ESTABLISHMENT OF LYSOGENY IN BACTERIOPHAGE 186 – DNA BINDING AND TRANSCRIPTIONAL ACTIVATION BY THE CII PROTEIN

Keith E. Shearwin# and J. Barry Egan

Biochemistry Department, University of Adelaide, Adelaide, Australia 5001.

Running Title: DNA binding and transcriptional activation by 186 CII

#Corresponding author
Telephone 61-8-8303 5362
Fax 61-8-8303 5348
Email keith.shearwin@adelaide.edu.au
Summary

The CII protein of bacteriophage 186 is a transcriptional activator of the helix-turn helix family required for establishment of the lysogenic state. DNA binding by 186 CII is unusual in that the invertedly repeated half sites are separated by 20 base pairs, or two turns of the DNA helix, rather than the one turn usually associated with this class of proteins. Here we investigate quantitatively the DNA binding properties of CII and its interaction with RNA polymerase at the establishment promoter, $p_E$. The stoichiometry of CII binding was determined by sedimentation equilibrium experiments using a fluorescein labeled oligonucleotide and purified CII. These experiments indicate that the CII species bound to DNA is a dimer, with additional weak binding of a tetrameric species at high concentrations. Examination of the thermodynamic linkages between CII self association and DNA binding shows that CII binds to the DNA as a preformed dimer (binding free energy - 9.9 kcal/mol at 4 °C) rather than by association of monomers on the DNA. CII binding induces in the DNA a bend of 41 ($\pm$ 5) degrees. The spacing between the binding half sites was shown to be important for CII binding, insertion or removal of just one base pair significantly reducing the affinity for CII. Removal of five or ten base pairs between binding half sites eliminated binding, as did insertion of an additional 10 base pairs. CII binding at $p_E$ was improved marginally by the presence of RNA polymerase ($\Delta \Delta G = -0.5 (\pm 0.3)$ kcal/mol). In contrast the binding of RNA polymerase at $p_E$ was undetectable in the absence of CII, but was improved markedly by the presence of CII. Thus, CII appears to recruit RNA polymerase to the promoter. The nature of the base pair changes in mutant phage, selected by their inability to establish lysogeny, are consistent with this mechanism of CII action.
Understanding the mechanisms involved in the control of gene expression requires knowledge of the physicochemical interactions occurring between the components. Even relatively simple prokaryotic systems usually involve several linked protein-protein and protein DNA interactions whose mechanistic and thermodynamic features must be defined in order to gain a full appreciation of the system (eg. 1-6). We were interested in understanding the nature of the interactions involved in the self-association and DNA binding of bacteriophage 186 CII protein, whose concentration inside the bacterial host cell ultimately decides the fate of that infected cell.

Upon infection of the host cell, coliphage 186 can pursue one of two independent but interchangeable life cycles: lytic development or lysogeny. The region of the 186 genome controlling the fate of the phage (and its host) consists of two face-to-face promoters, \( p_R \) and \( p_L \) (Figure 1), which control transcription of the lytic and lysogenic operons, respectively. As \( p_R \) is a much (approximately 200 fold) stronger promoter than \( p_L \) (8), the choice between lysis and lysogeny does not reflect simple competition between these promoters, as is thought to be the case in the related coliphage P2 (11). Rather, establishment of the lysogenic state in 186 requires a single additional promoter (\( p_E \)), activated by the CII protein, to transiently increase leftward (lysogenic) transcription and hence initiate production of the maintenance repressor, CI and the integrase protein, Int (10). CI then represses the lytic promoter \( p_R \) to maintain lysogeny and autogenously control its own transcription\(^1\), while Int catalyses integration of the phage DNA into the host chromosome. CII is required only for establishment of lysogeny but not its maintenance (10, 12).

\( cII \) is the second gene of the 186 lytic transcript (Figure 1). The 18.7 kD protein product of the \( cII \) gene contains a helix-turn-helix motif in the N-terminal region, identified by the method of
Dodd and Egan (13). The CII-DNA interaction is unusual in that the half sites for binding (seven base pair perfect inverted repeats) are centred at -38 and -58 relative to the start of transcription, and thus are two turns of the helix apart, rather than the one helical turn normally associated with the helix-turn-helix class of DNA binding protein. We have shown previously that CII associates in solution from monomers to tetramers (10). Although the symmetrical nature of the CII binding site suggests the involvement of multimeric proteins, one cannot simply extrapolate to conclude that tetramers are also the “active” DNA binding species. In addition, the question of whether monomeric proteins multimerise on the DNA or whether pre-existing multimers in solution bind DNA has important energetic implications for the mechanism of transcription activation.

DNA binding by CII has been investigated in terms of stoichiometry, bending, binding free energies, and spacing requirements. In addition, we show that CII and RNA polymerase are the only proteins required for efficient activation of $p_E$. RNA polymerase binding at $p_E$ in the absence of CII is negligible, but is much stronger when CII is present. Thus CII appears to activate transcription by recruiting RNA polymerase to the promoter.
**Materials and Methods**

**Bacteria.** HMS174 (λDE3) pLysS (14) was used as the host for expression of CII from the T7 promoter of the pET3a translation vector of Studier et al. (14). C600 was used for general plating of 186. KH124 (mut-D5) (15) was used for mutagenesis of 186.

**Bacteriophage.** The 186 strains used were 186+ (16) unless otherwise stated. cIV476 is a clear plaque mutant containing a T to A change at position 23440 (12).

**Plasmids.** pLH3 is a pET3a (14) based CII expression plasmid (10), and was used as a source of CII for the selection of pE defective mutants of 186.

pKES1 contains the CII gene of 186 inserted into the pET3a expression vector (14). This plasmid was used as a source of CII for protein purification (10).

pBend 5 (17) is a derivative of pBluescript SK-(Stratagene) containing a 238bp fragment with 17 duplicated restriction sites to allow generation circularly permuted DNA fragments. Between these repeats are restriction sites (XbaI, HpaI, SalI) to allow insertion of small DNA fragments containing the binding site of interest.

pBend5- pEwt was generated by inserting oligonucleotide pEwt (Figure 5) into the Hpa I site of pBend 5. The sequence and orientation of the insert was confirmed by DNA sequencing.
pMRR9 (10) is a derivative of the lacZ promoter assay plasmid pRS415 (18) containing translation stop codons from pKO2 (19) and the pUC19 polycloning site (20). The $p_E$ promoter region of 186 and its derivatives were inserted into pMRR9 using a PCR based strategy. The (23287-23520) region of wild type 186 or the $p_E$ deficient mutants was amplified by PCR using primers 168 and 169, designed to introduce XbaI (23287 end) and KpnI (23520 end) restriction sites. The XbaI/KpnI digested PCR products were then ligated into XbaI/KpnI digested pMRR9, such that transcription from $p_E$ reads into lacZ.

pPN72 (10) is the T7 promoter and ribosome binding site and the cII gene (from pLH3) cloned into the large NruI-BamHI fragment of pACYC184 (21). This low copy number plasmid, which produces functional CII as tested by complementation, was used as the source of CII in the in vivo β-galactosidase assays.

Oligonucleotides. Oligonucleotides were purchased from Geneworks (Adelaide) or Gibco (USA). The sequences of oligonucleotides used for DNA binding studies are shown in Figure 5. Other oligonucleotides, used as sequencing and/or PCR primers, are as follows:

Apl control (top strand): 5′-TTGATGGCAAGTGTTGGCAAACAGAGTCAAATCA-3′
Annealed to its complement, used as a negative control in the sedimentation equilibrium experiments. CII does not recognise this sequence, but the 186 helix-turn-helix protein Apl does bind specifically.

$p_E$ wt-F (top strand): 5′-AGGATGCAACATTTTGTGATTTCAGGTTTCCAACATCCCACCTATGA-3′
Annealed to its complement (below) used in sedimentation equilibrium experiments. The CII binding site is shown in
p_E wt-F (bottom strand): 5’-Fluorescein-

TCATAGTTGGGATTTGGAAACCTGAAAAATCAAACATGTTGCATCCT-3’.

