*Published online 17 July 2014* *Nucleic Acids Research, 2014, Vol. 42, No. 14* ***9209–9216***
*doi: 10.1093/nar/gku625*

# Spacing requirements for Class I transcription activation in bacteria are set by promoter elements

**Yi Zhou[1], Annie Kolb[2], Stephen J.W. Busby[3] and Yi-Ping Wang[1,*]**

[1]State Key Laboratory of Protein and Plant Gene Research, College of Life Sciences, Peking University, Beijing 100871, PR China, [2]Molecular Genetics Unit and CNRS URA-2172, Institut Pasteur, 25 rue du Dr Roux, 75724 Paris Cedex 15, France and [3]School of Biosciences and Institute of Microbiology and Infection, University of Birmingham, Birmingham B15 2TT, UK

Received April 10, 2014; Revised June 16, 2014; Accepted June 28, 2014

## ABSTRACT

**The *Escherichia coli* cAMP receptor protein (CRP) activates transcription initiation at many promoters by binding upstream of core promoter elements and interacting with the C-terminal domain of the RNA polymerase α subunit. Previous studies have shown stringent spacing is required for transcription activation by CRP. Here we report that this stringency can be altered by the nature of different promoter elements at target promoters. Several series of CRP-dependent promoters were constructed with CRP moved to different upstream locations, and their activities were measured. The results show that (i) a full UP element, located immediately downstream of the DNA site for CRP, relaxes the spacing requirements for activation and increases the recruitment of RNAP and open complex formation; (ii) the distal UP subsite plays the key role in this relaxation; (iii) modification of the extended −10 element also affects the spacing requirements for CRP-dependent activation. From these results, we conclude that the spacing requirements for CRP-dependent transcription activation vary according to the sequence of different promoter elements, and our results are important for understanding the organization of promoters in many different bacteria which are controlled by transcription factors that use activatory mechanisms similar to CRP.**

## INTRODUCTION

Promoter recognition by the multi-subunit bacterial RNA polymerase holoenzyme (RNAP) is the first step in the pathway to transcript initiation in bacteria, and, for many bacterial genes, this is the primary target where their expression is regulated. This recognition involves interactions between different promoter elements and determinants in the RNAP σ and α subunits (1–3). In *Escherichia coli*, the primary σ factor, $\sigma^{70}$, carries four conserved domains and has been said to 'orchestrate' transcript initiation (2). Determinants in $\sigma^{70}$ domains 2 and 4 respectively recognize promoter −10 and −35 elements, whilst some promoters lack a specific −35 element, but bear an extended −10 element immediately upstream of the −10 element, which is recognized by domain 3 (4). Concerning the RNAP α subunit, its primary role is in RNAP assembly, but the C-terminal domain of each α subunit (αCTD) binds to short sequence elements known as UP elements that are found upstream of the −35 element at many promoter (5). Previous reports have shown that αCTD is connected to the α N-terminal domain (αNTD), hence the rest of the RNAP, by a flexible linker, and this appears to permit some flexibility in where UP elements can be placed (1,5). Interestingly though, to be functional, UP elements must be placed such that αCTD is bound to the same face of the promoter DNA as the rest of the RNAP, and, at some promoters, there is a direct interaction between αCTD and domain 4 of $\sigma^{70}$ (6).

At many bacterial promoters, the base sequences of the different elements are such that RNAP recruitment is inefficient, and an activator protein is required. Such activators provide a simple way to couple transcription to environmental signals. Although some bacterial transcription activators function by more complex mechanisms, many activators simply make a direct contact with RNAP that recruits RNAP to the target promoter (reviewed in 7). Essentially, the direct contact between the bound activator and RNAP compensates for defects in RNAP-promoter DNA interactions. The simplest scenario is found at so-called Class I activator-dependent promoters, where the activator binds to a target located upstream of the promoter and then makes a direct interaction with αCTD that recruits αCTD, and thereby the rest of the RNAP, to the promoter. At many of these promoters, the dependence of the promoter on the activator is due to the lack of an UP element, but in some cases, the activator and an UP element function synergically in the recruitment of RNAP to the promoter (5,8).

*To whom correspondence should be addressed. Tel: +86 10 6275 8490; Fax: +86 10 6275 6325; Email: wangyp@pku.edu.cn

The most studied bacterial activator is the *E. coli* cyclic AMP (cyclic adenosine monophosphate) receptor protein (CRP, also known as CAP, catabolite activator protein), which is a global transcription factor that modulates expression from hundreds of promoters (9). At many of these, it activates transcription through a Class I mechanism, binding to a target positioned 60–90 base pairs upstream from the transcript start point. Extensive studies have shown that, irrespective of the target site location, the same surface of bound CRP (known as Activating Region 1, AR1) interacts with the same surface of αCTD (known as the 287 determinant) and it has been supposed that this is due the flexible linker joining αNTD and αCTD (reviewed in 10). To investigate this further, several groups constructed families of promoter constructs carrying the same DNA site for CRP at different locations upstream of the same promoter elements, and revealed the importance of stereospecific positioning of CRP with respect to and RNA polymerase, with peaks of activation observed when the DNA site for CRP could activate transcription was centered near positions 61, 71, 82 or 93 base pairs upstream from the transcript start (11,12). These studies concluded that Class I activation at promoters required CRP and RNAP to be bound on the same face of the DNA helix, with activation becoming less as the DNA site for CRP is located further upstream, likely due to the energetic cost of stretching the RNAP α inter-domain linker to facilitate the interaction between the αCTD 287 determinant and AR1 of CRP. These studies also showed that CRP-dependent activation fell sharply as the DNA site for CRP was moved from the optimal positions, where CRP and RNAP are on the same face of the DNA helix, indicating a limited flexibility within the linker of the α subunit. This can be attributed either to the energy required for αCTD to bind to a different face of the DNA helix, or to the energy needed to distort the promoter DNA to bring bound CRP, αCTD and the rest of RNAP back to the same face of the DNA helix.

