Purification and Self-association Equilibria of the Lysis-Lysogeny Switch Proteins of Coliphage 186*

(Received for publication, November 22, 1995, and in revised form January 25, 1996)

Keith E. Shearwin and J. Barry Egan†

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

The CI repressor protein, responsible for maintenance of the lysogenic state, and the Apl protein, required for efficient prophage induction, are the two control proteins of the lysis-lysogeny transcriptional switch of coliphage 186. These proteins have been overexpressed, purified, and their self-association behavior examined by sedimentation equilibrium. Phage 186 CI dimers self-associate in solution through tetramers to octamers in a concerted process. The Apl protein of 186 is monomeric in solution.

Coliphage λ has served as a model system for describing mechanisms of gene control in both prokaryotes and higher organisms. In particular, the means by which the λ CI and Cro proteins interact with their operator sites to control transcription, and so foster either lysogenic or lytic development, has been studied extensively by a range of genetic, biochemical, and physiochemical approaches. These studies have contributed enormously to our understanding of genetic control mechanisms (1). Bacteriophage 186 from the P2 family of phages has a completely different nucleotide sequence to λ and has evolved a different set of mechanisms for controlling expression of its genome. Coliphage 186, like λ, is able to replicate its genome through one of two independent but interchangeable pathways. The lytic pathway results in lysis of the host cell and production of progeny phage, while the lysogenic pathway involves integration of the phage genome into the chromosome of the bacterial host where it is replicated along with the host chromosome in subsequent generations. The lysogenic state in 186 is an extremely stable one: the frequency of uninduced transition from lysogeny to lytic development approaches the mutation rate (2). Despite the stability of the lysogenic state, the activation from lysogeny to lytic development approaches the mutation rate (2). Hence, as a first step in understanding the molecular mechanisms by which CI is able to repress transcription from pR and pL, and thereby efficiently maintain the lysogenic state, we have examined the ability of CI and Apl to self-associate and found that in solution CI higher order self-assembly proceeds in a concerted manner from dimer through tetramer to octamer, while Apl remains monomeric over the concentration range examined.

MATERIALS AND METHODS

Radiolabeled nucleotides, acrylamide solutions, and oligonucleotide primers were purchased from Bresatec (Adelaide), while restriction enzymes were from New England Biolabs. All chemicals were of reagent grade or better.

Cloning and Expression

The d gene from 186 was amplified by the polymerase chain reaction using primers designed to introduce an Nde restriction site (underlined) at the 5′-end of the gene (5′-GGTTTTATCCATATGAGAATA 3′) and a BamHI restriction site (underlined) at the 3′-end of the gene (5′-CAGGGATCCAGCGCCGAGCGC 3′). This fragment was ligated into a pET3a vector backbone (10) to give pET3aCI, which was then transformed into Escherichia coli strain BL21 (DE3) pLYS5 (10) and the insert was sequenced to ensure no base pair changes had been introduced (constructed by I. Dodd). The Apl protein was also overexpressed using the pET system. The apl gene was ligated into pET3a to give pET3aApl and transformed into BL21 (DE3) pLYS5.

For expression of CI, BL21 (DE3) pLYS5 pET3aCI cells were grown at 37 °C in 2-liter flasks containing Luria Broth (500 ml), 100 μg ml⁻¹ carbenicillin, and 30 μg ml⁻¹ chloramphenicol. When the culture had reached an optical density of 0.6–0.8 at 600 nm, isopropyl-1-thio-β-D-galactopyranoside was added to a final concentration of 0.4 mm, and unknown, one (FL), located within the d gene, the other (FR), found at the 5′-end of the apl gene (7). The Apl protein of 186, produced from the first gene of the rightward early lytic transcript, has no apparent role in lytic development after infection but is required for efficient prophage induction (6). It functions both at the level of derepression and of prophage excision (6).1

1 M. R. Reed, K. E. Shearwin and J. B. Egan, manuscript in preparation.
growth was continued for an additional 3 h. Cells were collected by centrifugation, washed once with 50 mM Tris-HCl, 0.1 mM EDTA, 10% glycerol, 150 mM NaCl, pH 7.5 (TEG 150) buffer, and stored at -70 °C in approximately 10 ml of the same buffer. Expression of Apl was the same as that described for CI, except that growth was at 30 °C in order to maximize the fraction of soluble protein.

