Deconstructing Ihf-Mediated Inhibition of the Complex acs Promoter

David S. Thach
Loyola University Chicago

Follow this and additional works at: https://ecommons.luc.edu/luc_diss

Part of the Microbiology Commons

Recommended Citation
Thach, David S., "Deconstructing Ihf-Mediated Inhibition of the Complex acs Promoter" (2010). Dissertations. 36.
https://ecommons.luc.edu/luc_diss/36

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License.
Copyright © 2010 David S. Thach
LOYOLA UNIVERSITY CHICAGO

DECONSTRUCTING IHF-MEDIATED INHIBITION OF THE COMPLEX αcs PROMOTER

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE GRADUATE SCHOOL
IN THE CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

PROGRAM IN MICROBIOLOGY AND IMMUNOLOGY

BY

DAVID SOKHA THACH

DIRECTOR: ALAN J. WOLFE, PH.D.

CHICAGO, ILLINOIS

MAY 2011
Copyright by David Sokha Thach, 2010
All rights reserved
ACKNOWLEDGEMENTS

I would like to thank my mentor, Dr. Alan Wolfe, for his help, support, and infinite patience. Without his guidance and daily reassurance I would no doubt have suffered a breakdown in sanity. I am also very grateful for the masterful maneuvering of my committee members Dr. Gayatri Vedantam, Dr. Nancy Zeleznik-Le, Dr. David Keating, and Dr. Thomas Gallagher. Their advice, confidence, and honesty guided me to become a better student and scientist. A very special thank you to Dr. Karen Visick for not accepting less than my best effort during joint lab meetings.

I would like to thank my former lab mates for their enjoyable company and interesting conversations about all things science as well as all things mundane. I would like to especially thank Dr. Christine Beatty for imparting all of her worldly knowledge onto a second year grad student while making science look easy.

Thank you to my friends and classmates for the good memories of graduate school as well as their support. I would like to thank Guy Townsend for his unrelenting friendship and support. I would like to say a special thank you to my dearest friend, Maike Müller, my confidant, my trusted colleague, and my closest friend.

Finally, but with no lesser degree of importance, I would like to thank my family. They are my motivation to become a better person today than the day before.
For my parents who gave everything they could to their children.
Failure is the opportunity to begin again more intelligently.

- Henry Ford
TABLE OF CONTENTS

ACKNOWLEDGEMENTS                       iii
LIST OF TABLES                           ix
LIST OF ILLUSTRATIONS                   x
ABSTRACT                                 xii

CHAPTER ONE: LITERATURE REVIEW
   Introduction                           1
   The process of transcription            4
     Transcription and Human disease       5
     Transcription and bacterial pathogenesis 8
     Mutations that alter transcription    11
     Bacterial transcription: the textbook view 13
   Adenine, Cytosine, Guanine, Thymine, and base pair stacking 15
     Base Pair Steps                       19
     The propeller twist                   21
   Genome compaction                      23
     Compaction of the Eukaryotic genome  24
     Compaction of the Bacterial genome   26
   Bacterial nucleoid associated proteins 30
     IHF                                   31
     NAPs and transcription: location, location, location 34
     CRP-dependent activation               38
   Acetyl-CoA synthetase                   40
     acs operon and promoter architecture  43
     Activation of acsP2                   44
     Modulation of CRP-dependent activation by nucleoid proteins 46
     Negative regulation by FIS            46
     Negative regulation by IHF            47
   Concluding remarks                     48

CHAPTER TWO: MATERIALS AND METHODS       50

CHAPTER THREE: EXPERIMENTAL RESULTS
THE MULTIPLE ROLES OF CRP AT THE COMPLEX ACS PROMOTER DEPEND
ON ACTIVATION REGION 2 AND IHF           vi
Introduction 62
Inhibition of CRP-dependent transcription by IHF III depends on CRP II 66
Inhibition of transcription involves K101 of CRP 68
Characterization of protein binding 71
  CRP 71
  IHF 73
  CRP and IHF 74
  RNAP and CRP 74
  RNAP, CRP and IHF 75
Characterization of open complex formation 77
CRP stabilizes the OC in the presence of binding to lower affinity sites by IHF 79
Discussion 81
Concluding thoughts 93
Fis-dependent inhibition of transcription 95

CHAPTER FOUR: EXPERIMENTAL RESULTS
EVIDENCE THAT IHF III IS A COMPOSITE SITE
Introduction 99
Shared protein binding sites in the acs-nrf intergenic region 100
An argument that IHF III is a composite site 100
Preliminary evidence supporting the composite site hypothesis 102
Systematic mutations of IHF IIIa and IHF IIIb sites 104
  Mutations in the 3'-TTR motifs 106
  Mutations that disrupt one site or the other but not both 108
Replacement of IHF IIIa with IHF IIIb and vice versa 110
Activity from the 1xIHF-VF and 1xIHF promoters involves the same activation regions 113
Discussion 114

CHAPTER FIVE: EXPERIMENTAL RESULTS
EVIDENCE THAT THE UPSTREAM SINGLE IHF II SITE ANTAGONIZES
  IHF IIIa/b- MEDIATED INHIBITION
Introduction 118
The palindrome does not modulate IHF-mediated inhibition 119
Evidence that IHF II is a positive modulator of IHF III-mediated inhibition 122
Individual analysis of IHF IIIa and IHF IIIb 126
Summary 129

CHAPTER SIX: DISCUSSION
Prior studies 135
IHF mediates formation of a stalled open complex 141
Countering IHF-mediated inhibition 144
IHF III is a composite site 147
IHF III-mediated inhibition of CRP-dependent transcription involves dual overlapping IHF III binding sites 153
Tipping the Scales of acs Promoter Regulation by IHF IIIa/b 158
IHF II is a Positive Modulator of IHF-Mediated Inhibition 160
The Dynamics of Fis- and IHF-dependent Inhibition on CRP-dependent Activation 163
Relevance and contribution to the field 165

REFERENCE LIST 171

VITA 191
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Strains, plasmids and phage not listed in Tables 2-4</td>
<td>57</td>
</tr>
<tr>
<td>2. Promoter fragments used with corresponding plasmid, phage, and strains</td>
<td>59</td>
</tr>
<tr>
<td>3. Promoter fragments not listed in Tables 1 or 2, used in Chapter 4</td>
<td>60</td>
</tr>
<tr>
<td>4. Promoter fragments not listed in Tables 1-3, used in Chapter 5</td>
<td>61</td>
</tr>
</tbody>
</table>
# LIST OF ILLUSTRATIONS

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Dinucleotide Slide and Twist</td>
<td>19</td>
</tr>
<tr>
<td>2.</td>
<td>Histone-Mediated DNA Compaction</td>
<td>26</td>
</tr>
<tr>
<td>3.</td>
<td>Crystallographic Model of IHF Bound to the λH’ Site</td>
<td>38</td>
</tr>
<tr>
<td>4.</td>
<td>CRP-Dependent Activation</td>
<td>39</td>
</tr>
<tr>
<td>5.</td>
<td>Acetyl-CoA Synthetase (Acs) and Transcription of its Gene (acs)</td>
<td>40</td>
</tr>
<tr>
<td>6.</td>
<td>The acs Operon and the nrf-acs Intergenic Region</td>
<td>43</td>
</tr>
<tr>
<td>7.</td>
<td>CRP Activates acsP2 Transcription Using the Class III Mechanism that Involves Tandem Class I Sites</td>
<td>45</td>
</tr>
<tr>
<td>8.</td>
<td>FIS-Dependent Anti-Activation of CRP-Dependent acs Transcription</td>
<td>48</td>
</tr>
<tr>
<td>9.</td>
<td>The organization of the acs promoter</td>
<td>64</td>
</tr>
<tr>
<td>10.</td>
<td>IHF III-dependent inhibition involves AR2 and AR3 of CRP</td>
<td>68</td>
</tr>
<tr>
<td>11.</td>
<td>Simultaneous binding of IHF, CRP, and RNAP results in protection between the tandem CRP sites</td>
<td>72</td>
</tr>
<tr>
<td>12.</td>
<td>Simultaneous binding of IHF, CRP, and RNAP results in protection between the tandem CRP sites</td>
<td>73</td>
</tr>
<tr>
<td>13.</td>
<td>IHF and CRP influence open complex formation</td>
<td>76</td>
</tr>
<tr>
<td>14.</td>
<td>CRP influences the structure of the open complex independently of IHF</td>
<td>78</td>
</tr>
<tr>
<td>15.</td>
<td>Binding of IHF to low affinity sites inhibits RNAP</td>
<td></td>
</tr>
</tbody>
</table>
binding in the absence of CRP

16. Model of the complexes formed at the acsP2 promoter

17. IHF-mediated inhibition requires Fis II

18. Both Fis and IHF reorganize RNAP binding from P2a to the productive P2U promoter

19. The putative IHF IIIa and IHF IIIb sites

20. Initial mutagenesis of IHF III

21. Mutations reported to cause defects in IHF binding

22. Systematic mutagenesis of the putative composite IHF III site

23. Swapping IHF IIIa and IHF IIIb

24. 1xIHF and 1xIHF-VF activities involve AR2 and AR3 of CRP

25. A folding model for IHF IIIa versus IHF IIIb in the context of 1xIHF

26. Presence of the complete palindrome sequence increases transcription from 1xIHF+5 promoter

27. Mutating IHF IIIa in the context of 2xIHF alleviates inhibition

28. IHF II is a positive modulator of acs transcription

29. Proposed model of simultaneous occupation of IHF III and IHF II

30. IHF bound at the IHF IIIa or IHF IIIb positions can be a negative or positive effector, respectively
ABSTRACT

acs encodes a high affinity enzyme that permits survival during carbon starvation. As befits a survival gene, its transcription is subject to complex regulation. Previously, the Wolfe lab reported that CRP activates acs transcription by binding tandem DNA sites located upstream of the major acsP2 promoter and that the nucleoid protein IHF binds three specific sites located just upstream. The most proximal site (IHF III) exhibits reduced transcription compared to the full-length promoter or to a construct lacking all three IHF sites. The goal of my research was to understand how IHF III inhibits CRP-dependent acs transcription. First, I helped define the minimal system required for this IHF-dependent inhibition, showing it requires the promoter-distal CRP site and an amino acid residue located within a surface determinant of CRP that interacts with RNAP. Surprisingly, for a Class III promoter, disruption of this surface determinant caused significant changes in the activity and structure of both the full-length promoter and the construct with the single proximal IHF site. My collaborator, Dr. Bianca Sclavi (Laboratoire de Biotechnologies et Pharmacologie génétique Appliquée, Paris, France) showed that occupancy of IHF III mediates formation of a stalled unproductive transcription complex. This work was published in Molecular Microbiology.

I furthered this research, obtaining evidence that IHF III is actually a composite site consisting of two overlapping IHF sites that sit on opposing faces of the DNA helix. This composite appears to behave as a transcriptional regulatory switch. If IHF occupies
one site, acs transcription occurs. If IHF occupies the opposing site, acs transcription is
inhibited. Which site becomes occupied appears to involve occupancy of IHF II, which is
located just upstream of IHF III.

This work demonstrates that the typical textbook view bacterial transcription is
overly simplistic. In fact, bacterial transcription can be quite complex.
CHAPTER ONE
INTRODUCTION AND LITERATURE REVIEW

Introduction

Molecular archeologists have uncovered trace evidence to suggest that early during evolution the first form of genetic information was the self-replicating RNA molecule (Doudna and Cech, 2002). Despite the lack of direct physical evidence, this is not hard to believe especially given the wide variety of RNA tertiary structures and their ability to carry out heterogeneous enzymatic reactions. As early RNA molecules self-replicated, mutations must have been introduced that improved the fitness of the molecules (Joyce, 2002; Doudna and Cech, 2002).

It is estimated that the dominance of the RNA world ended almost four billion years ago (Joyce, 2002). The exact selective pressure that favored the evolution of the heritable genetic material from RNA to DNA is unknown. Perhaps as the environmental milieu continued to change so too did the chemistry of RNA until it evolved into its more stable relative, DNA. A possible mechanism of DNA evolution may be based on an RNA template, as is the case in ciliate
genome replication. Studies on genomic replication in the ciliate *Oxytricha trifallax* researchers have uncovered a unique role for two forms of RNA: small RNAs and long macronuclear RNA. Two nuclei exist that house the genomic DNA, albeit in different forms. The micronucleus is small in volume and envelops germline DNA. The larger macronucleus contains fragmented genomic DNA, or somatic DNA, in copies as high as 1000 copies per of individual gene. Genes in the micronucleus are “scrambled” whereas DNA in the macronucleus is organized to code for complete genes. It is postulated that small RNAs in the micronucleus bind to short repeat sequences (or “pointers”) that aid in correctly joining the gene fragments resulting in sense DNA. During development DNA in the parental macronucleus is transcribed to produce long macronuclear RNAs, which then provides the template for unscrambling the DNA strands in the developing macronucleus (Mochizuki, 2010 and Nowacki et. al., 2008). The process of RNA-directed reorganization of DNA also occurs in *Tetrahymena thermophila* and the mitochondrion of *Diplonema papillatum*. Circumstantially, the RNA-directed editing of genomic DNA may be the same mechanism of how nucleotides evolved into the more stable DNA strand of ancestral simple eukaryotes and, thus, multicellular organisms. Though the RNA-dominant world has long been extinct, the importance of RNA in biological processes remains evident (Joyce, 2002).

Today, the Central Dogma of biology states that the genetic content of DNA is converted, or transcribed, into messenger RNA (mRNA) (Crick, 1970).
The information contained within the mRNA is then translated into protein. Regulation of the transcription process is essential to ensure that genes are appropriately expressed, i.e. at the right time, in the correct location, in the required amount, and in coordination with the proper subset of associated genes. (Lewin, 1997; Little, 1999; Alberts, 1994)

Why study transcription? First, the inability to properly regulate transcription often disrupts cellular homeostasis, which can lead to undesirable events, such as disease or the loss of organismal fitness (Lewin, 1997; Alberts, 1994; Calkhoven, 1996). To understand how improper transcription impacts the etiology of disease or the ability of cells to survive adverse environments, we must understand how cells normally regulate transcription. Furthermore, the exploration of transcription in both eukaryotes and bacteria has implications for understanding evolution gene regulation (Ooi and Wood 2008; Lee et al, 2007; Cattaneo, 2005; and Dozois, et al, 2002). The more we learn about both the common themes and the differences of transcriptional control in both kingdoms, the closer we come to uncovering clues about transcription in our common ancestors. For example, the principles that govern the initial barrier to transcription are the same – in both kingdoms, increases or decreases of DNA twisting are required for gene silencing or gene transcription, respectively (Travers and Muskhelishvili, 2007). Finally, the more we learn about the rules of gene regulation, the better we are able to design transcription-based therapies to treat devastating diseases such as cancer, to devise strategies to defeat
antibiotic resistant bacteria, or to construct semi-synthetic gene circuits to perform desired functions. For example, one might modulate access to a promoter region or use a small molecule to disrupt a critical protein-DNA interaction (Hsu et al., 2007; Pandolfi, 2001).

**The Process of Transcription**

In general, the process of transcription is the same for eukaryotes, archaea, and bacteria. Transcription takes place in four stages. The first step requires that RNA polymerase (RNAP) bind to promoter DNA, forming a binary protein/DNA complex known as the closed complex (CC). Once RNAP binds to the promoter, the DNA strands unwind (a process known as promoter melting or isomerization). The result is known as the open complex (OC). Now, ribonucleotides enter the RNAP, which begins to synthesize RNA using the sense strand of DNA as the template. The result is a short RNA transcript. This protein/DNA/RNA ternary complex is called the initiation complex (IC). Such short transcripts often abort. If so, the initiation process re-starts. If, however, transcription continues, the nucleoprotein complex dramatically alters its conformation, becoming the extremely processive form known as the elongation complex (EC). Polymerization of mRNA, or transcription, continues until termination signals/structures are encountered (Wolfe, 2008; Higgins, 2005) (reviewed by Rosenberg and Court, 1979).
Transcription can be enhanced or inhibited at any of these steps, but most often regulation occurs at one or more of the first three steps and is most often mediated through the binding of positive or negative regulators. These regulators can bind to sequences that overlap the promoter or bind to sequences positioned either upstream or downstream of the promoter. As such, transcription regulation depends on two distinct features: 1) sequences within the DNA (known as cis elements) and 2) the proteins that bind to those sequences (known as trans factors). Mutations that affect either cis elements or trans factors can result in debilitating defects in eukaryotes and bacteria alike (Ooi, 2008; Kadonaga, 2004).

**Transcription and Human Disease**

A disruption in transcription can result in an imbalance in cellular homeostasis, which can lead to metabolic inefficiency, arrest of cell maturation, reduced fitness, and cell or organismal death. Many severe human diseases result from aberrant gene expression (Pandolfi, 2001). Some of these diseases are attributed to mutations in cis elements, while others are attributed to mutations in trans factors. I will now describe a few examples of both.

Mutations in the gene that encodes a trans factor can disrupt the normal function of that protein. For example, Huntington disease is a debilitating neurodegenerative disorder that manifests itself when the normal huntingtin
HTT) gene acquires mutations that magnifies its endogenous CAG repeat 
(Scherzer, 2008; Cattaneo, 2005). The resulting mutant huntingtin gene (mHTT) 
contains greater than 37 CAG repeats. Since CAG encodes glutamine, this 
magnification results in an extended polyglutamine or polyQ tract in the mutant 
protein, known as mHTT. Normal HTT functions in numerous cellular processes, 
such as protein degradation and gene transcription. The extended polyQ tract of 
mHTT interferes with these processes. For example, the extended polyQ tract of 
mHTT disrupts its own N-terminal nuclear export sequence (NES). This 
interference results in a buildup of mHTT proteins in the nucleus. The nuclear 
localization of mHTT, specifically its extended polyQ domain, disrupts 
transcription. This is because many transcriptional regulators contain glutamine- 
rich domains that mediate the interaction between transcription factors and other 
transcription-associated proteins. mHTT proteins disrupts these interactions, 
possibly by sequestering the transcription factor (Kadonaga, 2004; Imarisio, 
2008). For example, CBP [CREB binding protein] is an important bridging 
protein (co-transcriptional activator) that is proposed to link the DNA-bound 
CREB transcription factor to the general transcriptional machinery. CBP activity is 
required for expression of neuron survival signals through its histone 
acetyltransferase function (Lu et. al., 2003). This activity decondenses the 
histone/DNA complex, allowing RNAP and transcription factors to gain access to 
the promoter region. The C-terminal region of CBP contains a polyQ stretch that 
mediates the interaction between CBP and CREB. mHTT protein sequesters
CBP, inhibiting DNA access and preventing transcription (Scherzer, 2008; Cattaneo, 2005).

Mutations in the promoter sequence may also alter gene expression. The etiology of Alzheimer’s Disease (AD) is attributed to the buildup of a spliced version of the full-length amyloid precursor protein (APP) – the 44 amino acid long amyloid β (Aβ) variant. Many tissues express APP, with the highest concentrations found in CNS neurons. Buildup of the Aβ variant form aggregates in brain tissue, creating plaques and causing mental deterioration of the patient (Sakanyan, 2005). Typically, the plaques are the result of increased splicing of the full-length APP protein mediated by a mutation in the gene changing a key amino acid from valine 717 to an isoleucine 717. Recently, however, Theuns et al. (Theuns et al, 2006) identified a mutation in the APP promoter region of patients exhibiting early onset stages of AD. They report that this mutation results in increased APP gene expression, leading to increased APP protein production, and a greater concentration of the Aβ variant. This mutation in the promoter region leads to AD disease progression, mirroring that of the disease mediated by the V717I mutation (Theuns et al, 2006; Theuns et al, 2000). Thus, this study has identified an alternate source of AD etiology that arises from a mutation in the app promoter region rather than from a mutation in the gene itself.
Transcription in Bacterial Pathogenesis

Likewise, bacteria regulate gene expression at the level of transcription. This regulation occurs in response to both external and internal signals that converge to aid in directing the transcription of genes needed to adapt to and survive changes in the bacterium's environment, such as the depletion of nutrients, the presence of antibiotics, and exposure to the host immune system. For example, the resultant alteration in gene program can result in the formation of differentiated bacterial entities, e.g. biofilms, fruiting bodies, or spores, formed to withstand harsh environmental conditions (Dorman and Corcoran, 2008; Kroos, 2007).

Of particular interest is the survival of pathogenic bacteria in an animal host. Entry into a host exposes the bacterial pathogen to nutrient-limiting conditions, a low-pH environment, and antigen exposure to the immune system. To survive in the host, the pathogen must express genes that encode virulence proteins used to alter its immediate environment for survival such as permeabilizing the membrane (membrane ruffling) or immunosuppression. Such virulence genes must be precisely regulated to ensure proper expression. To achieve this goal, bacteria have evolved sophisticated signaling networks and transcriptional machinery that permits them to differentiate between their *in vitro* and *in vivo* environments and to properly regulate the genes required for each environment (Kroos, 2007).
The machinery required to process the multitude of environmental signals and to properly transcribe a given gene often involves many signaling pathways that pass signals to multiple transcription factors that control gene expression from multiple promoters. Pivotal to this process is the constellation of DNA remodeling proteins, also known as histone-like proteins or nucleoid-associated proteins (NAPs) (Higgins, 2005; Cotter, 2000). Like eukaryotic histones, these NAPs bind, fold, and package the DNA. As such, they participate intimately in conformation-dependent processes including genome replication and the regulation of transcription. Of particular interest is the small heterodimeric DNA remodeling NAP, IHF (Integration Host Factor; ~21 KDa). IHF appears to be a key regulator of genes located on pathogenicity islands of Enteropathogenic *Escherichia coli* (EPEC) and Enterohemorrhagic *Escherichia coli* (EHEC), as well as the invasion-specific genes of *Shigella flexneri* (Baron, 2007; Tseng et al, 2009). The ability of IHF to remodel DNA and its affects on the function of proteins bound to the promoter region is the major focus of my dissertation.

Often, the remodeling of DNA topology by NAPs provides the appropriate DNA conformation to allow or disallow the binding of a transcription factor to its DNA site. Alternatively, NAPs can mediate the bringing of a transcription factor into close proximity with RNAP to permit a required protein-protein contact (Higgins, 2005; Ohniwa, 2006). Two well-studied transcription factors, FNR and CRP, regulate the transcription of genes that encode proteins responsible for survival under very specific conditions. For example, the homodimeric FNR
(fumarate and nitrate reduction; ~30 KDa) is a homodimeric transcription factor that responds to the environmental concentration of oxygen and, in response, regulates genes required for cell survival under anaerobic conditions. The functional status of FNR depends on the state of its oxygen labile [4Fe-4S]^{2+} cluster (Kiley and Beinert, 1998). Not only is the [4Fe-4S]^{2+} cluster important to the dimerization of FNR subunits, but it also behaves as the protein’s oxygen sensor. In the presence of oxygen (i.e., an aerobic environment), the [4Fe-4S]^{2+} cofactor undergoes oxidation to [2Fe-2S]^{2+}, losing two Fe and S^{2-} atoms, leading to a conformational change that eliminates DNA binding to its site (Kiley and Beinert, 1998). The reduction of FNR to its protein subunits thus results in deregulation of all FNR-dependent genes/operons. The complete depletion of oxygen from the environment (i.e., an anaerobic environment) restores the [4Fe-4S]^{2+} cofactor and thus the ability of FNR to bind to its DNA sites in the promoter regions of its regulon members. Expression of these genes permits the cell to adapt to an anaerobic environment, in part, by facilitating fermentation and anaerobic respiration (Ratledge and Dover, 2000). Additionally, FNR has been implicated as a virulence factor. During a Neisseria meningitidis serotype B (MenB) infection, the host immune system attempts to clear the invading pathogen by cell-mediated and humoral-mediated immunity. Cell-mediated clearance of a MenB infection involves phagocytosis by macrophages and the production of small molecule antimicrobials such as nitric oxide (NO). It is under these harsh conditions that the pathogen encounters environments with varying
concentrations of oxygen. To withstand the armament of host weaponry, *N. meningitidis* must switch from aerobic metabolism to anaerobic metabolism to survive within the host. This transition from aerobic to anaerobic respiration is mediated by FNR. Deleting *fnr* attenuates MenB virulence in mice, suggesting that successful MenB disease progression requires FNR and thus at least some members of its regulon (Bartolini et. al. 2006).

Like FNR, the closely related transcription factor, CRP (Cyclic AMP Receptor Protein, also known as Catabolite Activator Protein or CAP), controls the expression of greater than 150 genes in response to the availability of catabolite-repressing carbon sources, e.g. glucose. In the absence of glucose, the concentration of cyclic AMP (cAMP) rises (Lewis, M. 2005; Wolfe, 2008; Kumari et al, 2000). This second messenger binds to the CRP dimer, altering its conformation, and thus activating it for binding to its DNA site. Thus, the binding of cAMP permits CRP to participate in important protein-protein interactions (Reddy et al, 2009; Tan et al, 1991). These interactions are central to this dissertation.

**Mutations that Alter Transcription**

As described previously, mutations that alter transcription may occur in either the DNA sequence (*cis* elements) or the proteins that bind those sequences (*trans* factors). Mutations in the promoter can specifically affect the
binding and/or function of RNAP. Mutations in a DNA binding (DNAB) site can specifically influence the binding or function of an associated transcription factor. Finally, mutations can alter the overall structure of DNA and indirectly impact the binding of both RNAP and transcription factors (Ohniwa, 2006; Higgins, 2005; Dorman and Corcoran; 2008; El Sharoud, 2008; Hardy, 2005). Mutations in these cis elements generally affect only the transcription of the downstream gene or operon. For this reason, mutations in cis elements are preferable when studying the function of a global regulator at a particular promoter, such as the focus of this dissertation, the acs promoter.

In contrast to mutations in cis elements, mutations in trans factors may influence the transcription of multiple genes. This is particularly true if the trans factor functions globally. For example, mutations in RNAP that influence its ability to properly move through the steps of transcription are expected to exert an adverse effect on global transcription. Depending on the severity of the defect, such a mutation could be lethal (Hermsen, 2006; Hardy, 2005). Similarly, a mutation in a transcription factor (e.g. CRP) or a NAP (e.g. IHF) that exerts its influence globally would be expected to cause a global defect. Because such a mutation would be expected to impact a large number of genes directly and/or through its ability to alter the overall structure of the nucleoid, dissection of a given promoter is best performed with cis mutations that specifically alter the binding of a trans factor to the promoter region itself and not by deletion of the global trans factor itself.
In contrast, a mutation in a transcription factor (e.g., the Lac repressor), whose activity is limited to one or a small subset of promoters, would likely exert a more local effect (Fried and Crothers, 1981). For example, a mutation in a local negative regulator might be expected to result directly in constitutive expression of its regulated gene(s). Conversely, a mutant local activator that has lost its ability to stably bind to its DNA site within a given promoter would be expected to result in the loss of transcription from that promoter.

Finally, mutations in a transcription factor or a NAP could affect any of its multiple functions, including the binding of a ligand (e.g., cAMP) or covalent modification (e.g. phosphorylation or acetylation), oligomerization, DNA binding or bending, and/or the interaction of the transcription factor with RNAP (Martinez-Antonio, 2003). The severity of any particular mutation will depend upon the importance of the affected gene(s).

**Bacterial Transcription: the Textbook View**

The typical textbook shows bacterial transcription to be rather uncomplicated, especially in contrast to eukaryotic transcription (for example, see Lewin, 1996). The structure of bacterial DNA is often depicted as a simple linear “naked” structure, primarily unbound by proteins. However, this is not the biologically relevant conformation. The binding of proteins to DNA to form nucleoprotein complexes is essential to change the “linear” state of DNA to a
biologically relevant conformation is required for such processes as DNA replication, repair, recombination, and transcription (Higgins, 2005; Rice, 2008). The topological changes of DNA can take many conformations due to the pliability of its phosphate backbone and the base-stacking properties of its nucleosides (Sinden, 1994; Rice, 2008). DNA can be contorted into a variety of topologies by DNA binding (DNAB) proteins that package DNA by bending, twisting, and spooling the overwhelmingly long genomic DNA (1 m in eukaryotes and 1.7 mm in eubacteria) into the eukaryotic nucleus or the bacterial nucleoid. A similar process permits the packaging of viral DNA into the space-limiting nucleocapsid (Campos, 2009). The bending, twisting and spooling of DNA by DNAB proteins do not just mechanically package DNA into a small space but also protects the genome from nucleases and other destructive forces. Additionally, these protein-induced alternative states of DNA participate in the regulation of other DNA metabolic processes, including transcription (Lewin, 1996; Campos, 2009). Some of the proteins involved in DNA compaction and transcription regulation in bacteria and eukaryotes are quite similar; for example, HU of bacteria is a functional homolog of the histone H1 of eukaryotes. Most, however, differ quite dramatically. Yet, the general themes of DNA binding and of base pair stacking require the same four nucleotides: adenine, cytosine, guanine, and thymine.
Adenine, Cytosine, Guanine, Thymine, and Base Pair Stacking

All processes that involve DNA depend on its primary structure. This is particularly true for the mechanisms that regulate transcription. The fundamental impact on the structure and function of DNA involves the atoms that make up the bases, the base pair interactions, and the sugar-phosphate backbone. The structure, elasticity, polarity, and durability of DNA are due to the base pair stacking of the adenine (A), cytosine (C), guanine (G), and thymine (T) nucleotides in conjunction with the sugar-phosphate backbone. The overall stability of DNA, provided by the nitrogenous bases and the sugar-phosphate polymer, make it the ideal carrier of genetic information (Rice, 2008; Benham and Mielke, 2005).

