DNA Binding by the Coliphage 186 Repressor Protein CI*

(Received for publication, November 21, 1995, and in revised form, January 25, 1996)

Ian B. Dodd and J. Barry Egan‡

From the Department of Biochemistry, University of Adelaide, Adelaide 5005, Australia

The cl gene of coliphage 186 maintains lysogeny and confers immunity to 186 infection by repressing the major early promoter, pR, and the promoter for the late transcription activator gene, pL. Gel mobility shift and DNase I footprinting show that CI protein binds to the DNA at pR and pL and also to sites –300 base pairs upstream and downstream of pR, called FL and FR. Mutations which cause virulence reduce CI binding to pR. The biochemical and genetic data identify three CI operators at pR, two at pL, and single operators at FL and FR. The operators at the pL, FL, FR, and central pR sites are inverted repeat sequences, separated by 5 base pairs (Type A) or, in the case of pR, by 4 base pairs (Type A'). A different inverted repeat operator sequence (Type B) is proposed for the binding sites on each side of the central site at pR. Thus, CI appears to recognize two distinct DNA sequences. CI binds cooperatively to adjacent operators, and binding at pR is strongly dependent on these cooperative interactions. A high order CI multimer appears to be the active DNA binding species, even at single operators.

Coliphage 186 shares with the well characterized phage λ the ability to achieve a lysogenic state that is extremely stable yet which can efficiently switch to the lytic state in response to activation of the host SOS system (1). Elucidation of the gene control mechanisms and strategies used in λ’s genetic switch (2, 3) has been of profound value in molecular biology and has informed thinking about the ways in which higher organisms utilize alternative stable developmental states. 186 is a member of the P2 phage family (4) and, since it shows very little similarity at the DNA or protein sequence level to λ, appears to represent a relatively independent solution to the requirements of such a genetic switch.

The 186 cl gene is the central player in the maintenance of the stable lysogenic state (5). The cl gene product represses two promoters: pR, the promoter for the early lyric operon, including the apl, cl, and replicase genes, and pL, the promoter for the late promoter activator gene, B (Refs. 6 and 7; see Fig. 1). CI, thus, directly or indirectly, represses all the lymphocytic genes of the prophage and also blocks lytic development of 186 phage that infect the lysogen (8). Expression of cl is maintained during lysogeny by transcription from the cl promoter. The leftward pR promoter and the rightward pL promoter are arranged face-to-face, with their transcripts overlapping by 62 bases (6). This is quite different from the analogous pR and pL promoters of λ, which are arranged back-to-back.

The processes of establishment, stable maintenance, and efficient breakdown of the lysogenic transcriptional state in 186 display some unusual features, and characterization of the activity of the CI protein is needed to understand these processes. During lysogeny, the face-to-face promoter arrangement would seem to create difficulties for CI in maintaining repression of pR. If the CI protein binds at pR, then RNA polymerase from pL must pass through the CI-pR complex in order to transcribe the cl gene. This passage of RNA polymerase seems likely to remove CI from the DNA and thus make pR accessible for RNA polymerase binding. Whether this situation is a problem and, if so, what special mechanisms are used to cope with it, are questions that are relevant to the important topic of how mobile protein-DNA complexes, such as polymerases, interact with each other (9) and with static protein-DNA complexes, such as repressors and nucleosomes (10). The strategy for establishment of lysogeny in 186 appears very similar to λ, with lysogenic transcription requiring CI and the initial production of CI being dependent on another phage protein, CI1 (5). However, the activation of lysogenic transcription by 186 CI seems to be indirect, with CI repression of pR, removing an inhibition of pL caused by converging transcription from pL (6). Efficient breakdown of CI repression during SOS induction of the prophage is initiated not by RecA, as in λ, but by a phage protein, Tum, that antagonizes CI repression of pR and pL (11). Efficient derepression may also require repression of CI transcription from pL by the Api protein, which binds between pR and pL (11). Interactions between Api and CI at pR are likely to be critical in the operation of the lysis-lysogeny switch.

