Check for updates

Structural Basis of Dual Specificity of *Sinorhizobium meliloti* Clr, a cAMP and cGMP Receptor Protein

Laura Werel,[a] Neda Farmani,[b,c] Elizaveta Krol,[b,c] Javier Serrania,[b,c] ☉ Lars-Oliver Essen,[a] ☉ Anke Becker[b,c]

[a]Department of Chemistry, Philipps-Universität Marburg, Marburg, Germany
[b]Department of Biology, Philipps-Universität Marburg, Marburg, Germany
[c]Center for Synthetic Microbiology (SYNMIKRO), Philipps-Universität Marburg, Marburg, Germany

Laura Werel and Neda Farmani contributed equally to this article. Author order was drawn by lot.

ABSTRACT In bacteria, the most prevalent receptor proteins of 3′,5′-cyclic AMP (cAMP) and 3′,5′-cyclic GMP (cGMP) are found among transcription factors of the Crp-Fnr superfamily. The prototypic *Escherichia coli* catabolite activator protein (CAP) represents the main Crp cluster of this superfamily and is known to bind cAMP and cGMP but to mediate transcription activation only in its cAMP-bound state. In contrast, both cyclic nucleotides mediate transcription activation by *Sinorhizobium meliloti* Clr, mapping to cluster G of Crp-like proteins. We present crystal structures of Clr-cAMP and Clr-cGMP bound to the core motif of the palindromic Clr DNA binding site (CBS). We show that both cyclic nucleotides shift ternary Clr-cNMP–CBS-DNA complexes (where cNMP is cyclic nucleotide monophosphate) to almost identical active conformations, unlike the situation known for the *E. coli* CAP-cNMP complex. Isothermal titration calorimetry measured similar affinities of cAMP and cGMP binding to Clr in the presence of CBS core motif DNA (equilibrium dissociation constant for cNMP ($K_D{}^{cNMP}$], ∼7 to 11 μM). However, different affinities were determined in the absence of this DNA ($K_D{}^{cGMP}$, ∼24 μM; $K_D{}^{cAMP}$, ∼6 μM). Sequencing of Clr-coimmunoprecipitated DNA as well as electrophoretic mobility shift and promoter-probe assays expanded the list of experimentally proven Clr-regulated promoters and CBS. This comprehensive set of CBS features conserved nucleobases that are consistent with the sequence readout through interactions of Clr amino acid residues with these nucleobases, as revealed by the Clr–cNMP–CBS-DNA crystal structures.

IMPORTANCE Cyclic 3′,5′-AMP (cAMP) and cyclic 3′,5′-GMP (cGMP) are both long known as important nucleotide secondary messengers in eukaryotes. This is also the case for cAMP in prokaryotes, whereas a signaling role for cGMP in this domain of life has been recognized only recently. Catabolite repressor proteins (CRPs) are the most ubiquitous bacterial cAMP receptor proteins. *Escherichia coli* CAP, the prototypic transcription regulator of the main Crp cluster, binds both cyclic mononucleotides, but only the CAP-cAMP complex promotes transcription activation. In contrast, Crp cluster G proteins studied so far are activated by cGMP or by both cAMP and cGMP. Here, we report a structural analysis of the cAMP- and cGMP-activatable cluster G member Clr from *Sinorhizobium meliloti*, how binding of cAMP and cGMP shifts Clr to its active conformation, and the structural basis of its DNA binding site specificity.

KEYWORDS nucleotide second messenger, CRP-like protein family, allosteric control, Clr regulon, root nodule symbiosis

Address correspondence to Anke Becker, anke.becker@synmikro.uni-marburg.de, or Lars-Oliver Essen, essen@chemie.uni-marburg.de.

One of the most ubiquitous nucleotide second messengers in prokaryotes and eukaryotes is 3′,5′-cyclic AMP (cAMP). It was the first second messenger to be described, initially for its role in hormone-dependent signal transduction by eukaryotes

(1, 2) and later for catabolite repression in bacteria like *Escherichia coli* (3). Since then, a myriad of biological processes regulated by cAMP have been reported in bacteria, such as carbon metabolism, biofilm formation, type III secretion, virulence, and symbiosis (4, 5). Another cyclic mononucleotide second messenger widespread in eukaryotes is 3′,5′-cyclic GMP (cGMP) (6). Involvement of cGMP in bacterial regulation has been recognized only recently (7).

In bacteria, cyclic AMP receptor proteins (CRPs) are among the best characterized transcription factors and model for allosteric regulation (3, 8–10). Physiological effects of gene regulation by CRP-like proteins have been characterized in various bacterial species. They are highly versatile, controlling expression of more than 200 genes in *E. coli* and almost as many in *Mycobacterium tuberculosis* and *Corynebacterium glutamicum* (11–15).

CRPs belong to the superfamily of Crp-Fnr transcription regulators, which are composed of an N-terminal nucleotide binding domain and a C-terminal helix-turn-helix (HTH) motif (16). These domains are each conserved in eukaryotes and prokaryotes (4, 17). Upon DNA binding, the C-terminal domain of CRPs is able to interact with the bacterial RNA polymerase (RNAP) and to modulate promoter activation (18–20). Most of the CRPs act as transcription activators, while few act as repressors (6).

A phylogenetic analysis of CRP-like proteins from the Crp-Fnr superfamily grouped these proteins into a main Crp cluster, containing *E. coli* catabolite activator protein (CAP), and in addition suggested two associated CRP-like clusters, named F and G (16). Each of the clusters of CRP-like proteins mostly contains members from a broad distribution of proteobacterial genera, with some exceptions from other bacterial phyla. So far, all structurally characterized CRPs belong to the main Crp cluster, including *E. coli* CAP (21). These proteins are activated by cAMP (16). A prominent member of cluster F is *Synechocystis* sp. SyCRP1, a membrane-bound transcription factor, which is suggested to be released from the membrane through cAMP upon sensing of inorganic carbon (22). CgrA from *Rhodospirillum centenum* and Clr from *Sinorhizobium meliloti*, two functionally studied CRPs, can be assigned to cluster G by applying an approach described previously (16). CgrA appeared to be solely activated by cGMP with very low cAMP-binding affinity, whereas Clr is distinguished from any of the other known CRP homologues by its ability to be activated by both cAMP and cGMP (23, 24). cGMP binding has been described for *E. coli* CAP as well, albeit without stimulating DNA binding (25).

The soil-dwelling alphaproteobacterium *S. meliloti* is capable of fixing atmospheric nitrogen in a symbiotic relationship with leguminous plants of the genera *Medicago*, *Melilotus*, and *Trigonella*. The symbiotic program involves root hair infections and formation of root nodules, harboring the nitrogen-fixing bacteria (26, 27). The *S. meliloti* Rm1021 and Rm2011 genomes encode 13 proteins belonging to the CRP-Fnr superfamily, including the cAMP-independent FixK regulators of microoxic respiration and nitrogen fixation genes (3) and cAMP- and cGMP-controlled Clr (28). Eight of those are homologous to CRP-like proteins. Activation of Clr both by cAMP and cGMP is especially interesting given that these *S. meliloti* genomes contain an exceptionally high number of 28 putative adenylate cyclase/guanylate cyclase (AC/GC) genes (29). Clr and the AC/GCs, CyaD1, CyaD2, and CyaK, are implicated in repression of secondary infections of *Medicago sativa* roots (5). Moreover, Clr overexpression resulted in reduced swimming motility and increased succinoglycan production, which are traits relevant at early stages of the root nodule symbiosis (24).

In our study, we present crystal structures of Clr-cNMP-DNA complexes (where cNMP is cyclic nucleotide monophosphate), which to our knowledge is the first report of a cluster G CRP structural analysis. Our analysis unravels the molecular basis of Clr specificity for cAMP and cGMP. Furthermore, we expand the set of experimentally proven Clr-regulated promoters and Clr DNA binding site (CBS). These feature conserved nucleobases which are in accordance with the sequence readout revealed by the Clr–cNMP–CBS-DNA crystal structures.
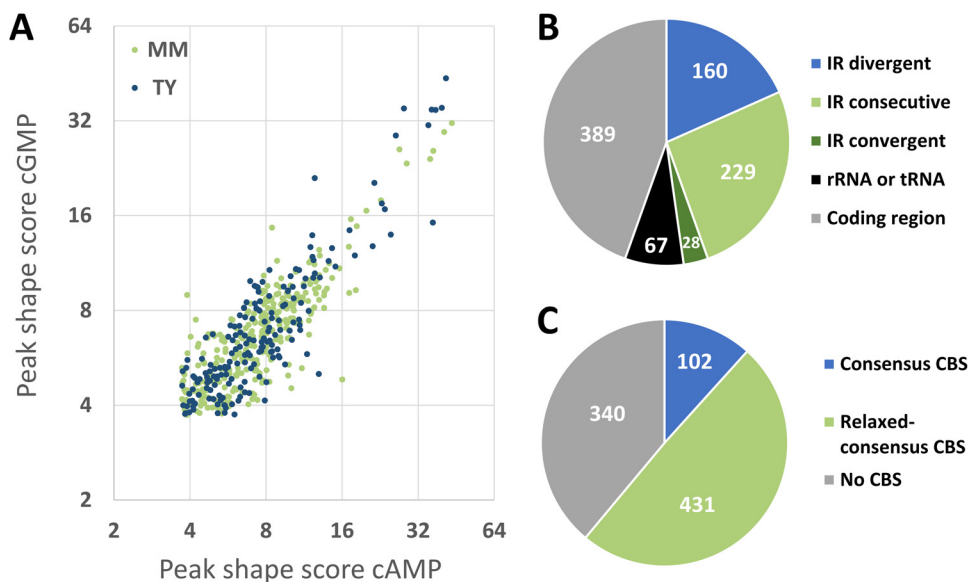
FIG 1 ChIP-seq-assisted screening for Clr-cAMP and Clr-cGMP binding sites. (A) Correlation between peak shape score values of ChIP-seq peaks detected in samples of cells grown in TY or MM medium, both supplemented with either cAMP or cGMP. (B) Location of Clr ChIP-seq peaks within genomic features. IR, intergenic region; IR divergent, IR between two genes transcribed in divergent directions; IR consecutive, IR between two genes transcribed in the same direction; IR convergent, IR between two genes transcribed in converging directions; rRNA or tRNA, rRNA or tRNA encoding region; coding region, protein coding region. (C) Presence of putative Clr binding sites (CBSs) in the vicinity of ChIP-seq peaks. Consensus CBS, GTNNCNNNNGNNAC. Relaxed-consensus CBS, one mismatch with the consensus was allowed.

RESULTS

ChIP-seq-assisted screening for candidate Clr binding sites. To identify Clr-cAMP and Clr-cGMP binding sites genome-wide in the DNA of *S. meliloti* Rm2011, we used chromatin immunoprecipitation coupled with deep sequencing (ChIP-seq). For this purpose, we first generated plasmid pWBT-Clr-CF and introduced it into *S. meliloti* strain Rm2011 Δ*clr*. pWBT-Clr-CF carries IPTG (isopropyl-β-D-thiogalactopyranoside)-inducible *clr-3xflag*, encoding Clr C-terminally fused to a 3×FLAG tag. *In vivo* functionality of Clr-3×FLAG was confirmed by its ability to activate plasmid-borne *egfp* reporter gene fusions to promoters of the previously identified target genes *SMc02178* and *SMb20906* (24) in Rm2011 Δ*clr* in the presence of externally added cAMP (Fig. S1). For anti-FLAG antibody-mediated chromatin immunoprecipitation, Rm2011 Δ*clr* carrying pWBT-Clr-CF was cultivated in tryptone-yeast extract (TY) complex medium or morpholinepropanesulfonic acid (MOPS)-buffered minimal medium (MM) and supplemented with either cAMP or cGMP. Deep sequencing of DNA enriched for Clr-bound regions and control samples representing sheared DNA isolated from cell lysates prior to immunoprecipitation yielded 3.3 to 3.9 million reads per sample (see Table S1A in the supplemental material). ChIP-seq data were processed using the CLC Genomics Workbench to determine peak locations and peak shape score values, which reflect enrichment of the respective regions (Table S1B to E). Consolidation of our data obtained for the four samples yielded 873 distinct ChIP-enriched genomic regions, further referred to as ChIP peaks (Table S1F). Consistent with our previous observation that both cAMP and cGMP promote Clr binding to DNA (24), 40 out of 50 top-score ChIP-enriched regions were detected in samples from cultures supplemented with either of the two cNMPs (Table S1F). Moreover, peak shape score values correlated well between cAMP and cGMP samples (correlation factors of 0.92 and 0.91 in TY and MM media, respectively) (Fig. 1A).

Out of 417 ChIP peaks identified within intergenic regions, 160 were located between two divergently transcribed genes, 229 between two genes transcribed in the same direction, and 28 between two convergently transcribed genes. The remaining ChIP peaks

were situated within rRNA or tRNA genes and protein coding regions (Fig. 1B). To search for Clr binding sites (CBSs) in ChIP-enriched regions, we extracted 400-bp DNA sequences surrounding the peak centers. Using the FIMO online tool (30), these regions were scanned for the core Clr binding site motif GTNNCNNNNGNNAC. This core motif was derived from the consensus sequence HGTYHCNNNNGRWACA, which represents six experimentally verified CBSs (24). Sequences matching the core motif were found in 102 ChIP-enriched regions, corresponding to a total of 107 putative CBSs, of which 87 were located in intergenic regions (Fig. 1C; Table S1G). A FIMO search with relaxed settings (P value, 0.001) identified additional 693 putative CBSs with one mismatch in the core motif (relaxed-consensus CBSs) in 431 ChIP-enriched regions (Table S1H). Of these, 89 putative CBSs were found in 51 ChIP-enriched regions, which already contained a strict consensus sequence.

