An acetylatable lysine controls CRP function in *E. coli*

Robert Davis [iD],[1†] Ana Écija-Conesa [iD],[2†]
Julia Gallego-Jara [iD],[2] Teresa de Diego [iD],[2]
Ekaterina V. Filippova,[3] Gina Kuffel,[4]
Wayne F. Anderson,[3] Bradford W. Gibson,[5]
Birgit Schilling [iD],[5] Manuel Canovas [iD][2*] and
Alan J. Wolfe [iD][1*]

[1]*Department of Microbiology and Immunology, Stritch School of Medicine, Health Sciences Division, Loyola University Chicago, Maywood, IL 60153, USA.*
[2]*Department of Biochemistry and Molecular Biology (B) and Immunology, Faculty of Chemistry, University of Murcia, Campus of Espinardo, Regional Campus of International Excellence "Campus Mare Nostrum", Murcia E-30100, Spain.*
[3]*Department of Biochemistry and Molecular Genetics, Center for Structural Genomics of Infectious Diseases, Northwestern University Feinberg School of Medicine, Chicago, IL 60611, USA.*
[4]*Loyola Genomics Facility, Stritch School of Medicine, Health Sciences Division, Loyola University Chicago, Maywood, IL 60153, USA.*
[5]*Buck Institute for Research on Aging, Novato, CA, 94945, USA.*

Summary

Transcriptional regulation is the key to ensuring that proteins are expressed at the proper time and the proper amount. In *Escherichia coli*, the transcription factor cAMP receptor protein (CRP) is responsible for much of this regulation. Questions remain, however, regarding the regulation of CRP activity itself. Here, we demonstrate that a lysine (K100) on the surface of CRP has a dual function: to promote CRP activity at Class II promoters, and to ensure proper CRP steady state levels. Both functions require the lysine's positive charge; intriguingly, the positive charge of K100 can be neutralized by acetylation using the central metabolite acetyl phosphate as the acetyl donor. We propose that CRP K100 acetylation could be a mechanism by which the cell downwardly tunes CRP-dependent Class II promoter activity, whilst elevating CRP steady state levels, thus indirectly increasing Class I promoter activity. This mechanism would operate under conditions that favor acetate fermentation, such as during growth on glucose as the sole carbon source or when carbon flux exceeds the capacity of the central metabolic pathways.

Introduction

Seven global transcription factors (TFs) combine to directly regulate expression of over 50% of all *Escherichia coli* genes (Martínez-Antonio and Collado-Vides, 2003). Given that they also regulate expression of other TFs, the number of genes influenced by these global TFs is staggering. Thus, to properly understand gene regulation at the local and global levels, it is critical to understand how cells regulate the activities of each global TF.

cAMP receptor protein (CRP, also known as the catabolite activator protein [CAP]) is arguably the best-studied of these global TFs (Busby and Ebright, 1999; Lawson *et al.*, 2004). CRP alone is responsible for regulating at least 283 operons (Ishihama *et al.*, 2016). Upon binding cAMP, the CRP dimer binds DNA and directly interacts with RNA polymerase (RNAP). At Class I promoters, CRP binds near positions −61, −71, −81 or −91 relative to the transcription start site (TSS), and the downstream subunit contacts the carboxy-terminal domain of the RNAP α subunit (RNAP α-CTD). CRP residues 156–164 (known as Activating Region 1 [AR1]), located on the surface of the DNA-binding domain, mediate this interaction with RNAP (Zhou *et al.*, 1993; Niu *et al.*, 1994). At Class II promoters, CRP binds near position −41 relative to the TSS and makes three separate interactions with RNAP. An interaction between AR1 of the upstream CRP subunit and the RNAP α-CTD domain is made, similar to the interaction at Class I promoters. A second interaction is made between Activating Region 2 (AR2) on the surface of the downstream CRP subunit and the amino-terminal domain of the RNAP α subunit (RNAP α-NTD). AR2 is a patch of primarily positively charged residues (H19, H21, E96 and K101) that interact with negatively charged residues (E162, E163, D164, E165) on the surface of RNAP α-NTD (Niu *et al.*, 1996). Recent

structural modeling based on the crystal structure of the *Thermus thermophilus* CRP and RNAP homologs bound to promoter DNA in a Class II configuration suggest CRP AR2 may also contact the RNAP β flap (Feng *et al.*, 2016). A third interaction is made between Activating Region 3 (AR3) on the downstream CRP subunit and the RNAP σ^{70} region 4 domain, although the contribution of this interaction to Class II promoter activation appears to be minimal (Rhodius and Busby, 2000a, 2000b).

The existence of two different classes of CRP-dependent promoters suggests the possibility that each class could be regulated separately. The current model of CRP function identifies two factors that affect CRP activity: cAMP concentration and the sequence of each CRP binding site (Lawson *et al.*, 2004). However, altering cAMP concentration would affect both classes equally, and the binding site sequences are hardwired into the DNA and unable to respond to environmental changes. This led us to question the existence of other mechanisms that could regulate CRP activity.

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 (Karve and Cheema, 2011). One of these modifications, Nε-lysine acetylation, has drawn significant attention in recent years, due in large part to the sheer number of acetylated sites in diverse bacterial species (Yu *et al.*, 2008; Zhang *et al.*, 2009, 2013; Wang *et al.*, 2010; Crosby *et al.*, 2012; Okanishi *et al.*, 2013; Wu *et al.*, 2013; Kim *et al.*, 2013; Weinert *et al.*, 2013; Lee *et al.*, 2013; Baeza *et al.*, 2014; Castaño-Cerezo et al., 2014; Liao *et al.*, 2014; Liu *et al.*, 2014, 2016; Pan *et al.*, 2014; Kuhn *et al.*, 2014; Mo *et al.*, 2015; Schilling *et al.*, 2015; Meng *et al.*, 2016; Carabetta *et al.*, 2016; Guo *et al.*, 2016; Ishigaki *et al.*, 2017; Weinert *et al.*, 2017; Chen *et al.*, 2017). Acetylome analyses consistently detect lysine 100 (K100) of CRP as acetylated (Zhang *et al.*, 2009, 2013; Weinert *et al.*, 2013, 2017; Castaño-Cerezo et al., 2014; Kuhn *et al.*, 2014; Schilling *et al.*, 2015). They also reveal that this acetylation depends on the acetyl donor acetyl phosphate (acP) (Weinert *et al.*, 2013; Kuhn *et al.*, 2014). Since K100 is adjacent to the positively charged AR2, we hypothesized that the K100 positive charge may also play a role in CRP function. We further hypothesized that acetylation of K100 would neutralize this positive charge, leading to a potential regulatory mechanism.

In this paper, we provide evidence that K100 exhibits a function that (i) requires its positive charge, (ii) acts independently of K101 and (iii) plays an important role at some native Class II promoters. Unexpectedly, loss of the K100 positive charge also increases CRP steady state levels, which could account for observed increases in transcription from some native Class I promoters. We further show that CRP K100 can be acetylated by acP *in vitro*, complementing the previous observations that K100 is acetylated in an acP-dependent manner *in vivo*. Finally, we discuss the potential for acetylation to inversely regulate CRP-dependent Class II and Class I promoter activities through direct and indirect mechanisms.

Results

CRP K100 positive charge is required for efficient class II transcription

We reasoned that the positive charge of K100 might be required for efficient transcription from Class II promoters, and that the neutralization of the K100 positive charge by acetylation would reduce Class II promoter activity. In support of this hypothesis, a previous report had shown that the K100A mutation reduced Class II promoter activity approximately twofold (Niu *et al.*, 1996). To investigate the role of the K100 positive charge on CRP-dependent promoter activity, we made substitutions at position 100 that mimic a lysine in its unacetylated (K100R) or acetylated (K100Q) states. The arginine residue carries a positive charge, as does an unmodified lysine residue. The glutamine residue is neutral and is more hydrophobic than a lysine residue, much like an acetyllysine. We also included the K101A mutation as a negative control, since loss of K101 results in a significant decrease specifically in Class II promoter activity (Niu *et al.*, 1996). We expected that the K100R mutant would have approximately wild-type (WT) Class II activity, whilst the K100Q mutant would have decreased Class II activity, similar to a K100A mutant. Starting with a Δ*lac* Δ*crp* strain (Supporting Information Table S1), we introduced two plasmids: (i) pDCRP, which carries *crp* alleles encoding either WT CRP or one of the K100/K101 mutants (or pBR322 as vector control) and (ii) pRW50 *CC(-41.5)*, which carries the semi-synthetic Class II CRP-dependent promoter *CC(-41.5)* fused to *lacZ*. These strains were grown in tryptone broth buffered at pH 7 (TB7) supplemented with 22 mM glucose to measure β-galactosidase activity as a readout for *CC(-41.5)* promoter activity.

