A precise DNA bend angle is essential for the function of the phage φ29 transcriptional regulator

Laura Pérez-Lago, Margarita Salas and Ana Camacho*

Instituto de Biología Molecular 'Eladio Viñuela' (CSIC), Centro de Biología Molecular 'Severo Ochoa' (CSIC-UAM), Universidad Autónoma, Canto Blanco, 28049 Madrid, Spain

ABSTRACT

Bacteriophage φ29 protein p4 is essential for the regulation of the switch from early to late phage transcription. The protein binds to two regions of the phage genome located between the regulated promoters. Each region contains two inverted repeats separated by 1 bp. We used circular permutation assays to study the topology of the DNA upon binding of the protein and found that p4 induced the same extent of bending independent of the topology of the binding region. In addition, the results revealed that the p4-induced bending is not dependent on the affinity to the binding site but is intrinsic to p4 binding. Independent binding sites were identified through the characterization of the minimal sequence required for p4 binding. The protein has different affinity for each of its binding sites, with those overlapping the A2c and A2b promoter cores (sites 1 and 3), having the highest affinity. The functionality of the p4 binding sites and the contribution of p4-mediated promoter restructuring in transcription regulation is discussed.

INTRODUCTION

Most transcription factors affect the topology of the DNA upon interaction with their binding sequence. DNA bending by transcription factors is an important facet of their function (1) and participates in transcription regulation by allowing factors binding to distal sites to act synergistically, by permitting correct interactions between regulatory factors and the transcription machinery, and by providing the necessary promoter conformation for its interaction with the RNA polymerase (RNAP). It can be anticipated that a protein-induced bend may synergize or antagonize the interactions with the required bend of the sequence for optimal transcription complex formation. This effect, called structural synergy (2), was first shown

experimentally with CAP by placing an intrinsic bend in different helical phases relative to a CAP-binding site (3). The dependency of the proper helical phasing of two operators on the repression of the *gal* promoters (4) and the effect of the DNA-induced bending on gene regulation has been documented (5–7). In addition, protein-directed DNA bending has been proposed to regulate the assembly of higher order DNA–multiprotein complexes and to modulate gene expression through the lifetime of such complexes, since transcription initiation is a non-equilibrium process.

The relationship between transcription factor-mediated bending and formation of competent transcription factor complexes has also been shown in eukaryotes. The yeast MADS-box transcription factor Mcm1p elicits substantial bending that determines local promoter architecture, allowing the formation of ternary transcription factors complexes that maintain correct gene expression profiles (8). On the other hand, it has been demonstrated that different bend angles induced by the mouse Sox2 factor, *trans*-activator of the *Fgf4* gene, elicited differences on the level of promoter activation (9).

Bacteriophage φ29 protein p4 is a homodimeric regulator that plays a critical role in transcription control. In fact, in the presence of the viral histone-like protein p6 (10), it is essential for the synchronic transition from early promoters A2c and A2b transcription to late transcription from promoter A3 (11–13). High-resolution chemical footprinting has shown that p4 binds to two regions (named 1 and 2) of 63 bp expanding from nucleotides −25 to −84 relative to promoter A2c and from position −27 to −87 relative to promoter A2b transcription start sites (14). Each region has two inverted repeats close to the consensus sequence TTGAAAAA-15bp-TTTTACAA, defined as essential for binding of p4 [see Figure 1 and (15)]. At each region, protein p4 binding results in the protection of six sets of nucleotides, named 1′–6′ in region 1 and 7′–12′ in region 2 (14). The two sets of protections are separated by 10 bp, thus, on the same DNA face, and four helix-turns separate one group of six protections from the other. Positions 1′–6′ and 7′–12′ encompass two inverted repeats each, named sites 1 and 2, and 3 and 4, respectively. Hence, p4 contacts the DNA backbone of regions 1 and 2 in three homologous positions of each
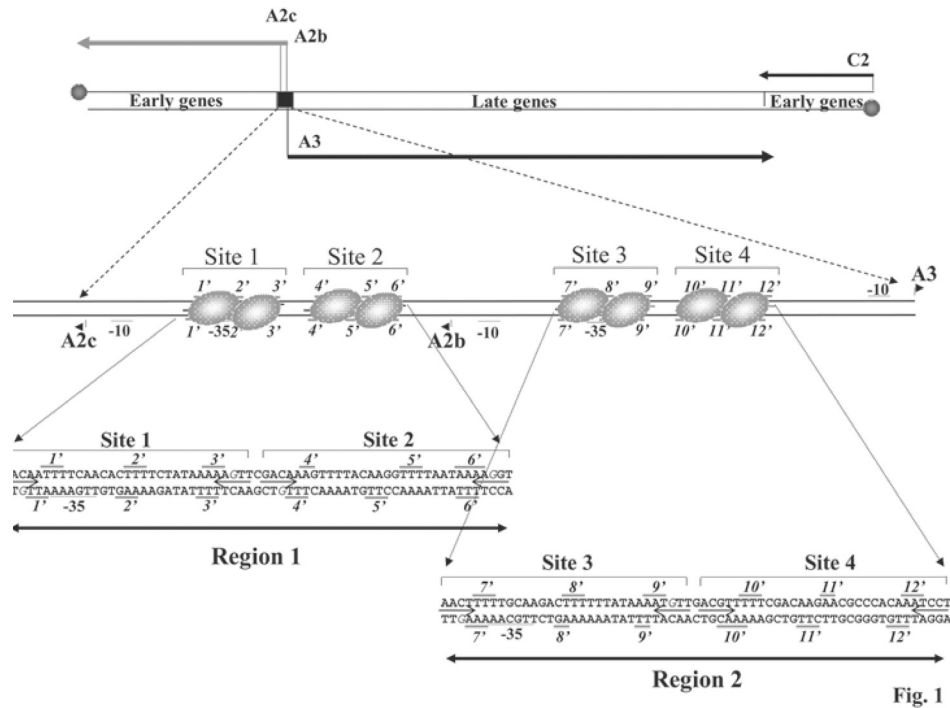

Figure 1. Protein p4 binding regions. The diagram shows the organization of the phage φ29 genome with the divergent transcription of promoters A2c, A2b and A3 denoted by head arrows. Protein p4 binding region 1 located between promoters A2c and A2b includes two inverted repeats (sites 1 and 2), and region 2, located between promoters A2b and A3, includes also two inverted repeats (sites 3 and 4). Inverted repeats are denoted by arrows. Sets of nucleotides protected by p4 from hydroxyl radical attack, named 1′–12′, are indicated by bars. Guanines whose methylation interfered with p4 binding are denoted in italics.

inverted repeat (Figure 1). This symmetric organization must be related to the function of the protein. Until now, assays on binding and function of protein p4 have been performed with sequences containing four or two of the inverted repeats; hence, the individual contribution of each inverted repeat to p4 function is unknown.

