2017

Mutual Regulation of CRP and N(epsilon)-Lysine Acetylation in Escherichia Coli

Robert James Davis
Loyola University Chicago

LOYOLA UNIVERSITY CHICAGO

MUTUAL REGULATION OF CRP AND Nε-LYSINE ACETYLATION IN *ESCHERICHIA COLI*

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE GRADUATE SCHOOL

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

PROGRAM IN BIOCHEMISTRY AND MOLECULAR BIOLOGY

BY

ROBERT JAMES DAVIS

CHICAGO, IL

DECEMBER 2017

ACKNOWLEDGEMENTS

I would like to thank my mentor, Dr. Alan Wolfe, for his guidance during my time at Loyola. He encouraged me to think deeply about science, taught me to become a better writer, and provided emotional support when I needed it. I would like to thank Dr. Mitchell Denning, my committee chair, for accepting me into the Molecular Biology program and guiding me through the academic side of graduate school. I would like to thank Dr. Karen Visick, who provided insightful discussions during lab meetings. I would like to thank Dr. Richard Schultz, who often provided interesting conversation, even when not directly related to my project. I would like to thank Dr. Adam Driks, whose constructive criticisms were invaluable for honing my presentation skills. I would also like to thank Dr. Francis Alonzo, who graciously stepped in at the last minute to round out my dissertation committee.

I would like to thank my family, who supported me both emotionally and financially throughout my entire education. I am very lucky to have loving parents who worked hard to ensure their children had every available opportunity. I would like to thank the friends I made at Loyola, who provided much needed relief from the stresses of graduate school. Finally, I would like to thank my future spouse, Cory. I cannot thank her enough for her patience and support during my time in graduate school. I don't think I would have had the strength to complete my degree without her, which is one of a million reasons I will always love her.

For Cory, who provides a never-ending supply of motivation, inspiration, and love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

RNAP	RNA polymerase
TSS	Transcription start site
CTD/NTD	C-terminal domain/N-terminal domain
TF	Transcription factor
cAMP	Cyclic AMP
CRP	cAMP receptor protein
HTH	Helix-turn-helix
AR	Activating region
PTS	Phosphotransferase system
TCA	Tricarboxylic acid cycle
TCST	Two component signal transduction
NAT	N-acetyltransferase
KAT	Lysine acetyltransferase
KDAC	Lysine deacetylase
acCoA	Acetyl CoA
acP	Acetyl phosphate
VC	Vector control
WT	Wild type
TB7	Tryptone broth pH 7

ABSTRACT

Post-translational modifications, such as Nε-lysine acetylation, are known to alter the behavior of transcriptional regulators in eukaryotes, but very little is known about the consequences of acetylation on transcriptional regulation in bacteria. Here, I provide evidence that a global transcriptional regulator of carbon metabolism, cAMP Receptor Protein (CRP), promotes both enzymatic and non-enzymatic lysine acetylation in *E. coli*. Non-enzymatic lysine acetylation occurs when cells ferment acetate, such as during growth on high concentrations of glucose. Intriguingly, CRP can be non-enzymatically acetylated on several lysines, including lysine 100 (K100). I provide evidence that neutralization of the K100 positive charge, as would occur upon K100 acetylation, has a dual effect on CRP activity. First, K100 neutralization decreases CRP activity at some Class II promoters. This decreased activity likely results from disruption of the interaction between Activating Region 2 (AR2) of CRP and the RNA polymerase α subunit N-terminal domain. Second, K100 neutralization increases the CRP half-life, leading to increased CRP steady state levels. Due to increased steady state levels, CRP activity is increased at some Class I promoters, in which CRP does not require AR2. Taken together, I propose that CRP promotes global acetylation, including CRP K100 acetylation, when cells are grown on glucose by positively regulating non-enzymatic acetylation. A consequence of K100 acetylation may be inverse regulation of Class II and Class I promoters under these conditions. This mechanism may help regulate carbon flux though central metabolism.

CHAPTER ONE

INTRODUCTION

Introduction

Bacteria can be found thriving almost everywhere in nature. To do so, they must be able to monitor their environment, both extracellular and intracellular, and alter their behavior in response to changes in their environment. For example, if there is a lack of nutrients in the immediate vicinity, a bacterium must move itself to a new location with greater nutrient availability, or limit its growth and energy expenditure until nutrients become available.

Bacterial behavior is governed in part by the proteins that are expressed within the cell. The cell must control the timing of the synthesis and degradation of each protein to ensure the correct behaviors are active at the right time. To this end, the cell encodes hundreds of transcription factors, which are responsible for regulating the expression of different subsets of genes. Some of these transcription factors are directly sensitive to the environment, becoming activated or inactivated through direct interaction with small molecules, such as oxygen, nitrogen, or carbon sources. Other transcription factors respond to changes in the environment through post-translational modifications: covalent modifications to proteins that often alter their activity. These post-translational modifications are indirect environmental signals, often passed down by proteins that more directly sense environmental cues.

This dissertation investigates a novel relationship between transcriptional regulation and post-translational modification. There is evidence that the global transcription factor CRP enhances the abundance of Nε-lysine acetylation throughout the cell, altering the activities of the modified proteins. Furthermore, CRP itself is modified by Nε-lysine acetylation, and therefore its activity may be affected by N_ε-lysine acetylation. The work presented in this document is aimed at investigating the mechanisms by which CRP influences Nε-lysine acetylation and *vice versa*.

To better understand the rationale and data presented in this dissertation, I will begin by describing the process of transcription. I will then describe the regulation of transcription, first generally and then with a specific focus on CRP. Next, I will discuss catabolite repression and overflow metabolism, both of which are mechanisms the cell uses to alter its metabolism when sufficient glucose is available. Finally, I will go into detail about the various types of protein acetylation and their effects on protein function, with a specific focus on Nε-lysine acetylation.

Transcription

Transcription is the process of generating RNA from a DNA template. In bacteria, transcription is performed by a single complex of proteins called RNA polymerase (RNAP) (1, 2). The RNAP core complex is made up of five subunits: α (x2), β, β', and ω. These five subunits are highly conserved across all domains of life and are sufficient for RNA synthesis; however, another subunit, σ, is required to bind to the RNAP core complex to direct the assembled RNAP to specific promoters. This RNAP-σ complex is called the RNAP holoenzyme. Binding of the

RNAP holoenzyme to DNA results in the formation of what has been termed the closed complex. In most cases, conformational changes (isomerization) within the RNAP holoenzyme provide the energy required to melt and unwind the promoter DNA, which forms the open complex. Once transcription begins, the first few ribonucleotides are ligated together within the initiation complex. At this point, the RNAP holoenzyme is still anchored to the promoter DNA via σ. Instead of moving forward down the DNA strand to synthesize RNA, RNAP pulls the DNA into itself, which has a two-fold effect. First, it allows for the synthesis of a short (9-11 nucleotides) initial fragment of RNA. Second, the DNA pulled into RNAP "scrunches", providing energy to allow RNAP to escape the promoter (3). In many cases, the DNA "slips" out of RNAP, causing a loss of the accumulated energy and release of the initial fragment of RNA, called an abortive transcript. If abortive transcription occurs, RNAP reverts to an open complex conformation and attempts to initiate transcription again. Upon successful initiation of transcription, RNAP uses the energy stored in the "scrunched" DNA to escape the promoter by dissociating from σ, forming the elongation complex, and continuing RNA synthesis from the initial fragment.

There are seven known σ factors in *E. coli*, six of which belong to the σ^{70} family of σ factors (4, 5). The seventh σ factor, σ^{54}, is structurally and functionally distinct from the σ^{70} family members and will not be discussed here (6). Each σ factor is comprised of four conserved major domains (regions 1-4), and each major domain can be broken down into minor domains. Although each σ factor recognizes a consensus sequence that is specific for that σ factor, there is some tolerance for degenerate sequences, which allows the cell to control the probability

and strength of the RNAP holoenzyme interaction with the promoter. Additionally, this tolerance can result in an overlap in the regulons of σ factors (7, 8).

The RNAP holoenzyme makes several contacts with nucleotide sequences within promoter DNA to ensure transcription begins at the correct location (1). The primary sequences that regulate σ factor specificity are called the -35 and -10 elements, since the sequences are located approximately 35 and 10 nucleotides upstream of the transcription start site (TSS) respectively. The -35 element interacts with σ region 4.2, and the -10 element interacts with σ region 2.4. A third contact is made between the promoter DNA and the spacer region between σ regions 2.4 and 3.1 (referred to as either region 2.5 or 3.0) when an extended -10 element is present in the promoter (9, 10). The extended -10 element is a two-nucleotide sequence located just upstream of the -10 element, and is generally present only when the -35 element is absent.

In addition to the contacts made between σ and the promoter DNA, the C-terminal domain (CTD) of the RNAP α subunit also contacts the promoter DNA at sequences called UP elements (11). These sequences are located upstream of the -35 element and are rich in adenosines (A) and thymidines (T). Binding of the α-CTD to the UP element provides an additional point for promoter recognition and stabilization of the RNAP closed complex.

Transcriptional Regulation

There are currently 4627 recognized genes in *Escherichia coli* (12). To ensure the proper timing and level of expression for each gene, the regulation of transcription must be tightly controlled. There are several layers of transcriptional regulation; some of which are hardwired

into the DNA (*cis*-acting), while others can respond to changes in the external or internal environment (*trans*-acting). Together these layers of regulation allow fine-tuning of gene transcription based on the changing needs of the cell.

In bacteria, genes are often organized in groups called operons. Genes within an operon generally encode proteins that perform similar functions or operate within similar pathways. At the most simplistic level, genes within an operon are transcribed together from a single promoter into a single messenger RNA (mRNA) transcript, and are therefore all under the same transcriptional regulation. This allows for efficient expression of many or all of the genes required for a particular function with minimal coordination required. Many operons have additional complexities, including internal promoters, to fine-tune the expression of genes within the operon. This means that while a single mRNA can encode all the genes of an operon under certain conditions, other conditions may favor separating transcription of the operon into multiple mRNAs.

There are approximately 300 DNA binding proteins that can bind to promoter DNA and regulate transcription (13). The majority of these DNA binding proteins fall into the category of transcription factors (TFs), which directly impact the recruitment or isomerization of RNAP. Much like the RNAP holoenzyme, TFs interact with the DNA at specific binding sites unique for each TF. By analyzing all known binding sites for a particular TF, a consensus sequence can be obtained that represents the sequence for which the TF has highest affinity. Deviation from the consensus sequence allows the cell to control the affinity a TF has for a particular binding site, allowing a more nuanced regulation of transcription initiation. Relatively weak affinity of a TF

for its binding site can be enhanced through interactions with other TFs, helping to keep a promoter turned off until the right conditions are met.

TFs regulate RNAP activity by two broad mechanisms: inhibition and activation. Many, but not all, TFs can act as either an inhibitor or an activator depending on the context of the promoter. Modes of transcriptional inhibition can be sorted into three categories. The first is repression by steric hindrance. This type of repression occurs when the TF binds at a site in the promoter that is very near or overlapping the RNAP binding site, preventing RNAP from binding. This is one of the mechanisms used by the *lac* operon repressor LacI (14). The second mode of repression is by DNA looping. This type of repression occurs when one or more TFs bind to multiple sites around the promoter. These TFs then interact with each other, causing the DNA to bend and form a loop around the promoter, preventing access to RNAP. This mechanism is used by GalR to repress the *gal* operon (15), and by LacI to enhance its repression of the *lac* operon (16). The third mode of repression is by disrupting the activity of an activator, or anti-activation. At the *deoP2* promoter, two CRP dimers bind to separate locations and activate transcription of the *deo* operon. The CytR repressor is capable of binding between the two CRP dimers and interacting with the promoter-proximal dimer, resulting in repression of *deo* transcription. The proposed mechanism involves a CytR-induced conformational change within CRP, preventing CRP-induced activation of *deo* transcription (17).

As with transcriptional repression, there are three categories of transcriptional activation. The first two categories operate through direct interaction with RNAP. Class I activation involves interaction between a TF and the RNAP α subunit CTD. This interaction aids in the recruitment of RNAP to the promoter and formation of the closed complex, usually in

cases where the RNAP binding site is weak (18, 19). Several activators employ this mechanism, including Ada at the *ada* promoter (20), OmpR at the *ompC* promoter (21), OxyR at the *katG* and *ahpC* promoters (22), and CRP at the *uxuAB* and *lac* promoters (23). Class II activation involves interaction between a TF and other regions of RNAP, generally the α subunit NTD or σ subunit. Unlike Class I activation, Class II activation primarily promotes the isomerization of RNAP from the closed to open complex (18). Examples of Class II activators include cI of the λ bacteriophage at the λP_{RM} promoter, which interacts with σ subunit region 4 (24), and CRP at the *galP1* promoter, which interacts with the α subunit NTD (25). The third category of transcriptional activation involves distorting the DNA to enhance RNAP binding. At the *soxS* promoter, there are 19 base pairs in the space between the -10 and -35 regions of the RNAP binding site instead of the ideal 17 ± 1 base pairs, preventing efficient RNAP binding. SoxR binds in the space between the -10 and -35 regions and twists the DNA, bringing the -10 and -35 regions closer together to allow RNAP binding. SoxR binding occurs on the opposite side of the DNA, allowing SoxR and RNAP to bind simultaneously (26, 27).

Almost half of all TFs are directly activated or inhibited by small molecules (28). For example, LacI becomes inactivated upon binding to allolactose, relieving repression of the *lac* operon, which encodes the genes required for the transport and metabolism of lactose (29). Fumarate and nitrate reductase (FNR) becomes inactivated through interaction with O_2, preventing expression of genes involved with anaerobic metabolism (30). Not all small molecule interactions inactivate TFs. Melibiose binding activates MelR, which activates transcription of genes involved in melibiose metabolism (31). Direct regulation by small molecules allows TFs to

sense the environmental or nutritional status of the cell, and in turn alter gene expression to mount an appropriate response.

cAMP Receptor Protein (CRP)

CRP (also known as catabolite activator protein [CAP]) is one of seven global TFs that together regulate transcription of over 50% of the *E. coli* genome (32). Alone, CRP directly regulates the transcription of 430 genes, or almost 10% of the genome (12). The CRP regulon primarily consists of genes involved in the catabolism of secondary carbon sources (see Chapter One – Catabolite Repression), including most famously the *lac* operon (33). However, CRP regulates genes involved in a wide variety of other pathways as well. These pathways include biofilm formation (34), translation (35), persistence (36), antibiotic resistance (37), and virulence (38). In addition, CRP regulates the transcription of genes encoding other transcriptional regulators (39–42) and sigma factors (43, 44), allowing CRP to indirectly influence the transcription of genes well outside of its own regulon. Altogether, the evidence suggests that the influence of CRP is wide-reaching and impacts nearly every aspect of the cell.

Regulation of Transcription by CRP

CRP is a dimer consisting of two identical monomers (45). Each monomer is made up of two domains: a large NTD and a smaller CTD. The CRP NTD (residues 1-133) is responsible for binding cyclic AMP (cAMP) and dimerization, while the CTD (residues 139-209) contains a helix-turn-helix (HTH) motif that is responsible for DNA binding. These two domains are connected by a short flexible hinge region (residues 134-138). In the absence of cAMP, the DNA recognition helices within the HTH motifs are buried within the CRP dimer (46, 47). Upon binding cAMP,

structural changes within CRP cause the DNA recognition helices to swing out, allowing access to promoter DNA. These DNA recognition helices insert into two consecutive major grooves in the DNA at the CRP binding site, making both specific and non-specific interactions with the DNA (48).

CRP activates transcription primarily through direct interaction with RNAP. The interactions made between CRP and RNAP depend on the distance between the CRP and RNAP binding sites (**Fig. 1**). At Class I promoters, the CRP binding site is centered around 61.5, 71.5, 82.5, or 92.5 nucleotides upstream of the TSS. The spacing between potential CRP binding sites is consistent with one complete turn of DNA (49), highlighting the fact that CRP is required to bind on the same face of the DNA as RNAP for productive interaction (50). The interaction between CRP and RNAP requires surface-exposed residues on each complex. On CRP, the critical residues are 156-164, also known as activating region 1 (AR1) (51) (**Fig. 2**). Although each CRP monomer contains an AR1, it is AR1 on the promoter-proximal monomer that is required at Class I promoters (52, 53). AR1 is located within a surface loop in the CRP CTD, though not near the HTH motif required for DNA binding. Residues within AR1 interact with residues 285-288 and 317 (also known as the 287 determinant) within the RNAP α subunit CTD (54). This interaction promotes the recruitment and stabilization of the RNAP closed complex (18).

At Class II promoters, CRP binds directly upstream of RNAP at a site centered around 41.5 nucleotides upstream of the TSS. Like at Class I promoters, there is a productive interaction between CRP AR1 and the RNAP 287 determinant at some, but not all, Class II promoters (55, 56). At these promoters, it appears that the RNAP α subunit CTD extends over

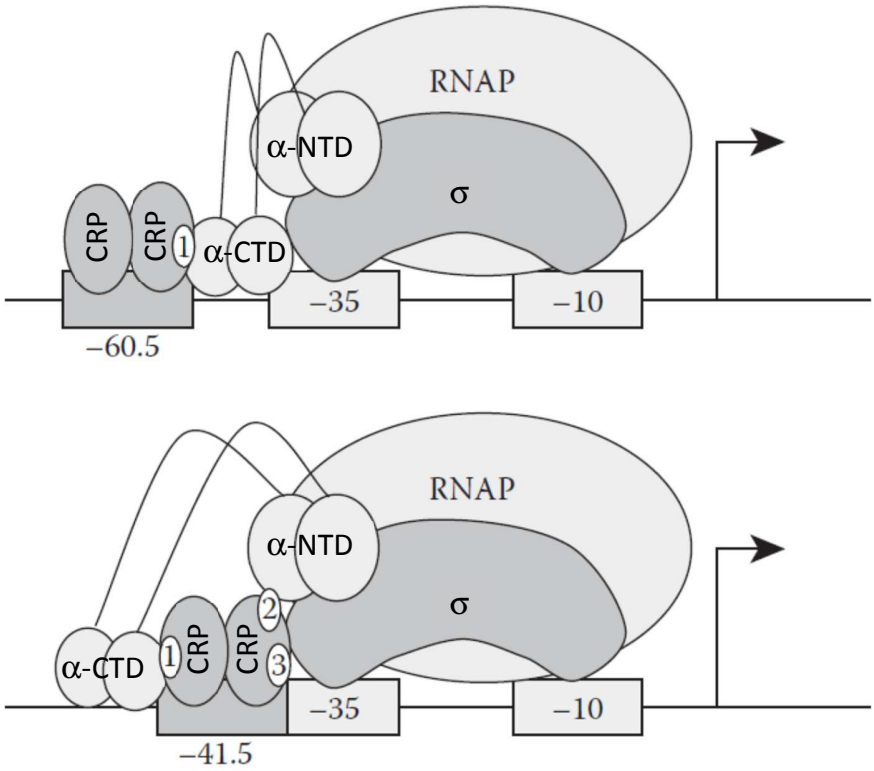

Figure 1. CRP-RNAP interactions at Class I and Class II promoters. (Top) At Class I promoters, CRP binds around 60.5, 70.5, 81.5, or 91.5 nucleotides upstream of the transcription start site (TSS). CRP AR1 interacts with the RNAP α subunit C-terminal domain (CTD). (Bottom) At Class II promoters, CRP binds around 41.5 nucleotides upstream of the TSS. CRP AR2 interacts with the RNAP α subunit N-terminal domain (NTD). The CRP AR1-RNAP α subunit CTD interaction also occurs at some, but not all, Class II promoters. CRP AR3 interacts with the RNAP σ subunit, although this interaction is not considered to be significant. This figure was adapted from (236).

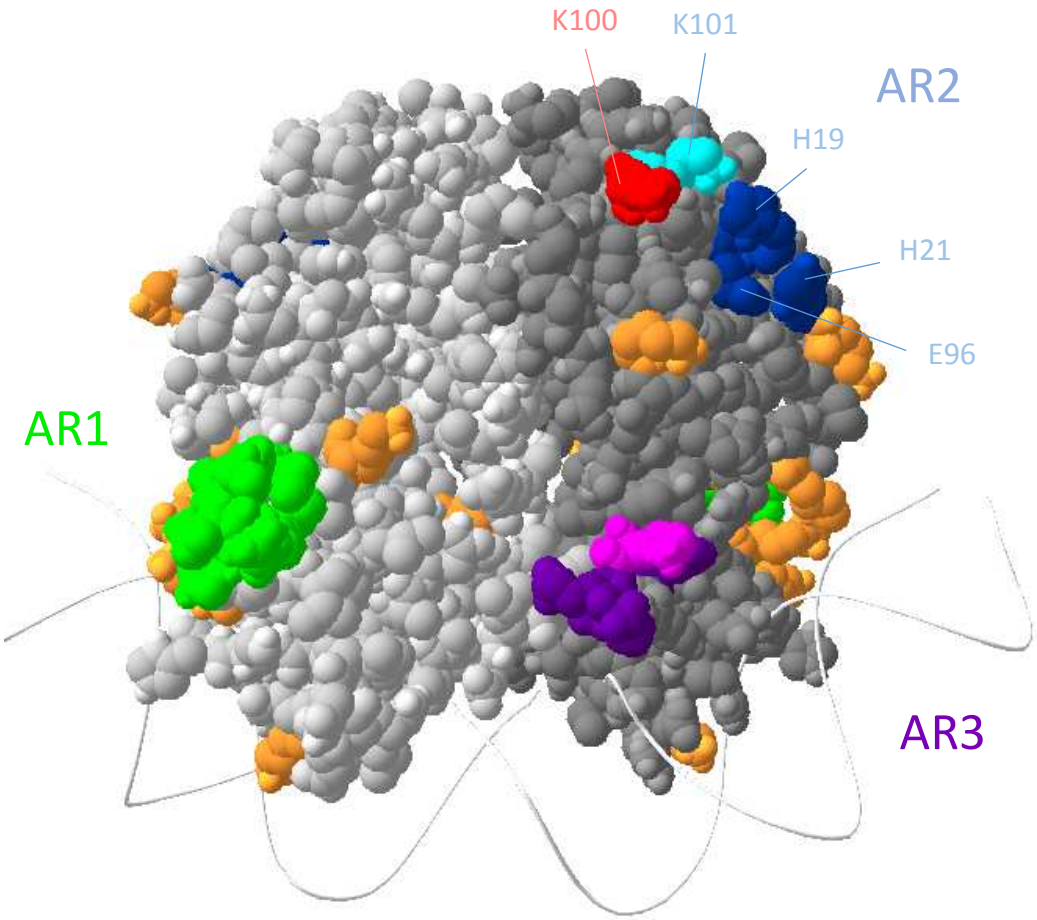

Figure 2. Crystal structure of CRP. A cAMP-bound CRP homodimer (dark grey and light grey) is seen bound to DNA (thin lines). Residues that make up the three activating regions are highlighted in green (AR1), blue (AR2), and purple (AR3). Lysines which have been detected as acetylated are shown in tan. K101, shown in light blue, and K52, shown in light purple, are both acetylated and components of AR2 and AR3 respectively. K100, a subject of this dissertation, is shown in red. The structure was determined by x-ray diffraction at a resolution of 3 angstroms (59). Figure was generated in Swiss-PdbViewer 4.1.0. (PDB: 1CGP)

the CRP dimer and interacts with AR1 of the promoter-distal monomer (53). However, the critical interaction between CRP and RNAP involves residues H19, H21, E96, and K101 (AR2) of CRP (**Fig. 2**) and residues E162, E163, D164, and E165 of the RNAP α subunit NTD (25). The positively charged residues of AR2 form ionic bonds with the negatively charged residues of the RNAP α subunit NTD, which primarily promotes the isomerization of the RNAP closed complex to the open complex (18). The negatively charged E96 residue has a small inhibitory effect on Class II transcription activation (25).

A third surface on CRP, AR3, can interact with the RNAP σ subunit region 4 at Class II promoters (57). AR3 is composed of an activating determinant consisting of residues D53, E54, E55, and E58, and an inhibitory determinant consisting of residue K52 (58). The negatively charged activating determinant interacts with positively charged residues K593, K597, and R599 within the RNAP σ subunit to promote the isomerization of the RNAP closed complex to the open complex at Class II promoters, much like the AR2-RNAP α subunit NTD interaction (57, 58). The positively charged inhibitory determinant likely clashes with the same positively charged RNAP σ subunit surface, disfavoring the AR3-RNAP σ subunit interaction (58). Interestingly, mutation of all AR3 residues to alanine results in nearly WT CRP activity at Class II promoters, suggesting the contributions of the activating and inhibitory determinants to promoter activation negate each other. Therefore, AR1 and AR2 are considered to be the only significant interactions between CRP and RNAP at Class II promoters.

In addition to making direct contact with RNAP, CRP regulates transcription by bending the surrounding DNA. Positively charged residues on the sides of the CRP dimer opposite the dimerization domains interact non-specifically with the negative phosphate groups in DNA,

causing the DNA to partially wrap around the CRP dimer. As a result, the DNA is bent approximately 80° around CRP (59–61) . This bend can bring relatively distant transcriptional regulators closer to RNAP, or it can prevent RNAP from binding to the promoter.

Regulation of CRP Expression

As one of the seven global regulators that control the transcription of half of the genome (32), it is not surprising that the regulation of CRP expression is complex. In this section, I will describe in detail the regulators that are responsible for this pattern of expression, including regulators of *crp* transcription and CRP stability.

Transcription of the *crp* gene occurs from two different promoters: *crp1* and *crp2*, located approximately 79 base pairs downstream of *crp1* (62). *In vivo*, the majority of *crp* transcription comes from the *crp1* promoter. There are two CRP binding sites within the *crp* promoter region, indicating that *crp* transcription is autoregulated (63). The first CRP site is centered 60.5 nucleotides upstream of *crp1* (139.5 nucleotides upstream of *crp2*) and promotes *crp* transcription, indicating that *crp1* is a Class I promoter (64). The second CRP site is centered 41.5 nucleotides downstream of *crp1* (37.5 nucleotides upstream of *crp2*) and inhibits *crp* transcription by recruiting RNAP to initiate transcription in the opposite direction from a divergent promoter (63, 65). *In vivo,* CRP predominately inhibits *crp* transcription; β-galactosidase activity from a *crp'-lacZ* promoter fusion is increased in stains lacking *crp* relative to strains that express *crp* (62, 66).

Transcription of *crp* is regulated by two other transcription factors, Cra and Fis. Little is known about the role of Cra in *crp* transcription, except that Cra activates *crp* transcription by a mechanism that requires CRP (67). Fis inhibits *crp* transcription by binding to the promoter and

overlapping the RNAP binding sites (62). Fis levels are high at the beginning of growth (68), repressing the transcription of both *crp* and the divergent gene (62). As Fis levels begin to decline over growth, the divergent promoter, but not the *crp* promoter, becomes available for RNAP to initiate transcription and maintain the repression of *crp*. Once Fis levels are at a minimum in stationary phase, Fis is no longer bound to the *crp* promoter and *crp* transcription increases. As an added complexity, CRP regulates Fis expression, leading to cyclical regulatory crosstalk between the two regulators (41).

Catabolite Repression

Bacteria utilize a variety of carbon sources for growth and energy production. While the bacteria will eventually use all available carbon sources, most species have a preferred order in which they use the carbon sources based on how readily the carbon sources can be catabolized. This hierarchy of carbon preference is maintained through a process called catabolite repression (69), which will be discussed in this section.

Catabolite repression was first formally identified when it was observed that *E. coli*, when grown in a medium with both glucose and lactose as carbon sources, will grow by exclusively consuming glucose until the glucose is depleted, pause growth, then resume growth by consuming lactose (70). This phenomenon became known as the “glucose-lactose diauxie”. Other “preferred” carbohydrates were identified as being able to repress the ability of *E. coli* to simultaneously consume “secondary” carbohydrates, but since glucose has been the best studied, this section will focus on glucose as the preferred carbohydrate.

The division of carbohydrates into “preferred” and “secondary” groups is primarily achieved by a transport system called the phosphotransferase system (PTS) (71, 72). The PTS is responsible for both repressing the expression of genes required for the catabolism of secondary carbohydrates and inhibiting the uptake of secondary carbohydrates, as long as the cell takes up the preferred carbohydrates. Each of these functions revolves around PTS facilitating the transfer of phosphoryl groups from PEP, the phosphoryl donor, to the incoming carbohydrate molecules.

The PTS consists of three components: EI, HPr, and EII (72) (**Fig. 3**). The EI and HPr components are generic to the PTS, serving only to transfer phosphoryl groups from PEP to EI to HPr to EII. The EII component is specific for each preferred carbohydrate and serves as the transporter and kinase for that carbohydrate. There are at least 22 known or putative EII systems in *E. coli*, each capable of transporting a different substrate or set of substrates (73). In the case of glucose, EII is composed of $EIIA^{(Glc)}$, which is phosphorylated by HPr, and $EIICB^{(Glc)}$, which is phosphorylated by $EIIA^{(Glc)}$. As glucose enters through the phosphorylated $EIICB^{(Glc)}$ subunit, $EIICB^{(Glc)}$ transfers its phosphoryl group to the glucose molecule, initiating glycolysis (72). During glycolysis, two molecules of PEP are formed; the PEP is converted to pyruvate by transferring the phosphoryl group to either ADP, forming ATP, or by donating its phosphoryl group to EI, which will pass that phosphoryl group down the PTS and on to the next incoming glucose molecule.

In the absence of glucose, $EIICB^{(Glc)}$ and $EIIA^{(Glc)}$ remain phosphorylated, since there are no glucose molecules to accept the phosphoryl groups. The phosphorylated $EIIA^{(Glc)}$ activates adenylate cyclase, the enzyme responsible for cAMP synthesis (74) (**Fig. 4**). Adenylate cyclase

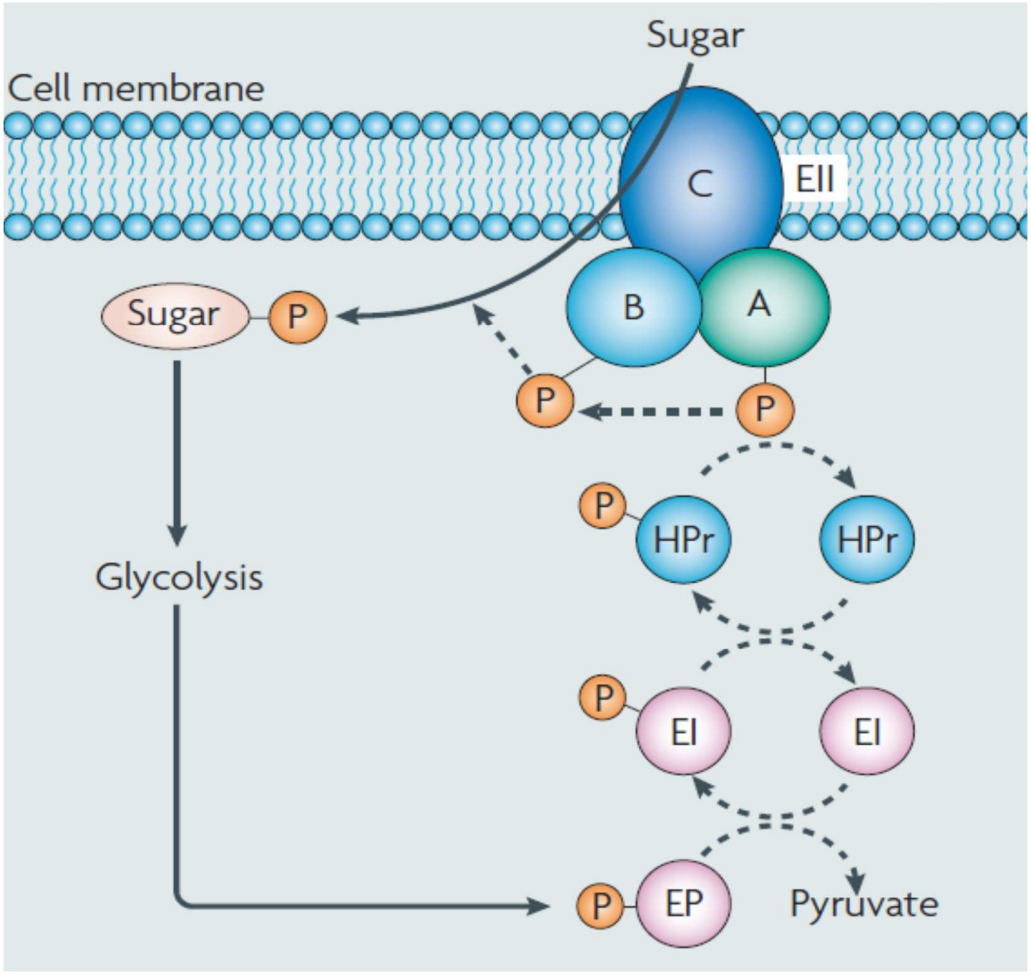

Figure 3. The phosphotransferase system. The PTS is phosphorylated by PEP, which is generated by glycolysis. PEP donates it phosphoryl group to EI, which passes the phosphoryl group to HPr, which passes the phosphoryl group to EIIA, which passes the phosphoryl group to EIIB. As a PTS carbohydrate is transported across the membrane by EIIC, EIIB transfers the phosphoryl group to the carbohydrate, initiating glycolysis. Glycolysis generates two molecules of PEP, which can provide the phosphoryl group for the next incoming carbohydrate molecule. This figure is taken from (71).

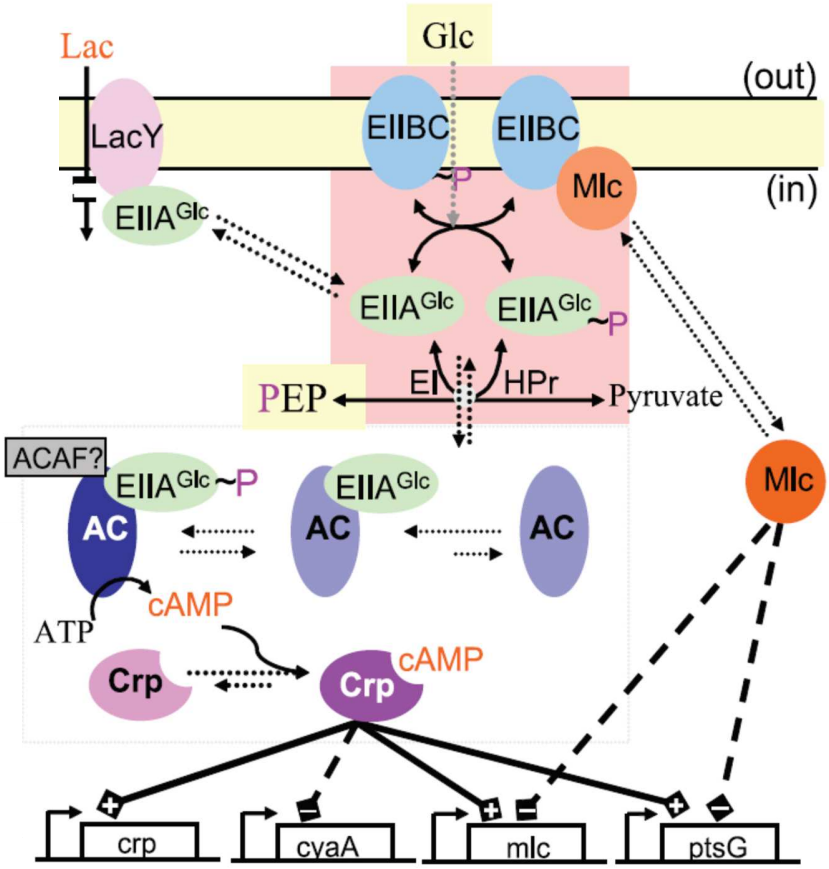

Figure 4. Catabolite repression/inducer exclusion. When glucose enters the cell via the PTS, members of the PTS pass their phosphoryl groups to the incoming glucose molecules, leaving the PTS components primarily unphosphorylated. Unphosphorylated $EIIA^{(Glc)}$ inhibits the transporters of secondary carbon sources such as lactose (LacY). This "inducer exclusion" prevents lactose from entering the cell and inducing transcription of the *lac* operon. At the same time, unphosphorylated $EIICB^{(Glc)}$ sequesters Mlc, a transcriptional inhibitor of the PTS, promoting PTS expression. Once the glucose is consumed, the PTS no longer has a substrate to transfer phosphoryl groups to, so it remains phosphorylated. Phosphorylated $EIIA^{(Glc)}$ (along with an unknown adenylate cyclase activation factor [ACAF]) activates adenylate cyclase, which makes cAMP. cAMP activates CRP, promoting transcription of proteins involved in the catabolism of secondary carbohydrates. Figure adapted from (72).

activation also requires an unknown adenylate cyclase activation factor (ACAF). The increased cAMP levels increase CRP activity, leading to increased expression of genes required for the import and catabolism of secondary carbohydrates. In the presence of glucose, $EIIA^{(Glc)}$ and $EIICB^{(Glc)}$ readily pass on their phosphoryl groups to the incoming glucose molecules, therefore spending much of their time unphosphorylated. The unphosphorylated $EIIA^{(Glc)}$ cannot activate adenylate cyclase, causing a reduction in cAMP levels and CRP activity (72).

The unphosphorylated form of $EIIA^{(Glc)}$ also interacts with and inhibits the activity of certain carbohydrate permeases, including permeases for lactose (75), melibiose (76), and maltose (77). This inhibition suppresses secondary carbohydrate catabolism by a mechanism referred to as “inducer exclusion” (78). Many carbohydrates act as transcriptional inducers of genes involved in their own uptake and catabolism, either by activating a transcriptional activator or inhibiting a transcriptional repressor. By inhibiting the activity of the carbohydrate permeases, unphosphorylated $EIIA^{(Glc)}$ prevents these carbohydrates from inducing gene transcription, and prevents the cell from using these carbohydrates as carbon source until the glucose is depleted and $EIIA^{(Glc)}$ is predominately phosphorylated.

Finally, the unphosphorylated form of $EIICB^{(Glc)}$ also participates in catabolite repression by promoting expression of the PTS (79). Mlc is a transcriptional inhibitor that targets the genes encoding EI, HPr, $EIIA^{(Glc)}$, $EIICB^{(Glc)}$, and a few EII subunits specific for other carbohydrates (80). Unphosphorylated $EIICB^{(Glc)}$ binds to and sequesters Mlc, relieving repression of PTS (81).

In the absence of preferred carbohydrates, the phosphorylation state of the PTS is in equilibrium with the PEP:pyruvate ratio (82). A high PEP:pyruvate ratio favors the phosphorylation of the PTS; a low PEP:pyruvate ratio favors its dephosphorylation. Carbon

sources that decrease the PEP:pyruvate ratio, such as glucose 6-phosphate, can cause the dephosphorylation of $EIIA^{(Glc)}$ and $EIICB^{(Glc)}$ even though the carbon sources are not transported or phosphorylated by the PTS. As a result, certain non-PTS carbohydrates can still cause catabolite repression and inducer exclusion by manipulating the phosphorylation state of the PTS.