To investigate CI repression, Lamont et al. (8) isolated and examined a number of virulent (vir) mutants of 186. These are phage mutants that are insensitive to lysogenic immunity and are able to develop lytically in a 186 lysogen. It was expected that these mutants would carry mutations at pR, which interfered with CI repression. Indeed, in all 19 vir mutants, mutations were found within the –49 to –3 region of pR. These mutations are clustered into three sites. All of the vir mutants carry at least one mutation in the central site (Site I), and most carry additional changes in the leftmost site (Site II). Two mutants have changes in the rightmost site (Site III). The three mutants with changes at Site I only are poorly virulent, forming plaques with low efficiency (10–15%) on lysogens, and not forming plaques on a strain carrying the cl gene on a multicopy plasmid. Mutants carrying additional changes at Sites II or III plate with higher efficiency on lysogens (22–52%), and most are able to form plaques on the strain with the cl plasmid. It was expected that these mutations disrupt CI operators and thus act by reducing CI binding to pR. However, no DNA sequence element that was conserved between the three sites and which could serve as a

1 P. Neufing, K. E. Shearwin, J. Camerotto, and J. B. Egan, manuscript in preparation.
2 M. R. Reed, K. E. Shearwin, and J. B. Egan, manuscript in preparation.
3 A. M. Brumby, K. E. Shearwin, and J. B. Egan, manuscript in preparation.

* This work was supported by a grant from the Australian Research Council (to J. B. E.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed. Tel.: 61-8-303-5361; Fax: 61-8-303-4348; E-mail: jegan@biochem.adelaide.edu.au.
likely CI operator could be found (8). In this number scheme, the pD +1 site is at position 270, and the pD +1 site is at 2747. Plasmids constructed as follows.

**Plasmids**

Plasmids pUC18, pUC19 (13), and the pBluescript plasmids (Stratagene) were used for general cloning. pEC456 (7) is pUC19 carrying the 2.1-kb 186 HindIII (61%)–XhoI 629 fragment from 186 BamB7. pEC625 (11) is pBluescript SK (−) carrying the 186 Mael (2666–MaelI 2688) pD fragment, with the MaelI end nearest the USP (universal sequencing primer: M13–20 primer) priming site. Numbering of 186 restriction sites is from the first base of the PstI site at 65.5% (14). In this number scheme, the pD +1 site is at position 270, and the pD +1 site is at 2747. Plasmids constructed as follows.

**Plasmid pEC629**—pBluescript KS(−) with the 186 BamB7 HindII-93–331 pD fragment from pEC456 cloned into the CiaI site with the HindII 93 end nearest the USP priming site.

**Plasmid pEC627**—pBluescript KS(−) with the 186 TaqI 2267–2451 FL fragment cloned into the CiaI site with the TaqI 2267 end nearest the USP priming site.

**Plasmid pEC631**—This plasmid contains the XmnI 2688–MunI 2756 186 fragment (minimal pD fragment) inserted into the EcoRV and EcoRI sites of pBluescript SK(−) with the XmnI end nearest the RSP (reverse sequencing primer) priming site. It was constructed by deletion of the EcoRV (polylinker)–XmnI 2688 portion of pEC630. pEC630 carries the EcoRI (polylinker)–MunI 2756 fragment from pEC625, inserted into the EcoRI site of pBluescript SK(−).

**Plasmid pj C521**—pBluescript SK(−) with the 186 BspMI 2893 (end-filled)–SnaBI 3238 FR fragment cloned into the SmaI site, with the SnaBI end nearest the USP priming site (constructed by J. Camerotto).

**General DNA Manipulations**

Plasmid preparations were by the alkaline lysis method, with a single CsCl gradient purification (15). Restriction digests were performed as specified by the suppliers (New England Biolabs, Pharmacia Biotech Inc.). HeLa cells (Boehringer Mannheim). The Klenow fragment of Escherichia coli DNA polymerase I (Bresatec, Australia) was used for endfilling. DNA fragments for ligation were isolated from agarose using the Genecon procedure (Bio 101, La Jolla, CA). T4 DNA ligase, T4 poly nucleotide kinase, and radiolabeled DNA from Bresatec. Bovine pancreatic DNase I was from Boehringer Mannheim.

**Radiolabeling of Oligonucleotides**

Oligonucleotides were end-labeled in 10-μl reactions containing 50 ng of oligonucleotide, 1.25 μM [γ-32P]ATP (5 μCi μl−1), 10 mM MgCl2, 70 mM Tris-HCl, pH 7.5, 5 mM dithiothreitol, and 10 units of T4 poly nucleotide kinase with incubation at 37 °C for 30–60 min, followed by 70 °C for 20 min to inactivate the enzyme. These reactions were used directly in PCRs and DNA sequencing reactions.

**Polymerase Chain Reactions**

The polymerase chain reaction (PCR) was used to prepare radiolabeled DNA fragments for gel retardation and DNA sequencing footprinting. Each reaction (10–30 μl) was in PCR reaction buffer (2.5 mM MgCl2, 67 mM Tris-HCl, pH 8.8, 16 mM (NH)4SO4, 0.05% Triton X-100, 0.2 mg ml−1 gelatin) with 0.05 unit μl−1 Taq polymerase and 1–2 ng μl−1 of each primer, dNTPs were present at 0.2 mM. Plasmid templates were at 0.05 ng μl−1, phage templates were at approximately 109 plaque-forming units μl−1. Thermal cycling was: 2 min 94 °C, 30 × (5 s 94 °C, 5 s 50 °C, 45 s 74 °C). DNA fragments from PCR were gel-purified before use.