To further explore the relationships between p4 and the DNA, we analysed DNA bending upon p4 binding to region 1. In addition, we defined the minimal sequences required for stable binding of the protein and characterized the relative affinity of the protein for each of the minimal binding sequences. The relevance of these findings to p4 function is discussed.

MATERIALS AND METHODS

Oligonucleotides and plasmids constructs

The oligonucleotides designed to construct p4-binding sites and the oligonucleotides used for amplification of region 1 wild-type and Mut1, Mut2 and Mut3 are depicted in Table 1. Mut1 sequence corresponds to A2c-X (16). Mu1A and Mu1B are the oligonucleotides of the upper and lower strands of site 1 containing the mutation described in Figure 3. Oligonucleotides were purified on polyacrylamide gels using standard procedures (17). One oligonucleotide from each pair was radioactively labelled by treatment with polynucleotide kinase and [γ-^{32}P]ATP for 45 min at 37°C. Unincorporated label was separated from the oligonucleotide by Sephadex-G50 column chromatography. Each fragment was constructed by overnight hybridization of pairs of oligonucleotide mixtures in a relation of 2 cold:1 labelled, and previously heated to 90°C for 2 min.

Construction of pBend2 derivatives and circular permutation gel retardation assays

Region 1 was amplified from φ29 DNA by PCR, using oligonucleotides pBend-1-SalI and pBend-2-SalI. The PCR product was cloned at the SalI site of DNA bending vector pBend2 (18). The recombinant plasmid was prepared by alkaline lyses procedure and sequenced to confirm the size and orientation of the insert. Purified plasmids were digested with BamHI, SspI, NruI, StuI, SmaI, PvuII, EcoRV, DraI, XhoI, SpeI, NheI or BglII restriction enzymes to generate DNA fragments having a circular permutated p4 binding region 1. The digested fragments were purified on a low melting agarose gel using QIAquick Gel extraction kit (Qiagen). Both cohesive and blunt-ended fragments were dephosphorylated with calf intestinal alkaline phosphatase and then labelled with [γ-^{32}P]ATP and T4 polynucleotide kinase. The free label was eliminated by Sephadex-G50 column chromatography. The labelled DNA fragments were mixed with 1 μg of poly(dI–dC) and 1 μg of BSA and incubated with purified p4 in 20 μl of DNA binding buffer (25 mM Tris–HCl, pH 7.5 and 10 mM MgCl$_2$) for 20 min at 4°C. After addition of 2 μl of 30% (v/v) glycerol, samples were loaded onto a non-denaturing 6% polyacrylamide gel. Electrophoresis was run at 25 mA/gel for 4 h at 4°C in a buffer containing 12 mM Tris–acetate, pH 7.5 and 1 mM EDTA. Gels were dried and quantified using a Fuji Bas-IIIs Image analyser.

Table 1. Oligonucleotides designed for the reconstruction of the sequences required in Figure 5

Oligonucleotides designed to construct p4 binding sites
    5HA3A: 5′-gctcagatctAACTTTTTGCAAGACTTTTTTATAAAATGTTGACGTTTTTCGACAAGAACGcgctagctcc-3′
    4HA3A: 5′-gccagatctAACTTTTTGCAAGACTTTTTTATAAAATGTTGACGTTTTTCGGctagctcc-3′
    3HA3A-site3: 5′-gcactacagatctcAACTTTTTGCAAGACTTTTTTATAAAATGTTGAgctagctacatgg-3′
    3HA3A-site4: 5′-gcactacagatctcGACGTTTTTCGACAAGAACGCCCACAAATCCTagctagctacatgg-3′
    2HA3A: 5′-ccatgtcgatacagatctAACTTTTTGCAAGACTTTTTTATAtatcgctagcactatcgc-3′
    3HA2cA-site2: 5′-gcactacagatctcACCTTTTATTAAAACCTTGTAAAACTTTGTCGAgctagctacat-3′
    4HA2cA: 5′-gctcagatctAACTTTGTCGAACTTTTTATAGAAAAGTGTTGAAAATTGTCGctagctcc-3′
    3HA2cA-site1: 5′-gcacagatctTCTCGAACTTTTTATAGAAAAGTGTTGAAAATTGTCagctagctacatgg-3′
    2HA2cA: 5′-ccatgtcgatacagatcTTATAGAAAAGTGTTGAAAATTGTCtatcgctagcactatcgc-3′

Oligonucleotides designed for amplification of region 1
    pBend-1-SalI, 5′CACCCTGGTCGACAATTTTC 3′ and
    pBend-2-SalI, 5′TAAATAGAGAACGTCGACAACAACC 3′

5HA3A and 4HA3A were the oligonucleotides of the upper strand of region 2 containing five and four hydroxyl radical sets of protections, respectively. Oligonucleotides 3HA3A-site3 and 3HA3A-site4 correspond to the upper strands of region 2 containing the three hydroxyl radical sets of protections included in sites 3 and 4, respectively. 2HA3A was the oligonucleotide for the upper strand of region 2 containing only two of the three hydroxyl radical sets of protections corresponding to site 3. 3HA2cA-site2 and 4HA2cA were the oligonucleotides for the upper strands of region 1 containing the three and four hydroxyl radical sets of protections including site 2. 3HA2cA-site1 and 2HA2cA were the oligonucleotides for the upper strands of region 1 containing the three or only two hydroxyl radical sets of protections corresponding to site 1, respectively. Each oligonucleotide was hybridized with its complementary counterpart to generate the double-stranded sequence used in the p4 binding assays.

Calculation of the p4-induced DNA bend angles and bend centre

The relative mobility (R_F) of the p4–DNA complex was defined as follows:

$$R_F = \text{Cm complex/Cm free DNA}$$

where Cm represents the migration, in cm, of fragments in the gel. The lengths of the different probes varied from 187 to 191 bp due to the sequence of the pBend polylinker. To adjust for the slight differences in the probe length, the relative mobility of the p4–DNA complexes was normalized against the highest mobility probe. The average relative mobility (three experiments for each mutant and eight for the wild-type sequences) of different experiments was plotted. These data were plotted measuring the relative mobility of the free DNA versus p4–DNA complex at each sample. The magnitude of the protein-induced DNA bending was calculated by the equation described by Thompson and Landy (19). The mobility of a rigid DNA fragment is related to its end-to-end distance and equal to:

$$L \cdot cos\alpha/2$$

where L is the length of the unbent DNA and α the angle by which the DNA is bent from linearity. The end-to-end distance of a fragment with the bend at an end will be virtually the same as L. Thus,

$$\mu_M/\mu_E = (L \cdot \cos \alpha/2)/L = \cos \alpha/2$$

$$\mu_M/\mu_E = \cos[k(\alpha/2)]$$

where μ_M is the mobility of the complex with protein bound centrally and μ_E is the mobility of the complex with the protein bound at the end of the DNA, k is a coefficient that depends on the electrophoresis conditions and α is the angle of the protein-induced DNA bend.