First, we quantified the steady state level of each CRP mutant and found that the mutants lacking a positive charge at K100 (K100A and K100Q) were present at elevated levels relative to the WT (Supporting Information Fig. S1A). This could not be explained by differential transcription (Supporting Information Fig. S1B). Despite different steady state levels, it has been reported that promoters containing a consensus CRP binding site, including *CC(-41.5)* (and later *CC(-61.5)*),
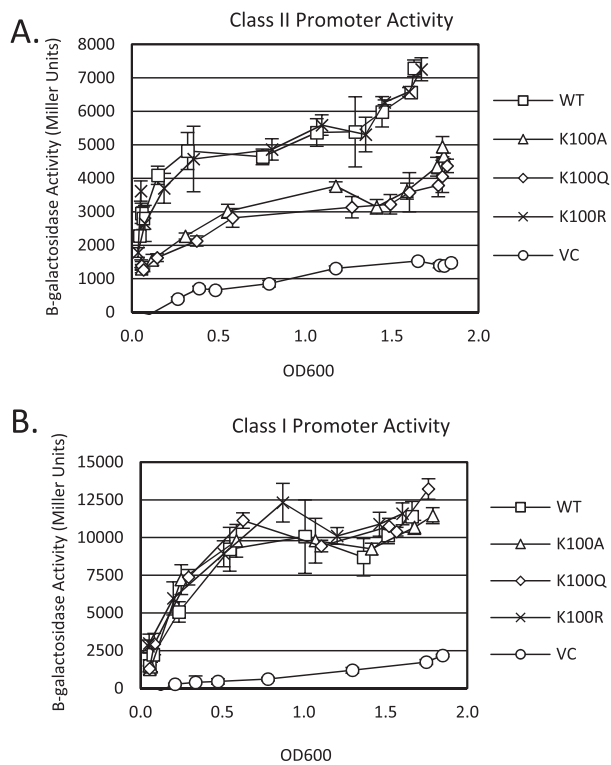
A.

B.

Fig. 1. Loss of K100 positive charge decreases Class II promoter activity. Δ*crp* Δ*lac* strains transformed with pBR322 (VC) or pDCRP encoding either WT CRP or one of the indicated mutants, along with pRW50 carrying the *lac* operon fused to either (A) the Class II promoter *CC(-41.5)* or (B) the Class I promoter *CC(-61.5)*, were grown in TB7 plus 22 mM glucose. Samples were collected at the indicated OD_{600} for β-galactosidase activity measurement. For both A and B, data are representative of two independent experiments performed in triplicate. Error bars represent standard deviation.

are insensitive to CRP concentration due to the extremely high affinity of CRP for the consensus site (Gaston *et al.*, 1988, 1989, 1990). Therefore, we do not expect the different steady state levels to affect the following promoter activity assays.

At *CC(-41.5)*, we observed a decrease in Class II promoter activity in strains expressing K100A or K100Q compared to the strains expressing WT or K100R as expected (Fig. 1A), though not to the same extent as the K101A mutant (Supporting Information Fig. S2). These data suggest that the K100 positive charge is required for efficient Class II promoter activation, though less so than the K101 positive charge.

To determine if the effect of the K100 positive charge was specific to Class II promoters, we performed a similar experiment, except we used pRW50 *CC(-61.5)*, carrying the Class I CRP-dependent promoter *CC(-61.5)* fused to *lacZ* (Fig. 1B). In contrast to the Class II promoter, the K100 mutations had little effect on CRP activity at *CC(-61.5)*, indicating that the K100 positive charge specifically promotes Class II transcription.

K100 regulates class II activity independently of K101

We next asked if K101 was required for the K100-dependent Class II promoter activity. To address this question, we introduced the K100Q and K100R mutations into pDCRP-K101A and examined Class II promoter activity. As expected, the strain expressing K101A CRP had a significant reduction in Class II promoter activity compared to the strain expressing WT CRP (Fig. 2A). However, the strain expressing the K101A
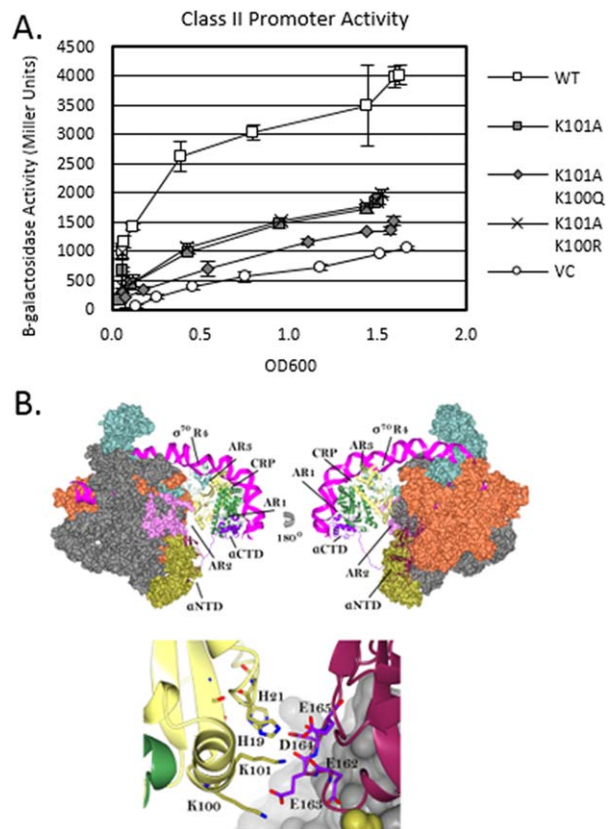
A.

B.

Fig. 2. K100-mediated regulation of Class II promoter activity is independent of K101.
A. Δ*crp* Δ*lac* strains transformed with VC or pDCRP encoding either WT CRP or one of the indicated mutants, along with pRW50 carrying the *lac* operon fused to the Class II promoter *CC(-41.5)*, were grown in TB7 plus 22 mM glucose. Samples were collected at the indicated OD_{600} for β-galactosidase activity measurement. Data are representative of two independent experiments performed in triplicate. Error bars represent standard deviation.
B. Top, model structure of the ternary complex between CRP, RNAP and a Class II promoter. Bottom, interactions between CRP 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 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. Non-carbon atoms of interacting residues are colored in red (oxygen), blue (nitrogen) and yellow (sulfur).

K100Q double mutant exhibited an additional decrease in promoter activity compared to the strains expressing either the K101A single mutant or K101A K100R double mutant, suggesting the K100-dependent regulation of Class II promoter activity is independent of K101. Given that K100 promotes CRP-dependent Class II transcription and acts independently of K101, it is likely K100 makes its own contact with the RNAP α-NTD. We thus modeled CRP and RNAP bound to DNA in a Class II configuration to determine if K100 could get close enough to the surface comprised of residues 162–165 of RNAP α-NTD to make direct contacts (Fig. 2B). Indeed, the model identifies E163 of RNAP α-NTD as a putative interaction partner for CRP K100. Further testing is required to confirm the interaction between E163 of RNAP α-NTD and K100 of CRP. However, these data, along with the transcription data, suggest that K100 is a *bona fide* member of AR2.

K100 positive charge influences global transcription

We next asked if K100's positive charge was required for proper transcription from endogenous promoters. To this end, we performed RNAseq on strains overexpressing WT CRP, K100Q, or K100R variants grown in TB7 supplemented with 22 mM glucose. Relative to WT CRP, a total of 86 and 67 genes were differentially expressed in the K100Q and K100R mutants respectively (Supporting Information Tables S2 and S3). These data support the hypothesis that K100 plays a role in regulating global transcription.