Sitting on the outer portion of the DNA helix is the sugar-phosphate backbone. Due to the phosphate moiety that covalently links the 5’-OH of one nucleotide sugar to the 3’-OH of the next nucleotide sugar (the so-called phosphodiester bond), DNA has an overall negative charge (Sinden, 1994; Rice, 2008). The negative charge of the sugar-phosphate backbone renders it conducive for attracting positively charged DNA binding proteins, such as eukaryotic histones and bacterial NAPs. These proteins, when bound to DNA, introduce the structural bends required for most DNA metabolic events (Rice, 2008; Higgins 2005). Though the phosphate moiety is a rigid tetrahedron in nature, the binding of carbon to the oxygen atoms provides free rotation about the phosphorous-oxygen bond. The individual nucleotide is rigid, but in a polymer
the nucleotides have “indirect rotational flexibility” (Travers and Thompson, 2009). Therefore, the longer the DNA strand, the more overall is its flexibility. In addition to providing the backbone for base stacking and general elasticity, the alternating 5’ and 3’ sugar phosphate moieties give DNA directionality, or polarity (Sinden, 1994; Travers and Thompson, 2009).

The phosphate moiety is bound to a sugar, which is then bound to a nitrogenous base. Unlike the sugars and phosphates, the nitrogenous bases of the nucleotides are insoluble in water. Thus, the hydrophilic sugar-phosphate backbone of DNA is exposed to water, while the hydrophobic bases are hidden away (Sinden, 1994). The insoluble nature, or hydrophobicity, of the nitrogenous bases is the driving force behind the antiparallel DNA helix formation. The Watson-Crick base pairing between A:T, T:A, G:C, and C:G bases fit into the DNA helix much like stairs of a winding staircase. The twisting staircase structure is the consequence of the hydrophobic bases shying away from the aqueous environment in an effort to exclude water molecules from the DNA core (Sinden, 1994; Benham and Mielke, 2005). Since the base pairs cannot stack directly on top of one another the bases twist slightly. The slight rotation of the DNA ladder occurs because the length between phosphate moieties along the sugar-phosphate chain is twice as long as the thickness of the bases. Though the sugar-phosphate chain mediates the slight twist in base pair stacking, it also limits the degree of rotation by providing a scaffold that holds the bases in place (Sinden, 1994).
The nucleic acid sequence of DNA plays a role in determining its overall structure, specifically the effects of base pair stacking. Since the nitrogenous bases attempt to minimize contact with water, they adopt the propeller twist conformation (Rice, 2008; Sinden, 1994; Hassan and Calladine, 1996). The DNA helix structure is determined at the atomic level by several parameters of base stacking within them. Of these parameters, three are considered when analyzing B form DNA, i.e. flexibility, translation, and rotation (Sinden, 1994). The flexibility of DNA owes itself to the movement of the nucleotide bases and their stacking in the helix. Bases encapsulated within the helix have six variables of positioning or ‘degrees of freedom’. This positioning falls under two categories: translation and rotation. Translation is the movement of a base in any direction without rotation. Discussion about translation movement is commonly limited to the X, Y, and Z planes. In contrast, the rotation of a base involves twisting, rolling, and tilting about the long and short axes of the base. To describe these movements and their relation to a base pair step, the nitrogenous bases will be referred to as blocks for simple illustration. Also, the sugar-phosphate backbone will not be considered, but serves as the framework for positioning the bases (Hassan and Calladine, 1996) (see Figure 1).

Translation positioning is the movement of a block in an X, Y, and Z three-dimensional grid. The block may move vertically along the Y-axis. The base may move into and out of the 2-D plane along the Z-axis. And finally it may move laterally along the X-axis. The movement along the X-axis is described as the
sliding (slide = S) of one or both blocks in relation to one another. Since block movement in both the Y and Z directions is impeded through their covalent attachment to the sugar-phosphate backbone, the current discussion will involve only sliding (Rice, 2008; Bloomfield, 2000).

The three parameters of rotation used to describe the relative positions of bases are twist, roll, and tilt. Blocks may twist about their short axis – the axis running from top to bottom when the block is viewed from the broad side facing into the hydrophobic core of the helix. The effect of rolling describes the position of the side edge of the base, again when viewed on the broad side, in relation to the horizontal axis. When the broadside of the base moves above or below, there is either a positive or negative change in the angle relative to the horizontal plane. Tilting is an angular change of a base relative to the vertical plane. Though these parameters are real considerations when studying the position of base steps relative to one another, the real world situation introduces constraints to tilting. Therefore, further discussions of base pair step rotation will only be done so using twist (T) and roll (R) (Sinden, 1994; Hassan and Calladine, 1996; Bloomfield, 2000).

**Base Pair Steps**

So far, the translation and rotation parameters for block orientation have been described for a single base. However, in the DNA helix, the bases are
ordered one after another. Therefore, the change in position of one base affects the neighboring bases. To better understand these relationships, we will consider two base pairs (four bases in all or two pairs of complementary bases) aligned one above the other, or a base pair step. The sugar-phosphate backbone will not be considered in this discussion, but the parameters are applied to the confines of the double helix. In addition to the backbone limiting base movement, the effort to exclude water between bases or avoid the creation of a vacuum also restricts movement and degree of this motion. The movement of the base pair step only encompasses three of the six parameters mentioned above: the rotations roll (R) and twist (T), and the translation slide (S). When the R, S, and T values are zero (i.e., when there are no changes in the position of either base), then the bases overlap perfectly in the dinucleotide step as depicted in Figure 1A. Any permutations in the R, S, or T parameters, thus, affect the position of the neighboring base (Postow, 2005; Patel et al, 1983; Delmonte, 2003; Sponer et al, 1997).

**Figure 1. Dinucleotide Slide and Twist.** (A) Depicts the dinucleotide Slide movement in the direction of the long axis of each base relative to its immediate base neighbor. (B) The Propeller-Twist conformation of dinucleotide base pairs, placing one base out of plane relative to its base partner. (Modified from Hassan and Calladine, 1996).
The twist parameter is properly defined as the rotation of a base about the local twist axis running vertically through two nearby bases. Base twisting provides the helical ladder of DNA, contributing an estimated 34.3° per base for 10.5 bases, or one turn of B form DNA. Since B form DNA is not the only structure present within a cell, the number of bases per turn of DNA is considered to be 11±1 bases. However, with no changes in the R or S values, each base is estimated to “twist” by 36° (i.e., as you move one base at a time up the DNA helix, the consecutive bases will be off-angle by 36°) (Sinden, 1994, Bloomfield, 2000). The value of the twist parameter is not uniform for the entire molecule. For example, the T value for folded DNA will differ from circular DNA. The twist angle is sensitive to base stacking and other helical forces (Bauer, 1993; Bauer, 1995). Importantly for this dissertation, protein binding to DNA also affects the twisting of DNA in order to accommodate the protein/DNA interaction.

The second rotation important for base motion with the DNA helix is the roll parameter. Rolling results in the movement of a base that changes the position of the broad side of the base in relation to the horizontal plane, or long axis. The roll angles may vary from +20° to -10° in the B form DNA helix. Any rotation upward towards the minor groove is considered positive, as depicted in the right block of Figure 1B, and a rotation downward towards the major groove is negative (Sinden, 1994; Rice, 2008).

The single translation parameter discussed here is base slide. This is the lateral motion of a base along the X-axis relative to a neighboring base. If the
The upper pair of bases in a dinucleotide step moves further to the left of the vertical axis than does the lower base pair, then this movement is considered positive. Base pair movement towards the right of the vertical axis is given a negative value. The values for these types of base shifts may range from +3 Angstrom to -2 Angstrom. The sugar-phosphate chains limit the slide for a base pair in a dinucleotide step (Sinden, 1994; Rice, 2008).

**The Propeller Twist**

DNA base pairs rarely occupy the same plane. Therefore, the roll angles of the aromatic rings of the base pairs are slightly different. This gives them a conformation similar to an airplane's propeller-twist (Sinden, 1994; Hassan and Calladine, 1996). This propeller-twist base conformation is the most common base position in a DNA helix. The propeller-twist maximizes base stacking interactions with the neighboring base, but lengthens the DNA strand and weakens Watson-Crick hydrogen bonds. Because base stacking tends to assume the propeller-twist conformation, it has been suggested that the sturdiness of the DNA duplex can be attributed more to the stabilization energy provided by base stacking than by the energy provided by Watson-Crick hydrogen bonds (Hassan and Calladine, 1996; Mohan, 1992). Importantly for this dissertation, the propeller-twist base pair conformation plays a key role in protein/DNA interactions because it best accommodates the changes in base pair
stacking and the torsional strain introduced by DNA binding proteins without completely disrupting the helix.

The DNA molecule must be flexible and elastic enough to withstand these changes without unraveling (Sinden, 1994; Calladine, 2004). Take for instance the binding of the minor groove binding protein IHF. This NAP protrudes a proline into the DNA, thus disrupting the local base pair stacking. If the bases were unable to alter their stacking positions and depended solely on Watson-Crick hydrogen bonds, then this proline-mediated disruption would most likely induce strand separation, leading to inefficient protein-DNA interactions, and limiting the effects of IHF binding. Fortunately, base pair stacking is the major stabilizing force and the bases are able to twist, roll, and slide to adjust (Sinden, 1994; Ramakrishnan, 1993). Finally, it should be noted that translation and rotation values differ in regions with different sequences. For example, a region of DNA containing a stretch of A's will have a sequence-induced bend, whereas a unique sequence containing a mixture of all four bases will not (Patel, 1993; Haran, 2009).

Genome Compaction

The ability of DNA molecules to absorb topological changes are credited to the sugar-phosphate backbone and base stacking. Though the shape of DNA is easier to appreciate as a 2-D diagram or a linear ball and stick model, its true
nature within the cell is a folded, more compact, supercoiled structure that contains many regions of looped domains (secondary structure). The folding and unfolding of DNA in both bacteria and eukaryotes determines whether the transcriptional machinery gains access to a promoter region. Promoter access is governed not only by the topology of naked DNA, but also by the proteins that bind to DNA and alter its configuration (Higgins, 2004; Akyol, 2009). In eukaryotes, proteins called histones fold the genome, compacting it into a form that can fit within the nucleus. To achieve compaction, these large protein complexes spool DNA around themselves and interact with each other. By condensing DNA, histones physically hinder the ability of RNAP to bind to the promoter and transcription factors to bind to their respective DNA sites. By controlling the degree of compaction, therefore, histones function as gatekeepers that determine access to the promoter.

Bacteria also must package their genome into a limited space. To achieve this, they use NAPs, which possess many of the same properties of histones and function in a similar, but not identical, manner (Higgins, 2004; Sherrat, 2003). Below, I will describe the process by which histones and NAPs compact their respective genomes.
Compaction of the Eukaryotic Genome

The eukaryotic genome undergoes many different levels of compaction. The most basic level of compaction consists of the natural structural folds constituted by intrinsic base pair interactions, forming the B-DNA double helix, and by the intracellular density (also called molecular crowding – a physiological phenomenon important for all cellular processes). Base pair stacking, as a result of hydrophobic interactions between the nucleic acid bases, along with the phosphate backbone interaction with environmental water molecules, give rise to the B-DNA structure. Human DNA measures over 2 meter in length when all 46 chromosomes are stretched out from end to end. But the maximum cell size measures 100 μm, with the nucleus measuring 10 μm. To compact the 2 meter DNA into the cell nucleus, the genome undergoes a 10,000-fold reduction (20,000 fold is the most) in volume.

The eukaryotic cell overcomes the size and volume barrier to package an enormous amount of DNA into the confined space of a cell by first wrapping the thin strands of negatively charged DNA around basic proteins (Campos, 2009). These conserved basic DNAB proteins (histones) package the DNA into a tight structure described as “beads on a string”. This term describes the wrapping of the DNA strand around the histone octamers – a structure unique to eukaryotes – at 150 bp intervals. The histone proteins associate with DNA in a sequence-independent manner to form a nucleosome core particle (Fig. 2). Each octamer core, composed of histones H2A, H2B, H3 and H4, is tethered by 50 bp of
“linker” DNA, hence the “beads-on-a-string” description. This DNA structure is then further compacted by creating loops of the protein-DNA beads.

The continuous spooling of DNA around core histones decreases the total length by 7-fold, while increasing the diameter by 5-fold (the diameter of linear naked DNA is 2 nm, while the diameter of the ‘beads-on-a-string’ structure is 10 nm) (Figure 2) (Lodish, 2003). This primary nucleosome-DNA structure is further compacted through the interaction between the histone-associated protein H1 and the DNA strand that exits the H2A+H2B complex, creating a 30 nm hollowed-solenoid structure (Fig 2). Together, the 7-fold DNA compaction of the primary nucleosome-DNA structure and the 6-fold solenoid compaction compacts naked DNA by about 40-fold.

The final nucleoid structure places the 30 nm DNA-histone solenoid onto a chromosome scaffold consisting of protein lamins and topoisomerase. Lamins function as the structural foundation onto which the solenoid DNA lays. In contrast, topoisomerases associated with the chromatin relieve any supercoiling tension introduced by packaging and keeps the DNA solenoid intact during mitosis (Lodish, 2003). Thus, ultimately, the eukaryotic DNA undergoes a reduction in size of approximately 10,000-fold. This extreme packaging is necessary to fit into the double membrane nucleus, a structure that also limits protein/transcription factor access to the nucleoid.
Compaction of the Bacterial Nucleoid

Early TEM studies made clear the evident discrepancies between eukaryotes and bacteria (e.g. the size differential and the presence of a double membrane enclosure for the eukaryote’s genome and lack thereof for that of the bacterium) (Higgins, 2004). However, these studies also made clear the curious fact that the bacterial genome displayed some degree of order despite the absence of a dedicated organelle to enforce organization. Current images rebuffed earlier expectations of a widely dispersed bacterial genome with no sense of morphology (Eltsov and Zuber, 2006).

The bacterial genome (or nucleoid) is localized in a structurally undefined region within the cell. It is not encapsulated within a membrane, a defining hallmark of eukaryotic cells. The length of the bacterial genome measures about 1.7 mm, while the length of a typical bacterial cell, e.g. *E. coli*, is 3 μm long. This is a size discrepancy of approximately 1000-fold, which is similar to that of eukaryotes. As with compaction of the eukaryote genome, macromolecular crowding contributes to the maintenance of bacterial...
genome compaction (Higgins, 2004; Azam, 1999). Gentle lysis of *E. coli* cells releases the nucleoid, revealing strands of DNA 10 times the length of its intracellular size (Kavenoff and Bowen, 1976). More recent studies have shown that nucleoid-associated proteins (NAPs) mediate much of the rest of the compaction (Higgins, 2004; Swinger and Rice, 2007).

It is now clear that the bacterial genome is organized much like the eukaryotic genome and that it undergoes morphological changes similar to that of the eukaryote genome (Thanbichler et. al., 2005). Just as the density of the eukaryotic genome changes as its proceeds through the cell cycle (G1, S, G2, and M), the density of the bacterial genome correlates with the different phases of bacterial growth. For example, the bacterial nucleoid of actively growing and dividing cells (e.g. those in the exponential phase of growth) looks like a collection of loose 30 nm fibers, not unlike the 30 nm DNA-histone solenoid of the typical eukaryote. As cells slow their rates of growth and division (i.e., during the transition to stationary phase), the nucleoid begins to condense, forming a structure that resembles coral. In contrast, the nucleoid of relatively inert cells (e.g., those in late stationary phase) is an extremely compact electron dense structure that in many ways resembles the condensed mitotic chromosome (Kim et al., 2004; Travers and Muskehelishvili, 2005). Like the eukaryotic genome, the bacterial nucleoid forms these increasingly compact structures by the formation of topologically isolated DNA loops. These loops are negatively supercoiled, a fact of vital importance for DNA metabolic events, e.g. transcription. Contained
within these loops is stored free energy introduced by the torsional tension that results from supercoiling. Relaxation of this tension, by single-strand nicking or double-strand breaks, renders the nucleoid unable to perform its metabolic processes and as such becomes functionally inert. There are an estimated 500 supercoiled domains per chromosome, reducing the radius of gyration – defined as the root mean square distance of the supercoiled domains to the center of the chromosome. Obviously, the formation of these supercoiled loops also contributes to compaction of the nucleoid. Thus, looping not only maintains free energy but also facilitates the repair process by localizing nicked strands to one region (Postow et al., 2004) (Travers and Muskhelishvili, 2007) (Dorman, 1991) (Azam et al., 1999) (Azam et al., 1999) (Azam et al., 2000) (Takeyasu et al., 2004).

Compaction of bacterial nucleoid is mediated by NAPs, small basic proteins analogous to eukaryotic histone proteins. Unlike histones, the concentrations of NAPs change according to the phase of growth. During exponential phase, for example, FIS is the predominant NAP, while Dps is numerically dominant during stationary phase. The physical and chemical properties of FIS homodimer tends to facilitate the decondensation characteristic of the exponential phase nucleoid, while the properties of the Dps dodecamer facilitate compaction of the stationary phase nucleoid (Nystrom, 2004; Ohniwa, 2006; Roy, 2008; Jeong, 2008).
Like eukaryotic histones, NAPs do more than compact the genome. They also provide protection from nucleases and regulate DNA metabolism, including transcription. However, due to the limitation of bacterial cell volume, the genomic content must be limited to those genes that are required for survival. Reserving genomic space to accommodate the genes that encode the histone octamers, while also carrying genes that encode transcription factors for gene expression, would severely decrease metabolic efficiency. Therefore, the constitutive interaction between DNA structural proteins (like histones) and DNA is not seen in bacteria for two reasons: 1) these proteins play multiple roles in the life of the cell and 2) the dynamic interplay between DNA and NAPs provides the cell with the ability to adjust more rapidly to the ever-changing extracellular milieu. Though the use of DNA binding proteins is absolutely required for the compaction of DNA into the very limited intracellular space, maintaining compaction is a difficult task since the protein-DNA interaction is very dynamic.

Today, as technology provides us with the resolution necessary to clearly see within the tiny bacterial cell, the long-held belief that the bacterium is a simple organism without sophisticated spatial and temporal organization has begun to rapidly change. The bacterial genome clearly undergoes organized morphological changes similar to that of the eukaryote genome (Thanbichler et. al., 2005). From the decondensed DNA fibers of exponential phase to the electron dense nucleoid of late stationary phase, these changes are mediated by NAPs.
**Bacterial Nucleoid Associated Proteins**

*Escherichia coli* cells possess twelve nucleoid proteins. Five of these NAPs are thought to be the major protein components of the bacterial nucleoid: factor for inversion stimulation (Fis), histone-like nucleoid structuring protein (HNS), heat-stable nucleoid protein (HU), integration host factor (IHF), and stationary phase-expressed DNA binding protein from starved cells (Dps) (Azam and Ishihama, 1999; Ishihama, 1999).

These NAPs have structural similarities to some eukaryotic histones and/or transcription factors. For example, they possess DNA binding domains structurally similar to those found in histone proteins. Note, though, that they possess little or no sequence similarity. Unlike histones, some bacterial NAPs actively participate in multiple functions within the cell. For example, Fis and IHF both package the genome and function as pleiotropic regulators that bind specific DNA sequences. As such, they play specific roles in the regulation of DNA replication and segregation, recombinatorial events, and transcription. In contrast, the sequence-independent DNA binding protein, Dps, appears to function, like histones, strictly as a packaging protein (Nystrom, 2004; Ohniwa, 2006; Roy, 2008; Jeong, 2008; Higgins, 2004). Because NAPs perform such central and diverse roles, deletion of the gene that encodes a particular NAP almost certainly will have pleiotropic effects and the study of a physiological behavior in a cell that lacks a given NAP will be riddled with secondary effects.
Thus, to study the role of a NAP (e.g. IHF), it is preferable to simply disrupt binding to a specific DNA site.

Though the characteristic motions of the bases contribute flexibility and adaptability to the change in the overall structure of the DNA, short regions of DNA are actually intrinsically stiff. NAPs can be utilized to overcome this rigidity, permitting genome manipulation and transcription regulation. For example, the binding of DNA by NAPs introduces the bends needed to promote the binding of transcription factors and other proteins required for the regulation of transcription. Below, I will describe the major NAP most central to this dissertation, IHF.

**IHF**

Integration Host Factor (IHF) initially was identified as an *E. coli* protein required for λ phage site-specific recombination. Since its discovery, we now better understand the functional significance and importance of IHF to the life cycle of *E. coli*. The requirement of IHF for cell survival and normal biological processes is not limited to *E. coli*, as its homologues, identified in many other proteobacterial species such as *Salmonella typhimurium*, *Shigella flexinerii*, *Sinorhizobium meliloti*, and *Neisseria gonorrhoeae*, function in the same capacity (Hill et. al., 2002; Higgins, 2005). Universal identification of IHF, and its homologues, underscores the importance of protein mediated manipulation of DNA topology during the life cycle of the bacterium. As is the case with *E. coli*,
the concentration of IHF varies with the stage of cell development. In exponential phase the estimated concentration of IHF is 6,000 dimers per cell whereas approximately 27,000 dimers are present within stationary-phase cells (roughly five times the number of IHF dimers measured by Ditto et al., 1994). Not surprisingly the concentration of IHF dimers correlate with the status of DNA shape: loose, beads on a string during growth phase and tightly compacted under non-growth, nutrient limiting conditions. (Azam, 1999; Talukder, 1999).

IHF, a member of the DNA binding II (DNAB II) family of DNA binding proteins, is a heterodimeric protein that binds the minor groove of DNA (Fig. 3). Unlike the structurally similar homologous protein, HU, IHF binds DNA in a site-specific manner, although the 30-35 bp consensus sequence is degenerate (WATCARXXXXTTR where W = A or T; X = A, T, C OR G; R = A OR G). In addition to the 13 bp core sequence, IHF sites also contain an AT-rich region (depicted as WWWWWW) located 8 bases 5’ of the core 13 bases. However, sequence specific binding, or direct recognition, is not the only factor that contributes to IHF binding. Indirect recognition based on DNA structure (i.e. DNA bend) also provides a signal for IHF affinity (Bonnefoy and Rouviere-Yaniv, 1991; Thompson and Landy, 1988; Teter et al., 2000). Direct and indirect DNA recognition are inherent features of IHF and its natural interaction with DNA. This built-in versatility allows IHF to actively participate in transcription at specific promoters when cells are grown in nutrient-rich media while non-specifically binding DNA, and thus compacting it, when those nutrients are used up.
Furthermore, it has been postulated that high affinity sites, such as *yjbE* ORF and BIME *gyrB* sites, are nearly completely saturated with IHF dimers in exponential phase but the percentage of saturation changes to somewhere between 50 to 70% in stationary phase. The low-affinity IS1 site displayed near 70% saturation in stationary phase, while displaying only 25 to 30% saturation in exponential phase, emphasizing that cell cycle phase and the proteins associated with cellular adjustment to the nutrient status affect the function and binding affinity of IHF (Murtin e. al., 1998).

Two genes encode the IHF heterodimer: *himA*, the alpha subunit (Mr 11350) and *himD*, the beta (Mr 10650) subunit. The alpha and beta subunits are similar in sequence (25% homology), as well as structure. IHF possesses two distinct conformational features: 1) the alpha helical body and 2) two beta sheet arms, one from each subunit. The alpha helical domain mediates IHF dimerization, while the flexible beta sheet arms are responsible for DNA binding. IHF-DNA interaction results in DNA bend angles of up to 180°. The folding of DNA is induced through the intercalation of proline tips (Pro<sub>alpha</sub> and Pro<sub>b eta</sub>) located in the beta sheet arms between AT-bases of the minor groove of DNA. The disruption of the base pair stacking upon proline insertion introduces an 80° bend at each site. The DNA bend angle is further exacerbated through electrostatic interactions between the mostly positive charged body of IHF with the negative phosphate DNA backbone. For example, the residues Arg<sup>60</sup> and Arg<sup>63</sup> of the alpha subunit and Arg<sup>46</sup> of the beta subunit make direct contact with
specific bases whereas the Ser\textsuperscript{47} residue of the alpha subunit binds the dA/dT tract. Due to the variation of genomic IHF sequences, the DNA bend angles differ. However, the closer a site is to consensus the more severe the contortion. This severe conformational change increases the accessibility of the bound DNA to enzymes, such as recombinases and endonucleases (Higgins, 2004; Mengeritsky, 1993).

The major topological changes in DNA induced by IHF, and other DNA binding proteins, are required for all DNA metabolic events and are universal themes across all bacteria, archaea, and eukaryota. Specific to this dissertation is the mechanism of IHF-mediated negative regulation of transcription. For example, promoters \textit{syrR}, \textit{galR}, and \textit{Pu} are all dependent on IHF binding to fold DNA for either activation, by bringing distant activators into close proximity with the RNAP holoenzyme, or repression, by folding the promoter DNA into a complex that traps RNAP and/or transcription factors. The protein content bound to a given promoter region often dictates how IHF will affect transcription.

**NAPs and Transcription: Location, Location, Location**

Direct sequence recognition by proteins is an effective way to precisely direct specific protein binding and thus to regulate DNA processes. DNA binding sequences are found at origins of replication, DNA recombination sites, and promoters. The precise alignment of protein binding sites in relation to the origin
of replication or transcription start site determines the functional behavior of the protein within a given DNA region. However, the structural features of DNA also contribute to protein recognition.

The interactions between DNA and NAPs are subcategorized into two groups: indirect and direct. Some NAPs (e.g. HU) bind DNA nonspecifically. They do not recognize a specific sequence, but rather structural features of DNA such as the major or minor grooves, DNA backbone, hydration shell, flexibility, and intrinsic curvature. The inherent bend is mostly found to be A:T-rich DNA regions. This recognition of and binding to DNA based on its bent structure is termed indirect. Other NAPs (e.g. Fis and IHF) bind DNA specifically. They recognize and bind to stretches of DNA that contain their consensus sequence. Such NAPs are called direct binders (Higgins, 2004; Rice, 1997). However, the mechanisms of DNA binding by direct or indirect recognition are not mutually exclusive. For example, despite the identification of an IHF sequence motif, IHF also binds DNA non-specifically throughout the genome, an important function for DNA compaction. The presence of an A:T-rich region, perhaps inherently curved, 5' of the consensus site, may provide the initial recognition signal for IHF binding. As demonstrated by Hales et al. (1994), mutations introduced into the A:T-rich region of DNA increases the efficiency of IHF binding (Hales, 1994).

The direct binding NAP IHF can either enhance or inhibit transcription, depending upon the location of its binding site. Like many DNA binding proteins, IHF can compete directly with RNAP for promoter access and thus repress
transcription. This repression occurs if an IHF site is positioned such that it overlaps the promoter sequences that recruit RNAP. IHF also can activate transcription. Unlike many transcription factors, activation by IHF is not attributed to direct protein-protein interactions. Instead, IHF’s importance in activation appears to stem from its ability to deform DNA (Rice, 1997; de Lorenzo, 1991). For example, the severe bend mediated by IHF binding can bring two proteins bound to distant sites into close proximity. Such a scenario occurs at the Pu promoter of Pseudomonas putida, which transcribes genes that control the degradation of xylene and toluene. Activation of this promoter requires a transcription factor that binds an upstream activating sequence (UAS). It also requires that IHF bind a DNA site positioned between the UAS and the promoter. The binding of IHF performs two functions: it recruits RNAP to the promoter and it induces a sharp turn in the DNA. This severe bend brings the UAS and its associated activator into close proximity to the bound RNAP, thereby activating transcription. Eliminating IHF binding to the Pu promoter abrogates transcription at this promoter (Perez-Martin and de Lorenzo, 1997). Further evidence that IHF functions primarily to deform DNA comes from experiments in which an IHF site has been replaced by an HU site. At certain IHF-dependent promoters, researchers introduced base pair mutations within the IHF consensus site. These mutations ablated transcription. Promoter activity, however, was restored when those researchers substituted the IHF site with an A:T-rich sequence that conferred a bent structure to the DNA and thereby promoted the indirect binding
of HU (Swinger and Rice, 2003; Swinger and Rice, 2004; Swinger and Rice, 2007). The ability to induce dramatic changes in DNA conformation makes IHF a versatile protein that can be used in a variety of scenarios. In the specific case of transcription, as described above, the placement of the IHF binding site within a promoter region can dictate how IHF regulates transcription from that promoter. However, the binding of other proteins (e.g. CRP) in the vicinity of the promoter can modulate the function of IHF.