168: 5’-GCTCTAGAGCCGTAAGGCTGCAAGAA-3’ (anneals to 23287-23304, contains Xba I site shown in bold);

169: 5’-GGGTACCCTCGGCAGCTTCGCCATGTTG-3’ (anneals to 23503-23520, contains Kpn I site shown in bold);

174: 5’-GCTCTAGAAAGAAGAGCAGTTGCG-3’ (anneals to 23341-23358, contains an XbaI site. Used as PCR primer in DNase I footprinting;

Biotin-34: 5’-BiotinCGGTCAAGTACATCCACGTT-3’ (anneals to 22965-22984, used as a PCR primer in DNase I footprints);

SK: 5’-CGCTCTAGAAGGACTTAGTG-3’ (PCR primer for DNase I footprints);

Biotin-RSP: 5’-BiotinGGAAACAGCTATGACCATG3’ (PCR primer for DNase I footprints);

CII Bam biotin: 5’-BiotinCGGGATCCGTCAGACGCTCATCAGA-3’ (anneals to 23768-23785, PCR primer for DNase I footprints);

pBend SK: 5’-TAGTGGATCCCCCGGCTGCA-3’ (used to amplify by PCR the circularly permuted region of pBend- p_E wt);

pBend KS: 5’-CGACGGTATCGATAAGCTTG-3’ (used to amplify by PCR the circularly permuted region of
pBend5-\(p_E\) wt).

**Preparation of proteins.** CII purified as described by Neufing *et al.* (10). CII concentrations were measured spectrophotometrically, using an extinction coefficient of 7900 M\(^{-1}\) cm\(^{-1}\) at 280 nm (10). It was assumed that CII was 100% active in DNA binding. *E. coli* RNA polymerase was purchased from Roche-Boehringer.

**Analytical Ultracentrifugation.** Sedimentation experiments were performed with a Beckman XL-A analytical ultracentrifuge equipped with absorbance optics and a four hole An60Ti rotor. Approximately 100 µL of sample and 105 µL of reference solution were loaded in the sectors of the aluminium centrepieces. Following 24 hours of centrifugation, scans were compared at 3 hour intervals to ensure that equilibrium had been reached. Data were collected at 260 nm and 490 nm and at a spacing of 0.001 cm as the average of three absorbance measurements per radial position. The buffer for all experiments was 50 mM Tris-HCl, 0.1 mM EDTA, 150 mM NaCl, 10% glycerol, pH 7.5 (TEG 150). DNA was prepared for centrifugation as follows. Equimolar amounts of the complementary single stranded oligonucleotides were annealed by heating to 90 °C for 2 minutes, followed by slow cooling to room temperature. The annealed double stranded oligonucleotides were then ethanol precipitated, purified by polyacrylamide gel (12%) electrophoresis, ethanol precipitated again after elution from the gel and resuspended in TEG150 buffer. Concentrations of the oligonucleotides were determined spectrophotometrically using extinction coefficients of 8.3 x 10\(^5\) M\(^{-1}\) cm\(^{-1}\) (22) at 260 nm and 6.8 x 10\(^4\) M\(^{-1}\) cm\(^{-1}\) at 490nm for \(p_E\) wt-F. Protein was prepared for centrifugation by exhaustive dialysis against TEG 150 and the dialysate used as the reference solution for centrifugation. Buffer density (\(\rho\)) was measured in an Anton-Paar precision density meter to be 1.03953 g/ml at 5 °C and 1.03644 g/ml at 20 oC. The partial specific volumes (\(\bar{\rho}\)) of the predicted CII-DNA complexes were calculated as a weight average value of the for CII (0.724 ml/g) (10).
and DNA (0.55 ml/g) (23) according to eq 1

\[ \text{Eq. 1} \]

where \( M_{\text{CII}} \) and \( M_{\text{DNA}} \) are the molecular weights of CII and oligonucleotide, respectively. Thus, the partial specific volume for the protein-DNA complex was 0.644 ml/mg for a CII dimer bound to \( p_E \) wt F and 0.672 ml/mg for a CII tetramer bound to \( p_E \) wt F.

Sedimentation data was analysed, using Sigmaplot 4.0 for Windows (SPSS), by globally fitting equilibrium distributions (absorbance versus radial distance) to equation 2, the basic sedimentation equilibrium equation.

\[ \text{Eq. 2} \]

In equation 2, \( A_r \) and \( A_{r,0} \) are the absorbances at radial distance, \( r \), and \( r_0 \). \( M \) is apparent molecular weight, \( \nu \) is the partial specific volume, \( \rho \) is the solution density, \( \omega \) is the rotor speed in radians per second, \( T \) is the temperature in degrees Kelvin, \( R \) is the gas constant and \( e \) is a baseline error term. \( M, A_0 \) and \( e \) were fitting parameters.

**Preparation of DNA.** Double stranded DNA fragments with one strand \(^{32}\)P end-labeled for gel shift assays or DNase I footprints were generated by PCR in which one of the PCR primers had been \(^{32}\)P end-labeled using polynucleotide kinase. One microlitre of the appropriate phage stock was used as the template for the PCR reactions. For gel shift assays, the 252 bp double stranded
\( ^{32}P \)-labeled PCR product (generated using primers 168 and 169) was purified by polyacrylamide gel electrophoresis, the DNA eluted from the gel slice overnight at 37 °C, ethanol precipitated and resuspended in binding buffer before use. For DNase I footprints, the PCR products, generated using one \( ^{32}P \) end-labeled primer and one biotinylated primer, were passed over a PCR purification spin column (Geneworks, Adelaide) and attached to paramagnetic streptavidin-beads (Dynal), as described previously (24).

Unlabeled template DNA for the \textit{in vitro} transcription assay was generated in a PCR reaction (20 \( \mu \)L) using primers 168 and 169 and containing 1\( \mu \)L of 186+ phage stock as template. The 252 bp PCR product was purified by polyacrylamide gel electrophoresis before use.

\textit{Gel mobility shift assays}. Binding reactions (10 \( \mu \)L) were prepared by addition of DNA (~50 cpm), CII (exhaustively dialysed against TEG 150) and binding buffer (TEG 150). Each reaction also contained 50 ng of sheared salmon sperm DNA to minimise non specific binding. Reactions were left on ice for at least 30 minutes to allow attainment of equilibrium, and 6 \( \mu \)L loaded onto running 15% polyacrylamide (1 X TBE) gels containing 10% glycerol. Gels were electrophoresed at 4 °C at constant current (20mA) for approximately 2 hours. Upon completion of electrophoresis, gels were dried, exposed to a phosphorimager screen and quantitated using the volume integration feature of the Imagequant software package (Molecular Dynamics).

The fraction of DNA bound in each lane was calculated as (counts for the retarded band) / (counts for the whole lane), and corrected for a small degree of protein independent smearing using a no protein control. Results were fitted to the appropriate binding equation using Sigmaplot4 (SPSS). The values of fractional saturation were then corrected for the small amount of incompetent (non binding) DNA (2). The DNA concentration was sufficiently low that total protein concentration
could be substituted for free protein concentration.

**DNase I footprinting**

Footprinting experiments were performed essentially according to Sandaltzopoulos and Becker (24). This method uses magnetic bead technology to facilitate sample preparation. Double stranded DNA fragments for footprinting were prepared by PCR using a $^{32}$P end labeled primer and a biotinylated unlabeled primer. Following PCR, the PCR reaction (20 µL) was passed over a PCR purification spin column (Geneworks, Adelaide) to remove any unincorporated biotinylated primer which would compete with full length product for binding to the beads. The eluate from the spin column (60 µL) was added to 75 µL of streptavidin coated magnetic beads (Dynabeads), prepared according to the manufactures recommendations, and incubated for one hour at room temperature to allow the biotinylated, radiolabeled PCR product to bind. The beads were then washed several times, resuspended in 50 - 100 µL binding buffer and stored at 0 °C for up to 1 week.