We chose to focus on the comparisons between the K100Q and K100R variants because they represent the two extremes: complete K100 acetylation and no K100 acetylation. Although we do not know the stoichiometry of WT K100 acetylation, we know it changes over time and under different growth conditions (Kuhn *et al.*, 2014; Schilling *et al.*, 2015), making the interpretation of comparisons with WT CRP difficult. Comparing the K100Q and K100R variants, a total of 386 genes were differentially regulated (Supporting Information Tables S2 and S3). Eighteen promoters with a Class II CRP binding site (of 57 total [32%] in the *E. coli* genome [Gama-Castro *et al.*, 2016]) drive transcription of 18 of these genes; genes whose transcription is driven by twelve of these 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 (12/57, 21%) of all known Class II promoters. Twenty-three promoters with at least one Class I CRP binding site, but no Class II site (of 169 total [14%] in the *E. coli* genome [Gama-Castro *et al.*, 2016]), drive transcription of twenty-six of the

differentially regulated genes. Genes whose transcription is driven by fourteen of these promoters exhibited increased transcription in the K100Q mutant relative to the K100R mutant. This implies that the positive charge of K100 does impact transcription of some Class I genes (14/169, 8%) 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 transcribed by 179 promoters with significantly increased expression and 148 genes transcribed by 137 promoters with significantly decreased expression (Supporting Information Tables S2 and S3). CRP is known to regulate transcription of sigma factors and other transcription factors; 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.

To verify that neutralization of the K100 positive charge reduces endogenous Class II promoter activity and to quantify the changes, we performed qRT-PCR on the three Class II-regulated genes with the greatest differential expression between the K100Q and K100R strains identified in the RNAseq experiment. There were fewer *mglC* (Fig. 3A) and *exuT* (Fig. 3B) transcripts in the strains expressing K100A and K100Q relative to strains expressing WT CRP or K100R, in good agreement with what we observed with RNAseq and the *CC(-41.5)* reporter assay. There was no significant difference in *tdcE* (Fig. 3C) transcripts between strains expressing any of the CRP variants, suggesting the *tdcE* promoter may not be regulated by K100. In summary, we were able to verify that at least some of the Class II-regulated genes identified by RNAseq are in fact regulated by K100.

To achieve a more physiologically relevant view of the role of K100, we chose to insert K100Q and K100R mutations into the chromosome. Since we were unable to insert these mutations into the native *crp* locus, we placed *crp* alleles (along with the native promoter region) into the *paaH* gene locus in a Δcrp background. We grew cells in minimal media supplemented with either glucose (10 mM) or acetate (30 mM) as the sole carbon source; exposure to glucose reduces CRP activity by inhibiting cAMP synthesis, whilst growth in acetate increases CRP activity by favoring cAMP synthesis (Deutscher, 2008).

First, we compared the behavior of the new reference (N) strain to that of its WT parent. In minimal medium supplemented with glucose, the rates of growth, glucose consumption and acetate excretion were similar between the two strains (Supporting Information
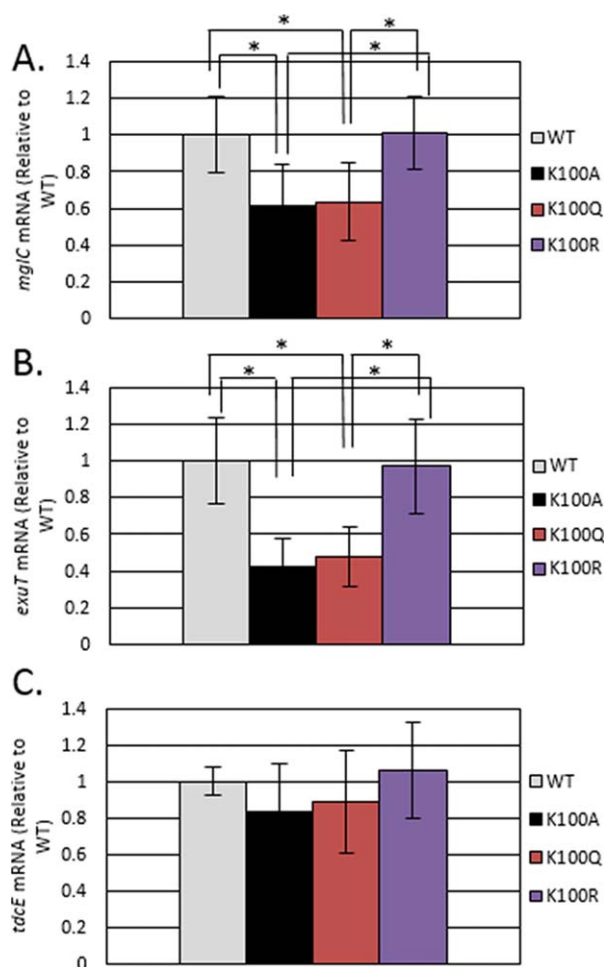
Fig. 3. K100-mediated regulation of endogenous Class II genes. Δ*crp* Δ*lac* strains transformed with pDCRP encoding either WT CRP or one of the indicated mutants were grown in TB7 plus 22 mM glucose to $OD_{600}\sim1.3$. Cells were harvested and subjected to qRT-PCR using primers specific for (A) *mglC*, (B) *exuT* or (C) *tdcE*. Primers specific for 16S rRNA were used as a loading control. Data are a combination of three independent experiments performed in triplicate. Error bars represent standard deviation. (*) indicates $p < 0.05$ as determined by Student's t-test.

Fig. S3A and B). In minimal medium supplemented with acetate (Supporting Information Fig. S3C and D), the rates of growth and acetate consumption were approximately similar. Comparative transcriptome analysis revealed few genes ($< 1.5\%$) with significant differences in gene expression between cells of the reference strain N and WT strains harvested during exponential growth or during stationary phase in minimal medium supplemented with either glucose or acetate (Supporting Information Table S4). We conclude that insertion of the native *crp* gene into the *paaH* gene exerts minimal effects upon global transcription and thus this strategy can be used to assess global transcription in cells of the reference strain N that express either the K100Q or K100R mutant proteins.

Next, we compared the growth characteristics of the reference strain N and its K100Q and K100R mutants. On both carbon sources, they grew at nearly similar rates (Supporting Information Fig. S3A and C); during exponential growth, they consumed glucose (Supporting Information Fig. S3A) and acetate (Supporting Information Fig. S3D) at similar rates. In contrast, the K100Q and K100R mutants excreted less acetate than did their parent N (Supporting Information Fig. S3B).

Finally, we performed a comparative transcriptome analysis using a DNA microarray. We grew the reference strain N and its K100Q and K100R mutants in minimal media supplemented with either glucose (Gl) or acetate (Ac) as the sole carbon source and harvested during exponential growth (X) or after entry into stationary phase (S). Using this strategy, we detected many more differentially regulated genes than we had previously using RNAseq, highlighting the impact of K100 on global gene regulation (Supporting Information Tables S3 and S4 and Fig. S4). To avoid complications in interpretation due to variable levels of K100 acetylation within WT CRP, we compared gene expression between the K100Q mutant and the K100R mutant for each growth condition (Supporting Information Tables S3 and S4). 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. 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 (Supporting

Information Tables S3 and S4). 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.

Using DAVID (Huang *et al.*, 2009a, 2009b), we performed an enrichment analysis of the genes that were differentially regulated between the K100Q and K100R mutants to identify pathways that might be particularly affected by the K100 charge status (Supporting Information Table S5). With 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. In contrast, DAVID did detect pathways that were significantly enriched during growth on acetate. From stationary phase cultures, biosynthesis of amino acids, metabolic pathways and ABC transporters were significantly enriched above background. From exponential phase cultures, the flagellar regulon was significantly enriched above background. For each differentially regulated gene in this regulon, 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.

