BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

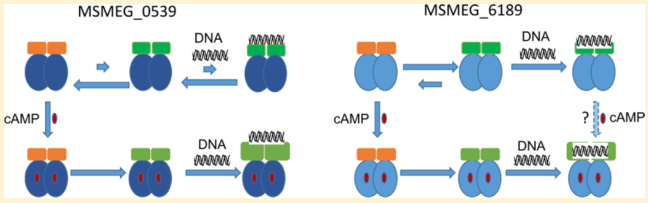
pubs.acs.org/biochemistry

# Paralogous cAMP Receptor Proteins in *Mycobacterium smegmatis* Show Biochemical and Functional Divergence

Ritu Sharma,[†,▽] Anisha Zaveri,[†] Jayashree Gopalakrishnapai,[†] Srinath Thiruneelakantan,[‡] Umesh Varshney,[‡] and Sandhya S. Visweswariah*[,†]

[†]Department of Molecular Reproduction, Development and Genetics and [‡]Department of Microbiology and Cell Biology, Indian Institute of Science, Bangalore 560012, India

**ABSTRACT:** The cyclic AMP receptor protein (CRP) family of transcription factors consists of global regulators of bacterial gene expression. Here, we identify two paralogous CRPs in the genome of *Mycobacterium smegmatis* that have 78% identical sequences and characterize them biochemically and functionally. The two proteins (MSMEG_0539 and MSMEG_6189) show differences in cAMP binding affinity, trypsin sensitivity, and binding to a CRP site that we have identified upstream of the *msmeg_3781* gene. MSMEG_6189 binds to the CRP site readily in the absence of cAMP, while MSMEG_0539 binds in the presence of cAMP, albeit weakly. *msmeg_6189* appears to be an essential gene, while the Δ*msmeg_0539* strain was readily obtained. Using promoter−reporter constructs, we show that *msmeg_3781* is regulated by CRP binding, and its transcription is repressed by MSMEG_6189. Our results are the first to characterize two paralogous and functional CRPs in a single bacterial genome. This gene duplication event has subsequently led to the evolution of two proteins whose biochemical differences translate to differential gene regulation, thus catering to the specific needs of the organism.




Cyclic AMP, via the cAMP receptor protein (CRP), plays an important role in the regulation of gene expression in diverse bacteria.[1] CRP is a transcription factor that belongs to the CRP/FNR family, which includes members that are global regulators associated with functions that are highly diverse in different bacterial species.[2] The biological processes under the regulation of CRP include adaptation to starvation, extreme temperature, low oxygen energy metabolism, cell division, toxin production, quorum sensing, and cellular motility.[3−7]

Mycobacteria, including *Mycobacterium tuberculosis*, which is the causative agent of tuberculosis, contain extraordinarily high levels of cAMP[8,9] produced by a large number of adenylyl cyclases encoded in their genomes.[10] The genome of *M. tuberculosis* H37Rv contains two CRP-like genes, encoded by ORFs Rv1675c and Rv3676.[9,11,12] The proteins are only 25% identical at the amino acid level and 46% similar at the sequence level. Both CRPs have a distinct set of regulons and are therefore suggested to be functionally divergent. Rv1675c regulates genes in *M. tuberculosis* upon macrophage infection and in response to elevated cAMP levels, but a strain from which this gene has been deleted does not exhibit attenuated virulence in mice.[13] Deletion of the Rv3676 gene from *M. tuberculosis* led to attenuation of the strain in the macrophage and mouse model, and Rv3676 directly regulates the *rpfA* gene that is considered to be important for persistence and reactivation of the pathogen during infection.[14,15] Rv3676 also regulates a large variety of genes involved in important cellular functions, suggesting its role as a global regulator in *M. tuberculosis*.[14,16,17]

Rv3676 differs in its cAMP binding, DNA binding, and transcriptional activity from the *Escherichia coli* paradigm.[18] Rv3676 is a homodimer that binds to two molecules of cAMP,[19] but unlike *E. coli* CRP, the cAMP-binding sites in Rv3676 do not show cooperativity and have a relatively weak affinity for cAMP (with a $K_b$ of ∼$1.7 \times 10^4$ $M^{-1}$).[18] There are reports that suggest that binding of cAMP to CRP results in only a small enhancement in specific DNA binding,[14,20] unlike in *E. coli* CRP where cAMP binding is an obligatory step for DNA binding.[21] The DNA-binding sites for Rv3676 were identified by a SELEX-based technique, and the 16 bp consensus sequence obtained [C/T G T G/C A/G $(N)_6$ T/C C/G A C G/A] was similar to the *E. coli* motif.[20,22] Computational predictions and experimental evidence from chromatin immunoprecipitation and sequencing allowed the identification of CRP-binding sites in the *M. tuberculosis* and *Mycobacterium smegmatis* genomes.[16,20,23]

In this study, we characterize two CRPs encoded in the genome of *M. smegmatis* that are 78% identical in terms of sequence yet show significant differences in their biochemical properties in terms of cAMP binding. We identify a CRP site upstream of the *msmeg_3781* gene and show differential regulation of the promoter, thus exemplifying the natural evolution of two proteins that, despite extensive similarity at the primary amino acid sequence, are likely to perform different functions in the cell.

**Received:** July 28, 2014
**Revised:** October 12, 2014
**Published:** November 17, 2014

   dx.doi.org/10.1021/bi500924v | *Biochemistry* 2014, 53, 7765−7776

## MATERIALS AND METHODS

**Bioinformatics.** Sequences of the 16S rRNA genes of the indicated species were aligned using ClustalW[24] implemented in MEGA (version 4) with default settings. An unrooted neighbor-joining (NJ) tree was constructed using MEGA.[25]

**Cloning and Expression of MSMEG_0539 and MSMEG_6189.** *msmeg_0539* and *msmeg_6189* were amplified from *M. smegmatis* genomic DNA using specific primers, sequences of which are available on request. The amplicon of *msmeg_0539* was digested with EcoRI and XhoI and cloned into a similarly digested pGEX-6P2 vector to generate pGEX_MS0539. The *msmeg_6189* amplicon was digested with EcoRI and NotI and cloned into similarly digested pGEX-6P3 to yield pGEX_MS6189. The sequence of the insert was confirmed (Macrogen, Seoul, South Korea). The sequences of all inserts were verified (Macrogen).

The SP850 strain of *E. coli*[26] (deficient in the adenylyl cyclase gene) was transformed with either pGEX_MS0539 or pGEX_MS6189. Protein production in the cells was induced using 500 $\mu$M IPTG for 3 h at 37 °C. The cells were lysed by sonication (Branson model 450 sonifier) for 3−5 min, in lysis buffer [50 mM Tris (pH 7.5), 200 mM NaCl, 5 mM DTT, 5 mM EDTA, 1 mM benzamidine, and 2 mM PMSF]. Following sonication, the lysate was centrifuged at 30000*g* for 30 min at 4 °C. The supernatant interacted with GSH beads for 1 h at 4 °C. Beads were then washed with a buffer containing 50 mM Tris (pH 8.0), 200 mM NaCl, 5 mM DTT, 5 mM EDTA, and 0.1% Triton X-100 and later with buffer containing 50 mM Tris (pH 8.0), 200 mM NaCl, 5 mM DTT, 5 mM EDTA, and 10% glycerol. The GST tag was cleaved using PreScission protease (1 $\mu$g for 10 $\mu$g of the fusion protein) with the immobilized protein on GSH beads, and the tag-less CRP protein was obtained from the supernatant. Protein estimation was performed using the modified Bradford method.[27]

**Partial Proteolysis of Mycobacterial CRPs.** Mycobacterial CRPs (10 $\mu$M) were treated with trypsin (1 $\mu$M) (Sigma) in a final volume of 25 $\mu$L, in buffer containing 50 mM HEPES and 100 mM NaCl for 1 h at 4 °C. Where indicated, the protease digestion reaction mixtures were supplemented with cAMP (1 mM) or cGMP (1 mM). The reaction was stopped by adding soy bean trypsin inhibitor (1 $\mu$g/mL) and heating the samples at 95 °C in SDS loading dye for 10 min. The samples were resolved on a 15% polyacrylamide gel and stained with Coomassie for visualization of cleaved peptides.

CRP peptide bands were cut and subjected to tryptic digestion by procedures reported previously.[28] Tryptic peptides were mixed with an equal volume of matrix solution ($\alpha$-cyano-4-hydroxycinnamic acid), spotted for analysis by a MALDI-TOF-TOF (Ultra Flex TOF-TOF, Bruker Daltonics) instrument equipped with a $N_2$ laser, and analyzed in reflectron mode using a time delay of 90 ns and an accelerating voltage of 25 kV in positive ion mode. The protein mass fingerprint (PMF) data were used to map the peptide on the sequence of the protein using MASCOT on the *M. smegmatis* proteomics database (www.proteomics.mbu.ernet.in), with a tolerance of 1 Da.

To determine the size of proteolytic fragments and thus identify the region of the proteins that they encompassed, mycobacterial CRPs (10 $\mu$M) were treated with porcine trypsin (1 $\mu$M) (Sigma) in a final volume of 25 $\mu$L in 20 mM Tris-HCl (pH 8.0), 20 mM NaCl, and 5 mM DTT at 25 °C for 3 h for MSMEG_0539 and 1 h for MSMEG_6189. The reaction was stopped by adding 1 mM PMSF. An aliquot was subjected to sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE) analysis to confirm the digestion, and the remaining mixture was used for MALDI-TOF analysis as described above. The obtained masses were used to map the peptides on the sequence of the protein using FindPept[29] (tolerance of 3 Da; methionines and tryptophans optionally oxidized; cysteines in the reduced form).

