Title
Translational control of fis

Permalink
https://escholarship.org/uc/item/9659r72t

Author
Nafissi, Maryam

Publication Date
2013

Peer reviewed|Thesis/dissertation
Translational control of $fis$

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Biological Chemistry

by

Maryam Nafissi

2013
In order to translate an mRNA into a protein, the ribosome has to proceed through initiation, elongation, and termination steps guided by mRNA signals and auxiliary protein factors. Translation initiation is highly regulated in prokaryotes and requires the step-wise assembly of ribosomal subunits, mRNA, initiator tRNA, and initiation factors. Even though binding between the 16S rRNA (anti-Shine Dalgarno) and its complementary sequence on the mRNA (Shine Dalgarno) is thought to be the major interaction driving translation initiation, additional cis- or trans-acting factors can also provide control. In the current work, we describe novel mRNA elements important for efficient translation initiation of the fis mRNA, which encodes a highly abundant Escherichia coli DNA binding protein. The fis mRNA is part of a bicistronic message that contains dusB at the 5’ end encoding a tRNA modifying enzyme. Even though the cellular mRNA levels are comparable, Fis protein levels are much higher than DusB. Moreover, the translation of the downstream fis coding region does not depend on translation of
Rather than relying on a SD sequence, fis translation requires upstream RNA elements located within the 3’ end of dusB. These noncanonical elements include a conserved AU-rich sequence (AU) and a putative RNA stem-loop structure (SL). We propose that the AU element is involved in recruiting the 30S ribosome to the translation start site, whereas the SL indirectly enhances translation initiation via an anti-inhibitory mechanism. Previous studies have shown that cellular Fis levels are determined by multiple mechanisms of transcriptional regulation, and the present work now demonstrates that multiple and unusual mechanisms of translational control are also operating.

Fis functions as a global transcriptional regulator and has also been proposed to play a role in chromosome organization and compaction. To investigate Fis binding to the E. coli chromosome in vivo, genome-wide ChIP-chip binding studies were performed under different growth and mutant conditions. Fis was found to bind prolifically throughout the chromosome, and as expected, genome-wide Fis binding decreases under growth conditions that reduce fis expression. Surprisingly, Fis binding was found not to be generally influenced by other abundant nucleoid-associated proteins like HU and H-NS. We analyzed Fis binding in vitro to subclasses of Fis binding regions identified by the ChIP-chip profiles. Narrow Fis binding peaks (<500 bp) were found to primarily consist of a single high-affinity Fis binding site in vitro whereas broad binding peaks (≥1.5 kb) often contained 3-4 localized high-affinity Fis binding sites. These broad Fis binding regions are good candidates for chromosome organizing centers. In summary, work in this thesis advances our understanding of noncanonical translation initiation in prokaryotes and the targeted binding of the nucleoid-associated protein Fis to the E. coli chromosome.
The dissertation of Maryam Nafissi is approved.

    Feng Guo
    Beth Lazazzera
    Alexander van der Bliek
    Reid C. Johnson, Committee Chair

University of California, Los Angeles

2013
DEDICATION

To Reid and my family
## CONTENT

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Committee</td>
<td>iv</td>
</tr>
<tr>
<td>Dedication</td>
<td>v</td>
</tr>
<tr>
<td>List of Figures</td>
<td>vii</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>ix</td>
</tr>
<tr>
<td>Vita</td>
<td>xi</td>
</tr>
<tr>
<td>Chapter I: Introductions</td>
<td>1</td>
</tr>
<tr>
<td>References</td>
<td>19</td>
</tr>
<tr>
<td>Chapter II: Robust translation of the nucleoid protein Fis requires a remote upstream AU element and is enhanced by RNA secondary structure</td>
<td>31</td>
</tr>
<tr>
<td>References</td>
<td>58</td>
</tr>
<tr>
<td>Chapter III: A closer look at factors involved in fis mRNA translation</td>
<td>63</td>
</tr>
<tr>
<td>References</td>
<td>82</td>
</tr>
<tr>
<td>Chapter IV: Fis binding in vivo to the <em>E. coli</em> chromosome compared to naked DNA in vitro</td>
<td>85</td>
</tr>
<tr>
<td>References</td>
<td>105</td>
</tr>
<tr>
<td>Chapter V: Conclusion</td>
<td>109</td>
</tr>
<tr>
<td>References</td>
<td>112</td>
</tr>
<tr>
<td>Figure Number</td>
<td>Description</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Figure 1-1</td>
<td>Structures of prokaryotic 30S and 50S</td>
</tr>
<tr>
<td>Figure 1-2</td>
<td>Translation initiation in <em>Escherichia coli</em></td>
</tr>
<tr>
<td>Figure 1-3</td>
<td>Factors affecting translation initiation in trans</td>
</tr>
<tr>
<td>Figure 2-1</td>
<td>Structure and expression of the <em>dusB-fis</em> operon</td>
</tr>
<tr>
<td>Figure 2-2</td>
<td>Sequences influencing <em>fis</em> translation</td>
</tr>
<tr>
<td>Figure 2-3</td>
<td>A putative stem-loop region is important for high <em>fis</em> translation</td>
</tr>
<tr>
<td>Figure 2-4</td>
<td>Importance of AU sequence elements upstream of <em>fis</em></td>
</tr>
<tr>
<td>Figure 2-5</td>
<td>Effects of chromosomal mutations on Fis protein levels</td>
</tr>
<tr>
<td>Figure 2-6</td>
<td>Deletion and insertion mutations within the <em>fis</em> translation sequence</td>
</tr>
<tr>
<td>Figure 2-7</td>
<td>Phylogenetic conservation within the AU and stem-loop regions upstream of <em>fis</em></td>
</tr>
<tr>
<td>Figure 3-1</td>
<td>Effect of a stronger Shine Dalgarno sequence on <em>fis</em> translation</td>
</tr>
<tr>
<td>Figure 3-2</td>
<td>Effect of M24 and M25 mutations on <em>fis</em> translation</td>
</tr>
<tr>
<td>Figure 3-3</td>
<td>The experimental set-up of toeprinting experiments</td>
</tr>
<tr>
<td>Figure 3-4</td>
<td>30SIC formation on various <em>fis</em> RNAs assessed by toeprinting analysis</td>
</tr>
<tr>
<td>Figure 3-5</td>
<td>Model for translation control of <em>fis</em></td>
</tr>
<tr>
<td>Figure 4-1</td>
<td>Cellular Fis levels under different growth conditions</td>
</tr>
<tr>
<td>Figure 4-2</td>
<td>Fis binding throughout the <em>E. coli</em> chromosome</td>
</tr>
<tr>
<td>Figure 4-3</td>
<td>Fis binding under slower-growth conditions</td>
</tr>
<tr>
<td>Figure 4-4</td>
<td>Fis binding in the absence of H-NS and HU</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------------------------------------------------</td>
</tr>
<tr>
<td>4-5</td>
<td>Comparing Fis binding in vitro and in vivo</td>
</tr>
<tr>
<td>4-6</td>
<td>Dissecting Fis binding to broad chromosomal regions</td>
</tr>
<tr>
<td>5-1</td>
<td>5’UTRs of major nucleoid-associated proteins</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

I foremost thank Reid for his continuous guidance and patience. I admire his dedication to science and the high standards he applies to everything he does. I am also greatly indebted to many former and present Johnson lab members, in particular Meghan McLean, Jeannette Chau, Yana Bernatavichute, Noah Dowell, John Heiss, Stefano Stella, Dan Yoo, and Tracy Ying. I also greatly value the advice I received from my thesis committee members Feng Guo, Beth Lazazzer, and Alex van der Bliek. Feng Guo’s expertise in RNA-related work has been especially helpful and much appreciated. I thank Reid Johnson, Yana Bernatavichute and Stephen Hancock for help with the preparation of this manuscript.

Chapter 2 of this manuscript is a version of the following article, and I would like to thank Jeannette Chau and Jimin Xu for their contributions: Nafissi, M., Chau, J., Xu, J., and Johnson R.C. (2012) “Robust translation of the nucleoid protein Fis requires a remote upstream AU element and is enhanced by RNA secondary structure” Journal of Bacteriology, (2012) 194: 2458-2469.

The toeprinting data presented in chapter 3 was obtained at the IBMC institute at the University of Strasbourg/France and I thank the members of the Romby group for all their help, in particular Stefano Marzi.

Yana Bernatavichute and Jeannette Chau made considerable contributions to data presented in chapter 4 by performing the ChIP and growth rate/Western blotting experiments, respectively.

On a personal note, I am thankful for having met so many wonderful people during my graduate career here at UCLA. I especially appreciate all the love and support I received from
my good friends Anna D’Souza, Sabine Gorynia, Peggy Vorwald, Celine Marban, Estelle Toulme, and Pierre-Yves Yanni.

This research was supported by the 2012/2013 dissertation year fellowship, short-term EMBO fellowship, and USPHS grant GM038509.
VITA

2005-2006 Undergraduate Researcher – Dr. Feng Guo
Department of Biological Chemistry
University of California, Los Angeles, CA

2006 B.S., Biochemistry, Cum Laude
University of California, Los Angeles
Los Angeles, California

2006-2013 Graduate Student Researcher
Department of Biological Chemistry
University of California, Los Angeles

2008-2009 Teaching Assistant
Department of Microbiology, Immunology and Molecular Genetics
University of California, Los Angeles

2012-2013 Dissertation Year Fellowship
University of California, Los Angeles

2012 Short-term EMBO fellowship

Publications and Presentations


Chapter I

Introduction
It is a truth universally acknowledged that DNA makes RNA makes protein. This ‘Central Dogma’ highlights the importance of replication, transcription, and translation in all living things and a need for their tight regulation. Despite being well conserved in all kingdoms of life, these basic processes have varying degrees of complexity. Less evolved model organisms like the prokaryote *Escherichia coli* have been instrumental in studying their mechanistic details and regulation.

As single-cell organisms bacteria are at the mercy of their ever-changing environments and need to constantly adapt their gene expression, which can be energetically costly. Since transcription is the first step in the process of making a protein, its regulation seems the most cost-effective way to adjust their gene expression profile. However, many genes are regulated at the level of both transcription and translation, most likely because this allows for incorporation of multiple environmental signals and faster responses. For this reason, there are increasing numbers of mechanistic studies of post-transcriptional gene expression in bacteria (Picard et al., 2009). My thesis identifies previously unknown translational cis elements affecting post-transcriptional expression of the highly abundant nucleoid associated protein Fis. In order to provide some context for these findings, important aspects of bacterial translation will be reviewed with an emphasis on mechanisms enhancing bacterial translation initiation.

The basic process of translation has been the subject of much research and was revolutionized in 2000 when the first high-resolution atomic structures of the prokaryotic ribosome subunits became available (Ban et al., 2000; Harms et al., 2001; Schluenzen et al., 2000; Wimberly et al., 2000). The impact of these and subsequent structures of the translation apparatus by Venkatraman Ramakrishnan, Thomas Steitz, and Ada Yonath was acknowledged by the Nobel Prize in Chemistry to these scientists in 2009. The combination of structural and
functional data collected by these and other groups have provided great insight into how this complex macromolecular machine works. The most interesting discovery was probably that the ribosome is a ribozyme since the catalytic center is made up almost entirely of RNA (Nissen et al., 2000). At the core of translation lie two basic processes decoding and peptide formation, which are both conserved among all three kingdoms of life. During decoding an mRNA codon is matched up with its cognate tRNA that delivers an amino acid to the ribosome, thereby extending the growing polypeptide chain. In this manner, the ribosome converts the mRNA sequence into a polypeptide sequence using the three-letter code which gives rise to a folded protein.

**Fig. 1-1** Structures of prokaryotic 30S and 50S. The 30S subunit is from *T. thermophilus* (Ban et al., 2000) and the 50S from *H. marismortui* (Schluenzen et al., 2000). All ribosomal proteins are shown in grey, whereas ribosomal RNAs of the small and large subunit are shown in purple and red, respectively. The ribosomial subunits mostly consist of RNA highlighting the dominant role of ribosomal RNAs in translation. The different components of the 30S subunit are marked.

**Basic composition and organization of the *E. coli* ribosome**

All ribosomes contain a small and large subunit. The small subunit contains the decoding center and the large subunit the peptidyl transferase center. In *E. coli* the small subunit (30S) and large subunit (50S) associate to form the 70S subunit (S refers to the Svedberg constant which represents the sedimentation coefficient of the subunits). The 30S consist of a ~1,500-nt-long
16S rRNA and 21 ribosomal proteins. The majority of the small subunit consists of RNA with the proteins serving as architectural factors that hold the structure together (see Fig. 1-1). Since the first cryo-EM images of the 30S subunit resembled a chick, its different parts were named ‘head’, ‘body’, and ‘platform’ (Lata et al., 1996). The mRNA wraps around the cleft between ‘head’ and ‘body’, which is also called the ‘neck’ (Yusupova et al., 2001). This marks the location of the decoding center where the mRNA codons interact with the anti-codon of the aminoacylated tRNAs. The three tRNA sites (aminoacyl [A], peptidyl [P] and exit [E]) are located at the subunit interface. The flat top surface is where the 50S subunit binds thereby forming the 70S ribosome.

**The three phases of translation**

The overall process of bacterial translation is split up into three parts: initiation, elongation, and termination. Each step is also assisted by non-ribosomal proteins. For initiation of translation in bacteria three initiation factors (IF1, IF2, and IF3) are required (Fig. 1-2). In order for translation to occur, the small (30S) and large (50S) subunit are initially kept apart by IF3. First a 30S translation initiation complex consisting of mRNA, a special initiator tRNA (fMet-tRNA\textsubscript{Met} or iRNA), and 30S is formed with the help of all three IFs. This complex is stabilized by several important contacts. The iRNA is bound at the P site of the 30S and its anticodon is basepaired with the start codon of the mRNA. The direct interaction between the mRNA and 30S is mediated by a purine-rich sequence (Shine-Dalgarno or SD) right upstream of the start codon and the 3’ end of the 16S rRNA (anti-Shine-Dalgarno or aSD). IF2 aids in recruiting the 50S subunit to form an active 70S translation initiation complex, which is now ready for entering the elongation phase.
The rate of translation initiation is much slower than elongation (Lovmar and Ehrenberg, 2006; Rodnina et al., 2007; Wintermeyer et al., 2004). tRNAs carry the amino acids, bind to specific mRNA codon sequences, and rotate through three ribosomal binding sites giving rise to a growing peptide chain. Decoding initiates at the A site where an mRNA codon base pairs to its cognate aminoacylated tRNA, which is delivered by elongation factor Tu (EF-Tu). Once the correct codon-anticodon interaction is formed, EF-Tu hydrolyzes GTP and leaves the ribosome. A new peptide bond is formed when the α-amino group of the A-site tRNA then attacks the
carbonyl group of the P-site peptidyl group. This is followed by a translocation step supported by EF-G, which causes tRNA and mRNA to move in a synchronous fashion along the ribosome. The deacylated tRNA of the P-site moves to the E site from which it exits the ribosome. The A site is now empty and with the help of EF-Tu the aminoacylated tRNA that matches the next mRNA codon gets recruited.

This cycle will continue until a stop codon is encountered and translation gets terminated. Once a stop codon enters the A site it gets recognized by release factors RF1 and RF2 that then cleave the polypeptide chain from the P site tRNA. During ribosome recycling, the two subunits dissociate with the help of ribosome recycling factor (PRF), EF-G, and IF3 until the next round of translation begins (Hirokawa et al., 2002; Karimi et al., 1999).

**Comparison of canonical translation initiation in prokaryotes and eukaryotes**

Just by comparing the high-resolution X-ray structures of ribosomes from prokaryotes (70S) and eukaryotes (80S), it becomes apparent that eukaryotic translation is much more complicated (Ban et al., 2000; Ben-Shem et al., 2011; Klinge et al., 2011; Rabl et al., 2011; Schluenzen et al., 2000; Wimberly et al., 2000). Both complexes are composed of unique and conserved ribosomal proteins and RNA segments. The proteins and RNA bases shared by both make up the common structural core where conserved processes such as peptide formation and decoding are carried out (Smith et al., 2008; Spahn et al., 2001). Most of the kingdom-specific proteins and RNA edits are at the periphery, which explains why there is bigger variation in processes that require interactions with other cellular proteins (Smith et al., 2008; Spahn et al., 2001).

