Interaction of the *Escherichia coli* Gal Repressor Protein with Its DNA Operators in Vitro†

Michael Brenowitz,*,‡ Elizabeth Jamison,‡ Alokes Majumdar,§ and Sankar Adhya§

Department of Biochemistry, The Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461, and Laboratory of Molecular Biology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892

Received September 27, 1989; Revised Manuscript Received November 28, 1989

ABSTRACT: The binding of *Escherichia coli* Gal repressor to linear DNA fragments containing two binding sites (O_E and O_I) within the *gal* operon was analyzed in vitro with quantitative footprint and mobility-shift techniques. In vivo analysis of the regulation of the *gal* operon [Haber, R., & Adhya, S. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 9683-9687] has suggested the role of a regulatory "looped complex" mediated by the association of Gal repressor dimers bound at O_E and O_I. The binding of Gal repressor to a single site can be described by a model in which monomer and dimer are in equilibrium and only the dimer binds to DNA. At pH 7.0, 25 mM KCl, and 20 °C, the binding and dimerization free energies are comparable, suggesting that the equilibrium governing the formation of dimers may be important to regulation. The two intrinsic binding constants, ΔG_I and ΔG_E, and a constant describing cooperativity, ΔG_{IE}, were determined by footprint titration analysis as a function of pH, [KCl], and temperature. Only at 4 and 0 °C was ΔG_{IE} negative, signifying cooperative binding. These results are thought to be due to a weak dimer to tetramer association interface. ΔG_E and ΔG_I had maximal values between pH 6 and pH 7. The dependence of these constants on [KCl] corresponded to the displacement of approximately 2 ion equiv. The temperature dependence could be described by a change in the heat capacity, ΔC_p, of −2.3 kcal mol^{-1} deg^{-1}. Mobility-shift titration experiments conducted at 20 and 0 °C yielded values for ΔG_{IE} that were consistent with those resolved from the footprint analysis. Unique values of ΔG_{IE} were determined by analysis of mobility-shift titrations of Gal repressor with wild-type operator subject to the constraint that $\Delta G_E = \Delta G_I$: a procedure that eliminates the need to simultaneously analyze wild-type titrations with titrations of O_E^- and O_I^- operators.

Developmental and metabolic pathways are regulated at several levels, including transcription, RNA processing, translation, posttranslational modifications, transport, and degradation. However, in many cases, the key regulatory step occurs at the level of transcriptional initiation. It is becoming increasingly clear that the mechanism of regulation of transcriptional initiation is a complex and varied process. There exists a variety of repressors and activators of transcription initiation (as well as their allosteric effectors) and a complex set of interactions among them. In addition, DNA structure and conformation play an important role in the regulation of many systems. The presence of such a variety of regulatory components adds to the complexity of the mechanisms of gene regulation which must be understood.

A feature common to the regulatory systems of many prokaryotic and eukaryotic genes is cooperative interaction between proteins bound at DNA recognition sequences at distant locations. This phenomenon, originally proposed in prokaryotic systems (Irani et al., 1983; Majumdar & Adhya, 1984; Dunn et al., 1984), is believed to be caused by protein-protein interactions between two (or more) proteins bound to specific sites on the DNA extruding the intervening DNA as a "loop". Examples of such interactions in prokaryotes include transcription [cf. Haber and Adhya (1988), Lee and Shleif (1989), and Hochschild and Ptashne (1986)], replication [cf. Echols (1986)] and recombination [cf. Moitoso de Vargas et al. (1989)]. In eukaryotes, such interactions are believed to occur at enhancers and other regulatory regions [cf. Garcia-Blanco et al. (1989), Keegan et al. (1986), and Bell et al. (1988)].

Because of the ubiquity of these interactions, we have initiated a detailed thermodynamic analysis of the *gal* operon of *Escherichia coli*. The *gal* operon is a good model system for these studies because it contains both positive and negative regulatory elements, while remaining relatively simple. In the absence of galactose, Gal repressor binds to two palindromic operators, O_E and O_I,[1] whose centers of symmetry are separated by 114 bp (Majumdar & Adhya, 1984). *Neither* site overlaps the promoters [see Adhya (1987) for review; Figure 1]. In addition, a site for CRP is present, partially overlapping with the promoters (Majumdar & Adhya, 1984). Transcription is initiated from promoter P_1 in the presence of bound CRP-cAMP and from P_2 in its absence (Musso et al., 1977). The fact that Gal repressor does not prevent binding of either RNA polymerase or CRP argues strongly against steric hindrance as the mechanism of negative regulation in the *gal* operon [see Adhya (1989)]. Studies of regulation in vivo (Haber & Adhya, 1988) support a mechanism that involves interaction of the bound Gal repressor molecules, giving rise to a "looped complex", as was proposed originally.

As the first step in the quantitative characterization of the interactions occurring among the proteins regulating transcription of the *gal* operon, the interaction of Gal repressor with O_E and O_I (Figure 1) has been analyzed by quantitative

† This work was supported by NIH Grant GM39929 and BRSG funds to the College of Medicine.

* Author to whom correspondence should be addressed.
‡ The Albert Einstein College of Medicine.
§ National Institutes of Health.

[1] Abbreviations: CRP, cyclic AMP receptor protein; DNase I, bovine pancreas deoxyribonuclease I (EC 3.1.21.1); BSA, bovine serum albumin; Tris, tris(hydroxymethyl)aminomethane; Bis-tris, [bis(2-hydroxyethyl)-amino]tris(hydroxymethyl)methane; DTT, dithiothreitol; BIS, *N,N'*-methylenebis(acrylamide); O_E, upstream, "external", operator site; O_I, "internal" operator site located within the *galE* gene.

0006-2960/90/0429-3374$02.50/0

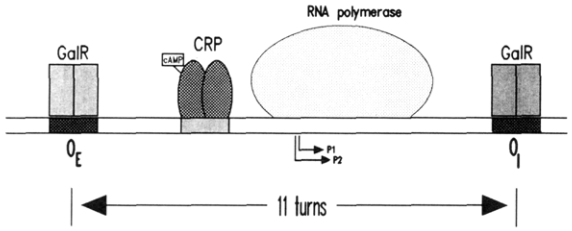

FIGURE 1: Schematic representation of the operator of the *gal* operon (not to scale).

Table I: Sequences of *gal* Operator Mutations Used in this Study[a]

plasmid	binding sites	O_E[b]	O_I[b]
pI24	Gal/Gal	GTGTAAAC|GATTCCAC	GTGGTAGC|GGTTACAT
pH104	Lac/Gal	tTGTgAgC|GcTcaCAa	GTGGTAGC|GGTTACAT
pH106	Gal/Lac	GTGTAAAC|GATTCCAC	tTGtgAGC|GcTcACAa

[a] Haber and Adhya, 1988. [b] Bold face letters indicate the bases that differ between O_I and O_E. Lower-case letters indicate the bases that were mutated to produce sites recognized by Lac repressor. The *lac* recognition sites produced in O_I and O_E are identical.

footprint and mobility-shift titration techniques. The results of these studies show that (1) Gal repressor binds cooperatively to linear DNA only at low temperatures and (2) Gal repressor self-association and DNA binding are linked equilibria. No direct evidence of DNA loop formation was obtained under "physiological" conditions. We suggest that the inability of Gal repressor to form a cooperative looped complex is due to the temperature dependence of the enthalpy of the dimer–tetramer association reaction, which must be compensated by some other process in vivo.

Materials and Methods

Protein and DNA Preparation. Gal repressor was purified as has been described (Majumdar et al., 1987). The Gal repressor preparations used in this study were >95% pure as judged by SDS electrophoresis. The assumption has been made the specific DNA-binding activity of the Gal repressor preparations used in these studies is 100%.

The sequences of the DNA operators are shown in Table I. The conversion of the O_I and O_E *gal* recognition sequences to a *lac* recognition sequence was accomplished by seven and six base pair substitutions, respectively (Haber & Adhya, 1988). Cells containing the plasmids were grown following standard procedures (Maniatis et al., 1982). The plasmid DNA was purified by the alkaline lysis method (Birnboim & Doly, 1979) followed by two sequential CsCl centrifugations (Maniatis et al., 1982). Agarose electrophoresis was used to verify that the plasmid DNA preparations were predominantly in the supercoiled form. The plasmid preparations were essentially free of protein as determined from the A_{260}/A_{280} ratio (Maniatis et al., 1982).