Protein Purification

CI—All steps in the purification of CI were performed at 4 °C. Cells (from 2 × 500-ml cultures) were thawed, and phenylmethylsulfonyl fluoride was added immediately to a final concentration of 0.1 mM, in order to minimize proteolytic activity. The cell suspension, already partially lysed as a result of the activity of lysozyme produced from the cell wall by the cell lysis procedure, was sonicated for 5 min. The mixture was centrifuged (10 min, 300 × g) and the supernatant was discarded. CI was released from the PEI pellet by resuspension in TEG buffer containing 500 mM NaCl. The mixture was centrifuged again and the supernatant was retained. This high salt extract was brought to 60% (NH₄)₂SO₄ saturation by the addition of solid (NH₄)₂SO₄. Following stirring for 30 min, precipitated proteins were removed by a brief centrifugation and solid (NH₄)₂SO₄ was added, with stirring, to give a final concentration of 0.5% (v/v). Precipitated nucleic acids and unwanted proteins were removed by brief centrifugation and solid (NH₄)₂SO₄ was added (3–7). Sequencenumberingbeginsatthe

Fig. 1. Organization of the major control region of 186. The map of the early region of 186 from the PstI site (65.5%) to the BssHI site (76.8%) is shown (3–7). Sequence numbering begins at the PstI site at 65.5%. Genes are shown as boxes (rightward genes above the line, leftward genes below), promoters as arrowheads, their transcripts as arrows, and terminators as stem-loops. p₁ is the lyogenic promoter, p₂ is the early lytic promoter, and p₃ is the promoter for the B gene, the product of which activates transcription of late genes. CI is the gene required for establishment of lysogeny, int is the integrase, and 69 is of unknown function. The phage attachment site for integration into the host chromosome, attP (Reed et al.) is shown. CI binding sites (7) are indicated by solid boxes, while Apl binding sites are shown as cross-hatched boxes. The Maed 2666 to Maed II 2668 (switch) region is enlarged in order to present the relative arrangement of the Apl and CI binding sites. The -10 and -35 regions of the p₁, p₂, and p₃ promoters are indicated, as are the start sites for transcription (+1).

Analytical Ultracentrifugation

All experiments were performed on a Beckman Optima XL-A analytical ultracentrifuge equipped with absorbance optics and a four-hole An60Ti rotor. Sedimentation equilibrium experiments were done at 5 °C using double sector centerpieces. Data were collected at 230 or 280 nm with a spading of 0.001 cm as the average of three absorbance readings (12), and, after 24 h of centrifugation, scans were compared at 3-h intervals to ensure that equilibrium had been reached. All experiments were done in 50 mM Tris-HCl, 0.1 mM EDTA, 150 mM NaCl, 10% (v/v) glycerol, 50 ng ml⁻¹ sonicated salmon sperm DNA, and the reactions (10 µl) were incubated for 30 min on ice. The binding reactions, containing -150 rpm, were loaded on non-denaturing 6% polyacrylamide gels (39.1 acrylamide/bisacrylamide) containing 10% glycerol. A separate lane of tracking dye was used. Gels were pre-electrophoresed at 4 °C for at least 30 min prior to loading the samples. Electrophoresis was carried out for approximately 2 h, the gels were dried under vacuum, and the distribution of labeled DNA was recorded on a phosphor screen. The phosphor screen was analyzed with the Imagequant program on a Molecular Dynamics PhosphorImager.

Substitution into Equation 1 and rearrangement gives, for i = 2,

\[
C_{r_i} = \left( \frac{\exp \left[ \ln c_i + AM_i \left( \frac{r_i^2 - r_1^2}{2} \right) \right]}{2} \right) + e \tag{Eq. 3}
\]
where equilibrium constants are fitted as $\ln K$ to constrain them to positive values. Additional terms (e.g. $i = 4$ for tetramer) can be added to Equation 3 for more complex association schemes. Hence, on the basis of data obtained from the ultracentrifuge (total concentration, on an absorbance scale, as a function of radial distance), equilibrium constants describing a given association scheme can be obtained by fitting the data to Equation 3. These fitted values of $K$ are then converted from an absorbance scale to a molar scale, based on the degree of polymerization and the appropriate extinction coefficient, $\varepsilon$ (corrected for the pathlength of the centerpiece).

$$K(M)^{-1} = K(Abs)^{-1} \left(\frac{M}{l}\right)^{-1} \varepsilon^{-1}$$  

(Eq. 4)

Data sets used in the analysis were truncated to include only absorbance values below 1.2, to ensure absorbance is linear with respect to protein concentration. Nonideality was not considered. Data analysis was done using a commercial graphics/curve-fitting program (Sigmaplot 5.1, Jandel Scientific, Corte Madera, CA) or by NONLIN (13). The partial specific volumes of CI and Apl were calculated using the amino acid partial specific volume values of Zamyatnin (14). These calculations gave $\rho = 0.725$ ml g$^{-1}$ for CI and $\rho = 0.735$ ml g$^{-1}$ for Apl. Buffer density at 5°C was measured in an Anton-Paar precision density meter to be 1.0378 g ml$^{-1}$.