Translation initiation is one such process since it involves only three bacterial IFs in contrast to the 9-13 eukaryotic counterparts (eIFs) needed. The mRNA substrates also differ
tremendously because the 5′ and 3′ ends of eukaryotic mRNAs are modified (5′ cap and poly A tail, respectively), whereas prokaryotic mRNA ends are not. Moreover, prokaryotic mRNAs can contain multiple open reading frames (polycistronic), but eukaryotic messages usually only have one (monocistronic). The prokaryotic 30S can bind to the translation initiation region (TIR) independently of tRNA, but the 40S subunit has to form a pre-initiation complex (43S) first; this occurs when the eIF2-GTP-tRNA complex associates with the 40S guided by various eIFs (eIF1, eIF1A, eIF3, FeIF5). As mentioned earlier, the 30S binds directly to the TIR whereas the 43S binds to the cap structure at the 5′ end. This initial binding event is mainly driven by a series of protein-protein interactions in eukaryotes but RNA-RNA interactions in prokaryotes (Aitken and Lorsch, 2012; Jackson et al., 2010). Bacterial mRNAs recruit the 30S by base pairing to the aSD of the 16S rRNA and IFs help in recruiting the tRNA to complete the 30S initiation complex. Forming a stable eukaryotic translation initiation complex requires more eIFs as well as additional steps. Several factors (eIF3, PABP, eIF4B, eIF4E, eIF4A, eIF4G) are needed to recruit the 43S complex to the 5′ end. From here the 43S complex starts scanning the mRNA in a 5′ to 3′ direction to search for a start codon. Once it encounters an AUG with favorable flanking sequences (… RNNAUGG… for vertebrates) (Kozak, 1984, 1986, 1991), it pauses and eIF2-GDP as well as eIF5 get released. Although these flanking sequences are reminiscent of prokaryotic SDs their affinity for the 43S is not strong enough to directly position it next to the start codon. In both kingdoms the final step of translation initiation entails GTP hydrolysis and association of the large subunit. This description of events represents the most common translation initiation pathway in eukaryotes, but simpler mechanisms exist that require fewer steps and initiation factors. These alternative pathways are more similar to prokaryotic translation initiation and worth a more thorough look.
Non-canonical translation initiation in eukaryotes driven by IRES sequences

Under normal conditions cap-dependent translation drives the translation of most eukaryotic housekeeping genes. Although this pathway gets compromised during stress-inducing conditions such as viral infection, some cellular mRNAs with long and highly structured 5′ UTRs can still be translated (Johannes and Sarnow, 1998). This phenomenon was attributed to Internal Ribosome Entry Sites or IRESs, which were first discovered in picornaviruses (Jang et al., 1988; Pelletier and Sonenberg, 1988). Some host mRNAs were also found to contain IRES sequences, but they seem less efficient than their viral equivalents at initiating translation (Andreev et al., 2009; Bert et al., 2006; Young et al., 2008). Because of frequent experimental artifacts in experiments identifying these cellular IRESs, their abundance is probably lower than initially anticipated (Gilbert, 2010; Komar et al., 2012; Kozak, 2005b; Shatsky et al., 2010). Furthermore, these mRNAs still have 5′ caps so cap-dependent translation is still conceivable (Johannes and Sarnow, 1998; Lang et al., 2002). In contrast to their viral counterparts, the secondary RNA structures of cellular IRESs seem more diverse and unstable (Baird et al., 2007; Komar and Hatzoglou, 2011; Xia and Holcik, 2009). For this reason, viral IRESs have been studied more thoroughly and their cap-independent translational initiation pathway will be described in some detail.

When a picornavirus infects a host cell it immediately inhibits the host cap-dependent translation pathway by inactivating an essential component such as eIF4 (Etchison Hershey 1982). Unlike most host mRNAs, viral uncapped mRNAs can still be efficiently translated thanks to IRES elements in their 5′ UTRs. These highly structured regions can directly recruit the ribosome to the translation start site and were found to be common in other RNA virus families as well (Brown et al., 1994; Kuhn et al., 1990; Molla et al., 1992). Despite some variation in
their RNA structures, sequences, and interaction partners, all viral IRESs have one thing in common: a specific RNA segment directly interacts with a major component of the translation apparatus. For example, IRES elements found in encephalomyocarditis (EMCV) and food-and-mouth disease virus (FMDV) have high affinity for eIF4G and eIF4A, which usually assist in cap-binding (Kolupaeva et al., 1998; Lomakin et al., 2000). Another group of viral IRES elements (hepatitis C virus HCV, Classical Swine Fever Virus CSFV, porcine teschovirus 1 PTV1) were shown to interact directly with a specific portion of the 40S forming a stable binary complex (Pestova et al., 1998; Pisarev et al., 2004). Additional cellular proteins (e.g. pyrimidine binding protein for EMCV and FMDV) are also often required, but they seem to favor 40S binding indirectly by stabilizing secondary and tertiary RNA structures (Yu et al., 2011). RNA from dicistrovirus is a well-studied example of a specific RNA structure that directly recruits the 40S without the help of additional protein factors (Deniz et al., 2009; Jan and Sarnow, 2002; Pestova and Hellen, 2003; Pestova et al., 2004). Its 180-nt long IRES sequences folds into three pseudoknots (Costantino et al., 2008; Pfingsten et al., 2006), which interact with the 40S and cause significant conformational changes by opening the mRNA binding channel (Spahn et al., 2004). One IRES domain seems to bind to the E site while another one interacts with the P site (Pestova et al., 2004; Spahn et al., 2004). This is quite unusual because an itRNA normally binds to the P site during translation. Thus the folded IRES domain seems to act as an itRNA substitute allowing the first tRNA to enter the A site and peptide elongation to commence. Put differently, the dicistrovirus IRES allows translation initiation to be skipped altogether and to immediately start off with elongation. This is one of many interesting ways to initiate translation in a cap-independent manner, and it highlights that more simple methods for initiating translation can be employed in eukaryotes as well.
**Bacterial translation initiation factors**

Bacteria use a more simplified mechanism than eukaryotes to initiate translation, but assembly of the correct 70S translation initiation complex is still a complicated process. IF1 and IF2 have archaeal and eukaryotic counterparts, whereas IF3 is unique to certain bacterial species. IF3 promotes ribosome recycling and is the first IF to bind to the 30S and prevents the subunits from re-associating (Grunberg-Manago et al., 1975; Subramanian and Davis, 1970) (Fig. 2-1). It also acts as a fidelity factor that monitors the binding between the correct codon and anti-codon at the P-site (Guenneugues et al., 2000). IF2 is involved in this important interaction as well and docks the initiator tRNA at the P site of the 30S prior to 50S association (Simonetti et al., 2008). IF2 GTPase activity is triggered upon subunit association followed by IF2 ejection from the assembled 70S (Grigoriadou et al., 2007; Tomsic et al., 2000). IF1 promotes activity of IF2 and IF3 by increasing their binding to the 30S ribosome (Hartz et al., 1989). Together with IF2 it also ensures that only the itRNA binds to the P-site. By binding near the A site IF1 physically blocks (Moazed et al., 1995) any tRNAs from entering the A site until 70S formation, which causes its release.

**mRNA elements affecting translation initiation in cis:**

Since initiation is thought to be the rate-limiting step of translation, it is also highly regulated (Lovmar and Ehrenberg, 2006; Rodnina et al., 2007; Wintermeyer et al., 2004). Factors that decide whether a protein is translated can be classified as cis-acting (sequence or structure elements that are part of the mRNA) or trans-acting (external proteins, RNAs, or small molecules). A thorough description of these factors with some known examples will be given with an emphasis on positively-acting factors. In order to recruit the 30S ribosome to the translation initiation region, several specific mRNA contacts have to be formed. The strength of
these interactions is affected by several cis-acting elements, which include the SD sequence, AU-rich RNA enhancer sequences, and upstream RNA secondary structures.

**Shine Dalgarno sequences**

One of the most important steps during translation initiation is the binding between a purine-rich sequence immediately upstream of the start codon (SD), and the 3′ end of the 16S rRNA (aSD). Having a strong SD/aSD interaction is generally thought to be advantageous to translation (Hui and de Boer, 1987). The three factors that affect the strength of SD/aSD base pairing are sequence complementarity, accessibility of the SD sequence, and spacing of the SD to the start codon (de Smit and van Duin, 1994b; Ringquist et al., 1992). *E. coli* genes on average have SD sequences consisting of six bases (Vimberg et al., 2007). In general, the closer a SD is to the consensus sequences (AGGAGGU) the stronger the protein expression (Ma et al., 2002; Ringquist et al., 1992). However, stable secondary structures near the translation start site can prevent the ribosome from accessing strong SD sequence (de Smit and van Duin, 1994a, b). Therefore, a lack of secondary structure near the SD sequence is important. The optimal spacing between SD and AUG is specific to each SD sequence, most likely because of an ideal spatial orientation of the 3′ end of the 16S rRNA and anti-codon of tRNA bound in the ribosomal P-site (Gualerzi and Pon, 1990). Spacing of 3-8 nt are most common since longer distances seem to weaken translation initiation (La Teana et al., 1995; Ringquist et al., 1992). The choice of start codon also plays a role in assembly of translation initiation complexes. There is a preference for AUG, followed by GUG and UUG (Ringquist et al., 1992). In summary, a strong and accessible SD sequence upstream of an AUG seems to direct strong translation, but suboptimal variants are quite common since they leave room for regulation.

**AU-rich enhancer sequences**
In addition to having SD sequences, the 5′ UTRs of some *E. coli* genes also contain AU-rich sequences that have been proposed to enhance translation initiation. These kind of sequences are characterized by being rich in As and Us and are commonly located upstream of the SD. However, only a few native and phage-derived AU-rich leader sequences have been shown to enhance translation in *E. coli* (Komarova et al., 2005; Komarova et al., 2002; Olins and Rangwala, 1989; Zhang and Deutscher, 1992). The mechanism by which these sequences enhance translation remains unclear, but two modes of action seem possible. AU-rich sequences could either activate translation by serving as binding sites for ribosomal protein S1 (Boni et al., 1991; Tzareva et al., 1994) or by creating a translation initiation region (TIR) free of strong secondary structures (Kozak, 2005a).

S1 is the largest protein component of the 30S subunit and is absolutely required for translation (Sorensen et al., 1998). Cryo-EM structures of the 30S subunit in complex with S1 show that S1 directly contacts an 11-nt long RNA stretch right upstream of the SD (Sengupta et al., 2001). This is the strongest piece of evidence supporting S1 binding to mRNA leader sequences and thereby aiding 30S recruitment to the TIR. Nevertheless, a direct connection between S1 binding to native enhancer sequences and high levels of translation initiation remains elusive. Interestingly, S1 has recently been shown to possess RNA chaperone activity, which could present an alternative explanation for its positive role on translation initiation (Qu et al., 2012). Upon binding to AU-rich single-stranded upstream regions, S1 is in a good position to melt any potential inhibitory RNA secondary structures that form over the TIR. It is also possible that S1 binding plays no part in the positive effect of AU-rich sequences. Instead, AU-rich sequences by themselves tend to have little RNA secondary structure thereby providing easy access to the TIR. In this manner, the RNA sequence by itself without the help of a protein factor
like S1 can create favorable environments for translation. Taken together, AU-rich sequences seem to act as translational enhancers (Boni et al., 1991; Komarova et al., 2002), but their ability to compensate for weak SD sequences by cooperating with S1 remains to be firmly established (Ma et al., 2002; Vimberg et al., 2007).

**RNA secondary structures**

There are three reported cases of 5′ UTR RNAs folding into stable secondary structures that enhance translation initiation (Boni et al., 2001; Holmqvist et al., 2010; Zhang and Deutscher, 1989). These structures are located well upstream of the TIR and it is not obvious how they can activate translation. One example is the *rpsA* gene, which encodes the highly expressed ribosomal protein S1. Surprisingly, *rpsA* translation is not dependent on a SD, but relies on the formation of three successive hairpin structures upstream of the start codon. Skorski et al. (2006) exclude the presence of alternative SD sequences in the loop regions and instead propose that S1 binds to the loops to recruit the *rpsA* mRNA to the ribosome. A second example is the regulation of *csgD* translation, which is more complicated since small noncoding RNAs (OmrA/B) are involved, in addition to an upstream hairpin structure. When OmrA or OmrB bind, the structure opens up causing a decrease in *csgD* translation. The hairpin does not contain an alternative ribosome binding site, and it also does not prevent formation of downstream inhibitory structures, so its mechanism of action remains elusive. Thirdly, the translation of *rnd*, which encodes RNaseD, relies on a stable stem-loop structure followed by a stretch of U’s. It is possible that these elements compensate for the weak SD sequence and the rare start codon (UUG), but this has not been tested and so their functions are still unknown. Overall, secondary structures that stimulate translation are rare and act in some still unspecified manner.

**Leaderless mRNAs**
It is important to point out that even though some bacterial and phage messages have either no leader sequence or just a few nucleotides upstream of the start codon, they can still be translated in vitro and in vivo. This suggests the existence of a completely SD-independent translation initiation pathway, but this non-canonical mechanism applies only to a small set of *E. coli* genes. Under normal growth conditions most native leaderless (LL) mRNAs are translated quite inefficiently, but stressful conditions such as cold shock change that (Moll et al., 2002). Vesper et al. (2011) describe selective translation of LL mRNAs by specialized ‘stress-ribosomes’ under amino acid starvation. This system is set up by MazF, a sequence-specific endoribonuclease, that cleaves 3′ ends of 16S rRNAs and 5′ ends of specific mRNAs. Despite the lack of any SD/aSD interactions, some LL mRNAs can be translated quite efficiently under these conditions.

So far only a few mechanistic details have been obtained from studying interactions of LL phage-derived mRNAs (e.g. λ cI) and *E. coli* ribosomes. Without a SD the start codon becomes the main docking point for the ribosome making the codon/anti-codon interaction crucial. Therefore, a 5′ AUG start codon is absolutely required for translating LL mRNAs and no other start codon can be tolerated (Brock et al., 2008; O'Donnell and Janssen, 2001; Van Etten and Janssen, 1998). In order to initiate translation at a 5′ AUG, a novel pathway driven by a ternary 30S-IF2-itRNA complex or the intact 70S has been proposed (Balakin et al., 1992; Moll and Blasi, 2002; Moll et al., 2004; O'Donnell and Janssen, 2001; Udagawa et al., 2004). In conclusion, even in the absence of SD sequences and enhancer elements, ribosomes get recruited to correct translation start sites, but the efficiency seems low unless stress conditions are induced.
Factors affecting translation initiation in trans:

RNA-binding proteins

RNA binding proteins can also affect translation initiation. The majority of these RNA binding proteins inhibit translation by competing directly with 30S for binding to the TIR of a given mRNA (Fig. 1-3A). Ribosomal proteins, for example, autorepress their own translation by inhibiting 30S binding (Haentjens-Sitri et al., 2008; Yates and Nomura, 1980). Simultaneous binding is also possible because two ribosomal proteins trap the ribosome in an inactive complex (Marzi et al., 2007; Schlax et al., 2001). Most of these RNA binding proteins interact with hairpin (Cerretti et al., 1988; Moine et al., 1990) or pseudoknot structures (Haentjens-Sitri et al., 2008; Schlax et al., 2001). There are, however, also examples of proteins that positively influence translation, such as the two nucleoid-associated proteins HU and H-NS (Balandina et
H-NS promotes translation of six genes (*malT, dpiA, ynfF, irhA, yhbW, znuA*) that have weak SDs and unstructured leader sequences in common. H-NS binding to the AU-rich *malT* 5′ UTR has been shown to reposition the ribosome on the mRNA suggesting an auxiliary role for H-NS (Park et al., 2010). HU, on the other hand, has only been shown to enhance *rpoS* translation so far (Balandina et al., 2001), but this alternative sigma factor (σ^S^) for RNA polymerase is very important since it is the master regulator of stress response. HU usually binds DNA nonspecifically, but has high affinity for the *rpoS* 5′ UTR whereby it increases σ^S^ protein levels by some unknown mechanism. Overall, there are only a few examples of RNA binding proteins acting positively on translation, but it is possible that other nonspecific DNA binding proteins also have RNA regulatory functions that have not yet been discovered.

**Small noncoding RNAs**

Small noncoding RNAs (sRNAs) regulate gene expression by base pairing to their target mRNAs in trans. These sRNA can have varying lengths (40 to 400) and usually base pair with mRNAs in a discontiguous fashion. Because they only have partial complementarity to their targets they often require the assistance of the RNA chaperone Hfq (Masse and Gottesman, 2002; Sledjeski et al., 2001; Valentin-Hansen et al., 2004; Zhang et al., 2002; Zhang et al., 2003). One sRNA can have multiple targets and a single gene can be regulated by multiple sRNAs. The expression of most sRNAs is induced under specific growth conditions, which enables rapid responses to external stimuli (Altuvia et al., 1997; Andersen and Delihas, 1990; Chen et al., 2004; Chou et al., 1993; Coyer et al., 1990; Sledjeski et al., 1996; Vanderpool and Gottesman, 2004). Most sRNAs inhibit translation by blocking ribosome access to the TIR (Fig. 1-3B), but some stimulate mRNA degradation by RNaseE (Masse et al., 2003; Pfeiffer et al.,
There are a few examples of sRNA that enhance translation, and they seem to follow different pathways. Some mRNAs are in naturally repressed conformations that prevent ribosome binding and require sRNA binding to overcome this auto-inhibition (Lease et al., 1998; Majdalani et al., 1998; Prevost et al., 2007; Soper et al., 2010; Urban et al., 2007). One sRNA (SgrS) activates translation of its target (vigL) by stabilizing the transcript (Papenfort et al., 2013). Thus, depending on how sRNA binding affects the stability or ribosome access of a message, it can either activate or repress translation with the latter being more common.