Footprint Titration Analysis. Quantitative DNase I footprint titration experiments were conducted as has been described (Brenowitz et al., 1986a; Brenowitz & Senear, 1989). Densitometric analysis of autoradiograms was accomplished with a microcomputer-based video densitometer (M. Brenowitz, P. Reiner, and B. Turner, unpublished computer programs). The *assay buffers* in which protein and DNA were equilibrated were 25 mM Tris (pH 8 and 9), Bis-tris (pH 6 and 7), or acetate (pH 5), 5 mM $MgCl_2$, 1 mM $CaCl_2$, 2 mM DTT, 50 μg/mL BSA, 2 μg/mL calf thymus DNA, and 25, 50, 100, 150, or 200 mM KCl. The pH of each buffer was adjusted with HCl and determined at the temperature at which the binding studies were to be conducted. All reagents were reagent or molecular biology grade. The pH studies were conducted with 25 mM KCl. The temperature studies were conducted with 100 mM KCl.

Table II: Binding Site Configurations and Associated Energy States of Protein Binding to Two Sites

species	binding sites O_I	binding sites O_E	free energy contributions[a]	total free energy[a]
1			ref state	ΔG_{s1}
2	P		ΔG_I	ΔG_{s2}
3		P	ΔG_E	ΔG_{s3}
4	P----	P	$\Delta G_I + \Delta G_E + \Delta G_{IE}$	ΔG_{s4}

[a] Free energies are in kcal/mol, where $\Delta G_{IE} = \Delta G_{total} - (\Delta G_I + \Delta G_E)$; i.e., the cooperative free energy ΔG_{IE} is the difference between binding to both sites simultaneously and binding to each site independently. P represents the oligomeric state of the protein which binds DNA, i.e., protein dimer.

A statistical–thermodynamic model that describes the interaction of Gal repressor with its two-site operator is shown in Table II. The *individual-site* binding equations for sites O_I and O_E, constructed by considering the relative probability of each configuration (Ackers et al., 1982, 1983), are

$$\bar{Y}_I = \frac{k_I[P] + k_I k_E k_{IE}[P]^2}{1 + (k_I + k_E)[P] + k_I k_E k_{IE}[P]^2} \quad (1a)$$

and

$$\bar{Y}_E = \frac{k_E[P] + k_I k_E k_{IE}[P]^2}{1 + (k_I + k_E)[P] + k_I k_E k_{IE}[P]^2} \quad (1b)$$

where the microscopic equilibrium constants have been substituted for the Gibbs free energies (Table II) from the relation $\Delta G = -RT \ln k$. The ΔG's are of two types: (1) binding of ligand to one site in the absence of binding to others ("intrinsic" binding free energies, ΔG_i); (2) excess (or cooperative) energy for binding to two sites simultaneously (ΔG_{ij}). [P] represents the free concentration of Gal repressor dimer. The assumption has been made that $[P]_{total} \approx [P]_{free}$ because the concentration of operator sites in the equilibrium mixtures (≤20 pM) is low relative to the equilibrium binding constants. The concentration of Gal repressor dimer was calculated on the basis of a monomer–dimer equilibrium dissociation constant of 2.2 nM (see Results).

For operators in which either O_E or O_I was mutated to prevent Gal repressor binding ("reduced-valency mutants"; Table I) the *individual-site* binding equations reduce to the *single-site* binding expressions

$$\bar{Y}_I = \frac{k_I[P]}{1 + k_I[P]} \quad (2a)$$

and

$$\bar{Y}_E = \frac{k_E[P]}{1 + k_E[P]} \quad (2b)$$

for O_E^- and O_I^- operators, respectively. Simultaneous analysis of eqs 1a, 1b, 2a, and 2b was used to resolve unique values of ΔG_I, ΔG_E, and ΔG_{IE} from the binding data (Ackers et al., 1982; Senear et al., 1986). The data were fit to the appropriate equations by methods of nonlinear least-squares parameter estimation to determine the best-fit values of the parameters,

their 65% confidence limits, and the variance of the fit (Johnson & Frasier, 1985).

Because the observed values of fractional protection of the binding sites frequently do not span the entire range of 0–1, the binding data are treated as transition curves with the upper and lower curve end points included as adjustable parameters in the nonlinear least-squares parameter estimation procedure. For example, the expression used to analyze the titration data for an O_E^- operator is

$$\bar{Y}_I = m\frac{k_I[P]}{1 + k_I[P]} + b \quad (3)$$

where $m = 1/(u - l)$ and $b = l/(l - u)$, u and l being the upper and lower end points, respectively. The need to treat quantitative footprinting data as transition curves is a critical part of the analysis procedure.

Mobility-Shift Titrations. Mobility-shift titration experiments were conducted in 12 × 16 × 0.15 cm, 6% polyacrylamide gels (1:37 acrylamide:BIS) with 1× TAE as the electrophoresis buffer (Maniatis et al., 1982). Samples were prepared and equilibrated as described for the footprint titration studies except that 100- rather than 200-μL reaction volumes were used; 75 μL of each sample was loaded onto the gel. Gels were preelectrophoresed for 15 min and run at a constant current of 19 mA cm^2. Buffers used in the mobility-shift assays were identical with those used for the footprint studies except that 3% glycerol was present to facilitate loading the samples on the gel. Samples were loaded with the current running to ensure that each sample was resident in the well for the same length of time. The gels were dried, autoradiographed, and densitometrically analyzed as described for the footprint titration studies. Densitometry was conducted with the same methods and programs developed to analyze footprint titration autoradiograms.

The experiments conducted at low temperature were performed in Hoeffer SE 600 gel boxes cooled by a temperature-regulated circulating water bath. The temperature of the buffer surrounding the gel was maintained at 2 ± 1 °C throughout electrophoresis. The gels were thermally equilibrated and preelectrophoresed for 30 min in the apparatus prior to loading the samples. The samples were equilibrated as described above and quickly transferred to the gel to minimize perturbation of the protein–DNA equilibrium.

The mobility-shift assay measures the fraction of DNA molecules with i ligands bound, where i varies from zero to the number of binding sites. For the model shown in Table II the fractions of molecules θ_i that are unliganded, singly liganded, and doubly liganded can be written respectively as

$$\theta_0 = \frac{1}{1 + (k_I + k_E)[P] + k_I k_E k_{IE}[P]^2} \quad (4a)$$

$$\theta_1 = \frac{(k_I + k_E)[P]}{1 + (k_I + k_E)[P] + k_I k_E k_{IE}[P]^2} \quad (4b)$$

$$\theta_2 = \frac{k_I k_E k_{IE}[P]^2}{1 + (k_I + k_E)[P] + k_I k_E k_{IE}[P]^2} \quad (4c)$$

where [P] is the concentration of free Gal repressor dimer (Senear et al., 1986). The assumption that $[P]_{total} \approx [P]_{free}$ was also made in these experiments.

The fraction of molecules in each configuration was calculated at each protein concentration as $\theta_{i,lane} = D_i/D_{total}$ where D_i is the density of band i and D_{total} is the sum of the densities of all the bands in a lane (LeBowitz et al., 1989). Numerical fitting of transition curve end points was not used in the analysis of the mobility-shift data. For the operators containing two binding sites, eqs 4a–4c were simultaneously fit to the titration data for unliganded, singly liganded, and doubly liganded molecules as described above. Equations 4a–4c reduce to

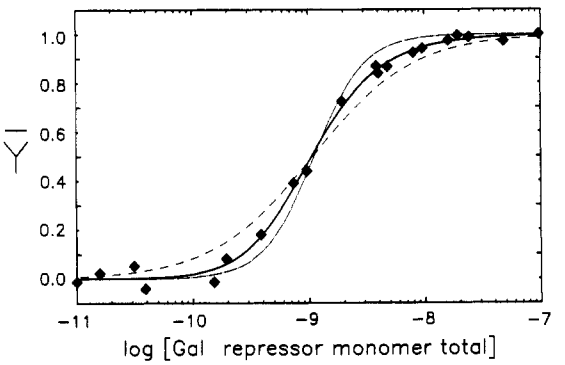

FIGURE 2: Individual-site binding curves for Gal repressor binding to site O_I in the O_E^- operator (Table I). Solution conditions are assay buffer (see Materials and Methods) at pH 7.0, 25 mM KCl, and 20 °C. The solid curve represents the best-fit values of ΔG_I and ΔG_{dimer} (Table III, line 1a). The dot-dashed line represents the limit where $\Delta G_{dimer} \gg \Delta G_I$. The dashed line represents the limit where $\Delta G_{dimer} \ll \Delta G_I$, a nondissociable dimer. $\bar{Y}$ represents fractional saturation.

$$\theta_0 = \frac{1}{1 + k_j[P]} \quad (5a)$$

and

$$\theta_1 = \frac{k_j[P]}{1 + k_j[P]} \quad (5b)$$

for analysis of single-site reduced-valency operators where θ_0 and θ_1 represent the fractions of unliganded and singly liganded molecules, respectively, and j refers to either site O_I or O_E. Simultaneous analysis of the wild-type and reduced-valency mutant operators was conducted as described for the footprint titration studies.