The following oligonucleotide primers (Bresatec) were used. Any 186 sequence is underlined, with the first and last 186 sequence positions given.

USP (universal sequencing primer: M13 –20 primer), GTAAAAGGAGGCGACGTTGC; RSP (M13 reverse sequencing primer), CACACAGAAAAACGTATGACCATG; Primer 34 (2666–2689), ACATTACGCTTGGCTCATCTACAAAGATC; Primer 68 (2907–2894), CCCCGCAATCGTGAATCCAAAGACACCCCTA; Primer 71 (2653–2670), CGGACC-TGGTACGGTCAAGTCATCCAGCTT; Primer 89 (2907–3070), TC-CCCGCGTACGAGACGCCGTACGAGTTG; Primer 87 (tum left), CGTAGTGAGGCTCATATGGATAGAGACT; Primer 9 (tum right), CACCTCCTGAAGGATGC.

**CI Preparations**

The preparation and purification of CI is described in full in the accompanying paper (16). Briefly, sonicated lysates of 186 overexpressing cells were polyethyleneimine-precipitated, resuspended, reprecipitated with ammonium sulfate, fractionated on an Affi-Gel Blue column, fractionated on a heparin column, and dialyzed against 50 mM Tris-HCl, 0.1 mM EDTA, 10% glycerol, 150 mM NaCl, pH 7.5 (TEG150). The fraction from the Affi-Gel Blue column (in 50 mM Tris-HCl, pH 8, 250 mM NaCl, 1 mM EDTA, 10% glycerol) was judged to be ≥95% pure CI and was used for the DNase I footprinting experiments of Fig. 3. The heparin column purified CI, judged to be ≥98% pure, was used in all the other experiments.

**Gel Mobility Shift Assays**

Radiolabeled DNA fragments for gel retardations were prepared by PCRs. CI binding reactions (10 μl) were in TEG150 with 0.1–0.3 ng μl−1 radiolabeled fragment (0.33–1 nM) and 5 ng μl−1 competitor DNA (sheared salmon sperm DNA). The reactions were incubated on ice for 15 min and were loaded onto 6% nondenaturing polyacrylamide gels (19:1 acrylamide:bisacrylamide) containing 20% glycerol, with electrophoresis at 4 °C in TBE. Gels were vacuum-dried and analyzed by phosphor autoradiography (Molecular Dynamics). Computer images of gels were contrast/brightness-adjusted using Adobe Photoshop.

**RESULTS**

CI Binding to Four DNA Regions—The CI gene was cloned into a protein expression plasmid, and crude cell extracts containing high levels of CI protein were obtained. Gel mobility shift assays were carried out using this extract, and a control extract was made from the same strain carrying the parent expression vector (data not shown) (29). Using various DNA fragments from the PstI-1.BglII 4244 early control region of the 186 chromosome (which contains pD, the lysogenic operon, pR, and the first four genes of the early lytic operon), four DNA regions that showed CI-specific binding were located (Fig. 1): a region spanning pD (termed the pR site), a region spanning pR (termed pB), a region upstream of pD (termed FL, for far left region), and a region downstream of pR (termed FR, for far right region).

Fig. 2 shows gel shift assays with CI purified to at least 98%
homogeneity (16) and with DNA fragments containing each of the four binding regions. In each case, there was a single major retarded species at each concentration tested. Thus, only one major complex appeared to be formed at each concentration. The mobility of the retarded species was similar for the different fragments, implying binding of a similar number of protein subunits in each case. A minor, less retarded species was seen in some experiments (see the FR fragment in Fig. 2). The lack of binding to the control DNA fragment in Fig. 2 showed that CI binding to the pB, FL, pR, and FR DNA fragments was sequence-specific.

The gel retardations showed an unusual effect that we are not able to explain: the mobility of the retarded species decreased in small steps with increasing CI concentration (Fig. 2). This effect was seen with all CI binding DNA fragments and occurred over increments of CI concentration even smaller than shown in Fig. 2. The phenomenon was not affected by the order in which samples were prepared or loaded. It is unlikely that the effect is caused by a nonspecific DNA-binding contaminant in the CI preparation because the binding is both sequence-specific, as shown with the control DNA fragment (Fig. 2), and CI-specific, as crude cell extracts not containing CI show no such binding to these fragments (data not shown) (29). Presumably, the decreasing mobility reflects increasing numbers of CI subunits per complex. However, if this were the case, one would expect both a greater magnitude of shift and the appearance of multiple bands in at least some tracks instead of the single band seen.