**Riboswitches**

Riboswitches are widespread genetic switches that can affect transcription or translation. In this case, the trans-acting factor is a small molecule (cellular metabolite, ion, nucleotide, amino acid, cofactor, or tRNA) that binds to the 5′ UTR of a gene causing a change in the local secondary structure and affecting expression of the downstream gene (Baker et al., 2012; Cromie et al., 2006; Mironov et al., 2002; Nahvi et al., 2002; Wang et al., 2008; Winkler and Breaker, 2003; Winkler et al., 2002b). Metabolites often regulate genes that are involved in their transport or biosynthesis. Most RNAs turn off transcription or translation in their ligand-bound state as part of a feedback mechanism. Ligand-binding to 5′ UTRs usually inhibits translation by favoring hairpin structures that sequester the SD and AUG (Fig. 1-3C) (Miranda-Rios et al., 2001; Nou and Kadner, 2000; Winkler et al., 2002a). In the absence of the ligand an alternative structure (anti-sequestering helix) is preferred, which allows access to the TIR. There is only one example of a riboswitch acting positively on translation (Lemay et al., 2011) making inhibition their major route of action.
fis expression

The nucleoid-associated protein Fis is an important regulator of gene expression (Bradley et al., 2007; Cho et al., 2008; Kahramanoglou et al., 2011) and is implicated in playing a role in chromosome compaction (Schneider et al., 2001; Skoko et al., 2006). Fis synthesis is rapidly induced upon nutrient upshifts and is highly abundant under rapid growth conditions (Ball et al., 1992). Transcription of fis mRNA has been shown to be regulated at multiple levels (Ball et al., 1992; Mallik et al., 2006; Mallik et al., 2004; Nasser et al., 2001; Ninnemann et al., 1992; Pratt et al., 1997; Walker et al., 2004), but prior to this work, no translational regulatory mechanisms had been reported. The fis gene is in the same operon as dusB, which encodes a tRNA modifying enzyme. Transcription of the dusB-fis message is driven by a single promoter that is upstream of dusB (Ball et al., 1992; Mallik et al., 2004; Ninnemann et al., 1992), but we find that their translation is not coordinated. My dissertation will discuss new findings about fis mRNA translation and what questions remain open.


Lang, K.J., Kappel, A., and Goodall, G.J. (2002). Hypoxia-inducible factor-1alpha mRNA contains an internal ribosome entry site that allows efficient translation during normoxia and hypoxia. Molecular biology of the cell 13, 1792-1801.


Robust translation of the nucleoid protein Fis requires a remote upstream AU element and is enhanced by RNA secondary structure
Synthesis of the Fis nucleoid protein rapidly increases in response to nutrient upshifts, and Fis is one of the most abundant DNA binding proteins in *E. coli* under nutrient-rich growth conditions. Previous work has shown that control of Fis synthesis occurs at transcription initiation of the *dusB-fis* operon. We show here that while translation of the dihydouridine synthase gene *dusB* is low, unusual mechanisms operate to enable robust translation of *fis*. At least two RNA sequence elements located within the *dusB* coding region are responsible for high *fis* translation. The most important is an A + U element centered 35 nts upstream of the *fis* AUG, which may function as a binding site for ribosomal protein S1. In addition, a 44 nt segment located upstream of the AU element and predicted to form a stem-loop secondary structure plays a prominent role in enhancing *fis* translation. On the other hand, mutations close to the AUG, including over a potential Shine-Dalgarno sequence, have little effect on Fis protein levels. The AU element and stem-loop regions are phylogenetically conserved within *dusB-fis* operons of representative enteric bacteria.

Jimin Xu initiated this work by cloning and characterizing *fis-lacZ* deletion constructs, of which some are presented in Figure 2-2. Jeannette Chau performed Fis western blotting experiments shown in Figure 2-5 and I carried out all remaining experiments.

This chapter is a version of a published article: Copyright© American Society for Microbiology [Journal of Bacteriology, 194, 2012, 2458-2469, doi:10.1128/JB00053-12].
Introduction

Fis is an abundant small DNA binding protein in *Escherichia coli* that regulates diverse transcription, recombination, and replication reactions (Dillon and Dorman, 2010; Finkel and Johnson, 1992; Johnson et al., 2005). Fis is also proposed to participate in nucleoid compaction, a function it shares with other small nucleoid-associated proteins like H-NS and HU (Dame, 2005; Schneider et al., 2001; Skoko et al., 2006). The expression pattern of Fis is unique among nucleoid proteins, being high when cells are growing rapidly in nutrient-rich conditions, but low when cells are growing slowly in poor medium or in stationary phase (Ball et al., 1992; Mallik et al., 2006; Ninnemann et al., 1992; Osuna et al., 1995). Moreover, Fis synthesis is rapidly induced upon nutrient upshift leading to transient accumulations of very high cellular Fis levels, which decrease well before cells enter stationary phase (Ball et al., 1992; Mallik et al., 2006). Thus, the regulation of Fis synthesis is nearly opposite to that of another nucleoid protein Dps, which is induced as cells approach stationary phase (Ali Azam et al., 1999).

The gene encoding Fis is positioned second within a bicistronic operon present in *enterobacteriaceae* family members of the *gammaproteobacteria* (Fig. 2-1A) (Beach and Osuna, 1998). The *dusB* gene is located upstream of *fis* in the operon and encodes one of three synthases in *E. coli* that modify tRNAs by converting uridine to 5,6-dihydrouridine within D-loops (Bishop et al., 2002). The *dusB-fis* operon is transcribed from a single promoter that initiates RNA synthesis 32 bp upstream of *dusB* (Ball et al., 1992; Mallik et al., 2004; Ninnemann et al., 1992). In addition to the full length transcript of 1,382 nts, which terminates 62 bases downstream of *fis*, Northern blots have detected an ~860 nt transcript encoding *fis* that is believed to reflect the product of an RNA cleavage event within *dusB* (Ball et al., 1992; Ninnemann et al., 1992). This smaller *fis* transcript is present in similar amounts as the full
length transcript under most conditions but is undetectable in a relA spoT deletion strain upon amino acid starvation (Mallik et al., 2004). The functional role of this cleavage is unknown, but we show here that removal of the dusB coding segment around the cleavage site has no significant effect on Fis expression.

Transcription rates from the dusB-fis promoter vary tremendously with respect to growth conditions and closely follow Fis protein levels (Ball et al., 1992; Ninnemann et al., 1992; Osuna et al., 1995). The growth phase and growth rate regulated transcription is largely controlled by a nucleotide sensing mechanism and by DksA and guanosine tetraphosphate levels in a manner resembling the control of rRNA gene P1 promoters (Mallik et al., 2006; Mallik et al., 2004; Paul et al., 2004; Walker et al., 2004). Additional regulation is mediated by Fis itself, which binds to multiple sites within the promoter region to repress transcription, and by CRP and IHF that bind to upstream sites to moderately inhibit and activate transcription, respectively (Ball et al., 1992; Nasser et al., 2001; Pratt et al., 1997). The dusB-fis mRNA exhibits half-lives of 2-3 min under different growth stages where mRNA and Fis protein synthesis vary widely (Pratt et al., 1997).

Although changes in transcription can account for the growth phase and growth rate patterns of Fis expression, we wondered if atypical translation mechanisms may contribute to the rapid increase in Fis protein levels in response to nutrient upshifts and the high levels of Fis protein expressed under rapid growth rates. Related to this, we wondered if DusB was also synthesized in high amounts from the dusB-fis mRNA. Surprisingly, we find that DusB is synthesized in very low levels relative to Fis. A dissection of sequences upstream of the fis coding region reveals that efficient initiation of fis translation occurs by unconventional mechanisms whereby a canonical Shine-Dalgarno sequence is not operating but where unusually far upstream AU-rich sequences and RNA secondary structure play important roles.
Results

**Fis protein levels are much higher than DusB.** Quantitative immunoblotting was performed to compare the levels of DusB and Fis protein synthesized from the *dusB-fis* mRNA. For detecting DusB, a 10 amino acid residue FLAG tag was added to the 3′ end of the gene, which was then allele-replaced into the *E. coli* chromosome. Overnight LB cultures of RJ3454 (MG1655 *pyrE*⁺Δ*lacI*) and RJ3812 (RJ3454 *dusB::FLAG*) were diluted 1/50 into fresh LB and aliquots were subjected to immunoblotting with anti-Fis or anti-FLAG antibody after 60-100 min of growth, corresponding to the time frame of maximal Fis levels. Whereas Fis levels peaked at ≥ 50,000 dimers per cell, DusB protein was undetectable (Fig. 2-1B). Based on DusB-FLAG standards electrophoresed on the same gel, <500 monomers of DusB are estimated to be present per cell, which corresponds to <0.5% the levels of Fis protomers.

As an independent estimation of relative abundance, LB cultures of RJ3454 containing *dusB*-lacZ and *fis*-lacZ gene fusions on pACYC184 plasmids were assayed for beta-galactosidase activities. Ninety minutes after subculturing, corresponding to the peak time of Fis-LacZ activity (21,900 Miller units), the *dusB*-lacZ translation reporter generated only 155 Miller units or about 0.7% of the Fis-LacZ activity (Fig. 2-1C). The activity of the *dusB*-lacZ translation reporter reflects a low amount of translation initiated from *dusB* because a UAA introduced in place of *dusB* codon 6 reduced beta-galactosidase activity to <0.5 units. Transcription reporter activities were high for both *dusB* and *fis*, although beta-galactosidase units programmed by the *dusB*-lacZ fusion were 40% that of the *fis*-lacZ fusion (Fig. 2-1C).

To more directly measure RNA levels over the *dusB-fis* operon, RT-qPCR assays were performed on RNA isolated 60 min after subculturing into fresh LB. Primer pairs were designed
Fig. 2-1 Structure and expression of the dusB-fis operon. (A) Schematic diagram of the dusB-fis operon. The starting and ending sites of the dusB and fis coding regions are given relative to the primary mRNA start site from the promoter. N-terminal sequencing has confirmed the start site for the fis coding region (Johnson et al., 1988, Koch et al. 1988). Naturally occurring restriction sites used to construct several lacZ fusion reporters and PCR primer pairs for qRT-PCR used in panel B are denoted. (B) DusB and Fis protein and mRNA levels. DusB monomer and Fis dimer levels per cell were determined by quantitative Western blotting of RJ3812 (dusB::FLAG) with anti-FLAG antibody and RJ3454 (wild-type) with anti-Fis antibody. Quantitative RT-PCR was performed to measure absolute dusB and fis mRNA levels in RJ3454. dusB mRNA was probed with primer pairs a and b and fis mRNA with primer pair c (panel A). Protein levels are given for 90 min after 1/50 subculture in LB, corresponding to the peak Fis protein expression, and mRNA levels are given for 60 min, near the peak fis RNA expression. (C) Translation and transcription reporter activities for dusB and fis. For these and all subsequent reporter assays, saturated overnight cultures containing the pACYCl84-derived lacZ fusion plasmids were subcultured 1/50 into fresh LB, and beta-galactosidase activity was measured after 90 minutes of growth and reported in Miller units (mean and standard deviations). (D) Translation of dusB does not affect Fis protein levels. Mutant M20 has a UAA introduced at codon 6 within dusB in the context of the fis-lacZ reporters. Beta-galactosidase activities of the M20 translation (TL) and transcription (Txn) reporters are given relative to the wild-type (WT) reporters, which are set to 100.
to cover a region in the 5' end of dusB that is prior to the cleavage site detected by Northern blots along with a 3' segment of dusB and a segment within fis (Fig. 2-1A). RNA levels were quantified based on standards derived from cloned DNA of the dusB-fis operon. Very similar amounts of RNA were obtained over all three segments of the operon (Fig. 2-1B). We conclude that whereas RNA levels are high over the coding regions of dusB and fis, Fis protein synthesis is much greater than DusB.

Fig. 2-2 Sequences influencing fis translation. The importance of potential fis SD sequences were evaluated by mutating sequences preceding the fis translation start site (A) and measuring fis'-lacZ translation (black bars) and fis'-lacZ transcription (grey bars) reporter activities relative to wild-type (B). Complementary matches to the 16S rRNA anti-SD sequence in the wild-type sequence are underlined, and mutations introduced are denoted in grey. (C and D) Sequences within dusB affecting Fis protein expression. Segments deleted in Mdel1-6 are denoted by diagonal lines (C) and their respective fis-lacZ reporter activities are given (D). Except for Mdel5 and Mdel6, the boundaries of the deleted regions are at the given native restriction sites.
A nonsense mutation introduced into the N-terminal end (codon 6) of the *dusB* coding sequence had no effect on *fis*-′*lacZ* translation and *fis*-′*lacZ* transcription reporter activities (Fig. 2-1D). Therefore, translation of *dusB* mRNA is not important for stabilizing the *dusB*-fis message or for translation of *fis*.

**A putative Fis Shine-Dalgarno sequence is not important for high Fis synthesis.**

Shine-Dalgarno (SD) sequences are usually located 4-9 nts upstream from the translation start codon (Chen et al., 1994; Ringquist et al., 1992). A potential SD sequence (AAGAGCU, where underlined bases are a complementary match to the 16S rRNA sequence) is located 9 nts upstream from the *fis* start (AUG) codon (Fig. 2-2A). To test the importance of this sequence for *fis* translation, the GAG center was mutated to ATA and CTC. Surprisingly, beta-galactosidase activities of these mutant *fis*-′*lacZ* translation reporters (M18 and M39) remained high at 69% and 63% of wild-type levels (Fig. 2-2B). Substitution of the two Gs between the putative SD sequence and the AUG with Cs (M40) had no effect on reporter activity. The results of these mutants imply that a SD-like sequence is not functionally important for efficient translation initiation of *fis*.

**The 3′ end of *dusB* contains sequences that are needed for high *fis* translation.** To test whether upstream elements function in *fis* mRNA translation, various regions of the *dusB* coding region were deleted (Fig. 2-2C), and beta-galactosidase activities were measured for *fis*-′*lacZ* translation and *fis*-′*lacZ* transcription reporters (Fig. 2-2D). Deleting a major portion of *dusB* (Mdel1, +320 to +956, numbered relative to the *dusB*-fis transcription start site) had a large impact on *fis* translation, decreasing the activity of the translation reporter 13-fold with little effect on the transcription reporter. Subdividing this region showed that sequences upstream of the NruI site (+625), which is predicted to include the site of RNA cleavage
(~+525), have little effect on fis translation (Mdel2). However, sequences downstream of the NruI site had an increasingly large effect on fis translation. In particular, a 101 bp deletion (Mdel4) corresponding to -168 to -67 with respect to the fis ATG, decreased fis'-lacZ translation activity 9-fold, and a 76 bp deletion (Mdel6) corresponding to -100 to -24 relative to the fis ATG decreased fis'-lacZ translation by >150-fold with little effects on transcription. Removal of sequences between -400 and -100 also caused modest decreases in fis'-lacZ translation activities as evidenced by Mdel3 and Mdel5. Taken together, the results of these deletions imply that the segment from -100 to -24 with respect to the fis initiating ATG has a large positive effect on fis translation, encouraging the more refined dissection of this region discussed below.

An RNA stem-loop predicted to form at the 3' end of dusB is important for efficient fis translation. The 3' end of dusB was probed for sequences contributing to fis translation by performing a mutagenesis scan consisting of a series of 3-4 tandem changes to the complement base pair beginning at the Mdel6 upstream deletion boundary at -100 with respect to the fis ATG (Fig. 2-3A). The translational activities of these fis'-lacZ reporters are represented as bar graphs in Fig. 3B. Six of the 16 mutants caused greater than 2-fold decreases in fis translation with some (M2, M3, M16, and M21) resulting in >5-fold decreases. RNA secondary structure analysis of this segment using Mfold (Zuker, 1989) and Sfold (Ding et al., 2004) predict the discontinuous stem-loop structure depicted in Fig. 3C (ΔG = -18.5 kcal/mol). All the mutations that decrease fis'-lacZ reporter activity are located within base-paired regions of this structure; bases within the loop regions appear unimportant for translation. The mutations covering the lower stem have the largest effects, but the triple mutations M24 and M25, which should disrupt base pairing of the lower stem, do not significantly decrease fis'-lacZ reporter activities.
Fig. 2-3 A putative stem-loop region is important for high fis translation. (A) The 3' end of the dusB mRNA sequence is shown (-100 to -26 relative to the fis AUG). The black arrow marks the 5' boundary of the Mde6 construct. The locations of mutations with 3-4 tandem substitutions to the complementary sequence are marked with their names on top of the sequence. The black boxes highlight regions that upon mutation cause a 2- to 10-fold decrease in fis'·lacZ translation reporter activities. Arrows underline RNA sequences that are predicted to be involved in base-pairing. (B) Bar graph showing the relative activities of the fis'·lacZ translation reporters containing the mutations denoted in panel A. The black bars highlight mutants with ≥ 2-fold decreases in translation activities relative to WT. Transcription reporter activities for these mutants were 61-107% of WT. (C) Secondary structure of the 3' end of the dusB mRNA from M23 to M26 as predicted by M-fold and S-fold. Nucleotides implicated as being important for fis translation (panel B) are highlighted with black circles. (D) The same RNA secondary structure is shown as in panel C, but the wild-type sequence is replaced with the M21, M16, M3, and M4 mutations (black boxes), which disrupt the base-paired stems. These mutations are combined with mutations in the opposing strand (grey boxes) creating rescue mutants M21:25, M16:24, M3:2, and M4:1. (E) Beta-galactosidase activities of the mutant fis'·lacZ translation reporters together with the rescue mutants. Activities are relative to WT. Transcription reporter activities for these mutants were 65-125% of WT.
To evaluate the importance of this stem-loop structure for fis translation, mutations on either side of the stems were combined such that base pairing will be restored (Fig. 2-3D). In all cases, the compensating double mutations rescue fis-lacZ activities back to wild-type levels (Fig. 2-3E). These include combining the strong mutations M16 and M21 with the phenotypically silent mutations M24 and M25. Taken together, the results of the scan and combinatorial mutagenesis provide evidence that an RNA secondary structure located -50 to -94 nts from the Fis AUG directly or indirectly enhances fis translation.

