doi:10.1016/j.jmb.2006.09.021 J. Mol. Biol. (2006) 364, 434–448

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Transcriptional Repressor CcpN from *Bacillus subtilis* Compensates Asymmetric Contact Distribution by Cooperative Binding

Andreas Licht* and Sabine Brantl

AG Bakteriengenetik
Friedrich-Schiller-Universität
Jena, Philosophenweg 12
D-07743 Jena, Germany

Carbon catabolite repression in *Bacillus subtilis* is carried out mainly by the major regulator CcpA. In contrast, sugar-dependent repression of three genes, *sr1* encoding a small untranslated RNA, and two genes, *gapB* and *pckA*, coding for gluconeogenic enzymes is mediated by the recently identified transcriptional repressor CcpN. Since previous DNase I footprinting yielded only basic information on the operator sequences of CcpN, chemical interference footprinting studies were performed for a precise contact mapping. Methylation interference, potassium permanganate and hydroxylamine footprinting were used to identify all contacted residues in both strands in the three operator sequences. Furthermore, ethylation interference experiments were performed to identify phosphate residues essential for CcpN binding. Here, we show that each operator has two binding sites for CcpN, one of which was always contacted more strongly than the other. The three sites that exhibited close contacts were very similar in sequence, with only a few slight variations, whereas the other three corresponding sites showed several deviations. Gel retardation assays with purified CcpN demonstrated that the differences in contact number and strength correlated well with significantly different K_D values for the corresponding single binding sites. However, quantitative DNase I footprinting of whole operator sequences revealed cooperative binding of CcpN that, apparently, compensated the asymmetric contact distribution. Based on these data, possible consequences for the repression mechanism of CcpN are discussed.

Keywords: CcpN; transcriptional repressor; chemical footprinting; carbon catabolite repression; DNA-protein-interaction

*Corresponding author

Introduction

Although many bacteria, including *Bacillus subtilis*, are able to utilise a vast number of other nutrients,[1,2] glucose is their preferred carbon source.[3] Therefore, cells need to shut-down other catabolic pathways in the presence of glucose to maximise the energy yield.[4] This is accomplished by so-called catabolite repression. In *Escherichia coli*, catabolite repression is mediated by the central signalling molecule cAMP and its receptor protein CRP.[5,6] By contrast, *B. subtilis* does not encode a CRP homologue nor does it produce detectable amounts of cAMP under aerobic conditions.[7] Instead, catabolite repression in *B. subtilis* is carried out mainly by the concerted action of CcpA and HPr-Ser46-P, which can interact to form a transcriptional repressor or activator, regulating genes involved in carbon catabolism.[8] However, it has been shown recently that at least two genes, *gapB* and *pckA*, are down-regulated in the presence of glucose, independent of CcpA.[9,10] Instead, they are regulated by a novel transcriptional repressor found by transposon mutagenesis screening for derepression of *gapB* and, therefore, named CcpN (for control catabolite protein of gluconeogenic genes).[11] The *gapB* gene encodes the rare isotype B of glyceraldehyde-3-phosphate dehydrogenase, and its gene product catalyses the conversion of 1,3-bisphosphoglycerate to glyceraldehyde 3-phosphate, but only during gluconeogenesis.[9] The *pckA* gene codes for another enzyme required for the synthesis of glucose from

E-mail address of the corresponding author: andreas.licht@uni-jena.de

Krebs cycle intermediates, PEP carboxykinase, which catalyses the conversion of oxaloacetate to phosphoenolpyruvate.[12]

The *ccpN* gene is cotranscribed with the *yqfL* gene, resulting in a bicistronic mRNA. It was shown that this operon is not autoregulated, but constitutively expressed under both glycolytic and gluconeogenic conditions.[11] Homologues of CcpN have been found in the genomes of other Bacilli, e.g. *B. halodurans*, *B. cereus*, *B. anthracis* and *Geobacillus stearothermophilus*, and in different Firmicutes.[11]

Recently, a third gene regulated by CcpN, *sr1*, has been discovered. This gene codes for a small untranslated RNA, SR1, which has been identified by a systematic search for small RNAs within intergenic regions of the *B. subtilis* genome.[13] *sr1* was expressed during gluconeogenesis, but repressed under glycolytic conditions. The *trans*-acting factor responsible for sugar-mediated repression was identified as CcpN.[13] Previous DNase I footprinting experiments for all three known CcpN operators indicated different locations of the binding regions relative to the transcription start site.

The aim of the present work was to investigate the interaction between CcpN and its operator regions in more detail using chemical interference footprinting. These experiments showed that contact strength varied greatly, depending on the sequence of a given site. Gel retardation assays with single binding sites confirmed these observations. However, quantitative DNase I footprinting experiments with DNA fragments of all three genes spanning the corresponding complete operator sequences indicated cooperative binding of CcpN. The possible impact of these results on the repression mechanism is discussed.

Results

Chemical interference footprinting experiments were performed with CcpN-His_5 (containing five additional C-terminal histidine residues) purified from an *E. coli* over-expression strain. Electrophoretic mobility shift assays (EMSAs) have verified that His-tagged CcpN shows the same binding properties as wild-type CcpN and Northern blots showed that it can exert the function of wild-type CcpN in a *ccpN* knockout strain (data not shown). All nucleotide numbers in the following paragraphs refer to the transcription start sites. The coding strand is always termed the top strand, and the non-coding strand is always termed the bottom strand.

Methylation interference

Methylation interference experiments were performed to determine guanine and adenine bases contacted by CcpN. DNA fragments were modified at purine residues by dimethyl sulphate before CcpN binding. Adenine is methylated at position N3 in the minor groove and guanine is methylated at position N7 in the major groove. Figure 1 shows the positions of the methyl groups interfering with CcpN binding. The top strand of the *sr1* operator exhibited interference at positions $G_{(-53)}$ and $G_{(-51)}$ in site I and, to a weaker extent, at $G_{(-21)}$ and $G_{(-19)}$ in site II. Adenine residues with a major contribution to CcpN binding were found only in site I at the top strand ($A_{(-48)}$ and $A_{(-46)}$), whereas in site II only less close contacts were detected. At the bottom strand, methylation of $G_{(-45)}$ in site I and $G_{(-17)}$ in site II interfered with CcpN binding. Only less close contacts to adenine residues have been found in the bottom strand: $A_{(-52)}$ and $A_{(-47)}$ in site I and $A_{(-20)}$ in site II were contacted by CcpN.

Since the contacts to guanine residues were in all cases closer than those to adenine, these results indicate that CcpN contacts the DNA mainly via the major groove with some auxiliary contacts in the minor groove. Furthermore, the contacts in site II were generally less close than those in site I.

In the *pckA* operator, only three contacts to guanine residues have been observed: $G_{(-38)}$ in site I and $G_{(-15)}$ in site II at the top strand as well as $G_{(-9)}$ in site II at the bottom strand. Binding site II was found to be contacted much more strongly than site I. The same was true for contacts to adenine. Whereas there were some significant contacts in binding site II ($A_{(-17)}$ and $A_{(-12)}$ at the top strand and $A_{(-11)}$ at the bottom strand), only one contacted adenine was detected in site I ($A_{(-36)}$ at the top strand). No significant contact was found in binding site I on the bottom strand.

The *gapB* operator showed a similarly asymmetric contact distribution, but here, contacts were concentrated in binding site I: Close contacts to guanine ($G_{(-17)}$ and $G_{(-15)}$ at the top strand and $G_{(-9)}$ on the bottom strand) were found, whereas in site II only one less closely contacted guanine ($G_{(+22)}$ on the bottom strand) was detected. The same contact distribution was found for adenine residues. Close and medium contacts were observed only in binding site I ($A_{(-19)}$ at the top strand and $A_{(-11)}$, $A_{(-8)}$ at the bottom strand). In binding site II, no close contact to adenine was observed.

Interestingly, contacts in the *sr1* operator were concentrated upstream of the −35 region and, to a lesser extent, in the spacer between −35 and −10, whereas almost no contact was observed directly within the −35 and −10 regions of p_{SR1}. In contrast, in both the *pckA* and *gapB* operator, close contacts were found only within the −10 region, whereas weak binding sites covered the −35 region and the region downstream from the transcription start site in the *pckA* and *gapB* operator, respectively. Moreover, the interference footprinting revealed that each of the three operators had two CcpN binding sites, although they appeared, due to the short spacer region, as one extended site in the previous DNase I footprints of the *pckA* operator.[11]

Potassium permanganate footprinting

Potassium permanganate footprinting was performed to determine contacts of CcpN to thymine