Gel retardation assays

Binding of p4 to DNA was performed in reactions (20 μl) containing the corresponding 5′-end-labelled fragment of ϕ29 DNA, 25 mM Tris–HCl, pH 7.5, 10 mM MgCl$_2$, 100 mM KCl, 1 μg of poly(dI–dC) and 10 μg of BSA. Protein p4 was added in the amount indicated in each experiment and incubated for 20 min at 4°C. After addition of 2 μl of 30% (v/v) glycerol, samples were loaded onto a non-denaturing 6% polyacrylamide gel. Electrophoresis was run for 6 h at 4°C at 25 mA/gel. Gels were dried and quantified using a Fuji Bas-IIIs Image analyser.

RESULTS

The protein p4 recognition region 1 is not statically curved

Protein p4 binds to two regions of the ϕ29 genome: region 1 located between promoters A2c and A2b and region 2 between promoters A2b and A3 (Figure 1). Each region includes two 31 bp inverted repeats with the consensus sequence (A/G)AC(T/A)$_3$-17bp-(A/T)$_4$GNT (see Figure 1). The region 1 inverted repeats are designated as sites 1 and 2, and the region 2 inverted repeats are named as sites 3 and 4 (14). Previous studies have shown a static DNA bend in region 2 and that this curvature is enhanced upon p4 binding (20); however, the structure of region 1 with or without p4 bound to it was unknown. To study if a bendable sequence is a requirement for protein p4 binding, the presence of a sequence-directed curvature was analysed in region 1. Intrinsically curved DNA has a preferred conformation and thus is more rigid than a random DNA molecule, a property that causes this DNA to migrate more slowly than non-bent DNA during low-temperature PAGE (21–23). The conformation of a 200 bp DNA fragment containing p4 binding region 1 (R1) was studied by comparing its electrophoretic mobility with that of two other DNA fragments of known topology, the 195 bp intrinsically curved fragment containing p4 binding region 2 (R2) and a 198 bp control fragment (C) lacking any detectable curvature (15). The mobilities of the fragments were compared by PAGE at two temperatures, 4 and 65°C. When the electrophoresis was performed at 65°C, all fragments had mobilities expected for their sizes (Figure 2,

lanes a–c). When the assay was carried out at 4°C, the migration of the fragment containing region 1 and the control straight sequence (lanes d and e) had the same migration as they did at 65°C; in contrast, the region 2-containing fragment exhibited an anomalous slow migration (lane f) in agreement with its intrinsic curvature of 45° (20). Therefore, region 1 most probably contains a straight sequence. Since p4 binds both regions 1 and 2, this result indicates that a sequence-directed curvature is not an essential requirement for p4 binding.

Protein p4 binding induces a bend in region 1

Protein p4 generates a strong bend upon binding to region 2, which was interpreted as reflecting an increase in the

Figure 2. Temperature dependency of the electrophoretical mobility of φ29 DNA fragments containing region 1 or 2. Region 1 (R1) was a 200 bp PCR fragment with the sequence from position +98 of promoter A2c to position −6 of promoter A2b start sites. Region 2 (R2) was a 195 bp PCR fragment with the sequence from position +5 of promoter A2c to position −69 of promoter A2b start sites. Control (C) was a 198 bp fragment containing from positions +102 to −96 of early promoter C2 start site (12). Lanes a–c and d–f, correspond to the native 4% acrylamide gel electrophoresis run at 65 and 4°C, respectively.

sequence-directed intrinsic curvature present in this region. Since region 1 does not seem to be intrinsically curved, we wondered whether p4 were capable of bending the DNA upon binding to a straight sequence and, if so, what the final angles were, when the protein binds to a straight versus a curved sequence. To investigate whether p4 binding to region 1 induces bending, we employed circular permutation assays. Plasmid pBend2wt contains a 72 bp DNA fragment with p4 binding region 1, with sites 1 and 2 present in the order described in Figure 3, cloned at the SalI restriction site. DNA fragments of 191 (blunt ends) or 187 bp (cohesive-ends) were produced by restriction enzyme digestions that cleaved twice in the plasmid, resulting in fragments that have the centre of region 1 (the central base pair between sites 1 and 2) located at different positions with respect to the DNA ends. Free DNA migrated with similar mobility, concurring with the absence of intrinsic curvature as demonstrated above (Figure 4A). In contrast, fragments with a gradient of mobilities were generated upon p4 binding, with the StuI fragment–p4 complex having the slowest mobility followed closely by the NruI-generated fragment complex, while the BglII fragment–p4 complex presents the fastest mobility. From an average of eight experiments, we could define StuI-derived complex as the one with slowest mobility. The variation of the electrophoretical mobility reflects the DNA conformation in the complex, indicating the p4-induced curvature of the region upon binding to the sequence. Extrapolation of the gel mobility data to the maximum allows the identification of the centre of the curvature. If StuI were the lowest mobility complex, the centre of bending will be at bp 25/26 (counting
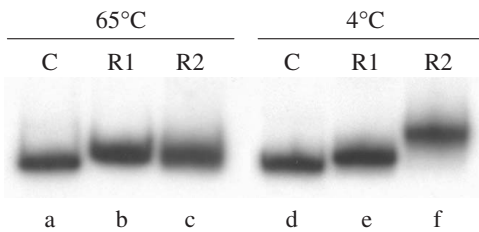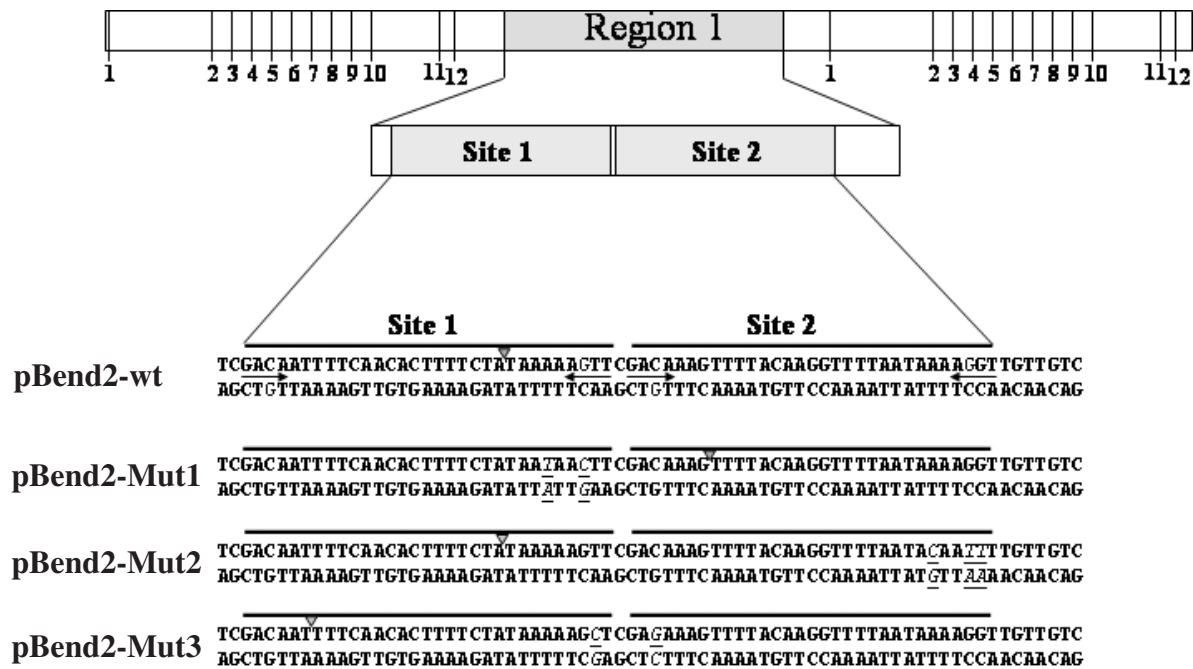