Bead DNA (5 µL, ~6000 cpm) was added to binding buffer containing appropriate CII concentrations, in a total volume of 40 µL. The footprint binding buffer consisted of 50 mM Tris-HCl (pH 7.5), 0.1 mM EDTA, 10% (v/v) glycerol, 75 mM NaCl, 10mM MgCl$_2$, 1.5 mM CaCl$_2$, 1 µM bovine serum albumin (BSA). These binding reactions were incubated at 37 °C for at least thirty minutes to allow attainment of equilibrium, prior to addition of DNase I (5 ng). The DNase I reaction was allowed to proceed for exactly two minutes at 37 °C before being stopped with 50 µL of stop solution (4M NaCl, 100 mM EDTA). The beads were washed once with 100 µL of 2M NaCl, 20 mM EDTA, once with 100 µL of 10 mM Tris-Cl, 1 mM EDTA, pH 8.0 and resuspended in 6 µL of loading buffer (90% formamide, 10 mM EDTA). The reactions were heated to 90 °C for 3 minutes and 5 µL loaded immediately onto a 6% denaturing polyacrylamide gel. Electrophoresis was at 1500 V (constant voltage) for approximately 2 hours. Following electrophoresis, gels were
dried onto filter paper, exposed overnight to a phosphorimager screen and viewed using Imagequant. Binding affinities were quantitated according to Brenowitz et al. (25). The upper and lower end points of the transition curves were included as adjustable parameters in the curve fitting procedure. DNA concentrations were sufficiently low that total protein concentration could be equated to free protein concentration.

*Bending assay*

A 48 base pair DNA fragment containing the CII binding site (pE wt, Figure 5) was prepared by annealing complementary oligonucleotides and was ligated into the *Hpa* I site of pBend 5 (17) to give pBend-pE wt. A 325 bp region from this plasmid containing 17 circularly permuted restriction sites and the CII binding site was then amplified by PCR using primers pBend SK and pBend KS. This fragment was labeled by inclusion of $^{32}$P $\alpha$-ATP (10 $\mu$Ci) in the PCR reaction. The PCR product was purified by polyacrylamide gel electrophoresis and an aliquot digested in 10 $\mu$L reactions with *Mlu*I, *Nhe*I, *Cla*I, *Xho*I, *EcoRV*, *Sma*I, *Nru*I, *Rsa*I or *Bam*HI. Digestion produced three fragments, including a 168 bp fragment containing the CII binding site located at various positions within the fragment. Two $\mu$L of this digest was used in binding reactions for determination by gel shift assay of the electrophoretic mobility of the protein-DNA complex. Binding reactions were in TEG 50 buffer and contained 500 nM CII. Samples were loaded on 8% polyacrylamide, 0.5X TBE gels containing 10% glycerol and run at 20 mA constant current. Loading dye was run in a separate lane. Following electrophoresis, gels were dried, exposed to a phosphorimager screen, and mobilities quantitated using Imagequant software (Molecular Dynamics) and the bend angle quantitated according to equation 3.
where $\alpha$ is the bend angle and $\mu_1$ and $\mu_2$ are the relative mobilities of DNA fragments containing the binding site at the middle and at the end of the fragment, respectively.

*Chromosomal single copy lacZ fusions*

$\lambda$RS45 and pMRR9 share portions of the N-terminus of both the ampicillin gene and the lacZ gene thus allowing any promoter insert in pMRR9 to be recombined into the phage. Lysogenisation of this recombinant phage gives a single copy chromosomal fusion. Strain NK7049 (18) transformed with the appropriate pMRR9 derivative was used as the host for growth of the $\lambda$RS45 phage vector.

Phage stocks obtained were plated on NK7049 and single recombinant plaques selected on the basis of colour in the presence of X-gal and purified once by streaking across a lawn of NK7049. Independent blue lysogens from at least two recombinant plaques were purified by restreaking. Single copy status of these lysogens was confirmed by PCR (26). For assay of $pE$ $\beta$-galactosidase activity, the appropriate CII expression plasmid (pPN72 or its parent pACYC184) was transformed into these lysogens and liquid cultures started by picking single colonies directly from the transformation plates. $\beta$-Galactosidase activity was assayed as described (27) with the exception that the cultures were grown in Luria-Bertani (LB) medium and permeabilised with 30 $\mu$L of chloroform.

*Transcription run off assays*

Run off assays (20 $\mu$L reaction volume) were performed at 37 °C in footprint buffer containing CTP, GTP and ATP (all at 200 $\mu$M), 20 $\mu$M unlabeled UTP, 5 $\mu$Ci (83nM) $\alpha^{32}$P UTP, 150 nM RNAP, 250 nM CII, 5 nM template DNA and 4 units of RNase inhibitor (Promega). Reaction
mixtures without template DNA were incubated for 15 minutes at 37 °C, prior to starting the reactions by the addition of template DNA. Transcription was allowed to proceed for 20 to 40 minutes, before 8 µL aliquots were removed and added to an equal volume of 90% formamide loading buffer. Samples were heated for 2 minutes at 90 °C prior to separation of the products by electrophoresis on a 6% polyacrylamide gel containing 8 M urea. 32P end-labeled HpaII digested pUC DNA (Geneworks, Adelaide) was used as a marker to estimate the size of the RNA transcripts. Gels were dried and exposed overnight to a phosphorimager screen. Images were analysed using Imagequant software. Control experiments showed that the generation of run off transcripts was dependent upon the addition of both RNAP and template DNA.
Results

Stoichiometry of CII binding. In previous work, analytical ultracentrifugation of purified CII protein revealed the existence of a monomer-dimer-tetramer self association equilibrium in solution (10). It is worth emphasising at this point that the oligomeric form of a DNA binding protein determined at one particular concentration in solution does not necessarily correlate with the oligomeric form of the protein bound to DNA. For example, bacteriophage lambda CII has been shown to form tetramers in solution (28) but it has not to our knowledge been demonstrated that tetramers are the active DNA binding species, although this is often assumed. The symmetrical nature of the 186 CII binding site suggests the involvement of an oligomeric species in binding and thus the possibility of thermodynamic linkage between protein self association and DNA binding, a common theme in the regulation of gene expression. As a first step in investigating any such linkage, we wished to determine the stoichiometry of CII binding to its operator. Sedimentation equilibrium, using a fluorescein-labeled oligonucleotide, was the method employed (6, 29). The use of absorbance optics in sedimentation equilibrium to determine the stoichiometry of a protein-DNA interaction is complicated by the significant spectral overlap between protein and DNA in the 260 -280 nm range (30). Use of a fluorescein-labeled oligonucleotide however allows separation of the protein and DNA signals. By monitoring the equilibrium distribution at 490nm, the absorption maximum of fluorescein, the apparent molecular weight of the oligonucleotide can be determined in the presence and absence of the protein in question. In the presence of an interacting protein, the apparent molecular weight will reflect that of the protein-DNA complex (provided the DNA is fully occupied) and thus the stoichiometry of the interaction can be estimated (6, 29).

A 48bp oligonucleotide, containing a centrally positioned CII binding site and a fluorescein moiety at the 5' end of the bottom strand (see experimental procedures), was sedimented to equilibrium in the absence or presence of a 6.6 fold excess of CII. This concentration of CII was
sufficient to saturate all of the available binding sites. The equilibrium distributions, obtained at two rotor speeds, were fitted globally to equation 2, with the apparent molecular weight as one of the fitting parameters. The results of these experiments and of the corresponding control experiments, in which an oligonucleotide without a CII binding site was used, are summarised in Table 1. The apparent molecular weights of the two oligonucleotides (pE wt-F, control) in the absence of CII are close to those calculated on the basis of sequence. The slight underestimation of the oligonucleotide molecular weight may be due to uncertainty in the partial specific volume of DNA, or to the presence of small amounts of dissociated single stranded DNA. In the presence of an eight fold molar excess of CII, the apparent molecular weight of the control oligonucleotide was unchanged, indicating no interaction with CII. However, in the presence of a 6.6 fold molar excess of CII, the Mw\text{app} of pE wt-F increased to approximately 79,000, indicating a CII (Mw 18,694) to DNA ratio of ~ 2.5. Thus, CII binds to DNA as a dimer, with a possible contribution of tetramer binding at this high protein concentration.