As expected, all previously characterized CBS sequences in the promoter regions of *SMc02178*, *SMc00653*, *SMc01136*, *SMc04190*, *cyaF2*, and *SMb20906* (5, 24) were found within the ChIP-enriched regions. Among genes adjacent to ChIP peak-containing intergenic regions, 35 genes were found (Table S1I) that were previously identified in transcriptome experiments to be upregulated upon overexpression of *clr* or adenylate cyclase gene *cyaJ* (24, 31). Twenty-five of these genes contained putative consensus and/or relaxed-consensus CBSs in their upstream noncoding regions. The remaining 10 genes lacked putative CBSs in their vicinity. Remarkably, these 10 genes were only detected as being transcriptionally activated upon *cyaJ*, not *clr*, overexpression.

Clr-cAMP and Clr-cGMP specifically bind DNA target sequences *in vitro*. To characterize putative CBS sequences identified within ChIP-seq-enriched DNA regions and located in upstream noncoding regions, we performed electrophoretic mobility shift assays (EMSAs) for 62 genomic regions (Fig. 2A; Fig. S2A to D). The corresponding DNA fragments that produced a complete electrophoretic mobility shift in the presence of Clr and cAMP were defined as regions with a strong CBS, whereas DNA fragments that were only partially shifted were defined as containing a weak CBS (Fig. 2B). Of the 62 CBS candidates tested, 24 were classified as strong and 24 as weak CBSs, while the remaining failed to mediate a detectable band shift in the EMSAs (Fig. 2A; Fig. S2A to D). Furthermore, Clr-cAMP produced stronger band shifts than Clr-cGMP. Not all DNA regions containing putative CBSs matching the consensus produced strong band shifts, which indicates that nonconserved nucleotides within the CBS and its sequence environment might contribute to Clr-cNMP binding. For strong binding sites, we noted that the left and right halves of the palindrome tended to be enriched for pyrimidine (C/T) and purine (G/A) base nucleotides, respectively (Fig. 2C; Table S1J). Conversely, putative CBSs mediating a weak or no band shift lacked a clear nucleotide preference at the nonconserved positions. Almost all DNA fragments causing a strong band shift included a CBS sequence that matched the core motif. A notable exception is the DNA fragment covering the *SMc05008* promoter region, as this region contains three overlapping putative CBSs, each with one mismatch in the core motif (Fig. 3A).

To verify the CBS location within DNA regions that showed strong Clr-cNMP binding, we exchanged the conserved GT in the left half of the palindrome to CA. This modification dramatically reduced Clr-cNMP binding, thus confirming the predicted CBS within these regions (Fig. S2A to D). Of the three overlapping putative CBSs in the *SMc05008* promoter region, only mutations in conserved nucleotide positions of the overlapping putative second and third CBSs inhibited Clr-cNMP binding (Fig. 3A). Band shift assays with synthetic DNA fragments containing either one of the putative CBSs identified in the *SMc05008* promoter region suggested that only the second putative CBS conferred Clr-cNMP binding (Fig. S2D). This indicates that a C→T transition at position 6 of the CBS core motif is not detrimental *per se* for Clr-cNMP binding to the DNA.

Target promoters are regulated by both Clr-cAMP and Clr-cGMP *in vivo*. All tested CBS candidates were further investigated for Clr-cNMP-dependent promoter activation *in vivo*. In cases where the CBS was located in noncoding regions between two opposingly transcribed genes, both putative promoter regions were tested. Plasmid-borne copies of promoter regions followed by up to 75 bp of the downstream
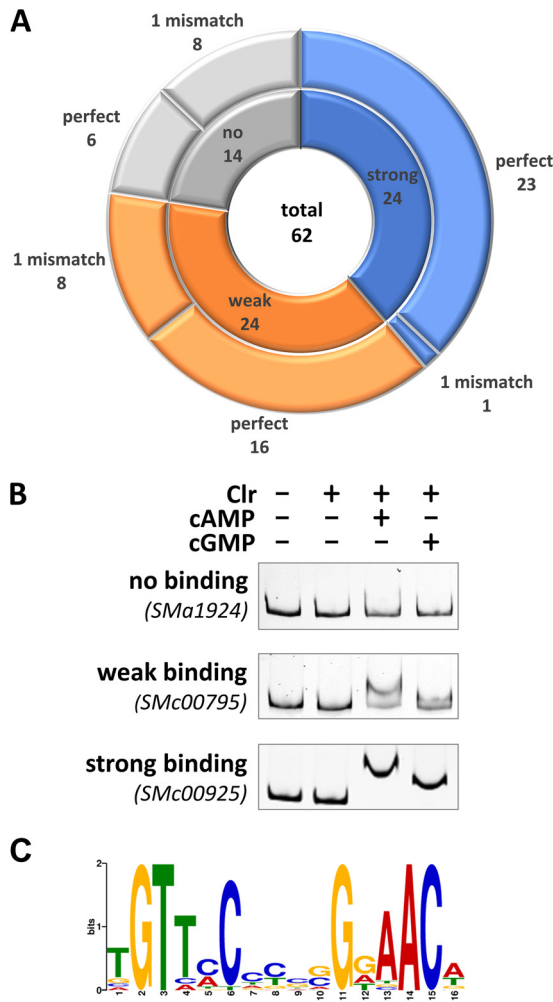
FIG 2 EMSA-based subclassification of CBS. All DNA fragments used for EMSAs contained at least one putative CBS matching the consensus binding site sequence either perfectly or with one mismatch. (A) EMSAs were subclassified into strong (blue), weak (orange), and no binding (gray). Numbers of perfect and imperfect binding sites are indicated for each subclass. (B) Representative EMSAs are shown for each subclass. These were performed with upstream regions of *SMa1924* (no binding), *SMc00795* (weak binding), and *SMc00925* (strong binding), which each contained a perfect-match binding site. (C) MEME logo of 24 Clr binding sites, conferring strong Clr binding in EMSAs.

coding regions or noncoding RNA genes fused to *egfp* were introduced into *S. meliloti* or *E. coli* for promoter activity assays.

Reporter construct-mediated fluorescence was determined in *S. meliloti* Rm2011 wild type upon addition of either cAMP or cGMP (Fig. 4A; Fig. S2E and Table S2). This analysis confirmed Clr-cNMP-dependent activation of 20 promoters by at least 1.5-fold either with cAMP or cGMP added. These included only promoter regions that showed a strong band shift in the EMSAs. To estimate the effect of CBS location relative to the transcription start site (TSS) on promoter activation, we gathered *S. meliloti* TSS positions, determined previously in Rm2011 and Rm1021 wild-type strains (32, 33) and Rm1021 derivative RFF625c lacking all *ecf* sigma factors (34), and determined the CBS-TSS distances (Table S1G and H). For most promoters activated by Clr-cNMPs, CBS-TSS distances of -42 to -76 were found (Fig. S2F). These promoters are thus comparable to CAP-dependent class I and class II promoters in *E. coli* (35–37). In *E. coli*, these promoters are subdivided into three classes, i.e., promoters with a CAP binding site roughly 60 to 90 bp upstream of the TSS (class I), close to the -35 element (class II), and at position -91 or further upstream (class III), with the latter requiring additional
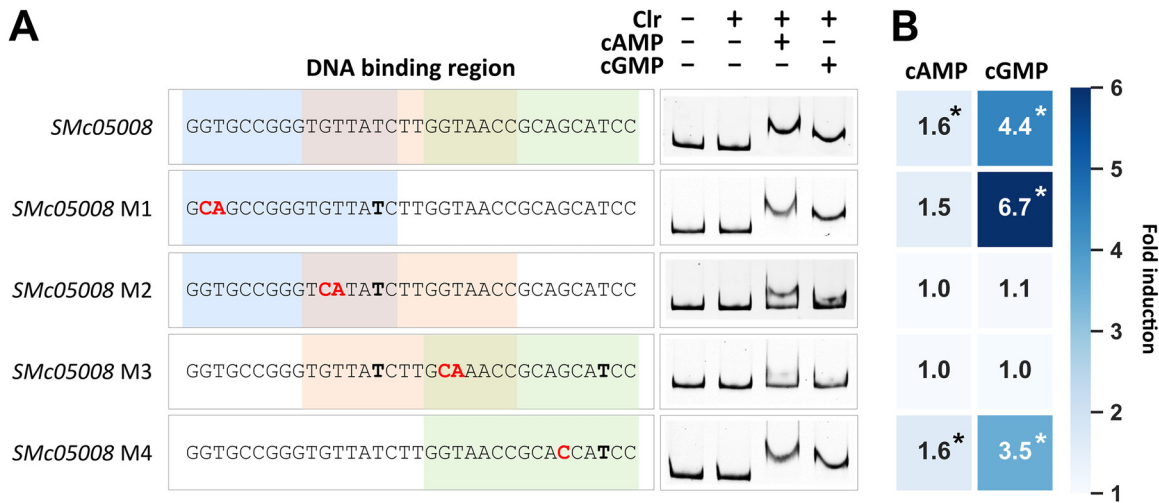
FIG 3 DNA binding of Clr-cAMP and Clr-cGMP at the *SMc05008* promoter. (A) The promoter region of *SMc05008* has three possible DNA binding sites (colored), with one mismatch each. Mutations indicated in red were introduced at four different positions (M1 to M4). Nucleotides not matching the consensus Clr binding sequence are shown in bold. (B) A heat map shows the cAMP- or cGMP-induced fold change of EGFP-mediated fluorescence compared to that of noninducing conditions (no cNMP added). Asterisks indicate significant changes in the EGFP-mediated fluorescence (Table S2).

binding of another transcription regulator closer to the promoter region for activation (35–37).

As expected, GT→CA mutation of the first two nucleotides of the core CBS motif abolished or substantially decreased Clr-cNMP-dependent promoter activation (Fig. S2E and Table S2). Regarding the three overlapping CBSs in the *SMc05008* promoter region, only nucleotide exchanges that diminished Clr-cNMP binding in the EMSA also abolished promoter activation by Clr-cNMP (Fig. 3B).

For the majority of CBS-containing promoter fragments tested, induction by cGMP was up to four times higher than that by cAMP. In total, we identified 11 novel Clr-cNMP-activated genes in addition to nine Clr-cNMP-activated genes reported previously (24, 31). The majority of the 20 confirmed Clr-cNMP-regulated genes encode hypothetical proteins,
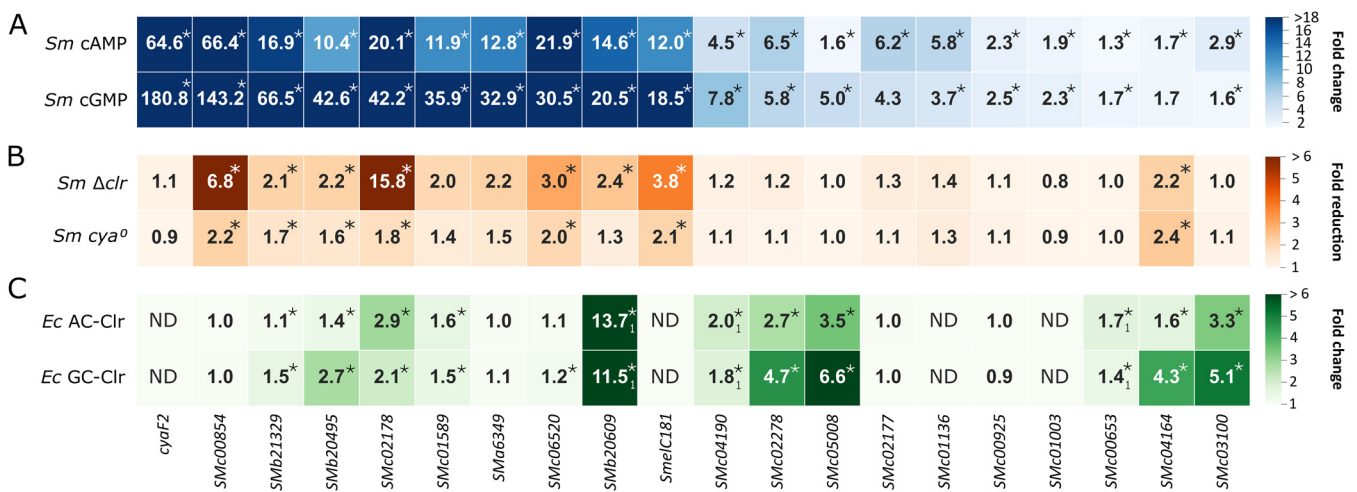

FIG 4 Promoter-probe measurements in different hosts. (A) Fold change of EGFP fluorescence signal mediated by promoter-probe constructs in *S. meliloti* Rm2011 (*Sm*) in cultures supplemented with cAMP or cGMP. Ratios derive from relative fluorescent units (RFU) of cNMP-induced against uninduced cultures. (B) Fold reduction in EGFP fluorescence signal mediated by promoter-probe constructs in *S. meliloti* Rm2011 Δ*clr* and Rm2011 *cya⁰* compared to that of Rm2011 wild type. Asterisks indicate statistically significant reductions in fluorescence ($P < 0.05$) in combination with a wild-type RFU threshold of ≥80. (C) Fold change of EGFP fluorescence signal mediated by promoter-probe constructs in *E. coli* BTH101 (*Ec*) upon induced production of an AC (CyaG1 of *S. meliloti*) or a GC (Cya2 of *Synechocystis* sp.) together with Clr (*S. meliloti*) in LB. Ratios shown derive from fluorescence units of strains under AC/GC-Clr production conditions relative to those of control strains producing the AC or GC but not Clr. Subscript 1 indicates values derived from reference 24. ND, data not determined. Asterisks indicate significant changes in the EGFP-mediated fluorescence (Tables S2 to S4).

and, notably, 8 of the 20 genes encode small proteins of less than 100 amino acids. Ten Clr-cNMP-activated genes encode proteins with an N-terminal signal peptide. Moreover, Clr-cNMP was found to directly induce transcription of *SMelC181* encoding a small non-coding RNA of unknown function (Fig. 4A and C). This RNA was previously identified in a transcriptome sequencing (RNA-seq) study (38).