DNase I Footprinting of CI Binding Sites—To locate the CI binding sites more precisely, DNase I protection studies were carried out. CI purified to at least 95% homogeneity (16) was used for these footprint experiments; similar results were originally obtained using crude cell lysates containing CI (29). The CI footprint results for one strand of the pR binding region is shown in Fig. 3. Fig. 4 summarizes the footprinting data for both strands at all four CI binding sites, showing the assignments of DNase I cleavages that were protected, exposed, or enhanced in the presence of CI. The lengths of the arrows in Fig. 4 indicate the strength of the CI effect, as judged by the lowest CI concentration at which the effect was observable.

At each site there was a particularly dense region of relatively strong protections (indicated by the brackets in Figs. 3 and 4). Within this region there were a few positions that were still sensitive or showed an enhanced sensitivity to DNase I attack in the presence of CI.

Extending beyond these dense footprint regions, usually on one side only, was a less dense region of enhancements and decreases in small stepwise increments of CI concentration (Fig. 2).

This effect was seen with all CI binding DNA fragments and occurred over increments of CI concentration even smaller than shown in Fig. 2. The phenomenon was not affected by the order in which samples were prepared or loaded. It is unlikely that the effect is caused by a nonspecific DNA-binding contaminant in the CI preparation because the binding is both sequence-specific, as shown with the control DNA fragment (Fig. 2), and CI-specific, as crude cell extracts not containing CI show no such binding to these fragments (data not shown) (29). Presumably, the decreasing mobility reflects increasing numbers of CI subunits per complex. However, if this were the case, one would expect both a greater magnitude of shift and the appearance of multiple bands in at least some tracks instead of the single band seen.

DNase I Footprinting of CI Binding Sites—To locate the CI binding sites more precisely, DNase I protection studies were carried out. CI purified to at least 95% homogeneity (16) was used for these footprint experiments; similar results were originally obtained using crude cell lysates containing CI (29). The CI footprint results for one strand of the pR binding region is shown in Fig. 3. Fig. 4 summarizes the footprinting data for both strands at all four CI binding sites, showing the assignments of DNase I cleavages that were protected, exposed, or enhanced in the presence of CI. The lengths of the arrows in Fig. 4 indicate the strength of the CI effect, as judged by the lowest CI concentration at which the effect was observable.

At each site there was a particularly dense region of relatively strong protections (indicated by the brackets in Figs. 3 and 4). Within this region there were a few positions that were still sensitive or showed an enhanced sensitivity to DNase I attack in the presence of CI.

Extending beyond these dense footprint regions, usually on one side only, was a less dense region of enhancements and decreases in small stepwise increments of CI concentration (Fig. 2).
weaker protections that tended to be interspersed with posi-
tions where DNase I cleavage was not affected by CI. Most of
the effects in this region were apparent only at higher CI
concentrations. This extended portion of the footprint was often
quite large; in the case of the pR site, this portion of the
footprint was almost 60 bp long and reached the −14 position
of pL. The less dense region effects were not intensified in
footprints with crude CI extracts (29). The accessible and pro-

**Fig. 4.** Summary of DNase I footprinting with CI. The effects of CI on DNase I cleavage at the pR (a), pB (b), FL (c), and FR sites (d) were judged from various footprint experiments, including that shown in Fig. 3. Protected bonds are indicated by vertical lines, bonds whose cleavage was enhanced by CI are indicated by arrows. The lengths of these lines or arrows represent the sensitivity of the effect, that is, the CI concentration at which the effect was apparent (the long, intermediate, and short lines correspond to the effects being apparent at 130 nM, 260 nM, and 520 nM, respectively). The small open circles denote bonds that remained DNase I-accessible in the presence of CI. In some cases, bonds were protected and remained somewhat accessible; these bonds are indicated by a line capped by a small open circle. The effects were judged by visual inspection of phosphorimages at high magnification; these effects are not always apparent from the printed images of Fig. 3. No effects were seen outside the regions shown. The brackets above and below each sequence indicate the dense portion of the footprint believed to contain the major binding determinants, with the shaded arrows between the strands showing the putative CI operator sequences (see text). The FL, pR, and FR sequences are numbered from the pR start site (186 position 2747; Ref. 6). The pB sequence is numbered from the pB start site (186 position 270; Ref. 28). The −35, −10, and +1 promoter sequences (including the −10 sequence of pL) are shown in bold, as is the stop codon of the apl gene.
DNA Binding by the Coliphage 186 Repressor Protein CI