Interestingly, there are many examples of promoter DNA that contain A-tract (or T-tracts) sequences and, therefore, curved DNA regions. Curved DNA is most commonly found at two locations in core promoter regions: 1) immediately upstream of the -35 hexamer and 2) further upstream beyond the -35, such as -120 of the *ilvI*H operon. These regions of curvature have been shown to enhance RNAP binding to promoter DNA as well as recruit transcription factor binding. The presence of curved DNA on its own affects transcription activation, both positively and negatively, based solely on its structure. Likewise, transcription factors, such as IHF, remodel promoter DNA to regulate transcription. However, the effects of curved DNA and transcription factor binding are not mutually exclusive. The *ilvP*G2 promoter, which drives the expression of the *ilvGMEDA* operon, contains both intrinsically bent DNA at -50 and an IHF binding site at -90, both of which activate transcription (Pagel et. al., 1992).
CRP-Dependent Activation

Most mechanistic studies of bacterial transcription activation have focused on simple situations in which a single transcription factor can activate a given promoter. Most of these activators bind upstream of the transcription start site, make direct contacts with RNAP, and recruit it to the target promoter. These activators partition into two classes based on the location of their binding sites. Class I activators bind to upstream locations near position $-61$, $-71$, $-81$ or $-91$ relative to the transcription start site. They activate transcription by making direct contacts with the carboxy-terminal domain of the $\alpha$ subunit ($\alpha$-CTD) of RNAP, stabilizing the initial steps in the pathway of open complex formation. In contrast, Class II activators bind sites that overlap the $-35$ hexamer of the target promoter and, typically, activate transcription by contacting domain 4.2 of the $\sigma^{70}$ subunit of RNAP and the $\alpha$-NTD, the amino-terminal domain of $\alpha$. In some cases, Class II activators also make productive contacts with the $\alpha$-CTD.
which often binds upstream of the bound activator. Arguably, the best-studied simple transcription factor is CRP (cAMP receptor protein, also known as CAP). Over 200 genes are regulated by CRP. This transcription factor becomes activated to bind DNA only when bound by cAMP. The active cAMP–CRP complex (hereafter referred to as CRP) can function either as a Class I activator that contacts the α-CTD via a well-characterized surface determinant called activation region 1 (AR1) (Fig. 4A), or as a Class II activator that interacts with RNAP via AR1 and AR2, which contacts RNAP via the α-NTD (Fig. 4B). CRP also can function in tandem combinations of both, either Class II–Class I or Class I–Class I. These more complex arrangements are often called Class III (Busby and Ebright, 1999).

FNR is a paralog of CRP; in addition to AR1 and AR2, it possesses a third surface (AR3), which permits it to also contact region 4.2 of σ^70 (Fig. 4B). In CRP, this surface is masked; certain mutants remove the mask, exposing this non-native surface (Busby and Ebright, 1999; Lawson et al., 2004; Rhodius, 2000).
Acetyl-CoA Synthetase

Acetyl coenzyme A synthetase (Acs) catalyzes the conversion of acetate to acetyl-CoA through an enzyme-bound acetyladenylate (acetyl-AMP) intermediate (Fig. 5A) (Berg, 1956). This activity permits the use of acetate to obtain energy through the tricarboxylic acid cycle and biosynthetic subunits via the glyoxylate bypass (Cozzone, 1998). *E. coli* cells require Acs to scavenge acetate from the environment (Brown et al., 1977; Kumari et al., 1995). This delays entry into stationary phase and allows cells to compete successfully during periods of carbon starvation. Thus, the induction of Acs represents a survival response for cells entering a nutrient poor environment (reviewed by Wolfe, 2005).

![Figure 5. Acetyl-CoA Synthetase (Acs) and Transcription of its Gene (acs).](image-url)

(A) Acs catalyzes the activation of acetate to acetyl-CoA in an ATP- and HS-CoA-dependent manner. The by-products are pyrophosphate (PPI) and adenosine monophosphate (AMP). The resultant acetyl-CoA enters the TCA cycle and the glyoxylate shunt (GS). (B) *acs* transcription is induces as glucose ([glc]) levels approach depletion and acetate ([ace]) concentration begins to peak. Peak transcription occurs as cells enter stationary phase.
Like many other genes required for the use of secondary carbon sources, \acs\ is subject to catabolite repression (Kumari et al., 2000): cells growing on their preferred carbon source (e.g. glucose for \emph{E. coli}) inhibit \acs\ transcription (Kumari et al., 2000; Oh et al., 2002; Shin et al., 1997). However, the major byproduct of glucose metabolism is acetate, which cells excrete into their environment. Thus, the extracellular concentration of acetate rises as the concentration of glucose falls (Fig. 5B). Just prior to exhaustion of the glucose, cells induce \acs\ transcription. This event occurs during the transition from exponential growth to stationary phase (Kumari et al., 2000; Shin et al., 1997). As the culture enters stationary phase, transcription peaks and then decreases rapidly. A similar expression pattern happens during growth in tryptone broth, a mixture of amino acids. As \emph{E. coli} cells exhaust L-serine, their preferred amino acid, they induce \acs\ transcription (Pruss et al., 1994; Kumari et al., 2000; Browning et al., 2004). As with glucose, transcription peaks as the culture enters stationary phase (Kumari et al., 2000), decreases rapidly, and then slowly increases - this time, in parallel with the concentration of extracellular acetate (Wolfe, 2005).

The enzymatic process catalyzed by Acs improves fitness. When first grown on glucose or acetate as the sole carbon source and then permitted to starve, cells that retain Acs out-compete those that lack Acs. The implication of these results is that starving cells rely on the ability of Acs to scavenge for scarce acetate in the environment (Wolfe, unpublished data). However, if left unregulated, Acs can be dangerous. Overexpression of Acs is lethal (Kumari et
al., 2000). This is likely due to its use of CoA as a substrate. The unregulated activity of Acs would deplete the cell’s store of CoA, which is maintained in limiting quantities. Thus, cells regulate the activity of Acs (via N-lysine acetylation) and the transcription of acs, the gene that encodes it (reviewed by Wolfe, 2005).

Regulation of acs transcription occurs primarily at the level of initiation (Kumari et al., 2000). As befits a critical survival gene, this regulation is quite complex. What follows is a description of the machinery that controls the initiation of acs transcription.
acs Operon and Promoter Architecture

acs is the first gene in an operon that includes yjcH, a hypothetical gene, and actP (formerly known as yjcG), which encodes an acetate permease (Fig. 6A) (Gimenenz et al., 2003) whose physiological purpose remains unknown. No evidence exists for internal promoters; thus, transcription of the acs operon
apparently initiates only from the regulatory region of acs, located upstream (5') of the acs open reading frame.

The regulatory region of acs includes three σ70-dependent promoters: the proximal major promoter acsP2; a minor promoter called acsP2A located 18-bp upstream of acsP2, and the distal minor promoter acsP1 located some 200-bp upstream of acsP2 and acsP2A (Fig.6B) (Beatty et al., 2003; Browning et al., 2002; Kumari et al., 2000a, Kumari et al., 2000b).

**Activation of acsP2**

In vitro, RNAP alone does not efficiently transcribe the proximal acsP2. It does, however, bind and melt an extensive region of DNA that includes both acsP2 and acsP2A (Beatty et al., 2003). For efficient transcription, it requires CRP. Because acs transcription depends on CRP, it is subject to catabolite repression (Beatty et al., 2003; Kumari et al., 2000). This well-studied regulatory mechanism results from the normal action of the phosphoenolpyruvate (PEP)-dependent sugar phosphotransferase system (PTS). Using PEP as the phosphoryl donor, the PTS simultaneously phosphorylates certain sugars (for example, glucose) and transports them across the cytoplasmic membrane into the cytosol, where the phosphorylated sugars enter glycolysis. As the sugars become scarce, the PTS instead activates adenylate cyclase, which synthesizes
cAMP. The cAMP can then bind the CRP homodimer, altering its conformation so that it can bind to its DNA sites (reviewed by Deutscher et al., 2006).

CRP can activate the major promoter acsP2 using a variant of the Class III mechanism (Fig. 7C). At this promoter, two CRP homodimers bind in tandem at DNA sites centered at Class I positions: the higher affinity and proximal site (CRP I) is centered at position -69.5, while the lower affinity and distal site (CRP II) is centered at -122.5. The binding of each CRP homodimer is predicted to incur a bend in the DNA up to $45^\circ$. This bending helps bring each homodimer closer to the promoter and thus to the $\alpha$-CTD. Activation absolutely requires that one dimer bind the proximal CRP I. To achieve maximal transcription, however, a second dimer also must bind to the distal CRP II (Beatty et al., 2003).
Most often CRP functions from Class I sites by stabilizing the closed complex (CC). Using the surface determinant AR1, the CRP dimer contacts the 287 surface determinant of the α-CTD (Fig. 4A). This interaction stabilizes the CC, permitting the subsequent formation of a open complex (OC), the next intermediate in the transcription initiation pathway (Browning and Busby, 2004; Busby and Ebright, 1999). At acsP2, the mechanism is a bit more complex. Although CRP I, AR1 and determinant 287 help to stabilize the TCC, their contribution cannot be limited to this role: RNAP alone can bind both acsP2A and acsP2 and form TOCs, albeit in inactive forms (Fig. 7A, B) (Beatty et al., 2003).

Modulation of CRP-Dependent Activation by Nucleoid Proteins

Of the twelve nucleoid proteins of *E. coli*, both FIS and IHF can inhibit transcription of acs. Within the acs promoter region, FIS and IHF can each bind to three sites (Fig. 6B). Furthermore, all six sites can be occupied simultaneously. Finally, both nucleoid proteins function independently to antagonize CRP-dependent activation of acsP2 (Browning et al., 2004).

Negative Regulation by FIS

FIS can inhibit CRP-dependent acsP2 transcription directly by a mechanism called anti-activation. This mechanism relies on the ability of FIS to bind to two higher affinity sites: *FIS II* and *FIS III* (Fig. 8). *FIS II* (centered at -98)
lies between CRP II (-122.5) and CRP I (-69.5), while FIS III (-59) overlaps CRP I. Competitive DNase I footprint and electrophoretic mobility shift analyses, and in vitro transcription assays indicate that FIS can displace CRP from both its sites. In vivo, a mutation in FIS II that diminishes its affinity for FIS by more than ten-fold increases acs transcription two- to three-fold during growth in tryptone broth. A similar increase results from a mutation that favors the binding of CRP over that of FIS to their overlapping sites (CRP I and FIS III). Thus, the competition between FIS and CRP for binding to their overlapping and tandem sites helps to keep acsP2 transcription low (Browning et al., 2004). Because FIS levels rise dramatically during outgrowth in rich medium from stationary phase and then progressively diminish throughout exponential growth (Azam et al., 1999; Ball et al., 1992). FIS appears to be responsible for maintaining acs transcription at basal levels during rapid growth conditions when the activity of acs is unnecessary.

**Negative Regulation by IHF**

The concentration of IHF in cells harvested from early exponential growth has been estimated at about 0.7 nM (Murtin et al., 1998). This concentration increases progressively throughout growth until, during stationary phase, it becomes the second most abundant nucleoid protein (Azam et al., 1999; Ishihama, 1999). The binding of IHF to its specific sites causes the DNA to bend up to 180 degrees and to wrap around the protein (Lynch et al., 2003; Rice et al.,
1996; Travers, 1997). IHF affects *acs* transcription by binding to three high affinity sites (*IHF I - III*) located between positions -140 and -240 relative to the +1 of *acsP2* (Fig. 6B) (Browning et al., 2004). Of particular importance to this dissertation is the observation that the 31-bp sequence *IHF III* causes a severe inhibition of CRP-dependent activation.

![Figure 8. FIS-Dependent Anti-Activation of CRP-Dependent *acs* Transcription.](image)

FIS out-competes CRP binding to their overlapping DNA sites. Because CRP cannot bind, transcription does not occur.

**Concluding Remarks**

The *acs* promoter exemplifies the complexity of bacterial promoters. At first glance, transcription from the *acs* promoter appears to be simple enough: *acsP2* is a weak CRP-dependent promoter activated by the tandem Class I variant of the Class III mechanism. Deeper study, however, shows that this promoter is subject to considerably more sophisticated regulatory mechanisms
that involve intricate competition between multiple transcription factors and multiple NAPs.

It was the goal of my dissertation to understand how \textit{IHF III} inhibits CRP-dependent \textit{acs} transcription. The initial studies performed in collaboration with Bianca Sclavi led to the recognition that occupancy of \textit{IHF III} causes the formation of a stalled nucleoprotein complex comprised of RNAP, two CRP homodimers, and IHF. Further investigation led to the discovery that \textit{IHF III} is a composite sequence comprised of two overlapping DNA sites for IHF. One, orientated towards the promoter, sits on one face of the DNA helix orient. The other, orientated away from the promoter, sits on the other. Together, they control CRP-dependent \textit{acs} transcription.
CHAPTER TWO
MATERIALS AND METHODS

Chemicals and Biological Reagents

Chemical reagents were purchased from Fisher Scientific (Pittsburgh, PA) and Sigma Chemical Company (St. Louis, MO). β-galactosidase assay reagents were purchased from Pierce Biochemicals (Rockford, IL). Restriction endonucleases and modifying enzymes were purchased from Promega Corp. (Madison, WI), New England Biolabs (Beverly, MA) or Gibco BRL (Gaithersburg, MD). Primers were purchased from Integrated DNA technologies (Iowa). Media components were obtained from Difco.

Bacterial strains, plasmids, bacteriophage, and promoter fragments

All bacterial strains, plasmids, bacteriophage, and promoter fragments used in this study are listed in Tables 1-4. Derivatives of the *E. coli* K-12 strain, AJW678, were used for all experiments.
By convention, all locations within the *nrfA - acs* intergenic region (Fig. 9A) are numbered relative to the transcription start site of *acsP2* (+1) with upstream and downstream locations prefixed ‘-’ and ‘+’, respectively.

The constructions of the promoter fragments FL, 1xIHF, and 0xIHF (*pacs444, pacs236, and pacs205* respectively) were described previously (Beatty *et al.*, 2003; Browning *et al.*, 2004). Promoter fragment 1xCRP (*pacs155*) was constructed by digestion of pCB63 (Browning *et al.*, 2004), followed by re-ligation to form a promoter fragment that lacked sequences upstream of *CRP II*. This construct was digested with *EcoRI* and *BamHI*, subcloned into the *lacZ* fusion vector pRS415 to generate pCB91, and recombined with the *lacZ* fusion vector *λRS88*, using the method of Simons et al. (Simons *et al.*, 1987). The constructions of the mutations G74C and G126C, which disrupt the binding of CRP to *CRP I* and *CRP II*, respectively, have been described previously (Beatty *et al.*, 2003). The double mutation C150G/C153G, which disrupts binding of IHF to *IHF III*, was generated using the Gene Editor site-directed mutagenesis kit (Promega, Madison WI). The primer IHFIII (CTGCATGTTCGTGAAG) was used to mutagenize the template pCB26 (Beatty *et al.*, 2003), generating pCB92. Successive promoter fragments were generated using either pCB26 (the full-length promoter cloned into pGEMT) as a template or existing promoter mutations cloned into pGEMT (e.g. pCB26-F2m; *Fis II* mutation). For example, pCB92 was used as the template to generate 1xIHF-I3m using this plasmid as a template the primer pairs P2 (GAATTCCATAACTGCATGTTTC) and 2926. After
amplification, all promoter fragments were cloned into TOPO2.1 vector (Invitrogen) and sequenced to identify any present mutations. All verified promoters were digested with EcoRI and BamHI and cloned into pRS415. Following this cloning step, all promoters were sequenced once more and then recombined with λRS88. A similar strategy was used to construct 2xIHF-I3m variants carrying the G74C and G126C mutations. Single lysogens of strain AJW678 (Δlac) were constructed and verified, as described previously (Simons et al., 1987). Primers specific for the λ attachment sites were used to determine whether the resultant lysogens were in single or multiple copies (as described by Sowell et al., 1994):

λ.attB (GAGGTACCAGCGCGGTTTGATC)
λ.attP (TTTAATATATTGATATTTATATCATTTTACGTCTTCGTTTC)
λ.int (ACTCGTCGCGAACCCTGCTTC)

The Δcrp::Km allele was transferred from strain CB369 by generalized transduction with P1kc (Silhavy et al., 1984).

For in vitro studies, we used the plasmids pSRacsFL and pSR1xIHF. The construction of pSRacsFL was described previously (Beatty et al., 2003). pSR1xIHF was constructed similarly. pSR1xIHF was PCR amplified using the primer pair P2 (GAATTCCATAACTGCA-TGTCC) and 2926HindIII (AAGCTTGTTGGTCTGCG-ATGTTG). This facilitated subcloning of the
fragment with an upstream EcoRI linker and a downstream HindIII linker into the
in vitro transcription vector pSR (Kolb et al., 1995).

All PCR amplifications were performed in a 50 µl reaction volume. The reactions contained 1X PCR buffer (Promega), 3 mM MgCl₂, 20 nM each primer, 0.2 mM dNTPs and 1 U Taq polymerase. Reactions were subjected to 1 cycle of 95°C for 5 minutes, then 30 cycles 95°C for denaturing, 68°C for annealing, and 72°C for extension. Site-directed mutations were generated using the Gene Editor site-directed mutagenesis kit (Promega), following the protocol provided. The resultant PCR products were subcloned into pGEM-t and the successful incorporation of mutations was verified by sequencing.

For cloning and transformations, plasmids were prepared using the Wizard miniprep kit (Promega), following the protocol provided. All restriction digestions and ligations were performed according to standard methods, as described (Sambrook et al., 1989).

**Media and growth conditions**

For strain construction procedures, cells were grown in Luria Broth (LB; 1% (wt/vol) tryptone, 0.5% (wt/vol) yeast extract and 0.5% (wt/vol) sodium chloride). For β-galactosidase assays, cells were grown at 37°C in tryptone broth (TB; 1% (wt/vol) tryptone and 0.5% (wt/vol) sodium chloride). Media were
supplemented with antibiotics or 5-bromo-D-chloro-galactopyranoside (X-gal), as needed.

Promoter activity assays

β-galactosidase activity was determined quantitatively using the Y-PER β-galactosidase assay kit (Pierce Biochemical). Cells were harvested at approximately one-hour intervals over a growth curve. The activities were plotted against optical density to standardize for subtle differences in growth rates. From these plots, the peak activity was identified and standardized to that of the AR2 mutant for each promoter construct tested. The values expressed are the mean and the standard error of the mean of three independent measurements. Each experiment was performed in triplicate and assayed three to five times to determine the reproducibility of the data. Some reporter truncations were reconstructed to ensure reproducibility from beginning to end.

Protein preparation

Purified RNA polymerase was purchased from Epicentre Technologies. Purified WT CRP was a kind gift from Annie Kolb (Institute Pasteur, Paris, France). The plasmids that encode the mutant proteins (CRP KE101 and CRP HL159) and the crp mutant cells were a kind gift of Nigel Savery (University of Bristol, UK). The mutant CRP proteins (CRP KE101 and CRP HL159) were
prepared, as described previously (Ghosaini et al., 1988). Purified IHF protein was prepared by the method of Nash et al. (Nash et al., 1987).

**DNase I footprint analysis**

Footprint analysis was a modification of the protocol described by Gralla (Gralla, 1985). 0.1 nM supercoiled plasmid DNA was incubated with varying concentrations of IHF, CRP, and/or RNAP in 10 µl total volume at 37°C in footprinting buffer (20mM Hepes pH 8, 5 mM MgCl₂, 50 mM potassium glutamate, 1 mM DTT, 500 µg/ml deacetylated BSA). After 20 minutes, 1 µl of 5 µg/ml DNase I (Worthington) was added. After 20 seconds, the reaction was stopped by adding 10 µl of stop solution (2X Taq Platinum buffer (Invitrogen), 2.5 nM labeled oligonucleotide, 2 µg/ml ctDNA, 240 µM dNTPs). One unit of Platinum Taq DNA polymerase (Invitrogen) was added. Primer extension was carried out in a Robocycler PCR machine (Stratagene) for 30 cycles. Line graphs of band intensities were obtained using ImageQuant Software (Molecular Dynamics) and standardized by the band at -35, and, when possible, bands upstream and downstream of the footprint.
Potassium permanganate footprint analysis

1 nM supercoiled plasmid DNA was incubated with varying concentrations of IHF, CRP, and/or RNA polymerase in 10 µl total volume at 37°C in footprinting buffer (20mM Hepes pH 8, 5 mM MgCl₂, 50 mM potassium glutamate, 1 mM DTT, 500 µg/ml deacetylated BSA). After 20 minutes, 1 µl of a frozen aliquot of 200 mM potassium permanganate was added. After 20 seconds, the reaction was stopped by the addition of 50 µl of stop solution (4% β-mercaptoethanol, 1.6 M sodium acetate, 4 µg/ml ctDNA). The samples were then cleaned by a phenol/chloroform extraction followed by ethanol precipitation. They were then resuspended in 10 µl of water to which 10 µl of primer extension solution was added (2X Taq Platinum buffer (Invitrogen), 2.5 nM labeled oligonucleotide, 5 mM MgCl₂, 240 µM dNTPs, 1 unit of Platinum Taq (Invitrogen)). Primer extension was carried out in a Robocycler PCR machine (Stratagene) for 20 cycles. Bar graphs were obtained using ImageQuant software to quantify the band intensities, Excel (Microsoft) to calculate the ratio of P2A, P2UP, or P2DOWN relative to total intensity, and Origin (Originlab) to produce the plots.
Table 1. Strains, plasmids and phage not listed in Tables 2-4.

<table>
<thead>
<tr>
<th>Strain, plasmid, phage</th>
<th>Relevant characteristics</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AJW678</td>
<td><em>thi-1 thr-1(Am) leuB6 metF159(Am) rpsL136 ΔlacX74</em></td>
<td>(Kumari et al., 2000a)</td>
</tr>
<tr>
<td>CB369</td>
<td>MG1655 Δcrp::Km</td>
<td>C. Bausch</td>
</tr>
<tr>
<td>P90C</td>
<td><em>ara Δ(pro-lac) thi</em></td>
<td>(Simons et al., 1987)</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pRS415</td>
<td><em>bla</em> <em>lacZ</em>+, transcriptional (operon) fusion vector</td>
<td>(Simons et al., 1987)</td>
</tr>
<tr>
<td>pCB26</td>
<td>pGEM-t carrying -411 <em>acs</em> through the <em>acs</em> ORF, template for site-directed mutants</td>
<td>(Beatty et al., 2003)</td>
</tr>
<tr>
<td>pDCRP</td>
<td>pBR322 derivative carrying wild-type <em>crp</em> gene (1^+2^+3^o)</td>
<td>(Rhodius et al., 1997)</td>
</tr>
<tr>
<td>pDCRP/H159L</td>
<td>pDCRP derivative carrying <em>crpH159L</em> allele (1^2^+3^o)</td>
<td>(Rhodius et al., 1997)</td>
</tr>
<tr>
<td>pDCRP/K101E</td>
<td>pDCRP derivative carrying <em>crpK101E</em> allele (1^+2^+3^o)</td>
<td>(Rhodius et al., 1997)</td>
</tr>
<tr>
<td>pDCRP/K52N</td>
<td>pDCRP derivative carrying <em>crpK52N</em> allele (1^+2^+3^o)</td>
<td>(Rhodius et al., 1997)</td>
</tr>
<tr>
<td>pDCRP/H159L/K101E</td>
<td>pDCRP derivative carrying <em>crpH159L K101E</em> allele (1^2^+3^o)</td>
<td>(Rhodius et al., 1997)</td>
</tr>
<tr>
<td>Vector/Phage</td>
<td>Description</td>
<td>Source/Reference</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>pDCRP/H159L/K101E/K52N</td>
<td>pDCRP derivative carrying crpH159L K101E K52N allele (1·2·3+)</td>
<td>(Rhodius and Busby, 2000)</td>
</tr>
<tr>
<td>pDCRP/H159L/K52N</td>
<td>pDCRP derivative carrying crpH159L K52N allele (1·2·3+)</td>
<td>S. J. W. Busby</td>
</tr>
<tr>
<td>pDCRP/K101E/K52N</td>
<td>pDCRP derivative carrying crpK101E K52N allele (1·2·3+)</td>
<td>S. J. W. Busby</td>
</tr>
<tr>
<td>pGEM-t</td>
<td>pUC19-derived TA cloning vector</td>
<td>Promega</td>
</tr>
<tr>
<td>pSR</td>
<td>in vitro transcription vector used for all in vitro work</td>
<td>(Kolb et al., 1995)</td>
</tr>
<tr>
<td>pRW50</td>
<td>promoterless lacZYA</td>
<td>(Lodge et al., 1992)</td>
</tr>
<tr>
<td><strong>Bacteriophages</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>λ.RS88</td>
<td>bla= lacZ imm\textsuperscript{434} ind, transcriptional (operon) fusion vector</td>
<td>(Simons et al., 1987)</td>
</tr>
</tbody>
</table>
### Table 2. Promoter fragments used in Chapter 3, with corresponding plasmid, phage, and strains.

<table>
<thead>
<tr>
<th>Promoter</th>
<th>Description</th>
<th>Plasmid</th>
<th>Phage</th>
<th>WT strain</th>
<th>delta crp strain</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pacs444</td>
<td>AatII-HindIII fragment carrying acs sequences from position -379 to +65</td>
<td>pCB33</td>
<td>λCB12</td>
<td>AJW1941</td>
<td>AJW2183</td>
<td>(Beatty et al., 2003)</td>
</tr>
<tr>
<td>pacs236</td>
<td>PCR product carrying acs sequences from position -171 to +65</td>
<td>pCB60</td>
<td>λCB22</td>
<td>AJW2026</td>
<td>AJW2086</td>
<td>(Browning et al., 2004)</td>
</tr>
<tr>
<td>pacs205</td>
<td>PCR product carrying acs sequences from position -140 to +65</td>
<td>pCB61</td>
<td>λCB42</td>
<td>AJW2179</td>
<td>AJW2198</td>
<td>(Browning et al., 2004)</td>
</tr>
<tr>
<td>pacs155</td>
<td>PCR product carrying acs sequences from position -90 to +65</td>
<td>pCB91</td>
<td>λCB39</td>
<td>AJW2203</td>
<td>AJW2199</td>
<td>This study</td>
</tr>
<tr>
<td>pacs236</td>
<td>CRP I mutant of pacs236</td>
<td>pCB71</td>
<td>λCB29</td>
<td>AJW2080</td>
<td>AJW2192</td>
<td>This study</td>
</tr>
<tr>
<td>G74C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pacs236</td>
<td>CRP II mutant of pacs236</td>
<td>pCB96</td>
<td>λCB45</td>
<td>AJW2202</td>
<td>AJW2812</td>
<td>This study</td>
</tr>
<tr>
<td>G126C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pacs236</td>
<td>IHF III mutant of pacs236</td>
<td>pCB96</td>
<td>λCB44</td>
<td>AJW2813</td>
<td>AJW2814</td>
<td>This study</td>
</tr>
<tr>
<td>C150G/C152G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3. Promoter fragments not listed in Tables 1 or 2, used in Chapter 4.