---

**Fig. 2-4** Importance of AU sequence elements upstream of fis. (A) The 55 nt sequence immediately upstream of the fis AUG, which includes the intergenic region and the 3' end of the dusB message to the uaa termination codon, is given. The underlined arrow marks the lower right arm of the putative stem-loop structure shown in Fig. 3C. The four tandem base changes representing mutants M43, M27, M28, M30, and M29 within or adjacent to the two highlighted AU-rich segments are denoted. (B) Bar graphs showing the relative beta-galactosidase activities of the fis-lacZ translation and fis-lacZ transcription reporters containing the mutations introduced into the fis leader.
An upstream AU segment is important for high fis translation. In some systems that lack a strong SD sequence AU-rich sequences proximal to the translational start codon have been found to be important for efficient translation (see Discussion). Two AU-rich segments consisting of 7 and 8 consecutive As and Us are centered 18 (proximal) and 35 (distal) nts upstream of the fis translation start site, respectively (Fig. 4A). Mutational analysis of these two segments revealed that the distal element (AAAAUUUU) is important for fis translation. Replacing the 4 As or 4 Us of this sequence with Gs (M27 and M28) reduced the relative beta-galactosidase activities of the fis’-lacZ translation reporters 15- to 33-fold but had no effect on transcription (Fig. 2-4B). By contrast, substitution of 4 Gs within the proximal sequence (M30) did not significantly reduce fis translation. Switching As and Us within the center of the distal AU element from AAUU to UUAA (M29) had no effect, suggesting that the A+U composition rather than primary sequence or RNA secondary structure is the key determinant. An additional 4 bp substitution beginning one nt upstream of the distal AU element (M43) makes the region G-rich but has little effect on fis’-lacZ reporter activity. We conclude that the AU-rich element at the 3’ end of the dusB coding sequence is of critical importance for robust translation of fis mRNA.

Chromosomal mutations that disrupt the stem-loop and AU elements decrease Fis levels. In order to confirm that the disruption of the putative stem-loop and AU elements reduce Fis synthesis, representative mutations were recombined into the dusB-fis locus on the E. coli chromosome using the λ-RED system. As shown in Figure 2-5A, 4 mutations located in the base-paired regions of the putative stem-loop (M1-M4), two rescue mutants (M3:2 and M4:1), one mutation within the internal loop (M10), and a G substitution mutation within the AU
element (M27) were evaluated. Direct measurements of Fis protein levels were consistent with the results of the plasmid-based reporter assays.

Fig. 2-5 Effects of chromosomal mutations on Fis protein levels. (A) Schematic diagram depicting the stem-loop and functionally-important AU elements upstream of the fis AUG. The locations of mutations recombined into the chromosome are highlighted. (B and C) Representative anti-Fis Western blots of mutants grown in LB (60 min after subculture, panel B) and MOPS-rich glucose media where Fis levels are constitutively high (C). (D) Fis protein levels in MOPS-rich glucose media were quantified from standards on the same gel and plotted relative to WT. (E) qRT-PCR analysis of fis mRNA in representative mutant strains 60 min after subculture into fresh LB. fis mRNA levels were compared to a control gene (pppA) and reported relative to WT, which was set to 1. Black bars represent single mutants and grey bars represent rescue mutants. (F) Fis protein levels generated during LB batch culture by M2 (disrupts the predicted stem-loop structure, grey symbols and lines) and WT (black symbols and lines). Open symbols with solid lines depict the number of Fis dimers per cell (numbers are averages of 2 parallel cultures for each strain) and solid symbols with dashed lines represent the growth curves.
The chromosomal mutants were grown together with the otherwise isogenic parent strain (RJ3454) to directly compare Fis protein levels by Western blotting. Representative immunoblots of Fis protein levels from cells grown in LB (Fig. 2-5B) or MOPS-rich glucose (Fig. 2-5C) medias are shown, and amounts quantified relative to the wild-type from MOPS-rich glucose media are reported in Fig. 2-5D. As previously observed with the lacZ fusion constructs, all of the mutations that disrupt the stem loop and AU element have lower Fis protein levels, whereas the rescue mutants (M3:2, and M4:1) and M10 with changes in the internal loop exhibit near wild-type levels. The fold differences between the mutants and wild-type, however, are not as large as measured by lacZ translation reporter activities. Quantitative RT-PCR analysis of mRNA from the mutants revealed no significant differences in fis transcript levels (Fig. 2-5E).

Fis protein levels over the course of a batch culture cycle in LB were analyzed for one of the strong mutants (M2, Fig. 2-5A) within the putative stem-loop region to determine whether the growth-phase expression pattern was altered. Fis dimers per cell were quantified every 15 min after subculturing and compared to parallel wild-type cultures, which exhibited a typical pattern of Fis expression with respect to the growth cycle (Fig. 2-5F) (Ball et al., 1992). Fis levels in the mutant peaked at about 12% of the wild-type levels. The similar overall expression pattern of the mutant suggests that the stem-loop element is enhancing fis translation to a similar extent throughout the growth cycle. It is notable that the growth curve of the M2 mutants slightly lags behind the wild-type strain. This trend is reproducible and comparable to the increased lag observed with a fis null mutant.

**Further dissection and spatial relationships of the functional upstream elements on fis translation.** Our results implicate at least two determinants located upstream of the fis coding region that are required for high translation. As noted above, removal of both the
putative stem-loop structure and the AU element in deletion Mdel6 (-100 to -24 relative to the initiating ATG) results in very low levels of Fis protein synthesis (<1% of wild-type) (Fig. 2-2C and Fig. 2-6B). Most of the stimulatory effect promoted by this region is lost in Mdel7, which removes the AU element but retains the stem-loop (Fig. 2-6A and B). When the stem-loop segment is precisely excised (Mdel8) or excised leaving the sequences over the M24 and M25 substitutions (Mdel9), activities of fis\(^{\prime}\)-lacZ translation reporters are only reduced by about 3-fold. These results imply that the AU element is the most critical determinant responsible for high fis translation. However, as described above, localized changes within the stems of the stem-loop region can have large negative impacts on Fis synthesis.

In other systems where AU-rich sequences contribute to translation initiation, the AU segment is proposed to facilitate ribosome loading through the action of ribosomal protein S1 (see Discussion). However, in most of these systems the AU segment is close to the initiating codon whereas in fis it is centered 35 nts upstream. We asked whether the AU element in the fis upstream region could function at even greater distances from the AUG. Insertions that increased the separation between the fis AUG and the AU element plus upstream DNA were added by first introducing a XhoI site 18 bp upstream of the start codon (M33 in Fig. 2-6C). The M33 substitution decreased fis\(^{\prime}\)-lacZ translation to 43% of wild-type levels (Fig. 2-6D). Addition of 20 nts into M33 resulted in only a small additional decrease (Fig. 2-6D). Forty and 60 nt insertions resulted in \(\geq\)10-fold decreases in Fis expression fis\(^{\prime}\)-lacZ translation activity (Fig. 2-6D). Fis protein levels were also quantified by Western blotting of the chromosomal dusB-FLAG strain, which inserts 30 bp at the C-terminal end of dusB, 3 bp 3’ to the AU element. Fis levels were reduced to about 50% of wild-type in the dusB-FLAG strain. These data suggest
that the AU element can function relatively effectively to enhance translation with modest increases (20 to 30 nts) in spacing from the AUG but that large increases are poorly tolerated.

Fig. 2-6 Deletion and insertion mutations within the fis translation sequence elements. (A) Sequence of the fis leader extending to -100 with the putative stem-loop structure drawn. Deletion endpoints are marked, and nucleotides within the stems where substitutions cause a decrease (black) or little change (grey) in Fis protein levels are highlighted. Mdel6 and Mdel7 have a 6 bp XhoI site in place of the deleted sequence. (B) Relative activities of translation and transcription reporters by the deletions designated in panel A are given. SL refers to the stem-loop and AU to the distal AU element deleted in the respective mutants. (C) fis leader sequence demarking the changes used to create XhoI sites (M33 and M34) where DNA insertions were introduced that alter the spacing between functional elements. (D and E) Relative activities of translation and transcription reporters by the M33 and M34 XhoI mutants and respective insertions.
The importance of spacing between the AU element and putative stem-loop was also investigated by adding DNA into a XhoI site introduced 42 bp upstream of the fis AUG (M34; Fig. 2-6C). Insertions of 15 and 40 nts caused a progressive decrease in fis′-lacZ translation activity to 42% and 23% of wild-type levels, respectively (Fig. 2-6E). Thus, RNA sequences upstream of the AU element that include the putative stem-loop region appear to enhance translation less effectively at increased distances.

Discussion

The Fis nucleoid protein is distinguished by its fast increase in synthesis rates following nutrient upshifts and its abundance in rapidly growing E. coli cells. Previous studies have shown that much of this regulation occurs at transcription initiation, but we report here that the high synthesis rates of Fis are further supported by distinct mechanisms operating at the translation level. Whereas a SD-like sequence seems to play little role in fis mRNA translation, we find that sequences from -32 to -94 nts upstream of the fis AUG are responsible for increasing fis′-lacZ translation reporter activities over 100-fold. Within this region an A+U sequence element centered 35 nts upstream of the fis AUG increases fis translation by as much as 15-fold. Formation of a putative RNA secondary structure element beginning 50 nts upstream of the AUG also positively affects fis translation by up to 10-fold.

The fis gene is co-transcribed with the upstream dusB gene encoding a tRNA modifying enzyme. DusB protein levels are very low even under conditions when there is high transcription of the operon and high levels of Fis protein. The 32 nt 5′ leader of dusB has no recognizable SD sequence, which presumably contributes to its poor translation. The poor translation of dusB and the lack of effects by nonsense (M20, Fig. 2-1D) or frameshift (data not shown) mutations in dusB suggests that there is insignificant translational coupling between dusB
and fis. Although the dusB-fis operon may be an extreme example, recent global profiling of translating ribosomes in E. coli has shown that differences in translational efficiencies of polycistronic messages is not uncommon (Oh et al., 2011). Even though dusB is translated poorly, mRNA levels over its length are similar to those over the fis coding region (Fig. 2-1B and RNA tiling array hybridization data of J. Peters, J. Grass, and R. Landick, personal communication). The stability of the untranslated dusB message is important since sequences within the 3’ third of the dusB gene are required for high fis translation, with the 3’ 70 nts being of greatest importance.

![Phylogenetic conservation within the AU and stem-loop regions upstream of fis](image)

**Fig. 2-7** Phylogenetic conservation within the AU and stem-loop regions upstream of fis. (A) Sequence conservation within the distal AU element region from nine representative enteric bacteria. The % identity with the E. coli MG1655 sequence over the dusB gene for each organism is given along with the % nucleotide identities over the stem-loop region shown in panel B. (B) Sequence conservation within the stem-loop region. (C) Sequence variations in panel B mapped onto the E. coli putative stem-loop structure. Coloring of varying sites is coordinated with panel B.
A sequence separated by nine nts from the \textit{fis} AUG exhibits moderate complementarity to the anti-SD sequence at the 3\textquotesingle end of 16S rRNA. However, mutagenesis of this sequence or other Gs proximal to the AUG provides no evidence for a functionally important SD-like sequence. On the other hand, comprehensive mutagenesis of the upstream region points to an 8 nt AU sequence (AAAAUUUU, where the central AAUU sequence can be switched to UUAA) centered 35 nts upstream of the AUG as the dominant element responsible for efficient Fis synthesis. As shown in Figure 2-7A, the AU sequence found in the \textit{E. coli} fis leader is conserved in location and sequence in seven other \textit{gammaproteobacteria}, with the exception that one of the As is replaced with a C in \textit{S. marcescens}, \textit{E. carotovora}, and \textit{Y. enterocolitica}.

AU-rich sequences have been noted to be overrepresented in 5\textquotesingle UTRs and have been experimentally shown to enhance translation of the downstream gene in a number of cases (Boni et al., 1991; Hook-Barnard et al., 2007; Komarova et al., 2002; O'Connor and Dahlberg, 2001; Olins and Rangwala, 1989; Vimberg et al., 2007; Zhang and Deutscher, 1992). Their positive effect on translation has been proposed to be correlated with a general lack of secondary structure, thereby facilitating ribosome accessibility of the region (Kozak, 2005), or by functioning as a binding site for the largest protein component of the ribosome S1 (Boni et al., 1991; Komarova et al., 2002; Zhang and Deutscher, 1992). S1 is an essential \textit{E. coli} protein that is weakly and reversibly associated with the ribosome and can directly contact mRNA sequences upstream of the start codon (Ringquist et al., 1995; Sengupta et al., 2001; Sorensen et al., 1998; Subramanian, 1983). S1 has been shown to promote ribosome binding by preferentially interacting with mRNA segments rich in As, Us, and Cs and notably lacking Gs (Mogridge and Greenblatt, 1998; Subramanian, 1983). Moreover, S1 has been found to be absolutely required in vitro for ribosome binding and translation of mRNAs lacking a SD sequence (Farwell et al.,
1992; Roberts and Rabinowitz, 1989; Tzareva et al., 1994). These functional studies, together with the cryoEM structures of Frank and co-workers (Sengupta et al., 2001), suggests that S1 is able to compensate for weak interactions between the SD and the anti-SD on the 16S rRNA by tethering AU-rich mRNA segments to the ribosome near the P site. For \textit{fis}, we propose that S1 binding to an 8 nt AU element functions to recruit the mRNA to the ribosome in the absence of a functionally important SD sequence. However, we note the study of Vimberg et al. who concluded that AU-rich translational enhancers function synergistically with strong SD sequences because they had little positive effect in the context of weak SD sequences in their assay system (Vimberg et al., 2007).

Interestingly, \textit{fis} contains two AU-rich elements centered 18 and 35 nts upstream of the AUG, but only the distal segment is important for translation. In most of the experimentally-supported examples cited above, the AU-rich translational enhancer is located close to the translation start site. Surprisingly, further increasing the spacing from the AUG to the functional AU element by 20-30 nts only decreases \textit{fis} translation 2- to 3-fold, but larger insertions have a severe effect. We conclude that the AU element on the \textit{fis} mRNA functions at an unusually remote position from the AUG to facilitate translation. In vitro experiments will be needed to confirm that the \textit{fis} AU segment functions as a binding site for ribosomal protein S1 and that this interaction promotes ribosome loading.

Secondary structure programs predict that the local region extending up to 50 nts upstream from the AUG (including the AU element) is likely to be free of stable secondary structures. However, sequences further upstream are predicted to fold into a 44 nt secondary structure that consists of two stem regions interrupted by a 10 nt unpaired segment on the 5’ side (Fig. 3C; \(\Delta G = -18.5\) kcal/mol). Mutations that disrupt the bottom stem decrease \textit{fis-lacZ}
translation reporter activities up to 10-fold. These include a single A to T change at -91, which reduces fis'-lacZ activity to 14% of wild-type (data not shown). Mutations disrupting the top stem have milder effects, and mutations in the unpaired regions have no demonstrable effects. Evidence for the importance of both stem structures over the primary sequence in facilitating fis translation is provided by combining mutations that restore base pairing together with high Fis synthesis (Fig 2-3D and E). Further support comes from the strong phylogenetic conservation of the stem regions (Fig. 2-7B). Each are predicted to fold into an identical structure and the only deviations within the stems occur in Klebsiella pneumoniae where a G to A transition converts the wobble base pair into a Watson-Crick base pair and Salmonella enterica where a GU wobble base pair replaces a GC base pair (Fig. 2-7C). Although variation occurs within the unpaired regions, the degree of primary sequence conservation is greater over the 44 nt stem-loop region (95-100%) than over the entire 966 nt dusB coding sequence (77-94%) (Fig. 2-7A). Data regarding the predicted secondary structure that is difficult to reconcile, however, is the lack of effect by mutations M24 and M25, which are located on the 3' side of the bottom stem (Fig. 2-3A, B). Moreover, a complete or partial deletion of the stem-loop segment, as represented by Mdel8 or Mdel9, respectively (Fig. 2-6A), only decreases the activities of fis'-lacZ translation reporters 3-fold. mRNA sequences within dusB that are even further upstream may also modulate fis translation, potentially by forming secondary and tertiary interactions. For example, Mdel5, which removes 65 nt beginning 3 nt upstream of the stem-loop structure discussed above, exhibits a 3.5-fold decrease in fis'-lacZ activities (Fig. 2-2D). This region is predicted to fold into a structure with a calculated ΔG = -18.2 kcal/mol. The absence of ribosomes over dusB due to its low translation may make the mRNA particularly prone to forming structures. Probing the structure of the upstream RNA in vitro may be informative with respect to the formation of these
predicted structures, but equilibrium structures present in purified systems may not necessarily reflect the in vivo situation.