**RESULTS**

Protein Purification—The $cI$ and $apl$ genes of the temperate coliphage 186 were cloned into the T7 expression system of *E. coli*. Expression of the proteins had little effect on the growth of the host cells, and almost all of the protein was in a soluble form.

SDS-PAGE of samples taken at various stages of the purification of CI are shown in Fig. 2A. In purifying CI, PEI precipitation results in the coprecipitation of CI with nucleic acids. CI is then efficiently separated from the nucleic acids by extracting the PEI pellet with a buffer of moderate salt concentration. Ammonium sulfate precipitation serves to separate CI from any remaining PEI and some of the protein contaminants. CI bound strongly to the Affi-Gel Blue column and was eluted as a broad peak to give a protein pool containing only a few contaminants. These were removed by chromatography on a heparin affinity column. An additional chromatography step on a Superdex 75 HR 10/30 fast protein liquid chromatography gel filtration column showed no evidence of contaminating species (not shown).

Fig. 2B shows SDS-PAGE of samples taken during the purification of Apl. As was the case for CI purification, PEI precipitation was used to remove nucleic acids and some contaminating proteins. However, Apl did not coprecipitate with the nucleic acids but remained in the supernatant, behavior consistent with its predicted isoelectric point of 10. Ammonium sulfate precipitation was used to concentrate the protein and separate it from PEI for the gel filtration procedure. The small size of the Apl protein (9.6 kDa) allowed purification in a single step on Sephacryl S200, under denaturing conditions. With conservative pooling of Apl-containing fractions, Apl was obtained with >95% purity. Refolding was performed by dialysis against progressively lower concentrations of urea. Since Apl contains only a single cysteine (3), refolding was straightforward and very little precipitate was observed.

The UV spectra of purified CI and Apl are typical of those obtained for tryptophan-containing proteins (not shown). The $A_{280}/A_{260}$ ratios are 1.69 for CI and 1.84 for Apl, indicating the absence of significant quantities of contaminating nucleic acids.

Activity—Gel mobility shift assays were used to follow the purification of CI and Apl by quantitating the ability of the proteins to bind specifically to their operator sites. A $^{32}$P-labeled 437-bp DNA fragment containing the $p_{A22}$ switch region (5) was used as the target for binding. This region contains binding sites for both CI and Apl (Fig. 1). One unit of activity was defined as the dilution at which 50% of the labeled DNA is retarded, under the standard conditions of the assay. Tables I and II show the yields and activities obtained throughout the purification processes. Assays could not be performed on PEI-containing fractions as this led to precipitation of the labeled DNA fragment. For CI, a 13-fold purification is achieved, with a yield of 0.7 mg of protein per liter of culture, while Apl underwent a 9-fold purification for a yield of approximately 1 mg/liter of culture.

**Table I**

| Step          | Volume | Total protein | Total activity$^a$ | % activity recovered | Specific activity | Purification |
|---------------|--------|--------------|--------------------|---------------------|------------------|-------------|
| Lysate        | 21 ml  | 130 mg       | $2.9 \times 10^6$  | 100                 | $2.2 \times 10^4$ | 1.0         |
| PEI/(NH$_4$)$_2$SO$_4$ | 10 ml  | 23.5 mg      | $1.1 \times 10^6$  | 38                  | $4.7 \times 10^4$ | 2.1         |
| Affi-Gel Blue | 28 ml  | 2.8 mg       | $3.4 \times 10^6$  | 12                  | $1.2 \times 10^5$ | 5.5         |
| Heparin       | 7.5 ml | 0.7 mg       | $2.0 \times 10^6$  | 7                   | $2.8 \times 10^4$ | 12.7        |

$^a$ Measured by gel mobility shift assay; arbitrary units defined as described in the text.
Self-association—On SDS-PAGE (Fig. 2), CI migrates as a single molecular species with an apparent molecular weight of 24,000, in reasonable agreement with the molecular weight (21,160) based on its gene sequence. Apl migrates with an apparent molecular weight of 10,000, consistent with its predicted molecular weight of 9652.