I will summarize this section in the context of glucose-lactose diauxie. As long as glucose is available, the PTS quickly transfers phosphoryl groups from PEP to glucose, leaving the PTS primarily unphosphorylated. Unphosphorylated $EIICB^{(Glc)}$ sequesters the transcriptional repressor Mlc, increasing expression of the PTS. Unphosphorylated $EIIA^{(Glc)}$ inhibits the uptake of lactose by direct interaction with the lactose permease LacY. In the absence of lactose, allolactose is not generated, so expression of the *lac* genes remains inhibited by LacI. Additionally, unphosphorylated $EIIA^{(Glc)}$ cannot activate AC, so cAMP production and thus CRP activity decreases. This further reduces expression of the *lac* genes, along with genes involved in the catabolism of other secondary carbohydrates. Once the glucose is consumed, phosphoryl transfer ceases and the PTS remains phosphorylated. Mlc is released and inhibits expression of the PTS genes. Lactose enters the cell though LacY and is converted to allolactose, relieving LacI repression of the *lac* operon. Phosphorylated $EIIA^{(Glc)}$ stimulates cAMP production, and cAMP binds to and activates CRP, which drives expression of the *lac* genes.

Overflow Metabolism

Under ideal conditions, bacteria that can tolerate aerobic conditions undergo aerobic respiration to metabolize carbon sources into carbon dioxide, ATP, and other cellular

components. However, if the flux of carbon through glycolysis is greater than the flux of carbon through the TCA cycle, some of the pyruvate and acCoA generated from glycolysis is diverted away from the TCA cycle and into aerobic fermentation in a process called overflow metabolism. By doing so, the cell can avoid many problems that come with a less-than-optimal TCA cycle flux at the expense of a significant amount of energy and biomass. In this section, I will discuss potential causes for a reduction in TCA cycle flux, leading to an imbalance between glycolysis and the TCA cycle. Additionally, I will discuss some of the potential problems that can arise due to this imbalance, along with some of the strategies *E. coli* uses to deal with these problems.

In general, there are two non-mutually exclusive causes for a reduction in TCA cycle flux: 1) reduced expression of TCA cycle proteins, and 2) reduced availability of factors that drive the TCA cycle forward. For example, low molecular oxygen levels prevent the expression of several TCA cycle enzymes, most notably 2-ketoglutarate dehydrogenase (237). Additionally, low molecular oxygen levels prevent NAD^+ re-oxidation via aerobic respiration. The TCA cycle requires 50% more NAD^+ than glycolysis when glucose is the carbon source (**Fig. 5A**), so NAD^+ limitation will cause a net decrease in TCA cycle flux relative to glycolytic flux. Low nitrogen levels prevent the synthesis of glutamate and glutamine from 2-ketoglutarate (95). Removal of some 2-ketoglutarate from the TCA cycle helps to drive the cycle forward, so in the absence of glutamate/glutamine synthesis the TCA cycle flux is decreased. It was recently shown that low magnesium levels also drive aerobic fermentation, as observed by decreased biomass and increased global acetylation (212). This could be due to a decrease in TCA cycle flux, though this hypothesis has not been tested.

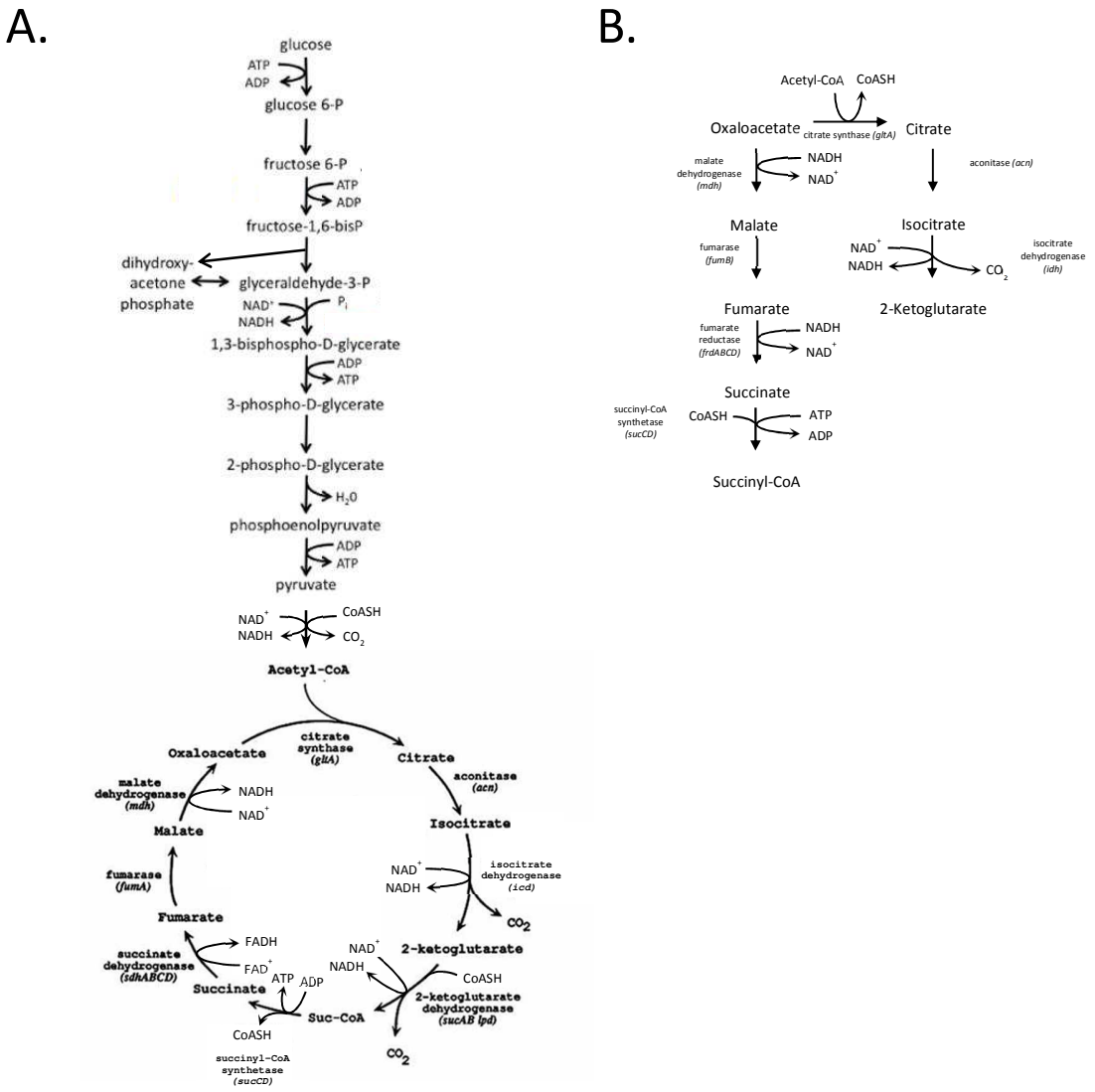

Figure 5. Glycolysis and the TCA cycle. (A) Glucose enters glycolysis at the top of the diagram. After phosphorylation and isomerization events, fructose-1,6-bisphosphate is cleaved into glyceraldehyde-3-phosphate (GAP) and dihydroxyacetone phosphate, which is isomerized to GAP. The two GAP molecules are oxidized into two molecules of acCoA, generating a net gain of two ATP and reducing four NAD^+ per glucose molecule. AcCoA enters the TCA cycle through covalent linkage to oxaloacetate to form citrate. Over several steps, citrate is further oxidized, and the two carbons gained from acCoA are released as CO_2, restoring the original oxaloacetate. Two additional ATP are generated and six additional NAD^+ are reduced per glucose molecule. In the presence of oxygen, the reduced NAD+ can be re-oxidized in the electron transport chain to generate additional ATP. This figure was taken and modified from (90) and (237). (B) In the absence of oxygen, or in the presence of non-limiting catabolite-repressing carbon sources such as glucose, expression of certain enzymes in the TCA cycle is significantly reduced. The most significantly reduced enzyme is 2-ketoglutarate dehydrogenase (KGDH), whose activity is essentially absent. The absence of KGDH breaks the TCA cycle into two branches, a reductive branch (left) and an oxidative branch (right). Neither branch generates ATP and is used only to generate biosynthetic precursors.

Growth on non-limiting concentrations of glucose also reduces carbon flux through the TCA cycle. Expression of the TCA cycle enzymes is reduced in the presence of glucose (84). Most significantly reduced are the 2-ketoglutarate dehydrogenase (KGDH), succinyl-CoA synthetase (SCS), and succinate dehydrogenase (SDH) enzymes encoded in the *sdh-suc* operon, with KGDH activity being essentially non-existent (83). At least some of this downregulation can be attributed to catabolite repression; CRP positively regulates expression of the *sdh-suc* operon (85). Similar downregulation of the *sdh-suc* operon occurs in the absence of oxygen, though different regulators are responsible (86). Reduction in KGDH, SCS, and SDH levels result in a disconnection of the TCA cycle into an oxidative branch and a reductive branch (**Fig. 5B**). Unlike the cyclical version, this branched version of the TCA cycle does not generate ATP and only operates biosynthetically (83) at a much-reduced rate (87).

The absence of a fully functioning TCA cycle can have several negative consequences for the cell. First, carbon sources like glucose are rapidly converted to acCoA (88). The amount of CoA in the cell is relatively constant (89), so it does not take long before all the CoA is sequestered in the form of acCoA. CoA is also required for other biological processes such as fatty acid synthesis, so the acCoA must continually be processed to release free CoA. Another problem for the cell is that in the absence of the TCA cycle, the NADH generated by glycolysis cannot be re-oxidized to NAD^+ through oxidative phosphorylation (90). Glyceraldehyde-3-phosphate dehydrogenase requires NAD^+ to function, so NADH must be re-oxidized by some other means to allow continued glycolytic flux. Finally, the lack of oxidative phosphorylation severely reduces the amount of ATP obtained from carbon sources.

To compensate for these issues, the cells undergo fermentation: anaerobic fermentation when oxygen is limiting, aerobic fermentation when something other than oxygen is limiting. Aerobic fermentation is also referred to as “overflow metabolism”, a seemingly wasteful strategy in which the pyruvate and acCoA that accumulates due to reduced TCA cycle flux “overflows” into fermentation instead of generating energy through respiration, despite the presence of oxygen (243). However, fermentation allows the cell to consume the accumulated pyruvate and acCoA, either generating ATP or re-oxidizing NADH to NAD^+ in the process. Which path the cell takes depends on the current redox and energy status of the cell (90). Converting pyruvate into lactate or acCoA into ethanol re-oxidizes NADH into NAD^+, which is required for glycolysis, but produces no energy. Converting acCoA into acetate generates two ATP per glucose molecule, doubling the amount of ATP generated from acCoA synthesis, but does not recycle NAD^+. Therefore, cells must balance lactate, ethanol, and acetate excretion to get the most energy out of the available glucose, while replenishing NAD^+ for glycolysis. Additionally, ethanol and acetate synthesis releases free CoA back to the cell to use for continued glycolytic flux or for cell growth.

The ATP-forming pathway that generates acetate in *E. coli* involves two enzymes, phosphotransacetylase (Pta) and acetate kinase (AckA) (**Fig. 6**). Pta converts acCoA into acP using inorganic phosphate, and AckA converts acP into acetate by transferring the phosphoryl group to ADP, forming ATP (91). In the presence of high concentrations of acetate (>10 mM), this pathway is reversed, consuming ATP to generate acP, and then acCoA (92, 93). At lower concentrations of acetate, the acetate is converted to acCoA by a single enzyme, acCoA synthetase (Acs), without an acP intermediate. A more detailed description of Acs activity and

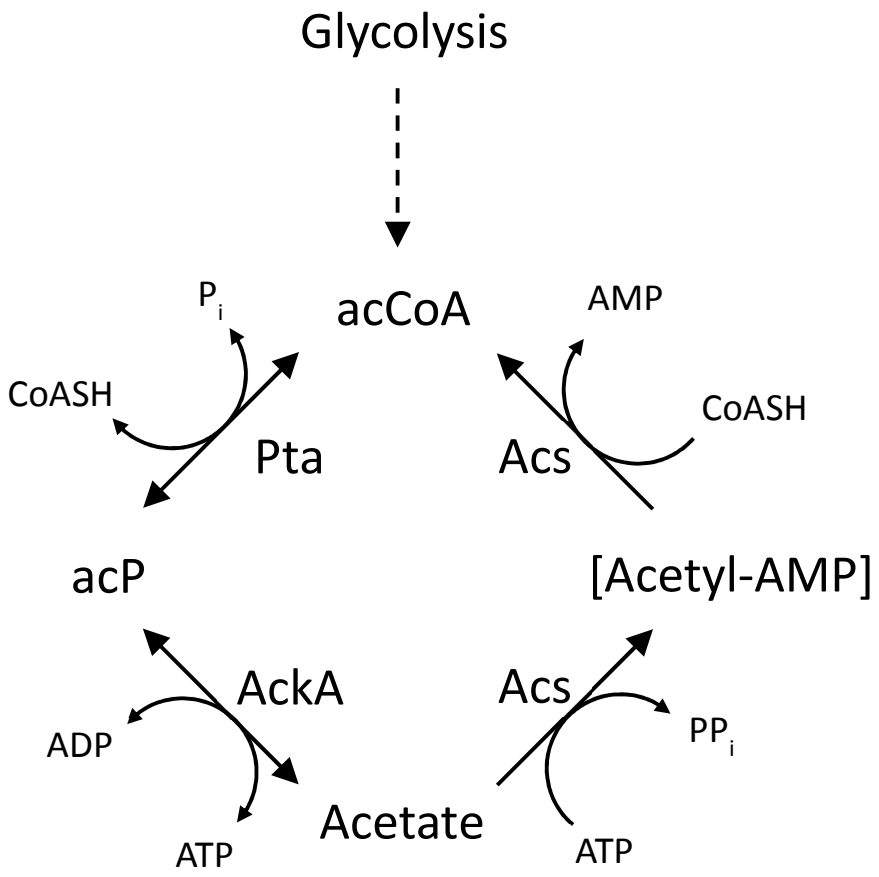

Figure 6. The acetate activation pathway. Acetyl CoA produced from glycolysis is converted to acetyl phosphate by Pta. Acetyl phosphate is then converted to acetate, producing one ATP, and then excreted from the cell. This half of the pathway is reversible in high concentrations of acetate. In low concentrations of acetate, acetate is converted to acetyl CoA by Acs via an acetyl-AMP intermediate. This process requires the hydrolysis of ATP to AMP and is irreversible.

its regulation will be presented in the next section (see Chapter One – Protein Acetylation – Nε-Lysine Acetylation).

Besides simply acting as an intermediate between acCoA and acetate, acP can post-translationally modify proteins by direct transfer of either its acetyl group or phosphoryl group. Protein acetylation is a recently discovered function of acP and will be discussed in detail in the next section (see Chapter One – Protein Acetylation – Nε-Lysine Acetylation). Protein phosphorylation by acP has been best studied in the context of two-component signal transduction (TCST) systems. In TCST systems, a histidine kinase senses a signal and autophosphorylates using ATP. The phosphoryl group is then passed to a response regulator, either directly or through a series of intermediate phosphorylation events. Phosphorylation of the response regulator alters its behavior, culminating in a cellular response to the original signal. AcP can directly phosphorylate many response regulators, providing an alternative activation mechanism independent of their associated histidine kinases (90, 94). The ability of acP to post-translationally modify proteins suggests acP levels are a signal sensed by the cell, perhaps to monitor a reduction carbon flux through the TCA cycle relative to carbon flux through glycolysis (95).

Protein Acetylation

Post-translational modifications permit cells to respond to changing environmental conditions by modifying existing proteins, altering their activity, localization, interactions with other proteins or nucleic acids, and/or stability (96). Many modifications are reversible, enabling a temporal adjustment to a protein's function. A single protein can carry multiple

different modifications, and cross-talk can occur between modifications (97). Each combination of modifications on a particular protein could be considered a unique isoform, greatly expanding the proteome beyond the constraints of the genome (98, 99).

One of the most common post-translational modifications in *E. coli*, and a subject of this dissertation, is protein acetylation. In this section, I will discuss protein acetylation in three parts: Nα-acetylation, Nε-lysine acetylation, and O-acetylation.

Nα-Acetylation

Nα-acetylation is the irreversible transfer of the acetyl group of acCoA to the α-amino group in the N-terminal amino acid of a protein. This modification is extremely common in eukaryotes, occurring on 60-80% of proteins (100). In bacteria, Nα-acetylation is much more rare. Only 47 *E. coli* proteins have been detected as Nα-acetylated, comprising only 1% of its genome (101). In *Pseudomonas aeruginosa*, the number of identified Nα-acetylated proteins is 117, totaling about 2% of its genome (102).

Eukaryotes possess up to six different N-terminal acetyltransferases (NATs), though three of these NATs (NatA-C) catalyze as many as 90% of the Nα-acetylation reactions in eukaryotes (103). These NATs recognize their targets based on the sequence of the two initial amino acids residues in the target protein, and whether the initial methionine has been cleaved by an aminopeptidase. NATs associate with ribosomes, allowing the acetylation reaction to occur co-translationally in most cases (104). Nα-acetylation regulates protein activity in many ways, including protein-protein interactions (105), localization (106), sorting (107), folding (108), and proteasomal degradation (109), and has been implicated in human disease (110, 111) and cancer (112, 113).

E. coli encodes three NATs, RimJ, RimL, and RimI. Despite having related functions, the bacterial NATs are not related to the eukaryotic NATs, suggesting that each group evolved independently (114). Unlike in eukaryotes, Nα-acetylation is performed post-translationally in bacteria (103). RimJ, RimL, and RimI Nα-acetylate ribosomal proteins S5 (115), L12 (116), and S18 (117), respectively. No other endogenous targets of the Rim proteins have been identified, and it is not known which proteins acetylate the other Nα-acetylated proteins in *E. coli*, suggesting the existence of additional NATs. Loss of either RimI or RimL results in no known phenotype (116, 117). RimJ appears to play a role in ribosome assembly (118) and transcription of the *pap* operon (119), though these roles may be independent of the acetyltransferase activity of RimJ. As a result, the significance of Nα-acetylation in *E. coli* has yet to be determined.

Nε-Lysine Acetylation

In contrast with Nα-acetylation, Nε-lysine acetylation is the transfer of an acetyl group specifically onto the ε-amino group of a lysine residue (**Fig. 7**). Using acCoA as the acetyl group donor, this reaction is catalyzed by lysine acetyltransferases (KATs), and in some cases can be reversed by lysine deacetylases (KDACs). AcP can also function as a non-enzymatic acetyl group donor in some bacteria (120, 121). Nε-lysine acetylation results in the neutralization of the lysine's positive charge, which can affect the protein's activity, localization, interactions with other proteins or nucleic acids, and/or stability (96).

Nε-lysine acetylation was first identified as a post-translational modification of histones in eukaryotes (122). Since then, it has been realized that Nε-lysine acetylation is one of many histone modifications that make up a "histone code," in which combinations of post-

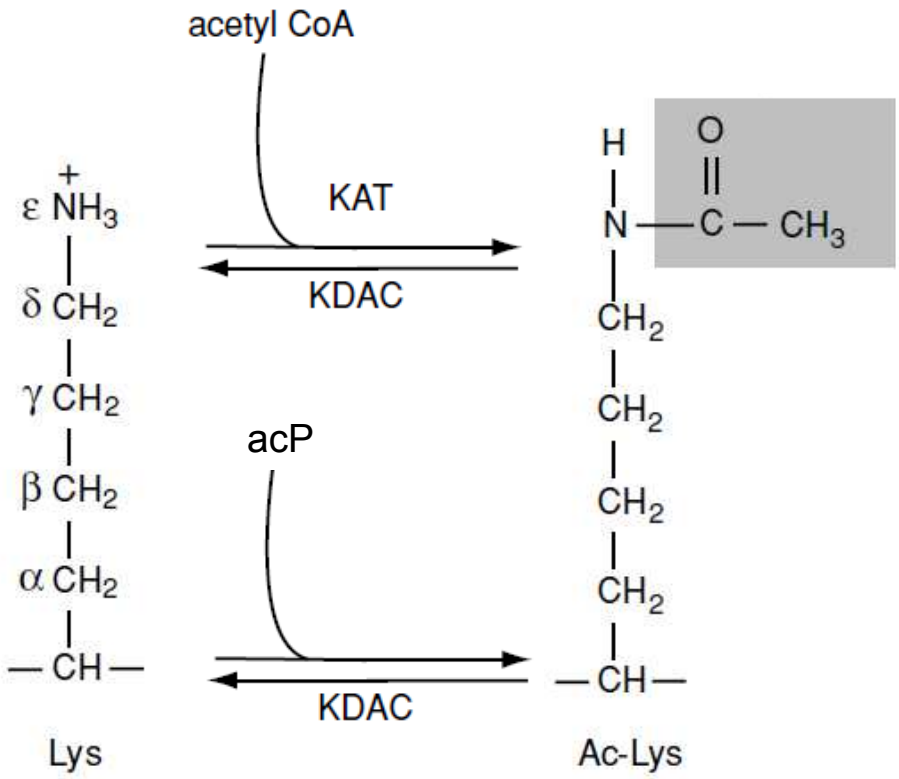

Figure 7. Nε-lysine acetylation. The acetyl group from acCoA can be transferred to the ε-amino group of a lysine sidechain through the action of a lysine acetyltransferase (KAT). Some of these acetylations can be reversed by the action of a lysine deacetylase (KDAC). In some bacteria, acP can also be an acetyl group donor. Acetylation by acP does not appear to require a KAT. Much like KAT-dependent acetylation, some of the acP-dependent acetylations can be reversed by KDACs. Figure adapted from (238).

translational modifications on histone tails have specific effects on local transcription (123). Nε-lysine acetylation affects histones by loosening their interaction with DNA (124) and recruiting other histone-modifying enzymes (125) or DNA binding proteins (126).

Histones are not the only eukaryotic proteins that undergo Nε-lysine acetylation. Thousands of human proteins have been detected as acetylated in the nuclear, cytoplasmic, and mitochondrial compartments (127, 128). For example, acetylation of the transcription factor p53 promotes p53 activity by recruiting additional coactivators and histone acetyltransferases, enhancing the activation of p53-dependent promoters (129). Stable α-tubulin polymers are highly acetylated, while overexpression of the KDAC HDAC-6 induces depolymerization (130). Almost every enzyme in glycolysis, gluconeogenesis, the TCA cycle, the urea cycle, fatty acid metabolism, and glycogen metabolism is acetylated at some level, and the acetylation pattern changes based on nutrient availability, though the function of most of these acetylations remains unknown (128). However, since acCoA is the acetyl group donor in eukaryotes and acCoA is the keystone intermediate of central metabolism, acetylation of metabolic enzymes may allow the cell to monitor its metabolic status.

Nε-lysine acetylation is also a highly abundant modification in bacteria, occurring on thousands of lysines spanning hundreds of proteins of diverse function. It has been identified in *E. coli* (120, 121, 131–137), *S. enterica* (138), *B. subtilis* (139, 140), *B. amyloliquefaciens* (141), *G. kaustophilus* (142), *E. amylovora* (143), *M. tuberculosis* (144), *S. eriocheiris* (145), *T. thermophilus* (146), and *C. glutamicum* (147). Despite the vast number and diversity of acetylated proteins, very few KATs have been identified. *E. coli* encodes one known KAT, YfiQ (148) (also called PatZ or Pka in *E. coli* and Pat in *S. enterica*). Some, but not most, of the YfiQ-

mediated acetylations are reversed by the NAD^+-dependent KDAC CobB (149). A recently discovered NAD^+-independent KDAC YcgC deacetylates a set of proteins that is at least partially distinct from CobB (150), suggesting protein deacetylation is highly regulated.

In addition to KAT-mediated Nε-lysine acetylation, non-enzymatic Nε-lysine acetylation can occur in bacteria that synthesize acP, a metabolic intermediate that can function as either a phosphoryl or acetyl group donor (90, 120, 121, 139). In *E. coli* and *B. subtilis*, acP-dependent Nε-lysine acetylation appears to be the dominant acetylation mechanism. While there is no clear consensus sequence to identify lysines susceptible to acP-mediated acetylation, susceptible lysines appear to have two characteristics in common: nearby positively charged or hydrogen bond-donating residues to interact with the negatively charged phosphoryl group, and nearby negatively charged residues to deprotonate and enhance the reactivity of the target lysine (121).

Much like in eukaryotes, metabolic enzymes are highly acetylated in bacteria (131, 138, 151). This has led to the hypothesis that metabolism and Nε-lysine acetylation are linked. Glucose greatly enhances the number of Nε-lysine acetylations in both *S. enterica* and *E. coli* (138, 152). In *S. enterica*, it is suggested that Nε-lysine acetylation by Pat and deacetylation by CobB promote glycolysis and gluconeogenesis respectively (138). In *E. coli*, CRP, a major regulator of secondary metabolism, is also required for the majority of these acetylations (153). The mechanism by which CRP promotes Nε-lysine acetylation is not established; however, cAMP, which is required for CRP activity, enhances expression of the KAT YfiQ (151). YfiQ-dependent Nε-lysine acetylation only makes up a fraction of total Nε-lysine acetylation in *E. coli* (120, 121), so other mechanisms by which CRP regulates Nε-lysine acetylation must exist.

Whether CRP regulates YfiQ expression, and whether CRP contributes to glucose-dependent Nε-lysine acetylation in other ways, has yet to be determined.

One of the first and best-studied example of Nε-lysine acetylation is the regulation of acetyl-CoA synthetase (Acs) activity (154). Acs synthesizes acCoA from acetate through the formation of an acetyl-AMP intermediate. In *S. enterica*, Pat acetylates Acs on K609, which inhibits the formation of the acetyl-AMP intermediate, but has no effect on the conversion of acetyl-AMP to acCoA (154, 155). Acetylation of K609 by Pat requires Acs L641, which is predicted to be involved in the Pat-Acs interaction (156). The NAD^+-dependent deacetylase activity of CobB reverses K609 acetylation, restoring Acs activity (154). Thus, when the cell has plenty of energy, high concentrations of acCoA provide the acetyl group donor to inhibit Acs activity and acetate uptake through Nε-lysine acetylation. On the other hand, in the absence of glycolytic carbon sources, oxidized NAD^+ accumulates. This provides a substrate for CobB to deacetylate Acs and activate acetate metabolism. Regulation of Acs activity by Nε-lysine acetylation is similar in *E. coli* (151) and *B. subtilis* (157), though activation of Acs in *B. subtilis* requires two deacetylases.

RNase R is an exoribonuclease in *E. coli* responsible for the degradation of highly structured RNAs (158). In exponential phase growth, RNase R is Nε-lysine acetylated by YfiQ on K544 (148). Acetylated RNase R is bound by both tmRNA and its accessory protein SmpB, members of the *trans*-translation pathway that rescues stalled ribosomes (159, 160). This binding induces RNase R degradation by the Lon protease (161). In stationary phase, RNase R is no longer acetylated by YfiQ (162). The absence of K544 acetylation prevents the interaction between RNase R and the tmRNA-SmpB complex, saving RNase R from Lon-dependent

degradation. Unlike Acs acetylation, RNase R K544 acetylation appears to be irreversible. Instead, the growth phase-dependent acetylation pattern of RNase R follows the expression pattern of YfiQ; YfiQ steady state levels are high in early exponential phase and gradually decrease into stationary phase (162). This is one example of Nε-lysine acetylation regulating the stability of a protein.

There is evidence that transcriptional regulators may also be regulated by Nε-lysine acetylation. The regulator of capsule synthesis B (RcsB) is the response regulator of the Rcs phosphorelay system (163). RcsB becomes phosphorylated and binds to DNA, regulating the transcription of genes involved with biofilm maturation, motility, and cell division. AcP-dependent acetylation of RcsB K154 reduces RcsB-dependent *rprA* promoter activity to baseline levels without interfering with RcsB DNA binding (164). YfiQ also acetylates RcsB *in vitro* on K180, located within the RcsB DNA binding domain, which does reduce binding to the *flhDC* promoter (165). However, whether YfiQ acetylates K180 *in vivo* has been called into question (164).

Nε-lysine acetylation of CRP has been reported on several occasions (120, 121, 131–137, 166) , but the role of Nε-lysine acetylation on CRP function has not been investigated. One of the first and most consistently-identified acetylated lysines within CRP is K100 (120, 121, 131–134, 136), which is adjacent to K101 and AR2, a region critical for CRP activity at Class II promoters (25). Several other lysines have been identified as acetylated, some of which are also very near or within regions critical for CRP function (**Fig. 2**). For example, K188 (120, 121, 133, 134) is very close to the DNA-binding HTH motif. K166 (120, 121, 132, 133, 136) and K201 (120, 121, 133, 134, 166) are near AR1. K101 (120, 133) and K52 (120, 133, 135) are components of

AR2 and AR3, respectively. Acetylation of any of these lysines could have a significant impact on CRP function. Considering the scope of the CRP regulon (12), CRP acetylation could also have a significant impact on cell physiology.

It is clear that Nε-lysine acetylation is a widespread phenomenon, and it has been compared to protein phosphorylation in terms of abundance, mechanism, and regulatory capacity (167, 168). Unlike protein phosphorylation, very little is known about the effect of individual acetylations, due to the infancy of the field. With such a limited understanding, the effect of Nε-lysine acetylation on a protein's function currently must be studied on a case-by-case basis.

O-Acetylation

In addition to the better-known protein acetylations discussed above, some species of pathogenic bacteria are capable of acetylating the hydroxyl sidechains of internal serine and threonine histidine residues within eukaryotic proteins. One proposed function of serine/threonine acetylation is to prevent the phosphorylation of these regulatory serines and threonines in the target proteins. For example, *Yersinia* species encode a Ser/Thr acetyltransferase YopJ, which is secreted into mammalian host cells to block innate immune signaling (169). Once inside the host cell, YopJ acetylates lysine, serine, and threonine residues within MAPKKs and IKKβ. The serine and threonine acetylations prevent phosphorylation of these key residues, inhibiting the activation of the MAPKKs and IKKβ. YopJ also acetylates serine, threonine, and lysine residues within TGFβ-activated kinase (TAK1) in both *Drosophila* and mammals, resulting in TAK1 inactivation (170).

Another YopJ family member, HopZ3 from *Pseudomonas syringae*, also prevents immune signaling in *Arabidopsis* through protein acetylation (171). HopZ3 acetylates several immunity-related proteins in *Arabidopsis* on lysine, serine, and threonine residues, along with the acetylation of a single histidine residue within AvrB3. While serine, threonine, and even tyrosine acetylation has been observed in histones (172), YopJ family members are the only known mediators of stable histidine acetylation. The continued discovery of novel protein acetylation modifications and their crosstalk with other modifications implies that protein acetylation plays a much bigger regulatory role than previously thought.

Summary

Bacteria must regulate their proteome to ensure the right proteins are not only expressed but also active at the right time. Some of this regulation takes the form of PTMs. These PTMs alter the function of proteins, either as a binary "on-off" switch or as a subtler rheostat. Another part of this regulation is in the form of transcriptional regulation: the control of the timing of gene expression, which ultimately leads to protein expression. In this dissertation, I propose an intimate relationship between these two modes of regulation.

In Chapter Three, I investigate the involvement of CRP in Nε-lysine acetylation. CRP is required for the majority of glucose-induced acetylation. cAMP, the allosteric activator of CRP, promotes the expression of *yfiQ*, the only reported KAT in *E. coli*. Does CRP promote *yfiQ* expression as well? Regardless, the relatively few YfiQ-mediated acetylations cannot account for all CRP-dependent acetylation. Could CRP also promote acP-dependent acetylation?

In Chapter Four, I investigate the involvement of Nε-lysine acetylation in CRP activity. Several lysines in CRP can be acetylated, including K100. The proximity of K100 to AR2 suggests K100 acetylation could influence Class II transcription, while having little impact on Class I transcription. The existence of two different classes of CRP-dependent promoters suggests the possibility that each class could be regulated separately, though no mechanism of differential regulation has been proposed. Could K100 acetylation be a mechanism to separate CRP-dependent Class II activity from Class I activity?

CHAPTER TWO

MATERIALS AND METHODS

Bacterial Strains, Plasmids, and Primers

All bacterial strains, plasmids, and primers used in this dissertation are listed in **Table 1**. Mutant strains were constructed by generalized transduction using P1kc (for details, see "Generalized P1 Transduction").

Culture Conditions

For strain construction, cells were grown in Luria Broth (LB) containing 1% (w/v) tryptone, 0.5% (w/v) yeast extract, and 0.5% (w/v) sodium chloride; LB plates also contained 1.5% agar. Unless otherwise noted, for all assays that required culturing bacteria in liquid media, cells were grown in tryptone broth buffered to pH 7 (TB7), containing 1% (w/v) tryptone, 61.5 mM potassium phosphate dibasic, and 38.5 mM potassium phosphate monobasic. Cultures were grown with aeration by shaking at 225 rpm at 37°C unless otherwise noted. Cell growth was monitored spectrophotometrically (DU640; Beckman Instruments, Fullerton, CA) by determining the optical density at 600 nm (OD_{600}).

Antibiotics were prepared as a stock solution 1000 times the working concentration. Working concentrations were as follows: ampicillin, 100 μg/mL; kanamycin, 40 μg/mL; tetracycline, 15 μg/mL; and chloramphenicol, 25 μg/mL. Stock solutions of ampicillin and

kanamycin were dissolved in water and filter sterilized. Stock solutions of tetracycline and chloramphenicol were dissolved in 50% and 100% ethanol, respectively, and filter sterilized. All antibiotic stocks were stored at -20°C.

Generalized P1 Transduction

1 mL of an overnight culture grown in 5 mL LB (plus antibiotic if appropriate) was pelleted and resuspended in 1 mL TBT (TB with 0.2% [w/v] glucose, 10 mM calcium chloride, 10 mM magnesium sulfate, 0.04 mM ferric chloride). 100 µL phage lysate was added to the resuspended cells and incubated statically at 37°C for 30 minutes. 200 µL of 1 M sodium citrate, pH 5.5 was added to the infected cells to prevent further infection. The infected cells were pelleted and resuspended in 500 µL LB, then 200 µL of 1 M sodium citrate, pH 5.5 was added to the resuspension. The cells were incubated statically at 37°C for 70 minutes. The cells were pelleted and resuspended in 100 µL sodium citrate, pH 5.5. The entire 100 µL of transduced cells was plated onto LB plates containing the appropriate antibiotics and incubated at room temperature for 48 hours, or until colonies appeared.

Plasmid Construction

Plasmid p770-*yfiQ'* is a derivative of p770 (a generous gift from Rick Gourse at the University of Wisconsin-Madison). It is an ampicillin-resistant ColE1-based plasmid that has the tandem *rrnB* operon terminators T1 and T2 located downstream of the insertion sites for the promoter. The *yfiQ* promoter region was amplified from strain AJW678 with primer pair *yfiQ*' *Eco*RI F and *yfiQ*' *Hin*DIII R to insert the EcoRI site upstream of the promoter region and the

*Hin*DIII site downstream. The resulting amplicon was ligated into pJET1.2, using the CloneJET PCR Cloning Kit (Fermentas) following kit instructions. The ligated *yfiQ'* insert was sequenced, digested out of pJET1.2 with *Eco*RI and *Hin*DIII, and gel-extracted. The gel-extracted fragment was cloned into previously digested p770 via overnight ligation at room temperature. Ligated plasmids were transformed into *E. coli* DH5α cells. Plasmids were recovered from transformants and restriction digested to screen for the presence of the *yfiQ* promoter. The presence of the promoter was confirmed via sequencing analysis. The plasmid containing the correct insert was purified, stored at -20°C and named p770-*yfiQ'*.

Site-Directed Mutagenesis

Site-directed mutagenesis of *crp* in pDCRP was performed using either the QuikChange II Site-Directed Mutagenesis Kit or the QuikChange Lightning Multi Site-Directed Mutagenesis Kit (Agilent Technologies) per the manufacturer's instructions. Primers used for site-directed mutagenesis can be found in **Table 1**. Mutations were confirmed by sequence analysis of the purified mutagenized plasmid.

Transformation

Three distinct transformation methods were used: transformation buffers (TBF), transformation storage solution (TSS), and electroporation.

TBF

Overnight cultures grown in LB (plus antibiotic if appropriate) were subcultured into 10 mL LB (plus antibiotic if appropriate) and shaken at 225-250 rpm at 37°C until the OD_{600} reached

0.4-0.6. The culture was cooled on ice for 5 minutes and then pelleted. The supernatant was removed and the pellet was resuspended with 4 mL cold TBF1 (30 mM potassium acetate, 100 mM potassium chloride, 10 mM calcium chloride, 50 mM manganese chloride, 15% glycerol, pH to 5.8 with acetic acid). The cells were pelleted, supernatant removed, then the pellet resuspended with 0.4 mL cold TBF2 (10 mM MOPS, 75 mM calcium chloride, 10 mM potassium chloride, 15% glycerol, pH to 6.5 with potassium hydroxide). The cells were chilled on ice for 30-60 minutes before use. 1 µL plasmid DNA was added to 50 µL chemically competent cells and chilled on ice for 10 minutes. The cell-DNA mixture was heat shocked in a waterbath for 45 seconds at 42°C, then chilled on ice for 2 minutes. 1 mL LB was added to the transformed cells, and the cells were shaken at 225-250 rpm at 37°C for one hour. 100 µL of the transformed cells were plated onto LB plates containing the appropriate antibiotics and incubated at 37°C overnight, or until colonies appeared.

TSS

Overnight cultures grown in LB (plus antibiotic if appropriate) were subcultured into 5 mL LB (plus antibiotic if appropriate) and shaken at 225-250 rpm at 37°C until the OD_{600} reached 0.3-0.4. 1 mL of the culture was pelleted and resuspended in 100 µL TSS (LB containing 10% [w/v] PEG, 5% (v/v) DMSO, 50 mM magnesium sulfate). 1 µL plasmid DNA was added to the cells and chilled on ice for 30 minutes. The cell-DNA mixture was heat shocked in a waterbath for 2 minutes at 42°C. 900 µL LB was added to the transformed cells, and the cells were shaken at 225-250 rpm at 37°C for one hour. 100 µL of the transformed cells were plated onto LB plates containing the appropriate antibiotics and incubated at 37°C overnight, or until colonies appeared.

Electroporation

Overnight cultures grown in LB (plus antibiotic if appropriate) were subcultured into 35 mL LB (plus antibiotic if appropriate) and shaken at 225-250 rpm at 37°C until the OD_{600} reached 0.4-0.6. 30 mL of culture (divided into two 15 mL aliquots) were centrifuged at 6700 rpm for 7 minutes. The supernatant was removed and each pellet was resuspended in 1 mL cold water. An additional 9 mL cold water was added to each cell suspension, bringing the volume to 10 mL each. The cell suspensions were centrifuged at 6700 rpm for 7 minutes. The supernatant was removed and each pellet was resuspended in 1 mL cold water. Each cell suspension was pelleted, the supernatants were removed, and each pellet was resuspended in 100 µL cold water for a total of 200 uL electrocompetent cells. 1 µL plasmid DNA was added to 50 µL electrocompetent cells, which was then transferred to a chilled electroporation cuvette with a 0.1 cm gap and electroporated at 1.8 kV. 1 mL LB was added to the transformed cells, and the cells were shaken at 225-250 rpm at 37°C for one hour. 100 µL of the transformed cells were plated onto LB plates containing the appropriate antibiotics and incubated at 37°C overnight, or until colonies appeared.