Most RNA secondary structures that control translation act as negative regulators by preventing the ribosome from accessing the translation start site (de Smit and van Duin, 1994; Kozak, 2005). Nevertheless, there are a limited number of examples of upstream secondary structures that function to enhance translation. The best studied examples are the stem-loop structures increasing translation of *rnd, rpsA*, and *csgD*, but their mechanisms of action remain poorly understood (Boni et al., 2001; Holmqvist et al., 2010; Skorski et al., 2006; Zhang and Deutscher, 1989). The prominent effects of RNA sequences that are likely to form secondary structures upstream of the AU element in *fis* raise the possibility of riboswitch-like mechanisms or binding sites for trans-acting sRNA or protein factors influencing *fis* translation. Although supporting evidence is currently lacking, such mechanisms cannot be ruled out at this point. Small metabolites like CTP and ppGpp have a strong effect on *fis* transcription (Ball et al., 1992; Mallik et al., 2006; Mallik et al., 2004; Ninnemann et al., 1992; Walker et al., 2004), and it could be imagined that small molecule(s) indicative of the cell’s physiological state could also be influencing *fis* translation. Candidate RNA chaperones and translational regulators like Hfq, CspE, StpA, CsrA, and BipA/TypA have not been found to associate with the *dusB-fis* mRNA or significantly affect Fis levels (our unpublished data and (Edwards et al., 2011; Sittka et al., 2008)), but a renewed search for trans-acting regulators may be prudent.

In summary, we find that Fis has evolved an unconventional mechanism to promote efficient translation of the *fis* mRNAs that are quickly induced under rapid growth conditions. We propose that canonical SD and anti-SD interactions are not the main route to recruiting the *fis* mRNA to the ribosome. Rather, initiation of *fis* translation largely depends on an AU-rich
element, which is located unusually far upstream of the initiating AUG as compared to other systems where AU-rich elements have been shown to function as translational enhancers. We postulate that the fis AU element functions as a binding site for the ribosomal protein S1 to efficiently recruit the fis mRNA to ribosomes. Further mechanistic work will be required to elucidate the steps of ribosome binding by the dusB-fis mRNA and the role(s) by which remote RNA secondary structures in dusB influence fis translation initiation.

Methods

Bacterial strains

A list of strains used in this work is provided in Table 1. Plasmids containing the lacZ fusions were transformed into the E. coli MG1655 derivative RJ3454. The λ RED system was used for recombining fis mutations and dusB-FLAG (see below) into the chromosome (Datsenko and Wanner, 2000; Yu et al., 2000). Initially, the rpsL-neo (Heermann et al., 2008) cassette was inserted between 986 and 1,325 bp downstream of the dusB-fis promoter start site in RJ3791 (RJ3454 rpsL150) to generate the kanamycin-resistant and streptomycin-sensitive strain RJ3802. Recombineered alleles were then identified by streptomycin resistance and kanamycin sensitivity and confirmed by DNA sequencing after PCR amplification of chromosomal DNA.

Construction of lacZ reporters and mutants

pRJ1772 contains the dusB-fis operon from 376 bp (HindIII site) upstream of the transcription start site to 259 bp (BglII site) downstream of the fis gene inserted between the HindIII and BamHI sites in pACYC184. LacZ translation and transcription fusions to dusB were generated by substituting the SmaI-Sal fragments from pRS414 and pRS415 (Simons et al., 1987), respectively, into pRJ1772 between the EcoRV (dusB codon 95) and Sall sites to generate the in-
frame protein fusion pRJ2922 and operon fusion pRJ2923 (see Fig. 2-1A and C). Likewise, lacZ translation and transcription fusions to fis were generated by transferring the SmaI-Sal fragments from pRS414 and pRS415, respectively, between the HpaI (fis codon 16) and SalI sites of pRJ1772. The resulting fis'-lacZ translation reporter (pRJ2617) and fis'-lacZ transcription reporter (pRJ2618) serve as the parent constructs from which other point, deletion, or spacing mutants were derived. Deletion mutants 1-4 (Mdel1-4) were generated between native restriction sites within the dusB coding region (see Fig. 2-1A) such that all created in-frame internal deletions. For Mdel5 the 65 bp sequence downstream of the dusB KpnI site was removed by amplifying the downstream region (66 bp after Kpn I) with a forward oligo that has a KpnI site at its 5' end. This fragment was then ligated between the KpnI and BamHI (at the lacZ fusion joint) sites of pRJ2617 and pRJ2618. In order to construct Mdel6 and Mdel7, pairs of XhoI sites were first introduced (M36 and M33 for Mdel6; M34 and M33 for Mdel7, Table S1, and see Fig. 2-6A and C) using a modified QuikChange method (Stratagene), and the intervening DNA segments were subsequently removed by XhoI digestion and ligation. The spacing mutants were constructed by inserting duplex oligonucleotides representing the sequence beginning 11 bp upstream of the lacZ ATG in the trp-lac W200 fusion (Yu et al., 1984) into the M33 and M34 XhoI sites (see Fig. 2-6C). This DNA segment from the 3’ end of trpB was utilized because it has been shown to not have a stimulatory effect on translation (Yu et al., 1984). Mdel8-10 were constructed using a two-step overlap PCR method (Horton et al., 1990). Specific nucleotide changes in the dusB coding region and intergenic regions were introduced by a modified QuikChange protocol. For the mutant scans, changes were typically made to complementary base pairs. The locations of the mutations are given in the relevant figures and provided in the plasmid list in Table S1. All constructs were confirmed by DNA sequencing.
**beta-galactosidase activity assay**

Fifteen hour overnight cultures were diluted 1/50 into fresh 5 ml LB cultures and typically grown for 1.5 hr, which corresponds to the peak time of fis′-lacZ activity. Cells were permealized by incubation with 0.005% SDS and chloroform for 10 minutes at 28°C, and beta-galactosidase assays were initiated by addition of o-nitrophenyl-β-D-galactopyranoside according to Miller (Miller, 1992). For each mutation, translation and transcription reporters were grown and assayed in parallel with the wild-type constructs (pRJ2617 and pRJ2618, respectively).

**Purification of recombinant proteins**

The dusB gene was amplified from RJ3454 chromosomal DNA and cloned into pET11a with a single FLAG tag (GADYKDDDDK) added to its 3′ end generating pRJ2720. BL21 cells containing pRJ2720 were induced with IPTG at 37°C for 2 hrs, and cells were harvested and lysed afterwards by a French Press in buffer containing 50 mM HEPES (pH 7.5), 15 mM EDTA, 10% glycerol, 5 mM DTT, 0.3 M NaCl. DusB-FLAG was highly overproduced but present in inclusion bodies, which were isolated by centrifugation at 8,000 x g for 20 min after an initial clearing spin at 1,600 x g for 15 min. DusB-FLAG was solubilized in 2% SDS, and SDS-PAGE showed that the resulting preparation was about 90% pure. The DusB-FLAG concentration was determined by comparing dilutions of solubilized DusB-FLAG protein to BSA standards of known concentration on SDS-PAGE. Recombinant Fis was purified as described previously (Stella et al., 2010).

**Western blotting**

Cells analyzed by anti-Fis Western blotting were grown in either LB media or MOPS-rich media with 0.3% glucose (Neidhardt et al., 1974). When LB medium was used, the 15 hour overnight
cultures were diluted 1/50 into fresh LB medium and cells were collected at various time points afterwards. To assay cells under more steady-state growth conditions, 15 hour cultures grown in MOPS-rich media were first diluted 1/200 into 5 ml of fresh MOPS-rich medium. At an OD$_{600}$ ∼ 1.0, cultures were chilled, stored at 4°C overnight, and diluted 1/50 into fresh media the next day. Cells were then collected at OD$_{600}$ = 0.4. Under these conditions, Fis levels remain high between OD$_{600}$ 0.1 and 1.0, (Y. Bernatavichute and R.C.J., unpublished data), unlike the rapidly changing Fis levels observed in LB batch cultures. The amount of cells equivalent to 1 ml of OD$_{600}$ of 1.0 were collected, resuspended in SDS loading buffer, boiled for 5 minutes, and the equivalent of OD$_{600}$ = 0.1 of cells were subjected to SDS-PAGE. Protein was transferred to a PVDF membrane (Immobilon, Millipore), which was incubated with rabbit polyclonal Fis antibody followed by mouse anti-rabbit IgG DyLight™ 800 conjugated (Thermo Scientific). The blots were visualized and quantified by an Odyssey imaging system (LI-COR). The amount of Fis present in each sample was calculated based on standard curves derived from purified Fis electrophoresed on the same gel. Fis protein levels are reported relative to wild-type cells (set to 1) or as Fis dimers per cell if viable cell counts (colony forming units) were determined. Western blots for DusB-FLAG were performed similarly except that up to 16 times more cells were loaded onto the SDS-PAGE gels. Membranes were probed with mouse monoclonal anti-FLAG M2 antibody (Sigma) followed by goat anti-mouse antibody (IgG DyLight™ 800 conjugated, Thermo Scientific) and imaged with the Odyssey system. Because no bands of the appropriate size for DusB-FLAG were visible, the lowest amount of purified DusB-FLAG (5 ng) that could be visualized by the antibody was considered to be the upper limit of DusB-FLAG levels in the cell.
Quantitative RT-PCR

Fifteen hour overnight cultures of RJ3454 and the *fis* chromosomal mutant strains (RJ3802-3808, RJ3817, and RJ3818) were diluted 1/50 into fresh LB media. After 60 minutes of growth cells were collected, the RNA was extracted using hot acid phenol (Case et al., 1988), and further purified by using the Qiagen RNeasy kit following the manufacturer’s directions. cDNA was prepared from 1 µg of total RNA and 250 ng of random hexamers (Qiagen) using Superscript II Reverse Transcriptase (RT, Invitrogen). The cDNA was diluted 5-fold and used as the template for quantitative PCR reactions with SYBR green. For each reaction a -RT control was performed, whose signal threshold appeared at least 10 cycles after the +RT reaction. Moreover, no correct product bands upon gel electrophoresis were present in the -RT controls. Primers were *fis* FP (oRJ343): TCGAACAACGCGTAAATTCTG, *fis* RP (oRJ344): TTGCATCACCATGTCCAACA, 5’ *dusB* FP (oRJ339): AGACCTTTTCGGACGTTGTG, 5’ *dusB* RP (oRJ340): TTCTTTCGGATCGCTACCAG, 3’ *dusB* FP (oRJ341) ACGACAGTATTCCGGGCAGTT, 3’ *dusB* RP (oRJ342): CAACTCCCCAGTGCCAGAT, and *rpoA* FP (oRJ313): AATTGACTGTCCGCTCTGCT, *rpoA* FP (oRJ314): TTACTCGTCAGCGATGCTTG. *dusB* and *fis* mRNA copy numbers were determined from standard curves obtained using serial dilutions of the template pRJ753, which contains the *dusB*-*fis* coding region (Johnson et al., 1988). *fis* mRNA levels in the different chromosomal *fis* mutants were normalized to *rpoA* mRNA levels assayed in parallel and reported relative to wild-type by using the $2^{-\Delta\Delta C_{t}}$ method (Livak and Schmittgen, 2001).


61


Chapter III

A closer look at factors involved in fis mRNA translation
**Introduction**

Initiation is thought to be the rate-limiting step of translation. In prokaryotes strong and accessible Shine Dalgarno (SD) sequences have high potential to base pair with the 3' end of the 16S rRNA (anti-Shine Dalgarno or aSD), which makes up the majority of the small ribosomal subunit (30S) (Ban et al., 2000; Harms et al., 2001; 2000; Wimberly et al., 2000). In general, open reading frames (ORFs) that have strong SD sequences in their 5' untranslated regions (5'UTRs) are expected to be translated at high levels (Ma et al., 2002; Ringquist et al., 1992), but there are exceptions. One example is the \textit{rpsA} mRNA, which has a poor SD sequence and yet translates into the highly abundant ribosomal protein S1 (Skorski et al., 2006).

We have recently described another example of a highly abundant protein that has weak canonical translational signals (Nafissi et al., 2012). The nucleoid-associated protein Fis has a poor SD sequence that is dispensable. However, mutating sequences further upstream showed that \textit{fis} translation is dependent on a putative stem-loop secondary structure (SL) and an AU-rich sequence (AU). There are other reported examples of AU-rich sequences serving as translational enhancers, but their mechanism of action are not well understood. It has been proposed that they keep the translation initiation region (TIR) free of strong secondary structures (Kozak, 2005) or that they serve as binding sites for ribosomal protein S1 (Boni et al., 1991; Tzareva et al., 1994). 5' UTRs that fold into stable RNA secondary structures, on the other hand, are usually thought to inhibit translation by preventing the ribosome from binding to the TIR. Since the \textit{fis} SL structure is located far upstream of the start codon (-50), a positive effect seems possible. However, there are only three other known examples of positively-acting structures and little is known about their underlying mechanisms (Boni et al., 2001; Holmqvist et al., 2010; Zhang and Deutscher,
Therefore, it appears that robust fis translation is driven by non-canonical elements that are still poorly understood.

In order to gain mechanistic insights, in vivo expression studies of fis-lacZ constructs that combine SL, AU, and SD mutations were performed. These studies were complemented with in vitro toeprinting experiments measuring assembly of a minimal 30S translation initiation complex (30SIC). The results of the expression experiments provide evidence that the SL and AU sequence aid in recruiting the 30S subunit, but the SL appears to act through more indirect means. Surprisingly, in vitro toeprinting experiments thus far failed to provide evidence for a role of these elements in assembly of the 30SIC. However, the toeprinting conditions do not seem optimal since they support only weak interactions between 30S and the fis TIR. We discuss possible models for how both elements function and how to test them.

Results

An enhanced SD sequence partially restores the reduced fis translation levels caused by disrupting the AU and SL element. We showed in Chapter 2 that mutating the SL and AU sequence can each decrease fis translation up to 10-fold. Because disruption of the putative weak SD had no effect, SD/aSD interactions seem to play no role in the efficiency of fis translation. We wondered if the AU element and/or the putative stem-loop structure provide alternative means for 30S recruitment. To test this hypothesis we coupled an improved SD sequence (SD\textsuperscript{up}, Fig. 3-1A) to the previously described AU-M27 and SL-M3 mutants (Fig. 3-1B), which both have weak translational activity (Nafissi et al., 2012). We measured activities of fis-lacZ transcriptional and translational fusion constructs encompassing the single and double mutations (Fig. 3-1B) by the beta-galactosidase activity assay and report them relative to wild-type, which is set to 100. An enhanced SD sequence by itself has no effect on fis-lacZ expression (Fig. 3-1C),
most likely because translation is already extremely efficient. However, in an AU-M27 mutant background the SD^{up} mutation improves translational activity by a factor of eight. With regards to the SL- M3 mutant, SD^{up} causes only a 2-fold increase in fis-lacZ translation. Thus, a stronger SD sequence can recover fis translation in an AU mutant to about half of wild-type levels, but only up to a quarter of wild-type levels for an SL mutant. Although this suggests that both elements are important for recruiting the 30S ribosome, the SL most likely supports fis translation in some other way, which needs further investigation.
Mutations within the lower right arm of the SL structure suppress the detrimental effect of stem mutations in the opposing strand. Previous base substitutions of the SL revealed that most base-pairing interactions are required for high fis-lacZ translational activity (Fig. 3-2A). M24 and M25 are the only exceptions since they have no effect on their own. Nevertheless, they effectively suppress mutations in the opposing strand (M16 and M21, respectively; Fig. 3-2A), which was interpreted by the restoration of base pairing (Nafissi et al., 2012). In order to
dissect this further, we individually combined both M24 and M25 mutations with mutations of regions they are not predicted to base pair with (M3 and M21). The translational activities of the corresponding double fis-lacZ mutant constructs were measured and reported relative to wild-type. Surprisingly, both M24 and M25 restore translation of the M3 and M21 mutants back to wild-type levels even though the M25:M21 pair is the only one predicted to do so. It is important to point out that this compensatory effect cannot be observed for all mutations in the lower right arm since M2 does not rescue M21 (Fig. 3-2B). One explanation for this result is that sequences disrupted by the M24 and M25 mutations may target an inhibitory factor whose binding or activity is normally masked by formation of the SL.