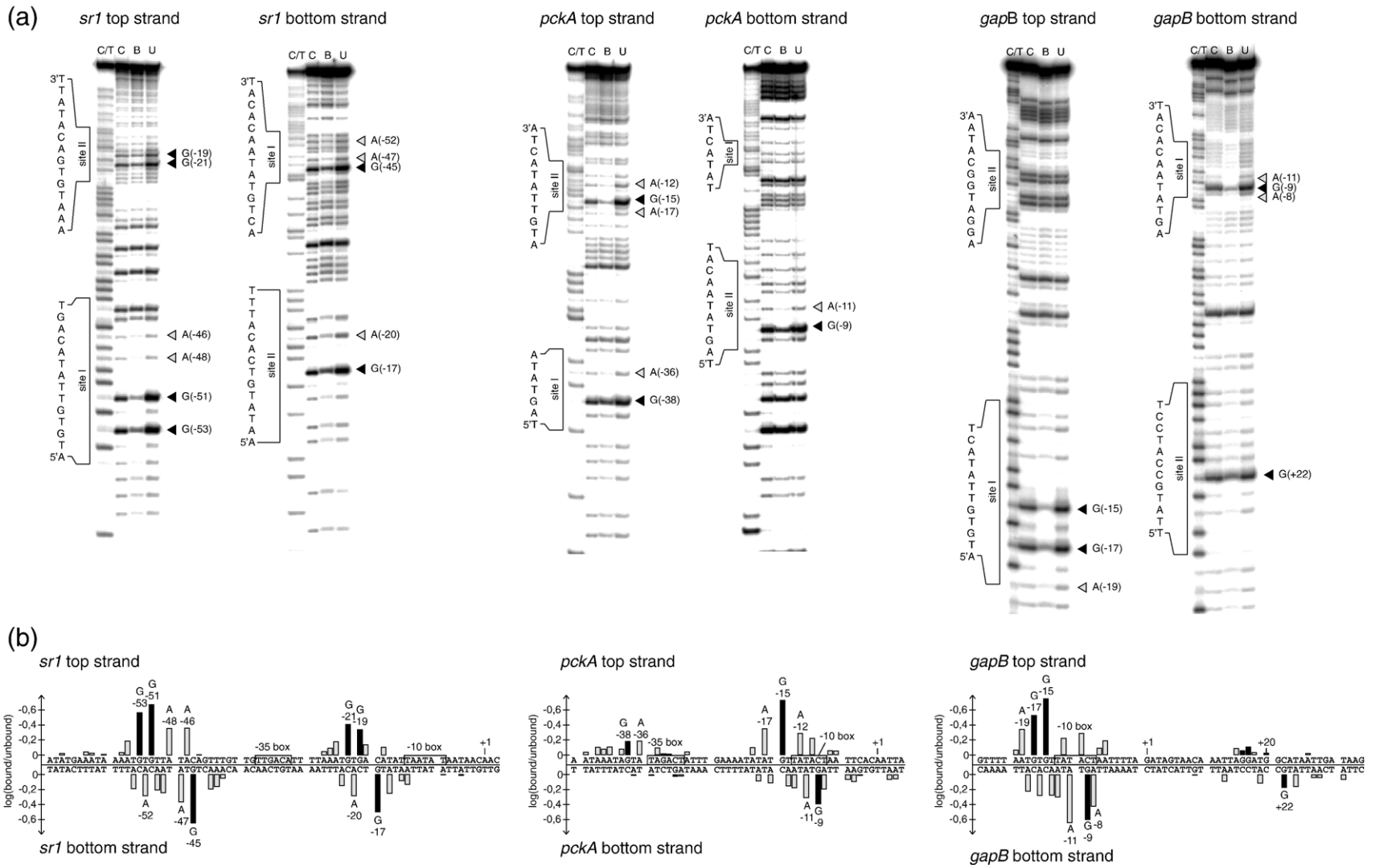

Figure 1. Methylation interference of the *sr1*, *pckA*, and *gapB* operators. (a) C/T, Maxam Gilbert C+T sequencing reaction; C, control (protein-free methylated DNA, this lane is equivalent to a Maxam–Gilbert G > A sequencing reaction); B and U, bound and unbound fraction of methylated DNA subjected to binding with CcpN-His_5. The numbers in the gels and column diagrams show the positions of the corresponding nucleotides relative to the transcription start site. Binding sites I and II for CcpN have been denoted according to interference footprinting experiments. Close contacts are indicated by black and grey triangles for G and A, respectively. (b) Column diagrams indicating the relative strength of interference signals for both strands of the three operators. Only positive signals, i.e. signals that indicate contacts, are shown. Measured values are averaged from four independent experiments.

residues within the three operators. $KMnO_4$, a strong oxidising agent, specifically oxidises thymine, thus impeding protein contacts. In addition to thymine, guanine is modified by $KMnO_4$, which results in bands for guanine residues in the gels. Figure 2 shows the positions of the modified thymine interfering with CcpN binding. In general, contact distribution correlated well with that found by methylation interference footprinting. The *sr1* operator exhibited the following strong interference signals in binding site I: $T_{(-52)}$ and $T_{(-50)}$ at the top strand and $T_{(-48)}$, $T_{(-46)}$, $T_{(-44)}$ at the bottom strand. However, in contrast to the contacts to guanine and adenine, contacts to thymine ($T_{(-22)}$ and $T_{(-20)}$ at the top strand) were slightly closer in binding site II. Significant contacts in binding site II have not been found on the bottom strand.

Both in the *pckA* and *gapB* operators, the positions of contacted thymine corresponded perfectly to those identified for guanine and adenine by methylation interference, too. The focus of contacts was in site II in the case of *pckA* (five close contacts, see Figure 2, at the top strand and $T_{(-12)}$ and $T_{(-10)}$ at the bottom strand), while only less close contacts were found in binding site I at the top strand and no significant contact at the bottom strand. The *gapB* operator exhibited strong interference signals only in site I (mainly $T_{(-16)}$ and $T_{(-14)}$ on the top strand and $T_{(-12)}$ and $T_{(-10)}$ on the bottom strand), whereas only less close contacts were found in binding site II at the top strand and no significant contact at the bottom strand. In all three operators, thymine bases that showed the strongest interference signals were located next to contacted guanine bases, together forming the contact center within each binding site.

Hydroxylamine footprinting

NH_2OH footprinting was used to analyse CcpN contacts to cytosine in the three operators. Hydroxylamine, a strong reductive agent, causes ring opening specifically at cytosine bases and, in this way, interferes with contact formation between protein and DNA. Figure 3 shows that contacts to cytosine bases were found in all three operators; however, the contacts were less close compared to the three other bases. Interference signals of almost equal intensity were found in the *sr1* operator in site I ($C_{(-45)}$ at the top strand and $C_{(-53)}$ and $C_{(-51)}$ on the bottom strand) and in site II ($C_{(-17)}$ at the top strand and $C_{(-21)}$ and $C_{(-19)}$ on the bottom strand). The *pckA* operator showed only two contacted cytosine bases in binding site II and no contact in binding site I. Interestingly, in the *gapB* operator, three contacted

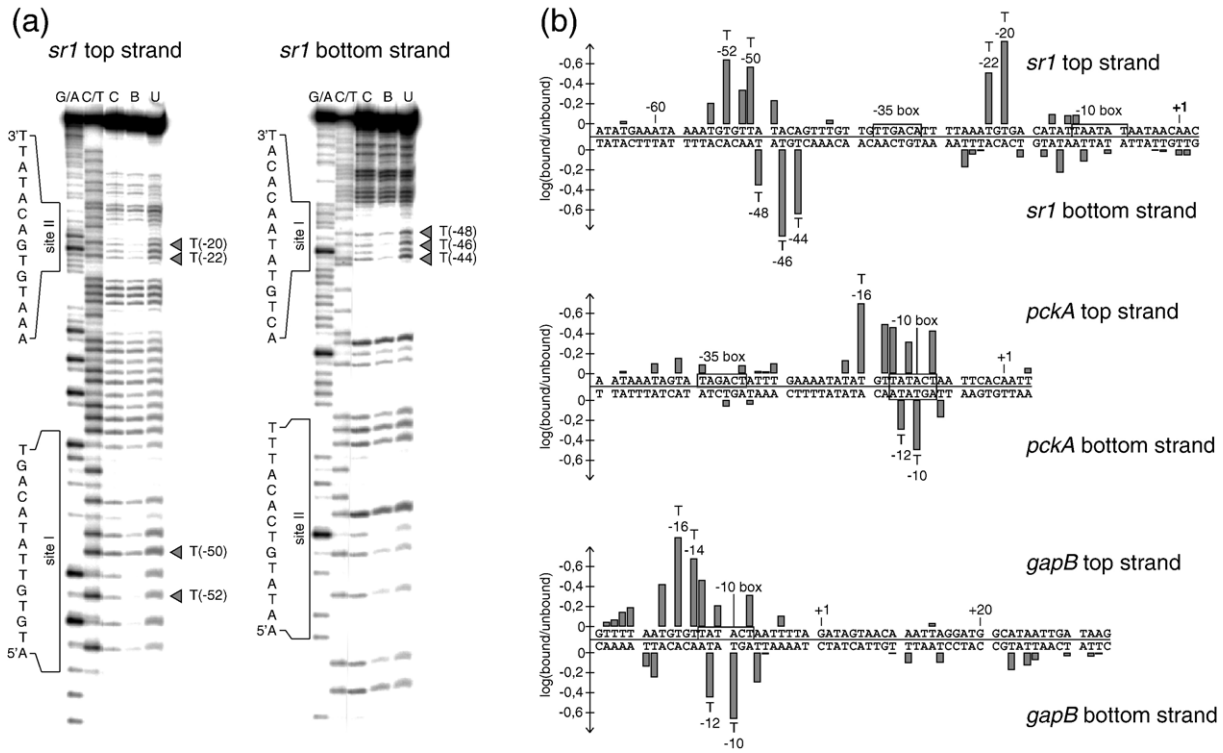

Figure 2. $KMnO_4$ interference of the *sr1*, *pckA* and *gapB* operators. (a) G/A, Maxam–Gilbert G>A sequencing reaction; C/T, Maxam–Gilbert C+T sequencing reaction; C, control (protein-free $KMnO_4$-treated DNA); B and U, bound and unbound fraction of methylated DNA subjected to binding with CcpN-His_5. The numbers in the gels and column diagrams show the positions of the corresponding nucleotides relative to the transcription start site. Binding sites I and II for CcpN are designated as in Figure 1. Close contacts are indicated by dark grey triangles. Only the gels for top and bottom strand of the *sr1* operator are shown. (b) The column diagrams present the relative strength of interference signals for both strands of the three operators. Only positive signals, i.e. signals that indicate contacts, are shown. Measured values are averaged from four independent experiments.