**Thermodynamics of binding.** Knowledge of stoichiometry alone does not provide quantitative information on linkages between protein association and DNA binding, just as knowledge of the solution state of the protein does not imply stoichiometry of binding. Complete analysis of such an interacting system requires examination of both protein self association in solution and protein binding to its DNA acceptor.

Sedimentation equilibrium of purified CII (20 µM loading concentration) (10) indicated a weight average apparent molecular weight of ~68,000, demonstrating that CII can associate to at least tetramers in solution (tetramer molecular weight = 74,780). Any model for CII assembly and DNA binding (Figure 2) must therefore be consistent with (i) the self association data for CII in
solution (monomer-dimer-tetramer) (ii) the stoichiometry of CII binding (dimeric) and (iii) the binding data for the association of CII with DNA. The binding scheme shown in Figure 2 is based on each of the possible CII species having the ability to bind DNA. In this scheme for example, CII could assemble as monomers at the individual half sites, perhaps cooperatively to give bound dimers ($K_3K_4$ pathway), or alternatively could bind to the DNA as the preformed dimeric species ($K_1K_6$ pathway). Consideration of the equilibria of Figure 2 in terms of the fraction of DNA bound (the parameter obtained in gel mobility shift assays) gives equation 4.

\[
\theta_{\text{bound}} = \frac{[MK_3 + K_1K_6][M] + K_1^2K_2K_7[M]^3}{1 + K_3[M] + K_1K_6[M]^2 + K_1^2K_2K_7[M]^4}.
\]  
(Eq. 4)

where $\theta_{\text{bound}}$ is the fraction of DNA bound. $[M]$, the concentration of unbound monomer, is related to its base-molar counterpart (weight concentration of free CII divided by monomer molecular weight), $[\text{CII}]_{\text{Free}}$, by the expression

\[
(M)
\]  
(Eq. 5)

The interaction of CII with its binding site at $p_E$, was quantitated by gel mobility shift assay. The DNA used was a 48 base pair duplex of the same sequence used in the stoichiometry experiments but lacking the attached fluorescein moiety. The DNA concentration used was sufficiently low that the concentration of CII bound to DNA was considered negligible and total CII concentrations could be substituted for free CII concentrations. As shown in Figure 3a, only one retarded species was observed, which on the basis of the stoichiometry experiments we assign as the dimer. The combined results from three independent CII titrations, plotted as the fraction of DNA bound as a
function of total CII concentration (in terms of monomer) are shown as solid circles in Figure 3b.

The data of Figure 3 were fitted to equation 4 (Table 2) using two alternative pairs of values for $K_1$ and $K_2$. These values were obtained from the earlier sedimentation equilibrium experiments (10) which did not allow us to describe a unique association model, due in part to the difficulty in obtaining data at sufficiently low CII concentrations. $K_3$, the binding constant describing the binding of a CII monomer to an isolated half site was estimated from gel shift experiments using DNA fragments containing only the left half site of $p_E$. No retardation was observed even at 10 µM CII, thus a conservative upper limit for $K_3$ of $1 \times 10^5$ M$^{-1}$ was assigned. $K_6$ and $K_7$ were the fitted parameters, with [M] the concentration of monomer calculated for each total CII concentration from equation 5. The line of best fit (SSR = 0.734) for model 1 is shown in Figure 3b. Clearly this model does not provide a satisfactory description of the data. Using fixed values of $K_7$, the association constant describing the binding of tetramer to $p_E$, and fitting only for $K_6$ did not improve the quality of the fit.

In model 2, the monomer-dimer association is strong, such that the fraction of monomeric CII is insignificant at the concentrations under consideration. The best fit of the gel shift data to model 2, again with $K_6$ and $K_7$ as fitting parameters, is shown as a solid line in Figure 3c. This curve describes the data well (Table 2). The fitted values obtained were $K_6 = 7.1 (\pm 1) \times 10^7$ M$^{-1}$ (in units of CII dimers), while $K_7$ was not well defined (the error in $K_7$ was larger than the value of $K_7$ itself). Fixing $K_7$ at zero did not alter the quality of the fit. Thus, this description of CII binding, that is preformed dimers associating with the DNA, is consistent with both the CII self association data (although not the preferred model of CII self association in Neufing et al., 1996), the stoichiometry results and the DNA binding data. This result emphasizes the importance of
quantitating both association and ligand binding in studying interacting systems.

In order to consider further the possibility of CII tetramers binding to the DNA, the effect of \( K_7 \) on the shape of the binding curve was simulated. As shown in Figure 3c, only values of \( K_7 \) greater than \( 10^9 \) M\(^{-1}\) gave significant deviations from the best fit curve. This prediction makes sense when it is realised that given the moderate value of the tetramerisation constant, \( K_2 \), the fraction of tetrameric species present only becomes significant at high protein concentrations. Thus in order for DNA binding by tetramers to influence the shape of the binding curve, the affinity of the poorly populated tetramer for the DNA would have to be high. In terms of the experimentally derived binding curve, binding of tetrameric CII to DNA with an affinity of less than \( 10^9 \) M\(^{-1}\) would not have been detected and therefore cannot be ruled out. It should be noted however that no second retarded species was observed in gel shift assays even at the highest concentrations employed. We expect therefore that tetramer binding to DNA is not physiologically significant.

**DNase I footprints.** A prediction of this preformed dimer model of CII assembly is that CII binding to individual half sites (i.e. the singly liganded species) will not be observed, since preformed dimers are the only 'active' binding species. This prediction was tested by DNase I footprinting using wild type and mutated (KS54) \( p_E \) templates (Figure 4). KS54 is a CII binding site mutant with a single A to G base pair change in the right half site which reduces the ability of CII to activate transcription to 63% that of wild type (described below). A DNase I footprint titration at 37 °C of the wild type template (Figure 4a) gave a clear footprint (10) in which both half sites were protected equivalently as CII concentration was increased. Quantitation of the titrations confirmed this observation (\( \Delta G_{\text{left}} = -11.0 \) kcal/mol, \( \Delta G_{\text{right}} = -11.0 \) kcal/mol).
In the titration of the KS54 template (Figure 4b), the overall binding affinity for CII was reduced, but like the wild type situation both half sites were protected equivalently as protein concentration was increased ($\Delta G_{\text{left}} = -8.1 \text{ kcal/mol}$, $\Delta G_{\text{right}} = -8.1 \text{ kcal/mol}$). Preferential protection of the wild type half site was never observed. This footprint data is consistent with the preformed dimer mode of binding, although it cannot formally exclude the possibility of highly cooperative binding of monomers, such as that seen for BirA binding (5).