Results

DNA Binding and Gal Repressor Self-Association. As the first step in our analysis, a series of titrations of Gal repressor and the O_E^- and O_I^- operators were conducted in *assay buffer* (see Materials and Methods) at pH 7.0, 25 mM KCl, and 20 °C. Since Gal repressor appeared to be primarily dimeric in solution (Majumdar et al., 1987), the titration data were first fit to either eq 2a or eq 2b with the assumption that the Gal repressor existed entirely as dimers in solution. As can be seen in Figure 2 (dashed line), the data are of sufficient quality to suggest that there are systematic deviations of the data from this model.

The simplest model that describes the data is one in which protein self-association and DNA binding are coupled equilibria. The model assumes that Gal repressor monomers and dimers are in equilibrium and that dimers site specifically bind to DNA with much greater affinity than monomers (Figure 2). The results of nonlinear least-squares analysis where both ΔG_I (or ΔG_E) and ΔG_{dimer} were simultaneously minimized are shown in Table III for two autoradiographs of each of four experiments. The data are highly reproducible. As can be seen from the high numerical correlation between ΔG_I (or ΔG_E) and ΔG_{dimer}, unique values of the two parameters could not be obtained from all experiments. However, for those experiments in which s, the square root of the variance, is less than 0.040, correlations of 0.92–0.96 indicate some measure of reliability of the average value of −11.6 kcal/mol obtained

Table III: Gal Repressor Binding to a Single Site: Determination of the Monomer–Dimer Association Free Energy[a]

operator[b]	ΔG_I	ΔG_E	ΔG_{dimer}[c]	s[d]	corrl[e]
O_E^-					
1a	−13.1 ± 0.3		−11.6 ± 0.4	0.029	−0.95
	−12.6 ± 0.1		(−20)[f]	0.042	
1b	−13.2 ± 0.3		−11.5 ± 0.6	0.034	−0.96
2a	−13.4 ± 0.8		−10.3 ± 1.3	0.069	−0.99
2b	−12.6 ± 0.3		−11.5 ± 0.8	0.044	−0.95
O_I^-					
1a		−12.8 ± 0.2	−12.9 ± 0.7	0.038	−0.94
1b		−13.0 ± 0.2	−12.3 ± 0.7	0.025	−0.92
2a		−13.0 ± 0.2	−11.6 ■ 0.5	0.023	−0.94
2b		−13.3 ± 0.5	−11.0 ± 0.9	0.042	−0.98

[a] Experiments were conducted in assay buffer at pH 7.0, 25 mM KCl, and 20 °C. Free energies are in kcal/mol. [b] Two experiments with each operator were conducted with two exposures of each gel analyzed. [c] Free energy of the association of Gal repressor monomers to dimers. [d] Square root of the variance of the fitted curve. [e] Numerical correlation between ΔG_{dimer} and ΔG_I or ΔG_E. [f] Value in parenthesis is fixed during the least-squares minimization. The value of −20 for ΔG_{dimer} is below the limit of $\Delta G_{dimer} \ll \Delta G_I$ and represents a nondissociable dimer (see text).

for ΔG_{dimer} (corresponding to a dimer dissociation constant of 2.2 nM). Inspection of Figure 2 reveals the cause of the high numerical correlation; the shape of the binding curve varies only over a fairly narrow range from the limit of $\Delta G_{dimer} \ll \Delta G_I$ (dashed line) to the limit of $\Delta G_{dimer} \gg \Delta G_I$ (dot-dashed line). It is only the very high quality of the data and the fact that the measured binding curves under these solution conditions are intermediate to the two limiting conditions that allow resolution of both ΔG_I (or ΔG_E) and ΔG_{dimer}. Analysis of the titration data with $\Delta G_{dimer} = -11.6$ kcal/mol rather than a value exceeding the limit where $\Delta G_{dimer} \ll \Delta G_I$ results in ΔG_I (or ΔG_E) changing by approximately −0.5 kcal/mol (Table III, lines 1 and 2).

That the monomer–dimer dissociation constant of 2.2 nM is unchanged as a function of pH, monovalent ion concentration, and temperature is assumed in the following studies. This assumption is made because ΔG_{dimer} could not be determined at each condition of pH, salt, and temperature (data not shown). However, it was noted that binding curves at pH values other than 7.0 tended toward the limit where $\Delta G_{dimer} \gg \Delta G_I$ (indicating a decrease in the stability of dimers), conditions under which simultaneous determination of ΔG_I (or ΔG_E) and ΔG_{dimer} is not possible. ΔG_{dimer} did not appear to vary as a function of [KCl], although the data are not conclusive (data not shown). The binding data at high and low temperatures were not of sufficient quality to suggest a dependence of ΔG_{dimer} on temperature.

Gal Repressor Binding to Wild-Type Operator. Titrations of the wild-type operator were also conducted under the conditions described above. The concentration of Gal repressor dimer was calculated from the value of 2.2 nM determined for the dimer dissociation constant from the analysis of the single-site mutant operators. The titration data for the wild-type operator (Figure 3, upper panel) were fit to eqs 1a and 1b to determine the intrinsic binding free energies, ΔG_E and ΔG_I, and the cooperative free energy, ΔG_{EI} (Table IV). These data indicate that under these experimental conditions Gal repressor binding is not cooperative within the confidence limits of the data.

The broad confidence limits of the resolved free energies when ΔG_I, ΔG_E, and ΔG_{IE} are simultaneously minimized is due to the numerical correlation between the intrinsic and cooperative free energies; a phenomenon that has been discussed in detail elsewhere (Senear et al., 1986). An alternative

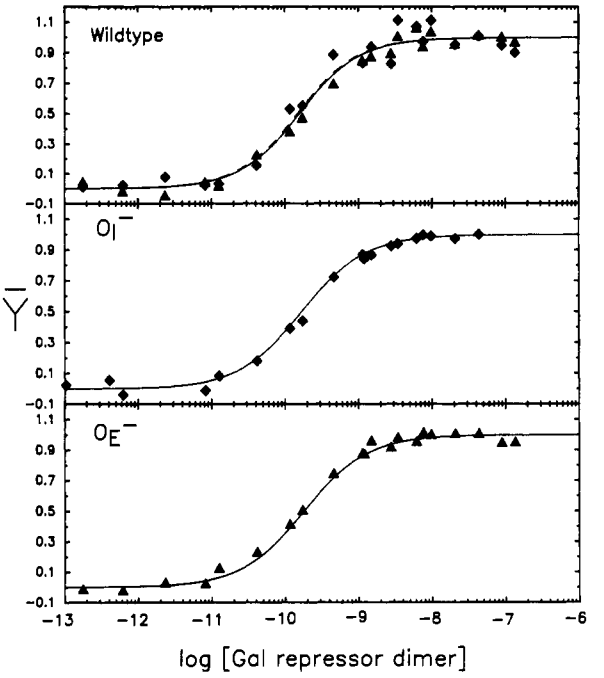

FIGURE 3: Examples of the individual-site binding curves for Gal repressor binding to the wild-type (upper), O_I^- (middle), and O_E^- (lower) operators used for the simultaneous analysis procedure to resolve ΔG_{IE}. Solution conditions are assay buffer (see Materials and Methods) at pH 7.0, 25 mM KCl, and 20 °C. A value of ΔG_{dimer} of −11.6 kcal/mol is assumed. $\bar{Y}$ represents fractional saturation.

Table IV: Gal Repressor Binding to Wild-Type Operator: Determination of Intrinsic and Cooperative Free Energies[a]

expt[b]	ΔG_I	ΔG_E	ΔG_{IE}[c]	s[d]
1a	−13.0 ± 0.2	−13.3 ± 0.2	(0)	0.052
	−12.9 ± 0.4	−13.2 ± 0.3	−0.3 ± 0.7	0.051
1b	−13.1 ± 0.2	−13.0 ± 0.2	(0)	0.059
	−12.7 ± 0.8	−12.6 ± 0.8	−0.8 ± 1.5	0.058
2a	−12.7 ± 0.2	−13.1 ± 0.3	(0)	0.044
	−12.7 ± 0.4	−13.0 ± 0.2	−0.1 ± 0.5	0.043
2b	−12.8 ± 0.2	−13.1 ± 0.2	(0)	0.046
	−12.7 ± 0.4	−13.0 ± 0.3	−0.2 ± 0.6	0.045

[a] Experiments were conducted in assay buffer at pH 7.0, 25 mM KCl, and 20 °C. [b] Two experiments were conducted with two exposures (a and b) of each gel analyzed. [c] Number in parentheses indicates that that value is fixed during the nonlinear parameter estimation. [d] Square root of the variance of the fitted curve.

method to determine cooperative free energies is the simultaneous analysis of wild-type titration data with titration data for each of the single-state (reduced-valency) mutants (Ackers et al., 1982; Brenowitz et al., 1986a; Senear et al., 1986). However, this analysis procedure makes the crucial assumption that the mutations used to eliminate binding at one site do not affect the intrinsic binding at the other site. This assumption has been shown to fail for deletion mutations, but not single base pair mutations, in the binding of the phage cIλ repressor to its operator (Brenowitz et al., 1989). As will be shown below, the reduced-valency mutant operator sequences used in these studies (Table I) appear to be consistent with the basic assumption of the analysis procedure.