CRP Expression and Purification

CRP expression was carried out in BL21(DE3) cells carrying pET15b-crp. Overnight cultures grown in LB with ampicillin were subcultured into 500 mL LB with ampicillin and shaken at 225-250 rpm at 37°C until the OD_{600} reached 0.4-0.6. 500 µL of 1 M IPTG was added to the culture to induce CRP expression, then the culture was allowed to shake for 3 more hours. The cells were divided into two 250 mL aliquots and centrifuged at 5000 rpm for 15

minutes to pellet the cells. The supernatant was removed and the pellets were stored at -80°C overnight. After thawing, each pellet was resuspended in 5 mL BugBuster (Novagen), 5 µL lysozyme, and 5 µL Benzonase. The resuspensions were incubated with shaking at room temperature for 30-40 minutes to lyse the cells. 15 mL Resuspension Buffer (50 mM disodium phosphate, 1.4 M sodium chloride, 20 mM imidazole, 0.1% [v/v] Tween-20, 5% [v/v] ethanol, 15 mM β-mercaptoethanol, pH 8.0) was added to each resuspension, then the resuspensions were centrifuged at 11500 rpm for 30 min at 4°C. The supernatant was collected and loaded onto a 500 µL Ni-NTA His-Bind Resin (Novagen) column. The column was washed with NiNTA Wash buffer (50 mM disodium phosphate, 0.3 M sodium chloride, 30 mM imidazole, 0.1% [v/v] Tween-20, 5% [v/v] ethanol, pH 8.0). CRP was eluted from the column by passing 1 mL volumes of NiNTA Wash buffer containing increasing concentrations of imidazole (50 mM, 75 mM, 100 mM, 150 mM, and 200 mM). The 50 mM, 75 mM, and 100 mM imidazole elutions were pooled and dialyzed in 1 L Storage Buffer (10 mM Tris-HCl pH 8.0, 100 mM potassium chloride, 50% [v/v] glycerol, 10 mM magnesium chloride, 0.1 mM EDTA, 1 mM DTT) overnight. The purified CRP was divided into aliquots and stored at -80°C for future use. This protocol was modified from (173).

In Vitro **Transcription**

In vitro transcription from the *yfiQ* promoter was executed by multiple round transcriptions. The final volume of the *in vitro* transcription reaction was 25 µL and contained 1x *in vitro* transcription buffer (40 mM Tris-HCl pH 8.0, 10mM MgCl2, 50 mM KCl, 1mM DTT and 25 ng BSA), 1X NTP mixture (200 µM ATP, CTP and GTP, and 10 µM UTP), 2 µCi alpha [32P]UTP,

50 ng p770-*yfiQ'* plasmid, purified CRP at concentrations ranging from 2.5-40 µM, 200 µM cAMP, and 10 nM of purified RNA polymerase-containing sigma 70 (Eσ70). When multiple reactions were performed per experiment, they were started at 30 second intervals by the addition of Eσ70 and incubated at 30°C for 15 minutes. After incubation, the reactions were stopped (also at 30-second intervals) by the addition of 25 µl of 2x stop solution (7 M Urea, 10 mM EDTA, 1% SDS, 2X TBE and 0.05% bromophenol blue). Immediately following the addition of the stop solution, the reactions were transferred onto ice.

The transcript fragments were separated using a denaturing acrylamide gel containing (6% acrylamide [19:1 acryl.:bis.], 7 M urea and 1x TBE). Polymerization was initiated by the addition of 0.01% of freshly made 10% APS solution and 0.001% TEMED and allowed to polymerize for 1 hour and 30 minutes undisturbed at room temperature.

23 µL of the *in vitro* transcription/stop solution reaction was loaded into the gel wells. The samples were allowed to run at approximately 200V for approximately 1.5-2 hours until the dye reached a position approximately 3 inches from the bottom of the gel, which keeps the unincorporated radioactivity on the gel and makes cleaning up the radioactive waste much easier. 1x TBE was used as the running buffer. Prior to phosphor imaging, the gels were dried using a gel drier for 2 hours. Images were obtained by exposing a storage phosphor screen to the dry gel overnight. The screen was read using a Typhon phosphoimager. The unused *in vitro* transcription reaction was stored at -20°C.

qRT-PCR

For quantification of *yfiQ* and *crp* mRNA, cells were grown in TB7 supplemented with 22 mM glucose at 37°C and harvested in late exponential phase (OD_{600} ~1.5). For quantification of *flhD* mRNA, cells were grown in 1% (w/v) tryptone and 0.5% sodium chloride at 30°C and harvested in mid-exponential phase (OD_{600} ~0.5).

1 mL culture was added to 2 mL RNAProtect Bacterial Reagent (Qiagen), and vortexed. After centrifugation, the cell pellet was frozen at -80°C and stored overnight. RNA was extracted using the MasterPure RNA Purification kit (EpiCentre). To extract RNA, the cell pellets were resuspended by adding 300 µl Tissue and Cell lysis solution and 2 µl proteinase K. The resuspensions were incubated at 65°C for 15 minutes, vortexing every 5 minutes. After cooling the resuspensions to room temperature, protein was precipitated by adding 175 µL MPC Protein Precipitation Reagent, briefly vortexing, then centrifuging for 20 minutes at 13500 rpm and 4°C. The supernatant was transferred to a new tube and the nucleic acid was precipitated by adding 500 µL isopropanol. The mixture was centrifuged for 30 minutes at 13500 rpm and 4°C. After discarding the supernatant, the pellet was resuspended in 195 µL 1x DNase Buffer and 5 uL RNase-free DNase I, and incubated for 30 minutes at 37°C. After incubation, 200 µL 2x T and C Lysis solution was added, and the mixture was immediately vortexed. 200 µL MPC Protein Precipitation reagent was added, and the mixture was immediately vortexed again. The mixture was centrifuged for 20 minutes at 13500 rpm and 4°C. The supernatant was transferred to a new 1.5 mL Eppendorf tube. The ribonucleic acid was precipitated by adding 500 µL isopropanol, mixing by inversion, and centrifuging for 30 minutes at 13500 rpm and 4°C. The supernatant was discarded, and the pellet was rinsed twice with 200 µL 75% ethanol. After

removing any residual ethanol, the pellet was resuspended in 30 uL RNase-free water. 1 µL RiboGuard RNase inhibitor (EpiCentre) was added to the RNA. RNA concentration was measured using a NanoDrop 2000 spectrophotometer (Thermo). RNA was stored at -80°C until cDNA synthesis.

1 µg RNA was used to synthesize cDNA, using the iScript cDNA Synthesis Kit (BioRad), according to the manufacturer's protocol. Quantitative reverse transcription PCR (qRT-PCR) was performed using iTaq Universal SYBR Green Supermix (BioRad), according to manufacturer's protocol, on a CFX96 Real-Time System (BioRad) with a C1000 Thermal Cycler (BioRad) and the following conditions: 95°C for 5 min, then 37 cycles of (95°C for 15 sec, 60°C for 20 sec, 72°C for 30 sec). Quantitation of 16S rRNA was used to normalize the data. The list of primers used can be found in **Table 1**. Each experiment included at least two replicates and was performed at least twice.

Western Immunoblot Analysis

Cultures were grown at 37°C in TB7 supplemented with 22 mM glucose. At the indicated timepoints cells were harvested, pelleted, and lysed with Bugbuster protein extraction reagent (Novagen 70584) using the manufacturer's protocol. The whole cell protein lysate was normalized to protein concentration by BCA assay (Pierce 23225), and equal amounts of protein were boiled in 2x loading buffer [0.1 M Tris pH 6.8, 4% (w/v) SDS, 12% (v/v) β-mercaptoethanol, 20% (v/v) glycerol, 0.001% (w/v) bromophenol blue] for 5 minutes. The samples were separated by SDS-polyacrylamide gel electrophoresis with 4.6 M urea in 1x running buffer [25 mM Tris, 200 mM glycine, 3.5 mM SDS]. The gel was transferred to a nitrocellulose membrane in 1x

transfer buffer [25 mM Tris, 200 mM glycine, 3.5 mM SDS, 20% (v/v) methanol] for 1.5 hr at 100V at 4°C. The blot was blocked with 5% (w/v) BSA prepared in PBST [PBS containing 0.05% (v/v) Tween-20] for 30 min at room temperature. To detect and quantify CRP, monoclonal mouse primary antibodies against CRP (BioLegend 664304, 1:2000) and RNAP α subunit (BioLegend 663102, 1:2000) were diluted in 5% BSA in PBST at 4°C overnight. To detect and quantify acetylation, a polyclonal rabbit primary antibody against acetyllysine (Cell Signaling #9441L, 1:2000) was diluted in 5% BSA in PBST at 4°C overnight. The blot was washed three times for 5 min each with TBST [TBS containing 0.05% (v/v) Tween-20] and then incubated with either HRP-conjugated goat anti-mouse secondary antibody (to detect CRP/RNAP α subunit, Millipore AP503P, 1:2500) or HRP-conjugated goat anti-rabbit secondary antibody (to detect acetylation, Cell Signaling #7074, 1:5000) in 5% milk in TBST for 2 h at room temperature. The blot was washed three times for 5 min each with TBST and exposed using 20X LumiGLO Reagent and Peroxide (Cell Signaling 7003) and detected using a FluorChem E imager (ProteinSimple). All experiments were performed at least twice.

To quantify CRP steady state levels, the intensities of the bands corresponding to CRP and RNAP α subunit were quantified from the digital image obtained from the FluorChem E imager using ImageJ version 1.47 (imagej.nih.gov/ij). For each sample, the raw intensity of the band corresponding to CRP was divided by the raw intensity of the band corresponding to RNAP α subunit to normalize CRP values across samples for each experiment.

Promoter Activity Assays

Cells were grown at 37°C in TB7 supplemented with 0.4% glucose. To monitor promoter activity from *CC(-41.5)* and *CC(-61.5)*, 50 µl culture aliquots were harvested at regular intervals and added to 50 µl of All-in-One β-galactosidase reagent (Pierce Biochemical). β-galactosidase activity was determined quantitatively using a microtiter format, as described previously (174). As a blank, 50 µl of sterile TB7 was used. Each experiment included three biological replicates. All experiments were performed at least twice.

Model Characterization

To visualize interactions between CRP and RNAP at Class II promoter, we created a structural model of CRP-RNAP-promoter complex. The structure of the ternary complex of CRP and RNAP at a Class II promoter was modeled in Coot (175) based on the three-dimensional EM structure of a complex comprising *E. coli* CRP, RNAP and a DNA fragment from a Class I CRP-dependent promoter (176). The crystal structure of CRP crystallized in presence of αCTD and DNA suggests how the αCTD could interact with AR1 at a Class II promoter and provides details of αCTD binding at a non-inhibitory location (177). Interactions between CRP AR2 and RNAP involve residues 162-165 located on the flexible loop of N-terminal domain of RNAP αNTD (25). To model these interactions, we used a crystal structure of RNAP in complex with squaramide compound (178), where the position of the flexible loop is structurally well defined. Interactions between the CRP AR3 region and RNAP were modeled based on data described previously (57). The figures of model were created in the CCP4 molecular graphics (CCP4mg) program (179).

Motility Assays

Overnight cultures grown in LB supplemented were diluted to OD_{600} 3.0-3.5. For each strain, 5 µL of diluted culture was spotted onto semi-solid agar plates (1% [w/v] tryptone, 0.5% [w/v] sodium chloride, 0.25% [w/v] agar) and allowed to dry for 5 minutes. Once dry, the plates were incubated at 30°C for 10 hours. At the end of 10 hours, a ruler was used to measure the diameter of each spot. Each experiment included at least eight replicates and was performed twice.

In Vitro Acetylation

Purified CRP (0.5 g/L) and the indicated concentrations of acP were added to acetylation buffer (150 mM Tris pH 7.3, 10 mM magnesium chloride, 150 mM sodium chloride, 10% glycerol). The acetylation reaction was incubated for either 15 minutes or 120 minutes at 37°C. Immediately following incubation, an equal volume of 2x loading buffer [0.1 M Tris pH 6.8, 4% (w/v) SDS, 12% (v/v) β-mercaptoethanol, 20% (v/v) glycerol, 0.001% (w/v) bromophenol blue] was added to each reaction, and the reactions were incubated at 95°C for 10 minutes. After incubation, the reactions were cooled to room temperature and were separated by SDS-polyacrylamide gel electrophoresis and subjected to Western immunoblot analysis using a polyclonal anti-acetyllysine antibody as described above.

Quantitative Mass Spectrometry

In vitro acetylation of CRP was performed in duplicate as described above by incubating 63.5 ng μl^{-1} purified CRP with either 0 mM or 12.8 mM acP in acetylation buffer for 15 minutes. After SDS-polyacrylamide gel electrophoresis, the gel was stained by incubation with SimplyBlue SafeStain (Invitrogen) and destained using deionized water according to the manufacturer's protocol. The prominent band from each sample was cut out of the gel and placed in a 1.5 mL Eppendorf tube with deionized water for storage. Gel fragments were further destained and dehydrated with ACN. Subsequently, proteins were reduced with 10mM dithiothreitol in 25 mM NH_4HCO_3 at 56 °C for 1 h and alkylated with 55 mM iodoacetamide in 25 mM NH_4HCO_3 at room temperature for 45 min. Samples were incubated overnight with trypsin (125 ng, 37 °C). The resulting proteolytic peptides were subjected to aqueous (100 µl H_2O, sonication, 10 min) and hydrophobic extraction (2x 50 µl of 50% ACN, 5% formic acid). Extracts were combined, concentrated and desalted using C18 zip-tips (Millipore, Billerica, MA). Finally, samples were analyzed by mass spectrometry after concentration under vacuum to a 10-15 µl final volume.

For each CRP protein sample, technical duplicates (MS injection duplicates) were acquired to assess technical variability. All samples were analyzed by reverse-phase HPLC-ESI-MS/MS using an Eksigent UltraPlus nano-LC 2D HPLC system (Dublin, CA) connected to a quadrupole time-of-flight TripleTOF 5600 mass spectrometer (AB SCIEX), as previously described in detail (121). Briefly, the resulting peptides were chromatographically separated on a C18 Acclaim PepMap100 reversed-phase analytical column (75 µm I.D.) at a flow rate of 300 nL/min with a total runtime of 90 min including mobile phase equilibration. The nanoLC system

was directly connected to the TripleTOF 5600 operating in data dependent mode with 1 MS1 survey scan (250 msec) followed by 30 MS/MS scans (50 msec each) per 1.8 second acquisition cycle. Mass spectrometric raw data and annotated MS/MS spectral libraries can be accessed at massive.ucsd.edu, (MassIVE ID: MSV000080568; password: winter); and at ProteomeXchange under PXD005965. The processed MS data is provided in **Table 4** to show comprehensive lists of un-modified and acetylated CRP peptides that were identified with all their mass spectrometric information.

Protein Identification and Quantification

For protein identification, all data were searched using Protein Pilot v. 4.5 beta (180), using a false discovery rate (FDR) of 1%. A SwissProt *E. coli* database (SwissProt fasta version 2013_07) was searched. The following sample parameters were used: trypsin digestion, cysteine alkylation set to iodoacetamide, acetylation emphasis, and species *E. coli*. Trypsin specificity was set at C-terminal lysine and arginine. Processing parameters were set to "Biological modification" and a thorough ID search effort was used.

MS1 chromatogram based quantification was performed in Skyline 2.5 an open source software project (http://proteome.gs.washington.edu/software/skyline) as described in detail (181). Detailed MS1 Filtering quantification results for acetylated CRP peptides are provided in **Table 4**. Quantification of potential protein level changes upon acP incubation analyzing 23 robustly observed, non-acetylated peptides by MS1 Filtering is shown in **Table 5**.

RNAseq

Cells were grown in TB7 supplemented with 22 mM glucose until OD_{600} ~1.8. 1 mL culture was added to 2 mL RNAProtect Bacterial Reagent (Qiagen 76506), and vortexed. After centrifugation, the cell pellets were frozen at -80°C and stored overnight. The cell pellets were resuspended by adding 300 µl Tissue and Cell lysis solution (EpiCentre) and 2 µl proteinase K. The resuspensions were incubated at 65°C for 15 minutes, vortexing every 5 minutes. After cooling the resuspensions to room temperature, RNA was isolated using the RNeasy Kit (Qiagen) using the manufacturer's protocol. After isolation, 1 µl RiboGuard RNase inhibitor (Epicentre) was added to the purified RNA. RNA purity was determined using a NanoDrop 2000 spectrophotometer (Thermo Scientific). The amount of total RNA in each sample was quantified using the Qubit 2.0 Flurometer (Life Technologies) and quality was assessed using the RNA6000 Nano Chip on the Bioanalyzer 2100 (Agilent).

A hybridization-capture process was carried out to remove ribosomal RNA using the Ribo-Zero Magnetic Kit (Epicentre) per manufacturer instructions. The rRNA-depleted samples were purified using Agencourt RNAClean XP beads (Beckman Coulter) and quality was assessed using the RNA6000 Pico Chip and the Bioanalyzer. Using the ScriptSeq v2 Complete Kit (Epicentre), the rRNA depleted RNA was fragmented and reverse transcribed using random primers that included a unique 5' tagging sequence. The resulting tagged RNA was tagged once more at its 3' end by the terminal-tagging reaction yielding single-stranded cDNA. The cDNA was then amplified using limited cycle PCR, incorporating sequencing adapters and barcodes to create a final double-stranded directional cDNA library ready for sequencing. The samples were sequenced on the Illumina MiSeq platform rendering 250 bp paired-end reads.

To perform data analysis, adapter sequences were removed from raw reads and low quality reads were trimmed using a Python-based tool, Cutadapt (182). The resulting reads were then mapped to the reference genome of *Escherichia coli* str. K-12 substr. MG1655, using Bowtie2 v. 2.1.0. The aligned sequencing reads and a list of genomic features were used as input for the Python package HTSeq to count the mapped genes and generate a table of raw counts. The R package, DESeq2 (183), was used to determine differential expression between sample groups using the raw count table by fitting the negative binomial generalized linear model for each gene and then using the Wald test for significance testing (244). Count outliers were detected using Cook's distance and were removed from further analysis. The Wald test p-values from the subset of genes that passed an independent filtering step were then adjusted for multiple testing using the Benjamin-Hochberg procedure (245). Genes that are up- or down-regulated with an adjusted p-value of <0.05 AND a fold difference in gene expression >=1.5-fold were identified as significantly modulated genes. The global gene expression data discussed in this work have been deposited in NCBI's Gene Expression Omnibus (184) and are accessible through GEO Series accession number GSE97406 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE97406).

DNA Microarray

Global gene expression was assessed in M9 minimal medium with 10 mM glucose or 30mM acetate as the sole carbon source. Aerobic 50 ml batch cultures were grown in 500 mL flasks at 37°C on a rotary shaker at 250 rpm. These cultures were inoculated to an initial optical density (OD_{600}) of 0.05 units with exponentially growing pre-cultures. Samples for RNA

extraction were taken in middle exponential phase (OD_{600}≈0.5) and in stationary phase (OD_{600}≈1.5). RNA was purified using VantageTM Total RNA Purification Kit (Origene, Rockville, MD, USA). Purity and concentration of isolated RNA were assessed in a NanoDrop One spectrophotometer (Thermo Scientific Incorporated, WI, USA). Quality was evaluated by microfluidic capillary electrophoresis on an Agilent 2100 Bioanalyzer (Agilent Technologies, Inc, USA). Gene expression profiles were analyzed using commercially available oligonucleotide microarrays GeneChip *E. coli* Genome 2.0 Arrays (P/N 900550, Affymetrix, Incorporated, Santa Clara, CA, USA). For each microarray, 5 μg of total RNA were used per the manufacturer's instructions (P/N 702232 Rev.3, Affymetrix).

Data analysis was performed using Partek Genomics Suite 6.6 and Partek Pathway (Partek Incorporated, St. Louis, MO, USA), with statistical methods and thresholds described by the manufacturer. Data were further analyzed using Affymetrix Transcriptome Analysis Console 3.0 (TAC). Unless otherwise indicated, genes that are up- or downregulated with a p value <0.05 were identified as significantly modulated genes. The global gene expression data discussed in this work have been deposited in NCBI's Gene Expression Omnibus (184) and are accessible through GEO Series accession number GSE96955 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE96955).

CRP Stability Assay

Cultures were grown in TB7 supplemented with 22 mM glucose at 37°C until OD_{600} ~0.5. Chloramphenicol was added to each culture at a concentration of 200 μg/mL to inhibit protein synthesis. At regular intervals, 2 mL samples were collected and centrifuged at 3000 rpm for 12

minutes. After removing the supernatant, the cell pellets were stored at -80°C. Western immunoblot analysis was performed on the cell pellets as described above using anti-CRP and anti-RNAPα antibodies.

Table 1. Strains, plasmids, and primers.

Strain	Description	Reference/Source
AJW678	*thi-1 thr-1(Am) leuB6 metF159(Am) rpsL136 lacX74*	(187)
CB369	MG1655 Δ*crp*::kan	Christoph Bausch (University of Oklahoma)
AJW2179	AJW678 λ42(*acs205'-lacZ*)	(239)
AJW2198	AJW678 λ42(*acs205'-lacZ*) Δ*crp*::kan	(241)
AJW4344	AJW678 λ42(*acs205'-lacZ*) Δ*yfiQ*::kan	P1: AJW2568 x AJW2179
AJW5105	AJW678 λ42(*acs205'-lacZ*) Δ*ackA-pta*::cm	P1: AJW2571 x AJW2179
AJW5174	AJW678 λ42(*acs205'-lacZ*) Δ*ackA*::frt	P1: JW2293 (242) x AJW5173, then removed antibiotic cassette
AJW5179	AJW678 λ42(*acs205'-lacZ*) Δ*ackA*::frt Δ*crp*::kan	P1: AJW2033 x AJW5174
AJW5246	AJW678 Δ*crp*::kan	P1: CB369 x AJW678
AJW5445	AJW678 Δ*crp*::frt	removed antibiotic cassette from AJW5246
AJW5587	AJW678 Δ*crp*::frt Δ*dnaJ*::kan	P1: JW0014 (242) x AJW5445
AJW5619	AJW678 Δ*crp*::frt Δ*dnaJ*::frt	removed antibiotic cassette from AJW5587
AJW5621	AJW678 Δ*crp*::frt Δ*dnaJ*::frt Δ*clpX*::kan	P1: JW0428 (242) x AJW5619
AJW5622	AJW678 Δ*crp*::frt Δ*dnaJ*::frt Δ*clpP*::kan	P1: JW0427 (242) x AJW5619
AJW5624	AJW678 Δ*crp*::frt Δ*dnaJ*::frt Δ*clpY*::kan	P1: JW3902 (242) x AJW5619
AJW5625	AJW678 Δ*crp*::frt Δ*dnaJ*::frt Δ*lon*::kan	P1: JW0429 (242) x AJW5619
AJW5626	AJW678 Δ*crp*::frt Δ*dnaJ*::frt Δ*pepE*::kan	P1: JW3981 (242) x AJW5619
AJW5628	AJW678 Δ*crp*::frt Δ*dnaJ*::frt Δ*ptrB*::kan	P1: JW1834 (242) x AJW5619
AJW5629	AJW678 Δ*crp*::frt Δ*dnaJ*::frt Δ*glpG*::kan	P1: JW5687 (242) x AJW5619
AJW5630	AJW678 Δ*crp*::frt Δ*dnaJ*::frt Δ*clpB*::kan	P1: JW2573 (242) x AJW5619
AJW5631	AJW678 Δ*crp*::frt Δ*dnaJ*::frt Δ*ycaC*::kan	P1: JW0880 (242) x AJW5619
AJW5632	AJW678 Δ*crp*::frt Δ*dnaJ*::frt Δ*yifB*::kan	P1: JW3738 (242) x AJW5619
AJW5633	AJW678 Δ*crp*::frt Δ*dnaJ*::frt Δ*ycbZ*::kan	P1: JW0938 (242) x AJW5619
AJW5680	AJW678 Δ*crp*::frt Δ*dnaJ*::frt Δ*prlC*::kan	P1: JW3465 (242) x AJW5619
AJW5740	AJW678 Δ*dnaJ*::kan	P1: JW0014 (242) x AJW678
AJW5784	AJW678 Δ*dnaJ*::frt	removed antibiotic cassette from AJW5740
AJW5792	AJW678 Δ*dnaJ*::frt Δ*clpX*::kan	P1: JW0428 (242) x AJW5784
AJW5793	AJW678 Δ*dnaJ*::frt Δ*clpP*::kan	P1: JW0427 (242) x AJW5784
AJW5794	AJW678 Δ*dnaJ*::frt Δ*clpQ*::kan	P1: JW3903 (242) x AJW5784
AJW5795	AJW678 Δ*dnaJ*::frt Δ*clpY*::kan	P1: JW3902 (242) x AJW5784

AJW5796	AJW678 Δ*dnaJ*::frt Δ*lon*::kan	P1: JW0429 (242) x AJW5784
AJW5797	AJW678 Δ*dnaJ*::frt Δ*pepE*::kan	P1: JW3981 (242) x AJW5784
AJW5798	AJW678 Δ*dnaJ*::frt Δ*clpA*::kan	P1: JW0866 (242) x AJW5784
AJW5799	AJW678 Δ*dnaJ*::frt Δ*ptrB*::kan	P1: JW1834 (242) x AJW5784
AJW5800	AJW678 Δ*dnaJ*::frt Δ*glpG*::kan	P1: JW5687 (242) x AJW5784
AJW5801	AJW678 Δ*dnaJ*::frt Δ*clpB*::kan	P1: JW2573 (242) x AJW5784
AJW5802	AJW678 Δ*dnaJ*::frt Δ*ycaC*::kan	P1: JW0880 (242) x AJW5784
AJW5803	AJW678 Δ*dnaJ*::frt Δ*yifB*::kan	P1: JW3738 (242) x AJW5784
AJW5804	AJW678 Δ*dnaJ*::frt Δ*ycbZ*::kan	P1: JW0938 (242) x AJW5784
AJW5805	AJW678 Δ*dnaJ*::frt Δ*clpS*::kan	P1: JW0865 (242) x AJW5784
AJW5806	AJW678 Δ*dnaJ*::frt Δ*prlC*::kan	P1: JW3465 (242) x AJW5784
AJW5074	BL21(DE3) pET15b-crp	pET15b-crp -> AJW4860

Plasmids	**Description**	**Reference/Source**
pBR322	Ap^R	(190)
pDCRP	pBR322 carrying *crp* cloned under the control of the *crp* promoter, Ap^R	(190)
pDCRP-crpK100A	pDCRP with K100A mutation in *crp*	This work
pDCRP-crpK100Q	pDCRP with K100Q mutation in *crp*	This work
pDCRP-crpK100R	pDCRP with K100R mutation in *crp*	This work
pDCRP-crpK101E	pDCRP with K101E mutation in *crp*	(25)
pDCRP-crpK101A	pDCRP with K101A mutation in *crp*	This work
pDCRP-crpK101AK100A	pDCRP with K101A K100A double mutation in *crp*	This work
pDCRP-crpK101AK100Q	pDCRP with K101A K100Q double mutation in *crp*	This work
pDCRP-crpK101AK100R	pDCRP with K101A K100R double mutation in *crp*	This work
pDCRP-crpH159L	pDCRP with H159L mutation in *crp*	(240)
pRW50 *CC(-41.5)*	pRW50 carrying the *E. coli melR* promoter derivative with a consensus DNA site for CRP centered at -41.5, Tet^R	(18)
pRW50 *CC(-61.5)*	pRW50 carrying the *E. coli melR* promoter derivative with a consensus DNA site for CRP centered at -61.5, Tet^R	(18)
p770-*yfiQ*'	p770 derivative carrying *yfiQ* promoter, Ap^R	This work
pET15b-crp	pET15b derivative carrying *crp* gene with IPTG-inducible promoter, Ap^R	(173)

Primers	**Sequence (5'->3')**	**Purpose**
yfiQ' EcoRI F	CCTGCGTGAAGCCCAGGCTGAATTCCAATATTGTACTGC	Generation of EcoRI site for insertion into p770

yfiQ' HinDIII R	CGATTTTGGTCGCAGTAAAGCTTCCAGTCCTCGCTG	Generation of HinDIII site for insertion into p770
K100A F	GTGAAGTGGCTGAAATTTCGTACGCAAAATTTCGCCAATTGATTCAGG	Mutagenesis of pDCRP to generate K100A allele
K100A R	CCTGAATCAATTGGCGAAATTTTGCGTACGAAATTTCAGCCACTTCAC	Mutagenesis of pDCRP to generate K100A allele
K100Q F	GTGAAGTGGCTGAAATTTCGTACCAGAAATTTCGCCAATTGATTCAGGT	Mutagenesis of pDCRP to generate K100Q allele
K100Q R	ACCTGAATCAATTGGCGAAATTTCTGGTACGAAATTTCAGCCACTTCAC	Mutagenesis of pDCRP to generate K100Q allele
K100R F	GAAGTGGCTGAAATTTCGTACAGAAAATTTCGCCAATTGATTCAG	Mutagenesis of pDCRP to generate K100R allele
K100R R	CTGAATCAATTGGCGAAATTTTCTGTACGAAATTTCAGCCACTTC	Mutagenesis of pDCRP to generate K100R allele
K101A F	GTGAAGTGGCTGAAATTTCGTACAAAGCATTTCGCCAATTGATTCAGGT	Mutagenesis of pDCRP to generate K101A allele
K101A R	ACCTGAATCAATTGGCGAAATGCTTTGTACGAAATTTCAGCCACTTCAC	Mutagenesis of pDCRP to generate K101A allele
K101A K100A F	GTGGCTGAAATTTCGTACGCAGCATTTCGCCAATTGATTCAGGT	Mutagenesis of pDCRP to generate K101A K100A allele
K101A K100A R	ACCTGAATCAATTGGCGAAATGCTGCGTACGAAATTTCAGCCAC	Mutagenesis of pDCRP to generate K101A K100A allele
K101A K100Q F	GTGGCTGAAATTTCGTACCAGGCATTTCGCCAATTGATTCAGGT	Mutagenesis of pDCRP to generate K101A K100Q allele
K101A K100Q R	ACCTGAATCAATTGGCGAAATGCCTGGTACGAAATTTCAGCCAC	Mutagenesis of pDCRP to generate K101A K100Q allele
K101A K100R F	GAAGTGGCTGAAATTTCGTACAGAGCATTTCGCCAATTGATTCAGGTAAA	Mutagenesis of pDCRP to generate K101A K100R allele
K101A K100R R	TTTACCTGAATCAATTGGCGAAATGCTCTGTACGAAATTTCAGCCACTTC	Mutagenesis of pDCRP to generate K101A K100R allele
q16S F	CGGTGGAGCATGTGGTTTA	qPCR quantification of 16S RNA
q16S R	GAAAACTTCCGTGGATGTCAAGA	qPCR quantification of 16S RNA
qCrp F	TCATTGCCACATTCATAAGTACCC	qPCR quantification of crp mRNA
qCrp R	TCATTTCTTTACCCTCTTCGTCTTT	qPCR quantification of crp mRNA
qFlhD F	TACTACTTGCACAGCGTTTGATTGT	qPCR quantification of flhD mRNA
qFlhD R	TTGTCGCCATTTCTTCATTTATGCC	qPCR quantification of flhD mRNA
qYfiQ F	CGTAGTGCCTCGCGTAATAAA	qPCR quantification of yfiQ mRNA
qYfiQ R	AAACCGGCACGCTGAATA	qPCR quantification of yfiQ mRNA

CHAPTER THREE

RESULTS

Regulation of Acetylation by CRP

There are two known mechanisms of Nε-lysine acetylation in *E. coli* (which I will refer to as simply "acetylation" for the rest of the dissertation): 1) Enzymatic acetylation by acetyltransferases using acetyl-CoA as the acetyl group donor, and 2) Spontaneous non-enzymatic acetylation using acP as the acetyl group donor. While these two mechanisms differ greatly in how they function and the number of lysines they target, both utilize metabolic intermediates as acetyl group donors. Thus, it is not surprising that there is some evidence both mechanisms are positively regulated by CRP, a transcriptional regulator of carbon metabolism. In this Chapter, I will provide additional evidence for the hypotheses that both YfiQ-dependent and acP-dependent acetylation are regulated by CRP.

CRP-Dependent Regulation of *yfiQ* Transcription

YfiQ is the only reported lysine acetyltransferase in *E. coli*. Before the discovery of acP-dependent acetylation, much work was done to identify the targets of YfiQ-dependent acetylation, determine the catalytic mechanism and, relevant to this dissertation, identify mechanisms by which cells regulate YfiQ. *In silico* analysis of the promoter region of the *yfiQ* gene identified several putative CRP binding sites (151). The two sites with the highest scores,

indicating the most likely CRP binding sites, were located at positions -41.5 and -81.5 relative to the TSS, suggesting the existence of both Class II and Class I binding sites. Transcription from the *yfiQ* promoter was shown to require cAMP in a concentration-dependent manner, further supporting regulation by CRP (151). However, there was a lack of direct evidence that CRP itself is required for *yfiQ* transcription. In this section, I will test the hypothesis that CRP regulates transcription of *yfiQ* and, if so, whether that regulation is direct or indirect.

To determine if CRP regulates the transcription of *yfiQ*, with the assistance of Bozena Zemaitaitis, I use qRT-PCR to quantify *yfiQ* transcripts in Δ*crp* strains transformed with either pDCRP (AJW2313), which carries the WT *crp* gene driven by the native *crp* promoter, or pBR322 (AJW4524) as vector control (VC) (**Fig. 8A**). As a negative control, I used a Δ*yfiQ* strain transformed with pBR322 (AJW5204). These strains were grown in TB7 until they reached stationary phase. Presence of *yfiQ* mRNA was detected in the strain expressing WT CRP, but not in the strain lacking *yfiQ*, as expected. In the absence of CRP, *yfiQ* mRNA levels were about 9% of those found in the strain expressing WT CRP. These data support the hypothesis that CRP is required for full *yfiQ* expression.

To determine precisely which activating regions are required for CRP-dependent *yfiQ* transcription, I included Δ*crp* strains transformed with either pDCRP-crpK101E (AJW5224) or pDCRP-crpH159L (AJW3660), encoding CRP variants defective in Class II and Class I transcription respectively, in the experiment described above (**Fig. 8A**). In the strain expressing the AR2 mutant CRP K101E, *yfiQ* mRNA levels were about 6% of those found in the strain expressing WT CRP, comparable to the results seen in the absence of CRP. This suggests that a fully functional AR2 is required for *yfiQ* transcription. Unfortunately, the strain expressing the AR1 mutant CRP

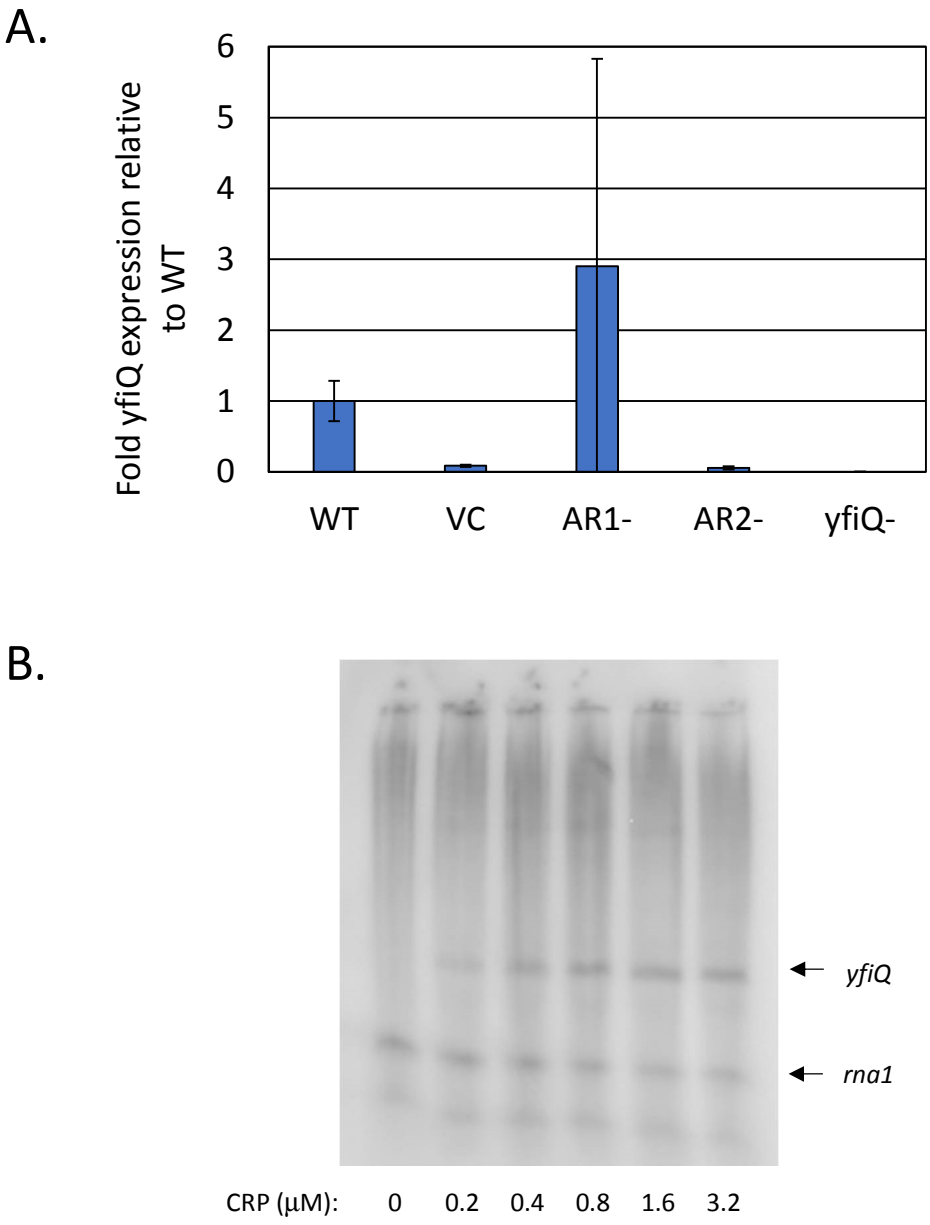

Figure 8. Transcription of *yfiQ* requires CRP. (A) Δ*crp* strains carrying pDCRP expressing WT CRP (WT), pBR322 (VC), pDCRP expressing CRP H159L (AR1-), or pDCRP expressing CRP K101E (AR2-), as well as a Δ*yfiQ* strain carrying pBR322 (yfiQ-) were grown in TB7 supplemented with 22 mM glucose at 37°C. Cells were harvested in stationary phase and subjected to qRT-PCR analysis using *yfiQ*-specific primers. Data represents the average and standard deviation of two independent experiments: one performed in triplicate and the other performed in singlet. All values were standardized to 16S rRNA. (B) *In vitro* transcription from the *yfiQ* promoter with increasing concentrations of purified CRP. Each reaction contains 10 nM RNAP-σ^{70} and 200 μM cAMP. The reactions were incubated for 15 minutes at 30°C, and transcripts were visualized by incorporation of ^{32}P-labelled UTP. The figure shown here represents the single successful attempt with usable data. *rna1* transcription served as an internal control.