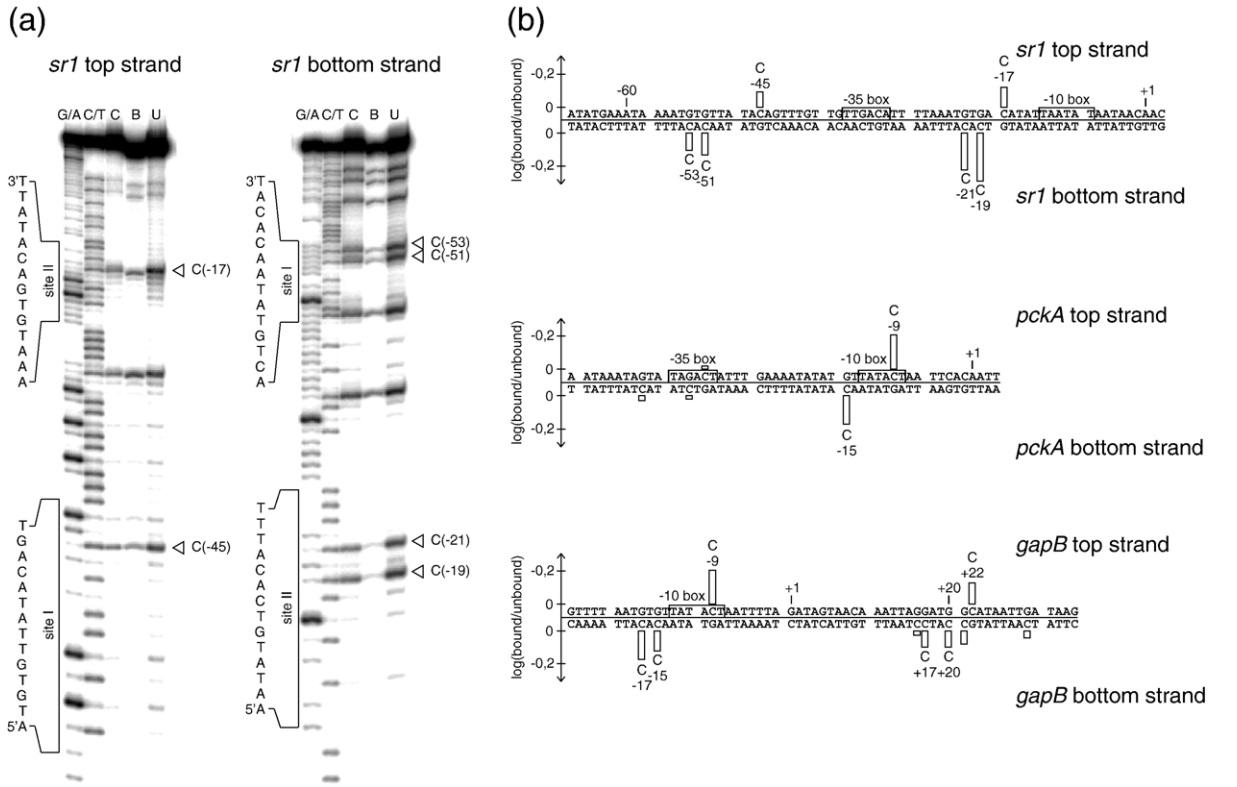

Figure 3. NH_2OH interference of the *sr1*, *pckA* and *gapB* operators (a) G/A, Maxam–Gilbert G>A sequencing reaction; C/T, Maxam–Gilbert C+T sequencing reaction; C, control (protein free NH_2OH-treated DNA); B and U, bound and unbound fraction of methylated DNA subjected to binding with CcpN-His_5. The numbers in the gels and column diagrams show the positions of the corresponding nucleotides relative to the transcription start site. Binding sites for CcpN are designated as in Figure 1. Close contacts are indicated by white triangles. Only the gels for top and bottom strand of the *sr1* operator are shown. (b) The column diagrams present the relative strength of interference signals for both strands of the three operators. As above, only positive signals are shown. Measured values are averaged from four independent experiments.

cytosine bases were found in both site I ($C_{(-9)}$ at the top strand and $C_{(-15)}$ and $C_{(-17)}$ at the bottom strand) and site II ($C_{(+22)}$ at the top strand and $C_{(+17)}$ and $C_{(+20)}$ at the bottom strand). However, due to the weak nature of these interference signals, contacts to cytosine do not seem to play an important role in the CcpN–DNA interaction.

Ethylation interference footprinting

To determine phosphate groups of the DNA backbone contacted by CcpN, ethylation interference experiments were carried out. Figure 4 presents the positions at which ethylation interfered with CcpN binding. Both binding sites in the *sr1* operator showed only two interference signals: In site I, $T_{(-50)}$ at the top strand and $A_{(-47)}$ at the bottom strand were contacted, and in site II, $A_{(-12)}$ and $A_{(-15)}$ at the bottom strand were contacted. In the *pckA* operator, contacts to the sugar-phosphate backbone were detected only in binding site II. Here, $T_{(-14)}$ at the top strand and $A_{(-8)}$ and $T_{(-10)}$ at the bottom strand exhibited interference signals. The same was found for the *gapB* operator, where only binding site I showed two contacts, to $T_{(-14)}$ at the top strand and $T_{(-10)}$ at the bottom strand.

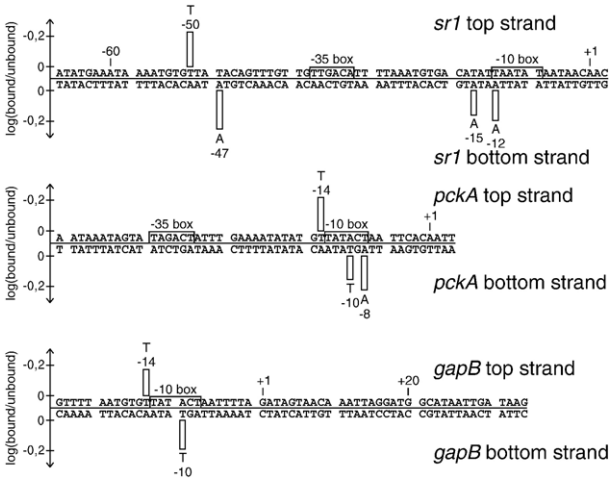

Figure 4. Ethylation interference of the *sr1*, *pckA* and *gapB* operators. The column diagrams present the relative strength of interference signals for both strands of the three operators. Only positive signals, i.e. signals that indicate contacts, are shown. Numbers in the column diagrams designate the positions of the corresponding nucleotides relative to the transcription start site. Measured values are averaged from three independent experiments.

Interestingly, the few DNA-backbone contacts were observed in most cases next to a contacted guanine residue. Obviously, these contacts play only a minor role in the binding of CcpN to its operators. Figure 5 summarises all probed contacts for the three operators.

EMSA

To determine the apparent equilibrium dissociation constants K_D for the CcpN-DNA complex, 23 bp double-stranded oligonucleotides containing a single CcpN binding site were incubated with increasing concentrations of CcpN-His_5 (Figure 6(a)). K_D values were estimated by non-linear regression using the average data from three independent experiments as described in Materials and Methods. The calculated K_D values as well as the binding energy ΔG for the CcpN–DNA interaction for the single sites are summarised in Table 1. Binding energy was calculated with the help of Van't Hoff's reaction isobare $\Delta G = -RT \ln(K)$, where R is the universal gas constant, T is the absolute temperature (in Kelvin) and K is the determined equilibrium association constant. The calculated K_D values for the single binding sites corresponded very well to the contacts that were observed by interference footprinting: Binding site I of the *gapB* operator, the one with the most and closest contacts (see Figure 1), showed the lowest K_D value, indicating a tight protein–DNA interaction, whereas binding site II of *gapB* or site I of *pckA*, both with less close contacts, exhibited high K_D values. Determined dissociation constants ranged from as low as 98 nM (*gapB*, site I) till 4.4 μM (*gapB*, site II).

To test whether the equilibrium dissociation constants differ when using whole operators, double-stranded oligonucleotides containing both CcpN binding sites were incubated with increasing concentrations of CcpN-His_5 (Figure 6(b)). K_D values were estimated by non-linear regression using the average data from three independent experiments as described in Materials and Methods. The apparent equilibrium dissociation constants were determined to be 19.3 nM, 15.5 nM and 12.8 nM for the *sr1*, *pckA* and *gapB* operator, respectively, and correspond well to the values determined by Servant *et al.*[11] All operators showed significantly lower K_D values than the single sites alone. At the *sr1* operator, the average K_D was decreased 30-fold, while the K_D of site I of the *pckA* and site II of the *gapB* operators was decreased 160-fold and 340-fold, respectively.