In the simultaneous analysis, a single nonlinear least-squares minimum is determined for the wild-type titration data fit to eqs 1a and 1b (sites O_I and O_E, respectively), O_E^- titration data fit to eq 2a, and O_I^- titration data fit to eq 2b. Simultaneous analysis of the titrations presented in Figure 3 shows that $\Delta G_I = -13.1 \pm 0.3$, $\Delta G_E = -13.1 \pm 0.2$, and $\Delta G_{IE} = -0.2 \pm 0.2$ ($s = 0.052$). Binding is clearly noncooperative under these solution conditions. The validity of the assumption

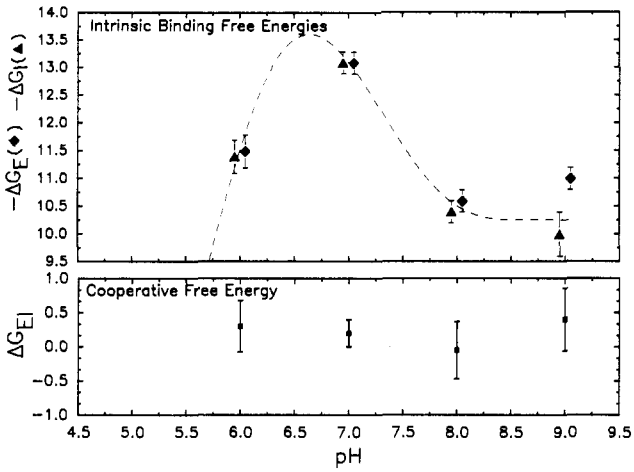

FIGURE 4: Intrinsic (ΔG_I and ΔG_E) and cooperative (ΔG_{IE}) free energies resolved from the simultaneous analysis of wild-type and reduced-valency mutant operators as a function of solution pH. Site O_I is represented by triangles (▲) and site O_E by diamonds (◆). Solution conditions are assay buffer (see Materials and Methods) at the designated pH, 25 mM KCl, and 20 °C. The dashed line in the upper panel is a smooth curve drawn to emphasize the changes in the data and does not represent any theoretical model. The dotted line in the lower panel shows $\Delta G_{IE} = 0$. Error bars indicate 65% confidence limits. The two sets of data are offset for clarity; determinations for both sites were made at the same pH.

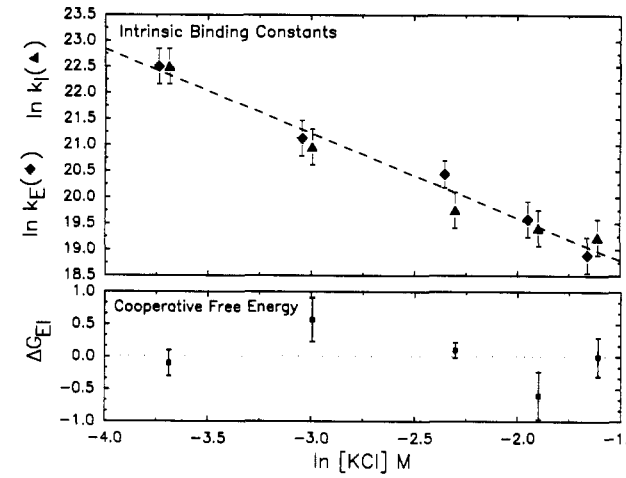

FIGURE 5: Intrinsic (k_I and k_E) and cooperative (k_{IE}) equilibrium binding constants resolved from the simultaneous analysis of wild-type and reduced-valency mutant operators as a function of solution KCl concentrations. Site O_I is represented by triangles (▲) and site O_E by diamonds (◆). Solution conditions are assay buffer (see Materials and Methods) at pH 7.0 and 20 °C with added KCl from 25 to 200 mM. The dashed line in the upper panel was determined by linear regression to the data points of both binding sites. The dotted line in the lower panel indicates ln $k_{IE} = 0$. Error bars indicate 65% confidence limits. The two sets of data are offset for clarity; determinations for both sites were made at the same salt concentration.

described above is indicated by the consistency of these results and the results obtained from the analysis of the wild-type operator alone (Table IV). Simultaneous analysis was used to determine the values of ΔG_I, ΔG_E, and ΔG_{IE} as a function of solution pH, [KCl], and temperature in the studies described below.

pH Dependence. The intrinsic binding free energies, ΔG_E and ΔG_I, and the cooperative free energy, ΔG_{EI}, were determined by simultaneous analysis of titration data for Gal repressor binding to wild-type and reduced-valency mutant operators as a function of solution pH (Figure 4). No repressor binding could be detected at pH 5. The fact that the resolved values of ΔG_{EI} are zero, within experimental error, shows that Gal repressor binding is noncooperative between pH 6 and pH 9. While the intrinsic binding energies ΔG_I and ΔG_E are equal, within experimental error, ΔG_E was consistently about 0.1 kcal/mol more negative than ΔG_I (Figures 4–6). This observation is consistent with the in vivo measurements of Haber and Adhya (1988); it is possible that this small difference in binding affinity is amplified in vivo. The binding affinity appears to reach a maximum between pH 6 and pH 7.

Salt Dependence. The binding of Gal repressor to operator was measured in assay buffer at pH 7.0 containing from 25 to 200 mM added KCl. A plot of ln k_I and ln k_E vs ln [KCl] is linear with a slope of 2.0 (Figure 5, upper panel), suggesting that approximately 2 ion equiv is displaced upon Gal repressor binding (Record et al., 1977, 1978). The plot of ΔG_{IE} versus ln [KCl] (Figure 5, lower panel) shows that Gal repressor binding is noncooperative.

Temperature Dependence. van't Hoff plots of the binding data obtained from 0 to 37 °C are shown in Figure 6. The data were fit to quadratic equations, shown as dashed lines. The intrinsic binding constants, k_I and k_E, reach a maximum at approximately 17.6 °C; the intrinsic free energies, ΔG_I and ΔG_E, reach a minimum at 22.7 °C (Figure 6, insert, upper panel). These results are strikingly reminiscent of the results obtained by Record and co-workers for the binding of Lac repressor and *Eco*RI endonuclease to their specific recognition sites (Ha et al., 1989). By use of their derivation, a value of −2.3 kcal mol^{-1} deg^{-1} was calculated as the change in heat

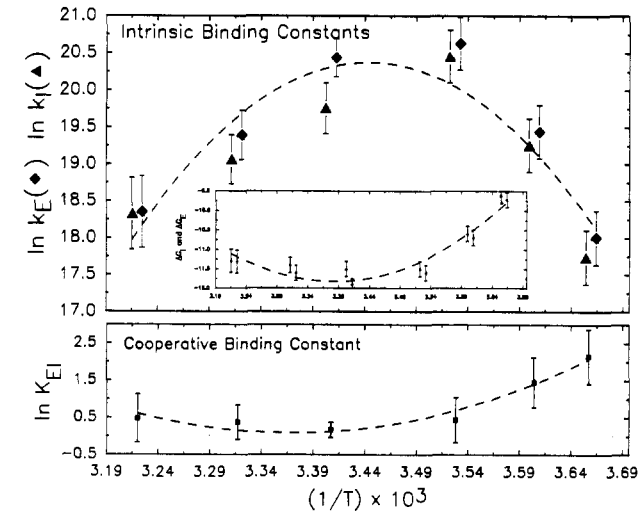

FIGURE 6: van't Hoff analysis of the intrinsic (k_I and k_E) and cooperative (k_{IE}) equilibrium binding constants resolved from the simultaneous analysis of wild-type and reduced-valency mutant operators. Site O_I is represented by triangles (▲) and site O_E by diamonds (◆). Solution conditions are assay buffer (see Materials and Methods) at pH 7.0 and 100 mM KCl at the indicated temperature. The dashed lines represent the best-fit quadratic equation to the data (including the data for both sites for the intrinsic binding constants). The dotted line in the lower panel indicates ln $k_{IE} = 0$. Error bars indicate 65% confidence limits. The temperature is in Kelvin. The two sets of data are offset for clarity; determinations for both sites were made at the same temperature.

capacity, ΔC_p, for the intrinsic binding of Gal repressor to its specific sites.