<table>
<thead>
<tr>
<th>Promoter</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>pacs1xIHF</em> A164T</td>
<td>predicted to affect both IHF IIIa and IIIb</td>
<td>This study</td>
</tr>
<tr>
<td><em>pacs1xIHF</em> T158G</td>
<td>predicted to affect IHF IIIb</td>
<td>This study</td>
</tr>
<tr>
<td><em>pacs1xIHF</em> A157G</td>
<td>predicted to affect IHF IIIb</td>
<td>This study</td>
</tr>
<tr>
<td><em>pacs1xIHF</em> A156G</td>
<td>predicted to affect IHF IIIb</td>
<td>This study</td>
</tr>
<tr>
<td><em>pacs1xIHF</em> C150A</td>
<td>predicted to affect IHF IIIa</td>
<td>This study</td>
</tr>
<tr>
<td><em>pacs1xIHF</em> A149C</td>
<td>predicted to affect IHF IIIa</td>
<td>This study</td>
</tr>
<tr>
<td><em>pacs1xIHF</em> A148C</td>
<td>predicted to affect IHF IIIa</td>
<td>This study</td>
</tr>
<tr>
<td><em>pacs1xIHF</em> T142A</td>
<td>predicted to affect both IHF IIIa and IIIb</td>
<td>This study</td>
</tr>
<tr>
<td><em>pacs1xIHF</em> A148C</td>
<td>predicted to affect IHF IIIa</td>
<td>This study</td>
</tr>
<tr>
<td><em>pacs1xIHF-VF</em></td>
<td>swaps <em>IHF IIIa</em> with <em>IHF IIIb</em></td>
<td>This study</td>
</tr>
<tr>
<td>Promoter</td>
<td>Relevant characteristics</td>
<td>Source or reference</td>
</tr>
<tr>
<td>------------------------</td>
<td>-------------------------------------------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>pacs1xIHF+5</td>
<td>complete palindrome sequence</td>
<td>This study</td>
</tr>
<tr>
<td>pacs1xIHF+5S</td>
<td>upstream 5 bases (palindrome half-site) scrambled</td>
<td>This study</td>
</tr>
<tr>
<td>pacs1xIHF C152G C150G</td>
<td>IHF III weakened</td>
<td>This study</td>
</tr>
<tr>
<td>pacs1xIHF C152A C153A</td>
<td>IHF III strengthened</td>
<td>This study</td>
</tr>
<tr>
<td>pacs2xIHF</td>
<td>contains IHF II and IHF III</td>
<td>This study</td>
</tr>
<tr>
<td>pacs2xIHF-A170C T169C</td>
<td>contains IHF II and IHF III; IHF II weakened</td>
<td>This study</td>
</tr>
<tr>
<td>pacs2xIHF-A170T</td>
<td>contains IHF II and IHF III; IHF II strengthened</td>
<td>This study</td>
</tr>
<tr>
<td>pacs2xIHF-G134C</td>
<td>contains IHF II and IHF III; CRP II mutated</td>
<td>This study</td>
</tr>
<tr>
<td>pacs2xIHF-C177T</td>
<td>contains IHF II and IHF III; IHF II mutated</td>
<td>This study</td>
</tr>
<tr>
<td>pacs2xIHF C152G C150G</td>
<td>contains IHF II and IHF III; IHF III mutated</td>
<td>This study</td>
</tr>
</tbody>
</table>
CHAPTER THREE

THE MULTIPLE ROLES OF CRP AT THE COMPLEX ACS
PROMOTER DEPEND ON ACTIVATION REGION 2 AND IHF

Introduction

Previously, the Wolfe lab reported that *E. coli* cells control Acs activity primarily at the level of transcription initiation (Kumari *et al.*, 2000a). The regulatory region of *acs* includes three $\sigma^{70}$-dependent promoters: the proximal major promoter *acs*P2; a minor promoter called *acs*P2A located 18-bp upstream of *acs*P2, and the distal minor promoter *acs*P1 (Figs. 6 and 9A) (Beatty *et al.*, 2003; Browning *et al.*, 2002; Kumari *et al.*, 2000a; Kumari *et al.*, 2000b). Transcription is induced during mid-exponential phase and peaks as cells enter stationary phase, whereupon it diminishes (Kumari *et al.*, 2000a) (Fig. 5B). While multiple factors influence transcription (Browning *et al.*, 2002; Kumari *et al.*, 2000a), CRP appears to function directly as the most critical transcription factor (Beatty *et al.*, 2003; Kumari *et al.*, 2000a).

CRP activates the major promoter *acs*P2 using a variant of the Class III mechanism. Two CRP homodimers bind in tandem at DNA sites centered at the
Class I positions -69.5 and -122.5 (Figs.7C and 9A). Using the surface determinant AR1, each dimer contacts the $\alpha$-CTD, recruiting RNAP to $acsP2$ and, thus, facilitating activation of $acs$ transcription. Activation absolutely requires that one dimer binds $CRP I$, the proximal, higher affinity site. To achieve maximal transcription, however, a second dimer also must bind to $CRP II$, the distal, lower affinity site (Beatty et al., 2003).

Similar to transcription regulation in eukaryotes, bacteria transcription is also modulated by histone-like proteins that fold the genome into the highly compacted, highly organized nucleoid (Kim et al., 2004). In both kingdoms, changing the DNA architecture provides a structural barrier that limits the access of transcription factors and RNAP. However, unlike eukaryotes, the interaction between DNA and bacterial histone-like proteins is dynamic – the same nucleoid protein may not remain stationary to the same DNA region (The Bacterial Chromosome, Higgins, 2005). *E. coli* possesses twelve nucleoid proteins whose functions are as diverse as their structures (Azam et al., 1999; Ishihama, 1999). Of these twelve proteins, Fis and IHF inhibit transcription of $acs$ (Browning et al., 2004).

Within the $acs$ promoter region, Fis and IHF each bind to three sites (Fig. 6B). Furthermore, each nucleoid protein independently antagonizes CRP-dependent activation of $acsP2$. The mechanism of this antagonism is largely unknown save for the direct competition between Fis and CRP at their
Figure 9. The organization of the \textit{acs} promoter. (A) This schematic shows the \textit{acs} promoter region from positions -379 to +65 relative to the \textit{acs}P2 transcription start site (+1). It shows the locations of each promoter, each CRP and IHF binding site and the extent of the promoter fragments. The bent arrows indicate the approximate location of each promoter and its direction of transcription. Inverted arrows indicate the locations of CRP I and CRP II, as determined by similarity to consensus. Hatched boxes designate the locations of IHF I-III, also determined by similarity to consensus. (B) Binding site mutations used in this study. The core consensus sequence for each binding is underlined and the substituted nucleotide(s) positioned above the mutated one. (C) \textit{\beta}-galactosidase activity from promoter fragments \textit{pacs444} (FL), \textit{pacs236} (1xIHF), \textit{pacs205} (0xIHF), \textit{pacs155} (1xCRP), or \textit{pacs236} (1xIHF) variants defective for CRP I, IHF III, or CRP II fused to \textit{lacZYA}. The wild type strain AJW678 was lysogenized with the hybrid phages \textit{\lambda}CB12, \textit{\lambda}CB22, \textit{\lambda}CB42, \textit{\lambda}CB39, \textit{\lambda}CB29, \textit{\lambda}CB44, or \textit{\lambda}CB45, respectively (Table 2). The resultant strains were grown in TB, samples harvested at one-hour intervals, the \textit{\beta}-galactosidase activity determined, and the peak activity expressed as a percentage of wild-type. Each value represents the mean ± SEM of at least three independent measurements. The percentage activities in relation to \textit{pacs444} (FL) are noted within each histogram.
overlapping sites, *Fis III* and *CRP I* (Fig. 8) (Browning *et al.*, 2004). Because its levels peak dramatically during outgrowth from stationary phase and then progressively diminish throughout exponential growth (Azam *et al.*, 1999; Ball *et al.*, 1992), *Fis* likely maintains *acs* transcription at basal levels under rapid growth conditions. The role of *IHF* has remained less clear: *in vitro* it is known to bind simultaneously with *CRP* and to cause inhibition of *acsP2* transcription; *in vivo*, however, the upstream sequence, which contains the two distal *IHF* sites, the minor promoter *acsP1*, and the divergent promoter *Pnrf*, has a positive effect on transcription (Browning *et al.*, 2004).

In this chapter, I will describe genetic and biochemical experiments designed to answer the following question: How does *IHF* inhibit transcription? We anticipated that the underlying mechanism would be novel because, unlike *Fis*, *IHF* does not compete directly with *CRP* for binding. To answer this question, we set out to: 1) test whether the 31-bp sequence that includes *IHF III* was responsible for the lowered transcriptional profile exhibited by the *acsP2* promoter; 2) identify the factors involved in this *IHF*–mediated inhibition of *acsP2* transcription; 3) determine how these factors work together to inhibit *acs* transcription; and 4) determine whether *Fis* plays a role in this inhibitory mechanism. This study demonstrates that interactions between well-characterized regulators can, within the context of complex promoters, result in novel mechanisms of gene regulation that remain to be explored.
Inhibition of CRP-dependent transcription by *IHF III* depends on *CRP II*

Previous results suggested the existence of multiple IHF-mediated nucleoprotein complexes that exert differential effects upon *acs* transcription (Browning et al. 2004; Browning et al. 2002). To simplify the testing of this hypothesis, we first sought to define the minimal system required for IHF-dependent inhibition. We began with 1xIHF (also known as *pacs236*), an *acs::lacZYA* promoter fusion that includes the inhibitory 31-bp sequence (Browning *et al.*, 2004). Since the 31-bp sequence includes *IHF III* (Fig. 9A), we determined whether it participates in the inhibition of *acs* transcription by introducing a double base pair replacement mutation (Fig. 9B) predicted to diminish affinity for IHF (Browning *et al.*, 2002; Lee *et al.*, 1991). The mutant promoter fusion, 1xIHF-I3m, and all others used in this study, was recombined into a hybrid λ bacteriophage and introduced as a monolysogen into AJW678, a Δlac strain that is wild-type (WT) for both acetate metabolism and *crp* (Table 1). We grew the resulting monolysogen in tryptone broth (TB) at 37°C, harvested cells at approximately one hour intervals, and compared its β-galactosidase activity to that of monolysogens that carried 1) its parent 1xIHF, 2) the smaller promoter fragment, 0xIHF (also known as *pacs205*), that includes *acsP2, CRP I*, and *CRP II*, but lacks *IHF III* (Table 2, Fig. 9A), or 3) the full-length *acs::lacZYA* fusion (FL) from which 1xIHF and 0xIHF were derived (Table 2, Fig. 9A). The mutant 1xIHF-I3m exhibited about two-fold more activity than its parent 1xIHF, but only about a third as much as FL and only about half as much as 0xIHF (Fig.
These results support the supposition that *IHF III* contributes to inhibition of *acs* transcription and that sequences upstream of *IHF III* contribute to full activation.

To determine whether inhibition by IHF also requires the integrity of the DNA sites for CRP, we introduced mutations (Fig. 9B) previously shown to reduce affinity for CRP (Ebright *et al.*, 1984). We incorporated these mutations into 1xIHF and monitored the activity of the resultant constructs (Fig. 9C). 1xIHF-C1m displayed about half the activity of its parent 1xIHF, approximately equivalent to basal levels (data not shown). In contrast, 1xIHF-C2m exhibited as much activity as 0xIHF, which lacks *IHF III* altogether. These results verify that activation of *acsP2* transcription requires *CRP I* (Beatty *et al.*, 2003) and shows that IHF-dependent inhibition requires *CRP II*.

This latter result came as a surprise, because Beatty and co-workers had previously reported that disruption of the *CRP II* site in the full-length promoter fragment (FL-C2m) exhibited two-fold less promoter activity than did its full-length, WT FL parent (Fig 9C) (Beatty *et al.*, 2003). To test our previous observation, we constructed a promoter fusion that includes *CRP I* but lacks *CRP II* (1xCRP; *pacs155*) (Fig. 9A), and monitored the activity of the resultant construct (Fig. 9C). As predicted if *CRP II* enhanced transcription, 1xCRP exhibited about one-half the activity displayed by 0xIHF, which includes both *CRP I* and *CRP II*. In the context of 1xIHF, therefore, *CRP II* has the opposite effect, acting negatively instead of positively.
Inhibition of transcription involves K101 of CRP

Previously, the Wolfe lab had shown that CRP-dependent activation of acs transcription from FL promoter depends on AR1 but not on AR2 (Beatty et al., 2003). To explore the mechanism of inhibition observed for 1xIHF, we performed similar studies on the smaller promoter constructs (Fig. 9A; Table 2). By generalized transduction, we introduced a crp deletion allele into monolysogens of 0xIHF and 1xIHF.

We then transformed the resulting strains with a set of plasmids, each of which expresses a CRP variant (Table
1). Each variant possesses one of the eight possible combinations of WT or mutant versions of three surfaces known to be involved in transcription activation. HL159 is defective for the AR1 surface that contacts the α-CTD of RNAP, KE101 is defective for the AR2 surface that interacts with the α-NTD, and the KE52 mutation exposes a non-native AR3 surface that contacts region 4.2 of σ^70. Activation of Class I promoters, such as acsP2, absolutely requires AR1 (Fig. 4A). In contrast, activation of Class II promoters requires AR2 and either AR1 or AR3 (Fig. 4B) (reviewed by (Busby and Ebright, 1999)).

For both 0xIHF and 1xIHF and for all other promoter constructs tested in this study, activation required AR1. Without it, as we previously observed with the FL promoter construct (Beatty et al., 2003), promoter activity was reduced to basal levels, regardless of the status of the other two surfaces (Fig. 10 and data not shown). When AR1 was left intact, activity from 1xIHF increased in the absence of AR2. This effect was reversed by the presence of AR3, while the effect of having both the AR2 and AR3 surfaces was additive (Fig. 10). In contrast, activity from 0xIHF was generally unaffected by either the status of AR2 or the addition of the AR3 surface to the AR2-less mutant. However, the presence of both AR2 and AR3 inhibited transcription. We conclude that AR2 contributes to IHF-dependent inhibition and propose that the additional presence of AR3 (when both AR1 and AR2 are intact) over-stabilizes RNAP on the DNA.

Since inhibition depends upon both CRP bound at CRP II and an intact AR2, we predicted that AR2 would have no effect in the absence of CRP II. To
test this prediction, we introduced the *crp* deletion allele into a monolysogen that carries 1xIHF-C2m, and transformed the resulting strain with plasmids that express wild-type CRP or mutants defective for AR1, AR2, and/or AR3. As a control, we did likewise with a monolysogen that carried 1xCRP, the promoter construct that includes only the promoter proximal *CRP I*. The major difference between the 1xCRP and 1xIHF-C2m was the presence of *IHF III*. As observed with all *acs* promoter fragments, both these promoter constructions required AR1 for activation. With 1xCRP, like with 0xIHF, neither the lack of AR2 nor the addition of AR3 to the AR2 mutant exerted a significant effect. In combination, however, these surfaces inhibited transcription (data not shown). Surprisingly, the loss of AR2 decreased transcription from 1xIHF-C2m, while the presence of AR3 suppressed this defect. As observed with 1xIHF, the combination of AR2 and AR3 quite severely inhibited transcription, resulting in levels comparable to that of mutants that lack AR1 (Fig. 10). Thus, on 1xIHF-C2m, the presence of IHF makes AR2 a positive component of CRP-dependent activation above that provided by AR1.
Characterization of protein binding

To understand the unexpected behaviour caused by disruption of AR2 and/or the presence of IHF III, we characterized the binding of each protein to the acs promoter region in vitro. Our collaborator, Dr. Bianca Sclavi, performed DNase I footprint analyses of both 1xIHF (Fig. 11) and FL (Figure 12) to determine whether the presence of IHF bound at IHF III and CRP bound to CRP II can affect the binding of RNAP and the structure of the open complex. To best mimic in vivo conditions, I constructed supercoiled DNA templates.

CRP

As observed previously with the linearized FL template (Beatty et al., 2003; Browning et al., 2004), WT CRP elicited patterns of protection and hypersensitive (HS) sites characteristic of binding to both CRP I and CRP II (Fig. 11A, lane 6; Fig. 11B; Fig. 12, lane 8). Surprisingly, CRP also elicited a pattern of protection and HS sites upstream of CRP II. This pattern appears consistent with the existence of a third, previously unidentified CRP binding site (CRP III). Because this “site” bears only a vague resemblance (CTGCA n6 TCAAA) to the consensus sequence (TGTGA n6 TCACA, (Ebright et al., 1989; Gaston et al., 1989)) and because we did not observe this site when using the linearized FL template (Beatty et al., 2003; Browning et al., 2004), we suspect that supercoiling might increase the affinity of the template for CRP at this site.
Figure 11. Simultaneous binding of IHF, CRP, and RNAP results in protection between the tandem CRP sites. (A) 0.1 nM pSRacs236 (carrying the 1xIHFacs promoter sequences from positions -171 to +65) was incubated with purified IHF (20 nM), either wild-type CRP or the AR2 mutant KE101 (100 nM), and/or increasing concentrations of RNAP, subjected to DNase I footprinting. The concentrations of RNAP in each reaction was as follows: lanes 1-2, 6, 13-15, 19 (0 nM); lanes 3, 7, 10, 16, 20, 23 (10 nM); lanes 4, 8, 11, 17, 21, 24 (20 nM); lanes 5, 9, 12, 18 (50 nM). The stars and diamonds denote RNAP- and CRP-associated HS sites, respectively. C, no proteins; U, uncut DNA. (B) Line graph representing the data from lanes 12 (RNAP only, dotted line) and 9 (RNAP plus CRP, solid line). The stars and diamonds denote RNAP- and CRP-associated HS sites, respectively. The binding sites for CRP and RNAP are underlined. (C) Line graph representing the data from lanes 2 (IHF and WT CRP, solid line) and 5 (IHF, WT CRP and RNAP, dotted line). The stars denote RNAP-associated HS sites. The grey bar denotes the intersite region.
On the linearized (Browning et al., 2004) or on the supercoiled FL template, IHF protected an extensive stretch that corresponds to IHF I, IHF II and IHF III (Fig. 12, lanes 3,16). In contrast, because 1xIHF does not include the sequences for the two upstream sites, IHF protected only sequences that...
correspond to IHF III (Fig. 11A, lanes 1, 14). With both templates, the binding of IHF was characterized initially by saturation of its specific binding site(s) followed, at higher concentrations (>20 nM), by propagation of the footprint to adjacent regions of the DNA (data not shown), a behaviour indicative of binding to lower affinity sites. This was followed (>100 nM) by the appearance of hypersensitive sites (data not shown), indicative of a reorganization of the DNA, possibly due to the formation of a regular, compact structure.

**CRP and IHF**

The presence of 20 nM IHF (Fig. 11A, lanes 2, 15; Fig. 11C; Fig. 12, lanes 4, 17) caused the loss of the CRP III-associated HS sites. In contrast, IHF exerted no substantial effect on the affinity of CRP to either CRP I or CRP II (data not shown).

**RNAP and CRP**

The binding of RNAP alone provided weak protection from about -30 to about -18 and induced a quadruplet of HS sites at -9 to -6 (Fig. 11A, lanes 10-12; Fig. 11B; Fig. 12, lanes 12-14). Since RNAP alone can bind either to acsP2A or to acsP2 (Beatty *et al.*, 2003), we interpret this pattern as the average of the
protection produced by the RNAP bound to \textit{acsP2} with that produced by the RNAP bound at \textit{acsP2A}.

The presence of CRP altered the pattern of protection and HS elicited by RNAP at multiple locations (Fig. 11A, compare lanes 7-9 to lanes 10-12; Fig. 11B; Fig. 12, compare lanes 9-11 to lanes 12-14). For example, the quadruplet of HS sites disappeared. Instead, a doublet of HS sites appeared (at positions -24 and -26) in the 17-bp spacer region between the -10 and -35 hexamers, indicative of a distortion of the double helix. Just downstream of the \textit{CRP I} site, an additional protection was observed, possibly due to binding of the \(\alpha\)-CTD as a result of its interaction with AR1. The reverse also was true; the presence of RNAP influenced the binding of WT CRP, substantially increasing the HS associated with all three CRP binding sites and with the HS sites between \textit{CRP I} and \textit{CRP II} (Fig. 3A, lanes 7-9; Fig. 12, lanes 9-11). Taken together, these results are consistent with the formation of a cooperative interaction between CRP and RNAP resulting in a structural rearrangement of the open complex.

\textbf{RNAP, CRP and IHF}

The combination of RNAP, IHF, and CRP (Fig. 11A, lanes 3-5; Fig. 11C; Fig. 12, lanes 5-7) increased protection throughout the region downstream of \textit{CRP II} to about -40, especially between \textit{CRP II} and \textit{CRP I}. This inter-site protection required IHF: in its absence, RNAP (Fig. 11A, lanes 7-9; Fig. 12, lanes
9-11) exerted considerably less influence. The replacement of WT CRP with the KE101 mutant in the presence of IHF resulted in a significant decrease in the protection between CRP I and CRP II (Fig. 11A, compare lanes 16-18 to lanes 3-5; Fig. 12, compare lanes 18-20 to lanes 5-7). Similar results were observed with the FL template (data not shown). Taken together, these data show that RNAP, CRP, and IHF can bind
simultaneously at saturating concentrations to their respective specific binding sites. Furthermore, they show that RNAP, IHF, and the AR2 surface of CRP are required for the protection observed between the two CRP binding sites. This inter-site protection may be interpreted as resulting from the formation of a more compact higher order nucleo-protein complex.

**Characterization of open complex formation**

Since disruption of the lower affinity CRP binding site, CRP II, increased *in vivo* promoter activity from 1xIHF, we predicted that within this context low concentrations of CRP would favor transcription, while high concentrations would inhibit it. To test this prediction, our collaborator, Bianca Sclavi, used potassium permanganate, which modifies thymine residues found within single-stranded DNA, to probe the supercoiled 1xIHF template for strand separation by RNAP, a step in the process of transcription initiation called transcription open complex (OC) formation (Record *et al*., 1996). Potassium permanganate modifies single-stranded DNA that results either from OC formation or from severe bends or kinks in the DNA (Sasse-Dwight and Gralla, 1989).

In the presence of RNAP alone (Fig. 13, lane 2), modification of the 1xIHF template occurred over an extensive region, as observed previously with a linearized template of FL (Beatty *et al*., 2003). These modifications can be sorted into two groups: P2A (-30/-29/-28, -22/-19) and P2, which can be further
subdivided into P2UP (-13/-12, -10/-9, -7/-6) and P2DOWN (-1/+3). In the presence of both RNAP and 20 nM IHF (Fig. 13, lane 3), a decrease in the amount of modifications at P2A occurred. Since the frequency of modification at P2UP did not change, we conclude that occupancy of IHF III by IHF does not affect the ability of RNAP to form an OC at acsP2. In contrast, the titration of WT CRP (Fig. 13, lanes 4-7) distinctively influenced each grouping. Increasing CRP concentration progressively decreased modification at P2A, while generally increasing modification at P2UP, as we observed previously (Beatty et al., 2003),

![Bar graphs representing the intensity of the bands at P2DOWN compared to the intensity of the band at -10. The hatched bars represent the pacs236 (1xIHF) data shown in Fig. 13, lanes 2-9; the other bars represent data from a similar analysis (performed and analysed simultaneously) of pSRacs444 (FL) performed in the absence of IHF (black) or in the presence of IHF (grey). The numbers on the x-axis denote the concentration in nM of either WT CRP or the mutant CRP KE101. Two pairs of reactions with no CRP are represented: the first was performed in the absence of IHF and the second was...](image)
especially at -6 and -7 (Fig. 13B). Strikingly, the effect at P2DOWN was non-linear (Fig. 13B and 14): modifications increased up to 12 nM CRP, above which they decreased. Since low CRP concentrations favoured occupancy of CRP I, while high concentration favoured binding of both CRP sites (Beatty et al., 2003); data not shown), bubble extension to P2DOWN thus inversely correlates with occupancy of CRP II. Strikingly, the OC formed in the presence of the KE101 mutation was more open at its downstream end independently of the CRP concentration. Under these conditions, we could observe an increase in the intensity of the bands at -1 and +3 and a decrease of the intensity of the bands at -9, -7 and -6 (Fig 13B and 14). This observation supports the hypothesis that the AR2 surface is involved in an interaction between RNAP and CRP bound at CRP I and that the binding of WT CRP to CRP II can stabilize this interaction and affect the structure of the resulting OC. Since we observed similar behaviour with FL in either the absence or presence of IHF (Fig. 14), this effect of CRP at the acs promoter appears to be general.

**CRP stabilizes the OC in the presence of binding to lower affinity sites by IHF**

Since 20 nM IHF did not exert a large effect on the RNAP footprint, we speculated that the negative effect of IHF observed previously on FL in vitro (Browning et al., 2004) might be due to its binding to lower affinity sites.
Therefore, we followed the formation of the OC at increasing RNAP concentrations at a moderately higher concentration of IHF (30 nM). At this concentration, IHF inhibited OC formation at both acsP2A and acsP2, even at the highest RNAP concentration tested. However, the addition of either WT CRP (Fig. 15A) or KE101 (Fig. 15B) following the addition of IHF permitted RNAP to form a OC at acsP2, but not at acsP2A. This was true even at the lower concentration of WT CRP (12 nM). Thus, the CRP-RNAP complex can successfully compete with this concentration of IHF for access to the acsP2 promoter, while RNAP alone cannot.
Discussion

The multiple roles of CRP during activation of acs transcription

Previously, we reported that in vitro RNAP alone could form an open complex (OC) at both acsP2A and acsP2. These OCs, however, were poorly productive, as determined by in vitro transcription assays. In contrast, the presence of CRP favoured OC formation at acsP2 and inhibited OC formation at acsP2A and, as a result, activated transcription in vitro. Similarly, activation of transcription in vivo required CRP, its AR1 surface, and an intact CRP I. On the basis of these observations, we proposed that CRP “focused” RNAP to acsP2 (Beatty et al., 2003). In the current study, we show that the binding of CRP does not substantially increase the extent of OC formation by RNAP on the acsP2 promoter. Instead, it causes a rearrangement of the interactions that RNAP makes with that promoter and thus stabilizes a different form of the OC (see below). The increased stability of this ternary CRP-RNAP-acsP2 complex activates transcription at multiple levels. First, it inhibits the formation of nascent binary RNAP-acsP2A complexes, but cannot successfully compete with pre-existing ones (data not shown). Second, it results in a conformation of the open complex that permits RNAP to initiate transcription and to overcome barriers to elongation. Third, it stabilizes the OC in the presence of non-specific IHF binding. This third effect might be important in vivo when IHF accumulates, which occurs during entry into stationary phase (see below).
The following evidence supports this model. (1) On naked DNA, in the absence of CRP, the majority of OC are formed at acsP2A (Figs. 13, 14 and 15). (2) The presence of even small amounts of CRP dramatically diminished OC formation at acsP2A, without significantly altering the amount of OC formation at acsP2, although it did affect its structure (Figs. 13, 14 and 15). (3) The presence of small amounts of RNAP influenced the pattern of CRP-dependent DNase I protection and HS and vice versa (Figs. 11 and 12), consistent with a cooperative interaction between RNAP and CRP. (4) Disruption of AR1 reduced the CRP-dependent inhibition of RNAP-acsP2A complex formation (data not shown), most probably by decreasing the cooperativity of the CRP-RNAP interaction. This reduced cooperativity would diminish the stability of the ternary CRP-RNAP-acsP2 complex and, thus, provide an opportunity for unbound RNAP to bind to free acsP2A. (5) In vivo activation of acs transcription depends upon both AR1 (Beatty et al., 2003) (Fig. 10) and determinant 287 (Beatty et al., 2003), a surface patch of the α-CTD known to interact with AR1 (Busby and Ebright, 1999).