**Half site spacing.** 186 CII is unusual in that two turns of the helix separate the inverted repeat sequences involved in binding. A number of DNA binding proteins such as AraC (31, 32) and CytR (33) are quite flexible in spacing requirements, while others are not (e.g. 2). These requirements for a particular spacing can reflect the degree of flexibility inherent in the protein structure. We wished to determine if half site spacing was critical for DNA binding by CII. A series of oligonucleotides were designed to remove or introduce bases between the half sites (Figure 5a) and were used in DNase I footprints to determine how CII binding was affected. Removal of ten base pairs from between the half sites ($p_{\text{E}10}$) resulted in complete loss of CII binding at the highest CII concentration tested (5 $\mu$M). Similarly, when five bases were removed ($p_{\text{E}15}$), positioning the half sites on opposite sides of the helix, no binding was observed. Insertion of ten bases into the spacer region ($p_{\text{E}30}$), giving 30 bp between half sites resulted in weak but detectable binding in mobility shift assays, however no DNase I footprint could be obtained (not shown). Removal of just one base pair between the half sites ($p_{\text{E}19}$) gave a distinct footprint, while insertion of one base pair ($p_{\text{E}21}$) gave a weaker but still distinct footprint at the same concentration of CII (Figure 5b). Quantitation of CII binding to the spacing templates was attempted by gel shift assays conducted at 4 °C. Removal of one base pair between the CII half
sites reduced the dissociation constant from 14 nM (-9.8 kcal/mol) to 90 nM (-8.8 kcal/mol). Binding to the other templates, including \( pE \) 21, was sufficiently weak that reliable dissociation constants could not be obtained.

**DNA bending.** Some prokaryotic transcriptional activators such as the catabolite activator protein (CAP) induce large (> 90 degree) bends in the DNA upon binding (34). Given that a dimer of the relatively small CII protein must stretch over two turns of the DNA helix, we wished to determine whether this is achieved at least in part by DNA bending, thus reducing the distance between the half sites. The circular permutation assay of Wu and Crothers (35) was used to estimate the degree of DNA bending. This method relies on the observation that the electrophoretic mobility of a protein-DNA complex will vary depending upon the position of the binding site within the fragment. A series of gel mobility shift assays were performed using pBend5 (17) as described in experimental procedures and the degree of bending estimated from equation 3 as 41 (±5) degrees (n=3). Thus CII appears to induce a moderate bend in the DNA. This is consistent with the footprints shown in Figure 4, the obvious enhancement in the centre of the half sites being indicative of DNA deformation.

**Requirements for \( pE \) activity**

Although strong activation of \( pE \) by CII has been measured in vivo (10), we had no evidence that CII and RNAP are the only factors required for activation of \( pE \). An in vitro transcription assay was performed to determine whether \( pE \) requires any additional host or phage derived factors for activity (Figure 6). In vitro transcription reactions were initiated by the addition of PCR derived 252 bp template DNA to a reaction mix containing nucleotides (including \(^{32}\text{P} \)}
UTP), RNAP and CII. Transcription from the +1 of $p_E$ (Figure 1) would generate an RNA product of 122 bases. Figure 6 shows that in addition to end to end transcripts observed both in the presence and absence of CII, a run off transcript of the expected size was observed specifically in the presence of CII. Some smaller premature termination products were also observed. Thus, $p_E$ requires no other factors besides RNAP and the CII activator protein. Comparison of the amount of $p_E$ transcript produced in vitro with that produced from a control promoter of similar strength (the constitutive 186 $p_R$ promoter) indicated that $p_E$ activity was of the expected magnitude.

**RNA polymerase binding**

The nature of the protein-DNA interactions required for $p_E$ activation was examined by DNase I footprinting. We have shown previously that purified 186 CII gives a strong DNase I footprint covering -28 to -68 of $p_E$ (ref 10; Figure 4, lane 2). The binding characteristics of RNAP at $p_E$ were similarly investigated. The solution conditions for the footprint experiments were identical to those employed in the in vitro transcription assay except that no nucleotides were present. Figure 7 shows that no protection from DNase I cleavage was observed on either top or bottom strand of $p_E$ in the presence of RNA polymerase alone. In the presence of both CII (250 nM) and RNAP (150 nM), the CII footprint was only slightly changed, the strong enhancement in the centre of the bottom strand CII footprint being shifted one base in the 3’ direction. In addition, there was a weak pattern of protections which extended from the edge of the CII footprint to +17 on the bottom strand and +22 on the top strand. This is consistent with the size of the RNAP footprint observed for many prokaryotic promoters (e.g. 36). Complete protection of the $p_E$ region by RNAP was never observed, indicating that even with assistance from CII, RNAP does not stably
occupy \( p_E \).

**Cooperativity between CII and RNAP.** The existence of two (or more) adjacent protein binding sites on DNA presents the opportunity for cooperative interactions, either favourable or unfavourable, between those proteins. Any change (\( \Delta \Delta G \)) in the intrinsic free energy of binding of one protein (\( \Delta G_{\text{intrinsic}} \)) in the presence of a constant saturating concentration of a second protein reflects a cooperative interaction energy (\( \Delta G_{\text{coop}} \)) (37) where \( \Delta \Delta G = (\Delta G_{\text{intrinsic}} + \Delta G_{\text{coop}}) - (\Delta G_{\text{intrinsic}}) \).

While CII enhances the binding of RNAP (Figure 7), the extent of this assistance could not be quantitated due to the absence of observable binding of RNAP alone. In contrast, the reverse situation whereby RNAP may provide additional free energy to CII binding is amenable to further analysis. Quantitation of CII binding in the absence of RNAP and separately in the presence of a constant saturating RNAP concentration (150 nM) was performed by DNase I footprinting. Fitting of the titration data to the Langmuir binding isotherm indicates that at 37 °C CII alone binds with free energy of -10.4 (± 0.1) kcal/mol under these conditions. In the presence of 150 nM RNAP, the apparent CII binding was slightly (approximately 2 fold) enhanced at -10.9 (± 0.2) kcal/mol. Thus, RNAP assists the binding of CII (\( \Delta G_{\text{coop}} \)) by - 0.5 (± 0.3) kcal/mol. This is similar to the slight cooperativity between P22 Cl and RNAP (38) but less than the 10 fold enhancement of lambda CII binding by RNAP (39).

**Selection of activation deficient \( p_E \) mutants.** In order to understand further the protein-DNA and protein-protein interactions involved in CII activation of the \( p_E \) promoter, a genetic approach was
devised to isolate 186 mutants unable to efficiently establish lysogeny. It was hoped that such mutants would be informative with respect to the base sequences required for effective activation of \( p_E \). These activation defective mutants were selected by plating stocks of 186\(^+\) phage, either wild type or mutagenised by passage through a mutD strain, on \( E. coli \) strain HMS174 (\( \lambda \)DE3), carrying plasmids pLH3 and pLyS. pLH3 is a CII expression plasmid derived from pET3a (14). Under the conditions employed, wild type plaques are not produced, presumably because infection is driven 100\% to lysogeny. Clear plaque mutants (presumptive \( cI \) mutants unable to maintain lysogeny) arose at a frequency of \( \sim 10^{-5} \), while turbid plaques (presumptive \( p_E \) activation mutants) arose approximately 100 fold less frequently. The clear plaque class of mutants would also be expected to contain \( p_E \) null mutants at a low frequency, however these would be indistinguishable from the more frequent \( cI^- \) mutants.

Turbid plaques were picked and the phage purified twice by replating. Plaque morphology varied from quite turbid to almost clear, indicating a range of \( p_E \) activities. All mutants were clear when plated on the parental HMS174 pET3a pLysS strain. The \( p_E \) region of these mutants was sequenced. The mutations obtained are shown in Figure 8, several of the mutations arising a number of times. At least three classes of mutants were observed; (i) two mutants (KS11 and KS61) mapped to the -10 region of \( p_E \), (ii) a mutant containing a one base pair deletion between -10 and the CII binding site (KS10) was obtained, while (iii) the majority of mutations were located within or adjacent to the inverted repeats of the CII binding site.

The effects of these mutations on the \textit{in vivo} activity of \( p_E \) were measured by generating single copy chromosomal \( p_E \)-lacZ fusions using the lambda reporter system of Simons \textit{et al.} (18). These results are shown in Table 3. Wild type \( p_E \) activity increased from less than 1 Miller unit in
the absence of CII (pACYC184 only) to 1288 units in the presence of CII, expressed from pPN72. This strong activation is consistent with the results of Neufing et al. (10). The mutations described in Figure 8 had a wide range of effects on $p_E$ activity. In the absence of CII (pACYC184 only), all $p_E$ mutants displayed background activity of less than 1 unit. In the presence of CII, the activity of the mutants ranged from less than 1 unit (0.05% of wild type) to 850 units (63% wild type) for KS54. The -10 mutants, KS11 and KS61, had the most severe effect on $p_E$ activity.