To evaluate promoter activation by Clr dependent on endogenously produced cNMPs, we compared the promoter-probe constructs conferring enhanced green fluorescent protein (EGFP) fluorescence between Rm2011 and its derivatives lacking either all annotated AC/GC genes (cya^0) or *clr* (Fig. 4B; Table S3). To generate the cya^0 strain, all 28 annotated class III AC/GC-encoding genes (29) were sequentially deleted from the Rm2011 genome. The strains were grown in medium without added cNMPs. Under these conditions, eight and seven of the 20 promoter-probe constructs mediated at least 1.5-fold ($P < 0.05$) reduced fluorescence levels in the Δclr and cya^0 strains, respectively, compared to the wild type (Fig. 4B; Table S3). Lack of the *clr* gene resulted in a much stronger reduction in the promoter activities of *SMc00854*, *SMc02178*, and the sRNA gene *SMelC181* than absence of the 28 AC/GC genes.

Furthermore, Clr-cNMP-mediated transcription activation of, overall, 17 promoter regions with confirmed and putative Clr-binding motifs was analyzed in *E. coli* as a heterologous host (Fig. 4C; Table S4). We used *E. coli* strain BTH101, which is unable to produce cAMP. This strain was equipped with plasmid-borne *clr*, coexpressed with either the *S. meliloti* adenylate cyclase gene *cyaG1* or the *Synechocystis* sp. guanylate cyclase gene *cya2* under the control of an IPTG-inducible promoter. Promoter-reporter constructs were introduced on a second plasmid. We determined previously that in *E. coli* BTH101, cAMP/cGMP production alone was insufficient to activate the *S. meliloti* Clr-cNMP target promoters, and therefore *E. coli* CRP did not replace Clr (24). Out of the 17 promoters tested, eight were activated by Clr-cNMP in *E. coli* by at least 1.5-fold either by cAMP or cGMP production. Including the three previously described candidates, *SMb20906*, *SMc00653*, and *SMc04190* (24), a total of 11 promoters were activated by Clr-cNMP in *E. coli*. Promoters that were not clearly activated by Clr-cNMP in *S. meliloti* Rm2011 also showed no activation in *E. coli* BTH101 (Tables S2 and S4). In contrast, five promoters (*SMc00854*, *SMa6349*, *SMc06520*, *SMc02177*, and *SMc00925*) that were activated in *S. meliloti* Rm2011 were not activated in *E. coli* BTH101, suggesting that these promoters might require additional *S. meliloti*-specific factors for Clr-cNMP-mediated activation.

Clr affinity for cGMP increases in the presence of DNA. As EMSA and promoter activation studies suggested cNMP-inducible binding of CBS-DNA by Clr, we further analyzed the binding behavior of Clr using isothermal titration calorimetry (ITC) (Fig. S3). A synthetic 33-nucleotide DNA sequence was designed based on the core CBS motif as a binding target for Clr. For ITC assays and subsequently for crystallization experiments (see below), this sequence was split into two identical synthetic double-stranded oligonucleotides, which dimerize via their palindromic sticky ends (Fig. 5A). The obtained thermograms are consistent with a simple model assuming one DNA binding site per Clr-cAMP complex. A similar behavior has been described for *E. coli* CAP (18, 39, 40) and *M. tuberculosis* CRP (40). The "active" form of CAP is represented by a CRP dimer-$cAMP_2$ complex (39, 41). The secondary binding site proposed for *E. coli* CAP and *M. tuberculosis* CRP located in the C-terminal HTH DNA binding domain is not conserved in Clr, and no binding could be observed in ITC, hydrogen/deuterium exchange mass spectrometry (HDX-MS), or crystallographic experiments (42). As the affinity for the secondary binding site is considerably above physiological concentrations, its biological relevance has recently been dismissed (9, 41, 43, 44).

The affinity of cAMP binding itself by Clr is mostly not affected by the presence of cognate CBS-DNA with equilibrium dissociation constant (K_D) values of 6 to 7 μM (Table 1). Formation of the ternary Clr–cAMP–CBS-DNA complex is enthalpically driven with an opposed loss of entropy, whereas that of the binary Clr-cAMP complex relies both on enthalpic and entropic contributions. In contrast, cGMP binding by Clr alone proceeds endothermically and with apparently lower affinity (K_D, ~24 μM). Compared to
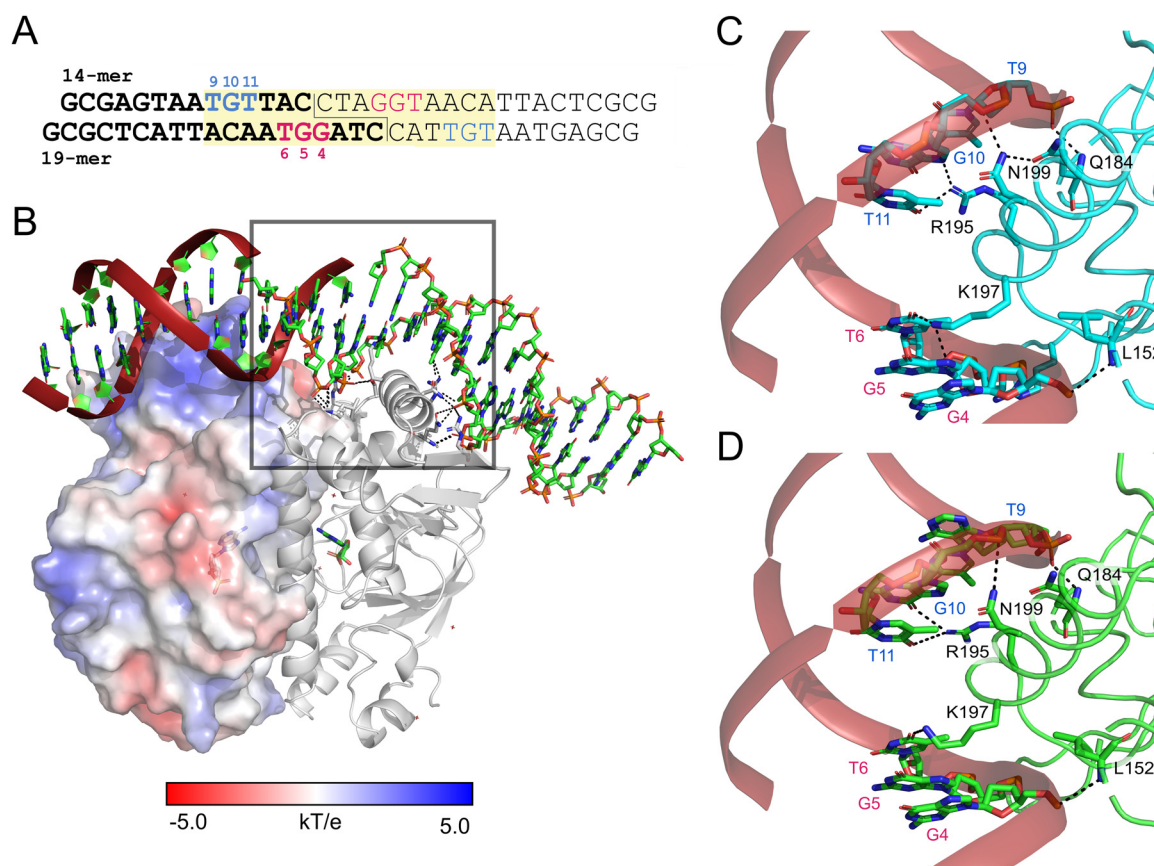
FIG 5 Interaction of Clr with its DNA recognition site. (A) Synthetic DNA fragment containing binding target sequence for Clr, used in ITC and crystallization. It is composed of two 14-mer and two 19-mer oligonucleotides. The two identical double-stranded fragments are shown in bold (left arm) and regular font (right arm). They are joined by a 4-bp sticky end, located right above the subunit interface. The design was based on the consensus motif GTNNCNNNNGNNAC. The 16-bp palindrome sequence containing the Clr binding consensus motif is highlighted in yellow. The nucleotide numbering is in the 5′-to-3′ orientation. (B) The surface of the helix-turn-helix DNA binding domain is positively charged to interact with the phosphate backbone. Helix α8 is inserted into the major groove and provides base readout via direct hydrogen bonding interactions. The staggered single-strand breaks above the subunit interface facilitate DNA kinking of about 40°. The 19-mer oligonucleotides belong to chains C and E, and the 14-mer oligonucleotides belong to chains D and F of the structure, respectively. (C and D) Direct interactions of Clr with the DNA major groove upon cAMP (C) and cGMP (D) binding. The DNA bases involved in the interaction are highly conserved. The color coding and the numbering of the DNA bases correspond to those shown in panel A.

the Clr–cAMP–CBS-DNA complex, ternary complex formation with cGMP gives similar affinities and also relies on both contributions. For Clr alone, i.e., without cAMP or cGMP, no DNA binding could be detected under ITC conditions. In the apo state of *E. coli* CAP, the cAMP and CBS-DNA binding domains appear as quasi-independent dynamic units, while the intra- and intersubunit correlations increased in the holo state (45). Our ITC data indicate a profound difference only for the binary Clr-cNMP complexes, which may affect formation of ensembles capable of binding to CBS-DNA. According to size

TABLE 1 ITC measurements for complex assembly[a]

Cell	Injectant	K_D (µM)	n	ΔH (kcal/mol)	ΔG (kcal/mol)	−TΔS (kcal/mol)
Clr	cAMP	5.8 ± 1.4	1.02 ± 0.016	−2.22 ± 0.05	−7.16 ± 0.12	−4.94 ± 0.22
Clr, DNA	cAMP	6.7 ± 2.7	0.57 ± 0.019	−11.5 ± 0.77	−7.08 ± 0.16	4.42 ± 0.36
DNA	Clr, cAMP	6.0 ± 4.7	0.39 ± 0.201	−11.0 ± 7.30	−7.13 ± 0.10	3.87 ± 1.37
Clr	cGMP	23.8 ± 14.5	1.49 ± 0.081	0.29 ± 0.02	−6.31 ± 0.05	−6.60 ± 0.07
Clr, DNA	cGMP	10.7 ± 3.9	0.57 ± 0.022	−4.91 ± 0.02	−6.79 ± 0.09	−1.88 ± 0.19
DNA	Clr, cGMP	4.5 ± 12.4	0.45 ± 0.408	−3.47 ± 5.33	−7.30 ± 0.02	−3.83 ± 0.02

[a]The measurements were conducted at 25°C. The titrant was injected stepwise from a syringe to the titrand solution (Clr, DNA or Clr-DNA) in the cell. ITC profiles were fit to a model with a single binding site (1:1); values for the equilibrium dissociation constant (KD), the enthalpy (ΔH) and the stoichiometry (n) were determined by curve fitting using Origin 7 software; values for free enthalpy (ΔG) and the entropic contribution (-TΔS) are therefrom calculated by the thermodynamic equations. Data are obtained from the average of three titrations. The corresponding raw data can be found in Fig. S3.

TABLE 2 Data collection and refinement statistics for cNMP-bound Clr-DNA complex crystal structures[a]

Parameter	Value for:	
	Clr-cAMP-DNA complex (7PZA)	Clr-cGMP-DNA complex (7PZB)
Data collection statistics		
X-ray source	DESY P13	SLS PXI
Space group	P2$_1$2$_1$2$_1$	P2$_1$2$_1$2$_1$
Unit cell dimensions		
a, b, c (Å)	50.1, 69.9, 200.4	50.7, 72.9, 200.0
$\alpha = \beta = \gamma$ (°)	90	90
Wavelength (Å)	0.976	1.000
Resolution range (Å)	57.34–2.72 (2.82–2.72)	49.15–3.12 (3.23–3.12)
Completeness (%)	74.89 (6.72)	70.19 (5.53)
No. of observed reflections	137,679 (927)	41,837 (146)
No. of unique reflections	14,734 (129)	9,719 (74)
Multiplicity	9.3 (7.2)	4.3 (2.0)
Wilson B factor (Å^2)	64.80	100.50
R_{merge} (%)	0.145 (1.223)	0.056 (0.481)
R_{meas} (%)	0.154 (1.315)	0.064 (0.621)
Mean $I/\sigma(I)$	11.35 (1.96)	11.08 (1.87)
$CC_{1/2}$[b]	0.994 (0.788)	1.000 (0.610)
Refinement statistics		
Resolution range (Å)	57.34–2.72	49.15–3.12
R_{work}/R_{free} (%)	0.271/0.303	0.289/0.333
Avg B factor (Å^2)	68.16	96.83
No. of atoms		
Total	4,588	4,694
Protein/DNA	4,537	4,627
Ligand	46	46
Solvent	5	3
RMSD		
Bond length (Å)	0.021	0.005
Bond angle (°)	2.34	1.09
Ramachandran plot (%)		
Favored	88.81	88.20
Allowed	6.26	8.45
Outliers	4.92	3.34

[a]Values for the highest-resolution shell are given in parentheses.
[b]$CC_{1/2}$ is the correlation coefficient between two random half data sets.

exclusion chromatography, apo-Clr forms a monomer-dimer equilibrium in solution that favors the monomeric state, with 86%. For comparison, the best studied Clr homolog, *E. coli* CAP, is able to bind both cAMP and cGMP but is only activated by cAMP for DNA binding (46). The lack of activation in response to cGMP has also been found for *M. tuberculosis* Rv3676 and has been assumed to be a generic trait of CRP-like proteins (43).