Fig. 3-2 Effect of M24 and M25 mutations on fis translation. (A) The exact base composition of the SL structures is given. The black spheres denote all the bases that upon introduction of complementary bases decrease fis-lacZ translation. All other RNA base changes including M24 and M25 have no effect on fis-lacZ expression. (B) Relative beta-galactosidase activity of mutated fis-lacZ translational constructs. The coloring of the bar graphs follows the parameters set in (A), the only addition is that wild-type is highlighted with light grey and dark grey marks the double mutants. As previously published, the stem mutants M3, M2, and M21 all inhibit fis-lacZ translation to about 10% of wild-type, whereas M24 and M25 produce no change in expression (Naftissi et al. 2012). Combining M24 and M25 with both M3 and M21 restores translation back to wild-type levels. As shown in (A) only M24 is a true compensatory mutant for M16, and the other pairs are not predicted to basepair with one another. The M2 mutation also modifies the same side of the stem as M24 and M25 but in contrast to them it does not rescue the M21 mutant. Therefore, mutating six nts at the 3' end of the SL can mask the effect of any mutations in the opposing stem. This implies that this region has an inhibitory role that only gets exposed when the SL is disrupted.
In vitro assembly of a minimal 30S initiation complex at the fis translation TIR is weak and absolutely depends on ribosomal protein S1. In vivo expression experiments have been crucial for identifying important determinants for efficient fis translation, but in vitro experiments provide a more controlled environment for uncovering mechanistic features. In vitro assembly of 30SICs has traditionally been studied by toeprinting assays. The minimum
components required to form a stable 30SIC are an RNA with an accessible TIR, purified 30S subunit, and uncharged initiator tRNA (itRNA). This ternary complex can block reverse transcriptase (RT) driven primer extension reactions. As a result, the RT stops cDNA synthesis when it encounters the 3’ end of the 30SIC, giving rise to a so-called ‘toeprint’ (Fig. 3-3A). For most E. coli mRNAs this toeprint usually occurs 16 nt downstream of the A of the AUG (+16).

In order to measure 30SIC formation on fis mRNA, toeprinting experiments were performed with short (152 nts) (Fig. 3-3B) and long (1,071 nts) (Fig.3-3C) versions of the fis RNA. Both substrates contain the SL, AU element, and beginning fis coding sequences (Fig. 3-3B), but the long substrate extends 5′ to the dusB-fis transcription start site. As expected, both wild-type constructs give a +16 stop in the presence of 30S and itRNA. However, a large excess (four- and eight-fold in comparison to fis RNA) of the 30S is required to obtain this toeprint, and the signal relative to background RT stops is low (Fig. 3-4A, B, C). These results suggest inefficient 30S binding to the fis TIR. For the short RNA all additional RT stops are 30S-independent, but for the long RNA some are 30S-dependent and of even greater intensity than the +16 stop (Fig. 3-4C). This indicates that 30SICs can form at multiple positions upstream of the fis translation start site which are not normally observed (Stefano Marzi, personal communication). Short open reading frames (ORFs) could explain these additional toeprints, but the only alternative start codon is found 18 nts upstream of the -60 stop (Fig. 3-4D), and substitutions over this AUG have no effect on fis translation in vivo (Nafissi et al., 2012). Therefore, these additional stops most likely represent nonspecific 30S binding that is observed because of the large 30S excess added and the weak binding affinity of 30S for the fis TIR.

Ribosomal protein S1 is an important component of the 30S ribosome and seems to interact with RNA sequences upstream of the SD (Sengupta et al., 2001). S1 has been proposed
to bind AU-rich sequences that are often found in 5'UTRs, thereby assisting 30S binding to TIRs (Boni et al., 1991; Tzareva et al., 1994). Messenger RNAs with weak translational elements, such as poorly conserved SD sequences, seem to rely heavily on S1 assistance (Farwell et al., 1992; Roberts and Rabinowitz, 1989; Tzareva et al., 1994). In order to test if the *fis* 30SIC also depends on S1, toeprinting experiments with S1-depleted (30S-S1) and wild-type 30S (30S) were compared (Fig. 3-4A, lanes 5,6 and 2,3, respectively). Even with up to eight times more 30S-S1 than *fis* mRNA no toeprint is observed (Fig. 3-4A, lanes 5,6), unless purified S1 was added back to the binding reaction (30S+S1) (Fig. 3-4A, lane 7). Duval et al. (manuscript in preparation) show that this preparation of 30S-S1 is competent for producing strong toeprinting signals for mRNAs with well exposed and conserved SD sequenced. We tried to increase the *fis* toeprint by supplying excess S1 in trans, but achieved no improvement (data not shown). From these results we can conclude that 30S binding to the *fis* TIR is S1 dependent, but that levels of S1 do not seem to be limiting complex formation.

**The AU and SL element are not required for forming a minimal 30SIC at the *fis* TIR in vitro.** We performed toeprinting experiments on *fis* mRNA with disrupted SL (M21) or AU (M27, M28) sequences to assess their role in 30SIC formation at the *fis* TIR. To our surprise, the short and long SL-M21, AU-M27, and AU-M28 mutant substrates displayed +16 toeprints of similar intensity as wild-type (Fig. 3-4B, lanes 5,7, 10,11; Fig. 3-4C, lanes 6,7, 10,11, 14,15). It is noteworthy that for some of the long RNA mutants (AU-M28 in particular) a somewhat weaker toeprint is observed (Fig. 3-4C, compare lanes 2,3 with 14,15). However, this result does not match the dramatic reduction of *fis-lacZ* translational activity observed for these mutations in vivo. For the mutant that disrupts the SD sequence, on the other hand, 30S in vitro binding (Fig. 3-4B, lanes 14,15) and in vivo expression data are in good agreement (Nafissi et al., 2012). It
appears that under the current toeprinting conditions, a simple translation initiation complex consisting of 30S, itRNA, and fis mRNA forms inefficiently regardless of whether the SD, AU, and SL elements are intact.

**Introducing AU and SL mutations into the long fis RNA may induce structural changes.** As mentioned earlier, long fis RNAs give rise to multiple primer extension stops, of which some are 30S-independent (Fig. 3-4C). These RT stops, which are not equally distributed among wild-type and mutant constructs, most likely result from stable secondary structures. It follows that the mutations introduced into the fis mRNAs may cause differential RNA folding. As shown in Figure 3-4D, the -31 and -34 stops are within or close to the AU-rich region. The -31 stop only occurs in the AU-M28 mutant suggesting that G substitutions within the AU sequence generate an alternative RNA structure. It is also possible that the -31 stop is simply due to the M28 base changes because a stretch of five consecutive G’s could pose a hindrance to RT elongation. Since a strong -34 stop is found in all constructs except SL-M21, it appears that disrupting the SL structure causes the RNA to adopt a different conformation. Moreover, the -71 stop is the only one ascribed to the SL structure, and it is only present in the SL-M21 mutant. RNA secondary structure programs do not predict the formation of alternative structures for SL-M21 and AU-M28, but that does not exclude their existence. To summarize, primer extension patterns suggest that fis RNA mutants may fold differently from wild-type, but this needs to be confirmed by thorough experimental analysis of fis RNA secondary structure using SHAPE and RNase probing techniques (Daou-Chabo and Condon, 2009; Low and Weeks, 2010).
Fig. 3-4 30SIC formation on different βs RNAs assessed by toeprinting analysis. Ternary complex assembly depends on the presence of ribosomal protein S1, but not the AU sequence or SL structure. (A). Toeprinting analysis of short wild-type βs RNA. 30SIC assembly on short wild-type βs RNA is assessed by toeprinting experiments. Equal amounts of radioactivity were loaded into every lane of a 10% polyacrylamide/8M urea sequencing gel. U, C, G, and A indicate the deoxy sequencing reactions. Lane 1 contains the RNA alone control, whereas all subsequent lanes contain 30S ribosome of different composition and concentration. For lanes 2 – 4 wild-type 30S (30S) was used, whereas 30S without S1 (30S*) was utilized for lanes 5, 6, and 30S reconstituted with purified S1(30S*) for lane 7. The 30S subunit is present in either 4- or 8-fold excess over βs mRNA (lanes 2, 3, and lanes 4, 6, 7, respectively). Furthermore, all lanes except 1 and 4 contain excess tRNA. Every lane (2, 3, 7) that includes 30S with S1 and tRNA produces a primer extension stop at +16 ( toeprint). If 30S is used no such band is observed suggesting that this ribosomal protein is absolutely required for establishing a 30SIC at the βs start codon. (B) Toeprinting experiments with short wild-type and mutant βs RNAs. Experiments were carried out as in (A) except that only wild-type 30S is used and three short mutated βs RNAs (see Fig. 3-3B) are assayed in parallel to wild-type. The mutants disrupt the SL (M21s), AU (M28s), or SD (SD) sequence and for each one a control primer extension reaction was done that did not contain any 30S (lanes 1, 5, 9, 13). The lowest ribosome to RNA ratio is 1:4 (lanes 2, 7, 10, 14) and was increased up to 1:8 (lanes 3, 4, 5, 6, 11, 12, 15, 16). As long as both 30S and tRNA were added to reaction mixtures, a +16 toeprint was observed, even for the mutants. It appears that SL, AU, and SD are not needed to achieve accumulation of 30SICs on βs RNA.
(C) Toeprinting analysis of long *fis* RNA in its native and mutated state. The experimental set-up is identical with (B) only that long RNA substrates were assayed. As shown in figure 4B, one SL mutant (M21) and two AU mutants (M27, M28) were analyzed. All reactions comprised of 3OS and iRNA yield a +16 stop that appears slightly weaker in the mutants (Wt > AU-M27 > SL-M21 > AU-M28). However, there are many additional bands present and some are of even higher intensity than the +16 stop. Black arrows highlight the most prominent primer extension stops. The asterisks denote the 3OS- and iRNA-dependent toeprints whereas the other stops arise due to inherent RNA elements that inhibit RT and seem to be construct specific. The presence of so many additional strong 3OS-dependent and independent stops is quite unusual and suggests weak 3OS binding. (D) Results of toeprinting assays shown in (C) mapped to the *fis* RNA sequence. A region containing sequences from -100 to +3 relative to the A of the *fis* AUG is depicted in its predicted secondary structure conformation. Grey boxes mark the *fis* start codon (AUG), the *duB* stop codon (UAA), and another putative AUG codon in the loop of the SL. Regions mutated in the various mutant constructs (SL-M21, AU-M27, AU-M28) are highlighted in grey. Circles and squares draw attention to the most prominent primer extension stops. Circles and asterisks mark the 3OS/iRNA-dependent toeprints whereas squares denote the stops that are observed regardless of whether 3OS and iRNA are present or not. Even though band intensities vary slightly, all four constructs share the same 3OS/iRNA-dependent stops. These additional toeprints presumably mark the 3'ends of stable initiation complexes that form further upstream of the *fis* AUG. Since only the +60 stop is close enough to an AUG, it appears to be the only potential candidate for an alternative translation start site. However, the distance between the toeprint and AUG is a little too big (+18) and its mutagenesis does not change *fis* expression (Narissa et al., 2012). The mutated constructs display variable patterns of 3OS/iRNA-independent stops, suggesting that their sequence differences result in alternative RNA conformation that inhibit RT processivity.
Discussion

The in vivo and vitro experiments performed here provide additional information on how non-canonical elements support fis translation. Rather than relying on a SD sequence, efficient fis translation in E. coli depends on an upstream AU-rich sequence and a putative secondary structure. We show here that strengthening SD/aSD interactions can compensate for the loss of AU and (albeit to a lesser extent) the SL element. This in turn implies that the AU sequence aids 30S recruitment and may provide alternative ways to establish contacts with the 30S. However, the AU and SL elements were not functional in toeprinting experiments employing just 30S subunits (including ribosomal protein S1), itRNA, and fis mRNA. We propose that this minimal experimental system for ribosome binding needs further refinement, and may be missing one or more factors that are crucial for efficient initiation of fis mRNA translation.

Experimental approaches to refine the role of SL and AU in translation initiation.

Additional experiments will be required to understand the discrepancies between the in vivo and in vitro results; in particular, why the fis in vivo translation depends on intact AU and SL elements, but in vitro neither element is required for forming a translation initiation complex. We propose that converting the putative fis SD into an even stronger consensus sequence might result in an even better rescue of the AU and SL mutations. This would make a stronger argument for AU and SL helping in 30S recruitment. To test if we can observe a similar effect in our in vitro experiments, the experimental design of the toeprinting experiments needs to be adjusted. Under current toeprinting experiments 30S binding to the fis TIR is quite weak, and we hope that can be overcome by using different fis RNA constructs and by including additional known protein factors involved in 30SIC formation. Candidate known auxiliary proteins that may enhance Fis translation include bacterial translation initiation factors IF1 and IF2 (Simonetti
et al., 2008) or other auxiliary proteins like HU and H-NS (Balandina et al., 2001; Park et al., 2010). Increasing the \textit{fis} TIR 30SIC signal may help reveal in vitro phenotypes of the AU and/or SL mutants.

The \textit{fis} RNA substrates used in the in vitro analyses need to be optimized in several ways. The short and long \textit{fis} RNA substrates yield different 30S-independent RT stops. The long RNA seems to be the better substrate because it closely resembles the constructs used for expression studies. However, producing 1kb-long RNAs of good quality and stability is challenging, so different truncations of the 5′ end of the long RNA should be tested. Sequences at the 3′ end, on the other hand, need to be extended to ensure sufficient spacing between RT primer and 30S binding sites. In general, it seems best to characterize all new RNA substrates in vivo prior to using them in vitro. Moreover, routinely including a control RNA with a good accessible SD would enable more quantitative comparisons of the toeprinting signals under different conditions.

\textbf{Possible mechanistic models.} We expect that these afore mentioned experiments will provide additional evidence for a direct or indirect role of the AU and SL in 30S binding to the \textit{fis} TIR. There are several possible models that can explain their positive contribution, so we present our favorite ones and how they can be tested. The AU element could act similarly to the mechanism proposed for other known AU-rich enhancers by serving as a direct binding site for protein components of 30S (e.g., S1) (Boni et al., 1991; Tzareva et al., 1994) or other auxiliary proteins of translation (e.g., HU) (Fig. 3-5A) (Balandina et al., 2001). This can be tested by measuring binding affinities of wild-type versus mutant \textit{fis} RNAs and 30S in the absence and presence of such protein factors by toeprinting and filter-binding assays. It is also possible that the purpose of the AU-rich sequence is simply to keep the \textit{fis} TIR free of secondary structures,
which are introduced by AU mutations M27 and M28. RNA secondary structure prediction programs S-fold and M-fold do not support this second model, but performing secondary structure probing techniques will be a more thorough test.

![Diagram](image)

Fig. 3-5 Model for translational control of fis mRNA. (A) Since an improved SD sequence can partially compensate for a mutated AU sequence, the AU element is likely involved in direct or indirect recruitment of the 30S subunit. Direct binding of 30S most likely occurs via ribosomal protein S1, which has a preference for AU-rich sequences. Alternatively, other auxiliary proteins (designated as ‘factor X’) can bind to the AU sequence and thereby recruit 30S to the fis translation start site indirectly. (B) Mutating regions in the lower stem region (red bar) of the SL has no effect on fis translation and also neutralizes the effects of other stem mutants that otherwise decrease fis translation. For this reason, we propose that the SL enhances translation indirectly by deactivating an inhibitory sequence (red bar). Factors that compromise SL stability like the point mutations we introduced in stem regions (other than M24/M25) or trans-acting protein/RNA factors (designated as ‘factor X’) would expose this sequence and thus inhibit fis translation.

In comparison to the AU element, the SL appears to be less important for initiating fis mRNA translation in vivo. Even though there are other examples for RNA secondary structures acting positively on translation, in no case is their mechanism of action understood (Boni et al., 2001; Holmqvist et al., 2010; Zhang and Deutscher, 1989). Our thorough mutagenesis analysis
of the *fis* SL implies that it has an anti-inhibitory rather than activating function because mutations in the lower right stem have no effect on their own and mask any defect the other stem mutants cause. We propose that as long the regions mutated in M24 and M25 (called inhibitory or ‘Inh’ element from here on) engage in base-pairing interactions, their inhibitory nature is masked (Fig. 4-5B). There are several possible mechanisms by which this inhibition can occur and further characterization should help in identifying the most probable one. The ‘Inh’ element could act in cis by causing changes in the RNA secondary structures of the TIR and thereby blocking 30S binding. It also seems plausible that other trans-acting factors like sRNAs, small metabolites, or RNA binding proteins are involved in this inhibition. A screen for cis- or trans-acting suppressors that rescue the strong phenotype of a stem mutation may reveal the inhibitory mechanism. There is precedence for sRNA binding to upstream secondary structures (Holmqvist et al., 2010), but it is not obvious how this would inhibit translation initiation at downstream regions. Riboswitch-like mechanisms often entail the formation of ligand-induced inhibitory RNA structures (Serganov and Nudler, 2013). Small molecules, such as CTP and ppGpp, regulate *fis* transcription (Ninnemann et al., 1992; Walker et al., 2004), and it is possible that metabolites indicative of physiological states may also control *fis* translation. Ribosomal proteins can autorepress their own translation (Babitzke et al., 2009), but it is not obvious what RNA binding protein could regulate *fis* translation.

In summary, in spite of the need for an intact SL and AU element for high *fis* translation in vivo, biochemical experiments performed so far have not elucidated their mechanistic function. However, this is work in progress, and the in vitro conditions employed need to be refined. At the moment our favored models envision the AU-rich sequence functioning to directly or indirectly recruit the 30S and the SL acting in as yet undefined anti-inhibitory fashion.
(Fig. 3-5). Additional in vivo and in vitro experiments will help to gain a better mechanistic understanding of how these two non-canonical elements function to generate high Fis protein levels.