The affinities of CII for these mutants were measured by gel shift assay in order to determine whether the deficiency in $p_E$ activation was due to loss of CII binding or elsewhere. The dissociation constant for the binding of CII to the wild type template, determined using a 253bp PCR product was 11 ($\pm$2) nM, in agreement with the gel shifts of Figure 3 which were obtained using a 48 bp $p_E$ containing oligonucleotide. As expected, the binding of CII to the -10 mutants KS11 and KS61 was indistinguishable from CII binding to wild type $p_E$, indicating that these mutations affect only the RNA polymerase - DNA interaction. The KS10 mutant which contains a one base pair deletion between the -10 region and the CII binding site also displayed wild type affinity for CII, suggesting that the relative spacing and/or orientations of the CII activating surface and RNA polymerase are critical for $p_E$ activation. In the case of the mutants located in or near the inverted repeats, one base pair changes in the right half site were sufficient to dramatically reduce binding affinity. Only KS54, which gave 63% of wild type activation in vivo, displayed a slight retardation and then only at the highest CII concentration examined, consistent with the DNase I footprints of Figure 4 which show that CII binds to the KS54 mutant ~ 100 fold more weakly than to the wild type template. No retardation of the DNA was observed for any of the other right half site mutants at 3.2 $\mu$M CII. Partial activation of these $p_E$ mutants in vivo was observed presumably
because the plasmid pPN72 supplies a sufficiently high intracellular concentration to allow weak occupation of the mutated CII binding sites.

The results for the left half site were more complex. The cIV mutation (12), a T to A change located in the 6th base of the left half site, did not bind DNA at the concentrations examined, in agreement with the DNase I footprint results of Neufing et al (10). The KS55 mutant gave poor binding in gel shift assays, the retarded band being quite diffuse, suggesting a more rapid dissociation during electrophoresis. Unexpectedly, the KS5 and KS9 mutants despite being deficient in the ability to activate $p_E$ (Table 3), bound CII with wild type affinity, indicating that this CG base pair at -34 of $p_E$ does not contribute to the CII-DNA interaction.
Discussion

The thermodynamic linkage of protein association and DNA binding is a common theme in the control of gene expression in both prokaryotes *e.g.* lambda repressor (40, 41), P22 Arc repressor (2, 42), LexA (1), and eukaryotes *e.g.* Doublesex (43), CREB (6), Bicoid (44). Protein association and DNA binding can be further linked to phosphorylation *e.g.* NtrC (45), to the binding of small ligands *e.g.* TyrR (46) or to interactions with additional DNA binding proteins (coactivators) *e.g.* Gal/CAP (37). For any regulatory system it is important to dissect these linkages in order to understand how changes in protein concentrations lead to changes in the level of transcriptional regulatory complex and hence levels of gene expression.

We have shown previously that 186 CII associates in a monomer-dimer-tetramer equilibrium (10) and here we demonstrate that it is the dimeric species which is the multimeric form bound to the DNA. Examination of the linkage between binding and self association indicates that the correct model for CII self association is that of strong dimer formation, with weaker association to tetramers (model 2). Thus, formation of CII dimers is not dependent on DNA binding. Tetramers of CII may also bind the DNA, but this association is sufficiently weak that we do not expect the DNA bound tetramer to be biologically relevant. Cooperativity does occur at the level of the CII –RNA polymerase interaction, such that CII ‘recruits’ polymerase to the DNA, or in other words provides additional free energy to strengthen the RNA polymerase-DNA interaction. An alternative mechanism of activation, that RNA polymerase may interact with $p_E$ in the absence of CII to form an inactive closed complex, and that CII brings about an isomerisation to the active open complex, is not supported by the DNase I footprint data at $p_E$ which show no change in the presence of RNAP only.

In the absence of CII, the 186 $p_E$ promoter is devoid of activity (Table 3), and no binding of
RNAP is observed in vitro. CII activates transcription from the establishment promoter $p_E$ by assisting RNAP binding. This mechanism is similar to phage lambda CII mediated activation of the pRE promoter (47) and is consistent with the role these promoters play in the phage lifecycle. Establishment of the prophage state requires a rapid, strong burst of lysogenic transcription in response to the CII signal, and this transcription ceases once repressor has been made and the CII signal has decayed. Several observations suggest however that the details of activation are different in 186 and lambda. Although 186 $p_E$ is considered to have no -35 consensus (10), located the optimal 17 bases from the -10 box there is a 2 out of 6 match to the E. coli -35 consensus (TTGACA) (Figures 1 and 8) overlapping the left inverted repeat by 2 bases. In contrast, the lambda CII binding site sits directly over the -35 position TTGCN₆TTGC/T, where N₆ is the expected position of the -35 consensus sequence (39). Thus although the unactivated 186 $p_E$ promoter has no transcriptional activity (Table 3), perhaps 186 CII assists a weak pre-existing RNAP / -35 sequence interaction, rather than directly substituting for it as appears to be the case in lambda. Consistent with this, the KS5 and KS9 $p_E$ activation deficient mutants reduce the -35 match to 1 out of 6, without reducing CII binding. The KS55 mutant also changes the -35 sequence to a 1/6 match, but alters the CII site sequence as well. The KS10 mutant isolated in this study is a one base pair deletion between the -10 and the CII binding site, reducing the -10 to -35 spacer length from the optimal 17 bp (48) and resulting in a dramatic reduction in $p_E$ activation. All these mutations may reduce further the promoter’s affinity for RNAP, making it more difficult for CII to recruit polymerase to the promoter. Alternatively, these mutations may simply misalign RNA polymerase and the activating surface of CII, making it more difficult for polymerase to access the activating residues on CII, as has been proposed for mutations which decrease activation of the lambdoid phage 434 pRM promoter by the 434 repressor (49). This misalignment may be
the consequence of underwinding or overtwisting of the DNA, or altered distances between critical bases/amino acids in the protein-DNA complex (49, 50).

There are other significant differences between 186 and lambda CII DNA binding mechanisms. Lambda CII binds (tetramer binding has been assumed but not directly demonstrated) to direct repeat sequences centred one turn of the helix apart (39), while 186 binds as a dimer to inverted repeats located two turns of the helix apart. The *E. coli rpoA341* mutation prevents activation of the three CII responsive promoters in lambda, indicating a contact with the \( \alpha \) subunit of RNAP (51). We do not know which RNAP subunit is contacted by 186 CII. The location of the CII half sites relative to the promoter may be pertinent to this question. The promoter proximal half site partly overlaps the -35 region of \( pE \). Other activators located close to the -35 region activate through contact with the sigma subunit of RNA polymerase (52). The promoter distal half site is centred at -58, a position where other activators are found to interact with the alpha subunit of RNAP. It will be of interest to determine whether CII activates through contacts with alpha, sigma or both subunits of RNAP.