The cooperative binding constant, k_{EI}, appears to show a similar relationship to temperature, with a minimum at roughly 20 °C (Figure 6, lower panel). While it is clear that cooperative binding of Gal repressor appears only at low temperatures ($\Delta G_{IE} = -0.2 \pm 0.5$, -0.8 ± 0.4, and -1.2 ± 0.4 at 10.0, 4.0, and 0.0 °C, respectively), the magnitude of the change in k_{EI} is insufficient relative to the error in the data to allow determination of a maximum in ΔG_{IE}. Thus, ΔC_p for the cooperative binding component of the reaction was not calculated.

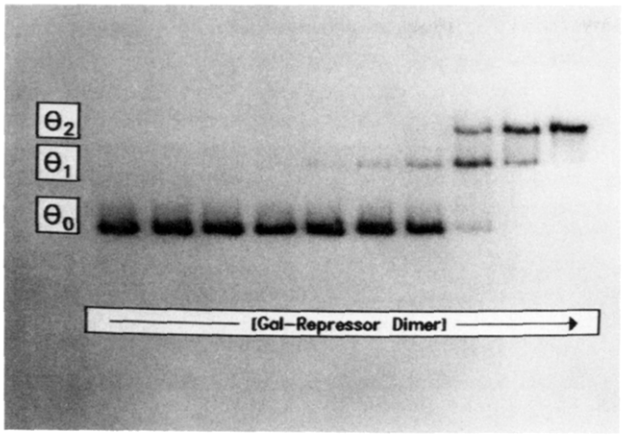

FIGURE 7: Example of an autoradiogram of a mobility-shift titration of Gal repressor binding to wild-type operator. θ_0, θ_1, and θ_2 represent unliganded, single liganded, and doubly liganded DNA; the relative intensities of the bands were fit to eqs 4a–4c, respectively, as described in the text.

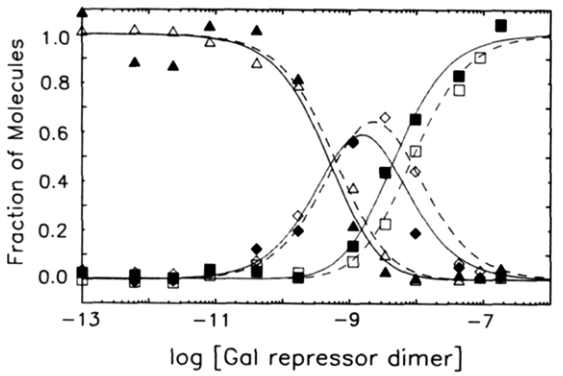

FIGURE 8: Titration curves resolved from mobility-shift experiments of Gal repressor binding to wild-type operator with the samples loaded with the gel running (closed symbols) and with the samples allowed to stand in the wells for 15 min prior to initiating electrophoresis (open symbols). Triangles represent the unliganded band, diamonds the singly liganded band, and squares the doubly liganded band. The lines represent the binding curves generated from the best-fit values to eqs 4a–4c (Table VIII); the solid lines correspond to the closed symbols and the dashed lines to the open symbols.

Mobility-Shift Titration Analysis. Mobility-shift assays were conducted in assay buffer at pH 7.0, 50 mM KCl, to which 3% glycerol was added. Footprint titration studies of the O_E^- operator in the presence and absence of 3% glycerol yielded values of ΔG_I of −12.4 ± 0.1 and −12.5 ± 0.1, respectively, suggesting that low concentrations of glycerol have no measurable effect on the intrinsic binding affinity of Gal repressor for operator.

Well-defined bands, without appreciable smearing, were observed under the electrophoresis conditions described (Figure 7). This is requisite for critical quantitative analysis since the relative proportion of a band is directly determined as its fraction of the total density within a lane. Numerical fitting of end points of the transition curves, as used in the footprint titration analysis, is not used in this analysis. Measurement of the sum of the band densities of each lane varied by no more than 5% within each of several autoradiograms (data not shown), suggesting the validity of this analysis procedure.

The mobility-shift assay differs from the footprint assay in that in the former technique the equilibrium mixture of protein and DNA is loaded directly onto the gel whereas in the latter technique the DNase I reaction is quenched prior to gel loading. It is possible that changes in the distribution of the protein–DNA complexes might occur when the sample is deposited onto the sample well. To test for this potential artifact, a set of titration samples was divided as follows: half was loaded onto a gel while the current was applied while the other half was loaded onto a separate gel and allowed to stand for 15 min before the current was applied. The gel on which the samples sat for 15 min before current application produced binding curves that were shifted slightly to the right (Figure 8). The total energy of binding ($\Delta G_I + \Delta G_E + \Delta G_{IE}$) to the operator decreased by ~0.6 kcal/mol. These results indicate that some dissociation of the protein–DNA complexes occurs upon prolonged incubation of the samples on the gel prior to electrophoresis.

Table V: Mobility-Shift Analysis of Gal Repressor Binding to Wild-Type and Reduced-Valency Mutant Operators

expt[a]	ΔG_I[b]	ΔG_{IE}	s
wt (a)	−11.7 ● 0.1	0.2 ± 0.1	0.037
wt (b)	−11.9 ± 0.2	−0.1 ● 0.1	0.043
O_E^-	−11.7 ± 0.2		0.053

[a] wt designates wild-type operator. In experiment a the samples were left in the wells for 15 min prior to electrophoresis. In experiment b the samples were loaded with the current on. In the O_E^- experiment the samples were loaded with the current on. Experimental conditions are assay buffer at pH 7.0, 50 mM KCl, and 20 °C. [b] Data obtained for the wild-type operator was analyzed subject to the constraint that $\Delta G_I = \Delta G_E$ (see text).

Table VI: Mobility-Shift Titrations of Gal Repressor Binding to Wild-Type and Reduced-Valency Mutant Operators at 0 °C[a]

operator	ΔG_I	ΔG_E	ΔG_{IE}	s
O_E^-	−9.3 ± 0.2			0.099
O_I^-		−9.2 ± 0.1		0.055
wt (a)	−9.2 ± 0.2	(=ΔG_I)	−0.6 ± 0.2	0.110
wt (b)	−9.3 ± 0.2	(=ΔG_I)	−0.7 ● 0.2	0.098

[a] The wild-type (wt) operator was analyzed subject to the constraint that $\Delta G_I = \Delta G_E$. Experimental conditions are assay buffer at pH 7.0, 100 mM KCl, and 0 °C. (a) and (b) are replicate experiments.

Since the intrinsic binding constants ΔG_I and ΔG_E were shown to be approximately equal by footprint titration analysis (Figures 4–6), the mobility-shift experiments were analyzed subject to the constraint that $k_E = k_I$ (eqs 4a–4c). This constraint allows the simultaneous determination of k_I *and* k_{IE} from the wild-type operator alone (unpublished numerical simulations). Analysis of the experiments shown in Figure 8 reveals that binding is noncooperative (Table V), providing an independent confirmation of the noncooperative binding obtained from the footprint titration experiments. The values of ΔG_I determined from the wild-type and O_E^- operators are in good agreement with each other (Table V) and differ by approximately 0.5 kcal/mol from the value obtained from the footprint titration studies under the same experimental conditions (Figure 5).

Mobility-shift experiments were conducted to verify the cooperative binding observed by footprint analysis at 0 °C. The temperature of the gel was maintained at 2 ± 1 °C with a temperature-regulated circulating water bath. The results of these experiments are shown in Table VI. As indicated in Table V, the resolved intrinsic free energies were approximately 0.5 kcal/mol less negative than those resolved from the footprint titration experiments (Figure 6). While the value of ΔGI_E of −0.7 kcal/mol resolved from the analysis of these experiments is clearly cooperative, it is somewhat lower than the value resolved from the footprint titration studies. This difference may be due to thermal reequilibration of the samples during transfer to the gel or dissociation within the gel of the more weakly bound protein–DNA complexes.

Despite the differences in the results obtained from the two experimental techniques, the fact that both titration techniques

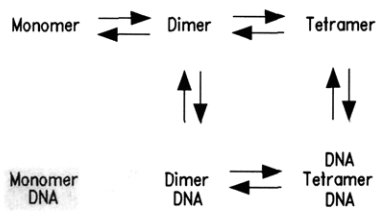

FIGURE 9: Schematic diagram showing the thermodynamic linkage of Gal repressor self-assembly and DNA binding.

report cooperative binding at low temperature supports the reliability of the observation.