Figure 3. Diagram of plasmid pBend2 with the φ29 sequence of region 1 and the mutated sequences Mut1, Mut2 and Mut3 cloned. Top lane, enzymes organization on the polylinker of pBend2 around the cloned region 1, where the number corresponds to the restriction enzymes targets of the plasmid: BamHI (1), SspI (2), NruI (3), StuI (4), SmaI (5), PvuII (6), EcoRV (7), DraI (8), XhoI (9), SpeI (10), NheI (11) and BglII (12). Scheme of the 72 bp DNA fragments cloned, with the two inverted repeats indicated in boxes as sites 1 and 2. The nucleotide sequence of wild-type (pBend2-wt) region 1 and the modifications of the sequence on Mut1 (pBend2-Mut1), Mut2 (pBend2-Mut2) and Mut3 (pBend2-Mut3) are shown. It should be pointed out that Mut3 insert was opposite in orientation, relative to the other three constructions but here, to make clear its sequence characteristics, has been drawn anti-sense. The inverted repeats at the edges of sites 1 and 2 are denoted by horizontal arrows. Vertical arrow-heads denote the centre of the curvature induced upon p4 binding.
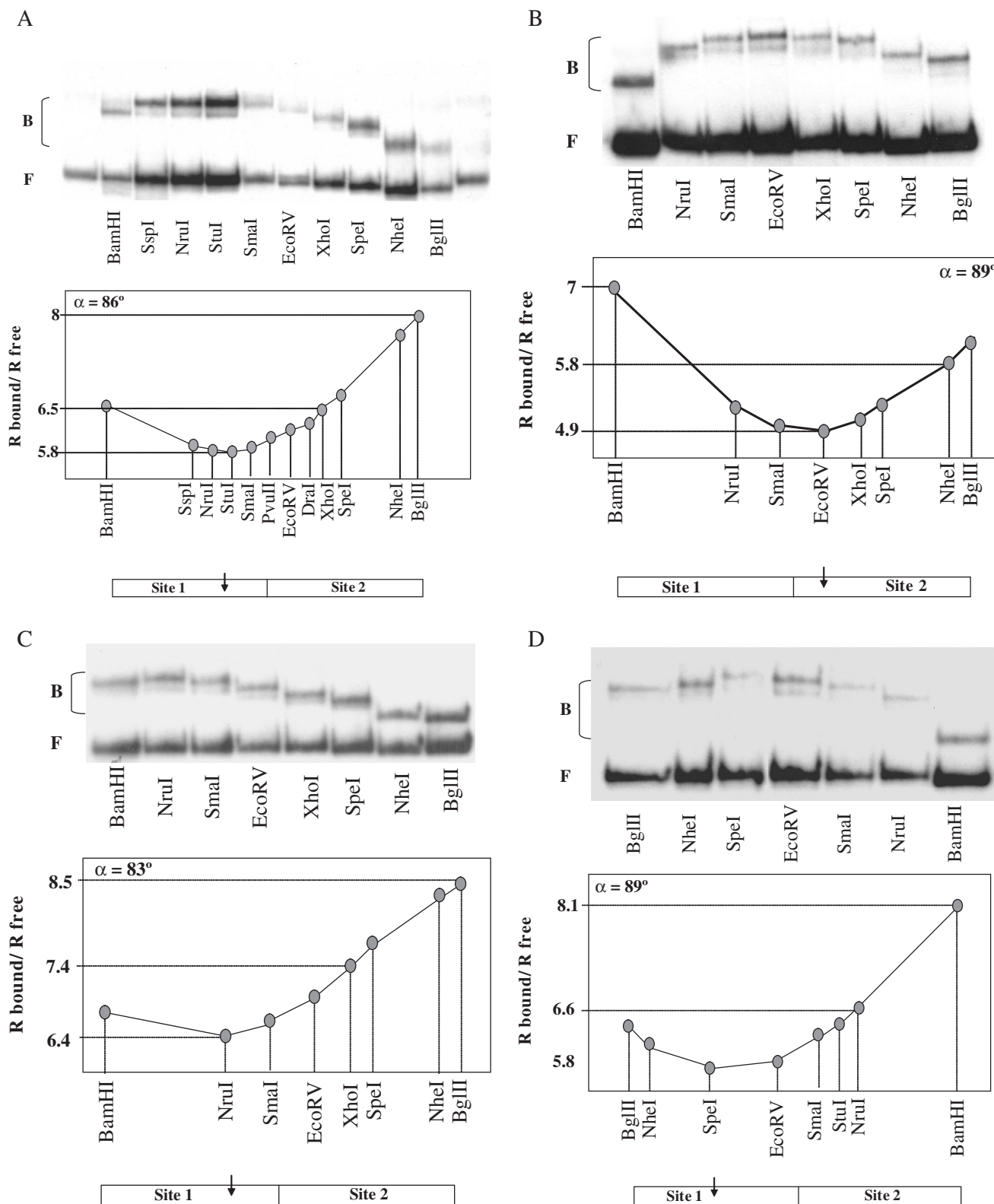
Figure 4. Band-shift assay of the p4–DNA complex of the circularly permuted DNA fragments, which contain p4 binding region 1 wild-type (**A**), Mut1 (**B**), Mut2 (**C**) or Mut3 (**D**). Protein p4 (160 nM) and purified circularly permuted DNA fragments from plasmids pBend2-wt, pBend2-Mut1, pBend2-Mut2 and pBend2-Mut3 were incubated for 15 min at 4°C and loaded directly onto the gel. Bands marked B denoted the p4–DNA complexes and bands labelled F, free DNA. The relative gel mobility of the p4 complexes was plotted against the position of restriction sites in the DNA fragment and expressed as the distance along the 5′–3′ (upper) strand from the left-hand BamHI site except in (D), where due to the orientation of the 72 bp insert, the distance was measured along the 5′–3′ lower strand from the BglII site. For the curve representations, data obtained from several experiments were used. Bending angles (α) are shown. Bend centre is shown as a vertical arrow. Gels show only some of the fragments analysed; other data are plotted in each figure.