What is the significance of the unusual 20 base pair spacing between the CII half sites? This two helical turn spacing is critical for CII binding. CII is unable to bind DNA if the half site spacing is changed to one or three turns of the helix. The addition or deletion of just one base pair between the half sites, which in addition to bringing the sites closer together also rotates the half sites 36° relative to each other, has a deleterious effect on binding, although a 19 base pair spacing is relatively well tolerated. These results indicate that the CII dimer cannot compensate for these altered spacings and may be somewhat inflexible. In order to span the two turns of the helix between operators, the 37 kDa CII dimer may have an elongated rather than globular structure. CII binding induces in the DNA a moderate (approximately 40°) bend. Bending may be one
mechanism for bringing the two operator half sites closer together, but only at the energetic cost of maintaining the strained DNA conformation. In addition bending may be necessary for correct contact of CII with RNA polymerase. There are few other examples of activators that bind over two turns of the helix. One example is the p4 protein of \textit{Bacillus subtilis} bacteriophage $\Phi$29. It recognises 8 bp inverted repeats separated by a 15bp spacer, (i.e. the centres of the operators are two turns of the helix apart) and activates transcription by stabilising RNA polymerase binding as the closed complex (53). Given the weak protection of $p_E$ in DNase I footprint experiments afforded by RNA polymerase, it is tempting to speculate that polymerase, even with assistance from CII, does not bind strongly to the promoter and that promoter clearance may be the favourable step in activation compared to a constitutive promoter of equivalent intrinsic strength.
References


**Acknowledgments**

KES was supported by an ARC Australian Research Fellowship and JBE by grants from the Australian Research Council. Analytical ultracentrifugation experiments were performed at the University of Queensland under the guidance of the late Michael Jacobsen, whose assistance is gratefully acknowledged. We also thank Dr. S. Adhya for the gift of pBend5 and members of the Egan lab for helpful discussions.

**Footnotes**

Figure 1. **Map of the control region of coliphage 186.**

The organization of the 186 genome is shown, with the early region from the t1 terminator (sequence position 20930) to the tR1 terminator (position 24412) enlarged to show details (7-9). Sequence numbering starts at the left cos end (0%) of the 186 genome. Genes are shown as grey boxes: 69, unknown function; int, integrase; cI, immunity repressor; apl, excisionase and transcriptional control; cII, establishment of lysogeny; fil, inhibitor of cell division; dhr, inhibitor of host replication. Promoters are shown as arrowheads, their transcripts as arrows, terminators as stem loops and the phage attachment site attP as a solid box. The pE region is further enlarged to show sequence details (10). The stop codon of the apl gene (TAA) and the start codon of the cII gene (ATG) are in bold. The +1 and -10 box of pE are indicated, while the seven base pair inverted repeats of the CII binding site, located within the cII gene, are shown as unshaded boxes.

Figure 2. **Scheme for CII binding to DNA.**

Sedimentation equilibrium experiments on CII in solution (10) indicate that CII exists in solution in an equilibrium between monomers, dimers and tetramers. The equilibrium constants describing this equilibrium (K1 and K2) could not be uniquely resolved from these experiments alone. The stoichiometry experiments described in Table 1 indicate that the CII species bound to DNA is a dimer, which may arise from direct binding of preformed dimer (K1K6 pathway) or by assembly of two monomers on the DNA (K3K4 pathway). Some degree of binding by tetramer may also be possible, based on the stoichiometry experiments. Analysis of DNA binding by CII in terms of this scheme and its associated binding equation (equation 4) allows further analysis of the details of the interaction of CII with its binding site.
Figure 3. **CII binding to** $p_E$ **wt DNA.**

a) Results of a gel mobility shift assay of CII binding to the $^{32}$P end labeled 48 base pair $p_E$ wt duplex. CII concentrations ranged from 3.6 nM to 5.3 µM. The positions of the free and protein bound DNA $p_E$ are indicated.

b) Binding data from three separate gel shift assays. The data (solid circles) are plotted as the fraction of DNA bound as a function of total CII concentration. The line shows the best fit to the whole data set according to CII self association model 1, a monomer-dimer-tetramer association where $K_1 = 5 \times 10^4$ M$^{-1}$, $K_2 = 1 \times 10^6$ M$^{-1}$ and $K_3 = 1 \times 10^5$ (Table 2). $K_6$ and $K_7$ were the fitted parameters. Clearly, this model does not provide a satisfactory fit to the data.

c) The same gel shift data (now plotted in terms of CII dimer units) were fitted to CII self association model 2, a dimer-tetramer association where the monomer-dimer association (described by $K_1$) is essentially complete over this concentration range. The value of $K_2$, the dimer-tetramer association constant ($8.7 \times 10^5$ M$^{-1}$) was that obtained from sedimentation equilibrium data, with $K_6$ and $K_7$ as the fitted parameters. The solid line shows the best fit of the data ($K_6 = 7.1 (\pm 1) \times 10^7$ M$^{-1}$, $K_7 = 0$) to this model. The effects of $K_7$ (CII tetramer binding to DNA) on the binding curves were calculated. The central dashed line represents the predicted binding curve for the case where $K_7 = 10^9$ M$^{-1}$, while the leftmost dashed line represents the predicted curve for $K_7 = 10^{10}$ M$^{-1}$. Values of $K_7$ less than $10^8$ M$^{-1}$ gave curves indistinguishable from those calculated using $K_7 = 0$.

Figure 4. **DNase I footprint at** $p_E$ **wt and** $p_E$ **KS54.**
A prediction of the K1K6 pathway, wherein CII binds to DNA as a preformed dimer, is that both CII half sites should always be occupied by CII, even when binding affinity for one half site has been reduced by mutation. DNase I footprints at wild type pE, and pE KS54, in which the 6th base of the right half site (asterisk) has been mutated are shown. Double stranded DNA fragments for footprinting (-120 to +434 of pE) were generated by PCR using 32P labeled primer 169 and biotin labeled primer 34. One µL of the appropriate phage stock was used as the template for PCR. The PCR products were purified using streptavidin coated magnetic beads and DNase I reactions performed as described in experimental procedures. The CII concentration ranges used were 2.5 nM to 250 nM for pE wt and 75 nM to 3.9 µM for pE KS54. The seven base pair half sites are indicated by solid lines.

Figure. 5 Effect of half site spacing on CII binding.
The effect on CII binding of spacing between operator half sites was studied using oligonucleotide duplexes designed to vary the centre to centre spacing of the CII binding sites. (a) The sequences of the DNA fragments used for mobility shift experiments are shown. Bases shown in bold indicate the CII binding half sites. The base pair spacing from the centre of the left half site to the centre of the right half site is indicated. (b) DNase I footprints on the pE 19 and pE 21 templates. The fragments for footprint analysis was generated by ligating the pE 19 or pE 21 fragments into the EcoRV site of pBluescript SK (Stratagene) and using this construct as a template for a PCR reaction with a 32P labeled SK primer and a biotin labeled RSP primer. Fragment preparation, CII binding reactions and DNase I treatment were as described in experimental procedures. The final concentration of CII in the reactions was 6 µM. Lanes labeled ACGT represent dideoxy
sequencing reactions performed on the $p_E$ 19 template using labeled SK primer.

Figure 6. **RNA polymerase and CII are sufficient to activate $p_E$.**

Results of an *in vitro* transcription run off assay using as a template a polyacrylamide gel purified 252 bp PCR product containing the $p_E$ promoter sequence (-130 to +122 of $p_E$), generated from a wild type phage stock using PCR primers 168 and 169. Run off reaction mixtures (20 µL) contained 150 nM RNA polymerase and either 0 or 250 nM CII. Following a 15 minute incubation at 37 °C, transcription was initiated by the addition of template DNA (5 nM final concentration). Reactions were allowed to proceed for 40 minutes. The positions of *Hpa*II digested pUC19 DNA size markers are indicated to the left of the gel. The 122 nucleotide $p_E$ specific run off product is indicated, as is the larger end to end transcript.

Figure 7. **RNA polymerase and CII binding at $p_E$.**

DNase I footprints of the $p_E$ region of 186 on the top and bottom strands were performed using the solid phase method (24). DNA templates for footprinting were generated by PCR from a wild type 186 phage stock. The top strand template was produced using $^{32}$P 5’ end-labeled primer 174 and biotinylated primer CII Bam biotin, while the bottom strand template was produced using $^{32}$P labeled primer 169 and biotinylated primer 34 biotin. RNA polymerase where present was at a final concentration of 150 nM, while CII where present was at 250 nM. Lanes labeled ACGT represent dideoxy sequencing reactions of each strand generated using the corresponding $^{32}$P labeled primer. The regions protected from cleavage by CII are indicated to the left of each gel as a solid line, while regions protected by RNA polymerase are indicated by a dashed line. Sequence
positions are shown relative to +1 of $p_E$. In some images, the sequencing tracks have been contrast adjusted differently from the footprint tracks.