Figure 5. Overview of the contacts in all three operators. (a) Overview of all contacts. Symbols used to indicate contacts to the bases are shown in the box below. Filled symbols denote close contacts (50%–100% compared to the strongest signal), open symbols represent medium or less close contacts (15%–50% compared to the strongest signal). For clarity, contacts with less than 15% relative strength are not shown. The −35 and −10 regions are boxed, and the transcription start site is indicated. Binding sites I and II are designated on the basis of all interference footprinting experiments. (b) Alignment of the core sequences of all binding sites. Positions that coincide with the consensus are shown in bold.

Quantitative DNase I footprinting

Since occupancy of single sites is not detectable in EMSA, the affinity of CcpN to the single sites within the complete operator was measured by quantitative DNase I footprinting. This technique allows us to determine K_D values for site I and site II separately, even if they are located on one DNA

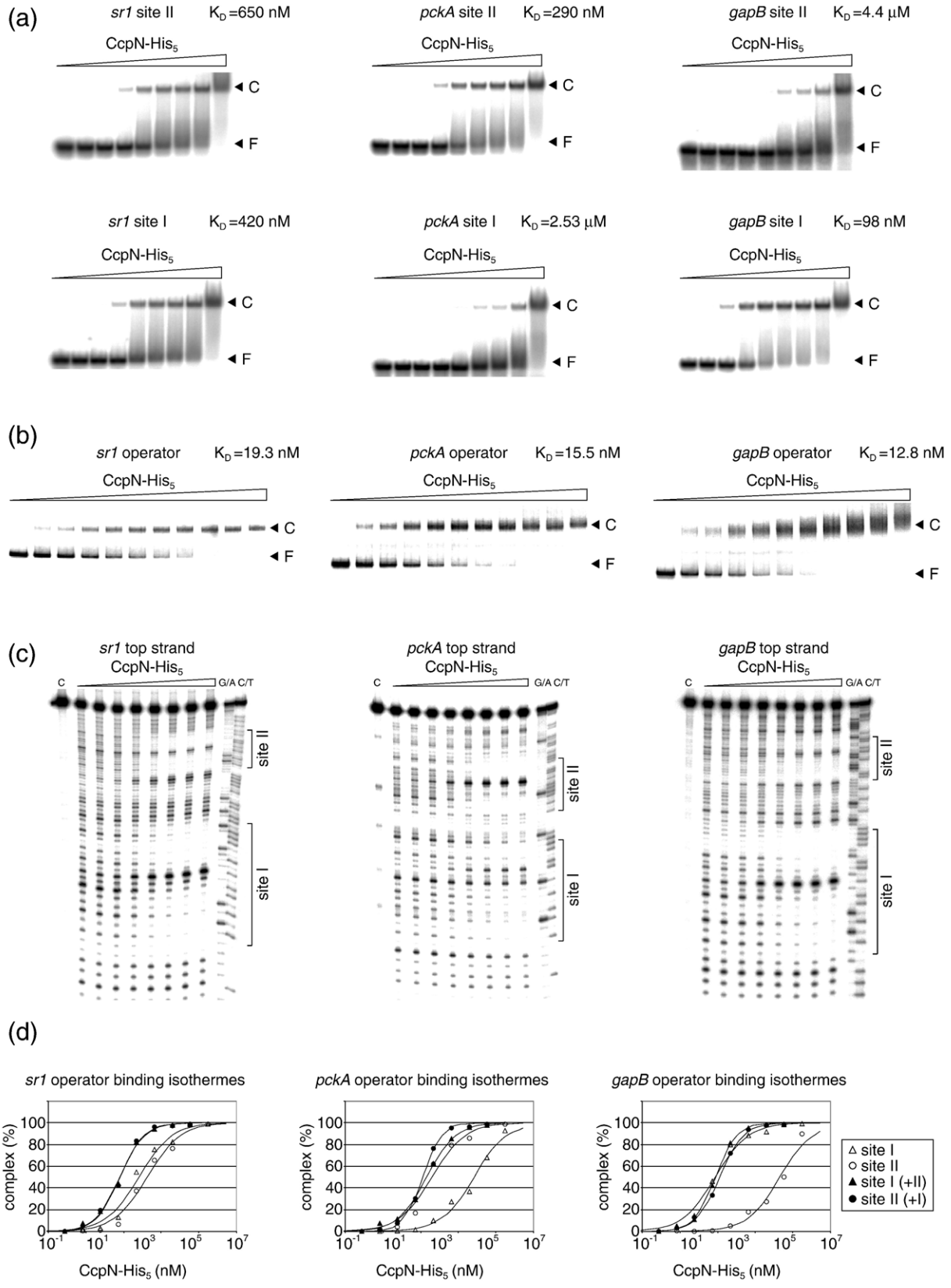

Figure 6 (*legend on opposite page*)

Table 1. Apparent dissociation constants and free energies for all CcpN binding sites

Single sites			Whole operator				
Binding site	K_D (nM)	ΔG (kJ/mol)	Binding site	K_D (nM)	ΔG (kJ/mol)	$\Delta\Delta G$ (kJ/mol)	h
sr1, site I	420 (±100)	−37.9 (±0.7)	*sr1*, site I	80 (±6)	−42.1 (±0.2)	−9.5	1.37
sr1, site II	650 (±130)	−36.8 (±0.5)	*sr1*, site II	81 (±16)	−42.1 (±0.5)		
pckA, site I	2530 (±310)	−33.2 (±0.3)	*pckA*, site I	145 (±32)	−40.6 (±0.6)	−9.8	1.45
pckA, site II	290 (±60)	−38.8 (±0.5)	*pckA*, site II	115 (±13)	−41.2 (±0.3)		
gapB, site I	98 (±12)	−41.6 (±0.3)	*gapB*, site I	89 (±6)	−41.9 (±0.2)	−9.7	1.49
gapB, site II	4400 (±15)	−31.8 (±0.0)	*gapB*, site II	114 (±48)	−41.2 (±1.0)		

Values were derived from three independent experiments. $\Delta\Delta G$ ($\Delta G_{complete} - \Delta G_{single}$) is the extra free energy that is gained when the two occupied sites are together on one DNA molecule, and h is the Hill coefficient.

fragment. To this end, 89 bp double-stranded oligonucleotides were incubated with increasing concentrations of CcpN-His_5 and, after equilibrium was reached, subjected to cleavage with DNase I (Figure 6(c)). In all experiments, the top strand was labelled, since the DNase I cleavage pattern of this strand was more homogeneous than that of the bottom strand. The degree of protection observed corresponded directly to the occupancy of the DNA by CcpN and allowed us to calculate the amount of complex formed. Apparent equilibrium dissociation constants were estimated by non-linear regression using the average data from three independent experiments. The calculated K_D values, the Hill coefficients and the binding energy ΔG for the CcpN–DNA interaction for all single sites are summarised in Table 1. Interestingly, the apparent dissociation constants for the complete operator sequences differed from those found for the investigated single sites and from the results of the footprinting experiments.

Cooperativity of CcpN binding

The K_D values for each side in the context of the whole operator were in all cases lower than for the corresponding single sites alone (see Table 1). This was especially true for binding site I of *pckA* and binding site II of the *gapB* operator. The gain in free energy upon CcpN binding to two separated single sites was lower than to two sites in a complete operator, i.e. the occupation of both sites in the operator is cooperative. This was verified by the finding that the K_D values obtained with DNA fragments spanning the whole operator are significantly lower than those obtained with single binding sites. Furthermore, when the values for the single site isotherms in the context of the whole operator were fit to the Hill equation (see Materials and Methods), the shape and the slope of the isotherms changed to a characteristic form for cooperative interactions. Moreover, the Hill coefficient h is in each case >1 (see Table 1), which is a reliable sign for cooperativity. In the case of *sr1*, where the two binding sites have nearly identical K_D values, the affinity of each site was increased by approximately equal amounts. By contrast, when one binding site was much stronger than the other, as in the case of *pckA* and *gapB*, the K_D value for the weaker binding site was improved dramatically (from 4.4 μM to 114 nM for *gapB* site II), but the K_D for the stronger binding site was mostly unaffected. A comparison of the binding isotherms for the single sites and the single-site isotherms for the complete operators that can be found in Figure 6(d) corroborates this conclusion.