Most often the role of CRP at a Class I promoter (e.g. lac) is to stabilize interactions of RNAP with the promoter. Stabilization of the closed complex permits the subsequent formation of a OC (Busby and Ebright, 1999). At the acsP2 promoter, this is still true (Fig. 16A) and is important for the CRP-RNAP complex to compete with IHF and possibly other nucleoid proteins (see below). However, RNAP alone can bind the acsP2 promoter, albeit in an inactive open conformation (Beatty et al., 2003). This inactivity might result from improper
Forward-leaning TOCA

\begin{align*}
&\text{A}\quad +\text{IHF} \\
&\quad \downarrow \text{+ CRP}
\end{align*}

\begin{align*}
&\text{B}\quad +\text{IHF} \\
&\quad \downarrow \text{+ CRP}
\end{align*}

Backward-leaning TOCI

\begin{align*}
&\text{C}\quad +\text{IHF} \\
&\quad \downarrow \text{+ CRP}
\end{align*}

\begin{align*}
&\text{D}\quad +\text{IHF} \\
&\quad \downarrow \text{+ CRP}
\end{align*}

\begin{align*}
&\text{E}\quad +\text{CRP} \\
&\quad \downarrow \text{+ CRP}
\end{align*}

\begin{align*}
&\text{F}\quad +\text{IHF} \\
&\quad \downarrow \text{CRP/IHF}
\end{align*}

Figure 16. Model of the complexes formed at the acsP2 promoter. RNAP alone can form an open complex at the acsP2 promoter. This complex, however, is inactive (Beatty et al., 2003). The occupation of CRP I by a CRP dimer mediates a change in the structure of the open complex by making contact between its AR1 surface and that of the $\alpha$-CTD of RNAP. This results in a deformation of the spacer, stabilization of the P2 TOC at the expense of the P2A TOC, and transcription from the former. Depending upon the concentration of CRP, this interaction can lead to two distinct open complexes: TOCA (top) and TOCI (bottom). These two complexes differ in the extent of opening at the downstream end of the bubble and in the level of transcription \textit{in vivo}. Concentrations of CRP that lead to the occupation of CRP I alone favor the formation of the forward-leaning TOCA, in which the DNA is open to the -1 to +3 positions (A). Concentrations that favor the occupation of both CRP sites favor the formation of the backward-leaning TOCI (B). TOCI is stabilized by an interaction with the AR2 region of CRP. The increased number of interactions formed in this complex, compared to those formed in TOCA, increases the stability of RNAP on the promoter, which results in increased transcription. The additional occupation of IHF III either increases or decreases transcription, depending upon the status of CRP II. When CRP II is not occupied, the binding of IHF enhances transcription, perhaps by increasing the stability of the open complex (C). When both CRP sites are occupied, the binding of IHF to IHF III results in significantly reduced transcription, possibly because compaction of the upstream elements overstabilizes the complex. The binding of IHF also causes increased protection between the two CRP sites, possibly through a rearrangement of the interaction of the $\alpha$-CTDs with the upstream elements (D and F). This protection depends upon the presence of AR2, and thus can only occur in the TOCI conformation. Occupation of the upstream IHF sites alleviates much of the IHF III-dependent compaction (F). The thickness of the arrows indicates the relative transcription from each complex. Note that although RNAP possesses only two $\alpha$-CTDs, in some cases three or more $\alpha$-CTDs are shown in order to represent a number of possible contacts.

positioning of the $\alpha$-CTD as observed at the malT promoter, whose binding site
for α-CTD (i.e. UP element) and single binding site for CRP resemble CRP I of acsP2 and the DNA sequence located immediately downstream (Tagami and Aiba, 1999). To activate acsP2 transcription (in the absence of IHF), CRP performs two roles. The first role is to inhibit the binding of RNAP to the overlapping promoter, acsP2A. The second is to restructure the OC. By analogy to the malT promoter, occupancy of CRP I by CRP might reposition the α-CTD to a more favourable position through an interaction with its AR1 surface (Tagami and Aiba, 1999), resulting in distortion within the spacer and stabilization of a OC open at its downstream end. The end result would be a complex that is more efficient in promoter escape.

The role of AR2

AR2, a well-characterized surface patch of CRP, has been shown to activate transcription of Class II promoters by interacting with the α-NTD of RNAP (Busby and Ebright, 1999; Lawson et al., 2004). To the best of our knowledge, AR2 has never been reported to play a role in the regulation of a Class I promoter or of a Class III promoter comprised of tandem Class I sites, such as acsP2. Yet here, we determined that in vivo AR2 could influence AR1-dependent activation of this Class III promoter, acting to either inhibit or activate transcription depending upon context. Furthermore, we showed in vitro that disruption of the AR2 surface results in a different structure of the OC, similar to
the one observed at low WT CRP concentrations, when CRP II is less likely to be occupied.

We propose that the presence of CRP bound at CRP II contributes to the recruitment of the RNAP to the promoter through an AR1-$\alpha$-CTD interaction (see below), causing enhanced transcription as expected for a Class III promoter comprised of tandem Class I sites (Fig. 16, compare panels A and B). This long-range interaction results in a bend in the upstream DNA that permits an additional stabilizing contact via the AR2 surface between RNAP and CRP bound to CRP I. We further propose that two distinct versions of the OC can form at acsP2 in the presence of CRP (Fig. 16). The first version, termed OC$_A$, is expanded at its downstream end (to include positions -1 and +3). It is associated with the occupancy of CRP I and/or the absence of the AR2 surface. The second version, termed OC$_I$, is more open at its upstream end (at positions -6 and -7). It is associated with the occupancy of both CRP sites by WT CRP and, because of the presence of additional upstream elements (e.g. CRP bound CRP II), with a more stable RNAP-DNA interaction.

In the context of the 1xIHF promoter construct, where the binding of IHF to IHF III accentuates the bending of upstream DNA, these additional upstream stabilizing interactions can result in a decrease in the amount of transcription (Fig. 16D). This model is supported by the following observations. First, the reduced transcriptional activity exhibited by 1xIHF required IHF, CRP II and an
intact AR2. Second, with IHF bound at its specific sites, additional protection was observed between the two CRP sites, but only in the presence of the intact AR2 surface. Third, the non-native AR3 surface could replace AR2. Intriguingly, evidence exists that AR3, a native surface of the CRP paralog FNR, inhibits transcription from certain semi-synthetic FNR-dependent Class III promoters (Wing et al., 1995). Remarkably, an AR2-dependent interaction also could exert a positive role. Maximal transcription from 1xIHF-C2m (which includes CRP I and IHF III, but not CRP II) depended upon the presence of both IHF and AR2 as well as the inability of CRP to bind at CRP II (Fig. 16C).

The mechanism by which the AR2 and AR3 surfaces of a CRP dimer bound at a Class I position can interact with RNAP remains unclear; however, it seems reasonable to propose that they do so by interacting with their known RNAP targets, i.e. the α-NTD and region 4.2 of σ70, respectively. These interactions could take place through a bend in the DNA region between the binding sites for the two proteins (i.e. CRP I and acsP2) or, alternatively, by displacing RNAP from the promoter to bind next to CRP.

Whether OCi and OCA correspond to RP_{O1} and RP_{O2} as defined by Record and co-workers (Record et al., 1996) or the moribund and active complexes described by Shimamoto and co-workers (Kubori and Shimamoto, 1996; Susa et al., 2006) remains to be determined (Hsu, 2002). In the first case,
OC₁ would be a necessary intermediate in the formation of OC₂. In the second case, these complexes would be the products of a branched pathway.

**The roles of IHF**

IHF is an abundant nucleoid protein that contributes to the organization and the condensation of the chromosome. Estimated to be about 0.7 nM during early exponential growth (Murtin et al., 1998), its concentration further increases throughout growth to become the second most abundant nucleoid protein during stationary phase (Azam et al., 1999; Ishihama, 1999). When IHF binds to its specific sites, it causes a large distortion of the DNA where the double helix bends up to 180 degrees and wraps around the protein (Lynch et al., 2003; Rice et al., 1996; Travers, 1997). On the acs promoter, IHF can affect transcription from three high affinity sites (IHF I - III) located between −140 and −240 upstream of the transcription start site (Browning et al., 2004).

A major goal of this investigation was to understand the mechanism by which IHF binding could influence the activity of RNAP at this promoter. On the basis of our results, we propose a model by which IHF uses two distinct mechanisms to affect transcription. The first mechanism depends on the binding of IHF to its specific sites and results in the formation of a wrapped nucleoprotein complex that also includes both CRP and RNAP. In contrast, the second mechanism depends on the role of IHF as a nucleoid protein that, when present
at higher concentrations, binds to lower affinity, less specific, sites and, in the process, inhibits open complex formation.

The first mechanism is supported by the following observations: (1) the presence of IHF on its three specific sites in the FL construct or on the IHF III site of the truncated 1xIHF does not inhibit the formation of an open complex in vitro in either the presence or the absence of CRP (Figs. 13-15). However, in the presence of both RNAP and CRP, IHF does mediate increased protection of the region between the two CRP sites (Figs. 11 and 12). (2) IHF-dependent transcription inhibition in vivo requires both IHF III and CRP II (Fig. 9). It also requires that at least one CRP dimer possess an intact AR2 (Fig. 10). Thus, this mode of IHF-dependent inhibition likely results from the formation of a multi-protein complex where a multitude of stabilizing contacts made by RNAP results in the formation of a poised OC as described above (Fig. 16D).

That over-stabilization of the OC by IHF could contribute to negative regulation in vivo is supported by observations made previously in the context of the FL WT acsP2 promoter construct. In this context, specific single alanine substitutions in the α-CTD subunit of RNAP resulted in an increase in the amount of transcription (Beatty et al., 2003). These mutations were located at R265, a residue that enhances interaction with DNA; L260 and L262, residues located close to and perhaps part of determinant 261, known to interact with region 4 of sigma; and T301 and I303, which are either part of the DNA interaction determinant or of some previously unknown surface. Mutation of these amino
acids could either decrease the stability of the open complex allowing for a faster escape or result in a less tightly wrapped multi-protein structure.

A second mode of IHF-dependent repression results from binding to lower affinity sites in the promoter region. While the three specific sites become protected with an affinity of about 2 nM, we can observe the appearance of additional DNase I protection due to binding to these lower affinity sites beginning at 30 nM (data not shown). Under these conditions, IHF inhibits OC formation by RNAP, which can be overcome by the presence of low concentrations of CRP (Fig. 15). With IHF bound to its lower affinity sites, RNAP cannot form the closed complex, presumably due to the low stability of these initial interactions. In contrast, IHF at this higher concentration does not inhibit the binding of RNAP in the presence of CRP, even when the transcription factor is present at the lower concentration (12 nM) tested (Fig. 15). It appears that CRP can stabilize the closed complex enough to compete successfully with IHF and, thus, form an OC. This mechanism might be essential for transcription to proceed in the face of the increased amounts of IHF as the cells enter stationary phase (Azam et al., 1999).

Intriguingly, IHF also enhanced CRP-dependent activation in both the 1xIHF and FL constructs (Fig. 16C). In the 1xIHF construct, this was observed in vivo only when CRP II remained unoccupied, exemplified by the mutant 1xIHF-C2m construct (Fig. 10). In vitro results suggest that it could also occur on the WT promoter whenever CRP concentrations are low and, hence, CRP II remains
unoccupied. Since the CRP-RNAP complex cannot compete with preformed OC at acsP2A, the presence of 12 nM CRP only partially inhibited OC formation at acsP2A. In contrast, the combined presence of IHF and 12 nM CRP resulted in the formation of virtually no open complex at acsP2A (Fig. 15). Thus, the occupation of IHF III both increases the probability that an OC will form at acsP2 and, as described above, the probability that the OC eventually will become poised, i.e. when the CRP concentration rises to levels that favour binding at CRP II.

In vivo, the upstream sequence that includes IHF I and IHF II exerts an overall positive effect on transcription ((Browning et al., 2004) and Fig. 9). It still cannot be excluded that this positive effect could result from the binding to this region of some as yet unidentified transcription factor; however, we currently favour the possibility that it results from the ability of the multi-protein complex to stabilize a DNA geometry that favours transcription initiation. The high affinity of IHF for these sites (about 2 nM, data not shown) suggests that they might be partially occupied most of the time in vivo and, thus, might play a role in the local organization of the nucleoid (Murtin et al., 1998), especially under growth conditions that favour strong activation of the acs promoter. It is anticipated that large differences in the local structure of the DNA and its degree of supercoiling would occur when IHF binds to all three sites (Fig. 16F) relative to when it binds only IHF III (Fig. 16D). These differences could directly affect the efficiency of transcription initiation (Travers, 2007).
While we did not observe a substantial difference in the overall affinity of RNAP for \textit{acsP2}, or of CRP for its sites, we cannot exclude the possibility that the presence of IHF exclusively at \textit{IHF III} and the resulting changes in DNA structure might affect the on and off rates of binding. An increased stability of the complex would result in a slower $k_{\text{off}}$. When combined with a decrease in $k_{\text{on}}$, possibly caused by reduced accessibility to the promoter, this would result in an unchanged affinity constant. We are currently investigating this possibility.

While investigation of the 1xIHF promoter construct permitted identification of AR2 as a potential mediator of stabilizing interactions with upstream elements, it should not be ignored that the \textit{CRP II-}, \textit{IHFIII-} AR2-dependent inhibition observed with this truncated promoter construct could occur within the context of the FL WT promoter. The two upstream IHF sites co-localize with binding sites for other transcription factors (NarL/P and FNR) and with at least one promoter, \textit{acsP1}. Competition has been reported between IHF bound at \textit{IHF I} and both RNAP and FNR for their overlapping sites and between IHF bound at \textit{IHF II} and NarL/P for their overlapping sites (Browning et al., 2002; Browning et al., 2005). One could easily imagine that competition between IHF and these other proteins might result in incomplete saturation of the two upstream IHF sites, especially when faster growth rates keep the IHF pool small enough to only half-saturate its highest affinity sites (Murtin et al., 1998). If the degree of wrapping of the nucleo-protein complex depends inversely on the number of IHF sites that become filled, then at low IHF concentrations, a tightly wrapped complex would result in over-
stabilization, poising the RNAP on the promoter (Fig. 16D). As growth slows and more IHF sites become filled, the structure would become “loosened”, allowing for more efficient escape of the poised RNAP (Fig. 16F). Because the RNAP would already be present at acsP2 in the poised OC, it could respond rapidly to the need to utilize acetate as a carbon source, a requirement that occurs as cells approach stationary phase and the IHF pool approaches its maximum.

The global regulator CRP, like the nucleoid protein IHF, has many low affinity sites on the chromosome (Grainger et al., 2005). CRP III may be one of them. However, it is contained completely within a high affinity IHF site (IHF III) that is likely to be filled as the cells reach stationary phase. Since this low affinity CRP site also appears to be dependent on the supercoiling of the DNA and the negative supercoiling of the DNA decreases as the cells reach stationary phase (Travers and Muskhelishvili, 2005), the affinity of CRP III would be expected to be at its lowest when IHF approaches its highest concentration. Thus, the question remains whether this low affinity CRP site functions in vivo and, if so, whether that function is positive or negative. In either case, the presence of low affinity sites for both IHF and CRP raises a much larger issue. Although the primary role of nucleoid proteins is to fold and organize the chromosome and the primary role of global regulators is to modulate transcription, their roles often overlap. The exact role in both cellular processes will depend upon the interplay between high and low affinity sites of any given protein and the competition for overlapping sites.
Concluding thoughts

It is typically thought that the regulation of transcription depends on the ability of one or a small set of transcription factors to recruit RNAP to the promoter and to form a complex active for transcription. These interactions must stabilize the active complex on the promoter long enough to allow for RNAP to begin transcription.

The regulation of the acsP2 promoter depends on a delicate balance of protein-DNA and protein-protein interactions that permit integration of several different signals, each of which reflects different aspects of the metabolic state of the cell. For acsP2, the most important signals appear to be those represented by the concentrations of...
CRP, IHF and FIS. The specific pattern of *acs* transcription results from the requirement for a combination of specific amounts of these three global regulators. Activation occurs through rearrangements of the OC in addition to stabilization against binding to lower affinity sites by IHF and other nucleoid proteins. In contrast, inhibition can occur by over-stabilization of the OC by CRP as in the context of 1xIHF. The resulting complex (comprised of IHF bound at *IHF III*, CRP dimers bound at *CRP II* and *CRP I*, and RNAP bound at *acsP2*) could fold into a wrapped conformation that inhibits escape due to numerous interactions between RNAP and several upstream elements (Fig. 16D). Inhibition of transcription at such a late stage in the process of forming a transcriptionally active complex could facilitate a rapid response upon a change in nutritional status.

In summary, during the early stages of exponential phase, FIS inhibits *acs* transcription by steric hindrance of CRP binding (Browning *et al.*, 2004). Acetyl-CoA synthetase is a high affinity enzyme that permits the scavenging for and metabolism of small amounts of acetate previously excreted into the medium; thus as the primary carbon source becomes depleted, the activation of the *acs* promoter by small amounts of CRP-cAMP would delay entrance into stationary phase. Thus, the proposed mechanism would link activation of *acs* transcription to this transitional period. Furthermore, once the primary and secondary carbon sources become depleted, the further increase in cAMP concentration would result in formation of the inhibitory complex, reflecting the reduced metabolic
rate. Finally, condensation of the nucleosome in stationary phase by IHF and other nucleoid proteins would result in inhibition of open complex formation at acsP2.

This study was published as:


What follows are experiments related to but not included in Sclavi et al. These experiments specifically explore the role of the NAP Fis and of *FIS II*, its DNA site that sits between but does not overlap the two CRP sites.

**Fis-Dependent Inhibition of Transcription**

Previously, we reported that the nucleoid protein, Fis, binds three sites within the *acs* promoter region (*FIS I, FIS II, and FIS III*) (Fig. 6B). We also reported that Fis inhibits *acs* transcription (Browning et al 2004). Part of this inhibition can be attributed to anti-activation, i.e. the ability of Fis to out-compete CRP for binding at their overlapping sites (*Fis III* and *CRP I*) (Fig. 8). In contrast, the contribution of *FIS II* to inhibition is unclear. That *FIS II* contributes to inhibition, however, is supported by the observation that its disruption in the WT FL promoter construct permits elevated *acs* transcription that approximately
equals that attained by disruption of *FIS III* (Browning et al. 2004). Since *FIS II* does not overlap either *CRP I* or *CRP II*, we sought a mechanism distinct from the anti-activation mechanism associated with *FIS III*. To seek its role, we disrupted *FIS II* in the context of the 1xIHF promoter fragment (1xIHF-F2m; Table 2).

The 1xIHF-F2m promoter was compared to 1) its WT parent 1xIHF, 2) FL, and 3) 0xIHF (Fig. 17). As we reported previously, 0xIHF displayed high promoter activity, while 1xIHF exhibited extremely low activity (Browning et al., 2004; Sclavi et al., 2007). Surprisingly, 1xIHF-F2m displayed activity similar to that of 0xIHF or approximately six-fold more activity than its WT parent (1xIHF). Whereas the activities of the other promoter fragments peaked near the transition from exponential phase into stationary phase then steadily decreased in stationary phase, the 1xIHF-F2m promoter activity remained high despite entry into

![Figure 18. Both Fis and IHF reorganize RNAP binding from P2a to the productive P2U promoter.](image)

Potassium permanganate footprint analysis was utilized to determine the location of open complex formation in the presence of constant concentrations of RNAP, CRP, and IHF. Increasing concentrations of Fis protein were titrated into the mix. P2D, P2 downstream; P2U, P2 upstream.
stationary phase (data not shown). I conclude that *FIS II* contributes to IHF-mediated inhibition of *acs* transcription.

To explore the role of Fis, my collaborator (Dr. Bianca Sclavi) performed a potassium permanganate analysis on supercoiled 1xIHF promoter fragment (Fig. 18). As observed previously (Fig. 13A), modification of the 1xIHF template occurred at both P2A and P2 (Fig. 18, lane 1), while the addition of a high concentration of CRP (200 nM) decreased modification at P2A, while increasing modification at P2UP (indicative of the unproductive transcription complex) and P2DOWN (indicative of the productive complex) (Fig. 18, lane 2). Increasing concentrations of Fis (up to 60 nM) correlated with increased modification at P2UP (Fig. 18, lanes 3-6). The presence of IHF inhibited this effect and further reduced the modification at P2A (Fig. 18, lanes 7-10).

We conclude that Fis does not inhibit the binding of RNAP. Furthermore, because Fis does not shift the modification back to P2A, it likely does not inhibit the binding of CRP. These results contrast with those obtained in the absence of RNAP, a condition under which Fis outcompetes CRP regardless of the order of addition. Yet, in vitro, Fis inhibits CRP-dependent *acs* transcription (Browning et al., 2004). Thus, we propose that Fis-mediated inhibition does not rely exclusively on the competition between the overlapping *Fis III* and *CRP I* sites and raises the possibility that *FIS II* is indeed involved (Fig. 8).

The genetic evidence (Fig. 17) clearly implicates *FIS II* in IHF-mediated
inhibition. However, the potassium permanganate footprint analysis of the 1xIHF template does not help to explain how. Intriguingly, the genetic evidence conflicts with the previously reported conclusion - formed on the basis of DNase I footprint and in vitro transcription analyses - that Fis and IHF function independently to inhibit acs transcription (Browning et al., 2004). Efforts to identify the role of FIS II in IHF-mediated inhibition are clearly warranted.
CHAPTER FOUR

EVIDENCE THAT IHF III IS A COMPOSITE SITE

Introduction

In the previous chapter, I presented evidence that IHF-mediated inhibition relies upon (1) the ability of IHF to bind IHF III, (2) the occupation of tandem CRP sites (CRP I and CRP II) by CRP, and (3) the surface determinant AR2 of CRP. On the basis of my genetic evidence and the biochemical evidence of my collaborator, I proposed that interactions between these three proteins and RNAP resulted in the formation of a stalled open complex (Fig. 16) (Sclavi et al., 2007). In brief, our model predicts that IHF-induced DNA bending causes the tandem CRP dimers to trap RNAP at the promoter in an unproductive, open complex.

In this chapter, I will provide evidence that IHF III is actually a complex sequence that consists of two overlapping sites for IHF (denoted IHF IIIa and IHF IIIb) offset by 2 bp, oriented in opposite directions, and positioned on opposing faces of the DNA helix.
Shared Protein Binding Sites in the \textit{acs-nrf} Intergenic Region

The approximately 300-bp \textit{nrf-acs} intergenic region includes the divergently transcribed \textit{nrf} and \textit{acs} promoters (Fig. 6). This region also appears to include all of the necessary regulatory elements for both promoters (Wolfe, 2009). As both promoters share many of these regulatory elements, this presents a logistical problem. The solution to this problem, however, may be simple. By evolutionary design or by a fortuitous coincidence, cells express the \textit{acs} and \textit{nrf} operons under vastly different, extreme conditions: \textit{acs} is required for survival under carbon starvation (Wolfe, 2009), while \textit{nrf} is required under anaerobic-nitrogen limitation (Browning et al., 2004). Yet, transcription from the \textit{acs} promoter does not seem to influence expression from the \textit{nrf} promoter and vice versa. It is proposed that this lack of competition between promoters in such close proximity results from the infrequent firing of both promoters (Browning et al. 2002). Thus, it is possible to understand the regulatory control of one promoter in relative isolation from the other promoter. However, because many of the regulatory elements are the same, lessons learned from the study of one promoter can be productively used to understand the other.

An Argument that \textit{IHF III} is a Composite Site

The central location of the three DNA sites for IHF lends support for the hypothesis that they play critical roles in regulating the transcription from both the
nrf and acs promoters. This hypothesis is further supported by the fact that they are highly conserved across all sequenced E. coli genomes (Browning et al., 2006). The experimental data make a compelling case for IHF I: by steric hindrance, it represses transcription from both nrf and the minor promoter acsP1 (Browning et al., 2002; Wolfe, 2009). The experimental data also provide evidence that IHF III plays a pivotal role. The binding of IHF to this site helps promote transcription from Pnrf by a presently uncharacterized mechanism (Browning et al., 2006). Meanwhile, it also mediates the formation of an unproductive stalled open complex at the acs promoter through a mechanism that also requires tandem CRP dimers and RNAP (Fig. 16) (Sclavi et al., 2007). The role of IHF II remains unclear (see Chapter 5).

Intriguingly, the mutations used to demonstrate the involvement of IHF III in the regulation of Pnrf and acsP2 implicate two different DNA sites for IHF (Browning et al., 2006; Sclavi et al., 2007). That these two sites might exist came from the observation that mutations introduced into the central motif of IHF III did not fully relieve IHF-mediated inhibition of acsP2 transcription (Fig. 9) (Sclavi et al., 2007). At first, I thought that incomplete relief of inhibition might result from incomplete loss of IHF binding. However, this conclusion seemed unlikely, as mutations of this kind in the central motif had been shown to cause severe defects in binding (Lee et al., 1991; Hales et al, 1994; Yang and Nash, 1995; and Aeling et al, 2006). A clue to this somewhat surprising result came from my realization that IHF III might actually be a composite site. The IHF III site reported
by our long time collaborators, the Busby lab, activated *nrf* transcription (Browning et al., 2006); however, it was not the same *IHF III* site that we had reported (Sclavi et al., 2007). The two sites overlapped, sat on opposite faces of the DNA helix and were offset by 2 bp (Fig. 19). Could both sites exist?

Preliminary Evidence Supporting the Composite Site Hypothesis

To answer this question, I first tested the alternative hypothesis: that *IHF III* is a single site consisting of only *IHF IIIa*. If this hypothesis is correct, then mutations that alter this site should affect 1xIHF promoter activity in a predictable fashion. For example, introduction of the double mutation C150G C152G into the central motif of the putative *IHF IIIa* (CCTCAA to CGTGA) reduces the site’s similarity to consensus (WATCAA). As described above, this mutation resulted in moderately enhanced transcription from the *acsP2* promoter (Sclavi et al., 2007), as would be predicted if the binding of IHF to *IHF III* were an inhibitory event. If so, I reasoned that mutations that brought the central motif of *IHF III* closer to consensus would decrease transcription from *acsP2*. Therefore, I constructed the

![Figure 19. The putative *IHF IIIa* and *IHF IIIb* sites. *IHF IIIa* is on top reading from left to right with the consensus sequence below. *IHF IIIb* is on the bottom reading from right to left with the consensus sequence above. W, A or T; R, A or G; n, any nucleotide. The 5’-W6, central WATCAA, and 3’-TTR motifs are denoted by a light blue, yellow, and green rectangles, respectively.](image)
C152A C153A double mutant, which converts the native central motif from CCTCAA to AATCAA, which is identical to central motif of the phage IHF site H’, considered by many to be the ideal IHF sequence (Lee et al., 1991).

Both mutations were introduced into the 1xIHF promoter fragment (Table 3). The mutant promoter fusions were then recombined into a hybrid bacteriophage and introduced as a monolysogen into AJW678, a lac strain that is WT for both acetate metabolism and crp (Table 1). The resulting monolysogens were grown in TB at 37°C, cells were harvested at approximately one hour intervals, and their β-galactosidase activity compared to that of monolysogens that carried the parent promoter fragment 1xIHF. As a control, I also tested the 0xIHF promoter fragment, which does not carry IHF III (Table 3).

As observed previously (Fig. 9) (Sclavi et al., 2007), the C150G C152G mutant (also known as 1xIHF-I3m) yielded activity intermediate to the WT 1xIHF and 0xIHF promoter fragments (Fig. 20). Contrary to our prediction, the C152A C153A mutant alleviated inhibition. Thus, I concluded that the alternative hypothesis was probably incorrect.

Coincidentally, the C153A mutation in the IHF IIIa site is predicted to strengthen the central motif (from GAACAT to TAACAT) of the putative IHF IIIb site reported by Browning et al. (2006). This led us to propose and then test the hypothesis that both sites existed and that they both influence acsP2 transcription.
**Systematic Mutation of IHF IIIa and IHF IIIb Sites**

Since the discovery of IHF, extensive work has been done to understand the mechanism by which it recognizes and binds to DNA. Like its homologue, HU, IHF has the propensity to bind to bent DNA structures. However, IHF does so in a sequence specific manner. There are two distinct DNA regions in an IHF site. The first being the core sequence WATCAA-n4-TTR (W = T or A; R = A or G; n = any base). These core bases are sufficient for IHF binding; however, binding is further enhanced when a dA-dT rich region is present 5’ to the core sequence. Because the bases of the dA-dT rich region vary, it is simply depicted as: 5’-W6-n8 (Hales et al, 1996; Goodrich et al, 1990).

To identify bases of the core sequence and dA-dT rich region important for IHF binding, Lee et al. (1991) generated single base pair mutations in three phage IHF sites - H’, H1 and H2 – and tested IHF affinity using a challenge phage assay. The resultant mutations were sequenced and categorized as weak,

---

**Figure 20.** Initial mutagensis of IHF III. β-galactosidase activity from promoter fragments 0xIHF, 1xIHF, and the 1xIHF variants C152G C150G and C153A C152A fused to lacZYA. The wild type strain AJW678 was lysogenized with the hybrid phages. The resultant strains were grown in TB, samples harvested at one-hour intervals, the β-galactosidase activity determined, and the peak activity expressed as a percentage of wild-type. Each value represents the mean ± SEM of at least three independent measurements. Promoter activity is relative to that of the 0xIHF construct.
moderate, or strong mutants (Fig. 21). Briefly, these studies revealed that single base pair mutations in the C of the central WATCAA motif and in the middle T of the TTR severely reduced IHF binding. Based on this study, I introduced a series of single base pair replacement mutations into the putative *IHF IIIa* site or the putative *IHF IIIb* site or both to favor IHF binding to one site or the other. The goal was to: 1) determine if *IHF III* is a dual overlapping IHF site; 2) assuming that both sites do exist, then determine if either or both sites affect *acs* transcription; and 3) finally determine the effect (either positive or negative) on *acsP2* transcription of IHF when bound to either *IHF IIIa* or *IHF IIIb*.