**Methods**

**Construction of fis-lacZ reporters**

A detailed explanation of how the fis-lacZ fusion is constructed is given in Nafissi et al. (2012). In essence, almost the entire dusB-fis operon (-376 to +1071 relative to the transcription start site) together with the lacZ gene are encoded by a low-copy plasmid derived from pACYC184. Thus, the entire dusB coding sequence as well as the first 48 nts of fis are fused the lacZ transcriptional and translational reporters. Nucleotide changes were introduced into the fis-lacZ fusions using a modified QuikChange protocol.

**beta-galactosidase activity assay**

MG1655-derived E. coli cells harboring the plasmid encoded fis-lacZ fusion constructs were assayed for beta-galactosidase activity as described earlier (Nafissi et al., 2012). In short, saturated overnight cultures were diluted 50-fold into 5-ml LB cultures containing 12.5 μg/ml Chloramphenicol, and cells were collected after 1.5 hours of growth since this corresponds to peak fis expression. Wild-type and mutant constructs were run in parallel, and their beta-galactosidase activities were determined according to the method of Miller (Miller, 1992).

**RNA preparation**

All RNA substrates were in vitro transcribed by T7 RNA polymerase. The DNA templates were produced by PCR using a forward oligo containing the T7 promoter to amplify plasmid encoded wild-type or mutant fis constructs. For short fis RNAs the forward oligo anneals 104 nt upstream of the fis AUG, which corresponds to ten nt upstream of the 5′ end of the SL. Long fis RNAs
begin at the transcription start site, which is at +1023 relative to the fis AUG. The common reverse oligo anneals 48 nts downstream of the fis translation start site. PCR products were purified by agarose gel electrophoresis before using them for in vitro transcription reactions. In general, standard T7 transcription reactions were used for making the short and long RNA constructs, but in order to improve the yield of the long RNA some adjustments were needed. Instead of the 2.5 µg used for the short constructs, 8 µg of DNA template were used for the long construct. Furthermore, the incubation times for in vitro transcription reactions were extended from 2.5 hours to overnight. This was followed by DNase I digestion, phenol-chloroform extraction, and ethanol precipitation. Both short and long RNAs were purified by denaturing polyacrylamide/8M urea slab gel electrophoresis, but their extraction methods varied. Whereas passive overnight elution at 4°C in a solution containing 0.5 M NH₄Acetate and 1 mM EDTA sufficed for the short constructs, the longer ones required two rounds of one hour long electroelution into TBE buffer at 4°C. After collecting the RNA it was ethanol precipitated again, and its concentration was determined by a NanoDrop spectrophotometer. The stability of the SD/aSD interaction was predicted by the RNAcofold module of the Vienna package (Gruber et al., 2008).

(forward primer for long RNAs: GATTCCTAATACGACTCA CTATAGGGCTTTGCAGTCACAGTATGG; forward oligo for short RNAs: GATTCTAATACGACTCACTATAGGGTCGGCGCAGCATC TCAA; reverse oligo for short and long RNAs: AACGGTAGAAACGGGTACG)

Toeprinting assays

The E. coli 30S subunits were purified on sucrose gradients as outlined in Fechter et al. (2009). In order to prepare S1-deprived 30S subunits (30S⁻S1), polyU Sepharose chromatography was
applied according to Duval et al. (manuscript in preparation). The oligo used for the RT-driven primer extension reactions was 5’end-labeled with $\gamma$-$^{32}\text{P}$-ATP by T4 PNK. Before assaying the formation of minimal 30S translation initiation complexes (30SICs) on different fis RNA constructs, the individual components had to be activated. The RNA needed to be melted first so it could fold properly and the $^{32}\text{P}$-labeled primer could anneal to it. Therefore, both RNA and primer were incubated at 90°C for exactly 1 minute, then on ice for 2 minutes, and finally at 20°C for 20 minutes in toeprinting buffer (20 mM Tris-HCl, pH 7.5, 60 mM KCl, 1 mM DTT) together with 9 mM MgCl$_2$. In the meantime, 30S subunits were pre-incubated in toeprinting buffer with 10 mM MgCl$_2$ at 37°C for 15 min. For each toeprinting reaction the reaction volume prior to primer extension analysis was 15 µl and contained 0.5 pmol of RNA substrate together with 200,000 cpm of $^{32}\text{P}$-labeled primer. The 30SIC was assembled in step-wise fashion starting with the binary complex, that formed when RNA (0.5 pmol/reaction) and 30S (4 or 8 pmol/reaction) were incubated in toeprinting buffer (total MgCl$_2$ concentration should be $\leq$ 8 mM) at 37°C for 15 min. Once excess itRNA (1 µM) and more MgCl$_2$ is added (final MgCl$_2$ concentration is 18 mM), an 30SIC could begin forming. Primer extension reactions were performed with 20 units of MMLV-RT (Moloney Murine Leukemia Virus Reverse Transcriptase) at 37°C and stopped after 15 min by ethanol precipitation. The radioactively labeled cDNAs were loaded on 10% polyacrylamide/8M urea gels along with the corresponding dideoxy sequencing reactions.


Chapter IV

Fis binding in vivo to the *E. coli* chromosome compared to naked DNA in vitro
Introduction

The 4.6 Mb-long circular *E. coli* chromosome has to undergo a high degree of compaction in order to fit into the small bacterial cell (2 x 1 µm). DNA supercoiling, macromolecular crowding, and small basic DNA binding proteins (nucleoid-associated proteins or NAPs) are thought to promote this compaction (Browning et al., 2010). Moreover, the nucleoid is believed to be folded into a rosette-like structure (Delius and Worcel, 1974; Kavenoff and Bowen, 1976), that is composed of ~10 kb-long topologically independent domains (Postow et al., 2004; Scheirer and Higgins, 2001; Sinden and Pettijohn, 1981; Stein et al., 2005; Worcel and Burgi, 1972). NAPs play an important role in establishing boundaries of individual loops because they are highly abundant (Ali Azam et al., 1999), bind long patches of chromosomal DNA (Kahramanoglou et al., 2011; Noom et al., 2007; Prieto et al., 2012; Vora et al., 2009), and possess the ability to bridge or loop DNA in vitro (Dame et al., 2006; Dame et al., 2000; Skoko et al., 2006). In addition to their architectural functions, the major NAPs Fis, H-NS, HU, IHF, and Dps also act as transcriptional regulators. Interestingly, most NAPs are differentially expressed depending on the growth stage (Ali Azam et al., 1999), but deleting a single NAP gene does not cause a strong phenotype (Dri et al., 1991; Filutowicz et al., 1992; Huisman et al., 1989; Paull et al., 1994). This suggests that there is some redundancy among NAPs and that the composition of the domain boundaries is fluid.

Out of all the NAPs, *fis* expression varies the most with growth phase and rate (Ball et al., 1992). Fis is likely to be involved in chromosome organization because it is a major component of the nucleoid (Murphy and Zimmerman, 1997) and can promote DNA branching (Schneider et al., 2001) and looping in vitro (Skoko et al., 2006). For this reason, we also believe that long Fis-bound DNA tracks give rise to the chromosome organizing centers and seek to provide
additional experimental support (Kahramanoglou et al., 2011; Schneider et al., 2001; Skoko et al., 2006; Vora et al., 2009). Chromatin immunoprecipitation (ChIP) can be very helpful in this regard because when combined with whole genome microarrays or DNA sequencing it identifies all protein-bound regions on a global scale. Because genome-wide Fis binding has already been determined in three independent studies (Cho et al., 2008; Grainger et al., 2006; Kahramanoglou et al., 2011), we decided to employ Fis ChIP-chip experiment in a slightly different manner. Each of the previous studies utilized different growth conditions, so we chose to use three types of growth media which cause cellular steady state Fis levels to vary 30-fold. We also used different genetic backgrounds to test the role of H-NS and HU in Fis binding and complemented our Fis ChIP-chip binding data with in vitro binding studies on selected regions. Even though our Fis binding data has not yet undergone thorough statistical analysis, we can make some qualitative conclusions based on Fis binding peaks visualized in a genome browser. We find that the Fis binding profile of cells grown under poor growth conditions where Fis levels are low displays peak-specific differences from that of rapidly growing cells where Fis levels are high. Surprisingly, we also observe that Fis binding is remarkably similar between data sets from fast growing cells deprived of H-NS or HU. In general, our in vitro binding results correlate well with the ChIP data since regions with strong binding peaks always display high-affinity Fis binding in vitro. Narrow binding peaks (≤ 500 bp) contain one high affinity site whereas broader binding peaks (≥ 1 kb) contain multiple (2-3) high-affinity Fis sites. However, the correlation between in vitro binding affinities and the occupancy of peaks in nutrient-poor media is less pronounced than in nutrient-rich media.
Results

Genome-wide Fis binding under three growth conditions that yield different fis expression levels. For our ChIP-chip experiments we used custom-generated high-density tiling arrays (Affymetrix) that probe both DNA strands with a probe separation of 4 bp. Affinity-purified polyclonal Fis antibody was used to immunoprecipitate Fis-bound chromosome segments. The data has been normalized to the total input DNA, which went through all processing steps except immunoprecipitation, and smoothed using a sliding window of 200 bp. Since fis expression is dependent on growth rate and growth phase (Ball, Osuna et al. 1992), Fis binding is greatly affected by growth conditions. A previous study already measured Fis binding
during different growth stages in LB batch cultures where Fis levels change dramatically during cell growth (Kahramanoglou et al., 2011). Therefore, we decided to measure Fis binding under near steady-state growth conditions utilizing MOPS-based defined media that generate different growth rates (Neidhardt et al., 1974). Where measured, cellular Fis levels under these growth conditions remain constant over ranges of OD$_{600}$ from $\leq$0.1 to $\geq$1.0 (Y. Bernatavichute and J. Chau, unpublished). MOPS media supplemented with glucose, amino acids, bases and vitamins is called ‘MOPS rich’; it supports the fastest growth rate and highest Fis protein levels (both similar to turbidistat LB cultures) (Fig. 4-1). For the other two growth conditions we omit the supplements and use glucose or alanine as sole carbon source (‘MOPS Glu’ and ‘MOPS Ala’, respectively). As shown in Figure 4-1 these medias give rise to slower growth rates and lower fis expression. In comparison to ‘MOPS Rich’ (25,000 Fis dimers/cell), cells grown in ‘MOPS Glu’ contain about 1/3 and ‘MOPS Ala’ only about 1/30 of the cellular Fis levels.

**Fis binding profile under nutrient-rich and nutrient-poor growth conditions exhibits differences.** In order to illustrate the Fis ChIP-chip results, images of the binding signals displayed in the IGB genome browser will be used. All images are set to the same scale and black blocks represent genes. Some representative regions are chosen and compared to one another in a qualitative manner. Although this initial analysis does not allow us to draw solid conclusions, it enables preliminary evaluation.

As was observed by other groups, Fis binding occurs throughout the *E. coli* genome (Cho et al., 2008; Grainger et al., 2006; Kahramanoglou et al., 2011). Figure 4-2 shows five 10 kb regions that contain numerous Fis peaks of varying heights and widths under ‘MOPS Rich’ conditions. To see if this binding pattern changes when cells are grown in MOPS media of lower nutritional value and less cellular Fis, we compared Fis binding signals within six representative
20 kb regions (Figure 4-3). As highlighted by the black dotted boxes, there are several peaks that remain occupied regardless of the media used. Nevertheless, there is frequent differential binding observed, and, as expected, most changes lead to decreased peak height and width under the slower growth conditions where cellular Fis levels are lower. The binding profile for ‘MOPS
ALa’, differs the most from ‘MOPS rich’, probably because it only supports about 1/30 of the Fis dimers produced with MOPS rich. There are also examples of binding peaks gaining in strength or appearing for the first time under slow-growth conditions, but they are less common (Fig. 4-3C-F). Because the apparent binding sites appearing in the ‘MOPS Ala’ media may reflect the noisier data under the low Fis conditions, these will need to be confirmed by additional ChIP and by in vitro experiments. In general, however, Fis targets the same sites in the different media, but the degree of occupancy and height of the binding region is often reduced.

Fig 4-3 Fis binding decreases when slower-growth conditions are used. Six representative 20 kb regions were randomly chosen. The Fis binding profiles under 'MOPS rich', 'MOPS Glu', and 'MOPS Ala' media conditions are displayed simultaneously. The black dotted boxes mark peaks that basically remain the same, whereas alterations are highlighted by arrows. Downward pointing arrows indicate that peaks decrease as cell growth decreases and upward pointing arrows emphasize peaks that form only or increase in poor media. Many peaks decrease in height and width (D, E, F), but most broad peaks remain strong (A, B, C). Binding profiles in 'MOPS Ala' media are quite different from the other two by displaying the biggest reduction in peak heights and the most unique binding events, which may reflect noise.
Genome-wide Fis binding remains largely unchanged in the absence of the major nucleoid-associated proteins H-NS and HU. Fis binding was measured in an *hns* and *hupAhupB* deletion strain to see if there is competition or cooperation between the three major NAPs. Surprisingly, Fis binding peaks in these mutant strains are very similar to wild type as demonstrated by the six representative regions in Figure 4-4. Although there are some subtle differences in peak height and width, there is no major rearrangement of Fis binding sites. Thus, it seems that on a global scale Fis binds its target sites independently of H-NS and HU.
In vitro Fis binding studies on selected *E. coli* chromosome regions. Fis binding sites cannot be easily recognized by DNA sequence alone since the 15 bp consensus sequences is highly degenerate (G__Y R W W W Y R _ _ C, where Y is C or T, R is A or G, and W is A or T) and other parameters such as DNA shape, particularly that of the minor groove, are also important recognition signals (Hancock et al., 2013; Stella et al., 2010). Nevertheless, Fis in vitro binding to strong sites can be sufficiently specific and stable to result in low nanomolar binding constants (Stella et al., 2010). As shown in Figure 4-2, genome-wide Fis binding profiles result either in broad (≥ 1 kb) or narrow (≤ 500 bp) peaks. We performed Fis EMSA experiments on 18
genomic loci to address the following questions. Do regions with strong in vivo signals display strong binding affinities in vitro? Does the strength of in vitro Fis binding correlate with whether a region remains occupied when less Fis protein is available? Do broad peaks represent an array of Fis binding sites or do they result from localized high affinity sites that recruit more Fis protein to the region? Moreover, in vitro binding data will be invaluable for helping with the computational analysis of Fis binding peaks, particularly with respect to assigning peak cut-off values.

In vitro Fis binding studies show that narrow Fis binding peaks contain at least one high-affinity Fis site, but the strength of these sites does not predict occupancy with decreasing Fis levels. We randomly chose nine narrow (≤500 bp) Fis-bound regions that give strong binding signals in ‘MOPS rich’ media. Six of them exhibit a decreased signal under slow-growth conditions (Fig. 4-5A) and three remain the same (Fig. 4-5B). As a control we used the F site at attR (region #10), which is part of the λ prophage and a known strong single Fis site in vitro (Papagiannis et al., 2007; Thompson et al., 1987). This region contains a strong peak as expected, and it remains high even when Fis levels are low (Fig. 4-5B). As negative controls we picked four regions that are not bound by Fis under any conditions and one region that only appears occupied under minimal media conditions (#17) (Fig. 4-5C). All regions were amplified as 500 bp DNA probes and radioactively labeled at the 5’ end for Fis EMSA experiments. We determined apparent $K_d$ values for Fis binding to each probe in the presence of 50 µg/ml competitor salmon sperm DNA. $K_d$ values of ≤ 5 nM were chosen as the cutoff for high-affinity Fis binding complexes.

In general, all of the ten regions that show positive Fis binding in vivo also display high affinity for Fis in vitro (1.1- 4.3 nM). For the λ F site we obtain binding affinities similar to the
ones reported earlier (1.3 nM vs. 2.2 nM) (Papagiannis et al., 2007). As predicted, regions that are not Fis-bound in vivo weakly attract Fis in vitro (11-23 nM) (Fig. 4-5C). Region #17, which has the unusual property of only being Fis-bound under minimal media conditions, exhibits intermediate affinity (5.7 nM). Furthermore, all of the ten peaks contain one high-affinity site, but according to the EMSAs they also contain additional lower affinity sites so more than one Fis dimer could potentially be binding in vivo. Based on the regions we sampled, we can conclude that regions with strong in vivo Fis binding also display strong in vitro binding. Using media conditions that decrease cellular Fis levels can have a significant effect on Fis binding (Figure 4-3). Six of the small binding peaks we selected have lower heights in the minimal medias (Fig. 4-5A) whereas four remain unaltered (Fig. 4-5B). The Fis binding peaks from both sets contain a single high-affinity in vitro Fis site and their range of $K_d$ values is the same (1.1/1.3-4.3 nM). This suggests that $K_d$ values measured on naked DNA do not directly correlate with in vivo binding affinity and cannot be used to predict Fis occupancy under low $fis$ expression conditions.