In order to gain a qualitative estimate of the extent of any self-association, purified CI and Apl were subjected to gel filtration chromatography on a column of Sephacryl S200. By calibrating the column with a series of proteins of known molecular weight, the elution volume of the protein of interest can be used to infer an apparent molecular weight. It should be emphasized that for small zone experiments, the technique is not used to infer an apparent molecular weight. It should be noted that CI self-associates and that the self-association is of species larger than octamer. Fitting the data sets into the global fit did not justify inclusion of additional species in the association scheme (dimer-tetramer-hexamer-octamer) did not improve the fit above that expected solely on the basis of fitting to an additional parameter.

The sedimentation data sets were then analyzed globally in terms of various assembly schemes. Insufficient data could be obtained at the low concentration end (even when distributions were recorded at 230 nm) to satisfactorily describe the monomeric interaction and so, in subsequent analyses, the smallest species in the association scheme was set at that of a dimer (M_r = 42,320). When constraining M_1 (Equation 3) to this value, the best fit to the data, as judged by the sum of squares of residuals (SSR), was to a dimer-tetramer-octamer equilibrium (ΔG^o_2,4 = −7.0 ± 0.1 kcal mol⁻¹, ΔG^o_2,8 = −21.3 ± 0.6 kcal mol⁻¹, SSR = 0.007). Consistent with this result, when the molecular weight of the smallest species (M_1) was allowed to float in the calculation, essentially the same free energies of association were obtained and the resulting fitted value for M_1 was, within experimental error, that of a CI dimer (43,390 ± 1340). Including additional species in the association scheme (dimer-tetramer-hexamer-octamer) did not improve the fit above that expected solely on the basis of fitting to an additional parameter.

Short pathlength cells (2.5 mm) were used with a higher loading concentration of CI (20 μM) in order to better define the high end of the association scheme. However, inclusion of these data sets into the global fit did not justify inclusion of additional species larger than octamer. Fitting the data sets to a dimer-octamer equilibrium (i.e. no formation of tetramer, K_2,4 = 0) resulted in a fit similar to that obtained for the dimer-tetramer-octamer scheme, reflecting the difficulty in defining the association constant for species which do not accumulate to a significant degree. Hence, the self-association of the 186 CI repressor protein under the conditions studied is best described by a dimer-tetramer-octamer equilibrium, the dimer-octamer transition being a concerted process (see "Discussion").

Apl self-association was also studied by sedimentation equilibrium. Four runs were performed, employing two different rotor speeds (16,000 rpm and 24,000 rpm) and two different loading concentrations (16 μM and 32 μM) (Fig. 5). Data sets from all four runs were fitted globally to various self-association schemes. Insufficient data could be obtained at the low concentration end (even when distributions were recorded at 230 nm) to satisfactorily describe the monomeric interaction and so, in subsequent analyses, the smallest species in the association scheme was set at that of a dimer (M_r = 42,320). When constraining M_1 (Equation 3) to this value, the best fit to the data, as judged by the sum of squares of residuals (SSR), was to a dimer-tetramer-octamer equilibrium (ΔG^o_2,4 = −7.0 ± 0.1 kcal mol⁻¹, ΔG^o_2,8 = −21.3 ± 0.6 kcal mol⁻¹, SSR = 0.007). Consistent with this result, when the molecular weight of the smallest species (M_1) was allowed to float in the calculation, essentially the same free energies of association were obtained and the resulting fitted value for M_1 was, within experimental error, that of a CI dimer (43,390 ± 1340). Including additional species in the association scheme (dimer-tetramer-hexamer-octamer) did not improve the fit above that expected solely on the basis of fitting to an additional parameter.

The ability of 186 CI and Apl to self-associate in solution was investigated in a quantitative manner by sedimentation equilibrium. All experiments were performed at 5 °C in TEG 150 buffer. For both proteins, different combinations of loading concentration and rotor speed were used. Fig. 4 shows the results of three of the sedimentation equilibrium experiments performed on CI with a loading concentration of 9 μM (in terms of total repressor subunits) at rotor speeds of 12,000, 16,000 and 24,000 rpm. Initially, the individual concentration distributions were fitted to the equation for a single species (i = 1) in order to obtain whole cell average molecular weights. These ranged from 91,000 (±1000) to 144,000 (±2300). Thus, the greatest whole cell average molecular weight is approximately 6.8 times that of the monomeric species. These results confirm that CI self-associates and that the self-association is of species larger than hexamer.