CRP K100 regulates flagellar motility

Based on the enrichment analyses, we asked if the K100 positive charge regulates flagellar motility. Several flagellar genes were upregulated in the K100Q mutants compared to the K100R mutants in both the RNAseq and DNA microarray analyses (Supporting Information Tables S2 and S4). The flagellar master regulator

FlhDC, encoded by *flhDC*, directly or indirectly regulates transcription of all flagellar genes. Whereas CRP does not directly regulate transcription of individual flagellar genes, it does directly control *flhDC* transcription via a Class I mechanism (Zhao *et al.*, 2007). Since *flhDC* transcription and that of other flagellar genes increased in a K100Q mutant, we might expect a parallel increase in flagellar-based motility when the positive charge is eliminated. We first verified that strains overexpressing CRP variants that lack the positive charge at position 100 (K100A or K100Q) express greater levels of *flhD* transcripts than strains overexpressing CRP variants that retain the positive charge (WT CRP or K100R) (Fig. 4A). Whilst the differences in *flhD* transcription did not reach statistical significance, the pattern of gene expression was consistent with enhanced *flhD* transcription in the mutants lacking the K100 positive charge.

Next, we assessed flagellar motility. The strains overexpressing K100A and K100Q migrated farther than the strains overexpressing either WT CRP or K100R on semi-solid agar plates (Fig. 4B). These results correlate with the *flhD* expression data and support the hypothesis that K100 regulates flagellar-based motility, likely by regulating *flhDC* transcription.

CRP 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*) (Kuhn *et al.*, 2014) and when WT cells were grown under conditions that either favor high acP levels (TB7 supplemented with 22 mM glucose) or low acP levels (TB7 with no supplementation) (Schilling *et al.*, 2015). Taken together, these results support the hypothesis that acP is the acetyl donor for CRP K100. To eliminate the possibility of an indirect effect, we incubated purified CRP with increasing concentrations of acP (0–12.8 mM) for 15 or 120 min and quantified the relative increase in CRP acetylation by western immunoblot analysis using an anti-acetyllysine antibody. We found that CRP acetylation increased in a time- and acP concentration-dependent manner. At the highest acP concentration, this resulted in a ∼twofold increase at 15 min and a ∼12-fold increase at 120 min (Fig. 5). These data clearly show that acP can act as an acetyl donor for CRP.

We next performed quantitative mass spectrometric analysis on the CRP samples incubated for 15 min 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. We identified
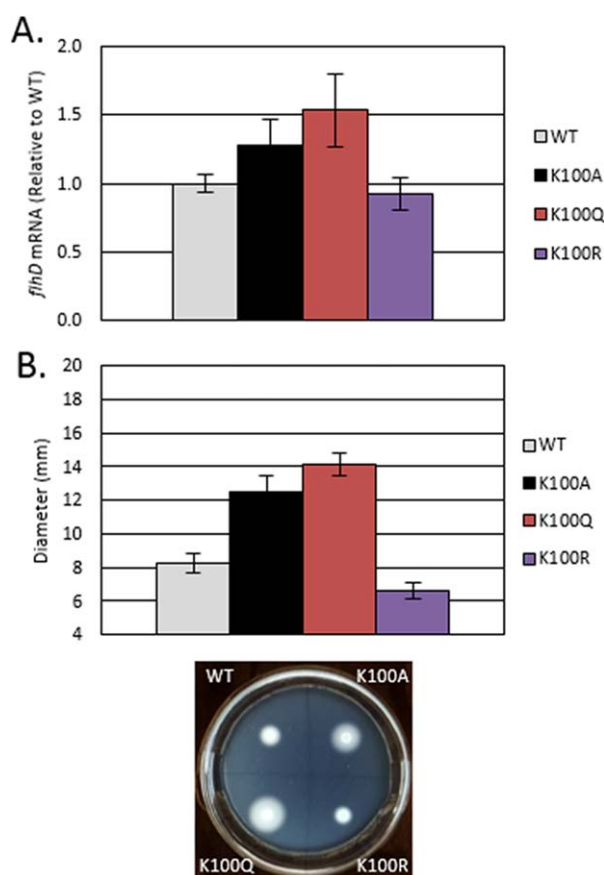
Table S7). These data suggest that several specific lysine residues on CRP are sensitive to acP-dependent acetylation, including K100.

Discussion

Three positively charged residues are traditionally thought to comprise CRP AR2 (H19, H21 and K101) due to their large impact on activity from the semi-synthetic Class II promoter *CC(-41.5)*. Other CRP residues, including K100, have been shown to have minor effects at this promoter (Niu *et al.*, 1996). Here, we confirmed a role for K100 at *CC(-41.5)* and provided additional evidence that its positive charge is essential for full promoter activity (Fig. 1A). The fact that the positive charge at residue 100 promotes transcription activation at a Class II promoter should not be too surprising; K100 is directly adjacent to AR2, a positively charged surface that directly interacts with a negatively charged surface on the RNAP α-NTD at Class II promoters, aiding in the recruitment, stabilization and isomerization of RNAP at such promoters (Niu *et al.*, 1996). We further showed that the influence of the positive charge of K100 extends genome-wide, including to genes that are not directly CRP-dependent (Supporting Information Tables S2–S4). Among the phenotypes regulated by K100 is flagellar-based motility, whose enhancement by the loss of the K100 positive charge correlated with increased

Fig. 4. K100-mediated regulation of motility.
A. Δ*crp* Δ*lac* strains transformed with pDCRP encoding either 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. Pairwise comparisons do not reach statistical significance using Student's *t*-test ($p > 0.05$).
B. (Top) Δ*crp* Δ*lac* strains transformed with pDCRP encoding either WT CRP or one of the indicated mutants were spotted onto semi-solid agar plates and incubated at 30°C. After 10 h, 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. For all pairwise comparisons, $p < 0.01$ using Student's *t*-test. (Bottom) Representative image of motility plate quantified above.

seven lysines that were acetylated in either sample, including K100 (Supporting Information Table S6). Acetylation of K100 and K26 increased 2.1- and 2.4-fold respectively, in the presence of acP (Table 1, Supporting Information Fig. S5A and B). 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 (Supporting Information Fig. S5C and
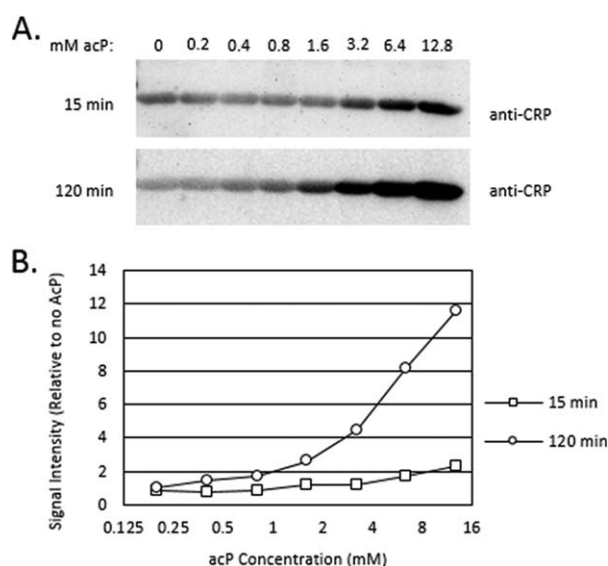

Fig. 5. CRP K100 can be acetylated by acP *in vitro*. Purified CRP (10 μg) was incubated with the indicated concentrations of acP for either 15 or 120 minutes at 37°C. The samples were then subjected to (A) Western immunoblot analysis using anti-acetyllysine antibodies and (B) the relative level of acetylation was quantified. Data are representative of two independent experiments.

Table 1. Fold change in acetylation of CRP lysines that could be quantified by MS1 filtering

Modified Peptide Sequence	Acetyl Site	Precursor M/Z	Precursor Charge	Ratio Incub. Repl. P1 (12.8 mM/0 mM)	*p*-Value P1 ratio	Ratio Incub. Repl. P2 (12.8 mM/0 mM)	*p*-Value P2 ratio	Mean Ratio (12.8 mM/0 mM)
TACEVAEISY**Kac**K	**K-100**	720.86	2	**2.0**	0.013	**2.3**	0.002	**2.1**[a]
TACEVAEISY**Kac**K	**K-100**	480.91	3	**2.0**	0.013	**2.4**	0.044	**2.2**[a]
YPS**Kac**STLIHQGEK	K-26	510.6	3	2.0	0.031	2.9	0.012	2.4[a]
GSVAVLI**Kac**DEEGK[b]	K-52	462.59	3	1.1	0.387	1.5	0.333	1.3
A**Kac**TACEVAEISYK	K-89	756.37	2	1.6	0.084	1.5	0.332	1.6

a. Statistically significant ($p < 0.05$).
b. Deamidated C-terminus.

flhDC transcription (Fig. 4). Finally, we showed that K100 is sensitive to direct acetylation by acP (Table 1 and Supporting Information Table S6).

