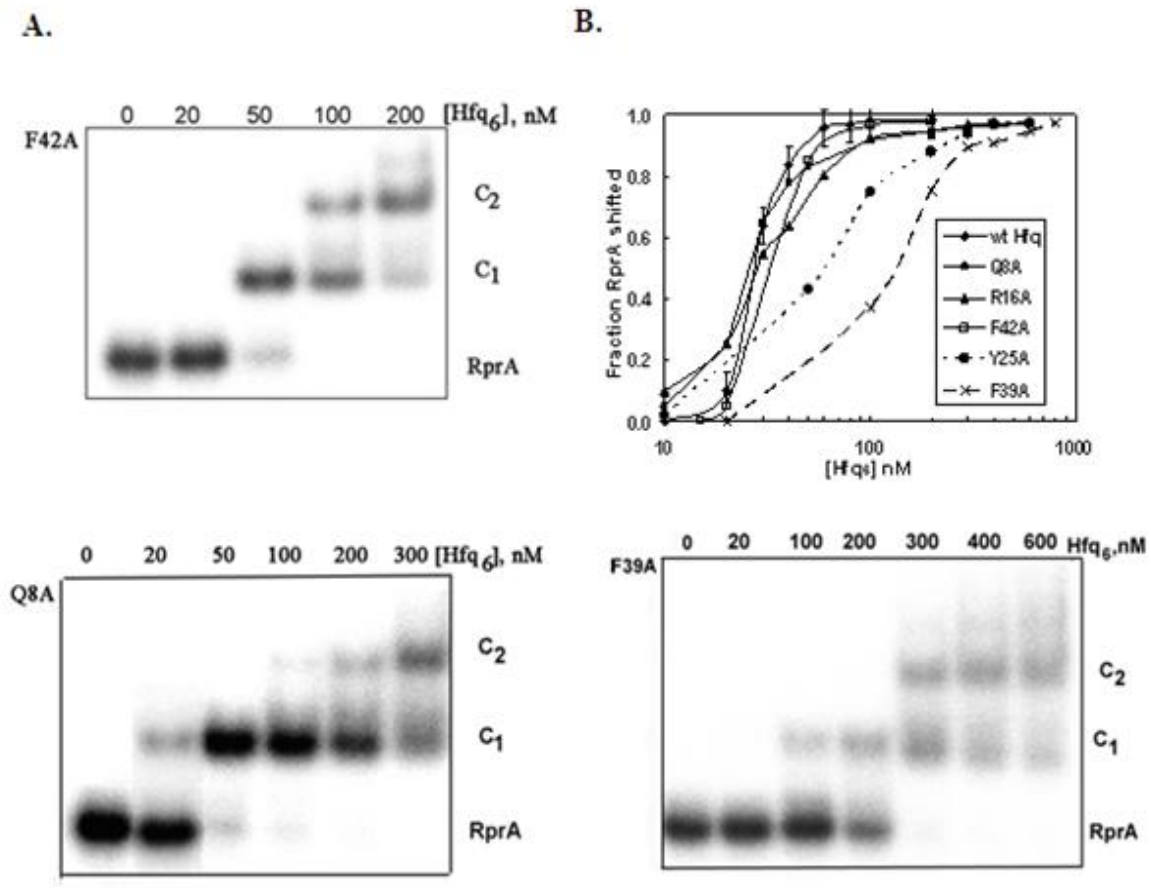


Figure 2.8: (**A**.) Binding of three Hfq mutants to RprA assessed by the gel shift assay. (**B**) Fraction of total RprA shifted from free to bound complexes as a function of Hfq_6 concentration for wt and five Hfq mutants. Error bars shown only for wt Hfq data for clarity.

Previous work showed that poly(A) and oligomers A_n (n = 18 or 27) bind to the distal Hfq surface (Mikulecky et al. 2004; Sun and Wartell 2006). Competition experiments were carried out in which poly(A) or A_{18} was added to the preformed C_1 complex of Hfq•RprA. Both A_{18} and poly(A) displaced Hfq from RprA (Figure 2.9A). This result contrasts with previous

work that showed that the addition of poly(A) or A_{27} to the Hfq•DsrA complex did not release DsrA and produced "supershifted" gel bands indicative of simultaneous binding of poly(A) and DsrA to Hfq (Brescia et al. 2003; Mikulecky et al. 2004). Competing poly(A) or A_{18} with Hfq•DsrA under our experimental conditions confirmed this behavior (Figure 2.9B). DsrA was not released in either case. The Hfq•DsrA band was shifted to lower mobility by poly(A), although not noticeably so by A_{18} which is shorter than the A_{27} used previously (Brescia et al. 2003).

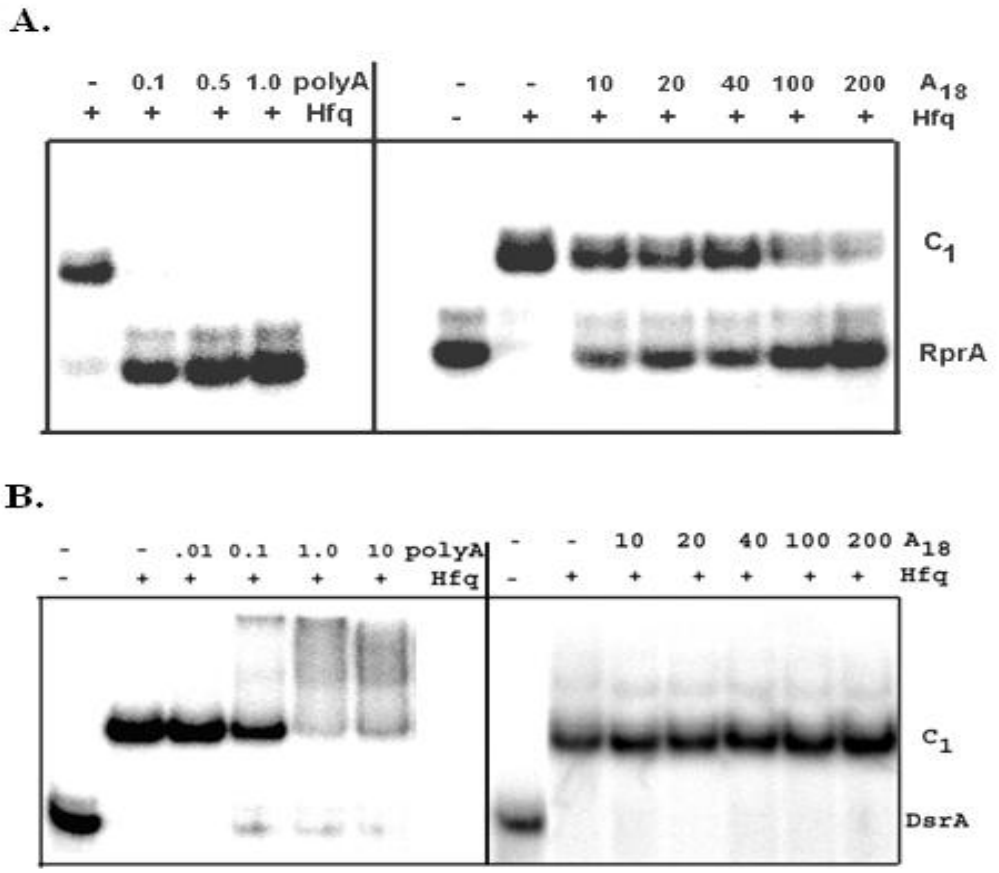


Figure 2.9: (**A.**) Effect of poly(A) and A_{18} on the Hfq-RprA complex. Hfq_6 (50 nM) was incubated with 4 nM ^{32}P-labeled RprA to form C_1 complex (lanes + for Hfq). Poly(A) or A_{18} was then added, and solutions were incubated for an additional 10 min prior to being run on 5% polyacrylamide gels. Four lanes on the left: C_1 complex with no poly(A) (-), and with 0.1, 0.5, and 1.0 ng/µl poly(A) (final concentrations). Seven lanes on the right: RprA alone, C_1 complex, and C_1 complex with A_{18} added to give a concentration of 10 - 200 nM. (**B.**) Same experiments as described in panel **A.** with DsrA replacing RprA.

Figure 2.10 shows three representative gel shift experiments of the mutant Hfq proteins binding to *rpoS*-S and plots of the experimental fraction of *rpoS*-S shifted as a function of Hfq concentration. Apparent dissociation constants of the mutant Hfq were evaluated using the simplified binding model described by eqs. (5) and (6) (Figure S2.4). Hfq-F42A, Hfq-Q8A and Hfq-Y25A exhibited relatively well defined shifted bands and their apparent K_d values were within two-fold of the wt Hfq value for *rpoS*-S. Gel shift experiments with the other Hfq mutants produced broader protein-RNA bands and weaker binding. This is illustrated in Figure 2.10A for Hfq-R16A. The F39A and R16A mutations reduced the apparent K_d by 3.2 fold compared to wt Hfq. We note that our results for Hfq-Q8A and Hfq-Y25A differ somewhat from a previous study of these mutant Hfqs that reported affinities for *rpoS* RNA slightly higher than wt Hfq (Mikulecky et al. 2004). These differences may be due to differences in reaction conditions or protein preparations and are in any case relatively small.

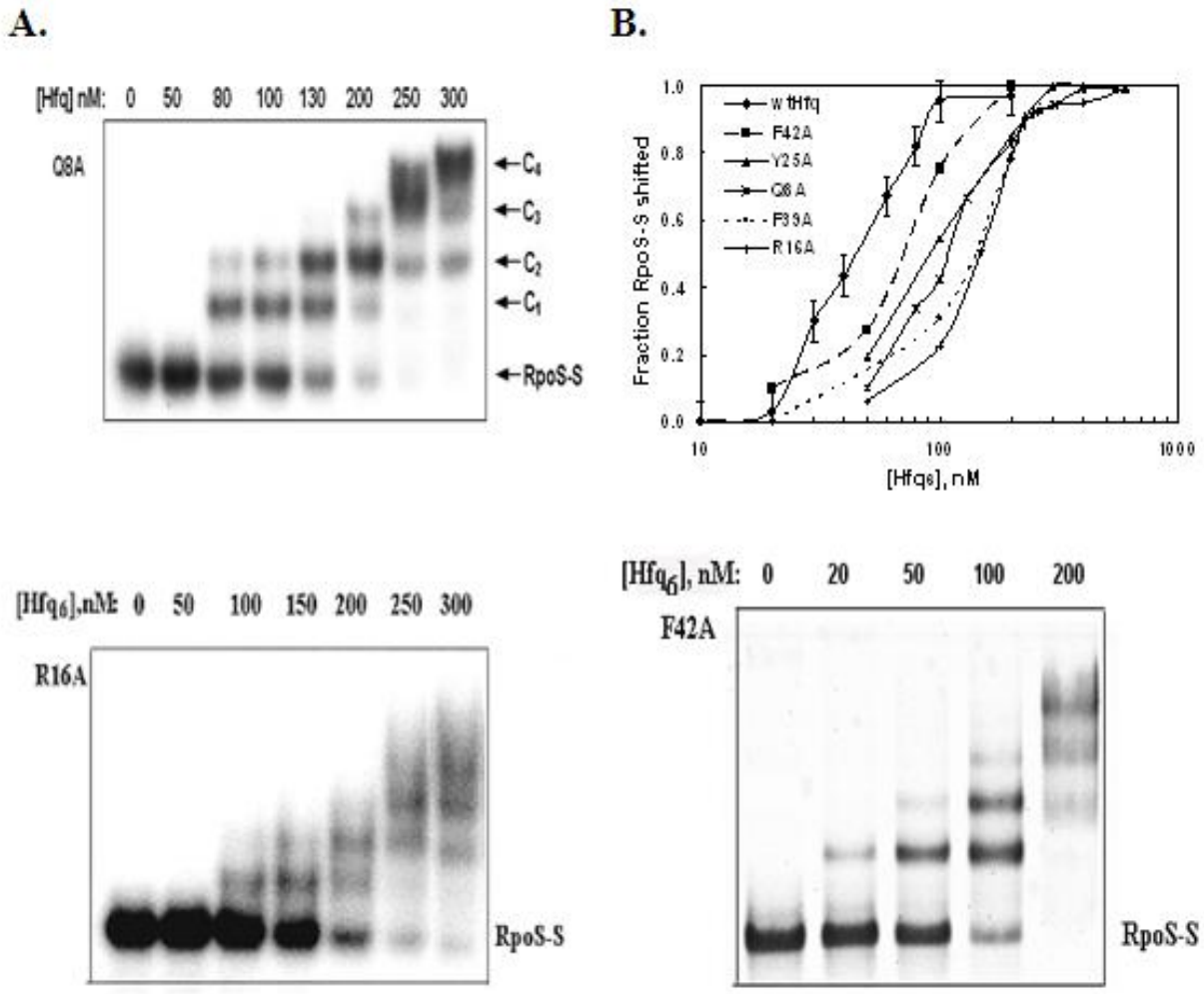


Figure 2.10: (**A.**) Binding of three mutant Hfq forms to *rpoS*-S assessed by a gel shift assay. (**B.**) Fraction of total *rpoS*-S shifted from free band to bound complexes as a function of Hfq_6 concentration for wt and five mutant Hfq's. Error bars shown only for wt Hfq data.

DISCUSSION

Genetic studies have shown that RprA and DsrA enhance translation of the *rpoS* mRNA in the presence of Hfq. Mutations in the sRNAs and compensating mutations in the *rpoS* mRNA provide compelling evidence that both sRNAs bind *rpoS* mRNA between nucleotides 452 and 473 paired to the RBS in Figure 2.2 (Majdalani et al. 1998; Majdalani et al. 2002). The sequence depicted in Figure 2.2 is undoubtedly important for *in vivo* regulation of *rpoS.* However, as previously noted (Majdalani et al. 2005), addition of Hfq to this *rpoS* RNA region and DsrA enhanced association of these RNAs by only 1.8 fold (Lease and Woodson 2004) compared to the 30 fold effect of Hfq on DsrA mediated regulation of *rpoS in vivo* (Sledjeski et al. 2001). This implies additional factors may be involved in the regulation of *rpoS* translation.

Our *in vitro* results point to an additional factor in the regulation of *rpoS* by Hfq and RprA. Hfq enhanced RprA binding to the 210 nt *rpoS*-S RNA by about 1.5 fold, similar to the enhancement observed for DsrA and a 140 nt *rpoS* RNA (Lease and Woodson 2004). However, when the entire *rpoS* leader region (*rpoS*-L) was examined, Hfq enhanced RprA binding by ~30-fold. This strong influence of the 5' end of the *rpoS* transcript is qualitatively consistent with *in vivo* results of Cunning *et al.* (Cunning et al. 1998). These authors showed that when the 5' end of the *rpoS* transcript was deleted such that it still retained 220 nt upstream of the start codon, which includes the RprA/DsrA binding region, it was not regulated by Hfq. They proposed that a site near the 5' end of the *rpoS* transcript may be involved in translational regulation.

The difference in affinity of RprA for *rpoS*-S vs. *rpoS*-L RNA in the absence of Hfq indicates that the initial ~ 430 nt of the *rpoS* transcript inhibits RprA binding over and above the inhibitory secondary structure around the RBS. Since the RNase H degradation assay did not

indicate a major difference in secondary structure in the vicinity of the RBS for *rpoS*-S vs. *rpoS*-L RNAs (Figure 2.3), the 5'end of this transcript does not appear to propagate a major rearrangement of secondary structure in the RBS region. We hypothesize that intramolecular tertiary interaction involving a site near the 5' end of the *rpoS* transcript and a site near the RBS inhibits RprA pairing and Hfq overcomes this inhibition. The effect of Hfq on DsrA binding to *rpoS*-S and *rpoS*-L RNAs give results similar to that observed with RprA (Figure S2.5). DsrA binds *rpoS*-S RNA with a K_d of ~100 nM and to *rpoS*-L RNA with a K_d of ~ 750 nM. 100 nM Hfq_6 enhanced the apparent K_d of DsrA binding to *rpoS*-L RNA to ~ 25 nM, a 30 fold increase in affinity. 20 nM Hfq_6 decreased the apparent K_d of DsrA for *rpoS*-S RNA by only about 1.5-fold.

The 30-fold enhancement of RprA•*rpoS*-L RNA binding induced by Hfq *in vitro* is considerably larger than the 6.7 fold enhancement by Hfq determined *in vivo* for RprA mediated translation of a *rpoS-lac* fusion gene (Majdalani et al. 2002). However, the temperature of our experiments, 25 ^{o}C, differs from the conditions of the *in vivo* study (37 ^{o}C). When we reexamined RprA binding to *rpoS*-L RNA at 37 ^{o}C, the K_d was ~300 nM and the addition of Hfq enhanced RprA•*rpoS*-L formation by 9 fold (data no shown). This improved agreement with the *in vivo* data lends supports to the relevance of the binding studies on the full leader region in understanding the role of Hfq in RpoS regulation *in vivo*. We note that Hfq enhancement in the stability of the RprA•*rpoS* RNA complexes reflects enhanced association rates (Figure S2.2). Both kinetic and thermodynamic perspectives may be needed to explain translation regulation of *rpoS* mRNA *in vivo*.

A comparison of our results for the interaction of Hfq with RprA with work on Hfq and DsrA reveals several common characteristics as well as differences. The similarities and differences may help identify what is important and what is not with regard to the interaction of

Hfq with these sRNAs and their common *rpoS* target. For example, gel shift characteristics of Hfq binding to RprA and DsrA are quite similar in spite of the difference in RNA sequences and predicted secondary structures. Two gel complexes are observed for both sRNAs, the Hill cooperativity parameter of Hfq binding to RprA and DsrA was similar, and the saturation point of Hfq to sRNA at high concentrations imply 2:1 stoichiometry for both strong Hfq•sRNA complexes (Figure S2.1, (Lease and Woodson 2004)).

K_d values evaluated for the Hfq-RprA complexes, 24 nM and 96 nM, are the same as the values determined by Mikulecky *et al.* (Mikulecky et al. 2004) for the Hfq-DsrA complexes observed in a gel shift assay. The affinity determined by Lease and Woodson (Lease and Woodson 2004) for the strong Hfq-DsrA complex (37 nM) is comparable to the corresponding Hfq-RrpA complex; however, the K_d for the second Hfq-DsrA complex (667nM) indicates a significantly lower affinity. The presence of non-specific tRNA in the reactions of the latter study may account for the higher K_d value.

We note that Hfq binding to DsrA by isothermal titration calorimetry indicates a 1:1 Hfq_6•DsrA stoichiometry (Mikulecky et al. 2004). This difference from the gel results may reflect a number of factors including technical complications inherent in either approach. The change in solvent conditions that occurs during the onset of electrophoresis before macromolecules enter the gel may alter the distribution of Hfq bound to the RNA. Higher concentrations needed for calorimetry may promote DsrA dimers. While determining the stoichiometry under different conditions is important, the experiments indicate that under similar conditions Hfq forms complexes of similar stability and characteristics with these two sRNAs.

The above comparison emphasizes similarities in Hfq binding to DsrA and RprA. We note two significant differences. First, the two sRNAs have different affinities for the full length

rpoS RNA leader region. The affinity of RprA for *rpoS*-L RNA (~2500 nM) is weaker than DsrA binding to this *rpoS* RNA (~750 nM). A possible explanation is the number of complimentary base pairs between RprA and DsrA for the *rpoS* target region. Optimum alignment of RprA with the RBS region shows 11 of 13 consecutive nucleotides compared to 21 out of 23 nucleotides for DsrA (Majdalani et al. 2002).

A second distinction between DsrA and RprA is how their binding to Hfq is influenced by amino acid mutations and the competition of $(A)_n$ sequences. Mutations to three residues on the distal surface of Hfq (Y25, I30, and K31) had little or no effect on the strong binding complex of Hfq with DsrA or to domain II of DsrA, but they drastically lower binding to poly(A) sequences (Mikulecky et al. 2004; Sun and Wartell 2006) . This together with the observation that poly(A) does not displace DsrA from Hfq (Brescia et al. 2003) implies that Hfq binds DsrA primarily with its proximal surface. In contrast, poly(A) or A_{18} displaced RprA from Hfq (Figure 2.9A). A mutation on the distal as well as proximal surface weakened binding of Hfq to RprA. We also note that approximately equimolar addition of DsrA displaces RprA from Hfq (data not shown). Reconciling these observations will require additional structural information on the two complexes.

Acknowledgment

I am grateful to Xueguang Sun, Adam Leon, Lively Lie, Sean Murray, and Sarah Paglioni for assisting in plasmid construction and wild type and mutant Hfq protein purification. I acknowledge Nabil Wilf for his contribution towards obtaining data from the RNase H assay and the binding studies of RprA and *rpoS*-S to wild type and mutant Hfq's.

REFERENCES

Afonyushkin, T., Vecerek, B., Moll, I., Blasi, U., and Kaberdin, V.R. 2005. Both RNase E and RNase III control the stability of sodB mRNA upon translational inhibition by the small regulatory RNA RyhB. *Nucleic Acids Res* **33**(5): 1678-1689.

Altuvia, S., Weinstein-Fischer, D., Zhang, A., Postow, L., and Storz, G. 1997. A small, stable RNA induced by oxidative stress: role as a pleiotropic regulator and antimutator. *Cell* **90**(1): 43-53.

Antal, M., Bordeau, V., Douchin, V., and Felden, B. 2005. A small bacterial RNA regulates a putative ABC transporter. *J Biol Chem* **280**(9): 7901-7908.

Arluison, V., Derreumaux, P., Allemand, F., Folichon, M., Hajnsdorf, E., and Regnier, P. 2002. Structural Modeling of the Sm-like Protein Hfq from Escherichia coli. *J Mol Biol* **320**(4): 705-712.

Arluison, V., Hohng, S., Roy, R., Pellegrini, O., Regnier, P., and Ha, T. 2007. Spectroscopic observation of RNA chaperone activities of Hfq in post-transcriptional regulation by a small non-coding RNA. *Nucleic Acids Res* **35**(3): 999-1006.

Bossi, L. and Figueroa-Bossi, N. 2007. A small RNA downregulates LamB maltoporin in Salmonella. *Mol Microbiol* **65**(3): 799-810.

Brescia, C.C., Mikulecky, P.J., Feig, A.L., and Sledjeski, D.D. 2003. Identification of the Hfq-binding site on DsrA RNA: Hfq binds without altering DsrA secondary structure. *Rna* **9**(1): 33-43.

Brown, L. and Elliott, T. 1996. Efficient translation of the *RpoS* sigma factor in Salmonella typhimurium requires host factor I, an RNA-binding protein encoded by the hfq gene. *J Bacteriol* **178**(13): 3763-3770.

Cunning, C., Brown, L., and Elliott, T. 1998. Promoter substitution and deletion analysis of upstream region required for *rpoS* translational regulation. *J Bacteriol* **180**(17): 4564-4570.

Franze de Fernandez, M.T., Eoyang, L., and August, J.T. 1968. Factor fraction required for the synthesis of bacteriophage Qbeta-RNA. *Nature* **219**(154): 588-590.

Fried, M.G. and Crothers, D.M. 1984. Equilibrium studies of the cyclic AMP receptor protein-DNA interaction. *J Mol Biol* **172**(3): 241-262.

Geissmann, T.A. and Touati, D. 2004. Hfq, a new chaperoning role: binding to messenger RNA determines access for small RNA regulator. *Embo J* **23**(2): 396-405.

Gottesman, S. 2004. The small RNA regulators of Escherichia coli: roles and mechanisms. *Annu Rev Microbiol* **58**: 303-328.

Gottesman, S., Storz, G., Rosenow, C., Majdalani, N., Repoila, F., and Wassarman, K.M. 2001. Small RNA regulators of translation: mechanisms of action and approaches for identifying new small RNAs. *Cold Spring Harb Symp Quant Biol* **66**: 353-362.

Kawamoto, H., Koide, Y., Morita, T., and Aiba, H. 2006. Base-pairing requirement for RNA silencing by a bacterial small RNA and acceleration of duplex formation by Hfq. *Mol Microbiol* **61**(4): 1013-1022.

Lange, R., Fischer, D., and Hengge-Aronis, R. 1995. Identification of transcriptional start sites and the role of ppGpp in the expression of *rpoS*, the structural gene for the sigma S subunit of RNA polymerase in Escherichia coli. *J Bacteriol* **177**(16): 4676-4680.

Lease, R.A. and Woodson, S.A. 2004. Cycling of the Sm-like protein Hfq on the DsrA small regulatory RNA. *J Mol Biol* **344**(5): 1211-1223.

Lenz, D.H., Mok, K.C., Lilley, B.N., Kulkarni, R.V., Wingreen, N.S., and Bassler, B.L. 2004. The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in Vibrio harveyi and Vibrio cholerae. *Cell* **118**(1): 69-82.

Majdalani, N., Cunning, C., Sledjeski, D., Elliott, T., and Gottesman, S. 1998. DsrA RNA regulates translation of *RpoS* message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. *Proc Natl Acad Sci U S A* **95**(21): 12462-12467.

Majdalani, N., Hernandez, D., and Gottesman, S. 2002. Regulation and mode of action of the second small RNA activator of *RpoS* translation, RprA. *Mol Microbiol* **46**(3): 813-826.

Majdalani, N., Vanderpool, C.K., and Gottesman, S. 2005. Bacterial small RNA regulators. *Crit Rev Biochem Mol Biol* **40**(2): 93-113.

Mathews, D.H., Disney, M.D., Childs, J.L., Schroeder, S.J., Zuker, M., and Turner, D.H. 2004. Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc Natl Acad Sci U S A* **101**(19): 7287-7292.

Mikulecky, P.J., Kaw, M.K., Brescia, C.C., Takach, J.J., Sledjeski, D., and Feig, A.L. 2004. *Escherichia coli* Hfq has distinct interaction surfaces for DsrA, *rpoS* and poly(A) RNAs. *Nat Struct Mol Biol* **11**(12): 1206-1214.

Mohanty, B.K., Maples, V.F., and Kushner, S.R. 2004. The Sm-like protein Hfq regulates polyadenylation dependent mRNA decay in Escherichia coli. *Mol Microbiol* **54**(4): 905-920.

Moll, I., Leitsch, D., Steinhauser, T., and Blasi, U. 2003. RNA chaperone activity of the Sm-like Hfq protein. *EMBO Rep* **4**(3): 284-289.

Moller, T., Franch, T., Hojrup, P., Keene, D.R., Bachinger, H.P., Brennan, R.G., and Valentin-Hansen, P. 2002a. Hfq: a bacterial Sm-like protein that mediates RNA-RNA interaction. *Mol Cell* **9**(1): 23-30.

Moller, T., Franch, T., Udesen, C., Gerdes, K., and Valentin-Hansen, P. 2002b. Spot 42 RNA mediates discoordinate expression of the E. coli galactose operon. *Genes Dev* **16**(13): 1696-1706.

Nikulin, A., Stolboushkina, E., Perederina, A., Vassilieva, I., Blaesi, U., Moll, I., Kachalova, G., Yokoyama, S., Vassylyev, D., Garber, M. et al. 2005. Structure of Pseudomonas aeruginosa Hfq protein. *Acta Crystallogr D Biol Crystallogr* **61**(Pt 2): 141-146.

Sauter, C., Basquin, J., and Suck, D. 2003. Sm-like proteins in Eubacteria: the crystal structure of the Hfq protein from Escherichia coli. *Nucleic Acids Res* **31**(14): 4091-4098.

Schumacher, M.A., Pearson, R.F., Moller, T., Valentin-Hansen, P., and Brennan, R.G. 2002. Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: a bacterial Sm-like protein. *Embo J* **21**(13): 3546-3556.

Sledjeski, D.D., Gupta, A., and Gottesman, S. 1996. The small RNA, DsrA, is essential for the low temperature expression of *RpoS* during exponential growth in Escherichia coli. *Embo J* **15**(15): 3993-4000.

Sledjeski, D.D., Whitman, C., and Zhang, A. 2001. Hfq is necessary for regulation by the untranslated RNA DsrA. *J Bacteriol* **183**(6): 1997-2005.

Storz, G., Opdyke, J.A., and Zhang, A. 2004. Controlling mRNA stability and translation with small, noncoding RNAs. *Curr Opin Microbiol* **7**(2): 140-144.

Sun, X. and Wartell, R.M. 2006. Escherichia coli Hfq binds A18 and DsrA domain II with similar 2:1 Hfq6/RNA stoichiometry using different surface sites. *Biochemistry* **45**(15): 4875-4887.

Sun, X., Zhulin, I., and Wartell, R.M. 2002. Predicted structure and phyletic distribution of the RNA-binding protein Hfq. *Nucleic Acids Res* **30**(17): 3662-3671.

Tsui, H.C., Leung, H.C., and Winkler, M.E. 1994. Characterization of broadly pleiotropic phenotypes caused by an hfq insertion mutation in Escherichia coli K-12. *Mol Microbiol* **13**(1): 35-49.

Udekwu, K.I., Darfeuille, F., Vogel, J., Reimegard, J., Holmqvist, E., and Wagner, E.G. 2005. Hfq-dependent regulation of OmpA synthesis is mediated by an antisense RNA. *Genes Dev* **19**(19): 2355-2366.

Vecerek, B., Rajkowitsch, L., Sonnleitner, E., Schroeder, R., and Blasi, U. 2008. The C-terminal domain of Escherichia coli Hfq is required for regulation. *Nucleic Acids Res* **36**(1): 133-143.

Wassarman, K.M., Repoila, F., Rosenow, C., Storz, G., and Gottesman, S. 2001. Identification of novel small RNAs using comparative genomics and microarrays. *Genes Dev* **15**(13): 1637-1651.

Zhang, A., Altuvia, S., Tiwari, A., Argaman, L., Hengge-Aronis, R., and Storz, G. 1998. The OxyS regulatory RNA represses *rpoS* translation and binds the Hfq (HF-I) protein. *Embo J* **17**(20): 6061-6068.

Zhang, A., Wassarman, K.M., Ortega, J., Steven, A.C., and Storz, G. 2002. The Sm-like Hfq protein increases OxyS RNA interaction with target mRNAs. *Mol Cell* **9**(1): 11-22.

SUPPLEMENTARY MATERIAL

Binding of Hfq to RprA at 100 nM

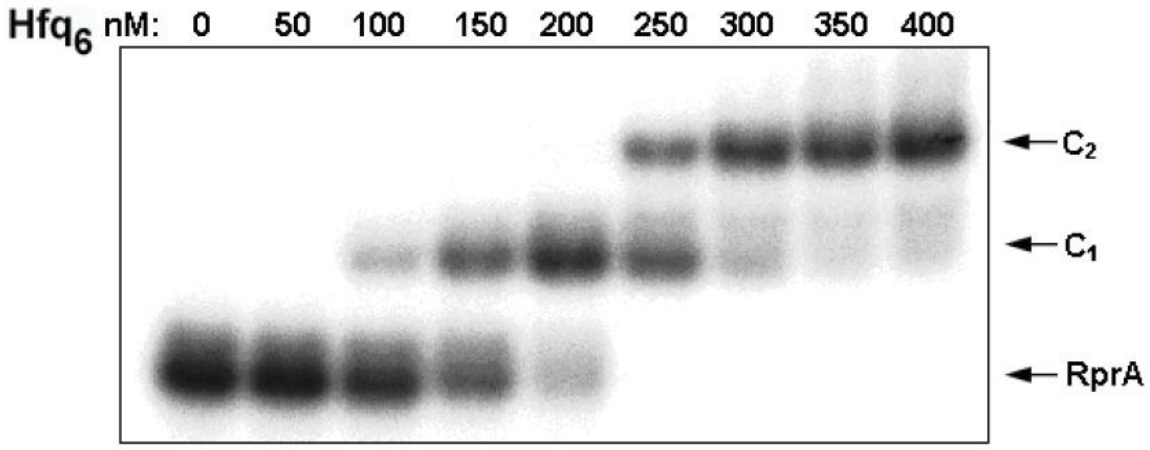


Figure S2.1: Hfq binding to RprA by gel mobility shift assay. Reaction carried out as described in Materials and Methods with ^{32}P-labeled RprA at concentration of 100 nM. Concentration of Hfq is in moles hexamer/liter. C1 complex reaches maximum at a 2:1 ratio of Hfq_6 to RprA.

Effect of Hfq on the rates of formation of RprA•*rpoS*-S and RprA•*rpoS*-L

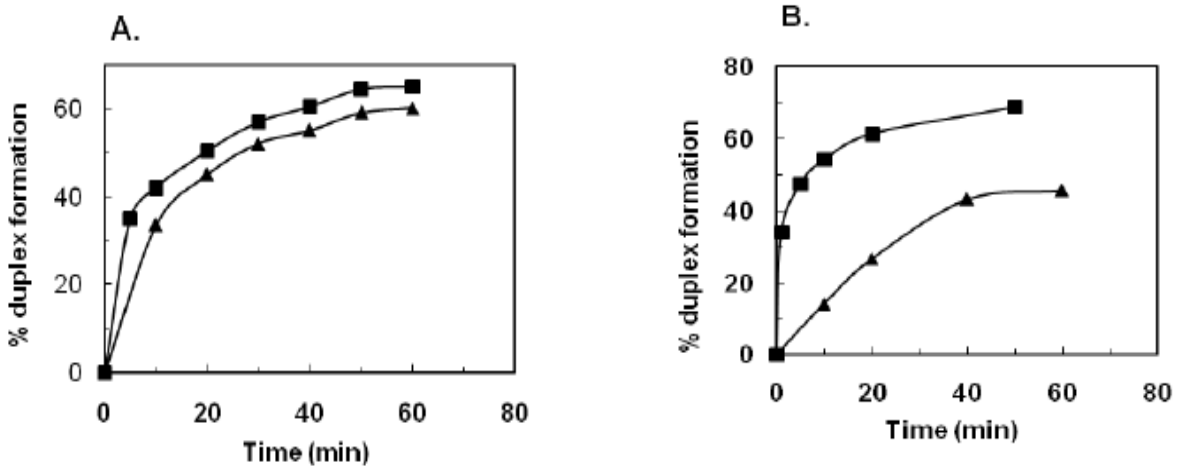


Figure S2.2: (A.) Rate of formation of RprA•*rpoS*-S RNA complex in the absence of Hfq (▲) and in the presence of 40 nM Hfq_6 (■) with 4 nM RprA and 200 nM *rpoS*-S RNA using conditions in Figures 1.5A and 1.5C. (**B.**) Rate of formation or RprA•*rpoS*-L RNA complex in the absence of Hfq (▲) and in the presence of 200 nM Hfq_6 (■) with 4 nM RprA and 3000 nM *rpoS*-L RNA.

Binding analysis of mutant Hfq proteins to RprA

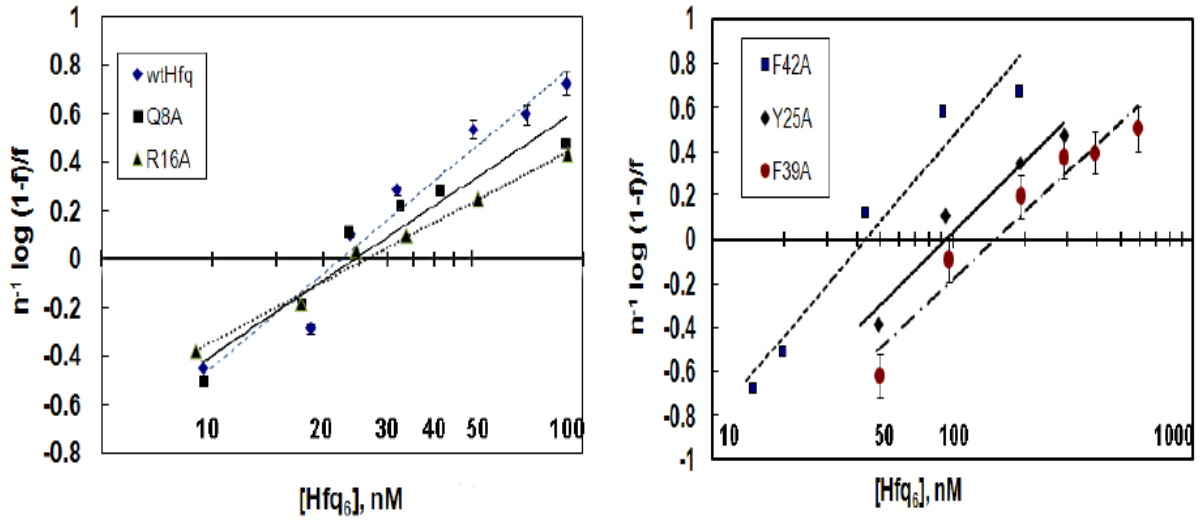


Figure S2.3: Graphical analysis of mutant Hfq proteins binding to RprA. Data from Figure 2.8B are plotted according to eqs. (5) and (6) with n= 2.5. Lines are least-squares fit to data ($R^2 \geq$ 0.95). Two panels are used in order to better separate and display plots.

Binding analysis of mutant Hfq proteins to *rpoS*-S

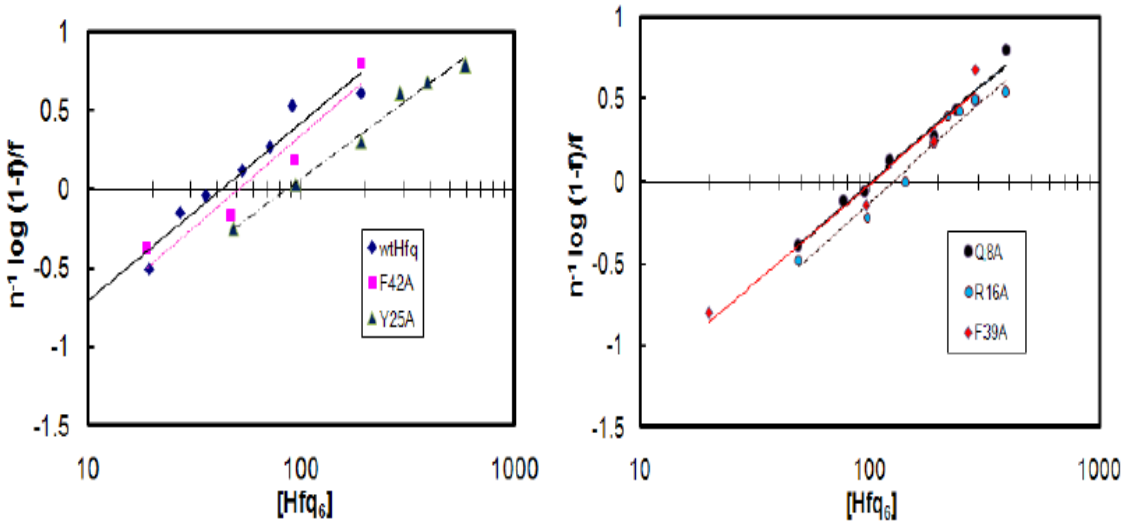


Figure S2.4: Graphical analysis of mutant Hfq proteins binding to *rpoS*-S RNA. Data from Figure 2.10B are plotted according to eqs. (5) and (6) with n= 2.5. Least-squares fit of lines to the data are shown ($R^2 \geq 0.95$).