The sedimentation data sets were then analyzed globally in terms of various assembly schemes. Insufficient data could be obtained at the low concentration end (even when distributions were recorded at 230 nm) to satisfactorily describe the monomeric interaction and so, in subsequent analyses, the smallest species in the association scheme was set at that of a dimer (M_r = 42,320). When constraining M_1 (Equation 3) to this value, the best fit to the data, as judged by the sum of squares of residuals (SSR), was to a dimer-tetramer-octamer equilibrium (ΔG^o_2,4 = −7.0 ± 0.1 kcal mol⁻¹, ΔG^o_2,8 = −21.3 ± 0.6 kcal mol⁻¹, SSR = 0.007). Consistent with this result, when the molecular weight of the smallest species (M_1) was allowed to float in the calculation, essentially the same free energies of association were obtained and the resulting fitted value for M_1 was, within experimental error, that of a CI dimer (43,390 ± 1340). Including additional species in the association scheme (dimer-tetramer-hexamer-octamer) did not improve the fit above that expected solely on the basis of fitting to an additional parameter.

Short pathlength cells (2.5 mm) were used with a higher loading concentration of CI (20 μM) in order to better define the high end of the association scheme. However, inclusion of these data sets into the global fit did not justify inclusion of additional species larger than octamer. Fitting the data sets to a dimer-octamer equilibrium (i.e. no formation of tetramer, K_2,4 = 0) resulted in a fit similar to that obtained for the dimer-tetramer-octamer scheme, reflecting the difficulty in defining the association constant for species which do not accumulate to a significant degree. Hence, the self-association of the 186 CI repressor protein under the conditions studied is best described by a dimer-tetramer-octamer equilibrium, the dimer-octamer transition being a concerted process (see "Discussion").

Apl self-association was also studied by sedimentation equilibrium. Four runs were performed, employing two different rotor speeds (16,000 rpm and 24,000 rpm) and two different loading concentrations (16 μM and 32 μM) (Fig. 5). Data sets from all four runs were fitted globally to various self-associa-
Fig. 4. Sedimentation equilibrium of 186 CI at 5 °C. Experiments were performed at 12,000 rpm (open squares), 20,000 rpm (solid circles), and 24,000 rpm (open circles). The data are shown as the concentration distribution as a function of radial distance. The lines represent the best fit of the data to a dimer-tetramer-octamer association scheme, while the lower residuals plot shows the difference between the experimental and fitted values. For clarity, only every third data point is shown.

Fig. 5. Sedimentation equilibrium of 186 Apl at 5 °C. Experiments were performed at 16,000 rpm (solid symbols) and 24,000 rpm (open symbols) with loading concentrations of 16 μM (circles) or 32 μM (squares). The data are presented as concentration distributions as a function of radial distance. The lines represent the best fit of the data according to a single monomeric species, while the lower residuals plot presents the difference between the experimental data and the fitted values. For clarity, only every third data point is shown.

Fig. 6. Distribution of CI oligomeric species as a function of total CI concentration, in total repressor subunits. The weight fraction of each species present was calculated using the free energies of association obtained from the sedimentation equilibrium experiments (ΔG_{s,CI}^0 = -7.0 kcal mol^{-1}, ΔG_{o,CI}^0 = -21.3 kcal mol^{-1}). A value of -10.2 kcal mol^{-1} was used for the monomer-dimer equilibrium (ΔG_{m,d}^0, see “Discussion”).

DISCUSSION

Two of the major control proteins from bacteriophage 186, CI and Apl, have been expressed, purified, and their self-association properties examined. CI reversibly self-associates in solution, and this self-assembly is best described by a monomer-dimer-tetramer-octamer equilibrium. Fig. 6 shows the distribution of CI species calculated from the free energies of association obtained from fitting the sedimentation equilibrium data to a dimer-tetramer-octamer association scheme. Even at the lowest concentration of CI used in the sedimentation equilibrium experiments, there was insufficient monomer present to characterize the monomer-dimer equilibrium. In order to permit inclusion of the calculated distribution of monomer and dimer in Fig. 6, a value of K_{1,2} has been estimated. This estimate, 1 × 10^8 M^{-1} (ΔG_{1,2}^0 = -10.2 kcal mol^{-1}), is based on the detection limit of the ultracentrifuge; that is, any value of K_{1,2} lower (weaker) than 1 × 10^8 M^{-1} would have been resolved in the fitting procedure. Based on this value of K_{1,2}, CI exists primarily as a mixture of monomers (21 kDa) and dimers (42 kDa) between 10^{-10} and 10^{-7} M (in terms of monomer). Below 10^{-10} M, CI is essentially monomeric. Use of a 10-fold higher value of K_{1,2} results in the monomer-dimer curves shifting one log unit to the left, without significant change to the tetramer and octamer curves (not shown). Between 10^{-6} and 10^{-4} M, CI exists in solution as a mixture of dimer (42 kDa), tetramer (84 kDa), and octamer (168 kDa), with octamer being the predominant species at concentrations above 10 μM. The tetrameric species exists only as an intermediate during the assembly of dimers to octamers, never reaching more than 35% of the total. The distribution of tetramer is subject to some uncertainty given the difficulty in precisely defining K_{2,4}. Within the concentration range examined, there was no evidence for formation of polymers higher than octamers. That octamer formation is a concerted (favored) process can be seen by calculating the free energy per dimer required for formation of the higher species, for tetramer formation, ΔG_{2,4}^0 = -3.5 kcal mol^{-1} per dimer, while for octamer formation, ΔG_{4,8}^0 = -5.3 kcal mol^{-1} per dimer. Thus, the free energy per dimer for octamer formation is more negative than the free energy per dimer of tetramer formation, and assembly of dimers to octamer is the energetically favored process.