from the left end of region 1), while if NruI were the lowest mobility fragment, the centre of bending will be at position 19/20. Based on the slightly lower mobility of the StuI- with respect to the NruI-generated fragments, we approximated bp 24 on site 1 as the apparent centre of curvature and estimate an error limit of ±2 bp (Figure 3). Therefore, the bending is not symmetric with respect to sites 1 and 2, but within site 1 towards the inverted repeat proximal to site 2. This result could reflect that p4-induced bending at sites 1 and 2 are in a different plane. To quantify the apparent p4-induced bending angle, we used the method of Thompson and Landy (19) and the value was 86°.

Protein p4 binding and bending to region 1 with only one intact inverted repeat

There is no evidence for the binding of p4 to the sequence containing only one of the inverted repeats. To investigate the requirements for protein p4 binding and bending, we analysed p4-binding to three region 1 derivatives with mutations at positions of the inverted repeats 1 and 2 involved in p4–DNA interaction (Figure 3). Mut1 has a G to C change on the proximal part (with respect to site 2) of the site 1 inverted repeat and a second change located three base pairs upstream; both positions are involved in p4–DNA interactions as shown by methylation interference (16) and by hydroxyl radical protection experiments (14). Mut2 has two G to T changes on the distal part (with respect to site 1) of the site 2 inverted repeat, and a second change also located three base pairs upstream, both positions being involved in p4–DNA interaction. Mut3 has a G to C change at the proximal part (with respect to site 1) of the site 2 inverted repeat, which is involved in p4–DNA interaction, and an additional base pair change on site 1, which does not have a defined function on p4–DNA binding. P4 binding to the Mut2 and Mut3 sequences was 95 and 90%, respectively, compared with the binding to the wild-type sequence; however, p4 bound 5-fold less efficiently to Mut1 DNA than to the wild-type sequence (data not shown). These results suggest that one intact inverted repeat could be sufficient for p4 binding provided the inverted repeat is the one corresponding to site 1.

To analyse the bend induced by p4 binding to region 1 with only one of the inverted repeats intact, and to determine the effects of the mutations both on the position of the curvature centre and on the final degree of p4-induced curvature, the 72 bp sequences corresponding to Mut1, Mut2 and Mut3 region 1 were cloned in pBend2 (pBend2-Mut1, pBend2-Mut2 and pBend2-Mut3, Figure 3). DNA fragments were produced by restriction enzyme digestion as described above. In the case of Mut1, the BamH1 fragment had the fastest mobility while the EcoRV fragment showed the slowest mobility (Figure 4B). Hence, the mutation at site 1 modified not only the p4 binding affinity but also the mobility of the generated restriction fragments. Using the extrapolation of the gel mobility data led us to identify base pair 41 (counting from the left of the fragment) on site 2 as the bending centre. The apparent bending angle induced by p4 bound to region 1 Mut1 was 89°. As shown in Figure 4C, Mut2 NruI fragment showed the slowest mobility and the BglII fragment showed the fastest mobility; however, further comparative assays of the fragments generated with NruI, and StuI from pBend2-wt and pBend2-Mut2 (data not

shown) suggested that Mut2 has the centre of curvature close to that of the wild-type sequence. The apparent bending angle for this mutant was 83°. Within the pBend2-Mut3-fragments, the p4 complex with the SpeI fragment showed the slowest mobility and the BamHI fragment showed the fastest mobility (Figure 4D). We identified the centre of the bending Mut3 sequence also at site 1, but at base pair 8 (counting from the left of the fragment). The apparent bending angle induced by p4 bound to Mut3 was 89°.

Thus, mutations at the distal part of site 2 (pBend2-Mut2) did not significantly affect the gradient of mobility of the restriction fragments or the location of the curvature centre, while the location of the curvature centre was affected when the proximal part of site 1 (pBend2-Mut1) or the proximal part of site 2 (pBend2-Mut3) were mutated. Therefore, only mutations at the inverted repeats abutting sites 1 and 2 affect the centres of curvature, which led us to speculate that dimers bound to each inverted repeat could interact. On the other hand, since the apparent bending angle induced by p4 bound to wild-type region 1 was 86° and p4-induced bending angles of 89°, 83° and 89° to region 1 of Mut1, Mut2 and Mut3, respectively, were found, both wild-type and mutants sequences exhibit close bending angles (83–89°), values similar to the 85° bend angle found when p4 is bound to region 2 (20). Thus, p4 produced a similar final angle of bending whether or not the sequence has an intrinsic curvature.