Binding of DsrA to *RpoS*-S and *RpoS*-L in the absence and presence of Hfq

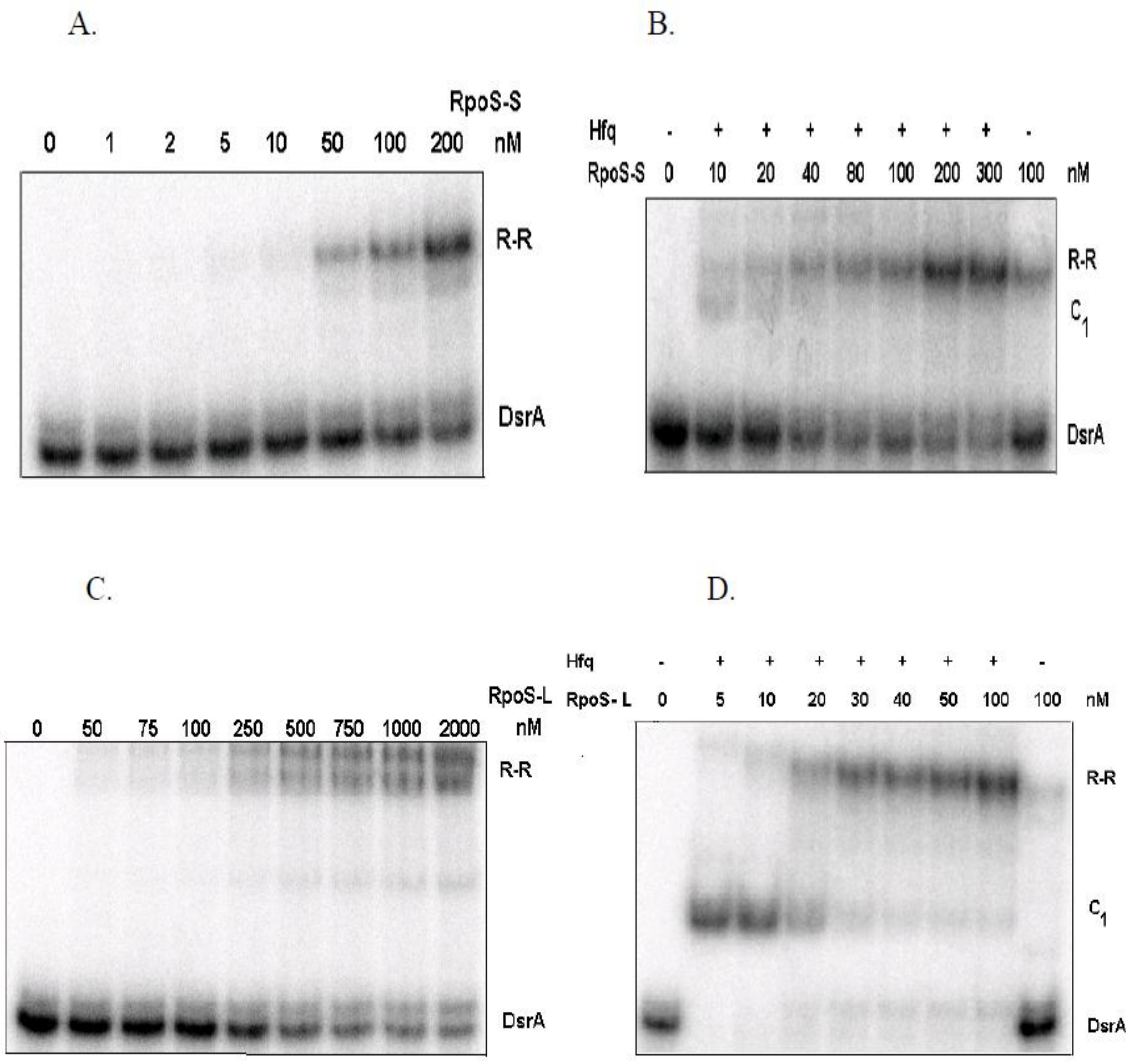


Figure S2.5: Panel **A.** shows binding of *rpoS*-S to 4 nM ^{32}P-DsrA in absence of Hfq, Kd ~ 100 nM based on three experiments. Panel **B.** shows binding of *rpoS*-S to DsrA in absence (-) and presence (+) of 20 nM Hfq_6. This Hfq_6 concentration reduced the apparent K_d to ~70 nM. Panels **C.** and **D.** show similar experiments in which *rpoS*-L was added to ^{32}P-DsrA in the absence and presence (+) of Hfq_6. 100 nM Hfq_6 was employed in + lanes of panel **D.**. Apparent K_d in absence of Hfq was ~750 nM, and ~20 nM in presence of Hfq. The Hfq concentrations used provided the largest reduction in apparent K_d among several tested.

CHAPTER 3

The influence of *Escherichia coli* mutations on RNA binding and mRNA•sRNA duplex formation in *rpoS* riboregulation

INTRODUCTION

Hfq, also known as Host Factor I, is a heat-stable RNA binding protein that has been shown to be an important regulator of gene expression in many bacterial species. Null mutants of the *hfq* gene typically display pleiotropic phenotypes such as decreased growth rate, altered patterns of protein synthesis, and increased sensitivity to environmental stress (Tsui et al. 1994). The broad impact of Hfq appears to stem from its role in regulating the expression of mRNAs in concert with small non-coding RNAs (sRNA) (Valentin-Hansen et al. 2004; Aiba 2007; Waters and Storz 2009). Hfq has been shown to stimulate the pairing of specific sRNAs to their targeted mRNAs which in turn repress or enhance translation of the mRNAs (Moller et al. 2002a; Zhang et al. 2002; Geissmann and Touati 2004; Majdalani et al. 2005).

A number of the pleiotropic phenotypes observed for *hfq* null mutants in *E. coli* can be attributed to a defect in the regulation of the *rpoS* gene that encodes the sigma factor involved in environmental stress adaptation and stationary phase growth (Hengge-Aronis 2002; Repoila et al. 2003). Hfq regulates expression of *rpoS* mRNA by facilitating the interaction of several sRNAs with this mRNA. RprA and DsrA enhance translation of *rpoS* mRNA by hybridizing to a segment within the leader region and disrupting a base paired structure that inhibits ribosome binding. OxyS inhibits translation, possibly by interacting with a sequence within the *rpoS* mRNA (Zhang et al. 1998; Zhang et al. 2002). While Hfq plays other roles in RNA metabolism, such as modulating mRNA degradation through its influence on 3' polyadenylation of mRNAs

(Hajnsdorf and Regnier 2000; Le Derout et al. 2003), the focus of the current work is on Hfq's role in facilitating sRNA• mRNA interaction, in particular its interactions with DsrA, RprA, OxyS, and the *rpoS* mRNA leader region.

Crystal structures of Hfq from *E. coli*, *Staphylococcus aureus*, and *Pseudomonas aeruginosa* show that the core of the protein forms a toroid comprised of six identical subunits. The monomer subunit adopts a common OB-like fold consisting of an N-terminal alpha helix followed by five bent beta strands running anti-parallel followed by a variable length C-terminal segment of unknown structure (Schumacher et al. 2002; Sauter et al. 2003; Nikulin et al. 2005). Studies employing mass spectroscopy, electron microscopy, and analytical ultracentrifugation indicate *E coli* Hfq also forms a hexamer structure in solution (Moller et al. 2002a; Zhang et al. 2002; Updegrove et al. 2011). The core toroidal structure of Hfq displays three surface regions which will be referred to as the proximal, distal, and outer-circumference or side-view (Figure 3.1).

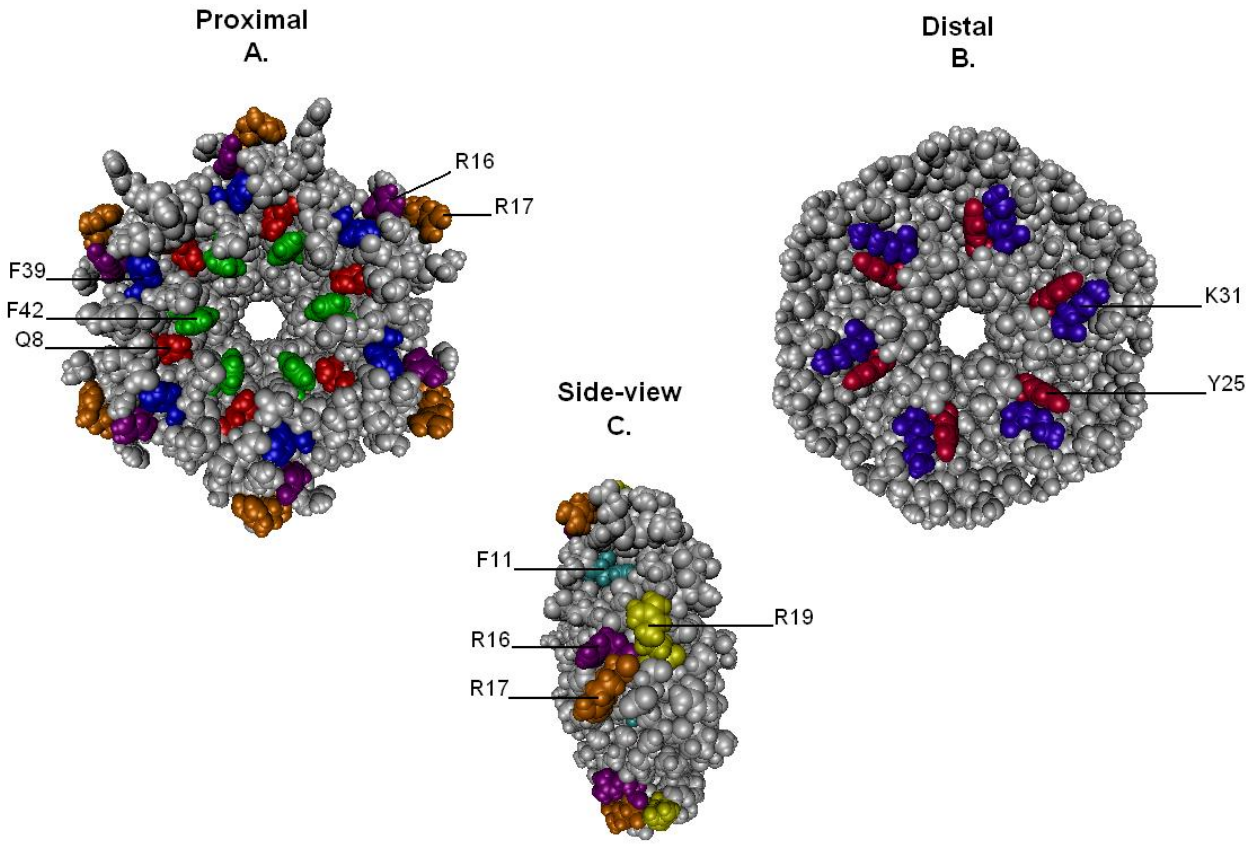


Figure 3.1: Space filling model of the toroidal part of *E. coli* Hfq hexamer showing the locations of the nine single amino acid mutations viewed from: proximal side (**A.**); distal side (**B.**); and side view or outer circumference (**C.**).

Solution studies have shown that wt Hfq binds to DsrA and RprA with an apparent equilibrium dissociation constant, K_d, in the range of 20 to 40 nM (Lease and Woodson 2004; Mikulecky et al. 2004; Updegrove et al. 2008). Utilizing mutant Hfqs with residue changes on the proximal and distal surfaces, binding studies have indicated that Hfq interacts with DsrA primarily on its proximal surface (Mikulecky et al. 2004; Sun and Wartell 2006), while Hfq interaction with RprA involves both proximal and distal surfaces (Updegrove et al. 2008). Mutations along the outer circumference of the Hfq toroid have not been tested against these or other sRNAs. Binding studies of Hfq to portions of OxyS suggest that Hfq requires a single

stranded segment and an adjacent stem-loop for strong binding (Zhang et al. 2002), an observation similar to that of other RNAs that bind Hfq (Moller et al. 2002b; Brescia et al. 2003). The effect of Hfq mutations on its interaction with OxyS has yet to be characterized.

The untranslated leader region of *rpoS* mRNA also binds Hfq with high affinity (Mikulecky et al. 2004). One Hfq binding site is located approximately 65 to 80 nt upstream of the *rpoS* mRNA start codon (Lease and Woodson 2004), and another is ~170 nt upstream of the start codon (Soper and Woodson 2008). The latter site contains a sequence motif $(ARN)_4$ that can bind to Hfq's distal surface (Link et al. 2009). Mutations that alter this repeat sequence reduce Hfq binding affinity and reduce DsrA and RprA mediated RpoS expression *in vivo* (Soper et al. 2010).

The current study examined the binding affinity of wild-type and eleven mutant Hfq proteins to DsrA, RprA, and OxyS, and several fragments of the *rpoS* mRNA. We also evaluated the influence of these mutations on Hfq's ability to facilitate DsrA•*rpoS* mRNA hybridization. Several residues on the proximal surface affected Hfq binding to the three sRNAs although in different ways. The outer-circumference mutations and C-terminal region do not appear to play a significant role in binding the above sRNAs or in stimulating DsrA•*rpoS* mRNA hybridization. Hfq with mutations on the proximal or distal surface that reduced binding to the *rpoS* RNA segment containing the RBS site or the segment containing the $(ARN)_4$ sequence were defective in DsrA•*rpoS* RNA hybridization. Thus, binding of Hfq(s) to both *rpoS* RNA sites appears to be required to maximize DsrA•*rpoS* RNA formation. Finally, Hfq did not significantly enhance OxyS binding to *rpoS* mRNA. This result does not support the idea that Hfq stimulated annealing of OxyS to *rpoS* RNA is the major mechanism for OxyS repression of *rpoS* translation.

MATERIALS AND METHODS

Purification and characterization of wild-type and mutant Hfq

The Impact-CN intein system (New England Biolabs) was used to purify Hfq proteins as previously described (Sun and Wartell 2006). The plasmids used to over express the Hfq proteins contained the *E. coli hfq* gene inserted into SapI-SmaI digested pTYB11 plasmid (pEcHfq) or mutant derivatives (see below). Protein purification was carried out according to the recommendation of the manufacturer using strain ER2566. Cell lysis was carried out using a french press. The cell lysate was centrifuged and the supernatant loaded onto a chitin column. The column was extensively washed with the lysis/wash buffer of 20 mM Tris (pH 8.3) and 1 M NaCl prior to incubation of the column with this buffer plus 40 mM dithiothreitol. The eluted protein was concentrated and buffer-exchanged to 0.5 M NaCl and 20 mM Tris at pH 8.3 using centrifugation filtration units.

To enhance removal of contaminating nucleic acids, Hfq preparations were subjected to a micrococcal nuclease treatment. 25 μl of 300 units/ml Micrococcal nuclease (Worthington Biochemical Corporation) was added to 1 ml of 0.3 to 0.4 OD_{274nm} Hfq in 0.2 M NaCl, 20 mM Tris (pH 8.5) and 5 mM $CaCl_2$ and incubated at 37^{o}C for 45 min. 10 μL of 0.5 M Na_2EDTA was added and sample was washed and concentrated in 15 ml of 0.5 M NaCl and 20 mM Tris at pH 8.3 using 30 kD MWCO Amicon centrifugal filter. The truncated Hfq-65 and Hfq-75 proteins were concentrated using an Amicon Ultrafiltration cell with a 2000 MWCO filter.

Plasmids containing mutant *hfq* genes were generated from pTYB11-wt Hfq using the QuikChange Mutagenesis Kit from Stratagene Inc. In addition to the previously described single amino acid change mutations F42A, F39A, Q8A, R16A, K31A, Y25A, and double amino acid

change mutations L12F/F42A and F39A/F42A (Sun and Wartell 2006), *hfq* genes with single residue mutations R19A, R17A, and F11A were constructed and their proteins expressed. Two additional mutant *hfq* genes were constructed by creating stop codons at residue positions 76 and 66 respectively. These plasmids yielded truncated Hfq designated Hfq-65 and Hfq-75. The wt Hfq and mutant Hfq's were characterized for purity by SDS-PAGE and UV spectroscopy.

Plasmid construction and transcription of *rpoS*, RprA, DsrA and OxyS

The *rpoS* size variants *rpoS*-S and *rpoS*-L, RprA and DsrA was constructed, transcribed and purified as previously described (Updegrove et al. 2008). *rpoS*$323_{254\text{-}457}$ RNA was constructed using PCR with the primers 5'-GTAGTAATACGACTCACTATA GGCCGCGTTGTTTATGCTG -3' and 5'-TAACGAATTTCAAAATGCAA GCGTGTTGAACTGG -3', and the plasmid bearing *rpoS*-L as the template. PCR amplicons of *rpoS*$323_{254\text{-}457}$ were purified and used directly as templates for transcription reactions. *rpoS*-F RNA contains the entire 565 nt untranslated leader region of *rpoS*, 200 nt beyond the AUG site, and 24 extra nucleotides at the 3' terminus. This RNA was constructed using PCR with the primers 5'-ATGGAATTCTAATACGACTCACTATAGGGCGGGTGAAC AGAGTGCTAACAAAATGTTGCCG -3', 5'-TATATGGATCCTTACTACTTAT CGTCGTCATCCTTGTAATCATAACCAATCTCACCAAGGTAAAGC -3', and *E. coli* st. K12 genomic DNA as template. *rpoS*-F PCR amplicons were cloned into pUC19 plasmid at the EcoRI and BamHI site. The constructed plasmid pUC19-*rpoS*-F was linearized at the BamHI site and used as template for transcription reactions of *rpoS*-F RNA. The OxyS DNA sequence was cloned using procedures similar to those used for RprA and DsrA. Primers used for OxyS amplification were 5'-GTAGAATTCTAATACGACTCACTAT -3' and 5'-

GTAGGATCCAAGCGGATCCTGGAG-3', and *E. coli* st. K12 genomic DNA was used as template. OxyS run off transcripts were generated from plasmids constructed from pUC19 and the OxyS gene. All RNAs were synthesized using the MEGAscript T7 kit (Ambion) according to manufacturer's protocol. RNAs were purified by ammonium acetate precipitation after digestion of template with DNase (Epicentre). RNAs were characterized by native and denaturing gel electrophoresis, and their concentrations determined by UV absorbance. ^{32}P-labeling of the RNAs at their 5' end was carried out using standard protocols (Sambrook & Russell, 2001). RNA was dephosphorylated with shrimp alkaline phosphatase, radioactively labeled at the 5' end with [γ-32]ATP and T4 polynucleotide kinase, and purified by NUCAWAY spin column (Ambion).

Electrophoretic gel mobility shift assay

Binding reactions between wt Hfq and mutant Hfq to ^{32}P labeled DsrA, RprA, OxyS, *rpoS*-L, *rpoS*-S or *rpoS*$323_{254\text{-}457}$ were prepared in 15 µL volumes. 5 µL of 12 nM RNA (4 nM final concentration) was mixed with 7.5 µL of an Hfq solution to give the appropriate Hfq concentration and 2.5 µL of loading buffer added (0.25% bromophenol blue, 30% glycerol). The reaction solvent consisted of 20 mM Tris–HCl (pH 8.0), 100 mM NH_4Cl, 50 mM NaCl, 50 mM KCl, and 5% glycerol. Reactions were incubated for at least ten minutes at room temperature prior to loading in gel lanes. Samples were analyzed by electrophoresis on native 5% polyacrylamide gels (acrylamide 29:1 (w/w) /bisacrylamide) gels with 3% glycerol in 0.5X TBE. Gels were run at 80 to 100V. The apparent equilibrium dissociation constants (K_d) were determined from the interpolated concentration of Hfq_6 required to shift 50% of the total RNA from the free band to a complex.

Association of *rpoS*-L with ^{32}P-labeled DsrA or OxyS in the absence or presence of wt Hfq or mutant Hfq was evaluated using the same reaction volume and buffer described above. Varying concentrations of unlabeled *rpoS*-L were added to 4 nM DsrA or 4 nM OxyS and incubated for 60 minutes at 25 ^{o}C prior to electrophoresis at 4 ^{o}C on a native 5% polyacrylamide gel. The lower temperature used during electrophoresis helped preserve the RNA-RNA complexes. Gels were run at 100-115 V for 60 to 90 minutes and analyzed using a Fujifilm Image Reader FLA-3000.

RESULTS

The influence of single amino-acid mutations and truncation of the C-terminal end on the binding of *E. coli* Hfq to OxyS, RprA, and DsrA

Immunoprecipitation and gel mobility shift experiments have shown that Hfq binds to OxyS *in vivo* and *in vitro* (Zhang et al. 1998; Zhang et al. 2002); however, evaluation of Hfq binding affinity for this sRNA has not been previously reported. Figure 3.2 shows an electrophoretic gel mobility shift assay (EMSA) of 4 nM ^{32}P 5' labeled OxyS with increasing amounts of Hfq. As observed previously with this (Zhang et al. 2002) and other similar length sRNAs (Lease and Woodson 2004; Updegrove et al. 2008), two discrete Hfq-OxyS complexes are observed (C1 and C2). The relative mobility of the bands as a function of Hfq concentration implies more Hfq are bound in the C2 complex. The amount of Hfq_6 required to shift 50% of free OxyS to a complex was estimated to be 55 ± 10 nM. This apparent K_d is approximately twice the value determined for Hfq binding to DsrA and RprA under identical conditions (Table 3.1).

Table 3.1: Apparent equilibrium dissociation constants (K_d) for wt and mutant Hfq binding to DsrA, RprA, OxyS and *rpoS* fragments by gel mobility shift assay[a]. Apparent binding constant for DsrA binding *rpoS*-L in the presence of 100 nM wt or mutant Hfq[b]. Values reported in nM.

Hfq	DsrA[a]	RprA[a]	OxyS[a]	*rpoS*-S[a]	*rpoS*323[a]	DsrA• r*poS*-L[b]	*rpoS*-F[a]
wt	23 ± 5	25 ± 5	55 ± 5	45±10	20± 5	25 ± 10	35±10
Proximal							
Q8A	19 ± 10	30[c]	175 ± 15	100[c]	20 ± 5	125 ± 25	15±10
R16A	47 ± 10	35[c]	155 ± 10	160[c]	20 ± 5	150 ± 25	150±40
F39A	50 ± 10	150[c]	130 ± 5	160[c]	20 ± 5	200 ± 50	25 ± 5
F42A	15 ± 10	35[c]	67 ± 5	75[c]	20 ± 5	75 ± 10	25±5
L12F/F39A	21 ± 10	63±10	105±10	150±20	NA	125 ± 25	35±5
F39A/F42A	NA	NA	NA	NA	NA	300 ±50	25 ±5
Distal							
Y25A	22 ± 5	80[c]	71 ± 5	90[c]	75 ± 10	250 ± 50	150±50
K31A	16 ± 10	100± 10	80 ± 5	75±10	175 ± 25	250 ± 50	150±50
Outer-rim							
F11A	40 ± 10	45 ± 5	74 ± 5	40 ± 10	25 ± 5	50 ± 10	15 ±5
R17A	16 ± 15	20 ± 10	66 ± 10	40 ± 10	20 ± 5	60 ± 20	15 ± 5
R19A	35 ± 15	35 ± 20	50 ± 5	25 ± 10	20 ± 5	15 ± 10	60 ±10
C-terminal							
Hfq-65	39 ± 10	37 ± 15	66 ± 5	50 ± 10	35 ± 10	20 ± 10	> 200
Hfq-75	33 ± 10	35 ± 10	47 ± 5	50 ± 10	20 ±5	20 ± 10	> 200

[a]Apparent binding constants determined from the amount of Hfq (moles hexamer) needed to promote 50% of free RNA into protein-RNA complex by gel mobility analysis.
[b]Apparent binding constants evaluated from amount of *rpoS*-L needed to promote 50% of free DsrA into DsrA•*rpoS*-L complex with Hfq. [c]Values previously reported in Updegrove et al. 2008. Values and error based on three independent experiments. NA refers to Not Assessed

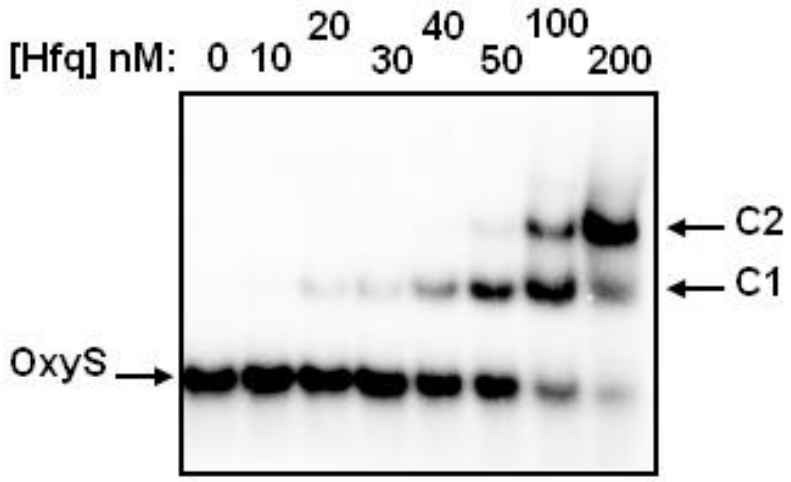


Figure 3.2: Binding of wt Hfq to OxyS assessed by the EMSA. 4 nM OxyS was used for each sample. Hfq concentrations in this and other experiments are given in moles hexamer/L.

The effect of mutating Hfq on its affinity to OxyS was assessed using the gel shift assay. Figure 3.3 shows typical experiments with Hfq-F39A, Hfq-Q8A, and Hfq-R19A binding to OxyS, and Table 3.1 shows the apparent K_d for all mutant Hfq tested against OxyS. Three proximal surface mutations, Hfq-F39A, Hfq-R16A, and Hfq-Q8A, exhibited a 2.3 to 3.2-fold decrease in affinity compared to wt Hfq, while another proximal surface mutation, Hfq-F42A, was similar to wt Hfq. The two truncated Hfq as well as Hfq with mutations on the distal surface and the outer- circumference surface produced affinities comparable to wt Hfq.

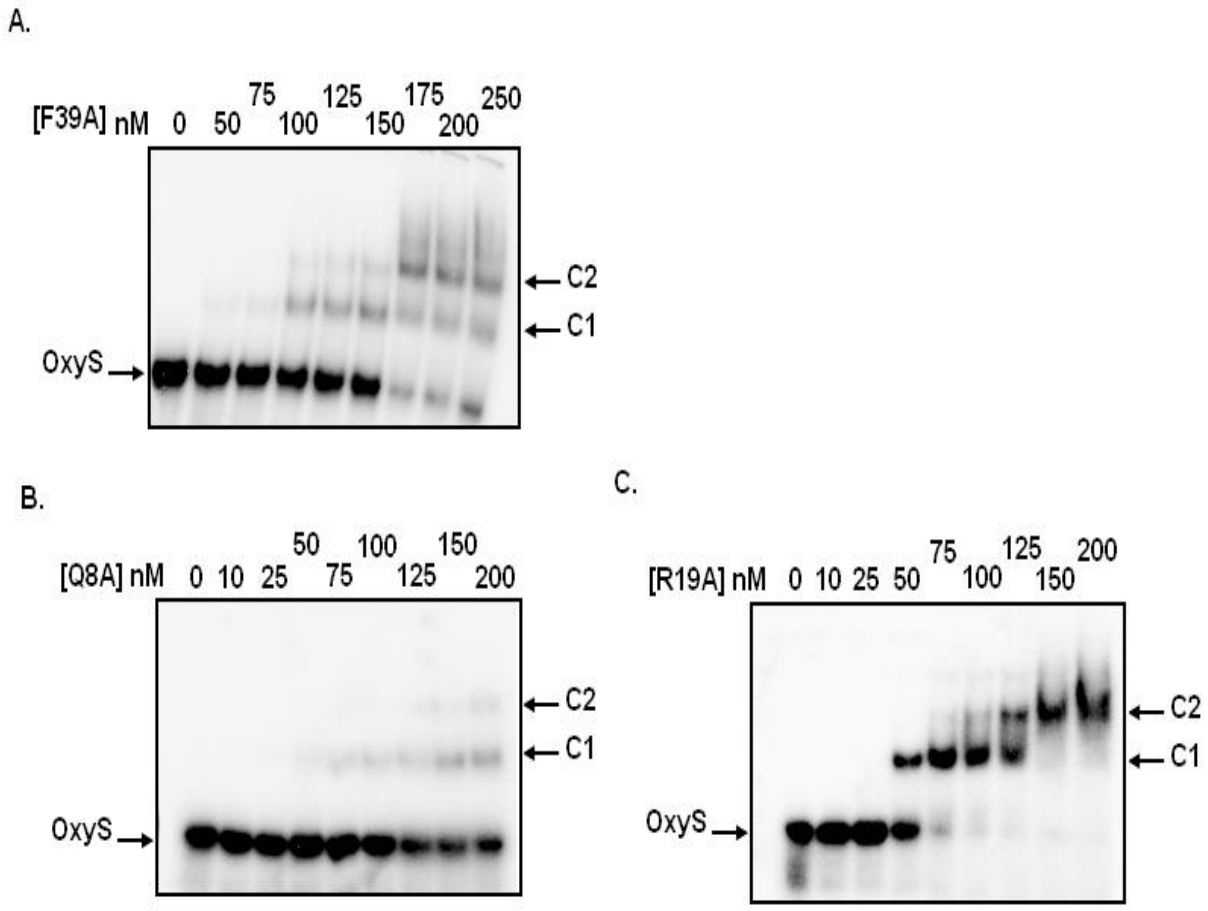


Figure 3.3: Binding of Hfq-F39A, Hfq-Q8A, and Hfq-R19A to 4 nM OxyS assessed by EMSA.

The binding of mutant Hfq to OxyS differ somewhat from data obtained with RprA (Table 3.1). For RprA, the only mutations that showed more than two-fold decrease in binding were Hfq-F39A, Hfq-Y25A and Hfq-K31A. Table 3.1 lists RprA data obtained previously as well as new data obtained with the truncated Hfq and the mutations to outer-circumference residues. Figure 3.4 shows gels of RprA binding Hfq-65, Hfq-F11A, and Hfq-R19A. Although the outer circumference mutant Hfq-F11A showed a slightly reduced affinity for RprA than wt Hfq, the other outer circumference mutations and the C-terminal truncated Hfq displayed affinities similar to wt Hfq (Table 3.1).

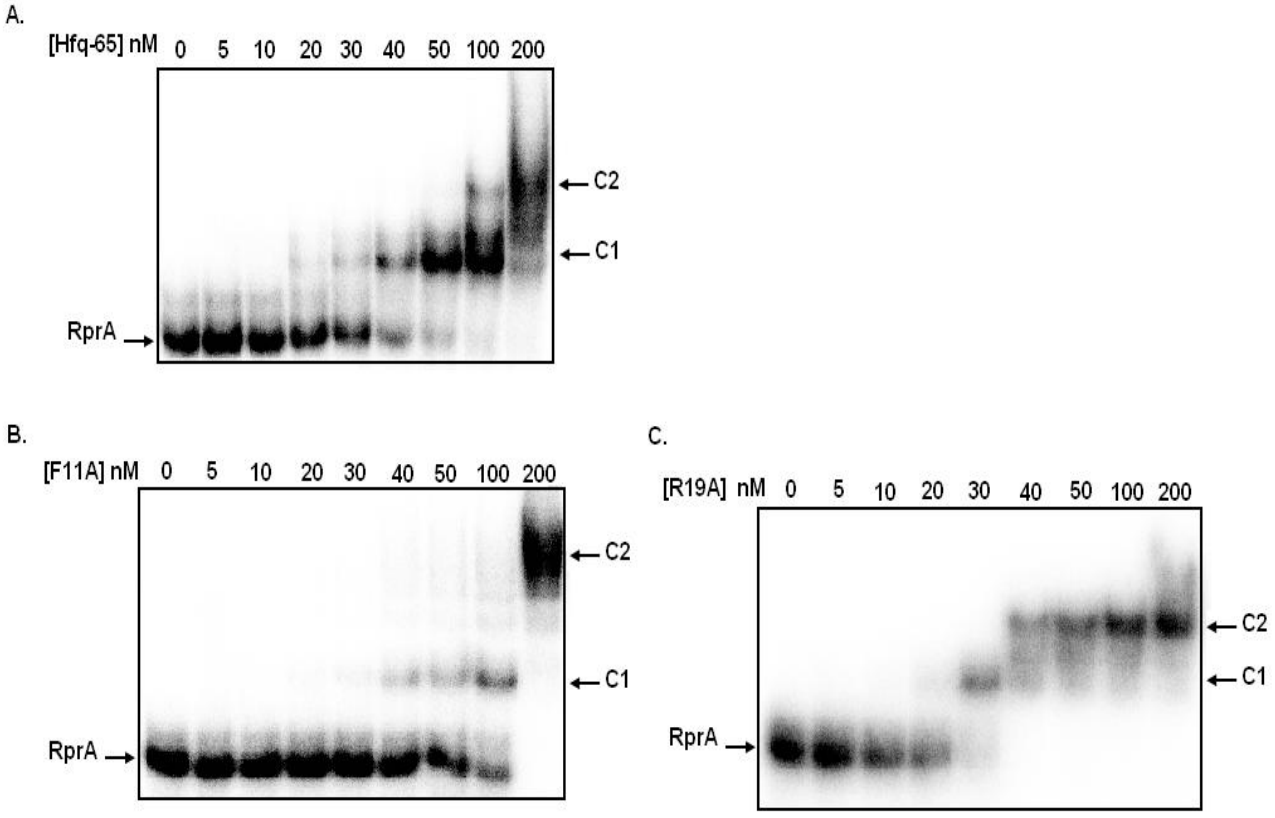


Figure 3.4: Binding of Hfq-65, Hfq-F11A, and Hfq-R19A to 4 nM RprA assessed by EMSA.

Data on the binding of wild-type and mutant Hfq to DsrA is also listed in Table 3.1. The results supplement previous work that evaluated the binding of mutant Hfq to DsrA (Mikulecky et al. 2004) and $DsrA_{DII}$ (Sun and Wartell 2006) - a 38 nt long 5'-fragment of DsrA that competes with full length DsrA in binding Hfq (Brescia et al. 2003). Figure 3.5 shows three representative gels of DsrA binding to Hfq-F39A, Hfq-Y25A, and Hfq-F42A. Similar to OxyS, proximal surface mutations R16A and F39A showed greater than two fold decrease in binding to DsrA compared to wt Hfq. These same two mutations were previously shown to reduce binding of Hfq to $DsrA_{DII}$ by about seven fold (Sun and Wartell 2006). This suggests that full length DsrA possess additional Hfq contacts not present with $DsrA_{DII}$, and is therefore not as sensitive

to these amino-acid changes. We also note that similar to the previous study with $DsrA_{DII}$ (Sun and Wartell 2006), the F39A/L12F double mutation regained binding affinity to DsrA lost by the F39A change. Since residues 12 and 39 are adjacent in the 3-D structure (Sauter et al. 2003), it reinforces the notion that a phenylalanine residue at this location is important for maximum binding. Similar to OxyS and RprA, DsrA binding to the C-terminal truncated Hfq, and Hfq with mutations on the outer circumference showed small changes (less than two-fold) relative to wt Hfq. A two-fold decrease in affinity of Hfq-65 to DsrA was previously reported (Sonnleitner et al. 2004).

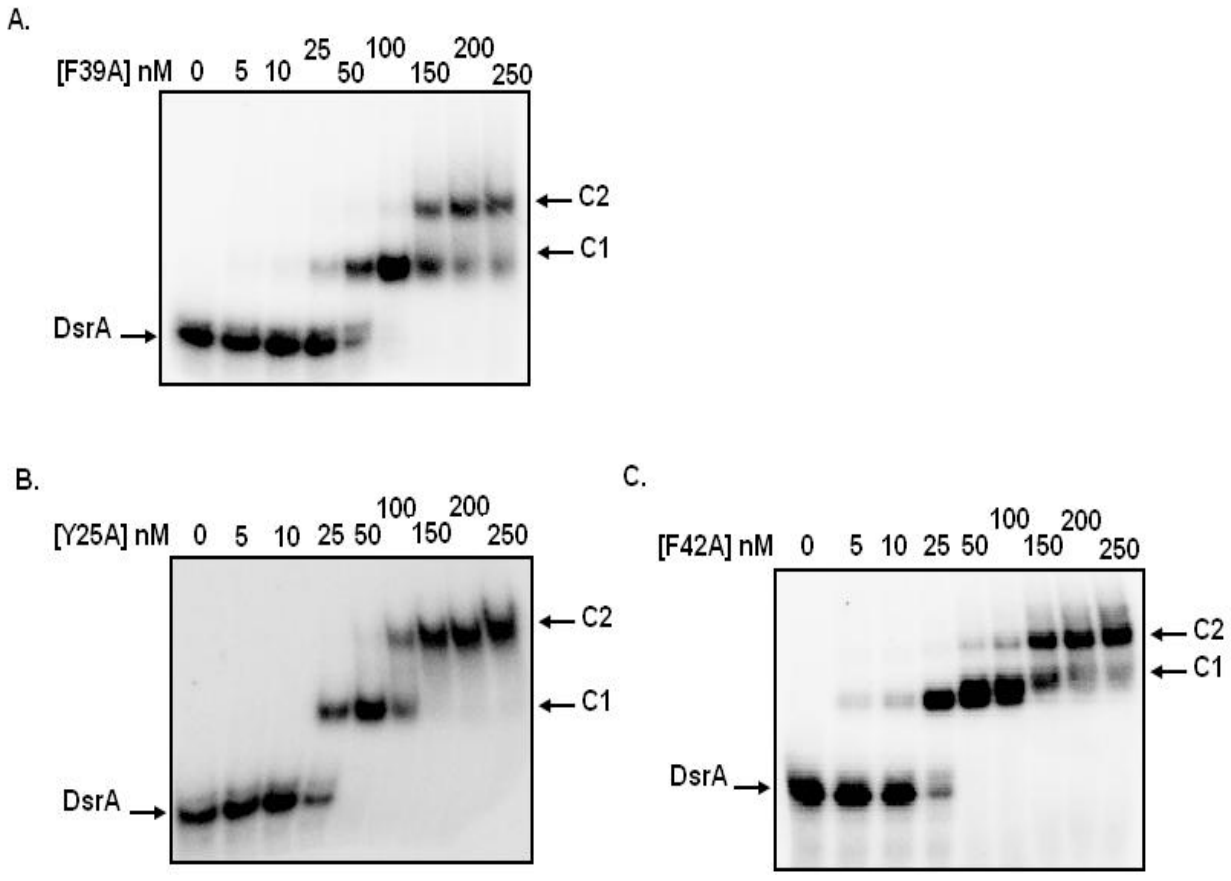


Figure 3.5: Binding of Hfq-F39A, Hfq-Y25A, and Hfq-F42A to 4 nM DsrA assessed by EMSA.

The binding of wild-type and mutant Hfq to fragments of the *rpoS* mRNA

The Hfq mutants Hfq-F42A, Hfq-Y25A, Hfq-Q8A, Hfq-F39A, and Hfq-R16A were previously tested for their ability to bind a fragment of *rpoS* mRNA referred to as *rpoS*-S (Updegrove et al. 2008). This fragment contains 127 nt upstream of the AUG start codon and 75 nt of the coding sequence beyond the start codon (Figure 3.6). It is predicted to form an intramolecular stem loop structure that sequesters the ribosome binding site (RBS) into a duplex segment that is opened by hybridization with RprA or DsrA (Majdalani et al. 1998; Majdalani et al. 2002). The proteins Hfq-F42A, Hfq-Y25A, and Hfq-Q8A produced discrete shifted bands with *rpoS*-S and an apparent K_d 1.5- to 2-fold higher than wt Hfq. Titration of Hfq-F39A and Hfq-R16A with *rpoS*-S exhibited weaker binding and apparent K_d values 3.5-fold higher than wt Hfq (Table 3.1). Gel shift experiments that assessed *rpoS*-S binding to Hfq with mutations on the outer-circumference surface and to the two truncated Hfq proteins produced K_d values similar to wt Hfq (Table 3.1).

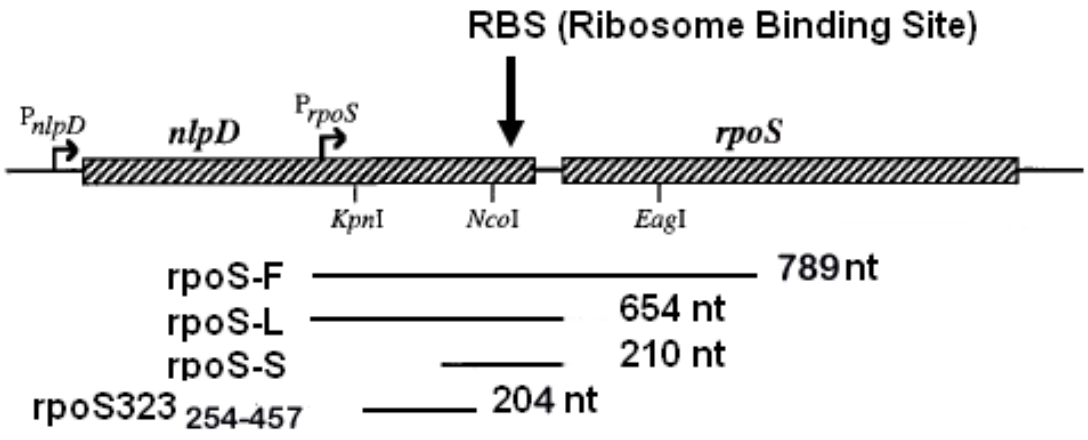


Figure 3.6: Map of the *rpoS* gene and the relative position of each generated *rpoS* construct.