Energetic calculations on CcpN-DNA interactions

Quantitative footprinting experiments like those described above were performed at 37 °C and 52 °C. The free energy ΔG was calculated on the basis of three independent experiments. Equation (1) describes the correlation between free energy (G), enthalpy (H) and entropy (S) and can be rearranged to yield equation (2), because ΔH and ΔS are independent of temperature. Thereby, T_1 and T_2 are

Figure 6. Determination of the binding isotherms for the *sr1*, *pckA* and *gapB* operators. (a) EMSAs of single CcpN binding sites. The 23 bp oligonucleotides were incubated with increasing concentrations of purified CcpN-His_5 (CcpN concentration from left to right: 0; 8.1 nM; 27.3 nM; 72.9 nM; 219 nM; 656 nM; 1.97 μM; 5.90 μM; 17.7 μM). F, free DNA; C, CcpN–DNA complex. To allow for a direct comparison with (c), EMSAs for binding sites I and II are shown in the same order as the binding sites appear in the DNase I footprinting gels. (b) EMSAs of whole CcpN operators. The 400 bp oligonucleotides were incubated with increasing concentrations of purified CcpN-His_5 (CcpN concentration from left to right: 0; 5.2 nM; 7.8 nM; 11.7 nM; 17.6 nM; 26.3 nM; 39.5 nM; 59.3 nM; 88.9 nM; 133 nM; 200 nM). F, free DNA; C, CcpN–DNA complex. The determined K_D values are given in each diagram. (c) Quantitative DNase I footprinting: C, control (uncleaved DNA); G/A, Maxam-Gilbert G > A sequencing reaction; C/T, Maxam-Gilbert C + T sequencing reaction. The 89 bp oligonucleotides containing both CcpN-binding sites were incubated with increasing amounts of CcpN-His5 (CcpN concentration from left to right: 0; 8.1 nM; 27.3 nM; 72.9 nM; 219 nM; 656 nM; 1.97 μM; 5.90 μM). Protected regions aredenoted site I and site II. (d) Binding isotherms of the single CcpN-binding sites and single-site isotherms for the whole operator sequence. Single sites I and II are represented by open triangles and circles, respectively. Filled triangles (site I) and circles (site II) designate single-site isotherms of the complete operator. The trend curves shown are averaged from three independent experiments.

310.15 K and 325.15 K, respectively, and G_1 and G_2 the free energies at the corresponding temperatures:

$$\Delta G = \Delta H - T\Delta S \quad (1)$$

$$\Delta S = \frac{\Delta G_1 - \Delta G_2}{T_2 - T_1} \quad (2)$$

Enthalpic and entropic contributions to CcpN-DNA binding were calculated using equation (2) and are summarised in Table 2. The CcpN–DNA interaction shows a small but unfavourable change in entropy that is overcome by a strong enthalpic contribution. This combination of enthalpy and entropy ensures that the CcpN–DNA interaction has nearly the same efficiency at all temperatures that are tolerated by *B. subtilis*.

Discussion

CcpN binds asymmetrically to its two consecutive binding sites in all three operators

Here, we report the high-resolution contact probing of the transcriptional repressor CcpN bound to its operator sites. CcpN, which has been identified recently as a repressor active under glycolytic conditions, is known to regulate three genes in *B. subtilis: sr1*, encoding a small untranslated RNA,[13] and genes for two gluconeogenic enzymes, *pckA* and *gapB*.[11] Using chemical interference footprinting with different chemical probes, we determined the bases contacted by CcpN in all three operators (summarised in Figure 5).

In all cases, two binding sites were identified, one of which was always contacted more strongly than the other. In the following, this site is referred to as the strong site, whereas the other is designated the weak site. Within all binding sites, core regions can be defined that resemble the consensus binding sequence TGTG(Y/A)YATAC that was previously determined for CcpN.[13] A comparison of all core regions with this consensus is presented in Figure 5(b).

In the *sr1* operator, the upstream binding site (site I, the strong site) was found to be contacted in a slightly stronger manner, but both binding sites showed extensive contacts especially to guanine and thymine residues (Figure 5) and less close contacts to adenine and cytosine residues, and to the sugar–phosphate backbone (Figures 1, 3 and 4). Moreover, both core regions conform well to the consensus. By contrast, contact distribution was found to be completely different in the other two operators. In the case of *pckA*, the majority of contacts were concentrated in the downstream binding site (site II, the strong site), where close contacts to all bases except cytosine were found (Figure 5). At site I, the weak site, only few and less close contacts were detected. Whereas the core region of the strong site again corresponded well to the consensus sequence, the core of the weak binding site deviated significantly from the consensus. Furthermore, although only one extended site appeared in the *pckA* DNase I footprint,[11] chemical interference revealed that the *pckA* operator consists of two binding sites, too.

Similar results have been found for the *gapB* promoter, except that the upstream site (site I) proved to be the strong site. As in the case of *pckA*, the most and the closest contacts were found in the strong site in the consensus-like core region, whereas site II showed only a low level of similarity to the consensus sequence. A series of *gapB* operator mutants tested by Servant *et al.*[11] can be evaluated in the light of the data published here: They found that a $T_{(-11)}\rightarrow A$ mutation, located in the strong site, severely inhibited CcpN binding, which can be explained by the close contact to the adenine residue on the complementary strand that we observed. Moreover, this position was shown to be invariant in the previously determined consensus sequence.[13] This holds true for the $A_{(23)}\rightarrow G$ mutation too, which concerns an invariant base in the weak site. However, the observed effect was not that pronounced, since the contribution of this position is not that great in this case. By contrast, the $T_{(-14)}\rightarrow G$ mutation showed almost no effect on the CcpN–DNA interaction, despite the close contacts that we mapped for this position. However, this site has been shown to be more variable in the consensus sequence,[13] and one could imagine that a mutation at this site is compensated by the surrounding sequence.

In all three operators, the major contacts determined with interference footprinting were contacts to guanine and thymine residues, and all focused within a core binding region. Since guanosine is methylated at N7 in the major groove, one can conclude that CcpN contacts its operator sequences primarily, but not exclusively, through contacts in the major groove, as found previously for many other proteins, e.g. RhaS from *E. coli*.[14] Like transcription factor TyrR from *E. coli*,[15] CcpN contacts its target through a large number of bases. Contacts to the sugar–phosphate backbone make only minor contributions to the CcpN–DNA interaction and, thus, do not seem to play an important role. Most probably, extended contacts to bases relieve the necessity to interact with the sugar–

Table 2. Reaction enthalpy and entropy for the CcpN–DNA interaction

Binding site	ΔG 37 °C (kJ/mol)	ΔG 52 °C (kJ/mol)	ΔH (kJ/mol)	ΔS (kJ/(mol K))
sr1, site I	−42.3	−41.8	−51.5	−0.03
sr1, site II	−42.7	−41.5	−68.4	−0.08
pckA, site I	−40.4	−39.5	−59.2	−0.06
pckA, site II	−41.7	−41.4	−48.6	−0.02
gapB, site I	−41.7	−41.3	−49.8	−0.03
gapB, site II	−41.2	−38.4	−100.1	−0.18

Quantitative footprinting was performed at 37 °C and 52 °C with DNA fragments carrying the whole operator sequence. The values were derived from three independent experiments.

phosphate backbone. Interestingly, contacts to the sugar–phosphate backbone were found mostly downstream from one of the guanine bases that provided one of the main contacts.

The occurrence of two binding sites with different contact strengths within one operator is rather peculiar, as many proteins with two binding sites bind these sites with more-or-less equal affinity.[16–18] In this regard, however, CcpN shows similarities with PurR,[19,20] whose operators have one strong and one weak binding site, too, although the differences are not as pronounced as in the case of CcpN.

CcpN binding sites are located at different positions at each operator

Previous DNase I footprinting experiments indicated that the binding site distribution is different among the three CcpN-regulated promoters.[11,13] Here, we substantiated these findings and determined the exact borders of the single binding sites using chemical interference footprinting experiments. Figure 5 shows that in all three operators, CcpN binding sites are located at different positions relative to the transcription start site.

At the *sr1* operator, site I was found to be centred upstream of the −35 box, around −48, and site II centred around −19. Bases within the −35 box were not contacted by CcpN, and only one base of the −10 box exhibited one less close contact. In contrast, in the *pckA* operator, site I overlapped the −35 box partially and site II the −10 box completely. The *gapB* operator revealed yet another positioning of the binding sites. Here, binding site I covered the −10 box as does site II in the case of *pckA*, and site II was located downstream from the transcription start site with its centre at position +19.

Diverse distribution of operator sites is not an uncommon feature. Beside transcription factors that show conserved binding site positioning, like CytR from *E. coli*,[21] numerous transcription factors bind to operators that are located at varying positions with regard to the promoter, as does CcpN. One example is CcpA, the major factor for carbon catabolite repression in *B. subtilis*, whose binding sites, termed *cre* elements, can be positioned differently relative to the transcription start site: Depending on their regulated gene, they are found at e.g. −33, −3 or +37.[22,23] Interestingly, all these *cre* elements mediate transcriptional repression, although their respective repression mechanism has not been elucidated.

Based on the distribution of the CcpN binding sites at the three different promoters, it is tempting to speculate about different repression mechanisms.[24] In the case of *sr1*, neither the −35 nor the −10 box are covered or contacted by CcpN. This might allow RNA polymerase to bind simultaneously with CcpN to the *sr1* promoter, which would exclude repression by steric hindrance and could result in inhibition of open complex formation, e.g. as found for the MerR repressor of *E. coli*.[25] Another conceivable mechanism is inhibition of promoter clearance, as shown for protein P4 of phage ϕ29 at the viral A2c promoter.[26] In contrast, at the *pckA* and *gapB* promoters, inhibition of transcription might occur by steric hindrance of RNAP binding, since at least one binding site of these promoters completely covers the −10 box, as it is the case for the Fur protein from *E. coli* as well as many other transcriptional repressors.[27] Future experiments will focus on the elucidation of the repression mechanism of CcpN at all three operators, for which the ligand that modulates CcpN activity[13] still needs to be identified.