H159L gave conflicting results. In one experiment, the strain expressing CRP H159L had about 6% of the *yfiQ* transcripts found in the strain expressing WT CRP, much like the strains expressing CRP K101E or no CRP, suggesting both AR1 and AR2 are required for full *yfiQ* transcription. In another experiment, the strain expressing CRP H159L had almost four times the level of *yfiQ* transcripts found in the strain expressing WT CRP, suggesting AR1 is inhibitory to *yfiQ* transcription. Further work will be required to determine the requirement of AR1 for *yfiQ* transcription. However, these results support the hypothesis that at least AR2 is required for CRP-dependent *yfiQ* transcription.

Although CRP is required for *yfiQ* transcription, it is not clear if the role of CRP is direct, through binding to one of the putative CRP binding sites, or indirect, through regulation of other transcriptional regulators. Given the presence of two high quality putative CRP binding sites, I hypothesized that at least some of the CRP-dependent regulation of *yfiQ* transcription is direct. To test this hypothesis, I performed *in vitro* transcription from the *yfiQ* promoter using purified WT CRP. The *yfiQ* promoter (from -177 to +44) was cloned into plasmid p770 to generate p770-*yfiQ*'. Purified RNAP holoenzyme containing σ^{70} was added to a mixture of p770-*yfiQ*', ribonucleotides containing [^{32}P]-labelled UTP, cAMP, and various concentrations of purified CRP in *in vitro* transcription buffer. After a 15-minute incubation at 30°C, transcripts were separated on an acrylamide gel and visualized using a phosphoimaging screen (**Fig. 8B**). Transcription from the *yfiQ* promoter only occurred in the presence of CRP, and appeared to reach maximal levels around 400 nM CRP. Unfortunately, after several attempts, I was unable to reproduce these data. The likely cause was RNase contamination, since I was unable to visualize transcription from the *rna1* control promoter. This experiment will need to be

reproduced to confirm the results. However, if the initial results were correct, these data would support the hypothesis that at least some of the CRP-dependent regulation of *yfiQ* transcription is through direct binding of CRP to the *yfiQ* promoter.

On the basis of these results, assuming successful reproduction of the *in vitro* transcription data, I conclude that CRP directly regulates transcription of *yfiQ*. CRP-dependent regulation of *yfiQ* transcription requires AR2, and possibly AR1. Given what is known about the way CRP activates transcription (see Chapter One – cAMP Receptor Protein (CRP)), and the existence of both Class II and Class I putative CRP binding sites in the *yfiQ* promoter region, it is likely that at least one of these ARs is responsible for the direct regulation of *yfiQ* transcription by CRP. However, this question was not directly addressed in these studies, and additional work is needed to confirm this conclusion. Additionally, this work does not directly test the hypothesis that CRP regulates YfiQ-dependent acetylation, though this work does provide evidence for this hypothesis. Additional work is required to show that CRP influences the acetylation of YfiQ-regulated lysines.

CRP-Dependent Regulation of Glucose-Induced Acetylation

Global acetylation is dramatically increased in *E. coli* grown in media containing non-limiting concentrations of glucose compared to media alone (120, 121, 134). Excess glucose promotes the synthesis of the acetyl group donor acP through overflow metabolism, where the rate of glucose uptake and catabolism exceeds the rate at which the cell can utilize acCoA in the TCA cycle (see Chapter One – Overflow Metabolism). AcP is generated as an intermediate between acCoA and acetate (**Fig. 6**). Synthesis of acP is required for the increase in glucose-

induced acetylation. CRP is also required for the increase in glucose-induced acetylation (153). CRP increases the rates at which glucose is consumed and acetate is produced, presumably increasing the rate at which acP is produced as well (185). I hypothesize that CRP regulates glucose-induced acetylation by promoting the synthesis of acP. In this section, I will address this hypothesis by determining if CRP and acP function through the same pathway to promote glucose-induced acetylation.

Epistasis analysis is a powerful tool used to determine the interactions between genes (186). By combining two alleles with differing phenotypic consequences (e.g. loss of acetylation by *crp* deletion and increase in acetylation by *ackA* deletion), it is possible to 1) determine if two genes function within the same pathway to regulate a phenotype, and 2) determine the order in which the two genes are required for the phenotype. To determine if CRP and acP function through the same pathway to promote glucose-induced acetylation, I began by asking if CRP-dependent acetylation requires synthesis of acP. I started with a strain wild-type for the Pta-AckA pathway (AJW2179) and its isogenic Δ*ackA-pta* mutant (AJW5105), which does not synthesize acP (**Fig. 6**). I next transformed each strain with either pDCRP or vector control (VC, pBR322). The resulting transformants were grown in TB7 supplemented with 22 mM glucose (to induce acetylation) and 3 mM cAMP (to overcome catabolite repression). Upon entry into stationary phase, samples were collected and subjected to Western immunoblot analysis using an anti-acetyllysine antibody (**Fig. 9A**). Overexpression of CRP (pDCRP) in the parental strain (AJW4526) resulted in an increase in global acetylation relative to the parental strain transformed with the VC (AJW4525). Loss of acP synthesis (Δ*ackA-pta* transformed with VC, AJW5111) resulted in a decrease in global acetylation relative to the parental strain

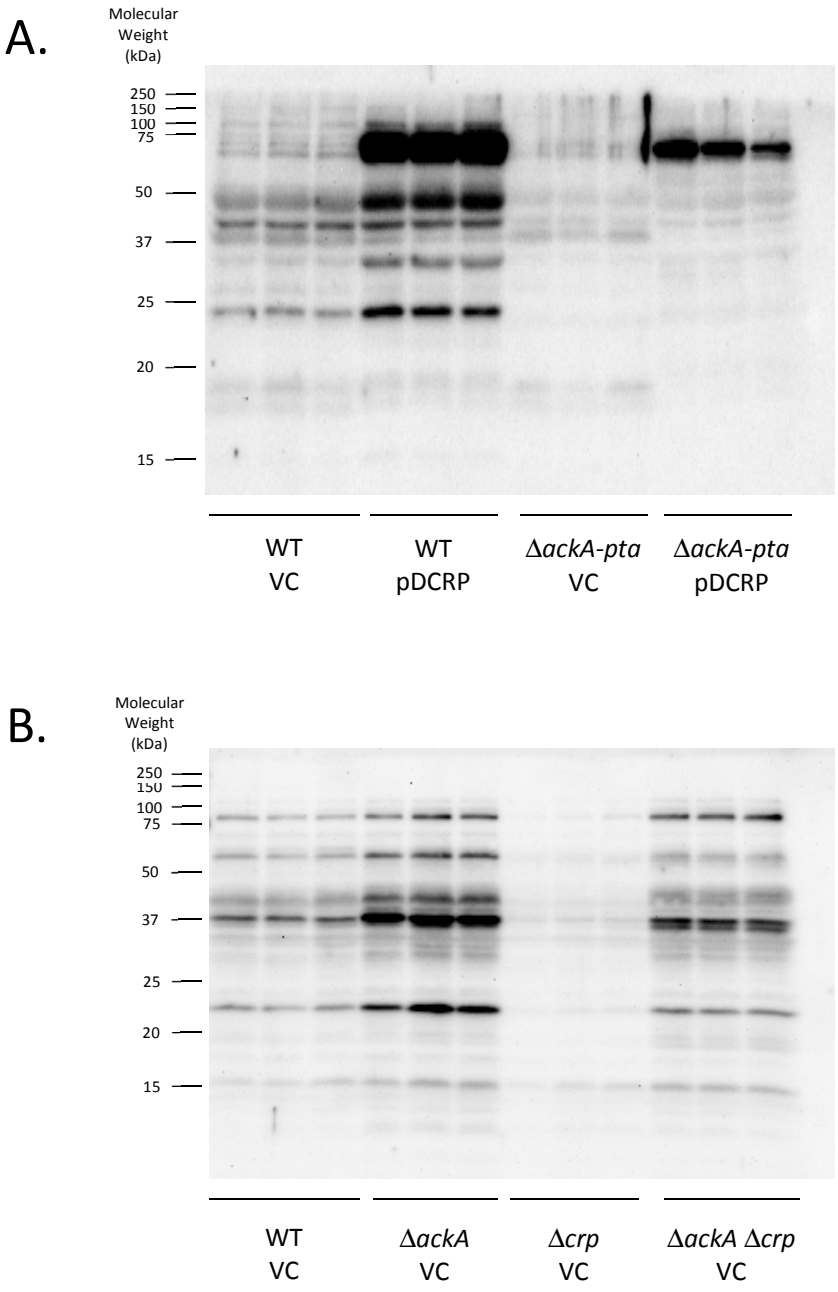

Figure 9. Epistasis between CRP and acP regarding glucose-induced acetylation. (A) WT or Δ*ackA-pta* strains each transformed with pDCRP or pBR322 (VC) were grown in TB7 supplemented with 22 mM glucose and 3 mM cAMP (to overcome catabolite repression) at 37°C. Cells were harvested upon entry into stationary phase and subjected to Western immunoblot analysis using an anti-acetyllysine antibody. Data shows results from cultures grown in triplicate and represents two independent experiments. (B) Same as (A), except WT, Δ*ackA*, Δ*crp*, and Δ*ackA* Δ*crp* strains all transformed with pBR322 (VC) were used.

transformed with the VC, as expected (121). Loss of acP synthesis prevented most of the CRP-dependent increase in global acetylation in the double mutant transformed with pDCRP (AJW5112). The one prominent band observed in the double mutant is likely Acs; Acs is approximately 72 kDa, is positively regulated by CRP, and is acetylated by YfiQ in an acetyl CoA-dependent manner (93, 151, 187). These data suggest that acP is required for glucose-mediated CRP-dependent acetylation.

I next asked if acP-dependent acetylation requires CRP. I used the same parental strain as above (AJW2179) and three isogenic mutants: Δ*ackA* (AJW5174, which accumulates acP [**Fig. 6**]), Δ*crp* (AJW2198), and Δ*ackA* Δ*crp* (AJW5179). Each of these strains contained the VC plasmid for consistency with the previous experiment. The resulting transformants were grown in TB7 supplemented with 22 mM glucose (to induce acetylation) and 3 mM cAMP (to overcome catabolite repression). Upon entry into stationary phase, samples were collected and subjected to Western immunoblot analysis using an anti-acetyllysine antibody (**Fig. 9B**). Accumulation of acP (Δ*ackA*) resulted in an increase in global acetylation relative to the parental strain, and deletion of *crp* resulted in a decrease in global acetylation relative to the parental strain as expected. Loss of CRP caused only a minor reduction in acP-dependent acetylation in the double mutant. These data suggest that through the accumulation of acP, the cell is able to bypass the requirement of CRP for glucose-mediated acetylation.

Together, these data support the hypothesis that CRP promotes glucose-induced acetylation by promoting acP synthesis. However, the mechanism by which CRP promotes acP synthesis was not addressed by this work. David Christensen, another student in the lab, began to address the mechanism after it was clear that our projects would intersect. He discovered

that CRP requires both AR1 and AR2 to promote glucose-induced acetylation; mutation of either AR alone had little effect on global acetylation (134). Additionally, he found that glucose-induced acetylation requires the high velocity glucose transporter PtsG, whose transcription is driven from a promoter with Class II and Class I CRP binding sites (134, 188, 189).

Summary

Previous studies identified potential roles for CRP in the regulation of both YfiQ-dependent and acP-dependent acetylation. However, these studies stopped short of directly linking CRP to acetylation. In this Chapter, I provided additional evidence that CRP positively regulates both mechanisms of acetylation. In the case of YfiQ-dependent acetylation, qRT-PCR analysis revealed that CRP promotes *yfiQ* transcription in an AR2-dependent, and possibly AR1-dependent, manner (**Fig. 8A**). *In vitro* transcription analysis suggests that at least some of this CRP-dependent regulation is direct (**Fig. 8B**), but these results are preliminary and need to be verified. In the case of acP-dependent acetylation, Western immunoblot analysis using an anti-acetyllysine antibody showed that CRP is not required for glucose-induced acetylation if acP levels are genetically increased (**Fig. 9**). This suggests the role of CRP in glucose-induced acetylation may be to promote acP synthesis. Further analysis done by David Christensen revealed that glucose-induced acetylation requires the high velocity glucose transporter PtsG, whose transcription is CRP-dependent (134). Taken together, these results support a hypothesis in which acetylation is tightly linked to carbon metabolism. As such, the acetylation of metabolic enzymes could be a way the cell monitors the rate of carbon metabolism, and represent an additional layer to the regulation of metabolic activity.

CHAPTER FOUR

RESULTS

Regulation of CRP by Acetylation

In the previous chapter, I provided evidence that both known mechanisms of acetylation, enzymatic and non-enzymatic, are positively regulated by CRP, either directly or indirectly. For the vast majority of the resultant acetylations, it is unknown how their presence affects the acetylated protein. However, there are examples of specific acetylations that have a profound impact on the acetylated protein's activity, protein-protein interactions, or stability (see Chapter One – Protein Acetylation – Nε-Lysine Acetylation). Intriguingly, CRP itself can be acetylated on several lysines *in vivo*, and some of these lysines are located near residues that are critical for CRP function (**Fig. 2**). This led me to hypothesize that acetylation may play a role in CRP activity.

When a lysine becomes acetylated, the lysine's positive charge is neutralized. In this Chapter, I investigated whether the positive charge of CRP K100, an acetylatable lysine, plays a role in CRP activity. I provide evidence that the K100 positive charge is required for full CRP activity at a certain class of CRP-dependent promoters. I provide additional evidence that loss of the K100 positive charge, which would occur upon K100 acetylation, promotes CRP activity at a different class of promoters.

The Role of the K100 Positive Charge at Class II and Class I Promoters

When I began this dissertation work, only two lysines within CRP had been identified as acetylated: K52 (135) and K100 (131). Previous work by A.J. Walker-Peddakotla, a former student in the lab, did not identify any significance to K52 acetylation (A.J. Walker-Peddakotla, unpublished data). As a result, I chose to focus on a potential role for K100 acetylation. Quantitative mass spectrometric analysis identified that K100 is acetylated in an acP- and glucose-dependent manner (121, 134). K100 is adjacent to AR2 (**Fig. 2**), a positively charged region on the surface of CRP that interacts with a negatively charged region on the RNAP α subunit NTD (25). This interaction is critical for activation of transcription specifically from Class II CRP-dependent promoters; single mutation of any of the positively charged AR2 residues to alanine results in 5-13 fold decreases in transcription from the semi-synthetic Class II promoter *CC(-41.5)*, but no defect in transcription from the semi-synthetic Class I promoter *CC(-61.5)* (25). Mutation of K100 to alanine results in a smaller (~2-fold) decrease in transcription from *CC(-41.5)*, and has no effect on transcription from *CC(-61.5)*. These data indicate that K100 also plays a positive, though smaller, role specifically in Class II transcription. In this section, I will test the hypothesis that K100 requires a positive charge to promote Class II transcription.

To determine if the K100 positive charge is required, I utilized plasmids pDCRP (190) and pRW50 *CC(-41.5)* (18). pDCRP is a derivative of pBR322, a multicopy plasmid, that carries the WT *crp* gene driven by the native *crp* promoter. pRW50 *CC(-41.5)* is a derivative of pRW50, a low copy number plasmid, which carries the semi-synthetic Class II CRP-dependent promoter *CC(-41.5)* fused to *lacZ*. The *CC(-41.5)* promoter is a semi-synthetic promoter; it consists of the CRP-dependent *melR* promoter with the portion containing the native CRP binding site replaced

by a consensus CRP binding site centered at position -41.5 relative to the TSS, creating a Class II promoter.

I generated mutations in pDCRP that would result in substitutions within CRP at position 100, leading to a retention (K100R) or loss (K100Q and K100A) of the positive charge. I also included the K101A substitution as a control, since loss of K101 results in a significant decrease specifically in Class II promoter activity (25). Starting with a Δ*lac* Δ*crp* strain, I introduced two plasmids: 1) pDCRP, which carried *crp* alleles encoding either WT CRP or one of the K100/K101 mutants (or pBR322 as vector control), and 2) pRW50 *CC(-41.5)*. These strains were grown in TB7 supplemented with 22 mM glucose to measure β-galactosidase activity as a readout for *CC(-41.5)* promoter activity. I expected that the K100R mutant would have approximately wild-type (WT) Class II activity, while the K100Q mutant would have decreased Class II activity, similar to a K100A mutant.

First, I quantified the steady state level of each CRP mutant by western immunoblot analysis using an anti-CRP antibody. I found that any mutant lacking a positive charge at either K100 or K101 (K100A, K100Q, and K101A) was present at elevated levels relative to the WT (**Fig. 10**). This difference in CRP steady state levels was investigated further and will be described later in this Chapter (see Chapter Five – CRP Stability).

Despite the differences in CRP steady state levels, previously published work indicates that it is unlikely that these differences impact the activity of the *CC(-41.5)* promoter, and later, the *CC(-61.5)* promoter. These promoters are insensitive to changes in CRP concentrations except at very low CRP concentrations (18, 191). Both *CC(-41.5)* and *CC(-61.5)* contain consensus CRP binding sites. Mutation of CRP binding site within the *galP1* promoter to match

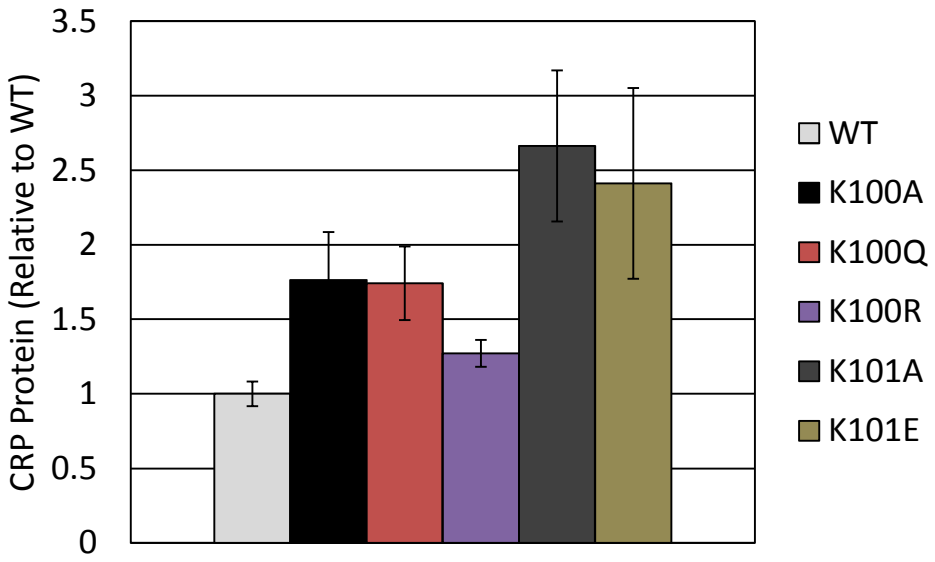

Figure 10. Quantification of CRP steady state levels. Δ*crp* strains carrying both pDCRP expressing the indicated CRP variants and pRW50 carrying the *CC(-41.5)* promoter fused to *lacZ* were grown in TB7 supplemented with 22 mM glucose at 37°C. Cells were harvested at OD_{600} 1.0 and subjected to Western immunoblotting using an anti-CRP antibody. Data shows average results from cultures grown in triplicate and represents multiple independent experiments. Error bars represent standard deviation.

the consensus sequence resulted in a 100-fold increase in the affinity of CRP for the consensus site over the original site (192). The addition of glucose to the media, which decreases cAMP synthesis and therefore the ability of CRP to bind DNA, had no effect on the activity of the *galP1* promoter containing the consensus sequence, suggesting the concentration of cAMP required for CRP binding to the consensus sequence is very low. *In vitro* DNA binding experiments were in agreement that the concentration of cAMP required for CRP to bind the consensus sequence was below the intracellular cAMP concentration found in cells grown in the presence of glucose (191). Finally, the dissociation time of CRP bound to the consensus sequence was greater than two hours, compared to a dissociation time of about eight minutes for CRP bound to the non-consensus CRP binding site at the *lac* promoter. Taken together, the authors concluded that CRP is permanently anchored to the consensus site *in vivo*. This makes CRP steady state levels irrelevant in the *CC(-41.5)* and *CC(-61.5)* assays, and therefore normalization of the β-galactosidase activity to CRP steady state levels is unnecessary.

Using *CC(-41.5)*, I observed a decrease in Class II promoter activity in strains expressing K100A or K100Q compared to the strains expressing WT or K100R, though not to the same extent as the K101A control (**Fig. 11**). These data suggest that the K100 positive charge is required for efficient Class II promoter activation.

To determine if the effect of the K100 positive charge was specific to Class II promoters, I performed a similar experiment using pRW50 *CC(-61.5)* (18). pRW50 *CC(-61.5)* is identical to pRW50 *CC(-41.5)*, with the exception of an additional insert that shifts the center of the consensus CRP binding site to position -61.5 relative to the TSS, creating a Class I promoter. Since AR2 does not play a role at Class I promoters, and K100 is not near AR1, I did not expect

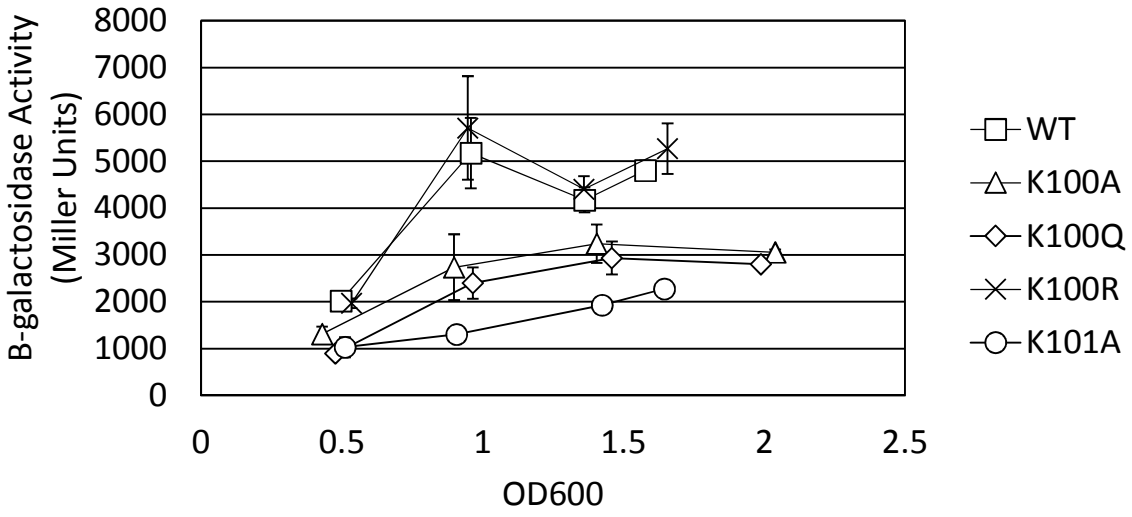

Figure 11. Effect of K100 positive charge on Class II promoter activity. Δ*crp* strains carrying both pDCRP expressing the indicated CRP variants and pRW50 carrying the *CC(-41.5)* promoter fused to *lacZ* were grown in TB7 supplemented with 22 mM glucose at 37°C. Samples were collected at the indicated OD_{600} and β-galactosidase activity was measured. Data shows average results from cultures grown in triplicate and represents multiple independent experiments. Error bars represent standard deviation.

the K100 positive charge to play a role at *CC(-61.5)*. Indeed, there was little difference in *CC(-61.5)* promoter activity between any of the CRP variants (**Fig. 12**), indicating that the K100 positive charge status has no impact on Class I promoter activation. Together, these data suggest that the role of the K100 positive charge in transcription activation is specific for Class II promoters.

Interaction Between K100 and RNAP α Subunit NTD

Given that 1) the positively charged AR2 makes contact with a negatively charged region on the RNAP α subunit NTD (25), 2) K100 requires a positive charge to promote Class II transcription (**Fig. 11**), and 3) K100 is positioned close to AR2 (**Fig. 2**), I hypothesized that K100 may make an independent contact with the RNAP α subunit NTD to promote Class II transcription. To address this hypothesis, I began by determining if a functional AR2 was required for K100-dependent regulation of Class II transcription. I introduced the K100A, K100Q, and K100R mutations into a K101A mutant background within pDCRP and tested the ability of these mutants to activate Class II transcription from *CC(-41.5)*. The K101A single mutant was used as a control for the status of K101. If K100 makes an independent contact with the RNAP α subunit NTD, I would expect K100-dependent regulation to occur even in the absence of a fully functional AR2. Promoter activity from *CC(-41.5)* was reduced considerably in the strain expressing CRP K101A compared to the strain expressing WT CRP, as expected (**Fig. 13**). The *CC(-41.5)* activity of the K101A K100R double mutant was nearly identical to the activity in the K101A single mutant, consistent with previous data showing the K100R mutation has little effect on *CC(-41.5)* activity (**Fig. 11**). Strains expressing either the K101A K100A or

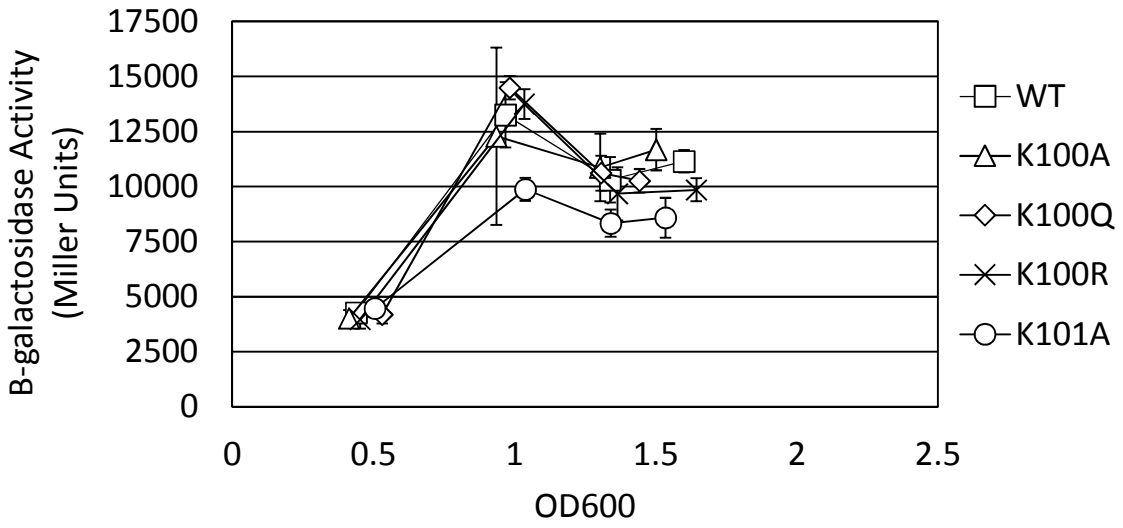

Figure 12. Effect of K100 positive charge on Class I promoter activity. Δ*crp* strains carrying both pDCRP expressing the indicated CRP variants and pRW50 carrying the *CC(-61.5)* promoter fused to *lacZ* were grown in TB7 supplemented with 22 mM glucose at 37°C. Samples were collected at the indicated OD_{600} and β-galactosidase activity was measured. Data shows average results from cultures grown in triplicate and represents multiple independent experiments. Error bars represent standard deviation.

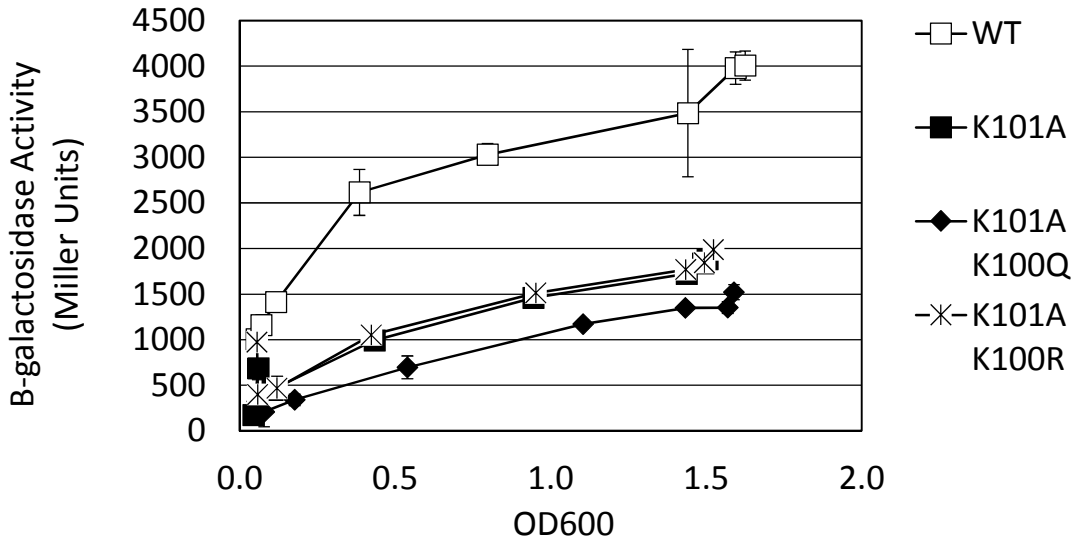

Figure 13. Determination of the independence of the K100 positive charge from K101. Δ*crp* strains carrying both pDCRP expressing the indicated CRP variants and pRW50 carrying the *CC(-41.5)* promoter fused to *lacZ* were grown in TB7 supplemented with 22 mM glucose at 37°C. Samples were collected at the indicated OD_{600} and β-galactosidase activity was measured. Data shows average results from cultures grown in triplicate and represents multiple independent experiments. Error bars represent standard deviation.

K101A K100Q double mutants displayed a further decrease in *CC(-41.5)* activity over the K101A single mutant. These data suggest K100-dependent regulation of Class II promoter activity does not require that K101 be intact, and support the hypothesis that K100 may make an independent contact with the RNAP α subunit NTD.

To further investigate the potential interaction between K100 and the RNAP α subunit NTD, I asked if K100 is physically close enough to any negatively charged residues in the RNAP α subunit NTD to make an ionic interaction at a Class II promoter. I worked with Dr. Ekaterina Filippova in Dr. Wayne Anderson's lab to model the interaction between CRP and RNAP at a Class II promoter. There are currently no crystal structures of CRP and RNAP bound at a Class II promoter, but there is a crystal structure of CRP and RNAP bound at a Class I promoter (176). This promoter consists of the Class I *lac* promoter, with the exception that all CRP and RNAP binding sequences have been replaced with consensus sequences. Dr. Filippova replaced the Class I promoter with the *CC(-41.5)* Class II promoter and adjusted the positions of CRP and RNAP to their relative binding sites, ensuring all of the known contacts between CRP and RNAP were made. The new model identified RNAP α E163 as a potential interaction partner with CRP K100 (**Fig. 14**). The model predicts an average distance of 2.75 angstroms between the CRP K100 ε-nitrogen atom and each of the RNAP α E163 γ-carboxyl group oxygen atoms. This value is well within the four angstrom range for salt bridge formation (193). These data further support the hypothesis that CRP K100 makes an independent contact with RNAP α, and suggest K100 should be considered a *bona fide* member of AR2.

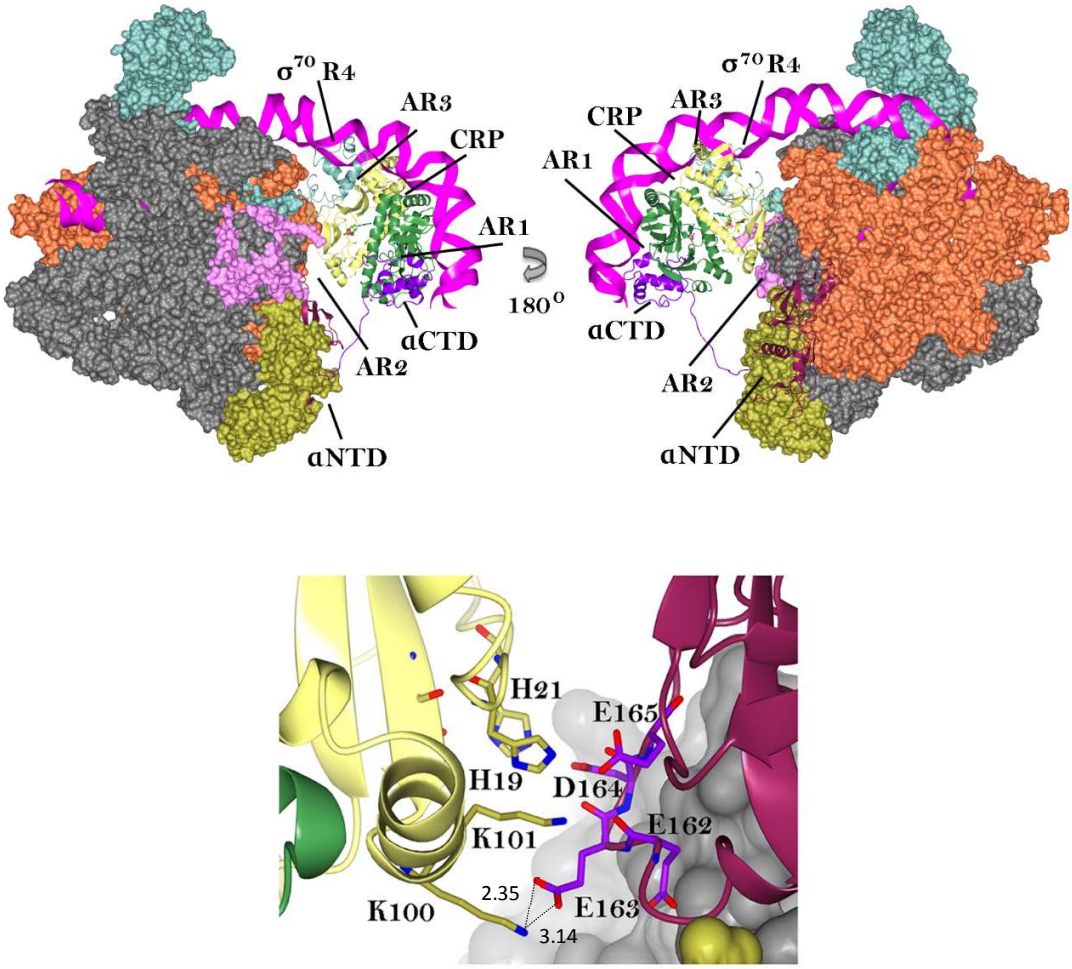

Figure 14. Model of the interaction between K100 and RNAP at a Class II promoter. Top, model structure of the ternary complex between CRP, RNAP and Class II promoter. Bottom, closeup of the interactions between CRP AR2 and the RNAP α subunit at a Class II promoter. α^{I} subunit of RNAP, σ^{70} region 4 domain of RNAP, CRP dimer and DNA are shown as ribbon diagram. α^{II}, β, β', ω subunits and σ^{70} non-conserved region (NCR) of RNAP are displayed as surface model. RNAP subunits are shown as follows: α^{I} C-terminal domain, purple; α^{I} N-terminal domain, dark purple; α^{II}, gold; β, coral; β', grey; ω, pink; σ^{70}, sea green. Subunits of the CRP dimer are shown in lemon and lawn green. DNA is colored in magenta. Activation region AR1 of CRP at class I and class II promoter composing residues 157-160 and 164 are colored in gold and light green, respectively. Activation regions AR2 and AR3 of CRP at class II promoter composing residues 19, 21, 100, 101, and 52-55, 58, respectively, are colored in gold. Corresponding residues of α^{I} C-terminal, N-terminal domain, and σ^{70} region 4 domain of RNAP are colored in lilac, purple, and light blue, respectively. Non-carbon atoms of interacting residues are colored in red (oxygen), blue (nitrogen), and yellow (sulfur).

The Role of the K100 Positive Charge in Global Transcription

In the previous sections, I found that the K100 positive charge promotes transcription from the Class II promoter *CC(-41.5)*, while having little effect on transcription from the Class I promoter *CC(-61.5)* (**Fig. 11**). However, these two promoters do not naturally exist in *E. coli*, and therefore do not necessarily represent the role of the K100 positive charge at all CRP-dependent promoters. CRP activity can vary greatly depending upon the position and sequence of the CRP binding site, as well as presence of other promoter binding elements, that presumably affect the interactions between CRP and RNAP (18, 194). Additionally, the information gathered from *CC(-41.5)* and *CC(-61.5)* does not address the role of the K100 positive charge in cell physiology.

CRP Overexpression in a Complex Media

To address these issues, I determined the effect of the K100 positive charge on global transcription using RNAseq analysis. I started by transforming a Δ*crp* strain (AJW5246) with pDCRP encoding the WT (AJW5248), K100Q (AJW5250), or K100R (AJW5251) CRP variants. I grew these strains in triplicate in TB7 supplemented with 22 mM glucose until early stationary phase (OD_{600} 1.8). RNA was extracted and submitted for RNAseq analysis, performed by Gina Kuffel in the Loyola Genomics Facility. The RNAseq data have been deposited in NCBI's Gene Expression Omnibus (184) and are accessible through GEO Series accession number GSE97406 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE97406). Processing and statistical analysis of the data was performed by Gina; I performed the analysis of the genes identified as differentially regulated in the pairwise comparisons between the strains expressing the WT, K100Q, and K100R CRP variants. I expected a general decrease in transcription of genes driven

by Class II CRP-dependent promoters in the strain expressing K100Q relative to the strains expressing either WT CRP or K100R, much like I saw at *CC(-41.5)* (**Fig. 11**). Furthermore, in contrast to *CC(-61.5)*, I expected a general increase in transcription of genes driven by Class I CRP-dependent promoters in the strain expressing K100Q relative to the strains expressing either WT CRP or K100R for the following reasons: 1) endogenous CRP-dependent promoters do not contain consensus CRP binding sites (191), and are therefore sensitive to CRP concentration, and 2) strains expressing K100Q have greater CRP steady state levels than strains expressing WT CRP or K100R (**Fig. 10**), so CRP-dependent promoters should be occupied by CRP more often in the strain expressing K100Q relative to strains expressing WT CRP or K100R.

Relative to the strain expressing WT CRP, a total of 86 and 67 genes were differentially expressed in the strains expressing the K100Q and K100R variants, respectively (**Table 2A**). These data support the hypothesis that K100 plays a role in regulating global transcription.

However, I chose to focus on comparing the data between the strains expressing the K100Q and K100R variants, since these variants represent the complete presence or absence of the K100 positive charge on all CRP molecules in the cell. This is in contrast to the strain expressing WT CRP, which likely consists of a mixture of CRP molecules that are either acetylated or unacetylated on K100. Comparing strains expressing the K100Q and K100R variants, a total of 386 genes were differentially regulated (**Table 2A**). Twelve out of eighteen (67%) differentially regulated genes driven by Class II CRP-dependent promoters displayed lower expression in the K100Q mutant relative to the K100R mutant, suggesting the positive charge is important for increased transcription from at least a subset of Class II promoters.