Each mutation was introduced into the 1xIHF promoter fragment. The mutant promoter fusions were then recombined into the hybrid bacteriophage and introduced as a monolysogen into AJW678 (Table 1). Each resulting monolysogen was aerated in TB at 37°C, cells were harvested at approximately

---

**Figure 21. Mutations reported to cause defects in IHF binding.** The 5'-W6, central WATCAA, and 3'-TTR motifs of the consensus and H' sites are upper case, bold and blue. The non-consensus nucleotides are lower case and black. Substitution mutations reported to cause binding defects are shown below. The length of the arrow denotes the strength of the defect. (adapted from Lee et al. 1991)
one hour intervals, and their β-galactosidase activity compared to that of monolysogens that carried the parent 1xIHF, 0xIHF, and the FL promoter fragments (Table 3). Figure 22 summarizes the results of these studies.

Mutations in the 3'-TTR motifs

**A164T T142A.** If both sites exist, I predicted that disrupting the second T in each TTR, and thus disrupting the ability of IHF to bind both sites, would result in behavior that resembles the 0xIHF promoter fragment. Thus, I substituted a T for the native A at bp 164 and an A for the native T at bp 142. Each of these mutations targets the middle T of the respective TTR, which is a mutation shown
to severely reduce IHF binding to lambda phage H' (Fig. 21) (Lee et al., 1991). Note that the each mutation also alters the opposing site's dA-dT rich region in a manner that severely diminished IHF binding (Lee et al., 1991). Thus, both A164T and T142A are predicted to reduce the consensus of both their respective 3'-TTR motif and the opposing site's 5'-W6 motif. The result should be a severe reduction in the ability of IHF to bind either site. Indeed, the double mutant A164T T142A performed well, resulting in activity that sometimes resembled that of the 0xIHF promoter fragment (Fig. 22) and sometimes that of the FL promoter (data not shown).

**A164T.** Since the A164T mutation reduces the consensus of the 3'-TTR motif of *IHF IIIb* and the 5'-W6 motif of *IHF IIIa* (Fig. 22), I chose to test this single mutant. As predicted, this mutation in the context of the 1xIHF promoter fragment increased promoter activity to a level that resembled that of the 0XlIHF promoter fragment.

**T142A.** In contrast, the T142A mutation reduces the consensus of the 3'-TTR motif of *IHF IIIa* and the 5'-W6 motif of *IHF IIIb* (Fig. 22) and, as such, is also predicted to reduce IHF binding to both *IHF III* sites. Unlike A164T, the T142A mutant promoter fragment exhibited activity intermediate to the 1xIHF and 0xIHF promoter fragments.
A141T. To be thorough, I also mutated the R of the *IHF IIIa* TTR motif, substituting a T for the native A, reducing consensus and thus likely binding affinity. According to Lee (Fig. 21), this mutation also should slightly increase consensus of the *IHF IIIb* 5'-W6 motif, by substituting the more consensus A for the native T on the complementary strand. The result of this mutation was promoter activity at or above that observed with the 0xIHF construct (Fig. 22).

Together, these results confirm the key role that *IHF III* plays in the inhibition of CRP-dependent *acsP2* transcription. However, since these mutations are predicted to disrupt binding of IHF to both sites (Lee et al., 1991), I could not distinguish the influence of each individual site. To do that, I had to introduce mutations predicted to exert a major effect on only one site. For *IHF IIIb*, I introduced G156T, the mutation reported by Browning and co-workers (Browning et al., 2006) and T155G; for *IHF IIIa*, I introduced C150A and A149C.

Mutations that disrupt one site or the other but not both

G156T and T155G. The substitution of either a T for the native G at bp 156 or G for the native T at bp 155 reduces the consensus of the central WATCAA motif of *IHF IIIb* (located on the complementary strand) from GAACAT to GAA\_AT or GA\_CAT, respectively. They both are predicted to disrupt binding of IHF to *IHF IIIb* (Lee et al., 1991). Because they alter the n8 between the 5'-W8 and the central motif, however, these substitutions should exert little or no effect upon the
binding of IHF to *IHF IIIa*. The G156T mutation severely diminished *Pnrf* activity, leading to the hypothesis that *IHF III* played a role in *nrf* transcription (Browning et al., 2006). Both mutations resulted in *acsP2* transcription intermediate to that of the 1xIHF and 0xIHF promoter fragments (Fig. 22). On the basis of these observations, it appears that *IHF IIIb* exists, that it plays a positive role for *Pnrf* activity, and that it plays some inhibitory role at *acsP2*.

**C150A and A149C.** In contrast, replacement of either an A for the native C at bp 150 or C for the native A at bp 149 reduces the consensus of the central WATCAA motif of *IHF IIIa* from CCTCAA to CCTAA and CCTCCA, respectively. Both substitutions are predicted to disrupt binding to *IHF IIIa* (Lee et al., 1991). However, both mutations should be without substantial, if any, affect on the binding of IHF to *IHF IIIb* because both substitutions alter the n8 located between the central and 5’ motifs. Both mutations exerted strong effects on *acsP2* transcription, resulting in activity that matched or exceeded that of 0xIHF (Fig. 22). On the basis of these results, it appears that *IHF IIIa* exists and that it plays a major inhibitory role at *acsP2*.

**Replacement of *IHF IIIa* with *IHF IIIb* and Vice Versa**

The experiments described above are consistent with the hypothesis that *IHF III* is a composite site composed of two overlapping sites that are oriented in
opposite directions and that sit on opposing faces of the DNA. Mutations predicted to severely reduce the affinity of both IHF IIIa and IHF IIIb (A164T or T142A or both) resulted in elevated transcription from the 1xIHF promoter fragment. Mutations predicted to severely reduce the affinity of either IHF IIIa (C150A or A149C) or IHF IIIb (T155G or G156T) provided evidence that both sites exist. However, the data are not definitive. To perform a more rigorous test of the composite site hypothesis, I

Figure 23. Swapping IHF IIIa and IHF IIIb. (A) The native IHF III sequence with the conserved motifs shown in red. (B) The VF sequence, in which IHF IIIb is positioned in the native IHF IIIa location and vice versa. (C) β-galactosidase activity from promoter fragments FL, 1xIHF, 0xIHF, 1xIHF-VF fused to lacZYA. The wild type strain AJW678 was lysogenized with the hybrid phages and the resultant strains were grown in TB, samples harvested at one-hour intervals, the β-galactosidase activity determined, and the peak activity expressed as a percentage of wild-type. Each value represents the mean ± SEM of at least three independent measurements. Promoter activity relative to that of the FL construct.
constructed a promoter fusion (1xIHF-VF; Table 3) in which the IHF IIIa and IHF IIIb sites are swapped. If IHF III is a composite site and if both sites contribute to regulation of acsP2 transcription, then I reasoned that the activity of the mutant promoter 1xIHF-VF would exhibit behavior that is similar, if not identical, to the WT 1xIHF promoter.

I constructed the 1xIHF-VF promoter (see Figure 23B) using a 45 bp primer encoding the 5’ – IHF IIIb – 3’ aligning it on the same DNA strand as the tandem CRP sites (CRP I and CRP II) and the RNAP binding site. Thus, in 1xIHF-VF, IHF IIIb is positioned in the exact original location of IHF IIIa. Simultaneously, IHF IIIa is placed onto the opposite DNA strand in the original location of IHF IIIb.

I recombined the mutant promoter fusion, 1xIHF-VF into the hybrid bacteriophage and introduced it as a monolysogen into AJW678 (Table 1). I grew the resulting monolysogen in TB at 37°C, harvested cells at approximately one hour intervals, and compared its β-galactosidase activity to that of monolysogens that carried the 1xIHF, 0xIHF, or FL promoter constructs (Fig. 23C).

As shown previously (Fig. 9C), the FL promoter fragment exhibited the most activity, while the 0xIHF promoter fragment displayed somewhat less activity. In contrast, the 1xIHF promoter fragment exhibited even less. Strikingly, the 1xIHF-VF promoter fragment displayed activity that most closely resembles the weak activity of the 1xIHF promoter. Although this result is consistent with the hypothesis that IHF III is a composite site, alternative explanations could be
envisioned. It is unlikely, however, that other mechanisms would require all of the elements shown to be required for IHF-mediated inhibition of transcription from the 1xIHF promoter fragment. I reasoned that the more properties that 1xIHF-VF shares with 1xIHF, the more likely that the same mechanism regulates their transcription. Thus, I tested whether the poor activity elicited by 1xIHF-VF required the same surfaces of CRP that influence the activity exhibited by 1xIHF.

Figure 24. 1xIHF and 1xIHF-VF activities involve AR2 and AR3 of CRP. β-galactosidase activity from promoter fragments 1xIHF and 1xIHF-VF fused to lacZYA. Each lysogen was deleted for crp and transformed with plasmids expressing wild-type CRP or mutant variants. Wild-type CRP possesses AR1 and AR2, but not AR3 and is designated as 1+2−3+. The mutants either lack AR1 (1−) or AR2 (2−) or possess the non-native AR3 (3+) or combinations of each. The resultant transformants were grown in TB, samples harvested are regular intervals, and the peak β-galactosidase activity determined and expressed relative to the mutant that lacks AR2 1+2−3+. Each value represents the mean ± SEM of at least three independent measurements.
Activity from the 1xIHF-VF and 1xIHF Promoters Involves the Same Activation Regions

As described previously, CRP possesses three surfaces that interact with RNAP: AR1, AR2, and AR3 mediate protein-protein interactions with the -CTD, the -NTD, and region 4.2 of the sigma subunit, respectively (Fig. 4). Transcription activation from the 1xIHF promoter fragment requires AR1; without it, activation does not take place, regardless of the status of AR2 and AR3. In contrast, AR2 is a key component of IHF-mediated inhibition; eliminating AR2 relieves inhibition. In the absence of AR2, the presence of AR3 can compensate while the presence of both AR2 and AR3 causes an even more severe inhibition of transcription (Sclavi et al., 2007).

Monolysogens of either the parental 1xIHF or the mutant 1xIHF-VF promoter fusion, defective for endogenous CRP (crp::kan), were transformed with plasmids carrying either WT CRP, the AR2 mutant, the AR2/AR3 mutant, the AR3 mutant, or the AR1 mutant (Table 1). The resultant transformants were aerated in TB at 37°C, the cells were harvested at approximately one hour intervals, and their β-galactosidase activity measured.

The behavior of the 1xIHF-VF promoter fragment (Fig. 24B) completely resembled that of its 1xIHF parent (Fig. 24A). Transformants of either promoter fragment that expressed WT CRP (AR1⁺ AR2⁺ AR3⁺) exhibited poor transcriptional activity, while those that expressed the AR2 mutant (AR1⁺ AR2⁻ AR3⁺) displayed substantially more activity and those that expressed the AR2/
AR3 double mutant (AR1⁺ AR2⁻ AR3⁺) exhibited activity similar to that of WT. In contrast, transformants of either promoter fragment that expressed the AR3 mutant (AR1⁺ AR2⁺ AR3⁺) exhibited activity about as low as that displayed by the AR1 mutant (AR1⁻ AR2⁺ AR3⁰). Thus, for both promoters, disruption of AR2 caused relief of inhibition in the context of both promoter fragments, the unmasking of AR3 on a CRP protein that retains AR1 compensated for the disruption of AR2, the presence of all three surfaces caused extremely low transcription that resembled that of the AR1 mutant.

Thus, the requirements for IHF-mediated inhibition are the same for 1xIHF and 1xIHF-VF. The simplest explanation is that *IHF III* is indeed a composite binding site consisting of two overlapping IHF sites.

**Discussion**

*IHF III is a composite site*

The data obtained from the systematic mutational analysis is consistent with the hypothesis that transcription depends on more than one *IHF III* site. For example, A149C disrupts the central IHF binding motif of *IHF IIIa*. This mutation, predicted to reduce or eliminate IHF binding to *IHF IIIa*, resulted in activity that was higher than that of the 0xIHF promoter fragment. This behavior might be explained in terms of enhanced binding of IHF to the opposing site *IHF IIIb*, which would function in a positive manner. This scenario is unlikely, however,
because disruption of the central motif of \textit{IHF IIIb} also resulted in enhanced promoter activity. Furthermore, these central motif mutations resulted in activity similar to that obtained when both sites were disrupted. I therefore proposed that both sites exist and that they both contribute to IHF-mediated inhibition of \textit{acs} transcription. My proposal was validated by the behavior of the 1xIHF-VF promoter construct that, in all tested characteristics, mirrored that of the parental 1xIHF promoter fragment.

The degree to which each mutation altered inhibition varied from mutation to mutation. A simple explanation for the different degrees of promoter activities may be attributed to the relative importance of certain bases in protein-DNA interactions between IHF and the two \textit{IHF III} binding sites. Mutations in single IHF sites show profound differences in protein binding and thus activity (Lee et al., 1991). This is not to suggest that this is true for all IHF sites. It is entirely possible that mutations made in overlapping IHF sites may be masked by the presence of an additional site. IHF binding to the weakened \textit{IHF III} sequence may be compensated by the second IHF site or by preserving the general DNA secondary structure or through sequence recognition. Additionally, the requirement for specific bases may vary from binding site to binding site and is probably dictated, to some degree, by upstream and downstream sequences as well as the proteins present at any given time. Taken together, however, the results presented in this chapter support our hypothesis that both \textit{IHF IIIa} and \textit{IHF IIIb} exist.
Both IHF IIIa and IHF IIIb Contribute to Inhibition

On the basis of these results, I also propose that both sites contribute to IHF-mediated inhibition. In support of this hypothesis, Dr. Cathy Lawson (Rutgers University) used in silico molecular modeling to determine the overall structure of an IHF/IHF IIIa- or IHF/IHF IIIb-mediated complex in the context of CRP dimers bound to the downstream CRP II and CRP I and the -CTD bound to the region just adjacent to CRP I (Fig. 25). The modeled output for both complexes shows almost identical DNA bends when IHF is placed on either IHF IIIa or IHF IIIb. Thus, when the downstream sites are occupied, IHF is predicted to bend the DNA in the same manner regardless of the IHF III site that it binds. Though the modeling program generated structures that are consistent with available crystallography data, these predictions were performed out of the context of the upstream sites (e.g. IHF II) and some of the downstream sequences (e.g. the -35 and -10 hexamers). Though IHF interaction with either site may result in similar bend angles, it should be noted that the DNA structure in vivo might prove to be very different when bound along side the proteins that bind these additional sites.

It is curious that both IHF III sites seem to mediate inhibition. Originally, we had predicted that the opposing IHF site, IHF IIIb, would counterbalance the negative affects of IHF IIIa in order for transcription to proceed. However, according to my data, both sites are required for inhibition – mutating either site resulted in promoter activity higher than that of the 1xIHF fragment and often as high as that of the 0xIHF construct. Based on these results, I conclude that
inhibition is not advanced by the reduction of binding affinity of either site. Perhaps IHF oscillates between IHF IIIa and IHF IIIb, interrupting the interaction between CRP and RNAP and thereby disrupting transcription. For transcription to occur, it is possible that proteins bound to upstream sequences insulate the transcription apparatus from the effects of the IHF/IHF III complex (see Chapter Five). Further investigation is warranted to identify the exact mechanism of IHF III-mediated inhibition by this complex site.

Figure 25. A folding model for IHF IIIa versus IHF IIIb in the context of 1xIHF. IHF (purple) bound to either IHF IIIb or IHF IIIa, CRP (blue) bound to CRP II and CRP I, and one α-CTD (green) bound to the region directly adjacent to CRP I. The DNA double helix is in white and cream. Note the similarity of the bends caused by the binding of IHF to IHF IIIb and to IHF IIIa and the lack of difference of effect on the downstream structure. Also note the lack of space between the CRP dimers for the binding of the other α-CTD (modeling performed by C. Lawson, Rutgers University).
CHAPTER FIVE

EVIDENCE THAT THE UPSTREAM SINGLE IHF II SITE
ANTAGONIZES IHF IIIa/b-MEDIATED INHIBITION

Introduction

In the previous chapter, I presented evidence that IHF III is a composite site, comprised of inverted dual-overlapping sites (denoted IHF IIIa and IHF IIIb). I proposed that when IHF binds to either IHF IIIa or IHF IIIb, it severely inhibits CRP-dependent transcription, at least in the context of the 1xIHF promoter fragment. This severe inhibition appears to require both sites through some currently unknown mechanism.

However, IHF-mediated inhibition is relieved by DNA sequences located upstream of IHF III (Browning et al., 2004; Sclavi et al., 2007). Since IHF is always present at moderately high levels in the cell (Azam, 2004), the low level of transcription exhibited by the 1xIHF promoter fragment is what we would expect from the FL promoter in vivo. Instead, the FL promoter exhibits considerably higher activity.

To explain this dichotomy, we propose that a cis element and/or a trans-
factor must antagonize IHF-dependent inhibition - otherwise acs transcription might never take place. Interestingly, the region encompassing both IHF III and a short sequence immediately upstream includes two sequences that may offer an explanation for how IHF-mediated inhibition is counteracted: a perfect palindrome and IHF II.

In this chapter, I will provide evidence that the palindrome does not appear to influence acs transcription. I will then present evidence that supports the hypothesis that IHF II can antagonize the negative effects of IHF III on acsP2 transcription and that occupancy of IHF II could guide IHF binding to IHF IIIb and, as a consequence, activate transcription from acsP2.

**The Palindrome does not Modulate IHF-Mediated Inhibition**

Experimental evidence verified the existence of a DNA site for IHF (IHF II) located immediately upstream of IHF III (Browning et al., 2004; Sclavi et al., 2007). Bioinformatic analysis also revealed the existence of a perfect palindrome that overlaps the junction between IHF III and IHF II.

To determine if either sequence influenced acs transcription, I constructed two promoter constructs: 1xIHF+5 and 2xIHF (Fig. 26A; Table 4). The former adds 5 bp to the 5’ end of the 1xIHF promoter fragment; thus, reconstituting the entire palindrome. The latter adds 25 bp; thus, reconstituting both the palindrome and IHF II.
Both promoter fragments were recombined into a hybrid bacteriophage and introduced as a monolysogen into AJW678 (Table 1). The resulting monolysogens were grown in TB at 37°C, cells were harvested at approximately one hour intervals, and their β-galactosidase activity compared to that of monolysogens that carried the FL, 1xIHF and 0xIHF promoter fragments.

As shown previously, the FL promoter exhibited the highest activity, the 1xIHF construct yielded the least activity, and the 0xIHF construct displayed an intermediate activity (Fig. 26B). Interestingly, the 1xIHF+5 promoter exhibited somewhat more activity than 1xIHF, while the 2xIHF promoter fragment yielded activity that resembled that of 1xIHF.

On the basis of these results, I hypothesized that the palindrome was an antagonist of IHF III-mediated inhibition. A palindromic sequence can function either as a binding site for a homodimeric transcription factor or as a region of DNA that can form a secondary structure, called a cruciform, that could influence protein binding or some aspect of DNA metabolism. If the increased transcription exhibited by 1xIHF+5 resulted from the presence of the palindrome, then disruption of one of its half sites would be predicted to prevent protein binding or efficient formation of cruciform DNA. Thus, we expected to observe reduced promoter activity to levels similar to that exhibited by the 1xIHF promoter. Because the downstream half-site overlaps IHF IIla and IHF IIlb, I scrambled the 5 bps of the upstream half-site. Unexpectedly, this construct (1xIHF+5S; Table 4) exhibited activity similar to that of the 1xIHF+5 promoter (data not shown).
Based on this result, I conclude that the palindrome does not affect *acs* transcription, either as a binding site for a protein homodimer or as a site for hairpin loop formation. I do not understand why the 5 bp adjacent to *IHF III* in the context of the 1xIHF promoter fragment enhances transcription; however, the negative result in response to such a massive alteration in sequence led me to end this line of investigation and instead focus on *IHF II*.

**Figure 26. Presence of the complete palindrome sequence increases transcription from 1xIHF+5 promoter.** (A) This schematic shows the *acs* promoter region from positions -379 to +65 relative to the *acsP2* transcription start site (+1). It shows the locations of each promoter, each CRP and IHF binding site and the extent of the promoter fragments. The bent arrows indicate the approximate location of each promoter and its direction of transcription. Inverted arrows indicate the locations of CRP I and CRP II, as determined by similarity to consensus. Hatched boxes designate the locations of IHF I-III, also determined by similarity to consensus. The orange rectangle indicates the approximate location of the palindrome. (B) β-galactosidase activity from promoter fragments 0xIHF, 1xIHF, 1xIHF+5, 2xIHF, or FL fused to *lacZYA*. The wild type strain AJW678 was lysogenized with the hybrid phages carrying the corresponding *acs* promoter fragment. The resultant strains were grown in TB, samples harvested at one-hour intervals, the β-galactosidase activity determined, and the peak activities compared. Each value represents the mean ± SEM of at least three
Evidence that *IHF II* is a Positive Modulator of *IHF III*-Mediated Inhibition

The weak promoter activity exhibited by the 2xIHF construct (Fig. 26B) is consistent with a model in which IHF bound at *IHF II* plays a minor role at best. Indeed, a mutation in its central motif, changing the sequence from TTTCAA to TTTGAA such that the binding of IHF to *IHF II* should be substantially diminished (Table 4), resulted in activity similar to that of the 1xIHF and 2xIHF promoter fragments (data not shown).

To determine if this weak promoter activity resulted from the same mechanism as observed in the context of the 1xIHF promoter construct, I tested the involvement of the activation surfaces of CRP. To do so, I used P1 transduction to introduce *crp* into the monolysogen that carried the 2xIHF promoter fragment. The resultant strain was then transformed with plasmids carrying various mutant CRP genes (Table 1). The resultant transformants were grown in TB at 37°C, the cells were harvested at approximately one hour intervals, and their β-galactosidase activities were compared. Like the 1xIHF promoter fragment, promoter activity from the 2xIHF construct required AR1 and was inhibited by AR2 and AR3 (data not shown). On the basis of these results, it would seem that *IHF II* plays no significant role in *acs* transcription; however, subsequent studies revealed that this conclusion was untrue.

In an effort to thoroughly examine the role of *IHF II* in the context of *IHF III*-mediated inhibition, I disrupted *CRP II* and *IHF IIIa* in the context of the 2xIHF
construct. The resulting mutant promoter constructs (denoted 2xIHF-IHF3m and 2xIHF-C2m; Table 4) were recombined into the hybrid bacteriophage and introduced as monolysogens and compared to that of monolysogens that carried the FL, 2xIHF, 1xIHF and 0xIHF promoter constructs. The controls all behaved as expected based on the results of previous experiments. The mutant 2xIHF-C2m exhibited activity similar to that of 0xIHF (Fig. 27), a result that also suggests that the weak promoter activity of 2xIHF results from a mechanism similar or identical to the one that maintains low transcription from 1xIHF. Intriguingly, however, the activity from the 2xIHF-I3m promoter fragment was substantially higher than both the 0xIHF and full-length promoter constructs. On the basis of this result, we proposed that IHF II can play a positive role in the regulation of acs transcription and that IHF IIIa might antagonize this effect.

Using EMSA and DNase I footprint analyses, Drs. Douglas Browning and Bianca Sclavi obtained evidence that IHF binds IHF III with higher affinity than it does IHF II (Browning, Sclavi, and Wolfe, unpublished data). Thus, it is possible that occupation of the higher affinity IHF III might interfere with occupation of the lower affinity IHF II. Armed with this information, I used genetics to direct IHF/DNA interaction to either IHF II or IHF III by simply strengthening or weakening IHF II (i.e. mutating the native sequence to bring it closer to or further from consensus). The promoter fragments, 2xIHF-I2wk (IHF II weakened) and 2xIHF-I2st (IHF II strengthened), were cloned into pRW50, a low copy plasmid that permits more rapid analysis of promoter-lacZ fusions. Similar constructs were
built for FL, 1xIHF and 2xIHF (Table 4). The resultant plasmids were introduced into AJW678 by transformation, the resultant transformants were grown in TB at 37°C, cells were harvested at approximately one hour intervals, and their β-galactosidase activity compared.

As expected, the FL promoter displayed high promoter activity whereas 1xIHF was low (Fig. 28C). However, 2xIHF showed an unusually high level of transcription unlike that observed when 2xIHF was assayed as a monolysogen (Fig. 26). Thus, despite the previously described evidence to the contrary (in the phage system), the 1xIHF and 2xIHF promoter constructs did not behave identically (in the plasmid system). The simplest explanation is that the presence of IHF II can influence acs transcription, at least when the promoter is located on a plasmid. But why does the 2xIHF promoter fragment exhibit vastly different behaviors when carried by the two different reporter systems? The different behavior could be attributed to the difference in copy number. The plasmid pRW50 is present in about ~10 copies per cell (Lodge et al, 1992), whereas the monolysogen is by definition present in single copy. Another possibility involves differences in topology. Topological state can impact local gene expression either promoting transcription or inhibiting it. DNA supercoiling tends to promote protein binding and the binding of those proteins can influence transcription (Niehaus et al, 2008). Plasmids tend to be more supercoiled than the chromosome. Because the prophage that carries the acs::lacZ fusion integrated into the chromosome, the fusion is expected to take on the degree of supercoiling characteristic of the
loop-domain into which it has inserted.

To determine if the plasmid system is a valid tool for assessing the behavior of the acs promoter, I disrupted \textit{IHF III} in the context of the 1xIHF construct (Fig. 28B), introducing the C152G C150G and C153A C153A double mutations used in Chapter 4 to assess the behavior of the 1xIHF monolysogen (Table 4). For both mutations (here denoted I3wk and I3st), I observed the same behavior, i.e. alleviation of inhibition (Fig. 28C). Thus, the plasmid system appears to be valid.

Next, I introduced two different mutations into the TTR motif of \textit{IHF II} in the 2xIHF promoter construct (Fig. 28B). By substituting CC for the native AT, the first mutation (I2wk) should reduce affinity. By substituting a T for the native A, the second mutation (I2st) should increase affinity. In the context of 2xIHF, weakening of \textit{IHF II} resulted in transcription similar to that obtained with the 1xIHF promoter fragment, while strengthening of \textit{IHF II} resulted in a slight increase in promoter activity compared to that exhibited by the 2xIHF promoter.

Taken together, these data are consistent with the hypothesis that \textit{IHF II} can behave as an \textit{IHF III} antagonist. Thus, we now possess genetic evidence that each site can antagonize the other.

Because they are predicted to sit on the same face of the DNA helix, \textit{IHF II} and \textit{IHF IIIa} are unlikely to be bound by IHF simultaneously. In contrast, since \textit{IHF II} and \textit{IHF IIIb} are predicted to sit on opposing faces, it is possible for both of these sites to be occupied simultaneously (Fig. 29). This led me to re-investigate
the influence of *IHF IIIa* and *IHF IIIb* on *acs* transcription.

**Individual Analysis of *IHF IIIa* and *IHF IIIb***

I chose to perform this re-investigation in the plasmid context because it seems to enhance the effect of *IHF II*. To reduce the complexity of the analysis, I deleted the composite *IHF III* site and instead replaced it with two single IHF sites (Fig. 30A): one construct had the DNA site for IHF in the position of *IHF IIIa* (denoted 1xIHF-2R3a; Table 4) and one had the site in the position of *IHF IIIb* (denoted 1xIHF-2R3b). In this manner, I could determine how the individual sites might contribute to *acs* regulation. For the single site, I chose the confirmed
Intriguingly, 1xIHF-2R3a and 1xIHF-2R3b had opposite effects on their promoter activities (Fig. 30B). 1xIHF-2R3a resulted in extremely low activity. In contrast, 1xIHF-2R3b displayed much higher activity, even more than that exhibited by the 0xIHF promoter. Thus, it appears that IHF bound in the IHF IIIa position can exert a negative effect on acs transcription and that IHF bound in the IHF IIIb position can exert a positive effect.

**Figure 28. IHF II is a positive modulator of acs transcription.** (A) Alignment of the IHF consensus, IHF IIIa, IHF IIIb, and IHF II. The consensus nucleotides are in bold blue. (B) The native arrangement of IHF II and IHF IIIa. The consensus nucleotides are in bold blue. Mutations above the sequence improve consensus, while the mutations below the sequence diminish consensus. (C) β-galactosidase activity from promoter fragments FL, 1xIHF, 1xIHFwk, 1xIHFst, 2xIHF, 2xIHFwk, or 2xIHFst fused to lacZYA. The fusions were carried by the low copy reporter plasmid pRW50. The resultant constructs were transformed into the wild type strain AJW678 single site IHF II.
Notably, when carried by the phage system, the same constructs both displayed higher promoter activity relative to that exhibited by 1xIHF (data not shown). Such a result might be due to a less topologically constrained DNA template, which could partially negate the effects of IHF IIIa.