Both cAMP and cGMP are able to shift Clr to its active conformation. Unlike apo-Clr and its binary complexes, we found that ternary Clr–cNMP–CBS-DNA complexes crystallize within a few hours in the same, orthorhombic crystal form. Given their intrinsic high anisotropic diffraction characteristics, we employed extensive optimization screening to yield crystals diffracting to 2.8 and 3.1 Å for the cAMP- and cGMP-bound Clr–CBS-DNA complex, respectively (Protein Data Bank [PDB] accession codes 7PZA and 7PZB) (Table 2). The structures show a Clr dimer that is bound to CBS-DNA. The individual subunits fold into two domains linked by helix α5 (residues P123 to L149), which is the Clr dimerization interface. The α5 helices of the dimer form a coiled-coil mainly driven by hydrophobic interactions, with T140 and T141 as notable exceptions. Residues M1 to Q121 belong to the cNMP binding domain, the larger of the two. It is formed by four α helices and seven β strands in an antiparallel orientation of a β-sandwich encasing the cNMP binding site. An HTH DNA binding domain consisting of four
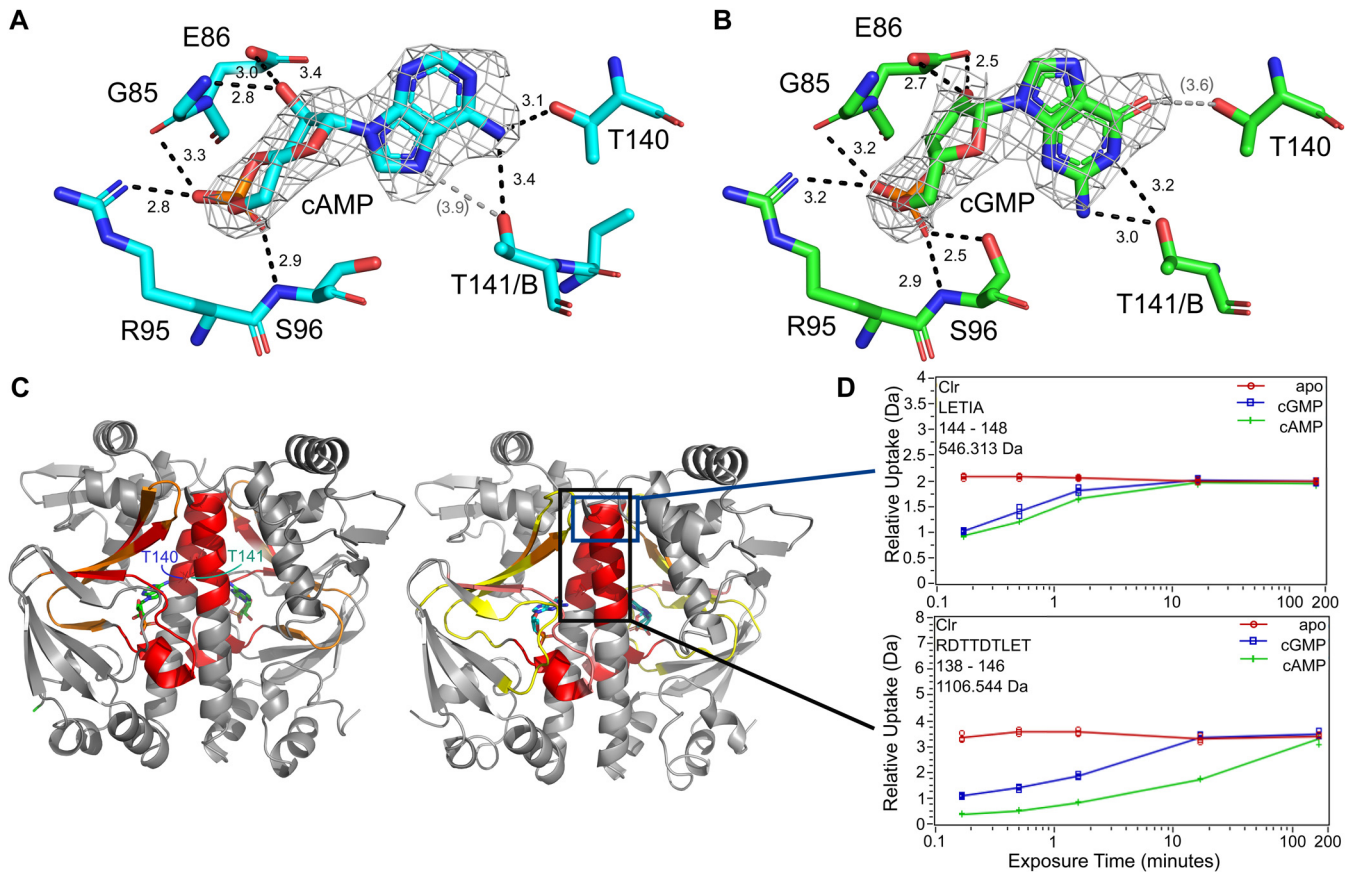
FIG 6 Characteristics of the Clr-ligand interaction. (A and B) Nucleotide binding environment in the cAMP- and cGMP-bound crystal structures. The sugar and phosphate moiety of the ligand interact with G85 and E86 of helix α4, as well as R95 and S96. The side chain of Ser96 is tilted toward the ligand in the cAMP-bound structure, and Thr141 of the adjacent subunit is in closer proximity to it. The primary amine of the adenosine is coordinating Thr140 from α5 in the same subunit and Thr141 from the adjacent one. The ligand 2 mF_o − DF_c experimental electron density maps contoured at 1.3 σ are shown in gray. Predicted hydrogen bonds (black dotted lines) are shown in angstroms; values in parentheses are too distant for indicating proper H-bonds. (C) Areas of reduced relative hydrogen-deuterium exchange upon binding of cAMP (left) and cGMP (right) colored from 5 to 20% in 5% increments (yellow, orange, red). In addition to the expected displacement of solvent in the binding cavity, the C-terminal half of helix 5 (black box) shows a reduction in exchange that can be attributed to a coil-to-helix transformation. (D) Time-dependent deuterium exchange graphs for two key peptides within helix 5 show that the shielding cannot be due solely to interaction with T140 and T141 but is the result of a change in secondary structure. Peptides not including both T140 and T141, exemplified by [144]LETIA (upper box), also show a decrease in hydrogen/deuterium exchange.

α helices and two β strands is located at the C terminus (residues Y150 to D234). The cNMP binding domain is universally conserved throughout cNMP receptor proteins, as well as protein kinases and ion channels, with sequence identities above 20% even for mammal representatives (17). The HTH motif of the DNA binding domain is prevalent in 116 structures so far deposited in the PDB according to InterPro classifications (47, 48). Thus, both domains perform their function as core components of a variety of proteins adapted to fulfill different biological tasks. Accordingly, the overall fold of Clr, which belongs to the more diverse cluster G of the Crp/Fnr superfamily (Fig. S5), is highly similar to that of members of the main Crp cluster, namely, the catabolite activator protein (CAP) from *E. coli* (PDB accession code 1CGP) (49) and even more so to the *M. tuberculosis* (3MZH) and *C. glutamicum* (4CYD) CRPs (50, 51). Accordingly, the latter two served as the templates for molecular replacement.

The CBS-DNA interacts with its symmetry mate in the crystal lattice and is therefore well resolved. When the ternary complexes are compared, the Clr-cNMP structures as well as the CBS-DNA show a high conformational congruity with pairwise root mean square deviation (RMSD) values below 0.6 Å. This further supports our notion that both cyclic nucleotides shift ternary Clr–cNMP–CBS-DNA complexes to almost identical active conformations. The cNMP binding site located in the N-terminal domain is stoichiometrically occupied by cAMP and cGMP, respectively (Fig. 6A and B; Fig. S4). The

binding modes of the phosphate and sugar moieties are identical, involving G85 and E86 within α4 as well as R95 and S96 of the following loop. The side chain of S96 is flipped toward the phosphate only in the cGMP-bound complex. The nucleobase of cGMP is flipped to the *syn* conformation, while it occupies the *anti* conformation in the Clr-cAMP complex. Both nucleobase conformations allow for interaction with helix α5 residue T140 of the same Clr subunit and T141 of the adjacent Clr subunit. While cAMP interacts via its primary N^6-amine with both T140 and T141, cGMP forms hydrogen bonds between T140 and its C-6 carbonyl group as well as between T141 and the N^2-amine. The latter interaction between the N^2-amine of cGMP and T141 of the adjacent molecule is a substantial difference between Clr and the inactive cGMP complex of *E. coli* CAP (52). Interaction of the cAMP nucleotide with both α5 helices has been described to trigger a coil-to-helix transition in the C-terminal part of the hinge and to shift the HTH domain toward its active conformation for CAP (53). So far, cGMP appeared to be incapable of such an interaction with the second helix for promoting CRP activation (46). In the case of Clr, both the interaction and arrangement of the full α5 helix are resolved by the structure, which is in an agreement with Clr-DNA binding *in vitro* and activation of promoter activity *in vivo* by both cNMPs. Moreover, mapping the relative hydrogen-deuterium exchange between apo and holo states by HDX-MS experiments shows that upon nucleotide addition, the C-terminal half of the helix gets shielded from exchange. This further supports the notion that both nucleotides are capable of triggering the coil-to-helix transition in the linker to the C-terminal helix-turn-helix domain (Fig. 6C and D). This shielding is not due only to nucleotide interactions with T140 and T141, as even fragments, containing the last five residues of the α5 helix, show a clear decline in solvent accessibility for nucleotide-bound states. Helix α6 is shortened in response to the hinge rearrangement and interacts with the former at an angle of about 40° in the active conformation (11). This interaction is stabilized in Clr by a formerly unknown chloride-binding site. The relative hydrogen-deuterium exchange comparing the apo and holo states also shows a slightly stronger decline in deuterium uptake within the binding pocket for the cAMP-bound state than for the cGMP-bound state. Some uptake differences within helix α7 and α8 of the HTH domain can be observed (Fig. S4). Upon nucleotide binding, the deuterium exchange increases in this region, indicating some degree of conformational change. This agrees with structures published for *E. coli* CAP, where the active conformation is achieved by a 60° rotation of α8 (46, 54, 55).

The crystal structures of nucleotide-bound Clr shed light on its DNA specificity. Clr binds to the target DNA via electrostatic interactions and hydrogen bonding to the major groove. The surface of the HTH DNA binding domain has a strongly positively charged patch with which it interacts with the DNA phosphate backbone (Fig. 5B). Clr has a likewise positively charged cavity in each subunit for accommodating cNMP. Helix α8 is inserted in the major groove and interacts with hydrogen bonds over six bases, as depicted in Fig. 5C and D. The binding results in a 40° and a 32° kinking of the DNA helix at the longer and shorter DNA oligomers, respectively, which is eased by the engineered strand breaks at the 2-fold symmetry center. The entropic cost associated with DNA bending is overcompensated by establishing further interactions. As a result, the flexibility of the flanking regions of the recognition sequence might be distinguished by Clr, adding another layer of specificity (56, 57).

The electron density of the duplex DNA that was annealed from four oligonucleotides corresponds to the 14 bp of the 5′ oligonucleotides of the core as well as 19 bp of the two 3′ oligonucleotides. The base readout occurs at most conserved positions of the core CBS motif (Fig. 5A) (24). Q184 of α7 and R195 and N199 of α8 are interacting with the [9]TGT triplet of the 14-mers, whereas K197 and S194 of the same helix, as well as L152 within the loop next to the C-terminus of α5 are hydrogen bridged to the [4]GGT triplet of the 19-mer. For *E. coli* CAP, mainly R180, E181, and R185 have been identified to interact with the recognition sequence (58, 59). In Clr, R180 corresponds to R195, E181 is not conserved, and R185 is conserved as R200 but fails to interact with

the DNA as it adopts a different rotamer. Accordingly, proteins in the phylogenetic Crp main cluster, like *E. coli* CAP, have a highly conserved RE(T/M)VGR motif in their α8 helix, whereas members of cluster G, to which Clr belongs, are significantly more diverse in this region (16). Sequence differences between the main Crp cluster and the G cluster are not restricted to the DNA binding site but also occur in the motifs surrounding the cNMP-binding site, including the hinge/α5 region and the α6 helix. The higher sequence variation found for members of cluster G than for the Crp main cluster (Fig. S5) may correlate with the lower specificity for activation by cyclic nucleotides.