Determination of the minimal sequence required at the p4 recognition sites

High-resolution chemical footprinting has shown that protein p4 binds to two regions (regions 1 and 2) of 63 bp expanding from nucleotides −25 to −84 and −27 to −87 relative to promoters A2c and A2b transcription start sites, respectively. By hydroxyl radical footprinting, protein p4 binding results in the protection of six sets of bands at each region separated by 10 bp (14). To define the minimal sequence required for p4 binding, assays were performed with fragments with decreasing numbers of the sequences protected by p4 from hydroxyl radical cleavage. In all fragments, except for the ones containing the sequence corresponding to positions 1′–6′ and 7′–12′, the decrease of φ29 DNA sequence was compensated by adding the same amount of unspecific sequence. Protein p4 and ^{32}P-labelled DNA fragments were incubated and the p4–DNA complexes formed were analysed using band-shift gel electrophoresis (Figure 5). The results obtained with region 1 indicated that: (i) all fragments except the one containing the sequence corresponding to only two of the hydroxyl radical-generated sets of protections were able to support p4 binding; (ii) formation of the p4–DNA complex did not require the whole set of protected sequences (12 sets) simultaneously, but required at least the sequences corresponding to three sets included in one inverted repeat; (iii) region 1 contains two independent p4 binding sites (marked sites 1 and 2; Figure 5, region 1); and (iv) a higher amount of complex was formed when the fragment contained site 1 than with the one containing site 2, suggesting different affinities of p4 for each site. On the other hand, at region 2, p4 bound only when the site 3 inverted repeat was included. However, an efficient p4–DNA complex was obtained with either this inverted repeat alone or with the entire region 2; surprisingly, the p4 binding affinity
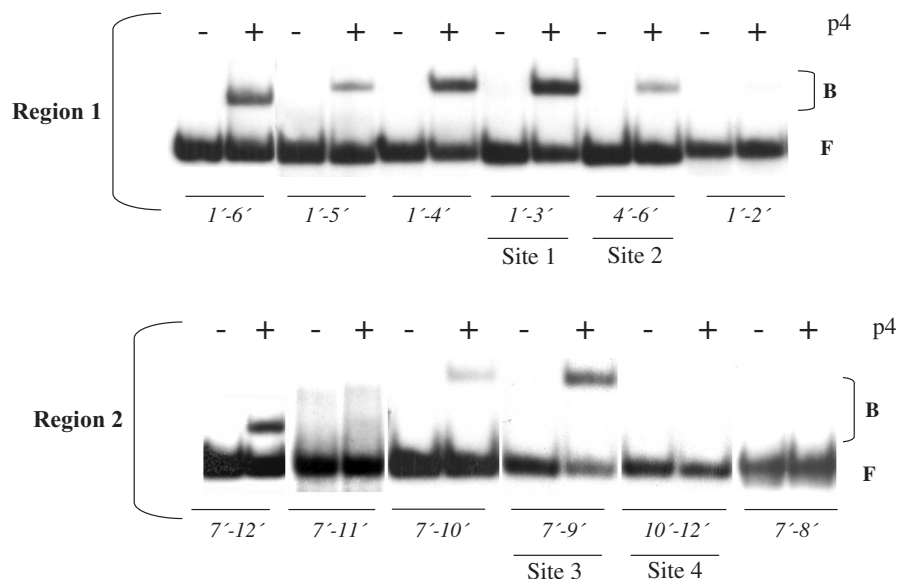
Figure 5. Determination of the minimal sequence required for p4 binding at regions 1 and 2. Fragments containing the sequence corresponding to the entire region 1 ($1'$–$6'$) or diminished progressively by one ($1'$–$5'$) or two ($1'$–$4'$) hydroxyl radical-protected positions, or fragments containing one inverted repeat ($1'$–$3'$ or $4'$–$6'$), or a fragment with the sequence containing only two hydroxyl radical-protected positions ($1'$–$2'$) were assayed in binding of protein p4. Fragments containing the sequence corresponding to the entire region 2 ($7'$–$12'$) or diminished progressively by one ($7'$–$11'$) or two ($7'$–$10'$) hydroxyl radical-protected positions, or fragments containing one inverted repeat ($7'$–$9'$ or $10'$–$12'$), or a fragment with the sequence containing only two hydroxyl radical-protected positions ($7'$–$8'$) were assayed in binding of protein p4.
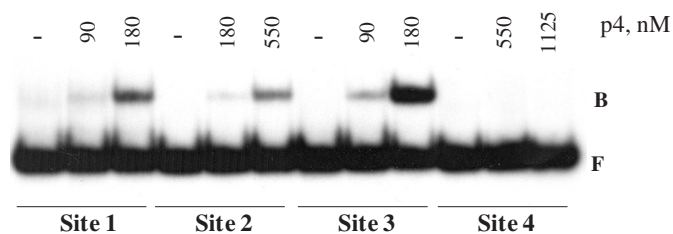

Figure 6. Binding of p4 to each of its binding sites. Increasing amounts of p4 were incubated in 20 µl of binding buffer with a fixed amount of 60 bp fragments (4 nM) containing the 31 bp of the inverted repeats corresponding to site 1, site 2, site 3 or site 4 sequence surrounded by identical but unspecific double-strand DNA. B denotes the protein–DNA complexes and F, free DNA.

for site 3 diminished with the progressive addition of specific ϕ29 sequences (compare $7'$–$9'$ with $7'$–$10'$ and $7'$–$11'$), which might be explained by modifications of the intrinsic curvature through the progressive diminution of the sequence. No binding of p4 to the site 4 inverted repeat was detected; however, p4 bound to site 4 was demonstrated by DNase I and hydroxyl radical footprinting (14,15) using DNA fragments containing both sites 3 and 4.

Based on the results presented, we conclude that one inverted repeat is the minimal sequence required for p4 binding. However, the p4 affinity for each binding site seems different. To quantify the p4 binding affinity for sites 1, 2, 3 and 4, identical amounts of same size fragments containing each site and the same non-specific sequence at both ends were used (Figure 6). Each fragment was incubated with the indicated concentrations of p4, and the p4–DNA complexes were resolved from unbound DNA in band-shift assays. The data indicated that protein p4 binds those fragments containing sites 1, 2 and 3, but not to DNA containing site 4. These results suggest that either the p4–DNA complex at site 4 is not strong enough to resist the electrophoresis force, or the p4 dimer bound to site 4 requires further stabilization through interaction with the dimer at site 3. The amount of protein required for p4–DNA complex formation at sites 1, 2 and 3 and the amount of complex formed, varied depending on the binding site analysed. Protein p4 binds 2-fold more efficiently to site 3 than to site 1, and 5-fold better to site 1 than to site 2.

Protein p4 has higher affinity for site 1 than for site 2, and p4 bound to region 1 bends the DNA with the centre of the curvature located at site 1, indicating that p4 bends the DNA toward the higher affinity binding site. When p4 was bound to the Mut1 sequence, the curvature centre was present at site 2, suggesting that in this case, the latter is the higher affinity site. To determine if the changes present in the Mut1 sequence impeded p4 binding to site 1, we analysed binding of the protein to a fragment containing only site 1 with the Mut1 sequence (Figure 7). The results showed that 8-fold more p4 was needed to form a complex, in comparison to wild-type DNA.