**Isothermal Titration Calorimetry (ITC).** The two CRPs were cloned into the pPRO vectors and expressed as N-terminal hexahistidine-tagged proteins. Details of cloning strategies can be provided upon request. Plasmids were transformed into SP850 cells, and protein expression was induced following the addition of 500 $\mu$M IPTG at 16 °C overnight. Pelleted cells were resuspended in lysis buffer containing 20 mM Tris (pH 7.5), 100 mM NaCl, 10% glycerol, 5 mM $\beta$-mercaptoethanol, 2 mM PMSF, and 1 mM benzamidine and lysed by sonication. After sonication, all purification steps were conducted at 20 °C as high concentrations of the proteins precipitate at 4 °C. The sonicated cell lysate was centrifuged at 30000*g*, and the supernatant was loaded onto a 3 mL Ni-NTA agarose column previously equilibrated with lysis buffer. The bound protein was washed with 3−5 bed volumes of wash buffer [20 mM Tris (pH 7.5), 500 mM NaCl, 20 mM imidazole, 10% glycerol, and 5 mM $\beta$-mercaptoethanol], and bound CRPs were eluted in a buffer containing 20 mM Tris (pH 7.5), 100 mM NaCl, 300 mM imidazole, 5 mM $\beta$-mercaptoethanol, and 10% glycerol. Proteins were concentrated using a 10 kDa cutoff Centricon Concentrator (Amicon) followed by dialysis against 20 mM Tris (pH 7.5), 100 mM NaCl, 10% glycerol buffer.

ITC experiments were conducted using a VP-ITC calorimeter from MicroCal Inc. (Northampton, MA); 1.8 mL of 60−100 $\mu$M protein solutions was titrated with a 6−10-fold molar excess of 3′,5′-cAMP (present in the same buffer as the protein) from a 300 $\mu$L syringe, while being stirred at a speed of 386−400 rpm at 20 °C. Each injection (10 $\mu$L) occurred over 20 s, with an interval of 3 min between each injection. The heats of dilution of the 3′,5′-cAMP solution were subtracted from titration heats, and the data so obtained were fit via a nonlinear least-squares minimization method to determine the binding stoichiometry ($n$), the binding constant ($K_b$), the change in the enthalpy of binding ($\Delta H_b$), and the change in entropy ($\Delta S_b$), using MicroCal-enabled Origin software. Each titration experiment was repeated at least twice.

**Electrophoretic Mobility Shift Assay (EMSA).** A double-stranded oligonucleotide harboring the CRP site present upstream of *msmeg_3780/3781* genes (encompassing the 16 bp CRP site and including an additional 7 bp on either side) was end labeled by $T_4$ polynucleotide kinase (PNK) using [$\gamma$-$^{32}$P]ATP. The labeled double-stranded oligonucleotide was purified from the unincorporated nucleotide by passing it through a Sephadex G-25 spin column equilibrated with sterile Milli-Q water.

Indicated amounts of CRP were incubated in buffer containing 25 mM Tris (pH 7.8), 50 mM KCl, 5 mM $MgCl_2$, 10 $\mu$M EDTA, 1 $\mu$g of poly-dI/dC (for nonspecific binding), 10% glycerol, and 50 $\mu$g/mL BSA in a 20 $\mu$L reaction volume, along with the labeled oligonucleotide (50 fmol) (~10000 cpm per reaction tube). Where indicated, cAMP (1 mM) was added in the reaction mix prior to incubation. In some cases, cytosolic fractions (10 $\mu$g) prepared from *M. smegmatis* cells as described previously[30] were used. Samples were incubated at 37 °C for 20 min, in the absence or presence


dx.doi.org/10.1021/bi500924v | *Biochemistry* 2014, 53, 7765−7776

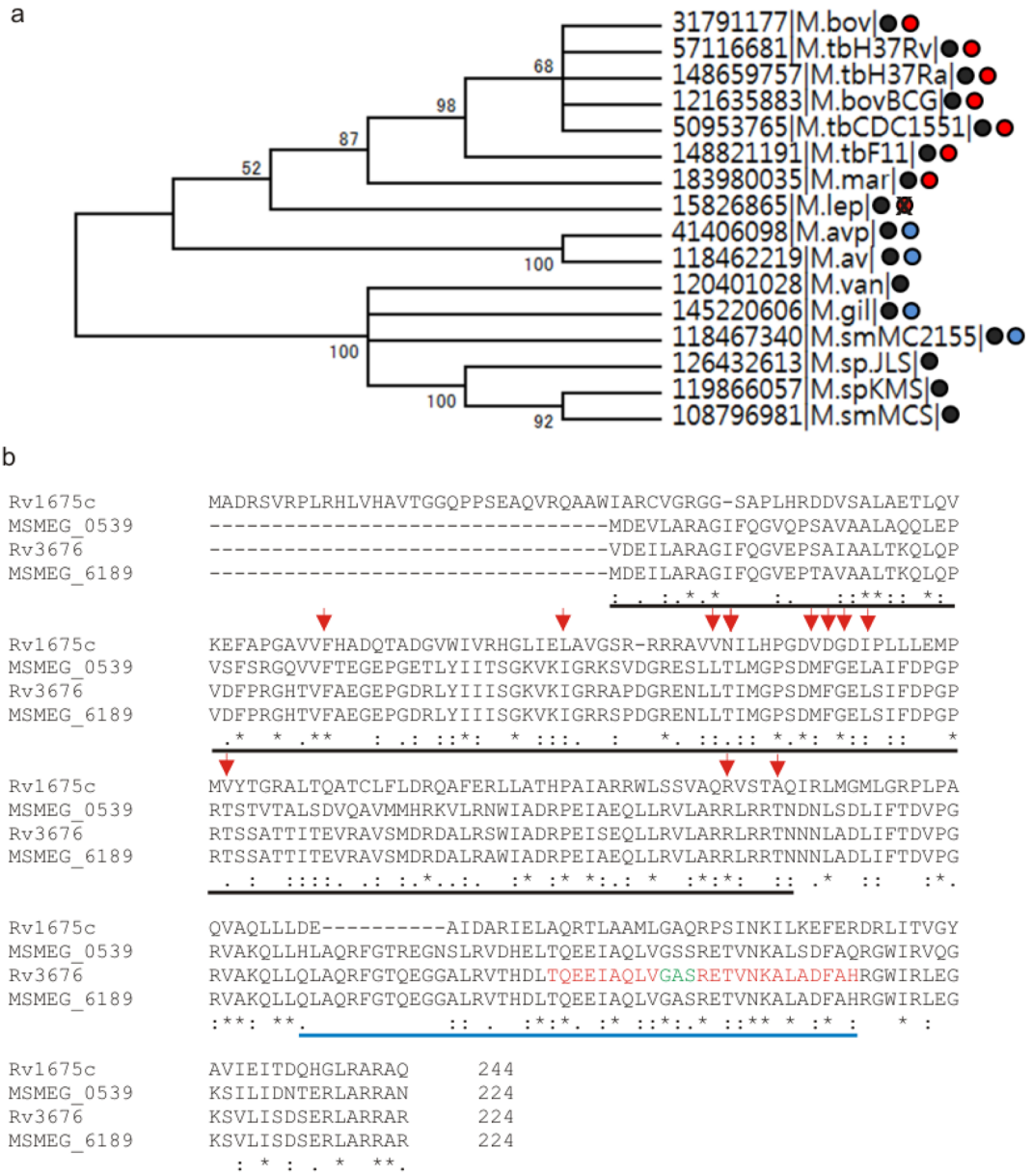



**Figure 1.** CRPs in mycobacteria. (a) The unrooted, NJ tree with bootstrap values is shown. Dots next to each member of the tree represent a CRP gene in the genome. Dots colored black and red indicate the presence of two CRPs in the genome that are only 23−25% identical in sequence. Dots colored gray and blue indicate the presence of CRPs that are 68−78% identical in sequence. A cross on the dot refers to the presence of a pseudogene. The members of the tree are identified by the gi number, the name of the host bacterium, the strain designation where applicable, and their primary annotation. (b) Multiple-sequence alignment of Rv3676, Rv1675c, and the *M. smegmatis* CRPs (MSMEG_0539 and MSMEG_6189), generated by ClustalW.[24] Regions underlined with black or blue indicate the cyclic AMP-binding domain or the DNA-binding domain, respectively. Red arrows indicate the residues that are important for cAMP binding in Rv3676, and the red and green residues in the Rv3676 sequence represent the helices and "turn" in the HTH motif, respectively.

of an unlabeled double-stranded CRP site oligonucleotide (5 pmol). Purified anti-Rv3676 IgG (1 $\mu$g) was added to the reaction mix for supershift assays. The complexes were electrophoresed on 5% polyacrylamide gels in Tris-borate/EDTA (45 mM Tris-borate and 1 mM EDTA) buffer at 100 V and 4 °C for 1.5 h to separate the bound complex from the unbound probe. Subsequently, the gels were dried and scanned using a phosphoimager (Bass 1800, Fuji).