Recent studies indicate that sequences upstream of *rpoS*-S are crucial for Hfq's ability to produce 20-30 fold enhancement of intermolecular pairing between RprA or DsrA with *rpoS* mRNA (Soper and Woodson 2008; Updegrove et al. 2008). This upstream region contains $(ARN)_4$ and A_6 elements that are potential binding sites for Hfq. The $(ARN)_4$ site appears to be responsible for the stronger binding of Hfq to the full length leader sequence of *rpoS* mRNA compared to *rpoS*-S and similar core sequences (Soper and Woodson 2008). A model was proposed in which Hfq binding to this upstream region alters the *rpoS* mRNA structure that inhibits DsrA and RprA pairing to the RBS site and also helps recruit an sRNA to the RBS site (Soper and Woodson 2008). The binding of the mutant Hfq proteins to this upstream region was examined.

The *rpoS* mRNA fragment $rpoS323_{254\text{-}457}$ was previously shown to bind wt Hfq with strong affinity (Soper and Woodson 2008). This 204 nt RNA encompasses nucleotides 254 to 457 numbered from a *rpoS* mRNA transcription start site (Lange et al. 1995) and contains the $(ARN)_4$ and A_6 elements but is missing the RBS site. Among the eleven mutant Hfq proteins examined only Hfq-K31A and Hfq-Y25A exhibited binding to $rpoS323_{254\text{-}457}$ that was weaker than wt Hfq (Table 3.1, Figure S3.1). This result provides evidence that only the distal surface of Hfq is involved in binding to this portion of the *rpoS* leader region.

The ability of the mutant Hfq proteins to bind to the entire leader region of *rpoS* was explored. The RNA *rpoS*-L is a construct of *rpoS* that contains the entire untranslated leader region of *rpoS* mRNA and 75 nt of the coding sequence (Figure 3.6). The apparent K_d of wt Hfq binding to this RNA was previously estimated to be ~35 nM using the gel shift assay (Updegrove et al. 2008). We verified wt Hfq binding *rpoS*-L, and tested the binding of several mutant Hfq to this RNA (data not shown). However, we found *rpoS*-L to be intrinsically difficult to use in the

gel shift assay. This RNA was more susceptible to degradation than other RNAs during the process of radioactive labeling. Additionally, free *rpoS*-L formed a doublet band making it difficult to assess small shifts in mobility. Other lengths of *rpoS* mRNA that contained the leader region and a portion of the coding region were examined for binding studies with mutant Hfq proteins.

An *rpoS* mRNA construct, denoted as *rpoS*-F, was prepared that possessed the entire 565 nt leader region, 200 nt of the coding region and additional 3'segment (Figure 3.6). *rpoS*-F was amendable to 5' end labeling with ^{32}P with little degradation and appeared as a single band in the gel. Binding of wt Hfq and Hfq mutants to *rpoS*-F was examined. Figure 3.7A shows the titration of wt Hfq to 4 nM of ^{32}P-labeled *rpoS*-F. An apparent K_d of ~35 nM was estimated from the lowest Hfq concentration that noticeably retarded the mobility of the free RNA, in good agreement with the value obtained with *rpoS*-L (Updegrove et al. 2008).

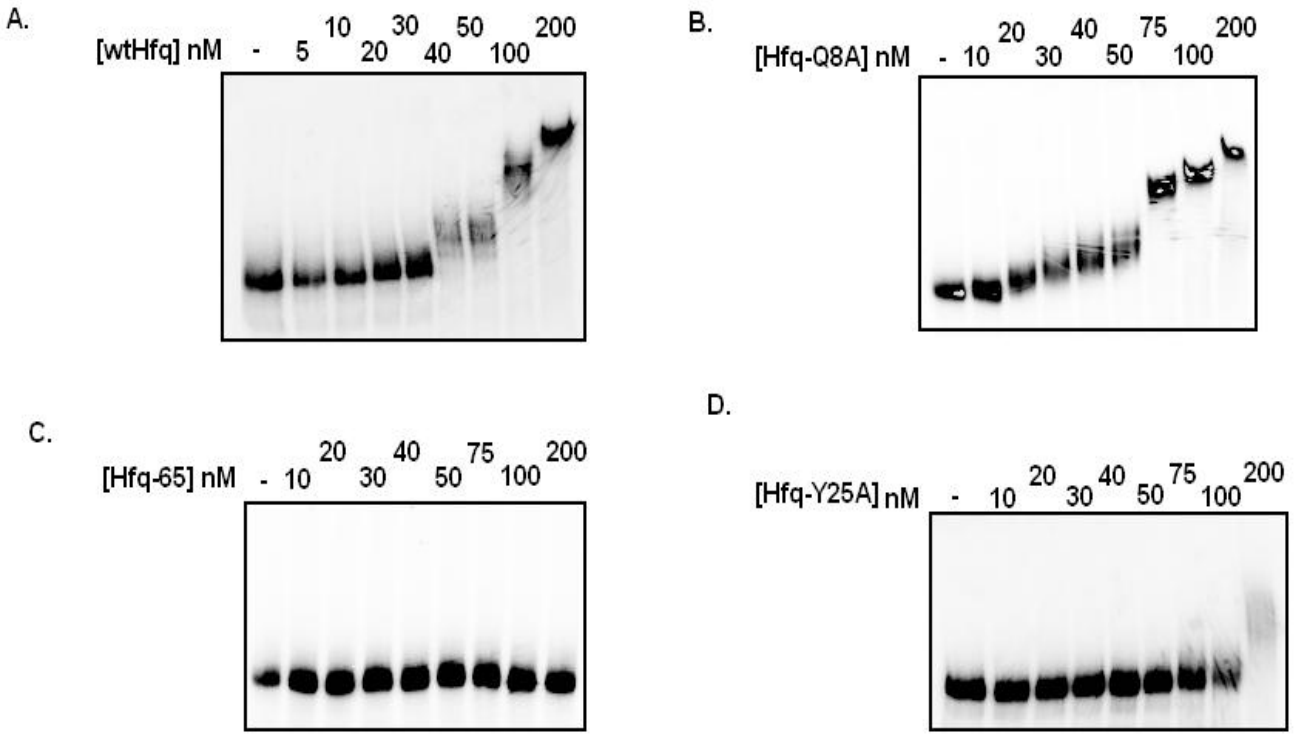


Figure 3.7: Binding of wt Hfq, Hfq-Q8A, Hfq-65, and Hfq-Y25A to 4 nM *rpoS*-F assessed by EMSA.

Representative gel mobility shift assays of mutant Hfq binding to *rpoS*-F are shown in Figure 3.7. Hfq-65 (and Hfq-75) had no apparent effect on *rpoS*-F mobility up to 200 nM, and Hfq-Y25A did not indicate binding until 200 nM. Hfq-Q8A was similar to wt Hfq producing a gradual decrease in *rpoS*-F mobility starting from 20 nM followed by a shift to a lower mobility complex at 75 nM. Among the other mutant Hfqs, Hfq-31A and Hfq-R16A displayed weak binding, while Hfq-F39A, Hfq-F42A and the three Hfq with outer-circumference mutations behaved similar to wt Hfq (Table 3.1). The apparent lack of binding of Hfq-65 and Hfq-75 to *rpoS*-F was unexpected since these truncated Hfq displayed an affinity to *rpoS*-S and $rpoS323_{254\text{-}457}$ similar to wt Hfq (Figure 3.8, Table 3.1). This apparent incongruity is explored further in the

Discussion. We note that Hfq-65 induces a smaller shift in the mobility to *rpoS*-S and *rpoS323*$_{254\text{-}457}$ than wt Hfq (Figure 3.8).

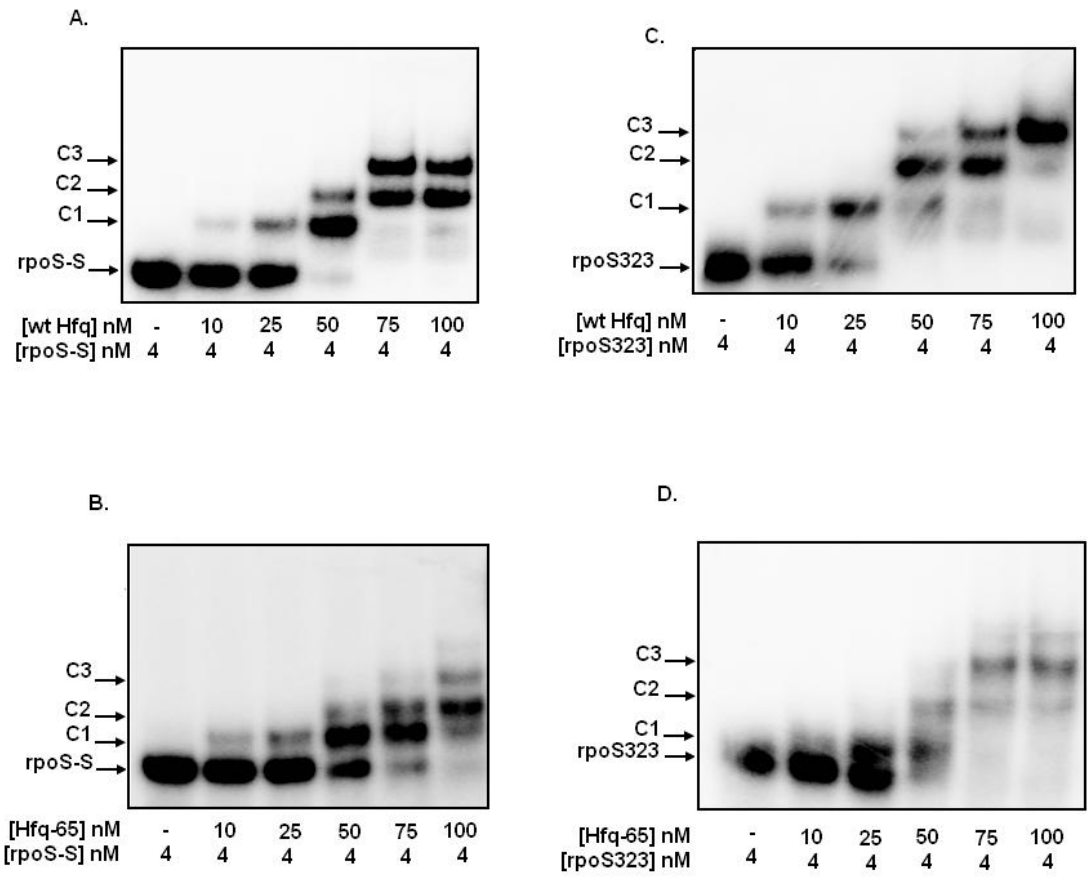


Figure 3.8: Gel mobility shift assay of the binding of wt Hfq to *rpoS*-S (**A.**), Hfq-65 to *rpoS*-S (**B.**), wt Hfq to *rpoS323*$_{254\text{-}457}$ (**C.**), and Hfq-65 to *rpoS323*$_{254\text{-}457}$ (**D.**). 4 nM RNA was used in each experiment.

Effect of Hfq mutants on stimulating DsrA•*rpoS* annealing *in vitro*

To determine the influence of Hfq mutants on their ability to enhance DsrA•*rpoS* annealing we utilized the *rpoS*-L RNA. DsrA binding to *rpoS*-L was determined by titrating 4 nM ^{32}P labeled DsrA with increasing amount of unlabeled *rpoS*-L in the absence or presence of

Hfq. Apparent K_d values were estimated from the *rpoS*-L concentration required to sequester 50% of the free DsrA into the *rpoS*-L•DsrA complex. Under our reaction conditions, the apparent K_d of DsrA binding *rpoS*-L was ~750 nM (Updegrove et al. 2008). The addition of 100 nM wt Hfq_6 promoted *rpoS*-L•DsrA duplex formation and reduced the apparent K_d to 25 nM (Table 3.1, Figure S3.2).

The ability of the mutant Hfqs to stimulate DsrA•*rpoS*-L annealing was examined under the conditions employed for wt Hfq. Figure 3.9 shows three representative experiments with Hfq-65, Hfq-Y25A, and Hfq-L12F/F39A. Binding isotherms for all mutant Hfqs tested are shown in Figure S3.3. The mutant Hfqs that displayed the greatest defects in stimulating sRNA•mRNA duplex formation were Hfq-K31A, Hfq-Y25A, Hfq-Q8A, Hfq-R16A, Hfq-F39A, Hfq-L12F/F39A, and Hfq-F39A/F42A (Table 3.1). This result supports the notion that both the proximal and distal surfaces of Hfq are important for DsrA•*rpoS*-L annealing. Hfq mutants that showed weak binding to either *rpoS*-S or $rpoS323_{254\text{-}457}$ also produced weak stimulation of DsrA• *rpoS*-L annealing (Table 3.1).

Under the gel electrophoresis conditions employed the mobility of the binary DsrA•*rpoS*-L complex could not be distinguished from the ternary complex involving Hfq and the two RNAs (Figure S3.2A, right-most two lanes). To assess if Hfq forms a ternary complex with DsrA and *rpoS*-L in the gel, an Hfq derivative carrying a C-terminal Flag-tag sequence (Hfq- Flag) was employed in an assay using the conditions of Figure S3A. Due to differing gel mobility properties of Hfq-Flag compared to wt Hfq, the complex formed by Hfq-Flag, DsrA, and *rpoS*-L could now be distinguished from the DsrA•*rpoS*-L complex (Figure S3.2B, lanes 8 & 9). A western blot using anti-Flag primary antibodies also demonstrated the presence of Hfq-Flag in this ternary complex (Figure S3.2C).

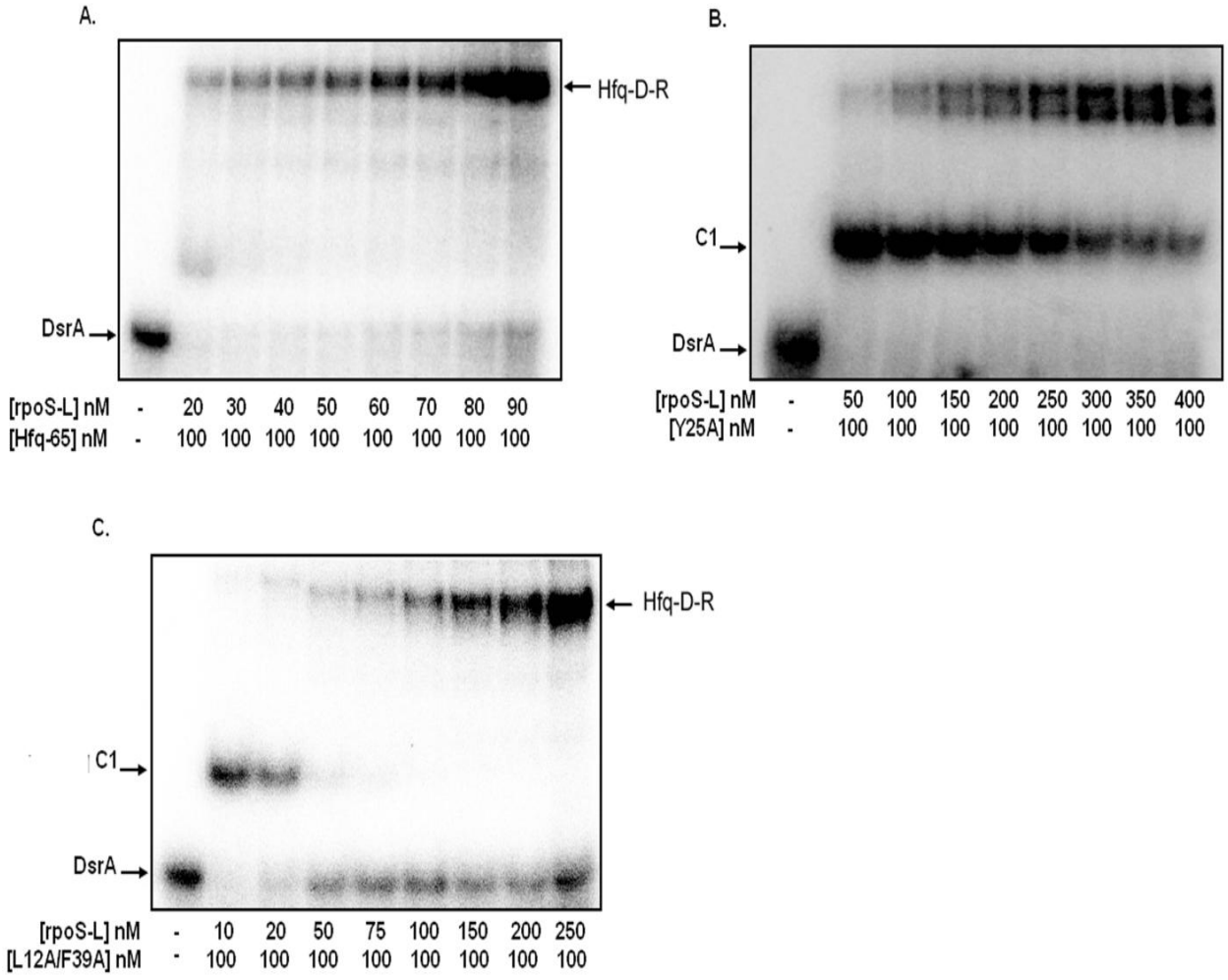


Figure 3.9: Gel shift assay of the binding of *rpoS*-L to 4 nM ^{32}P-labeled DsrA in the presence of 100 nM of Hfq-65, Hfq-Y25A, and Hfq-F39A/L12F. C_1 is the Hfq•DsrA complex, H-D-R is the ternary complex involving Hfq, *rpoS*-L, and DsrA.

Binding of OxyS to *rpoS* RNA and the influence of Hfq

In vivo studies indicate that OxyS negatively regulates *rpoS* gene expression posttranscriptionally and depends on Hfq (Zhang et al. 1998). One model of how Hfq influences OxyS regulation of *rpoS* assumes Hfq facilitates intermolecular pairing of OxyS to a *rpoS* mRNA site and this complex inhibits *rpoS* mRNA translation (Zhang et al. 2002). Evidence supporting this model include the observations that OxyS and *rpoS* mRNA are found associated

with immuneprecipitated Hfq (Zhang et al. 1998), and *in vitro* studies that indicate Hfq can form a ternary complex with OxyS and an *rpoS* mRNA segment (Zhang et al. 2002). In order to assess this model further, a quantitative evaluation of the formation of the OxyS•*rpoS* RNA complex and the influence of Hfq on its formation were made.

When 4 nM of ^{32}P-labeled OxyS was titrated with increasing amount of unlabeled *rpoS*-L, very little OxyS•*rpoS*-L complex was observed using the gel mobility shift assay (Figure 3.10A). Concentrations up to 4 μM *rpoS*-L showed less than 10% of the total OxyS shifted to a OxyS•*rpoS*-L complex after a 1 hr equilibration. Under the same reaction conditions, the addition of ~750 nM and 2.5 μM *rpoS*-L to DsrA and RprA respectively shifted over 50% of these sRNAs to a complex with *rpoS*-L (Updegrove et al. 2008). Since DsrA and RprA bind *rpoS*-S significantly better than to *rpoS*-L (Updegrove et al. 2008), OxyS binding to *rpoS*-S was examined. As Figure 3.10B shows, the gel shift assay also indicated a weak affinity of OxyS for *rpoS*-S. Less than 10% of the OxyS shifted to the OxyS•*rpoS*-S complex in the presence of 4 μM *rpoS*-S.

Unlike DsrA and RprA, the addition of wt Hfq_6 had relatively little effect on OxyS binding to *rpoS*-L by the gel shift assay (Figure 3.10C). A similar result was obtained with *rpoS*-S (data not shown). This amount of Hfq_6 was sufficient to enhance DsrA and RprA binding to *rpoS*-L 25 to 30 fold. A range of Hfq concentrations was examined to determine if a different amount would optimize OxyS binding to *rpoS*-L. Increasing amounts of wt Hfq were added to 4 nM ^{32}P-labeled OxyS mixed with 1 μM *rpoS*-L. The reactions were then run on a gel to determine OxyS•*rpoS*-L formation. Hfq had a very small effect even at concentrations approaching 2 μM Hfq_6 (Figure 3.10D). Only the highest Hfq concentration employed induced a

ternary complex as previously observed (Zhang et al. 2002). Hfq does not enhance binding of OxyS to *rpoS* mRNA to the same extent as DsrA and RprA.

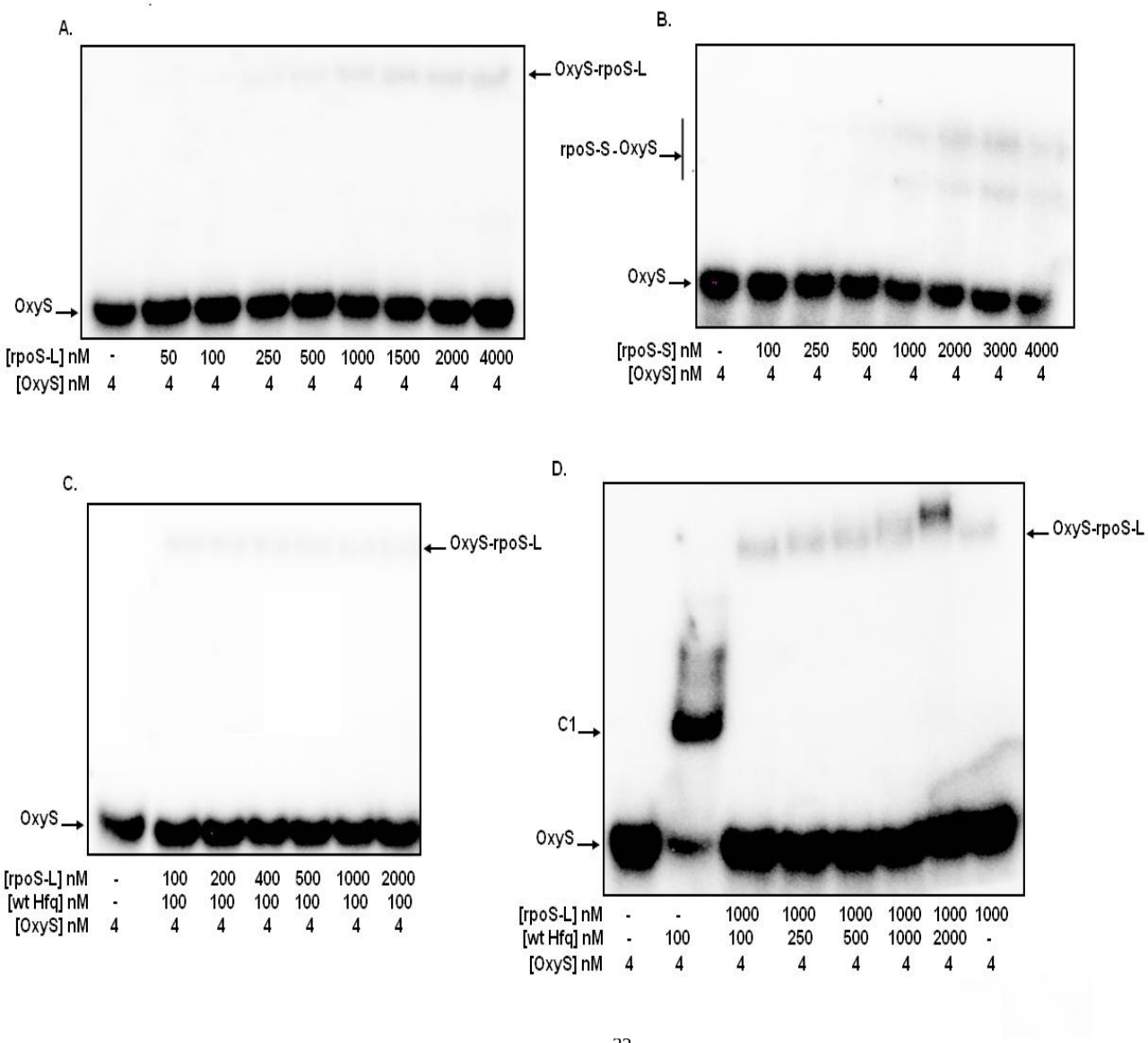


Figure 3.10: Gel shift assays of the binding of 4 nM ^{32}P-labeled OxyS to: (**A.**) increasing (50-4000 nM) *rpoS*-L; (**B.**) increasing (100-4000 nM) *rpoS*-S; (**C.**) increasing (100-2000 nM) *rpoS*-L in the presence of 100 nM wt Hfq_6; and (**D.**) 1000 nM *rpoS*-L in the presence of increasing (100-2000 nM) wt Hfq_6.

DISCUSSION

The sRNAs DsrA, RprA, and OxyS require Hfq in order to elicit riboregulation of *rpoS* mRNA (Brown and Elliott 1996; Sledjeski et al. 1996; Zhang et al. 1998; Majdalani et al. 2002). DsrA and RprA bind to wt Hfq with an apparent K_d of ~25 nM under the conditions of our assay, while OxyS exhibited a lower affinity with a K_d ~ 55 nM. To ascertain if the above sRNAs interact with similar sites on Hfq, we examined their ability to bind mutant Hfq with single residue changes on the protein's three main surface regions – proximal, distal, and outer circumference, and Hfq proteins that had portions of their C-terminal ends deleted.

Mutations on the distal surface of Hfq had little effect on the protein's ability to bind DsrA and OxyS, but reduced the apparent K_d for RprA by 3 to 4-fold. Two lines of evidence support interpreting this observation as a distal surface contact with RprA: (i) poly(A) displaces RprA but not DsrA from wt Hfq in a gel shift assay (Updegrove et al. 2008), and (ii) RprA has a $(ARNN')_4$ repeat from positions 47 to 62, a sequence suggested from structural and biophysical analysis to be capable of strong distal side binding (Link et al. 2009). Single site mutations on the outer-circumference surface, and deleting portions of the C-terminal end had little impact on Hfq binding to all three sRNAs. Mutations on the proximal surface produced varying effects for the sRNAs: F42A had no significant effect on Hfq binding to all three sRNAs, Q8A reduced Hfq's affinity for OxyS by 3-fold while exhibiting no effect on its affinity for DsrA and RprA, R16A reduced Hfq affinity to DsrA and OxyS but not RprA, while F39A reduced binding to all three sRNAs.

Our results indicate that the C-terminal ends, the outer circumference mutations examined, and Phe42 do not play a significant role in binding the three sRNAs. Residues on the

proximal surface did influence Hfq-sRNA binding, however no mutation except for F39A affected Hfq binding to all three sRNAs, and the effects of all mutations on their apparent K_d were small. These results reinforce the notion of multiple weak interactions previously inferred by Feig and colleagues in their study of mutant Hfq binding to RNAs (Mikulecky et al. 2004). The strong affinity of Hfq for sRNAs appears to involve multiple sites on the proximal and sometimes distal surface that vary for different sRNAs. The current results together with previous mutational studies (Mikulecky et al. 2004) suggest that a segment of DsrA and perhaps OxyS as well, track along a path from the sm2 motif residues Tyr55 and Lys56 to Phe39 and Arg16 (Figure S3.4). Contacts along this path may vary depending on the RNA sequence. This speculative model is consistent with the broad sequence specificity Hfq appears to exhibit for sRNAs.

Chemical and nuclease protection assays indicate that the leader portion of *rpoS* mRNA has several Hfq binding sites (Lease and Woodson 2004; Soper and Woodson 2008; Updegrove et al. 2008). Our results on the binding of mutant Hfqs to r*poS*-S and $rpoS323_{254\text{-}457}$ are consistent with the notion that Hfq's proximal surface sites bind to the *rpoS*-S segment, while only distal surface sites interact with $rpoS323_{254\text{-}457}$ (Table 3.1). Mutations along the outer circumference of Hfq and the truncation of the C-terminal end had little impact on binding either *rpoS* segment.

The ability of mutant Hfq to stimulate annealing of DsrA to *rpoS*-L provided a more functional assessment of the mutations. DsrA hybridization to *rpoS*-L was ≥ 5-fold lower in the presence of Hfq-Q8A, Hfq-R16A, Hfq-F39A, Hfq-Y25A, and Hfq-K31A compared to wt Hfq. The reduced effectiveness of these mutants correlates with their >2-fold lower affinity to either *rpoS*-S or $rpoS323_{254\text{-}457}$. The mutant Hfq with outer-circumference mutations and C-terminal

deletions stimulated DsrA•*rpoS*-L annealing similar to wt Hfq and bound *rpoS*-S and $rpoS323_{254\text{-}457}$ with affinities similar to wt Hfq. Strong Hfq binding to both the RBS region and $(ARN)_4$ site appears necessary to maximize DsrA annealing to its *rpoS* RNA target. We also note that Hfq-L12F/F39A displayed a greater enhancement of DsrA•*rpoS*-L hybridization (1.6 fold increase) than Hfq-F39A. This is consistent with the regained affinity of Hfq-L12F/F39A for domain II of DsrA (Sun and Wartell 2006) and DsrA (Table 3.1) compared to Hfq-F39A, and supports a functional significance for the covariance at neighboring positions 12 and 39 inferred from a comparison of Hfq sequences of bacterial species (Sun and Wartell 2006).

Hfq-65 was previously shown to be less effective than wt Hfq in binding a ~575 nt *rpoS* RNA containing the entire leader region in a gel shift assay(Vecerek et al. 2008), yet Hfq-65 facilitated sRNA regulation of *rpoS in vivo* (Olsen et al. 2010). We found Hfq-65 and Hfq-75 displayed wt Hfq like effectiveness in promoting DsrA•*rpoS*-L binding and high affinity for the ~ 200 nt *rpoS*-S and $rpoS323_{254\text{-}457}$ RNAs in the gel shift assay, yet these truncated Hfq did not alter the mobility of the longer *rpoS*-F RNA (Table 3.1, Figure 3.7) or *rpoS*-L (data no shown). A change in mobility of a long RNA due to the binding of these smaller proteins is expected to be more difficult to detect than for a short RNA. However the large mobility shift of *rpoS*-F RNA induced by wt Hfq suggests other considerations may be needed to explain the dissimilar behavior of Hfq-65 and Hfq-75 with the 200 nt *rpoS* RNA segments vs *rpoS*-F RNA.

In order to detect a significant mobility shift of long RNAs such as *rpoS*-F it may be necessary for multiple Hfq hexamers to bind. One hypothesis is that wt Hfq can readily aggregate onto RNA in the gel environment while Hfq-65 and Hfq-75 is less able to do so. Although the biological role of the C-termini of *E. coli* Hfq remains uncertain, it has been suggested they could be involved in protein-protein interactions (Brennan and Link 2007). The

potential extent of the C-termini is illustrated in Figure S3.5. Hfq-65 and Hfq-75 may bind to a few sites on the long *rpoS* RNA but not induce a significant shift in mobility because they do not form larger aggregates. Using a two-color fluorescent staining method (Jing et al. 2003) we assessed this possibility in the bands of a gel shift assay in which *rpoS*-F was titrated with Hfq-65 or Hfq-75 (Figure S3.6). Due to the low sensitivity of the SYPRO Ruby red protein stain for Hfq, higher concentrations of *rpoS*-F (200 nM) and Hfq (0 to 2 μM) were employed. Both Hfq-65 and Hfq-75 bound to *rpoS*-F at concentrations that did not affect the mobility of the RNA. This observation lends support to the above ideas, and adds a cautionary note to interpreting gel mobility shift data with long RNAs.

Our results indicate that binding of OxyS to either *rpoS*-S or *rpoS*-L is weak relative to DsrA or RprA in the absence or presence of Hfq. A stretch of complementary bases (9 out of 10) occurs between positions 80-89 of the OxyS sequence and +19 to +28 of *rpoS* mRNA relative to the start codon. This potential hybridization site may be relevant to OxyS repression of translation; however, our results do not indicate a significant stimulation of OxyS binding to *rpoS* RNA by Hfq. Our findings are more consistent with recent *in vitro* as well as *in vivo* experiments that suggest OxyS may repress *rpoS* expression by displacing Hfq from sRNAs such as DsrA that enhance translation (Fender et al. 2010; Hussein and Lim 2011).

Acknowledgements

I thank Xueguang Sun and Charles Terry for assistance with the purification of mutant Hfq used in the study, and I am grateful for the support provided by the NASA Astrobiology Institute and the GAANN Fellowship awarded to me.

REFERENCES

Aiba, H. 2007. Mechanism of RNA silencing by Hfq-binding small RNAs. *Curr Opin Microbiol* **10**(2): 134-139.

Brennan, R.G. and Link, T.M. 2007. Hfq structure, function and ligand binding. *Curr Opin Microbiol* **10**(2): 125-133.

Brescia, C.C., Mikulecky, P.J., Feig, A.L., and Sledjeski, D.D. 2003. Identification of the Hfq-binding site on DsrA RNA: Hfq binds without altering DsrA secondary structure. *Rna* **9**(1): 33-43.

Brown, L. and Elliott, T. 1996. Efficient translation of the RpoS sigma factor in Salmonella typhimurium requires host factor I, an RNA-binding protein encoded by the hfq gene. *J Bacteriol* **178**(13): 3763-3770.

Fender, A., Elf, J., Hampel, K., Zimmermann, B., and Wagner, E.G. 2010. RNAs actively cycle on the Sm-like protein Hfq. *Genes Dev* **24**(23): 2621-2626.

Geissmann, T.A. and Touati, D. 2004. Hfq, a new chaperoning role: binding to messenger RNA determines access for small RNA regulator. *Embo J* **23**(2): 396-405.

Hajnsdorf, E. and Regnier, P. 2000. Host factor Hfq of Escherichia coli stimulates elongation of poly(A) tails by poly(A) polymerase I. *Proc Natl Acad Sci U S A* **97**(4): 1501-1505.

Hengge-Aronis, R. 2002. Signal transduction and regulatory mechanisms involved in control of the sigma(S) (RpoS) subunit of RNA polymerase. *Microbiol Mol Biol Rev* **66**(3): 373-395, table of contents.

Hussein, R. and Lim, H.N. 2011. Disruption of small RNA signaling caused by competition for Hfq. *Proc Natl Acad Sci U S A* **108**(3): 1110-1115.

Jing, D., Agnew, J., Patton, W.F., Hendrickson, J., and Beechem, J.M. 2003. A sensitive two-color electrophoretic mobility shift assay for detecting both nucleic acids and protein in gels. *Proteomics* **3**(7): 1172-1180.

Lange, R., Fischer, D., and Hengge-Aronis, R. 1995. Identification of transcriptional start sites and the role of ppGpp in the expression of rpoS, the structural gene for the sigma S subunit of RNA polymerase in Escherichia coli. *J Bacteriol* **177**(16): 4676-4680.

Le Derout, J., Folichon, M., Briani, F., Deho, G., Regnier, P., and Hajnsdorf, E. 2003. Hfq affects the length and the frequency of short oligo(A) tails at the 3' end of Escherichia coli rpsO mRNAs. *Nucleic Acids Res* **31**(14): 4017-4023.

Lease, R.A. and Woodson, S.A. 2004. Cycling of the Sm-like protein Hfq on the DsrA small regulatory RNA. *J Mol Biol* **344**(5): 1211-1223.

Link, T.M., Valentin-Hansen, P., and Brennan, R.G. 2009. Structure of Escherichia coli Hfq bound to polyriboadenylate RNA. *Proc Natl Acad Sci U S A* **106**(46): 19292-19297.

Majdalani, N., Cunning, C., Sledjeski, D., Elliott, T., and Gottesman, S. 1998. DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. *Proc Natl Acad Sci U S A* **95**(21): 12462-12467.

Majdalani, N., Hernandez, D., and Gottesman, S. 2002. Regulation and mode of action of the second small RNA activator of RpoS translation, RprA. *Mol Microbiol* **46**(3): 813-826.

Majdalani, N., Vanderpool, C.K., and Gottesman, S. 2005. Bacterial small RNA regulators. *Crit Rev Biochem Mol Biol* **40**(2): 93-113.

Mikulecky, P.J., Kaw, M.K., Brescia, C.C., Takach, J.J., Sledjeski, D., and Feig, A.L. 2004. *Escherichia coli* Hfq has distinct interaction surfaces for DsrA, rpoS and poly(A) RNAs. *Nat Struct Mol Biol* **11**(12): 1206-1214.

Moller, T., Franch, T., Hojrup, P., Keene, D.R., Bachinger, H.P., Brennan, R.G., and Valentin-Hansen, P. 2002a. Hfq: a bacterial Sm-like protein that mediates RNA-RNA interaction. *Mol Cell* **9**(1): 23-30.

Moller, T., Franch, T., Udesen, C., Gerdes, K., and Valentin-Hansen, P. 2002b. Spot 42 RNA mediates discoordinate expression of the E. coli galactose operon. *Genes Dev* **16**(13): 1696-1706.

Nikulin, A., Stolboushkina, E., Perederina, A., Vassilieva, I., Blaesi, U., Moll, I., Kachalova, G., Yokoyama, S., Vassylyev, D., Garber, M. et al. 2005. Structure of Pseudomonas aeruginosa Hfq protein. *Acta Crystallogr D Biol Crystallogr* **61**(Pt 2): 141-146.

Olsen, A.S., Moller-Jensen, J., Brennan, R.G., and Valentin-Hansen, P. 2010. C-terminally truncated derivatives of Escherichia coli Hfq are proficient in riboregulation. *J Mol Biol* **404**(2): 173-182.

Repoila, F., Majdalani, N., and Gottesman, S. 2003. Small non-coding RNAs, co-ordinators of adaptation processes in Escherichia coli: the RpoS paradigm. *Mol Microbiol* **48**(4): 855-861.

Sauter, C., Basquin, J., and Suck, D. 2003. Sm-like proteins in Eubacteria: the crystal structure of the Hfq protein from Escherichia coli. *Nucleic Acids Res* **31**(14): 4091-4098.

Schumacher, M.A., Pearson, R.F., Moller, T., Valentin-Hansen, P., and Brennan, R.G. 2002. Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: a bacterial Sm-like protein. *Embo J* **21**(13): 3546-3556.

Sledjeski, D.D., Gupta, A., and Gottesman, S. 1996. The small RNA, DsrA, is essential for the low temperature expression of RpoS during exponential growth in Escherichia coli. *Embo J* **15**(15): 3993-4000.

Sonnleitner, E., Napetschnig, J., Afonyushkin, T., Ecker, K., Vecerek, B., Moll, I., Kaberdin, V.R., and Blasi, U. 2004. Functional effects of variants of the RNA chaperone Hfq. *Biochem Biophys Res Commun* **323**(3): 1017-1023.

Soper, T., Mandin, P., Majdalani, N., Gottesman, S., and Woodson, S.A. 2010. Positive regulation by small RNAs and the role of Hfq. *Proc Natl Acad Sci U S A* **107**(21): 9602-9607.

Soper, T.J. and Woodson, S.A. 2008. The rpoS mRNA leader recruits Hfq to facilitate annealing with DsrA sRNA. *RNA* **14**(9): 1907-1917.

Sun, X. and Wartell, R.M. 2006. Escherichia coli Hfq binds A18 and DsrA domain II with similar 2:1 Hfq6/RNA stoichiometry using different surface sites. *Biochemistry* **45**(15): 4875-4887.

Tsui, H.C., Leung, H.C., and Winkler, M.E. 1994. Characterization of broadly pleiotropic phenotypes caused by an hfq insertion mutation in Escherichia coli K-12. *Mol Microbiol* **13**(1): 35-49.

Updegrove, T., Wilf, N., Sun, X., and Wartell, R.M. 2008. Effect of Hfq on RprA-rpoS mRNA pairing: Hfq-RNA binding and the influence of the 5' rpoS mRNA leader region. *Biochemistry* **47**(43): 11184-11195.

Updegrove, T.B., Correia, J.J., Chen, Y., Terry, C., and Wartell, R.M. 2011. The stoichiometry of the Escherichia coli Hfq protein bound to RNA. *RNA*.

Valentin-Hansen, P., Eriksen, M., and Udesen, C. 2004. The bacterial Sm-like protein Hfq: a key player in RNA transactions. *Mol Microbiol* **51**(6): 1525-1533.

Vecerek, B., Rajkowitsch, L., Sonnleitner, E., Schroeder, R., and Blasi, U. 2008. The C-terminal domain of Escherichia coli Hfq is required for regulation. *Nucleic Acids Res* **36**(1): 133-143.

Waters, L.S. and Storz, G. 2009. Regulatory RNAs in bacteria. *Cell* **136**(4): 615-628.

Zhang, A., Altuvia, S., Tiwari, A., Argaman, L., Hengge-Aronis, R., and Storz, G. 1998. The OxyS regulatory RNA represses rpoS translation and binds the Hfq (HF-I) protein. *Embo J* **17**(20): 6061-6068.