CRP proteins interact with the bacterial RNAP via the stacked β-strands within the HTH domain that has been termed AR1 in class I promoters (58, 60) or the top of the β-barrel within the cNMP binding domain, which has been labeled AR2, in class II promoters (61, 62). For Clr AR1, three of the four β-strands are present, whereas the N-terminal loop of the motif is turned more toward the protein core (Fig. S4). Crystal structures of CAP in complex with DNA and the carboxyl-terminal domain (αCTD) of the alpha-subunit of RNAP show very little flexibility within the interface and surprisingly only one hydrogen bond between T158 of CAP and T285 of the αCTD (PDB accession codes 5CIZ, 3N4M, and 1LB2). The interface is largely hydrophobic, a feature that is retained in Clr. The activating region AR1 has a significantly divergent orientation in its N-terminal part for Clr. This does prevent interaction with the *E. coli* RNAP *in vivo*, which we infer from promoter activation by Clr in *E. coli* (see above). The overall conservation of the 22 residues comprising AR1 is rather low, with three identical and three similar amino acids. In class II promoters, the RNAP interacts with the CRP on two additional sites: AR1 of the second dimer symmetry unit and AR2 of this unit (3). The interactions with AR2 involve the beta and omega subunits as shown in crystal and cryo-electron microscopy (cryo-EM) structures from *Thermus thermophilus* and *E. coli* (62, 63). The interactions of Clr within the omega subunit are conserved as G47 and D48 (G14 and E15 in *T. thermophilus*), whereas the area involved in the interaction with the beta subunit in *T. thermophilus* (mainly through E8) deviates in a way that allows no deductions on whether it would interact in the same way.

DISCUSSION

Our study was focused on the dual specificity of Clr for cAMP and cGMP. We confirmed these binding specificities of Clr *in vitro* and characterized the binding properties of this transcription regulator for both these cNMPs and for its DNA binding site. In our ITC measurements, cAMP bound to Clr with about four times higher affinity than did cGMP ($K_D{}^{cGMP}$ = 24 μM; $K_D{}^{cAMP}$ = 6 μM). For CAP from *E. coli*, values in the same range have been reported, the affinity for cAMP (10 μM) also being greater than that for cGMP, at 16 μM (9, 42). However, we showed that the difference between the nucleotide affinities of Clr diminished as soon as cognate DNA was present, showing that ternary complex formation by both effector molecules is similarly driven.

The regulation of cAMP receptor proteins depends on an equilibrium between an active form that is capable of DNA interaction and an inactive form that is not. The equilibrium is shifted toward the active conformation by effector binding (18). For Clr, both our crystal structures and the HDX experiments clearly prove that Clr undergoes this switch independently of the purine base of cNMP. As the cNMP binding site does not deviate significantly from that of the *E. coli* homologue, Clr might have a lower energy barrier for the shift in population toward the active conformation. This is further supported by similar equilibrium shifts that have been introduced via mutations of *E. coli* CAP before (64). cAMP binding initiates the population shift by interacting with residues of the central helices of both subunits (CAP:T127 and CAP:S128) (65). This in turn leads to a coil-to-helix transition in the C-terminal part of the helix, resulting in a DNA binding domain in its active conformation (11, 54, 66, 67). In contrast to any previous findings in *E. coli*, we clearly established that Clr performs an identical binding movement via a ^{140}TT motif for both cNMPs. As cGMP is bound in a *syn* conformation, the interactions can be retained. We were able to show the resulting coil-to-helix

switch of the α5 helix by HDX-MS in solution due to a decline in deuterium uptake, further proving the dual cNMP specificity nature of Clr. Clr residues G85, E86, R95, and S96, which are also part of the binding pocket, are largely conserved in CAP as well and responsible for the high-affinity nucleotide binding (68).

The C-terminal HTH motif of CRPs recognizes its cognate consensus DNA motif via major groove readout (45). Binding of the Clr dimer to two symmetrically related motifs results in a kink of the DNA by 40° and 32°. For CAP, DNA kinks of even 110° have been reported and modified promoters containing an inherent kink have been shown to function independently of CRP, which highlights the importance of this particular feature (49, 59, 69–71). Our ITC measurements of DNA binding gave dissociation constants of 6.0 and 4.5 μM and free energy changes of -7.1 and -7.3 kcal/mol for cAMP and cGMP, respectively. This makes Clr a transcription factor on the weaker side of the binding spectrum. However, stronger does not always equal better, as the binding needs to be reversible as well (56). For duplex DNA of the similar-length CAP dimer·cAMP$_2$, binding affinities of about 10 μM have been reported, which suggests similar binding strengths for CAP and Clr (71).

Clr interacts with the nucleobases of the DNA core motif via hydrogen bonds. The residues involved in sequence readout are exactly contacting the bases within the CBSs as derived from genome-wide ChIP-seq and EMSA data of this work. The dual specificity of Clr for cAMP and cGMP was demonstrated *in vivo* as well. Virtually all promoters directly targeted by Clr were activated by Clr-cAMP and Clr-cGMP in both *S. meliloti* and *E. coli*. The arrangement of three overlapping imperfect CBS core motifs in the *SMc05008* upstream region is intriguing. It suggests a binding mode different from the binding of a single CBS or that differential binding of the individual sites might generate a different regulatory output. In *S. meliloti*, we found that externally added cGMP activated gene expression up to four times higher than did cAMP. Both the affinities and structures of cNMP-bound Clr showed no clear distinction, explaining the higher degree of cGMP-dependent promoter activation in *S. meliloti*. Moreover, stronger cGMP-mediated activation was not observed in *E. coli*. The difference is therefore likely rooted in higher cGMP levels within the *S. meliloti* cell under the *in vivo* assay conditions used in this study. Whether this is caused by endogenous AC/GCs, differences in uptake of exogenous cNMP, or different cAMP and cGMP degradation rates remains to be elucidated. Previously, intracellular cGMP was detected in *S. meliloti* wild-type cultures, but levels of cGMP were 20-fold lower than levels of cAMP (5). This cGMP may originate from side activities of enzymes with primary AC activity, since no dominant GC activity has yet been found for any of the putative AC/GC enzymes in *S. meliloti*. Small amounts of cGMP were detected in the culture supernatant of an *S. meliloti* CyaB overproduction strain (24). Which cNMP molecules are synthesized under native conditions and at what levels are not yet known, as concentrations seem to be rather low.

Except for *cyaF2*, which encodes a putative AC or GC, all other identified genes directly regulated by Clr-cNMP code for hypothetical proteins. Eight of the Clr-cNMP-regulated genes encode small proteins of up to 100 amino acids, six of these with an N-terminal signal peptide. *SMb20495*, *SMc02177*, and *SMc02178* predicted to encode hypothetical proteins were previously shown to participate in autoregulation of alfalfa root hair infections (5, 31, 72). With the exception of these three genes, the biological role of the other identified Clr-cNMP targets remains unclear. Because many of the directly Clr-cNMP controlled target genes encode proteins with an N-terminal signal peptide, we speculate that most of these proteins might be relevant for extracellular functions and possibly involved in the symbiotic interaction with the host plant.

Notably, Clr-cNMP was found to directly induce transcription of the small noncoding RNA *SMelC181* of unknown function. Only a few studies reported regulation of small noncoding RNAs by a CRP. Examples are small noncoding RNAs associated with the regulation of quorum sensing, nitrogen assimilation, and galactose utilization in *E. coli*, *Salmonella*, *Vibrionales*, and cyanobacteria (73–75). This makes *SMelC181* an interesting candidate for upcoming studies.

In the last 2 decades, cNMP signaling networks have been characterized in several bacterial species and different CRPs have been studied for their cNMP and DNA binding properties (3, 6, 10, 76). These studies imply that cNMP signaling networks and CRPs have evolved differently in the context of the different habitats of these bacteria. In this study, we showcased a CRP-like protein that has evolved to be activated both by cAMP and cGMP to perform its function in transcriptional regulation in a plant-symbiotic alphaproteobacterium. Clr is well conserved in the *Sinorhizobium* genus, including the cAMP-binding and the unusual DNA-binding amino acid motifs. In other closely related symbiotic nitrogen-fixing rhizobia, such as *Rhizobium leguminosarum* and *Rhizobium etli*, Clr homologs have not been found. However, similar proteins are present in individual *Rhizobium* species and *Bradyrhizobium diazoefficiens* isolates. The biological function of Clr dual specificity for activation by the two cNMPs remains to be discovered. It might have evolved as an adaptation to endogenous bacterial cGMP synthesis. However, conditions leading to decent levels of cGMP in *S. meliloti* are unknown. A role for cGMP in plant signal transductions related to development, stress response, ion homeostasis, and hormone function has been demonstrated, and long-distance cGMP signaling was suggested (77, 78). It is therefore tempting to speculate that *S. meliloti* may sense host plant-derived cGMP. Although the *S. meliloti* Clr regulon has been extensively characterized, current knowledge of its biological functions is limited to its role in early stages of the symbiosis (5). Deciphering the biological function of the responsiveness of Clr to cAMP and cGMP and whether cluster G CRP transcription factors are generally characterized by being activatable by cGMP or by both cGMP and cAMP is an exciting goal of future research.

MATERIALS AND METHODS

Strains, plasmids, and growth conditions. Unless specified otherwise, *E. coli* was grown at 37°C in LB-Lennox medium (1% tryptone, 0.5% yeast extract, 0.5% NaCl) with the addition of kanamycin (50 μg/mL), gentamicin (8 μg/mL), or tetracycline (10 μg/mL) as needed. *S. meliloti* was grown in TY medium (0.5% tryptone, 0.3% yeast extract, 2.7 mM CaCl$_2$) or MOPS minimal medium (1% MOPS, 1 mM MgSO$_4$, 20 mM sodium glutamate, 20 mM mannitol, 2 mM K$_2$HPO$_4$, 250 μM CaCl$_2$, 37 μM FeCl$_3$, 4.1 μM biotin, 45.8 μM H$_3$BO$_3$, 10 μM MnSO$_4$, 1 μM ZnSO$_4$, 0.5 μM CuSO$_4$, 0.27 μM CoCl$_2$, 0.5 μM Na$_2$MoO$_4$). If required, streptomycin (600 μg/mL), kanamycin (200 μg/mL), gentamicin (30 μg/mL), or tetracycline (10 μg/mL) was added to solid medium containing agar as a solidifying agent. For liquid media, the concentrations of kanamycin, gentamicin, or tetracycline were reduced by half.

Plasmids were transferred to *S. meliloti* by *E. coli* S17-1-mediated conjugation as previously described (79).

Promoter-probe constructs were generated by insertion of an *S. meliloti* gene upstream region of 300 to 735 bp and up to 75 bp of the associated protein or RNA coding region into the replicative low-copy-number plasmid pPHU231-EGFP or replicative medium-copy-number plasmid pSRKKm-EGFP for studies in *S. meliloti* strains. For the promoters of small noncoding RNA genes, a Shine-Dalgarno sequence was included downstream of the selected promoter sequence.

Site-directed mutagenesis of putative CBSs was performed by overlap extension PCR, and mutated promoter regions of the same size as the corresponding wild-type promoter region were inserted into pSRKKm-EGFP.

For cloning of CBSs, two complementary oligonucleotides (1 μM each) were hybridized in T4 DNA ligase buffer (NEB) containing T4 polynucleotide kinase (PNK; 0.2 μL) (NEB) in a total volume of 10 μL to obtain double-stranded DNA flanked by XbaI and HindIII overhangs. Phosphorylation of the oligonucleotides was performed at 37°C for 45 min, followed by a PNK deactivating step at 65°C for 20 min. The reaction mixture was then heated at 95°C for 5 min and slowly cooled down to 30°C. Thirty nanograms of the hybridized product was directly used as an insert for ligation with pSRKKm-EGFP.

For construction of the C-terminal 3×FLAG-tagged Clr overexpression construct, the full-length *clr* coding sequence was amplified from genomic DNA of *S. meliloti* Rm2011, digested with NdeI and XbaI, and inserted into plasmid pSRKKm-CF, yielding pSRKKm-*clr*-CF. Subsequently, *clr*-CF was amplified from this plasmid, digested with XbaI and HindIII, and inserted into plasmid pWBT, yielding pWBT-*clr*-CF. Expression plasmids pWBT-AC/GC and pWBT-AC/GC-*clr* are described in reference 24.

To obtain *S. meliloti* gene deletion mutants, gene-flanking regions of 350 to 750 bp were cloned into suicide vector pK18mobSacB. Following conjugation-mediated plasmid transfer to *S. meliloti* Rm2011 and plasmid integration into the genomic DNA by homologous recombination, transconjugants were subjected to sucrose selection as previously described (80). Gene deletions were verified by PCR. The *S. meliloti* *cya^0* multiple deletion mutant lacking the annotated 28 putative class III AC/GC genes was generated by sequential gene deletions. The genomes of *cya^0* strains as well as the parenting strain Rm2011 were sequenced. Total DNA was purified using a DNeasy blood and tissue kit (Qiagen). DNA sequencing libraries were generated by applying the Nextera XT DNA library preparation kit (Illumina). Sequencing was performed on a MiSeq desktop sequencer (Illumina), using the MiSeq reagent kit v2 for 2 × 250-bp paired-end reads (Illumina) for Rm2011, resulting in 4.45 × 10^6 reads, or the MiSeq reagent kit v3 for 2 × 75-bp paired-end reads for Rm2011 *cya^0*, resulting in 5.82 × 10^6 reads. Single-nucleotide polymorphism (SNP)

detection was performed by applying CLC Genomics Workbench (v10.1.1; Qiagen). Paired reads were mapped to the annotated reference genome of *S. meliloti* Rm1021 (chromosome, GenBank accession no. AL591688; pSymA, NC_003037; pSymB, AL591985). The minimum variation frequency was 50%, and at least 10 reads with a minimum mismatch coverage of 8 were considered. The *S. meliloti cya⁰* strain used in this study (Fig. 4; see Table S3 in the supplemental material) carries one SNP (Table S1K).

To construct the expression plasmid coding for C-terminally His_6-tagged Clr, the native encoding sequence was amplified by PCR and inserted into NdeI- and XhoI-digested pET36b(+) (Novagen). The fusion construct was verified by DNA sequencing.