Figure 29. Proposed model of simultaneous occupation of IHF III and IHF II. (A) Diagram of 2xIHF promoter fragment. Relative positions of IHF, CRP and Fis binding half-site positions are indicated above or below their respective sites. IHF III and IHF II are contiguous sites predicted to access the minor groove of the top DNA strand. (B) Proposed space-filling model of the DNA bend induced by the simultaneous occupation of IHF III and IHF II in the absence of CRP and RNAP. α and β subunits are depicted in teal and purple, respectively. Red and yellow DNA strands represent the coding strand for IHF III and IHF II, respectively. (Adapted from D. Browning; personal communication)
Discussion

A Proposed Model

We propose that the occupancy status of IHF II plays a major role in dictating whether IHF binds to IHF IIIa or to IHF IIIb and thus whether transcription from acsP2 becomes activated or becomes stalled by the mechanism described in Chapter 3.

This hypothesis is supported by the following observations: (1) in the context of the phage system, a mutation (C152G C150G, also known as I3m or I3wk) that reduces the consensus of IHF IIIa and enhances the consensus of IHF IIIb resulted in activity from the 2xIHF promoter construct greater than that exhibited by the 0xIHF or FL constructs (Fig. 27); (2) also in the context of the phage system, the 2xIHF promoter exhibited 1xIHF-like activity (Fig. 26B) that was unaltered by a mutation that reduces IHF II consensus (data not shown); (3) like the poor activity of 1xIHF, this poor activity of 2xIHF depends on AR2 (data not shown) and CRP II (Fig. 27); (4) in the context of the plasmid system, 2xIHF exhibited high activity, while 1xIHF exhibited weak activity (Fig. 28C); (5) also in the plasmid system, the I2wk mutation, which is predicted to decrease the affinity of IHF II for IHF, resulted in low levels of transcription resembling that of the 1xIHF construct (Fig. 28C); (6) also in the plasmid system, the placement of a simple site into the IHF IIIa position resulted in poor promoter activity from the 1xIHF construct, while placement of that same simple site into the IHF IIIb
position resulted in strong activity (Fig. 30B); and finally (7) IHF II and IHF IIIa are predicted to sit on the same face of the DNA helix, while IHF II and IHF IIIb are predicted to site on opposite faces (Fig. 29).

**Figure 30. IHF bound at the IHF IIIa or IHF IIIb positions can be a negative or positive effector, respectively.** β-galactosidase activity from promoter fragments 0xIHF, 1xIHF, 1xIHF-A164T/T142A, 2R3a, 2R3a-rev, or 2R3b fused to lacZYA. The fusions were carried by the low copy reporter plasmid pRW50. The resultant constructs were transformed into the wild type strain AJW678. The resultant strains were grown in TB, samples harvested at one-hour intervals, the β-galactosidase activity determined, and the peak activities compared. Each value represents the mean ± SEM of at least three independent measurements. 2R3a, 2R3a-rev, and 2R3b are described in the text.

IHF III *Competes with IHF II*

In the context of 2xIHF, the weakening of IHF IIIa coupled with the strengthening of IHF IIIb (C152G C150G) resulted in _acs_ transcription that was far greater than that of the 0xIHF construct (Fig. 27). This observation suggests that IHF II can play a positive role. It also supports the hypothesis that occupation
of *IHF IIIa* can inhibit the positive function of *IHF II*, at least when the promoter fusion is integrated into the chromosome.

When both IHF sites are available (as in the 2xIHF construct), perhaps IHF prefers to bind to *IHF IIIa* because of a more favorable consensus site. This preference would have to be due to specific nucleotide composition rather than divergence from consensus, because both *IHF II* and *IHF IIIa* have the same number of differences from consensus.

Alternatively, the preference for *IHF IIIa* may rely on the pre-existing DNA structure induced by the CRP/RNAP complex (Sclavi et al., 2007). CRP binding within a promoter region bends DNA up to 45 angles per homodimer. RNAP also induces DNA bending when bound to DNA. Together, the total DNA bend angle may exceed 90. Studies on the effects of DNA curvature (e.g. as AT-rich stretches) and DNA binding proteins (specifically the nucleoid protein IHF and its paralog HU) show that these proteins have a greater affinity for curved DNA (Bonnefoy and Rouvière-Yaniv, 1991). Therefore, the affinity for *IHF IIIa* may not be as dependent on sequence as it is on the proximity of *IHF IIIa* in relation to the CRP/RNAP complex.

In either case, the weak activity of the 2xIHF promoter may be the result of the inability of IHF to occupy both *IHF II* and *IHF IIIa* simultaneously. With a preference for *IHF IIIa*, IHF would mediate inhibition of CRP-dependent transcription. And, yet, DNase I analyses show that IHF can occupy both *IHF II*
and *IHF III* simultaneously (Browning et al., 2004; Sclavi et al., 2007). Because IHF should be unable to simultaneously bind *IHF II* and *IHF IIIa*, I propose that simultaneous binding occurs at *IHF II* and *IHF IIIb*.

**Why does the 2xIHF construct behave differently in the plasmid and phage systems?**

In the context of the phage reporter system, the 2xIHF promoter exhibits about as much activity as the 1xIHF promoter (Fig. 26B). Thus, when integrated into the chromosome, the natural IHF *II* site cannot overcome the inhibitory effects of *IHF IIIa*. In contrast, in the context of the plasmid system, the 2xIHF construct yielded activity that resembled that of the 0xIHF construct. As described above, the two obvious reasons for this dichotomy is copy number and superhelicity.

The copy number difference is about one order of magnitude. If some *trans* factor that helped determine whether occupancy of *IHF II* was favored was limiting, then the increase in copy number could have a significant impact on *acs* transcription. The simplest possibility is IHF itself. Although IHF is synthesized in large amounts, its DNA sites are numerous. Thus, titration could influence the occupancy of the three binding sites in the 2xIHF construct: *IHF II*, *IHF IIIa*, and *IHF IIIb*. Titration of IHF would likely favor the higher affinity *IHF III* site over the lower affinity *IHF II*. Thus, for titration to influence *acs* transcription, it would have
to favor occupancy of IHF IIIb over that of IHF IIIa.

Alternatively, the key factor might be a difference in the superhelicity of the plasmid relative to that of the integrated phage. This is important because the degree of superhelicity affects the binding of proteins and the binding of those proteins can influence transcription (Niehaus et al, 2008).

In either case, the plasmid system allowed me to demonstrate that IHF II can antagonize IHF III-mediated inhibition and might explain how (in the FL construct) the upstream sequences counteract IHF IIIa-mediated inhibition. In this model, the binding of a protein upstream would somehow help IHF to bind to IHF II. As a consequence, IHF would bind to IHF IIIb instead of IHF IIIa.

**Does IHF IIIb Function as a Positive Modulator of acs Transcription?**

If this model is correct, then occupancy of IHF IIIb must have a positive effect on acs transcription. In the context of the phage system, this is clearly not true: mutations predicted to favor binding to IHF IIIb did not enhance transcription from the 1xIHF construct. In contrast, such a mutation (C152G C150G or I3m) in the 2xIHF construct resulted in the highest transcription observed (Fig. 27). This led me to ask if IHF IIIb could favor acs transcription. In fact, in the plasmid system, it did. Placement of a simple IHF site (IHF II) in the location of IHF IIIb resulted in strong promoter activity in the 1xIHF construct (1xIHF-2R3b).
construct, placement of the same site in the position of *IHF IIIa* resulted in very weak activity (1xIHF-2R3a) (Fig. 30).
CHAPTER SIX

DISCUSSION

Prior Studies

To survive a variety of environmental conditions, a bacterium must adjust swiftly and efficiently to signals it perceives concerning its milieu (Roszak and Colwell, 1987) (Kussell et al., 2005). To adapt to harsh conditions, bacteria up- or downregulate genes and/or operons that are necessary for survival. Therefore for immediate reaction some transcription factors with global activities responsible for coordinating the response preexist in the cytoplasm eliminating some time required for transcription factor synthesis (Zhou et al., 2008) (Kort et al., 2008) (Kussell et al., 2005). The expression of proteins, such as acs during carbon starvation, that respond to the specific signal/stress are upregulated to help the bacterium adapt to the altered milieu. This network of effectors and sensors are precisely coordinated to reduce or eliminate any inefficiency (Slauch et al., 1997) (Khmel, I. A., 2005). Removal drastically alters the cells survivability under normal conditions and spell disaster in challenging ones.

Due to the large variety of environmental stresses encountered by a bacterium, many genes are needed to survive. However, the volume of the
absolutely required. To conserve space, bacterial genes are often organized as operons, with many genes or operons overlapping, while others are encoded on opposing DNA strands.

Much like this dense organization of genes on the chromosome, promoter elements also maximize the amount of regulatory information for precise gene expression. Missteps in transcription can lead to cellular death during times of carbon limitation. At the most basic promoter, DNA sequence information directs RNAP to bind to the -35 and -10 sequences, hallmarks of any promoter. At these “simple” promoters, the stability of RNAP binding does not require a transcription factor to stably bind the promoter region. However, not all promoters allow for stable RNAP binding. To more stably anchor RNAP to the DNA some promoters have an additional cis element located just upstream of the -35 hexamer, called an UP element. The UP element is an A-T rich sequence that is recognized and bound by the RNAP α-CTD, thereby providing a more stable RNAP/DNA interaction. Not only does promoter sequence provide a means for RNAP recognition for higher affinity binding but also DNA structure. The UP element and -10 hexamer (TATAAT) are A-T rich regions that are prone to creating intrinsic DNA bends. This sequence-induced DNA secondary structure creates an environment that is more conducive for DNA recognition and protein docking to the promoter region*. However, the location of an UP element may exist 100 bp further upstream than ones found proximal to the -35. To activate and enhance transcription at these promoters, a transcription factor solely utilized for DNA
remodeling is required to bring this distant UP element into close proximity to the α-CTD. Despite the straightforward regulation of simple promoters, the protein/DNA dynamics at the majority of bacterial promoters involve more than one activator, more than one inhibitor, and many relevant DNA sequences within the same region of DNA.

The majority of bacterial promoters have many multiple relevant sequences where cis and trans factors contribute to regulating transcription. Due to the various DNA sequences, and the proteins that bind to them, within complex bacterial promoters the dynamics between transcriptionally relevant inputs complicates our understanding of the mechanisms by which transcription occurs. This limits our ability to utilize this knowledge for applications ranging from treating disease to enhancing biotechnology. At a complex promoter, multiple transcription factors bind to multiple sites to either activate or inhibit transcription. The nucleotide sequence arrangement of transcription factor and RNAP binding sites of complex promoters have evolved to optimally position DNA binding sites to ensure precise regulation of a given gene or operon. Deciphering the code of gene regulation is not a straightforward process because the positioning of DNA binding sites makes it rather complex to understand the precise role of inhibitors, activators, RNAP, and the contribution of promoter sequence and structure to regulation. Thus, to deconstruct the operating mechanisms of a given promoter region, we must study the individual and group contributions of any identified, putatively relevant sequence. My studies revolved
around the isolation of individual nucleotide elements and then introducing point mutations to each site within the whole to measure the transcriptional output of a reporter construct.

The *acs-nrf* intergenic region, our model promoter, was an ideal choice to study regulation of a complex promoter for many reasons. Among them are: 1) the *acs* operon and its protein products are not essential under our assay conditions and media (37°C with aeration in tryptone broth, or TB); therefore, reduced or disabled gene expression does not interfere with promoter analysis; 2) the complexity of the promoter region is sufficiently crowded – the *acs* promoter contains four RNAP binding sites, *Pnrf*, *acsP1*, *acsP2A*, and *acsP2*, two Fis binding sites, two CRP site, three IHF sites, and a putatative ArcA site whose function will not be discussed further. The protein binding sites identified as acting upon *acs* do not include the transcription factors NarL, FNR, and NarP that regulate transcription of the opposing, divergent *nrf* promoter. Interestingly, a few of the protein binding sites shared by both regulatory regions affect transcription of each operon, specifically those for the nucleoid protein IHF (Fig. 6) (reviewed by Wolfe, 2005; Wolfe, 2009).

The functions of these promoters and binding sites were identified through the use of nested 5′ truncations with the intent to identify those DNA sequences that most affected transcription from the major *acs* promoter, *acsP2*. To ensure that no relevant sequences were missed, Beatty and Wolfe designed a 444 bp ‘full-length’ promoter (FL, *acs444*) that included the entire *acs-nrf* intergenic
region and its flanking sequences (from +101 bases downstream of the \textit{nrf} transcription start site (+1) to +76 bases downstream of the \textit{acsP2 +1}). The merit of this approach became apparent when 3'-truncations (lacking the 76 bases downstream of the \textit{acsP2 +1}) showed noticeably different promoter behaviors than their more full-length variants (Thach, Beatty and Wolfe, unpublished data). Although this phenomenon supports the hypothesis that the downstream sequences contribute to the regulation of \textit{acs} transcription, the mechanism remains unknown and dissection of that mechanism remained outside the scope of this dissertation. To evaluate the function(s) of specific promoters and sites, these researchers also chose to mutate specific DNA sites rather than deleting the gene that encodes the transcription factor or NAP predicted or known to bind to that sequence. Beatty and Wolfe reasoned that the complete absence of a global regulator might result in formation of a non-native nucleoid and thus increase the probability that the resulting data would reflect a non-native mechanism. They also reasoned that deletion of a global regulator with multiple functions at a given promoter would reflect only one of those functions. The merit of this approach was demonstrated by the behavior of a \textit{himA} mutant, which exhibited substantially reduced \textit{acs} transcription despite the clear indication that its gene product, IHF, functions primarily as an inhibitor (Beatty and Wolfe, unpublished data). I chose to follow their lead, working with and constructing additional 5' nested truncations of the FL promoter and introducing specific and, wherever possible, well-characterized point mutations into those truncations.
Coupling the point mutation strategy with conventional in vitro transcription assays, Wolfe and co-workers demonstrated that acsP2 functioned as the primary promoter. They further showed that the ability of acsP2 to transcribe depends on the ability of CRP to bind to a high-affinity promoter-proximal site (CRP I) from where one of its surfaces (AR1) must make physical contact with a surface (determinant 287) of the α-CTD of RNAP (Beatty et al., 2003). Moreover, they showed that modulation of this CRP-dependent activation involves the binding of multiple proteins (CRP, IHF, and Fis) to multiple binding sites located 5’ to CRP I. For example, they showed that the binding of CRP to a low-affinity promoter-distal site (CRP II) enhanced transcription about two-fold (Beatty et al., 2003). They also demonstrated that Fis inhibits CRP-dependent transcription in part by binding to Fis III, which overlaps CRP I. Thus, the binding of Fis to FIS III hinders the required binding of CRP to CRP I. Since Fis interferes with the activator CRP and not RNAP itself, the inhibitory process is termed anti-activation (Browning et al., 2004). In contrast, the mechanism by which the binding of Fis to FIS II remained unclear.

Wedding the truncation strategy to in vitro transcription assays, Wolfe and co-workers further identified a 31-bp sequence located just upstream of CRP II that exerted a strong inhibitory effect on the CRP-dependent activation of acs transcription. They further showed that sequences located further upstream alleviated this inhibition. On the basis of their observation that the 31-bp sequence bound IHF, they proposed that the inhibition resulted from the binding
of IHF to this 31-bp sequence (Browning et al., 2004). It was the purpose of my dissertation to test this hypothesis and to dissect the underlying mechanism.

### IHF Mediates Formation of a Stalled Open Complex

First, I verified the genetic observations that led to this hypothesis. As reported previously (Browning et al., 2004), the FL construct displayed the highest promoter activity and the 0xIHF (acs205) promoter fragment exhibited about two-thirds the FL activity. In contrast, the 1xIHF (acs236) promoter fragment achieved very low activity, reaching levels less than one-fourth of 0xIHF and one-sixth that of FL (Fig. 9C). Thus, my co-authors and I concluded that the 31-bp sequence (denoted \textit{IHF III}) exerted a negative influence on acsP2 transcription.

I extended these observations by introducing individual mutations into the \textit{CRP I, CRP II, and IHF III} sites of the 1xIHF promoter fragment (Fig. 9). As expected, introduction into \textit{CRP I} of a mutation predicted to severely decrease its affinity for CRP essentially eliminated \textit{acs} transcription. In contrast, introduction into \textit{CRP II} of a similar mutation gave us a most peculiar result. This promoter, denoted 1xIHF-C2m, exhibited promoter activity similar to that of the 0xIHF promoter fragment. This result was consistent with an inhibitory role for \textit{CRP II}, a conclusion diametrically opposed to that obtained from the introduction of the identical mutation into the FL promoter (Beatty et al., 2003). To verify that \textit{CRP II}
could enhance CRP I-dependent transcription, I tested the promoter fragment 1xCRP (acs155), whose transcription is controlled by the single CRP I site. As expected if CRP II can function as a positive element, 1xCRP exhibited about half the activity of 0xIHF, which carries both CRP sites. We concluded that CRP II could function either as a positive or negative element and that the mechanism underlying IHF-mediated inhibition depends on the status of CRP II (Sclavi et al. 2007).

I next tested a mutant 1xIHF promoter fragment (1xIHF-I3m) in which a G was substituted for the native conserved C within the central WATCAA motif. This mutant fragment exhibited activity intermediate to 1xIHF and 0xIHF, suggesting that the 31-bp-dependent inhibition depends on IHF, as previously proposed (Fig. 9) (Browning et al., 2004; Sclavi et al., 2007). IHF is a protein demonstrated to affect DNA structure by inducing an IHF-DNA bend angle reaching up to 180°. Though hard biochemical evidence measuring the bend angle induced by IHF at the acs promoter in the presence of RNAP and CRP has not been reported here, we do provide genetic evidence to support the requirement of IHF-mediated DNA remodeling. This was a surprising result, as disruption of CRP II, in the context of 1xIHF, resulted in a transcription profile similar to 0xIHF. This result suggested that CRP II was more important for inhibition than IHF III. Yet, as I explained above, CRP II functions as a positive effector in the absence of the 31-bp sequence that includes IHF III. Thus, my co-authors and I concluded that IHF-mediated inhibition required the combined function of IHF III and CRP II. We
proposed that together these two sites determine whether CRP bound to $CRP_I$ can activate transcription by RNAP bound at the acsP2 promoter (Sclavi et al., 2007).

To gain insight into the mechanism by which this IHF-CRP tandem inhibits transcription, I identified the surfaces of CRP required for or involved in IHF-mediated inhibition of CRP-dependent acs transcription. By deleting the crp gene and complementing with either WT CRP or mutants of the surfaces known to interact with RNAP, I learned that IHF-mediated inhibition requires AR2 (Fig. 10). Interestingly, the non-native AR3 can compensate in the absence of AR2 (Sclavi et al., 2007). Armed with the results of my genetic analyses and plasmid constructs that I specifically constructed for in vitro analyses, my collaborator (Dr. Bianca Sclavi) performed DNase I and potassium permanganate footprint analyses. On the basis of her studies, we concluded that IHF mediates the formation of a complex that also includes RNAP and both CRP homodimers. This complex, we proposed, entraps RNAP in an unproductive open complex (Fig. 16). The specific role of AR2 remains unknown. Since this unproductive complex includes two CRP homodimers and since each subunit of the CRP dimer includes its own AR2 surface, there exist four distinct AR2 surfaces within the stalled complex. To identify the required AR2 surface(s) would require the use of oriented heterodimer analysis (Zhou et al., 1993), a genetic ‘trick’ that allows the investigator to direct a mutant surface to a specific binding site and to orient that surface either upstream or downstream.
How does IHF mediate the formation of this stalled complex? I propose that this is all made possible by the conformational and torsional strain induced on DNA by IHF when bound to IHF III. Upon binding to DNA, IHF inserts proline “fingers” into the minor groove of DNA. This insertion helps induce a bend but does not change the superhelicity of the immediate DNA strands of the IHF binding site. In contrast, DNA superhelicity is changed upstream (positive) and downstream (negative) (Aeling et al, 2006; Swinger and Rice, 2004; Swinger et al, 2003). The introduction of negative supercoiling in the region where CRP and RNAP bind may cause the CRP dimers to come into close contact with each other and RNAP. The additional protein-protein interaction then stabilizes the complex, thwarting promoter escape. Our genetic and footprint analyses support this conclusion, but analysis of the actual DNA structure of the inhibition complex remained beyond the scope of this thesis.

Countering IHF-Mediated Inhibition

The conclusions drawn from the study described in Chapter 3 (Sclavi et al., 2007) are limited to the 236-bp promoter fragment, 1xIHF. Although both genetic and biochemical analyses also were performed on the FL promoter, the differences in the architecture and behavior are substantial. The FL promoter measures 444 base pairs long and includes many additional protein binding sites that are not only relevant to acs regulation but also to the divergent promoter,
Pnrf. Furthermore, this construct exhibits six times more promoter activity than does 1xIHF (Fig. 9). This behavior is most interesting. Clearly, the upstream sequences somehow counterbalance the inhibitory effect caused by the binding of IHF to IHF III.

In Chapter 3, I proposed that the countering mechanism might revolve around the occupancy of CRP II (Fig. 16). If IHF III remains occupied during exponential growth, then the occupancy of CRP II becomes the critical factor. During mid-exponential growth, when cAMP begins to accumulate, the higher affinity CRP I can become occupied, but the lower affinity CRP II remains unoccupied, the CRP dimer bound at CRP I would be free to activate acs transcription by a typical Class I, AR1-dependent mechanism. As the concentration of cAMP increases, however, CRP II can become occupied and RNAP becomes stalled in an AR2-dependent mechanism.

Another intriguing possibility that is not mutually exclusive with the mechanism proposed in the previous paragraph is titration. IHF is a global regulator of many DNA-related events of which transcription is only one. Although IHF is made in moderately large amounts during exponential growth, the pool is limited. The chromosome possesses a large number of DNA sites for IHF. Perhaps the combination of limited IHF molecules and excess DNA sites titrates IHF away from IHF III, permitting some acs transcription even when both CRP sites are filled.
Together, these and other mechanisms could explain the expression profile of acs. Under the conditions tested (37°C in tryptone broth with aeration), acs transcription remains low during early exponential growth, gradually increases during mid-exponential growth, peaks upon the transition into stationary phase, and drops markedly as the culture progresses further into stationary phase (Kumari et al, 2000).

**Inhibition during early exponential growth.** During early exponential growth, acs transcription remains low most likely through a lack of CRP-dependent activation and Fis-dependent anti-activation. As described at the end of Chapter 3, if IHF were also bound to IHF III, the Fis-dependent anti-activation would be even stronger.

**Activation during mid-exponential growth.** As the favored carbon sources become depleted, cAMP becomes more prevalent and Fis becomes less so, shifting the balance of acs-bound proteins from Fis to CRP. Even if IHF is bound at IHF III, there should only be enough CRP to bind the higher affinity CRP I site and thus transcription would ensue.

**Inhibition following entry into stationary phase.** As the culture enters into stationary phase, the concentration of cAMP continues to increase and the concentration of IHF approaches its maximum, while Fis becomes undetectable. Under these conditions, the lower affinity CRP II site could become occupied simultaneously with IHF III. The binding of both RNAP and CRP should induce a
DNA bend. Since IHF targets bent DNA the secondary structure induced by the binding of both RNAP and CRP should facilitate the binding of IHF (Swinger and Rice, 2007; Teter et al, 2000).

If this were indeed the mechanism that underlies IHF-mediated inhibition of CRP-dependent acs transcription, then the 1xIHF promoter would follow a similar pattern of gene expression, albeit at a lower level due to the absence of upstream sequences required for the maximal transcription as seen with the full-length (FL) promoter. Because expression levels from 1xIHF reaches only 13-15% of FL during the course of an 8-10 hr experiment, we conclude that IHF-mediated regulation is not solely dependent upon the number of IHF molecules available for binding. If IHF-mediated inhibition of the 1xIHF promoter depended on only the number of IHF molecules present, then we should observe transcription to be higher that the reported level (13-15% of FL) when IHF levels are relatively low during log phase. A precipitous drop in transcription should occur when the cellular concentration of IHF reaches its peak. This led us to ponder how the inhibitory effect of IHF at acs is regulated.

**IHF III is a composite site**

Mutational analysis of the 31-bp sequence that includes *IHF III* restored 1xIHF promoter activity to approximately 50% of that exhibited by the 0xIHF promoter fragment. If the binding of IHF to *IHF III* was the underlying reason for
the 31-bp-dependent inhibition, we expected this mutation to restore transcription
to a level similar, if not identical, to that of the 0xIHF promoter fragment (Beatty et
al, 2004). The puzzling results provided me with a new focus: if the 31-bp
sequence includes a single 27-bp IHF III site, then why does a replacement
mutation predicted to severely decrease IHF binding fail to completely restore
transcription to the level exhibited by the 0xIHF promoter fragment?

As mentioned earlier, the intergenic region between the acs and nrf
operons share protein binding sites that regulate their expression, albeit under
different physiological conditions. acs transcription occurs under aerobic
conditions as preferred carbon sources diminish. In contrast, nrf expression is
activated under anaerobic, nitrogen limiting conditions (Wolfe, 2005). However, a
key factor in transcription at both promoters is IHF III – activating transcription at
nrf and inhibiting acs expression (Browning et al, 2004; Browning et al., 2006).
Initially, it was unclear how a single IHF III site could have contradictory functions
at divergent promoters. We imagined that the function of IHF at both promoters
was dependent upon the specific conditions under which transcription takes
place from both promoters, i.e. anaerobic in the presence of RNAP and FNR at
Pnrf and aerobic in the presence of CRP with RNAP at Pacs. But how does an
IHF binding site 127 base pairs upstream of the nrf +1 activate transcription from
one strand of DNA, while functioning as an inhibitor of acs expression on the
other DNA strand?
Theoretically, *IHF III* could exist as a single site and function as both an activator of transcription and inhibitor because of its ability to cause topological changes. IHF binding not only causes DNA bending and folding, but it also can cause changes in DNA supercoiling both upstream and downstream of its binding site and, as a consequence, affect gene expression at either long or short distances from that site. For example, at the $\sigma^{54}$ *Pu* promoter of *Pseudomonas putida*, IHF binding to an upstream region spanning -56 to -86 is crucial for activating transcription. The folding of DNA brings an enhancer region (surrounding position -104) into contact with the $\alpha$-CTD of RNAP (Macchi et al 2003). Not surprisingly, the direction of DNA bending is crucial to activation as insertion of an additional five bp, which places the IHF binding site on the opposite face of the DNA helix, causes a loss of transcription.

Our collaborators at the University of Birmingham (UK) have demonstrated that their predicted *IHF III* (denoted as *IHF IIIb*) sequence is physiologically relevant to *nrf* regulation (Browning et al., 2006), while we have published reports on the negative affects of our predicted *IHF III* sequence (referred to as *IHF IIIa*) on *acs* transcription (Browning et al., 2004; Sclavi et al., 2007). When we analyzed the sequence of both predicted sites, we concluded that they both could be real, as they each possess strong similarities to the IHF consensus site (WATCAANNNNTTR) with the same number of bases diverging from consensus. Furthermore, they overlap and are offset by two bases.
van Rijn and co-workers first reported the presence of two transcriptionally relevant, overlapping and inverted IHF sites (1991). The early promoter (Pe) and the repressor promoter (Pc) of bacteriophage Mu share the same intergenic region. These convergent promoters are both activated by IHF. These dual inverted overlapping IHF binding sites (designated ihfa and ihfb) are offset by three bases and are the only reported activators of transcription in this intergenic region. Thus, the mechanism of transcriptional activation appears to be solely dependent on the conformation of DNA induced by IHF.

Like van Rijn et al., we tested the hypothesis that IHF III is a composite site containing an IHF IIIa site relevant to acs transcription and an IHF IIIb site that controls nrf transcription. To determine whether two overlapping IHF binding sites existed and whether either or both sites were relevant to acs promoter regulation, I introduced single base pair mutations into the conserved motifs in accordance to results reported by Lee et al, (1991) and Hales et al (1994). Coincidentally, IHF IIIa and IHF IIIb sequences are offset by two base pairs. Each of their 3’ TTR motifs sits opposite the opposing dA-dT region. Whereas mutations in the dA-dT have been reported to cause weak to moderate defects in IHF binding (Lee et al, 1994; Hales et al, 1996; Fyfe and Davies, 1998), mutations in the TTR have been shown to cause severe defects in IHF binding. For example, the substitution of the second T with an A disrupts required hydrogen bonding between E44, R42, and R46 of the β subunit of IHF, resulting in a severe decrease of IHF/DNA interaction (Granston and Nash, 1993;
Mengeritsky et al., 1993). Using this information as my guide, I inserted mutations into the TTR and coincidentally into the opposing dA-dT region. Because of the overlap, these substitutions were predicted to reduce IHF binding to both sites and, as expected, they caused moderate to severe loss of IHF-mediated inhibition. I next inserted mutations into the central WATCAA motifs of either site. Since these mutations were predicted to disrupt binding to one site at a time, I used them to determine whether either or both sites existed. On this basis, I concluded that both sites likely existed and that they both exerted an inhibitory effect upon acs transcription.