**DNase I Footprint Assay.** MS3780_150promRvs primer (5′-ATATCGCCGCTGGTCACG-3′) was end labeled by $T_4$ polynucleotide kinase using [$\gamma$-$^{32}$P]ATP at 37 °C for 90 min. The labeled template DNA containing the CRP site was obtained following amplification using a MS3780_150promRvs (labeled) and MS3780promXbaStu primer (5′-ACT CTA GAG GCC TCA TTC GGC ATC CCT GTC CA-3′, 10 pmol each) and a pMV-MS3780prom construct[8] as the template. The labeled DNA template was purified using silica beads (Gene Clean kit, MP Biochemicals).

For the footprint assay, the labeled DNA template (0.1 $\mu$M) was incubated with purified *M. smegmatis* CRPs (1 $\mu$M) independently, in a total volume of 10 $\mu$L of binding buffer [50 mM Tris (pH 7.5), 10 mM $MgCl_2$, 10 mM $CaCl_2$, 1 mM DTT, 0.1 mg/mL BSA, and 20% glycerol]. Where indicated, cAMP (1 mM) was added to the reaction tube. The mixture was


dx.doi.org/10.1021/bi500924v | *Biochemistry* 2014, 53, 7765−7776

incubated for 20 min at 37 °C, and DNase I (0.1 milliunit/μL) was then added and digestion continued for 3 min at 37 °C. The nuclease digestion was stopped by adding EDTA (1 mM) and heating the sample in reaction termination dye (containing 90% formamide, 0.05% bromophenol blue, and 0.05% xylene cyanol) at 90 °C for 5 min. Subsequently, the samples were snap-chilled and loaded onto a 6.5% polyacrylamide−urea (8 M) denaturing gel and run at 1500 V for 3−4 h. The sequences of the protected and hypersensitive bands were identified by comparison with the sequencing ladder (G, A, T, C) of the same DNA fragment. The gel was dried, exposed to a phosphoimager plate, and scanned using a phosphoimager (Typhoon 9210) after overnight exposure.

**Western Blotting of *M. smegmatis* CRPs.** The polyclonal antibody to Rv3676 protein was available in the laboratory. Equal amounts of purified MSMEG_0539, MSMEG_6189 protein, or cytosolic fractions (30 μg) from *M. smegmatis* cells grown for 12 h (log phase) or 30 h (stationary phase) were subjected to Western blot analysis essentially as described previously.[30] Immunoreactive bands were visualized by enhanced chemiluminescence.

**Luciferase Reporter Assays.** The genomic region containing the putative promoter of MSMEG_3781 was amplified by PCR using primers MS3780prmScaSpe (5′-TACTAGTACTCAGCGCACGGCGGAAGC-3′) and MS3780prmXbaStu (5′-ACTCTAGAGGCCTCATTCGGCA-TCCCTGTCCA-3′). The product was digested with XbaI and SpeI and cloned into SpeI-digested plasmid pBKSII to generate plasmid pBKS-pMSMEG_3781. The sequence of the inset was verified (Macrogen). The XbaI−BamHI fragment from pBKS-pMSMEG_3781 and the BamHI−HindIII fragment from pBKS-luxAB[8] were cloned into XbaI−HindIII-digested vector pMV-10-25[8] to generate plasmid pMV-pMSMEG_3781-luxAB.

For the generation of a promoter construct with mutations in the CRP site, site-directed mutagenesis was conducted on template plasmid pBKS-pMSMEG_3780[8] using MS3780CRPmutf (5′-TCCATGTCCTGACGTGCATCTCA-AGTACCT-3′) and MS3780CRPmutr (5′-AGGTACTTGAG-ATGCACGTCAGGACATGGA-3′) primers. The primers included a G → C mutation at position 2 and an A → T mutation at position 14 of the CRP site, thereby disrupting the palindrome that is required for CRP binding. The sequence of the mutant promoter was confirmed by sequencing (Macrogen). This plasmid (pBKS-mutpMSMEG_3780) was used as a template for PCR with MS3780prmScaSpe and MS3780prmXbaStu. The resulting product was digested with XbaI and SpeI and cloned into SpeI-digested plasmid pBKSII to generate pBKS-mutpMSMEG_3781. The XbaI−BamHI fragment from pBKS-mutpMSMEG_3781 and the BamHI−HindIII fragment from pBKS-luxAB were cloned into XbaI−HindIII-digested vector pMV-10-25[8] to generate plasmid pMV-mutpMSMEG_3781-luxAB.

Mycobacterial strains were grown in 7H9TG (Middlebrook 7H9 broth supplemented with 0.2% glycerol, 50 μg/mL hygromycin, and 0.05% Tween 80) at 37 °C, while being shaken at 200 rpm. Promoter plasmids were electroporated into *M. smegmatis* cells, and colonies were picked. Cultures at an OD of ∼1.0 were diluted into fresh medium to an OD of 0.02. Aliquots of cultures at indicated time points were taken, and luciferase counts were measured as previously described.[8]

**Generation of the Δ*msmeg_0539* Strain.** A strain with the deletion of the *msmeg_0539* gene (Δ*msmeg_0539*) was generated using the suicide vector approach.[31] The 5′ knockout fragment was amplified by PCR using primers MS0539 5′ f (5′-GTCGTCGCGCTAGCGTGGCGTGTCAG-3′) and MS0539 5′ r (5′-CGGCCGAACTAGTCCTTCCCCGACGTGATG-3′) and cloned into the pBKSII vector. The 3′ knockout fragment was amplified by PCR using primers MS0539 3′ f (5′-CGCAAAACAATTGCTGCATCTCGC-3′) and MS0539 3′ r (5′-TGTTGAGGGATATCGCCATGAGCAGCGAC-3′) and cloned into the pBKSII vector. DNA sequencing confirmed the identity of the cloned sequences compared to that reported in the genome. The 5′ and 3′ knockout fragments were excised using restriction digestion from the respective pBKSII vectors and cloned into vector p2NIL to generate p2NIL-MSMEG_0539-5′3′ KO. The PacI fragment from pGOAL19 containing three markers (β-galactosidase, hygromycin resistance, and sucrose susceptibility) was cloned into p2NIL-MSMEG_0539-5′3′ KO to generate plasmid p2NIL-MSMEG_0539-5′3′ KO-PacI. Plasmid p2NIL-MSMEG_0539-5′3′ KO-PacI was electroporated into *M. smegmatis* mc$^2$155, and single crossovers and double crossovers were obtained essentially as described previously.[8,31] Double crossovers were further tested by genomic PCR and RT-PCR to obtain and confirm the Δ*msmeg_0539* strain.

**RNA Isolation and Reverse Transcriptase PCR.** RNA was isolated as previously described.[28] The transcript levels of MSMEG_0539, MSMEG_6189, MSMEG_3780, and MSMEG_3781 were assessed using specific primers (sequences available on request). Levels of 16S rRNA were assessed to normalize for RNA that had been reverse transcribed.

## RESULTS

**Identification of Two Similar CRPs in *M. smegmatis*.** The presence of two CRP-like genes is evident in many mycobacterial genomes (Figure 1a). One CRP-like gene is found in all species of mycobacteria sequenced so far (represented by a black dot), while divergent CRPs [similar to Rv1675c (red dot)] are present only in a subset of mycobacteria, including *M. tuberculosis*. The HTH motifs present in Rv3676-like and Rv1675c-like mycobacterial CRPs are distinct in terms of sequence, suggesting unique sets of regulons for the two CRP types, and in agreement with earlier reports.[13,20]

Interestingly, some mycobacteria contain two CRPs, both of which are Rv3676-like (represented by black and blue dots). *M. smegmatis* presents such an example, harboring two genes (*msmeg_0539* and *msmeg_6189*) that encode putative CRP-like proteins. The two genes are 78% identical in nucleotide sequence, suggesting the occurrence of a gene duplication event during the evolution of mycobacterial species. Figure 1b shows a sequence alignment of Rv3676 and Rv1675c (divergent CRPs) along with the two proteins from *M. smegmatis* with sequences that are highly similar to that of Rv3676. MSMEG_0539 shares 77% amino acid sequence identity with Rv3676, whereas MSMEG_6189 shares 97% sequence identity with Rv3676. The sequences of MSMEG_0539 and MSMEG_6189 are 78% identical and 90% similar. Residues implicated in cAMP binding in Rv3676 are conserved in both MSMEG_0539 and MSMEG_6189 (Figure 1b[19]). The close similarity in the sequences of the HTH domain suggests that the two proteins are likely to recognize similar DNA-binding sites in the genome. This therefore raises the question of whether the two CRPs present in *M. smegmatis* represent an example of gene redundancy or whether these two gene