The status of K100 affects global transcription

The status of K100 influenced many Class II promoters, but not all (Supporting Information Tables S2–S4). Promoter context is likely the major factor for determining which Class II promoters are sensitive to the status of the K100 positive charge. The distance between the CRP binding site and the RNAP binding site plays a significant role in the ability of CRP to activate transcription (Gaston *et al.*, 1990; West *et al.*, 1993). For Class II promoters, it is critical that CRP not only bind close enough to RNAP to make direct contact, but that AR2 aligns with the RNAP α-NTD. Shifting the CRP binding site one nucleotide in either direction not only alters the distance between CRP and RNAP, but also their relative orientation due to the helicity of the DNA. Other DNA-binding elements, such as transcription factors or nucleoid proteins, could also play a role, by bending the DNA or changing the helicity of the DNA, both of which could alter the orientation between CRP and RNAP. These factors may disrupt the interaction between the RNAP α-NTD and some of the other AR2 residues, but they also may reposition K100 to permit a more productive interaction with the RNAP α-NTD and allow it to have a stronger impact on transcription.

The transcription of genes driven by Class I promoters also was affected by the status of the K100 positive charge (Supporting Information Tables S2–S4). This result was unexpected, since AR2 is not thought to play a direct role in Class I transcription (Niu *et al.*, 1996). However, the observation that the K100 positive charge plays a role in CRP steady state levels could explain the observed alterations in Class I transcription. Whereas increasing CRP concentration caused by the K100Q mutation would increase the probability that CRP bound at any of its DNA sites, the effect should be promoter-specific, depending on the role played by CRP at each individual promoter. Because of the increased availability of Class I-competent CRP, the expectation would be increased transcription from Class I promoters.

Why was this global behavior not predicted by investigation of the *CC(-61.5)* promoter? Unlike the majority of native CRP binding sites, the CRP binding site of *CC(-61.5)* is consensus and thus possesses high affinity for CRP (Gaston *et al.*, 1989). At the same time, CRP is overexpressed in the tested strains due to the multicopy nature of pDCRP (Bell *et al.*, 1990). The combination of a high affinity binding site and high CRP concentration could mean the CRP binding site is saturated, leading to maximal *CC(-61.5)* activity in strains expressing WT CRP or K100R variants. Even though CRP levels increase in strains expressing K100A or K100Q, *CC(-61.5)* activation cannot increase any further, masking the expected increase in *CC(-61.5)* activation in these strains. In contrast, at some native promoters, lower CRP concentrations coupled with lower affinity for CRP could allow modulation of CRP concentration to have a greater impact on promoter activity.

Most genes influenced by the status of K100 are driven by promoters that are not directly regulated by CRP. There are at least two non-mutually exclusive mechanisms by which the K100 positive charge status could influence transcription of genes outside of the CRP regulon. First, CRP could be involved in a series of transcriptional hierarchies. Altering the transcription of a K100-dependent transcription factor would in turn alter transcription of its downstream targets, potentially giving rise to major changes in gene expression. This appears to be case for the flagellar gene network; loss of the K100 positive charge increases transcription of CRP-dependent *flhDC*, which in turn increases transcription of the CRP-independent FlhDC regulon, culminating in an increase in flagellar-based motility (Fig. 4). Second, the cell may be responding to changes in its physiology caused by alterations to K100-dependent transcription. The dysregulation of CRP-dependent genes, for example, may activate stress responses or alternative metabolic pathways in an attempt to retain homeostasis.

Examination of the differentially regulated genes within the pathways enriched during growth on acetate revealed 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 (Supporting Information Table S4). 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 (Weissbach and Brot, 1991; Maas, 1994). The apparent lack of activity by the repressors suggests these synthesis pathways are not active. The predicted lack of flux through these pathways may signal that the pathways are misregulated, or that the cell is trying to compensate for disruptions in other pathways.

K100 is a constituent of activation region 2

K101, the closest AR2 neighbor to K100, is not required for K100 to promote Class II activity (Fig. 2A). This leads us to hypothesize that K100 makes its own independent direct contact with the RNAP α-NTD that could work together with that of K101 to stabilize interactions between CRP and RNAP. Our model of CRP and RNAP bound to DNA in a Class II configuration supports this hypothesis (Fig. 2B). The model predicts that CRP K101 interacts with E165 of α-NTD, and that both CRP H19 and H21 interact with D164. This aligns with previous work on the CRP-α-NTD interaction (Niu *et al.*, 1996). Importantly, the model also predicts that CRP K100 interacts with α-NTD E163, which is located in a group of negatively charged residues (162–165) on the surface of the α-NTD known to be important for CRP-dependent Class II transcription (Niu *et al.*, 1996). The interaction between CRP K100 and α E163 must be formally tested, but if this predicted interaction is correct, and since CRP K100 plays a positive role in Class II transcription, we propose that K100 be included in the list of residues considered to comprise AR2.

Recently, another model of *E. coli* CRP and RNAP bound to DNA in a Class II configuration was reported based on the crystal structure of the *T. thermophilus* CRP homolog TAP and RNAP bound to DNA (Feng *et al.*, 2016). In that model, CRP AR2 is predicted to interact with both the α-NTD and β subunits of RNAP, and the authors provide experimental evidence that RNAP β D853 and E859 are important for Class II promoter activation. However, it is not clear from the model with which RNAP residues K100 could interact. Due to lack of an actual crystal structure of the *E. coli* CRP-RNAP complex bound in a Class II configuration,

additional work is needed to determine why K100 is important for Class II promoter activation.

The positive charge of K100 helps maintain CRP level

In addition to their role in Class II promoter activation, the K100 positive charge appears to be involved in maintaining proper CRP steady state levels (Supporting Information Fig. S1A). Loss of the K100 positive charge increased CRP steady state levels in the cell. CRP autoregulates its own transcription via both Class II and Class I mechanisms (Hanamura and Aiba, 1991, 1992), so it is reasonable to imagine that by neutralizing the K100 positive charge and altering CRP function at the *crp* promoter, a feedback loop could increase CRP steady state levels. However, loss of the K100 positive charge reduced *crp* transcript levels (Supporting Information Fig. S1B), indicating that differences in *crp* transcription cannot account for differences in CRP steady state levels. We infer that these positive charges impact CRP steady state levels downstream of transcription, possibly at the level of CRP stability. This hypothesis will be the subject of future study.

A mechanism to differentially regulate class II and class I promoters

We have shown the K100 positive charge plays a role both in CRP activity and in CRP steady state levels. Acetylation of K100 would allow the cell to manipulate the K100 charge status, taking advantage of the roles played by K100 for regulatory purposes. By mass spectrometry, several groups have reproducibly detected acetylated K100 in different *E. coli* strains grown under different conditions (Zhang *et al.*, 2009, 2013; Weinert *et al.*, 2013, 2017 Castaño-Cerezo *et al.*, 2014; Kuhn *et al.*, 2014; Schilling *et al.*, 2015). AcP, the predominant acetyl donor in *E. coli* (Weinert *et al.*, 2013), appears to be responsible for CRP K100 acetylation both *in vivo* (Weinert *et al.*, 2013; Kuhn *et al.*, 2014) and *in vitro* (Table 1 and Supporting Information Table S6).

