# Thermodynamic and alkylation interference analysis of the lac repressor-operator substituted with the analogue 7-deazaguanine

**ARTICLE** *in* BIOCHEMISTRY · NOVEMBER 1993

Impact Factor: 3.02 · Source: PubMed



**2 AUTHORS**, INCLUDING:

Philip A Gottlieb
University at Buffalo, The State University of …

**75** PUBLICATIONS **2,066** CITATIONS

# Thermodynamic and Alkylation Interference Analysis of the *lac* Repressor–Operator Substituted with the Analogue 7-Deazaguanine†

Xiaolin Zhang and Philip A. Gottlieb*

*Department of Chemistry and Biochemistry, University of Delaware, Newark, Delaware 19716*

*Received March 30, 1993; Revised Manuscript Received August 4, 1993®*

ABSTRACT: Guanine residues in the *lac* operator were substituted with the isosteric analogue of 7-deazaguanine. The observed equilibrium dissociation constants for *lac* repressor binding to substituted operators were measured in 10 mM Tris, 150 mM KCl, 0.1 mM EDTA, and 0.1 mM DTE, pH 7.6, at 25 °C, using either a standard equilibrium assay or a competition assay. Of the seven individual sites tested, only three significantly altered the complex affinity. Two of these sites are symmetrically related and are positioned in the major grooves that are known to interact directly with repressor, and the third site is located in the central core region of the operator. The alkylation interference assay, which identifies essential phosphate sites, was used to define the extent of perturbation on the protein–DNA complex by the modified nucleotide. Chemical footprinting data for the singly substituted operator done at a single concentration of protein demonstrated that these operators were similar to the wild-type pattern, indicating that the loss of affinity reflected the alignment of sterically incompatible groups or disruption in the local conformation, but did not appear to alter general phosphate backbone interactions. Similar experiments with a doubly substituted operator revealed that the free energy of binding was an additive combination of each of the individual sites, and the alkylation interference data were similar to the singly substituted and wild-type operators. A constitutive mutant which also had a higher binding constant had a similar alkylation interference pattern.

Control of gene expression is mediated in part by repressor proteins binding specifically to unique DNA operator sequences. A challenging biochemical problem is to determine how repressor proteins achieve specificity for such sequences. In this regard, two general interactive modes are thought to play crucial roles (Otwinowski et al., 1988). The first is commonly referred to as direct readout (Seeman et al., 1976). This involves the formation of intermolecular hydrogen bonds or alignment of hydrophobic regions to promote binding to a unique region of DNA. Specifically, base pairs from the DNA and amino acid residues from the protein are positioned within the bimolecular complex to create a compatible series of interactive sites along the DNA–protein interface. A second and more novel recognition mode called indirect readout uses water-mediated hydrogen bonds and is based on observations made in the *trp* protein–operator crystal structure (Otwinowski et al., 1988). The protein–DNA complex also uses general structural features such as induced bending, without regard to base-specific functional contact sites, and employs an extensive network of phosphate interactions to stabilize the complex. Although controversy remains regarding the validity of the structure derived as representative of specific binding (Staake et al., 1990; Carey et al., 1991), other crystal structures of repressors in complex with DNA support the idea that indirect effects play a role in the formation of protein–DNA complexes (Aggarwel et al., 1988; Jordan & Pabo, 1988; Harrison & Aggarwal 1990; Schultz et al., 1991).

By expanding the normal repertoire of bases that can be used, isosteric analogues have played an important role in dissecting protein–DNA interaction while preserving DNA structure. There are now numerous examples of modified nucleotides being used to analyze protein–DNA interaction (Aikens & Gumport, 1991). The effect of modified nucleotides on steady-state kinetics has been measured for *Eco*RI (Brennan et al., 1987a,b; McLaughlin et al., 1987; Seela & Kehne, 1987) and for the *Eco*RV restriction endonucelase system (Newmann et al., 1990a,b). Nucleoside analogues have also been used to probe protein–DNA interactions in the *lac* repressor–operator (Goeddel et al., 1977, 1978; Caruthers, 1980), the *trp* repressor–operator (Mazzarelli et al., 1992), and RNA polymerase (Dubbendorf et al., 1987). Recently, a more in-depth analysis has been performed on the *Eco*RI restriction endonuclease by Jen-Jacobson and co-workers (Lesser et al., 1990). Measuring the effect of both analogues and natural base substitutions on the acquisition of the transition-state complex, they showed that both direct and indirect readout modes combined to produce stringent discrimination in the endonuclease–DNA interaction.