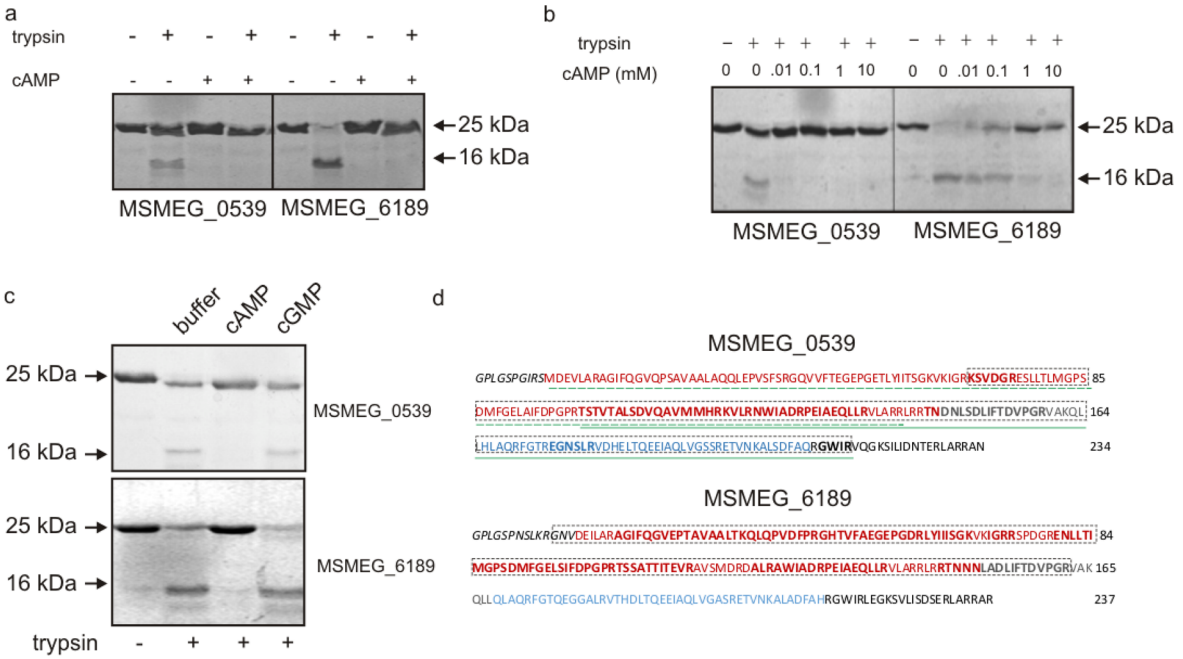



**Figure 2.** Trypsin sensitivity of MSMEG_6189 and MSMEG_0539. (a) Purified CRPs (10 $\mu$M) were incubated with trypsin in the absence or presence of 1 mM cAMP. Peptides were subjected to SDS−PAGE and bands visualized by Coomassie staining. (b) Reaction mixtures containing 10 $\mu$M CRPs (MSMEG_0539 and MSMEG_6189) were incubated with trypsin (1 $\mu$M) and varying concentations of cAMP as indicated and then subjected to SDS−PAGE. (c) MSMEG_0539 and MSMEG_6189 were incubated with trypsin (1 $\mu$M) in the presence of either cAMP or cGMP (1 mM each). Samples were then subjected to SDS−PAGE. (d) Peptide mixtures obtained from partial trypsin digests of MSMEG_0539 and MSMEG_6189 were taken for MALDI analysis. Sequences of the CRPs are shown as would be obtained after PreScission protease cleavage of the GST fusion proteins. The N-terminal residues originating from the linker between GST and CRP are shown in italics. The cAMP-binding domain is colored red, the linker/hinge region gray, and the DNA-binding domain blue. The boxed regions indicate the 16502 and 16630 Da peptides for MSMEG_0539 and MSMEG_6189, respectively. For MSMEG_0539, the dotted green underline indicates the 14317 Da peptide and the solid green underline the 14379 Da peptide. For both MSMEG_0539 and MSMEG_6189, the 16 kDa fragment was also subjected to peptide mass fingerprinting, and the peptides obtained are indicated as bold.

products have distinct biochemical properties that may reflect divergence in their biological roles.

**Cyclic AMP Binding to *M. smegmatis* CRPs.** Various lines of evidence indicate that cAMP allosterically regulates the DNA binding ability of CRP.[32−34] In *E. coli* CRP, as well as Rv3676, this conformational change was assayed by changes in the susceptibility to protease cleavage in the presence and absence of cAMP. We therefore cloned and expressed MSMEG_6189 and MSMEG_0539 to obtain pure protein, which as determined by gel filtration and dynamic light scattering experiments, formed dimers in the absence or presence of 1 mM cAMP (data not shown).

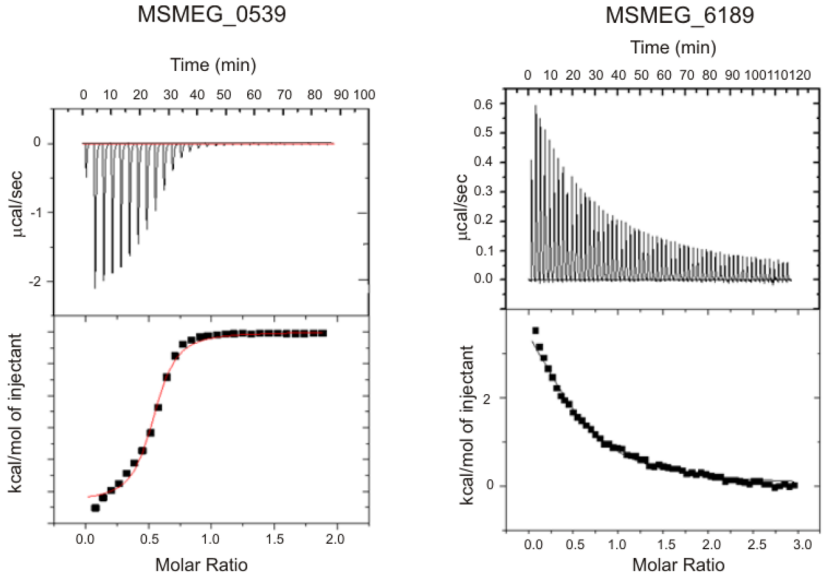
When MSMEG_0539 and MSMEG_6189 were treated with limiting amounts of trypsin, they both underwent proteolysis to generate a low-molecular mass (~16 kDa) peptide (Figure 2a). MSMEG_0539 was slightly more resistant to tryptic cleavage. In the presence of cAMP, however, both CRPs were rendered resistant to proteolytic digestion (Figure 2a). We then monitored this resistance to proteolysis in the presence of varying amounts of cAMP. MSMEG_0539 attained complete protection at 10 $\mu$M cAMP, while MSMEG_6189 required millimolar concentrations of cAMP (Figure 2b). This indicated that MSMEG_0539 had a higher affinity for cAMP. As shown in Figure 2c, protection was provided specifically by cAMP but not by cGMP.

We analyzed the 16 kDa peptide formed postproteolysis by mass spectrometry. In the case of MSMEG_6189, the peptides obtained from the 16 kDa fragment all mapped to the cAMP-binding domain (Figure 2d). However, in MSMEG_0539, peptides from the 16 kDa fragment originated from both the cAMP-binding domain and the DNA-binding domain (Figure 2d). The target residues for trypsin action (R and K amino acid) in the linker/hinge region as well as the flanking regions are well-conserved in both proteins (Figure 1b), suggesting that MSMEG_0539 adopted a conformation that significantly protected a trypsin-sensitive site in the cAMP-free form.

To identify the region of the proteins covered in the 16 kDa fragments and to identify the 14 kDa fragment seen following tryptic cleavage of MSMEG_0539, we subjected the peptide mixture postproteolysis to mass spectrometry. For MSMEG_0539, we obtained peptides with masses of 14317, 14379, and 16502 Da. The 14317 Da fragment mapped entirely to the cAMP-binding domain, while the 14379 Da fragment encompassed the cAMP-binding domain, the linker/hinge, and the DNA-binding domain (Figure 2d). We were able to map two regions of the protein that would generate a tryptic fragment of 16502 ± 3 Da, namely, residues 10−159 and residues 70−215. However, given the presence of peptides EGNSLR (residues 175−180) and GWIR (residues 212−215) in the peptide mass fingerprinting, we identified the 16501 Da fragment as being comprised of residues 70−215, which includes the cAMP-binding domain, the linker/hinge, and the DNA-binding domain.

For MSMEG_6189, we obtained a peptide with a mass of 16630 Da, which mapped to residues 12−162 of the protein (Figure 2d). This fragment originates from the cAMP-binding domain with the C-terminal cleavage site located precisely


dx.doi.org/10.1021/bi500924v | *Biochemistry* 2014, 53, 7765−7776




| Protein | N | Kb ($M^{-1}$) | ΔHb (kcal/mol) | ΔGb (kcal/mol) | TΔS (Kcal/mol/K) |
|---|---|---|---|---|---|
| MSMEG_0539 | 0.75 ± 0.02 | 3.2 ± 0.6 x $10^5$ | -3.4 ± 0.97 | -7.4 | 4 |
| MSMEG_6189 | 1.03±0.2 | 3.8 ± 2.4 x $10^4$ | 2.4 ± 0.7 | -6.1 | 8.5 |

**Figure 3.** Analysis of cAMP binding by ITC. The top panel shows the raw data and the bottom panel a nonlinear least-square fit of the heat released as a function of added ligand for the titration (binding isotherms). The data were fit to a single-site model. The table summarizes the thermodynamic binding parameters, and values are means ± the standard deviation of assays repeated thrice.

between the cAMP-binding domain and the DNA-binding domain, within the linker.

**Thermodynamic Parameters of Binding of cAMP to MSMEG_6189 and MSMEG_0539.** Proteolytic cleavage data demonstrated that the two CRPs from *M. smegmatis* differ in their affinity for cAMP as well as their conformation. To verify if these differences were reflected in changes in biochemical properties, we measured binding of cAMP to the individual proteins by ITC. Representative titrations are shown in Figure 3, and for both proteins, the data yielded good nonlinear least-squares fits to a single set of identical binding sites. Moreover, in contrast to the *E. coli* CRP dimer, where the two binding sites in a dimer show negative cooperativity, the two cAMP-binding sites in mycobacterial CRPs do not communicate with each other allosterically. A similar observation was made previously for Rv3676.[18]

MSMEG_6189 bound cAMP with moderate affinity ($K_d$ ~ 30 $\mu$M), with binding of cAMP being an endothermic reaction ($\Delta G_b$ ~ −6.1 kcal/mol) and with a positive binding enthalpy ($\Delta H_b$ ~ 2.4 kcal/mol). Therefore, the binding of cAMP appeared to be entropically driven and dominated by hydrophobic interactions. These results are similar to those reported earlier for Rv3676.[18] In contrast, we observed a higher binding affinity of cAMP for MSMEG_0539 ($K_d$ ~ 3 $\mu$M), and the binding of cAMP was an exothermic reaction ($\Delta H_b$ ~ −3.4 kcal/mol), suggesting that favorable hydrogen bonds along with hydrophobic interactions contributed to cAMP binding. Therefore, these two proteins, despite their overall similarity, appear to use distinct mechanisms to interact with cAMP, and perhaps DNA, as well.