CcpN binds cooperatively to its two binding sites

Our interference footprinting experiments indicated that CcpN contacts its respective binding sites with different strengths, especially at the *pckA* and *gapB* promoters (Figure 6). These results were confirmed by EMSAs using oligonucleotides carrying the single binding sites. The K_D values determined varied greatly from as low as 98 nM for the strong site of *gapB* to 4.4 μM for the associated weak site. The same was true for the strong and the weak site of the *pckA* operator, whereas the K_D values for the binding sites of the *sr1* operator did not differ much with 420 nM and 650 nM for site I and II, respectively, due to the only slight differences in contacts between these sites (Figure 6(a) and Table 1). Surprisingly, the K_D values obtained in EMSAs with DNA fragments containing the whole operators were greatly reduced, up to 340-fold for the weak site of *gapB*, compared to those for the single binding sites and correlated well to what was found by Servant *et al.*[11] Obviously, two binding sites on one DNA strand dramatically increase the binding efficiency of CcpN.

The determination of the K_D values for the single site in the context of the complete operators confirmed these results. Here, in all three operators, both sites were occupied with almost the same efficiency and showed only slight variations in K_D values between the strong and the weak sites. In addition, all K_D values were, partly significantly, decreased. In the *sr1* operator, both binding sites showed an almost equal increase in binding affinity, whereas at the *pckA* and *gapB* operators, only the weak binding sites exhibited a significantly lower K_D in the complete operator. Thereby, the affinity of the strong binding sites was mainly unchanged or increased only slightly. This increase in binding affinity leads to an increase in energy gain upon CcpN binding: Binding to two sites that are in close vicinity on one DNA strand is energetically more favourable than binding to two separate strands. Furthermore, a change in the shape and slope of the binding curves, resulting from a Hill coefficient $h > 1$, which indicates cooperativity, was observed. All this leads to the conclusion that CcpN apparently binds to its operators in a cooperative way, but this cooperativity is different for the three promoters. While the *sr1* operator shows two more-or-less equal

binding sites, the *pckA* and *gapB* operators are composed of one main and one auxiliary site, and binding to the auxiliary site was found to be greatly improved in the presence of the main site. Strong and weak binding sites were observed also for the DeoR repressor–operator system in *B. subtilis*.[28] The DeoR operator consists of one full and one half binding site but, unlike CcpN, DeoR does not bind single sites.

Cooperative binding suggests an interaction between the CcpN molecules bound to the stronger and weaker sites. Conspicuously, the spacer region between the two binding sites differs between the three operators (see Figures 5 and 7). Whereas in the *sr1* and *gapB* operator it comprises three helical turns, in the *pckA* operator, only two helical turns separate the two binding sites. This indicates that, at least in the case of *sr1* and *gapB*, CcpN most likely bends its operator DNA to enable a contact between the two binding entities.

CcpN binding is driven exclusively by a strong binding enthalpy

The determination of binding constants and free binding energy showed that CcpN binding to its operator is unfavourable in terms of entropy change, i.e. entropy decreases upon CcpN–DNA interaction. This effect is overcome by a strong favourable enthalpic contribution, likely due to the numerous and close contacts made to the bases in the operator sequence. This has a clear practical consequence for *B. subtilis*: Since this species tolerates temperatures from as low as 12 °C to as high as 52 °C, a strong binding enthalpy, which is temperature-independent, and a low binding entropy change, whose contribution to total binding energy depends on the temperature, ensures that CcpN retains its binding affinity and K_D value for its operators over a large temperature scale.

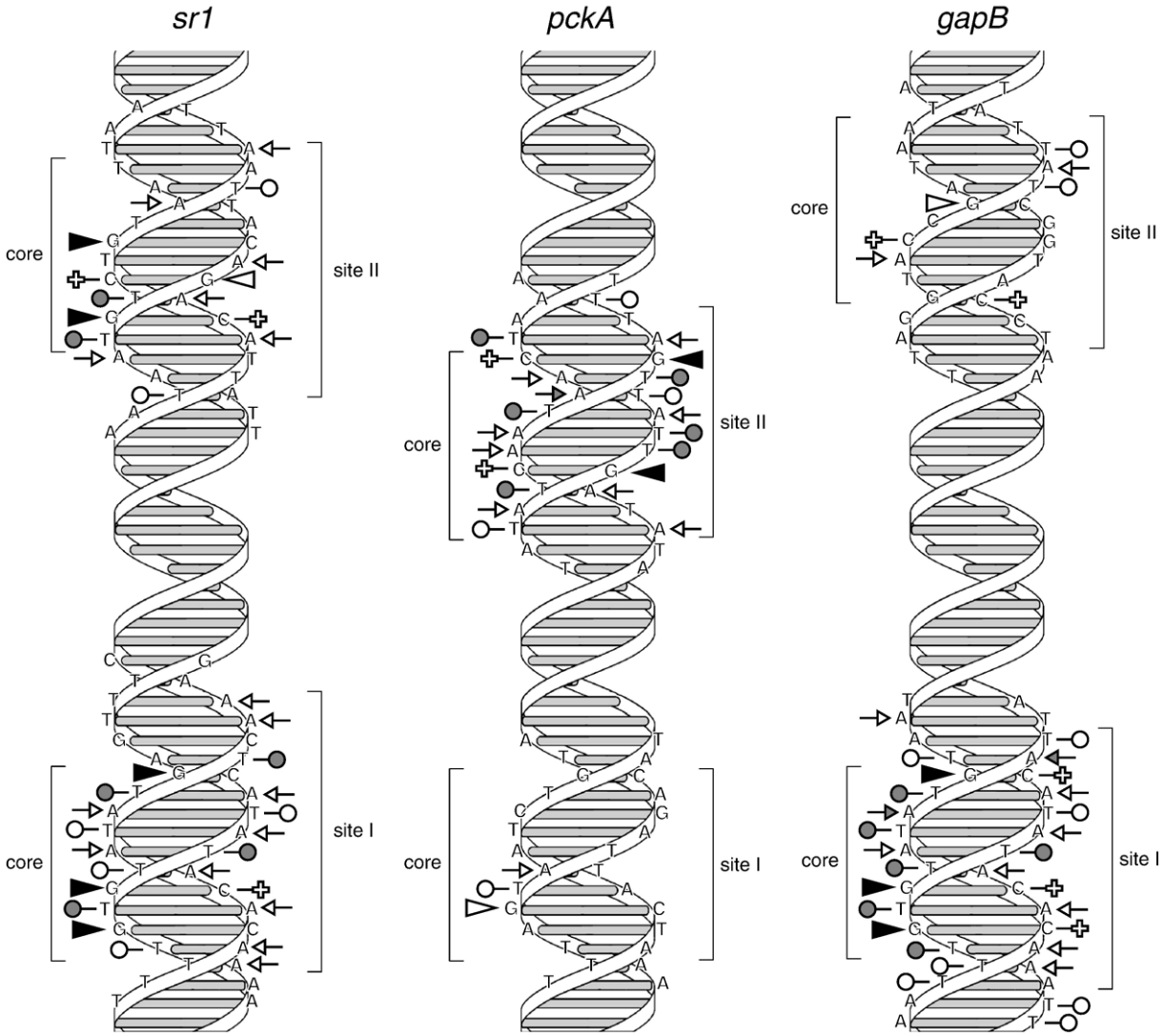

Figure 7. Ribbon model of all CcpN operators. Symbols used to indicate contacts between CcpN and the operator are the same as in Figure 5. The binding sites and the core binding regions are indicated by brackets.

Materials and Methods

Preparation of labelled CcpN targets

Oligonucleotides were purified by treatment with piperidine for 30 min at 90 °C to avoid contamination with depurinated DNA resulting from the removal of the protective groups during synthesis. Subsequently, purified oligonucleotides were 5′ end-labelled with [γ-^{32}P]ATP using bacteriophage T4 polynucleotide kinase (NEB) and purified from denaturing 15% (w/v) polyacrylamide gels.[29] Pairwise combinations of labelled and unlabelled oligonucleotides were annealed by incubation at 65 °C for 5 min and subsequent slow cooling to 37 °C. The top and bottom strand of all oligonucleotides carry two G or C residues, respectively, at each end to facilitate correct annealing and to promote additional stability. Labelled double-stranded DNA fragments for the EMSAs with the whole operator sequences were obtained by PCR using the appropriate primer pairs (all oligonucleotides used in this study are summarised in Table 3). The PCR products were purified from an ethidium bromide-stained native 6% polyacrylamide gel and 5′ end-labelled as described above. The DNA was then separated from unincorporated [γ-^{32}P]ATP by passage through a Sephadex column.

Over-expression and purification of CcpN

A *ccpN* over-expression strain was constructed by cloning a NcoI/BglII digested PCR fragment obtained with primers SB673 and SB674 on chromosomal DNA of *B. subtilis* DB104 into the pQE60 NcoI/BglII vector (Qiagen). The resulting vector was designated pQGDR. For cloning and subsequent purification of the C-terminally His-tagged protein, *E. coli* strain TG1(REP4) was used. The sequence was confirmed using a Sequenase kit (Amersham Bioscience).