Like 186 CI, the λ CI repressor also associates from dimer through tetramer to octamer in a concerted process (16), although at very high protein concentrations (up to 100 μM) the
Purification and Properties of Phage 186 CI and Apl

\( \lambda \) data were consistent with a further association to dodecamer. Not only do the \( \lambda \) and 186 repressors assemble in solution in the same manner, but comparison of the fitted free energy values for the various steps in the association process shows a remarkable similarity. For the dimer-octamer equilibrium, \( \Delta G^0 = -7.1 \text{ kcal mol}^{-1} \) for \( \lambda \) and \(-7.0 \text{ kcal mol}^{-1} \) for 186, while the free energies for the dimer-tetramer equilibrium, \( \Delta G^0 = -23.0 \text{ kcal mol}^{-1} \) for \( \lambda \) CI and \(-21.3 \text{ kcal mol}^{-1} \) for 186 CI. Given this correspondence between the self-association characteristics of the two proteins, one might expect their structures to show some similarity. While the domain structure of the \( \lambda \) repressor has been studied extensively, little is known about the tertiary structure of 186 CI. Several lines of evidence suggest that 186 CI consists of two domains (17). The N-terminal domain is responsible for DNA binding and the C-terminal domain contains the determinants for oligomerization and cooperativity. The N domain interacts with DNA via a helix-turn-helix motif, a motif common to many DNA-binding proteins. In the 186 CI repressor, a region in the N-terminal third of the amino acid sequence gives a weak match to a helix-turn-helix motif as judged by weight matrix analysis (18). We suspect that 186 CI contains a variant form of this DNA binding motif (7). Other than this, there is little similarity at the nucleotide or amino acid level between the \( \lambda \) and 186 repressors. The nature of the operator sites to which the two repressors bind do share some features. The repressors bind to three operators in both \( \lambda \) at \( O_4, O_5, 2, \) and \( O_6, 3 \) and 186 at \( p_6 \) (site II, site I, and site III), each operator being separated by approximately two turns of the helix (1, 7). In 186, however, the central operator (site I) has a consensus sequence unrelated to that of the other adjacent operators, indicating that 186 CI may contain more than one region capable of binding DNA (7). DNase I footprint analysis of mutated 186 CI operator regions show evidence of cooperativity (7).

What are the implications of these results for 186 CI binding to its operator sites? Protein-protein association in solution to provide multidentate ligands capable of binding multiple sites on DNA is a common mechanism for cooperative binding of regulatory proteins to DNA (16). Thus, in order to fully characterize a protein-DNA interaction, oligomerization of the protein must be considered. For example, as discussed by Seneer et al. (16), linkage between protein self-assembly and DNA binding may produce free energy changes for oligomerization which will differ depending on whether binding to DNA favors or disfavors protein self-assembly. In the case of \( \lambda \), Laue et al. (19) found that binding of \( O_6, 3 \) oligonucleotides to octameric CI repressor did not dissociate it to tetramers, and, therefore, any model for \( \lambda \) which proposes pairwise cooperativity between adjacent DNA-bound dimers (to give DNA-bound tetramers) must also consider the free energy required to destabilize the octamer. Similarly, in the case of 186, further studies of the protein-protein and protein-DNA interactions are required to delineate the contributions of cooperativity, linkage, and allosteric to the molecular mechanism by which CI binds to its operator sites and stably maintains the lysogenic state.

In considering the ability of CI to maintain the lysogenic state, one must also be aware of the relative arrangement of the lytic and lysogenic promoters (Fig. 1). An inevitable consequence of the overlapping face to face arrangement of the \( p_6 \) and \( p_4 \) operators in 186 is that RNA polymerase, in transcribing CI from \( p_6 \), must traverse \( p_4 \) in order to maintain the lysogenic state. In doing so it must presumably dislodge CI already bound at \( p_4 \), providing the opportunity for loss of repression. A possible mechanism for preventing this loss of repression involves the flanking sites FL and FR providing a locally high concentration of CI, allowing rapid rebinding of CI to \( p_4 \) following passage of RNA polymerase. Such a mechanism would require oligomerization of CI such that it could bind simultaneously to \( p_6 \) and either FL or FR. These studies have demonstrated that this oligomerization can occur, at least in solution.