**DNA Binding Ability of *M. smegmatis* CRPs.** The evidence provided so far suggests that despite the extensive similarity between MSMEG_6189 and MSMEG_0539, the two proteins differ in their biochemical properties, which could translate to changes in their DNA binding properties. We have previously characterized an adenylyl cyclase from *M. smegmatis*, MSMEG_3780, the activity of which contributes to levels of cAMP observed in the cell.[8] Sequence analysis of the region between *msmeg_3780* and the divergent *msmeg_3781* genes revealed the presence of a consensus sequence for CRP [$C_{/T}$ $GTG_{/C}$ $A_{/G}$ $N_6$ $T_{/C}$ $C_{/G}$ $ACG_{/A}$ (Figure 4a)]. We therefore proceeded to test if MSMEG_6189 and/or MSMEG_0539 bound to this palindromic sequence. Gel shift assays were performed using an oligonucleotide containing the putative CRP-binding site in the presence of cAMP. MSMEG_6189 bound the DNA efficiently at low protein concentrations, while the binding of MSMEG_0539 to the probe was barely detected in the EMSA (Figure 4b). The specificity of the CRP band shift was tested by adding excess unlabeled oligonucleotide, which inhibited interaction with the radiolabeled probe. A mutant CRP site oligonucleotide, where core residues critical for CRP binding were altered, did not inhibit binding of MSMEG_6189 to the radiolabeled probe (Figure 4b).

EMSAs were performed with purified MSMEG_0539 and MSMEG_6189 in the absence of cAMP. MSMEG_0539 showed no shift in the absence of cAMP (data not shown). However, significant binding of MSMEG_6189 was detected even in the absence of cAMP (Figure 4c). Thus, some CRPs from mycobacteria have a reduced dependency for cAMP binding to allow DNA interaction.[18]


dx.doi.org/10.1021/bi500924v | *Biochemistry* 2014, 53, 7765−7776

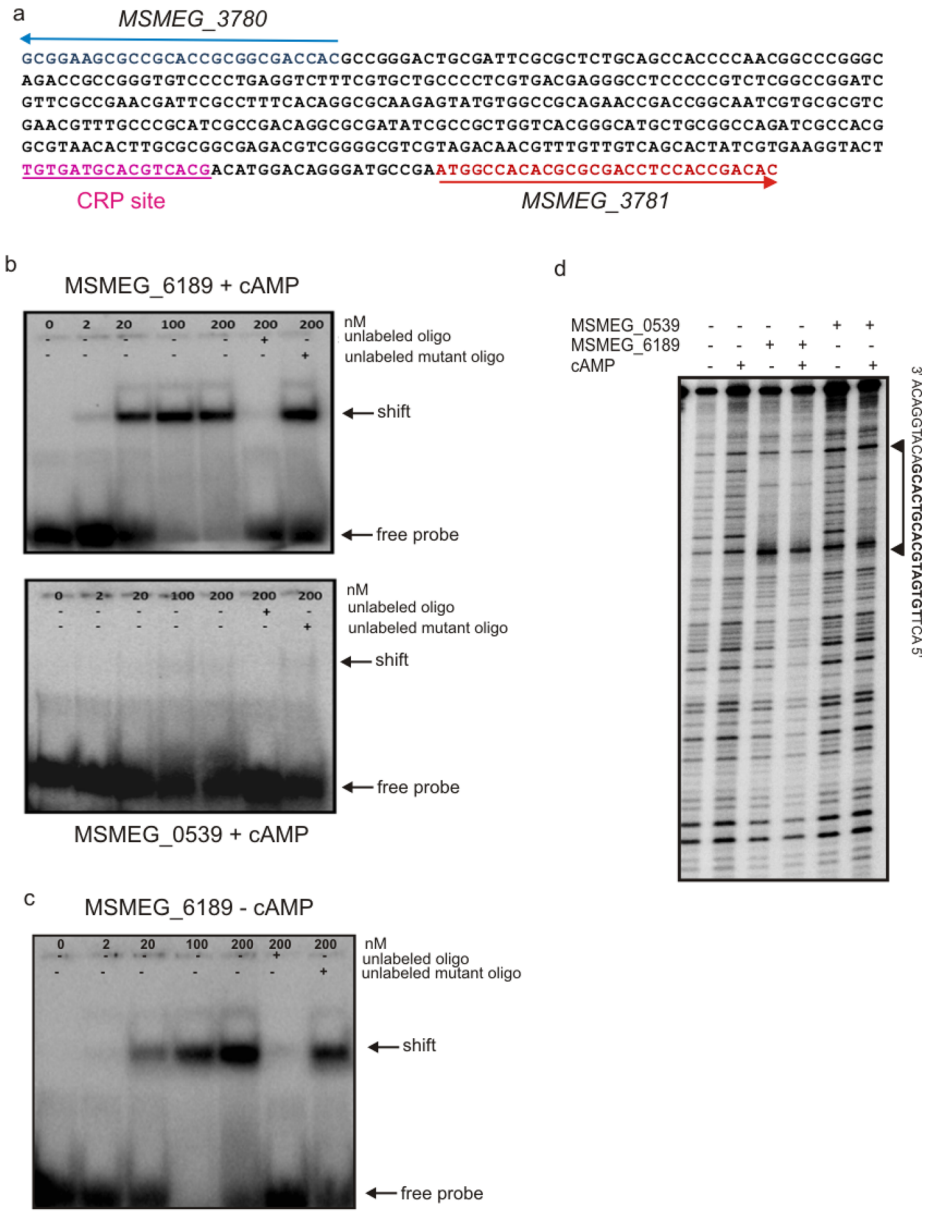



**Figure 4.** DNA binding ability of MSMEG_0539 and MSMEG_6189. (a) Sequence of the intergenic region of MSMEG_3780 and MSMEG_3781. The beginnings of the annotated genes are shown as blue and red letters, and the CRP site is colored pink and underlined. (b) Gel shift assays were performed with the indicated concentrations of the two CRPs in the presence of cAMP (1 mM). Excess unlabeled oligonucleotide and mutant oligonucleotide were used to test the specificity of the DNA−protein complex. Experiments were repeated at least twice, and a representative gel picture is shown. (c) Gel shift assay performed with MSMEG_6189 in the absence of cAMP. Data shown are representative of assays repeated thrice. (d) DNase I footprinting analysis with *M. smegmatis* CRPs in the presence or absence of cAMP. The DNA fragment (0.1 $\mu$M), end-labeled on one strand, was incubated with either MSMEG_0539 or MSMEG_6189 (1 $\mu$M each), in the absence of presence of cAMP (1 mM). The products were analyzed post-DNase I treatment. The protected region was read by the sequencing ladder shown on the left side of the figure. The protected region encompassing the CRP site is shown. Data shown are representative of assays repeated twice.

The interaction of MSMEG_0539 and the CRP site may have been weak and lost during the EMSA. We therefore performed DNase I footprint assays with both MSMEG_6189 and MSMEG_0539, in the presence and absence of cAMP. MSMEG_6189 could robustly protect a region in the DNA fragment in the absence of cAMP, in agreement with the results of the EMSAs (Figure 4d). The protected region encompassed the sequence containing the binding motif for CRP as well as flanking nucleotide sequences. Interestingly, MSMEG_0539 showed a marginal protection in the absence of cAMP and an increased level of protection in its presence. Thus, while both CRPs could recognize the same consensus sequence, MSMEG_0539 shows a significantly enhanced binding to DNA only in the presence of cAMP, in a manner similar to that seen with the *E. coli* CRP. In contrast, MSMEG_6189 was similar to Rv3676, in that it could bind to a CRP site efficiently, even in the absence of cAMP.