Fig. 5. Proposed CI operators at the pB, FL, and FR sites. The DNA half-sites of the inverted-repeat sequences proposed as CI operators at the pB, FL, and FR sites (see Fig. 4) are aligned. The diamond shows the center of symmetry of the sequences. The small triangles indicate bonds which remain DNase-accessible (unfilled) or whose DNase cleavage is enhanced (filled) in the presence of CI (Fig. 4); upward pointing triangles indicate the bond on the complementary strand. Conserved bases and their frequencies are indicated below the sequences. To quantitate the discriminative power of the putative CI binding sequence (see text), a full-length consensus sequence MTTC-WCWWWWWGWGAAK (M = A or C, K = T or G, W = A or T) was derived from the alignment. We scored the degree of match of each of the four putative binding sequences with the consensus sequence by assigning a match at a nonredundant consensus position (A, C, G, or T) a score of 1 and a match at a redundant position (W, M, or K) a score of 0.5. With this scheme, each binding sequence scores at least 12 points out of a maximum of 12.5. A computer program was written to scan random DNA sequence for matches to consensus sequences; this showed that a score of 12 or more on this consensus occurred by chance less than once per 1000 kb.

The location of the eight half-sites relative to the footprints is indicated in Fig. 4. The position of the sequences at pB, FL, and FR is consistent with the DNase I footprints in three ways, further strengthening their assignment as recognition sequences for CI. (i) The sequences lie totally within the dense portion of the footprints. (ii) The two operators at pB are positioned symmetrically on either side of the strong enhancements in the center of the pB footprint and are on the same face of the DNA helix as each other (having a center-to-center distance of 31 bp). The DNase-sensitive minor groove between these sites lies on the opposite face of the helix. Thus, the hypersensitivity of the DNA in the center of the pB footprint may be explained by DNA distortion, perhaps bending, induced by interactions between CI molecules bound to adjacent sites. (iii) There is a tendency for certain bonds in approximately equivalent positions in the half-sites of the sequences to remain exposed to DNase or to show enhanced cleavage in the presence of CI (Fig. 4: indicated by triangles in Fig. 5), suggesting that these sequences interact similarly with CI. The relationship between these cleavages on the two strands indicates sensitivity of the minor groove at a single region in each operator arm.

The Binding Determinants at pR. Lie within the Dense Footprint Region—With such strong evidence supporting the proposed operator sequence at pB, FL, and FR, we were surprised to find no significant matches to this sequence at the pR binding site.

In order to further examine CI binding at pR, we first narrowed down the location of the binding determinants, since the extended CI footprint at the pR site was very large, extending from positions −78 to +76. The dense portion of the footprint was considerably smaller, from −53 to +14, and contains the loci of the vir mutants. To test the idea that all of the CI binding determinants at pR are contained within the dense footprint region, we examined CI binding to a pR DNA fragment from which most of the 186 sequences around the dense footprint region were removed.

Using the gel shift assay, we compared CI binding to a DNA fragment containing the −58 to +14 sequence of pR, which carries little more than the dense footprint region (the minimal pR fragment), with binding to a DNA fragment containing the −81 to +126 sequence, which covers the entire CI footprint at the pR site (large pR fragment). One result is shown in Fig. 6, top. The amounts of bound and unbound DNA in the gel were quantitated and graphed in Fig. 6, bottom. This shows that the affinity of CI for the pR site was not affected by the replacement of sequences from the extended footprint region with non-186 DNA. The apparent dissociation constant for CI to pR, Kobs, was obtained as the CI concentration at which half the DNA was bound. This was 30 nM for the large pR fragment and 31 nM for the minimal pR fragment. Ratios close to one for the relative binding strengths of the two fragments were found in three experiments (average Kobs/minimal/ Kobs/large = 0.97 ± 0.05).

Putative CI Recognition Sequences at pR—We therefore re-
The sequence conservations between Site I and the pB-FL-FR recognition sequence Type A and the Site I recognition sequence Type A’ to indicate its different half-site spacing. All of the 19 vir mutants carry a change at Site I, and there are in total 29 base changes at this site (8). This large body of mutational data provides a stringent test of the proposed recognition sequence. The very location of the A’-type sequence at Site I supports its role in CI recognition. However, the fit between this sequence and the vir mutations is much more extensive than co-location and not only provides very strong evidence that it is a CI recognition sequence but also indicates those bases that are critical for binding. Firstly, the mutations lie in both arms of the sequence (Fig. 7), showing the importance of both half-sites, a feature expected for a symmetrical binding sequence. Secondly, in all of the 19 vir mutants, the match to the consensus sequence is worsened. Of the 29 mutations occurring at Site I, 27 involve one of the three fully conserved positions in the half-sites. Two mutations lie at less conserved positions, one slightly improving the match to the consensus. However, both these mutations occur in combination with changes at the fully conserved positions. There is therefore a remarkable agreement between the vir mutations and the consensus sequence. The mutations indicate that the two fully conserved C residues within the half-sites are critical for CI binding, since every one of the 19 vir mutants carries a change at one of these positions.