Turning now to Apl, this protein functions as a repressor during prophage induction and is involved in excision of the prophage from the bacterial host, roles performed by two proteins, Cro and XIS, in \( \lambda \). Apl binds to a set of seven direct repeats at \( p_6/p_6 \) and five direct repeats at the attachment site, attP (Fig. 1; Ref. 5). These sites have 10–11 base pair center to center spacing, indicating that Apl binds to the same face of the helix. Given the narrow range of concentration over which Apl fills the multiple operator sites within the \( p_6/p_6 \) region (6), cooperative interactions must be involved. Apl, like the homologous Cox proteins from P2 and HP1 phage, has a predicted helix-turn-helix DNA binding motif (18, 20). In general, helix-turn-helix proteins are trimers or tetramers in solution (for example, the majority of those listed in (21)) and it is these oligomers which interact with their operator sites, usually (but not always) inverted repeat sequences. In contrast, analytical ultracentrifugation of purified Apl shows that Apl remains monomeric in solution up to millimolar concentrations.

It could be argued that since Apl has been denatured and refolded during the course of the purification, the majority of the protein is in an inactive, nonassociating form and that the activity observed in gel shift assays arises from a small fraction of active associated protein. While we cannot completely rule this out, three lines of evidence argue against this possibility. Firstly, Apl purified to approximately 80% purity by ion exchange chromatography (without unfolding) eluted on a calibrated Sephacryl S200 column (Fig. 3) with the same elution volume as Apl purified by the unfolding/refolding procedure. This indicates that the “native” protein is the same size as the protein purified by unfolding/refolding. Secondly, this partially purified Apl fraction had approximately the same specific activity in gel shift assays as the pure refolded Apl. The yield and extent of purification of Apl (Table II) is consistent with retention of DNA binding activity throughout the purification procedure. Finally, while monomeric helix-turn-helix DNA binding proteins are unusual, there are examples in the literature. Thus, while a Cro is a dimer in solution, Cro protein from phage 434 remains monomeric, even at the high concentrations employed for crystallization (22). Again, the beton operon repressor (BirA) remains monomeric at concentrations 2–3 orders of magnitude higher than the concentration required for operator binding (23). Of particular interest are the Ner proteins of bacteriophage Mu and the closely related phage D108. Like apl, ner is functionally analogous to \( \lambda \) cro in that it is the first gene encoded by the early lytic operon and that its protein product...
undertaking further studies of both CI-DNA interactions and with their operators in the 186 control region. We are currently basemodels describing the interaction of these control proteins it involves both protein-protein contacts and DNA bending. unknown, but, given the periodic enhancements of DNaseI ciation. The structural basis of cooperativity in Apl binding is concentration of 25 a 30-bp oligonucleotide (26), yet is monomeric in solution at a monomers to the DNA (27). Similarly, D108 Ner forms dimers between the two promoters (25, 26). Mu Ner binds, with a dissociation constant in the nanomolar range, as a tetramer on a 30-bp oligonucleotide (26), yet is monomeric in solution at a concentration of 25 µM (27). Similarly, D108 Ner forms dimers on DNA but is monomeric in solution up to 200 µM (28). Although the tertiary structures of the Ner proteins are not known, they are selected by the weight matrix method of Dodd and Egan (18) as being potential helix-turn-helix proteins, albeit with a relatively weak score. Given the monomeric nature of the Apl protein in solution and the fact that Apl is expected to bind cooperatively to its recognition sequences, it follows that this cooperativity can only be mediated on the DNA. This is illustrated in Fig. 7 which shows a simple thermodynamic cycle for a protein (A) binding to multiple sites on DNA. The protein-DNA complex can be formed in two ways. The protein can either self-associate in solution and then bind to DNA (K1K2 pathway) or can bind as monomers to the DNA (K3K4 pathway) where the protein sub-units may or may not interact to give rise to cooperativity. A combination of these pathways is also possible, if both the monomeric and the self-associated forms of the protein have affinity for the DNA. The overall equilibrium will then reflect the relative affinities of the two forms of the protein for the DNA. For Apl, K1, the protein self-association constant in solution is 0. Apl binding therefore can only occur through the K3K4 pathway: DNA binding followed by protein-protein association. The structural basis of cooperativity in Apl binding is unknown, but, given the periodic enhancements of DNaseI cleavage noted in footprint experiments (6), we speculate that it involves both protein-protein contacts and DNA bending. The present results have provided a framework on which to base models describing the interaction of these control proteins with their operators in the 186 control region. We are currently undertaking further studies of both CI-DNA interactions and Apl-DNA interactions in order to dissect at the molecular and energetic levels the mechanisms by which these proteins control the lysis/lysogeny switch in bacteriophage 186.