**Regulation of Transcription by CRP.** To verify that endogenously expressed CRP(s) could bind to the CRP site, we performed gel shift assays with cytosol prepared from *M. smegmatis* cells harvested during the log phase of growth. As shown in Figure 5a, a complex of DNA and protein was formed, which was abolished when unlabeled oligonucleotide was included in the reaction mixture. Western blot analysis with


dx.doi.org/10.1021/bi500924v | *Biochemistry* 2014, 53, 7765−7776

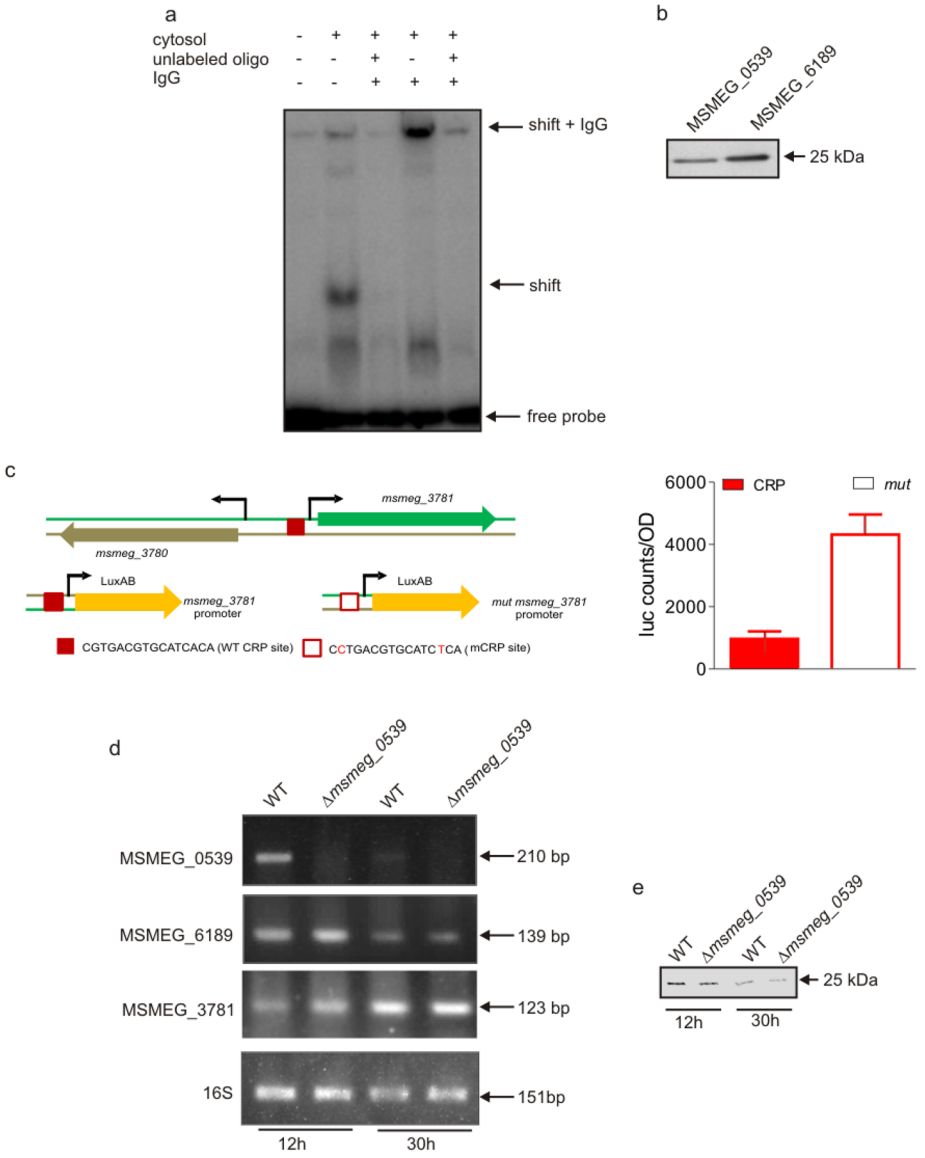



**Figure 5.** Regulation of MSMEG_3781 transcription by CRPs. (a) Gel shift assays were performed using a labeled CRP site oligonucleotide and 10 $\mu$g of *M. smegmatis* cytosolic fraction in the presence of cAMP (1 mM), and in the absence or presence of an unlabeled oligonucleotide. In some reactions, the antibody to Rv3676 (1 $\mu$g) was included. A distinct and specific supershift is seen. (b) Western blot analysis of purified MSMEG_0539 and MSMEG_6189 using antibodies raised against Rv3676. Purified CRPs (30 ng per lane) were subjected to Western blot analysis. (c) The left panel shows a schematic representation of the MSMEG_3780−MSMEG_3781 intergenic region indicating the position of the CRP site. Also shown are the promoter luciferase constructs. Red boxes represent the wild-type CRP sequence (*wt*), while the empty box represents the promoter with mutations in the CRP consensus sequence (*mut*). The right panel shows luciferase activity monitored at 12 h of growth of the wild-type or mutant MSMEG_3781 promoter. Data are means ± the standard error of the mean of assays repeated twice. (d) Steady state mRNA levels of the indicated genes. Reverse transcription PCR was performed on RNA isolated from either the wild type or the Δ*msmeg_0539* strain at the log (12 h) and stationary phase (30 h) of growth. Transcript levels were normalized to 16S rRNA. (e) Western blot analysis of CRP levels in the wild type and Δ*msmeg_0539* strain during the log (12 h) and stationary phase (30 h) of growth.

the CRP antibody (raised against Rv3676) using purified MSMEG_0539 and MSMEG_6189 showed that the antibody could react with both proteins, albeit with lower reactivity with MSMEG_0539 (Figure 5b). We therefore performed the DNA binding assays in the presence of the antibody and observed that the DNA−protein complex was supershifted by antibodies to CRP (Figure 5a), indicating that either of the CRPs or both could bind to the intergenic region of *msmeg_3780 msmeg_3781* and regulate transcription.

We have previously characterized the MSMEG_3780 promoter and demonstrated promoter activity in a fragment of DNA that was downstream of the annotated start site.[8] This 250 bp fragment did not contain the CRP site, suggesting that transcription of MSMEG_3780 may be independent of CRP binding. This information, coupled with the proximity of the CRP site to the predicted start site of *msmeg_3781*, suggested that CRP may regulate *msmeg_3781* gene expression. We therefore cloned the *msmeg_3780_msmeg_3781* intergenic region upstream of luciferase in an orientation that would represent the direction of *msmeg_3781* transcription. We also generated a mutation of the CRP-binding site in this putative promoter [designated *mut* (Figures 4b and 5c)] that would abolish CRP binding. We electroporated these reporter plasmids into wild-type *M. smegmatis* and monitored luciferase


dx.doi.org/10.1021/bi500924v | *Biochemistry* 2014, 53, 7765−7776

activity during the log phase of growth. As seen in Figure 5d, *msmeg_3781* promoter activity was enhanced following mutation of the CRP site, indicating that binding of CRP (either MSMEG_0589 or MSMEG_6189) repressed transcription of MSMEG_3781.

To determine which CRP regulated *msmeg_3781* expression, we attempted to generate *M. smegmatis* strains from which either *msmeg_0539* or *MSMEG_6189* had been deleted. We were unable to obtain a Δ*MSMEG_6189* strain but readily obtained a Δ*msmeg_0539* strain. RTPCR confirmed the absence of the *msmeg_0539* transcript in the knockout strain (Figure 5d). No appreciable difference in the *MSMEG_6189* transcript in the wild-type and knockout strains was seen (Figure 5d). We then measured transcript levels of *msmeg_3781* in the Δ*msmeg_0539* strain, thus elucidating the role of MSMEG_6189 in regulating *msmeg_3781* gene expression. As seen in Figure 5d, levels of *msmeg_3781* mRNA were similar in both the wild type and the Δ*msmeg_0539* strain, indicating that MSMEG_6189 is the sole repressing CRP that could suppress transcription of the *msmeg_3781* gene.

As seen in Figure 5d, the transcript levels of *msmeg_0539* and *msmeg_6189* were reduced following growth for 30 h (stationary phase). This resulted in lower CRP levels in comparison with the log phase of growth, as revealed by Western blot analysis (Figure 5e). Interestingly, levels of the *msmeg_3781* transcript increased at 30 h, suggesting that a reduced level of expression of MSMEG_6189 may contribute to the relief of the transcriptional repression observed during the log phase of growth, when MSMEG_6189 levels are high.

## DISCUSSION

We believe that this is the first study to extensively characterize two paralogous CRPs in the genome of an organism. We show that the presence of multiple CRPs in an organism may not just represent a simple case of gene redundancy but reflect the nuances in gene regulation that an organism requires for its survival. In a recent survey of CRP-like genes in Cyanobacteria, it was noted that *Synechocystis* spp. contain two CRP-like genes, which are ∼65% similar.[35] However, in the PCC 6803 strain of *Synechocystis*, one of the two CRP-like genes (*sll1924*) lacks several residues that are important for cAMP binding and did not bind to cAMP.[36] To the best of our knowledge, the existence of two functional CRPs in species of mycobacteria is unique and must cater to the requirements of these organisms in modulating gene expression in very specific ways.

While we have not directly shown that MSMEG_6189 is essential, the fact that a Δ*msmeg_6189* deletion strain was not readily obtained suggests that this protein must have a critical role to play in *M. smegmatis*. Interestingly, deletion of the ortholog of this gene, *rv3676*, from *M. tuberculosis* was achieved, though the strain grew poorly *in vitro*, as a result of dysregulation of serine biosynthesis.[37] It is therefore clear that MSMEG_6189 must play a critical role in the regulation of a number of genes required for cell viability, of which one of them may be *msmeg_3781*. In contrast MSMEG_0539 appears to be nonessential, at least under laboratory conditions tested here. We have provided evidence that transcripts of *msmeg_0539* can be detected in the cell under the experimental conditions we have used in the study. However, in the absence of an antibody that specifically recognizes only MSMEG_0539, or a strain from which *msmeg_6189* has been deleted, we do not have any direct evidence to indicate whether MSMEG_0539 protein is produced under these conditions.