Table 2. Number of genes and promoters whose regulation was significantly altered by the K100 charge status.

A.

		RNAseq[a]			DNA Microarray[b]											
		TB7 + 22 mM Glucose			M9 + 10 mM Glucose						M9 + 30 mM Acetate					
		QvWT	RvWT	QvR	QvN-X	RvN-X	QvR-X	QvN-S	RvN-S	QvR-S	QvN-X	RvN-X	QvR-X	QvN-S	RvN-S	QvR-S
Class II Regulated Genes	Up	2	6	6	3	3	2	7	4	6	1	0	1	2	1	3
	Down	2	0	12	6	14	3	2	15	1	32	6	31	16	1	36
Class I Regulated Genes	Up	6	3	16	27	14	9	33	7	34	8	9	8	38	0	45
	Down	4	1	10	7	32	4	10	67	2	25	19	25	25	0	42
Non-CRP Regulated Genes	Up	40	26	194	145	182	153	148	84	451	100	55	78	228	24	425
	Down	32	31	148	155	267	149	128	830	25	247	241	167	256	61	486
Total Genes	Up	48	35	216	175	199	164	188	95	491	109	64	87	268	25	473
	Down	38	32	170	168	313	156	140	912	28	304	266	223	297	62	564
Total Genes	**All**	**86**	**67**	**386**	**343**	**512**	**320**	**328**	**1007**	**519**	**413**	**330**	**310**	**565**	**87**	**1037**

B.

		RNAseq[a]			DNA Microarray[b]											
		TB7 + 22 mM Glucose			M9 + 10 mM Glucose						M9 + 30 mM Acetate					
		QvWT	RvWT	QvR	QvN-X	RvN-X	QvR-X	QvN-S	RvN-S	QvR-S	QvN-X	RvN-X	QvR-X	QvN-S	RvN-S	QvR-S
Class II Promoters	Up	2	3	6	3	3	2	7	2	6	1	0	1	2	1	3
	Down	2	0	12	4	12	3	2	11	1	23	5	21	13	1	22
Class I Promoters	Up	5	3	14	20	11	8	25	7	26	8	9	7	29	0	33
	Down	4	1	9	7	25	4	6	44	2	19	18	21	16	0	26
Non-CRP Promoters	Up	39	22	179	124	163	146	128	63	398	85	54	63	187	23	317
	Down	32	27	137	145	244	138	122	661	23	222	217	151	230	58	421
Total Promoters	Up	46	28	199	147	177	156	160	7	430	94	63	71	218	24	353
	Down	38	28	158	156	281	145	130	716	26	264	240	193	259	59	469
Total Promoters	**All**	**84**	**56**	**357**	**303**	**458**	**301**	**290**	**723**	**456**	**358**	**303**	**264**	**477**	**83**	**822**

[a] WT, Q, and R refer to Δ*crp* strains expressing the WT, K100Q, and K100R CRP variants respectively from a multicopy plasmid. A gene was considered "up" or "down" if there was a statistically significant difference in gene expression between the two strains being compared, and that difference was at least 1.5-fold.

[b] N, Q, and R refer to Δ*crp* strains expressing the WT, K100Q, and K100R CRP variants respectively from a *crp* insertion in the *paaH* locus in the chromosome. A gene was considered "up" or "down" if there was a statistically significant difference in gene expression between the two strains being compared, regardless of the magnitude of that difference.

Sixteen out of twenty-six (62%) differentially regulated genes driven by promoters with at least one Class I CRP binding site, but no Class II site, exhibited increased transcription in the K100Q mutant relative to the K100R mutant. This implies that the positive charge of K100 does impact transcription of Class I genes with a tendency to decrease transcription.

The positive charge of K100 exerted its biggest impact on transcription of genes not directly regulated by CRP. Relative to the K100R mutant, the K100Q mutant had a total of 194 genes with significantly increased expression and 148 genes with significantly decreased expression (**Table 2A**). CRP is known to regulate transcription of sigma factors and other transcription factors (39–44); each of these could influence transcription of genes beyond direct CRP control. These results further support the hypothesis that the positive charge of K100 can impact global transcription.

While the RNAseq data show that the CRP K100 positive charge affects the transcription of many genes, the data may be misleading because CRP does not regulate the transcription of individual genes directly; CRP regulates transcription from promoters, from which RNAP can transcribe multiple genes at once if these genes are organized within an operon. This could lead to an exaggeration of the role of the K100 positive charge; if multiple genes are regulated by the same CRP-dependent promoter, then alterations to CRP would affect transcription of all of these genes, even though only one promoter is affected.

To address this concern, I used the EcoCyc *E. coli* Database (ecocyc.org) to identify the promoters associated with each of the differentially regulated genes. Since operons can be regulated by multiple promoters, I tried to assign the same promoter for each gene within an operon. I then assessed the number of differentially regulated promoters for each of the CRP

variants, with the expectation that the overall numbers would go down since multiple genes may have the same promoter, but that I would observe the same pattern I saw analyzing individual genes.

Relative to the strain expressing WT CRP, a total of 84 and 56 promoters were differentially regulated in the strains expressing the K100Q and K100R variants, respectively (**Table 2B**). Comparing the strains expressing the K100Q and K100R variants, a total of 357 promoters were differentially regulated. Twelve out of eighteen (67%) of the differentially regulated Class II promoters displayed decreased activity in the K100Q mutant relative to the K100R mutant. Fourteen out of twenty-three (61%) of the differentially regulated Class I promoters displayed increased activity in the K100Q mutant relative to the K100R mutant. For promoters not directly regulated by CRP, 179 promoters displayed increased activity and 137 promoters displayed decreased activity in the K100Q mutant relative to the K100R mutant. These values are very close to those observed when analyzing individual gene expression, suggesting 1) many of the promoters affected by the K100 positive charge only regulate a single gene, 2) in many cases, only one gene within an operon was detected as differentially regulated by the K100 positive charge, or 3) a combination of both. Regardless, these data support the hypothesis that, in general, the CRP K100 positive charge promotes transcription from Class II promoters and reduces transcription from Class I promoters, which can influence transcription from promoters that are not directly CRP-dependent.

Single Copy Chromosomal Expression of *crp* in Minimal Media

While I was analyzing global gene expression in strains overexpressing CRP K100 acetylation mimics, we were informed that Ana Écija-Conesa from the lab of Dr. Manuel

Canovas was analyzing global gene expression in strains expressing CRP K100 acetylation mimics in single copy from the chromosome. Our labs decided to collaborate and share data.

To generate strains expressing CRP variants from the chromosome in single copy, Ana began with a Δ*crp* strain and used λRed recombineering (195) to replace a specific locus in the chromosome with the WT, K100Q, and K100R *crp* alleles. She was unsuccessful in replacing the native *crp* locus for unknown reasons, so she replaced the *paaH* locus instead. The *paaH* locus was chosen for three reasons: 1) The *paa* operon is not essential for *E. coli* growth on minimal media with glucose or acetate as the sole carbon source (196); 2) The *paa* operon is not expressed during growth on minimal media with glucose or acetate as the sole carbon source (Davis *et al.* in revision); and 3) The distance between the *paaH* and *crp* loci is large enough to avoid undesired recombination events during the gene insertion process. The newly recombined strain carrying the WT *crp* allele in place of *paaH* is referred to as the "N" reference strain, while the strains carrying the K100Q and K100R alleles are called the K100Q and K100R strains respectively.

To assess differences in global gene expression between the N, K100Q, and K100R strains, Ana grew these strains in minimal media with either 10 mM glucose or 30 mM acetate as the sole carbon source. These carbon sources were chosen to maximize or minimize global acetylation, respectively, including K100 acetylation in the N strain. Samples were collected in exponential phase (OD_{600} 0.5) and in stationary phase (OD_{600} 1.5) and subjected to DNA microarray analysis. The global gene expression data discussed in this work have been deposited in NCBI's Gene Expression Omnibus (184) and are accessible through GEO Series accession number GSE96955

(https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE96955). Processing and statistical analysis of the data was performed by members of the Canovas lab; I performed the analysis of the genes identified as differentially regulated in the pairwise comparisons between the N, K100Q, and K100R strains.

Relative to the reference strain N expressing WT CRP, several genes were differentially expressed in the strains expressing the K100Q and K100R variants, depending on media and growth phase (**Table 2A**). These data support the hypothesis that K100 plays a role in regulating global transcription.

To specifically assess the importance of K100's positive charge, I compared gene expression between the K100Q mutant and the K100R mutant for each growth condition (**Table 2A**). Regarding Class II promoters, few genes were differentially expressed during growth on glucose (exponential: 2 up/3 down; stationary: 6 up/1 down). In contrast, during growth on acetate (both exponential [1 up/31 down] and stationary phase [3 up/36 down]), transcription of many genes driven by promoters with a Class II CRP binding site was reduced in the K100Q mutant relative to the K100R mutant. These data agree with the RNAseq and promoter fusion assays, supporting the hypothesis that the K100 positive charge promotes Class II transcription.

The effect of the K100 positive charge on transcription of genes with Class I CRP binding sites varied, depending on the carbon source and the phase of growth (**Table 2A**). In stationary phase cultures grown in glucose, the majority of differentially regulated Class I genes displayed increased transcription in the K100Q mutant relative to the K100R mutant (34 up/2 down); in exponential phase cultures grown in acetate, the opposite was true (8 up/25 down). In the other two growth conditions, there was little difference between the two mutants in the

number of Class I genes with increased or decreased transcription (exponential glucose: 9 up/4 down; stationary acetate: 45 up/42 down). These results support the hypothesis that the K100 positive charge also plays a role in Class I transcription, though that role is dependent on carbon source and growth phase.

As with the RNAseq data, the largest impact of the K100 positive charge was on gene transcription that does not depend directly on CRP (**Table 2A**). Regarding non-CRP genes, for exponential phase cells grown in glucose (153 up/149 down) and stationary phase cells grown in acetate (425 up/486 down), approximately half of the differentially regulated genes had increased transcription in the K100Q mutant relative to the K100R mutant, and half had decreased transcription. For stationary phase cells grown in glucose, many more genes had increased transcription in the K100Q mutant relative to the K100R mutant (451 up/25 down); for exponential phase cells grown in acetate, the opposite was true (78 up/167 down). Furthermore, there were many more differentially regulated genes in stationary phase for each carbon source [519 (491 up + 28 down) in glucose and 1037 (473 up + 564 down) in acetate] than in exponential phase [320 (164 up + 156 down) in glucose and 310 (87 up + 223 down) in acetate], suggesting the K100 positive charge may play a more significant role in stationary phase than in exponential phase. Taken together, the RNAseq and DNA microarray data argue that the K100 positive charge plays a major role in global transcription.

As with the differentially expressed genes identified by RNAseq, I used the EcoCyc *E. coli* Database (ecocyc.org) to identify the promoters associated with each of the differentially regulated genes identified by DNA microarray. The number of promoters that drive transcription of the differentially regulated genes is less than the number of differentially

regulated genes in most of the conditions (compare Tables 2A and 2B), but for each condition the pattern of differentially regulated promoters matches the pattern of differentially regulated genes. These data support the hypothesis that the large number of genes differentially regulated by the K100 positive charge is not an artifact of measuring gene expression instead of promoter activity, but instead is indicative of the major role of the K100 positive charge in global transcription.

Pathway Enrichment Analysis of K100-Regulated Genes

RNAseq and DNA microarray analyses identified hundreds of genes that are regulated by the K100 positive charge under different media and growth conditions (**Table 2A**). I reasoned that because the manipulation of the K100 positive charge affected the expression of so many genes, the physiology of the cell would also be affected. However, because the K100 positive charge regulates different sets of genes under different media and growth conditions, the physiological consequences of K100 manipulation would also depend on media and growth conditions.

To identify pathways that might be particularly affected by the K100 charge status, I performed a pathway enrichment analysis of the genes that were differentially regulated between the K100Q and K100R mutants using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) (197, 198).

Using either the RNAseq or microarray data, growth in glucose did not result in any significantly enriched pathways as measured by the Benjamini-Hochberg corrected P-values for multiple tests (**Table 3**).

Table 3. DAVID analysis of genes significantly altered by the K100 charge status.

QvR Glucose (RNAseq):

Category	Term	Count	%	P-Value	Benjamini
KEGG_PATHWAY	Cysteine and methionine metabolism	8	2	1.20E-02	6.40E-01
KEGG_PATHWAY	Glycine, serine and threonine metabolism	8	2	3.80E-02	8.10E-01
KEGG_PATHWAY	Glutathione metabolism	5	1.2	7.20E-02	8.80E-01
KEGG_PATHWAY	2-Oxocarboxylic acid metabolism	6	1.5	7.80E-02	8.20E-01
KEGG_PATHWAY	Ribosome	12	3	8.50E-02	7.80E-01

QvR Exponential Glucose (Microarray):

Category	Term	Count	%	P-Value	Benjamini
KEGG_PATHWAY	Peptidoglycan biosynthesis	6	2	1.30E-02	6.30E-01
KEGG_PATHWAY	Metabolic pathways	56	18.7	2.10E-02	5.60E-01
KEGG_PATHWAY	One carbon pool by folate	4	1.3	5.60E-02	7.60E-01
KEGG_PATHWAY	Nucleotide excision repair	3	1	8.90E-02	8.30E-01

QvR Stationary Glucose (Microarray):

Category	Term	Count	%	P-Value	Benjamini
KEGG_PATHWAY	Two-component system	28	5.7	1.10E-02	6.00E-01
KEGG_PATHWAY	ABC transporters	31	6.4	1.40E-02	4.50E-01
KEGG_PATHWAY	DNA replication	6	1.2	4.40E-02	7.10E-01
KEGG_PATHWAY	Homologous recombination	7	1.4	9.50E-02	8.70E-01

QvR Exponential Acetate (Microarray):

Category	Term	Count	%	P-Value	Benjamini
KEGG_PATHWAY	**Bacterial chemotaxis**	**9**	**3.1**	**7.50E-05**	**5.00E-03**
KEGG_PATHWAY	**Flagellar assembly**	**10**	**3.4**	**1.50E-03**	**4.80E-02**
KEGG_PATHWAY	Phosphotransferase system (PTS)	9	3.1	1.80E-02	3.30E-01
KEGG_PATHWAY	Galactose metabolism	7	2.4	5.50E-02	6.10E-01
KEGG_PATHWAY	Two-component system	18	6.1	6.60E-02	6.00E-01
KEGG_PATHWAY	Fructose and mannose metabolism	7	2.4	9.40E-02	6.70E-01
KEGG_PATHWAY	Alanine, aspartate and glutamate metabolism	6	2	9.60E-02	6.20E-01

QvR Stationary Acetate (Microarray):

Category	Term	Count	%	P-Value	Benjamini
KEGG_PATHWAY	**Biosynthesis of amino acids**	**53**	**5.4**	**2.30E-07**	**2.30E-05**
KEGG_PATHWAY	**Metabolic pathways**	**198**	**20.2**	**4.60E-05**	**2.30E-03**
KEGG_PATHWAY	**ABC transporters**	**59**	**6**	**5.70E-04**	**1.90E-02**
KEGG_PATHWAY	2-Oxocarboxylic acid metabolism	14	1.4	2.70E-03	6.60E-02
KEGG_PATHWAY	Biosynthesis of secondary metabolites	89	9.1	2.90E-03	5.70E-02
KEGG_PATHWAY	Alanine, aspartate and glutamate metabolism	15	1.5	6.00E-03	9.80E-02
KEGG_PATHWAY	Propanoate metabolism	16	1.6	8.10E-03	1.10E-01
KEGG_PATHWAY	Cysteine and methionine metabolism	14	1.4	8.70E-03	1.10E-01
KEGG_PATHWAY	Arginine biosynthesis	10	1	1.20E-02	1.30E-01
KEGG_PATHWAY	Biosynthesis of antibiotics	59	6	1.90E-02	1.80E-01
KEGG_PATHWAY	Taurine and hypotaurine metabolism	5	0.5	3.00E-02	2.50E-01
KEGG_PATHWAY	Histidine metabolism	6	0.6	3.80E-02	2.80E-01
KEGG_PATHWAY	Phenylalanine, tyrosine and tryptophan biosynthesis	10	1	3.80E-02	2.60E-01
KEGG_PATHWAY	beta-Alanine metabolism	8	0.8	4.10E-02	2.60E-01
KEGG_PATHWAY	Butanoate metabolism	14	1.4	4.70E-02	2.80E-01
KEGG_PATHWAY	Valine, leucine and isoleucine biosynthesis	8	0.8	5.80E-02	3.20E-01
KEGG_PATHWAY	Carbon metabolism	33	3.4	7.70E-02	3.80E-01
KEGG_PATHWAY	Methane metabolism	11	1.1	7.90E-02	3.70E-01
KEGG_PATHWAY	beta-Lactam resistance	8	0.8	8.00E-02	3.60E-01
KEGG_PATHWAY	Lipopolysaccharide biosynthesis	12	1.2	8.80E-02	3.80E-01

In contrast with cultures grown in glucose, DAVID analysis did detect pathways that were significantly enriched during growth on acetate. From stationary phase cultures (**Table 3**), biosynthesis of amino acids, metabolic pathways, and ABC transporters were significantly enriched above background. Examination of the differentially regulated genes within these pathways identified several genes associated with the methionine/S-adenosyl-L-methionine (SAM) and arginine synthesis pathways whose transcription were significantly increased in the K100Q mutant relative to the K100R mutant and the reference strain N. However, the transcriptional repressors of these pathways (*metJ* and *argR*, respectively) also were highly upregulated. These repressors become active upon binding to SAM or arginine, respectively (199, 200). These results suggest that, while these pathways are highly expressed, it is likely that there is little flux through the pathways.

From exponential phase cultures grown in acetate, the flagellar regulon was significantly enriched above background. For each gene in this regulon that was differentially regulated, its transcription was elevated in the K100Q mutant relative to the K100R mutant. Many of these genes also were upregulated in the K100Q mutant relative to the reference strain N. Taken together, these observations suggest the K100Q mutant may be more motile than the K100R mutant and the reference strain N.

K100 Regulates Flagellar-Based Motility

E. coli can use flagella to navigate their environment, either toward nutrients or away from repellants (201). Flagella are extremely long, flexible protein polymers that extend out of the cell and rotate at their base, providing thrust to propel the cell through or across a medium.

Regulation of the synthesis of these complex machines is similarly complex, involving three tiers of gene expression (202). The Class I genes (not to be confused with Class I promoters) are *flhD* and *flhC*, which are expressed from a single promoter and encode transcriptional activators of the next tier of genes. Class II genes encode structural proteins of the flagellar base and export machinery, as well as FliA, a sigma factor that activates transcription of the third tier of genes. The Class III genes encode the motor proteins and the flagellar filament.

CRP is required for flagellar motility; it positively regulates transcription of the *flhDC* operon from a Class I promoter (203). Flagellar genes were identified as more highly expressed in strains expressing K100Q relative to strains expressing either WT CRP or K100R in all media and growth conditions tested in both the RNAseq and DNA microarray analyses (Davis *et al*. in revision). I previously observed that strains expressing CRP variants lacking a positive charge at position 100 displayed greater steady state levels than strains expressing WT CRP or K100R (**Fig. 10**), which should increase transcription from Class I promoters in the former strains over the latter strains. While I did not observe this increase in transcription from the Class I promoter *CC(-61.5)* (**Fig. 12**), the CRP binding site at the *flhDC* promoter is not consensus and thus would likely respond to changes in CRP steady state levels (191). I hypothesized that loss of the CRP K100 positive charge enhances expression of flagellar genes by increasing transcription of *flhDC*. If this is the case, then there may be a similar increase in motility in strains expressing K100Q or K100A over strains expressing WT CRP or K100R.

To determine if loss of the K100 positive charge promotes *flhDC* transcription, I transformed a Δ*crp* strain (AJW5246) with pDCRP encoding WT CRP (AJW5248), K100A (AJW5249), K100Q (AJW5250), or K100R (AJW5251). These strains were grown in TB

supplemented with 0.5% NaCl until mid-exponential phase. At this point, cells were harvested and subjected to qRT-PCR analysis using primers specific to *flhD* (**Fig. 15A**). The strains expressing WT CRP and K100R displayed approximately equivalent levels of *flhD* transcript. The strains expressing K100A and K100Q displayed about 50% more *flhD* transcript than the strains expressing either WT CRP or K100R, suggesting the loss of the K100 positive charge does enhance transcription of *flhDC*.

To determine if the enhancement in *flhDC* transcription correlates with an enhancement in flagellar motility, I spotted the same Δ*crp* strains expressing the CRP variants on motility plates (**Fig. 15B**). As the cells engage in flagellar motility, the diameter of the spot increases equally in all directions. After 10 hours at 30°C, the diameters of the spots from strains expressing K100A or K100Q were about 50% greater than the spots from strains expressing WT CRP or K100A. These results correlate well with the quantification of *flhD* transcripts, supporting the hypothesis that loss of the CRP K100 positive charge enhances expression of flagellar genes, and flagellar motility, by increasing transcription of *flhDC*.

K100 is Acetylated by AcP *In Vitro*

We previously reported that CRP K100 acetylation is elevated *in vivo* when acP levels are high. This was true when we compared isogenic mutants that either accumulate acP (*ackA*) or mutants that cannot synthesize acP (*pta ackA*) (121) and when WT cells are grown under conditions that either favor high acP levels (TB7 supplemented with 22 mM glucose) or low acP levels (TB7 with no supplementation) (134). Taken together, these results support the hypothesis that acP is the acetyl donor for CRP K100. However, it is not clear if K100 is

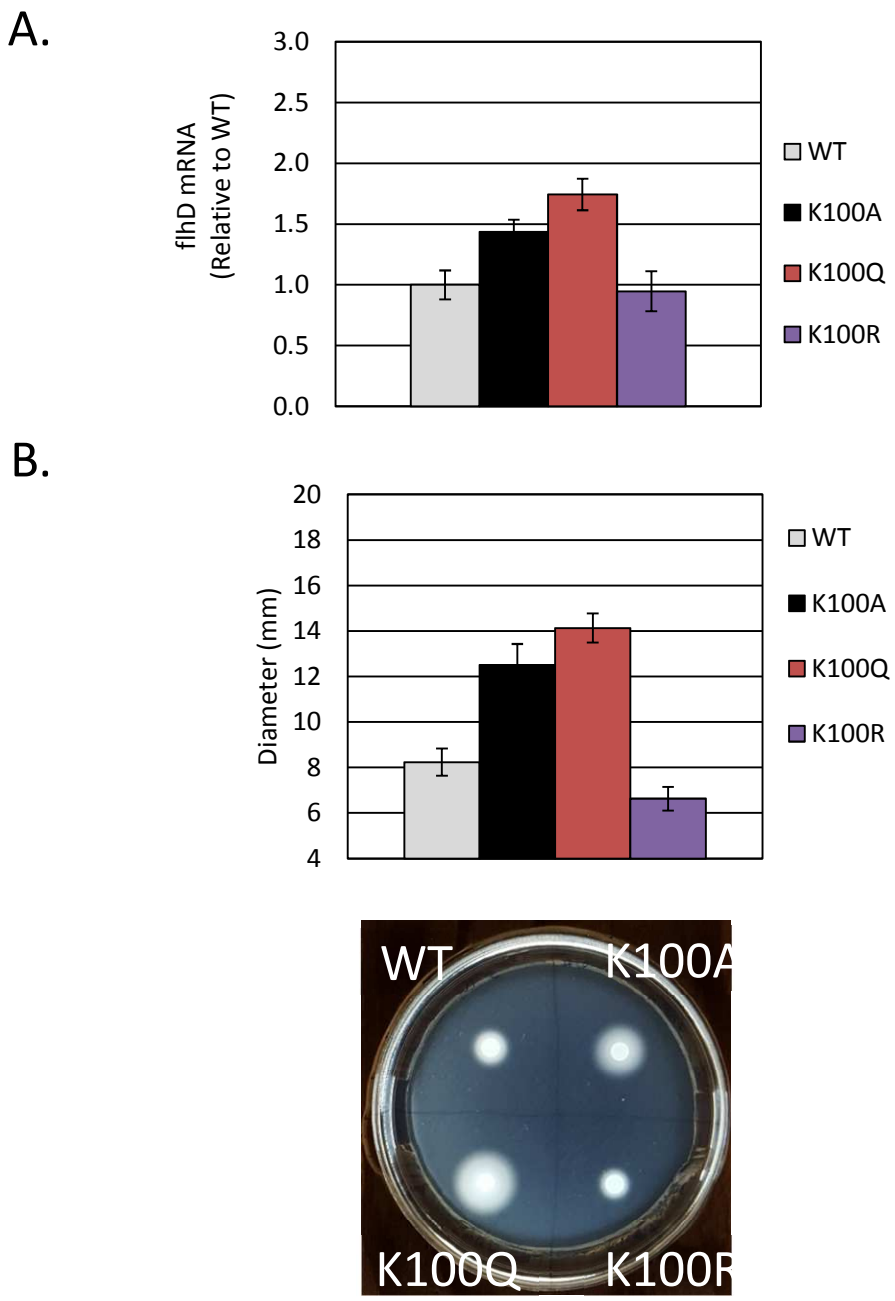

Figure 15. K100-mediated regulation of motility. (A) Strains expressing WT CRP or one of the indicated mutants were grown at 30°C until the cultures reached OD600 ~0.5. Samples were collected for qRT-PCR analysis using *flhD*-specific primers. Primers specific for 16S rRNA were used as a loading control. Data are representative of two experiments performed in duplicate. Error bars represent standard deviation. (B) (Top) Strains expressing WT CRP or one of the indicated mutants were spotted onto semi-solid agar plates and incubated at 30°C. After 10 hours, the diameter of each spot was quantified. Data are representative of two independent experiments performed with at least eight replicates. Error bars represent standard deviation. (Bottom) Representative image of motility plate quantified above.

acetylated by acP directly, or by an unknown acP-dependent acetyltransferase. Since other proteins and peptides can be acetylated directly by acP *in vitro* (120, 121), I expected K100 would also be acetylated directly by acP.

I first addressed if acP could directly acetylate CRP. I performed an *in vitro* acetylation reaction by incubating purified CRP with increasing concentrations of acP (0-12.8 mM) for 15 or 120 minutes at 37°C, and then stopped the reaction by adding 2x sample loading buffer and heating the sample to 95°C. I quantified the relative increase in CRP acetylation by western immunoblot analysis using an anti-acetyllysine antibody. CRP acetylation increased in a time- and acP concentration-dependent manner. At the highest acP concentration, this resulted in a ~2-fold increase at 15 minutes and a ~12-fold increase at 120 minutes (**Fig. 16**). These data support the hypothesis that acP can acetylate at least some lysines within CRP directly.

I next asked specifically which CRP lysines were sensitive to acetylation by acP *in* vitro. Dr. Birgit Schilling from the Buck Institute for Research on Aging performed quantitative mass spectrometric analysis on purified CRP samples incubated for 15 minutes with either no acP or 12.8 mM acP to identify acetylated lysines and to determine the relative increase in acP-dependent acetylation of each lysine. She identified seven lysines that were acetylated in either sample, including K100 (Davis *et al.*, in revision). Acetylation of K100 and K26 increased 2.1- and 2.4-fold, respectively, in the presence of acP (**Fig. 17A and 17B, Table 4**). Exposure to acP resulted in no significant change in the acetylation of K52 or K89; the change in acetylation of the other lysines (K152, K166, K188) could not be quantified due to relatively weak signal, interferences or methionine oxidation. CRP protein level itself did not change upon acP incubation as determined by quantifying 23 non-acetylated peptides (**Fig. 17C and Table 5**).

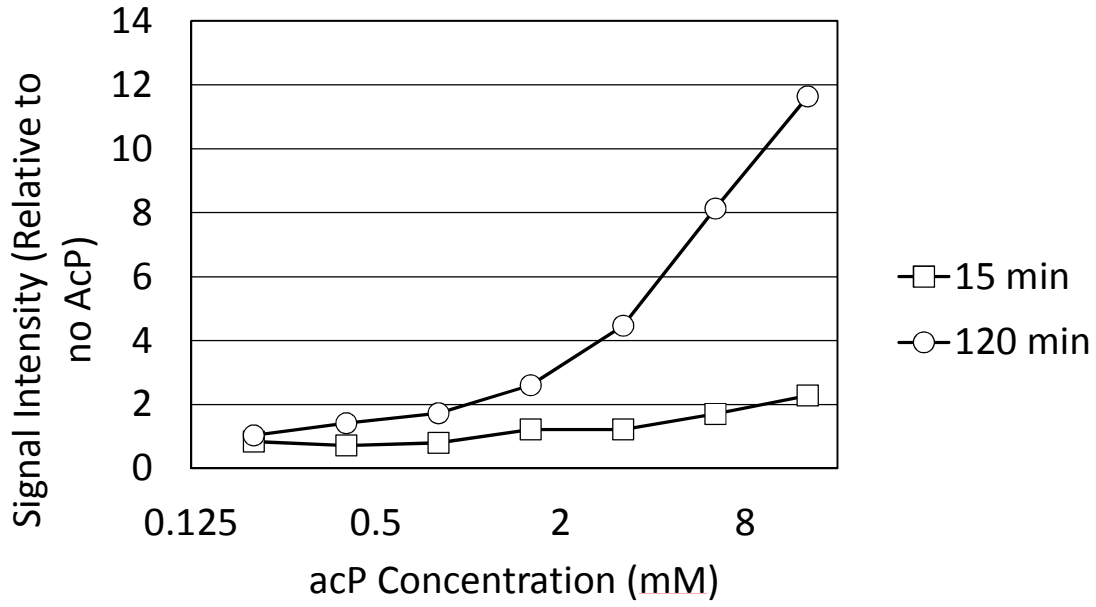

Figure 16. CRP acetylation by acP *in vitro*. Purified CRP (10 µg) was incubated with the indicated concentrations of acP for either 15 or 120 minutes at 37°C. The samples were then subjected to Western immunoblot analysis using anti-acetyllysine antibodies, and the relative level of acetylation was quantified. Data are representative of two independent experiments performed in singlet.

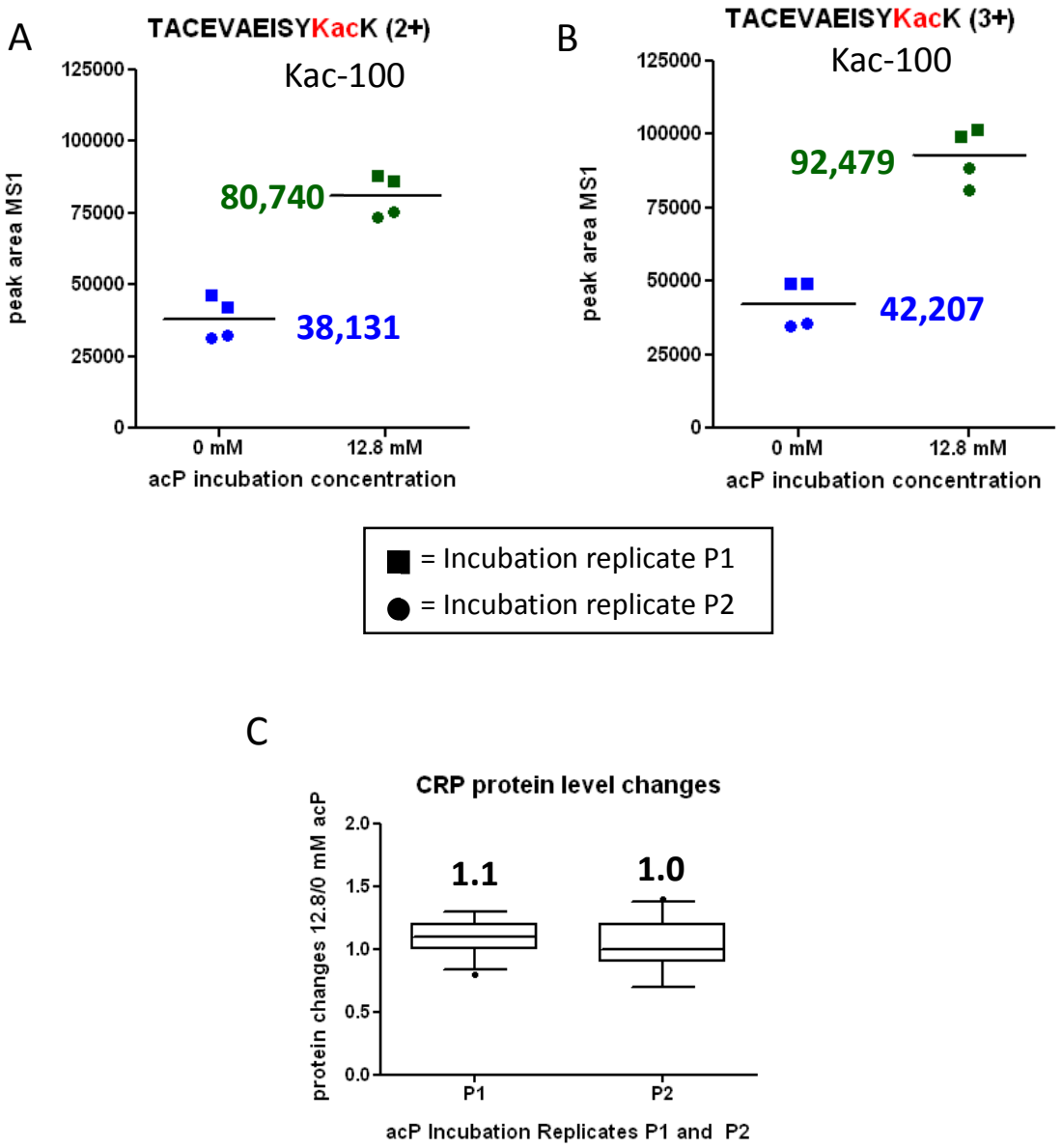

Figure 17. Analysis of K100 acetylation *in vitro*. Purified CRP (10 µg) was incubated with or without 12.8 mM acP for 15 minutes at 37°C. MS1 Filtering was used to quantify relative changes in the abundance of K100 acetylation for two incubation replicates (P1 and P2), each acquired in technical MS duplicates. Results for acetylated K100 showing measured peak areas are presented for acetyl peptide TACEVAEISYKacK in its (A) 2+ and (B) 3+ charge state. (C) CRP protein levels did not change upon acP incubation, as determined by monitoring 23 non-acetylated peptides. For (A), (B), and (C), data are from the lone experiment performed in doublet.

Table 4. Quantification of increases in CRP acetylation after incubation with acP.

Modified Peptide Sequence	Acetyl Site	Precursor M/Z	Precursor Charge	Ratio Incub. Repl. P1 (12.8m M/0m M)	P-Value P1 ratio	Ratio Incub. Repl. P2 (12.8m M/0m M)	P-Value P2 ratio	Mean Ratio (12.8m M/0m M)
TACEVAEISYKacK	**K-100**	720.86	2	**2**	0.013	**2.3**	0.002	**2.1** *
TACEVAEISYKacK	**K-100**	480.91	3	**2**	0.013	**2.4**	0.044	**2.2** *
YPSKacSTLIHQGEK	K-26	510.6	3	2	0.031	2.9	0.012	2.4 *
GSVAVLIKacDEEGK#	*K-52*	*462.59*	*3*	*1.1*	*0.387*	*1.5*	*0.333*	*1.3*
AKacTACEVAEISYK	*K-89*	*756.37*	*2*	*1.6*	*0.084*	*1.5*	*0.332*	*1.6*

* statistically significant ($P < 0.05$)

deamidated C-terminus

Table 5. Quantification of increases in CRP abundance after incubation with acP.