Figure 8. **Mutants defective for activation of $p_E$.**

Phage mutants deficient in $p_E$ activity were obtained by plating 186 (wild type or mutagenised) on a CII overexpressing strain (HMS174 pLH3 pLysS) and selecting turbid plaque formers. These plaques were picked, phage purified by replating and the $p_E$ region sequenced using primer 168 or 169. The sequence of the $p_E$ promoter and surrounding region is shown, with the sequence changes of the $p_E$ deficient mutants indicated. The inverted repeats of the CII binding site (ATGTTTG) are indicated with arrows, the -10 of $p_E$ is boxed, the ribosome binding site of the $cII$ gene is overlined, and the initiation codon of the $cII$ gene is in bold. The expected position of the $-35$ box is underlined. The transcriptional activation strengths and CII binding affinities of these mutants are shown in Table 3. The $cIV$ mutant was isolated in a previous study (12).
## Table 1. Determination of CII stoichiometry by sedimentation equilibrium.

<table>
<thead>
<tr>
<th>Samplea</th>
<th>Loading concentrations</th>
<th>Conditionsb</th>
<th>Mwapp c</th>
<th>Mw predictedd</th>
</tr>
</thead>
<tbody>
<tr>
<td>pE wt-F</td>
<td>3 µM DNA, 20,000, 25,000rpm, 20000rpm, 25000rpm, 490nm</td>
<td>30,000</td>
<td>167</td>
<td>31,540</td>
</tr>
<tr>
<td>pE wt-F + CII</td>
<td>3 µM DNA, 2 µM CII, 5 oC, 5,000, 10,000rpm, 490nm</td>
<td>78,890</td>
<td>1910</td>
<td>68,928 (dimer)</td>
</tr>
<tr>
<td>ApI control</td>
<td>0.5 µM DNA, 20 oC, 20,000, 25,000rpm, 260nm</td>
<td>21,170</td>
<td>182</td>
<td>22,000</td>
</tr>
<tr>
<td>ApI control + CII</td>
<td>0.5 µM DNA, 4 µM CII, 5 oC, 5,000, 10,000rpm, 260nm</td>
<td>22,250</td>
<td>1021</td>
<td>22,000 (if no binding)</td>
</tr>
</tbody>
</table>

a For sequence of oligonucleotides, see experimental procedures.

b Temperature, rotor speeds and detection wavelength.

c Values of Mwapp were obtained by fitting sedimentation data to equation 2. Errors are estimates obtained from the fitting procedure.

d Values indicate molecular weights calculated on basis of sequence.
Table 2. **Analysis of CII-DNA binding data**

<table>
<thead>
<tr>
<th></th>
<th>K₁ (M⁻¹)</th>
<th>K₂(M⁻¹)</th>
<th>K₃(M⁻¹)</th>
<th>K₆(M⁻¹)b</th>
<th>K₇(M⁻¹)b</th>
<th>SSRc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1d</td>
<td>5 x 10⁴</td>
<td>1 x 10⁶</td>
<td>1 x 10⁵ f</td>
<td>2.3(±0.1) x 10¹⁰</td>
<td>g</td>
<td>0.734</td>
</tr>
<tr>
<td>Model 2e</td>
<td>∞</td>
<td>8.7 x 10⁵</td>
<td>—</td>
<td>7.1 (±1) x 10⁷</td>
<td>g</td>
<td>0.106</td>
</tr>
</tbody>
</table>

a Based on the two best fits to the sedimentation equilibrium data of Neufing et al.(10).

b Fitted parameters.

c Sum of squares of residuals.

d Values given in terms of CII monomers.

e Values given in terms of CII dimers.

f An upper limit based on gel shift assays with an isolated pE half site.

g Fitted values for K₇ were ill defined with the errors in the fitted values larger than the value itself.
Table 3. Characterisation of $p_E$ deficient mutants.

<table>
<thead>
<tr>
<th>Mutant</th>
<th>% wild type activity $a$</th>
<th>$K_d$ (nM) $b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>wild type</td>
<td>100</td>
<td>11 (±2)</td>
</tr>
<tr>
<td>KS11</td>
<td>0.05</td>
<td>10 (±2)</td>
</tr>
<tr>
<td>KS61</td>
<td>0.5</td>
<td>9 (±1)</td>
</tr>
<tr>
<td>KS5</td>
<td>5</td>
<td>6 (±2)</td>
</tr>
<tr>
<td>KS9</td>
<td>1</td>
<td>10 (±2)</td>
</tr>
<tr>
<td>KS10</td>
<td>3</td>
<td>7 (±1)</td>
</tr>
<tr>
<td>KS55</td>
<td>2</td>
<td>$\uparrow$ $c$</td>
</tr>
<tr>
<td>cIV 476</td>
<td>43</td>
<td>&gt;3,200</td>
</tr>
<tr>
<td>KS1</td>
<td>41</td>
<td>&gt;3,200</td>
</tr>
<tr>
<td>KS2</td>
<td>36</td>
<td>&gt;3,200</td>
</tr>
<tr>
<td>KS6</td>
<td>7</td>
<td>&gt;3,200</td>
</tr>
<tr>
<td>KS23</td>
<td>23</td>
<td>&gt;3,200</td>
</tr>
<tr>
<td>KS24</td>
<td>39</td>
<td>&gt;3,200</td>
</tr>
<tr>
<td>KS54</td>
<td>66</td>
<td>&gt;3,200</td>
</tr>
</tbody>
</table>

$a$ Assayed using NK7049 $\lambda$ $p_E$ -lacZ as described in experimental procedures. CII was supplied from pPN72. The estimated error in the lacZ assays was ±15%. Wild type $p_E$ activity was 1,288 units.

$b$ Determined by gel shift assay from at least two experiments. Dissociation constants are in terms of dimer concentration. > 3 200 indicates no detectable binding at 3.2 µM CII.

$c$ A binding constant could not be calculated for this mutant due to increased dissociation during electrophoresis.
(a) **Wild type**

20 bp

AGGATGCCAACATGTTTGATTTTCAGGTTTTCCAAACATCCCACTATGA
TCCTACGGTGTACAACCTTAAACTCCAAAGGTTTGTAAGGGGTGATCT

10 bp

AGGATGCCAACATGTTTGATTTTCAGGTTTTCCAAACATCCCACTATGA
TCCTACGGTGTACAACCTTAAACTCCAAAGGTTTGTAAGGGGTGATCT

15 bp

AGGATGCCAACATGTTTGATTTTCAGGTTTTCCAAACATCCCACTATGA
TCCTACGGTGTACAACCTTAAACTCCAAAGGTTTGTAAGGGGTGATCT

19 bp

AGGATGCCAACATGTTTGATTTTCAGGTTTTCCAAACATCCCACTATGA
TCCTACGGTGTACAACCTTAAACTCCAAAGGTTTGTAAGGGGTGATCT

21 bp

AGGATGCCAACATGTTTGATTTTCAGGTTTTCCAAACATCCCACTATGA
TCCTACGGTGTACAACCTTAAACTCCAAAGGTTTGTAAGGGGTGATCT

30 bp

AGGATGCCAACATGTTTGATTTTCAGGTTTTCCAAACATCCCACTATGA
TCCTACGGTGTACAACCTTAAACTCCAAAGGTTTGTAAGGGGTGATCT

(b) $\rho_{E19}$ $\rho_{E21}$

CI\[\ + - + - \ A C G T \]
Establishment of lysogeny in bacteriophage 186 - DNA binding and transcriptional activation by the CI1 protein
Keith E. Shearwin and J. Barry Egan

J. Biol. Chem.  published online June 27, 2000

Access the most updated version of this article at doi: 10.1074/jbc.M004574200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts