doi:10.1016/j.jmb.2008.08.027 *J. Mol. Biol.* (2008) **383**, 313–323

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

CRP Binding and Transcription Activation at CRP-S Sites

Andrew D. S. Cameron and Rosemary J. Redfield*

Department of Zoology, University of British Columbia, Vancouver, British Columbia, Canada

Received 12 June 2008; received in revised form 11 August 2008; accepted 13 August 2008
Available online
22 August 2008

In *Haemophilus influenzae*, as in *Escherichia coli*, the cAMP receptor protein (CRP) activates transcription from hundreds of promoters by binding symmetrical DNA sites with the consensus half-site 5′-$A_1A_2A_3T_4G_5T_6G_7A_8T_9C_{10}T_{11}$. We have previously identified 13 *H. influenzae* CRP sites that differ from canonical (CRP-N) sites in the following features: (1) Both half-sites of these noncanonical (CRP-S) sites have C_6 instead of T_6, although they otherwise have an unusually high level of identity with the binding site consensus. (2) Only promoters with CRP-S sites require both the CRP and Sxy proteins for transcription activation. To study the functional significance of CRP-S site sequences, we purified *H. influenzae* (*Hi*)CRP and compared its DNA binding properties to those of the well-characterized *E. coli* (*Ec*)CRP. All *Ec*CRP residues that contact DNA are conserved in *Hi*CRP, and both proteins demonstrated a similar high affinity for the CRP-N consensus sequence. However, whereas *Ec*CRP bound specifically to CRP-S sites *in vitro*, *Hi*CRP did not. By systematically substituting base pairs in native promoters and in the CRP-N consensus sequence, we confirmed that *Hi*CRP is highly specific for the perfect core sequence $T_4G_5T_6G_7A_8$ and is more selective than *Ec*CRP at other positions in CRP sites. Even though converting $C_6 \rightarrow T_6$ greatly enhanced *Hi*CRP binding to a CRP-S site, this had the unexpected effect of nearly abolishing promoter activity. A + T-rich sequences upstream of CRP-S sites were also found to be required for promoter activation, raising the possibility that Sxy binds these A + T sequences to simultaneously enable CRP–DNA binding and assist in RNA polymerase recruitment.

Edited by R. Ebright

Keywords: CRP; CAP; Sxy; CRP-S

Introduction

The cAMP receptor protein (CRP; also called catabolite activator protein) regulates a global sugar starvation response in three γ-proteobacteria families, the Enterobacteriaceae, Pasteurellaceae, and Vibrionaceae.[1] CRP binds specific sites at gene promoters when activated by its allosteric effector cAMP (a signal of sugar depletion) and then recruits RNA polymerase (RNAP) through direct protein–protein contacts (reviewed in Ref. 2). Because only *Escherichia coli* CRP has been extensively characterized (reviewed in Refs. 2–4), its structure and function serve as models for understanding CRP in other bacteria. Recent results have also implicated CRP as a regulator of natural competence in several important pathogens, including *E. coli*, *Haemophilus influenzae*, *Vibrio cholerae*, and *Salmonella* sp.[1,5]

CRP homodimers bind 22-bp sequences with 2-fold symmetry (consensus half-site 5′-$A_1A_2A_3T_4G_5T_6G_7A_8T_9C_{10}T_{11}$) and cause DNA to bend 80–90°.[6] CRP's strong preference for the sequence $T_4G_5T_6G_7A_8$ in both half-sites arises primarily from hydrogen bonds formed between protein side chains and base pairs G_5·C, G_7·C, and A_8·T. In addition, nonclassical hydrogen bonds between the protein and thymine methyl groups at T_4·A, T_6·A, and A_8·T make minor contributions to specificity.[7–10] Although CRP contacts the thymine methyl group at T_6,[7] strong selection for T at position 6 arises through an 'indirect

**Corresponding author.* Mailing address: Life Sciences Centre (Zoology), University of British Columbia, 2350 Health Sciences Mall, Vancouver, British Columbia, Canada V6T 1Z3. E-mail address: redfield@zoology.ubc.ca.

Abbreviations used: CRP, cAMP receptor protein; RNAP, RNA polymerase; sBHI, supplemented brain heart infusion.

readout' mechanism because the T_6/G_7 base step favors the ~40° kink needed for DNA deformation by CRP.[9,11]

All characterized *E. coli* CRP sites differ from the consensus at one or more positions, as do the binding sites of other global regulators.[12] Until recently, this has been thought to only reflect selection for stronger or weaker binding, with no significance attached to the presence of specific nonconsensus bases at specific CRP site positions. However, we have found

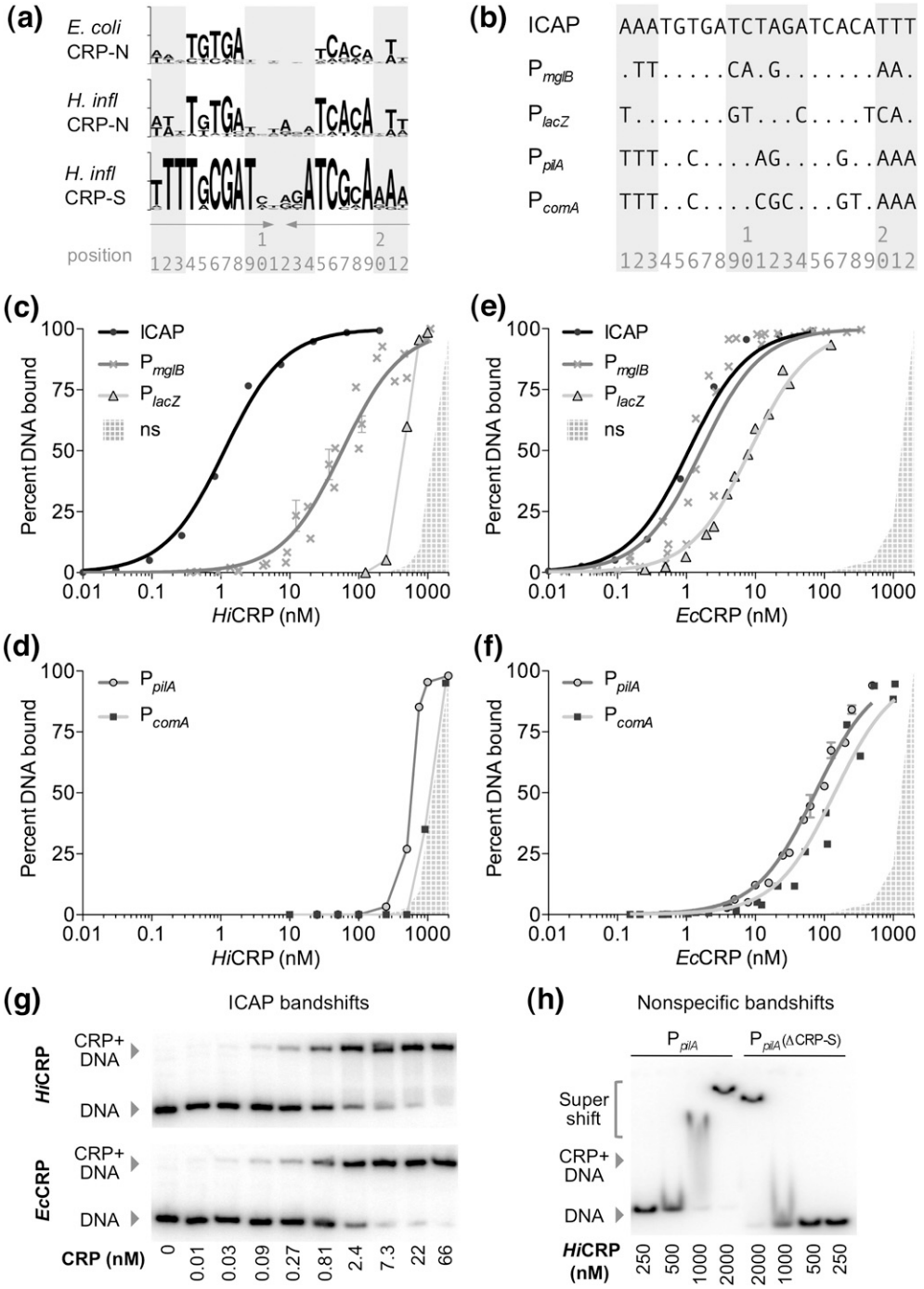

Fig. 1. *H. influenzae* and *E. coli* CRP binding to promoter DNA. (a) Sequence logos of *E. coli* CRP-N sites ($n = 49$), *H. influenzae* CRP-N sites ($n = 45$), and *H. influenzae* CRP-S sites ($n = 13$). The binding site cores (positions 4–8 and 15–19) are highlighted with white columns, and gray arrows at the bottom indicate the inverse palindrome of CRP sites. (b) Alignment of CRP binding sites in promoter DNAs used as bait in bandshift assays. Bases are numbered as in (a), and differences from the binding site consensus (ICAP) are shown. (c–f) Bandshift data. Binding curves were fit to the data except in cases where *Hi*CRP did not bind specifically to promoter DNA. Gray-shaded areas on graphs indicate the amount of nonspecific (ns) DNA binding at high CRP concentrations. (c) *Hi*CRP binding to CRP-N sites. (d) *Hi*CRP binding to CRP-S sites. (e) *Ec*CRP binding to CRP-N sites. (f) *Ec*CRP binding to CRP-S sites. (g) Representative bandshifts from which data in (c) and (e) were derived. (h) Bandshifts illustrating the supershifts characteristic of nonspecific DNA binding.

that the CRP sites of *H. influenzae* competence gene promoters consistently differ from canonical sites in having C_6 and never T_6 in both half-sites.[13] Bioinformatic analysis of other γ-proteobacteria genomes also revealed that the putative CRP sites of Pasteurellaceae, Enterobacteriaceae, and Vibrionaceae competence gene promoters have a consistent overrepresentation of C_6, suggesting that noncanonical (CRP-S) sites are not an *H. influenzae*-specific phenomenon.[1] However, the functional significance of C_6 in CRP-S sites is unclear because, in *E. coli*, a $T_6 \rightarrow C_6$ substitution in both halves of a canonical (CRP-N) site reduces CRP–DNA affinity 80-fold.[9]

All *H. influenzae* promoters with CRP-S sites also require the protein Sxy for transcription activation, raising the possibility that Sxy assists CRP binding to CRP-S sites and/or RNAP recruitment.[13] Although the abovementioned three bacterial families all have Sxy homologs, and extensive genetic studies have confirmed Sxy's role as a regulator of competence genes,[5,13–19] direct characterization of its action at CRP-S promoters has been stymied by the toxicity and intractability of this small protein. To better understand CRP binding to CRP-S sites, we have conducted a detailed analysis of *H. influenzae* (*Hi*) CRP and *E. coli* (*Ec*)CRP binding and transcriptional activation at native and synthetic CRP sites. Even though the proteins are 78% identical and all of *Ec*CRP's DNA-binding residues are conserved in *Hi*CRP, the latter is much more selective for binding sites. Nevertheless, two key features are shared: both proteins preferentially bind the same consensus

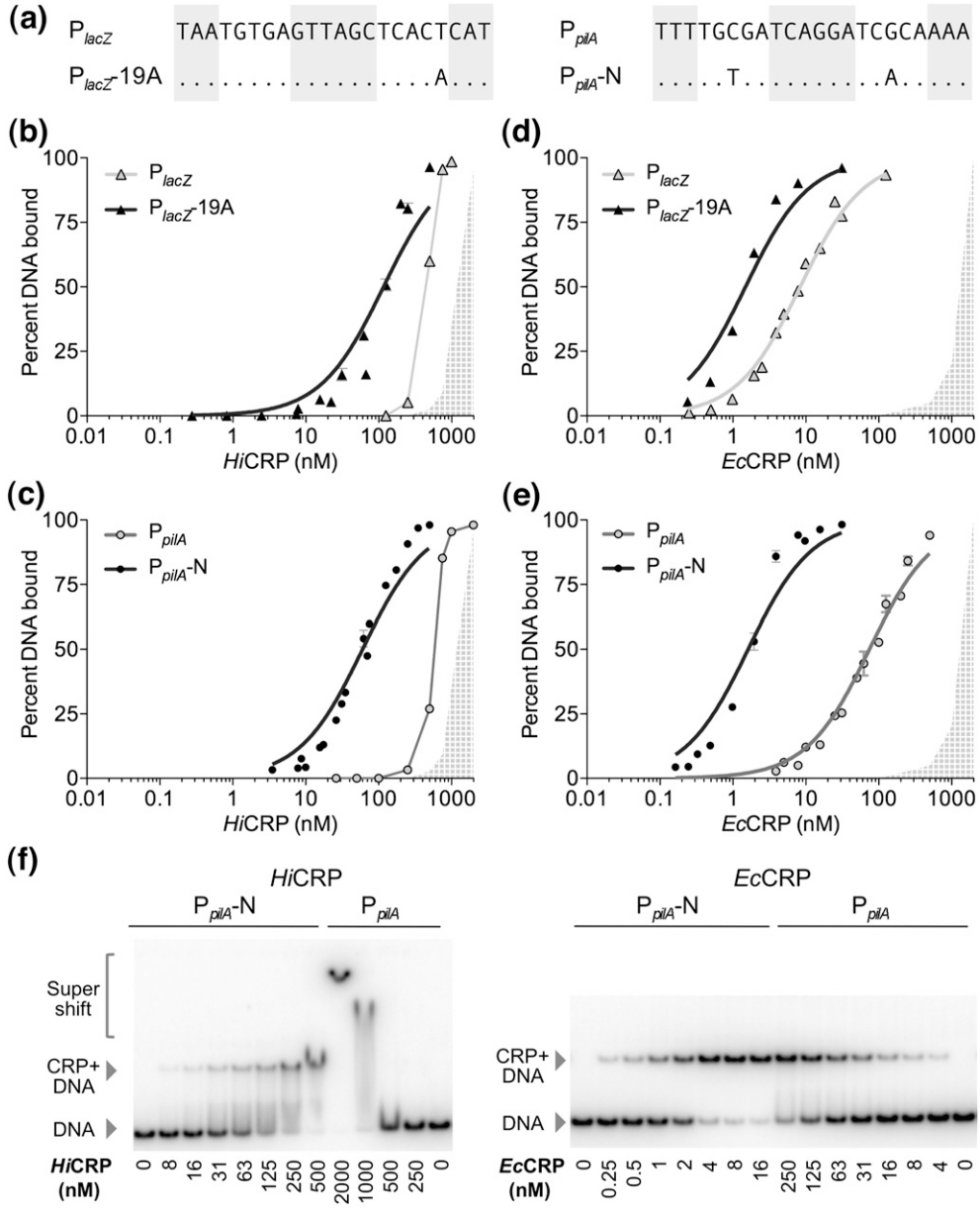

Fig. 2. CRP binding to P_{lacZ} and P_{pilA} and promoter mutants. (a) Alignment of CRP binding sites as in Fig. 1b. (b–e) Bandshift data plotted as in Fig. 1; P_{lacZ} and P_{pilA} binding data from Fig. 1c–f. (f) Representative bandshifts from which data in (c) and (e) were derived.

sequence, 5′-$A_1A_2A_3T_4G_5T_6G_7A_8T_9C_{10}T_{11}$, and neither protein can activate transcription from *H. influenzae*'s Sxy-regulated promoters in the absence of Sxy, even when bound to promoter DNA.

Results

H. influenzae and *E. coli* CRP have different affinities for promoter DNA *in vitro*

Transcriptome and genome analyses in *H. influenzae* have identified 54 CRP-regulated promoters: 41 promoters containing 45 CRP-N sites, which regulate genes for nutrient uptake and central metabolism, and 13 Sxy-dependent promoters containing 13 CRP-S sites.[13] Figure 1a compares a sequence logo of 49 experimentally determined *E. coli* CRP sites (the standard reference set, available at DPInteract[20]) to logos of *H. influenzae*'s CRP-N and CRP-S sites. The *E. coli* and *H. influenzae* CRP-N logos are very similar, showing consistent overrepresentation of the $T_4G_5T_6G_7A_8$ motif and the presence of A+T-rich sequences at positions 1 and 2 in both half-sites. The *H. influenzae* CRP-S logo differs from both CRP-N logos in several features: (i) consistent presence of C_6 in both half-sites; (ii) stronger core consensus; (iii) substantially stronger consensus at non-core positions, especially at T_3 and T_9 of both half-sites; and (iv) consistent presence of T rather than A/T at positions 1 and 2 of both half-sites.

To find out if these distinct characteristics of CRP-S sites are responsible for their regulatory differences, we first conducted a detailed analysis of *Hi*CRP binding to CRP-S and CRP-N sites using electrophoretic mobility shift (bandshift) assays; the well-characterized *Ec*CRP was used as a positive control and reference. Figure 1b compares the four CRP sites tested in these assays to the consensus CRP-binding site 'ICAP'.[21] Two of the promoters tested contained CRP-N sites, *H. influenzae* P_{mglB} (*mglBAC* operon) and *E. coli* P_{lacZ} (*lacZYA* operon), and two contained *H. influenzae* CRP-S sites, P_{pilA} (*pilABCD* operon) and P_{comA} (*comABCDE* operon). Binding reactions contained excess nonspecific competitor DNA to approximate *in vivo* conditions. Because this precluded measurement of equilibrium binding constants (K_{obs}) or dissociation constants (K_d), for Figs. 1 and 2, we report apparent dissociation constants (K_{app}) in terms of active protein. All protein dilution series, DNA preparations, and binding reactions were conducted at least twice and a single curve was fit to all data. Three independent preparations of *Hi*CRP and *Ec*CRP each gave very reproducible binding to P_{mglB}; minor differences between these preparations explain the scatter in the P_{mglB} binding data (Fig. 1c). Tests for nonspecific binding ('ns' in Figs. 1, 2, and 3) used three DNA sequences: P_{pilA} with its CRP-S site removed [P_{pilA}(ΔCRP-S)], coding sequence from the *E. coli hofB* gene (as used in Ref. 1), and the pUC19 multiple-cloning site.

*Hi*CRP's highest affinity was for ICAP (K_{app} = 1.1 ± 0.1 nM); binding curves are illustrated in Fig. 1c, and a representative bandshift is shown in Fig. 1g. Although P_{mglB} is a very close match to the consensus, *Hi*CRP demonstrated almost 50-fold lower affinity for this promoter (K_{app} = 57 ± 7 nM). *Hi*CRP did not bind P_{lacZ}, P_{pilA}, or P_{comA} until protein concentrations were high enough to elicit nonspecific DNA binding (>300 nM) (Fig. 1c and d). Moreover, binding to all

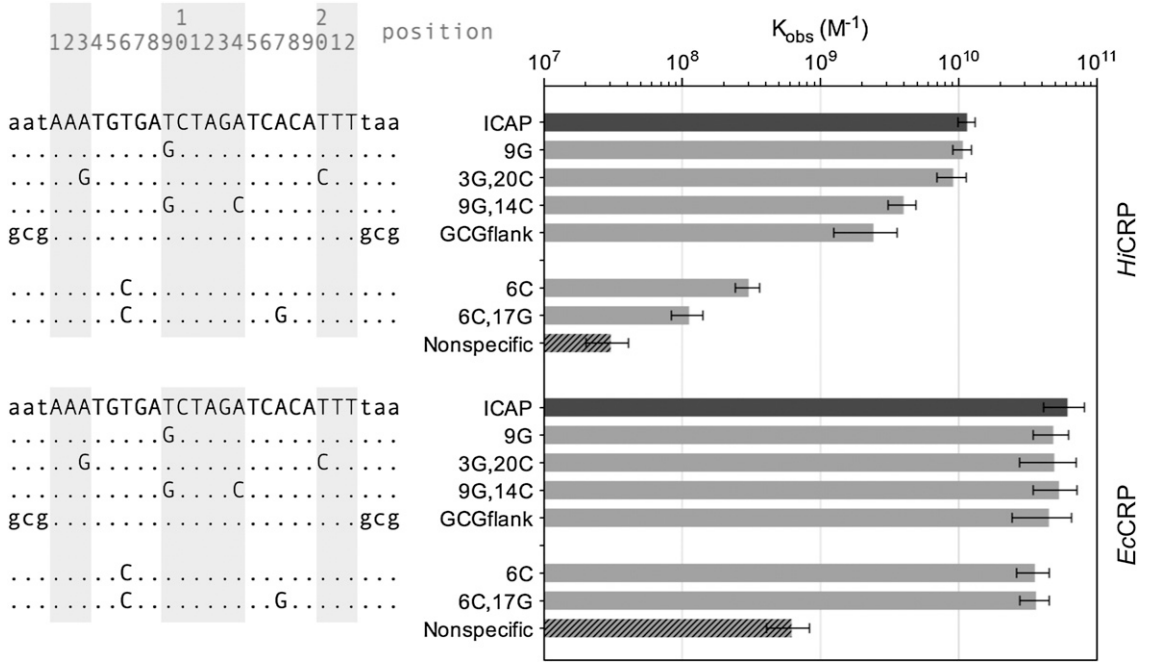

Fig. 3. Equilibrium binding constants (K_{obs}) measured for ICAP and variant sites using bandshift assays. ICAP and variant sites are aligned on the left as in Fig. 1b. K_{obs} for nonspecific DNA was measured using P_{pilA}(ΔCRP-S). The means and standard errors from replicate experiments are plotted.

three DNAs generated supershifts characteristic of multiple proteins binding to a single DNA molecule (Fig. 1h). The failure of CRP to significantly bind the P_{pilA} and P_{comA} sites is quite unexpected, given these promoters' almost complete dependence on CRP *in vivo*.[13]

*Ec*CRP demonstrated a high affinity for ICAP ($K_{app}=1.0\pm0.2$ nM) equal to that of *Hi*CRP. However, unlike *Hi*CRP, *Ec*CRP also had a high affinity for P_{mglB} ($K_{app}=1.6\pm0.3$ nM) and P_{lacZ} ($K_{app}=8.4\pm0.7$ nM) (Fig. 1e). *Ec*CRP demonstrated lower, but appreciable, affinities for the two CRP-S sites P_{pilA} ($K_{app}=75\pm10$ nM) and P_{comA} ($K_{app}=141\pm29$ nM) (Fig. 1f).

Together, these results show that *Hi*CRP and *Ec*CRP differ in at least two ways. First, because P_{lacZ}, P_{pilA}, and P_{comA} differ from the consensus at one, two, and three core positions, respectively (Fig. 1b), *Hi*CRP's 1000-fold preference for ICAP over these sites suggests that the protein is highly selective for the perfect core sequence $T_4G_5T_6G_7A_8$ in both half-sites. Second, *Hi*CRP greatly prefers the consensus sequence ICAP over P_{mglB}, indicating that *Hi*CRP affinity, unlike *Ec*CRP affinity, is sensitive to bases outside $T_4G_5T_6G_7A_8$.

DNA sequence determinants for *Hi*CRP binding to natural CRP sites

We tested whether the presence of nonconsensus bases at particular core positions was responsible for the poor *Hi*CRP binding to the P_{lacZ} and P_{pilA} sites by changing these positions to match the CRP-N core consensus (Fig. 2a). In P_{lacZ}, the nonconsensus base T_{19} was converted to A_{19} to generate P_{lacZ}-19A; in P_{pilA}, the CRP-S consensus bases C_6 and G_{17} were converted to the CRP-N consensus bases T_6 and A_{17} to generate P_{pilA}-N. Figure 2b and c shows that *Hi*CRP bound with significantly greater affinity to the mutant promoters: P_{lacZ}-19A ($K_{app}=116\pm37$ nM) and P_{pilA}-N ($K_{app}=60\pm9$ nM). Further, *Hi*CRP binding to P_{lacZ}-19A and P_{pilA}-N generated the discrete bandshifts characteristic of site-specific binding (Fig. 2f). Thus, in these reaction conditions, *Hi*CRP binding requires that DNA sites perfectly match the consensus $T_4G_5T_6G_7A_8$ core in both half-sites. Nevertheless, *Hi*CRP's affinity for P_{lacZ}-19A and P_{pilA}-N was much lower than that for ICAP despite their having identical perfect-consensus cores, confirming that bases outside of the $T_4G_5T_6G_7A_8$ core are important for *Hi*CRP binding.

Identical binding experiments revealed that *Ec*CRP has a 5-fold higher affinity for P_{lacZ}-19A than P_{lacZ} ($K_{app}=1.5\pm0.4$ nM *versus* 8.4 ± 0.7 nM) and a 44-fold higher affinity for P_{pilA}-N than P_{pilA} ($K_{app}=1.6\pm0.3$ nM *versus* 75 ± 10 nM) (Fig. 2d and e, respectively). Although these mutant promoters still differed from ICAP at six or eight non-core positions, *Ec*CRP exhibited almost identical affinities for them and for ICAP ($K_{app}=1.0\pm0.2$ nM). Thus, when core positions match the consensus, bases outside the core appear to make only a very minor contribution to *Ec*CRP–DNA binding.

DNA sequence determinants for *Hi*CRP binding to synthetic CRP sites

Two lines of evidence suggest that positions 3 and 9 (and the reciprocal positions 20 and 14) are important for *Hi*CRP binding. First, these positions have the most significant overrepresentation in the CRP-S logo in Fig. 1a. Second, these are three of the six nonconsensus positions in the low-affinity P_{lacZ}-19A site. To systematically address the importance of bases at these positions for *Hi*CRP binding, we constructed variants of ICAP and measured equilibrium binding constants using bandshift assays. These assays used low concentrations of bait DNA and no competitor DNA to allow comparison with established K_{obs} data for ICAP.

Figure 3 shows that *Hi*CRP's affinity for ICAP ($K_{obs}=1.2\pm0.2\times10^{10}$ M^{-1}) was fivefold less than *Ec*CRP's ($K_{obs}=6.1\pm2\times10^{10}$ M^{-1}); the *Ec*CRP binding constant measured here is consistent with the K_{obs} range previously measured using filter-binding assays ($4.2\pm0.3\times10^{10}$ M^{-1} to 7.1×10^{10} M^{-1}).[9,21] Changing ICAP at either $T_9\rightarrow G_9$ or $A_3\rightarrow C_3$ as well as $T_{20}\rightarrow C_{20}$ did not significantly change *Hi*CRP affinity. However, simultaneously substituting both $T_9\rightarrow G_9$ and $A_{14}\rightarrow C_{14}$ reduced *Hi*CRP affinity threefold. Thus, *Hi*CRP's low affinity for P_{lacZ}-19A is at least partly due to the nonconsensus bases immediately adjacent to the $T_4G_5T_6G_7A_8$ cores.

Although CRP binding sites are conventionally treated as 22 bp in length, we have noticed that CRP-S sites are usually flanked by additional A+T-rich sequence, as is ICAP.[21] We thus tested whether positions outside the 22-bp consensus sequence contribute to *Hi*CRP binding by converting ICAP's A+T-rich flanking sequence to G+C-rich sequences; this reduced affinity fivefold.

We also measured binding constants for ICAP variants containing one or both of the CRP-S-like bases C_6 and G_{17}. A single $T_6\rightarrow C_6$ substitution reduced *Hi*CRP affinity 38-fold, while having both C_6 and G_{17} reduced affinity 100-fold. Nevertheless, *Hi*CRP bound the doubly substituted ICAP with a 4-fold greater affinity than nonspecific DNA, indicating that having perfect matches to the consensus binding site at all other positions confers site specificity even in the presence of the unfavorable bases C_6 and G_{17}.

Unlike *Hi*CRP, binding constants of *Ec*CRP were not significantly affected by substitutions outside of the $T_4G_5T_6G_7A_8$ core (Fig. 3). Notably, substitutions at positions 6 and 17 did not significantly reduce *Ec*CRP affinity in these assays, although filter-binding assays by Chen *et al.* showed an 80-fold reduction for the double mutant.[9] This unexpected difference is unlikely to be due to our use of a shorter incubation time (20 min rather than 60 min) because the binding we see is unexpectedly high rather than unexpectedly low, as it would be if binding had not reached equilibrium. An alternative explanation may be that the energetics of DNA binding (and perhaps bending) differ between filter and gel assays.

Despite the 96% similarity between *Ec*CRP and *Hi*CRP's DNA-binding domains, all bandshift data indicated that *Hi*CRP is more selective than *Ec*CRP for DNA sites *in vitro*. We used SWISS-MODEL[22] to map *Hi*CRP residues on a crystal structure of *Ec*CRP bound to ICAP[9,10,23] and found the predicted *Hi*CRP–DNA structure to be completely congruous with the *Ec*CRP–DNA structure (Supplemental Fig. 1). The PROCHECK,[24] PROVE,[25] and WHAT IF[26] algorithms for validating protein structures confirmed that no *Hi*CRP residues are predicted to reshape the DNA-binding domain or to sterically interfere with the protein–DNA interactions that are known to occur in *E. coli*.

CRP binding to promoter DNA is insufficient to stimulate transcription in the absence of Sxy

The inability of *Hi*CRP to bind specifically to P_{comA} and P_{pilA} *in vitro* despite its strong regulatory effect *in vivo* raised the possibility that CRP requires Sxy to facilitate its binding at CRP-S sites. We had two ways to test this hypothesis *in vivo*: first, by testing whether *Ec*CRP's intrinsic affinity for CRP-S sites is sufficient for transcription activation and, second, by testing whether *Hi*CRP alone can activate transcription from the P_{pilA}-N promoter.

*Ec*CRP is known to restore natural transformability to an *H. influenzae crp* null mutant,[27] but the dependence of this complementation on Sxy has not been tested. The *E. coli crp* gene was expressed in *H. influenzae crp*− and *sxy*− mutants to test whether *Ec*CRP is able to activate the CRP-S sites of *H. influenzae* competence gene promoters in the absence of Sxy. Real-time (quantitative) PCR was used to measure transcription of *comA* 60 min after transfer to the competence-inducing medium MIV, when competence genes are usually maximally expressed.[13] Transformation frequency, which provides the most sensitive assay of competence gene induction, was measured 90 min after cells were transferred. As expected, cells were not transformable in the absence of *crp* or *sxy*, but *Ec*CRP restored transformability to a *crp*− strain (RR815) (Fig. 4a).[15,27] Also, as previously reported, *comA* expression was induced over 100-fold in MIV, but induction required *crp* and *sxy*.[13] However, even though *Ec*CRP binds P_{comA} with high affinity *in vitro*, it could not restore transformability or *comA* expression in *crp*+/*sxy*− cells (RR1130). To confirm that the absence of competence induction was not due to interference between *Ec*CRP and *Hi*CRP in the *sxy*− strain, we measured transformation in wild-type *H. influenzae* carrying *Ec*CRP (RR1129), and these cells were found to be fully transformable (last column, Fig. 4a). An alternative explanation of the Sxy-dependent competence of cells containing both *Hi*CRP and *Ec*CRP (strain RR1130) would be that all of the *Ec*CRP is sequestered in inactive heterodimers, with the observed expression of competence genes due entirely to active *Hi*CRP homodimers. However, this could only occur if *Hi*CRP were in excess to *Ec*CRP, whereas the reverse is much more likely because in these experiments, $crp_{E.coli}$ is plasmid-borne and driven by a constitutive P_{lacZ} promoter while $crp_{H.influenzae}$ is chromosomal and driven by its own CRP-repressible promoter.

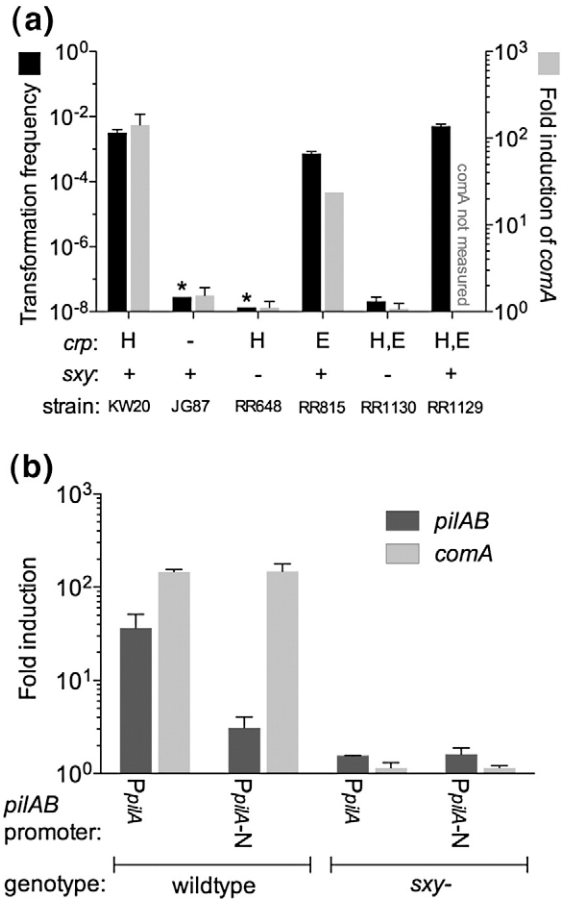

Fig. 4. Transformation and competence gene induction in different genetic backgrounds. (a) *Ec*CRP complementation of natural transformation and *comA* induction in *H. influenzae* cells. The *x*-axis indicates the source of the *crp* gene (H, *H. influenzae*; E, *E. coli*) and the presence or absence of the *sxy* gene. Transformation frequency (black bars) and *comA* expression (gray bars) were measured after 90 and 60 min, respectively, in competence-inducing medium (MIV). Asterisks on transformation frequencies denote upper limits of transformation assays that produced no transformants. Fold induction of *comA* was measured using real-time PCR and is plotted relative to the baseline expression of the gene in exponentially growing (uninduced) wild-type cells. The mean and range of induction in two independent cultures are plotted on a log scale; the exception was measurement of *comA* expression in only one culture of the control strain RR815, but replicated high transformation frequencies in RR815 cultures imply high *comA* expression levels. (b) P_{pilA}, P_{pilA}-N, and P_{comA} activity in wild-type and *sxy*− *H. influenzae* cells. Gene expression was measured before and 60 min after transfer of cells to MIV, using real-time PCR. The fold induction levels of *pilAB* (dark bars) and *comA* (chromosomal control gene; light gray bars) after 60 min are plotted relative to expression levels in the uninduced state before transfer to MIV. The mean and range of induction in two or three independent cultures are plotted on log scales. *pilAB* and *comA* baseline expression levels were consistent in all strains.

The experiments reported in Fig. 2 showed that replacing the CRP-S-site bases C_6 and G_{17} of P_{pilA} with their CRP-N-site counterparts (P_{pilA}-N) permitted *Hi*CRP to bind this site *in vitro* (in the absence of Sxy). To find out whether this change also allowed *Hi*CRP to bind P_{pilA}-N and stimulate transcription *in vivo* in the absence of Sxy, we used real-time PCR to measure *pilAB* transcription originating from the promoters of the engineered plasmids. Expression of *comA* served as a positive control to confirm the presence/absence of the competence-inducing signal in *sxy*+/− backgrounds, respectively. Because the $C_6 \rightarrow T_6/G_{17} \rightarrow A_{17}$ substitutions improved CRP binding *in vitro*, these changes were not predicted to affect the amount of transcription activation in wild-type cells. Surprisingly, P_{pilA}-N was induced only 3-fold in MIV, significantly less than the 35-fold induction of P_{pilA} (Fig. 4b). The conversion of $C_6 \rightarrow T_6/G_{17} \rightarrow A_{17}$ was hypothesized to relieve Sxy dependence, but even the modest 3-fold induction of P_{pilA}-N was dependent on Sxy. Together, the *Ec*CRP-P_{comA} and *Hi*CRP-P_{pilA}-N results provide complementary evidence that Sxy remains essential for transcription activation *in vivo* even when CRP can efficiently bind CRP-S-regulated promoters *in vitro*.

A+T runs upstream of CRP-S sites are required for promoter activation

The above analysis strongly suggested that CRP-S-regulated promoters contain additional elements outside of their CRP-S sites. Previous searches for putative Sxy binding sites in *H. influenzae* did not identify any conserved motifs other than CRP-S sites,[13] but the search algorithms were insensitive to short (<10 bp) motifs. To detect additional conserved elements in *H. influenzae*'s CRP-S-regulated promoters, we generated a sequence logo from alignment of the 13 promoter sequences at their CRP-S sites, including 200 bases upstream of predicted transcription start sites. Most CRP-S sites were predicted to be located around −61.5 from transcription start points,[13] and this numbering is used in Fig. 5. The σ70 −35 and −10 sites apparent in the logo validates our earlier analysis of CRP-S site location and provides strong evidence that competence genes are regulated by σ70 in *H. influenzae*. The sequence logo also revealed striking A+T runs at positions −79, −90, and −102 (runs II, III, and IV), with an additional run in the upstream end of the CRP-S site (run I, at position −71). These A+T runs resemble the RNAP αCTD binding sites called UP elements;[29] contact between αCTD and UP elements enhances transcription from 2- to >100-fold at many bacterial promoters.[29–31] Although UP elements are usually located between −40 and −60, DNA bending by CRP or IHF has been shown to allow αCTD to bind UP elements at −80 and −90.[32,33]

P_{pilA} variants in which the A+T runs were replaced or translocated were constructed to test whether these A+T runs are important for promoter activity. As shown in Fig. 5b, replacement of element II (at −79) with a G+C-rich SacII site (CCCGGG)

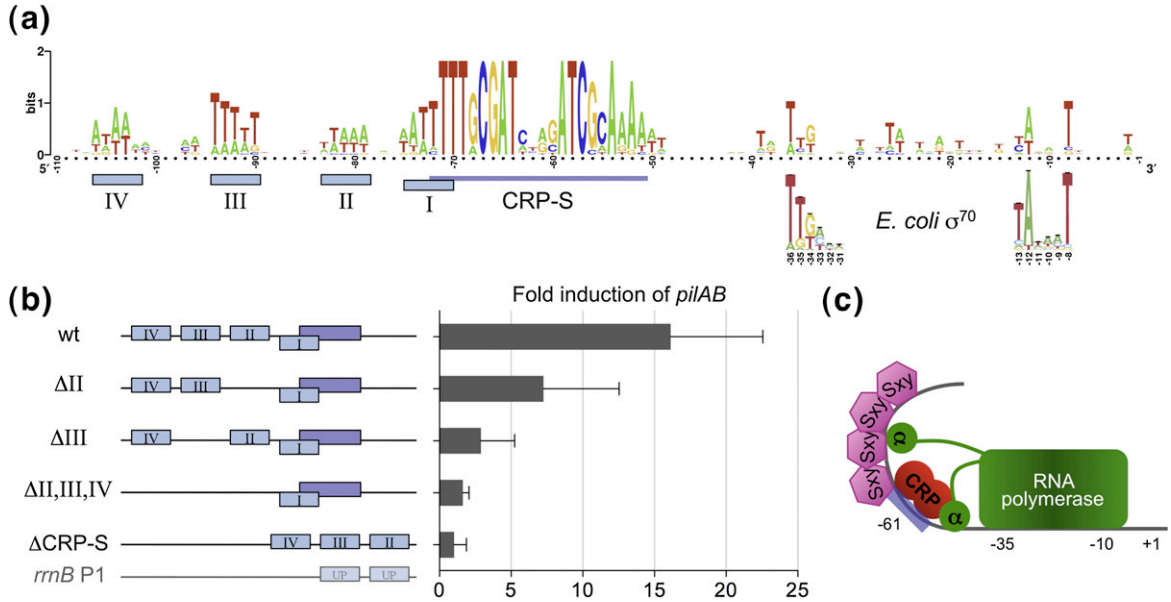

Fig. 5. A+T runs upstream of CRP-S sites are important for promoter activation. (a) Sequence logo generated from alignment of *H. influenzae*'s 13 CRP-S sites. Numbering indicates the average distance of CRP-S sites from predicted transcription start sites identified in Redfield *et al.*;[13] A+T runs are underlined blue and numbered. The sequence logo generated from alignment of 401 *E. coli* σ70 binding sites (copied from Ref. 28) facilitates comparison with similar motifs in the *H. influenzae* logo above. (b) Schematic of P_{pilA} mutants and *E. coli*'s *rrnB P1*. Fold induction of *pilAB* after 60 min in MIV is plotted as in Fig. 4b. Baseline *pilAB* expression in the uninduced state was the same in all strains. (c) Proposed model for the CRP–Sxy–RNAP nucleoprotein complex formed at promoters with CRP-S sites. By binding A+T runs, Sxy could increase DNA curvature, assist CRP binding, and directly assist in RNAP recruitment through contact with αCTD.

reduced expression by half. Replacement of element III (at −90) by a SacII site reduced expression by more than eightfold, whereas simultaneous deletion of elements II, III, and IV completely eliminated inducibility. The A+T runs might be important for transcription because CRP-mediated DNA bending allows RNAP's αCTD subunits to bind them. To test whether the A+T runs are enhancers of RNAP binding, elements II and III were moved to a position immediately adjacent to RNAP's −35 binding site, thus aligning them with the UP elements of the model UP-enhanced promoter *rrnB* P1[34] (Fig. 5b). This translocation did not alter or enhance promoter activity during exponential growth; neither did it allow the promoter to be induced in MIV. Transformation frequencies and *comA* expression confirmed that competence induction was normal in all strains (data not shown). These results suggest that the A+T runs cannot function as UP elements to recruit RNAP and also show that the CRP-S site is essential for promoter induction.

*Ec*CRP bound normally to the ΔII,III,IV promoter, indicating that the A+T runs are not important for CRP binding (data not shown). Unfortunately, Sxy remains recalcitrant to overexpression and purification; hence, bandshifts could not be used to investigate whether Sxy binds the A+T runs.

Discussion

This study of the molecular mechanisms regulating competence genes in *H. influenzae* is both the first detailed analysis of CRP binding to CRP-S promoters and the first biochemical analysis of CRP from a member of the Pasteurellaceae. Although *Hi*CRP, like *Ec*CRP, demonstrated a high affinity for the consensus CRP site (ICAP), it was much more sensitive than the promiscuous *Ec*CRP to base composition outside the $T_4G_5T_6G_7A_8$ core. This is surprising given the very high level of similarity of the sequence of its DNA-binding domain to that of *Ec*CRP. Indeed, homology modeling of *Hi*CRP bound to ICAP predicted a protein structure identical with that of *Ec*CRP. The explanation may be that *Hi*CRP and *Ec*CRP are both capable of forming the same protein–DNA contacts, but *Hi*CRP–DNA interactions are more transient because of features in the protein's less conserved N-terminal domain. For example, *Hi*CRP dimers may be less stable than *Ec*CRP dimers, thus requiring more optimal CRP sites for *Hi*CRP binding in the absence of other cellular proteins. Dimer stability has been previously measured for *Ec*CRP,[35] and similar experiments may help explain the DNA-binding characteristics of *Hi*CRP.

Unlike substitutions at other core positions, which eliminate important protein–DNA bonds, the $T_6 \rightarrow C_6$ substitution inhibits CRP binding by generating a C_6/G_7 base step that increases the free energy required for DNA kinking.[9] This unique feature of a $T_6 \rightarrow C_6$ substitution suggests a distinct mechanistic role for the C_6/G_7 base step in CRP-S function and promoter regulation. *H. influenzae*'s CRP-S sites are distinguished from CRP-N sites not only by having the unfavorable C_6/G_7 base step in both core half-sites but also by having a higher frequency of favorable bases at other core positions and non-core positions. Thus, CRP-S sites appear to have evolved to facilitate targeting of CRP (extensive hydrogen bonding between CRP and DNA) but prevent activation (CRP alone cannot form an activating complex).

DNA topology plays an important role in bacterial promoter regulation (reviewed in Refs. 36–38). The dramatic (~90°) deformation of DNA caused by CRP binding is thought to stimulate transcription both by bringing upstream promoter elements into contact with the polymerase and by facilitating DNA strand separation;[2,39,40] for example, because CRP sites at the *gal* and *malK* promoters can be functionally replaced by intrinsically bent DNA, CRP's primary role at these sites is to bend DNA.[41,42] The inflexibility of CRP-S sites may be specially selected to prevent upstream DNA and/or proteins from interacting with the promoter except in the presence of CRP-induced kinking. However, our results suggest that CRP binding alone cannot stimulate transcription in the absence of Sxy, indicating that CRP-induced kinking alone is insufficient for transcription activation. Furthermore, translocating upstream A+T runs to a position adjacent to the RNAP binding site did not stimulate transcription. Thus, we favor a model in which Sxy binds to the A+T runs and directly assists in RNAP recruitment (Fig. 5c). CRP-induced bending would be essential for this interaction because, in its absence, Sxy bound to A+T runs would be too far upstream to recruit RNAP to the −35 and −10 sites. CRP-S-regulated promoters are unusually strong—global analysis of the *H. influenzae* transcriptome revealed that, in starvation conditions, most promoters with CRP-S sites are induced much more strongly than CRP-N-regulated promoters[13]—suggesting that, once formed, the putative CRP–Sxy–DNA nucleoprotein complex is very efficient at recruiting RNAP.

We have previously found that *Ec*CRP has very low affinity for its cognate CRP-S promoters.[1] Thus, a hallmark of both Enterobacteriaceae and Pasteurellaceae CRP-S sites is that they are low-affinity binding sites for cognate CRP. This low affinity is achieved in a species-specific fashion that corresponds to CRP's site selectivity. In *H. influenzae*, CRP-S sequences are all strong matches to the CRP-binding site consensus but always include a stiff C_6/G_7 base step that prevents DNA binding by *Hi*CRP in the absence of other proteins. *E. coli* CRP-S sites (which were originally identified in homologs of *H. influenzae* CRP-S-regulated genes) differ from the CRP-binding site consensus at many positions and always have either a C_6/G_7 or a G_6/G_7 base step.[11] Thus, CRP-S sites, may serve to create tight regulation without requiring a repressor: basal levels of transcription are very low because the inflexible C_6/G_7 base step limits occupancy by CRP, but expression levels can be very high once activated with the assistance of Sxy.

Materials and Methods

Strains and culture conditions

H. influenzae cells were cultured at 37 °C in supplemented brain heart infusion (sBHI), BHI supplemented with NAD (2 μg/ml) and hemin (10 μg/ml). Novobiocin (2.5 μg/ml), kanamycin (7 μg/ml), or chloramphenicol (2 μg/ml) was added to sBHI when required. *H. influenzae* strains listed in Table 1 were constructed by transformation of competent cells with chromosomal or plasmid DNA as previously described.[48] To induce competence, exponentially growing (noncompetent) cells were transferred from sBHI to the defined starvation medium MIV.[49] Transformation frequency was measured as the ratio of novobiocin-resistant transformants to total cells after 24 h incubation on sBHI agar with novobiocin. *E. coli* was cultured in Luria–Bertani broth and made chemically competent with RbCl and transformed as previously described;[50] chloramphenicol (25 μg/ml) was added when required.

Protein purification and bandshifts

The histidine-tagged CRP proteins were constructed as follows: the $crp_{H.influenzae}$ open reading frame was PCR amplified using primers Hi-crp-F and Hi-crp-R (Supplemental Table 1) and cloned in the His-tag vector pQE-30UA (Qiagen) by directly ligating PCR amplicons into the pQE-30UA vector; the $crp_{E.coli}$ open reading frame was cloned by Peekhaus and Conway in pQE30 (Qiagen) by restriction digestion of $crp_{E.coli}$ amplicons and the pQE30 vector as described in Ref. 46. His-tagged proteins were expressed and purified as previously described.[1] The fraction of CRP active in sequence-specific DNA binding (usually ~20%) was assessed by titration of ICAP in the absence of competitor DNA.

CRP–DNA binding reactions (5 μl) contained 8 mM Tris (pH8.0), 30 mM KCl, 3% (v/v) glycerol, 250 μg/ml bovine serum albumin, 100 μM cAMP, and 1 mM dithiothreitol. Reactions were mixed on ice and then incubated at room temperature for 20 min before being loaded onto a 4 °C running polyacrylamide gel {30:1 acrylamide/bisacrylamide; 0.2×TBE [89 mM Tris, 89 mM borate, and 2 mM ethylenediaminetetraacetic acid (pH8.3)], 2% glycerol, and 200 μM cAMP; running buffer 0.2×TBE and 100 μM cAMP}. After electrophoresis for 2 h at 10 mA, the gel was dried and exposed (45 min to overnight) to a phosphor screen. Bands were visualized using a STORM 860 scanner (GE Healthcare). Shifted and unshifted radioactivity was quantified using Image Quant (GE Healthcare), and background radioactivity in gel lanes with no CRP was subtracted. K_{app} and K_{obs} values were calculated by Prism 5.0 (GraphPad Software) using nonlinear regression analysis for one-site (specific) binding with B_{max} constrained to ≤100% binding.

Bait DNAs were PCR amplified from chromosomal or plasmid templates (primers are listed in Supplemental Table 1) and ranged in size from 90 to 200 bp—these sizes of bait DNA enabled us to accurately measure the amount of bait DNA in bandshift reactions using real-time PCR. Amplicons were purified using polyacrylamide gel electrophoresis; bands were excised and DNA was eluted from macerated gel overnight in TE at 37 °C, followed by ethanol precipitation and resuspension in 10 mM Tris. DNA was end-labeled with T4 polynucleotide kinase using a 10-fold molar excess of γ-[^{32}]P ATP, and unincorporated label was removed using illustra ProbeQuant G-50 Micro Columns (GE Healthcare). Bandshifts contained 35 ng/μl poly(dI–dC) cold competitor DNA and 0.1–2 nM bait DNA (~2100 CPM/fmol) to measure K_{app}. For K_{obs} measurements, protein was in at least 5-fold excess over bait DNA (0.01–0.02 nM).

Table 1. Strains and plasmids used in this work

Strain or plasmid	Relevant genotype	Source or reference
H. influenzae		
KW20	Wild-type Rd; sequenced strain	43
MAP7	KW20 Nov^r	44
JG87	KW20 *crp*: :Kan^r	45
RR648	KW20 *sxy*: :Kan^r	15
RR815	KW20 *crp*: :Kan^r pXN13	27
RR1129	KW20 pXN13	This study
RR1130	KW20 *sxy*: :Kan^r pXN13	This study
E. coli		
DH5α	*F80lacZ Δ(lacIZYA-argF) endA1*	
M15	pREP4	Qiagen
RR1234	M15 pQE*Hi.crp*	This study
Plasmids		
pGEM-T Easy		Promega
pQE-30UA		Qiagen
pREP4	*lacI*q	Qiagen
pQE*Hi.crp*	$crp_{H.influenzae}$ in pQE-30UA	This study
pNP-52	$crp_{E.coli}$ in pQE30	46
pSU20	Cam^r, P_{lacZ}	47
p*pilA*	pSU20 containing P_{pilA}, *pilA*, and *pilB*	This study
p*pilA2*	P_{lacZ} removed from p*pilA*	This study
p*pilA2*: :ΔII	A+T run II replaced in p*pilA2*	This study
p*pilA2*: :ΔIII	A+T run III replaced in p*pilA2*	This study
p*pilA2*: :ΔII, III,IV	A+T runs II,III,IV removed from p*pilA2*	This study
p*pilA2*: :ΔCRP-S	CRP-S site removed from p*pilA2*	This study

Nov^r, novobiocin resistance; Kan^r, kanamycin resistance; Cam^r, chloramphenicol resistance.

Cloning and site-directed mutagenesis

Complementary oligonucleotides (2 μM) were annealed in 50 μl of buffer (10 mM Tris, 10 mM $MgCl_2$, 50 mM NaCl, and 1 mM dithiothreitol, pH7.9) by heating to 95 °C followed by gradual cooling (−0.1 °C every 1.5 min) to 4 °C to clone ICAP and variants. Double-stranded oligonucleotides were then cloned using a pTOPO TA cloning kit (Invitrogen) and sequenced.

The chromosomal region (coordinates 333193–335531) containing *ampD*, *pilA*, and the N-terminal half of *pilB* was PCR amplified and cloned in pGEM-T Easy (Promega) to clone P_{pilA}. An AccI digest was used to excise a fragment containing P_{pilA}, *pilA*, *pilB*, and 20 bp of the multiple-cloning site. This fragment was cloned in the AccI site in the *H. influenzae* cloning vector pSU20[47] to generate plasmid p*pilA*. *H. influenzae* does not have the *lacI* repressor gene, and we found pSU20's P_{lacZ} to be constitutively expressed; thus, the promoter was removed by an XmnI and XhoI double digest. After DNA was purified on an agarose gel, the sticky end generated by XhoI was filled using the Klenow fragment of DNA Polymerase I and then was ligated to the XmnI-cut blunt end to generate plasmid p*pilA2*.

Site-directed mutations were generated using Stratagene's QuikChange Site-Directed Mutagenesis Kit accord-

ing to the manufacturer's instructions (mutagenic primers listed in Supplemental Table 1). The conversion of P_{lacZ} to P_{lacZ}-19A was conducted by site-directed mutation of pSU20. The conversion of P_{pilA} to P_{pilA}-N was conducted by site-directed mutation of p*pilA*2, and all P_{pilA} variants listed in Fig. 5 were also constructed in p*pilA*2. The CRP-S site was removed from p*pilA*2 using inverse PCR with *pfu* DNA polymerase and primers ppilA/-35-F and ppilA/UPII-R, which were designed to bind sequence flanking the CRP-S site. Amplicons contained full p*pilA*2 sequence except for the CRP-S site and were circularized using blunt-end ligation. The ΔII,III,IV promoter (Fig. 5) was constructed by a SacII digest of p*pilA*2::ΔII plasmid, which cut both the SacII site used to replace element II and a second SacII site ~200 bp upstream; recircularizing the digested plasmid placed vector DNA adjacent to element I.

Real-time (quantitative) PCR

Total RNA was isolated from cultures using RNeasy Mini Kits (Qiagen), and purity and quality were assessed by electrophoresis on 1% agarose (1×TAE). RNA was DNase treated twice with a DNA Free kit (Ambion), followed by cDNA synthesis using the iScript cDNA synthesis kit (Bio-Rad). For each PCR primer set (Supplemental Table 1), reactions were carried out in duplicate on an Opticon 2 system (Bio-Rad) using iTaq SYBR Green Supermix (Bio-Rad). *murG* RNA served as an internal control for normalization of each sample because this gene's expression is constant in the culture conditions used in this study.[13] Standard curves were generated using six serial 10-fold dilutions of MAP7 chromosomal DNA or p*pilA*2 plasmid DNA. PCR primers pilB-RT-F and ppilA2-RT-R were designed to flank the junction of *pilB* and the multiple-cloning site in the plasmid p*pilA*2 in order to exclusively measure *pilAB* transcripts originating from engineered plasmid-borne promoters. This primer set did not generate amplicons from chromosomal DNA template, confirming that it targets only plasmid-encoded *pilAB* transcripts.

Acknowledgements

The authors thank Mike Nichols for technical support, Steven Hallam for use of a real-time PCR machine, and Tyrrell Conway for providing plasmid pNP-52. This work was supported by a University of British Columbia Graduate Research Fellowship to ADSC and an Operating Grant from the Canadian Institutes for Health Research to RJR.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2008.08.027

References

1. Cameron, A. D. & Redfield, R. J. (2006). Non-canonical CRP sites control competence regulons in *Escherichia coli* and many other gamma-proteobacteria. *Nucleic Acids Res.* **34** , 6001–6014.
2. Lawson, C. L., Swigon, D., Murakami, K. S., Darst, S. A., Berman, H. M. & Ebright, R. H. (2004). Catabolite activator protein: DNA binding and transcription activation. *Curr. Opin. Struct. Biol.* **14** , 10–20.
3. 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.
4. Busby, S. & Ebright, R. H. (1999). Transcription activation by catabolite activator protein (CAP). *J. Mol. Biol.* **293** , 199–213.
5. Meibom, K. L., Blokesch, M., Dolganov, N. A., Wu, C. Y. & Schoolnik, G. K. (2005). Chitin induces natural competence in *Vibrio cholerae*. *Science*, **310**, 1824–1827.
6. Schultz, S. C., Shields, G. C. & Steitz, T. A. (1991). Crystal structure of a CAP–DNA complex: the DNA is bent by 90 degrees. *Science*, **253**, 1001–1007.
7. Gunasekera, A., Ebright, Y. W. & Ebright, R. H. (1992). DNA sequence determinants for binding of the *Escherichia coli* catabolite gene activator protein. *J. Biol. Chem.* **267**, 14713–14720.
8. Mandel-Gutfreund, Y., Margalit, H., Jernigan, R. L. & Zhurkin, V. B. (1998). A role for CH.O interactions in protein–DNA recognition. *J. Mol. Biol.* **277** , 1129–1140.
9. Chen, S., Gunasekera, A., Zhang, X., Kunkel, T. A., Ebright, R. H. & Berman, H. M. (2001). Indirect readout of DNA sequence at the primary-kink site in the CAP–DNA complex: alteration of DNA binding specificity through alteration of DNA kinking. *J. Mol. Biol.* **314**, 75–82.
10. Parkinson, G., Wilson, C., Gunasekera, A., Ebright, Y. W., Ebright, R. E. & Berman, H. M. (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.
11. Chen, S., Vojtechovsky, J., Parkinson, G. N., Ebright, R. H. & Berman, H. M. (2001). Indirect readout of DNA sequence at the primary-kink site in the CAP–DNA complex: DNA binding specificity based on energetics of DNA kinking. *J. Mol. Biol.* **314**, 63–74.
12. Salgado, H., Gama-Castro, S., Peralta-Gil, M., Diaz-Peredo, E., Sanchez-Solano, F., Santos-Zavaleta, A. *et al.* (2006). RegulonDB (version 5.0): *Escherichia coli* K-12 transcriptional regulatory network, operon organization, and growth conditions. *Nucleic Acids Res.* **34** , D394–397.
13. Redfield, R. J., Cameron, A. D., Qian, Q., Hinds, J., Ali, T. R., Kroll, J. S. & Langford, P. R. (2005). A novel CRP-dependent regulon controls expression of competence genes in *Haemophilus influenzae*. *J. Mol. Biol.* **347**, 735–747.
14. Redfield, R. J. (1991). *sxy-1*, a *Haemophilus influenzae* mutation causing greatly enhanced spontaneous competence. *J. Bacteriol.* **173**, 5612–5618.
15. Williams, P. M., Bannister, L. A. & Redfield, R. J. (1994). The *Haemophilus influenzae sxy-1* mutation is in a newly identified gene essential for competence. *J. Bacteriol.* **176**, 6789–6794.
16. Zulty, J. J. & Barcak, G. J. (1995). Identification of a DNA transformation gene required for *com101A*+ expression and supertransformer phenotype in *Haemophilus influenzae*. *Proc. Natl Acad. Sci. USA*, **92**, 3616–3620.
17. Karudapuram, S. & Barcak, G. J. (1997). The *Haemophilus influenzae dprABC* genes constitute a competence-inducible operon that requires the product of the *tfoX*

(*sxy*) gene for transcriptional activation. *J. Bacteriol.* **179**, 4815–4820.
18. Bhattacharjee, M. K., Fine, D. H. & Figurski, D. H. (2007). *tfoX* (*sxy*)-dependent transformation of *Aggregatibacter* (*Actinobacillus*) *actinomycetemcomitans*. *Gene*, **399**, 53–64.
19. Cameron, A. D., Volar, M., Bannister, L. A. & Redfield, R. J. (2008). RNA secondary structure regulates the translation of *sxy* and competence development in *Haemophilus influenzae*. *Nucleic Acids Res.* **36**, 10–20.
20. Robison, K., McGuire, A. M. & Church, G. M. (1998). A comprehensive library of DNA-binding site matrices for 55 proteins applied to the complete *Escherichia coli* K-12 genome. *J. Mol. Biol.* **284**, 241–254.
21. Ebright, R. H., Ebright, Y. W. & 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.
22. Schwede, T., Kopp, J., Guex, N. & Peitsch, M. C. (2003). SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res.* **31**, 3381–3385.
23. Napoli, A. A., Lawson, C. L., Ebright, R. H. & Berman, H. M. (2006). Indirect readout of DNA sequence at the primary-kink site in the CAP–DNA complex: recognition of pyrimidine–purine and purine–purine steps. *J. Mol. Biol.* **357**, 173–183.
24. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thorton, J. M. (1993). PROCHECK—a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291.
25. Pontius, J., Richelle, J. & Wodak, S. J. (1996). Deviations from standard atomic volumes as a quality measure for protein crystal structures. *J. Mol. Biol.* **264**, 121–136.
26. Vriend, G. (1990). WHAT IF: a molecular modeling and drug design program. *J. Mol. Graphics*, **8**, 52–56, 29.
27. Chandler, M. S. (1992). The gene encoding cAMP receptor protein is required for competence development in *Haemophilus influenzae Rd*. *Proc. Natl Acad. Sci. USA*, **89**, 1626–1630.
28. Shultzaberger, R. K., Chen, Z., Lewis, K. A. & Schneider, T. D. (2007). Anatomy of *Escherichia coli* sigma70 promoters. *Nucleic Acids Res.* **35**, 771–788.
29. Gourse, R. L., Ross, W. & Gaal, T. (2000). UPs and downs in bacterial transcription initiation: the role of the alpha subunit of RNA polymerase in promoter recognition. *Mol. Microbiol.* **37**, 687–695.
30. Ross, W., Aiyar, S. E., Salomon, J. & Gourse, R. L. (1998). *Escherichia coli* promoters with UP elements of different strengths: modular structure of bacterial promoters. *J. Bacteriol.* **180**, 5375–5383.
31. Ross, W. & Gourse, R. L. (2005). Sequence-independent upstream DNA–alphaCTD interactions strongly stimulate *Escherichia coli* RNA polymerase-lacUV5 promoter association. *Proc. Natl Acad. Sci. USA*, **102**, 291–296.
32. Giladi, H., Koby, S., Prag, G., Engelhorn, M., Geiselmann, J. & Oppenheim, A. B. (1998). Participation of IHF and a distant UP element in the stimulation of the phage lambda PL promoter. *Mol. Microbiol.* **30**, 443–451.
33. Lee, D. J., Busby, S. J. & Lloyd, G. S. (2003). Exploitation of a chemical nuclease to investigate the location and orientation of the *Escherichia coli* RNA polymerase alpha subunit C-terminal domains at simple promoters that are activated by cyclic AMP receptor protein. *J. Biol. Chem.* **278**, 52944–52952.
34. Meng, W., Belyaeva, T., Savery, N. J., Busby, S. J., Ross, W. E., Gaal, T. *et al.* (2001). UP element-dependent transcription at the *Escherichia coli rrnB* P1 promoter: positional requirements and role of the RNA polymerase alpha subunit linker. *Nucleic Acids Res.* **29**, 4166–4178.
35. Brown, A. M. & Crothers, D. M. (1989). Modulation of the stability of a gene-regulatory protein dimer by DNA and cAMP. *Proc. Natl Acad. Sci. USA*, **86**, 7387–7391.
36. Perez-Martin, J. & de Lorenzo, V. (1997). Clues and consequences of DNA bending in transcription. *Annu. Rev. Microbiol.* **51**, 593–628.
37. McLeod, S. M. & Johnson, R. C. (2001). Control of transcription by nucleoid proteins. *Curr. Opin. Microbiol.* **4**, 152–159.
38. Barnard, A., Wolfe, A. & Busby, S. (2004). Regulation at complex bacterial promoters: how bacteria use different promoter organizations to produce different regulatory outcomes. *Curr. Opin. Microbiol.* **7**, 102–108.
39. Eichenberger, P., Dethiollaz, S., Buc, H. & Geiselmann, J. (1997). Structural kinetics of transcription activation at the malT promoter of *Escherichia coli* by UV laser footprinting. *Proc. Natl Acad. Sci. USA*, **94**, 9022–9027.
40. Coulombe, B. & Burton, Z. F. (1999). DNA bending and wrapping around RNA polymerase: a "revolutionary" model describing transcriptional mechanisms. *Microbiol. Mol. Biol. Rev.* **63**, 457–478.
41. 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.
42. Richet, E. & Sogaard-Andersen, L. (1994). CRP induces the repositioning of MalT at the *Escherichia coli malKp* promoter primarily through DNA bending. *EMBO J.* **13**, 4558–4567.
43. Alexander, H. E. & Leidy, G. (1951). Determination of inherited traits of *H. influenzae* by desoxyribonucleic acid fractions isolated from type-specific cells. *J. Exp. Med.* **93**, 345–359.
44. Barcak, G. J., Chandler, M. S., Redfield, R. J. & Tomb, J. F. (1991). Genetic systems in *Haemophilus influenzae*. *Methods Enzymol.* **204**, 321–342.
45. Tomb, J. F., Barcak, G. J., Chandler, M. S., Redfield, R. J. & Smith, H. O. (1989). Transposon mutagenesis, characterization, and cloning of transformation genes of *Haemophilus influenzae* Rd. *J. Bacteriol.* **171**, 3796–3802.
46. Peekhaus, N. & Conway, T. (1998). Positive and negative transcriptional regulation of the *Escherichia coli* gluconate regulon gene *gntT* by GntR and the cyclic AMP (cAMP)–cAMP receptor protein complex. *J. Bacteriol.* **180**, 1777–1785.
47. Bartolome, B., Jubete, Y., Martinez, E. & de la Cruz, F. (1991). Construction and properties of a family of pACYC184-derived cloning vectors compatible with pBR322 and its derivatives. *Gene*, **102**, 75–78.
48. Poje, G. & Redfield, R. J. (2003). General methods for culturing *Haemophilus influenzae*. *Methods Mol. Med.* **71**, 51–56.
49. Herriott, R. M., Meyer, E. M. & Vogt, M. (1970). Defined nongrowth media for stage II development of competence in *Haemophilus influenzae*. *J. Bacteriol.* **101**, 517–524.
50. Ausubel, F. M. (1995). *Current Protocols in Molecular Biology.* J. Wiley and Sons, Inc., Brooklyn, NY.