The almost identical sequences of the HTH domains of MSMEG_0539 and MSMEG_6189 indicate that they could bind to similar regulatory sequences. Both proteins are expressed maximally during the log phase of growth, which is also the time that cAMP levels are the highest in *M. smegmatis*[8] (Figure 5d,e). We have shown here that both proteins can indeed interact with the CRP site upstream of *msmeg_3781*, but significant DNA binding by MSMEG_0539 requires the presence of cAMP and is generally less robust than that of MSMEG_6189 (Figure 4d). Transcript levels of *msmeg_6189* are not altered in the Δ*msmeg_0539* strain (Figure 5d), suggesting that these two proteins may regulate a diverse set of genes, and expression of one may not compensate for that of the other. Therefore, despite the high level of conservation of the HTH domains in these proteins, functional DNA binding and gene regulation appear to be modulated by subtle but distinct features of the two proteins.

Conformational dissimilarities between the two CRPs are reflected in differential trypsin sensitivity. MSMEG_0539 is far less susceptible to tryptic cleavage in the absence of cAMP and readily attains a structure that protects protease-sensitive sites at low cAMP concentrations (Figure 2a,b). Therefore, during the log phase of growth when cAMP levels are high,[8] it is conceivable that almost all MSMEG_0539 present in the cell is bound by cAMP and therefore competent in binding to its target sites in the genome. It appears that MSMEG_0539 does not bind tightly to the CRP site in the *msmeg_3780 msmeg_3781* intergenic region (Figure 4b), in contrast to MSMEG_6189, again underscoring that similarity in the HTH domains is not the sole determinant for high-affinity and sequence specific binding.

The cAMP-binding domains of the individual proteins show dramatic differences in their thermodynamic parameters of cAMP binding (Figure 3). MSMEG_6189 binds cAMP with properties similar to those of Rv3676, demonstrating a comparatively low-affinity interaction with cAMP. MSMEG_0539 has a higher affinity than MSMEG_6189 for cAMP (Figures 2b and 3), and binding appears to be driven by both hydrophobically driven and hydrogen bond-driven interactions. In sharp contrast to that of *E. coli* CRP,[21] binding of cAMP to each site in a dimer of MSMEG_6189 and MSMEG_0539 is independent (Figure 3). Structural information available for Rv3676 reveals a change in conformation following binding of cAMP far less dramatic than that seen in the *E. coli* CRP,[19,38] perhaps accounting for the significant binding to DNA observed even in the absence of cAMP.

The role of mycobacterial CRPs in regulating transcription cannot be generalized. WhiB1 and CRP both bind to the *groEL* promoter and repress transcription.[39] *WhiB1* transcription is itself regulated by Rv3676 at intermediate cAMP concentrations, while at high concentrations, repression is observed.[18] This is suggested to be a consequence of occupation of a downstream site by CRP at high cAMP concentrations, which antagonizes activation from the upstream site by preventing open complex formation. Significant binding to CRP sites is observed for MSMEG_6189 in the absence of cAMP, but binding is enhanced in the presence of cAMP at lower protein concentrations (Figure 4b,c). Our data show that binding of CRP to the *msmeg_3781* promoter represses its transcription during the log phase of growth, when cAMP levels are high (Figure 5c). It is possible that cAMP-free MSMEG_6189 could

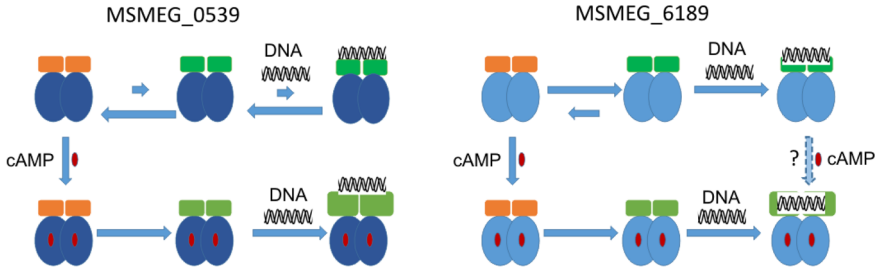



**Figure 6.** Conformational dynamics regulate the divergent properties of MSMEG_0539 and MSMEG_6189. MSMEG_0539 behaves like the *E. coli* CRP in its dependence on cAMP for specific and strong DNA binding. However, binding to the CRP site in the *msmeg_3781* promoter is weak, and the DNA readily dissociates from the protein in the absence or presence of cAMP. MSMEG_6189 equilibrates readily between its inactive and active conformations that can readily bind specific DNA sequences, even in the absence of cAMP. Binding of cAMP to MSMEG_6189 potentiates its DNA binding ability by stabilizing the DNA-binding conformation via conformational selection. It is also possible that post-DNA binding, MSMEG_6189 can bind cAMP, which stabilizes the interaction (indicated by a question mark in the left panel). The two CRPs are depicted in different shades of blue. The active conformation of the HTH region in the CRPs is colored green, while the inactive conformation is colored orange. The tightness of binding DNA is indicated by its position (either partially associated with the HTH or contained within the HTH).

bind to a distinct set of sites when cAMP are levels are low, such as in the stationary phase of growth or during conditions of nutrition starvation.[8]

We have studied a single CRP site in this study. It is therefore conceivable that the sequence of a high-affinity binding site for MSMEG_0539 may differ slightly from the one described here, which seems to show high-affinity binding to MSMEG_6189. This raises the possibility that optimal recognition sequences for MSMEG_0539 may be low-affinity sites for MSMEG_6189, and vice versa. Recently, a binding site for MSMEG_0539 was identified upstream of the cytochrome *bd* gene in *M. smegmatis*,[40] wherein seven of nine bases are identical between the *bd* site and the one we have identified in this study. Thus, differences in recognition sequences for the two CRPs, if present, are likely to be subtle. Analyses along these lines would be possible via the identification of all binding sites for individual CRPs using chromatin immunoprecipitation with protein specific antibodies. If the two CRPs recognize the same DNA sequence with equal affinity, it is possible that they may compete for binding to the same site on the DNA, given that they are co-expressed during the log phase of growth.

It is interesting to observe that mycobacteria have evolved to encode CRPs that are less dependent on cAMP binding for interaction with DNA.[1] We therefore propose that allostery in these proteins occurs through a process of conformational selection[41] (Figure 6). MSMEG_6189 demonstrates a high degree of conformational flexibility such that it can adopt a conformation that binds DNA even in the absence of cAMP. Binding of cAMP to the apo form stabilizes the DNA-binding conformation, thereby accounting for the slightly enhanced binding to DNA seen in the presence of cAMP. It is also conceivable that cAMP can bind to the MSMEG_6189−DNA complex and further stabilize the interaction. In contrast, MSMEG_0539 has a weak propensity to adopt a DNA-binding conformation in the absence of cAMP. Cyclic AMP binding now induces a conformational change, perhaps reflected in the negative enthalpy seen on binding cAMP that allows the HTH to adopt a conformation that binds to DNA (Figure 6). At least in the case of the CRP site upstream of *msmeg_3781*, this binding is weak and the DNA can readily dissociate.

CRPs from different bacteria have become specialized to serve roles in unique niches, but the study presented here is unique because the two CRP-like proteins in question are present simultaneously in an organism yet are likely to perform different functions. The biochemical differences between these two CRPs and their distinct properties that we have described here appear to translate to their functional roles, thus allowing mycobacterial species to adapt to as well as utilize the high cAMP levels in the cell.

## AUTHOR INFORMATION

### Corresponding Author

*Department of Molecular Reproduction, Development and Genetics, Indian Institute of Science, Bangalore, India 560012. Telephone: +91 80 22932542. Fax: +91 80 2360999. E-mail: sandhya@mrdg.iisc.ernet.in.

### Present Address

[▽]Department of Microbiology and Immunology, Weill Cornell Medical College, 1300 York Avenue, New York, New York, U.S.A.

### Author Contributions

A.Z. and J.G. contributed equally to this study.

### Funding

R.S. and J.G. were supported by fellowships from the Council of Scientific and Industrial Research, Government of India. Financial support from the Department of Biotechnology, Government of India, is also acknowledged (U.V. and S.S.V.). Support was also provided under the Department of Science and Technology FIST and IRHPA Programmes. U.V. and S.S.V. also acknowledge support from J. C. Bose National Professorships.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We thank Ann Terene Thomas for technical assistance and Nishad Matange for useful comments and discussions.

## ABBREVIATIONS

CRP, cyclic AMP receptor protein; FNR, fumarate and nitrate reductase; HTH, helix−turn−helix; ITC, isothermal titration calorimetry; SELEX, systematic evolution of ligands by exponential enrichment.

## REFERENCES

(1) Green, J., Stapleton, M. R., Smith, L. J., Artymiuk, P. J., Kahramanoglou, C., Hunt, D. M., and Buxton, R. S. (2014) Cyclic-AMP and bacterial cyclic-AMP receptor proteins revisited: Adaptation for different ecological niches. *Curr. Opin. Microbiol. 18*, 1−7.