Strains, plasmids, and oligonucleotides used in this study are described in Table S5A and B.

EGFP fluorescence measurements. *S. meliloti* strains were grown in 100 μL TY medium, and *E. coli* BTH101 was grown in 100 μL LB medium in 96-well polystyrene flat-bottom plates (Greiner) at 30°C with shaking at 1,200 rpm. For cultures in which cNMPs were added to the medium, stationary precultures of the respective strains grown in TY medium without cNMPs were diluted 1:100 in medium containing 400 μM cAMP or cGMP or as a control in medium without cNMPs, followed by incubation for 24 h. For promoter studies in *E. coli* BTH101, stationary precultures were diluted 1:100 in medium containing 100 μM IPTG, followed by incubation for 16 h.

EGFP fluorescence measurements were carried out as described previously (24). Relative fluorescent units (RFU) represent fluorescent values divided by the optical density at 600 nm (OD_{600}). The measurements were carried out in *S. meliloti* Rm2011, Rm2011 Δclr, and Rm2011 $cya⁰$ and in *E. coli* BTH101 strains. Processed data and underlying raw data of each experiment are listed in Tables S2 to S13.

In *S. meliloti* strains, the background fluorescence of the corresponding empty vector control was subtracted (Fig. 1A and C; Fig. S2E and Tables S2 and S3). In measurements carried out in *E. coli* BTH101 (Fig. 4B; Table S4), no background fluorescence was subtracted. Fluorescence fold change of *S. meliloti* strains carrying promoter-probe constructs was calculated from the RFU of the induced strains divided by the RFU of the noninduced strains or by the RFU of the wild type divided by the RFU of the mutants (Table S2). In *E. coli*, fluorescence fold change was calculated using the fluorescent values derived from the strains carrying pPHU231-promoter-probe-EGFP in combination with pWBT-AC/GC-*clr*-EGFP divided by the values derived from the strains carrying pPHU231-promoter-probe-EGFP in combination with pWBT-AC/GC-EGFP (Table S4). Three to four independent transconjugants and transformants of each strain containing the promoter-*egfp* constructs were used as biological replicates.

ChIP-seq. Cultures of Rm2011 Δclr carrying the *clr*-FLAG overexpression construct pWBT-clr-CF were grown in 60 mL of TY or MOPS medium in 500-mL flasks, each supplemented with either 400 μM 3′,5′-cAMP or 3′,5′-cGMP and 500 μM IPTG. Upon reaching an OD_{600} of 0.6, the cells were fixed with 1% formaldehyde for 20 min at room temperature. Fixation was quenched with 250 μM glycine for 20 min.

Cell pellets were resuspended in 300 μL immunoprecipitation (IP) buffer (50 mM HEPES-KOH [pH 7.8], 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS) with 1 mM phenylmethylsulfonyl fluoride (PMSF). Cells were lysed and DNA fragmented (~500 bp) by sonication in a Bioruptor Plus (Diagenode) by 48 cycles of 30 s of sonication and 30 s of cooling. Three hundred microliters of sonicated product was mixed with 3 mL of IP buffer and 1 mM PMSF and centrifuged at full speed at 4°C for 30 min. One hundred microliters of supernatant was taken and frozen at −20°C as a control. Clr-CF was immunoprecipitated with Anti-FLAG M2 affinity gel (Sigma-Aldrich). The protein was eluted with 100 μL of 3×FLAG peptide elution buffer for 1 h at 4°C. Cross-links of control and ChIP DNA were removed with 200 μM NaCl, and proteins and RNA were degraded with proteinase K and RNase A overnight at 65°C. Samples were checked by Western blotting with monoclonal anti-FLAG M2-horseradish peroxidase (HRP) antibody (1:1,000). DNA was purified with the QIAquick PCR kit. Ten nanograms of DNA was taken for ChIP-seq DNA library preparation and 2.5 ng for quantitative PCR (qPCR). The DNA library preparation was carried out as described previously (81) with the adapter sequences 2, 4 to 7, and 12 to 14 compatible with Illumina's TruSeq platform. Sequencing was performed on a MiSeq desktop sequencer (Illumina) using a MiSeq reagent kit v3 with 2 × 75 paired-end reads, and sequence analysis was performed using CLC Genomics Workbench (v10.1.1; Qiagen).

ChIP-seq reads 1 and 2 were paired with a distance between 1 and 1,000 bp, and failed reads were removed. Paired reads were mapped to the annotated *S. meliloti* 1021 genome (Chr1021 embl with annotation AL591688, pSymA1021 embl with annotation AE006469, and pSymB1021 embl with annotation AL591985). The following mapping parameters were used: match score = 1, mismatch cost = 2, linear gap cost = 3 (deletion/insertion), length fraction = 0.5, similarity fraction = 0.8 (at least 50% of the alignment has to have 80% sequence identity). Nonspecific matches were mapped randomly but also assigned as such. Peak calling was performed with the transcription factor ChIP-seq tool by comparing the mapped reads of IP and control samples (maximum P value = 0.05). Sequences flanking each center of a peak derived from the ChIP-seq analysis were defined and summarized as FASTA sequences using the R package (82). Motifs were generated with the MEME Suite online tools MEME (v5.1) and FIMO (v5.3) (30, 83).

EMSA. An EMSA reaction mixture contained 2 mM HEPES (pH 7.0), 30 mM NaCl, 50 mM KCl, 850 ng of sonicated salmon sperm DNA (GE Healthcare), 1 μg of bovine serum albumin (BSA) (Sigma), and 20 ng of Cy3-labeled DNA in a final volume of 10 μL. Native and synthetic Cy3-labeled DNA fragments (Table S5C and D) were obtained by PCR with Cy3-labeled primers (Table S5B) using the corresponding pSRKKm-promoter-EGFP and pSRKKm-CBS-EGFP constructs as the template, respectively. Synthetic Cy3-labeled DNA fragments of 592 bp were composed of the CBS flanked by pSRKKm-EGFP-derived sequence. The protein was added at 3 μg (111 μM) per reaction mixture, and 3′,5′-cAMP or 3′,5′-cGMP was added at 1 mM if indicated. The reaction mixtures were incubated at room temperature for 30 min in the dark. A mixture of 1 μL of 90% glycerol and 1.5 μL of 5× Tris-borate-EDTA (TBE) buffer was added to each reaction mixture, and 10 μL of this mixture was loaded onto a 10% polyacrylamide gel in 1× TBE. Following electrophoresis

at 9 V cm^{-1} at room temperature for 3.5 h, images were taken using a Typhoon 8600 variable mode imager (Amersham Bioscience).

Protein expression and purification. Clr was overproduced in *E. coli* BL21(DE3) cells in LB-Lennox medium containing kanamycin (30 μg/mL) at 37°C and induced at an OD$_{600}$ of 0.5 using IPTG (0.1 mM). Three hours after addition of 0.1 mM IPTG, Clr was harvested by centrifugation (15 min at 4°C; 4,000 rpm). The pellet was resuspended in 20 mL of binding buffer (50 mM sodium dihydrogen phosphate, 1,000 mM sodium chloride, 10 mM imidazole, pH 7.0) and disrupted by three passages through a cold French pressure cell press at 15.2 $\times$ 10^5 Pa cm^{-2}. The lysate was centrifuged at 18,000 rpm for 60 min at 4°C and filtered through 0.45-μm membrane filter. For Clr purification, the initial purification was conducted using a 5-mL Protino Ni-nitrilotriacetic acid (Ni-NTA) column (Macherey-Nagel). Size-exclusion chromatography over a 16/600 Superdex 200 column (GE) using elution buffer (20 mM HEPES, 300 mM NaCl, pH 7.0) was done as a polishing step. Preceding the purification, the size exclusion buffer was optimized as described using SYPRO orange dye (84).

Protein crystallization. Clr at 0.25 mM or 0.44 mM was crystallized in the presence of 25 mM cAMP or cGMP, respectively, 25 mM MgCl$_2$, and a 1.25-fold excess of 19-mer/14-mer duplex target DNA (19-mer sense strand, 5'-CTA GGT AAC ATT ACT CGC G-3'; 14-mer antisense strand, 5'-GCG AGT AAT GTT AC-3'). The oligonucleotide (BioCat) was prepared from single strands by heating equimolar amounts to 98°C for 5 min and slowly cooling to room temperature. The DNA oligonucleotides used for the crystallization of Clr were designed based on the genes *SMc04190* and *SMc00925*, which have previously been reported to be regulated by Clr (24).

Crystals of Clr in complex with cAMP were grown at 291 K by sitting-drop vapor diffusion in 0.09 M sodium fluoride, 0.09 M sodium bromide, 0.09 M sodium iodide, 0.1 M Tris base/bicine (pH 8.5), 12.5% (vol/vol) 2-methyl-2,4-pentanediol (MPD), 12.5% (wt/vol) polyethylene glycol (PEG) 1000, and 12.5% (wt/vol) PEG 3350. Crystals containing cGMP were grown at 281 K in sitting-drop vapor diffusion plates in 0.2 M 1,6-hexanediol, 0.2 M 1-butanol, 0.2 M (*RS*)-1,2-butanediol, 0.2 M 2-propanol, 0.2 M 1,4-butanediol, 0.2 M 1,3-propanediol, 0.1 M MOPS/HEPES-Na (pH 7.5), 12.5% (vol/vol) MPD, 12.5% (wt/vol) PEG 1000, and 12.5% (wt/vol) PEG 3350.

The crystals were flash-frozen in liquid nitrogen, and diffraction data were collected at 100 K at the EMBL/DESY P13 beamline (Hamburg) using a Pilatus 6M detector (Dectris) and at the SLS PXI X06SA beamline (Paul Scherrer Institute, Villigen, Switzerland) using an Eiger 16M detector (Dectris), respectively. At wavelengths of 0.976 and 1.000 Å, the crystals diffracted to 2.8 and 3.1 Å, respectively. The data were processed using XDS (85) in space group P2$_1$ 2$_1$ 2$_1$. Data reduction and scaling were done using CCP4i2 (v7.0.065) (86), AIMLESS in particular (version 0.7.3). The phases were solved by molecular replacement using a hybrid model derived from GlxR from *Corynebacterium glutamicum* (PDB code 4CYD) and Crp from *Mycobacterium tuberculosis* (PDB code 3MZH). The model was built using COOT (v0.8.9) (87) with refinement in Phenix (v1.11.1) (88). Final refinement statistics are given in Table 2. The coordinates and structure factors were deposited in the Protein Data Bank under PDB codes 7PZA and 7PZB.

ITC measurements. Clr was transferred to the binding buffer (20 mM HEPES, 300 mM NaCl, 20 mM MgCl$_2$, pH 7.0) via a PD-10 column (Cytiva Life Sciences), and the titration calorimetry measurements were performed using a Malvern MicroCAL ITC calorimeter as previously described (89). A typical titration consisted of injecting 2-μL aliquots of 5 mM ligand solution into 0.3 to 0.4 mM protein solution every 2.5 min to ensure that the titration peak returned to the baseline prior to the next injection. For the measurement of DNA affinities, the setup was reversed, with protein titration at 0.3 to 0.4 M containing 1 mM cyclic nucleotide into DNA at 25 μM. The cell was temperature controlled to 25°C. Titration curves for the dilution of the ligand solution were deducted from the data.

HDX-MS measurements. Sample preparation was automated with a two-arm robotic autosampler (LEAP Technologies). A 7.5-μL volume of 50 μM Clr with and without 1 mM cNMP was mixed with 67.5 μL of D$_2$O-based size exclusion chromatography (SEC) buffer. After 10, 30, 95, 1,000, and 10,000 s at 25°C, respectively, the hydrogen/deuterium exchange was quenched by the addition of equal parts quench buffer (400 mM KH$_2$PO$_4$/H$_3$PO$_4$, 2 M guanidine hydrochloride, pH 2.2) at 1°C. The solution was injected into an Acquity ultraperformance liquid chromatography (UPLC) M-class system with HDX technology (Waters). The protein was digested online with immobilized porcine pepsin at 12°C at 100 μL/min (H$_2$O, 0.1% formic acid), and the resulting peptides were collected on a trap column (2 mm by 2 cm) with POROS 20 R2 material (Thermo Scientific) at 0.5°C. After 3 min, the trap column was switched online with an Acquity UPLC BEH C$_{18}$ column (1.7 μm, 1.0 by 100 mm) (Waters), and the peptides were eluted at 0.5°C using a gradient of H$_2$O–0.1% formic acid (A) and acetonitrile–0.1% formic acid (B) at 30 μL/min (5% to 35% B in 7 min, 35% to 85% B within 1 min, and isocratic flow 85% B for 2 min). The column was washed for 1 min at 95% B and equilibrated at 5% B for 5 min after this. Peptides were ionized by electrospray ionization at a 250°C source capillary temperature and a spray voltage of 3.0 kV. Mass spectra were acquired on a G2-Si high-definition mass spectrometer (HDMS) with ion mobility separation (Waters) over a range of 50 to 2,000 *m/z* in HDMSE (high-definition MS^E) or HDMS mode for undeuterated and deuterated samples, respectively. Lock mass correction was performed with a [Glu1]-fibrinopeptide B standard (Waters). Between samples, the pepsin column was washed three times with 80 μL of 4% (vol/vol) acetonitrile and 0.5 M guanidine hydrochloride, and additionally, blank runs were performed between samples. Peptides were identified, and deuterium uptake was determined by employing the PLGS and DynamX 3.0 software suites (both from Waters) as described previously (90).