Discussion

An understanding of the formation and stability of protein–DNA regulatory complexes is fundamental for the determination of the molecular mechanism of gene regulation. When these regulatory proteins oligomerize or when multiple proteins are present, there exist two levels of interactions within these complexes: protein–DNA and protein–protein. Although most published studies assume that these two levels operate independently, linkage of these interactions is not only possible but probable.

Monomer–Dimer Self-Association. Analysis of the shape of the binding curves determined from the titrations of Gal repressor with single-site operators strongly suggests that protein association and DNA binding are linked in this system. Our ability to resolve a value of ΔG_{dimer} was a combination of having obtained a data set of exceptionally high quality and the fortuitous occurrence that under the experimental conditions chosen for study ΔG_I, ΔG_E, and ΔG_{dimer} are very similar. Unique values of ΔG_{dimer} can be obtained only when this occurrence of values occurs (unpublished numerical simulations). While it is clear that direct physical techniques must be used for a more detailed analysis of Gal repressor self-association, these data provide a starting point for such analyses.

Gal repressor, like other DNA binding proteins, possesses a "helix–turn–helix" motif and recognizes a DNA sequence with a 2-fold axis of symmetry (Majumdar & Adhya, 1984). The solution of cocrystal structures of several proteins with helix–turn–helix motifs with operator DNA obtained by X-ray diffraction analysis shows that each monomer recognizes one "half-site" (Jordan & Pabo, 1988; Aggarwal et al., 1988; Otwinowski et al., 1988). The observation that base-pair substitutions on opposite half-sites are nonadditive in some "helix–turn–helix" proteins (Benson et al., 1988; Sarai & Takeda, 1989) argues for the functional linkage of protein dimerization and DNA recognition in this class of repressor proteins.

Several lines of evidence suggest that the monomers of a Gal repressor dimer do not bind each half-site independently. Circular dichroism spectroscopy (Wartel & Adhya, 1988) and ethylation interference studies (Majumdar & Adhya, 1989) both indicate that the conformation of the DNA comprising the binding sites is altered upon interaction with Gal repressor. In addition, electrophoretic analysis shows that Gal repressor bends DNA upon site-specific binding (Zwieb et al., 1989). Exactly how these structural changes contribute to site-specific recognition and whether DNA structural changes are involved in potential cooperative interactions that produce DNA looping are unknown.

A proposed thermodynamic scheme for Gal repressor that describes the linkage of protein self-assembly and DNA binding is shown in Figure 9. It is clear from the results presented in Figure 2 and Table III that the formation of the dimer–DNA complex proceeds via the binding of "preformed" dimers rather than the "cooperative" interaction of bound monomers.

A consequence of the linkage of Gal repressor dimerization and DNA recognition is the potential for an additional level of regulation. Since protein dimerization is required for the binding of Gal repressor, then the observation that the free energy of dimerization is comparable to the free energy of binding (Table III) suggests that allosteric regulation by galactose could occur at the level of either protein self-assembly or DNA binding. Thus, the dimerization interface may be important for regulation. The presence of a second regulatory interface, with properties different from those of the protein–DNA interface, may allow subtle regulation of repressor binding and hence gene expression.

Interest in the possible regulatory role of the dimerization of DNA-binding proteins has been rekindled with the identification of the "leucine zipper" motif (Landschulz et al., 1988) and the observation that proteins bearing this motif can form heterodimers (Turner & Tjian, 1989; Gentz et al., 1989). These authors suggest that heterodimer formation may allow additional binding specificities to be created; that is, heterodimers would recognize binding sites composed of two different recognition half-sites. Support for this idea comes from the preference of peptides corresponding to the leucine zipper regions of Fos and Jun to form heterodimers (O'Shea et al., 1989). Weak heterodimer contacts would result in continuous disproportionation of the complexes in the cell, resulting in a mixed population of homo- and heterodimers. On the other hand, strong homodimer contacts would require that heterodimers form at synthesis, arguing for coordinated synthesis of these proteins. Thus, the balance between homodimer and heterodimer formation is critical to the regulatory properties of these proteins.

The effect of pH on the intrinsic binding constants of Gal repressor is unusual; site-specific DNA-binding proteins in general show monotonically decreasing binding affinity with increasing pH (Vershon et al., 1987a,b; Senear & Ackers, 1990). Although the observed binding affinity maximum might reflect the involvement in site-specific recognition of a group or groups with pK_a's between pH 6 and pH 7, an alternative interpretation is suggested by the fact that the binding curves at pH values other than 7.0 tended toward the limit where $\Delta G_{dimer} \gg \Delta G_I$, suggesting a pH dependence of ΔG_{dimer}. Thus, the pH profile may reflect the properties of protein self-association rather than the protein–DNA interaction. However, differentiation of the chemistry of protein self-association and DNA recognition will require direct analysis of protein self-association by physical methods.

In contrast to the unusual pH dependence observed for the intrinsic binding free energies, the monovalent ion concentration dependence is typical of many DNA-binding proteins and is consistent with the release of approximately 2 counterion equiv (Record et al., 1977, 1978). This value is significantly lower than the values of 7.9 and 8.2 reported for Lac repressor binding (Whitson et al., 1986), a result that may reflect significant differences in the mechanisms of DNA recognition of these proteins despite their high degree of sequence homology.

Interpretation of ΔC_p. The temperature dependence of the intrinsic binding free energies is indicative of a large change in the heat capacity, ΔC_p, as was shown for Lac repressor and *Eco*RI site-specific binding (Ha et al., 1989). This result indicates that the enthalpy of association, ΔH, is temperature dependent. ΔH is positive at low temperature, passes through zero when K is at a maximum, and becomes negative at higher

temperatures. The binding of Gal repressor to operator is entropy driven at low temperature and enthalpy driven at high temperature.

A large negative ΔC_p has been interpreted as evidence for a large hydrophobic driving force of DNA binding due to solvent exclusion (Ha et al., 1989). However, the fact the Gal repressor clearly undergoes monomer–dimer self-association reactions makes the interpretation of our results more problematic. It is possible that the ΔC_p observed for Gal repressor could be the result of opposing enthalpies for the monomer–dimer and dimer–DNA interactions (Figure 9). Since it was not possible to resolve unique values of ΔG_{dimer} from the single-site operators as a function of temperature, the temperature dependence of k_I and k_E may be due to compensatory changes in the dimerization and DNA-binding reactions.

Cooperativity and Formation of a Looped Complex. ΔG_{IE} is negative only below 10 °C (indicating cooperative binding of Gal repressor to its operator), a result confirmed by the low-temperature mobility-shift titrations. Since the cooperative free energy, ΔG_{IE}, is simply an energy difference (Table II), a negative cooperative energy does not necessarily demonstrate the existence of a looped complex. Conversely, a looped complex can form without a net negative ΔG_{IE}; its formation can be driven solely by mass action.

Is there independent evidence to suggest that cooperative binding by Gal repressor results in a looped complex? No additional bands were observed in the 0 °C mobility-shift titrations although a looped complex might be expected to have an altered electrophoretic mobility. However, a systematic study of the mobility of complexes of Lac repressor and operator as a function of the number of helical turns between binding sites (Kramer et al., 1987) showed that the looped and tandem (a separate protein molecule bound to each site) complexes have the same mobility when the binding sites are separated by 11 helical turns, the same as for the *gal* operon. Another diagnostic of loop formation is the presence of DNase I hypersensitive sites in the region between the binding sites (Hochschild & Ptashne, 1986). No hypersensitivity was observed in titrations of wild-type operator with Gal repressor under conditions other than 0 °C. Some hypersensitivity was observed at 0 °C, although its magnitude is much lower than that observed upon the binding of Lac repressor to a *gal* operator in which mutations change the recognition sequence from *gal* to *lac*. A looped complex is stably formed by Lac repressor and this operator (unpublished results; Mandal et al., 1990). While the low magnitude of hypersensitivity may reflect the low magnitude of ΔG_{IE} (hence the low concentration of the looped complex configuration at equilibrium), these results suggest that a looped complex is not readily formed by Gal repressor and the DNA restriction fragment containing the *gal* operators.

Regulation in Vivo versus Binding in Vitro. The noncooperative binding observed in vitro is not consistent with an extensive in vivo analysis of regulation of the *gal* operon (Haber & Adhya, 1988) which revealed that the simultaneous occupancy of both O_I and O_E by Gal repressor is required for regulation, taken as evidence for the formation of the regulatory looped complex. What might be the cause of the differences between the in vivo and in vitro studies? We consider below several explanations: (1) solution conditions and small molecules; (2) requirement for an additional macromolecule; (3) structure and topology of the DNA; (4) dimer–tetramer association of Gal repressor.