Acetylation by acP is concentration-dependent: as acP levels rise, acP-dependent acetylation of susceptible lysines increases as well (Schilling *et al.*, 2015). We would thus predict the proposed regulatory mechanism would be most relevant under conditions in which acP levels accumulate. Fermentation of acetogenic sugars (such as glucose) or growth in high concentrations of acetate result in acP synthesis (Wolfe, 2005; Klein *et al.*, 2007; Keating *et al.*, 2008). Supplementation of glucose to buffered tryptone broth leads to increased acetylation, including a 7.4-fold increase in K100 acetylation (Schilling *et al.*, 2015). Growth in minimal medium

supplemented with acetate also induces global acetylation (Weinert *et al.*, 2013; Castaño-Cerezo *et al.*, 2014). The global stoichiometry for acetylation under the tested growth conditions has not yet been determined; however, the microarray data suggest that growth in 10 mM glucose leads to more acetylation than growth in 30 mM acetate. If one assumes that the transcriptional profiles of the K100R and K100Q mutants generally reflects the profiles of cells in which all CRP molecules are either completely unacetylated or completely acetylated at K100 respectively, then we might expect the profile of cells expressing WT CRP to fall somewhere between these two extremes. Intriguingly, the profile of K100Q mutant most resembled the profile of its reference strain N during growth on glucose, whilst the profile of the K100R mutant most resembled the reference strain profile during growth on acetate. This was especially true during stationary phase (Supporting Information Tables S3 and S4 and Fig. S4). From this result, we cautiously infer that K100 of the reference strain N may be differentially acetylated during growth on these two different carbon sources.

In summary, we propose that K100 acetylation is a mechanism by which the cell could potentially turn down transcription from a subset of Class II CRP-dependent promoters, whilst simultaneously turning up transcription from a subset of Class I promoters. This mechanism would be most relevant when the cell ferments or consumes high concentrations of acetate. In the gut, concentrations of glucose and short-chain fatty acids such as acetate can be found at high enough concentrations that one would expect global acetylation to occur, including K100 acetylation (Ferraris *et al.*, 1990; Mortensen and Clausen, 1996). Whether K100 acetylation occurs in the gut and how differential regulation of CRP-dependent promoters might affect cell survival has yet to be explored.

Experimental procedures

Bacterial strains, bacteriophage, plasmids and culture conditions

All bacterial strains, bacteriophage and plasmids used in this study are listed in Supporting Information Table S1. Derivatives were constructed by generalized transduction with P1kc, as described previously (Silhavy *et al.*, 1984). Transformations were performed through the use of either transformation buffers 1 and 2 (Hanahan, 1983) or transformation and storage solution (Chung *et al.*, 1989).

For strain construction, cells were grown in 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. For Western immunoblot and transcription reporter analyses, cells were grown at 37°C in TB7 [1% (w/v) tryptone

buffered at pH 7.0 with potassium phosphate (100 mM)] supplemented with 22 mM glucose, unless otherwise noted. Cell growth was monitored spectrophotometrically (DU640; Beckman Instruments) by determining the optical density at 600 nm (OD_{600}). For microarray experiments, cells were grown in standard M9 minimal medium (pH 7.4) containing: 2.6 mg ml^{-1} $(NH_4)_2SO_4$, 1.0 mg ml^{-1} NH_4Cl, 0.5 mg ml^{-1} NaCl, 15.0 mg ml^{-1} $Na_2HPO_4 \cdot 12H_2O$, 3.0 mg ml^{-1} KH_2PO_4, 50 μg ml^{-1} $FeCl_3 \cdot 6H_2O$, 2.0 mM $MgSO_4$, 65.0 μg ml^{-1} EDTA Na_2, 1.8 μg ml^{-1} $ZnSO_4 \cdot 7H_2O$, 1.8 μg ml^{-1} $CuSO_4 \cdot 5H_2O$, 1.2 μg ml^{-1} $MnSO_4 \cdot H_2O$, 1.8 μg ml^{-1} $CoCl_2 \cdot 6H_2O$, 2.0 mM $MgSO_4$, 0.2 mM $CaCl_2$ and 0.3 mM thiamine·HCl. Either 10 mM glucose or 30 mM acetate were used as the sole carbon source.

Kanamycin (25, 40 or 50 μg ml^{-1}), ampicillin (100 μg ml^{-1}) and chloramphenicol (10 or 25 μg ml^{-1}) were added to growth media when needed.

Site-directed mutagenesis

Site-directed mutagenesis of *crp* in pDCRP was conducted using a QuikChange II Site Directed Mutagenesis Kit (Stratagene), in accordance with the manufacturer's instructions. The list of mutagenic primers used can be found in Supporting Information Table S1.

Western immunoblot analysis

Cells were grown at 37°C in TB7 supplemented with 22 mM glucose. At regular intervals, they 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 h at 100 V 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. 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. The blot was washed three times for 5 min each with TBST [TBS containing 0.05% (v/v) Tween-20] and then incubated with HRP-conjugated goat anti-mouse secondary antibody (Millipore AP503P, 1:2500) 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.

qRT-PCR. For quantification of *crp*, *mglC*, *exuT* and *tdcE* mRNA, cells were grown in TB7 supplemented with 22 mM

glucose at 37°C and harvested in exponential phase (OD_{600} ~1.3). 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). One milliliter culture was added to 2 mL RNAProtect Bacterial Reagent (Qiagen 76506), and vortexed. After centrifugation, the cell pellet was frozen at −80°C and stored overnight. RNA was isolated using MasterPure RNA Purification Kit (Epicentre MCR85102), using the manufacturer's protocol. RNA concentration/purity was determined using a Nano-Drop 2000 spectrophotometer (Thermo Scientific). 1 ug RNA was used to synthesize cDNA, using the iScript cDNA Synthesis Kit (BioRad 170–8890), according to the manufacturer's protocol. Quantitative reverse transcription PCR (qRT-PCR) was performed using iTaq Universal SYBR Green Supermix (BioRad 172–5120), according to manufacturer's protocol, on a CFX96 Real-Time System (Bio-Rad) with a C1000 Thermal Cycler (BioRad) and the following conditions: 95°C for 5 min, then 37 cycles of (95°C for 15 s, 60°C for 20 s, 72°C for 30 s). Quantitation of 16S rRNA was used to normalize the data. The list of primers used can be found in Supporting Information Table S1. Each experiment included at least two replicates and was performed at least twice.

Promoter activity assay. Cells were grown at 37°C in TB7 supplemented with 22 mM 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 (Beatty *et al.*, 2003). 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 (Emsley and Cowtan, 2004) 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 (Hudson *et al.*, 2009). In our model, CRP binds to DNA as a dimer of two identical protomers, centered around position −41 relative to the TSS. 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 (Benoff *et al.*, 2002). Interactions between CRP AR2 and RNAP involve residues 162–165 located on the flexible loop of N-terminal domain of RNAP αNTD (Niu *et al.*, 1996). To model these interactions, we used a crystal structure of RNAP in complex with squaramide compound (Molodtsov *et al.*, 2015), 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 (Rhodius and

Busby, 2000a). These interactions involve residues (52–55, 58) on the β-turn of the CRP DNA-binding domain and residues (593–596, 599) located on an α-helix of the RNAP σ^{70} region 4 domain. Therefore, taking into account flexibility of the DNA, the αNTD-AR2 interacting loop and geometric restraints of the linker between αNTD and αCTD of RNAP, we propose a structural model of the ternary complex of CRP-RNAP at Class II promoter. Figures of the model were created in the CCP4 molecular graphics (CCP4mg) program (Potterton *et al.*, 2004).

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 min. 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 (Martin, 2011). 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 (Love *et al.*, 2014), 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.

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 Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995). Unless otherwise indicated, genes that are up- or downregulated with an adjusted *p*-value < 0.1 were identified as 'significantly modulated genes.'

The global gene expression data discussed in this work have been deposited in NCBI's Gene Expression Omnibus (Edgar *et al.*, 2002) and are accessible through GEO Series accession number GSE97406 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc= GSE97406).