DISCUSSION

By DNase I and hydroxyl radical protection experiments, it was shown that protein p4 binds to two specific DNA recognition regions, each of them located between two of the three promoters regulated by the protein: one between promoters A2c and A2b (region 1) and the other between promoters A2b and A3 (region 2). Protein p4-binding regions differ in the specific base sequence, but contain two tandem elements, consisting of two inverted repeats separated by 1 bp (sites 1 and 2 in region 1, and 3 and 4 in region 2). P4 binds to region 1 and region 2 simultaneously (14) and independently (this paper).
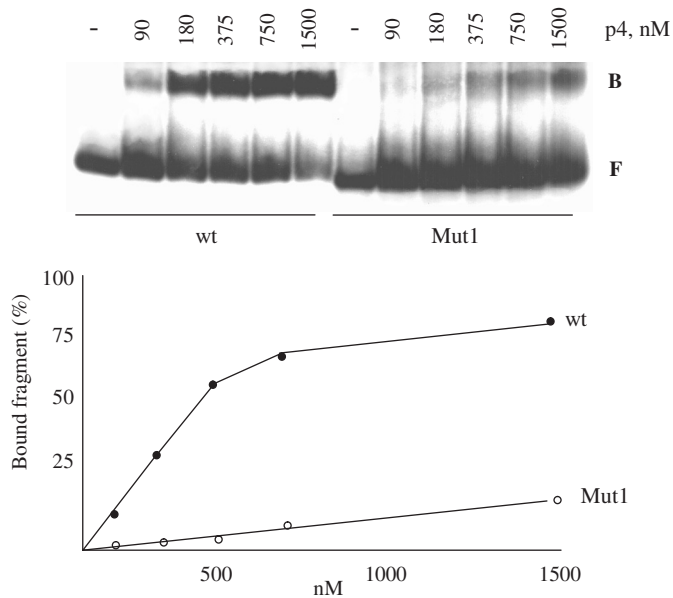
Figure 7. Binding of p4 to the Mut1 sequence. Increasing amounts of p4 were incubated in 20 μl of binding buffer with a fixed amount of fragment containing the 31 bp of the inverted repeat, corresponding to site 1 or Mut1 sequence surrounded by 31 bp of identical but unspecific double-stranded DNA. B denotes the protein–DNA complexes and F, free DNA. Below, representation of the percentage of p4–DNA complexes formed by increasing concentration of p4 with the wild-type (wt) or mutated site 1 (Mut1).

Here, we have shown that each inverted repeat is an independent binding site and defined p4 site as the 31 bp inverted repeat: AAC(T)$_5$-15bp-(A)$_5$GTT.

The following denote the sequence of each site:

site 3: TTGTAAAATATTTTTTCAGAACGTTTTTCAA;
site 1: TTGAAAAATATCTTTTCACAACTTTTAACAG;
site 2: TGGAAAATAATTTTGGAACATTTTGAAACAG; and
site 4: TCCTAAACACCCGCAAGAACAGCTTTTTGCA.

Site 3, with the sequence closest to the consensus inverted repeat is the highest affinity site followed by site 1, which lacks one of the external thymines. Site 2, of lower affinity than site 1, also lacks one of the external thymines but, in addition, has the thymine-track broken by a guanine. Interruption of the thymine-track with guanine resulted in a 3-fold decrease in p4 binding to site 3 (24,25). Site 4, with the most imperfect inverted repeat, is unable to bind p4 by itself. Thus, by measuring p4 binding affinities to each site independently, we conclude that inverted repeats with sequences closest to the consensus define the highest affinity site.

Previous work provided evidence for the existence of a 45° sequence-dependent DNA bend overlapping with region 2, and the enhancement of this curvature to 85° when protein p4 bound to this region (15,20). Here, we demonstrated that p4 binding to the non-curved sequence of region 1 resulted in a final DNA complex with similar bent angle (86°), as described for p4 bound to region 2. The accuracy of bend angle estimation using circular permutation assays has been occasionally questioned, as in the case of several bZip proteins (26). However, circular permutation assays estimate bending angles that are in good agreement with the results obtained from phasing

and cyclization experiments, and from X-ray crystallography for many globular proteins such as CAP and Cro (27,28).

We have shown that differences in the DNA sequence do not play a major role in dictating the magnitude of the p4-induced DNA bending, and that a curvature is induced whenever p4 is able to bind to its cognate site. In contrast, the final induced DNA bending varies with the affinity of the tumour suppressor protein p53 to site-specific response elements (29).

High-resolution chemical footprinting has shown that p4 binds simultaneously to regions 1 and 2. Protein p4 binding to sites 1, 2, 3 and 4 gives rise to a complex where the bound proteins are located on the same face of the DNA helix. The formation of an open loop could be the final result of the p4-induced DNA curvature. This loop could have a double function: (i) favouring progressive binding and stabilization of the p6 transcription regulator, by providing the appropriate structure of the double helix, and (ii) bringing the sequences of promoters A2b and A3, which are 123 bp apart, in close proximity and thereby, favouring the displacement of the RNAP from the high-affinity promoter A2b to the low-affinity promoter A3.

Protein p6 binding displaces p4 from binding site 2 and, concomitantly, stabilizes the binding of p4 dimers to sites 1 and 3 (14), most probably through p4–p6 interactions. High-resolution chemical footprinting protections and hypersensitivities due to p4 bound to sites 1 and 3, are identical in the presence or absence of p6. Since the p4-induced centre of curvature is maintained in the p4–p6 complex (data not shown), the centre of curvatures around position −47 of promoter A2c start site and at about position −60 (20) of promoter A3 start site would modify the topology of these promoters and consequently, RNAP promoter interactions. Binding of RNAP to promoter A2c induces DNase I hypersensitivity sites at position −37 on the non-template strand (14), which agrees with the DNase I hypersensitive sites data of the *Escherichia coli* RNAP transcription complex with 33 different promoters (30), and with the position of the 36° kink of the crystallographic open complex model (31). The p4-induced curvature at position −47 would contribute to destabilization of the promoter A2c transcription complex and consequently to the repression of promoter A2c, while the p4-induced curvature near position −60 of promoter A3 start site would contribute to place p4 and RNAP in close proximity, favouring the interaction between p4 and the CTD domain of the RNAP α subunit (32), required for the correct positioning of the enzyme at promoter A3 (14). In fact, by separation of promoter A3 from the p4 binding sites by the insertion of bent and non-bent DNA sequences of different lengths, it has been shown that p4 is capable of functioning at longer DNA distances from the RNAP binding site only if curved intervening sequences are provided (33). In summary, the results presented, which define the role of the individual p4 binding sites in p4 binding and bending, provide insights on the functional topology of the bacteriophage φ29 genome region, including promoters A2c, A2b and A3.