Zhang, A., Wassarman, K.M., Ortega, J., Steven, A.C., and Storz, G. 2002. The Sm-like Hfq protein increases OxyS RNA interaction with target mRNAs. *Mol Cell* **9**(1): 11-22.

SUPPLEMENTARY MATERIAL

Binding of *rpoS*323$_{254\text{-}457}$ to wild-type and mutant Hfq

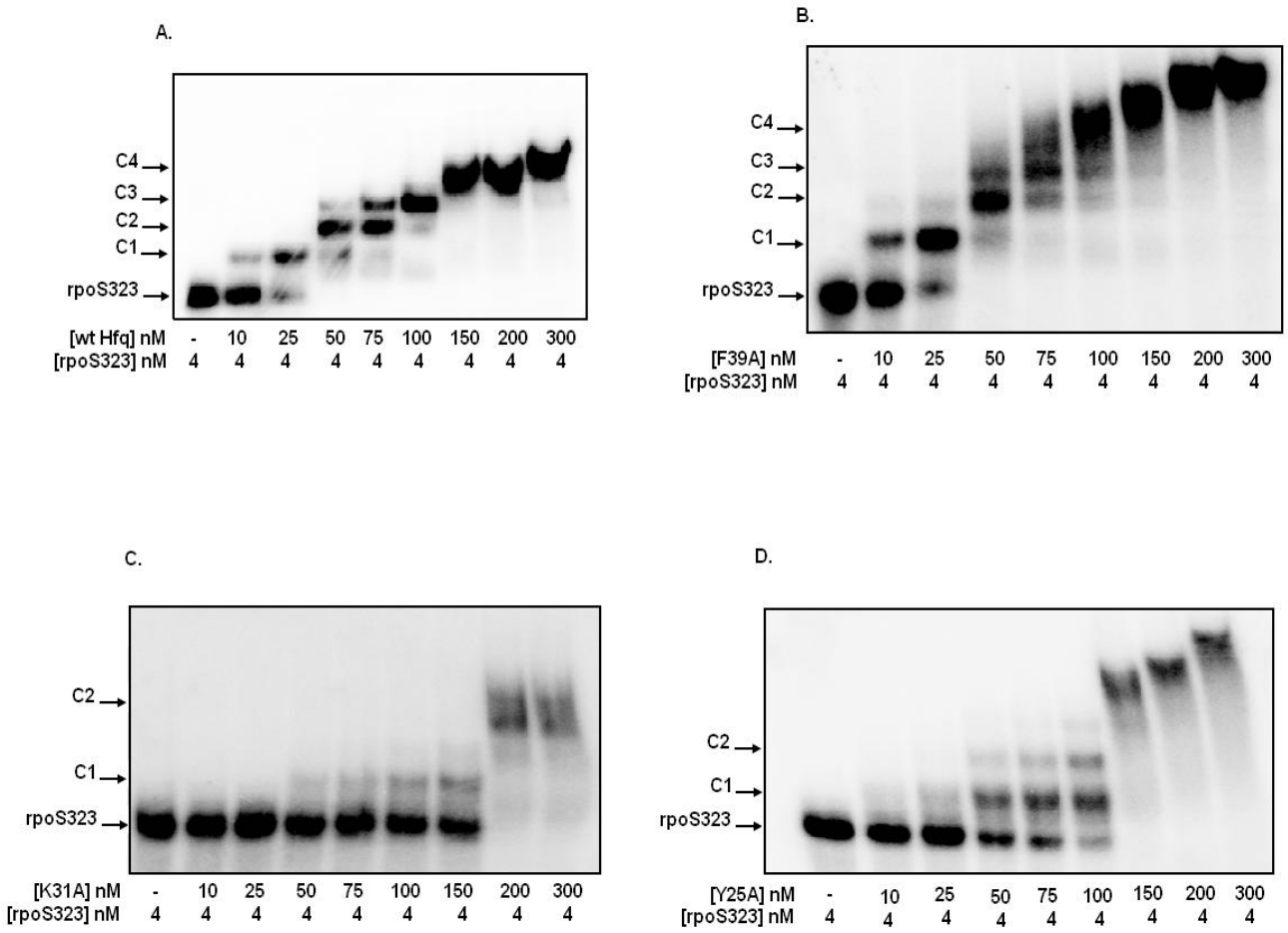


Figure S3.1: Binding of 4 nM ^{32}P-labeled *rpoS323*$_{254\text{-}457}$ to (**A.**) wt Hfq$_6$ and to (**B-D.**) three mutant Hfq$_6$ assessed by the gel shift assay.

Binding of DsrA to *rpoS*-L in the preseance of wild-type Hfq and Hfq-Flag

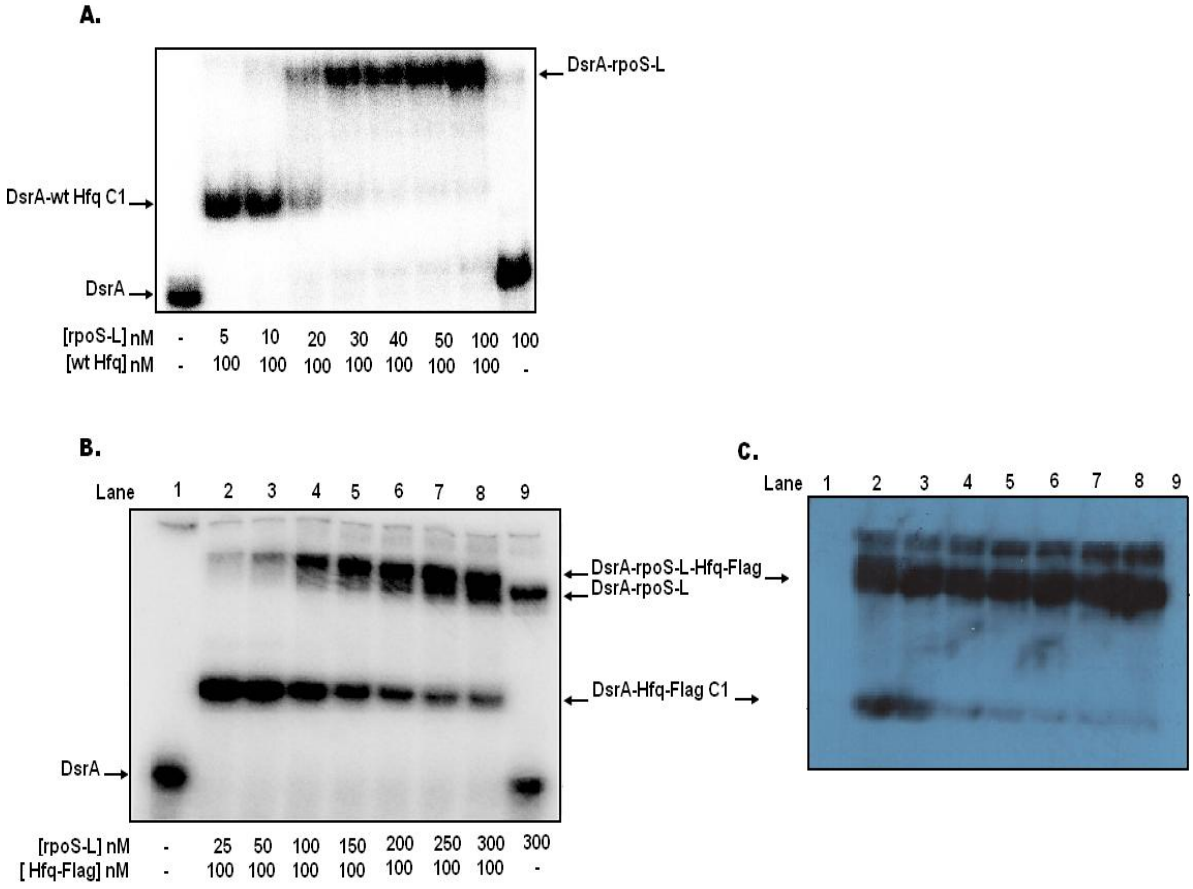


Figure S3.2: (**A.**) Binding of *rpoS*-L to 4 nM ^{32}P-labeled DsrA in the presence of 100 nM wt Hfq_6, (**B.**) Binding of *rpoS*-L to DsrA in the presence of 100 nM Hfq-Flag. (**C.**) Western blot of gel shown in **B.** using anti-Flag primary antibodies. Similar results were obtained using RprA in place of DsrA (data not shown).

Binding isotherms of DsrA to *rpoS*-L in the presence of mutant Hfq

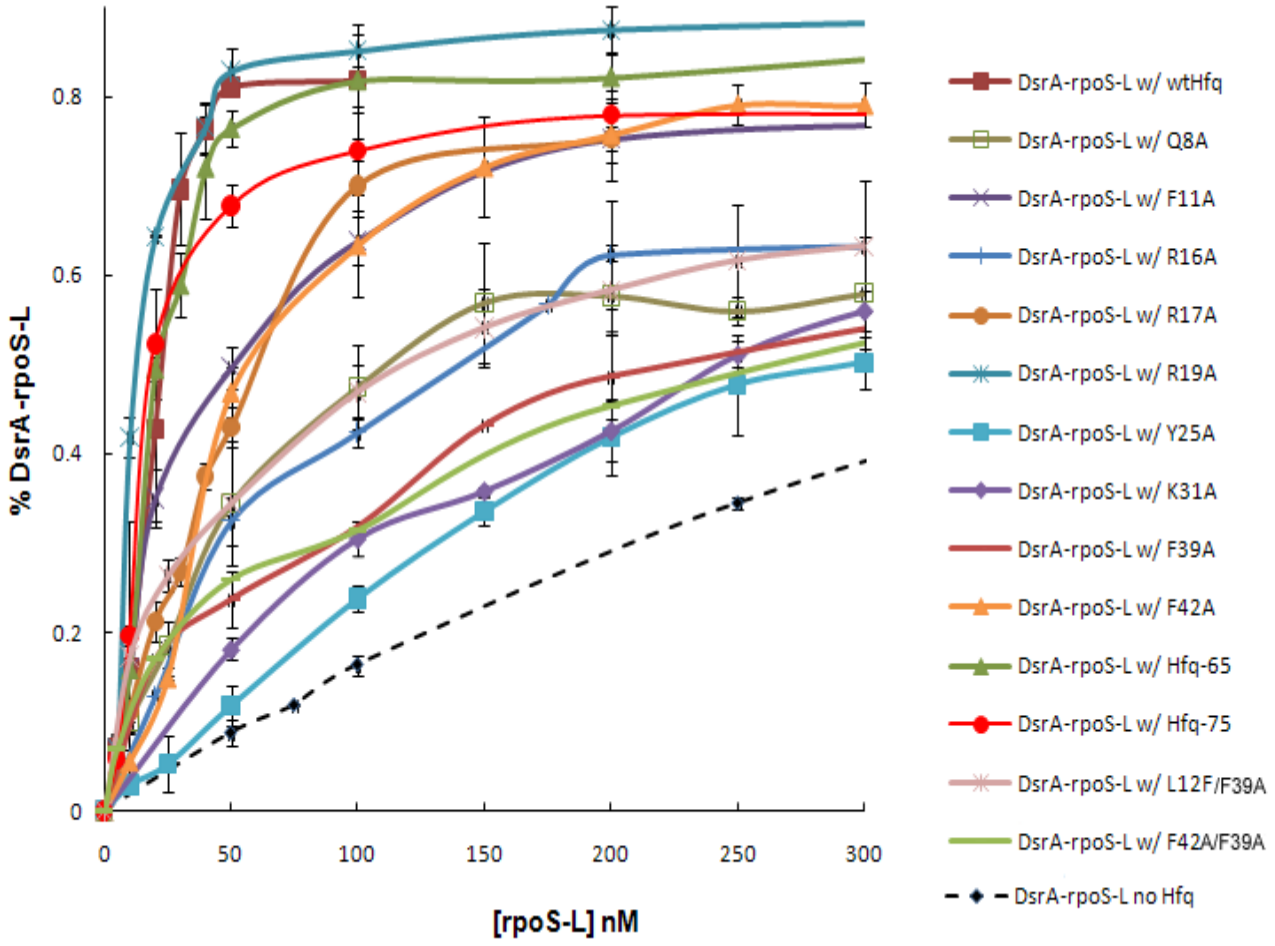


Figure S3.3: Binding isotherms of 4 nM ^{32}P-labeled DsrA with increasing amounts of *rpoS*-L in the presence of 100 nM wild-type and mutant Hfq. Each data point is the average of three values from independent experiments with error bars shown. Binding was assessed using the gel shift assay.

Model of RNA binding path along the proximal surface of Hfq

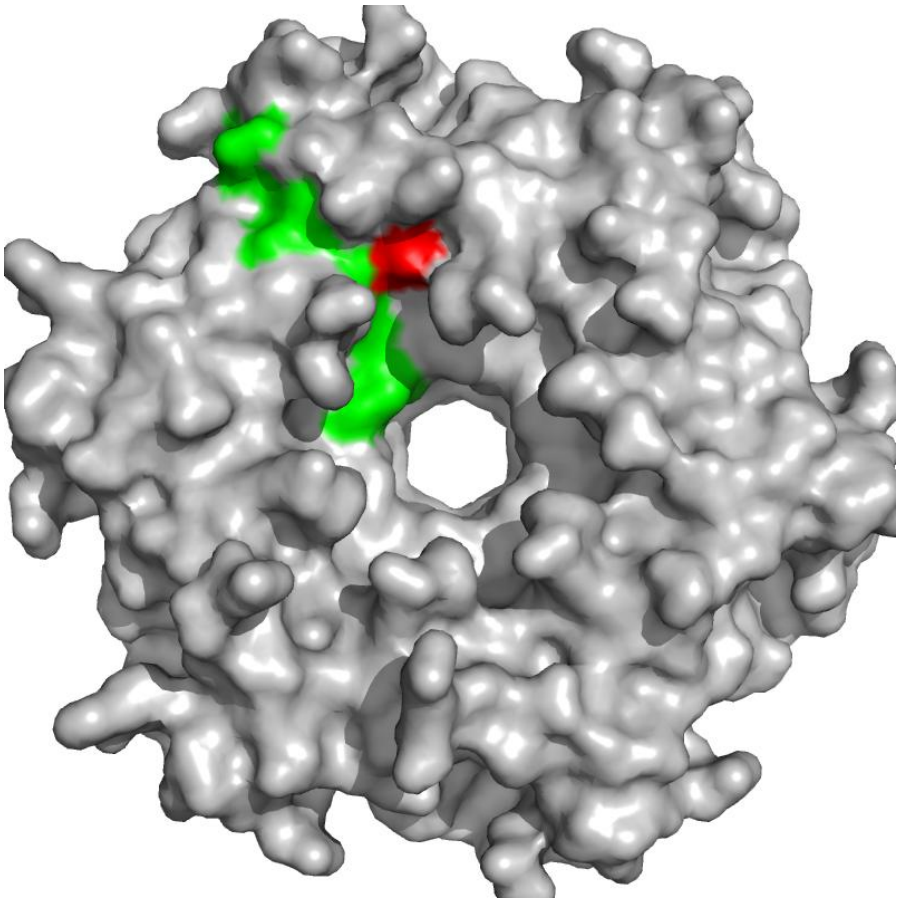


Figure S3.4: Representation of the proximal surface of *E. coli* Hfq hexamer with the residues Tyr55, Lys56, Phe39, and Arg16 in one subunit highlighted in green. The Gln8 residue is shown in red.

Model of *E. coli* Hfq hexamer with C-terminal domains

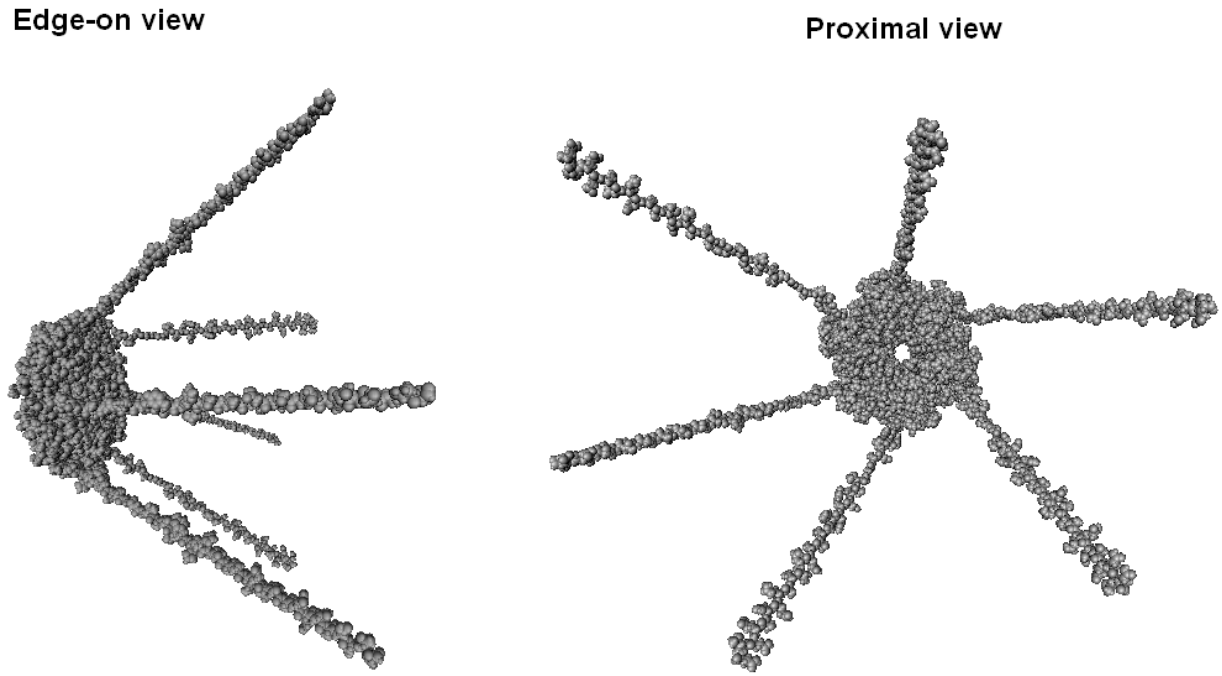


Figure S3.5: Space filling model of the toroidal part of *E. coli* Hfq_6 (Sauter *et al* (2003)) and a representation of the C-terminal domains. The (φ, ψ) torsion angles for residues 66 to 102 were initially assumed to be (-120^o, 130^o) and the structure was then energy minimized. The model illustrates the potential extent of these domains relative to the Hfq_6 toroid.

Binding of 200 nM *rpoS*-F to Hfq-65 and Hfq-75

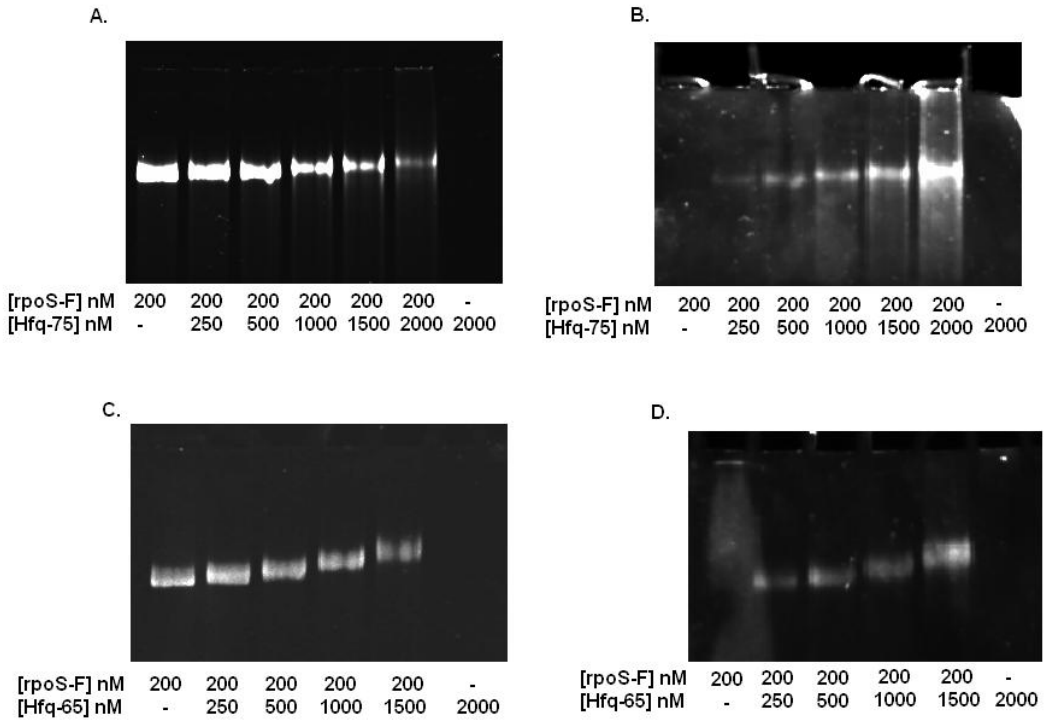


Figure S3.6: Gel assay in which Hfq-65 and Hfq-75 was added in increasing amounts to 200 nM *rpoS*-F RNA. Gels images **A** and **B** are the same 5% polyacrylamide gel of Hfq-75 and rpoS-F. The gel was first stained with SYBR Gold nucleic acid stain (**A**), and then stained using the SYPRO Ruby Red protein stain (**B**) using a method previously described (Jing *et al*. 2003). The latter stain is specific for protein and shows Hfq-75 in in the same position as the *rpoS*-F band. We note that Hfq-75 alone was added to the last lane but no band is observed implying the protein does not enter the gel by itself. Similar results are shown for Hfq-65 in gel images **C** and **D**. These results indicate Hfq-65 and Hfq-75 can bind *rpoS*-F without altering its gel mobility.

References

Sauter C, Basquin J, Suck D. 2003. Sm-like proteins in Eubacteria: the crystal structure of the Hfq protein from Escherichia coli. *Nucleic Acids Res 31*:4091-4098.

Jing D, Agnew J, Patton WF, Hendrickson J, Beechem JM. 2003. A sensitive two-color electrophoretic mobility shift assay for detecting both nucleic acids and protein in gels. *Proteomics 3*:1172-1180.

CHAPTER 4

The Stoichiometry of *Escherichia coli* Hfq protein bound to RNA

INTRODUCTION

The Hfq protein of *Escherichia coli* is an RNA binding protein and a key factor in post-transcriptional gene regulation (Valentin-Hansen et al., 2004; Majdalani et al., 2005; Brennan & Link, 2007; Waters & Storz, 2009). *E coli* Hfq and its bacterial homologues have been implicated in various facets of bacterial metabolism including stress induced sRNA regulation of mRNA translation as well as mRNA stability. In addition to its well documented interaction with RNA, Hfq has been found associated with DNA (Takada et al., 1997; Azam et al., 2000; Updegrove et al., 2010) as well as a number of proteins (Butland et al., 2005). The nature of Hfq's interactions with DNA and many of the proteins are not well understood; however, there is increasing recognition that they may reflect additional functions of Hfq (Le Derout et al., 2010).

Considerable attention has been focused on the role of Hfq in gene regulation by noncoding small RNAs (sRNA). A number of sRNAs, such as OxyS, SgrS, DsrA, RprA, Spot42, and Qrr1-4, require Hfq to facilitate their regulation of mRNA translation (Sledjeski et al., 2001; Majdalani et al., 2002; Moller et al., 2002a; Moller et al., 2002b; Zhang et al., 2002; Lenz et al., 2004; Kawamoto et al., 2006). *In vitro* studies suggest that Hfq's role is to enhance the association rate and/or stability of a sRNA to its mRNA target site near the start codon (Geissmann & Touati, 2004; Kawamoto et al., 2006; Soper & Woodson, 2008; Updegrove et al., 2008). The formation of a sRNA-mRNA hybrid can inhibit or enhance ribosome accessibility to mRNA thus providing either negative or positive regulation of translation (Majdalani et al.,

2005; Waters & Storz, 2009). Hfq's presence in the cell enhances sRNA stability and its capacity for functional interaction with mRNA targets. Hfq has also been shown to influence mRNA stability *in vivo* by enhancing sRNA-mRNA interaction or by binding mRNA directly (Tsui et al., 1997; Vytvytska et al., 1998; Masse et al., 2003; Morita et al., 2005).

In addition to its interactions with the translational initiation regions of mRNAs, Hfq also influences the stability of some mRNAs through its interaction with their 3' ends. It has been estimated that > 90% of the *E. coli* transcriptome possess posttranscriptionally added poly(A) tails (Mohanty & Kushner, 2006). Studies show that Hfq stimulates the addition of poly(A) tails to the 3' end of some mRNAs by poly(A) polymerase I (PAP) (Le Derout et al., 2003; Mohanty et al., 2004; Folichon et al., 2005). *In vivo,* inactivation of the *hfq* gene reduces the length of poly(A) tails synthesized at the 3' end of the *rpsO* mRNA by PAP, and *in vitro,* the addition of Hfq increases the processivity of PAP on *rpsO* mRNA. The addition of poly(A) tails has been shown to enhance mRNA decay in eubacteria (Steege, 2000). Studies also indicate that Hfq binding to poly(A) tails can prevent mRNAs from binding to enzymes involved in RNA degradation (Folichon et al., 2003; Mohanty et al., 2004; Folichon et al., 2005). Understanding the role of Hfq in the degradation of mRNAs requires understanding how Hfq binds to the 3' ends of mRNAs with poly(A) tails, as well as with PAP and possibly other RNA processing enzymes.

Initial studies on Hfq binding to RNA homopolymers and oligomers demonstrated that Hfq has a strong affinity for poly(A) and A_n oligomers with $n > 15$ (Carmichael et al., 1975; de Haseth & Uhlenbeck, 1980b). Studies on the binding of mutant Hfq to A_n oligomers indicated that the distal surface of the Hfq hexamer (Hfq_6) interacts with poly(A) sequences (Mikulecky et al., 2004; Sun & Wartell, 2006). A binding model proposed to accommodate information on the

complex (Brennan & Link, 2007), and a recent crystal structure of *E. coli* Hfq and A_{15} imply that the Hfq_6 forms a 1:1 complex with A_n oligomers. However, experimental studies employing several methodologies suggested different stoichiometries for Hfq and oligoriboadenylates. Isothermal titration calorimetry suggested one Hfq_6 bound to two A_{18} (Mikulecky et al., 2004), while fluorescence anisotropy, fluorescence quenching and a gel shift assay supported a model in which two Hfq_6 was bound to one A_{18} (Sun & Wartell, 2006).

DsrA is an 87- nucleotide (nt) sRNA that acts as a positive regulator for the translation of the stationary phase sigma factor RpoS. Hfq facilitates DsrA binding to the leader region of the *rpoS* mRNA and releases an inhibitory stem-loop that sequesters the Shine-Delgarno (SD) sequence (Cunning et al., 1998). Hfq binds both DsrA and *rpoS* mRNA with similar affinities (Soper & Woodson, 2008; Updegrove et al., 2008). Studies have explored the number of Hfq molecules binding to each RNA participant. Gel shift measurements yielded data supporting a 2:1 (Hfq_6:RNA) binding model for a 138 nt segment of *rpoS* mRNA, DsrA, (Lease & Woodson, 2004) and $DsrA_{DII}$ (Sun & Wartell, 2006), while isothermal titration calorimetry indicated a 1:1 complex for Hfq_6 binding to DsrA and a segment of *rpoS* mRNA (Mikulecky et al., 2004).

The ability of Hfq to stimulate sRNA-mRNA duplex formation has been observed under both *in vitro* and *in vivo* conditions. How Hfq recognizes and binds each of the RNAs and facilitates their pairing remains obscure. Evidence that Hfq can alter secondary and/or tertiary structure of some sRNAs and mRNAs lends support to the notion that Hfq acts as a chaperone and modulate the sRNA and/or mRNA structure, making one or the other RNA more amendable for heteroduplex formation. Another role ascribed to Hfq is an ability to bind and hold two pairing RNA molecules simultaneously, thus bringing them in close proximity and driving the reaction to favor sRNA-mRNA duplex formation. However, we note that the ability of Hfq to

separately bind two complementary RNAs is not always sufficient to promote RNA pairing (Arluison et al., 2007). Exactly how Hfq brings together two independent RNA molecules depends on the number of Hfq hexamers required to bind each RNA molecule and the number and type of RNAs that can simultaneously bind each Hfq hexamer. The stoichiometry of Hfq_6 binding to RNA is clearly pertinent to understanding the mechanism of how Hfq promotes ribo-regulation.

The focus of the current work was to determine the stoichiometry of the strong binding complexes of Hfq with A_{18} and $DsrA_{DII}$. The oligoriboadenylate A_{18} mimics the size and sequence of poly(A) tails at the 3' end of mRNAs, and results on how this oligonucleotide interacts with Hfq may be of functional significance in terms of Hfq's role and mechanism in facilitating polyadenylation by poly(A) polymerase. $DsrA_{DII}$, a 38 nt portion of DsrA (nucleotides 23-60), competes with DsrA for binding to Hfq (Brescia et al., 2003). It contains a stem loop and U rich segment of DsrA that binds Hfq. Mass spectrometry, fluorescence anisotropy, and analytical ultracentrifugation provide evidence supporting a 1:1 stoichiometry for Hfq_6 and oligo A_{18} as well as for Hfq_6 and $DsrA_{DII}$. A competition electrophoretic gel mobility shift assay also supports 1:1 complexes for Hfq_6 binding to A_{18} as well as to full length DsrA, RprA, and OxyS.

MATERIALS AND METHODS

Purification and characterization of wild-type and mutant Hfq

The Impact-CN intein system (New England Biolabs) was used to purify Hfq proteins as previously described (Sun & Wartell, 2006). The plasmids used to over express the Hfq proteins contained the *E. coli hfq* gene inserted into SapI-SmaI digested pTYB11 plasmid (pEcHfq) or mutant derivatives (see below). Protein purification was carried out according to the recommendation of the manufacturer using strain ER2566. Cell lysis was carried out using a french press. The cell lysate was centrifuged and the supernatant loaded onto a chitin column. The column was extensively washed with the lysis/wash buffer of 20 mM Tris (pH 8.3) and 1 M NaCl prior to incubation of the column with this buffer plus 40 mM dithiothreitol. The eluted protein was concentrated and buffer-exchanged to 0.5 M NaCl and 20 mM Tris at pH 8.3 using centrifugation filtration units.

To remove contaminating nucleic acids, Hfq preparations were subjected to a micrococcal nuclease treatment. 25 µl of 300 units/ml micrococcal nuclease (Worthington Biochemical Corporation) was added to 1 ml of 0.3 to 0.4 OD_{274nm} Hfq in 0.2 M NaCl, 20 mM Tris (pH 8.5) and 5 mM $CaCl_2$ and incubated at 37^{o}C for 45 min. This nuclease has a strict dependence on Ca^{2+}. 10 µL of 0.5 M Na_2EDTA was added and sample was washed and concentrated in 15 ml of 0.5 M NaCl and 20 mM Tris at pH 8.3 using 30 kD MWCO Amicon Utrafiltration cell.

The mutant Hfq protein, Hfq-65, was produced for this study from the plasmid pHfq-65 which was generated from pEcHfq using the QuikChange Mutagenesis Kit from Stratagene Inc (Sun & Wartell, 2006). Oligonucleotides employed placed a stop codon at position 66 of the *hfq*

gene: 5'- GCGATTTCTACTGTTGTC CCGTCTTAGCCGGTTTCTCATCACAG-3' and 5'-CTGTGATGAGAA ACCGGCTAAGACGGGAC AACAGTAGAAATCGC-3'. The plasmid construct was verified by DNA sequencing. The purification procedure for the mutant protein was similar to that used for wt Hfq. All proteins displayed expected molecular weights on a denaturing sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Concentrations were determined using an extinction coefficient of $\varepsilon = 2900\ M^{-1}\ cm^{-1}$ at 274 nm for the truncated protein and $4250\ M^{-1}\ cm^{-1}$ for wild-type Hfq (Gill & von Hippel, 1989). UV spectra showed absorbance ratios of A_{275nm}/A_{255nm} (peak to valley) of 1.8 or higher. Analysis of the spectra indicated less than 5 % contaminating nucleic acids (Sun & Wartell, 2006).

RNA synthesis and purification

The following RNAs were purchased commercially (Integrated DNA Technologies, Inc) and purified by HPLC: $DsrA_{DII}$ (AACGAAUUUUUUAAGUGCUUCUUGCUUAAGCAAGUUUC), OxyS-18 (GAAUAACUAAAGCCAACG) and A_{18}. $DsrA_{DII}$ and A_{18} were also purchased with 6-carboxyfluorescein (FAM) linked to their 5' end. The full length DsrA, OxyS, and RprA RNAs were cloned as described previously and transcribed using a T7 MEGAscript High Yield RNA transcription kit (Ambion®) (Updegrove et al., 2008). They were ^{32}P-labeled at their 5' end using standard phosphatase and kinase reactions and purified by gel extraction (Sambrook & Russell, 2001).

Mass spectrometry and cross-linking of Hfq to RNA

20 μL samples were prepared by adding Hfq to fixed amounts of A_{18}, $DsrA_{DII}$, or OxyS-18 in phosphate binding buffer (0.2 M NaCl, 20 mM Na_2HPO_4, pH 7.8). Concentrations are described in Results. For the Hfq-A_{18} mixture 10 μL of 0.2 M EDC (1-ethyl-3-3-dimethylaminopropyl carbodiimide hydrochloride, Pierce, Rockford, IL) was added and allowed to react for 4 hr at room temperature. For the other Hfq-RNA mixtures 2μL of a 3 % formaldehyde solution was added and allowed to react for 15 minutes at room temperature. 1 μl of 3 M glycine (in water) was then added to quench the reaction (Niranjanakumari et al., 2002). 20 μL of the Hfq-RNA solutions described above were then concentrated to 3 μL with a C4 ZipTip (Millipore, Billerica, MA) and then mixed with 3 μL of matrix solution. The matrix solution was prepared by adding 20 mg of sinapinic acid and 50 mg ammonium citrate in 500 μL of 18 MΩ deionized water. 1 μL of analyte-matrix mixture was then deposited onto a 100-well stainless steel MALDI plate. The MALDI-MS experiments were performed using a Voyager DE STR MALDI-TOF mass spectrometer (Applied Biosystems, Foster City, CA) equipped with a 337 nm N_2 laser (3 Hz). The accelerating voltage, grid voltage, and delay time were typically 25 kV, 91%, and 1500 ns, respectively. The laser intensity was checked daily to obtain the best signal-to-noise ratio. Mass spectra were obtained by averaging 10-50 laser shots.

Analytical ultracentrifugation: sedimentation velocity and equilibrium

Sedimentation studies were performed in a Beckman Optima XLA analytical ultracentrifuge equipped with absorbance optics and an An60 Ti rotor at 19.7^{o}C. Temperature was calibrated as described previously (Liu & Stafford, 1995). Velocity data were typically collected at the appropriate speeds using 274 nm for Hfq and 495 nm for FAM-A_{18} at a spacing

of 0.01 cm with one flash at each point in a continuous-scan mode. When collecting data at multiple wavelengths, care must be taken to collect data at peaks to avoid dramatic signal variations due to wavelength uncertainty (+/- 4 nm) with the XLA. All experiments were initially analyzed with Sedfit to produce c(s) distributions (Schuck et al., 2002) and with DCDT^{+2} to produce g(s) distributions and weight average S value (Philo, 2006). Direct boundary fitting of velocity data to discrete models can also be performed with the program Sedanal (Stafford & Sherwood, 2004). Analysis with Sedanal requires input of MW, extinction coefficients, and density increments (typically estimated from 1-vbar*rho values). The buffer solution density was estimated in Sednterp to be 1.01920 gm/ml at 19.7^{o}C. The vbar of Hfq was estimated with Sednterp (Laue, 1992) to be 0.7248. The vbar of FAM-A_{18} is assumed to be 0.55. The extinction coefficient of FAM-A_{18} at 495 nm is 75,000 $M^{-1}cm^{-1}$ or using a molecular weight of 6113 D, 12.269 ml/mg/cm. The extinction coefficient of Hfq at 274 nm is 0.400 ml/mg/cm (Stafford & Sherwood, 2004). Parameter uncertainty is calculated with an Fstat routine within Sedanal at the 95% confidence interval and reported in a <, > format.

Hfq alone (at 2, 4 and 8 μM) or mixed at a 1:1 ratio with FAM-A_{18} was spun at 19.7^{o}C and at 12K, 16K and 20K in six channel double sector cells. Data on Hfq alone was collected at 274 nm. Data with mixtures of Hfq and FAM-A_{18} were collected at 495 nm. Equilibrium at each speed was judged with the software utility WinMATCH (http://www.biotech.uconn.edu/auf/?i=aufftp). This program makes a least-square comparison of successive scans to establish that equilibrium has been achieved. Values for density, vbar, and extinction coefficients were as described under Sedimentation Velocity measurements. Nine data sets from three concentrations and three speeds were best fit to a single species model using Sedanal. Molecular weight uncertainty is calculated with Fstat as described above.

Polyacrylamide gel electrophoresis and gel mobility shift assay

Binding reactions of Hfq and FAM-A_{18} were carried out in the phosphate binding buffer (0.2 M NaCl, 20 mM Na_2HPO_4 pH 7.8). wt Hfq, Hfq-65 or both were added to FAM-A_{18} and the reactions allowed to equilibrate at 25^{o}C for 10 min prior to the addition of 3.2 µl of gel loading buffer (100 mM Tris-Cl (pH 6.8), 4 % (w/v) SDS, 0.2 % (w/v) bromophenol blue, 20 % (v/v) glycerol). Final reaction volumes were 20 µl and contained 0.6 % SDS. The SDS was added in order to enhance the negative charge of the Hfq •A_{18} complexes and enable them to migrate into the gel prior to the free A_{18} running out the bottom. The concentration of Hfq (moles hexamer) in each reaction varied between 2 to 3 µM, and the concentration of FAM-A_{18} was 1 µM. The total reaction volumes were electrophoresed into a 6 % polyacrylamide (29:1) gel with 4% glycerol that was layered onto a 2.5 cm bottom plug consisting of 15 % polyacrylamide (29:1). The latter was employed to slow and retain the free A_{18}. The gel was 20 cm x 20 cm x 1.5 mm. Electrophoresis was conducted at 120V at 4^{o}C using 0.5 X TBE buffer for approximately 8 hrs. Analysis of the gels used excitation and emission wavelengths of 473 and 520 nm, respectively, of the Fujifilm Image Reader FLA-3000.

Similar competition gel assays were carried out in which wt Hfq, Hfq-65, or both were bound with ^{32}P- labeled DsrA, RprA or OxyS in 15 µL binding solution (50 mM NaCl, 50 mM KCl, 100 mM NH_4Cl ,20 mM Tris-HCl (pH 8.0), 4 % glycerol). The indicated amounts of Hfq were added to the indicated amount of sRNA and the reaction allowed to equilibrate at 25^{o}C for 10 min prior to running on a 8 % polyacrylamide (29:1) gel with 3% glycerol. Electrophoresis was conducted at 120V at 4^{o}C using 0.5 X TBE buffer for approximately 2 hrs. Imaging and analysis of the gels were made using the Fujifilm Image Reader FLA-3000.

Fluorescence anisotropy measurements

Fluorescence anisotropy measurements of Hfq binding to FAM-A_{18} were carried out at room temperature in the 0.5 M NaCl and 20 mM Tris (pH 8.3) solvent as previously described (Sun & Wartell, 2006). The L-format was employed with the excitation monochromator at 490 nm and emission monochromator at 522 nm. Anisotropy values were obtained from the average of 10 iterations using an integration time of 4 to 8 s for each measurement depending on FAM-A_{18} concentration. The slits employed were set at 1 or 2 mm. Wild-type Hfq was serially titrated into fluorescence cells with a working volume of 1 ml or 0.5 ml for FAM-A_{18} at 2 nM. When 5 μM of FAM-A_{18} was employed a 50 μL micro-cell was employed. The fluorescence intensity of FAM-A_{18} showed a small decrease with Hfq binding after accounting for dilution (~2%). Similar anisotropy experiments were carried using $DsrA_{DII}$ with FAM attached to its 5' end. The solvent employed for the FAM-$DsrA_{DII}$ experiments was 0.1 M NaCl and 20 mM Tris (8.3) since Hfq affinity for $DsrA_{DII}$ increased with decreasing salt concentration and conditions favoring strong binding were sought (unpublished data). Unlike FAM-A_{18}, Hfq binding decreased the fluorescence intensity of FAM-$DsrA_{DII}$ indicating that the quantum yield of the bound fluorophore was less than the free molecule. The ratio of quantum yield for bound vs. free FAM-$DsrA_{DII}$, Q_b/Q_f, was determined to be 0.70 by saturating FAM- $DsrA_{DII}$. The change in anisotropy was corrected for this factor (Lundblad et al., 1996).