Genomic insertion

The protocol used to insert *crp* alleles into the *paaH* locus was specifically designed for this study, based on the methods of Ying *et al.* (Ying *et al.*, 2013) and Blank *et al.* (Blank *et al.*, 2011) with minor modifications. First, we used PCR to amplify the native *crp* gene including the promoter region (CRP N) from *E. coli* BW25113 genomic DNA. We then used the restriction enzymes *XbaI* and *PstI* to clone the amplicon into plasmid pBAD24 (Guzman *et al.*, 1995) to obtain the plasmid pBAD24-CRP N. Next, we used PCR to amplify the FRT-Km-FRT fragment from plasmid pKD13 (Datsenko and Wanner, 2000) and the restriction enzymes *EcoRI* and *KpnI* to clone the amplicon upstream of the CRP N allele to obtain the plasmid pBAD24-FRT-Km-FRT-CRP N, which carries the wild-type gene expressed from its native promoter linked to the kanamycin resistance cassette. Using this plasmid as the template, we performed site-directed mutagenesis using the Stratagene kit according to the manufacturer's instructions, to construct the plasmids pBAD24-FRT-Km-FRT-CRP K100R and pBAD24-FRT-Km-FRT-CRP K100Q, which carry the mutant alleles *crp*K100R (R) and *crp*K100Q (Q) expressed from the native *crp* promoter linked to the kanamycin resistance cassette. All the primers used are listed in Supporting Information Table S1.

These three alleles were inserted into the *paaH* locus. First, we used PCR amplification to obtain linear target sequences from the three plasmids described above. The resultant amplicon was digested with the restriction enzyme *DpnI* to eliminate the template plasmid DNA; this step reduced the likelihood of false positives. These restricted linear fragments contained the FRT-Km-FRT-CRP N, R or Q alleles flanked by 90 bp of sequence up- and downstream of the *paaH* gene. We chose the *paaH* locus for two reasons. First, the *paa* operon encodes enzymes involved in phenylacetate degradation, which is unessential during growth of *E. coli* on glucose or acetate as the sole carbon source (Ferrández *et al.*, 1998). Additionally, transcriptome analysis of WT cells revealed that this operon is not transcribed under these growth conditions (Supporting Information Table S4). Furthermore, the *paaH* locus is located $2 \cdot 10^6$ bp from *crp* locus, far enough away to avoid undesirable recombination events during the genomic insertion process between the FRT sequences in the target

sequences and the FRT scar in the Δ*crp* allele of the recipient strain JW5702 (Baba *et al.*, 2006).

We transformed this recipient strain with the plasmid pWRG730, which carries the temperature-inducible redαβ recombination system. Into this transformant, we introduced the linear target sequences by electroporation using 0.1 cm cuvettes in a Gene Pulser Xcell™ Electroporation System (Biorad, Munich, Germany), according to the manufacturer's instructions. The recombination system was induced with a heat-shock (12.5 min to 42°C).

The transformation mixture was plated onto LB-agar plates containing 25 μg ml^{-1} kanamycin to select for insertion of the *crp* allele. Kanamycin-resistant colonies were verified as bona fide recombinants by PCR using various primer pairs. Bona fide recombinants were cured of the plasmid pWRG730 and the kanamycin resistance cassette was eliminated from the genome using plasmid pCP20 (Datsenko and Wanner, 2000). The result was the set of three strains: the reference strain BW25113 Δ*crp* Δ*paaH::crpP-crp N* (N) and its mutants BW25113 Δ*crp* Δ*paaH::crpP-crp R* (R) and BW25113 Δ*crp* Δ*paaH::crpP-crp Q* (Q).

DNA microarray. Global gene expression was assessed in M9 minimal medium with 10 mM glucose or 30 mM acetate as sole carbon source. Aerobic 50 ml batch cultures were grown in 500 mL flasks at 37°C on a rotary shaker at 250 r.p.m. 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 Vantage™ Total RNA Purification Kit (Origene, Rockville, MD, USA). Purity and concentration of isolated RNA were assessed in a Nano-Drop 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, according to the manufacturer's recommendations (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 through the use of 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 (Edgar *et al.*, 2002) and are accessible through GEO Series accession number GSE96955 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE96955).

Motility assay

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 h. 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.

CRP purification

CRP purification was performed as previously described (Wickstrum and Egan, 2002).

In vitro acetylation

Purified CRP (0.5 μg μl^{-1}) 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 min or 120 min 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 min. After incubation, the reactions were cooled to room temperature and were separated by SDS-polyacrylamide gel electrophoresis and subjected to Western immunoblot analysis 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 min. 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 10 mM dithiothreitol in 25 mM NH$_4$HCO$_3$ at 56°C for 1 h and alkylated with 55 mM iodoacetamide in 25 mM NH$_4$HCO$_3$ 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$_2$O, 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 (Kuhn *et al.*, 2014). 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 ms 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 Supporting Information Table S6 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 (Shilov *et al.*, 2007) 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 (Schilling *et al.*, 2012). Detailed MS1 Filtering quantification results for acetylated CRP peptides are provided in Supporting Information Table S6. Quantification of potential protein level changes upon acP incubation analyzing 23 robustly observed, non-acetylated peptides by MS1 Filtering is shown in Supporting Information Table S7.

References

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., and Baba, M. (2006) Construction of *Escherichia coli* K-

12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2**: 2006. 0008.

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. *J Biol Chem* **289**: 21326–21328.

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**: 5148–5157.

Bell, A., Gaston, K., Williams, R., Chapman, K., Kolb, A., and Buc, H. (1990) Mutations that alter the ability of the *Escherichia coli* cyclic AMP receptor protein to activate transcription. *Nucleic Acids Res* **18**: 7243–7250.

Benjamini, Y., and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B* **57**: 289–300.

Benoff, B., Yang, H., Lawson, C.L., Parkinson, G., Liu, J., Blatter, E., *et al.* (2002) Structural basis of transcription activation: the CAP-alpha CTD-DNA complex. *Science* **297**: 1562–1566.

Blank, K., Hensel, M., Gerlach, R.G., and Xu, S-y. (2011) Rapid and highly efficient method for scarless mutagenesis within the *Salmonella enterica* chromosome. *PLoS One* **6**: e15763.

Busby, S., and Ebright, R. (1999) Transcription activation by catabolite activator protein (CAP). *J Mol Biol* **293**: 199–213.

Carabetta, V.J., Greco, T.M., Tanner, A.W., Cristea, I.M., Dubnau, D., and Overall, C.M. (2016) Temporal regulation of the *Bacillus subtilis* acetylome and evidence for a role of MreB acetylation in cell wall growth. *mSystems* **1**: e00005–16.

Castaño-Cerezo, S., Bernal, V., Post, H., Fuhrer, T., Cappadona, S., Sanchez-Diaz, N.C., *et al.* (2014) Protein acetylation affects acetate metabolism, motility and acid stress response in *Escherichia coli*. *Mol Syst Biol* **10**: 762.

Chen, Z., Zhang, G., Yang, M., Li, T., Ge, F., and Zhao, J. (2017) Lysine acetylome analysis reveals photosystem II manganese-stabilizing protein acetylation is involved in negative regulation of oxygen evolution in model *Cyanobacterium Synechococcus* sp. PCC 7002. *Mol Cell Proteomics* **16**: 1297–1311.

Chung, C.T., Niemela, S.L., and Miller, R.H. (1989) One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. *Proc Natl Acad Sci U S A* **86**: 2172–2175.

Crosby, H., Pelletier, D., Hurst, G., and Escalante-Semerena, J. (2012) System-wide studies of N-lysine acetylation in *Rhodopseudomonas palustris* reveal substrate specificity of protein acetyltransferases. *J Biol Chem* **287**: 15590–15601.

Datsenko, K.A., and Wanner, B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* **97**: 6640–6645.

Deutscher, J. (2008) The mechanisms of carbon catabolite repression in bacteria. *Curr Opin Microbiol* **11**: 87–93.

Edgar, R., Domrachev, M., and Lash, A.E. (2002) Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* **30**: 207–210.

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

Feng, Y., Zhang, Y., and Ebright, R.H. (2016) Structural basis of transcription activation. *Science* **352**: 1330–1333.

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–G837.

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**: 25974–25986.

Gama-Castro, S., Salgado, H., Santos-Zavaleta, A., Ledezma-Tejeida, D., Muñiz-Rascado, L., García-Sotelo, J.S., *et al.* (2016) RegulonDB version 9.0: high-level integration of gene regulation, coexpression, motif clustering and beyond. *Nucleic Acids Res* **44**: D133–D143.