(1) Of the wide range of solution conditions analyzed (Figures 4–6), only at low temperature is Gal repressor binding cooperative. Although simultaneous extremes in conditions were not analyzed, it seems unlikely that other combinations of solution conditions would produce cooperativity. The question of whether cooperative binding or looped complex formation requires a small effector molecule is more difficult to answer. Although our search has not been exhaustive, the results described herein, as well as studies in progress on the effect of specific mono- and divalent ions (unpublished data), suggest that specific ions do not induce cooperative binding. However, the possibility exists that a specific allosteric effector or cofactor is required.

(2) That an additional macromolecule may be required for cooperative binding and/or loop formation is suggested by recent studies on the IHF protein and Lambda Integrase (Moitoso de Vargas et al., 1989). It was shown that IHF facilitates the formation of a loop between a core site and an *att*P′ site by Lambda Integrase by bending the DNA. These authors argue that the principle function of IHF in vivo is to bend DNA and facilitate the formation of looped complexes. The presence of a CRP-binding site in the region between O_I and O_E (Figure 1), coupled with the documented ability of CRP to bend DNA upon site-specific binding (Gartenberg & Crothers, 1988), suggests the possibility of a similar interaction within the *gal* operon. Such a mechanism would imply a very subtle interplay of transcription activators and repressors.

(3) Two observations suggest that the formation of a looped complex is not inhibited by the binding sites being on opposite sides of the helix or unusual stiffness of the *gal* operon DNA. First, the Gal repressor binding sites have a center to center distance of 114 bp corresponding to a separation of 11 helical turns (assuming a B-DNA helical repeat of 10.4 bp/turn; Wang, 1979). Thus, the O_I and O_E are "in phase" with each other on the same side of the helix; no energetic penalty is incurred in order to juxtapose them [cf. Bellomy et al. (1988)]. Second, Lac repressor forms a stable looped complex with linear restriction fragments of the *lac* operator (Kramer et al., 1987). The binding of Lac repressor to a mutant *gal* operator in which O_I and O_E were altered to a *lac* recognition sequence is highly cooperative and also forms a stable looped complex (unpublished results; Mandal et al., 1990). Lac and Gal repressors are highly homologous proteins that bend DNA equivalently upon site-specific binding (Zwieb et al., 1989). Since Lac repressor is known to form tetramers at low protein concentrations and forms a looped complex with *gal* DNA, we conclude that the inability of Gal repressor to form a stable looped complex in vitro is an inherent property of the protein and not the result of an unusual structure of the operator DNA.

The ability of negative supercoiling to facilitate formation of looped complexes is well documented, particularly for Lac repressor binding to its primary-operator and pseudooperator sites (Whitson et al., 1987; Borowiec et al., 1987). It is generally believed that negative supercoiling facilitates loop formation by decreasing the energy required to bring the two binding sites into juxtaposition, particularly if the two binding sites are on opposite sides of the helix. The fact that O_E and O_I are in phase with each other suggests that supercoiling may not have a dramatic effect on the cooperative interaction of bound Gal repressor dimers.

(4) The thermodynamic linkage scheme connecting Gal repressor self-assembly and DNA binding is shown in Figure 9. (Note that this scheme does not include the allosteric effector galactose.) The data presented in Figure 2 and Table III suggest that the dimer–DNA complex forms predominantly via preformed dimers. There is currently no direct evidence

for the dimer–tetramer association reactions depicted in the linkage scheme. The formation of a looped complex can be argued only by analogy with Lac repressor and from in vivo analysis. However, this scheme is a plausible description of the binding of Gal repressor to its operator, suggesting that the noncooperative binding observed in vitro is the result of a small dimer–tetramer association constant. Direct protein self-association studies will be required to determine if this is, in fact, the case.

Resolving Cooperative Free Energies by the Mobility-Shift Assay. The mobility-shift assay (Garner & Revzin, 1981; Fried & Crothers, 1981) is frequently used to analyze site-specific protein–DNA interactions. Discussions of the experimental variables that must be considered for the determination of thermodynamically valid binding parameters have been published (Carey, 1987; Revzin, 1989; Fried, 1989; Cann, 1989). Despite the extensive use of the mobility-shift assay, quantitation of the proportion of the intermediately liganded bands to yield information concerning cooperative binding has been exploited on only a very limited basis (Senear et al., 1986; Fried, 1989; Tsai et al., 1989; LeBowitz et al., 1989). Both the intrinsic and cooperative free energies determined from the mobility-shift titrations of Gal repressor binding to its multisite operator were remarkably consistent with those obtained by the footprint titration analysis under two sets of experimental conditions. However, the generality of this result has not been determined; caution must be used in the analysis of other systems under different experimental conditions. Since measurement of cooperative free energies is critically dependent on the concentration of the intermediate species and mobility-shift titrations provide a direct measure of these concentrations, this assay may be a very useful tool in the quantitative analysis of cooperative binding.

References

Ackers, G. K., Johnson, A. D., & Shea, M. A. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 1129–1133.

Ackers, G. K., Shea, M. A., & Smith, F. R. (1983) *J. Mol. Biol. 170*, 223–242.

Adhya, S. (1987) in *Escherichia coli and Salmonella Hyphiourium: Cellular and Molecular Biology* (Neidhard, F., Ingraham, J., Low, B., Magasanik, B., Schalchrerau, M., & Umberger, E., Eds.) pp 1503–1512, American Society of Microbiology.

Adhya, S. (1989) *Annu. Rev. Genet. 23*, 207–230.

Aggarwal, A. K., Rodgers, D. W., Drottar, M., Ptashne, M., & Harrison, S. C. (1988) *Science 242*, 899–907.

Bell, S. P., Learned, M., Jantzen, H.-M., & Tjian, R. (1988) *Science 241*, 1192–1197.

Bellomy, G. R., Mossing, M. C., & Record, M. T. (1988) *Biochemistry 27*, 3900–3906.

Benson, N., Sugiono, P., & Youderian, P. (1988) *Genetics 188*, 21–29.

Birnboim, H., & Doly, J. (1979) *Nucleic Acids Res. 7*, 1513–1523.

Borowiec, J. A., Zhang, L., Sasse-Dwight, S., & Gralla, J. D. (1987) *J. Mol. Biol. 196*, 101–111.

Brenowitz, M., & Senear, D. F. (1989) in *Current Protocols in Molecular Biology* (Ausubel, F. M., et al., Eds.) Supplement 7, Unit 12.4, Greene Publishing Associates and Wiley-Interscience, New York.

Brenowitz, M., Senear, D. F., Shea, M. A., & Ackers, G. K. (1986a) *Methods Enzymol. 130*, 132–181.

Brenowitz, M., Senear, D. F., Shea, M. A., & Ackers, G. K. (1986b) *Proc. Natl. Acad. Sci. U.S.A. 83*, 8462–8466.

Brenowitz, M., Senear, D. F., & Ackers, G. K. (1989) *Nucleic Acids Res. 17*, 3747–3755.

Cann, J. R. (1989) *J. Biol. Chem. 264*, 17032–17040.

Carey, J. (1987) *Proc. Natl. Acad. Sci. U.S.A. 85*, 975–979.

Dunn, T. M., Hahn, S., Ogden, S., & Schleif, R. F. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 5017–5020.

Echols, H. (1986) *Science 233*, 1050–1056.

Fried, M. (1989) *Electrophoresis 10*, 366–376.

Fried, M., & Crothers, D. M. (1981) *Nucleic Acids Res. 9*, 6505–6525.

Garcia-Blanco, M. A., Clerc, R. G., & Sharp, P. A. (1989) *Genes Dev. 3*, 739–745.

Garner, M. M., & Revzin, A. (1981) *Nucleic Acids Res. 9*, 3047–3060.

Gartenberg, M. R., & Crothers, D. M. (1988) *Nature 333*, 824–829.

Gentz, R., Rauscher, F. J., Abate, C., & Curran, T. (1989) *Science 243*, 1695–1699.

Gralla, J. D. (1989) *Cell 57*, 193–195.

Ha, J.-H., Spolar, R. S., & Record, M. T. (1989) *J. Mol. Biol. 209*, 801–816.

Haber, R., & Adhya, S. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 9683–9687.

Hochschild, A., & Ptashne, M. (1986) *Cell 44*, 681–687.

Irani, M. H., Orosz, L., & Adhya, S. (1983) *Cell 32*, 783–788.

Johnson, M. L., & Frasier, S. G. (1985) *Methods Enzymol. 117*, 301–342.