	Both incubation replicates P1 and P2 which both comprised two technical MS injection replicates were quantified														
	peptide peak area ratios for incubations of 12.8 mM acP over 0 mM acP are formed														
Peptide	Transition	Rank	Modified Sequence	Precursor M/Z	Product M/Z	Protein Name	Protein Description	Precursor Charge	Ratio P1_12.8mM/P1_0mM	P-Value P1_12.8mM/P1_0mM	Ratio P2_12.8mM/P2_0mM	P-Value P2_12.8mM/P2_0mM		log ratio P1	log ratio P2
YPSKSTLIHQGEK	precursor	1	YPSKSTLIHQGEK	496.60	496.60	sp|P0ACJ8|CRP_ECOLI	Catabolite gene activator OS=Escheric	3	1.0	0.8494	0.7	0.3150		-0.0109	-0.1400
STLIHQGEKAETLY	precursor	1	STLIHQGEKAETLY	530.61	530.61	sp|P0ACJ8|CRP_ECOLI	Catabolite gene activator OS=Escheric	3	1.1	0.0818	1.3	0.0039		0.0577	0.0982
AETLYYIVK	precursor	1	AETLYYIVK	550.31	550.31	sp|P0ACJ8|CRP_ECOLI	Catabolite gene activator OS=Escheric	2	1.1	0.0789	0.7	0.0015		0.0453	-0.1480
GSVAVLIKDEEGK	precursor	1	GSVAVLIKDEEGK	672.87	672.87	sp|P0ACJ8|CRP_ECOLI	Catabolite gene activator OS=Escheric	2	1.1	0.2311	1.4	0.1409		0.0454	0.1434
GSVAVLIKDEEGK	precursor	1	GSVAVLIKDEEGK	448.92	448.92	sp|P0ACJ8|CRP_ECOLI	Catabolite gene activator OS=Escheric	3	1.2	0.0566	1.3	0.1663		0.0906	0.1057
AKTACEVAEISYK	precursor	1	AKTAC[+57]EVAEISYK	735.37	735.37	sp|P0ACJ8|CRP_ECOLI	Catabolite gene activator OS=Escheric	2	1.2	0.0171	1.1	0.1690		0.0886	0.0460
AKTACEVAEISYKK	precursor	1	AKTAC[+57]EVAEISYKK	533.28	533.28	sp|P0ACJ8|CRP_ECOLI	Catabolite gene activator OS=Escheric	3	1.3	0.1231	1.0	0.5204		0.1104	0.0177
AKTACEVAEISYKK	precursor	1	AKTAC[+57]EVAEISYKK	400.21	400.21	sp|P0ACJ8|CRP_ECOLI	Catabolite gene activator OS=Escheric	4	1.3	0.1103	1.0	0.6200		0.1117	-0.0159
TACEVAEISYK	precursor	1	TAC[+57]EVAEISYK	635.80	635.80	sp|P0ACJ8|CRP_ECOLI	Catabolite gene activator OS=Escheric	2	1.0	0.9602	1.2	0.0214		0.0019	0.0697
TACEVAEISYK	precursor	1	TAC[+57]EVAEISYK	424.20	424.20	sp|P0ACJ8|CRP_ECOLI	Catabolite gene activator OS=Escheric	3	1.0	0.9797	1.2	0.1046		-0.0013	0.0622
TACEVAEISYKK	precursor	1	TAC[+57]EVAEISYKK	699.85	699.85	sp|P0ACJ8|CRP_ECOLI	Catabolite gene activator OS=Escheric	2	1.2	0.0783	1.1	0.0432		0.0752	0.0291
TACEVAEISYKK	precursor	1	TAC[+57]EVAEISYKK	466.90	466.90	sp|P0ACJ8|CRP_ECOLI	Catabolite gene activator OS=Escheric	3	1.2	0.0185	1.0	0.2212		0.0704	0.0149
QLIQVNPDILMR	precursor	1	Q[-17]LIQVNPDILM[+16]R	719.89	719.89	sp|P0ACJ8|CRP_ECOLI	Catabolite gene activator OS=Escheric	2	1.1	0.8230	1.1	0.3774		0.0213	0.0442
LQVTSEK	precursor	1	LQVTSEK	402.73	402.73	sp|P0ACJ8|CRP_ECOLI	Catabolite gene activator OS=Escheric	2	1.0	0.2315	1.3	0.0225		-0.0192	0.1012
VGNLAFLDVTGR	precursor	1	VGNLAFLDVTGR	631.35	631.35	sp|P0ACJ8|CRP_ECOLI	Catabolite gene activator OS=Escheric	2	1.0	0.4261	0.7	0.0182		0.0072	-0.1254
VGNLAFLDVTGR	precursor	1	VGNLAFLDVTGR	421.23	421.23	sp|P0ACJ8|CRP_ECOLI	Catabolite gene activator OS=Escheric	3	1.0	0.6801	0.7	0.0060		-0.0040	-0.1326
IAQTLLNLAK	precursor	1	IAQTLLNLAK	542.84	542.84	sp|P0ACJ8|CRP_ECOLI	Catabolite gene activator OS=Escheric	2	1.1	0.0277	0.9	0.0311		0.0419	-0.0499
IAQTLLNLAK	precursor	1	IAQ[+1]TLLNLAK	543.33	543.33	sp|P0ACJ8|CRP_ECOLI	Catabolite gene activator OS=Escheric	2	1.1	0.0391	0.9	0.0120		0.0470	-0.0522
IAQTLLNLAK	precursor	1	IAQTLLN[+1]LAK	543.33	543.33	sp|P0ACJ8|CRP_ECOLI	Catabolite gene activator OS=Escheric	2	1.0	0.2252	1.0	0.2515		0.0062	-0.0164
QPDAMTHPDGMQIK	precursor	1	Q[-17]PDAM[+16]THPDGM[+	792.35	792.35	sp|P0ACJ8|CRP_ECOLI	Catabolite gene activator OS=Escheric	2	0.8	0.0402	1.1	0.1240		-0.1161	0.0346
QEIGQIVGCSR	precursor	1	Q[-17]EIGQIVGC[+57]SR	615.30	615.30	sp|P0ACJ8|CRP_ECOLI	Catabolite gene activator OS=Escheric	2	1.0	0.8367	1.2	0.0464		-0.0005	0.0619
MLEDQNLISAHGK	precursor	1	M[+16]LEDQNLISAHGK	491.25	491.25	sp|P0ACJ8|CRP_ECOLI	Catabolite gene activator OS=Escheric	3	1.0	0.3835	1.0	0.2325		0.0109	0.0119
TIVVYGTR	precursor	1	TIVVYGTR	454.76	454.76	sp|P0ACJ8|CRP_ECOLI	Catabolite gene activator OS=Escheric	2	1.0	0.5552	0.9	0.0316		0.0084	-0.0437
Note: mean ratios across all non-acetylated peptides for incubation replicates P1 and P2, respectively, are formed after log transformation.												Mean (log)		0.0299	0.0051
												Mean (Natural)		1.07	1.01
Protein level changes in P1 = mean ratio (12.8/0 mM acP) = 1.07															
Protein level changes in P2 = mean ratio (12.8/0 mM acP) = 1.01															

These data give further support to the hypothesis that CRP K100 is sensitive to acP-dependent acetylation, and that this acetylation is direct.

Stability of K100 and K101 Mutants

Earlier, I observed that strains expressing CRP variants lacking a positive charge at either position 100 or 101 displayed greater CRP steady state levels than strains expressing CRP variants with a positive charge at either position (**Fig. 10**). One attractive hypothesis is that loss of either positive charge at positions 100 or 101 promote the transcription of *crp* via a combination of two modes of autoregulation: positive autoregulation from the Class I *crp* promoter (64), and negative autoregulation by promoting transcription of a divergent gene (*yhfA*) from a Class II promoter, leading to inhibition of *crp* transcription through steric hindrance (65). Loss of the K101 positive charge is known to result in defective Class II transcription (25), and I showed that loss of the K100 positive charge does as well (**Fig. 11**). Additionally, loss of either positive charge results in an increase in CRP steady state level (**Fig. 10**), which increases transcription from at least some Class I promoters (**Table 2, Fig. 15A**). I expected that loss of either the K100 or K101 positive charges would 1) increase transcription from the Class I *crp* promoter, and/or 2) decrease transcription from the Class II *yhfA* promoter. Either outcome would result in an increase in transcription of *crp*.

To determine if the loss of the CRP K100 and K101 positive charges increase *crp* transcription, I quantified *crp* transcript levels in each of the strains expressing the CRP variants. I grew Δ*crp* strains carrying pDCRP encoding the CRP K100 and K101 variants in TB7 supplemented with 22 mM glucose until mid-exponential phase (OD_{600} ~1.0). After extracting

RNA from the cultures, I performed qRT-PCR using *crp*-specific primers. Contrary to my expectations, there was very little difference in *crp* transcript levels between strains expressing any of the CRP variants (**Fig. 18**). These data suggest that regulation of CRP steady state level by the K100 and K101 positive charges is not at the level of *crp* transcription, but somewhere downstream.

One possibility is that the loss of the K100 and K101 positive charges enhance CRP stability, resulting in greater CRP steady state levels. This hypothesis is supported by reports that the stability of another protein, RNase R, is regulated by the positive charge of an acetylatable lysine (148, 160, 162).

I tested the hypothesis that positive charges at positions 100 and 101 reduce CRP stability using a CRP stability assay. I grew Δ*crp* strains carrying pDCRP encoding the CRP K100 and K101 variants in TB7 supplemented with 22 mM glucose until early-exponential phase (OD_{600} ~0.5). I used the K101E variant, which results in a reversal of the K101 positive charge, whose steady state levels are elevated to levels similar to the K101A variant (**Fig. 10**). I then added chloramphenicol to each culture to prevent the synthesis of new proteins. At regular intervals after chloramphenicol addition, I collected samples of each culture to quantify the relative amount of remaining CRP by western immunoblot analysis. Based on the rates of CRP loss, WT CRP displayed a half-life of about 107 minutes (**Fig. 19A**). The K100R variant was less stable, displaying a half-life of about 77 minutes. In contrast, the K100Q variant was considerably more stable, displaying a half-life of 256 minutes. The K101E mutant, which previously displayed the greatest steady state levels, displayed no appreciable decrease in CRP levels over the course of the experiment, suggesting this variant is the most stable of those

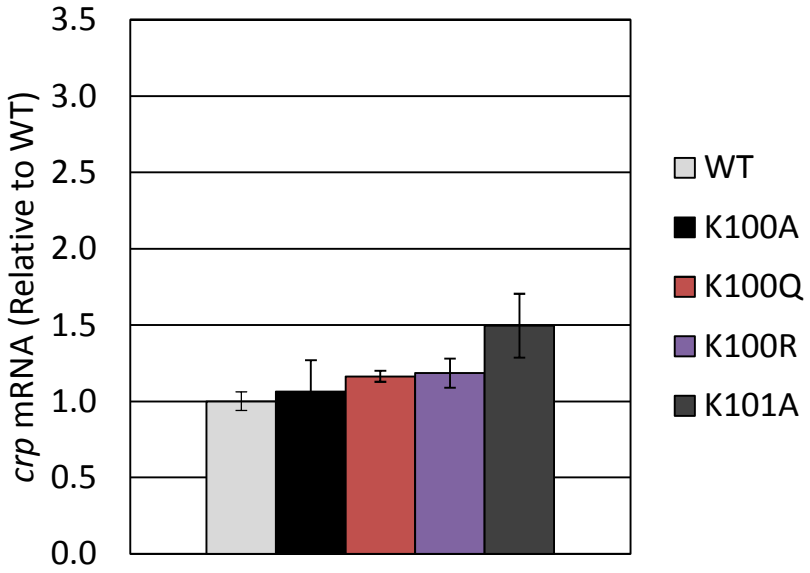

Figure 18. Quantification of *crp* mRNA. Δ*crp* strains carrying both pDCRP expressing the indicated CRP variants and pRW50 carrying the *CC(-41.5)* promoter fused to *lacZ* were grown in TB7 supplemented with 22 mM glucose at 37°C. Cells were harvested at OD_{600} 1.0 and subjected to qRT-PCR analysis using *crp*-specific primers. Data shows average results from cultures grown in triplicate and represents multiple independent experiments. Error bars represent standard deviation.

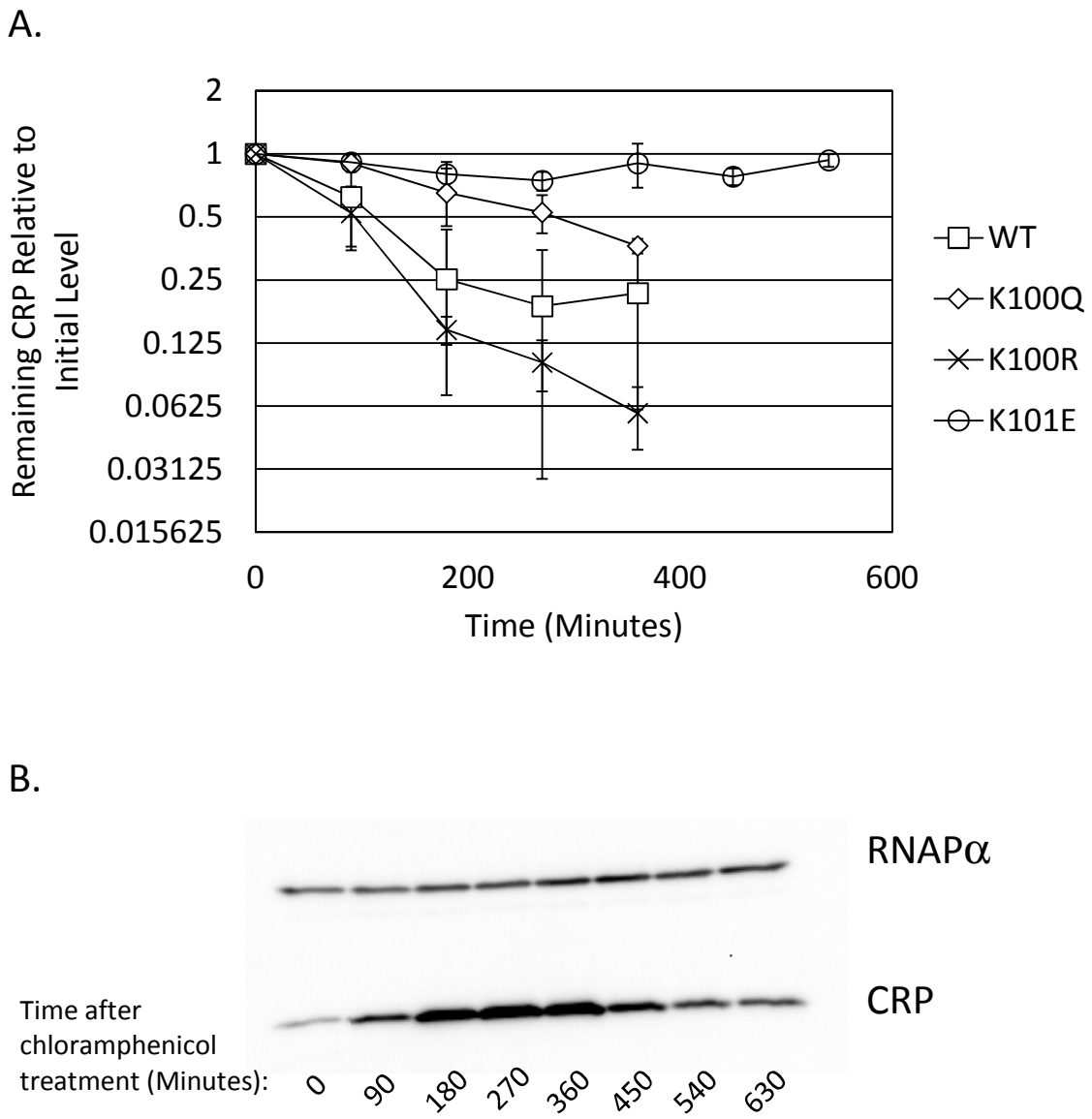

Figure 19. Stability of CRP variants. (A) Δ*crp* strains carrying pDCRP expressing the indicated CRP variants were grown in TB7 supplemented with 22 mM glucose at 37°C. At OD_{600} 0.5, chloramphenicol was added to inhibit protein synthesis. At regular intervals, samples were collected for Western immunoblot analysis using an anti-CRP antibody. Time 0 is the timepoint at which CRP concentrations were greatest for each variant. Data shows average results from cultures grown in triplicate and represents two independent experiments, with the exception that the data from K100R represents a single experiment performed in triplicate. Error bars represent standard deviation. (B) A Western blot from (A) showing the increase in CRP K100Q concentration after chloramphenicol addition prior to the decrease observed due to CRP degradation. This blot is representative of the strains expressing the WT, K100Q, and K100R variants.

tested. Together, these results support the hypothesis that the positive charges at positions 100 and 101 regulate CRP stability, and that the loss of either of these positive charges increases stability of the protein.

There are a few caveats to the above experiment. First, the experiment was ended before the complete degradation of CRP. Without following the degradation to its endpoint, I only know the shape of a section of the degradation curve, which may not represent the shape of the full degradation curve. As a result, the best-fit line from which the half-lives were calculated may not be accurate in the context of the entire degradation curve. Second, there was an issue with the inhibition of protein synthesis. After the addition of chloramphenicol, CRP levels would continue to rise for a few hours before finally starting to decrease (**Fig. 19B**). As a result of this issue, I considered the 0 timepoint to be the timepoint at which CRP levels were greatest for each replicate and calculated the degradation curve starting from this point. Both of these issues need to be addressed for any future studies of CRP half-life. However, the data still suggest that there is some difference in the stability of the CRP variants, and these differences warrant further investigation into the roles of the K100 and K101 positive charges in CRP stability.

Identification of Proteins Involved in CRP Degradation

In the previous section, I provided evidence that loss of the CRP K100 and K101 positive charges enhance the stability of CRP. The simplest explanation is that these positive charges are required for proper CRP degradation. The positive charges may be part of a motif that is recognized by a protease or chaperone protein, or the positive charges may comprise a

protease cleavage site, similar to trypsin-like proteases, which cleave after positively charged lysine and arginine residues. Loss of either positive charge may change the interaction between CRP and the relevant protease and/or chaperone.

The protease responsible for CRP degradation has not yet been identified. However, it has been reported that the chaperone protein DnaJ is required for CRP stability; loss of DnaJ reduces CRP half-life from >80 minutes to 3 minutes (204). The mechanism by which DnaJ protects CRP from degradation is not understood, but one possibility is that DnaJ interacts with CRP directly, shielding CRP from protease activity. If this is true, I hypothesize the CRP K100 and K101 positive charges are inhibitory to the DnaJ-CRP interaction. In the absence of the K100 and K101 positive charges, for example by neutralization with an acetyl group, there would be increased interaction between DnaJ and CRP, greater CRP stability, and greater CRP steady state levels. However, in the absence of DnaJ, the DnaJ-CRP interaction no longer exists, so the status of the K100 and K101 positive charges should have no effect on the stability of CRP, leading all CRP variants to be equally unstable.

To test this hypothesis, I generated a Δ*crp* Δ*dnaJ* double mutant strain (AJW5587) and transformed it with pDCRP encoding WT CRP (AJW5599), K100A (AJW5600), K100Q (AJW5601), K100R (AJW5602), K101A (AJW5603), or K101E (AJW5604). I grew each of these strain in TB7 supplemented with 22 mM glucose until early exponential phase (OD_{600} ~0.7) and performed Western immunoblot analysis using an anti-CRP antibody. CRP was not detectable in the strain expressing WT CRP (**Fig. 20A**), in line with the significantly reduced stability of CRP (204). CRP also was not detectable in the strains expressing K100A or K100R. However, CRP was still readily detectable in the strains expressing K100Q, K101A, or K101E. These results suggest that

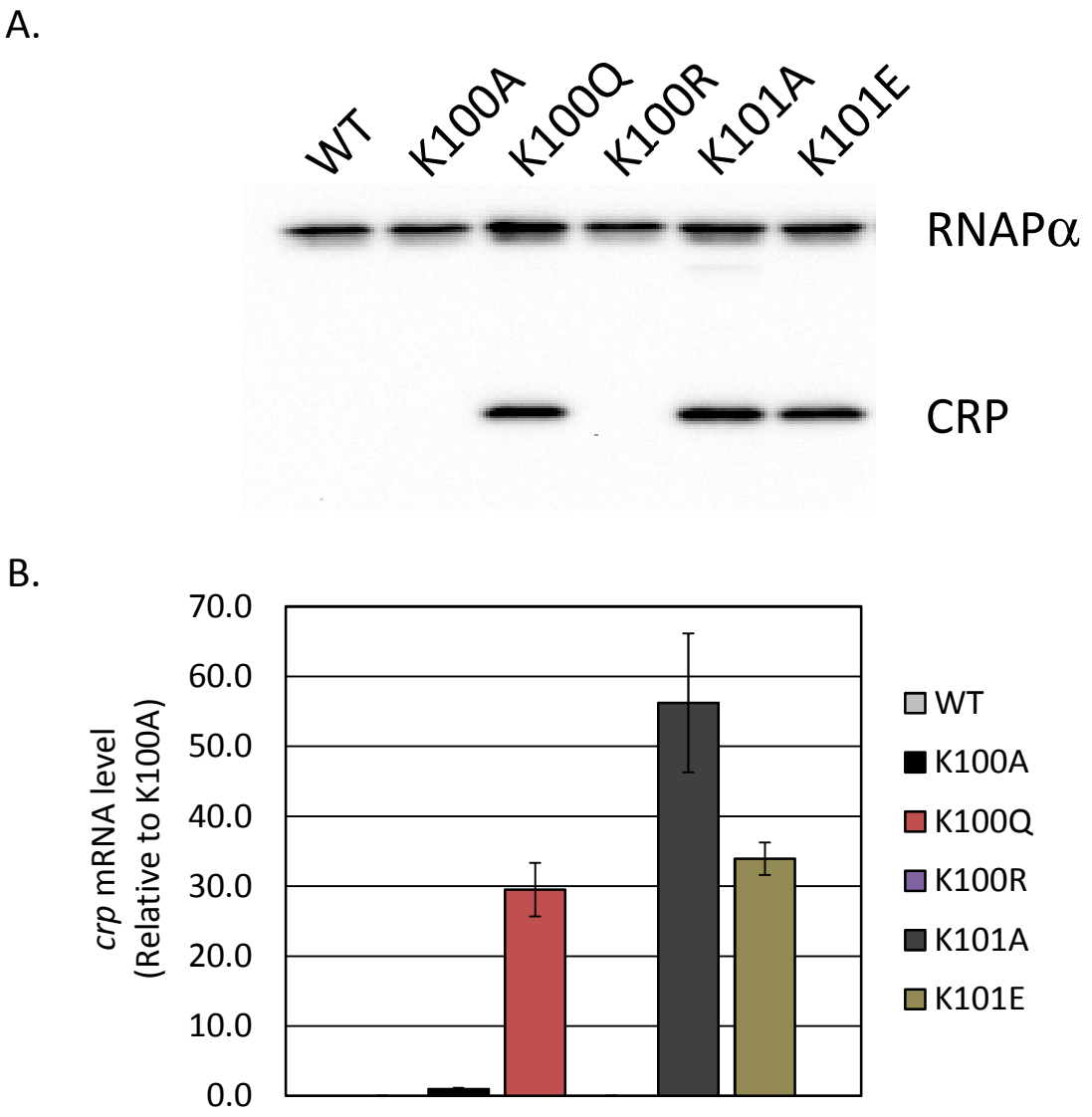

Figure 20. Role of K100 positive charge in DnaJ-mediated CRP stability. (A) Δ*dnaJ* Δ*crp* strains carrying pDCRP were grown in TB7 supplemented with 22mM glucose at 37°C until entry into stationary phase. The cultures were harvested and subjected to Western immunoblot analysis using an anti-CRP antibody. This blot is representative of two independent experiments performed in triplicate. (B) Δ*dnaJ* Δ*crp* strains carrying pDCRP were grown in TB7 supplemented with 22mM glucose at 37°C until entry into stationary phase. The cultures were harvested and subjected to qRT-PCR using *crp*-specific primers. Data shows average results from sole experiment performed in triplicate.

the CRP variants lacking the K100 or K101 positive charge (with the exception of K100A) are stable even in the absence of DnaJ.

To determine if the undetected CRP variants were still transcribed, I performed qRT-PCR using *crp*-specific primers on the samples from the above experiment. There were very low-to-undetectable levels of *crp* transcript in the strains expressing WT CRP, K100A, and K100R (**Fig. 20B),** consistent with the inability to detect CRP protein in these strains (**Fig. 20A**). These results are in opposition to previously published data showing that WT *crp* is still transcribed in a Δ*dnaJ* mutant (205). My data indicate I cannot make any conclusions regarding the role of the K100 or K101 positive charges in the potential DnaJ-CRP interaction, since the CRP variants with a positive charge at either position are not expressed in the absence of DnaJ. However, in my hands it appears DnaJ is required for *crp* transcription, and by some unknown mechanism, loss of either the K100 or K101 positive charge bypasses this requirement.

Since I was unable to determine the role of the K100 and K101 positive charges on the potential DnaJ-CRP interaction, I tried to determine the role of the K100 and K101 positive charges on the interaction between CRP and the protease responsible for CRP degradation. I hypothesized that the increase in CRP stability in mutants lacking the K100 and K101 positive charges may be due to a disruption in the interaction between CRP and this protease.

The protease that degrades CRP has not been identified, so I first attempted to identify this protease. To identify the protease responsible for CRP degradation, I set up a screen to identify mutants that were deficient in CRP degradation. I started with a Δ*crp* Δ*dnaJ* double mutant strain (AJW5619), since DnaJ is required for CRP stability (204). I then generated triple mutants by introducing deletion alleles of several genes encoding subunits of different

proteases. I also introduced deletion alleles of some genes encoding other chaperone proteins to broaden my search for proteins responsible for CRP degradation. Finally, I introduced pDCRP encoding WT CRP into each of the triple mutant strains. I used a Δ*crp* Δ*dnaJ* double mutant carrying pDCRP as a control. All of these strains will express CRP, but due to the instability of CRP it will be undetectable by Western immunoblot analysis. In the strains carrying a deletion allele for a gene involved in CRP degradation, CRP should become more stable and once again detectable by Western immunoblot analysis.

To identify proteases or chaperones required for CRP degradation, I grew the above triple mutants carrying pDCRP in TB7 supplemented with 22 mM glucose until entry into stationary phase. These strains were grown at 28°C due to the temperature sensitivity of many of the strains. I collected samples from each culture and performed Western immunoblot analysis using an anti-CRP antibody. There was no detectable CRP in the Δ*crp* Δ*dnaJ* double mutant carrying pDCRP as expected (**Fig. 21A**). There was also no detectable CRP in the strain lacking the Lon protease, suggesting Lon is not involved in CRP degradation. However, detection of CRP was restored in each of the other triple mutants. These results suggest one of two possibilities: 1) Each one of these proteases and chaperones are required to degrade CRP (which is very unlikely), or 2) A suppressor mutation arose in these strains, allowing CRP to be stable in the absence of DnaJ. Due to the temperature-sensitive nature of many of the strains, I decided to change the strategy for the screen. This time I started with a Δ*dnaJ* single mutant, introduced each of the deletion alleles for proteases and chaperones, and analyzed endogenous CRP levels. The goal was to eliminate the need to maintain the pDCRP plasmid and overexpress CRP, thereby reducing the burden on the cells and decreasing the likelihood of

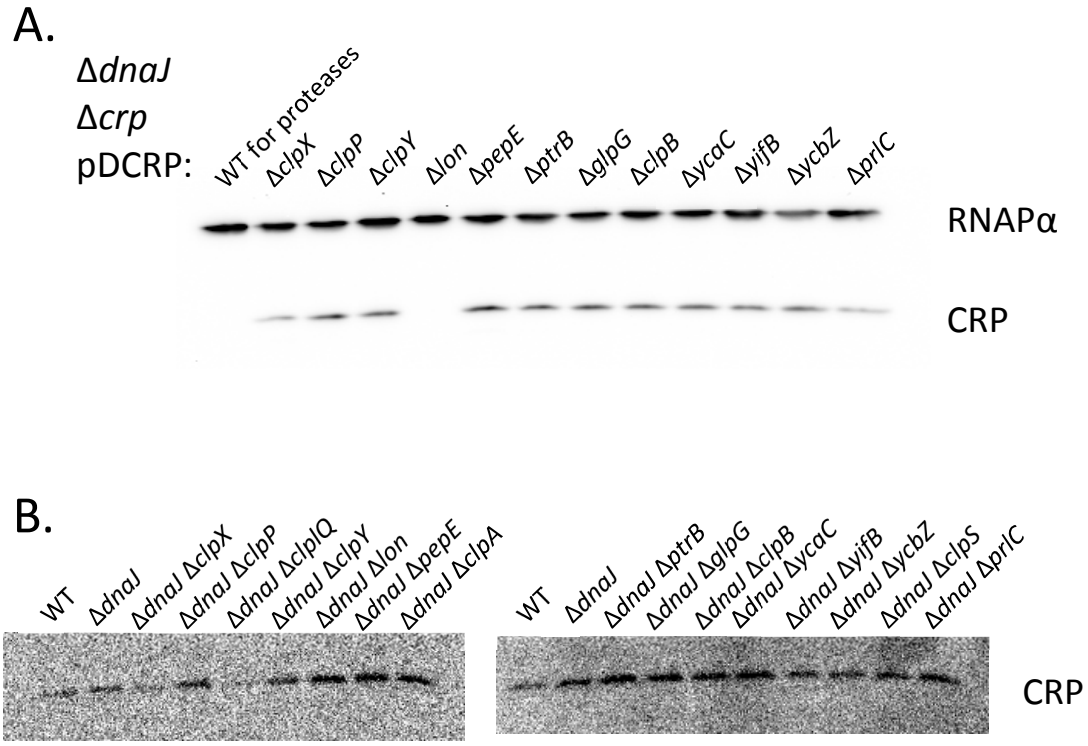

Figure 21. Determination of proteins involved in CRP degradation. (A) Δ*dnaJ* Δ*crp* strains carrying pDCRP and lacking the indicated genes encoding various proteases and chaperones were grown in TB7 supplemented with 22mM glucose at 30°C until entry into stationary phase. The cultures were harvested and subjected to Western immunoblot analysis using an anti-CRP antibody. The experiment was performed once in singlet. (B) Same as (A), except Δ*dnaJ* strains expressing endogenous *crp* and lacking the indicated genes were used. The experiment was performed once in singlet.

suppressor mutations. Western immunoblot analysis of these new strains again showed that CRP was detectable at varying levels in every strain, this time including the Δ*dnaJ* single mutant negative control (**Fig. 21B**). These results suggest another suppressor mutation arose in the strains, stabilizing CRP. Due to the technical difficulties of the experiment, I abandoned the search for proteases and chaperones involved in CRP degradation.

Summary

Acetylation is a PTM that can alter the function of the modified protein. CRP K100 has been identified as acetylated in numerous studies. Since K100 is adjacent to AR2, a critical region for CRP activity at Class II promoters, I investigated the importance of K100 acetylation on CRP activity. In this Chapter, I provided evidence that the K100 positive charge, which would be neutralized upon acetylation, was required for full CRP activity at an artificial Class II promoter, but not at an artificial Class I promoter (**Fig. 11 and Fig. 12**). The K100 positive charge promoted Class II promoter activity even in the absence of a functional AR2 (**Fig. 13**), suggesting K100 makes an independent contact with RNAP. Modeling the interaction between CRP and RNAP at a Class II promoter supported this hypothesis, highlighting RNAP α E163 as a potential interaction partner for CRP K100 (**Fig. 14**).

Intriguingly, loss of the K100 positive charge also increased CRP steady state levels (**Fig. 10**). This was not due to an increase in *crp* transcription (**Fig. 18**), but likely due to an increase in CRP stability (**Fig. 19**). The increase in CRP steady state levels correlated with an increase in both transcription from the Class I *flhDC* promoter (**Fig. 15A**) and flagellar motility (**Fig. 15B**). I hypothesized the K100 positive charge was important for an interaction between CRP and an

unknown protease responsible for CRP degradation, but I was unsuccessful in identifying the unknown protease.

Taken together, these results suggest a model in which the acetylation of CRP K100 directly reduces Class II promoter activity by disrupting the contact between CRP and RNAP, while simultaneously indirectly increasing Class I promoter activity by increasing CRP steady state levels and the occupancy of CRP binding sites. K100 acetylation could be a way for the cell to selectively enhance Class I promoter activity at the expense of Class II promoter activity under conditions in which acP levels are elevated.

CHAPTER FIVE

DISCUSSION

CRP Regulation of Acetylation

In Chapter Three, I provided evidence that CRP positively regulates both YfiQ-mediated and acP-mediated acetylation. Since both mechanisms of acetylation use metabolic intermediates as the acetyl group donors (acCoA and acP, respectively), it is not surprising that CRP, a major regulator of central metabolism, is involved in their regulation. However, the mechanisms by which CRP regulates each method of acetylation are very different. CRP appears to directly promote the transcription of *yfiQ* (**Fig. 8**) (151). In contrast, the role of CRP in acP-mediated acetylation is likely indirect, in part by increasing the rate of carbohydrate uptake, which promotes acP synthesis (134). Both mechanisms of acetylation could be used by the cell to monitor its energy status, but the two different mechanisms of CRP regulation suggest that YfiQ-mediated and acP-mediated acetylation monitor different aspects of the cell's energy status.

YfiQ-mediated acetylation requires acCoA as the acetyl group donor. As a result, YfiQ activity will be high when the cell is consuming carbon and producing significant amounts of acCoA. Maximal *yfiQ* transcription occurs: 1) when the carbon source being consumed is non-catabolite repressing, or 2) just after the catabolite repressing carbon source is depleted (151). This pattern of transcription is consistent with regulation by CRP. I hypothesize that YfiQ-

mediated acetylation senses when the cell's energy status is high under non-catabolite repression conditions. In support of this hypothesis, the acetate scavenging enzyme Acs is inactivated by YfiQ-mediated acetylation (154). Expression of *acs* is induced under non-catabolite repressing conditions by CRP (187). When acCoA levels are high, it is not necessary to scavenge acetate, so Acs is inactivated by YfiQ. Once acCoA levels drop, YfiQ-mediated acetylation decreases and CobB-mediated deacetylation restores Acs activity, enabling the conversion of acetate to acCoA (154).

Much like YfiQ-mediated acetylation, acP-mediated acetylation occurs when the cell is consuming carbon. AcP is generated when the carbon flux through glycolysis is too high, such as during growth on non-limiting concentrations of glucose. I hypothesize that acP-mediated acetylation senses when carbon flux is too high. However, growth on glucose causes catabolite repression, limiting CRP activity. This means CRP must act to promote acetylation prior to acP accumulation. CRP promotes the expression of *ptsG*, which encodes the major glucose transporter PtsG (also called $EIICB^{(Glc)}$) (188). PtsG is required for high glucose flux and acP-mediated acetylation (134). Therefore, it is likely CRP promotes acP-mediated acetylation by promoting high carbon flux through expression of *ptsG*. These acetylations could adjust the activities of metabolic enzymes to accommodate the high carbon flux. Many central metabolic enzymes are acetylated on multiple lysines in an acP-dependent manner, suggesting acetylation may impact their function (121).

Use of K100 Acetyllysine Mimics

The focus of Chapter Four is identifying potential effects of CRP K100 acetylation on CRP function and stability. One of the major difficulties in studying K100 acetylation was my inability to specifically acetylate K100. Until recently, it was thought acetylation required a KAT, so the only acetylated lysines studied were those that could be acetylated by YfiQ, the only known KAT in *E. coli*. This allowed fairly specific targeting of the lysine of interest, especially *in vitro*, simply by manipulating the presence or absence of YfiQ. In contrast with YfiQ-mediated acetylation, K100 appears to be non-enzymatically acetylated by acP (**Fig. 16, Table 4**) (121). Three issues complicated the specific acP-dependent acetylation of K100: 1) CRP K26 is also acetylated by acP (**Table 4**), making it difficult to attribute any acP-dependent alterations in CRP activity specifically to K100 acetylation; 2) At least 1309 lysines proteome-wide are acetylated in an acP-dependent manner (compared with 67 lysines for YfiQ) (121), making it impossible to specifically target CRP K100 by manipulation of acP-levels; 3) AcP also phosphorylates proteins, most notably response regulators (90), meaning any acP-dependent effects cannot necessarily even be attributed to acetylation.

In an attempt to avoid the issues described above, my work made use of mutations that mimic either an unacetylated lysine (K100R) or an acetylated lysine (K100Q). These mutations have been used extensively to study lysine acetylation because they allow the manipulation of a single lysine, are simple to implement, and retain the charge status of their mimicked counterparts (206). However, it is important to recognize that these substitutions are not lysine or acetyllysine and cannot be treated as perfect mimics. For example, the acetyllysine side chain is considerably longer than its glutamate counterpart, thus different interactions may be

made between an interacting protein and the acetylated variant. In fact, there is evidence that substitution of lysine with glutamate may overestimate the influence of acetylation *in vivo* (207).

A better, but considerably more complicated, method of studying K100 acetylation would be the use of the Schultz method to genetically encode the incorporation of acetyllysine at a specific site (208, 209). While the Wolfe lab has access to this powerful tool, it has yet to be optimized for use. This method makes use of a tRNA/tRNA synthetase pair from *Methanosarcina barkeri*. The tRNA recognizes the UAG amber codon, which is normally reserved as a signal to stop protein synthesis in *E. coli*. In *M. barkeri*, the tRNA synthetase covalently links pyrrolysine to the tRNA, and the pyrrolysine is then incorporated into nascent proteins when the amber codon is encountered by a ribosome. In the laboratory, this tRNA synthetase has been evolved many times to recognize and incorporate over 30 unnatural amino acids onto its cognate tRNA, including acetyllysine (209). The Schultz method is particularly useful for studying acetylated proteins *in vitro*, since one can enrich the acetylation of a specific lysine before purification. The major drawback of the Schultz method is the lack of a proper control for *in vivo* studies, i.e. a tRNA/tRNA synthetase pair that incorporates unacetylated lysine when it encounters the UAG amber codon.

Function of K100 Acetylation

I provided strong evidence that neutralization of the K100 positive charge has two distinct functions: 1) to decrease transcription from some CRP-dependent Class II promoters (**Fig. 11**), and 2) to increase the steady state levels of CRP (**Fig. 10**). A consequence of the

second function is an increase in transcription from some CRP-dependent Class I promoters. Together, these two functions provide a mechanism for the cell to inversely regulate Class II and Class I promoters.

The transcriptomics analyses largely support the hypothesis that neutralization of the K100 positive charge decreases activity from some Class II promoters, while increasing activity from some Class I promoters (**Table 2**). Not all conditions tested follow this pattern however, and even under conditions that do follow the pattern there were promoters that behaved opposite of expectations. Several possibilities exist for why some promoters might not behave as expected. First, many CRP-dependent promoters are also regulated by other TFs. In some cases, the activities of these other TFs may overshadow the change in CRP activity due to K100 neutralization. Second, the positioning of CRP and RNAP may not be ideal for contact between K100 and the RNAP alpha subunit at some Class II promoters. The position of the CRP binding site can deviate slightly from the ideal -41.5 position and still function as a Class II promoter (210, 211). This non-ideal spacing may prevent K100 from contacting the RNAP alpha subunit, reducing the significance of K100 neutralization. Third, CRP can act as a repressor at some promoters. This can occur by one of two mechanisms: 1) making too strong of an interaction with RNAP, overstabilizing RNAP and preventing it from escaping the promoter, or 2) steric hindrance. At promoters in which CRP overstabilizes RNAP, disruption of the CRP-RNAP interaction by K100 neutralization would increase activity from Class II promoters, and at Class I promoters, the increased occupancy of CRP binding sites would decrease activity from these promoters. At promoters in which CRP inhibits transcription through steric hindrance, K100

neutralization would increase the occupancy of these inhibitory binding sites, further increasing CRP-dependent transcription inhibition.

DAVID analysis of genes affected by K100 neutralization identified significantly enriched pathways from cultures grown in the presence of acetate (**Table 3**). The pathways enriched were dependent on growth phase. Flagellar motility, which was identified in exponential phase cultures, will be discussed in the next section (see Chapter Five – K100 Regulates Flagellar-Based Motility). In stationary phase, two examples of nearly complete biosynthetic pathways were found upregulated in the K100Q mutant, the methionine/S-adenosyl-L-methionine (SAM) and arginine synthesis pathways. Upregulation of the allosteric transcriptional inhibitor of each pathway suggests these pathways could be non-functional under these conditions. The expression of non-functional pathways indicates one of two possibilities: 1) K100 neutralization results in a need for the products of these pathways, but the cell lacks the appropriate precursors, or 2) K100 neutralization results in the aberrant expression of these pathways, wasting precious resources. These possibilities can be tested by growing the K100Q mutant in the presence and absence of precursors for each pathway. If the cell needs the products of the pathways for optimum growth, then the addition of the precursors should promote growth. If instead the expression of these pathways is aberrant, then the addition of precursors would not be beneficial and in fact may hinder growth due to the synthesis of unnecessary products. Either way, the results would support the hypothesis that K100 neutralization impacts cell physiology.

DAVID analysis did not identify any significantly enriched pathways from cultures grown in the presence of glucose (**Table 3**). This was somewhat surprising, since the number of

differentially regulated genes were comparable between cultures grown in glucose and cultures grown in acetate (**Table 2**). This suggests that in the presence of glucose, the genes affected by the K100 charge status do not encompass entire pathways, but instead comprise small sections of many different pathways. This could be the result of catabolite repression, reducing the number of CRP-dependent genes expressed in the presence of glucose. It could also be the result of a difference in the sensitivities of CRP-dependent promoters to K100 neutralization, due to differences in the activities of other TFs between the two growth conditions.