Analysis of fluorescence anisotropy data

The two models employed in the analysis of Hfq binding to FAM-A_{18} at low concentration (nM) were described in Sun and Wartell (Sun & Wartell, 2006). Both assume that Hfq exists only as hexamers. The first model assumes a one to one complex forms between the

Hfq hexamer and FAM-A_{18}. An equation describing the fluorescence anisotropy in terms of the dissociation constant K_d and other parameters of the experiment can be derived (Lundblad et al., 1996) and is given by eq. 1.

$$A = A_f + (A_b - A_f)\,[\beta - [\beta^2 - 4\,R_t\,P_t]^{1/2}]\,/2R_t \qquad (1)$$

where $\beta = R_t + P_t + K_d$. A is the measured anisotropy of FAM-A_{18} during the titration, A_f and A_b are the anisotropy of the free and bound FAM-A_{18} respectively, and R_t and P_t are the total concentrations of FAM-A_{18} and Hfq hexamer respectively. Non-linear least squares fit of the equation to data was made. For a situation where binding quenches the fluorescence of the RNA (i.e., $DsrA_{DII}$) eq. 1 has to be corrected for the difference in quantum yields for free and bound RNA (Q_f, Q_b). Defining $\alpha = [\beta - [\beta^2 - 4\,R_t\,P_t]^{1/2}]\,/2R_t$ one obtains

$$A = [\,A_f + (\,A_b(Q_b/Q_f) - A_f\,)\alpha]\,/\,[1 - (1 - (Q_b/Q_f)\,)\alpha\,] \qquad (2)$$

The second model assumed that Hfq hexamers bind FAM-A_{18} in a two-step reaction. The binding reaction is described by a dissociation constant K_1 for binding the first Hfq hexamer, and a dissociation constant K_2 for binding a subsequent Hfq hexamer;

$$R + P \leftrightarrow RP \qquad \text{with} \quad K_1 = (\,R\,)(\,P\,)/RP \qquad (3a)$$

$$RP + P \leftrightarrow RP_2 \qquad \text{with} \quad K_2 = (\,RP\,)(\,P\,)/RP_2 \qquad (3b)$$

P corresponds to Hfq hexamer and R is FAM-A_{18}. Data was fit to the second model using the BIOEQS program (Royer & Beechem, 1992; Royer, 1993). This algorithm performs a non-linear least squares fit of eq. 3 to the anisotropy data using parameters corresponding to the standard state free energies related to K_1 and K_2 , anisotropies of free RNA, RNA in the RP complex, and RNA in the RP_2 complex. The anisotropy of the free FAM-A_{18} was fixed to the experimental

value, and the remaining four parameters fit to the data. Supplementary information with Figures S4.1 to S4.3 is available upon request.

RESULTS

MALDI-TOF mass spectroscopy indicates Hfq_6 forms a 1:1 complex with $DsrA_{DII}$, A_{18} and OxyS-18

Matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectrometry was first used to examine the molecular mass of *E. coli* Hfq alone and then as a complex with A_{18}. These experiments were done in the absence of cross linking as well as after EDC cross-linking of the Hfq-A_{18} complex prior to mass spectrometric analysis. The MALDI-TOF spectrum of Hfq shown in Figure 4.1A was carried out with EDC crosslinking, and reveals discrete ions with m/z ratios corresponding to the Hfq monomer and multimers up to the hexamer (67,060 D, theoretical mass 66,998 D). This observation is in agreement with a previous study (Moller et al., 2002a) and illustrates that Hfq can stably exist as multimers up to the hexamer in the laser desorption ionization process. We note that macromolecules are generally expected to be singly charged ions in MALDI-TOF experiments (Karas et al., 2000).

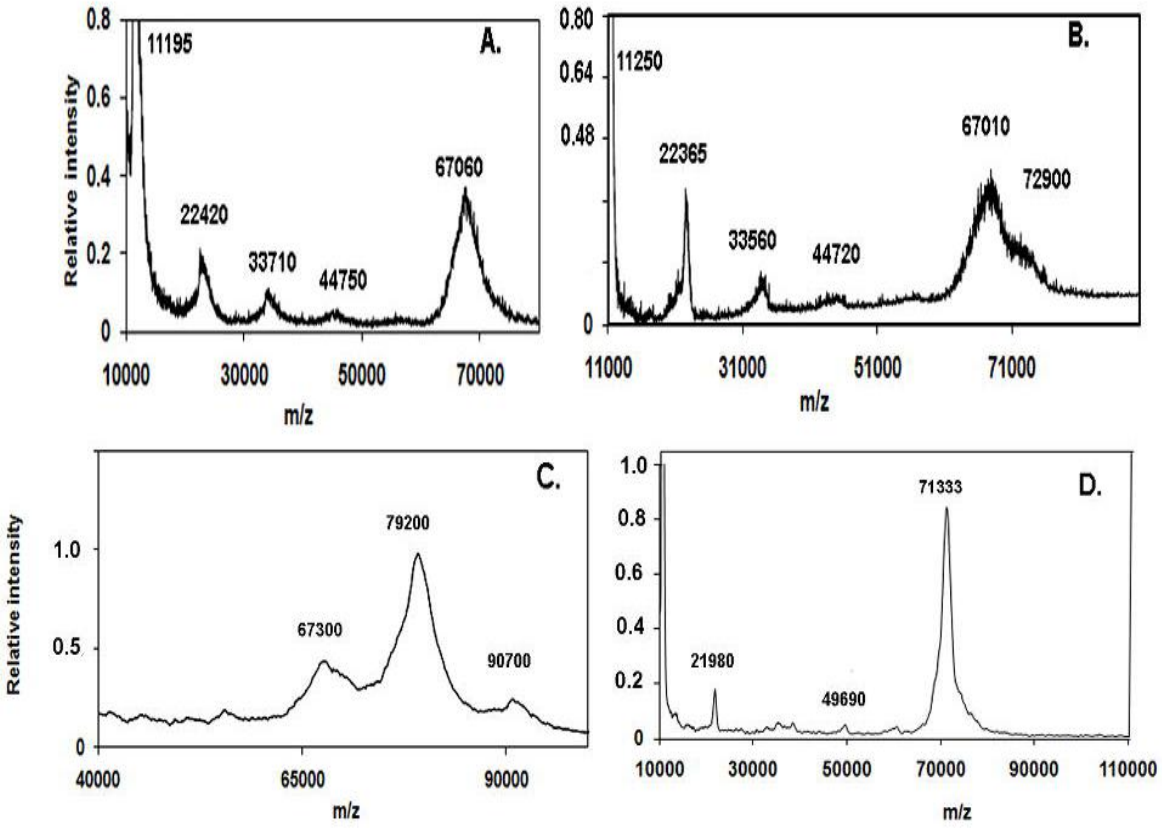


Figure 4.1: MALDI-TOF m/z spectra of 2 μM Hfq_6 (**A.**), 2 μM Hfq_6 and 0.7 μM A_{18} (**B.**), 8 μM Hfq_6 and 4 μM DsrA domain II (**C.**), and, 4 μM Hfq_6 and 2 μM OxyS-18 (**D.**). All samples were prepared in the 0.2 M Na^+ solvent and matrix solution as described in Materials and Methods.

The addition of 0.7 μM A_{18} to 2 μM Hfq_6 resulted in the formation of an additional peak corresponding to a molecular mass of 72,900 D (Figure 4.1B). Since the theoretical mass of A_{18} is 5,840 D, this new peak is very close to an expected complex with a 1:1 ratio of Hfq_6 to A_{18} (theoretical mass 72,839 D). No peaks were observed at the molecular mass corresponding to 2:1 or 1:2 ratios of Hfq_6 to A_{18}. Similar results were also obtained when 0.07 μM A_{18} and 0.2 μM Hfq_6 were employed with and without EDC cross-linking (data not shown). The addition of

EDC increased the relative signal intensities of Hfq_6 and $Hfq_6 \bullet A_{18}$ complex over the Hfq subunit multimers, consistent with suppression of hexamer dissociation.

The Hfq_6-$DsrA_{DII}$ complex required a more robust cross-linking agent to withstand the conditions imposed by the MALDI-TOF experiment. Formaldehyde proved to be an efficient cross-linker and allowed detection of the Hfq_6- $DsrA_{DII}$ complex. Figure 4.1C shows a spectrum resulting from a mixture of 8 µM Hfq_6 with 4 µM $DsrA_{DII}$. A pronounced peak occurs at a m/z ratio of 79,200 flanked by less pronounced peaks of 67,300 and 90,700. Since the theoretical molecular weight of $DsrA_{DII}$ is 12,031 D, the large middle peak is consistent with one Hfq_6 bound to one $DsrA_{DII}$. The smaller and larger molecular weight peaks are consistent with Hfq_6 and one Hfq_6 bound to two $DsrA_{DII}$ molecules, respectively. $DsrA_{DII}$ has been shown to form two bands at low µM concentrations in a polyacrylamide gel environment (Sun & Wartell, 2006). When 2 µM Hfq_6 was added to 1 µM $DsrA_{DII}$, only the 79,000 and 67,000 m/z peaks were observed (data not shown). Unfortunately a MALDI-TOF experiment with full length DsrA and Hfq gave weak or negligible signals barely above background at the m/z ratio expected for Hfq_6•DsrA or higher masses. The larger negative charge intrinsic to the full length DsrA molecule appears to compromise a study of this complex by this method.

OxyS is a 109 nt sRNA that was shown to bind Hfq *in vitro* and *in vivo*, and acts as a negative regulator for the translation of the *rpoS* mRNA. A 18-nt portion of OxyS sRNA that spans nucleotides 64 to 81 is thought to be critical for Hfq binding based on the observation that an oligonucleotide complementary to this region strongly inhibits Hfq from binding to the full-length OxyS molecule (Zhang et al., 2002). MALDI-TOF was used to assess the stoichiometry of Hfq binding to this segment of OxyS. When 4 µM of Hfq_6 was added to 2 µM OxyS-18 and formaldehyde used as the cross-linking agent, only one extremely large peak was observed at an

m/z ratio of 71333 (Figure 4.1D). With the theoretical molecular weight of OxyS-18 being 5769.6 D, the large peak in Figure 4.1D is in good agreement with one Hfq hexamer bound to one OxyS-18. No peak was detected at an m/z ratio corresponding to either 1:2 or 2:1 Hfq_6 to OxyS-18 stoichiometry.

Analytical ultracentrifugation analysis of Hfq•A_{18} complex in solution

Analytical ultracentrifugation analysis was employed to determine the stoichiometry of the Hfq-A_{18} complex in aqueous solution. Sedimentation velocity of Hfq alone in 0.5 M NaCl and 20 mM Tris (8.2) indicated a single major species with a sedimentation coefficient (s) of s = 3.42 sec <3.41, 3.44> and no more than 2 % of a higher molecular weight aggregate with s = 5.56 S. Figure 4.2A shows the results of a sedimentation velocity experiment of Hfq analyzed using the c(s) method (Schuck et al., 2002). The sedimentation coefficient distribution was independent of loading concentrations from 3.2 to 12.1 μM Hfq in mole hexamer. Direct boundary fitting of the sedimentation velocity data using SedAnal (Stafford & Sherwood, 2004) indicated a molecular weight for the 3.42 S species of 64,815 D <59,733, 70,301>. This value is slightly lower than the expected value of 66,998 D and is consistent with the hexamer being the dominant Hfq species at these concentrations. The slightly lower than expected value can be explained by uncertainty in the partial specific volume employed or the influence of the minor aggregate on the fit. (Traces of sed. velocity run and model fitting using SedAnal are given in Figure S4.1).

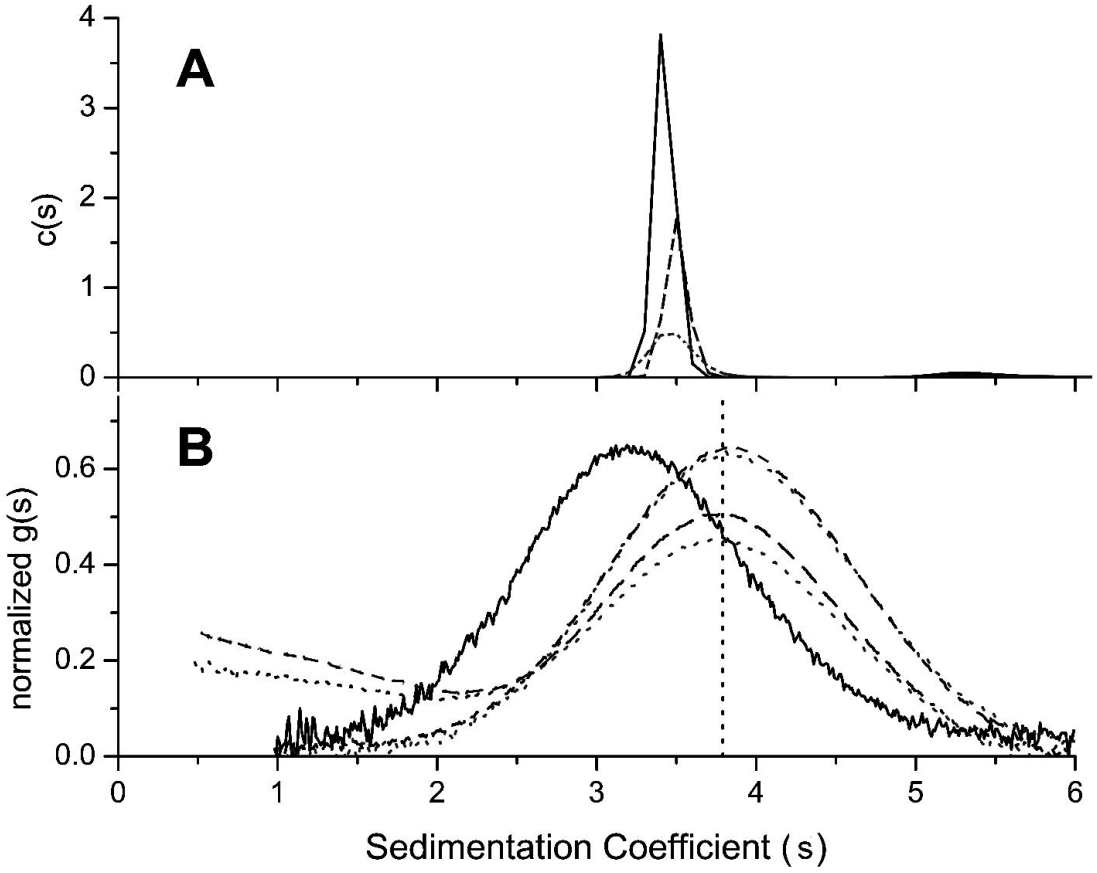


Figure 4.2: (A.) Sedimentation coefficient distribution c(s) determined by program Sedfit for three different Hfq_6 concentrations, 3.2, 5.7, and 12 μM shown as dotted, dashed, and solid lines, respectively. The average integrated value for these data is 3.51 ± 0.03 S. (**B.)** Sedimentation coefficient distribution displayed as normalized g(s) for Hfq_6 at 6.9 μM alone as solid line (average integrated value for Hfq data is 3.42 ± 0.04 S), and with 4.4 μM and 8.3 μM FAM-A_{18} added. Upper pair of overlapping dashed and dotted lines show the 6.9:4.4 mixture evaluated by absorbance at 274 nm and 495 nm, respectively. Lower pair of dashed and dotted lines display the 6.9:8.3 mixture evaluated at the same two wavelengths. Lack of alignment of c(s) and g(s) peaks for Hfq_6 alone (3.51 S vs 3.42 S) is attributed to minor components affecting the main c(s) peak.

Figure 4.2B shows the normalized g(s) distribution of concurrently run sedimentation velocity experiments which examined 6.9 μM Hfq_6 alone, 6.9 μM Hfq_6 with 4.4 μM FAM-A_{18}, and 6.9 μM Hfq_6 with 8.3 μM FAM-A_{18}. FAM-A_{18} binding increased the sedimentation coefficient of Hfq_6 from 3.3S to 3.9S. At the concentration ratio of [FAM-A_{18}]/[Hfq_6] of 1.2, a

trailing boundary of excess FAM-A_{18} is observed. Free FAM-A_{18} has a sedimentation coefficient of 1.355 S <1.345, 1.364> with no evidence of concentration dependence or additional species (data not shown). Using the SedAnal software a good fit to the Hfq-A_{18} data was obtained with a model that assumed Hfq hexamer binds A_{18} with a 1:1 stoichiometry. The best Sedanal fit (constraining S for Hfq and FAM-A_{18}) returned a K of 1.71×10^7 M^{-1} <0.85, UB>. The unbounded upper limit means all larger values of K are indistinguishable in the least squares sense. A subsequent run with a new protein sample returned a larger K ~10^{10} M^{-1} with 95% confidence limits of <7.5×10^7, UB>. Thus the data is consistent with a tight 1:1 Hfq-A_{18} complex with an affinity in excess of $10^7 M^{-1}$.

Sedimentation equilibrium runs of 2, 4, and 8 μM Hfq_6 alone and mixed with 1:1 molar ratios of FAM-A_{18} confirmed that the stoichiometry of the Hfq_6 • FAM-A_{18} complex in solution is not 2:1, but 1:1. The evaluated molecular weight of Hfq alone was 61.475 kD <58.8, 64.2> (rms = 0.00596) (Figure S4.2) similar to the value obtained from sedimentation velocity analysis. Analysis of the sedimentation equilibrium data of the Hfq •FAM-A_{18} mixtures, monitored at the FAM-A_{18} absorbance peak of 495 nm, yielded a molecular weight of 68929 D <67.4, 70.4> (rms = 0.00724) (Figure S4.2). This clearly does not correspond to a complex consisting of 2 Hfq_6 molecules and one A_{18} molecule but is consistent with a 1:1 complex.

Gel mobility shift study of wild-type Hfq and Hfq-65 binding to A_{18} and other RNAs

Previous gel mobility experiments in which A_{18} or other RNAs were titrated with Hfq at concentrations above apparent K_d values indicated 2:1 Hfq_6 to RNA stoichiometry (Lease & Woodson, 2004; Sun & Wartell, 2006; Updegrove et al., 2008) . Since these previous results conflict with the above findings, we examined the stoichiometry of Hfq•A_{18} complexes in the gel

environment using a different approach that relies on a qualitative comparison rather than quantative analysis of band intensities. The Hfq•A_{18} complexes that formed in the presence of wt Hfq and Hfq-65 were determined. Hfq-65 is a truncated variant of wt Hfq consisting of 65 residues from the N-terminal end. This truncated Hfq was previously shown to bind DsrA two to three-fold less well than wt Hfq, and to A_{27} with an affinity similar to wt Hfq (Vecerek et al., 2008). Lane 3 of Figure 4.3A shows the gel-shift of the Hfq-65•A_{18} complex in a 6% PAG. The Hfq-65•A_{18} complex migrates with a slower mobility than the wt Hfq•A_{18} complex (lane 2) in spite of its reduced size. A plausible explanation of this phenomenon is the increased positive charge of Hfq-65 compared to wt Hfq. Hfq-65 has four less negatively charged residues (Asp 97, Glu 99, Glu 100, and Glu 102) and one less positively charged residue (Arg 66) than each wt Hfq subunit. When equimolar amounts of wt Hfq and Hfq-65 were mixed with A_{18} for 5 minutes and run into the gel, two bands were observed corresponding to wt Hfq•A_{18} and Hfq-65•A_{18} (lane 4, Figure 4.3A). This result is consistent with a 1:1 stoichiometry for complexes of Hfq_6 and A_{18}. If the stoichiometry of the Hfq•A_{18} complexes were two Hfq_6 and one A_{18} a band of intermediate mobility would be expected in lane 4. Changing the ratio of wt Hfq and Hfq-65 concentrations altered the intensity of the two bands in direct proportion, but no additional band is observed (lanes 5, 6, Figure 4.3).

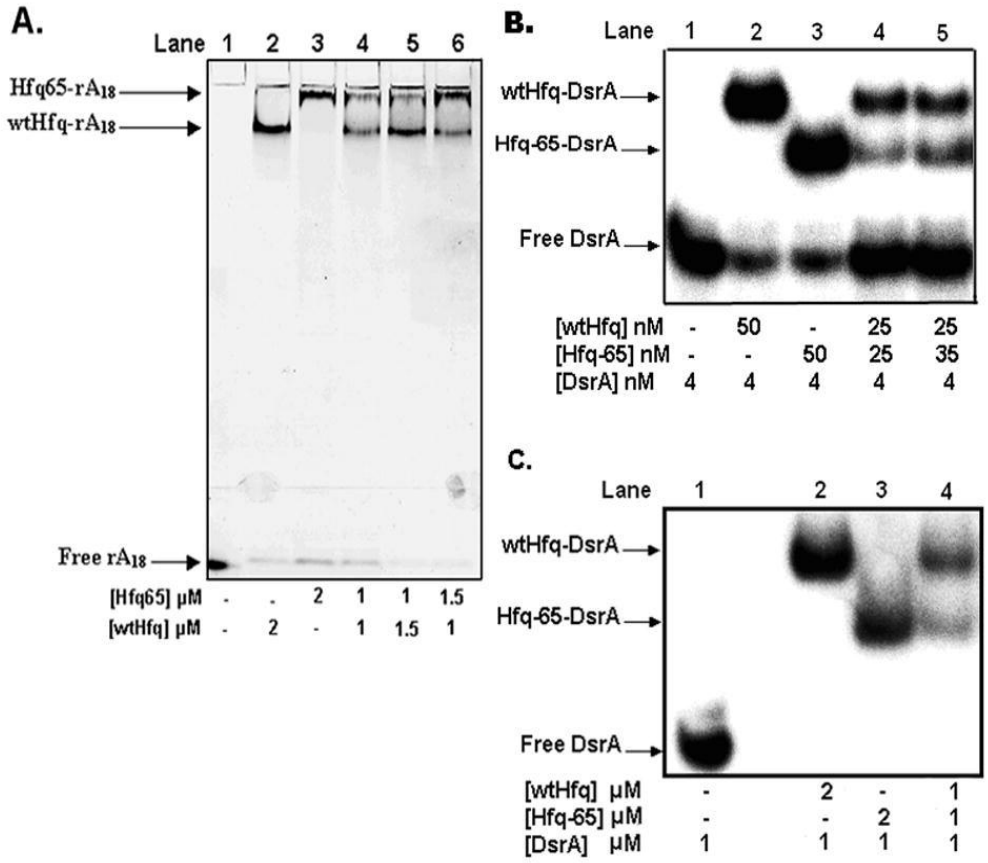


Figure 4.3: DsrA and A_{18} bind both wt Hfq and Hfq-65 in a 1:1 stoichiometry. Varying concentrations in moles hexamer/L of wt Hfq and Hfq-65 were added to 1 µM FAM-A_{18} (**A.**), 4 nM ^{32}P-DsrA **(B.**), and 1 µM ^{32}P-DsrA (**C.**). Similar results were obtained when ^{32}P end labeled RprA and OxyS sRNAs were added to both wt Hfq and Hfq-65.

When 4 nM ^{32}P-labeled DsrA was added to 50 nM of either wt Hfq or Hfq-65 (moles hexamer), most of the RNA was shifted to a slower moving complex. Under these conditions the DsrA•Hfq-65 complex migrates faster than the DsrA•wt Hfq complex (lanes 2, 3, Figure 4.3B). Since DsrA has considerably more negative charge than $A_{18,}$ it will likely dominate the charge differences between wt Hfq and Hfq-65. The size difference between wt Hfq and Hfq-65 rather than their intrinsic charge difference appears to be the governing factor in the migration of these Hfq•DsrA complexes. When 25 nM wt Hfq and 25 nM Hfq-65 (moles hexamer) were added to 4 nM DsrA, only two apparent slow migrating bands were evident; one corresponding to the

DsrA•wt Hfq complex and the other corresponding to the DsrA•Hfq-65 complex (lane 4, Figure 4.3B). Similarly, when 1 µM each of wt Hfq and Hfq-65 was added to 1 µM DsrA only two slow migrating bands were observed (Lane 3.4, Figure 4.3C). The outcome was the same when 25 nM wt Hfq and 25 nM Hfq-65 was added to 4 nM ^{32}P-labeled OxyS or RprA (data not shown). The results are consistent with a 1:1 stoichiometry for Hfq_6 binding to these RNAs.

Hfq binding to A_{18} or $DsrA_{DII}$ monitored by fluorescence anisotropy

Another experimental approach that suggested two Hfq_6 bound A_{18} was fluorescence anisotropy (Sun & Wartell, 2006). A model in which two Hfq_6 sequentially bound A_{18} gave a better fit to fluorescence anisotropy data than a model that assumed a 1:1 complex. We have re-examined and extended these measurements and the analyses in light of the above results. Figure 4.4A shows that the 2:1 binding model (solid line) does give the best fit to the titration of Hfq to 2 nM of FAM-labeled A_{18}. The dotted line is the nonlinear least squares fit of the 1:1 model (eq (1) in Materials and Methods) with K_d a variable parameter and the other parameters (A_f, A_b, $[R]_T$, $[P]_T$) determined from the experimental data. The A_b value of 0.166 was determined from the horizontal asymptote to the anisotropy values of the four highest Hfq_6 concentrations used in the experiment. If, however, one allows A_b to be somewhat flexible and assume a value of 0.185, the fit of the 1:1 model approaches that of the 2:1 model (dashed line). Considering that the 2:1 model has more variable parameters with which to fit the data, the difference between the two models no longer persuasively favors the 2:1 model. Both models indicate K_d values in the range of 5 to 10 nM.

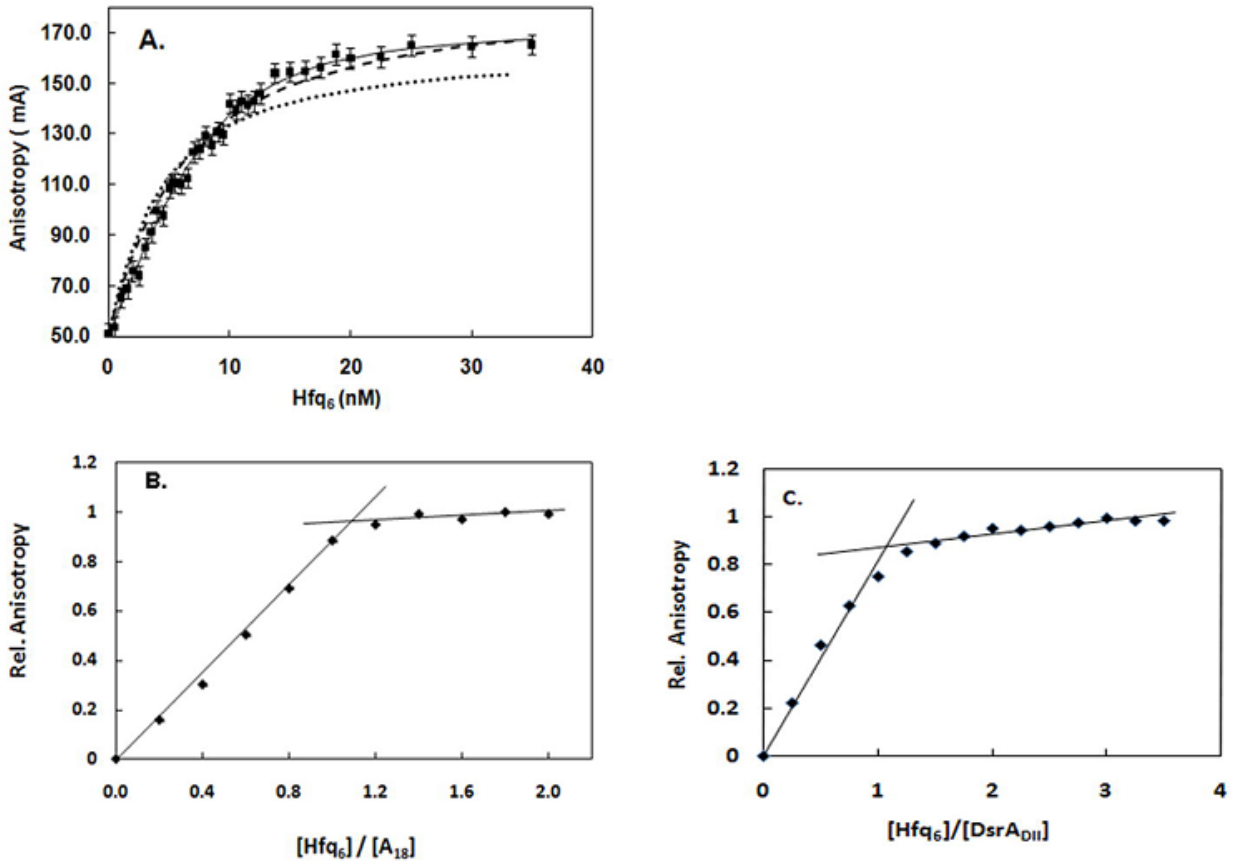


Figure 4.4: Fluorescence anisotropy titration of FAM-A_{18} with Hfq. (**A**.) Comparison of experimental data with 2 nM FAM-A_{18} (spuares) to best fit of 2:1 model (solid line), 1:1 model with K_d variable (circles), and 1:1 model with variable K_d and A_b (dotted line). Parameters for: 2:1 model; K_1 = 10.1 nM, K_2 = 5 nM, A_{b1} = 0.148, A_{b2} = 0.172. For 1:1 models; K_1 = 4.4 nM, A_b = 0.166 for dotted line, K_1 = 5 nM, A_b = 0.185 for dashed line. (**B.)** Experimental anisotropy measurements of 5 μM FAM-A_{18} titrated with Hfq_6. (**C.)** Experimental anisotropy measurements of 2 μM $DsrA_{DII}$ titrated with Hfq_6.

To further examine the stoichiometry of Hfq_6 binding to A_{18} using this experimental approach, titration of A_{18} with Hfq was carried out at concentrations well above the K_d (5 μM A_{18}) where stoichiometric binding is expected. Figure 4.4B shows that the anisotropy change of A_{18} saturates at a ratio of Hfq_6 and A_{18} consistent with a 1:1 stoichiometry. A similar experiment conducted with 2 μM $DsrA_{DII}$ also showed a break in the plot at a 1:1 molar ratio of Hfq_6 and

$DsrA_{DII}$ (Figure 4.4C). The K_d of Hfq_6 binding to $DsrA_{DII}$ under the conditions of the experiment (0.1 M NaCl + 20 mM Tris) was approximately 4 nM (Figure S4.3).

Hfq interaction with both A_{18} and $DsrA_{DII}$

Polyacrylamide gel mobility shift experiments have previously demonstrated that Hfq can form a complex with a poly(A) sequence and DsrA (Brescia et al., 2003). The observation of a "super shifted" gel band consisting of the above three components indicates a ternary complex but does not exclude the possibility that more than one Hfq hexamer is needed to form this complex. MALDI-TOF mass spectrometry was employed to examine if a mass could be detected consistent with a complex formed by Hfq_6, $DsrA_{DII}$, and A_{18}. 10 µM Hfq_6 was mixed with 5 µM $DsrA_{DII}$ and 5 µM A_{18} for 15 minutes and treated with formaldehyde as described in Materials and Methods. Figure 4.5 shows the MALDI-TOF spectrum of this sample. Peaks were observed corresponding to molecular masses very similar to Hfq_6 (66650 D, theoretical mass 66998 D), Hfq_6•A_{18} (72400 D, theoretical mass 72839 D), and Hfq_6•$DsrA_{DII}$ (78230 D, theoretical mass 79029 D). A small but reproducible peak was observed in the region corresponding to a mass of 84355 D, consistent with the combined mass of one Hfq_6, one $A_{18,}$ and one $DsrA_{DII}$ (theoretical mass 84869 D). We note that the lower observed masses compared to theoretical masses (by 350-700 D) appears to be due to external calibration error.

The small peak corresponding to a mass of 55530 D is consistent with five subunits of Hfq (theoretical mass 55832 D). Small nearby peaks were reproducibly observed and may be related to four or five subunits of Hfq with A_{18}, $DsrA_{DII}$, or both. The intensities of peaks corresponding to the unbound forms of four and five Hfq subunits were slightly higher (relative to the Hfq monomer peak) in the presence of both $DsrA_{DII}$ and A_{18} (~3 %) compared to when

only one RNA was present (~2%). Not surprisingly, the Hfq•A_{18} and Hfq•$DsrA_{DII}$ peaks were smaller by about 2.5 fold when both $DsrA_{DII}$ and A_{18} were present compared to spectra of Hfq and only one RNA. The small peak at 89515 is consistent with one Hfq_6 and a dimer of $DsrA_{DII}$ and is similar to the small peak observed with Hfq and $DsrA_{DII}$ (Figure 4.1C).

The intensity in the region of the 84355 D mass in Figure 4.5 is consistent with a 1:1:1 Hfq_6• A_{18}• $DsrA_{DII}$ ternary complex; however, this peak was considerably smaller than the peaks corresponding to Hfq_6•A_{18} or Hfq_6• $DsrA_{DII}$. This may reflect an intrinsic instability of this ternary complex or a limitation of the method in reporting complexes of Hfq_6 with two RNAs.

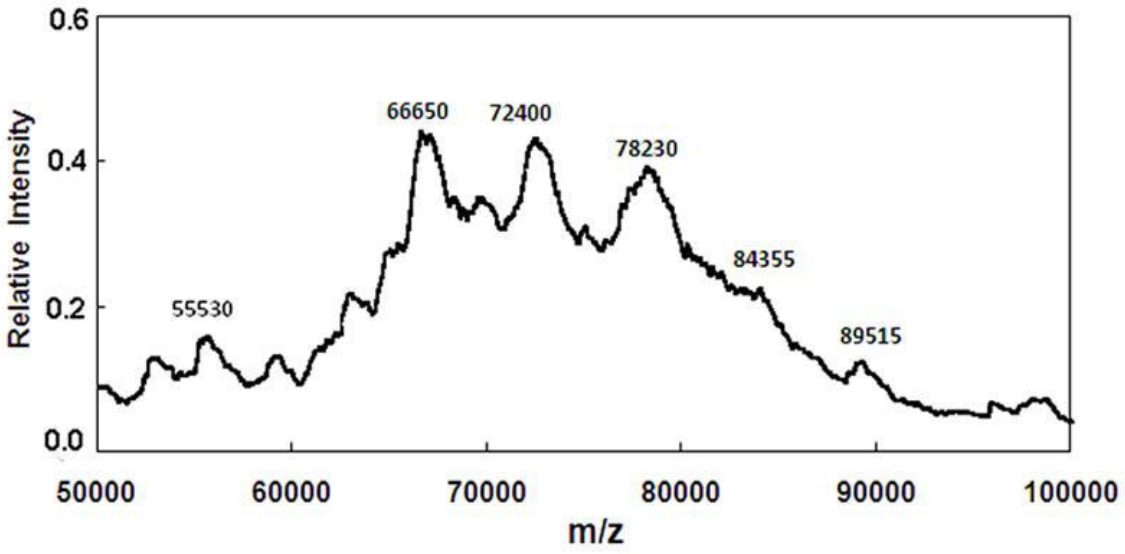


Figure 4.5: MALDI-TOF m/z spectrum of 10 μM Hfq_6 plus 5 μM $DsrA_{DII}$ and 5 μM A_{18} prepared in the 0.2 M Na^+ solvent, cross-linked with formaldehyde, and mixed with matrix solution as described in Material and Methods.

To explore this question in solution, we examined the effect of adding $DsrA_{DII}$ on the fluorescence anisotropy of a preformed complex of Hfq_6•FAM-A_{18}. Hfq_6 was added to 100 nM FAM-A_{18} in a solvent of 0.1 M NaCl+ 20 mM Tris (8.3), increasing the anisotropy from 0.037 to 0.080, approximately 45% of the maximum anisotropy change induced by saturating Hfq_6. Adding aliquots of $DsrA_{DII}$ to produce a final solution with 75 nM FAM-A_{18}, 65 nM $DsrA_{DII}$, and 63 nM Hfq_6 reduced the anisotropy by about 30% (Figure 4.6). If a ternary Hfq_6• A_{18}• $DsrA_{DII}$ complex is stable relative to the 1:1 Hfq_6•RNA complexes, an increase rather than decrease in anisotropy is expected. This experiment was repeated using the complete DsrA, surmising its higher molecular weight and strong binding to Hfq_6 may be required to observe the expected anisotropy increase resulting from formation of a ternary complex. However the outcome was similar (data not shown). When Hfq and FAM-$DsrA_{DII}$ were preformed and A_{18} or polyA added to the solution a similar decrease in anisotropy was observed (data not shown). The above results were surprising given the outcome of gel shift experiments (Brescia et al., 2003; Mikulecky et al., 2004; Updegrove et al., 2008) that clearly show complexes can form involving Hfq_6, DsrA, and a poly(A) sequence. The apparently disparate implications of the two types of experiments may, however, be reconcilable as discussed below.

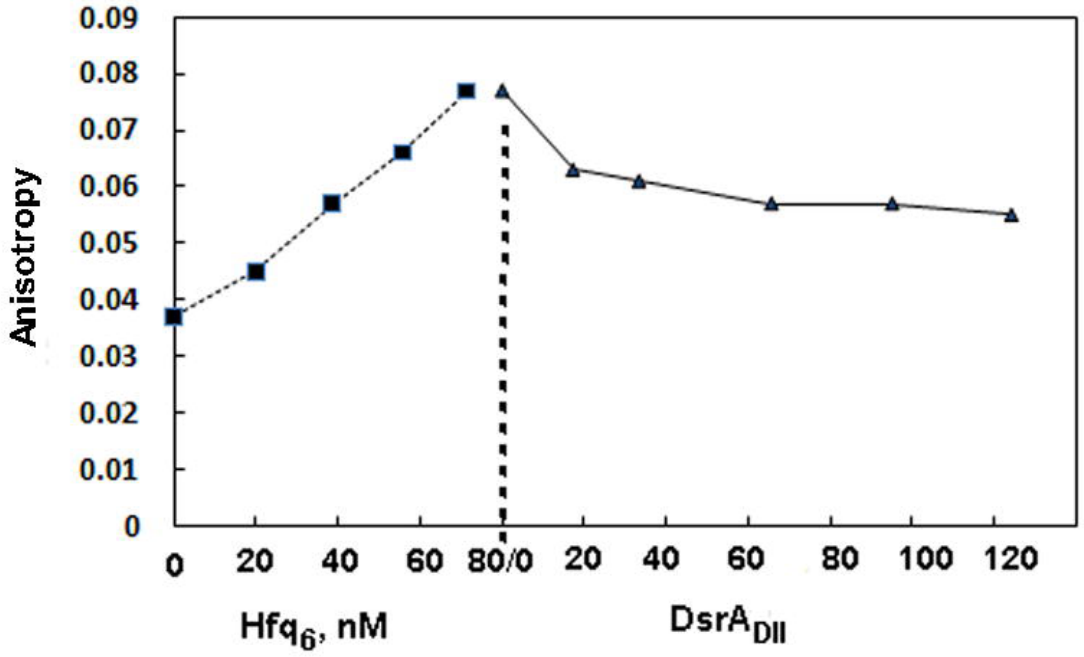


Figure 4.6: Fluorescence anisotropy experiment of FAM-A_{18} with Hfq and $DsrA_{DII}$. Hfq was titrated to 100 nM FAM-A_{18} to give 0.080, ~45% of the maximum anisotropy. Then aliquots of unlabeled $DsrA_{DII}$ were added to give the concentrations shown.

DISCUSSION

The results from mass spectrometry, analytical ultracentrifugation, fluorescence anisotropy, and competition gel mobility shift assay all point to a 1:1 stoichiometry for the Hfq_6• A_{18} and Hfq_6• $DsrA_{DII}$ complexes. The more limited studies on Hfq binding to the RNAs DsrA, RprA, OxyS and OxyS-18 support a similar conclusion. These experiments were carried out with RNA concentrations from 4 nM to 5 μM in solvents with 0.1 M - 0.5 M Na^+. The 1:1 stoichiometry is the same value determined by isothermal titration calorimetry measurements of Hfq_6 binding DsrA or a 140 nt *rpoS* mRNA segment (Mikulecky et al., 2004), but differs from the 2:1 (Hfq_6:RNA) stoichiometry inferred from gel shift assays of Hfq_6 binding to DsrA, 138 nt *rpoS* RNA (Lease & Woodson, 2004), $DsrA_{DII}$ (Sun & Wartell, 2006), and RprA (Updegrove et al., 2008), as well as the fluorescence anisotropy and fluorescence quenching study of Hfq_6 binding to A_{18} (Sun & Wartell, 2006). Since two methods used in the current work, mass spectrometry and sedimentation equilibrium, are robust model-independent approaches, our results raise the question why a 2:1 stoichiometry was inferred from previous investigations.