In this study, we report that the stringency of the spacing requirement for Class I activation by CRP can be relaxed by the introduction of an UP element immediately downstream of the DNA site for CRP. The previous study of Law *et al.* (13) showed that Class I activation by CRP could be enhanced by the presence of an UP element located immediately downstream of the DNA site for CRP, and the effects of positioning the UP element at different locations were investigated. Hence, here, we exploited the constructions where the UP element resulted in optimal enhancement, in order to measure its effects on the stringency of the requirements for positioning the DNA site for CRP. Involvement of the UP element significantly increases the recruitment of RNAP, and stabilizes its binding and we suggest that this allows stabilization of the less favorable (distorted) conformation needed for a productive CRP–RNAP|promoter complex to form at some promoters. We argue that, since many CRP-dependent promoters carry sequences resembling UP elements immediately downstream of the DNA site for CRP, and the location of the CRP binding site differs from promoter to promoter, our finding will be important for predicting the activity patterns of as yet uncharted promoters that are dependent on CRP or other Class I activators, and for understanding their architecture. In a second set of experiments, we show that the introduction of a consensus extended −10 element into a Class I CRP-dependent promoter produces a similar relaxation in the stringency of the spacing requirements for activation. Since promoters with consensus extended −10 elements are relieved of the requirement for anchoring of $\sigma^{70}$ region 4 to the −35 element, we argue that DNA flexibility is likely to be the predominant factor linking activation and binding site location.

## MATERIALS AND METHODS

### Strains, plasmids and promoters

All the CRP-dependent promoters described here were carried on EcoRI–HindIII fragments and are derivatives of the *E. coli melR* promoter, carrying a consensus DNA site for CRP (12). General cloning vector pAA121 was exploited for cloning and propagation. For measurement of promoter activities, the fragments were cloned into the low copy-number *lac* expression vector, pRW50 (14). Recombinant plasmids were propagated in the *E. coli* K12 Δ*lac* strain, VH1000, or in the Δ*crp* derivative, VH1000Δ*crp*. By convention, promoter sequences are numbered with the transcript start taken as +1, with upstream sequences prefixed '−'. Many of the promoters used in this work are denoted CC(−X)α(−Y) where −X is the position of the center of the DNA site for CRP, and −Y is the position of the last of four thymines on the non-template strand in the middle of the UP element as previously adopted (8,13). Derivatives of these promoters with varying distal or proximal UP subsite are denoted CC(−X)dis or CC(−X)pro (Figure 1). Promoters with the promoter-distal UP subsite replaced by the promote-proximal subsite sequence are denoted CC(−X)pd (Figure 1).

### Construction of promoters

The starting point was the EcoRI–HindIII fragment carrying the CC(−72.5) promoter (i.e. the center of the DNA site for CRP is located between base pairs 72 and 73, upstream from the transcript start). Polymerase chain reaction (PCR) mutagenesis was used to increase or delete DNA between the CRP binding site and transcription start point to yield a series of promoters with the binding sites for CRP shifted with and increment of 1 base pair. The relevant promoters are described in Figure 1a. Each CRP-dependent promoter was modified by introducing a full UP element 4 base pairs downstream of the DNA site for CRP and the resulting set of promoters are described in Figure 1b. A subset of these promoters was modified so that they carried either just the promoter-distal (Figure 1c) or promoter-proximal (Figure 1d) UP element subsite. Finally, a set of promoters was constructed by replacing the promoter-distal UP element subsite with the promoter-proximal subsite sequence (Figure 1e).

### Construction of CRP-dependent promoters with a modified 'extended −10' element

The MUT1 and MUT2 series of promoters were derived by PCR from certain of the promoters in Figure 1a by