Since K100 acetylation is acP-dependent (**Fig. 16**) (121), the role of K100 acetylation would be most significant under conditions that favor synthesis of acP. Fermentation occurs through overflow metabolism when carbon flux through glycolysis exceeds carbon flux through the TCA cycle (90). Of the potential fermentation products, acetate (and therefore acP) is synthesized when the cell requires energy, since acetate is the only fermentation product whose synthesis produces ATP. In other words, acP is generated when the cell needs energy but there is too much carbon available for the TCA cycle to process. CRP is a global regulator of carbon metabolism, so perhaps one function of K100 acetylation is to sense that the cell needs energy, but is unable to acquire it through the TCA cycle and respiration. As a result, CRP diverts the excess carbon toward acetate fermentation. CRP enhances acP-dependent acetylation in the presence of excess glucose (**Fig. 9**), a consequence of increased acP and acetate synthesis (185). In this way, K100 acetylation could be a positive feedback mechanism: acP synthesis promotes K100 acetylation, which in turn promotes acP synthesis. It is established that acP promotes K100 acetylation (**Fig. 16**) (121), but the hypothesis that K100 acetylation promotes acP synthesis has yet to be tested. To address this hypothesis, the acP concentration could be

measured in strains expressing K100Q or K100R grown in TB7 supplemented with glucose. As a more indirect measure of acP production, the relative levels of global acetylation could be compared between strains. If K100 neutralization promotes acP synthesis, both readouts would be greater in the strain expressing K100Q compared to the strain expressing K100R.

K100 Acetylation and Flagellar Motility

Transcription of *flhDC*, which encodes the master regulator of flagellar synthesis, was increased in strains expressing CRP variants lacking a positive charge at position 100 (K100Q and K100A) relative to strains with a positive charge at position 100 (WT and K100R) (**Fig. 15A, Table 3**) (Davis *et al.* in revision). This increase in *flhDC* transcription was associated with an increase in flagellar motility (**Fig. 15B**). Together, these data strongly suggest flagellar motility should be enhanced under conditions in which acP levels, and therefore K100 acetylation, are high. Contrary to this expectation, however, flagellar motility is significantly repressed in the Δ*ackA* mutant, which accumulates acP (214). This repression comes from the ability of acP to phosphorylate and thus activate the TCST response regulator RcsB. Phosphorylated RcsB directly inhibits *flhDC* transcription (215), while simultaneously promoting expression of genes required for biofilm formation, including those for colanic acid synthesis (216). If acP inhibits flagellar synthesis, what is the significance of the enhancement of flagellar motility by acP-dependent K100 acetylation?

The canonical mechanism of RcsB phosphorylation is through a phosphorelay involving the sensor kinase RcsC and an intermediate histidine phosphotransferase RcsD, which then phosphorylates RcsB (163). RcsC can also act as a phosphatase, draining phosphate groups from

RcsD and RcsB (217). In fact, it appears that RcsC is by default a phosphatase, and switches to a kinase upon receiving some environmental signal. The exact nature of the environmental signal is not known, but many conditions that activate the Rcs pathway and RcsB phosphorylation involve perturbations of the cell membrane, including osmotic shock, desiccation, and growth on a solid surface (218, 219). Activation of RcsB promotes the transition from attached cells to development of biofilm architecture; phosphorylated RcsB inhibits flagellar synthesis, motility, and the initial attachment to a surface (219), but is required for late-stage biofilm development and maturation (220).

In contrast to RcsB, CRP activates flagellar synthesis and motility (203). While it seems counterintuitive, motility is required for biofilm formation (221). It is speculated that motility is required for the initial attachment to a surface by allowing the cells to overcome repulsive forces at the surface-medium interface, as well as allowing cells to spread out on the surface. Thus, it makes sense that CRP is required during the early stages of biofilm formation (222).

One condition in which acP-dependent enhancement of motility may play a role is during biofilm formation in the gut. A combination of high glucose concentrations (0.4-24 mM) (223) and little oxygen (224) in the gut would promote significant acetate fermentation and acP synthesis. I envision a mechanism where acP contributes to biofilm formation in two steps. Prior to attachment, acP acetylates CRP K100, enhancing flagellar synthesis and motility to bring cells closer to the surface and to allow the cells to spread out across the surface. At the same time, acP phosphorylates RcsB, but the phosphatase activity of RcsC keeps RcsB phosphorylation low. Once cells are attached to the surface, RcsC switches from a phosphatase to a kinase, no longer antagonizing RcsB phosphorylation but instead enhancing it.

Phosphorylated RcsB now shuts off flagellar synthesis and promotes transcription of genes required for biofilm maturation.

To test this model, a CRP K100R mutant could be used to determine if enough flagella could be synthesized prior to RcsB activation to generate a successful biofilm without the enhancement of *flhDC* transcription by K100 acetylation. Additionally, it would be necessary to examine the kinetics of acP-dependent acetylation of CRP K100, phosphorylation of RcsB by acP and by RcsC, and dephosphorylation of RcsB by RcsC to determine if sufficient amounts of acetylated CRP can accumulate to activate flagellar synthesis before sufficient amounts of phosphorylated RcsB can accumulate to inhibit flagellar synthesis. This experiment would be more difficult to interpret, since it is unclear what "sufficient amounts" of either K100 acetylation or RcsB phosphorylation means in this context.

Stoichiometry of K100 Acetylation

The significance of K100 acetylation depends strongly on the stoichiometry of K100 acetylation. This stoichiometry is defined as the ratio of CRP molecules acetylated on K100 relative to total CRP molecules (137). K100 acetylation increases 7.4-fold in the presence of glucose relative to its absence (134) and 3.5-fold in a Δ*ackA* mutant that accumulates acP relative to its isogenic parent (121). However, since we do not know the initial or final stoichiometry, this information is not particularly useful in determining physiological relevance: in the case of acP accumulation, an increase from 0.1% to 0.35% K100 acetylation is much different than an increase from 20% to 70% K100 acetylation. The stoichiometry of lysine

acetylation on a global scale has been measured in *E. coli* (133, 137, 166), though none of these studies detected CRP K100 acetylation specifically.

While none of the work in this dissertation directly addressed the issue of K100 acetylation stoichiometry, there are patterns in the DNA microarray analysis that suggest a relatively high level of K100 acetylation in cultures grown in glucose as the sole carbon source and a relatively low level of K100 acetylation in cultures grown in acetate as the sole carbon source. In stationary phase cultures grown with glucose as the sole carbon source, the number of differentially regulated genes/promoters in the comparison between the N reference strain and K100Q was much smaller than in the comparison between the N and K100R strains (**Table 2**). This indicates that the N strain is transcriptionally more similar to K100Q than K100R. The only difference between the three strains is the status of K100, and since the K100Q mutation mimics an acetylated K100, the data may be interpreted to mean that, under these conditions, there may be more CRP molecules acetylated at K100 than unacetylated in the N strain. Conversely, in stationary phase cultures grown with acetate as the sole carbon source, the number of differentially regulated genes/promoters in the comparison between the N and K100R strains was smaller than in the comparison between the N and K100Q strains. These data support the hypothesis that the N strain is transcriptionally more similar to K100R under these conditions, and that there may be more CRP molecules acetylated on K100 than not acetylated on K100.

For cultures grown in glucose, the fact that the N reference strain behaves more like the K100Q strain than the K100R strain is not surprising. We know that glucose increases acetylation both globally (120, 121, 134) and specifically on K100 (121), and that acetylation

levels accumulate throughout growth, peaking in stationary phase (121). Therefore, we would expect the behavior of the N strain to gradually become more similar to the K100Q strain over time.

For cultures grown in acetate, the expectations were not as obvious. Acetate was originally chosen as a carbon source because it is not acetogenic, so no acP should be produced by way of acetate formation and K100 acetylation should be minimal. However, acetate taken up through the Pta-AckA pathway also generates acP (**Fig. 6**). At high (> 10 mM) concentrations of acetate, cells preferential reassimilate acetate by means of the Pta-AckA pathway over the Acs pathway (92, 93). The 30 mM acetate used in the DNA microarray experiment would be sufficient to produce acP, so we might expect high levels of acetylation, contrary to the original goal of using acetate to minimize acetylation. In fact, global acetylation is increased in cultures grown on 40 mM acetate as the sole carbon source (120).

Despite arguments for high K100 acetylation in acetate as the sole carbon source, the DNA microarray data suggest the stoichiometry of K100 acetylation is low under this condition. There are two possible reasons for the discrepancy. First, the degree of similarity between the N reference strain and either of the K100 mutant strains may not be a good indicator of the stoichiometry of K100 acetylation. The K100Q and K100R mutations may behave differently than acetylated and unacetylated lysines, and the transcriptional similarities between these strains and N may be coincidental. Second, despite high global acetylation in the presence of high concentrations of acetate, there may be relatively little K100 acetylation due to unknown KDAC activity that specifically targets K100 when acetate is present. This possibility would

suggest the role of K100 acetylation is dependent on more than simply the concentration of acP, but also the context under which that acP is formed.

It would be interesting to repeat the DNA microarray analysis using different carbon sources that induce more acP synthesis than glucose or less acP synthesis than acetate to determine if there is a correlation between acP synthesis and the degree to which the N reference strain is similar to either the K100Q or K100R strains. The amount of acetate a cell produces is correlated with the oxidation state of the carbon source (225). A highly oxidized carbon source does not require the reduction of NAD^+ to NADH to enter glycolysis. In fact, some carbon sources such as glucuronic acid must oxidize NADH to NAD^+ prior to glycolysis, balancing the reduction of NAD^+ during glycolysis. This means more carbon can be diverted to synthesizing ATP-generating acetate instead of synthesizing redox-balancing ethanol, lactate, or succinate. On the other hand, more reduced carbon sources such as sorbitol or glycerol require NAD^+ reduction prior to entering glycolysis, further disrupting the redox balance of the cell. This means more of the carbon must be used to generate the redox-balancing fermentation products, leaving less carbon for acetate synthesis (90).

In contrast with the DNA microarray data, expression from the Class II *CC(-41.5)* promoter in the presence of glucose suggests that the stoichiometry of K100 acetylation was very low. The promoter activity in the strain expressing WT CRP was indistinguishable from the promoter activity in the strain expressing K100R (**Fig. 11**), indicating the level of K100 acetylation was not sufficient to reduce the Class II promoter activity of WT CRP. A possible explanation for this discrepancy is the difference in growth media between the two experiments. The DNA microarray experiments used a minimal defined media supplemented

with glucose. When glucose is the sole carbon source, cells begin consuming the glucose immediately. Glucose consumption correlates with an increase in global acetylation, including K100 acetylation (134). By the time all the glucose was consumed (i.e. in stationary phase), a significant amount of K100 acetylation may have accumulated. Unlike the DNA microarray experiments, the promoter activity assays used TB7, a buffered tryptone broth composed primarily of peptide fragments, supplemented with glucose. When presented with both tryptone and glucose, cells consume the tryptone first, and do not begin consuming glucose until stationary phase (212). As a result, global acetylation does not accumulate significantly until stationary phase (121, 134). Since the promoter assay experiments were ended just as the cultures were entering stationary phase, it is possible that not enough K100 acetylation had yet accumulated to significantly affect the Class II promoter activity of WT CRP. Thus, extending the promoter activity assays past the point where the cultures enter stationary phase, allowing K100 acetylation to accumulate, may reduce the Class II promoter activity of WT CRP to the point where its activity begins to resemble the activity of the K100Q mutant.

Ultimately, the best way to determine the stoichiometry of K100 acetylation under any condition would be to measure it directly using previously established methods of mass spectrometry (133, 137, 166). These methods address an issue with traditional mass spectrometry, in which post-translationally modified peptides have different ionization efficiencies than their unmodified counterparts (137). This means values obtained for the modified and unmodified peptides cannot be directly compared, so stoichiometry cannot be determined. To overcome this issue, cell lysates containing some level of lysine acetylation are first incubated with labeled acetic anhydride, which chemically acetylates any unacetylated

lysines. Since there are no unmodified peptides, there is no discrepancy between modified and unmodified peptides. Mass spectrometry allows peptides acetylated by acetic anhydride to be distinguished from peptides acetylated endogenously based on a shift in mass equal to the mass of the acetic anhydride label, which varies depending on the specific method used. The abundance of unlabeled and labeled peptides can now be compared, and a stoichiometry value is obtained. It may be beneficial to first purify CRP from the cell lysates using anti-CRP antibodies or Ni^{+}-affinity resin (173) to enrich signals coming from CRP peptides over peptides from the rest of the proteome.

Acetylation of CRP K26

In addition to CRP K100, K26 was also identified as acetylated *in vitro* in an acP-dependent manner (**Table 4**). K26 has been identified as acetylated in previous *in vivo* studies (120, 121, 133), though its acetylation did not appear to require acP (121). The discrepancy between the *in vitro* and *in vivo* results regarding the acP-dependence of K26 acetylation could be due to one of three possibilities: 1) K26 is not acetylated by acP *in vivo*, and the *in vitro* results are merely an artifact of an artificial system; 2) Technical issues may have incorrectly identified K26 acetylation as acP-dependent *in vitro* or acP-independent *in vivo*; or 3) K26 acetylation is acP-dependent, but KDAC activity *in vivo* keeps K26 acetylation low in the presence of acP. This last possibility is especially intriguing, because it would indicate that while both K100 and K26 are non-enzymatically acetylated by acP, their acetylation status is differentially regulated by deacetylation. However, no evidence for such a KDAC presently exists.

K26 is located in the DNA binding region of CRP. While K26 is not involved in the recognition of the CRP consensus sequence, it does make non-specific contacts with negatively charged phosphate groups in the DNA (59, 226). These contacts contribute to the bending of DNA around CRP. Acetylation of K26, which would neutralize the lysine's positive charge, might weaken the interaction between CRP and the DNA. This would likely not have an effect on the recruitment of CRP to the promoter or on the interactions between CRP and RNAP, but it may affect transcription at CRP-dependent promoters by reducing the DNA bend angle around CRP, preventing distant transcriptional regulators from coming into close contact with RNAP. This hypothesis could be tested by measuring the DNA bend angle *in vitro* (61) using purified CRP that is specifically acetylated or unacetylated at K26 using the Schultz method (209).

CRP Stability

I provided evidence that both the K100 and K101 positive charges are important for proper stability of CRP (**Fig. 19A**). Very little is known concerning the factors that control the stability of CRP. The only reported regulator of CRP stability is the heat-shock chaperone protein DnaJ; in the absence of DnaJ, CRP stability decreases from >80 minutes to 3 minutes (204). However, the mechanism for DnaJ-mediated stability of CRP has not yet been established.

DnaJ is a component of the DnaK/DnaJ/GrpE chaperone complex, responsible for protein folding *de novo*, refolding, disaggregation, and secretion (227). DnaJ binds stretches of approximately eight amino acids made up of hydrophobic (I or L), aromatic (F, W, or Y), or arginine (R) residues within peptides (228). DnaJ then delivers these peptides to DnaK, and

promotes the ATPase activity of DnaK to drive the unfolding and refolding of the target peptide. While DnaJ promotes CRP stability, DnaK is not required, suggesting DnaJ may work with multiple chaperone complexes (204).

One possibility is that loss of either the K100 or K101 positive charge enhances the interaction between CRP and DnaJ, allowing DnaJ to better support CRP stabilization. There are a significant number hydrophobic, aromatic, and arginine residues surrounding K100 and K101 (90-TACEVAE**IS****Y**KK**FR**Q**LI**QVNP-110), suggesting this region could be a recognition site for DnaJ. While arginines are more likely to be found within a DnaJ binding region, lysines, alanines, and glutamates are less likely to be found within a DnaJ binding region, and glutamines are found in DnaJ binding regions as frequently as in non-DnaJ binding regions (228). This indicates arginines likely favor DnaJ binding, while lysines, alanines, and glutamates likely disfavor DnaJ binding. However, this pattern of DnaJ binding is not consistent with the patterns of CRP stability observed when K100 or K101 is mutated to alanine, arginine, glutamate, or glutamine (**Fig. 10, Fig. 19A**). These results argue that while DnaJ may interact with the CRP region containing K100 and K101, this interaction is probably not responsible for the observed variation in CRP stability.

Why were some of the CRP variants not transcribed in the Δ*dnaJ* Δ*crp* pDCRP strains (**Fig. 20**)? Even though DnaJ stabilizes CRP, and CRP regulates its own transcription, loss of *crp* increases transcription from a *crp-lacZ* promoter fusion (66). Additionally, previous work showed that WT *crp* is still transcribed in a Δ*dnaJ* mutant (205), suggesting there was a problem with my experiment. Mutations within *dnaJ* prevent growth at temperatures above 42°C and cause slowed growth at temperatures above 30°C (229). These mutants are also genetically unstable, and suppressor mutations frequently arise that allow faster growth at temperatures

above 30°C. Generation of the Δ*dnaJ* Δ*crp* and Δ*dnaJ* Δ*crp* pDCRP strains was performed at 37°C, so it is quite possible that suppressor mutations that prevented the expression of *crp* arose sometime during the generation of these strains. To test this hypothesis, pDCRP could be purified and sequenced from each of the Δ*dnaJ* Δ*crp* pDCRP strains to determine if the sequences have been altered. Additionally, each of the Δ*dnaJ* Δ*crp* pDCRP strains could be regenerated at a lower temperature to reduce the likelihood of suppressor mutations.

CRP Degradation

Changes in a CRP-DnaJ interaction are unlikely to explain the difference in stability between the CRP K100 and K101 mutants. Another possibility is that the K100/K101 region is recognized by a protease, either as a binding site or as a cleavage site, and the CRP-protease interaction is disrupted by loss of either the K100 or K101 positive charge. There are no reports indicating which protease or proteases are responsible for CRP degradation, but the fact that CRP variants containing positively charged residues at positions 100 or 101 are less stable than variants lacking a positive charge in these positions suggests a trypsin-like protease may be involved. Trypsin and trypsin-like proteases cleave peptides after positively charged lysine and arginine residues, unless they are followed by a proline (230). Post-translational modifications that alter the positive charge of these residues, including lysine acetylation, also prevent cleavage by trypsin and trypsin-like proteases (231). Acetylation of K100 or K101 could reduce CRP degradation by a trypsin-like protease, enhancing the stability of CRP.

PtrB (also called Protease II or oligopeptidase B) and PrlC (also called Protease In or oligopeptidase A) are the two known trypsin-like proteases in *E. coli* (232, 233). Each protease

cleaves after lysine and arginine residues, though neither appears to be structurally related to trypsin. Interestingly, PrlC does not efficiently cleave sites containing consecutive positively charged residues (KK or RR) (234), while PtrB cleaves sites with consecutive positively charged residues more efficiently than sites with only a single positively charged residue (235). Mutation of either K100 or K101 to a non-positively charged residue increased the stability of the CRP variant, suggesting PtrB may be involved in CRP degradation. This hypothesis could be tested by performing an *in vitro* proteolytic assay using purified PtrB and each of the CRP variants to determine if PtrB can cleave CRP, and if the charge status of K100 or K101 contributes to the efficiency of this cleavage.

Concluding Remarks

Transcriptional regulation is a complex process that integrates signals from both the environment and the intracellular space to ensure genes are expressed at the right time and in the right amounts. The work presented here provides evidence that CRP is involved in the production of one of these signals, Nε-lysine acetylation. This acetylation is a highly abundant post-translational modification in *E. coli*, although we are only beginning to appreciate its significance. Additional evidence supports the hypothesis that CRP itself is regulated by acetylation. Neutralization of K100 decreases CRP activity at Class II promoters, while indirectly increasing CRP activity at Class I promoters by increasing the CRP half-life. Transcriptomics analyses revealed K100 neutralization has wide-reaching effects on global transcription, affecting both CRP-dependent and CRP-independent genes. The work presented here highlights the impact a single acetylation could have on global gene expression. Dissecting the impact of

acetylation on the function of other transcriptional regulators will be crucial in understanding how the cell responds to changes in its environment.

REFERENCE LIST

1. Browning, D. F., and Busby, S. J. (2004) The regulation of bacterial transcription initiation. *Nat. Rev. Microbiol.* **2**, 57–65

2. Borukhov, S., and Nudler, E. (2008) RNA polymerase: the vehicle of transcription. *Trends Microbiol* **16**, 126–134

3. Revyakin, A., Liu, C., Ebright, R. H., and Strick, T. R. (2006) Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. *Science* **314**, 1139–43

4. Wösten, M. M. (1998) Eubacterial sigma-factors. *FEMS Microbiol. Rev.* **22**, 127–50

5. Borukhov, S., and Severinov, K. (2002) Role of the RNA polymerase sigma subunit in transcription initiation. *Res. Microbiol.* **153**, 557–62

6. Merrick, M. J. (1993) In a class of its own--the RNA polymerase sigma factor sigma 54 (sigma N). *Mol. Microbiol.* **10**, 903–9

7. Gruber, T., and Gross, C. (2003) Multiple Sigma Subunits and the Partitioning of Bacterial Transcription Space. *Annual Review of Microbiology* **57**, 441–466

8. Wade, J., Roa, D., Grainger, D., Hurd, D., Busby, S., Struhl, K., and Nudler, E. (2006) Extensive functional overlap between σ factors in Escherichia coli. *Nat Struct Mol Biology* **13**, 806–814

9. Barne, K. A., Bown, J. A., Busby, S. J., and Minchin, S. D. (1997) Region 2.5 of the Escherichia coli RNA polymerase 70 subunit is responsible for the recognition of the ‘extended- 10’motif at promoters. *The EMBO journal* **16**, 4034–4040

10. Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O., and Darst, S. A. (2002) Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* **296**, 1285–90

11. Gourse, R., Ross, W., and Gaal, T. (2000) UPs and downs in bacterial transcription initiation: the role of the alpha subunit of RNA polymerase in promoter recognition. *Mol Microbiol* **37**, 687–695

12. Keseler, I., Mackie, A., Peralta-Gil, M., Santos-Zavaleta, A., Gama-Castro, S., Bonavides-Martínez, C., Fulcher, C., Huerta, A., Kothari, A., Krummenacker, M., Latendresse, M., Muñiz-Rascado, L., Ong, Q., Paley, S., Schröder, I., Shearer, A., Subhraveti, P., Travers, M., Weerasinghe, D., Weiss, V., Collado-Vides, J., Gunsalus, R., Paulsen, I., and Karp, P. (2013) EcoCyc: fusing model organism databases with systems biology. *Nucleic Acids Res* **41**, D605–D612

13. Pérez-Rueda, E., and Collado-Vides, J. (2000) The repertoire of DNA-binding transcriptional regulators in Escherichia coli K-12. *Nucleic Acids Res.* **28**, 1838–47

14. Gilbert, W., Gralla, J., Majors, J., and Maxam, A. (1975) in *Protein-Ligand Interactions* (Sund, H., and Blauer, G., eds.) pp. 193–206, de Gruyter, Berlin

15. Semsey, S., Geanacopoulos, M., Lewis, D., and Adhya, S. (2002) Operator-bound GalR dimers close DNA loops by direct interaction: tetramerization and inducer binding. *Embo J* **21**, 4349–4356

16. Oehler, S., Eismann, E. R., Krämer, H., and Müller-Hill, B. (1990) The three operators of the lac operon cooperate in repression. *EMBO J.* **9**, 973–9

17. Shin, M., Kang, S., Hyun, S. J., Fujita, N., Ishihama, A., Valentin-Hansen, P., and Choy, H. E. (2001) Repression of deoP2 in Escherichia coli by CytR: conversion of a transcription activator into a repressor. *EMBO J.* **20**, 5392–9

18. Gaston, K., Bell, A., Kolb, A., Buc, H., and Busby, S. (1990) Stringent spacing requirements for transcription activation by CRP. *Cell* **62**, 733–743

19. Browning, D. F., and Busby, S. J. (2016) Local and global regulation of transcription initiation in bacteria. *Nat. Rev. Microbiol.* **14**, 638–50

20. Sakumi, K., Igarashi, K., Sekiguchi, M., and Ishihama, A. (1993) The Ada protein is a class I transcription factor of Escherichia coli. *J. Bacteriol.* **175**, 2455–7

21. Igarashi, K., Hanamura, A., Makino, K., Aiba, H., Aiba, H., Mizuno, T., Nakata, A., and Ishihama, A. (1991) Functional map of the alpha subunit of Escherichia coli RNA

polymerase: two modes of transcription activation by positive factors. *Proc National Acad Sci* **88**, 8958–8962

22. Tao, K., Fujita, N., and Ishihama, A. (1993) Involvement of the RNA polymerase α subunit C-terminal region in co-operative interaction and transcriptional activation with OxyR protein. *Mol Microbiol* **7**, 859–864

23. Igarashi, K., and Ishihama, A. (1991) Bipartite functional map of the E. coli RNA polymerase alpha subunit: involvement of the C-terminal region in transcription activation by cAMP-CRP. *Cell* **65**, 1015–22

24. Nickels, B., Dove, S., Murakami, K., Darst, S., and Hochschild, A. (2002) Protein–Protein and Protein–DNA Interactions of σ70 Region 4 Involved in Transcription Activation by λcI. *J Mol Biol* **324**, 17–34

25. Niu, W., Kim, Y., Tau, G., Heyduk, T., and Ebright, R. H. (1996) Transcription activation at class II CAP-dependent promoters: two interactions between CAP and RNA polymerase. *Cell* **87**, 1123–34

26. Hidalgo, E., Leautaud, V., and Demple, B. (1998) The redox-regulated SoxR protein acts from a single DNA site as a repressor and an allosteric activator. *EMBO J.* **17**, 2629–36

27. Philips, S., Canalizo-Hernandez, M., Yildirim, I., Schatz, G., Mondragón, A., and O'Halloran, T. (2015) TRANSCRIPTION. Allosteric transcriptional regulation via changes in the overall topology of the core promoter. *Sci New York N Y* **349**, 877–81

28. Madan Babu, M., and Teichmann, S. A. (2003) Evolution of transcription factors and the gene regulatory network in Escherichia coli. *Nucleic Acids Res.* **31**, 1234–44

29. Jobe, A., and Bourgeois, S. (1972) lac Repressor-operator interaction. VI. The natural inducer of the lac operon. *J. Mol. Biol.* **69**, 397–408

30. Lazazzera, B. A., Bates, D. M., and Kiley, P. J. (1993) The activity of the Escherichia coli transcription factor FNR is regulated by a change in oligomeric state. *Genes Dev.* **7**, 1993–2005

31. Webster, C., Gardner, L., and Busby, S. (1989) The Escherichia coli melR gene encodes a DNA-binding protein with affinity for specific sequences located in the melibiose-operon regulatory region. *Gene* **83**, 207–13

32. Marťínez-Antonio, A., and Collado-Vides, J. (2003) Identifying global regulators in transcriptional regulatory networks in bacteria. *Curr Opin Microbiol* **6**, 482–489

33. Majors, J. (1975) Specific binding of CAP factor to lac promoter DNA. *Nature* **256**, 672–4

34. Hufnagel, D. A., Evans, M. L., Greene, S. E., Pinkner, J. S., Hultgren, S. J., and Chapman, M. R. (2016) The Catabolite Repressor Protein-Cyclic AMP Complex Regulates csgD and Biofilm Formation in Uropathogenic Escherichia coli. *J. Bacteriol.* **198**, 3329–3334

35. Shimada, T., Yoshida, H., and Ishihama, A. (2013) Involvement of Cyclic AMP Receptor Protein in Regulation of the rmf Gene Encoding the Ribosome Modulation Factor in Escherichia coli. *J Bacteriol* **195**, 2212–2219

36. Uppal, S., Shetty, D., and Jawali, N. (2014) Cyclic AMP Receptor Protein Regulates cspD, a Bacterial Toxin Gene, in Escherichia coli. *J Bacteriol* **196**, 1569–1577

37. Hirakawa, H., Inazumi, Y., Senda, Y., Kobayashi, A., Hirata, T., Nishino, K., and Yamaguchi, A. (2006) N-Acetyl-d-Glucosamine Induces the Expression of Multidrug Exporter Genes, mdtEF, via Catabolite Activation in Escherichia coli. *J Bacteriol* **188**, 5851–5858

38. Bodero, M., and Munson, G. (2009) Cyclic AMP Receptor Protein-Dependent Repression of Heat-Labile Enterotoxin. *Infect Immun* **77**, 791–798

39. Lorenzo, V., Herrero, M., Giovannini, F., and Neilands, J. (1988) Fur (ferric uptake regulation) protein and CAP (catabolite-activator protein) modulate transcription of fur gene in Escherichia coli. *Eur J Biochem* **173**, 537–546

40. Zhao, K., Liu, M., and Burgess, R. (2007) Adaptation in bacterial flagellar and motility systems: from regulon members to “foraging”-like behavior in E. coli. *Nucleic Acids Research* **35**, 4441–4452

41. Nasser, W., Schneider, R., Travers, A., and Muskhelishvili, G. (2001) CRP Modulates fis Transcription by Alternate Formation of Activating and Repressing Nucleoprotein Complexes. *J. Biol. Chem.* **276**, 17878–17886

42. Claret, L., and Rouviere-Yaniv, J. (1996) Regulation of HU alpha and HU beta by CRP and FIS in Escherichia coli. *J. Mol. Biol.* **263**, 126–39

43. Guo, M., Wang, H., Xie, N., and Xie, Z. (2015) Positive Effect of Carbon Sources on Natural Transformation in Escherichia coli: Role of Low-Level Cyclic AMP (cAMP)-cAMP Receptor Protein in the Derepression of rpoS. *J Bacteriol* **197**, 3317–3328

44. Kallipolitis, B. H., and Valentin-Hansen, P. (1998) Transcription of rpoH, encoding the Escherichia coli heat-shock regulator sigma32, is negatively controlled by the cAMP-CRP/CytR nucleoprotein complex. *Mol. Microbiol.* **29**, 1091–9

45. Harman, J. (2001) Allosteric regulation of the cAMP receptor protein. *Biochimica Et Biophysica Acta Bba - Protein Struct Mol Enzym* **1547**, 1–17

46. Won, H. S., Yamazaki, T., Lee, T. W., Yoon, M. K., Park, S. H., Kyogoku, Y., and Lee, B. J. (2000) Structural understanding of the allosteric conformational change of cyclic AMP receptor protein by cyclic AMP binding. *Biochemistry* **39**, 13953–62

47. Sharma, H., Yu, S., Kong, J., Wang, J., and Steitz, T. (2009) Structure of apo-CAP reveals that large conformational changes are necessary for DNA binding. *Proc National Acad Sci* **106**, 16604–16609

48. Kolb, A., Busby, S., Buc, H., Garges, S., and Adhya, S. (1993) Transcriptional regulation by cAMP and its receptor protein. *Annu. Rev. Biochem.* **62**, 749–95

49. Levitt, M. (1978) How many base-pairs per turn does DNA have in solution and in chromatin? Some theoretical calculations. *Proc National Acad Sci* **75**, 640–644

50. Ushida, C., and Aiba, H. (1990) Helical phase dependent action of CRP: effect of the distance between the CRP site and the -35 region on promoter activity. *Nucleic Acids Res.* **18**, 6325–30

51. Niu, W., Zhou, Y., Dong, Q., Ebright, Y. W., and Ebright, R. H. (1994) Characterization of the activating region of Escherichia coli catabolite gene activator protein (CAP). I. Saturation and alanine-scanning mutagenesis. *J. Mol. Biol.* **243**, 595–602

52. Zhou, Y., Zhang, X., and Ebright, R. H. (1993) Identification of the activating region of catabolite gene activator protein (CAP): isolation and characterization of mutants of CAP specifically defective in transcription activation. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6081–5

53. Zhou, Y., Pendergrast, P. S., Bell, A., Williams, R., Busby, S., and Ebright, R. H. (1994) The functional subunit of a dimeric transcription activator protein depends on promoter architecture. *EMBO J.* **13**, 4549–57

54. Savery, N., Lloyd, G., Busby, S., Thomas, M., Ebright, R., and Gourse, R. (2002) Determinants of the C-Terminal Domain of the Escherichia coli RNA Polymerase α Subunit Important for Transcription at Class I Cyclic AMP Receptor Protein-Dependent Promoters. *Journal of Bacteriology* **184**, 2273–2280

55. Savery, N. J., Lloyd, G. S., Kainz, M., Gaal, T., Ross, W., Ebright, R. H., Gourse, R. L., and Busby, S. J. (1998) Transcription activation at Class II CRP-dependent promoters: identification of determinants in the C-terminal domain of the RNA polymerase alpha subunit. *EMBO J.* **17**, 3439–47

56. West, D., Williams, R., Rhodius, V., Bell, A., Sharma, N., Zou, C., Fujita, N., Ishihama, A., and Busby, S. (1993) Interactions between the Escherichia coli cyclic AMP receptor protein and RNA polymerase at class II promoters. *Mol. Microbiol.* **10**, 789–97

57. Rhodius, V. A., and Busby, S. J. (2000) Interactions between activating region 3 of the Escherichia coli cyclic AMP receptor protein and region 4 of the RNA polymerase sigma(70) subunit: application of suppression genetics. *J. Mol. Biol.* **299**, 311–24

58. Rhodius, V. A., and Busby, S. J. (2000) Transcription activation by the Escherichia coli cyclic AMP receptor protein: determinants within activating region 3. *J. Mol. Biol.* **299**, 295–310

59. Schultz, S. C., Shields, G. C., and Steitz, T. A. (1991) Crystal structure of a CAP-DNA complex: the DNA is bent by 90 degrees. *Science* **253**, 1001–7

60. Heyduk, T., and Lee, J. C. (1992) Solution studies on the structure of bent DNA in the cAMP receptor protein-lac DNA complex. *Biochemistry* **31**, 5165–71

61. Kapanidis, AN, Ebright, YW, and Ludescher, RD (2001) Mean DNA bend angle and distribution of DNA bend angles in the CAP-DNA complex in solution. *J Mol. Biol.* **312**, 453-468

62. González-Gil, G., Kahmann, R., and Muskhelishvili, G. (1998) Regulation of crp transcription by oscillation between distinct nucleoprotein complexes. *EMBO J.* **17**, 2877–2885

63. Aiba, H. (1983) Autoregulation of the Escherichia coli crp gene: CRP is a transcriptional repressor for its own gene. *Cell* **32**, 141–9

64. Hanamura, A., and Aiba, H. (1992) A new aspect of transcriptional control of the Escherichia coli crp gene: positive autoregulation. *Mol. Microbiol.* **6**, 2489–97

65. Hanamura, A., and Aiba, H. (1991) Molecular mechanism of negative autoregulation of Escherichia coli crp gene. *Nucleic Acids Res.* **19**, 4413–9

66. Cossart, P., and Gicquel-Sanzey, B. (1985) Regulation of expression of the crp gene of Escherichia coli K-12: in vivo study. *J. Bacteriol.* **161**, 454–7

67. Zhang, Aboulwafa, and Saier, M. H. (2014) Regulation of crp Gene Expression by the Catabolite Repressor/Activator, Cra, in Escherichia coli. *Journal of Molecular Microbiology and Biotechnology* **24**, 135–141

68. Ball, C. A., Osuna, R., Ferguson, K. C., and Johnson, R. C. (1992) Dramatic changes in Fis levels upon nutrient upshift in Escherichia coli. *J. Bacteriol.* **174**, 8043–56

69. MAGASANIK, B. (1961) Catabolite repression. *Cold Spring Harb. Symp. Quant. Biol.* **26**, 249–56

70. Lengeler, J. W. (2015) PTS 50: Past, Present and Future, or Diauxie Revisited. *J. Mol. Microbiol. Biotechnol.* **25**, 79–93

71. Görke, B., and Stülke, J. (2008) Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nat. Rev. Microbiol.* **6**, 613–24

72. Deutscher, J., Francke, C., and Postma, P. W. (2006) How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiol. Mol. Biol. Rev.* **70**, 939–1031

73. Tchieu, J. H., Norris, V., Edwards, J. S., and Saier, M. H. (2001) The complete phosphotransferase system in Escherichia coli. *J. Mol. Microbiol. Biotechnol.* **3**, 329–46

74. Park, Y.-H., Lee, B., Seok, Y.-J., and Peterkofsky, A. (2006) In Vitro Reconstitution of Catabolite Repression in Escherichia coli. *J Biol Chem* **281**, 6448–6454

75. Dills, S. S., Schmidt, M. R., and Saier, M. H. (1982) Regulation of lactose transport by the phosphoenolpyruvate-sugar phosphotransferase system in membrane vesicles of Escherichia coli. *J. Cell. Biochem.* **18**, 239–44

76. Mitchell, W. J., Misko, T. P., and Roseman, S. (1982) Sugar transport by the bacterial phosphotransferase system. Regulation of other transport systems (lactose and melibiose). *J. Biol. Chem.* **257**, 14553–64

77. Dean, D. A., Reizer, J., Nikaido, H., and Saier, M. H. (1990) Regulation of the maltose transport system of Escherichia coli by the glucose-specific enzyme III of the phosphoenolpyruvate-sugar phosphotransferase system. Characterization of inducer exclusion-resistant mutants and reconstitution of inducer exclusion in proteoliposomes. *J. Biol. Chem.* **265**, 21005–10

78. Saier, M. H., and Roseman, S. (1976) Sugar transport. Inducer exclusion and regulation of the melibiose, maltose, glycerol, and lactose transport systems by the phosphoenolpyruvate:sugar phosphotransferase system. *The Journal of biological chemistry* **251**, 6606–15

79. Plumbridge, J. (2001) DNA binding sites for the Mlc and NagC proteins: regulation of nagE, encoding the N-acetylglucosamine-specific transporter in Escherichia coli. *Nucleic Acids Res* **29**, 506–514

80. Plumbridge, J. (1999) Expression of the phosphotransferase system both mediates and is mediated by Mlc regulation in Escherichia coli. *Mol Microbiol* **33**, 260–273

81. Tanaka, Y., Kimata, K., and Aiba, H. (2000) A novel regulatory role of glucose transporter of Escherichia coli: membrane sequestration of a global repressor Mlc. *EMBO J.* **19**, 5344–52

82. Hogema, B., Arents, J., Bader, R., Eijkemans, K., Yoshida, H., Takahashi, H., Aiba, H., and Postma, P. (1998) Inducer exclusion in Escherichia coli by non-PTS substrates: the role of the PEP to pyruvate ratio in determining the phosphorylation state of enzyme IIAGlc. *Molecular Microbiology* **30**, 487498

83. Amarasingham, C. R., and Davis, B. D. (1965) Regulation of alpha-ketoglutarate dehydrogenase formation in Escherichia coli. *J. Biol. Chem.* **240**, 3664–8

84. Gray, C. T., Wimpenny, J. W., and Mossman, M. R. (1966) Regulation of metabolism in facultative bacteria. II. Effects of aerobiosis, anaerobiosis and nutrition on the formation of Krebs cycle enzymes in Escherichia coli. *Biochim. Biophys. Acta* **117**, 33–41

85. Zhang, Z., Gosset, G., Barabote, R., Gonzalez, C., Cuevas, W., and Saier, M. (2005) Functional Interactions between the Carbon and Iron Utilization Regulators, Crp and Fur, in Escherichia coli. *J Bacteriol* **187**, 980–990

86. Park, S. J., Chao, G., and Gunsalus, R. P. (1997) Aerobic regulation of the sucABCD genes of Escherichia coli, which encode alpha-ketoglutarate dehydrogenase and succinyl coenzyme A synthetase: roles of ArcA, Fnr, and the upstream sdhCDAB promoter. *J. Bacteriol.* **179**, 4138–42

87. Guest, J. R., and Russell, G. C. (1992) Complexes and complexities of the citric acid cycle in Escherichia coli. *Curr. Top. Cell. Regul.* **33**, 231–47

88. Chohnan, S., Furukawa, H., Fujio, T., Nishihara, H., and Takamura, Y. (1997) Changes in the size and composition of intracellular pools of nonesterified coenzyme A and coenzyme A thioesters in aerobic and facultatively anaerobic bacteria. *Appl. Environ. Microbiol.* **63**, 553–60

89. Chohnan, S., Izawa, H., Nishihara, H., and Takamura, Y. (1998) Changes in size of intracellular pools of coenzyme A and its thioesters in Escherichia coli K-12 cells to various carbon sources and stresses. *Biosci. Biotechnol. Biochem.* **62**, 1122–8

90. Wolfe, A. J. (2005) The acetate switch. *Microbiol. Mol. Biol. Rev.* **69**, 12–50

91. Rose, I. A., Grunberg-Manago, M., Korey, S. R., and Ochoa, S. (1954) Enzymatic phosphorylation of acetate. *The Journal of biological chemistry* **211**, 737–56

92. Brown, T. D., Jones-Mortimer, M. C., and Kornberg, H. L. (1977) The enzymic interconversion of acetate and acetyl-coenzyme A in Escherichia coli. *J. Gen. Microbiol.* **102**, 327–36

93. Kumari, S., Tishel, R., Eisenbach, M., and Wolfe, A. J. (1995) Cloning, characterization, and functional expression of acs, the gene which encodes acetyl coenzyme A synthetase in Escherichia coli. *J. Bacteriol.* **177**, 2878–86

94. Wolfe, A. (2010) Physiologically relevant small phosphodonors link metabolism to signal transduction. *Curr Opin Microbiol* **13**, 204–209

95. McCleary, W. R., Stock, J. B., and Ninfa, A. J. (1993) Is acetyl phosphate a global signal in Escherichia coli? *J. Bacteriol.* **175**, 2793–8

96. Karve, T., and Cheema, A. (2011) Small Changes Huge Impact: The Role of Protein Posttranslational Modifications in Cellular Homeostasis and Disease. *J Amino Acids* **2011**, 207691

97. Yang, X.-J. J., and Seto, E. (2008) Lysine acetylation: codified crosstalk with other posttranslational modifications. *Mol. Cell* **31**, 449–61

98. Hunter, T. (2007) The Age of Crosstalk: Phosphorylation, Ubiquitination, and Beyond. *Mol Cell* **28**, 730–738

99. Prabakaran, S., Lippens, G., Steen, H., and Gunawardena, J. (2012) Post-translational modification: nature's escape from genetic imprisonment and the basis for dynamic information encoding. *Wiley Interdiscip Rev Syst Biology Medicine* **4**, 565–583

100. Arnesen, T., Damme, P., Polevoda, B., Helsens, K., Evjenth, R., Colaert, N., Varhaug, J., Vandekerckhove, J., Lillehaug, J., Sherman, F., and Gevaert, K. (2009) Proteomics analyses reveal the evolutionary conservation and divergence of N-terminal acetyltransferases from yeast and humans. *Proc National Acad Sci* **106**, 8157–8162

101. Brown, C. W., Sridhara, V., Boutz, D. R., Person, M. D., Marcotte, E. M., Barrick, J. E., and Wilke, C. O. (2017) Large-scale analysis of post-translational modifications in E. coli under glucose-limiting conditions. *BMC Genomics* **18**, 301

102. Ouidir, T, Jarnier, F, Cosette, P, and Jouenne, T (2015) Characterization of N-terminal protein modifications in Pseudomonas aeruginosa PA14. *Journal of proteomics* **114**, 214-225

103. Drazic, A., Myklebust, L. M., Ree, R., and Arnesen, T. (2016) The world of protein acetylation. *Biochim. Biophys. Acta* **1864**, 1372–401

104. Polevoda, B., Brown, S., Cardillo, T. S., Rigby, S., and Sherman, F. (2008) Yeast N terminal acetyltransferases are associated with ribosomes. *Journal of cellular biochemistry* **103**, 492–508

105. Scott, D., Monda, J., Bennett, E., and Harper, J. (2011) N-terminal acetylation acts as an avidity enhancer within an interconnected multiprotein complex. *Science* **334**, 674-678

106. Behnia, R, Panic, B, Whyte, J., and Munro, S (2004) Targeting of the Arf-like GTPase Arl3p to the Golgi requires N-terminal acetylation and the membrane protein Sys1p. *Nature cell biology* **6(5)**, 405-413

107. Forte, G., Pool, M., and Stirling, C. (2011) N-terminal acetylation inhibits protein targeting to the endoplasmic reticulum. *PLoS biology* **9(5)**, e1001073

108. Holmes, WM, Mannakee, BK, and Gutenkunst, RN (2014) Loss of N-terminal Acetylation Suppresses A Prion Phenotype By Modulating Global Protein Folding. *Nat. Commun.* **5**, 4383

109. Varshavsky, A (2011) The N-end rule pathway and regulation by proteolysis. *Protein Science* **20**, 1298-1345

110. Arnesen, T, Starheim, KK, and Damme, V. P. (2010) The chaperone-like protein HYPK acts together with NatA in cotranslational N-terminal acetylation and prevention of Huntingtin aggregation. *Molecular and cellular biology* **30**, 1898-1909

111. Rope, AF, Wang, K, Evjenth, R, and Xing, J (2011) Using VAAST to identify an X-linked disorder resulting in lethality in male infants due to N-terminal acetyltransferase deficiency. *The American Journal of Human Genetics* **89**, 28-43

112. Kuo, H.-P. P., and Hung, M.-C. C. (2010) Arrest-defective-1 protein (ARD1): tumor suppressor or oncoprotein? *Am J Transl Res* **2**, 56–64

113. Lee, CF, Ou, D., Lee, SB, and Chang, LH (2010) hNaa10p contributes to tumorigenesis by facilitating DNMT1-mediated tumor suppressor gene silencing. *The Journal of Clinical Investigation* **120(8)**, 2920-2930

114. Vetting, M. W., Bareich, D. C., Yu, M., and Blanchard, J. S. (2008) Crystal structure of RimI from Salmonella typhimurium LT2, the GNAT responsible for N(alpha)-acetylation of ribosomal protein S18. *Protein Sci.* **17**, 1781–90

115. Cumberlidge, A. G., and Isono, K. (1979) Ribosomal protein modification in Escherichia coli. I. A mutant lacking the N-terminal acetylation of protein S5 exhibits thermosensitivity. *J. Mol. Biol.* **131**, 169–89

116. Isono, S., and Isono, K. (1981) Ribosomal protein modification in Escherichia coli. III. Studies of mutants lacking an acetylase activity specific for protein L12. *Molecular & general genetics : MGG* **183**, 473–7

117. Isono, K., and Isono, S. (1980) Ribosomal protein modification in Escherichia coli. II. Studies of a mutant lacking the N-terminal acetylation of protein S18. *Molecular & general genetics* **177**, 645–51

118. Roy-Chaudhuri, B., Kirthi, N., Kelley, T., and Culver, G. M. (2008) Suppression of a cold-sensitive mutation in ribosomal protein S5 reveals a role for RimJ in ribosome biogenesis. *Mol. Microbiol.* **68**, 1547–59

119. White-Ziegler, C. A., Black, A. M., Eliades, S. H., Young, S., and Porter, K. (2002) The N-acetyltransferase RimJ responds to environmental stimuli to repress pap fimbrial transcription in Escherichia coli. *J. Bacteriol.* **184**, 4334–42

120. Weinert, B. T., Iesmantavicius, V., Wagner, S. A., Schölz, C., Gummesson, B., Beli, P., Nyström, T., and Choudhary, C. (2013) Acetyl-phosphate is a critical determinant of lysine acetylation in E. coli. *Mol. Cell* **51**, 265–72

121. Kuhn, M. L., Zemaitaitis, B., Hu, L. I., Sahu, A., Sorensen, D., Minasov, G., Lima, B. P., Scholle, M., Mrksich, M., Anderson, W. F., Gibson, B. W., Schilling, B., and Wolfe, A. J. (2014) Structural, kinetic and proteomic characterization of acetyl phosphate-dependent bacterial protein acetylation. *PLoS ONE* **9**, e94816

122. Allfrey, V., Faulkner, R., and Mirsky, A. (1964) Acetylation and Methylation of Histones and their Possible Role in the Regulation of RNA Synthesis. *Proceedings of the National Academy of Sciences* **51**, 786–794

123. Strahl, B. D., and Allis, C. D. (2000) The language of covalent histone modifications. *Nature* **403**, 41–5

124. Wolffe, A., and Hayes, J. (1999) Chromatin disruption and modification. *Nucleic Acids Res* **27**, 711–720

125. Cieniewicz, A., Moreland, L., Ringel, A., Mackintosh, S., Raman, A., Gilbert, T., Wolberger, C., Tackett, A., and Taverna, S. (2014) The Bromodomain of Gcn5 Regulates Site Specificity of Lysine Acetylation on Histone H3. *Mol Cell Proteom* **13**, 2896–2910

126. Mujtaba, S., Zeng, L., and Zhou, M.-M. (2007) Structure and acetyl-lysine recognition of the bromodomain. *Oncogene* **26**, 5521–5527

127. Choudhary, C., Kumar, C., Gnad, F., Nielsen, M. L., Rehman, M., Walther, T. C., Olsen, J. V., and Mann, M. (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* **325**, 834–40

128. Zhao, S., Xu, W., Jiang, W., Yu, W., Lin, Y., Zhang, T., Yao, J., Zhou, L., Zeng, Y., Li, H., Li, Y., Shi, J., An, W., Hancock, S. M., He, F., Qin, L., Chin, J., Yang, P., Chen, X., Lei, Q., Xiong, Y.,

and Guan, K.-L. L. (2010) Regulation of cellular metabolism by protein lysine acetylation. *Science* **327**, 1000–4

129. Barlev, N. A., Liu, L., Chehab, N. H., Mansfield, K., Harris, K. G., Halazonetis, T. D., and Berger, S. L. (2001) Acetylation of p53 activates transcription through recruitment of coactivators/histone acetyltransferases. *Mol. Cell* **8**, 1243–54

130. Matsuyama, A., Shimazu, T., Sumida, Y., Saito, A., Yoshimatsu, Y., Seigneurin-Berny, D., Osada, H., Komatsu, Y., Nishino, N., Khochbin, S., Horinouchi, S., and Yoshida, M. (2002) In vivo destabilization of dynamic microtubules by HDAC6-mediated deacetylation. *EMBO J.* **21**, 6820–31

131. Zhang, J., Sprung, R., Pei, J., Tan, X., Kim, S., Zhu, H., Liu, C.-F., Grishin, N., and Zhao, Y. (2009) Lysine Acetylation Is a Highly Abundant and Evolutionarily Conserved Modification in Escherichia Coli. *Molecular & Cellular Proteomics* **8**, 215–225

132. Zhang, K., Zheng, S., Yang, J., Chen, Y., and Cheng, Z. (2013) Comprehensive profiling of protein lysine acetylation in Escherichia coli. *Journal of proteome research* **12**, 844–51

133. Weinert, B. T., Satpathy, S., Hansen, B. K., Lyon, D., Jensen, L. J., and Choudhary, C. (2017) Accurate quantification of site-specific acetylation stoichiometry reveals the impact of sirtuin deacetylase CobB on the E. coli acetylome. *Mol. Cell Proteomics* **16(5)**, 759-769

134. Schilling, B., Christensen, D., Davis, R., Sahu, A., Hu, L., Walker-Peddakotla, A., Sorensen, D., Zemaitaitis, B., Gibson, B., and Wolfe, A. (2015) Protein acetylation dynamics in response to carbon overflow in Escherichia coli. *Mol Microbiol* **98**, 847–863

135. Yu, B. J., Kim, J. A., Moon, J. H., Ryu, S. E., and Pan, J.-G. G. (2008) The diversity of lysine-acetylated proteins in Escherichia coli. *J. Microbiol. Biotechnol.* **18**, 1529–36

136. Castaño-Cerezo, S., Bernal, V., Post, H., Fuhrer, T., Cappadona, S., Sánchez-Díaz, N., Sauer, U., Heck, A., Altelaar, A., and Cánovas, M. (2014) Protein acetylation affects acetate metabolism, motility and acid stress response in Escherichia coli. *Molecular Systems Biology* **10**, 762

137. Baeza, J., Dowell, J., Smallegan, M., Fan, J., Amador-Noguez, D., Khan, Z., and Denu, J. (2014) Stoichiometry of site-specific lysine acetylation in an entire proteome. *Journal of Biological Chemistry* **289(31)**, 21326-38

138. Wang, Q., Zhang, Y., Yang, C., Xiong, H., Lin, Y., Yao, J., Li, H., Xie, L., Zhao, W., Yao, Y., Ning, Z.-B. B., Zeng, R., Xiong, Y., Guan, K.-L. L., Zhao, S., and Zhao, G.-P. P. (2010) Acetylation of metabolic enzymes coordinates carbon source utilization and metabolic flux. *Science* **327**, 1004–7

139. Kosono, S., Tamura, M., Suzuki, S., Kawamura, Y., Yoshida, A., Nishiyama, M., and Yoshida, M. (2015) Changes in the Acetylome and Succinylome of Bacillus subtilis in Response to Carbon Source. *PLoS ONE* **10**, e0131169

140. Kim, D., Yu, B. J., Kim, J. A., Lee, Y.-J. J., Choi, S.-G. G., Kang, S., and Pan, J.-G. G. (2013) The acetylproteome of Gram-positive model bacterium Bacillus subtilis. *Proteomics* **13**, 1726–36

141. Liu, L., Wang, G., Song, L., Lv, B., and Liang, W. (2016) Acetylome analysis reveals the involvement of lysine acetylation in biosynthesis of antibiotics in Bacillus amyloliquefaciens. *Sci Reports* **6**, 20108

142. Lee, D.-W. W., Kim, D., Lee, Y.-J. J., Kim, J.-A. A., Choi, J. Y., Kang, S., and Pan, J.-G. G. (2013) Proteomic analysis of acetylation in thermophilic Geobacillus kaustophilus. *Proteomics* **13**, 2278–82

143. Wu, X., Vellaichamy, A., Wang, D., Zamdborg, L., Kelleher, N. L., Huber, S. C., and Zhao, Y. (2013) Differential lysine acetylation profiles of Erwinia amylovora strains revealed by proteomics. *J Proteomics* **79**, 60–71

144. Xie, L., Wang, X., Zeng, J., Zhou, M., Duan, X., Li, Q., Zhang, Z., Luo, H., Pang, L., Li, W., Liao, G., Yu, X., Li, Y., Huang, H., and Xie, J. (2014) Proteome-wide lysine acetylation profiling of the human pathogen Mycobacterium tuberculosis. *Int J Biochem Cell Biology* **59**, 193–202

145. Meng, Q., Liu, P., Wang, J., Wang, Y., Hou, L., Gu, W., and Wang, W. (2016) Systematic analysis of the lysine acetylome of the pathogenic bacterium Spiroplasma eriocheiris reveals acetylated proteins related to metabolism and helical structure. *J Proteomics* **148**, 159–169

146. Okanishi, H., Kim, K., Masui, R., and Kuramitsu, S. (2013) Acetylome with Structural Mapping Reveals the Significance of Lysine Acetylation in Thermus thermophilus. *Journal of Proteome Research* **12**, 3952–3968

147. Mizuno, Y., Nagano-Shoji, M., Kubo, S., Kawamura, Y., Yoshida, A., Kawasaki, H., Nishiyama, M., Yoshida, M., and Kosono, S. (2016) Altered acetylation and succinylation profiles in Corynebacterium glutamicum in response to conditions inducing glutamate overproduction. *Microbiologyopen* **5**, 152–73

148. Liang, W., Malhotra, A., and Deutscher, M. (2011) Acetylation Regulates the Stability of a Bacterial Protein: Growth Stage-Dependent Modification of RNase R. *Mol. Cell* **44**

149. AbouElfetouh, A., Kuhn, M. L., Hu, L. I., Scholle, M. D., Sorensen, D. J., Sahu, A. K., Becher, D., Antelmann, H., Mrksich, M., Anderson, W. F., Gibson, B. W., Schilling, B., and Wolfe, A. J. (2015) The E. coli sirtuin CobB shows no preference for enzymatic and nonenzymatic lysine acetylation substrate sites. *Microbiologyopen* **4**, 66–83

150. Tu, S., Guo, S.-J., Chen, C.-S., Liu, C.-X., Jiang, H.-W., Ge, F., Deng, J.-Y., Zhou, Y.-M., Czajkowsky, D., Li, Y., Qi, B.-R., Ahn, Y.-H., Cole, P., Zhu, H., and Tao, S.-C. (2015) YcgC represents a new protein deacetylase family in prokaryotes. *Elife* **4**, e05322

151. Castaño-Cerezo, S., Bernal, V., Blanco-Catalá, J., Iborra, J. L., and Cánovas, M. (2011) cAMP-CRP co-ordinates the expression of the protein acetylation pathway with central metabolism in Escherichia coli. *Mol. Microbiol.* **82**, 1110–28

152. Lima, B. P., Antelmann, H., Gronau, K., Chi, B. K., Becher, D., Brinsmade, S. R., and Wolfe, A. J. (2011) Involvement of protein acetylation in glucose-induced transcription of a stress-responsive promoter. *Mol. Microbiol.* **81**, 1190–204

153. Walker-Peddakotla, A. (2012) Understanding the Regulation of Metabolic Enzyme Acetylation in E. coli. *Master's Theses* 839 [online] http://ecommons.luc.edu/luc_theses_839

154. Starai, V. J., Celic, I., Cole, R. N., Boeke, J. D., and Escalante-Semerena, J. C. (2002) Sir2-dependent activation of acetyl-CoA synthetase by deacetylation of active lysine. *Science* **298**, 2390–2

155. Starai, V. J., and Escalante-Semerena, J. C. (2004) Identification of the protein acetyltransferase (Pat) enzyme that acetylates acetyl-CoA synthetase in Salmonella enterica. *J. Mol. Biol.* **340**, 1005–12

156. Starai, V. J., Gardner, J. G., and Escalante-Semerena, J. C. (2005) Residue Leu-641 of Acetyl-CoA synthetase is critical for the acetylation of residue Lys-609 by the Protein acetyltransferase enzyme of Salmonella enterica. *J. Biol. Chem.* **280**, 26200–5

157. Gardner, J. G., and Escalante-Semerena, J. C. (2009) In Bacillus subtilis, the sirtuin protein deacetylase, encoded by the srtN gene (formerly yhdZ), and functions encoded by the acuABC genes control the activity of acetyl coenzyme A synthetase. *J. Bacteriol.* **191**, 1749–55

158. Cheng, Z.-F. F., and Deutscher, M. P. (2005) An important role for RNase R in mRNA decay. *Mol. Cell* **17**, 313–8

159. Withey, J. H., and Friedman, D. I. (2003) A salvage pathway for protein synthesis: tmRNA and trans-translation. *Annual Reviews in Microbiology* **57**, 101–123

160. Liang, W., and Deutscher, M. (2012) Transfer-messenger RNA-SmpB Protein Regulates Ribonuclease R Turnover by Promoting Binding of HslUV and Lon Proteases. *J Biol Chem* **287**, 33472–33479

161. Chen, C, and Deutscher, MP (2010) RNase R is a highly unstable protein regulated by growth phase and stress. *RNA* **16**, 667-72

162. Liang, W., and Deutscher, M. (2012) Post-translational modification of RNase R is regulated by stress-dependent reduction in the acetylating enzyme Pka (YfiQ). *RNA* **18**, 37–41

163. Majdalani, N., and Gottesman, S. (2005) The Rcs phosphorelay: a complex signal transduction system. *Annu. Rev. Microbiol.* **59**, 379–405

164. Hu, L. I., Chi, B. K., Kuhn, M. L., Filippova, E. V., Walker-Peddakotla, A. J., Bäsell, K., Becher, D., Anderson, W. F., Antelmann, H., and Wolfe, A. J. (2013) Acetylation of the response regulator RcsB controls transcription from a small RNA promoter. *J. Bacteriol.* **195**, 4174–86

165. Thao, S., Chen, C.-S. S., Zhu, H., and Escalante-Semerena, J. C. (2010) Nε-lysine acetylation of a bacterial transcription factor inhibits Its DNA-binding activity. *PLoS ONE* **5**, e15123

166. Meyer, J., D'Souza, A., Sorensen, D., Rardin, M., Wolfe, A., Gibson, B., and Schilling, B. (2016) Quantification of Lysine Acetylation and Succinylation Stoichiometry in Proteins Using Mass Spectrometric Data-Independent Acquisitions (SWATH). *J Am Soc Mass Spectr* **27**, 1758–1771

167. Kouzarides, T. (2000) Acetylation: a regulatory modification to rival phosphorylation? *The EMBO Journal* **19**, 1176-1179

168. Rao, R., Thelen, J., and Miernyk, J. (2014) Is Lys-N-acetylation the next big thing in post-translational modifications? *Trends in plant science* **19**, 550–3

169. Mukherjee, S., Keitany, G., Li, Y., Wang, Y., Ball, H. L., Goldsmith, E. J., and Orth, K. (2006) Yersinia YopJ acetylates and inhibits kinase activation by blocking phosphorylation. *Science* **312**, 1211–4

170. Paquette, N, Conlon, J, Sweet, C, and Rus, F (2012) Serine/threonine acetylation of TGFβ-activated kinase (TAK1) by Yersinia pestis YopJ inhibits innate immune signaling. *Proceedings of the National Academy of Science* **109**, 12710-12715

171. Lee, J., Manning, A. J., Wolfgeher, D., Jelenska, J., Cavanaugh, K. A., Xu, H., Fernandez, S. M., Michelmore, R. W., Kron, S. J., and Greenberg, J. T. (2015) Acetylation of an NB-LRR Plant Immune-Effector Complex Suppresses Immunity. *Cell Reports* **13**, 1670–1682

172. Britton, L.-M. P. M., Gonzales-Cope, M., Zee, B. M., and Garcia, B. A. (2011) Breaking the histone code with quantitative mass spectrometry. *Expert Rev Proteomics* **8**, 631–43

173. Wickstrum, J. R., and Egan, S. M. (2002) Ni+-affinity purification of untagged cAMP receptor protein. *BioTechniques* **33**, 728–30

174. Beatty, C. M., Browning, D. F., Busby, S. J., and Wolfe, A. J. (2003) Cyclic AMP receptor protein-dependent activation of the Escherichia coli acsP2 promoter by a synergistic class III mechanism. *J. Bacteriol.* **185**, 5148–57

175. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr Sect D Biological Crystallogr* **60**, 2126–2132

176. Hudson, B. P., Quispe, J., Lara-González, S., Kim, Y., Berman, H. M., Arnold, E., Ebright, R. H., and Lawson, C. L. (2009) Three-dimensional EM structure of an intact activator-dependent transcription initiation complex. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 19830–5

177. Benoff, B., Yang, H., Lawson, C. L., Parkinson, G., Liu, J., Blatter, E., Ebright, Y. W., Berman, H. M., and Ebright, R. H. (2002) Structural basis of transcription activation: the CAP-alpha CTD-DNA complex. *Science* **297**, 1562–6

178. Molodtsov, V., Fleming, P. R., Eyermann, C. J., Ferguson, A. D., Foulk, M. A., McKinney, D. C., Masse, C. E., Buurman, E. T., and Murakami, K. S. (2015) X-ray crystal structures of Escherichia coli RNA polymerase with switch region binding inhibitors enable rational design of squaramides with an improved fraction unbound to human plasma protein. *J. Med. Chem.* **58**, 3156–71

179. Potterton, L., McNicholas, S., Krissinel, E., Gruber, J., Cowtan, K., Emsley, P., Murshudov, G. N., Cohen, S., Perrakis, A., and Noble, M. (2004) Developments in the CCP4 molecular-graphics project. *Acta Crystallogr Sect D Biological Crystallogr* **60**, 2288–2294

180. Shilov, I. V., Seymour, S. L., Patel, A. A., Loboda, A., Tang, W. H., Keating, S. P., Hunter, C. L., Nuwaysir, L. M., and Schaeffer, D. A. (2007) The Paragon Algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. *Mol. Cell Proteomics* **6**, 1638–55

181. Schilling, B., Rardin, M., MacLean, B., Zawadzka, A., Frewen, B., Cusack, M., Sorensen, D., Bereman, M., Jing, E., Wu, C., Verdin, E., Kahn, C., MacCoss, M., and Gibson, B. (2012) Platform-independent and Label-free Quantitation of Proteomic Data Using MS1 Extracted Ion Chromatograms in Skyline: Application to Protein Acetylation and Phosphorylation. *Molecular & Cellular Proteomics* **11**, 202–214

182. Martin, M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal* 17.1

183. Love, M. I., Huber, W., and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550

184. Edgar, R., Domrachev, M., and Lash, A. E. (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **30**, 207–10

185. Perrenoud, A., and Sauer, U. (2005) Impact of Global Transcriptional Regulation by ArcA, ArcB, Cra, Crp, Cya, Fnr, and Mlc on Glucose Catabolism in Escherichia coli. *Journal of Bacteriology* **187**, 3171–3179

186. Phillips, P. (2008) Epistasis — the essential role of gene interactions in the structure and evolution of genetic systems. *Nat Rev Genet* **9**, 855–867

187. Kumari, S., Beatty, C. M., Browning, D. F., Busby, S. J., Simel, E. J., Hovel-Miner, G., and Wolfe, A. J. (2000) Regulation of acetyl coenzyme A synthetase in Escherichia coli. *J. Bacteriol.* **182**, 4173–9

188. Plumbridge, J. (1998) Expression of ptsG, the gene for the major glucose PTS transporter in Escherichia coli, is repressed by Mlc and induced by growth on glucose. *Molecular Microbiology* **29**, 1053–1063

189. Kimata, K., Inada, T., Tagami, H., and Aiba, H. (1998) A global repressor (Mlc) is involved in glucose induction of the ptsG gene encoding major glucose transporter in Escherichia coli. *Mol. Microbiol.* **29**, 1509–1519

190. Bell, A., Gaston, K., Williams, R., Chapman, K., Kolb, A., Buc, H., Minchin, S., Williams, J., and Busby, S. (1990) Mutations that alter the ability of the Escherichia coli cyclic AMP receptor protein to activate transcription. *Nucleic Acids Res.* **18**, 7243–50

191. Gaston, K., Kolb, A., and Busby, S. (1989) Binding of the Escherichia coli cyclic AMP receptor protein to DNA fragments containing consensus nucleotide sequences. *Biochem. J.* **261**, 649–53

192. Gaston, K., Chan, B., Kolb, A., Fox, J., and Busby, S. (1988) Alterations in the binding site of the cyclic AMP receptor protein at the Escherichia coli galactose operon regulatory region. *Biochem. J.* **253**, 809–18

193. Kumar, S., and Nussinov, R. (2002) Close-range electrostatic interactions in proteins. *Chembiochem* **3**, 604–17

194. Zhou, Y., Kolb, A., Busby, S. J., and Wang, Y.-P. P. (2014) Spacing requirements for Class I transcription activation in bacteria are set by promoter elements. *Nucleic Acids Res.* **42**, 9209–16

195. Thomason, L. C., Sawitzke, J. A., Li, X., Costantino, N., and Court, D. L. (2014) Recombineering: genetic engineering in bacteria using homologous recombination. *Curr Protoc Mol Biol* **106**, 1.16.1–39

196. Ferrández, A., Miñambres, B., García, B., Olivera, E. R., Luengo, J. M., García, J. L., and Díaz, E. (1998) Catabolism of phenylacetic acid in Escherichia coli. Characterization of a new aerobic hybrid pathway. *J. Biol. Chem.* **273**, 25974–86

197. Huang, D. W. a W., Sherman, B. T., and Lempicki, R. A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44–57

198. Huang, D. W. a W., Sherman, B. T., and Lempicki, R. A. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13

199. Weissbach, H., and Brot, N. (1991) Regulation of methionine synthesis in Escherichia coli. *Mol. Microbiol.* **5**, 1593–7

200. Maas, W. K. (1994) The arginine repressor of Escherichia coli. *Microbiol. Rev.* **58**, 631–40

201. Guttenplan, S. B., and Kearns, D. B. (2013) Regulation of flagellar motility during biofilm formation. *FEMS Microbiol. Rev.* **37**, 849–71

202. Fitzgerald, D., Bonocora, R., and Wade, J. (2014) Comprehensive Mapping of the Escherichia coli Flagellar Regulatory Network. *Plos Genet* **10**, e1004649

203. Soutourina, O., Kolb, A., Krin, E., Laurent-Winter, C., Rimsky, S., Danchin, A., and Bertin, P. (1999) Multiple control of flagellum biosynthesis in Escherichia coli: role of H-NS protein and the cyclic AMP-catabolite activator protein complex in transcription of the flhDC master operon. *J Bacteriol* **181**, 7500–8

204. Ohki, R., Kawamata, T., Katoh, Y., Hosoda, F., and Ohki, M. (1992) Escherichia coli dnaJ deletion mutation results in loss of stability of a positive regulator, CRP. *J. Biol. Chem.* **267**, 13180–4

205. Ohki, M., Uchida, H., Tamura, F., Ohki, R., and Nishimura, S. (1987) The Escherichia coli dnaJ mutation affects biosynthesis of specific proteins, including those of the lac operon. *J. Bacteriol.* **169**, 1917–22

206. Kamieniarz, K., and Schneider, R. (2009) Tools to tackle protein acetylation. *Chem. Biol.* **16**, 1027–9

207. Fujimoto, H., Higuchi, M., Koike, M., Ode, H., Pinak, M., Bunta, J., Nemoto, T., Sakudoh, T., Honda, N., Maekawa, H., Saito, K., and Tsuchida, K. (2012) A possible overestimation of the effect of acetylation on lysine residues in KQ mutant analysis. *J Comput Chem* **33**, 239–246

208. Xie, J., and Schultz, P. G. (2006) A chemical toolkit for proteins--an expanded genetic code. *Nat. Rev. Mol. Cell Biol.* **7**, 775–82

209. Neumann, H., Peak-Chew, S. Y., and Chin, J. W. (2008) Genetically encoding N(epsilon)-acetyllysine in recombinant proteins. *Nat. Chem. Biol.* **4**, 232–4

210. Hendrickson, W., Stoner, C., and Schleif, R. (1990) Characterization of the Escherichia coli araFGH and araJ promoters. *J. Mol. Biol.* **215**, 497–510

211. Hollands, K., Busby, S. J., and Lloyd, G. S. (2007) New targets for the cyclic AMP receptor protein in the Escherichia coli K-12 genome. *FEMS Microbiol. Lett.* **274**, 89–94

212. Christensen, D. G., Orr, J. S., Rao, C. V., and Wolfe, A. J. (2017) Increasing Growth Yield and Decreasing Acetylation in Escherichia coli by Optimizing the Carbon-to-Magnesium Ratio in Peptide-Based Media. *Appl. Environ. Microbiol.* **83**, e03034-16

213. Holms, W. H. (1986) The central metabolic pathways of Escherichia coli: relationship between flux and control at a branch point, efficiency of conversion to biomass, and excretion of acetate. *Curr. Top. Cell. Regul.* **28**, 69–105

214. Fredericks, C. E., Shibata, S., Aizawa, S.-I., Reimann, S. A., and Wolfe, A. J. (2006) Acetyl phosphate-sensitive regulation of flagellar biogenesis and capsular biosynthesis depends on the Rcs phosphorelay. *Mol. Microbiol.* **61**, 734–47

215. Francez-Charlot, A., Laugel, B., Gemert, A., Dubarry, N., Wiorowski, F., Castanié-Cornet, M., Gutierrez, C., and Cam, K. (2003) RcsCDB His-Asp phosphorelay system negatively regulates the flhDC operon in Escherichia coli. *Mol Microbiol* **49**, 823–832

216. Gottesman, S., Trisler, P., and Torres-Cabassa, A. (1985) Regulation of capsular polysaccharide synthesis in Escherichia coli K-12: characterization of three regulatory genes. *J. Bacteriol.* **162**, 1111–9

217. Clarke, D. J., Joyce, S. A., Toutain, C. M., Jacq, A., and Holland, I. B. (2002) Genetic analysis of the RcsC sensor kinase from Escherichia coli K-12. *J. Bacteriol.* **184**, 1204–8

218. Ferrières, L., and Clarke, D. (2003) The RcsC sensor kinase is required for normal biofilm formation in Escherichia coli K-12 and controls the expression of a regulon in response to growth on a solid surface. *Mol Microbiol* **50**, 1665–1682

219. Huang, Y.-H., Ferrières, L., and Clarke, D. (2006) The role of the Rcs phosphorelay in Enterobacteriaceae. *Res Microbiol* **157**, 206–212

220. Danese, P. N., Pratt, L. A., and Kolter, R. (2000) Exopolysaccharide production is required for development of Escherichia coli K-12 biofilm architecture. *J. Bacteriol.* **182**, 3593–6

221. Pratt, L., and Kolter, R. (1998) Genetic analysis of Escherichia coli biofilm formation: roles of flagella, motility, chemotaxis and type I pili. *Mol Microbiol* **30**, 285–293

222. Jackson, D., Simecka, J., and Romeo, T. (2002) Catabolite Repression of Escherichia coli Biofilm Formation. *J Bacteriol* **184**, 3406–3410

223. Ferraris, R. P., Yasharpour, S., Lloyd, K. C., Mirzayan, R., and Diamond, J. M. (1990) Luminal glucose concentrations in the gut under normal conditions. *Am. J. Physiol.* **259**, G822–37

224. Bäckhed, F., Ley, R., Sonnenburg, J., Peterson, D., and Gordon, J. (2005) Host-Bacterial Mutualism in the Human Intestine. *Science* **307**, 1915–1920

225. Alam, K. Y., and Clark, D. P. (1989) Anaerobic fermentation balance of Escherichia coli as observed by in vivo nuclear magnetic resonance spectroscopy. *J. Bacteriol.* **171**, 6213–7

226. Parkinson, G., Wilson, C., Gunasekera, A., Ebright, Y., Ebright, R., and Berman, H. (1996) Structure of the CAP-DNA Complex at 2.5 Å Resolution: A Complete Picture of the Protein-DNA Interface. *J Mol Biol* **260**, 395–408

227. Slepenkov, S., and Witt, S. (2002) The unfolding story of the Escherichia coli Hsp70 DnaK: is DnaK a holdase or an unfoldase? *Mol Microbiol* **45**, 1197–1206

228. Rüdiger, S., Schneider-Mergener, J., and Bukau, B. (2001) Its substrate specificity characterizes the DnaJ co-chaperone as a scanning factor for the DnaK chaperone. *EMBO J.* **20**, 1042–50

229. Sell, S. M., Eisen, C., Ang, D., Zylicz, M., and Georgopoulos, C. (1990) Isolation and characterization of dnaJ null mutants of Escherichia coli. *J. Bacteriol.* **172**, 4827–35

230. Vandermarliere, E., Mueller, M., and Martens, L. (2013) Getting intimate with trypsin, the leading protease in proteomics. *Mass Spectrom Rev* **32**, 453–465

231. Zee, B. M., and Garcia, B. A. (2012) Discovery of lysine post-translational modifications through mass spectrometric detection. *Essays Biochem.* **52**, 147–63

232. Kanatani, A., Masuda, T., Shimoda, T., Misoka, F., Lin, X. S., Yoshimoto, T., and Tsuru, D. (1991) Protease II from Escherichia coli: sequencing and expression of the enzyme gene and characterization of the expressed enzyme. *J. Biochem.* **110**, 315–20

233. Jiang, X., Zhang, M., Ding, Y., Yao, J., Chen, H., Zhu, D., and Muramatu, M. (1998) Escherichia coli prlC gene encodes a trypsin-like proteinase regulating the cell cycle. *J. Biochem.* **124**, 980–5

234. Kato, M., Irisawa, T., Ohtani, M., and Muramatu, M. (1992) Purification and characterization of proteinase In, a trypsin-like proteinase, in Escherichia coli. *European Journal of Biochemistry* **210**, 1007–1014

235. Polgár, L. (1997) A potential processing enzyme in prokaryotes: Oligopeptidase B, a new type of serine peptidase. *Proteins Struct Funct Bioinform* **28**, 375–379

236. Wolfe, A.J. (2009) Lessons from a feral promoter. In *The Metabolic Pathway Engineering Handbook: Fundamentals* (Smolke, C., ed.) Chapter 12-1, CRC Press, Stanford

237. Cronan, J., Laporte, D. (2006) Tricarboxylic Acid Cycle and Glyoxylate Bypass. In *EcoSal Plus* (Stewart, V., ed.) American Society for Microbiology, Davis

238. Pelletier, N., Grégoire, S., and Yang, X.-J. J. (2008) Analysis of protein lysine acetylation *in vitro* and *in vivo*. *Curr. Protoc. Protein Sci.* **54**, 14.11:14.11.1-14.11.17

239. Browning, D., Beatty, C., Sanstad, E., Gunn, K., Busby, S., and Wolfe, A. (2003) Modulation of CRP-dependent transcription at the Escherichia coli acsP2 promoter by nucleoprotein complexes: anti-activation by the nucleoid proteins FIS and IHF. *Molecular Microbiology* **51**, 241-254

240. Williams, R., Bell, A., Sims, G., and Busby, S. (1991) The role of two surface exposed loops in transcription activation by the Escherichia coli CRP and FNR proteins. *Nucleic Acids Res.* **19**, 6705–12

241. Sclavi, B., Beatty, C. M., Thach, D. S., Fredericks, C. E., Buckle, M., and Wolfe, A. J. (2007) The multiple roles of CRP at the complex acs promoter depend on activation region 2 and IHF. *Mol. Microbiol.* **65**, 425–40

242. Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., and Mori, H. (2006) Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**, 2006.0008

243. Basan, M., Hui, S., Okano, H., Zhang, Z., Shen, Y., Williamson, J. & Hwa, T. (2015). Overflow metabolism in *Escherichia coli* results from efficient proteome allocation. *Nature* **528**, 99–104

244. Chen, Z., Liu, J., Ng, H., Nadarajah, S., Kaufman, H., Yang, J. & Deng, Y. (2011). Statistical methods on detecting differentially expressed genes for RNA-seq data. *Bmc Syst Biol* **5**, 1–9

245. Benjamini, Y. & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society* **57**, 289–300

VITA

The author, Robert James Davis, was born in Waterloo, Iowa to Rose and Gene Davis. He attended the University of Iowa where he earned Bachelor of Science degrees in both Biology and Biochemistry in May 2005. During his undergraduate studies, he worked in the lab of Dr. Lori Wallrath, studying the effects of nuclear lamin disruption in *D. melanogaster*. After graduation, he spent time at Abbott Laboratories, working as a Senior Quality Technician.

In 2010, Robert matriculated into the Loyola University Chicago Stritch School of Medicine Biochemistry Master's program. After one year, he transferred into the Molecular Biology Ph.D. program and continued his graduate education under the mentorship of Dr. Alan J. Wolfe. Robert's doctoral work focused on the mutual regulation of the transcriptional regulator CRP and Nε-lysine acetylation in *E. coli*. He presented his work at several conferences, including an oral presentation at the Molecular Genetics of Bacteria and Phages meeting in 2014. He was awarded a travel grant from the American Society for Microbiology for Outstanding Student Poster in 2013.

After completion of his graduate studies, Robert will pursue a career in industry.