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**: 809–818.

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**: 649–653.

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

Guo, J., Wang, C., Han, Y., Liu, Z., Wu, T., Liu, Y., *et al.* (2016) Identification of Lysine Acetylation in *Mycobacterium abscessus* Using LC-MS/MS after Immunoprecipitation. *J Proteome Res* **15**: 2567–2578.

Guzman, L.M., Belin, D., Carson, M.J., and Beckwith, J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* **177**: 4121–4130.

Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**: 557–580.

Hanamura, A., and Aiba, H. (1991) Molecular mechanism of negative autoregulation of *Escherichia coli crp* gene. *Nucleic Acids Res* **19**: 4413–4419.

Hanamura, A., and Aiba, H. (1992) A new aspect of transcriptional control of the *Escherichia coli crp* gene: positive autoregulation. *Mol Microbiol* **6**: 2489–2497.

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

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

Hudson, B.P., Quispe, J., Lara-González, S., Kim, Y., Berman, H.M., Arnold, E., *et al.* (2009) Three-dimensional EM structure of an intact activator-dependent transcription initiation complex. *Proc Natl Acad Sci USA* **106**: 19830–19835.

Ishigaki, Y., Akanuma, G., Yoshida, M., Horinouchi, S., Kosono, S., and Ohnishi, Y. (2017) Protein acetylation

involved in streptomycin biosynthesis in *Streptomyces griseus*. *J Proteomics* **155**: 63–72.

Ishihama, A., Shimada, T., and Yamazaki, Y. (2016) Transcription profile of *Escherichia coli*: genomic SELEX search for regulatory targets of transcription factors. *Nucleic Acids Res* **44**: 2058–2074.

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.

Keating, D.H., Shulla, A., Klein, A.H., and Wolfe, A.J. (2008) Optimized two-dimensional thin layer chromatography to monitor the intracellular concentration of acetyl phosphate and other small phosphorylated molecules. *Biol Proced Online* **10**: 36–46.

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**: 1726–1736.

Klein, A.H., Shulla, A., Reimann, S.A., Keating, D.H., and Wolfe, A.J. (2007) The intracellular concentration of acetyl phosphate in *Escherichia coli* is sufficient for direct phosphorylation of two-component response regulators. *J Bacteriol* **189**: 5574–5581.

Kuhn, M.L., Zemaitaitis, B., Hu, L.I., Sahu, A., Sorensen, D., Minasov, G., *et al.* (2014) Structural, kinetic and proteomic characterization of acetyl phosphate-dependent bacterial protein acetylation. *PLoS One* **9**: e94816.

Lawson, C.L., Swigon, D., Murakami, K.S., Darst, S.A., Berman, H.M., and Ebright, R.H. (2004) Catabolite activator protein: DNA binding and transcription activation. *Curr Opin Struct Biol* **14**: 10–20.

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**: 2278–2282.

Liao, G., Xie, L., Li, X., Cheng, Z., and Xie, J. (2014) Unexpected extensive lysine acetylation in the trump-card antibiotic producer *Streptomyces roseosporus* revealed by proteome-wide profiling. *J Proteomics* **106**: 260–269.

Liu, F., Yang, M., Wang, X., Yang, S., Gu, J., Zhou, J., *et al.* (2014) Acetylome analysis reveals diverse functions of lysine acetylation in *Mycobacterium tuberculosis*. *Mol Cell Proteomics* **13**: 3352–3366.

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 Rep* **6**: 20108.

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.

Maas, W.K. (1994) The arginine repressor of *Escherichia coli*. *Microbiol Rev* **58**: 631–640.

Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* **17**: 10–12.

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

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.

Mo, R., Yang, M., Chen, Z., Cheng, Z., Yi, X., Li, C., *et al.* (2015) Acetylome analysis reveals the involvement of lysine acetylation in photosynthesis and carbon metabolism in the model *Cyanobacterium synechocystis* sp. PCC 6803. *J Proteome Res* **14**: 1275–1286.

Molodtsov, V., Fleming, P.R., Eyermann, C.J., Ferguson, A.D., Foulk, M.A., McKinney, D.C., *et al.* (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**: 3156–3171.

Mortensen, P.B., and Clausen, M.R. (1996) Short-chain fatty acids in the human colon: relation to gastrointestinal health and disease. *Scand J Gastroenterol* **31**: 132–148.

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.

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**: 1123–1134.

Okanishi, H., Kim, K., Masui, R., and Kuramitsu, S. (2013) Acetylome with structural mapping reveals the significance of lysine acetylation in *Thermus thermophilus*. *J Proteome Res* **12**: 3952–3968.

Pan, J., Ye, Z., Cheng, Z., Peng, X., Wen, L., and Zhao, F. (2014) Systematic analysis of the lysine acetylome in *Vibrio parahemolyticus*. *J Proteome Res* **13**: 3294–3302.

Potterton, L., McNicholas, S., Krissinel, E., Gruber, J., Cowtan, K., Emsley, P., *et al.* (2004) Developments in the CCP4 molecular-graphics project. *Acta Crystallogr Sect D Biol Crystallogr* **60**: 2288–2294.

Rhodius, V.A., and Busby, S.J. (2000a) 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**: 311–324.

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

Schilling, B., Rardin, M., MacLean, B., Zawadzka, A., Frewen, B., Cusack, M., *et al.* (2012) Platform-independent and label-free quantitation of proteomic data using MS1 extracted ion chromatograms in skyline: application to protein acetylation and phosphorylation. *Mol Cell Proteomics* **11**: 202–214.

Schilling, B., Christensen, D., Davis, R., Sahu, A., Hu, L., Walker-Peddakotla, A., *et al.* (2015) Protein acetylation dynamics in response to carbon overflow in *Escherichia coli*. *Mol Microbiol* **98**: 847–863.

Shilov, I.V., Seymour, S.L., Patel, A.A., Loboda, A., Tang, W.H., Keating, S.P., *et al.* (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**: 1638–1655.

Silhavy, T.J., Berman, M.L., and Enquist, L.W. (1984) *Experiments with gene fusions*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Wang, Q., Zhang, Y., Yang, C., Xiong, H., Lin, Y., Yao, J., *et al.* (2010) Acetylation of metabolic enzymes coordinates carbon source utilization and metabolic flux. *Science* **327**: 1004–1007.

Weinert, B.T., Iesmantavicius, V., Wagner, S.A., Schölz, C., Gummesson, B., Beli, P., *et al.* (2013) Acetyl-phosphate is a critical determinant of lysine acetylation in *E. coli*. *Mol Cell* **51**: 265–272.

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**: 759–769.

Weissbach, H., and Brot, N. (1991) Regulation of methionine synthesis in *Escherichia coli*. *Mol Microbiol* **5**: 1593–1597.

West, D., Williams, R., Rhodius, V., Bell, A., Sharma, N., Zou, C., *et al.* (1993) Interactions between the *Escherichia coli* cyclic AMP receptor protein and RNA polymerase at class II promoters. *Mol Microbiol* **10**: 789–797.

Wickstrum, J.R., and Egan, S.M. (2002) Ni+-affinity purification of untagged cAMP receptor protein. *BioTechniques* **33**: 728–730.

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

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.

Ying, B.-W.W., Akeno, Y., and Yomo, T. (2013) Construction of synthetic gene circuits in the *Escherichia coli* genome. *Methods Mol Biol* **1073**: 157–168.

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**: 1529–1536.

Zhang, J., Sprung, R., Pei, J., Tan, X., Kim, S., Zhu, H., *et al.* (2009) Lysine acetylation is a highly abundant and evolutionarily conserved modification in *Escherichia coli*. *Mol Cell Proteomics* **8**: 215–225.

Zhang, Q., Gu, J., Gong, P., Wang, X., Tu, S., Bi, L., *et al.* (2013) Reversibly acetylated lysine residues play important roles in the enzymatic activity of *Escherichia coli* N-hydroxyarylamine O-acetyltransferase. *FEBS J* **280**: 1966–1979.

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 Res* **35**: 4441–4452.

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 USA* **90**: 6081–6085.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web-site.