The results described by Figure 4.4A provide an explanation why a 2:1 stoichiometry was previously misinterpreted from the fluorescence anisotropy measurements of Hfq binding to FAM-A_{18} at low nM concentrations. The anisotropy of the fully bound FAM-A_{18}, A_b, appears to have been previously underestimated. Increasing the experimentally derived value of A_b by ~11% produced a much better fit to the data using the 1:1 model. Assuming some flexibility in the A_b value can be justified since there is uncertainty in the Hfq_6 concentration required to saturate binding of FAM-A_{18}. With this adjustment to A_b the difference between the predictions of the 2:1 model vs. the 1:1 model no longer persuasively favors the 2:1 model.

The 2:1 stoichiometry inferred from the gel shift assay was suggested by equilibrium binding analyses of gel shift data obtained using 2 to 4 nM RNA that indicated a Hill coefficient above 2, as well as from data obtained with 400 nM to 1.0 μM of RNA, concentrations above the K_d (Lease & Woodson, 2004; Sun & Wartell, 2006). Since similar outcomes came from different laboratories it seems unlikely that differences in binding activity of Hfq preparations influenced this outcome. Also, the Hfq used in the current experiments, which yield a 1:1 stoichiometry, reproduced the outcome of the gel shift assay (data not shown). While a definitive argument cannot yet be made why the gel shift assay yielded a 2:1 stoichiometry, several factors that might complicate interpretation of gel shift data may provide an explanation.

The equilibrium established in the sample solution may be altered as the low ionic strength buffer (0.5 X TBE) exchanges with the loading buffer as the macromolecules enter the gel or during electrophoresis (Bloomfield et al., 2000). Although a low ionic strength solution may stabilize Hfq•RNA complexes, it has also been shown to produce well ordered fibers of Hfq_6 (Arluison et al., 2006). If Hfq_6 aggregates in the gel environment it could alter the nature or amount of the Hfq•RNA complexes.

Factors governing the mass transport of Hfq•RNA complexes in a gel may also contribute to misleading interpretation of gel shift data, independent of the potential for Hfq_6 aggregation. Using a phenomenological theory of gel electrophoresis, Cann simulated the gel patterns produced by several protein-DNA interactions employing association and dissociation rate constants representative of the interactions and experimentally derived transport parameters (Cann, 1989). The simulations validated the application of the gel shift method for determining binding constants and stoichiometry for strong interactions with association (k_a) and dissociation

(k_d) rate constants of $k_a = 3 \times 10^9$ $M^{-1}s^{-1}$, $k_d = 1.3 \times 10^{-4}$ s^{-1}. However the simulation also showed that a significant amount of the initial protein-nucleic acid complex entering the gel can irreversibly dissociate during electrophoresis. When parameters mimicking an intermediate strength complex were used ($k_a = 1.3 \times 10^6$ $M^{-1}s^{-1}$, $k_d = 1.3 \times 10^{-4}$ s^{-1}) with 10 nM each of protein and nucleic acid, 49% of the initial protein-nucleic acid complex irreversibly dissociated from this band during electrophoresis. The extent of irreversible dissociation of the initial protein-nucleic acid complexes clearly depends on the concentrations used and the parameters of the system. The importance of these considerations has been demonstrated for properly interpreting gel shift data on a repressor-DNA operator system (Kleinschmidt et al., 1991).

It is worth noting that in the above example although electrophoresis depleted the amount of material in the nucleic acid-protein band, the unbound nucleic acid band could still be used to calculate the equilibrium dissociation constant to good accuracy (Cann, 1989). Thus gel shift data can be used to evaluate binding constants, even when the nucleic acid-protein bands do not accurately reflect the initial amount of these complexes. We note that interpretation of the competition gel shift experiment described in Figure 4.3 does not depend on a quantitative evaluation of band intensities. The absence of a band intermediate between the shifted bands corresponding to RNA bound to wt-Hfq or Hfq-65 is consistent with 1:1 complexes.

The third method that suggested a 2:1 stoichiometry for $Hfq_6 \bullet A_{18}$ was fluorescence quenching of Hfq's tyrosines by A_{18}. Quenching of Hfq fluorescence saturated when the amount of added A_{18} reached a molar ratio of 0.5:1 (A_{18}: Hfq_6) (Sun & Wartell, 2006). Controls indicated that the inner filter effect (Lakowicz, 2006) due to the absorbance of A_{18} at the excitation wavelength was negligible. We are currently unable to reconcile the apparent 2:1 stoichiometry implied from this experiment with the 1:1 stoichiometry determined in the current

work. It is possible that A_{18} binding has a complex effect on the fluorescence of Hfq's three tyrosines such that a straightforward interpretation of the data is quantitatively flawed.

Several lines of evidence have shown that Hfq_6 possesses two distinct RNA binding surfaces (Mikulecky et al., 2004). The proximal surface appears to be involved in Hfq binding to a single-stranded sequence with several uracils and/or adenines adjacent to one or more hairpins (Schumacher et al., 2002; Zhang et al., 2002; Geissmann & Touati, 2004). The distal surface of Hfq_6 binds to a repeated motif $(ARN)_n$, $n \geq 4$ (with R a purine, N any nucleoside) (Link et al., 2009). The latter motif includes the poly(A) sequence at the 3' ends of mRNAs, and segments found in the 5' leader region of at least two mRNAs (Soper & Woodson, 2008; Salim & Feig, 2010). With two distinct binding surfaces, a single Hfq hexamer has the potential to bind a mRNA and sRNA simultaneously.

The MALDI-TOF results suggest the existence of a $Hfq_6 \bullet A_{18} \bullet DsrA_{DII}$ complex; however, the small size of the peak does not support the notion that a 1:1:1 complex is very stable. The fluorescence anisotropy experiment in Figure 4.6 also does not provide evidence for a stable ternary complex in solution. DsrA and A_{18} do not appear to bind Hfq independently under the conditions of the experiment. This appears to contradict the observation that poly(A) sequences can form a ternary complex with Hfq_6 and DsrA in polyacrylamide gels. A possible explanation of these observations may be related to the low ionic strength solvent and cage effect of the gel environment. Studies by deHaseth and Uhlenbeck (de Haseth & Uhlenbeck, 1980a) as well as the more recent demonstration of Hfq fibers (Arluison et al., 2006) indicate that low ionic strength solutions promote Hfq aggregation. The gel environment may promote Hfq_6 aggregation and enable ternary complexes that involve more than one Hfq_6. These complexes may not form in the 0.1 M Na^+ solution employed in the anisotropy experiment.

A counter hypothesis that can explain why putative ternary complexes are not reported by fluorescence anisotropy is more difficult. If the dissociation lifetime of a ternary complex is shorter than its rotational correlation time (τ_c) it could go undetected. For a 1:1:1 complex of DsrA, FAM-A_{18}, and $Hfq_{6,}$ τ_c can be estimated to be ~60 ns (Serdyuk et al., 2007). A dissociation lifetime this short is inconsistent with a stable ternary complex. The total anisotropy reflects the sum of each anisotropic species. Binding of $DsrA_{DII}$ or DsrA to FAM-A_{18}•Hfq_6 is expected to slow the rotational correlation time and increase anisotropy. If binding also induces a conformational change that partially releases the FAM –A_{18} it may cancel the effect of the increased size on the rotational correlation time and in principle could reduce the anisotropy. In order to explain all of the results this would also have to be true for A_{18} binding to $DsrA_{DII}$•Hfq_6. This seems a less likely explanation of the data than displacement of the bound RNA from Hfq_6 by the other RNA.

Regardless of the uncertainty in a definitive explanation for the stoichiometry reported by the previous gel results and the nature of the polyA-Hfq-DsrA complex observed in gels, the major conclusion from this work, that Hfq_6 has a 1:1 binding stoichiometry with RNA at concentration and ionic strength conditions mimicking a cell environment, addresses a question important to understanding how Hfq facilitates interactions between RNAs.

Acknowledgements

I acknowledge John Correia for his contribution in conducting the sedimentation velocity and equilibrium experiments, and for the data analysis. I would also like to thank Yanfeng Chen for conducting the MALDI-TOF MS experiments and processing the data and Charles Terry for his assistance in purifying the mutant Hfq used in this study. I am grateful for the support

provided by a Georgia Tech/CDC seed grant and funding from the NASA Astrobiology Institute as well as a URS award to Charles Terry by the Institute of Bioengineering and Biosciences.

REFERENCES

Arluison V, Hohng S, Roy R, Pellegrini O, Regnier P, Ha T. 2007. Spectroscopic observation of RNA chaperone activities of Hfq in post-transcriptional regulation by a small non-coding RNA. *Nucleic Acids Res 35*:999-1006.

Arluison V, Mura C, Guzman MR, Liquier J, Pellegrini O, Gingery M, Regnier P, Marco S. 2006. Three-dimensional structures of fibrillar Sm proteins: Hfq and other Sm-like proteins. *J Mol Biol 356*:86-96.

Azam TA, Hiraga S, Ishihama A. 2000. Two types of localization of the DNA-binding proteins within the Escherichia coli nucleoid. *Genes Cells 5*:613-626.

Bloomfield VA, Crothers DM, Tinoco I. 2000. *Nucleic acids : structures, properties, and functions*. Sausalito, Calif.: University Science Books.

Brennan RG, Link TM. 2007. Hfq structure, function and ligand binding. *Curr Opin Microbiol 10*:125-133.

Brescia CC, Mikulecky PJ, Feig AL, Sledjeski DD. 2003. Identification of the Hfq-binding site on DsrA RNA: Hfq binds without altering DsrA secondary structure. *RNA 9*:33-43.

Butland G, Peregrin-Alvarez JM, Li J, Yang W, Yang X, Canadien V, Starostine A, Richards D, Beattie B, Krogan N, Davey M, Parkinson J, Greenblatt J, Emili A. 2005. Interaction network containing conserved and essential protein complexes in *Escherichia coli*. *Nature 433*:531-537.

Cann JR. 1989. Phenomenological theory of gel electrophoresis of protein-nucleic acid complexes. *J Biol Chem 264*:17032-17040.

Carmichael GG, Weber K, Niveleau A, Wahba AJ. 1975. The host factor required for RNA phage Qbeta RNA replication in vitro. Intracellular location, quantitation, and purification by polyadenylate-cellulose chromatography. *J Biol Chem 250*:3607-3612.

Cunning C, Brown L, Elliott T. 1998. Promoter substitution and deletion analysis of upstream region required for rpoS translational regulation. *J Bacteriol 180*:4564-4570.

de Haseth PL, Uhlenbeck OC. 1980a. Interaction of Escherichia coli host factor protein with oligoriboadenylates. *Biochemistry 19*:6138-6146.

de Haseth PL, Uhlenbeck OC. 1980b. Interaction of Escherichia coli host factor protein with Q beta ribonucleic acid. *Biochemistry 19*:6146-6151.

Folichon M, Allemand F, Regnier P, Hajnsdorf E. 2005. Stimulation of poly(A) synthesis by Escherichia coli poly(A)polymerase I is correlated with Hfq binding to poly(A) tails. *FEBS J 272*:454-463.

Folichon M, Arluison V, Pellegrini O, Huntzinger E, Regnier P, Hajnsdorf E. 2003. The poly(A) binding protein Hfq protects RNA from RNase E and exoribonucleolytic degradation. *Nucleic Acids Res 31*:7302-7310.

Geissmann TA, Touati D. 2004. Hfq, a new chaperoning role: binding to messenger RNA determines access for small RNA regulator. *Embo J 23*:396-405.

Gill SC, von Hippel PH. 1989. Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem 182*:319-326.

Karas M, Gluckmann M, Schafer J. 2000. Ionization in matrix-assisted laser desorption/ionization: singly charged molecular ions are the lucky survivors. *J Mass Spectrom 35*:1-12.

Kawamoto H, Koide Y, Morita T, Aiba H. 2006. Base-pairing requirement for RNA silencing by a bacterial small RNA and acceleration of duplex formation by Hfq. *Mol Microbiol 61*:1013-1022.

Kleinschmidt C, Tovar K, Hillen W. 1991. Computer simulations and experimental studies of gel mobility patterns for weak and strong non-cooperative protein binding to two targets on the same DNA: application to binding of tet repressor variants to multiple and single tet operator sites. *Nucleic Acids Res 19*:1021-1028.

Lakowicz JR. 2006. *Principles of fluorescence spectroscopy*. New York: Springer.

Laue TM, Shah, B.D., Ridgeway, T.M. and Pelletier, S.L. 1992. *Analytical Ultracentrifugation in Biochemistry and Polymer Sciences (Eds: Harding, S.E., Rowe, A.J., and Horton, J.C)* Royal Society of Chemistry, Cambridge, U.K.

Le Derout J, Boni IV, Regnier P, Hajnsdorf E. 2010. Hfq affects mRNA levels independently of degradation. *BMC Mol Biol 11*:17.

Le Derout J, Folichon M, Briani F, Deho G, Regnier P, Hajnsdorf E. 2003. Hfq affects the length and the frequency of short oligo(A) tails at the 3' end of Escherichia coli rpsO mRNAs. *Nucleic Acids Res 31*:4017-4023.

Lease RA, Woodson SA. 2004. Cycling of the Sm-like protein Hfq on the DsrA small regulatory RNA. *J Mol Biol 344*:1211-1223.

Lenz DH, Mok KC, Lilley BN, Kulkarni RV, Wingreen NS, Bassler BL. 2004. The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in Vibrio harveyi and Vibrio cholerae. *Cell 118*:69-82.

Link TM, Valentin-Hansen P, Brennan RG. 2009. Structure of Escherichia coli Hfq bound to polyriboadenylate RNA. *Proc Natl Acad Sci U S A.*

Liu S, Stafford WF, 3rd. 1995. An optical thermometer for direct measurement of cell temperature in the Beckman instruments XL-A analytical ultracentrifuge. *Anal Biochem 224*:199-202.

Lundblad JR, Laurance M, Goodman RH. 1996. Fluorescence polarization analysis of protein-DNA and protein-protein interactions. *Mol Endocrinol 10*:607-612.

Majdalani N, Hernandez D, Gottesman S. 2002. Regulation and mode of action of the second small RNA activator of RpoS translation, RprA. *Mol Microbiol 46*:813-826.

Majdalani N, Vanderpool CK, Gottesman S. 2005. Bacterial small RNA regulators. *Crit Rev Biochem Mol Biol 40*:93-113.

Masse E, Escorcia FE, Gottesman S. 2003. Coupled degradation of a small regulatory RNA and its mRNA targets in Escherichia coli. *Genes Dev 17*:2374-2383.

Mikulecky PJ, Kaw MK, Brescia CC, Takach JJ, Sledjeski D, Feig AL. 2004. *Escherichia coli* Hfq has distinct interaction surfaces for DsrA, rpoS and poly(A) RNAs. *Nat Struct Mol Biol 11*:1206-1214.

Mohanty BK, Kushner SR. 2006. The majority of Escherichia coli mRNAs undergo post-transcriptional modification in exponentially growing cells. *Nucleic Acids Research 34*:5695-5704.

Mohanty BK, Maples VF, Kushner SR. 2004. The Sm-like protein Hfq regulates polyadenylation dependent mRNA decay in Escherichia coli. *Mol Microbiol 54*:905-920.

Moller T, Franch T, Hojrup P, Keene DR, Bachinger HP, Brennan RG, Valentin-Hansen P. 2002a. Hfq: a bacterial Sm-like protein that mediates RNA-RNA interaction. *Mol Cell 9*:23-30.

Moller T, Franch T, Udesen C, Gerdes K, Valentin-Hansen P. 2002b. Spot 42 RNA mediates discoordinate expression of the E. coli galactose operon. *Genes Dev 16*:1696-1706.

Morita T, Maki K, Aiba H. 2005. RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes Dev 19*:2176-2186.

Niranjanakumari S, Lasda E, Brazas R, Garcia-Blanco MA. 2002. Reversible cross-linking combined with immunoprecipitation to study RNA-protein interactions in vivo. *Methods 26*:182-190.

Philo JS. 2006. Improved methods for fitting sedimentation coefficient distributions derived by time-derivative techniques. *Anal Biochem 354*:238-246.

Royer CA. 1993. Improvements in the numerical analysis of thermodynamic data from biomolecular complexes. *Anal Biochem 210*:91-97.

Royer CA, Beechem JM. 1992. Numerical analysis of binding data: advantages, practical aspects, and implications. *Methods Enzymol 210*:481-505.

Salim NN, Feig AL. 2010. An upstream Hfq binding site in the fhlA mRNA leader region facilitates the OxyS-fhlA interaction. *PLoS One 5*.

Sambrook J, Russell DW. 2001. *Molecular cloning : a laboratory manual*. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

Schuck P, Perugini MA, Gonzales NR, Howlett GJ, Schubert D. 2002. Size-distribution analysis of proteins by analytical ultracentrifugation: strategies and application to model systems. *Biophys J 82*:1096-1111.

Schumacher MA, Pearson RF, Moller T, Valentin-Hansen P, Brennan RG. 2002. Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: a bacterial Sm-like protein. *Embo J 21*:3546-3556.

Serdyuk IN, Zaccai NR, Zaccai G. 2007. *Methods in molecular biophysics : structure, dynamics, function*. Cambridge ; New York: Cambridge University Press.

Sledjeski DD, Whitman C, Zhang A. 2001. Hfq is necessary for regulation by the untranslated RNA DsrA. *J Bacteriol 183*:1997-2005.

Soper TJ, Woodson SA. 2008. The rpoS mRNA leader recruits Hfq to facilitate annealing with DsrA sRNA. *RNA 14*:1907-1917.

Stafford WF, Sherwood PJ. 2004. Analysis of heterologous interacting systems by sedimentation velocity: curve fitting algorithms for estimation of sedimentation coefficients, equilibrium and kinetic constants. *Biophys Chem 108*:231-243.

Steege DA. 2000. Emerging features of mRNA decay in bacteria. *RNA 6*:1079-1090.

Sun X, Wartell RM. 2006. Escherichia coli Hfq binds A18 and DsrA domain II with similar 2:1 Hfq6/RNA stoichiometry using different surface sites. *Biochemistry 45*:4875-4887.

Takada A, Wachi M, Kaidow A, Takamura M, Nagai K. 1997. DNA binding properties of the hfq gene product of Escherichia coli. *Biochem Biophys Res Commun 236*:576-579.

Tsui HC, Feng G, Winkler ME. 1997. Negative regulation of mutS and mutH repair gene expression by the Hfq and RpoS global regulators of Escherichia coli K-12. *J Bacteriol 179*:7476-7487.

Updegrove T, Wilf N, Sun X, Wartell RM. 2008. Effect of Hfq on RprA-rpoS mRNA pairing: Hfq-RNA binding and the influence of the 5' rpoS mRNA leader region. *Biochemistry 47*:11184-11195.

Updegrove TB, Correia JJ, Galletto R, Bujalowski W, Wartell RM. 2010. E. coli DNA associated with isolated Hfq interacts with Hfq's distal surface and C-terminal domain. *Biochim Biophys Acta*.

Valentin-Hansen P, Eriksen M, Udesen C. 2004. The bacterial Sm-like protein Hfq: a key player in RNA transactions. *Mol Microbiol 51*:1525-1533.

Vecerek B, Rajkowitsch L, Sonnleitner E, Schroeder R, Blasi U. 2008. The C-terminal domain of Escherichia coli Hfq is required for regulation. *Nucleic Acids Res 36*:133-143.

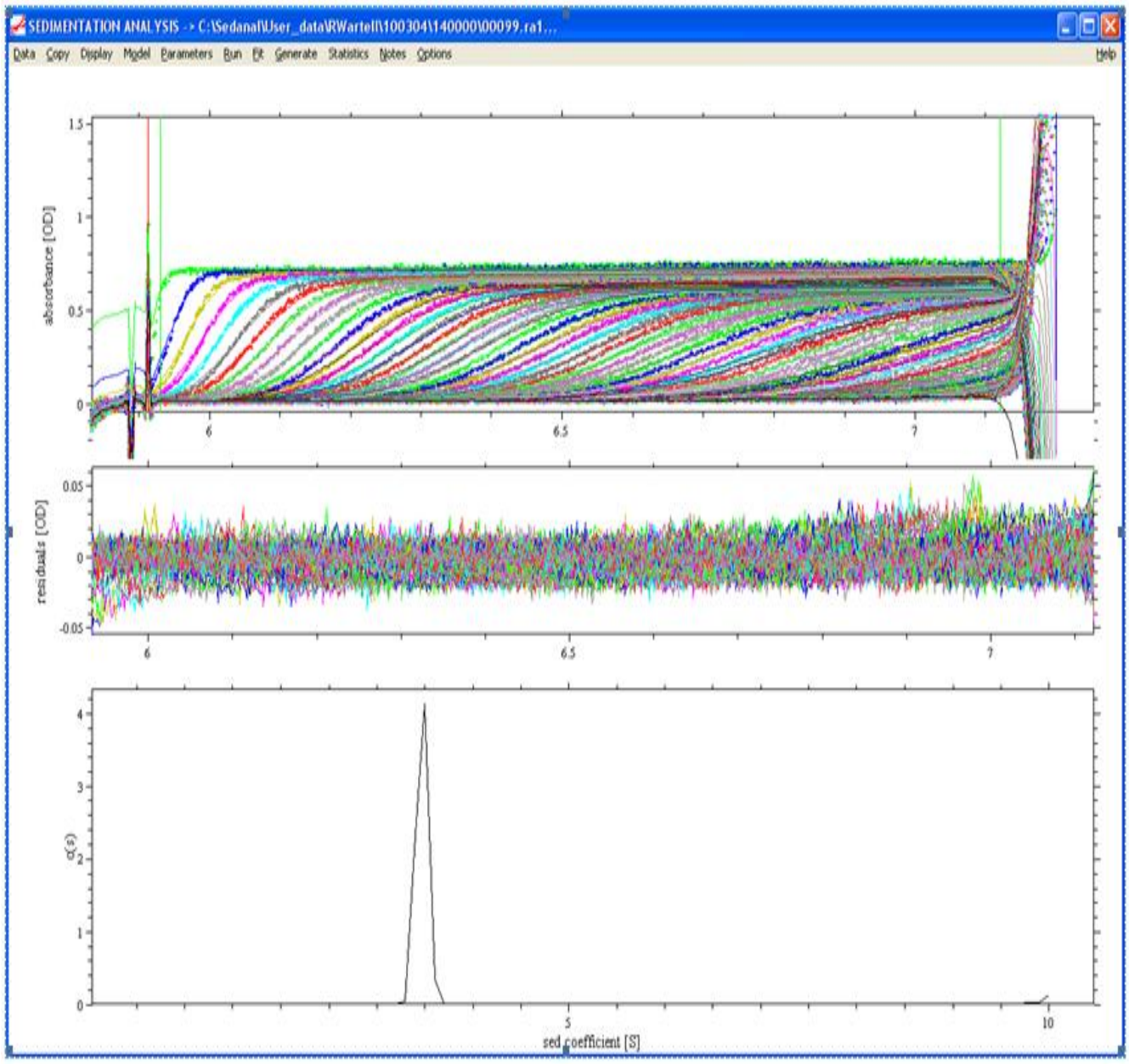
Vytvytska O, Jakobsen JS, Balcunaite G, Andersen JS, Baccarini M, von Gabain A. 1998. Host factor I, Hfq, binds to Escherichia coli ompA mRNA in a growth rate-dependent fashion and regulates its stability. *Proc Natl Acad Sci U S A 95*:14118-14123.

Waters LS, Storz G. 2009. Regulatory RNAs in bacteria. *Cell 136*:615-628.

Zhang A, Wassarman KM, Ortega J, Steven AC, Storz G. 2002. The Sm-like Hfq protein increases OxyS RNA interaction with target mRNAs. *Mol Cell 9*:11-22.

SUPPLEMENTARY MATERIAL

Sedimentation velocity runs of Hfq



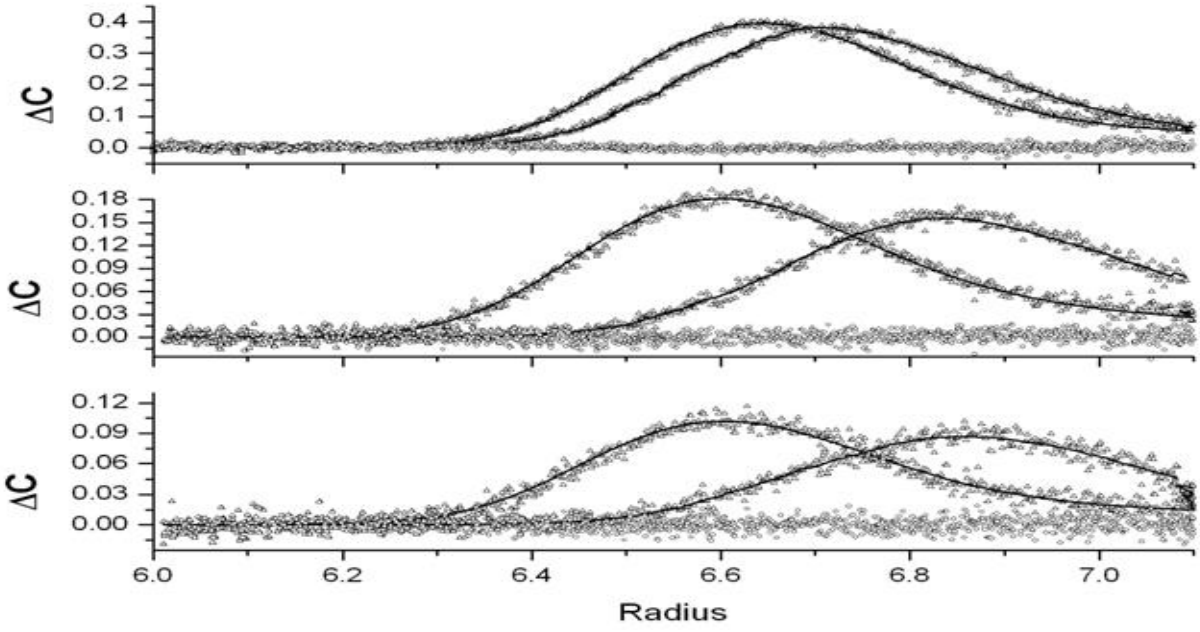


Figure S4.1: Traces from sedimentation velocity run of Hfq at 7.5 μM and Sedfit analysis that produced c(s) plot of Figure 4.2A followed by Sedanal single species fit of the three data sets for evaluation of molecular weight as described in the text under Analytical Ultracentrifugation Analysis.

Sedimentation equilibrium runs on Hfq alone and with FAM-A_{18}

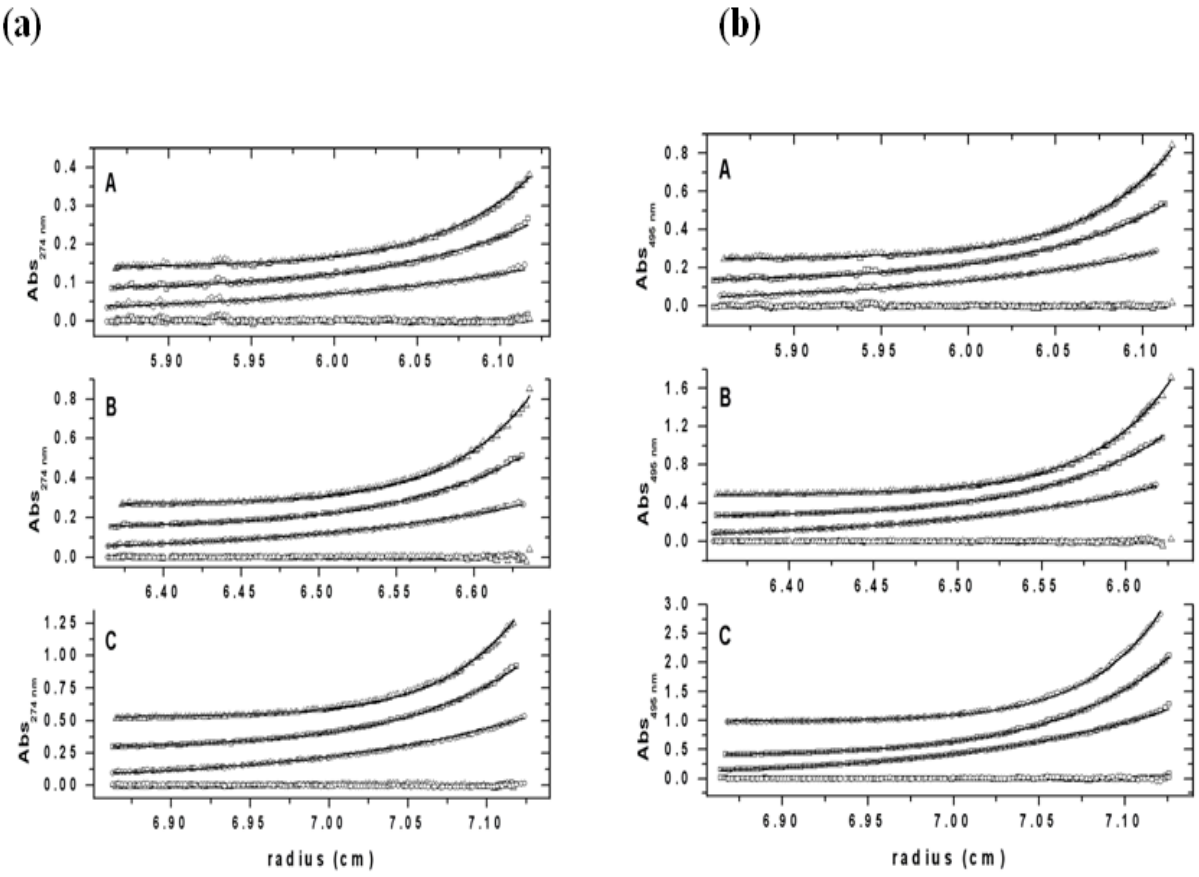


Figure S4.2: Sedimentation equilibrium runs on (**A.)** Hfq alone and **(B.)** with FAM-A_{18} at a 1:1 ratio. Loading concentrations are 2, 4 and 8 μM and samples were spun in an XLA at 12K, 16K, and 20K in a six channel centerpiece at 20°C. Data were collected at each speed after achieving equilibrium at 274 nm for Hfq samples and at 495 nm for the FAM-A_{18} samples at 0.001 cm spacing with nine flashes of the flash lamp. Nine data sets were globally fit in Sedanal to a single species model and 95% confidence intervals were estimated with an F-statistic procedure. Hfq gave a MW of 61.475 kD <58.5,64.2> and an rms = 0.00596; Hfq + FAM-A_{18} gave a MW of 68.929 kD <67.4,70.4> and an rms = 0.00724. This MW difference is consistent with a 1:1 binding model. The best global fits are plotted with A, B, C corresponding to channels A, B, C and the 2, 4, 8 μM samples. Within each panel data from different speeds are plotted as open symbols (12K (circles), 16K (squares) and 20K (triangles)) and the best fit as solid lines. The 16K and 20K data are offset for clarity. Residuals are superimposed near zero on the y-axis as open symbols.

Fluoresence anisotropy of Hfq binding to $DsrA_{DII}$

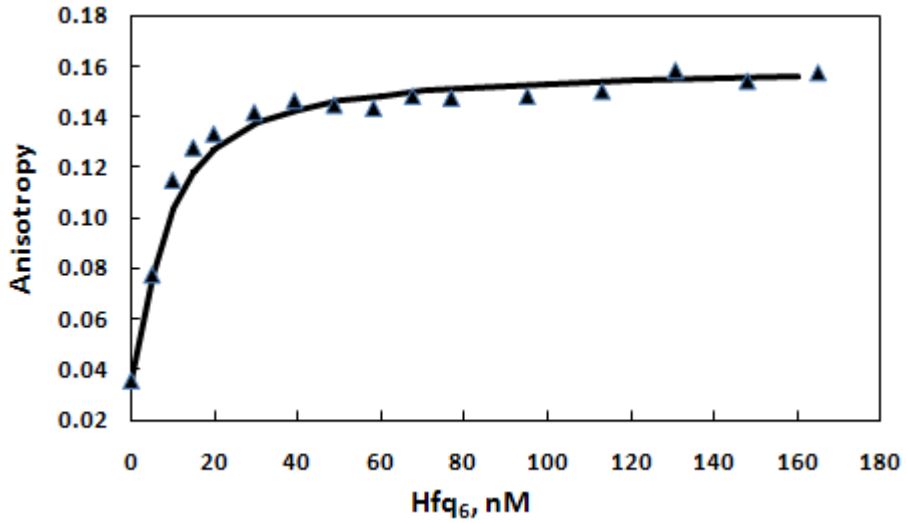


Figure S4.3: Fluorescence anisotropy of Hfq binding to 5 nM $DsrA_{DII}$. Data points are the experimental data, and the solid line shows the theoretical prediction using a K_d of 4 nM and A_b of 0.160 with a 1:1 model described in text and by eq. (2). Analysis fitting K_d and A_b to data indicates K_d is 4 ± 2 nM.

CHAPTER 5

Escherichia coli DNA associated with isolated Hfq interacts with Hfq's distal surface and C-terminal domain

INTRODUCTION

The Hfq protein of *E. coli* (also known as HF-1) is an RNA-binding protein and a key factor in posttranscriptional gene regulation (Valentin-Hansen et al. 2004; Majdalani et al. 2005; Brennan and Link 2007; Jousselin et al. 2009; Papenfort and Vogel 2009; Waters and Storz 2009) . The pleiotropic phenotypic effects that results from inactivating the Hfq gene in *E. coli* (Tsui et al. 1994) and other gram negative bacterial species has been linked to Hfq's role in facilitating the interaction of small regulatory RNAs (sRNAs) with messenger RNAs (mRNAs) and as a modulator of sRNA and mRNA stability. Hfq's role in regulating mRNA translation by sRNAs is an important feature of bacterial adaptation to stress and the establishment of virulence (Robertson and Roop 1999; Sonnleitner et al. 2003; Ding et al. 2004; Geng et al. 2009; Meibom et al. 2009) . Many studies of Hfq have been directed towards understanding its interaction with sRNAs and mRNAs in both *in vitro* and *in vivo* contexts. However, several studies on proteins associated with the *E. coli* nucleoid DNA suggested that Hfq binding to DNA may also have a functional role (Kajitani et al. 1994; Azam and Ishihama 1999).

Hfq was among the ten most prevalent proteins associated with nucleoid DNA isolated from *E. coli* (Azam and Ishihama 1999). In exponentially growing cells, it was the third most prevalent of the ten proteins. *In situ* immunofluorescence studies indicated that most Hfq appears to be in the cytoplasm (80 - 90 %), however a portion of this protein was found in the DNA nucleoid region of the cell (10 - 20 %) (Kajitani et al. 1994; Azam et al. 2000). Recent

electron microscopy studies have confirmed the presence of Hfq in the cytoplasm and nucleoid and demonstrated that Hfq is also localized close to the inner membrane (Diestra et al. 2009). Plasmid DNAs grown in *E. coli* were shown to bind Hfq *in vivo* and *in vitro* (Takada et al. 1997). Apparent equilibrium dissociation constants (K_d) for Hfq binding to two 60 bp DNA duplexes were reported as 125 and 250 nM (Azam and Ishihama 1999). Greater affinity was exhibited for the curved DNA than for the mixed sequence DNA. Although the above range of affinities is weaker than Hfq affinity for sRNAs (K_d ~ 20-50 nM), it is tight enough to suggest that Hfq binding to DNA may play a functional role *in vivo*, particularly in light of Hfq's μM-level cellular abundance (Valentin-Hansen et al. 2004).

In the current work, we have characterized *E. coli* genomic DNA fragments found associated with Hfq purified from lysed cells and investigated the nature of the Hfq - DNA interaction. Several lines of evidence indicate that Hfq binding to DNA involves the protein's distal surface and C-terminal domain. The sequences of amplified segments of the genomic DNA exhibit several interesting characteristics. Over half are predicted to have helical axis curvature and are predominantly from genes coding for membrane proteins.

MATERIALS AND METHODS

Purification and characterization of wild-type Hfq and mutant Hfq

The Impact-CN intein system (New England Biolabs) was used to produce and purify Hfq proteins as previously described (Sun and Wartell 2006). The *E. coli hfq* gene was cloned into the pTYB11 plasmid to create the expression plasmid pTYB11-wt Hfq. Hfq was expressed from this plasmid in *E. coli* strain ER2566 using the recommendations of the manufacturer. Cells were lysed using a French press in 0.5 M NaCl, 20 mM Tris (pH 8.3), 0.1 mM EDTA, 0.1% Triton X100, and 5% glycerol. The cell lysate was centrifuged and the supernatant loaded onto a chitin column. The column was extensively washed (15 to 20 bed volumes) with the wash buffer that consisted of 20 mM Tris (pH 8.3) and 0.5 or 1.0 M NaCl with or without 0.1% Triton X100 (all variations gave similar outcomes). The column was then incubated with 0.5 M NaCl and 20 mM Tris buffer plus 40 mM dithiothreitol. Eluted protein was concentrated and buffer-exchanged to 0.5 M NaCl and 20 mM Tris at pH 8.3 using 30 kD MWCO centrifugation filtration units. The protein preparation at this stage is referred to as Hfq-NA. Hfq was further purified by either a DEAE column or more commonly by a nuclease treatment to remove 250-260 nm absorbing material.

The nuclease treatment of Hfq-NA preparations was carried out by adding 7.5 U of micrococcal nuclease (Worthington Biochemical Corporation) to 1 ml of 0.2 - 0.4 OD_{274nm} units of the protein sample in a solvent of 0.2 M NaCl, 20mM Tris (pH 8.3), and 5 mM $CaCl_2$. We note that micrococcal nuclease activity is absolutely dependent on Ca^{2+}. Reactions were incubated at 37^{o}C for 1 hr and terminated by adding 10 µl of 0.5 M Na_2EDTA. Reactions were then extensively buffer exchanged with 0.5 M NaCl and 20mM Tris (pH 8.3) and their volumes

reduced to ~ 1 ml using a 15 ml 30 kD centrifugation filter. This approach was more consistent than a DEAE column in giving a high A_{275}/A_{250} absorbance ratio (Figure 5.1A).

Plasmids containing mutant *hfq* genes were generated from pTYB11-wt Hfq using the QuikChange Mutagenesis Kit from Stratagene Inc (Hemsley et al. 1989). In addition to the previously described mutations F42A, F39A, Q8A, R16A, K31A, and Y25A (Sun and Wartell 2006), *hfq* genes with single residue mutations R19A, R17A, and F11A were constructed and their proteins expressed. Two additional mutant *hfq* genes were constructed by creating stop codons at residues 76 and 66, respectively. These plasmids yielded truncated Hfq designated Hfq-65 and Hfq-75. The wt Hfq and mutant Hfq's were characterized for purity by SDS-PAGE and UV spectroscopy (Sun and Wartell 2006).

Analytical Ultracentrifugation

Sedimentation velocity studies were performed using a Beckman Optima XLA analytical ultracentrifuge equipped with absorbance optics and an An60 Ti rotor at 19.7^{o}C. Temperature was calibrated as described previously (Liu and Stafford 1995). Velocity data were typically collected at appropriate speeds using 274 nm to monitor Hfq and Hfq-NA, and 495 nm when FAM-A_{18} was added. Spacing of 0.002 cm was employed with one flash at each point in a continuous-scan mode. All experiments were initially analyzed with Sedfit to produce c(s) distributions (Schuck et al. 2002) and with $DCDT^{+2}$ to produce g(s) distributions and weight average S values (Philo 2006). Direct boundary fitting of velocity data to discrete models were also performed with the program Sedanal (Stafford and Sherwood 2004). Analysis with Sedanal requires input of MW, extinction coefficients, and density increments (typically estimated from 1-vbar*rho values). The buffer solution density for 0.5 M NaCl and 20 mM Tris (8.3) was

estimated in Sednterp to be 1.01920 gm/ml at 19.7°C. The vbar of Hfq was estimated with Sednterp (Laue 1992) to be 0.7248. The vbar of FAM-A_{18} is assumed to be 0.55. The extinction coefficient for FAM-A_{18} at 495 nm is 75,000 $M^{-1}cm^{-1}$; the extinction coefficient of Hfq at 274 nm is 4350 $M^{-1}cm^{-1}$ (Sun and Wartell 2006) . Parameter uncertainty is calculated with an Fstat routine within Sedanal at the 95% confidence interval and reported in a < , > format.