Data availability. The ChIP-seq data have been deposited in the ArrayExpress database at EMBL-EBI (https://www.ebi.ac.uk/arrayexpress) under accession no. E-MTAB-11788 (91). TSS data of *S. meliloti* strain RFF625c are available from ArrayExpress (E-MTAB-12127). The Clr protein crystal structure data have been deposited in the Protein Data Bank archive under accession codes 7PZA and 7PZB. All study data are included in the article and/or supplemental material.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.
FIG S1, PDF file, 0.1 MB.
FIG S2, PDF file, 0.7 MB.
FIG S3, PDF file, 0.5 MB.
FIG S4, PDF file, 0.6 MB.
FIG S5, PDF file, 0.2 MB.
TABLE S1, XLSX file, 0.9 MB.
TABLE S2, XLSX file, 0.7 MB.
TABLE S3, XLSX file, 0.1 MB.
TABLE S4, XLSX file, 0.1 MB.
TABLE S5, XLSX file, 0.1 MB.

REFERENCES

1. Berthet J, Rall TW, Sutherland EW. 1957. The relationship of epinephrine and glucagon to liver phosphorylase. IV. Effect of epinephrine and glucagon on the reactivation of phosphorylase in liver homogenates. J Biol Chem 224: 463–475.

2. Berthet J, Sutherland EW, Rall TW. 1957. The assay of glucagon and epinephrine with use of liver homogenates. J Biol Chem 229:351–361. https://doi.org/10.1016/S0021-9258(18)70622-4.

3. Busby S, Ebright RH. 1999. Transcription activation by catabolite activator protein (CAP). J Mol Biol 293:199–213. https://doi.org/10.1006/jmbi.1999.3161.

4. Soberón-Chávez G, Alcaraz LD, Morales E, Ponce-Soto GY, Servín-González L. 2017. The transcriptional regulators of the CRP family regulate different essential bacterial functions and can be inherited vertically and horizontally. Front Microbiol 8:451. https://doi.org/10.3389/fmicb.2017.00959.

5. Tian CF, Garnerone A-M, Mathieu-Demaziere C, Masson-Boivin C, Batut J. 2012. Plant-activated bacterial receptor adenylate cyclases modulate epidermal infection in the Sinorhizobium meliloti-Medicago symbiosis. Proc Natl Acad Sci U S A 109:6751–6756. https://doi.org/10.1073/pnas.1120260109.

6. Beavo JA, Brunton LL. 2002. Cyclic nucleotide research—still expanding after half a century. Nat Rev Mol Cell Biol 3:710–718. https://doi.org/10.1038/nrm911.

7. Gomelsky M. 2011. cAMP, c-di-GMP, c-di-AMP and now cGMP: bacteria use them all! Mol Microbiol 79:562–565. https://doi.org/10.1111/j.1365-2958.2010.07514.x.

8. Gosset G, Zhang Z, Nayyar S, Cuevas WA, Saier MH. 2004. Transcriptome analysis of Crp-dependent catabolite control of gene expression in Escherichia coli. J Bacteriol 186:3516–3524. https://doi.org/10.1128/JB.186.11.3516-3524.2004.

9. Harman JG. 2001. Allosteric regulation of the cAMP receptor protein. Biochim Biophys Acta 1547:1–17. https://doi.org/10.1016/s0167-4838(01)00187-x.

10. Kolb A, Busby S, Buc H, Garges S, Adhya S. 1993. Transcriptional regulation by cAMP and its receptor protein. Annu Rev Biochem 62:749–795. https://doi.org/10.1146/annurev.bi.62.070193.003533.

11. Won H-S, Lee Y-S, Lee S-H, Lee B-J. 2009. Structural overview on the allosteric activation of cyclic AMP receptor protein. Biochim Biophys Acta 1794:1299–1308. https://doi.org/10.1016/j.bbapap.2009.04.015.

12. Kohl TA, Tauch A. 2009. The GlxR regulon of the amino acid producer Corynebacterium glutamicum: detection of the corynebacterial core regulon and integration into the transcriptional regulatory network model. J Biotechnol 143:239–246. https://doi.org/10.1016/j.jbiotec.2009.08.005.

13. Knapp GS, McDonough KA. 2014. Cyclic AMP signaling in Mycobacteria. Microbiol Spectr 2:MGM2-0011-2013. https://doi.org/10.1128/microbiolspec.MGM2-0011-2013.

14. Kahramanoglou C, Cortes T, Matange N, Hunt DM, Visweswariah SS, Young DB, Buxton RS. 2014. Genomic mapping of cAMP receptor protein (CRP Mt) in Mycobacterium tuberculosis: relation to transcriptional start sites and the role of CRP Mt as a transcription factor. Nucleic Acids Res 42: 8320–8329. https://doi.org/10.1093/nar/gku548.

15. Jungwirth B, Sala C, Kohl TA, Uplekar S, Baumbach J, Cole ST, Pühler A, Tauch A. 2013. High-resolution detection of DNA binding sites of the global transcriptional regulator GlxR in Corynebacterium glutamicum. Microbiology 159:12–22. https://doi.org/10.1099/mic.0.062059-0.

16. Körner H, Sofia HJ, Zumft WG. 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. https://doi.org/10.1016/S0168-6445(03)00066-4.

17. Shabb JB, Corbin JD. 1992. Cyclic nucleotide-binding domains in proteins having diverse functions. J Biol Chem 267:5723–5726. https://doi.org/10.1016/S0021-9258(18)42609-9.

18. Youn H, Kerby RL, Koh J, Roberts GP. 2007. A C-helix residue, Arg-123, has important roles in both the active and inactive forms of the cAMP receptor protein. J Biol Chem 282:3632–3639. https://doi.org/10.1074/jbc.M606602200.

19. Lawson CL, Swigon D, Murakami KS, Darst SA, Berman HM, Ebright RH. 2004. Catabolite activator protein: DNA binding and transcription activation. Curr Opin Struct Biol 14:10–20. https://doi.org/10.1016/j.sbi.2004.01.012.

20. Benoff B, Yang H, Lawson CL, Parkinson G, Liu J, Blatter E, Ebright YW, Berman HM, Ebright RH. 2002. Structural basis of transcription activation: the CAP-alpha CTD-DNA complex. Science 297:1562–1566. https://doi .org/10.1126/science.1076376.

21. McKay DB, Steitz TA. 1981. Structure of catabolite gene activator protein at 2.9 Å resolution suggests binding to left-handed B-DNA. Nature 290: 744–749. https://doi.org/10.1038/290744a0.

22. Bantu L, Chauhan S, Srikumar A, Hirakawa Y, Suzuki I, Hagemann M, Prakash JSS. 2022. A membrane-bound cAMP receptor protein, SyCRP1 mediates inorganic carbon response in Synechocystis sp. PCC 6803. Biochim Biophys Acta Gene Regul Mech 1865:194803. https://doi.org/10 .1016/j.bbagrm.2022.194803.

23. Roychowdhury S, Dong Q, Bauer CE. 2015. DNA-binding properties of a cGMP-binding CRP homologue that controls development of metabolically dormant cysts of Rhodospirillum centenum. Microbiology (Reading) 161:2256–2264. https://doi.org/10.1099/mic.0.000172.

24. Krol E, Klaner C, Gnau P, Kaever V, Essen L-O, Becker A. 2016. Cyclic mononucleotide- and Clr-dependent gene regulation in Sinorhizobium meliloti. Microbiology (Reading) 162:1840–1856. https://doi.org/10.1099/mic.0.000356.

25. Ebright RH, Le Grice SF, Miller JP, Krakow JS. 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. https://doi.org/10.1016/0022-2836(85)90030-0.

26. Gourion B, Berrabah F, Ratet P, Stacey G. 2015. Rhizobium-legume symbioses: the crucial role of plant immunity. Trends Plant Sci 20:186–194. https://doi.org/10.1016/j.tplants.2014.11.008.

27. Jones KM, Kobayashi H, Davies BW, Taga ME, Walker GC. 2007. How rhizobial symbionts invade plants: the Sinorhizobium-Medicago model. Nat Rev Microbiol 5:619–633. https://doi.org/10.1038/nrmicro1705.

28. Batut J, Daveran-Mingot ML, David M, Jacobs J, Garnerone AM, Kahn D. 1989. fixK, a gene homologous with fnr and crp from Escherichia coli, regulates nitrogen fixation genes both positively and negatively in Rhizobium meliloti. EMBO J 8:1279–1286. https://doi.org/10.1002/j.1460-2075.1989.tb03502.x.

29. Galibert F, Finan TM, Long SR, Puhler A, Abola P, Ampe F, Barloy-Hubler F, Barnett MJ, Becker A, Boistard P, Bothe G, Boutry M, Bowser L, Buhrmester J, Cadieu E, Capela D, Chain P, Cowie A, Davis RW, Dreano S, Federspiel NA, Fisher RF, Gloux S, Godrie T, Goffeau A, Golding B, Gouzy J, Gurjal M, Hernandez-Lucas I, Hong A, Huizar L, Hyman RW, Jones T, Kahn D, Kahn ML, Kalman S, Keating DH, Kiss E, Komp C, Lelaure V, Masuy D, Palm C, Peck MC, Pohl TM, Portetelle D, Purnelle B, Ramsperger U, Surzycki R, Thebault P, Vandenbol M, et al. 2001. The composite genome of the legume symbiont Sinorhizobium meliloti. Science 293:668–672. https://doi .org/10.1126/science.1060966.

30. Grant CE, Bailey TL, Noble WS. 2011. FIMO: scanning for occurrences of a given motif. Bioinformatics 27:1017–1018. https://doi.org/10.1093/bioinformatics/ btr064.

31. Zou L, Gastebois A, Mathieu-Demazière C, Sorroche F, Masson-Boivin C, Batut J, Garnerone A-M. 2017. Transcriptomic insight in the control of legume root secondary infection by the Sinorhizobium meliloti transcriptional regulator Clr. Front Microbiol 8:292. https://doi.org/10.3389/fmicb .2017.01236.

32. Schlüter J-P, Reinkensmeier J, Barnett MJ, Lang C, Krol E, Giegerich R, Long SR, Becker A. 2013. Global mapping of transcription start sites and promoter motifs in the symbiotic α-proteobacterium Sinorhizobium meliloti 1021. BMC Genomics 14:156. https://doi.org/10.1186/1471-2164-14-156.

33. Sallet E, Roux B, Sauviac L, Jardinaud M-F, Carrère S, Faraut T, de Carvalho-Niebel F, Gouzy J, Gamas P, Capela D, Bruand C, Schiex T. 2013. Next-generation annotation of prokaryotic genomes with EuGene-P: application to Sinorhizobium meliloti 2011. DNA Res 20:339–354. https:// doi.org/10.1093/dnares/dst014.

34. Lang C, Barnett MJ, Fisher RF, Smith LS, Diodati ME, Long SR. 2018. Most Sinorhizobium meliloti extracytoplasmic function sigma factors control accessory functions. mSphere 3:e00454-18. https://doi.org/10 .1128/mSphereDirect.00454-18.

35. Zhou Y, Merkel TJ, Ebright RH. 1994. Characterization of the activating region of Escherichia coli catabolite gene activator protein (CAP). II. Role at class I and class II CAP-dependent promoters. J Mol Biol 243:603–610. https://doi.org/10.1016/0022-2836(94)90035-3.

36. Ushida C, Aiba H. 1990. Helical phase dependent action of CRP: effect of the distance between the CRP site and the −35 region on promoter activity. Nucleic Acids Res 18:6325–6330. https://doi.org/10.1093/nar/18.21.6325.

37. Ebright RH. 1993. Transcription activation at Class I CAP-dependent promoters. Mol Microbiol 8:797–802. https://doi.org/10.1111/j.1365-2958.1993 .tb01626.x.

38. Schlüter J-P, Reinkensmeier J, Daschkey S, Evgueniiva-Hackenberg E, Janssen S, Jänicke S, Becker JD, Giegerich R, Becker A. 2010. A genome-wide survey of sRNAs in the symbiotic nitrogen-fixing alpha-proteobacterium Sinorhizobium meliloti. BMC Genomics 11:245. https://doi.org/10 .1186/1471-2164-11-245.

39. Małecki J, Polit A, Wasylewski Z. 2000. Kinetic studies of cAMP-induced allosteric changes in cyclic AMP receptor protein from Escherichia coli. J Biol Chem 275:8480–8486. https://doi.org/10.1074/jbc.275.12.8480.

40. Stapleton M, Haq I, Hunt DM, Arnvig KB, Artymiuk PJ, Buxton RS, 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. https://doi.org/10.1074/jbc.M109.047720.

41. Tutar Y. 2012. Neglected role of cAMP receptor protein monomer. Mol Biol Rep 39:4261–4265. https://doi.org/10.1007/s11033-011-1212-5.

42. Scott S-P, Jarjous S. 2005. Proposed structural mechanism of Escherichia coli cAMP receptor protein cAMP-dependent proteolytic cleavage protection and selective and nonselective DNA binding. Biochemistry 44: 8730–8748. https://doi.org/10.1021/bi0479609.

43. Reddy MCM, Palaninathan SK, Bruning JB, Thurman C, Smith D, Sacchettini JC. 2009. Structural insights into the mechanism of the allosteric transitions of Mycobacterium tuberculosis cAMP receptor protein. J Biol Chem 284:36581–36591. https://doi.org/10.1074/jbc.M109.041343.

44. Epstein W, Rothman-Denes LB, Hesse J. 1975. Adenosine 3':5'-cyclic monophosphate as mediator of catabolite repression in Escherichia coli. Proc Natl Acad Sci U S A 72:2300–2304. https://doi.org/10.1073/pnas.72.6.2300.