In this work, we have investigated the effect of base substitutions on *lac* repressor binding to *lac* operators which had single or double guanine residues replaced with 7-deazaguanine in the *lac* operator. Coupling these results with those of the phosphate alkylation interference which was used to identify essential phosphate sites, we mapped phosphate sites required for *lac* repressor binding to modified operators. Our results indicated that the introduction of a single base substitution altered the observed equilibrium binding constant, but did not appear to alter the essential characteristics of the complex. The double-substituted operator and a constitutive mutant operator which exhibited a loss in bimolecular stability appear to be similar to the wild-type complex in their footprinting pattern.

## MATERIALS AND METHODS

*Synthesis of 7-Deaza-2′-deoxyguanosine Phosphoramidite.* Synthesis of 2-amino-7-(2′-deoxy-β-D-*erthyro*-pentofuranosyl)-1,7-dihydro-4*H*-pyrrolo[2,3-*d*]pyrimidin-4-one (7-

† This work was supported by the University of Delaware Research Fund.

* To whom correspondence should be addressed. Phone: (302) 831-2739. Fax: (302) 831-6335. Email: gottlieb@brahms.udel.edu.

® Abstract published in *Advance ACS Abstracts,* October 1, 1993.

0006-2960/93/0432-11374$04.00/0 

deazadeoxyguanosine) was accomplished according to procedures previously described (Seela et al., 1983a,b, 1987; Seela & Driller, 1986). The nucleoside compound was isolated by silica column chromatography and had physical and chemical characteristics identical to those published previously. The modified nucleoside was converted to the phosphoramidite using the published procedure of Seela and Kehne (1987).

*Synthesis and Purification of Oligonucleotides.* Chemical synthesis was carried out on an Applied Biosystems Inc. Model 381A DNA synthesizer with automated cycles for either 0.2- or 1.0-μmol scale. Removal of protecting groups was achieved by heating the material to 50 °C for 70 h in a 30% ammonia solution. Subsequently, the solvent was removed by evaporation, and the DNA was purified twice by denaturing gel electrophoresis (20% polyacrylamide, 7M urea). Purity of the oligonucleotide was assessed by phosphorylating the DNA with T4 polynucleotide kinase and [$\gamma$-$^{32}$P]ATP and then subjecting the products to denaturing gel electrophoresis (45 V/cm). The DNA purity was assessed by exposing the gel to X-ray film. The wild-type sequence was confirmed by chemical sequencing, and each analogue operator was subjected to purine sequencing reactions (Maxam & Gilbert, 1977). All sequencing reactions were done with duplex DNA.

*DNA Duplex Preparation.* DNA concentration was determined by UV absorbance at 260 nm and corrected for nearest neighbor effects(Brown & Brown, 1991). Radiolabeled duplex DNA was prepared by taking single-stranded DNA (4 pmol) and incubating it with T4 polynucleotide kinase and [$\gamma$-$^{32}$P]ATP at 37 °C for 30 min, followed by ethanol precipitation. Unlabeled complementary strand was added at a 20% excess to the labeled DNA in a final volume of 40 μL of 10 mM Tris, 150 mM KCl, and 0.1 mM EDTA at pH 7.6. The solution was heated to 90 °C and cooled slowly. The DNA was then stored at 4 °C overnight. The excess DNA was required for quantitative annealing of the labeled strand. The duplex DNA was assessed by 10% PAGE and was found to be greater than 99% in duplex form. The concentration of operator was based on the labeled strand.