The Type A recognition sequence cannot explain CI binding in the left (Site II) and right (Site III) regions of pR as we were unable to find such sequences in these regions, even when other alternative half-site spacings were tested. We had begun with the assumption that CI is able to recognize only one type of sequence. However, some DNA-binding proteins, for example, the integrase proteins of λ and other phages, have two
of CI-dependent DNase I enhancements that is very well conserved between all four half-sites (Fig. 4; marked by filled triangles in Fig. 8). These enhancements indicate minor groove sensitivity at the same position in each half-site and argue that CI interacts in a similar fashion with the four sequences.

Two further observations support our three-site model: (B-A'-B) at pR. Firstly, the location of the three sequences correlates very closely to the dense footprint region (see Fig. 4).

Secondly, the three sequences all lie on the same face of the DNA helix. The center-to-center spacing is 21.5 bp between the Site II and Site I sequences (B-A') and 20.5 bp between the Site I and Site III sequences (A'-B). Binding to the same face of the DNA helix is a feature of binding at the pB site and is also consistent with the interactions seen between CI bound at pR, as described below.

Binding Studies with pR from vir Mutants—We examined CI binding to pR fragments carrying mutations at the three vir sites, to firstly, confirm our assumption that the vir mutations disrupt CI binding at pR and, secondly, to investigate independent binding of CI to its individual operators at pR. The three vir pR fragments used in these experiments represent three classes of vir mutants. All fragments contain the same double mutation at Site I (see Fig. 9C). For the vir122 fragment (referred to as Site II 'A-'III'), this is the only change from wild-type. The vir97 fragment carries an additional change at Site II (Site II 'A-III'), and the vir121 fragment carries an additional change at Site III (Site II 'A-III').

Gel shift studies with these fragments showed that all three mutant fragments bound CI more weakly than wild-type (Fig. 9A). Furthermore, each of the three mutations weakened CI binding. From the Kobs values (Fig. 9A, legend), mutation at Site I reduced binding 2.8-fold and further mutation at Site II or Site III reduced binding an additional 2.9- and 2.3-fold, respectively. These effects on CI binding correlate with the degree of virulence shown by the mutants, weak for vir122 and strong for vir97 and vir121 (8), and correlate with the number of intact CI operators in these fragments. Although we have tested only three vir mutants, it is now reasonable to assume that CI binding is weakened in all of the vir mutants.

We noted previously that CI binding to pR, pB, FL, and FR DNA fragments produced, at each CI concentration, a single retarded species of similar mobility (Fig. 2), despite the fact each fragment contains a different number of CI operators. The identical gel shift patterns seen with pR fragments carrying one, two, or three intact operators confirms this result. Thus, at any one CI concentration, it appears that a similar number of CI subunits is binding to the DNA whether this DNA carries one, two, or three operators.

Fig. 9B shows the DNase I protection results with the vir mutant fragments at CI concentrations at which there was very little CI binding in the extended footprint region. We found no evidence for independent binding to individual operators at the pR site. Instead, each mutation affected binding not only to its own site but to the whole of the dense footprint region. With the wild-type fragment, cleavage at almost every position was strongly protected or enhanced by CI at 210 nm (indicated by dots to the right of the gel lanes). Most effects were also visible at 70 nm. The Site I mutation (vir122) strongly reduced binding to the central Site I region but also to the whole dense footprint area, with a subset of the wild-type effects remaining at 210 nm only. Further mutation at Site II or Site III (vir97 and vir121) eliminated all CI binding at these concentrations, despite the fact that there was an intact Type B site on these fragments. Again, the mutation at one site affected binding at the others: mutation at Site II removed the weak remaining binding at Sites I and III; mutation at Site III

Fig. 8. A second putative CI recognition sequence at Sites II and III. The half-site sequences of the operators proposed for CI binding at vir Sites II and III (Fig. 4) are aligned. The diamond shows the center of symmetry of the sequences. The small triangles indicate bases which remain DNase I-accessible (unfilled) or whose DNase cleavage is enhanced (filled) in the presence of CI (Fig. 4); upward pointing triangles indicate the bond on the complementary strand. Conserved bases and their frequencies are indicated below the triangles. The changes found in the left and right half-sites in the vir mutants (8) are indicated, with the subscripts denoting the frequency of each change. The asterisks indicate mutations that do not reduce the match to the consensus. One mutation, lying just left of the Site II sequence, is not shown. The consensus Type B sequence: TNGRYWWWRYCNA (W = A or T, R = A or G, Y = C or T, N = any base) was derived from the alignment. The match of the Site II or III sequences to this consensus was found to arise by chance once per 67 kb.