Jordan, S. R., & Pabo, C. O. (1988) *Science 242*, 893–899.

Keegan, L., Gill, G., & Ptashne, M. (1986) *Science 231*, 699–704.

Kramer, H., Niemoller, M., Amouyal, M., Revet, B., von Wilcken-Bergmann, B., & Muller-Hill, B. (1987) *EMBO J. 6*, 1481–1491.

Landschulz, W. H., Johnson, P. F., & McKnight, S. L. (1988) *Science 240*, 1759–1763.

LeBowitz, J. H., Clerc, R. G., Brenowitz, M., & Sharp, P. A. (1989) *Genes Dev. 3*, 1625–1638.

Lee, D.-H., & Schleif, R. F. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 476–480.

Majumdar, A., & Adhya, S. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 6100–6104.

Majumdar, A., & Adhya, S. (1989) *J. Mol. Biol. 208*, 217–223.

Majumdar, A., Rudikoff, S., & Adhya, S. (1987) *J. Biol. Chem. 262*, 2326–2331.

Mandal, N., Su, W., Haber, R., Adhya, S., Echols, H. (1990) *Genes Dev.* (in press).

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning—A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Moitoso de Vargas, L., Kim, S., & Landy, A. (1989) *Science 244*, 1457–1461.

Musso, R. E., Di Lauro, R., Adhya, S., & de Crombrugghe, B. (1977) *Cell 12*, 847–854.

O'Shea, E. K., Rutkowski, R., Stafford, W. F., & Kim, P. S. (1989) *Science 245*, 646–647.

Otwinowski, Z., Schevitz, R. W., Zhang, R.-G., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q., Luisi, B. F., & Sigler, P. B. (1988) *Nature 335*, 321–329.

Record, M. T., Jr., deHaseth, P. L., & Lohman, T. M. (1977) *Biochemistry 16*, 4791–4796.

Record, M. T., Anderson, C. F., & Lohman, T. M. (1978) *Q. Rev. Biophys. 11*, 103–178.

Revzin, A. (1989) *Biotechniques 7*, 346–356.

Sarai, A., & Takeda, Y. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 6513–6517.
Schevitz, R. W., Otwinowski, Z., Joachimiak, A., Lawson, C. L., & Sigler, P. B. (1985) *Nature 317*, 782–786.
Senear, D. F., & Ackers, G. K. (1990) *Biochemistry* (in press).
Senear, D. F., Brenowitz, M., Shea, M. A., & Ackers, G. K. (1986) *Biochemistry 25*, 7344–7354.
Tsai, S. Y., Tsai, M.-J., & O'Malley, B. W. (1989) *Cell 57*, 443–448.
Turner, R., & Tjian, R. (1989) *Science 243*, 1689–1694.
Vershon, A. K., Liao, S.-M., McClure, W. R., & Sauer, R. T. (1987a) *J. Mol. Biol. 195*, 311–322.
Vershon, A. K., Liao, S.-M., McClure, W. R., & Sauer, R. T. (1987b) *J. Mol. Biol. 195*, 323–331.
Wang, J. C. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 200–203.
Wartell, R. M., & Adhya, S. (1988) *Nucleic Acids Res. 16*, 11531–11541.
Whitson, P. A., Olson, J. S., & Matthews, K. S. (1986) *Biochemistry 25*, 3852–3858.
Whitson, P. A., Hsieh, W.-T., Wells, R. D., & Matthews, K. S. (1987) *J. Biol. Chem. 25*, 4943–4946.
Zwieb, C., Kim, J., & Adhya, S. (1989) *Genes Dev. 3*, 606–611.

A Single mRNA Encodes Multiple Copies of the Egg Peptide Speract[†,‡]

Chodavarapu S. Ramarao,[§] Deborah J. Burks,[§] and David L. Garbers*[,§,∥]
Department of Pharmacology and Department of Molecular Physiology and Biophysics and Howard Hughes Medical Institute, Vanderbilt University School of Medicine, Nashville, Tennessee 37232
Received August 11, 1989; Revised Manuscript Received November 27, 1989

ABSTRACT: A complementary DNA clone (2.3 kb) that encodes the egg peptide speract (Gly-Phe-Asp-Leu-Asn-Gly-Gly-Gly-Val-Gly) has been isolated from an ovary cDNA library of the sea urchin *Strongylocentrotus purpuratus*. The DNA sequence predicts an open reading frame of 296 amino acids. The likely site of initiation, however, is a downstream in-frame translation initiation codon that would result in a polypeptide of 260 amino acids containing 10 decapeptides, each separated by a single lysine residue. Four of the peptides are speract, and six have the predicted structures of Gly-Phe-Ala-Leu-Gly-Gly-Gly-Gly-Val-Gly (occurs twice), Gly-Phe-Asn-Leu-Asn-Gly-Gly-Gly-Val-Gly, Gly-Phe-Ser-Leu-Thr-Gly-Gly-Gly-Val-Gly, Gly-Thr-Met-Pro-Thr-Gly-Ala-Gly-Val-Asp, and Ile-Asp-His-Asp-Thr-Leu-Ala-Ser-Val-Ser. The isolated cDNA insert hybridized to two species of ovarian mRNA (1.2 and 2.3 kb) obtained from species known to produce speract or speract-like peptides, but failed to hybridize to RNA from other species. Subsequently, a second ovarian cDNA clone (1.2 kb) was isolated and sequenced; this clone contained two additional potential decapeptides: Ser-Phe-Asp-Leu-Asn-Gly-Gly-Gly-Val-Gly and Ser-Thr-Met-Pro-Thr-Gly-Ala-Gly-Val-Asp. The various speract and speract-like peptides found in egg-conditioned media, therefore, reflect, in part, variable structures within a single copy of mRNA. The peptides are likely processed by a combination of trypsin and carboxypeptidase B activity; although the peptides containing a Gly-Gly-Gly-Val-Gly carboxyl-terminal sequence are known to possess sperm respiration-stimulating activity, the function, if any, of the other potential peptides remains to be determined.

The egg-conditioned media of sea urchins contain peptides that are capable of stimulating sperm metabolism and motility; these peptides may also act as chemoattractants for spermatozoa (Hansbrough & Garbers, 1981a; Suzuki et al., 1981; Garbers et al., 1982; Ward et al., 1985). One such peptide, speract (Gly-Phe-Asp-Leu-Asn-Gly-Gly-Gly-Val-Gly), was initially isolated from egg-conditioned media of *Strongylocentrotus purpuratus* sea urchins based on its ability to stimulate sperm respiration at acidic extracellular pH values (Hansbrough & Garbers, 1981a; Suzuki et al., 1981). Subsequently, speract was shown to have several effects on spermatozoa including increased motility (Hansbrough & Garbers, 1981a; Suzuki et al., 1981), elevated cyclic nucleotide concentrations (Garbers et al., 1982), proton efflux leading to intracellular alkalinization (Hansbrough & Garbers, 1981b), dephosphorylation of guanylate cyclase, leading to markedly decreased enzyme activity (Ramarao & Garbers, 1985), and transient elevations of Ca^{2+} (Schackmann & Chock, 1986; Lee & Garbers, 1986). The egg peptides are species-specific, and speract, for example, fails to stimulate spermatozoa from the sea urchin *Arbacia punctulata* (Ward et al., 1985). Egg-conditioned media of *A. punctulata* contain a different peptide named resact (Cys-Val-Thr-Gly-Ala-Pro-Gly-Cys-Val-Gly-Gly-Gly-Arg-Leu-NH_2) that stimulates spermatozoa (Suzuki et al., 1984; Ward et al., 1985). A diversity of peptide structures exist across the species, which accounts for the failure of egg-conditioned media of a particular species to necessarily cross-react with spermatozoa from another species. However, even within the same species, many variants of the same basic peptide structure have been isolated (Shimomura et al., 1986; Suzuki et al., 1988). From the egg-conditioned media of *Hemicentrotus pulcherrimus*, for example, the following peptides have been chemically isolated: Gly-Phe-Asp-Leu-Asn-Gly-Gly-Gly-Val-Gly, Gly-Phe-Asp-Leu-Thr-Gly-Gly-Gly-Val-Gly, Gly-Phe-Ser-Leu-Asn-Gly-Gly-Gly-

[†] Supported by NIH Grant HD10254 and by a grant from the Andrew W. Mellon Foundation.
[‡] The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number J02896.
* Address correspondence to this author at the Howard Hughes Medical Institute.
[§] Department of Pharmacology.
[∥] Department of Molecular Physiology and Biophysics and Howard Hughes Medical Institute.

0006-2960/90/0429-3383$02.50/0