*Protein Activity.* The activity of *lac* repressor protein was measured at a concentration of protein and DNA above the equilibrium dissociation constant. Wild-type duplex DNA at a concentration of $1 \times 10^{-8}$ M was incubated with varying dilutions of repressor protein, whose concentration was previously determined by UV absorbance ($2.3 \times 10^{-4}$ M), in binding buffer and at 25 °C. All the samples were incubated for 30 min and then filtered through nitrocellulose, dried, and counted. Each protein concentration tested was done in duplicate. The total counts retained versus the dilution was plotted, and filter efficiency was determined by comparing the total amount of DNA added to each sample and the quantity retained after filtration at saturation. The activity of the protein was extracted from the linear portion of this graph, assuming a one to one complex. The *lac* repressor activity was between 80% and 100%. Repressor protein was a gift of Mitch Lewis, Helen Pace, and Ponzy Lu from the University of Pennsylvania. Purification procedures for the protein have been described (Lewis et al., 1990). The mutant repressor *lac*$^{adi}$ was a gift of Michael Brenowitz from the Albert Einstein School of Medicine.

*Equilibrium Dissociation Assays.* The equilibrium binding constant was measured using two different methods (Riggs et al., 1970; Lin & Riggs, 1972, 1975). The first was a standard concentration assay where complex formation was measured as a function of increasing protein concentration. Briefly, the wild-type or analogue operator (2 pmol) was labeled (see above), and radioactive ATP was removed by passing the reaction mixture through a Sephadex G-50 column equilibrated with 10 mM Tris, pH 7.6, 150 mM KCl, and 0.1 mM EDTA.

Equilibrium measurements were performed by incubating a constant concentration of operator ($1 \times 10^{-11}$ M) with an increasing concentration of protein, in a final volume of 100 μL of solution containing 10 mM Tris, pH 7.6, 150 mM KCl, 0.1 mM EDTA, 1 mM DTE, and 50 μg/mL BSA. After 20 min at ambient temperature, the mixture was filtered on presoaked nitrocellulose filters, washed once with 500 μL of wash buffer (10 mM Tris, pH 7.6, 150 mM KCl, and 0.1 mM EDTA), and dried. Binding assays were performed at various times up to 2 h without any change in the measured constant. Duplicate samples of protein concentrations were tested, and each assay was repeated at least twice. Background was determined by the addition of the inducer IPTG (isopropyl β-D-thiogalactoside, 1–10 mM) or by treating the operator DNA in the absence of protein identically to the binding reactions. The background counts were identical for both methods. The dried filters were placed in scintillation cocktail, and the amount of radioactivity was determined by scintillation counting. To preclude any effects by single-stranded DNA on complex formation, labeled DNA was incubated with various concentrations of protein and was found not to be retained on nitrocellulose filters. The results were fitted as a binding curve using nonlinear regression analysis with SIGMAPLOT software for the equation derived below:

$$K_D = \frac{[R]_f[O]_f}{[RO]} \quad (1)$$

Since

$$[R]_f = [R]_t - [RO]; \quad [O]_f = [O]_t - [RO] \quad (2)$$

then

$$K_D = \frac{([R]_t - [RO])([O]_t - [RO])}{[RO]} \quad (3)$$

where $[R]_t$ and $[O]_t$ are the total concentrations of repressor and operator, respectively. Solving the quadratic equation for a realizable root yields the following equation:

$$[RO] = [K_D + [O]_t + [R]_t - ((K_D + [O]_t + [R]_t)^2 - 4[R]_t[O]_t)^{1/2}]/2 \quad (4)$$

The maximum operator concentration available for binding was determined from the complex retained on nitrocellulose filters at saturating amounts of repressor protein. Filter efficiency was measured by comparing total input counts to the amount retained on the filter at saturation and was approximately 50% of the total counts.