**In vitro Fis binding studies show that broad Fis binding peaks are correlated with clusters of high affinity sites.** We have dissected two broad Fis binding regions by splitting them up into four 500 bp segments and testing their Fis binding affinities in vitro (Fig. 4-6A, B). Segment #2 in the *dusB-fis* promoter region is known to contain multiple Fis binding sites that function to repress transcription (Ball et al., 1992). Region #12 covers the divergently oriented *aer* and *patA* promoters; in vitro Fis binding studies over this region have not been reported. For ‘MOPS rich’ media the Fis binding signal is quite high throughout both extended regions, but shrinks in ‘MOPS Glu’ and ‘MOPS Ala’. The *dusB-fis* peak is more asymmetric in comparison
Fig. 4-6 Dissecting Fis binding to broad chromosomal regions. EMSA experiments are performed on regions shown in (A) and (B) by splitting them into four 500 bp sections as marked by vertical dotted lines. (A) This asymmetric binding peak forms over the fis promoter, which has previously been shown to contain multiple high-affinity Fis sites (Ball, Osuna et al., 1992). Segment #2 displays the highest in vivo signal under all growth conditions and contains three high affinity sites. (B) A Fis binding peak centers over the intergenic region between the aer and patd ORFs and has not been analyzed previously. Only segment #12 contains high affinity Fis sites (4) and remains occupied as fis expression drops. The K values for segment #2 (A) and #12 (B) are about the same as for the best small peak (1 nM vs. 1.1 nM for #22, Fig. 4-5A), but they contain more high-affinity sites (3-4). Moreover, the outermost segments exhibit weak in vitro binding and also experience a sharp decrease in Fis occupancy under slow-growth conditions. This suggests that at least for segments of broad peaks in vitro binding correlates well with in vivo binding under conditions producing low fis expression. Additional EMSA experiments on chromosome regions shown in (C, D, E) are planned.
to the *aer/patA* peak. Fis EMSA experiments show strong in vitro Fis binding only for inner segments (#2 and #12) of the broad binding peaks. Segments on the periphery display very weak binding (Fig. 4-6A, B) and give affinities similar to regions that exhibit background binding in vivo (Fig. 4-5C). For segment #2 we observe three Fis binding sites with $K_d$ values $\leq 5$ nM and none for adjacent segments. Earlier DNaseI footprinting studies over the *dusB-fis* promoter region detected up to six binding sites, although higher amounts of Fis and no competitor DNA was used (Ball et al., 1992). As for the segment #12 *aer-patA* intergenic region, we observe four high-affinity complexes, but none form over the adjacent coding regions (segments #11, #14, #15). Footprinting assays of this region using DNase I and Fis-chemical nuclease chimeras are currently in progress to determine the exact position of these Fis sites. In summary, the binding data collected for two broad Fis binding regions suggests a model whereby Fis binding to localized clusters of high affinity sites may function to nucleate binding into adjacent regions. More broad peaks like the ones shown in Figures 4-6D, E, and F need to be dissected by EMSA and footprinting experiments to further validate this model.

**Discussion**

**Previous genome-wide Fis binding studies.** Three other groups have previously published Fis binding data on a genome-wide scale, but there is remarkable variation in their binding sites (Cho et al., 2008; Grainger et al., 2006; Kahramanoglou et al., 2011). Kahramanoglou and coworkers (2011) report only a 67% and 31% overlap between their ChIP-seq data and the high and low resolution ChIP-chip data of Cho et al. (2008) and Grainger et al. (2006), respectively. Both ChIP-chip studies (Grainger et al., 2006) used minimal M9 growth media supplemented with fructose and glucose respectively, whereas Kahramanoglou et al.
(2011) used rich LB media. Despite data variability, all three groups, like us, find that Fis can cover long stretches of the *E. coli* chromosome.

**Fis binding under different expression levels.** Our ChIP-chip experiments demonstrate that Fis binds to numerous regions throughout the *E. coli* chromosome, and our in vitro experiments show that most of the binding peaks are composed of true high-affinity binding sites. Some peaks are quite narrow (≤500 bp, within the range of the sonicated DNA lengths) and have one high-affinity Fis site (Fig. 4-2, 4.5A, B), whereas broader peaks contain clusters of high affinity sites (Fig. 4-6A). For cells harvested from media supporting lower fis expression, narrow Fis binding peaks are much more affected than broad binding peaks (Fig. 4-3). Most narrow peaks become lower as Fis levels decrease (Fig. 4-3 Fig. 4-5A), but this behavior correlates poorly with in vitro binding affinities. For this reason, factors in addition to DNA sequence seem to be involved in Fis recruitment; these could be other DNA binding proteins, local effects of DNA topology and supercoiling, or differences within proposed chromosomal macrodomain structures (Boccard et al., 2005).

**Global Fis binding appears largely independent of the other major NAPs H-NS and HU.** In addition to Fis, H-NS and HU proteins are also proposed to be involved in organization of the *E. coli* chromosome. H-NS has been mainly implicated in forming the topologically-independent nucleoid loops (Dame et al., 2006; Dame et al., 2002; Noom et al., 2007; Schneider et al., 2001) and HU in compacting the nucleoid through DNA bending (Skoko et al., 2006; van Noort et al., 2004). H-NS and Fis can directly compete for overlapping binding sites in promoter regions in vitro (Falconi et al., 1996; Zacharias et al., 1991), and genome-wide studies have concluded that their binding profiles are mutually exclusive (Kahramanoglou et al., 2011). HU has been proposed to compete for binding sites with H-NS (Dame et al., 2002; Manna and
Gowrishankar, 1994; van Noort et al., 2004), but because HU and Fis are both DNA bending proteins, they could potentially act synergistically. In order to determine whether Fis binding is affected by H-NS and HU, we deleted their genes and measured global Fis binding. We conclude that neither of these abundant nucleoid proteins seems to influence global Fis binding under rapid growth conditions.

**Does Fis play a role in nucleoid organization?** In general, our Fis binding data provides some indirect support for a role for Fis in the organization of the nucleoid. We show that Fis binds to broad regions (Fig. 4-2) and that these regions remain occupied (albeit to smaller degree) even when much less Fis is present (Fig. 4-3, Fig. 4-6), suggesting that Fis binding at these specific loci is of importance. We show that these long Fis tracts arise from a few high affinity sites that function to nucleate spreading along the DNA (Fig. 4-6). Spreading of H-NS along DNA has also been proposed to nucleate at weakly specific binding sites (Bouffartigues et al., 2007; Rimsky et al., 2001). Because genome-wide Fis binding correlates weakly with differential gene expression in a *fis* deletion strain, most of the Fis binding events are probably involved in nucleoid organization (Cho et al., 2008; Kahramanoglou et al., 2011).

Caglieri et al. (2013) and Wang et al. (2011) recently questioned the role of Fis in chromosome organization in studies employing chromosome conformation capture (3C) and fluorescence microscopy, respectively. Both studies conclude that there is no spatial clustering of Fis binding sites, which contradicts ChIP and previously published immunofluorescence data showing Fis binding foci (Azam et al., 2000; Kahramanoglou et al., 2011). However, Caglieri et al. (2013) and Wang et al. (2011) both use data collected under sub-optimal growth conditions that keep Fis levels low. We think that comparing chromosome conformation capture profiles from wild type and *fis* deletion strains would be a better way to test if Fis drives nucleoid
compaction. In summary, several lines of evidence support a role for Fis in chromosome organization, but direct in vivo proof is still missing.

Methods

Bacterial strains and growth conditions

An *E. coli* MG1655 *pyrE* + strain containing a λ prophage (RJ3798) was used as the wild-type strain for the ChIP-chip experiments. All other strains are derivatives of this parent strain, with genes deleted by standard P1 bacteriophage transduction methods (RJ3905 *fis::kan*; RJ3833 *hupA::CAM, hupB::kan*; RJ3834 *hns::kan*). The mopholinepropanesulfonic acid (MOPS) media was prepared according to the instructions provided by Neidhardt and co-workers (Neidhardt et al., 1974). ‘MOPS rich’ media was supplemented with additional nutrients, which was purchased from Teknova (10X ACGU and 5X Supplement EZ solution). As carbon sources we used either 0.3% glucose or 0.3% L-Alanine. In order to achieve steady-state growth conditions an initial 15-hour overnight culture was diluted 1:100, grown to an OD$_{600}$ between 0.8 and 1.0, and stored overnight at 4°C. On the next day, the culture was diluted and grown in the same manner again. After overnight storage at 4°C, the subculture was diluted 1:1,000 into 20 ml of fresh media. Samples for Western blotting, viable cell counts, and OD$_{600}$ measurements were collected to determine Fis dimers/CFU and doubling times. Cells for chromatin immunoprecipitation were grown to an OD$_{600}$ between 0.3 and 0.4.

Western Blotting

Quantitative Fis Western blotting was essentially performed as described in Nafissi et al. (2012). Rabbit polyclonal Fis antibody and mouse-rabbit IgG DyLight 800 conjugate were used to visualize the blot by the Odyssey imaging system (LI-COR). Standard curves obtained from
purified Fis protein electrophoresed alongside the samples were used to determine the amount of Fis present in each sample.

**Chromatin Immunoprecipitation (ChIP)**

Cells were grown under steady-state growth conditions until they reach an OD$_{600}$ between 0.3 and 0.4. Cells (1 ml) were suspended in a M9 minimal salt solution (Miller, 1992) and subjected to crosslinking with 1% formaldehyde, which was quenched after 20 min with 55 mM glycine. Cells were collected by centrifugation, washed with PBS, and then stored at -80°C. Cells were lysed by sonication in 250 µl of lysis buffer composed of 10 mM Tris-HCl (pH 8.0), 20% sucrose, 50 mM NaCl, 10 mM EDTA, 0.2% Triton-X, and 1 mM PMSF. The DNA was sonicated to fragment sizes ranging from 100-500 bp with a median around 250 bp using a Diagenode Bioruptor for 1 hr (30 s sonication followed by 1 min break). The extract was incubated with 3 µl of affinity-purified rabbit polyclonal Fis antibody for one hr long at 4°C. Forty µl of dynabead protein G bead slurry (Invitrogen) was added to the antibody reaction and buffer conditions were adjusted to 50 mM HEPES, pH 7.5, 100 mM NaCl, 1mM PMSF. Following overnight incubation, the beads were washed twice with the following three buffers: low salt buffer (150 mM NaCl, 1 mM EDTA, 0.1% sodium deoxycholate, 0.1% Triton-X, 50 mM HEPES, pH 7.5), high salt buffer (same as low salt buffer except 500 mM NaCl), and LiCl buffer (250 mM LiCl, 1 mM EDTA, 0.5% sodium deoxycholate, 0.5% Triton-X, 10 mM HEPES, pH 7.5). After a brief wash in 10 mM Tris-Cl, pH 8.0 + 1mM EDTA, the protein-DNA complexes were eluted twice with 100 µl of a solution containing 1% SDS and 0.1 M NaHCO$_3$ at 65°C. The crosslinking was reversed by incubating the eluate in 200 mM NaCl at 65°C for 2 hrs or overnight. Each sample was treated with 40 µg of Proteinase K for 30 min at 37°C, followed by phenol extraction and ethanol precipitation. As an ‘input’ control we used an aliquot
of the lysate (1/10 of the volume) that was not incubated with beads and antibodies, but underwent all other processing steps. Data from two biological replicas were combined to generate the genome browser tracks for each ChIP experiment.

**Amplification of DNA**

We followed the guidelines provided by Affymetrix for amplifying the immunoprecipitated and input DNA. In general, we initially amplified ChIP-precipitated DNA with random primer A (GTTTCCAGTCACGGTC(N)9, HPLC purified) and Sequenase (#70775Y, USB-Affymetrix). Subsequent amplification was carried out with primer B (GTTTCCAGTCACGGTC) and Crimson Taq polymerase (NEB) for 25 or 30 cycles. ChIP samples prepared from wild-type cells grown in ‘MOPS rich’ were amplified for 30 and 25 cycles; these generated very similar binding profiles. The DNA collected from wild-type cells grown in ‘MOPS Glu’, and ‘MOPS Ala’ was amplified for 30 cycles, whereas the DNA obtained from the *hns* and *hupA-hupB* strains that were grown in ‘MOPS rich’ was amplified for only 25 cycles. Amplified DNA was purified with Alquick Qiagen columns, followed by fragmentation and labeling according to Affymetrix GeneChip WT Double-stranded Terminal Labeling Kit (# 900812).

**Whole-genome-tiled microarray analysis**

The high-density tiling arrays were custom-designed for *E. coli* K-12 strain MG1655 and bacteriophage λ (Vora et al., 2009). The arrays contain 2.3 million 25-mer oligonucleotides with a tiling density of an 8-base spacing per strand, which gives a 4-bp resolution across the entire genome.

**Data Analysis**

All microarray data was smoothed in 200-bp windows and normalized to input controls collected from wild-type and *fis* cells grown in ‘MOPS Rich’ and ‘MOPS Glu’ media. Because
some ChIP samples were amplified for 25 and others for 30 cycles, they were only normalized to input controls that underwent similar amplification.

**Electrophoretic Mobility Shift Assay**

The Fis protein used for the binding studies was purified according to Stella et al. (2010). The same Fis stock solution was used for all experiments, and its concentration was determined by the SDS-PAGE analysis with lysozyme standards. Amplification of *E. coli* chromosome regions was accomplished by using genomic DNA extracted from a MG1655 derivative strain (RJ3454) as template. EMSAs were performed in 20-µl reactions containing 0.5 mg/ml BSA (Roche), 20 mM HEPES (pH 7.5), 1 mM EDTA, 1 mM DTT, 5% glycerol, 150 mM NaCl, 0.05 µg/µl sonicated salmon sperm DNA (Rockland), and 3,000 cpm ³²P-end-labeled DNA probe. Binding reactions were incubated at room temperature for 10 min before electrophoresis on 5% polyacrylamide gels in 0.5 X TBE buffer. For most DNA probes, Fis concentrations ranging from 1 nM to 45 nM were used in 2-fold incremental steps. The Kₐ values were determined by estimating the amount of Fis needed to bind half the DNA probe, and each probe was assayed at least twice.


involvement of FIS and H-NS proteins in the transcriptional control of hns expression. Mol
Microbiol 19, 965-975.

Filutowicz, M., Ross, W., Wild, J., and Gourse, R.L. (1992). Involvement of Fis protein in

proteins with coding and non-coding segments of the Escherichia coli genome. Nucleic Acids
Res 34, 4642-4652.

of DNA minor groove width and Fis protein binding by the purine 2-amino group. Nucleic Acids
Res.

Holmqvist, E., Reimegard, J., Sterk, M., Grantcharova, N., Romling, U., and Wagner, E.G.
(2010). Two antisense RNAs target the transcriptional regulator CsgD to inhibit curli synthesis.
EMBO J 29, 1840-1850.


Kahramanoglu, C., Seshasayee, A.S., Prieto, A.I., Ibberson, D., Schmidt, S., Zimmermann, J.,
Benes, V., Fraser, G.M., and Luscombe, N.M. (2011). Direct and indirect effects of H-NS and


Gene 361, 13-37.

growth phase-dependent regulation, growth rate-dependent control, and stringent control of fis

Growth phase-dependent regulation and stringent control of fis are conserved processes in enteric


Harbor Press).


Chapter V

Conclusion
I wanted to find out what drives efficient *fis* translation and to my surprise stumbled upon two non-canonical RNA elements: an AU-rich segment (AU) and an RNA secondary structure (SL). Both types of elements have been described previously, but their method of action is not very well understood (Boni et al., 2001; Boni et al., 1991; Holmqvist et al., 2010; Olins and Rangwala, 1989; Skorski et al., 2006; Zhang and Deutscher, 1989). Preliminary results suggest that the AU segment enhances translation via 30S recruitment and the SL structure via some anti-inhibitory pathway opening up possibilities for novel mechanisms to control Fis levels.

Traditionally, transcription has been considered to drive most *fis* expression (Ball et al., 1992; Mallik et al., 2006; Mallik et al., 2004; Ninnemann et al., 1992; Walker et al., 2004), but translational regulation may be required for fine-tuning and/or faster response. The SL and AU elements are good candidates for translational regulators, especially under conditions that quickly induce (nutritional upshift) or turn off (amino acid starvation) *fis* expression.

These findings contribute valuable information to the field of bacterial translation since only a few non-canonical *E. coli* mRNAs have been characterized in detail. We report the fourth example of a positively acting RNA structure (Boni et al., 2001; Holmqvist et al., 2010; Zhang and Deutscher, 1989) and present only the second thorough analysis of a native AU-rich enhancer (Zhang and Deutscher, 1992). Further dissection of these elements should help to better

---

**Fig. 5-1** 5′UTRs of nucleoid-associated proteins. RNA sequences preceding the start codons of major nucleoid-associated proteins are aligned. The putative SD sequences are underlined and they are quite variable. All 5′UTRs contain AU-rich sequences (highlighted in grey) and they could potentially be involved in efficient translation of the downstream messages.
understand their mechanisms and this knowledge can be applied to the other related systems. In contrast to the three known examples of activating RNA secondary structures, AU-rich sequences likely enhance translation via a common underlying mechanism. Since the exact function of AU-rich segments is still unclear (Kozak, 2005), translational studies of other highly abundant proteins could be very insightful. As shown in Figure 5-1 all major nucleoid-associated proteins also have AU-rich sequences in their 5’UTRs and I am curious to see if they behave like the fis AU element. Altogether, the data presented in this thesis advances our understanding of noncanonical mechanisms directing bacterial translation. Given the wide variety of mechanisms controlling translation, I wonder if there are other novel pathways still waiting to be discovered.