A TG1(REP4, pQGDR) overnight culture grown in TY with 50 μg/ml of ampicillin and 25 μg/ml of kanamycin was diluted 100-fold, grown for an additional 3 h and induced with 1 mM IPTG. After 2.5 h, cells were harvested by centrifugation and sonicated for 10 min in sonication buffer (50 mM sodium phosphate (pH 8.0), 300 mM NaCl, 10 mM imidazole). The supernatant obtained by centrifugation was purified over a Ni-agarose column (Qiagen). The column was washed twice with washing buffer (50 mM sodium phosphate (pH 8.0), 300 mM NaCl, 20 mM imidazole) and, afterwards, CcpN was eluted with elution buffer (50 mM sodium phosphate (pH 8.0), 300 mM NaCl, 250 mM imidazole). Purification was followed by SDS-PAGE. In this way, approximately 80% pure CcpN-His_5 was obtained and stored with 50% (v/v) glycerol at −20 °C.

EMSA and determination of apparent equilibrium dissociation constant K_D

Binding reactions were performed in a final volume of 10 μl containing 0.5× TBE, 0.05 g/l of herring sperm DNA as non-specific competitor, 1 nM end-labelled DNA fragment and 5.2 nM to 17.7 μM CcpN-His_5. All CcpN-His_5 dilutions were made in storage buffer and the same volume of diluted protein was used in each sample to ensure an equal concentration of salt. After incubation at 37 °C for 15 min, the reaction mixtures were separated on native 6% (for whole operator DNA fragments) or 8% (for 23 bp DNA fragments) polyacrylamide gels run at room temperature for 1 h at 200 V. Visualisation and quantification of the bands were performed using a Fuji-PhosphorImager and the PCBAS 2.09 quantification software (Raytest). All autoradiograms were made from dried gels. The image data generated by scanning the gel are linear proportionally to the radiation intensity of the sample. The amount of CcpN–DNA complex relative to the concentration of CcpN was fit with the non-linear regression programme Solver (included in Microsoft® Excel) to the following equation:

$$[C] = \frac{[D][P]}{K_D + [P]}$$

where [C], [D] and [P] represent total concentrations of formed complex, DNA and protein, respectively, and K_D is the apparent equilibrium dissociation constant.

Methylation interference footprinting

The 5′ end-labelled DNA fragments were modified by dimethyl sulphate as described for the G > A reaction using the Merck oligonucleotide sequencing kit. Modified DNA was subjected to CcpN binding and EMSA as described above. Bound and unbound fractions were separated on a native 6% polyacrylamide gel and visualised by wet autoradiographic exposure. Bound and unbound DNA was cut out and eluted from the gel by diffusion (elution buffer: 50 mM Tris–HCl (pH 8.0), 10 mM EDTA, 0.15 M NaCl), treated with phenol/chloroform and precipitated in ethanol. DNA samples and protein-free DNA as control were depurinated for 15 min at 90 °C, cleaved by piperidine for 30 min at 90 °C, precipitated in ethanol twice, resuspended in formamide loading dye and separated on a 15% polyacrylamide sequencing gel.

Potassium permanganate interference footprinting

The 5′ end-labelled DNA fragments were modified by $KMnO_4$ as described.[30] Modified DNA was subjected to CcpN binding and EMSA as described above. Bound and unbound fractions were separated on a native 6% polyacrylamide gel and isolated as described above. DNA samples and protein-free DNA as control were cleaved by piperidine for 30 min at 90 °C, precipitated in ethanol twice, resuspended in formamide loading dye and separated on a 15% polyacrylamide sequencing gel.

Hydroxylamine interference footprinting

The 5′ end-labelled DNA fragments were modified by NH_2OH as dscribed.[30] Modified DNA was subjected to CcpN binding and EMSA as described above. Bound and unbound fractions were separated on a native 6% polyacrylamide gel and isolated as described above. DNA samples and protein-free DNA as control were cleaved by piperidine for 30 min at 90 °C, precipitated in ethanol twice, resuspended in formamide loading dye and separated on a 15% polyacrylamide sequencing gel.

Ethylation interference footprinting

The 5′ end labelled DNA fragments were modified by N-ethyl-N-nitrosourea (Sigma) as described.[31] Modified DNA was subjected to CcpN binding and EMSA as

Table 3. Oligonucleotides used in this study

Designation	Sequence	Purpose
SB499	5′ GGAAAATGTGTTATACAGTTTGG	*sr1*, site I, upper strand
SB500	5′ CCAAACTGTATAACACATTTTCC	*sr1*, site I, lower strand
SB964	5′ GGTAAATGTGACATATTAATAGG	*sr1*, site II, upper strand
SB965	5′ CCTATTAATATGTCACATTTACC	*sr1*, site II, lower strand
SB962	5′ GGAAATAGTATAGACTATTTGGG	*pckA*, site I, upper strand
SB963	5′ CCCAAATAGTCTATACTATTTCC	*pckA*, site I, lower strand
SB602	5′ GGAATATATGTTATACTAATTGG	*pckA*, site II, upper strand
SB603	5′ CCAATTAGTATAACATATATTCC	*pckA*, site II, lower strand
SB598	5′ GGTTAATGTGTTATACTAATTGG	*gapB*, site I, upper strand
SB599	5′ CCAATTAGTATAACACATTAACC	*gapB*, site I, lower strand
SB960	5′ GGAAATTAGGATGGCATAATTGG	*gapB*, site II, upper strand
SB961	5′ CCAATTATGCCATCCTAATTTCC	*gapB*, site II, lower strand
SB869	5′ GGATATGATGATATGAAATAAAATGTGTTATACAGTTTGTTGTTGACATTTTAAATGTGACATATTAATATAATAACAACAAAAGAAGG	*sr1*, complete operator, upper strand
SB870	5′ CCTTCTTTTGTTGTTATTATATTAATATGTCACATTTAAAATGTCAACAACAAACTGTATAACACATTTTATTTCATATCATCATATCC	*sr1*, complete operator, lower strand
SB886	5′ GGATGCACACACAAAAAACAATAAATAGTATAGACTATTTGAAAATATATGTTATACTAATTCACAATTAGCAAAACACAAAAAACGGG	*pckA*, complete operator, upper strand
SB887	5′ CCCGTTTTTTGTGTTTTGCTAATTGTGAATTAGTATAACATATATTTTCAAATAGTCTATACTATTTATTGTTTTTTGTGTGTGCATCC	*pckA*, complete operator, lower strand
SB894	5′ GGTACTGGCGAATTTGTTTTAATGTGTTATACTAATTTTAGATAGTAACAAATTAGGATGGCATAATTGATAAGGGGTGTCCAACATGG	*gapB*, complete operator, upper strand
SB895	5′ CCATGTTGGACACCCCTTATCAATTATGCCATCCTAATTTGTTACTATCTAAAATTAGTATAACACATTAAAACAAATTCGCCAGTACC	*gapB*, complete operator, lower strand
SB673	5′ GAATTCCCATGGGAAGTACGATCGAACTAAAT	Plasmid pQGDR
SB674	5′ CTGCAGAGATCTTTATTAGTGATGGTGATGGTGTAGGATTTCATTTTCAGA	Plasmid pQGDR
SB342	5′ CCCAGGAGAAATTATTACAG	*sr1* downstream primer
SB423	5′ TCGAGGATCCAACAAGGTGAATATGATGAT	*sr1* upstream primer
SB1027	5′ GAGGGCAGTCAGTGCGGAGC	*gapB* upstream primer
SB1028	5′ CAATAAAAAATAAAAAGCATGCGGCTTTAAGCCGCATGCTTTTTTAGCCACAACCTCTTTGTCGT	*gapB* downstream primer
SB1029	5′ AGAGTATCCGCTCAATGAAA	*pckA* upstream primer
SB1030	5′ CAATAAAAAATAAAAAGCATGCGGCTTTAAGCCGCATGCTTTTTTTGTTGTCGCGCGAACAGCAC	*pckA* downstream primer

Oligonucleotides SB673 and SB674 were used to construct plasmid pQGDR. SB342, SB423 and SB1027-SB1030 were used as primers for the amplification of whole operator fragments. All other oligonucleotides were annealed pairwise to create double-stranded targets for footprinting experiments and EMSAs.

described above. Bound and unbound fractions were separated on a native 6% polyacrylamide gel and isolated as described above. The DNA was cleaved with 143 mM NaOH at 90 °C for 30 min as described.[32] Protein-free DNA as a control was prepared by NaOH cleavage of an aliquot of the ethylated DNA. After precipitation in ethanol twice and resuspension in formamide loading dye, the samples were separated on a 15% polyacrylamide sequencing gel.

Densitometric quantification of the footprinting experiments

Band intensities were determined with quantification software (PCBAS 2.09, Raytest) and, afterwards, normalised by dividing them by the total band intensity of the same lane to correct for unequal loading. Data were plotted as logarithm (log) of the ratio of band intensity of bound DNA *versus* band intensity of the unbound DNA for each base position. Negative values were interpreted as interference signals.