A second putative CI recognition sequence at Sites II and III is apparent in the DNase footprint. There is a pattern of DNA Binding by the Coliphage 186 Repressor Protein CI
DNA Binding by the Coliphage 186 Repressor Protein CI

...removed the weak remaining binding at Sites I and II.

These results indicate a high degree of cooperativity in CI binding at pR, that is, the favorable interactions between bound CI subunits are strong compared with the interactions between the subunits and the DNA. Thus, the strong CI binding to pR appears to be a result of strong cooperativity between CI promoters at relatively weak DNA binding sites. Strong cooperativity in CI binding is supported by the gel shift results with pB and pR fragments. Noncooperative or weakly cooperative binding of a protein to fragments containing multiple operators should give rise, at any one protein concentration, to multiple retarded species, with the species differing in the number of protein subunits bound. The lack of such species (Figs. 2 and 9A), therefore, argues for strong cooperativity in CI binding to adjacent operators.

**DISCUSSION**

Gel shift and DNase I footprinting studies have confirmed that the 186 CI repressor is a sequence-specific DNA-binding protein and have identified four binding regions in the early control region of the phage genome. CI binds to the lytic promoters pR and pB that it represses and also binds to sites at the −330 and +350 positions of pR (FL and FR). The biological role, if any, of these flanking sites is not yet known. The CI binding region at pR is distinct from that of the Cro-like Apl protein, although there is some overlap between the two regions (11). We showed that the CI binding determinants at pR are located between the −58 and −14 positions of the promoter and confirmed for three vir mutants that the mutations at pR carried by these mutants reduced CI binding.

DNA recognition by CI is unusual. There appear to be two distinct recognition sequences, Type A and B. Furthermore, CI appears to be able to recognize a Type A half-site spacing variant, A'. This gives the structure AA-A-AB-A-B for CI binding to the pB-FL-pR-FR region. Three types of evidence provide strong support for our operator assignments. (i) There is good sequence conservation between the A/A'-type sequences and between the B-type sequences; with very strong conservation between the A-type sequences at pB, FL, and FR. (ii) The operators are consistent with the DNase I protection data, including many of the fine details of the footprints. (iii) The proposed operator sequences at the pR site can explain the large body of genetic data provided by the vir mutations. The A' sequence (Site I) is disrupted in all 19 of the vir mutants, invariably involving one of the fully conserved bases. Fifteen of these mutants also carry a disruption of the left B-type sequence (Site II), and one mutant carries a disruption of the right B-type sequence (Site III), with the most conserved bases often involved. The number of intact operators correlates with the strength of binding of CI to pR in the three mutants tested and with the sensitivity to immunity of all the vir mutants (8).

A number of proteins are known to recognize two different sequences or different half-site spacings. For example, the λ integrase protein is able to recognize distinct "core" and "arm" type DNA sequences, utilizing different regions of the protein (22), a property that seems to be general to the large family of "complex" integrases (see, for example, Ref. 23). An example of...
recognition of alternative half-site spacings is provided by the AraC protein, which has a flexible linker between the dimerization domain and the DNA binding domain of the protein (21). Further work is planned to define the spacing requirements within and between CI recognition sequences and to identify the DNA-binding regions of the protein.

Certain features of CI binding can be deduced from the data. In some of the A-type sites and in both the B-type sites, the minor groove in each arm of the operator remains sensitive or becomes hypersensitive to DNase I in the presence of CI. This observation is consistent with the rotationally symmetrical CI dimer being hypersensitive to DNase I in the presence of CI (see Figs. 5 and 8). Thus, the contacts made by CI with the bases in these operator arms must occur via the major groove. The inverted repeat nature of the operators and the spacing of the exposed bonds (9-11 bp at Type A sites and 8 bp at Type B sites) indicates that a rotationally symmetrical CI dimer is recognizing successive major grooves that are close to being on one face of the DNA helix.