Sedimentation equilibrium studies were carried out on Hfq-NA at 1.75, 3.5, and 7 μM. The solvent for most studies was 0.5 M NaCl and 20 mM Tris (8.3). Employing a buffer of 0.2M NaCl and 20 mM Tris (8.3) gave similar results. Samples were centrifuged at 12,000 rpm and 16,000 rpm at a temperature of 19.7°C in six channel double sector cells. Data were collected at 274 nm. Equilibrium at each speed was judged with the software utility WinMATCH (http://www.biotech.uconn.edu/auf/?i=aufftp). This program makes a least-square comparison of successive scans to establish that equilibrium has been achieved. Values for density, vbar and extinction coefficients were as described above for sedimentation velocity measurements. Non-linear least squares fit of sedimentation equilibrium profiles to a model of two independent non-interacting components gave a much better fit than to a single species model (Figure. 5.2B). The six data sets from three concentrations and two speeds were best fit to a two species model using Sedanal. Molecular weight uncertainty is calculated with Fstat as described above.

Characterization of nucleic acid in Hfq-NA

The nucleic acid associated with Hfq in Hfq-NA was characterized by examining the aqueous phase after phenol-chloroform extraction. 30 μl of ~10 μM Hfq-NA was phenol-chloroform extracted and then analyzed on a 1% agarose gel, staining with ethidium bromide.

The effect of RNase A on Hfq-NA was examined by adding 1 μg of RNase A (Promega) to 30 ul of Hfq-NA and incubating for 30 min at 37^{o}C prior to phenol-chloroform extraction. The influence of DNase on Hfq-NA was examined by adding 6 units of DNase I (Promega) and incubating for 30 min at 37^{o}C. Control experiments using either yeast RNA or *E coli* genomic DNA verified the effectiveness of these nucleases. The aqueous phases from the phenol extractions were loaded onto a 1% agarose gel along with Hfq-NA not treated with nuclease.

Isolation and cloning of DNA associated with Hfq

4 μg of proteinase K (Sigma Aldrich) was added to 80 μl of ~ 0.8 OD_{274} non-nuclease treated Hfq-NA and incubated for 2 hrs at 50^{o}C. 10 μl of 10X gel loading buffer (0.2 M NaCl, 20 mM Tris (pH 8.3), 40 % glycerol and 0.02 % (wt/vol) bromophenol blue) was then added to the sample and 45 μl loaded directly into two lanes of a 1 % agarose gel. Following staining of the gel with ethidium bromide, the smeared DNA fragments were excised from the gel under 350 nm UV light and the DNA purified using IsoPure™ DNA Purification Kit (Denville Scientific). PCR amplification of ~ 20 ng of this DNA or *E. coli* genomic DNA as a control was carried out with tagged random primers (Grothues et al. 1993) . The tagged random primer employed was 5'-GGTAATC GGATCCAAGCNNNNNN-'3. The six underlined bases denote the BamHI recognition site. After initial 1 minute incubation at 95^{o}C, samples were exposed to 45 cycles of 96^{o}C for 30 sec, 30^{o}C for 1 min, 40^{o}C for 1 min, and 72^{o}C for 2 min, followed by a final incubation at 72^{o}C for 5 minutes. PCR products were purified from primers and reaction buffer using QIAquick® PCR purification kit (Qiagen) and quantified by UV absorbance. A second PCR reaction used the tagged primer 5'-GGTAATCGGATCCAAGC - 3' and 20 ng from the first PCR reaction with conditions of 95^{o}C for 2 min, followed by 35 cycles of 95^{o}C for 1 min,

60°C for 1 min, 72°C for 2 min, and a final incubation at 72°C for 5 minutes. Product from the second PCR reaction was digested with BamHI and cloned into BamHI digested pUC19 plasmid treated with Antarctic Phosphatase (New England Biolabs). Blue-white screening was used to isolate transformants with inserts. The inserts of purified plasmids were sequenced and searched against the Microbial Genome database using BLAST at the NCBI website (http://www.ncbi.nlm.nih.gov/cig-bin/Entrez/genome_table_cgi). 20 nt primers were used to amplify several of the cloned DNA segments in Table 5.1 for Hfq binding and gel mobility studies (Table S5.1). Following PCR reactions, the products were purified from primers, enzymes, and buffer using QIAquik PCR Purification Kit (Qiagen) and analyzed on a 1% agarose gel to assess purity and later on a 10% polyacrylamide gel to assess their helical axis curvature (Diekmann 1992). DNAs used for binding studies were ^{32}P labeled at the 5' ends using standard protocols and purified from [γ-^{32}P]ATP and buffer exchanged using centrifugal filtration with a 10kD MWCO filter.

Electrophoretic Gel Mobility Shift Assay

Binding reactions were carried out in a solution composed of 0.2 M NaCl, 20 mM Tris-HCl (pH 8.0), 0.05% bromophenol blue, and 5% glycerol in a 15 µl volume. Hfq was mixed with nucleic acids and maintained at room temperature for at least 30 minutes prior to loading and running on the gel. The gels were 5 % polyacrylamide [29:1 (w/w) acrylamide / bisacrylamide] with 3 % glycerol in 0.5 X TBE. Electrophoresis was conducted at room temperature at 80-100 V. Gels were either stained after the run with SYBR® Gold nucleic acid stain (Invitrogen) or if ^{32}P-labeled DNA was employed were imaged using a Fujifilm Image Reader FLA-3000. The gel images were analyzed using the Fujifilm Multi Gauge Imaging

Software V3.0. The sRNAs DsrA and RprA employed in this assay were transcribed and purified from previously constructed plasmids (Updegrove et al. 2008) using a MEGAscript-T7 kit (Ambion Inc).

DNA sequence analysis

The DNA fragments associated with Hfq were examined for their propensity for helical axis curvature by calculating bendability/curvature propensity plots for their sequences using the BEND.IT server, (http://hydra.icgeb.trieste.it/dna/bend_it.html). The DNase I based bendability parameters (Brukner et al. 1995) and the consensus bendability parameters (Gabrielian and Pongor 1996) were employed. Curvature scores were calculated using the default window size of 31 bp and are given in degrees/helical turn.

A search for conserved DNA sequence motifs among the DNA fragments isolated from Hfq-NA was made using the Gibbs Centroid Sampler (Thompson et al. 2007). This method utilizes information from the full ensemble of possible solutions and has been shown to improve motif finding compared to single optimal alignment of motif sites. The probability that 'x' DNA fragments in a sample of 'n' cloned DNA fragments are membrane protein genes was estimated using the hypergeometric probability distribution (http://stattrek.com/Lesson2/ Hypergeometric.aspx). This probability distribution predicts the probability of sampling 'n' items without replacement and finding 'x' of the items with a specific property when the total population of N items has k items with the specific property. Assumptions made in applying this distribution to the cloned DNA fragments are discussed in Results.

RESULTS

Hfq protein produced by the pTYB11-wt Hfq expression plasmid in *E. coli* and purified using affinity column chromatography and centrifugal filtration (see Materials and Methods) consistently exhibited UV spectra that suggested the presence of a small amount of contaminating nucleic acid (Figure 5.1A). The peak of the UV spectrum was at the expected value of ~274 nm, but the absorbance ratio A_{274}/A_{250} of ~1.3 was lower than predicted for a protein with Hfq's amino acid composition (SEDNTERP ver.1.09, http://www.jphilo.mailway.com (Laue 1992)). Although UV absorbing contamination could be a by-product of the isolation procedure, the observation by other researchers of similar spectral characteristics using different methods of Hfq isolation (Carmichael et al. 1975; Kajitani et al. 1994) suggested nucleic acid contamination may be a common occurrence in Hfq purification.

A micrococcal nuclease treatment followed by centrifugal filtration was added as a final step of Hfq purification. This step resulted in Hfq spectra with A_{274}/A_{250} ratios ≥ 1.8 (Figure 5.1A) supporting the hypothesis that nucleic acids had been removed. Centrifugal filtration without the nuclease treatment was not effective. We refer to Hfq samples prior to removal of the contaminating nucleic acid as Hfq-NA. Changes in cell growth temperature or method of cell lysis (sonication) did not alter the UV spectral characterization or analytical ultracentrifugation characterization described below.

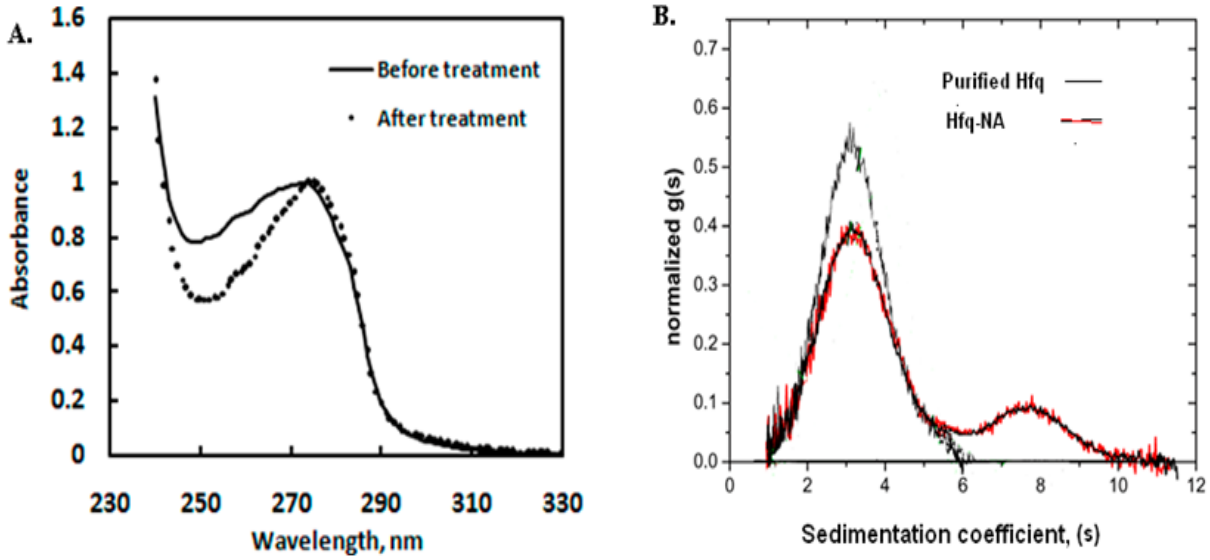


Figure 5.1: (**A**.) UV spectra of Hfq-NA, prior to micrococcal nuclease treatment (solid line), and after nuclease treatment (circles). (**B.**) Normalized g(s) distribution of sedimentation velocity runs monitored at 274 nm of purified Hfq (black line), and Hfq-NA samples (red/black line).

Analytical Ultracentrifugation Analysis

Sedimentation velocity experiments monitored at 274 nm showed a significant difference between purified Hfq and Hfq-NA (Figure 5.1B). The normalized g(*s*) distribution of purified Hfq exhibited a single major species with a sedimentation coefficient (*s*) of *s* = 3.42 S <3.41, 3.44>. Analysis of the data using Sedanal (Stafford and Sherwood 2004) yielded a molecular weight of 64,815 D <59733, 70301>. This value is slightly lower than the expected value of 66,998 D for the hexamer and is consistent with the hexamer as the dominant Hfq species. The slightly lower than expected value can be explained by uncertainty in the partial specific volume employed.

The normalized g(*s*) distribution of Hfq-NA showed two species. The dominant species has the same sedimentation coefficient as Hfq_6. The second component has a sedimentation coefficient of s_2 = 7.42 S <7.31, 7.53 >. A similar sedimentation velocity profile has been

previously observed for Hfq (Arluison et al. 2002). The sedimentation coefficient distributions of Hfq and Hfq-NA were independent of loading concentrations from 3 to ~8 μM Hfq (moles hexamer). The faster moving species in the g(*s*) distribution of Hfq-NA could be a multimer of Hfq_6 (Arluison et al. 2002), but given the different UV absorbance spectra and sedimentation profiles of Hfq-NA and Hfq, a nucleic acid or nucleic acid-Hfq complex seemed more likely. Experiments described below indicate this species is a Hfq-nucleic acid complex.

Previous work has shown that A_n with n > 18 bind Hfq_6 with high affinity (K_d ~ 10 nM) and interacts with Hfq_6's distal surface (Mikulecky et al. 2004; Sun and Wartell 2006). Figure 5.2A shows that the addition of 3 μM FAM-labeled A_{18} to 3 μM Hfq-NA disrupts the 7.4 S species and yields a new macromolecule with a sedimentation coefficient of 3.8 S monitored at 495 nm or 274 nm. FAM-rA_{18} alone has a sedimentation coefficient close to 1 S. Experiments on FAM-rA_{18} binding to purified Hfq_6 show that the 3.8 S species corresponds to a 1:1 Hfq_6• FAM-rA_{18} complex (data not shown). Thus, A_{18} binds Hfq_6 displacing it from the complex constituting the 7.4 S peak. The small shoulder from 6 S to 8 S in the g(*s*) distribution of Hfq-NA and FAM-rA_{18} monitored at 274 nm may reflect released nucleic acid.

Sedimentation equilibrium runs were next carried out with Hfq-NA at 12K and 16K rpm at concentrations of 1.75, 3.5, and 7 μM moles Hfq hexamer (assuming absorbance at 274 nm is solely due to Hfq). Globally fitting the six data sets using Sedanal to a two species model yielded molecular weights of 68,977 <57,830, 78,263>, and 509,932 <327,529, 761,844>. Figure 5.2B compares the non-linear least squares fits of the two-species model and one-species model to independent sedimentation equilibrium data. The model of two independent non-interacting components gave a much better fit than the single species model. The lower molecular weight species corresponds closely to the Hfq hexamer. The ~509Kd species

correlates with the complex corresponding to the 7.4 S peak in the sedimentation velocity experiment and is much larger than a dimer of Hfq_6. Table S5.2 in supporting information summarizes sedimentation coefficients and experimentally derived molecular weights of the different Hfq species used in this study.

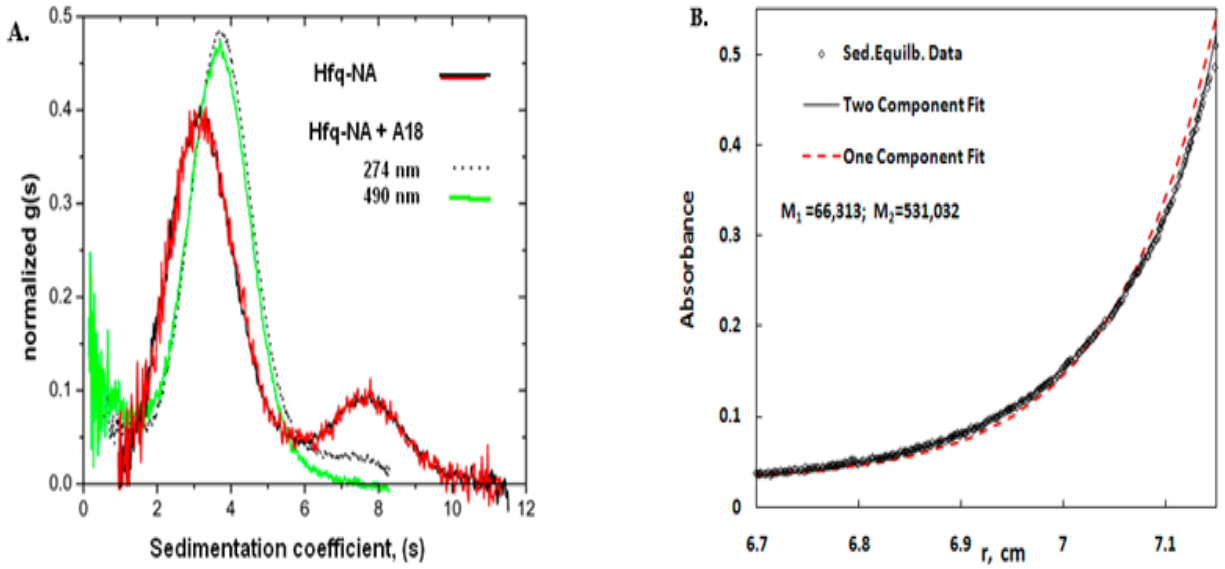


Figure 5.2: (**A.**) Effect of adding 3 μM FAM-A_{18} on sedimentation velocity distribution of 3 μM Hfq-NA; distribution of Hfq-NA alone monitored at 274 nm (red/black line), distribution of Hfq-NA and FAM-A_{18} monitored at 274 nm (black dotted line) and monitored at 495 nm (green line). (**B.**) Sedimentation equilibrium profile of Hfq-NA sample (circles) and non-linear least squares fit of one component (red dotted line) and two component models (solid line). The one component fit yielded a M.W. of 79842. The two component model yielded M.W. of 66,313 and 531,032 for the two species. The latter values are in very good agreement with a global fit to data obtained with another Hfq-NA sample (see text).

The nucleic acid contaminant is DNA

In order to determine the nature of the nucleic acid contributing to the 7.4 S peak, Hfq-NA was subjected to one of three treatments; i) phenol-chloroform extraction to remove Hfq, ii)

digestion with DNase I prior to phenol-chloroform extraction, or iii) digestion with RNase A prior to phenol-chloroform extraction. The aqueous phases from each of the above treatments were subjected to electrophoresis in a 1% agarose gel. The results indicate that DNA rather than RNA is the major nucleic acid component of Hfq-NA (Figure 5.3). The mobility of the DNA corresponded to the mobility of duplex DNAs ranging from ~400 to 1000 bp. A simplified analysis of the UV spectrum of Hfq-NA, assuming this spectrum is a linear combination of purified Hfq and 50% GC duplex DNA indicate a very small amount of DNA contamination (~ 0.3%).

DNA isolated from Hfq-NA following the phenol-chloroform extraction and agarose gel separation was PCR amplified using the tagged random primers as described in Materials and Methods. The amplified duplex DNA was digested with BamHI, cloned into pUC19 plasmid and sequenced. Table 5.1 lists sequence characteristics of 41 cloned DNAs. There were 24 different core sequences, 13 which were represented more than once, ranging in length from 60 to 567 bp. Most of the core sequences that were observed multiple times had slight differences at their ends (± 10 nts). This suggests they originated from different clones rather than a single plasmid. Four of the 24 fragments contained sequences from the expression plasmid. The others were from the *E coli* genome. The average GC content of the *E. coli* fragments was 52.6%, close to the average for the *E. coli* genome (50.8%). The DNA sequences were distributed throughout most of the *E. coli* genetic map.

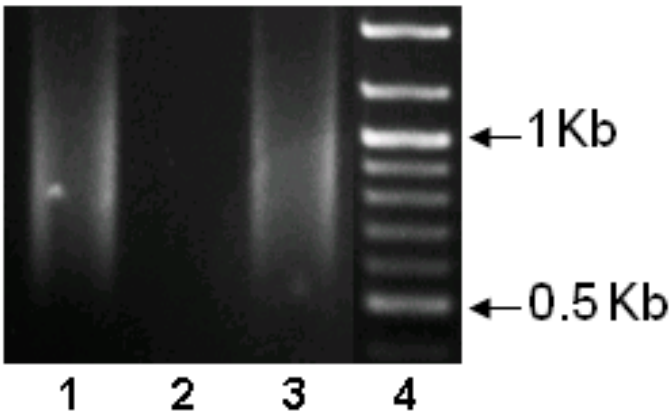


Figure 5.3: 1% agarose gel of aqueous phase of phenol extracted Hfq-NA samples first treated with RNase A (lane 1), DNase I (lane 2), or untreated (lane 3). DNA markers are in lane 4.

Thirteen of the 24 sequences in Table 5.1 had DNA helical axis curvature scores greater or equal to 10.0 using the BEND.IT algorithm (Munteanu et al. 1998; Vlahovicek et al. 2003). Previous studies have shown that DNA sequences known to be curved in solution have scores greater than 9 (Gabrielian et al. 1997; Munteanu et al. 1998). The average calculated curvature score over the *E. coli* DNA genome is 7.7 (Gabrielian et al. 1997). The DNA sequence F11which was found six times among the 41 clones had the highest curvature score, 15.5. The relatively high percentage of isolated segments with curvature scores 10 or higher (27/41) is consistent with the previous indication that Hfq has a binding preference for curved DNA (Azam and Ishihama 1999). An experimental assessment of helical axis curvature was carried out on six DNA sequences in Table 5.1 by comparing their electrophoretic gel mobility against standard DNA lengths (Figure S5.1). The ratio (L_R) of apparent length to actual length of four of the DNAs indicated significant helical axis curvature ($L_R > 1.15$). The other two DNAs examined had predicted curvature peaks near the ends of the duplexes which may explain their lower L_R values of 1.02 and 1.08.

Sequence characteristics of DNA segments amplified from DNA associated with Hfq.

DNA	Locus	Gene/functional description	Length (bp)	Copies	Curvature scores		% GC
					Dnase	Consensus	
G1	2434398..2433909	*nrdA*/ribonucleoside diphosphate reductase 1, α subunit	491	2	8.5	9	52
H1	2787761..2788044	*yfhB*/conserved hypothetical protein	283	1	11	10	53
H3	335274..335626	**sbmA*/ABC transporter trans-membrane	352	2	11.5	11	53
H5	4550393..4550611	**ytfR*/predicted sugar transporter subunit, ATP-binding	218	2	9	7	58
F1	3185823..3186066	*gshB*/glutathione synthetase	243	2	7.5	8	50
F2	1132955..1133187	**putP*/sodium/proline symporter	232	3	12	10	51
F5	4382281..4382555	**actP*/ acetate symporter (acetate transporter)	274	2	10	9	52
F8	3986385..3986781	**tnaB*/ tryptophan transporter of low affinity	396	2	10	9.5	49
B7	3830968..3831226	*xylR*/DNA-binding transcriptional activator; AraC family	258	2	13	9.5	52
B8	2325399..2325946	**mglC*/methyl-galactoside transporter subunit	547	2	13	11	51
B9	506621..507026	**emrE*/DLP12 prophage; multidrug resistance protein	405	1	15	11	43
D12	4361318..4361637	*qor*/quinone oxidoreductase, NADPH-dependent	319	1	10	9	54
F10	3938383..3938530	**yicO*/predicted xanthine/uracil permase	147	1	5.5	4.5	54
F11	3706138..3706379	**yhhS*/periplasmic protein; putative transporter	241	6	15.5	12	52
F15	4005299..4005492	**pstA*/phosphate transporter subunit	193	2	6.5	7	55
F16	1444080..1444470	**sapB*/antimicrobial peptide transporter subunit	390	1	11	8	53
H4	4673798..4673858	**yjjK*/ predicted transporter subunits of ABC superfamily	60	2	7	6	52
D11	2647738..2648025	*eutB*/ethanolamine ammonia-lyase	287	1	12	13	58
D11'	2657140..2657417	**eutH*/predicted inner membrane protein	277	1	9	8	53
C4	2980403..2980970	*cysJ*/sulfite reductase, alpha subunit, flavoprotein	567	1	8.5	8	56
H1'	Expression plasmid	pTYB11 sequence	200	3	13	12	54
F1'	Expression plasmid	pTYB11 sequence	183	1	9.5	7.5	49
D12'	Expression plasmid	pTYB11 sequence	230	1	9.5	7.5	52
F11'	Expression plasmid	pTYB11 sequence	199	1	9.5	7.5	49

Table 5.1: First column lists DNA designation. DNA segments from the same plasmid share the same letter, e.g. D11, D11'. Locus refers to the sequence of the genome of *E. coli* str. K12 substr.D10B given at NCBI. Some DNA sequences were found multiple times (copies). Curvature scores refer to the peak scores calculated with the two parameter sets available at the BEND.IT server (Munteanu et al. 1998; Vlahovicek et al. 2003). "Dnase" refers to scores obtained using the dinucleotide bendability parameters determined from DNase I digest data (Brukner et al. 1995), while "Consensus" refers to scores obtained using consensus dinucleotide bendability parameters (Dnase I digest and nucleosome positioning data (Gabrielian and Pongor 1996)). DNAs coding for transport or membrane proteins are noted with a *.

Additional characteristics of the cloned DNA sequences

It was noted that 13 of the 20 *E. coli* DNA sequences in Table 5.1 (65%) were from genes identified as transport or membrane proteins. The annotated *E. coli* K12 strain has a total of 4793 genes with 538 or 11.2% identified as genes coding for membrane proteins (i.e. transport, symport, antiport, export, efflux, permease or channel proteins) (Durfee et al. 2008). If the DNA associated with Hfq represented random fragments of the *E. coli* genome it is highly unlikely 65% would be from membrane protein genes. Given the above number of total genes (4793) and membrane protein genes (538), application of the hypergeometric probability distribution indicates a probability of ~10^{-8} for finding 13 membrane protein genes among 20 DNA segments by chance. Although this calculation assumes the cloned DNA segments are the size of genes, if one revisits the calculation assuming the DNA segments are 1/6 of an average gene length (as in this study), one obtains a similar probability.

A search was also made for conserved sequence elements among the *E coli* DNA segments using the Gibbs Centroid Sampler algorithm (Thompson et al. 2007). This method found one to three copies of an 8 base consensus motif in 16 of the 20 segments. The sequence was (A/T)T(A/G)TGCCG with 78% to 100% identity to the consensus observed at each position for the 24 occurrences. The relationship of this motif and the other DNA characteristics on Hfq binding affinity are uncertain, but they suggest that Hfq may have a sequence and/or structure preference in its interaction with DNA.

Binding affinity of wild-type Hfq to DNA

The binding affinity of wt Hfq to the 352 bp H3-DNA and 241 bp F11-DNA was examined using the electrophoresis gel mobility shift assay (Figure 5.4). The mobility of the

shifted DNA bands decreased with increasing Hfq concentration suggesting multiple Hfq bind to each DNA. This interpretation is supported by the observation that addition of unlabeled DNA to preformed DNA-Hfq complexes increased the mobility of the observed complex with a concomitant increase in free DNA (data not shown). The concentration of Hfq_6 that shifted 50% of the initial amount of free F11-DNA and H3-DNA to lower mobility complexes was ~ 400 nM. This value is higher than values previously reported for shorter DNA fragments (Azam and Ishihama 1999). We note that the PCR generated H3-DNA shows two bands in the absence of Hfq. Separate experiments showed that single-strand specific exonuclease I removed the slower moving band, while heating and quick-cooling the H3-DNA greatly enhanced this band. This suggests that the slower moving band is a folded form of one of the strands of the H3-DNA duplex. The observation that this band shifts with the addition of Hfq implies that Hfq can bind to DNA with single stranded character. Preliminary studies of Hfq binding to ten DNA single strands 60 to 80 nt long with various sequences indicate that Hfq has variable affinities for DNA strands with some showing little binding ($K_d > 2\mu M$) and one approaching an apparent $K_d \sim 200$ nM (unpublished data). We note that Mura *et al.* showed that Sm-like archeal proteins, which are related to Hfq, can bind supercoiled DNA and this interaction was inhibited by 20-30 nt single stranded DNAs (Mura et al. 2003).

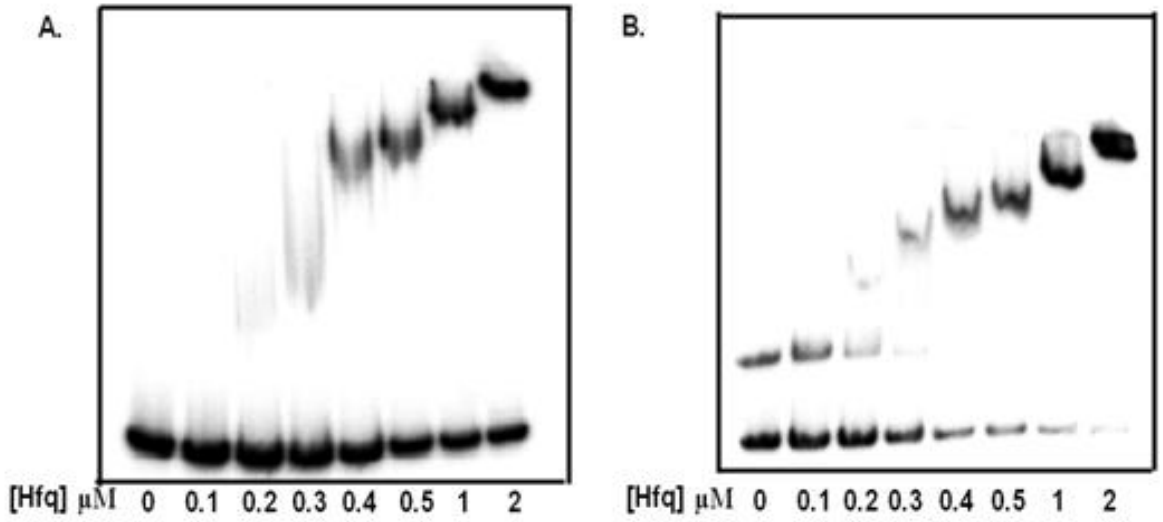


Figure 5.4: Hfq binding to ^{32}P-labeled DNA fragments (4 nM) assessed by the gel mobility shift assay. Hfq_6 concentrations added varied from 0 to 2 µM. (**A**.) Titration of the 241 bp F11-DNA. (**B.**) Titration of the 352 bp H3-DNA.

In order to confirm that the DNA associated with Hfq during the isolation procedure is due to a specific interaction with Hfq a competition assay was carried out. PCR products were generated using the conditions described in Materials Methods from 1) DNA extracted from the Hfq-NA sample, and 2) *E. coli* K12 genomic DNA isolated from cells. The effect of competing these two unlabeled pools of DNA segments against the preformed ^{32}P-labeled F11-DNA-Hfq complex was examined by a gel-shift assay. Figure S5.2 shows the results of this competition assay. 250 ng of PCR product from genomic DNA was needed to displace the F11-DNA-Hfq complex, while only 25 ng of product from the Hfq-NA DNA was required to displace the F11 DNA- Hfq complex. This result confirms that Hfq has a specific affinity for certain *E. coli* DNA sequences.

Hfq sites involved in binding DNA

The fact that A_{18} displaced Hfq from the Hfq-NA complexes in the sedimentation velocity experiment suggests that A_{18} and DNA interact with the same or overlapping Hfq binding sites. To help identify sites on Hfq that interact with DNA we examined the effect of altering Hfq residues on DNA binding. Nine mutant Hfq proteins each with a single residue changed to alanine and two truncated Hfq were expressed and purified as described in Materials and Methods. Locations of the nine single residue changes are shown in Figure 5.5. Hfq proteins designated Hfq- F39A, Hfq-F42A, Hfq-Q8A, Hfq-R16A, Hfq-R17A have single residue mutations on the proximal surface. Residues R16 and R17 also overlap onto the side or edge surface. The mutant Hfq designated Hfq-F11A and Hfq-R19A have mutations on the edge surface. Hfq-K31A and Hfq-Y25A have single residue mutations on the distal surface. Two mutant Hfq examined, Hfq-65 and Hfq-75, have all or part of the C-terminal domain deleted (Figure S5.3). The structure of the C-terminal domain is unknown. The amino acid sequence (residues 66-102) suggests they are structurally disordered (Sauter et al. 2003).

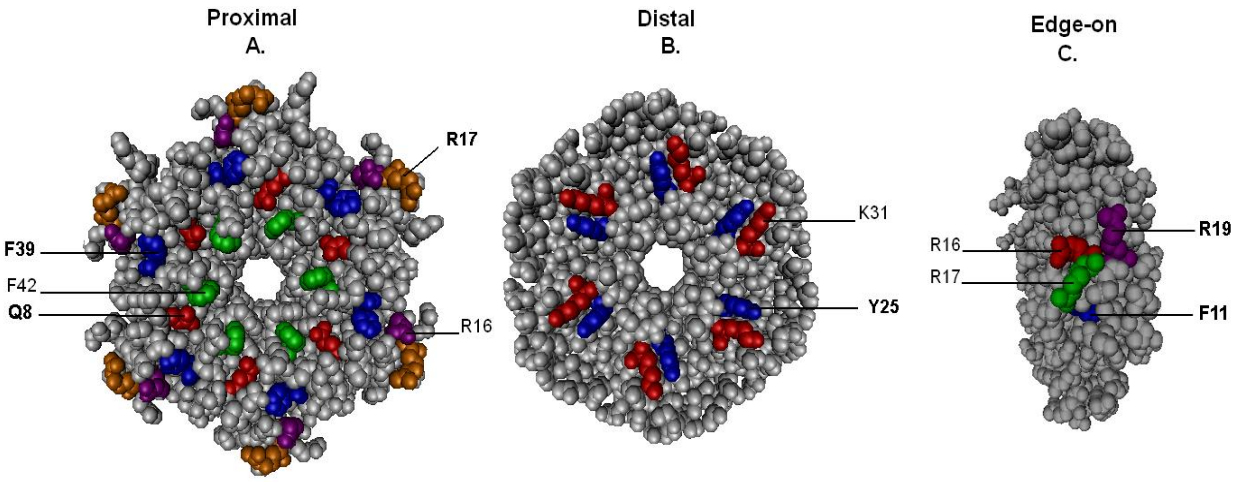


Figure 5.5: Space filling model of the toroidal part of *E. coli* Hfq_6 showing the locations of the nine single residue mutations examined from the proximal, distal and edge surfaces.

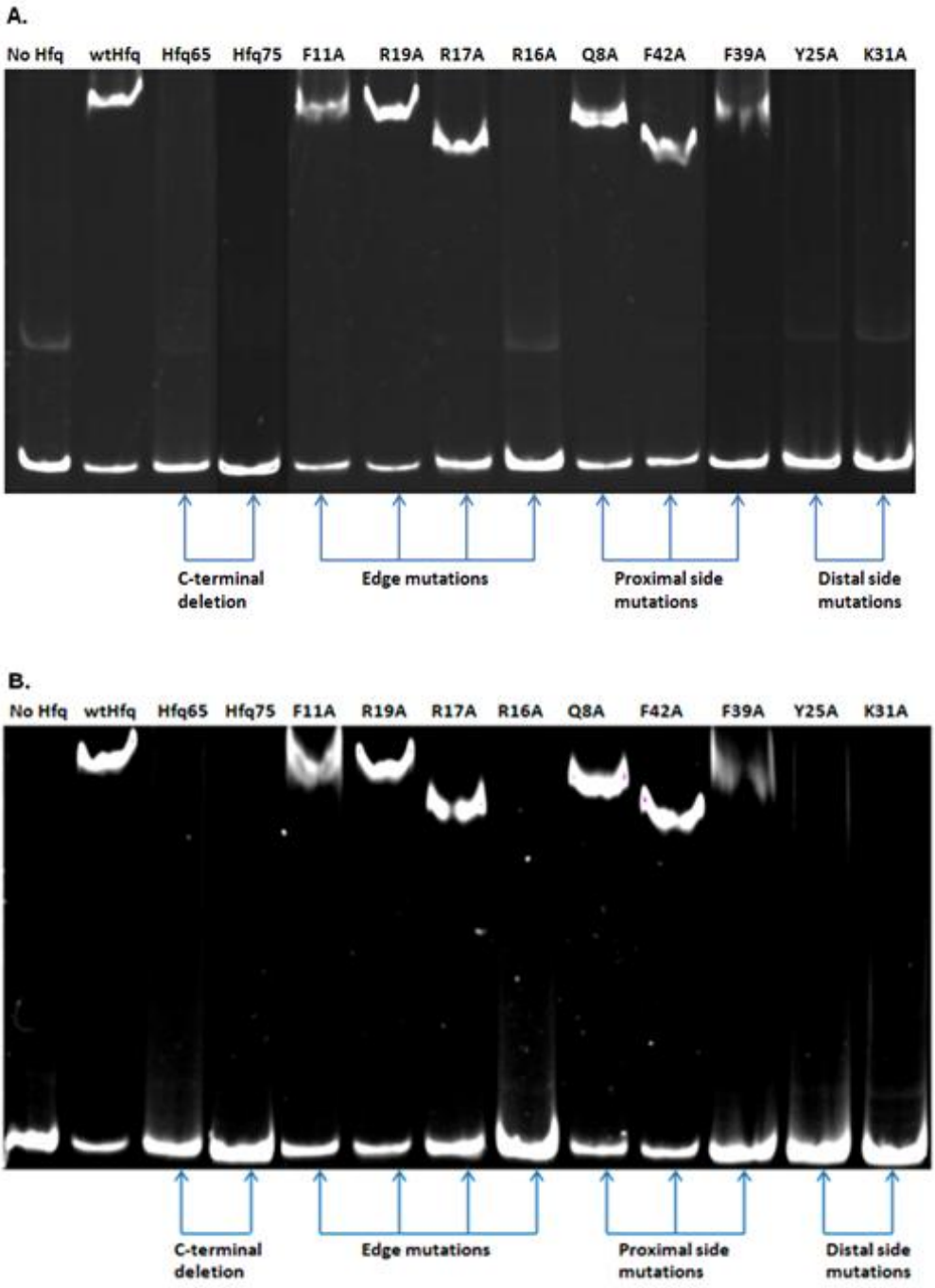


Figure 5.6: Gel shift assay of the relative affinity of wild-type and eleven mutant Hfq_6 for; (**A.**) H3-DNA, and F11-DNA (**B.**). 2 μM of each Hfq was added to 30 nM DNA in each lane except for the DNA only control lanes.

The relative affinity of the wt and mutant Hfq for two genomic DNA segments H3-DNA and F11-DNA was assessed using a gel shift assay (Figure 5.6). 2 μM of each Hfq was added to 30 nM of each DNA. When wt Hfq was added to either DNA, approximately 65 % of the free DNA band became a Hfq-DNA complex that migrated with low mobility. Hfq-F42A, Hfq-Q8A, Hfq-F39A and Hfq-R17A with alterations on the proximal surface also produced shifted bands although with varying characteristics. Hfq-F39A appears to bind both DNAs more weakly than wt Hfq since the shifted band has less intensity and is more diffuse. The mobility of the shifted bands for Hfq-42A and Hfq-R17A were faster than those for Hfq-39A, Hfq-Q8A and wt Hfq. Only the R16A mutation at the top of the proximal surface abolished Hfq binding to DNA.

Hfq with edge surface mutations F11A and R19A formed complexes with DNA. However both distal surface mutations, Y25A and K31A, eliminated Hfq binding to DNA. The latter result is consistent with the sedimentation velocity experiment in which A_{18} displaced DNA from Hfq. Both distal surface mutations were previously shown to affect Hfq binding to poly(A) sequences (Mikulecky et al. 2004; Sun and Wartell 2006; Arluison et al. 2007) .

A competition gel assay experiment in which DsrA or RprA is added to wt Hfq and F11-DNA provided additional evidence that Hfq's distal surface is involved in binding DNA. DsrA binds primarily to the proximal surface of Hfq (Brescia et al. 2003; Updegrove et al. 2008), while RprA interacts with both proximal and distal surfaces (Updegrove et al. 2008) . When DsrA is added to a F11-DNA-Hfq complex, a super shift in the gel mobility shift assay is observed (Figure S5.4A). This suggests Hfq can bind DsrA and DNA simultaneously. When RprA is added the F11-DNA-Hfq complex disappears and only free DNA is observed (Figure. S5.4B). This is consistent with RprA and F11-DNA sharing a common binding site on the distal surface.

Neither Hfq-65 nor Hfq-75 showed significant binding to either DNA in this assay (Figure 5.6). Previous studies show that Hfq-65 has similar, slightly reduced (2X) binding affinity to the RNAs DsrA and A_{27} compared to wt Hfq (Vecerek et al. 2008). In the above assay both Hfq-65 and Hfq-75 shifted all DsrA and A_{18} to a complex (data not shown). Thus the C-terminus appears to be more involved in binding DNA than these RNAs. We note that all mutant Hfq's used in this study including Hfq-K31A and Hfq-Y25A can bind DsrA and RprA with apparent K_d's varying from 20 to 150 nM (Mikulecky et al. 2004; Sun and Wartell 2006; Updegrove et al. 2008), and can shift all of the sRNA to a complex (unpublished data).