Quantitative DNase I footprinting

DNase I footprinting was performed in a final volume of 10 μl containing 0.5× TBE, 6.25 mM $MgCl_2$, 0.05 g/l of herring sperm DNA, 1 nM end-labelled DNA fragment and 8.1 nM to 5.9 μM CcpN-His_5. After incubation at 37 °C for 30 min, the samples were treated with 1 μl of DNase I (Roche, 0.05 U/μl) for 2 min at 37 °C. Two control samples, one without protein, one without DNase I, were treated in parallel. The reaction was stopped by extraction with phenol and subsequent precipitation in ethanol. The pellets were dissolved in 3 μl of formamide loading dye, denatured for 5 min at 90 °C and separated on a denaturing 15% polyacrylamide gel along with a Maxam–Gilbert sequencing reaction obtained from the same DNA fragment. The dried gel was analysed by PhosphorImaging. DNA occupancy by CcpN was determined by measuring the band intensity at the binding sites divided by the intensity at an unoccupied part of the DNA.

To ensure that the CcpN–DNA complex is at equilibrium, footprinting experiments with different incubation times before DNase I cleavage were carried out. Steady state was reached no later than after 5 min of incubation. To show that DNase I is not able to displace CcpN from its operator, footprinting experiments with different concentration of DNase I were performed. The amount of CcpN-DNA complex relative to the CcpN concentration was fitted with the non-linear regression programme Solver (included in Microsoft® Excel) to the Hill equation:

$$[C] = \frac{[D][P]^h}{K_D^h + [P]^h}$$

where [C], [D] [P] and h represent total concentrations of formed complex, DNA, protein and the Hill coefficient, respectively, and K_D is the apparent equilibrium dissociation constant.

Acknowledgements

We thank E. Birch-Hirschfeld (Institut für Virologie, Jena) for synthesising the oligodeoxyribonucleotides. This work was supported by grant BR1552/6-1 from Deutsche Forschungsgemeinschaft (to S. B). A. L. is financed by a scholarship from the "Fonds der chemischen Industrie".

References

1. Steinmetz, M. (1993). Carbohydrate catabolism: pathways, enzymes, genetic regulation, and evolution. In Bacillus subtilis *and Other Gram-positive Bacteria: Biochemistry, Physiology, and Molecular Genetics* (Sonenshein, A. L., Hoch, J. A. & Losick, R., eds), pp. 157–170, Am. Soc. Microbiol., Washington, DC.
2. Reizer, J., Bachem, S., Reizer, A., Arnaud, M., Saier, M. H., Jr & Stülke, J. (1999). Novel phosphotransferase system genes revealed by genome analysis: the complete complement of PTS proteins encoded within the genome of *Bacillus subtilis*. *Microbiology*, **145**, 3419–3429.
3. Monod, J. (1942). *Recherches sur la Croissance des Cultures Bacteriennes.* Hermann et Cie, Paris.
4. Chambliss, G. H. (1993). Carbon source mediated catabolite repression. In Bacillus subtilis *and Other Gram-positive Bacteria: Biochemistry, Physiology, and Molecular Genetics* (Sonenshein, A. L., Hoch, J. A. & Losick, R., eds), pp. 212–219, Am. Soc. Microbiol., Washington, DC.
5. Emmer, M., deCrombrugghe, B., Pastan, I. & Perlman, R. (1970). Cyclic AMP receptor protein of *E. coli*: its role in the synthesis of inducible enzymes. *Proc. Natl Acad. Sci. USA*, **66**, 480–487.
6. Eron, L., Arditti, R., Zubay, G., Connaway, S. & Beckwith, J. R. (1971). An adenosine 3′:5′-cyclic monophosphate-binding protein that acts on the transcription process. *Proc. Natl Acad. Sci. USA*, **68**, 215–218.
7. Mach, H., Hecker, M. & Mach, F. (1984). Evidence for the presence of cyclic adenosine monophosphate in *Bacillus subtilis*. *FEMS Microbiol. Letters*, **22**, 27.
8. Stülke, J. & Hillen, W. (2000). Regulation of carbon catabolism in Bacillus species. *Annu. Rev. Microbiol.* **54**, 849–880.
9. Fillinger, S., Boschi-Muller, S., Azza, S., Dervyn, E., Branlant, G. & Aymerich, S. (2000). Two glyceraldehyde-3-phosphate dehydrogenases with opposite physiological roles in a nonphotosynthetic bacterium. *J. Biol. Chem.* **275**, 14031–14037.
10. Yoshida, K., Kobayashi, K., Miwa, Y., Kang, C. M., Matsunaga, M., Yamaguchi, H. *et al.* (2001). Combined transcriptome and proteome analysis as a powerful approach to study genes under glucose repression in *Bacillus subtilis*. *Nucl. Acids Res.* **29**, 683–692.
11. Servant, P., Le Coq, D. & Aymerich, S. (2005). CcpN (YqzB), a novel regulator for CcpA-independent catabolite repression of *Bacillus subtilis* gluconeogenic genes. *Mol. Microbiol.* **55**, 1435–1451.
12. Goldie, H. & Medina, V. (1990). Physical and genetic analysis of the phosphoenolpyruvate carboxykinase (*pckA*) locus from *Escherichia coli* K12. *Mol. Gen. Genet.* **220**, 191–196.
13. Licht, A., Preis, S. & Brantl, S. (2005). Implication of CcpN in the regulation of a novel untranslated RNA (SR1) in *Bacillus subtilis*. *Mol. Microbiol.* **58**, 189–206.
14. Bhende, P. M. & Egan, S. M. (1999). Amino acid-DNA contacts by RhaS: an AraC family transcription activator. *J. Bacteriol.* **181**, 5185–5192.
15. Hwang, J. S., Yang, J. & Pittard, A. J. (1999). Specific contacts between residues in the DNA-binding domain

of the TyrR protein and bases in the operator of the *tyrP* gene of *Escherichia coli*. *J. Bacteriol.* **181**, 2338–2345.
16. Steinmetzer, K. & Brantl, S. (1997). Plasmid pIP501 encoded transcriptional repressor CopR binds asymmetrically at two consecutive major grooves of the DNA. *J. Mol. Biol.* **269**, 684–693.
17. Brenowitz, M., Senear, D. F., Shea, M. A. & Ackers, G. K. (1986). "Footprint" titrations yield valid thermodynamic isotherms. *Proc. Natl Acad. Sci. USA*, **83**, 8462–8466.
18. Lewis, M. (2005). The Lac repressor. *Crit. Rev. Biol.* **328**, 521–548.
19. Shin, B. S., Stein, A. & Zalkin, H. (1997). Interaction of *Bacillus subtilis* purine repressor with DNA. *J. Bacteriol.* **179**, 7394–7402.
20. Bera, A. K., Zhu, J. & Zalkin, H. (2003). Functional dissection of the *Bacillus subtilis* pur operator site. *J. Bacteriol.* **185**, 4099–4109.
21. Collado-Vides, J., Magasanik, B. & Gralla, J. D. (1991). Control site location and transcriptional regulation in *Escherichia coli*. *Microbiol. Rev.* **55**, 371–394.
22. Weickert, M. J. & Chambliss, G. H. (1990). Site-directed mutagenesis of a catabolite repression operator sequence in *Bacillus subtilis*. *Proc. Natl Acad. Sci. USA*, **87**, 6238–6242.
23. Grundy, F. J., Turinsky, A. J. & Henkin, T. M. (1994). Catabolite regulation of *Bacillus subtilis* acetate and acetoin utilization genes by CcpA. *J. Bacteriol.* **176**, 4527–4533.
24. Rojo, F. (2001). Mechanisms of transcriptional repression. *Curr. Opin. Microbiol.* **4**, 145–151.
25. Heltzel, A., Lee, I. W., Totis, P. A. & Summers, A. O. (1990). Activator-dependent preinduction binding of sigma-70 RNA polymerase at the metal-regulated *mer* promoter. *Biochemistry*, **29**, 9572–9584.
26. Monsalve, M., Mencia, M., Salas, M. & Rojo, F. (1996). Protein p4 represses phage phi 29 A2c promoter by interacting with the alpha subunit of *Bacillus subtilis* RNA polymerase. *Proc. Natl Acad. Sci. USA*, **93**, 8913–8918.
27. Escolar, L., Perez-Martin, J. & de Lorenzo, V. (1998). Coordinated repression *in vitro* of the divergent *fepA-fes* promoters of *Escherichia coli* by the iron uptake regulation (Fur) protein. *J. Bacteriol.* **180**, 2579–2582.
28. Zeng, X. & Saxild, H. H. (1999). Identification and characterization of a DeoR-specific operator sequence essential for induction of *dra-nupC-pdp* operon expression in *Bacillus subtilis*. *J. Bacteriol.* **181**, 1719–1727.
29. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor NY.
30. Rubin, C. M. & Schmid, C. W. (1980). Pyrimidine-specific chemical reactions useful for DNA sequencing. *Nucl. Acids Res.* **8**, 4613–4619.
31. Rimphanitchakit, V. & Grindley, N. D. F. (1991). The study of protein-DNA contacts by ethylation interference. In *A Laboratory Guide to in Vitro Studies of Protein-DNA Interactions* (Jost, J. P. & Saluz, H. P., eds), pp. 111–120, Birkhäuser Verlag, Basel.
32. Büning, H., Baeuerle, P. & Zorbas, H. (1995). A new interference footprinting method for analyzing simultaneously protein contacts to phosphate and guanine residues on DNA. *Nucl. Acids Res.* **23**, 1443–1444.

Edited by M. Gottesman

(Received 21 July 2006; received in revised form 4 September 2006; accepted 6 September 2006)
Available online 12 September 2006