Strong cooperative binding of CI to adjacent operator sequences was indicated by (i) the presence of only a single retarded species in gel shift experiments with fragments carrying multiple operators, and (ii) the finding that vir mutations at pR weakened binding to nonmutated operators. Cooperative binding is not surprising, as CI has been shown to exist in a monomer-dimer-tetramer-octamer equilibrium in solution (16) with interaction energies very similar to λ CI (24).

Although the data of Shearwin and Egan (16) show that CI in solution is predominantly dimeric at the concentrations used in our studies, the gel shift experiments suggest that a higher-order CI multimer is the active binding species, since the mobility of the major retarded species was very similar for DNA fragments with differing numbers of operators. This multimer seems able to occupy DNA containing up to two operators of any one type (two A-type operators at pB or two B-type operators at pR). Assuming that each operator is contacted by a CI dimer, then this species must be at least a tetramer. Studies of CI-DNA stoichiometries at the different operators are planned.

Once we have suitably characterized the relationship of type A, A’, and B sequences to CI binding, we will then be in a position to investigate the possible interaction between CI bound at FL, FR, and pR and its biological significance.

Acknowledgments—We thank Keith Shearwin for gifts of purified CI and Juliane Camerotto for plg C251.

REFERENCES

1. Lamont, I., Brumby, A. M., and Egan, J. B. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 5492–5496
2. Ptashne, M. (1986) A Genetic Switch: Gene Control and Phage λ, Blackwell Scientific Publications & Cell Press, Cambridge, MA
3. Johson, A. D., Poteete, A. R., Lauer, G., Sauer, R. T., Ackers, G. K., and Ptashne, M. (1981) Nature 294, 217–223
4. Bertani, L. E., and Egan, J. B. (1988) in The Bacteriophages (Calendar, R., ed) Vol. II, pp. 73–143, Plenum Publishing Corp., New York
5. Lamont, I., Richardson, H., Carter, D. R., and Egan, J. B. (1993) J. Bacteriol. 175, 5286–5288
6. Dodd, I. B., Kalironis, B., and Egan, J. B. (1990) J. Mol. Biol. 214, 27–37
7. Dibbens, J. A., Gregory, S. L., and Egan, J. B. (1992) Mol. Microbiol. 6, 2643–2650
8. Liu, B., Wong, M. L., Tinker, R. L., Gedshekek, E. P., and Alberts, B. M. (1988) Nature 366, 33–39
9. van Holde, K. E., Lohr, D. E., and Robert, C. (1992) J. Biol. Chem. 267, 2837–2840
10. Shearwin, K. E., Lohr, D. E., and Robert, C. (1992) J. Mol. Biol. 199, 379–382
11. Dodd, I. B., Reed, M. R., and Egan, J. B. (1993) Mol. Microbiol. 10, 1139–1150
12. Miller, J. H. (1972) Experiments in Molecular Genetics, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
13. Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) Gene (Amst.) 33, 103–119
14. Kalironis, B., Dodd, I. B., and Egan, J. B. (1993) J. Mol. Biol. 219, 199–209
15. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
16. Shearwin, K. E., and Egan, J. B. (1996) J. Biol. Chem. 271, 11525–11540
17. Maa, C., Carlson, N. G., and Little, J. W. (1994) J. Mol. Biol. 235, 532–544
18. Carlson, N. G., and Little, J. W. (1993) J. Mol. Biol. 230, 1108–1130
19. Drew, H. R. (1984) J. Mol. Biol. 176, 535–557
20. Drew, H. R., and Travers, A. A. (1985) J. Mol. Biol. 186, 773–790
21. Eustance, R. J., Bustos, S. A., and Schlief, R. F. (1994) J. Mol. Biol. 242, 330–338
22. Moisio de Vargas, L., Pargellis, C., Hasan, N. M., Bushman, E. W., and Landy, A. (1988) Cell 54, 923–929
23. Smith-Mungo, L., Chan, I. T., and Landy, A. (1994) J. Biol. Chem. 269, 27098–27080
24. Senear, D. F., Laue, T. M., Ross, J. B. A., Waxman, E., Eaton, S., and Rusinova, E. (1993) Biochemistry 32, 6179–6189
25. Richardson, H., Puppers, A., and Egan, J. B. (1989) J. Mol. Biol. 206, 251–255
26. Dibbens, J. A., and Egan, J. B. (1992) Mol. Microbiol. 6, 2629–2642
27. Richardson, H., and Egan, J. B. (1989) J. Mol. Biol. 206, 59–68
28. Kalironis, B., Pritchard, M., and Egan, J. B. (1988) J. Mol. Biol. 191, 211–220
29. Dodd, I. B. (1993) The Control of the LysI-Lysogeny System of Coliphage 186. Ph.D. thesis, University of Adelaide