Acknowledgments—We thank Prof. D. J. Winzor and M. Jacobsen (University of Queensland) for use of the XL-A.

REFERENCES
1. Ptashne, M. (1992) A Genetic Switch, 2nd Ed., Cell Press & Blackwell Scientific Publications, Cambridge, MA
2. Lamont, I., Brumby, A. M., and Egan, J. B. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 5492–5496
3. Kalionis, B., Dodd, I. B., and Egan, J. B. (1986) J. Mol. Biol. 191, 191–209
4. Richardson, H., and Egan, J. B. (1989) J. Mol. Biol. 206, 59–68
5. Dodd, I. B., Kalionis, B., and Egan, J. B. (1990) J. Mol. Biol. 214, 127–127
6. Dodd, I. B., Reed, M. R., and Egan, J. B. (1993) Mol. Microbiol. 10, 1139–1150
7. Dodd, I. B., and Egan, J. B. (1996) J. Biol. Chem. 271, 11332–11340
8. Dibbens, J. A., Gregory, S. L., and Egan, J. B. (1992) Mol. Microbiol. 6, 2643–2650
9. McKnight, S. L., and Yamamoto, K. R. (eds) (1992) Transcriptional Regulation, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
10. Studier, R. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) Methods Enzymol. 185, 60–89
11. Wetlaufer, D. B. (1962) Adv. Protein Chem. 17, 303–390
12. Henner, P. E., Radford, L. E., and Beams, J. W. (1961) Proc. Natl. Acad. Sci. U. S. A. 47, 1848–1852
13. Johnson, M. L., Coreira, J. J., Yphantis, D. A., and Halvorsen, H. R. (1981) Biophys. J. 36, 575–588
14. Zamyatin, A. A. (1984) Annu. Rev. Biophys. Bioeng. 13, 145–165
15. Flinta, C., Persson, B., Jornvall, H., and von Heijne, G. (1986) Eur. J. Biochem. 154, 193–196
16. Senear, D. F., Lave, T. M., Ross, J. B. A., Waxman, E., Eaton, S., and Rusinova, E. (1993) Biochemistry 32, 6179–6189
17. Burz, D. S., and Ackers, G. K. (1994) Biochemistry 33, 8406–8416
18. Dodd, I. B., and Egan, J. B. (1990) Nucleic Acids Res. 18, 5025–5026
19. Lave, T. M., Senear, D. F., Eaton, S., and Ross, J. B. A. (1993) Biochemistry 32, 2469–2472
20. Esposito, D., and Scocca, J. J. (1994) Microbiol. 13, 685–695
21. Brown, R., and Matthews, B. W. (1989) J. Biol. Chem. 264, 1903–1906
22. Harrison, S. C. (1992) in Transcriptional Regulation (McKnight, S. L., and Yamamoto, K. R., eds) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
23. Abbott, J., and Beckett, D. (1993) Biochemistry 32, 9649–9658
24. Gossen, N., and van de Putte, P. (1986) J. Bacteriol. 167, 503–507
25. Kukolj, G., and DuBow, M. S. (1989) J. Biol. Chem. 264, 17827–17835
26. Strzelecka, T. E., Hayes, J. J., Clore, G. M., and Gronenborn, A. M. (1995) Biochemistry 34, 2946–2955
27. Allet, B., Payton, M., Mattaliano, R. J., Gronenborn, A. M., Clore, G. M., and Wingfield, P. T. (1988) Gene (Amst.) 65, 259–268
28. Kukolj, G., Tolas, P. P., Auxtert, C., and DuBow, M. S. (1989) EMBO J. 8, 3141–3148
Purification and Self-association Equilibria of the Lysis-Lysogeny Switch Proteins of Coliphage 186
Keith E. Shearwin and J. Barry Egan

J. Biol. Chem. 1996, 271:11525-11531.
doi: 10.1074/jbc.271.19.11525

Access the most updated version of this article at http://www.jbc.org/content/271/19/11525

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

This article cites 25 references, 6 of which can be accessed free at http://www.jbc.org/content/271/19/11525.full.html#ref-list-1