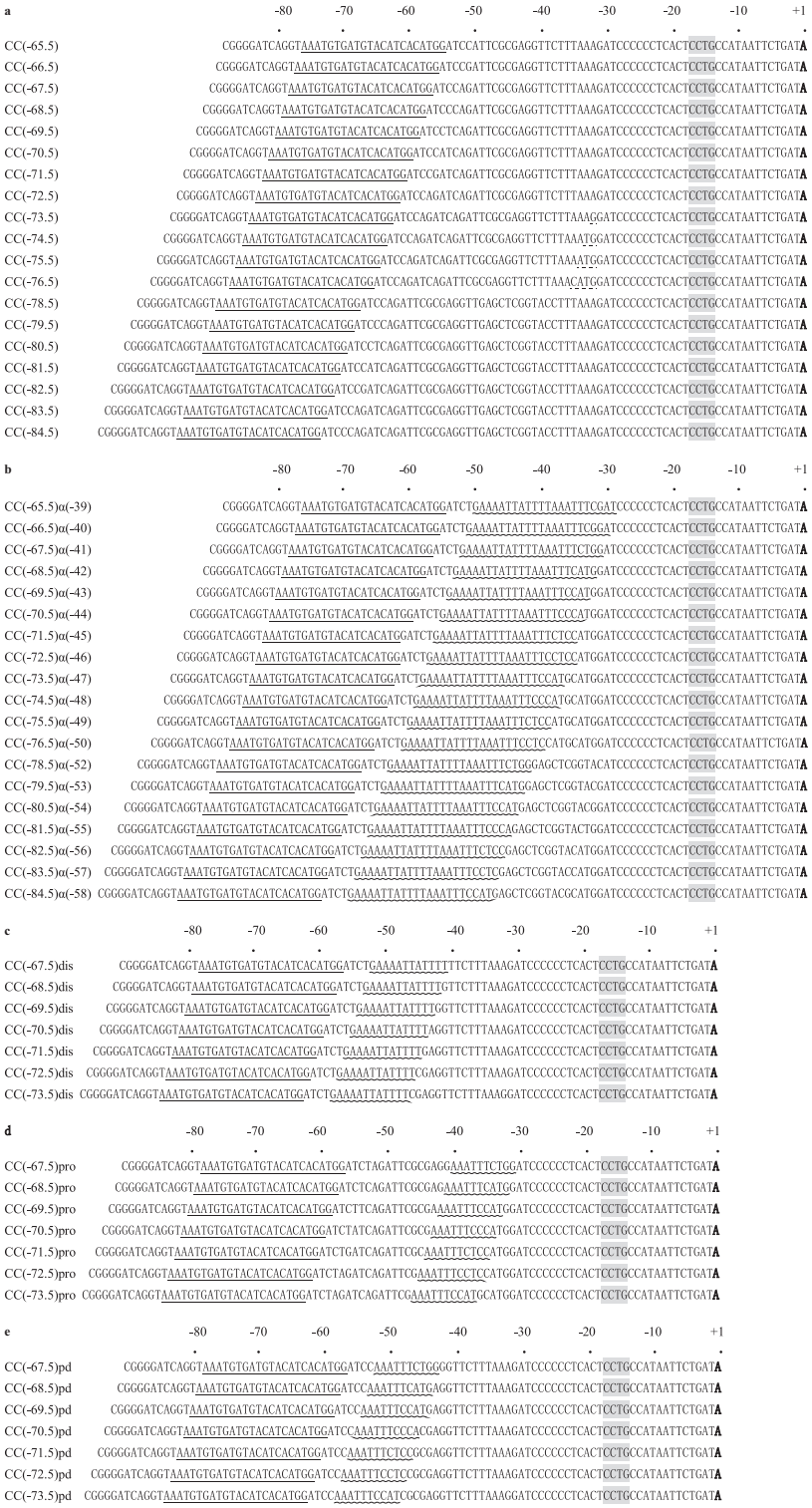


**Figure 1.** Base sequences of Class I CRP-dependent promoters. The sequences of the non-template strands of different promoters constructed in this study are shown. Sequences are numbered with the transcript start site as +1. The locations of the DNA sites for CRP are are single underlined. The UP elements/UP subsites are wavy underlined. The extended −10 motifs are shadowed. The different promoters are named according to the location of the UP element and the DNA site for CRP. (**a**) A set of promoters where the location of the DNA site for CRP was varied. (**b**) A set of promoters where both the location of the UP element and the DNA site for CRP were varied synchronously. (**c**) A set of promoters where both the location of the distal UP subsite and the DNA site for CRP were varied synchronously. (**d**) A set of promoters where both the location of the proximal UP subsite and the DNA site for CRP were varied synchronously. (**e**) A set of promoters where the sequence of the proximal UP subsite in (d) are placed to the position of distal UP subsite.

introducing a consensus extended −10 element (MUT2: $^{-17}$TGTGC$^{-13}$) or an extended −10 element that bore no relation to the consensus (MUT1: $^{-17}$CCACC$^{-13}$).

### Assays of promoter activity *in vivo*

To assay the activity of the different promoters, EcoRI–HindIII fragments were cloned into the broad host-range, low copy-number *lac* expression vector, pRW50, placing the *lac* genes under the control of the promoter. The resulting plasmids were transformed into the Δ*lac* strains, VH1000 or VH1000Δ*crp*, and transformants were grown in Luria–Bertani (LB) broth supplemented with appropriate antibiotics. β-galactosidase expression was measured exactly as described previously using the Miller method (15). The CRP-dependence of any promoter was judged by comparing its activity in the VH1000 and VH1000Δ*crp* backgrounds.

### Calculation of activation factors

To describe the relative dependence of any promoter on CRP, we defined an 'activation factor' as follows: (i) The β-galactosidase activities from the same series of promoters are measured in both strains VH1000 and VH1000Δ*crp*; 2) For each promoter, the β-galactosidase activity measured in strain VH1000 is then divided by the whole average β-galactosidase activities from the same series of promoters measured in strain VH1000Δ*crp* to get its activation factor. For example, for the promoters with both CRP binding site and UP element, CC(−65.5)α(39) to CC(−84.5)α(58), each plasmid is transformed into both strains VH1000 and VH1000Δ*crp*, and single colony from VH1000/VH1000Δ*crp* is incubated in LB broth until the $OD_{600}$ is around 0.4. Then the β-galactosidase activity for each promoter from both VH1000 and VH1000Δ*crp* background is measured according to Miller method (15). Afterwards, the whole average value for promoters from CC(−65.5)α(−39) to CC(−84.5)α(−58), measured in VH1000Δ*crp* background, is calculated.

$$\text{Activation factor}_{\text{CC(-xx.5)}\alpha\text{(-yy)}} = \frac{\beta - \text{gal activity}_{\text{CC(-xx.5)}\alpha\text{(-yy)}}}{\text{Average } \beta - \text{gal activity}_{\text{fromCC(-65.5)}\alpha\text{(-39) to CC(-84.5)}\alpha\text{(-58)}}}$$

## ELECTROPHORETIC MOBILITY SHIFT ASSAY

Reaction mixtures containing a 5 nM DNA fragment, 30 nM CRP, and 30 nM RNAP were incubated in binding buffer (20 mM Tris-HCl, pH 7.9, 100 mM NaCl, 3 mM $MgCl_2$, 0.1 mM EDTA, 0.1 mM DTT, 5% glycerol, 50 μM cAMP, 50 μg/mL BSA) and assayed as follows. Firstly, 5 nM 5' digoxin labelled DNA fragment was incubated with 30 nM CRP for 30 min at 22 °C in 16 μL reaction mixtures; then 30 nM RNAP was added and incubated for 30 min at 22 °C; finally, 3 μL of 50 μg/mL of heparin was added for 3 min to perform the heparin challenge. After this, the complexes were analyzed by native gel electrophoresis at room temperature. The gel contained 6% polyacrylamide, 7.5% glycerol, and 45 mM Tris-borate (pH 8.3), 1 mM EDTA (TBE). After electrophoresis, the steps remaining were carried out according to Roche DIG Shift Kit, 2nd Generation.

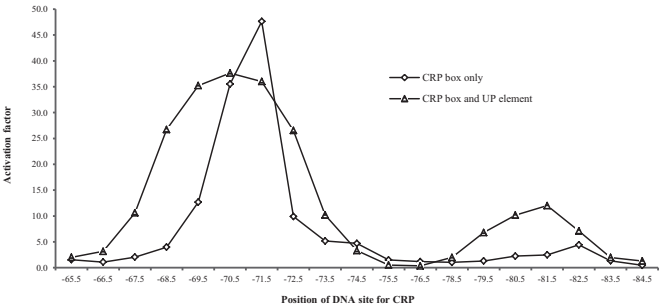


**Figure 2.** The activation of promoter activity by CRP-cAMP with the UP element. β-Galactosidase activities were measured in VH1000$crp^+$/VH1000Δ*crp* cells carrying different Class I CRP-dependent promoters cloned into the *lac* expression vector pRW50. Promoter activation by CRP-cAMP is calculated as the $crp^+$ value at each distance divided by the average Δ*crp* value. Cells were grown exponentially in Luria–Bertani broth containing 35μg/ml tetracycline, and β-galactosidase expression was measured by the Miller protocol; activities are expressed in Miller units of β-galactosidase activity. Each value is the mean ± S.D. of at least three independent assays.

## RESULTS AND DISCUSSION

### Spacing requirements for CRP-mediated activation, and its dependence on an UP element

The starting point of this work was the set of promoters illustrated in Figure 1a in which a consensus DNA site for CRP was positioned at different locations upstream of *E. coli melR* promoter, whose activity is completely dependent on CRP (12). The locations were varied from position −65.5 to position −84.5 by one base pair intervals. Each promoter was cloned on an EcoRI–HindIII fragment and, to measure its activity, the fragment was cloned into the pRW50 *lac* expression vector, and β-galactosidase activities were measured in *E. coli* VH1000 Δ*lac* and VH1000 Δ*lac* Δ*crp* host cells. The results, illustrated in Figure 2 and Supplementary Figure S1, confirm that promoter activity is CRP-dependent and that, as expected (12), 'peaks' of CRP-dependent activity are seen when the DNA site for CRP is located at position −71.5 or −82.5.

To evaluate the effect of an UP element on CRP-dependent activation, the 22 base pairs UP element from the *E. coli rrnB* P1 promoter (16) was introduced immediately downstream of the CRP binding site at each promoter and the assays were repeated. The design of these promoters, illustrated in Figure 1b, was informed by previous studies that had investigated the optimal juxtaposition between DNA sites for CRP and UP elements (8,13). By convention, the different promoters are named after the location of the DNA site for CRP, and the UP element, with respect to the transcript start site. We first used the Δ*crp* host cells to check whether the UP elements altered the very low promoter activity in the absence of CRP. The results, illustrated in Figure S1, show that the UP elements in the newly constructed promoters cause only a small stimulation in CRP-independent promoter activity, relative to the starting promoters. Next, the promoters were assayed in the

*crp*$^+$ background, and the results (Figure 2, Supplementary Figure S1) show that the UP element, in combination with CRP, causes a large increase in the activity of many of the promoters. For example, the CC(−70.5)α(−44) promoter is more active than the CC(−70.5) promoter, and large effect is seen when the CRP binding site is moved one helical turn upstream. On the other hand, for promoters where CRP is bound on the other face of the DNA (for instance, at position −75.5), the presence of the full UP element did not increase promoter activity.

The results in Figure 2 show that the presence of an UP element causes a more relaxed phasing-dependent pattern of CRP-dependent expression. For example, the promoters in the series from CC(−68.5)α(−42) to CC(−72.5)α(−46) exhibit similar activities in *crp*$^+$ cells. This contrasts with the 'peak' of CRP-dependent expression observed with the CC(−71.5) promoter lacking an UP element. To eliminate the possibility that the observed effects were due to the creation of alternative promoters, we followed the approach adopted previously to change the −10 hexamer from CATAAT to CGTAAT (8). Data presented in the Supplementary Material (Supplementary Figure S2) show that this change completely suppresses the activity of all the promoters tested, indicating that no alternative promoter had been created or unmasked during the introduction of the UP elements. A further control experiment showed that CRP-dependent activation of the promoters in the series from CC(−65.5)α(−39) to CC(−84.5)α(−58) was dependent on the interaction between AR1 of CRP and the 287 determinant of αCTD, as the data in Supplementary Figure S3 show that the H159L substitution in CRP that inactivates AR1 (9) completely suppresses CRP-dependent activation at the different promoters.

To explain why the presence of an UP element causes a more relaxed phasing-dependent pattern of CRP-dependent expression, we suggest that the synergistic binding of CRP and αCTD to their respective sites facilitates the flexibility needed to facilitate contact between AR1 of CRP and the 287 determinant of αCTD at the different promoters. The EMSA (Electrophoretic Mobility Shift Assay) experiments (Supplementary Figure S4) were performed to analyze the ternary RNAP–CRP–promoter complexes at the CC(−68.5) and CC(−68.5)α(−42), and demonstrated that the UP element significantly increased the recruitment of RNAP and open complex formation.

## UP element effects on CRP-mediated activation depend on the distal UP subsite

Previous studies have shown that UP elements consist of two 11 base pair subsites, each of which is recognized by the C-terminal domain of one of the two RNAP α subunits (17,18). To investigate whether the effects we had observed with the full 22 base pair *rrnB* P1 UP element were due one subsite or the other, we modified the CC(−67.5) to CC(−73.5) series of promoters to carry either the promoter-distal subsite (Figure 1c: dis series) or the promoter-proximal subsite (Figure 1d: pro series). These new promoters were cloned into *lac* expression vector, pRW50 and *β*-galactosidase activities were measured in the *crp*$^+$ strain, VH1000. The results (Figure 3a and Supplementary Figure S5) showed that the distal UP subsite caused an overall stimulation of CRP-dependent activation, and relaxed the spacing requirement for CRP-dependent activation, similar to the full UP element. In contrast, whilst the promoter-proximal UP subsite caused increases in promoter activity (Figure 3b), the profile of the dependence of activation by CRP on location of the DNA site for CRP was unaltered (Supplementary Figure S5).

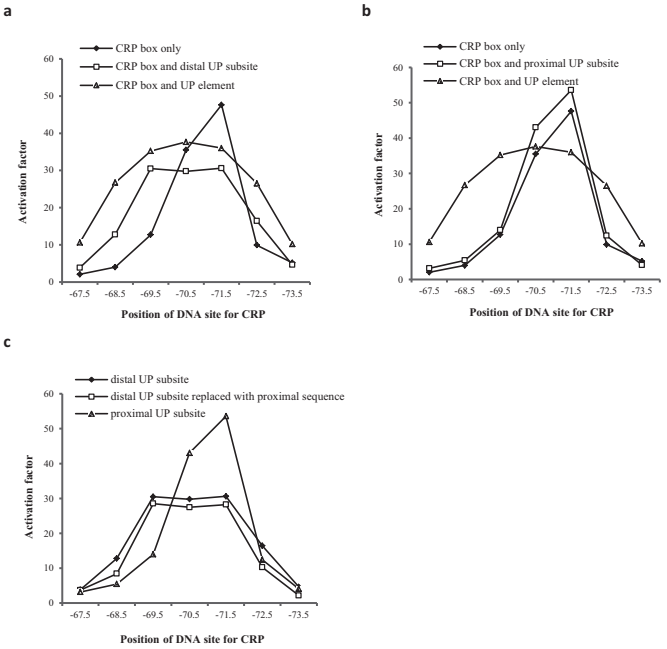


**Figure 3.** The activation of promoter activity by CRP-cAMP with a UP subsite. *β*-Galactosidase activities were measured in VH1000*crp*$^+$/VH1000*Δcrp* cells carrying different Class I CRP-dependent promoters cloned into the *lac* expression vector pRW50. The binding center for CRP was shown on the x-axis. (**a**) Promoters carried both a CRP binding site and a distal UP subsite. (**b**) Promoters carried both a CRP binding site and a proximal UP subsite. (**c**) Promoters carried a CRP binding site and a proximal UP subsite sequence on the position of the distal UP subsite. The activation factor in a, b and c was calculated as taking the promoter value in *crp*$^+$ background divided by average *Δcrp* value within the same set of promoters.

Taken together, our results indicate that it is the distal UP subsite that is responsible for relaxation of the spacing requirement for CRP-dependent activation. To check that these results were not simply due to a peculiarity of the base sequence of the upstream part of the *rrnB* P1 UP element, we constructed the CC(−67.5) to CC(−73.5) pd series of promoters, in which the promoter-distal UP element subsite sequence was replaced with the promoter-proximal subsite sequence (Figure 1e: pd series). Data illustrated in Figure 3c and Supplementary Figure S5 show that the spacing requirements for CRP-dependent activation are relaxed irrespective of the precise promoter-distal UP element subsite sequence. Hence, we propose that, for an UP subsite to exert its effects on CRP-dependent activation, it is its location, rather than its sequence, that matters.

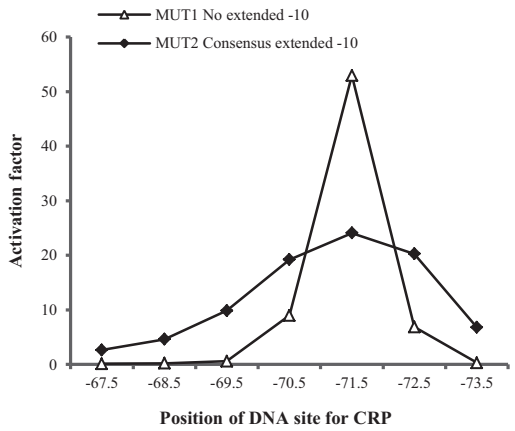


**Figure 4.** The activation of promoter activity by CRP-cAMP with a modified extended −10 element. Mut 1 and Mut 2 series of promoters are derivatives of the CC(−n) promoters listed in Figure 1a. The Mut 2 series, carries an extended −10 element ($^{-17}$TGTGc$^{-13}$) that resembles the consensus ($^{-17}$TRTGn$^{-13}$), whilst the Mut 1 carries a sequence ($^{-17}$CCACc$^{-13}$) that has no extended −10 element function (19). Activation factors were calculated by taking the measured promoter activity in the *crp*$^+$ background divided by activity in the *Δcrp* background.

### Effects of an extended −10 motif

To investigate any effects of the extended −10 motif on the profile of CRP-dependent activation, we modified the CC(−67.5) to CC(−73.5) series of promoters to carry either a consensus extended −10 element ($^{-17}$TGTGC$^{-13}$: MUT2 series) or an extended −10 element that bore no relation to the consensus ($^{-17}$CCACC$^{-13}$: MUT1 series) (19). The activities of the two series of promoters were measured in the VH1000 *crp*$^+$ strain and the results are illustrated in Figure 4. The results show that the presence of a consensus extended −10 element increases promoter activity and causes a relaxation in the spacing requirement for CRP-dependent activation, similar to that observed with the UP element in the 'dis' series of promoters (Figure 3, Supplementary Figure S6). Weakening of the extended −10 element of the MUT1 series leads to a more stringent spacing requirement for CRP-dependent activation (compare data in Figures 2 and 4). We also introduced a near-consensus −35 hexamer element ($^{-35}$TTAACA$^{-30}$) to the CC(−67.5) to CC(−73.5) series of promoters, as a control. Comparison of data in Figure 4, Supplementary Figure S6 and S7 shows that improvement of the promoter −35 element has little or no effect on the stringent spacing requirements, unlike the presence of a consensus extended −10 element (note that we were constrained here by the need to use −35 element that did not release the promoters' dependency on CRP).

## CONCLUSIONS

Studies of CRP and its activities have shown that it binds to hundreds of target sites in *E. coli* and is a global regulator of transcription (20–22). At many targets, bound CRP activates transcript initiation at a neighboring promoter by making a direct interaction with RNAP αCTD that recruits αCTD, and thereby the rest of the RNAP, to the promoter (10,23). At these promoters, known as Class I CRP-dependent promoters, CRP increases the initial binding of RNAP, and optimal CRP-dependent activation occurs when CRP is located on the same face of the DNA helix as the rest of RNAP (23–25). Hence, activation occurs when the DNA site for CRP is centered near positions −61.5, −71.5, −82.5 or −92.5 upstream from the transcript start, but, all other factors being equal, activation becomes progressively weaker as bound CRP moves further upstream, likely because of the energy cost of separating αCTD from the rest of the RNAP (10).

Previous reports have stressed the stringency of the spacing requirements for activation by CRP and activation falls off rapidly as the DNA site for CRP is moved from these activatory locations (11,12). This must be due to the prohibitive energy cost of distorting the RNAP α inter-domain linker to permit αCTD to track round the DNA to interact with CRP, or, alternatively, the energy cost of twisting the DNA to bring CRP, αCTD and RNAP back into register on the same face of the DNA.

In this study, first, we report that an UP element, juxtaposed downstream of a CRP binding site, relaxes the spacing requirement for CRP action (Figure 2), and the effect is due to the αCTD that is bound immediately adjacent to CRP. Since, the DNA-bound CRP must make a direct interaction with DNA-bound αCTD, we suggest that the energy of formation of the CRP–αCTD–DNA complex compensates for the energy needed either to distort the RNAP α inter-domain linker or to twist the promoter DNA.

Our second major result is that, in the absence of an UP element, a consensus extended −10 element also reduces the stringency of the spacing requirement for CRP-mediated activation (Figure 4a). To explain this, we suggest that it is DNA flexibility that is the paramount factor in accommodating Class I CRP-dependent activation at promoters with CRP bound at different locations. Recall that the presence of a consensus-like extended −10 element at a bacterial promoter aids the interaction of σ with the promoter DNA, and can make the −35 element redundant and negate the need for $\sigma^{70}$ region 4 to contact the −35 element (4,19). This will extend the length of the DNA sequence downstream of bound CRP that can be distorted to facilitate the interaction between CRP and αCTD, thereby reducing the energetic cost of the distortion. Consistent with this, the spacing requirement for CRP-mediated activation is unchanged by improvement of the −35 element.