(2) Korner, H., Sofia, H. J., and Zumft, W. G. (2003) Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: Exploiting the metabolic spectrum by controlling alternative gene programs. *FEMS Microbiol. Rev. 27*, 559−592.
(3) Green, J., Scott, C., and Guest, J. R. (2001) Functional versatility in the CRP-FNR superfamily of transcription factors: FNR and FLP. *Adv. Microb. Physiol. 44*, 1−34.
(4) Bhaya, D., Nakasugi, K., Fazeli, F., and Burriesci, M. S. (2006) Phototaxis and impaired motility in adenylyl cyclase and cyclase receptor protein mutants of *Synechocystis* sp. strain PCC 6803. *J. Bacteriol. 188*, 7306−7310.
(5) Yoshimura, H., Yoshihara, S., Okamoto, S., Ikeuchi, M., and Ohmori, M. (2002) A cAMP receptor protein, SYCRP1, is responsible for the cell motility of *Synechocystis* sp. PCC 6803. *Plant Cell Physiol. 43*, 460−463.
(6) Liang, W., Pascual-Montano, A., Silva, A. J., and Benitez, J. A. (2007) The cyclic AMP receptor protein modulates quorum sensing, motility and multiple genes that affect intestinal colonization in *Vibrio cholerae*. *Microbiology 153*, 2964−2975.
(7) Chandler, M. S. (1992) The gene encoding cAMP receptor protein is required for competence development in *Haemophilus influenzae* Rd. *Proc. Natl. Acad. Sci. U.S.A. 89*, 1626−1630.
(8) Dass, B. K., Sharma, R., Shenoy, A. R., Mattoo, R., and Visweswariah, S. S. (2008) Cyclic AMP in mycobacteria: Characterization and functional role of the Rv1647 ortholog in *Mycobacterium smegmatis*. *J. Bacteriol. 190*, 3824−3834.
(9) Shenoy, A. R., and Visweswariah, S. S. (2006) New messages from old messengers: cAMP and mycobacteria. *Trends Microbiol. 14*, 543−550.
(10) Shenoy, A. R., and Visweswariah, S. S. (2006) Mycobacterial adenylyl cyclases: Biochemical diversity and structural plasticity. *FEBS Lett. 580*, 3344−3352.
(11) Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., III, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M. A., Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J. E., Taylor, K., Whitehead, S., and Barrell, B. G. (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature 393*, 537−544.
(12) McCue, L. A., McDonough, K. A., and Lawrence, C. E. (2000) Functional classification of cNMP-binding proteins and nucleotide cyclases with implications for novel regulatory pathways in *Mycobacterium tuberculosis*. *Genome Res. 10*, 204−219.
(13) Gazdik, M. A., Bai, G., Wu, Y., and McDonough, K. A. (2009) Rv1675c (cmr) regulates intramacrophage and cyclic AMP-induced gene expression in *Mycobacterium tuberculosis*-complex mycobacteria. *Mol. Microbiol. 71*, 434−448.
(14) Rickman, L., Scott, C., Hunt, D. M., Hutchinson, T., Menendez, M. C., Whalan, R., Hinds, J., Colston, M. J., Green, J., and Buxton, R. S. (2005) A member of the cAMP receptor protein family of transcription regulators in *Mycobacterium tuberculosis* is required for virulence in mice and controls transcription of the rpfA gene coding for a resuscitation promoting factor. *Mol. Microbiol. 56*, 1274−1286.
(15) Russell-Goldman, E., Xu, J., Wang, X., Chan, J., and Tufariello, J. M. (2008) A *Mycobacterium tuberculosis* Rpf double-knockout strain exhibits profound defects in reactivation from chronic tuberculosis and innate immunity phenotypes. *Infect. Immun. 76*, 4269−4281.
(16) Akhter, Y., Yellaboina, S., Farhana, A., Ranjan, A., Ahmed, N., and Hasnain, S. E. (2008) Genome scale portrait of cAMP-receptor protein (CRP) regulons in mycobacteria points to their role in pathogenesis. *Gene 407*, 148−158.
(17) Agarwal, N., Raghunand, T. R., and Bishai, W. R. (2006) Regulation of the expression of whiB1 in *Mycobacterium tuberculosis*: Role of cAMP receptor protein. *Microbiology 152*, 2749−2756.
(18) Stapleton, M., Haq, I., Hunt, D. M., Arnvig, K. B., Artymiuk, P. J., Buxton, R. S., and Green, J. (2010) *Mycobacterium tuberculosis* cAMP receptor protein (Rv3676) differs from the *Escherichia coli* paradigm in its cAMP binding and DNA binding properties and transcription activation properties. *J. Biol. Chem. 285*, 7016−7027.
(19) Reddy, M. C., Palaninathan, S. K., Bruning, J. B., Thurman, C., Smith, D., and Sacchettini, J. C. (2009) Structural insights into the mechanism of the allosteric transitions of *Mycobacterium tuberculosis* cAMP receptor protein. *J. Biol. Chem. 284*, 36581−36591.
(20) Bai, G., McCue, L. A., and McDonough, K. A. (2005) Characterization of *Mycobacterium tuberculosis* Rv3676 (CRPMt), a cyclic AMP receptor protein-like DNA binding protein. *J. Bacteriol. 187*, 7795−7804.
(21) Harman, J. G. (2001) Allosteric regulation of the cAMP receptor protein. *Biochim. Biophys. Acta 1547*, 1−17.
(22) Ebright, R. H., Ebright, Y. W., and Gunasekera, A. (1989) Consensus DNA site for the *Escherichia coli* catabolite gene activator protein (CAP): CAP exhibits a 450-fold higher affinity for the consensus DNA site than for the *E. coli* lac DNA site. *Nucleic Acids Res. 17*, 10295−10305.
(23) Kahramanoglou, C., Cortes, T., Matange, N., Hunt, D. M., Visweswariah, S. S., Young, D. B., and Buxton, R. S. (2014) Genomic mapping of cAMP receptor protein (CRP Mt) in *Mycobacterium tuberculosis*: Relation to transcriptional start sites and the role of CRPMt as a transcription factor. *Nucleic Acids Res. 42*, 8320−8329.
(24) Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics 23*, 2947−2948.
(25) Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol. 24*, 1596−1599.
(26) Shah, S., and Peterkofsky, A. (1991) Characterization and generation of *Escherichia coli* adenylate cyclase deletion mutants. *J. Bacteriol. 173*, 3238−3242.
(27) Zor, T., and Selinger, Z. (1996) Linearization of the Bradford protein assay increases its sensitivity: Theoretical and experimental studies. *Anal. Biochem. 236*, 302−308.
(28) Nambi, S., Gupta, K., Bhattacharyya, M., Ramakrishnan, P., Ravikumar, V., Siddiqui, N., Thomas, A. T., and Visweswariah, S. S. (2013) Cyclic AMP-dependent protein lysine acylation in mycobacteria regulates fatty acid and propionate metabolism. *J. Biol. Chem. 288*, 14114−14124.
(29) Gattiker, A., Bienvenut, W. V., Bairoch, A., and Gasteiger, E. (2002) FindPept, a tool to identify unmatched masses in peptide mass fingerprinting protein identification. *Proteomics 2*, 1435−1444.
(30) Nambi, S., Basu, N., and Visweswariah, S. S. (2010) cAMP-regulated protein lysine acetylases in mycobacteria. *J. Biol. Chem. 285*, 24313−24323.
(31) Parish, T., and Stoker, N. G. (2000) Use of a flexible cassette method to generate a double unmarked *Mycobacterium tuberculosis* tlyA plcABC mutant by gene replacement. *Microbiology 146* (Part 8), 1969−1975.
(32) Eilen, E., Pampeno, C., and Krakow, J. S. (1978) Production and properties of the $\alpha$ core derived from the cyclic adenosine monophosphate receptor protein of *Escherichia coli*. *Biochemistry 17*, 2469−2473.
(33) Krakow, J. S., and Pastan, I. (1973) Cyclic adenosine monophosphate receptor: Loss of cAMP-dependent DNA binding activity after proteolysis in the presence of cyclic adenosine monophosphate. *Proc. Natl. Acad. Sci. U.S.A. 70*, 2529−2533.
(34) Ebright, R. H., Le Grice, S. F., Miller, J. P., and Krakow, J. S. (1985) Analogs of cyclic AMP that elicit the biochemically defined conformational change in catabolite gene activator protein (CAP) but do not stimulate binding to DNA. *J. Mol. Biol. 182*, 91−107.
(35) Ochoa de Alda, J. A., and Houmard, J. (2000) Genomic survey of cAMP and cGMP signalling components in the cyanobacterium *Synechocystis* PCC 6803. *Microbiology 146* (Part 12), 3183−3194.
(36) Yoshimura, H., Hisabori, T., Yanagisawa, S., and Ohmori, M. (2000) Identification and characterization of a novel cAMP receptor

protein in the cyanobacterium *Synechocystis* sp. PCC 6803. *J. Biol. Chem. 275*, 6241−6245.
(37) Bai, G., Schaak, D. D., Smith, E. A., and McDonough, K. A. (2011) Dysregulation of serine biosynthesis contributes to the growth defect of a *Mycobacterium tuberculosis* crp mutant. *Mol. Microbiol. 82*, 180−198.
(38) Kumar, P., Joshi, D. C., Akif, M., Akhter, Y., Hasnain, S. E., and Mande, S. C. (2010) Mapping conformational transitions in cyclic AMP receptor protein: Crystal structure and normal-mode analysis of *Mycobacterium tuberculosis* apo-cAMP receptor protein. *Biophys. J. 98*, 305−314.
(39) Stapleton, M. R., Smith, L. J., Hunt, D. M., Buxton, R. S., and Green, J. (2012) *Mycobacterium tuberculosis* WhiB1 represses transcription of the essential chaperonin GroEL2. *Tuberculosis 92*, 328−332.
(40) Aung, H. L., Berney, M., and Cook, G. M. (2014) Hypoxia-activated cytochrome bd expression in *Mycobacterium smegmatis* is cyclic AMP receptor protein dependent. *J. Bacteriol. 196*, 3091−3097.
(41) Changeux, J. P. (2012) Allostery and the Monod-Wyman-Changeux model after 50 years. *Annu. Rev. Biophys. 41*, 103−133.