45. Aykaç Fas B, Tutar Y, Haliloğlu T. 2013. Dynamic fluctuations provide the basis of a conformational switch mechanism in apo cyclic AMP receptor protein. PLoS Comput Biol 9:e1003141. https://doi.org/10.1371/journal .pcbi.1003141.

46. Popovych N, Tzeng S-R, Tonelli M, Ebright RH, Kalodimos CG. 2009. Structural basis for cAMP-mediated allosteric control of the catabolite activator protein. Proc Natl Acad Sci U S A 106:6927–6932. https://doi.org/10.1073/ pnas.0900595106.

47. Rohs R, Jin X, West SM, Joshi R, Honig B, Mann RS. 2010. Origins of specificity in protein-DNA recognition. Annu Rev Biochem 79:233–269. https:// doi.org/10.1146/annurev-biochem-060408-091030.

48. Finn RD, Attwood TK, Babbitt PC, Bateman A, Bork P, Bridge AJ, Chang H-Y, Dosztányi Z, El-Gebali S, Fraser M, Gough J, Haft D, Holliday GL, Huang H, Huang X, Letunic I, Lopez R, Lu S, Marchler-Bauer A, Mi H, Mistry J, Natale DA, Necci M, Nuka G, Orengo CA, Park Y, Pesseat S, Piovesan D, Potter SC, Rawlings ND, Redaschi N, Richardson L, Rivoire C, Sangrador-Vegas A, Sigrist C, Sillitoe I, Smithers B, Squizzato S, Sutton G, Thanki N, Thomas PD, Tosatto SCE, Wu CH, Xenarios I, Yeh L-S, Young S-Y, Mitchell AL. 2017. InterPro in 2017—beyond protein family and domain annotations. Nucleic Acids Res 45:D190–D199. https://doi.org/10.1093/nar/gkw1107.

49. Schultz SC, Shields GC, Steitz TA. 1991. Crystal structure of a CAP-DNA complex: the DNA is bent by 90 degrees. Science 253:1001–1007. https:// doi.org/10.1126/science.1653449.

50. Townsend PD, Jungwirth B, Pojer F, Bußmann M, Money VA, Cole ST, Pühler A, Tauch A, Bott M, Cann MJ, Pohl L. 2014. The crystal structures of apo and cAMP-bound GlxR from Corynebacterium glutamicum reveal structural and dynamic changes upon cAMP binding in CRP/FNR family transcription factors. PLoS One 9:e113265. https://doi.org/10.1371/journal .pone.0113265.

51. Akhter Y, Wilmanns M. 2011. Crystal structure of cAMP receptor protein from Mycobacterium tuberculosis in complex with cAMP and its DNA binding element. https://www.rcsb.org/structure.3mzh.

52. Seok S-H, Im H, Won H-S, Seo M-D, Lee Y-S, Yoon H-J, Cha M-J, Park J-Y, Lee B-J. 2014. Structures of inactive CRP species reveal the atomic details of the allosteric transition that discriminates cyclic nucleotide second messengers. Acta Crystallogr D Biol Crystallogr 70:1726–1742. https://doi .org/10.1107/S139900471400724X.

53. Sim D-W, Choi JW, Kim J-H, Ryu K-S, Kim M, Yu H-W, Jo K-S, Kim E-H, Seo M-D, Jeon YH, Lee B-J, Kim YP, Won H-S. 2017. C-terminal dimerization of apo-cyclic AMP receptor protein validated in solution. FEBS Lett 591: 1064–1070. https://doi.org/10.1002/1873-3468.12613.

54. Won HS, Yamazaki T, Lee TW, Yoon MK, Park SH, Kyogoku Y, Lee BJ. 2000. Structural understanding of the allosteric conformational change of cyclic AMP receptor protein by cyclic AMP binding. Biochemistry 39:13953–13962. https://doi.org/10.1021/bi000012x.

55. Tutar Y. 2008. Chemical linkage at allosteric activation of E. coli cAMP receptor protein. Protein J 27:21–29. https://doi.org/10.1007/s10930-007 -9104-1.

56. Harteis S, Schneider S. 2014. Making the bend: DNA tertiary structure and protein-DNA interactions. Int J Mol Sci 15:12335–12363. https://doi.org/10.3390/ijms150712335.

57. Dai J, Lin S-H, Kemmis C, Chin AJ, Lee JC. 2004. Interplay between site-specific mutations and cyclic nucleotides in modulating DNA recognition by Escherichia coli cyclic AMP receptor protein. Biochemistry 43:8901–8910. https://doi.org/10.1021/bi0499359.

58. Fic E, Bonarek P, Gorecki A, Kedracka-Krok S, Mikolajczak J, Polit A, Tworzydlo M, Dziedzicka-Wasylewska M, Wasylewski Z. 2009. cAMP receptor protein from Escherichia coli as a model of signal transduction in proteins—a review. J Mol Microbiol Biotechnol 17:1–11. https://doi.org/10.1159/000178014.

59. Parkinson G, Wilson C, Gunasekera A, Ebright YW, Ebright RH, Ebright RE, Berman HM. 1996. Structure of the CAP-DNA complex at 2.5 angstroms resolution: a complete picture of the protein-DNA interface. J Mol Biol 260:395–408. https://doi.org/10.1006/jmbi.1996.0409.

60. Hudson BP, Quispe J, Lara-González S, Kim Y, Berman HM, Arnold E, Ebright RH, Lawson CL. 2009. Three-dimensional EM structure of an intact activator-dependent transcription initiation complex. Proc Natl Acad Sci U S A 106:19830–19835. https://doi.org/10.1073/pnas.0908782106.

61. Liu B, Hong C, Huang RK, Yu Z, Steitz TA. 2017. Structural basis of bacterial transcription activation. Science 358:947–951. https://doi.org/10.1126/science.aao1923.

62. Feng Y, Zhang Y, Ebright RH. 2016. Structural basis of transcription activation. Science 352:1330–1333. https://doi.org/10.1126/science.aaf4417.

63. Shi W, Jiang Y, Deng Y, Dong Z, Liu B. 2020. Visualization of two architectures in class-II CAP-dependent transcription activation. PLoS Biol 18:e3000706. https://doi.org/10.1371/journal.pbio.3000706.

64. Youn H, Koh J, Roberts GP. 2008. Two-state allosteric modeling suggests protein equilibrium as an integral component for cyclic AMP (cAMP) specificity in the cAMP receptor protein of Escherichia coli. J Bacteriol 190:4532–4540. https://doi.org/10.1128/JB.00074-08.

65. Weber IT, Steitz TA. 1987. Structure of a complex of catabolite gene activator protein and cyclic AMP refined at 2.5 Å resolution. J Mol Biol 198:311–326. https://doi.org/10.1016/0022-2836(87)90315-9.

66. Tzeng S-R, Kalodimos CG. 2009. Dynamic activation of an allosteric regulatory protein. Nature 462:368–372. https://doi.org/10.1038/nature08560.

67. Youn H, Kerby RL, Conrad M, Roberts GP. 2006. Study of highly constitutively active mutants suggests how cAMP activates cAMP receptor protein. J Biol Chem 281:1119–1127. https://doi.org/10.1074/jbc.M509421200.

68. Gunasekara SM, Hicks MN, Park J, Brooks CL, Serate J, Saunders CV, Grover SK, Goto JJ, Lee J-W, Youn H. 2015. Directed evolution of the Escherichia coli cAMP receptor protein at the cAMP pocket. J Biol Chem 290:26587–26596. https://doi.org/10.1074/jbc.M115.678474.

69. Bracco L, Kotlarz D, Kolb A, Diekmann S, Buc H. 1989. Synthetic curved DNA sequences can act as transcriptional activators in Escherichia coli. EMBO J 8:4289–4296. https://doi.org/10.1002/j.1460-2075.1989.tb08615.x.

70. Pyles EA, Lee JC. 1998. Escherichia coli cAMP receptor protein-DNA complexes. 2. Structural asymmetry of DNA bending. Biochemistry 37:5201–5210. https://doi.org/10.1021/bi972451a.

71. Lin S-H, Lee JC. 2003. Determinants of DNA bending in the DNA—cyclic AMP receptor protein complexes in Escherichia coli. Biochemistry 42:4809–4818. https://doi.org/10.1021/bi027259.

72. Garnerone A-M, Sorroche F, Zou L, Mathieu-Demazière C, Tian CF, Masson-Boivin C, Batut J. 2018. NsrA, a predicted β-barrel outer membrane protein involved in plant signal perception and the control of secondary infection in Sinorhizobium meliloti. J Bacteriol 200:e00019-18. https://doi.org/10.1128/JB.00019-18.

73. Papenfort K, Pfeiffer V, Lucchini S, Sonawane A, Hinton JCD, Vogel J. 2008. Systematic deletion of Salmonella small RNA genes identifies CyaR, a conserved CRP-dependent riboregulator of OmpX synthesis. Mol Microbiol 68:890–906. https://doi.org/10.1111/j.1365-2958.2008.06189.x.

74. de Lay N, Gottesman S. 2009. The Crp-activated small noncoding regulatory RNA CyaR (RyeE) links nutritional status to group behavior. J Bacteriol 191:461–476. https://doi.org/10.1128/JB.01157-08.

75. Bækkedal C, Haugen P. 2015. The Spot 42 RNA: a regulatory small RNA with roles in the central metabolism. RNA Biol 12:1071–1077. https://doi.org/10.1080/15476286.2015.1086867.

76. McDonough KA, Rodriguez A. 2011. The myriad roles of cyclic AMP in microbial pathogens: from signal to sword. Nat Rev Microbiol 10:27–38. https://doi.org/10.1038/nrmicro2688.

77. Isner JC, Maathuis FJM. 2016. cGMP signalling in plants: from enigma to main stream. Funct Plant Biol 45:93–101. https://doi.org/10.1071/FP16337.

78. Gehring C, Turek IS. 2017. Cyclic nucleotide monophosphates and their cyclases in plant signaling. Front Plan Sci 8:1704. https://doi.org/10.3389/fpls.2017.01704.

79. Krol E, Becker A. 2014. Rhizobial homologs of the fatty acid transporter FadL facilitate perception of long-chain acyl-homoserine lactone signals. Proc Natl Acad Sci U S A 111:10702–10707. https://doi.org/10.1073/pnas.1404929111.

80. Schäfer A, Tauch A, Jäger W, Kalinowski J, Thierbach G, Pühler A. 1994. Small mobilizable multi-purpose cloning vectors derived from the Escherichia coli plasmids pK18 and pK19: selection of defined deletions in the chromosome of Corynebacterium glutamicum. Gene 145:69–73. https://doi.org/10.1016/0378-1119(94)90324-7.

81. Ford E, Nikopoulou C, Kokkalis A, Thanos D. 2014. A method for generating highly multiplexed ChIP-seq libraries. BMC Res Notes 7:312. https://doi.org/10.1186/1756-0500-7-312.

82. R Core Team. 2021. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

83. Bailey TL, Elkan C. 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol 2:28–36.

84. Ericsson UB, Hallberg BM, Detitta GT, Dekker N, Nordlund P. 2006. Thermofluor-based high-throughput stability optimization of proteins for structural studies. Anal Biochem 357:289–298. https://doi.org/10.1016/j.ab.2006.07.027.

85. Kabsch W. 2010. XDS. Acta Crystallogr D Biol Crystallogr 66:125–132. https://doi.org/10.1107/S0907444909047337.

86. Potterton L, Agirre J, Ballard C, Cowtan K, Dodson E, Evans PR, Jenkins HT, Keegan R, Krissinel E, Stevenson K, Lebedev A, McNicholas SJ, Nicholls RA, Noble M, Pannu NS, Roth C, Sheldrick G, Skubak P, Turkenburg J, Uski V, von Delft F, Waterman D, Wilson K, Winn M, Wojdyr M. 2018. CCP4i2: the new graphical user interface to the CCP4 program suite. Acta Crystallogr D Struct Biol 74:68–84. https://doi.org/10.1107/S2059798317016035.

87. Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. Acta Crystallogr D Biol Crystallogr 66:486–501. https://doi.org/10.1107/S0907444910007493.

88. Liebschner D, Afonine PV, Baker ML, Bunkóczi G, Chen VB, Croll TI, Hintze B, Hung LW, Jain S, McCoy AJ, Moriarty NW, Oeffner RD, Poon BK, Prisant MG, Read RJ, Richardson JS, Richardson DC, Sammito MD, Sobolev OV, Stockwell DH, Terwilliger TC, Urzhumtsev AG, Videau LL, Williams CJ, Adams PD. 2019. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr D Struct Biol 75:861–877. https://doi.org/10.1107/S2059798319011471.

89. Wiseman T, Williston S, Brandts JF, Lin L-N. 1989. Rapid measurement of binding constants and heats of binding using a new titration calorimeter. Anal Biochem 179:131–137. https://doi.org/10.1016/0003-2697(89)90213-3.

90. Osorio-Valeriano M, Altegoer F, Steinchen W, Urban S, Liu Y, Bange G, Thanbichler M. 2019. ParB-type DNA segregation proteins are CTP-dependent molecular switches. Cell 179:1512–1524.e15. https://doi.org/10.1016/j.cell.2019.11.015.

91. Athar A, Füllgrabe A, George N, Iqbal H, Huerta L, Ali A, Snow C, Fonseca NA, Petryszak R, Papatheodorou I, Sarkans U, Brazma A. 2019. ArrayExpress update—from bulk to single-cell expression data. Nucleic Acids Res 47:D711–D715. https://doi.org/10.1093/nar/gky964.