The main conclusion from this study is that the stringency of the spacing requirement for CRP-dependent activation is dependent on the nature of the promoter elements. Hence, location is not the sole factor determining whether a bound CRP molecule is competent for Class I-type transcription activation. This may provide a rationale for the existence of UP-like sequences, located immediately downstream of CRP sites at many different CRP-regulated promoters, and the apparent 'non-optimal' architecture of some of these promoters (Figure 5; see also Ref. 26). It may also explain the full variety of binding locations seen for many different transcription activators, as well as providing a framework for understanding promoter architectures in newly sequenced bacterial genomes.

| promoter | Location of DNA site for CRP | σ factor | Partial promoter sequence (5'→3') |
|---|---|---|---|
| *acs* P2 | -69.5 | $\sigma^{38}$, $\sigma^{70}$ | CTTTTGCGTGATCTGTCGCCCAAATACTAAACAAAACTGCCAATAC |
| *ascF* P | -70.5 | $\sigma^{70}$ | TATTCAGGTGACCGGTTTCACAAATATAAAAAATGAACAATTCACT |
| *csiD* P | -68.5 | $\sigma^{38}$ | TAATTTGTTGCTTTTGATCACAATAAGAAAACAATATGTCGCTTTT |
| *cstA* P1 | -89.5 | | TAACGGAGTGATCGAGTTAACATTGTTAAGTTAAATATTGGTTTCA |
| *exuT* P2 | -161.5 | | TATTTTCGTGAGTTAGATCAATAAACGTAGTTAAAAAAATTACTCT |
| *fixA* P | -126.5 | $\sigma^{70}$ | AATATTGGTGATCCATAAAACAATATTGAAAATTTCTTTTTGCTAC |
| *fur* P1 | -77.5 | $\sigma^{70}$ | TGTAAATGTAAGCTGTGCCACGTTTTTATTAACAATATTTGCCAGG |
| *glnA* P1 | -71.5 | $\sigma^{70}$ | GTCCCTTTGTGATCGCTTTCACGGAGCATAAAAAGGGTTATCCAAA |
| *glpABC* P | -90.5 | $\sigma^{70}$ | ATTAAATGTGAATTGCCGCACACATTATTAAATAAGATTTACAAAA |
| *hlyE* P | -61.5 | | ACATTGTTTGATATTTATCATATTAATAGAAATAAAGACATTGACG |
| *upB* P3 | -83.5 | $\sigma^{70}$ | CAAAATAGTGACCTCGCGCAAAATGCACTAATAAAAACAGGGCTGG |
| *kbaZ* P | -60.5 | $\sigma^{70}$ | TTCTTTTGTGAATCAGATCAGAAAACCATTATCTTTCGTTTTATTT |
| *malT* P | -70.5 | $\sigma^{70}$ | TGGAATTGTGACACAGTGCAAATTCAGACACATAAAAAAAACGTCAT |
| *malX* P | -101.5 | $\sigma^{70}$ | TGATTATGTGACAGATAAAACGTTTTACCTTTTATTTATCTTATAC |
| *mhpA* P | -95.5 | $\sigma^{70}$ | TTGTTCTGCATATTAATTGACATTTCTATAGTTAAAACAACGTGGT |
| *ompF* P | -92.5 | $\sigma^{70}$ | TTAAATTTTACTTTTGGTTACATATTTTTTCTTTTTGAAACCAAAT |
| *rhaS* P | -91.5 | $\sigma^{38}$, $\sigma^{70}$ | GCATTTCCTGAAAATTCACGCTGTATCTTGAAAAATCGACGTTTTT |
| *sdhC* P | -83.5 | $\sigma^{70}$ | TGTTATCGTGACCTGGATCACTGTTCAGGATAAAACCCGACAAACT |
| *ulaA* P | -60.5 | $\sigma^{70}$ | ATTATTTGCGGGTCGCGTCACATTTAATCATAAATAATCTTGTTGT |

**Figure 5.** Alignment of CRP-dependent promoters from *Escherichia coli* K-12. (taken from RegulonDB) The CRP binding sites are single underlined, whilst downstream sequences resembling the UP subsites are wavy underlined. The consensus sequence for distal UP subite is 5′-NNAWWWWWTTTTN-3′; The consensus sequence for proximal UP subsite is 5′-AAAAAARNR-3′.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGMENT

We thank Mark S. Thomas from the University of Sheffield Medical School for providing the *E. coli* VH1000 strain.

## FUNDING

NSFC of P. R. China [31270127 and 30830005 to Y.P.W.]; 973 National Key Basic Research Program [2010CB126503 to Y.P.W.]; Program of Introducing Talents of Discipline to Universities [B06001]; State Key Laboratory of Protein and Plant Gene Research [B02]. National Science Fund for Distinguished Young Scholars [39925017 to Y.P.W.]. Funding for open access charge: NSFC of P. R. China (31270127).
*Conflict of interest statement.* None declared.