To test this hypothesis, I swapped the two IHF III sites. I reasoned that if two IHF III sites existed, then re-positioning IHF IIIa to the native location of IHF IIIb on the opposite DNA strand, and vice versa, would result in promoter activity from this mutant construct (denoted 1xIHF-VF) that was similar, if not identical, to the wild-type 1xIHF promoter. Indeed, I observed virtually identical levels of promoter activity from both 1xIHF and 1xIHF-VF. Their near identical promoter activities led us to test whether the low transcription activity exhibited by 1xIHF-VF required the same distinguishing features for low transcription activity as displayed by 1xIHF. These include IHF, RNAP, both CRP dimers, and at least one of the four AR2 surfaces (Sclavi et al., 2007). Thus, the effects of IHF-mediated inhibition can be bypassed if: a) CRP cannot bind CRP II and/or b) the AR2 surface has been eliminated. I chose to test the involvement of the CRP surfaces, as they provide more information than simply disrupting CRP II.
In the context of the 1xIHF promoter fragment (Chapter 3), IHF-mediated inhibition requires a functional AR2 surface. In its absence, the non-native AR3 surface can compensate and the presence of both AR2 and AR3 results in extreme inhibition. Remember that all three activation surfaces (AR1, AR2, and the mutant AR3) mediate protein-protein interaction. Importantly, they are independent of the surface responsible for binding DNA. At Class III promoters of the Class I-Class I kind (e.g. acsP2), it is proposed that the α-CTD, linked to the body of RNAP by a flexible tether, contacts AR1 of both CRP dimers to activate transcription (Busby and Ebright, 1999). In contrast, AR2 and AR3 are reported to directly contact the body of RNAP. AR2 is reported to interact with the α-NTD (Busby and Ebright, 1999) and AR3 is reported to make contact with Region 4 of σ70 (Rhodius et al; 2000). For AR2 or AR3 to bind to the body of RNAP, a CRP site must be located near the RNAP binding site and yet the proximal site is located at a distance from the promoter, centered at -69.5. Given that promoter constructs that do not exhibit IHF-mediated inhibition (e.g. 1xIHF-C2m) are unaffected by AR2 or AR3, we surmise that IHF alters the local DNA topology enough to permit intimate contact between the CRP dimer bound at CRP I and either the α-NTD (mediated by AR2) or Region 4 of σ70 RNAP (mediated by AR3). Application of a similar analysis to 1xIHF-VF yielded results that were virtually identical to those obtained with the parental 1xIHF. Therefore, I conclude that IHF III is a composite site and that IHF IIIa and IHF IIIb both contribute to inhibition, at least in the 1xIHF context.
IHF III-mediated inhibition of CRP-dependent transcription involves dual overlapping IHF III binding sites

Previous work on acs regulation and the divergent nrf operon yielded contradictory information concerning the function and location of the IHF III site. According to data obtained through our studies, IHF bound to IHF III induces an inhibition loop if tandem CRP dimers bind to CRP I and CRP II and RNAP binds the promoter. A point mutation introduced into CRP II or the AR2 surface of CRP relieves IHF-dependent inhibition. Conversely, Browning et al. demonstrated that IHF occupation of IHF III helps promote transcription of the divergent promoter, nrf. Here, IHF assists activation in the absence of inhibitors (Browning et al., 2006). The IHF binding site relevant to acs (designated IHF IIIa) is offset by two base pairs compared to the IHF site identified as important for nrf activation (designated IHF IIIb). Point mutations in both predicted IHF binding sites do not completely abolish IHF binding (Sclavi et al., 2008; Browning et al., 2006), though this was not an unreasonable result given the IHF degenerate consensus sequence. Mutations in IHF III, within the confines of the 1xIHF promoter, decreased transcription as measured by expression of the reporter gene, lacZ. Total abolition of lacZ expression, or restoration of transcription to levels at or near that exhibited by the 0xIHF promoter, was never observed.

Further, my effort to generate proper mutations that would localize IHF heterodimers to either IHF IIIa or IHF IIIb, or to eliminate IHF binding altogether was met with less than definitive results. Using published work by Lee et al., I
designed a library of 1xIHF promoters that had mutations directed to either the A-T rich sequence, the central WATCAA motif, or the TTR motif. These mutations were constructed in the effort to definitively conclude whether \( IHF \text{ III}_a \), \( IHF \text{ III}_b \), or both sites exist. However, this systematic mutational analysis did not allow us to make strong conclusions as to whether two overlapping IHF sites that are offset by two base pairs, oriented in opposite directions, and positioned on opposing faces are a reality and have biological significance to \( acs \) regulation. Though we could not clearly assign function or relevance to either \( IHF \text{ III}_a \) or \( IHF \text{ III}_b \) using mutational analysis by base pair replacement, the varied results did suggest that \( IHF \text{ III} \) was not your typical IHF binding site. Mutations in the single IHF near-consensus sequence, H’ of \( \lambda \) phage, reveals specific nucleotide replacement at important binding motifs had measurable affects on IHF binding. They observed that specific bases strongly inhibited IHF binding, while some base pair replacements only mildly or moderately did so. In our study of \( IHF \text{ III} \), the same base pair replacement mutations that targeted either \( IHF \text{ III}_a \) or \( IHF \text{ III}_b \) only had weak to moderate effects. Furthermore, double nucleotide replacement meant to severely limit IHF binding to both predicted sites \( (A164T \ T142A; \text{ both replacements were done in the TTR motif}) \) restored promoter activity to levels matching 0xIHF. The single mutations in the TTR of either site, specifically \( T142A \) and \( A141T \), displayed only an intermediate effect on promoter activity and thus IHF binding.

However, replacing \( IHF \text{ III}_a \) with \( IHF \text{ III}_b \), and, consequently, \( IHF \text{ III}_b \) with
IHF IIIa, produces a promoter (designated 1xIHF-VF) whose activity strongly resembles the 0xIHF promoter. Furthermore, the requirements for inhibition are the same as the wild-type 1xIHF promoter, namely, AR2 of CRP and CRP bound to CRP II. Swapping single IHF sites from one DNA strand to the complementary strand would alter the effects of IHF binding on the downstream promoter. For example, if IHF IIIa were the only site pertinent to both acs and nrf transcription, then repositioning it to other DNA strand, while preserving the directionality of binding, IHF would not cause inhibition of acs. In fact, our promoters 1xIHF-2R3a and 1xIHF-2R3b confirm this prediction. Placement of the single IHF binding site, IHF II, in the position of IHF IIIa inhibited acs transcription. In contrast, replacement of IHF III with the single IHF II site did not cause inhibition.

Interestingly, the activities of these 1xIHF promoter variants behaved rather differently when they were assayed from the low-copy plasmid pRW50 relative to the single copy promoter fusions inserted into the chromosome using λ phage. As single fusions, 1xIHF-2R3a increased transcription beyond the activity of the 0xIHF promoter, whereas 1xIHF-2R3b displayed activity similar to that of 0xIHF. We postulate that the difference in promoter activity from the plasmid versus the single-copy fusion depends on the surrounding DNA and its protein-associated contents. pRW50 is not a typical plasmid as it is larger (~17 kb) than the everyday cloning plasmid (~3 kb) and, like all plasmid DNA, undergoes compaction and supercoiling much like genomic DNA. However, the number of genes on pRW50 (~6; Lodge et al., 1992) pales in comparison to the E. coli
genome, which contains over 4300 genes, thus providing more promoter DNA for transcription factors and RNAP to bind. Therefore, I suspect that the difference in reporter gene expression from the pRW50 plasmid and the integrated single-copy of the same promoter construct on the genome is due to a titration effect of available transcription factors and RNAP. To wit, not only do proteins specific for transcription affect gene expression but also local DNA structure, surrounding protein content, and relative transcription activity within the surrounding region influence transcription (Speck et al., 1999; Su’etsugu et al., 2001; Han et al., 2009). Thus, the depth of DNA decoration on the genome may cause the single 2R3a IHF site to behave as a co-activator of transcription rather than an inhibitor. On the other hand, repositioning a single IHF site to yield 2R3b might allow IHF to bind with minimal consequence to acs transcription due to the direction of the DNA bend. However, when both sites are layered, as they are within the natural promoter, IHF may dynamically change binding positions from IHF IIla to IHF IIlb. Finally, the consequence of IHF binding to either site is brought forth by the proteins that are bound downstream, and perhaps, upstream of IHF IIla/b. It is well documented that proteins affect DNA structure upon binding and the aftereffect of such interactions provides the groundwork for further protein binding (Han et al., 2009).

In future studies, we would need to address whether the dual-overlapping IHF sites could function at other promoters similar to the acs architecture (i.e. tandem activator sites)? How would IHF IIla/b affect a Class II promoter or a
complex Class II promoter? Can we find promoters in other organisms that have opposing dual-overlapping binding sites for the same transcription factor? Are there nucleotides outside of the consensus motifs that are critical to IHF binding that might be mutated to more completely reduce binding to either or both sites? This particular experiment would be useful to discern the effects of IHF choosing to bind to either **IHF IIIa** or **IHF IIIb** through the use of oriented heterodimer analysis. This type of analysis would provide us more control over the location of IHF binding and better understand **IHF IIIa** and **IHF IIIb** by dictating when occupation takes place. If we were able to isolate distinct mutations that affected only **IHF IIIa** or **IHF IIIb** then we could use the resultant promoters to generate IHF mutants that recognize the distinct mutants as was done by Busby and Ebright for CRP (1994). This would allow us to introduce mutant IHF heterodimers into single-copy phage lysogens without eliminating the native IHF gene. We would bypass the metabolic problems seen in cells lacking this global regulator and avoid competition for binding by the natural IHF heterodimer. Such a study would permit us to analyze the activity of the 1xIHF promoter in a more natural state while determining the functional properties of the individual **IHF IIIa** and **IHF IIIb** sites. Additionally, we would like to ask: does the current location of **IHF IIIa/b** in relation to the tandem CRP sites and RNAP allow it to function optimally? Or would shifting it up- or downstream change how **IHF IIIa** and **IHF IIIb** effect transcription? And finally, how does the swapping of **IHF IIIa** with **IHF IIIb** affect the opposing divergent promoter, **nrf**? Thorough investigation of **IHF**
IIIa/b is needed to understand how this unique transcription factor binding site influences transcription and the surrounding protein binding sites at divergent promoters, \textit{acs} and \textit{nrf}, under different metabolic conditions. Designing finely tuned future synthetic promoters may benefit from utilizing such as site for manipulation in drug production, metabolic enhancements, etc.

**Tipping the Scales of \textit{acs} Promoter Regulation by \textit{IHF IIIa/b}**

In this thesis, I provided evidence that the upstream, contiguous IHF binding site, \textit{IHF II}, may help determine the location of IHF binding to \textit{IHF IIIa/b} at a given time. Whereas we had seen severe IHF-mediated inhibition in the context of the 1xIHF promoter, the additional presence of \textit{IHF II} (in the promoter fragment denoted 2xIHF) modestly relieved inhibition and produced gene expression levels similar to the 1xIHF mutant constructed by Dr. Beatty (Sclavi et al., 2008). When I extended Dr. Beatty’s mutant by adding \textit{IHF II} expression levels of the reporter gene (\textit{lacZ}) soared past her mutant 1xIHF fragment and even the full-length promoter, a construct with the highest possible expression tested to date. We reason that the 1xIHF mutant weakens IHF binding to both \textit{IHF IIIa/b} but does not eliminate binding altogether; thus, it results in intermediate level of gene expression. The addition of \textit{IHF II} appears to have positioned an IHF heterodimer at either \textit{IHF IIIa} or \textit{IHF IIIb}. Based on their locations relative to \textit{IHF II}, we believe that two IHF heterodimers would be sterically hindered at \textit{IHF II-IHF IIIa}
positions. However, it would be more feasible for two IHF proteins to bind at an 
*IHF II*-IHF *IIIb* conformation. The latter scenario would not form an inhibition loop 
and may position all proteins in an optimal or a better than optimal position for 
transcription. Alternatively, IHF may bind to only *IHF II* as it would be a higher 
affinity site than the mutant composite *IHF III* site. This scenario may actually be 
more ideal because there would be one less DNA bending protein. A less 
crowded *acs* promoter would permit CRP dimers to fully wrap promoter DNA and 
extend the bend angle they impose positioning both DNA and CRP proteins in a 
more optimal location relative to RNAP.

To discern between the two different scenarios, promoter fragments with 
single *IHF IIIa* and *IHF IIIb* binding sites, in the context of the 2xIHF promoter, 
would be tested. The single IHF site, *IHF II*, would be used construct the 
promoters 2xIHF-2R3a and 2xIHF-2R3b to answer the questions: does IHF to 
*IHF II* and *IHF IIIa* or *b* simultaneously? If so, what are the affects on 
transcription? The proposed promoters would help determine the optimal 
orientation of IHF binding to two scenarios where dual IHF sites allows two 
heterodimers to bind either on the same DNA face or on opposing strands. 
Additionally, the same promoter fragments could be used to determine whether 
either 2xIHF-2R3a and 2xIHF-2R3b fragments have different IHF binding 
affinities.
IHF II is a Positive Modulator of IHF-Mediated Inhibition

I next reconsidered the mechanism underlying the vast difference in the weak activity exhibited by the 1xIHF promoter construct and that of the FL construct. Two sequences upstream of the composite IHF IIIa/b caught my interest: a perfect palindrome and IHF II. A construct that included the palindrome but not IHF II (1xIHF+5) reproducibly displayed more activity than 1xIHF (Fig. 26). Since perfect palindromes are exceedingly rare, I invested a lot of effort trying to understand this behavior. In the end, however, I had to conclude that the palindrome did impact acs transcription – at least not under the conditions that I tested. I therefore turned my attention to IHF II.

IHF II is contiguous with IHF III and is predicted to site on the same face of the DNA helix as IHF IIIa. Because of their proximity, we propose that the binding of IHF to IHF IIIa would sterically hinder the binding of IHF to IHF II and vice versa. Under some conditions, the stronger affinity IHF IIIa would become filled at the expense of the lower affinity IHF II. This would lead to IHF-mediated inhibition. Under other conditions, IHF II would be occupied at the expense of IHF IIIa. Since IHF IIIb is predicted to sit on the opposite face of the helix from IHF II, occupancy of IHF II would permit the simultaneous occupancy of IHF IIIb. This complex would mediate CRP-dependent activation of acs transcription. The following observations support these proposals. (1) Using the crystallographic data for IHF bound to H', a model was constructed for the simultaneous binding of two IHF heterodimers to IHF II and IHF IIIb (Fig. 29). A similar model for
simultaneous binding to *IHF II* and *IHF IIIa* was unsuccessful (Doug Browning, personal communication). (2) In the phage system, the 2xIHF construct behaved like the 1xIHF construct (Figs. 26-28). It exhibited low activity that depended on both *CRP II* and AR2. In contrast, weakening *IHF IIIa* and strengthening *IHF IIIb* relative to consensus (the I3m or C152G C150G mutation) caused a much more dramatic increase in transcription in the context of 2xIHF than in the context of 1xIHF (Fig. 27). Thus, the balance between *IHF IIIa* and *IHF IIIb* appears to be the fulcrum controlling transcription from the 2xIHF construct. (3) In the plasmid system, the weakening of *IHF II* in the 2xIHF construct resulted in low activity similar to that of the 1xIHF construct. In contrast, strengthening *IHF II* caused a small increase in activity. (4) Also in the plasmid system, the placement of a simple site (*IHF II*) into the native location of *IHF IIIa* resulted in poor transcription from the 1xIHF construct, while the placement of that same simple site into the native location of *IHF IIIb* resulted in strong transcription. On the basis of these observations, we hypothesize that binding of IHF to *IHF II* induces the switch from *IHF IIIa* to *IHF IIIb* and thus the transition from the stalled unproductive complex to a transcriptionally active complex. We further propose that, at the native *acs* promoter, some upstream *cis* element and/or *trans* factor helps *IHF II* to dictate whether IHF will bind to *IHF IIIa* or *IHF IIIb* and thus whether *acs* transcription will be inhibited or activated.

The underlying cause(s) of the difference in 2xIHF behavior when assayed in the phage and plasmid systems remain(s) unknown. The simplest explanation
is that the different behaviors result from some basic property of plasmids relative
to a prophage integrated into the chromosome. The two most obvious differences
are copy number and superhelicity. In the plasmid system, the copy number is
about one order of magnitude greater than in the phage system. This could result
in a titration effect that would alter the balance between the trans factors that
control acs transcription. Similarly, the superhelicity of a plasmid is likely different
than that of a phage genome that has integrated itself into the chromosome.
Such a difference in superhelicity could have definite consequences on
transcription because superhelicity influences the binding of transcription factors,
NAPs, and RNAP. By controlling copy number, one could test its role. There are
numerous ways to do this, but one of them is a temperature-sensitive expression
plasmid that permits control of its average copy number (Lin et al, 1994;
Hashimoto and Sekiguchi, 1976; Sheridan et al, 1999). We could test the role of
superhelicity in vivo by using combinations of gyrase and topoisomerase mutants
and in vitro using templates that possess different states of supercoiling. It should
be noted that the manipulation of topoisomerase and gyrase activity would affect
the overall health of the cell thus making interpretation difficult.

We are left questioning whether a competition takes place between IHF
Illα and IHF Illβ for IHF binding. Does Illα help promote transcription whereas Illβ
exists to regulate IHF binding to Illα? If Illα does indeed enhance transcription,
then does this mean transcription may be unregulated in the absence of Illβ?
Does the dynamic nature of IHF III, due to the dual-overlapping sites, help
determine when inhibition takes place in conjunction with the available components: CRP, protein contacts via AR2, DNA bending by IHF leading the “trapping” of RNAP at acsP2? Finally, are there fundamental differences in IHF binding affinities between IHF II and IHF IIIa/b? Though intriguing, the answers to these questions are beyond the scope of this dissertation.

The Dynamics of Fis- and IHF-dependent Inhibition on CRP-dependent Activation

Previous studies have determined that the binding of CRP to CRP I and CRP II activates transcription (REF). In contrast, two FIS binding sites, one (Fis II) that overlaps CRP I and one (Fis III) situated between both CRP sites, inhibit transcription by competing with the binding of the activator CRP (REF). Interestingly, a point mutation introduced into CRP I completely disables transcription, despite the presence of an intact CRP II. It is clear that acs transcription depends upon CRP I and the interaction between RNAP and CRP. Any point of regulation would begin with modulating CRP binding to CRP I, in this case by Fis occupation of the overlapping Fis II site. Therefore, inhibition of CRP-dependent activation takes place through competition when Fis occupies Fis II. In contrast, the contribution of Fis III is currently unknown; but, based on the influence of Fis on DNA topology (bend angles from 50 to 90°), the occupation of both Fis I and Fis II may incur a change in promoter structure not conducive to
CRP binding. As Fis and CRP compete for binding at their overlapping sites, the second Fis dimer could tip the balance in favor of inhibition when active CRP concentrations are low. This raises the question: if Fis is an effective inhibitor of CRP-dependent transcription, why would \textit{acs} regulation need IHF?

I speculate that fluctuations of intracellular concentrations of Fis and CRP, during the course of growth and glucose depletion, will reach a point where there are equimolar concentrations of both proteins and will result in an event where neither protein can influence promoter activity. This would lead to RNAP binding inducing only basal transcription. Possibly, to overcome this stalemate the presence of IHF, specifically, when bound to \textit{IHF III}, may be the key to deciding the fate of transcription. IHF, whose concentrations are never depleted from the cell, but increase during stationary phase, may be the protein that occupies the \textit{acs} promoter region at \textit{IHF III} more often than CRP, Fis, or RNAP. When Fis molecules outnumber active CRP dimers, IHF may remodel the promoter into a conformation that promotes Fis-mediated inhibition - possibly when bound to \textit{IHF IIIa} - thus preventing any competition from CRP. The presence of Fis may allow IHF to bind more stably to \textit{IHF IIIa} based on the downstream distance of both Fis sites. However, with IHF still bound, Fis molecules decrease as active CRP dimers increase, and with the constant presence IHF, a more active competition between Fis and CRP takes place at the overlapping \textit{Fis II/CRP I} sites, perhaps destabilizing IHF at \textit{IHF IIIa}. As IHF moves back and forth from \textit{IHF IIIa} and \textit{IHF IIIb} it now disfavors Fis and decreases the effect of \textit{Fis I}. As the rising
concentration of CRP outcompetes Fis, maximal transcription takes place, perhaps with IHF now bound to $IHF\text{ IIIb}$. However, transcription then comes to a halt when active CRP becomes unavailable and IHF concentrations increase thereby forming an inhibition loop preventing additional CRP dimer access and RNAP promoter escape.

**Relevance and Contribution to the Field**

The morphological line that distinguishes eukaryotes from bacteria is beginning to blur. The cellular organelles characteristic of eukaryotes might be absent in bacteria, but much of their basic organizations are similar. For example, it is well established that the cellular processes of eukaryotes tend to be compartmentalized and that specialized cytoskeletal proteins provide the scaffolds that allow these processes to take place without the cell collapsing. In the past, compartmentalization and cytoskeletal proteins were thought to be absent in bacteria. With the advancement of biological technology, however, the gap between the 'sophistication' of eukaryotes and the 'primitiveness' of bacteria has been closing rapidly. For example, early data mining of bacterial genomes revealed no obvious homologues of eukaryotic cytoskeletal proteins. Through the use of high resolution imaging methods and fluorescence technology, however, it was discovered that bacteria actually do possess cytoskeletal proteins, many of which turn out to be distantly related to eukaryotic cytoskeletal proteins.
Furthermore, these techniques have been able to resolve subcellular structures.

Likewise, the once defining characteristics of the eukaryotic chromosome (i.e. chromosomal multiplicity, ploidy, linearity, transcriptional silencing, partitioning, and packaging) also have now been described for bacteria (Bendich and Drlica, 2001). For example, it is now clear that the *E. coli* chromosome is a highly compact and organized structure. Despite being compact, it is very dynamic, undergoing massive reorganization to adjust to the environmental conditions. Both bacterial and eukaryotic chromosomes are bound by proteins that package these enormous structures into the relatively small confines of the cell. However, unlike the eukaryotic chromosome, which has stationary histone proteins to wrap and package DNA into the membrane-bound nucleus, the NAPs associated with the bacterial chromosome do not remain stationary. Furthermore, they perform functions above and beyond packaging. Many of these functions are critical for survival.

Interestingly, some NAPs share sequence and/or functional homology with eukaryotic histones. For example, HU shares amino acid homology, while H-NS shares functional activity with histone H2B. By regulating promoter access and DNA topology, histones and NAPs can regulate transcription (Higgins et al, 2005; Rouviere-Yaniv et al, 1979). For example, both histones and NAPs (together with topoisomerases) induce DNA looping. This important topological change in DNA has consequences for genomic compaction and thus transcription regulation. In transcription, these domains act as a barrier, confining gene expression to a
particular gene, operon, or region. They also provide a structure more conducive to transcription. In bacteria, these loop-domains also protect the integrity of a functional chromosomal structure. If a nick in the DNA double strand occurs within a negatively supercoiled loop, the rest of the chromosome is insulated from this potentially detrimental event. In contrast, a nick in plasmid DNA relaxes its structure, rendering it transcriptionally inactive (Postow et al, 2006).

Eukaryotic and bacterial promoters also share structural similarities, e.g. low stability, higher curvature, and bendability (Kanhere and Bansal, 2005). The transcriptional machinery, the RNAP holoenzyme, first must gain access to the promoter and then bind to the appropriate sequences. Promoter activity is dependent on cis and trans factors. The sequences of the -35 and -10 hexamers in bacteria and the TATA and CAAT boxes in eukaryotes are identified consensus sites for RNAP binding. The relative promoter strength correlates with the similarity of a given promoter sequence to the consensus sequence of these hexamers and boxes. For example, a weak promoter generally possesses poor resemblance to consensus. At such promoters, an activator is generally needed to stabilize the interaction of the RNAP with the promoter. In contrast, a strong promoter generally shares much similarity with consensus. Such promoters are often repressed by the binding of a transcription factor to a DNA site adjacent to or overlapping the transcription initiation site (+1), the -10 hexamer or the -35 hexamer.

In addition to the basic principles of promoter recognition and binding of
the RNAP holoenzyme, the next set of steps that lead to transcription are similar in all organisms. For example, in all organisms, the binding of RNAP results in closed complex formation followed by formation of an open complex around the -10 hexamer (the TATA box in eukaryotes). The low stability of the AT-rich sequence allows for the strand separation of the DNA duplex characteristic of open complex formation, which permits RNAP to begin ribonucleotide polymerization. An assessment of the region upstream of the transcription start site reveals a propensity for DNA curvature. This curvature promotes RNAP and transcription factor binding and thus transcription. Upon binding, RNAP, as well as some transcription factors, wraps the DNA around itself, perhaps for a more stable interaction. The bendability of the DNA strand also brings distant upstream enhancer sequences with their associated proteins into close proximity to the RNAP.

It is generally accepted that eukaryotic promoters are very complex, often encoding many sites for the binding of a multitude of transcription and accessory factors. In contrast, the early (and too often prevailing) view of bacterial promoters suggested a simple mechanism of activation, where one transcription factor (activator or inhibitor) influences RNAP function (see above). Yet, a survey of bacterial promoters demonstrates the existence of both simple and complex promoters. Within a complex bacterial promoter there can be many binding sites for a variety of DNA binding proteins, as well as multiple start sites, enhancer sequences, and regions of DNA recombination. The location of these features
within a promoter provides the platform that determines the mechanism of regulation.

The information gathered from the studies described in this dissertation expands our knowledge of the complex acs promoter and bacterial gene regulation, in general. My dissection of the complex nrf-acs intergenic region allows us to build on the general knowledge of bacterial transcription and contribute insight into the rules of promoter design that Nature has encoded into DNA regulatory regions.

Not surprisingly, the multifaceted nature of NAPs is location-dependent and is influenced by proteins bound upstream and downstream of the promoter. For example, IHF bound to a single site located midway between an enhancer activation sequence (UAS) 100 bp upstream and the downstream RNAP site in the Pu promoter behaves as an activator of transcription. It does so by folding the DNA, thereby bringing the UAS into close proximity to RNAP. Eliminating IHF binding to this site shuts off transcription. However, replacing this IHF site with a DNA sequence with an intrinsic bend compensates for the loss of IHF. Similarly, IHF activates transcription at the ilv promoter, also by bending the DNA and repositioning a UAS (Sheridan et al, 1999). Indeed, DNA remodeling by IHF is so profound that it has the capability of inducing V(D)J recombination in ΔRAG B-cells.
In contrast, IHF can also behave as an inhibitor. For example, at the \( ilvP_{G1} \) promoter, the IHF binding site at this promoter overlaps that of RNAP, these sites are not located on the same face of the DNA helix. Thus, although IHF occludes RNAP binding, it does not do so through physical occlusion. Instead, it remodels the DNA structure, preventing RNAP from gaining access to the promoter.

In this dissertation, I provided evidence that supports the existing notion that the location of an IHF protein binding is just as important as the affinity of that site for IHF. What is novel is the discovery of a composite IHF site (IHF III) comprised of two inverted and overlapping sites. This dissertation also reports the ability of an upstream, contiguous site (IHF II) that can counteract the inhibitory effect of IHF III, presumably by guiding IHF to bind to IHF IIIb, which appears to function as a positive effector of acs transcription, instead of IHF IIIa, which appears to be a negative effector. The most pressing issue is how IHF II becomes occupied. In the context of the 2xIHF promoter construct, activation required expression from a highly supercoiled multicopy plasmid or disruption of IHF IIIa. The mechanism by which IHF II becomes occupied in the native acs context requires further exploration.


VITA

In 1975 Heng and Sang Thach, along with their four year old daughter, Somaly, immigrated to the Chicagoland area to escape the repercussions of the Vietnam War. David Thach, the middle of three children (younger brother, Tony), was born in Glen Ellyn, IL. David graduated from Donald E. Gavit High School and entered college while working two jobs. After a stint studying Computer Science, David went on to finish his bachelor’s in Microbiology with a minor in Genetics at Purdue University. A yearlong hiatus in Minneapolis, MN pressed David to enter graduate school in the Department of Microbiology and Immunology at Loyola University Chicago in 2001.

David joined the laboratory of Dr. Alan J. Wolfe and began his investigation on transcriptional regulation of the acs complex promoter. Here he received excellent instruction and advice on the process of laboratory science. After pushing the project to an interesting end, David began the process of writing his dissertation before leaving for a postdoctoral position in Oakland, CA in the summer of 2009. David is currently continuing to apply his laboratory and professional skills learned under Dr. Wolfe’s direction to study gene regulation in Neisseria meningitidis.