ACKNOWLEDGEMENTS

REFERENCES

1. Pérez-Martín,J. and de Lorenzo,V. (1997) Clues and consequences of DNA bending in transcription. *Annu. Rev. Microbiol.*, **51**, 593–628.
2. Grosschedl,R. (1995) Higher-order nucleoprotein complexes in transcription: analogies with site-specific recombination. *Curr. Opin. Cell Biol.*, **7**, 362–370.
3. Kahn,J.D. and Crothers,D.M. (1998) Measurement of the DNA bend angle induced by the catabolite activator protein using Monte Carlo simulation of cyclization kinetics. *J. Mol. Biol.*, **276**, 287–309.
4. Lewis,D.E., Lee,H.J., Liu,M. and Adhya,S. (2003) Effect of varying the supercoiling of DNA on transcription and its regulation. *Biochemistry*, **42**, 10718–10725.
5. Kim,S.I., Jourlin-Caselli,C., Wellington,S.R. and Sonenshein,A.L. (2003) Mechanism of repression by *Bacillus subtilis* CcpC, a LysR family regulator. *J. Mol. Biol.*, **334**, 609–624.
6. Ross,E.D., Keating,A.M. and Maher,L.J. (2000) DNA constraints on transcription activation *in vitro*. *J. Mol. Biol.*, **297**, 321–334.
7. Bourgerie,S.J., Michan,C.M., Thomas,M.S., Busby,S.J. and Hyde,E.I. (1997) DNA binding and DNA bending by the MelR transcription activator protein from *Escherichia coli*. *Nucleic Acids Res.*, **25**, 1685–1693.
8. Lim,F.L., Hayes,A., West,A.G., Pic-Taylor,A., Darieva,Z., Morgan,B.A., Oliver,S.G. and Sharrocks,A.D. (2003) Mcm1p-induced DNA bending regulates the formation of ternary transcription factor complex. *Mol. Cell. Biol.*, **23**, 450–461.
9. Scaffidi,P. and Bianchi,M.E. (2001) Spatially precise DNA bending is an essential activity of the Sox2 transcription factor. *J. Biol. Chem.*, **276**, 47296–47302.
10. Serrano,M., Salas,M. and Hermoso,J.M. (1990) A novel nucleoprotein complex at a replication origin. *Science*, **248**, 1012–1016.
11. Elías-Arnanz,M. and Salas,M. (1999) Functional interactions between a phage histone-like protein and a transcriptional factor in regulation of φ29 early-late transcriptional switch. *Genes Dev.*, **13**, 2502–2513.
12. Camacho,A. and Salas,M. (2001) Mechanism for the switch of φ29 DNA early to late transcription by regulatory protein p4 and histone-like protein p6. *EMBO J.*, **20**, 6060–6070.
13. Calles,B., Salas,M. and Rojo,F. (2002) The φ29 transcriptional regulator contacts the nucleoid protein p6 to organize a repression complex. *EMBO J.*, **21**, 6185–6194.
14. Camacho,A. and Salas,M. (2004) Molecular interplay between RNA polymerase and two transcriptional regulators in promoter switch. *J. Mol. Biol.*, **336**, 357–368.
15. Barthelemy,I. and Salas,M. (1989) Characterization of a new prokaryotic transcriptional activator and its DNA recognition site. *J. Mol. Biol.*, **208**, 225–232.
16. Monsalve,M., Calles,B., Mencía,M., Rojo,F. and Salas,M. (1998) Binding of phage φ29 protein p4 to the early A2c promoter: recruitment of a repressor by the RNA polymerase. *J. Mol. Biol.*, **283**, 559–569.
17. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular cloning: A laboratory Manual*. 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 13.78–13.95.
18. Kim,J., Zwieb,C., Wu,C. and Adhya,S. (1989) Bending of DNA by gene-regulatory proteins: construction and use of a DNA bending vector. *Gene*, **85**, 15–23.
19. Thompson,J.F. and Landy,A. (1988) Empirical estimation of the protein-induced DNA bend to lambda site-specific recombination complexes. *Nucleic Acids Res.*, **16**, 9687–9705.
20. Rojo,F., Zaballos,A. and Salas,M. (1990) Bend induced by the phage φ29 transcriptional activator in the viral late promoter is required for activation. *J. Mol. Biol.*, **211**, 713–725.
21. Dieckman,S. (1989) The migration anomaly of DNA fragments in polyacrylamide gels allows the detection of small sequence-specific DNA structure variations. *Electrophoresis*, **10**, 354–359.
22. Kahn,J.D. (1999) Methods for analyzing DNA bending. *Methods Mol. Biol.*, **94**, 109–124.
23. Crothers,D.M. and Drak,J. (1992) Global features of DNA structure by comparative gel electrophoresis. *Methods Enzymol.*, **212**, 46–57.
24. Nuez,B., Rojo,F. and Salas,M. (1991) Identification of the sequences recognized by phage φ29 transcriptional activator: possible interaction between the activator and the RNA polymerase. *Nucleic Acids Res.*, **19**, 2337–2342.
25. Nuez,B., Rojo,F. and Salas,M. (1994) Requirement for an A-tract structure at the binding site of phage φ29 transcriptional activator. *J. Mol. Biol.*, **237**, 175–181.
26. Hagerman,P.J.Jr,II (1996) From sequence to structure to function. *Curr. Opin. Struct. Biol.*, **6**, 277–280.
27. van der Vliet,P.C. and Verrijzer,C.P. (1993) Bending of DNA by transcription factors. *BioEssays*, **15**, 24–32.
28. Starr,D.B., Hoopes,B.C. and Hawley,D.K. (1995) DNA bending is an important component of site-specific recognition by the TATA binding protein. *J. Mol. Biol.*, **250**, 434–446.
29. Nagaich,A.K., Appella,E. and Harrington,R.E. (1997) DNA bending is essential for the site-specific recognition of DNA response elements by DNA binding domain of the tumour suppressor protein p53. *J. Biol. Chem.*, **272**, 14842–14849.
30. Ross,W., Ernst,R. and Gourse,L. (2001) Fine structure of *E.coli* RNA polymerase–promoter interactions: alpha subunit binding to the UP element minor groove. *Genes Dev.*, **15**, 491–506.
31. Murakami,K.S., Masuda,S., Capbell,E.A., Muzzin,O. and Darst,S.T. (2002) Structural basis of transcription initiation: an RNA polymerase holoenzyme–DNA complex. *Science*, **296**, 1285–1290.
32. Mencía,M., Monsalve,M., Rojo,F. and Salas,M. (1996) Transcription activation by phage φ29 protein p4 is mediated by interaction with the α subunit of *Bacillus subtilis* RNA polymerase. *Proc. Natl Acad. Sci. USA*, **93**, 6616–6620.
33. Serrano,M., Barthelemy,I. and Salas,M. (1991) Transcription activation at a distance by phage φ29 protein p4. Effect of bent and non-bent intervening DNA sequences. *J. Mol. Biol.*, **219**, 403–414.