## REFERENCES

1. Busby,S. and Ebright,R. (1994) Promoter structure, promoter recognition, and transcription activation in prokaryotes. *Cell*, **79**, 743–746.
2. Gross,C. A., Chan,C. and Young,B. (1998) The functional and regulatory roles of sigma factors in transcription. *Cold Spring Harb. Symp. Quant. Biol.*, **63**, 141–155.
3. Browning,D. F. and Busby,S. J. (2004) The regulation of bacterial transcription initiation. *Nat. Rev. Microbiol.*, **2**, 57–65.
4. Murakami,K. S. and Darst,S. A. (2003) Bacterial RNA polymerases: the wholo story. *Curr. Opin. Struct. Biol.*, **13**, 31–39.
5. Gourse,R.L., Ross,W. and Gaal,T. (2000) UPs and downs in bacterial transcription initiation: the role of the alpha subunit of RNA polymerase in promoter recognition. *Mol. Microbiol.*, **37**, 687–695.
6. Ross,W., Schneider,D. A, Paul,B. J., Mertens,A. and Gourse,R. L. (2003) An intersubunit contact stimulating transcription initiation by *E. coli* RNA polymerase: interaction of the α C-terminal domain and σ region 4. *Genes Dev.*, **17**, 1293–1307.
7. Lee,D. J., Minchin,S. D. and Busby,S. J. (2012) Activating transcription in bacteria. *Annu. Rev. Microbiol.* **66**: 125–152.
8. Lloyd,G. S., Busby,S. J. W. and Savery,N. J. (1998) Spacing requirements for interactions between the C-terminal domain of the α subunit of *Escherichia coli* RNA polymerase and the cAMP receptor protein. *Biochem. J.*, **330**, 413–420.
9. Kolb,A., Busby,S., Buc,H., Garges,S. and Adhya,S. (1993) Transcriptional regulation by cAMP and its receptor protein. *Annu. Rev. Biochem.*, **62**, 749–795.
10. Busby,S. J. and Ebright,R. H. (1999) Transcription activation by catabolite activator protein (CAP). *J. Mol. Biol.*, **293**, 199–213.
11. Ushida,C. and Aiba,H. (1990) Helical phase dependent action of CRP-effect of the distance between the CRP site and the -35 region on promoter activity. *Nucleic Acids Res.*, **18**, 6325–6330.
12. Gaston,K., Bell,A., Kolb,A., Buc,H. and Busby,S. (1990) Stringent spacing requirements for transcription activation by CRP. *Cell*, **62**, 733–743.
13. Law,E. C., Savery,N. J. and Busby,S. J. (1999) Interactions between the *Escherichia coli* cAMP receptor protein and the C-terminal domain of the α subunit of RNA polymerase at Class I promoters. *Biochem. J.*, **337**, 415–423.
14. Lodge,J., Fear,J., Busby,S., Gunasekaran,P. and Kamini,N. (1992) Broad host range plasmids carrying the *Escherichia coli* lactose and galactose operons. *FEMS Microbiol. Lett.*, **95**, 271–276.
15. Miller,J. H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Press, New York.
16. Estrem,S. T., Gaal,T., Ross,W. and Gourse,R. L. (1998) Identification of an UP element consensus sequence for bacterial promoters. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 9761–9766.

17. Ross,W., Aiyar,S. E., Salomon,J. and Gourse,R. L. (1998) *Escherichia coli* promoters with UP elements of different strengths: modular structure of bacterial promoters. *J. Bacteriol.*, **180**, 5375–5383.
18. Estrem,S. T., Ross,W., Gaal,T., Niu,W., Ebright,R. H. and Gourse,R. L. (1999) Bacterial promoter architecture: subsite structure of UP elements and interactions with the carboxyl-terminal domain of the RNAP α subunit. *Genes Dev.*, **13**, 2134–2147
19. Mitchell,J. E., Zheng,D., Busby,S. J. and Minchin,S. D. (2003) Identification and analysis of 'extended -10' promoters in *Escherichia coli*. *Nucleic Acids Res.*, **16**, 4689–4695.
20. Zheng,D., Constantinidou,C., Hobman,J. L. and Minchin,S. D. (2004) Identification of the CRP regulon using *in vitro* and *in vivo* transcriptional profiling. *Nucleic Acids Res.*, **32**, 5874–5893.
21. Grainger,D. C., Hurd,D., Harrison,M., Holdstock,J. and Busby,S. J. (2005) Studies of the distribution of *Escherichia coli* cAMP-receptor protein and RNA polymerase along the *E. coli* chromosome. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 17693–17698.
22. Shimada,T., Fujita,N., Yamamoto,K. and Ishihama,A. (2011) Novel roles of cAMP receptor protein (CRP) in regulation of transport and metabolism of carbon sources. *PLoS One*, **6**, e20081.
23. Lawson,C. L., Swigon,D., Murakami,K. S., Darst,S. A., Berman,H. M. and Ebright,R. H. (2004) Catabolite activator protein: DNA binding and transcription activation. *Curr. Opin. Struct. Biol.*, **14**, 10–20.
24. Chen,H., Tang,H. and Ebright,R. H. (2003) Functional Interaction between RNA Polymerase α subunit C-Terminal Domain and $\sigma^{70}$ in UP-Element- and Activator-Dependent Transcription. *Mol. Cell*, **11**, 1621–1633.
25. Zhang,Y., Feng,Y., Chatterjee,S., Tuske,S., Arnold,E. and Ebright,R. H. (2012) Structural basis of transcription initiation. *Science*, **338**, 1076–1080.
26. Germer,J., Becker,G., Metzner,M. and Hengge-Aronis,R. (2001) Role of activator site position and a distal UP-element half-site for sigma factor selectivity at a CRP/H-NS-activated $\sigma^{S}$-dependent promoter in *Escherichia coli*. *Mol. Microbiol.*, **41**, 705–716.