DISCUSSION

Hfq has been shown to be involved in various facets of RNA metabolism in the cell including sRNA ribo-regulation, and modulation of mRNA and sRNA half-life. Previous studies have demonstrated that Hfq also binds to DNA but with a weaker affinity than to RNA targets (Takada et al. 1997; Azam and Ishihama 1999). The current study set out to characterize what was thought to be RNA contaminants and revealed the contaminant to be predominantly genomic DNA from the host strain. The results support previous work indicating Hfq binds DNA (Takada et al. 1997; Azam and Ishihama 1999) and provides new information on the nature of Hfq-DNA interaction.

The sedimentation velocity and sedimentation equilibrium data together with the DNase digestion results indicate that Hfq-DNA complexes are responsible for the 7.4S sedimentation species. Addition of Hfq to duplex DNAs caused a steady decrease in gel mobility of the Hfq-DNA complexes, consistent with multiple Hfq hexamers binding these DNA probes. The average molecular weight of the Hfq-DNA complexes of Hfq-NA estimated from sedimentation equilibrium analysis, 509 kD, also suggests several Hfq_6 per complex for DNAs ~ 500 bp in length.

The ability of A_{18} to disrupt the Hfq-DNA complex in the sedimentation velocity experiment suggests that DNA binds to the distal surface of Hfq. The observation that two amino acid residues on the distal side of Hfq are important for DNA binding, and the influence of DsrA and RprA on Hfq-DNA complexes supports this hypothesis. The distal surface of *E. coli* Hfq_6 is known to be positively charged (Brennan and Link 2007) and this property could help drive its association with DNA. We note however that isolation of Hfq-NA involved extensive

washing with 1 M NaCl buffer, and the sedimentation experiments were conducted in 0.2 M NaCl or 0.5 M NaCl solutions with similar outcomes. The stability of Hfq-DNA complexes under these conditions implies that non-specific electrostatic interactions do not dominate Hfq-DNA binding.

Most mutant Hfq with residue changes on the proximal and edge surfaces exhibited Hfq-DNA complexes with a gel mobility that was the same as the wild-type Hfq-DNA complex (Figure 5.6). The mobility of the Hfq-DNA complexes involving Hfq-R17A and Hfq-F42A differed. Since complexes formed by Hfq-R17A and Hfq-F42A with RprA and DsrA have gel mobilities identical to wt Hfq((Updegrove et al. 2008) and unpublished data), the faster mobility of these mutant Hfq-DNA complexes appears to reflect aspects of Hfq-DNA interactions rather than a general defect in multiple Hfq binding to nucleic acids. Interestingly, Hfq-R16A, a mutation at the interface of the proximal and edge surface, did not bind DNA. This mutation reduces Hfq binding to domain II of DsrA by 7 fold, but it has a minimal effect on Hfq binding to RprA (Sun and Wartell 2006; Updegrove et al. 2008).

Both truncated forms of Hfq, Hfq-65 and Hfq-75, did not bind DNA in the gel assay. The C-terminal domain of *E. coli* Hfq has been shown to stabilize its hexamer form (Arluison et al. 2004) and truncating them could influence DNA or RNA-binding indirectly by reducing the amount of Hfq in hexamer form. While this cannot be ruled out, the same gel shift assay carried out with these truncated Hfq and rA_{18} or DsrA showed both RNAs shifting completely to a complex (data not shown). These results suggest that residues of the C-terminus beyond position 75 may be involved directly or indirectly in binding DNA.

The sequence analysis of the DNA fragments associated with Hfq revealed several interesting characteristics. 27 of the 41 cloned DNA segments had a helical axis curvature score

of 10 or higher compared to an average of 7.7 for the *E. coli* genome. The gel mobility of several of the DNA fragments verified they had significant helical axis curvature. While helical axis curvature appears to be correlated with Hfq affinity to DNA, it is unlikely to be the sole determinant. Prediction of helical axis curvature in the *E. coli* DNA genome using the BEND.IT algorithm indicates ~20% of DNA segments have curvature scores of 10 or higher and 4.7% have scores of 15 or higher (Gabrielian et al. 1997; Bolshoy and Nevo 2000). It is nonetheless of interest to note that an analysis of the 5% most curved DNA regions of the *E. coli* genome shows they are preferentially located 100 to 200 bases upstream of the nearest start codon (Bolshoy and Nevo 2000). While the results of this work do not address the functional significance of Hfq-DNA interaction, one may speculate that for some genes Hfq binding to DNA regions near promoters could be coupled to Hfq binding to and regulation of the corresponding transcribed mRNA. A recent work implied that Hfq affects the transcription of genes as well as influences post-transcriptional events (Le Derout et al. 2010).

A second perhaps more significant feature of the DNA associated with Hfq was the observation that a majority of the *E. coli* DNA segments (13 of 20) was from genes coding for membrane proteins. The estimated probability of this occurring by chance is extremely low (~ 10^{-8}). A recent study showed Hfq is located close to the inner membrane of *E. coli* as well as in internal regions of the cell when expressed at levels corresponding to the stationary phase of growth (Diestra 2009). When overexpressed, Hfq is also located at the outer membrane. It will be of interest to determine if the later characteristic is related to Hfq overexpression or if it occurs during physiological levels of Hfq expression (Sauter et al. 2003).

Acknowledgments

I gratefully acknowledge the contributions made by John Correia, Roberto Galletto, and Wlodzimierz Bujalowski in conducting the sedimentation velocity and equilibrium experiments used in this study. I also thank Xueguang Sun and Charles Terry for their assistance in the purification of mutant and wild type Hfq proteins used in this study and King Jordan and Jianrong Wang for helpful discussions on sequence analysis. I am grateful for the funding provided by the NASA Astrobiology Institute.

REFERENCES

Arluison, V., Derreumaux, P., Allemand, F., Folichon, M., Hajnsdorf, E., and Regnier, P. 2002. Structural Modelling of the Sm-like Protein Hfq from Escherichia coli. *J Mol Biol* **320**(4): 705-712.

Arluison, V., Folichon, M., Marco, S., Derreumaux, P., Pellegrini, O., Seguin, J., Hajnsdorf, E., and Regnier, P. 2004. The C-terminal domain of Escherichia coli Hfq increases the stability of the hexamer. *Eur J Biochem* **271**(7): 1258-1265.

Arluison, V., Mutyam, S.K., Mura, C., Marco, S., and Sukhodolets, M.V. 2007. Sm-like protein Hfq: location of the ATP-binding site and the effect of ATP on Hfq-- RNA complexes. *Protein Sci* **16**(9): 1830-1841.

Azam, T.A., Hiraga, S., and Ishihama, A. 2000. Two types of localization of the DNA-binding proteins within the Escherichia coli nucleoid. *Genes Cells* **5**(8): 613-626.

Azam, T.A. and Ishihama, A. 1999. Twelve species of the nucleoid-associated protein from Escherichia coli. Sequence recognition specificity and DNA binding affinity. *J Biol Chem* **274**(46): 33105-33113.

Bolshoy, A. and Nevo, E. 2000. Ecologic genomics of DNA: upstream bending in prokaryotic promoters. *Genome Res* **10**(8): 1185-1193.

Brennan, R.G. and Link, T.M. 2007. Hfq structure, function and ligand binding. *Curr Opin Microbiol* **10**(2): 125-133.

Brescia, C.C., Mikulecky, P.J., Feig, A.L., and Sledjeski, D.D. 2003. Identification of the Hfq-binding site on DsrA RNA: Hfq binds without altering DsrA secondary structure. *Rna* **9**(1): 33-43.

Brukner, I., Sanchez, R., Suck, D., and Pongor, S. 1995. Sequence-dependent bending propensity of DNA as revealed by DNase I: parameters for trinucleotides. *EMBO J* **14**(8): 1812-1818.

Carmichael, G.G., Weber, K., Niveleau, A., and Wahba, A.J. 1975. The host factor required for RNA phage Qbeta RNA replication in vitro. Intracellular location, quantitation, and purification by polyadenylate-cellulose chromatography. *J Biol Chem* **250**(10): 3607-3612.

Diekmann, S. 1992. Analyzing DNA curvature in polyacrylamide gels. *Methods Enzymol* **212**: 30-46.

Diestra, E., Cayrol,B., Arluison, V., Risco, C. . 2009. Cellular electron Microscopy Imaging reveals the Loacalization of the Hfq Protein Close to the Bacterial Membrane *PLoS ONE* **4**(12): e8301.

Ding, Y., Davis, B.M., and Waldor, M.K. 2004. Hfq is essential for Vibrio cholerae virulence and downregulates sigma expression. *Mol Microbiol* **53**(1): 345-354.

Durfee, T., Nelson, R., Baldwin, S., Plunkett, G., 3rd, Burland, V., Mau, B., Petrosino, J.F., Qin, X., Muzny, D.M., Ayele, M. et al. 2008. The complete genome sequence of Escherichia coli DH10B: insights into the biology of a laboratory workhorse. *J Bacteriol* **190**(7): 2597-2606.

Gabrielian, A. and Pongor, S. 1996. Correlation of intrinsic DNA curvature with DNA property periodicity. *FEBS Lett* **393**(1): 65-68.

Gabrielian, A., Vlahovicek, K., and Pongor, S. 1997. Distribution of sequence-dependent curvature in genomic DNA sequences. *FEBS Lett* **406**(1-2): 69-74.

Geng, J., Song, Y., Yang, L., Feng, Y., Qiu, Y., Li, G., Guo, J., Bi, Y., Qu, Y., Wang, W. et al. 2009. Involvement of the post-transcriptional regulator Hfq in Yersinia pestis virulence. *PLoS One* **4**(7): e6213.

Grothues, D., Cantor, C.R., and Smith, C.L. 1993. PCR amplification of megabase DNA with tagged random primers (T-PCR). *Nucleic Acids Res* **21**(5): 1321-1322.

Hemsley, A., Arnheim, N., Toney, M.D., Cortopassi, G., and Galas, D.J. 1989. A simple method for site-directed mutagenesis using the polymerase chain reaction. *Nucleic acids research* **17**(16): 6545-6551.

Jousselin, A., Metzinger, L., and Felden, B. 2009. On the facultative requirement of the bacterial RNA chaperone, Hfq. *Trends Microbiol* **17**(9): 399-405.

Kajitani, M., Kato, A., Wada, A., Inokuchi, Y., and Ishihama, A. 1994. Regulation of the Escherichia coli hfq gene encoding the host factor for phage Q beta. *J Bacteriol* **176**(2): 531-534.

Laue, T.M., Shah, B.D., Ridgeway, T.M. and Pelletier, S.L. 1992. *Analytical Ultracentrifugation in Biochemistry and Polymer Sciences (Eds: Harding, S.E., Rowe, A.J., and Horton, J.C)* Royal Society of Chemistry, Cambridge, U.K.

Le Derout, J., Boni, I.V., Regnier, P., and Hajnsdorf, E. Hfq affects mRNA levels independently of degradation. *BMC Mol Biol* **11**: 17.

Liu, S. and Stafford, W.F., 3rd. 1995. An optical thermometer for direct measurement of cell temperature in the Beckman instruments XL-A analytical ultracentrifuge. *Anal Biochem* **224**(1): 199-202.

Majdalani, N., Vanderpool, C.K., and Gottesman, S. 2005. Bacterial small RNA regulators. *Crit Rev Biochem Mol Biol* **40**(2): 93-113.

Meibom, K.L., Forslund, A.L., Kuoppa, K., Alkhuder, K., Dubail, I., Dupuis, M., Forsberg, A., and Charbit, A. 2009. Hfq, a novel pleiotropic regulator of virulence-associated genes in Francisella tularensis. *Infect Immun* **77**(5): 1866-1880.

Mikulecky, P.J., Kaw, M.K., Brescia, C.C., Takach, J.J., Sledjeski, D., and Feig, A.L. 2004. *Escherichia coli* Hfq has distinct interaction surfaces for DsrA, rpoS and poly(A) RNAs. *Nat Struct Mol Biol* **11**(12): 1206-1214.

Munteanu, M.G., Vlahovicek, K., Parthasarathy, S., Simon, I., and Pongor, S. 1998. Rod models of DNA: sequence-dependent anisotropic elastic modelling of local bending phenomena. *Trends Biochem Sci* **23**(9): 341-347.

Mura, C., Kozhukhovsky, A., Gingery, M., Phillips, M., and Eisenberg, D. 2003. The oligomerization and ligand-binding properties of Sm-like archaeal proteins (SmAPs). *Protein Sci* **12**(4): 832-847.

Papenfort, K. and Vogel, J. 2009. Multiple target regulation by small noncoding RNAs rewires gene expression at the post-transcriptional level. *Res Microbiol* **160**(4): 278-287.

Philo, J.S. 2006. Improved methods for fitting sedimentation coefficient distributions derived by time-derivative techniques. *Anal Biochem* **354**(2): 238-246.

Robertson, G.T. and Roop, R.M., Jr. 1999. The Brucella abortus host factor I (HF-I) protein contributes to stress resistance during stationary phase and is a major determinant of virulence in mice. *Mol Microbiol* **34**(4): 690-700.

Sauter, C., Basquin, J., and Suck, D. 2003. Sm-like proteins in Eubacteria: the crystal structure of the Hfq protein from Escherichia coli. *Nucleic Acids Res* **31**(14): 4091-4098.

Schuck, P., Perugini, M.A., Gonzales, N.R., Howlett, G.J., and Schubert, D. 2002. Size-distribution analysis of proteins by analytical ultracentrifugation: strategies and application to model systems. *Biophys J* **82**(2): 1096-1111.

Sonnleitner, E., Hagens, S., Rosenau, F., Wilhelm, S., Habel, A., Jager, K.E., and Blasi, U. 2003. Reduced virulence of a hfq mutant of Pseudomonas aeruginosa O1. *Microb Pathog* **35**(5): 217-228.

Stafford, W.F. and Sherwood, P.J. 2004. Analysis of heterologous interacting systems by sedimentation velocity: curve fitting algorithms for estimation of sedimentation coefficients, equilibrium and kinetic constants. *Biophys Chem* **108**(1-3): 231-243.

Sun, X. and Wartell, R.M. 2006. Escherichia coli Hfq binds A18 and DsrA domain II with similar 2:1 Hfq6/RNA stoichiometry using different surface sites. *Biochemistry* **45**(15): 4875-4887.

Takada, A., Wachi, M., Kaidow, A., Takamura, M., and Nagai, K. 1997. DNA binding properties of the hfq gene product of Escherichia coli. *Biochem Biophys Res Commun* **236**(3): 576-579.

Thompson, W.A., Newberg, L.A., Conlan, S., McCue, L.A., and Lawrence, C.E. 2007. The Gibbs Centroid Sampler. *Nucleic Acids Res* **35**(Web Server issue): W232-237.

Tsui, H.C., Leung, H.C., and Winkler, M.E. 1994. Characterization of broadly pleiotropic phenotypes caused by an hfq insertion mutation in Escherichia coli K-12. *Mol Microbiol* **13**(1): 35-49.

Updegrove, T., Wilf, N., Sun, X., and Wartell, R.M. 2008. Effect of Hfq on RprA-rpoS mRNA pairing: Hfq-RNA binding and the influence of the 5' rpoS mRNA leader region. *Biochemistry* **47**(43): 11184-11195.

Valentin-Hansen, P., Eriksen, M., and Udesen, C. 2004. The bacterial Sm-like protein Hfq: a key player in RNA transactions. *Mol Microbiol* **51**(6): 1525-1533.

Vecerek, B., Rajkowitsch, L., Sonnleitner, E., Schroeder, R., and Blasi, U. 2008. The C-terminal domain of Escherichia coli Hfq is required for regulation. *Nucleic Acids Res* **36**(1): 133-143.

Vlahovicek, K., Kajan, L., and Pongor, S. 2003. DNA analysis servers: plot.it, bend.it, model.it and IS. *Nucleic Acids Res* **31**(13): 3686-3687.

Waters, L.S. and Storz, G. 2009. Regulatory RNAs in bacteria. *Cell* **136**(4): 615-628.

SUPPLEMENTARY MATERIAL

Experimental assessment of DNA curvature

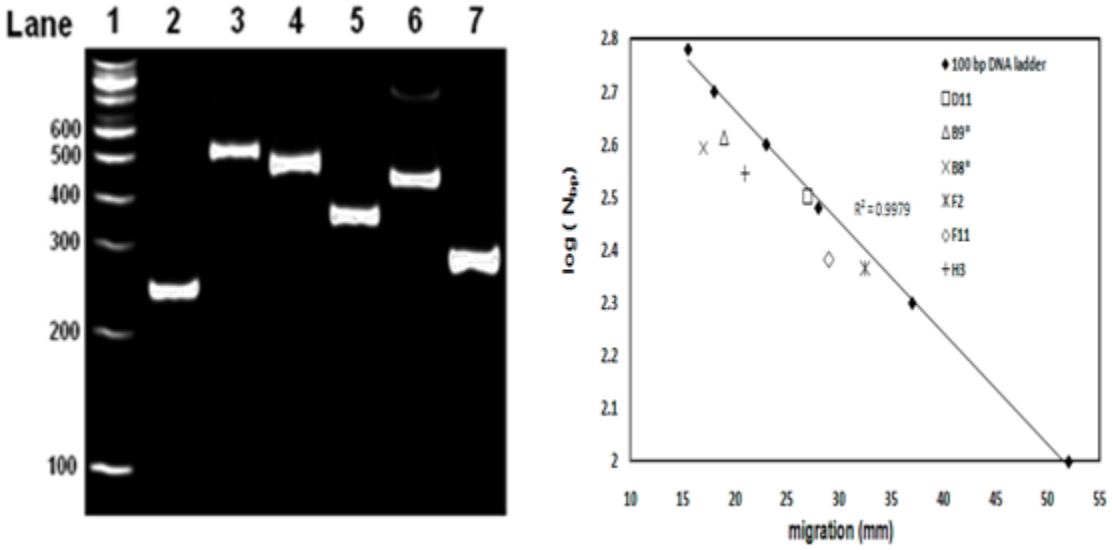


Figure S5.1: Helix axis curvature of six DNA segments was assessed by polyacrylamide gel electrophoresis[1]. DNA segments were electrophoresed in a 10% polyacryamide gel (30:1 acrylamide:bisacrylamide) in 0.5 X TBE buffer at 4^{o} C. A 100 bp DNA ladder was used as standard DNA lengths. A typical gel run is shown on the left with DNAs; F2 (lane 2), B8* (lane 3), B9* (lane 4), a 355 bp control (lane 5), H3 (lane 6) and F11 (lane 7). On the right is a plot of log of DNA length in bp vs. distance migrated for DNAs in the ladder and six DNA segments. Line is best fit through points of DNA ladder standards (♦). The ratio of the apparent DNA length to the actual DNA length, L_R, is listed below. The larger the deviation of L_R above 1.0 is indicative of increased DNA curvature. B9* differs slightly in length from the B9 segment in Table 5.1 due to primer design considerations. B8* spanned base pairs 62 to 455 of the 547 bp B8 segment placing the predicted curvature peak in the middle. The DNA segments D11 and F2 which do not show anomalous mobility have their predicted curvature peaks ≤ 40 bp from one end.

DNA segment (length)	L_R values (based on four experiments)
D11 (287)	1.02 ± 0.02
B9* (410)	1.19 ± 0.01
B8* (393)	1.33 ± 0.05
F2 (232)	1.08 ± 0.04
F11 (241)	1.26 ± 0.02
H3 (352)	1.26 ± 0.04

Primer sequences for amplifying DNA clones

DNA Clone	Primer Sequence	Size (bp)
B8*	Forward: 5' CGGTGGGTTGTATTGAGCTCGG 3'	393
	Reverse: 5' ACTGGCAGTACATCATCAAAGG 3'	
B9*	Forward: 5' AAGCGGGCAGGAATCCTGGTC 3'	410
	Reverse: 5' TCCAAGCCAGCGTCTGAGC 3'	
D11	Forward: 5' AGCCCGCCCGGTGC 3'	287
	Reverse: 5' AATACCACCATCGGTATTCCGGGC 3'	
F2	Forward: 5' CCAAGCCATCAGTGAGATAATGG 3'	232
	Reverse: 5' AGCTGCGGCGACGATCCTTTAC 3'	
F11	Forward: 5' AAGCGGCGGCGGGACTG 3'	241
	Reverse: 5' AGGCCAGACGGCGTACTCTTCCG 3'	
H3	Forward: 5' CCAAGCCCCAGAAAGACGCC 3'	352
	Reverse: 5' CCAAGCCGGTGGGGGAG 3'	

Table S5.1: Primer pairs used to amplify six of the DNA clones for gel mobility and Hfq binding studies.

Molecular weights derived from sedimentation velocity and equilibrium analysis

Specie	***s*, (S units)**	**M.W. from SV**	**M.W. from SE**	**Expected M.W.**
Hfq_6	3.42 <3.41,3.44>	64815<59730,70300>	61475 <58800,64200>	66998
FAM-A_{18}	1.36 <1.35,1.37>	6306 <5960,6670>	--------	6401
Hfq-NA	3.35 <3.30,3.41> 7.42 <7.31,7.53>	---------	68977 66313* 509932 531032*	------ ------
Hfq_6 • FAM-A_{18}	3.78 < 3.75,3.83>	----------	68930 <67400,70400>	73399
Hfq-NA+ FAM-A_{18}	3.83<3.81,3.90>	----------	----------	----------

Table S5.2: Sedimentation velocity coefficients (*s*) and experimentally derived and expected molecular weights of Hfq species from analytical ultracentrifugation analysis. Hfq concentrations were 3- 12 μM (moles hexamer). Solvent was 0.5 M NaCl + 20 mM Tris unless otherwise noted. SV is sedimentation velocity, SE is sedimentation equilibrium.

* experiment done in 0.2 M NaCl+ 20 mM Tris. Molecular weight derived by SV only when single specie was evident in g(s) distribution. Uncertainties in the two M.W.'s of the Hfq-NA specie were approximately < 57830, 78260> and <327530, 761840>.

Competition of PCR products from Hfq-NA DNA and genomic DNA with Hfq-F11-DNA complex.

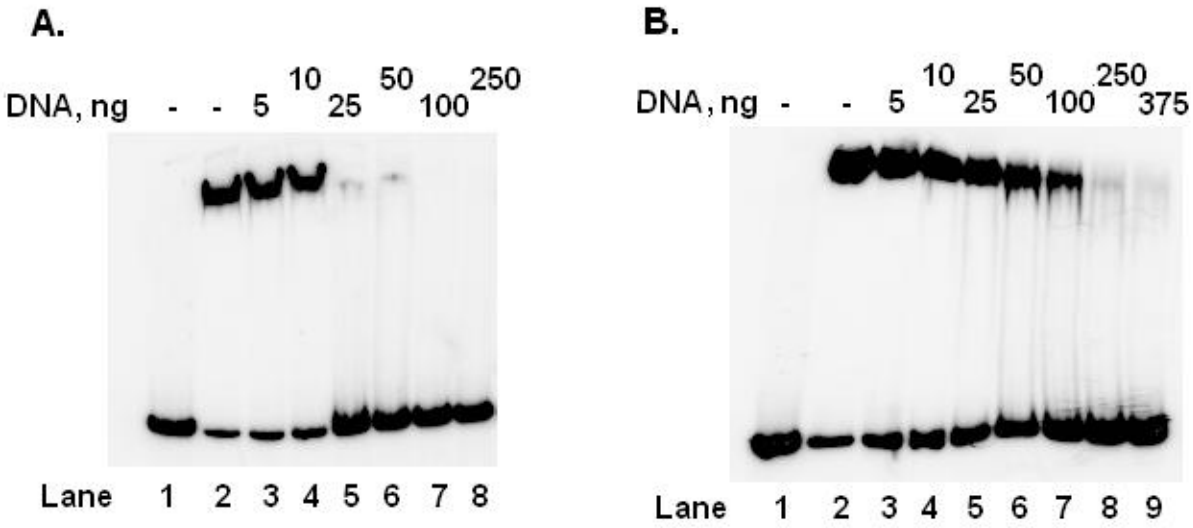


Figure S5.2: A 5% polyacrylamide gel was employed to assess the effect of adding PCR products amplified from DNA isolated from the Hfq-NA sample (**A**.) and PCR products amplified from *E. coli* K12 genomic DNA (**B**.) on preformed Hfq-F11-DNA complex. Each lane had 4 nM of ^{32}P-labeled F11-DNA. Lanes 2 through 8 (**A**.), or 2 through 9 (**B**.) contained 500 nM wt Hfq. The amount of the PCR products are given at the top in ng.

The 102 amino acid sequence of *E. coli* Hfq

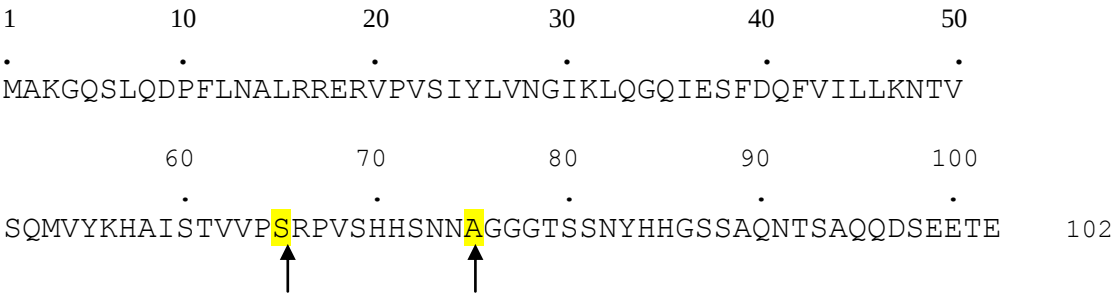


Figure S5.3: The last residue for Hfq-65 and Hfq-75 are illustrated with arrows and highlighted. We note that the crystal structure of the *E. coli* Hfq hexamer (Figure 5.5) was obtained with a truncated version of Hfq (residues 1 to 72). Residues 1 to 3 and 71 and 72 could not be located in the crystal structure [2].

Effect of adding DsrA and RprA on Hfq-F11-DNA complex

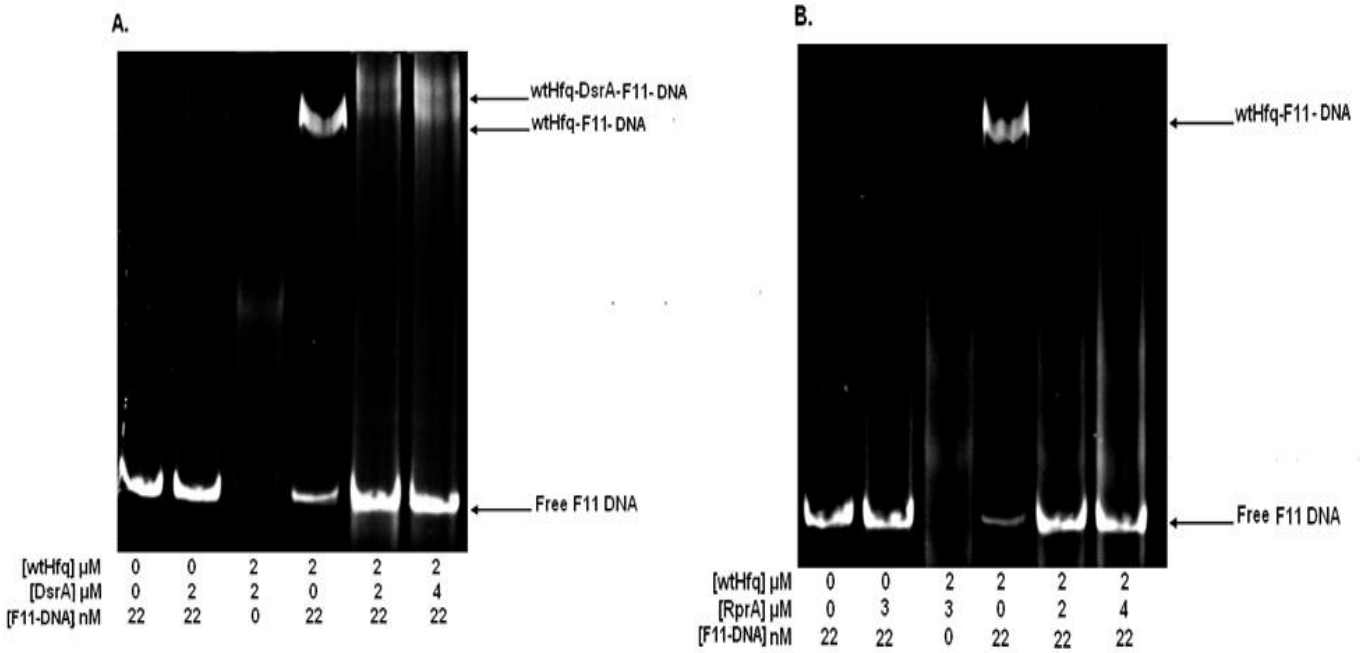


Figure S5.4: A 5% polyacrylamide gel was employed to assess the effect of adding DsrA (**A.**) and RprA (**B.**) on the wt Hfq-DNA complex. From left to right in each gel; 15 µl samples consisted of F11-DNA, F11-DNA and the sRNA, Hfq and sRNA, Hfq and F11-DNA, Hfq, F11-DNA and sRNA, and Hfq, F11-DNA and sRNA at higher concentration. The gel was run for 110 V for ~ 2.5 hrs to enable observation of the free F11 DNA at the bottom and its complex with Hfq at the top of the gel. The sRNA and the major Hfq-sRNA complexes ran off the gel. Faint bands corresponding to small amounts of a higher order Hfq-sRNA complex are observed in the third lane. Addition of DsrA results in a supershift of the Hfq-F11-DNA complex, while addition of RprA displaces the Hfq-F11-DNA complex. Both gels were stained with SYBR® Gold stain.

References

[1] S. Diekmann, Analyzing DNA curvature in polyacrylamide gels, Methods Enzymol 212 (1992) 30-46.

[2] C. Sauter, J. Basquin, D. Suck, Sm-like proteins in Eubacteria: the crystal structure of the Hfq protein from Escherichia coli, Nucleic. Acids Res. 31 (2003) 4091-4098.

CONCLUSIONS

This dissertation holds several chapters which have advanced the study of Hfq and its interaction with RNA targets and genomic DNA. Chapter 2-4 explores many questions regarding Hfq's role and mechanism for the riboregulation of the *rpoS* mRNA. Chapter 2 describes the kinetic and thermodynamic binding properties of Hfq to RprA and *rpoS* mRNA, and the effect of Hfq on enhancing RprA-*rpoS* and DsrA–*rpoS* binding. A major conclusion drawn from the study is the requirement of the entire untranslated leader region of *rpoS* mRNA for maximum Hfq enhancement of binding to RprA and DsrA. The amount of Hfq enhancement corresponded very well with what was observed *in vivo* in an *rpoS*-lacZ system involving Hfq and RprA. Moreover, *in vivo* studies also demonstrate that the entire leader region of *rpoS* needs to be present for maximum stimulation of *rpoS*-lacZ translation in the presence of Hfq and the sRNA. Thus we can gather that the *in vitro* binding studies of RprA and DsrA to *rpoS* presented in Chapter 2 recapitulate what is seen *in vivo* and can provide a model system to study the riboregulation of RpoS.

Although this study demonstrates the importance of the entire leader region of *rpoS* for Hfq enhancement of sRNA-*rpoS* binding, it did not reveal the exact mechanism for the enhancement. We proposed a model whereby Hfq binding to sites upstream of the RBS rearranges an inhibitory structure that prevents efficient sRNA binding. Indeed, more recent studies have localized two Hfq binding sites on the leader region of *rpoS* mRNA above the RBS and one of the sites is critical for enhancing sRNA binding. What needs to be investigated is if Hfq binding to these specific sequences rearranges the *rpoS* structure, or possibly destabilizes the intramolecular hairpin sequestering the RBS. To the latter end, our lab is engaged in UV-

absorbance melting studies of *rpoS* fragments possessing the region thought to be regulated by Hfq and RprA/DsrA to see if the presence of Hfq significantly destabilizes secondary structure(s) of the RNA. Likewise it remains to be determined if the binding of Hfq to DsrA or RprA is influencing the overall RNA structure in a way that makes it more amendable for *rpoS* binding. Recently our lab uncovered evidence that Hfq binding to DsrA does indeed alter the secondary structure of the RNA. UV-absorbance melting studies demonstrated that the presence of Hfq not only dramatically destabilizes a predicted hairpin structure that contains the *rpoS* binding region, but removes entirely another predicted hairpin at colder temperatures. It would be interesting to determine if this destabilization is critical for DsrA to bind *rpoS* mRNA.

Chapter 3 examined the binding of Hfq proteins with select mutations on all major surface regions of Hfq to RprA, DsrA, OxyS, and several *rpoS* fragments in an attempt to locate specific RNA binding surfaces on Hfq. We find that the proximal surface region of Hfq is likely to be involved in binding all three sRNAs and to a specific region on *rpoS* that is near the RBS. The distal side, however, is involved in binding a specific site on *rpoS* that is upstream of the RBS. Binding of Hfq to the distal specific and proximal specific sites on *rpoS* is required to get maximum Hfq enhancement of DsrA binding *rpoS*. The C-terminal domain and the outer circumference region seems to be less important for binding all of the above RNAs and in stimulating sRNA-mRNA binding. This begs the question: what is the exact function of the unstructured C-terminal domain?

Our results support a recent study showing *in vivo* that the expression of RpoS can occur when wt Hfq is replaced by Hfq-65; however, this conflicts with results obtained previously from the Bläsi lab showing that not only was Hfq-65 defective in promoting *rpoS* translation and facilitating sRNA riboregulation of other target mRNAs *in vivo*, but also failed to bind the *rpoS*

transcript and other mRNAs *in vitro*. As discussed in the Discussion section of Chapter 3, we provide a possible explanation for the latter observation, but the former still requires reconciliation. Ultimately the function of the C-terminal end of the *E. coli* Hfq (and Hfq from other γ- and β- proteobacteria) still requires further investigation, but would provide a greater overall understanding of the difference in the functional role Hfq plays in bacteria with and without the extended C-termini. It is worth mentioning that in Chapter 5 we found the C-termini are required for Hfq binding to genomic DNA; thus, this region could be important for whatever function Hfq is serving by binding to DNA. It would also be interesting to see if the C-termini are required for Hfq binding to other proteins. Recently Aiba's group showed a direct interaction of wild type *E. coli* Hfq to a specific region on the RNase E enzyme. It would be interesting to see if the Hfq C-termini are in anyway involved in that interaction.

In Chapter 3 we found that Hfq does not seem to enhance the binding of OxyS to *rpoS* mRNA. OxyS binding to *rpoS* in the presence and absence of Hfq is extremely weak relative to that seen with RprA and DsrA, while the affinity of Hfq for the three sRNAs is comparable. If the mechanism of OxyS repression of *rpoS* involves Hfq, as reported previously in multiple independent studies, but does not involve the enhancement of sRNA-mRNA binding, than how does Hfq and OxyS repress RpoS expression? A recent study showing that ectopic over expression of OxyS can disrupt DsrA activation of *rpoS* would suggest Hfq could be a limiting factor for sRNA signaling, and that OxyS could sequester Hfq from sRNAs involved in the activation of *rpoS*. The displacement of Hfq from DsrA to OxyS would lower the half-life of DsrA in the cell; and OxyS could be competing Hfq away from the essential sites on the *rpoS* leader region that seems to be necessary for the enhanced binding of DsrA and RprA to *rpoS*. However, to truly confirm that OxyS represses RpoS expression by disrupting the sRNA

signaling through the competition of Hfq would require accurate knowledge of the intracellular concentration of Hfq and the mRNA and sRNAs, and the affinity of Hfq for these RNAs; then thermodynamically one can predict (or estimate) the proportion of Hfq binding to each RNA. It is also worth noting that the presence of Hfq and OxyS was shown to reduce the stability of the RpoS protein, which hints at posttranslational regulation of RpoS by OxyS and Hfq. Regardless though, more studies will have to be made in order to really pin down the exact method of RpoS repression by OxyS and Hfq.

Chapter 4 dealt with determining the stoichiometry of Hfq binding to some of its RNA targets. This has recently been a contentious issue with multiple biophysical methods indicating different results. We show in this chapter using an array of different methods that Hfq binds to the tested RNAs in a 1:1 stoichiometry. What makes our result solid is that unlike previous methods used to determine stoichiometry, the methods we employed are more direct measurements of the molecular weight of protein-RNA complexes; as opposed to previous inferences from model fitting. However, we only show the stoichiometry of Hfq to the sRNAs DsrA, RprA, and OxyS, and to oligo A_{18}. The next step would be to determine the stoichiometry of Hfq to the leader portion of the *rpoS* mRNA. Since the study in Chapter 3 showed that Hfq binding to at least two sites on the *rpoS* transcript – the $(ARN)_4$ and the RBS sites- seems to be important for Hfq stimulation of DsrA binding, and that each site seems to be specific to a different surface region of Hfq, it would be important to see if only one Hfq hexamer binds both sites simultaneously, or if each site is occupied by one or more Hfq hexamers. Furthmore, since the RBS of *rpoS*, DsrA, and RprA share a common binding surface on Hfq's proximal side one may speculate that two different Hfq hexamers are needed for Hfq to bind one of these sRNAs and *rpoS*. Alternatively, the simultaneous binding of both RNAs to one Hfq hexamer may be part

of the driving force that pairs the two RNAs together. Quantitative immunoblotting may be a method that can shed light on the number of Hfq hexamers that binds *rpoS* mRNA alone and in complex with sRNA.

Chapter 5 really stood on its own from previous chapters; it involved characterizing the sequence and binding properties of genomic DNA fragments that were isolated from Hfq preps after the purification procedure. Surprisingly we found that a majority of the DNA sequences isolated were from genes encoding membrane proteins. However, this could be an artifact from over expressing the Hfq in *E. coli*. A recent study demonstrated that the over expression of Hfq in *E. coli* causes a considerable amount of Hfq to accumulate in and around the cell membrane, possibly as a stress response to over expression of the protein. Since it is well known in bacteria that transcription and translation can be coupled, and that nascent membrane polypeptides emerging from the ribosome are targeted to membrane bound receptors, the DNA of membrane genes may be tethered to the membrane and more able to interact with Hfq in the vicinity. Alternatively, Hfq may simply bind to regions on membrane protein genes and regulate transcription. What could be done to ascertain Hfq binding site on DNA would be to do immunoprecipitation experiments with anti-Hfq antibodies, remove RNA from the Hfq complex, and clone and sequence the eluted DNA. This would affirm actual binding of Hfq to these sequences under physiological Hfq concentrations and not concentrations affiliated with over expression.

What perhaps is not surprising about the DNA sequences we isolated from the Hfq preps is that most of them are predicted to display significant helical curvature in solution. It was previously shown that Hfq binds to a DNA fragment that was curved more strongly than a linear DNA of the same length, thus suggesting Hfq to bind curved regions affiliated with strong

promoters of genes. We further showed that Hfq binds single stranded DNA more strongly than double stranded DNA, and in a sequence specific manner. It would be interesting to determine what motif sequence Hfq is recognizing on single stranded DNA and to determine if the higher affinity for single stranded DNA would destabilize double stranded DNA, which could be a mechanism of influencing gene transcription.

VITA

TAYLOR B. UPDEGROVE

Taylor Updegrove was born in Fort Worth Texas, USA in 1980. He received his B.S. degree from Eastern Washington University in 2004 where he double majored in Chemistry and Biology, and minored in African American Studies. Upon graduation, Taylor Updegrove received the distinct academic honor of Summa Cum Laude, and received multiple awards from the Chemistry department and the Biology department. In 2005, Taylor Updegrove was awarded his M.S. degree from Eastern Washington University where he studied Microbiology under the advisement of Dr. Donald Lightfoot. From 2006 to 2011, Taylor Updegrove has been working under the supervision of Dr. Roger Wartell at the Georgia Institute of Technology, School of Biology, Institute of Bioscience and Bioengineering. His research has resulted in first author publications in the journals: *Biochemistry*, *Biochimica et Biophysica Acta*, and *RNA;* and middle author publications in the Journal *Neoplasma* and the *Journal of Natural Medicines*. At Georgia Tech, Taylor Updegrove was the recipient of the 2009-10 Molecular Biophysics Training Grant and the 2010-11 GAANN Doctoral Research Fellowship. He is also an active member of the Phi Sigma National Biological Honor Society, the National Honor Society of Phi Kappa Phi, and the American Association for the Advancement of Science. Taylor Updegrove expects to receive his Ph.D. from the Georgia Institute of Technology in August, 2011.