The second approach used a competition assay which is similar to a previously described method (Lin & Riggs, 1972). The equation for competition is

$$[RO] = \frac{[R]_t([O]_t - [RO])}{K_{RO}\left(1 + \frac{[C]_t}{K_{RC}}\right) + ([O]_t - [RO])} \quad (5)$$

where $[R]_t$, $K_{RO}$, $K_{RC}$, $[C]_t$, and $[O]_t$ are the total protein concentration, the equilibrium dissociation constant for wild-type operator, the equilibrium dissociation constant for competitor, the total competitor DNA concentration, and total labeled operator concentration, respectively. Solving the

```
                                  1    5    10   15   20
O+:   5' CTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCACACACAG      3'
      3'      ATATTACACACCTTAACACTCGCCTATTGTTAAAGTGTGTGTCCTTC 5'
                                       ⇧
                                  1    5    10   15   20
O8:   5' CTCGTATAATGTGTGGAATTATTGTTAGACAATAATTTCACACACAG      3'
      3'      ATATTACACACCTTAATAACAATCTGTTATTAAAGTGTGTGTCCTTC 5'
                                  1    5    10   15   20
OIN:  5' CTCGTATAATGTGTGGTTAACAATAGGCGAGTGTTAATCACACACAG      3'
      3'      ATATTACACACCAATTGTTATCCGCTCACAATTAGTGTGTGTCCTTC 5'
```

FIGURE 1: The 43-base-pair segments of the UVlac5 DNA sequence used for binding measurements with the *lac* repressor. The first sequence, $O^+$, is the wild-type *lac* operator, which has the *lac* operator sequence highlighted and the axis of symmetry marked by an arrow. The second sequence ($O^8$) is a duplex that contains all the known constitutive mutations, and the third operator is an inverted sequence for the 21-base-pair operator ($O^{IN}$). The latter two sequences were used to measure the level of nonspecific binding of *lac* repressor to DNA.

quadratic equation explicitly for [RO] yields

$$[RO] = \left[K_{RO} + \frac{K_{RO}}{K_{RC}}[C]_t + [R]_t + [O]_t - \left(\left(K_{RO} + \frac{K_{RO}}{K_{RC}}[C]_t + [R]_t + [O]_t\right)^2 - 4[O]_t[R]_t\right)^{1/2}\right]/2 \quad (6)$$

The competition curve of the complex formed as a function of increasing concentrations of competitor DNA is fit to the best value for the dissociation constant $K_{RC}$ using a nonlinear least-squares fit. The value $[O]_t$ (the total available operator for binding) was derived from the amount of complex available at saturating concentrations of protein. The concentrations of wild-type repressor and operator were chosen which set the initial value of [RO] ($[RO]_0$) between 50% and 80% of the total counts ($[O]_t$). Typically, wild-type operator DNA ($1.0 \times 10^{-11}$ M) and repressor ($5.0 \times 10^{-11}$ M) were incubated with varying concentrations of competitor DNA in the buffer described above. After 30 min at ambient temperature, samples were filtered and treated in a manner identical to that described above for the equilibrium assay. The starting concentration $[RO]_0$ can be varied with no effect on the derived equilibrium dissociation constant. Duplicate samples were performed for each concentration, and each operator sequence was tested at least three times. The average values for the observed equilibrium constant derived from the competition assay and the standard deviation of the mean are listed in Table I. Using the wild-type operator as competitor DNA, the observed equilibrium dissociation constant was measured and was in good agreement with values obtained for the standard equilibrium assay at $1.2 \times 10^{-11}$M. Single-stranded DNA was also tested as competitor and found unable to affect the $[RO]_0$ complex, even at concentrations of single-stranded DNA in the micromolar range.

*Phosphate Alkylation Interference.* A modified version of a protocol previously described allowed DNA to be alkylated randomly (Siebenlist & Gilbert, 1980). Basically, duplex DNA (4 pmol) was labeled on one strand (see above) and then alkylated by the addition of a saturated solution of *N*-ethylnitrosourea at 70 °C for 10 min. The material was then precipitated once by the addition of 25 μL of 3 M NaOAC and 700 μL of ethanol (80%), cooled to −20 °C for 15 min, centrifuged for 15 min, decanted, and rinsed twice with ethanol. The material was allowed to dry, and 40 μL of binding buffer was added. Repressor protein was added and allowed to equilibrate for 20 min. The bound and the unbound complexes were separated using non-denaturing (10%, 99:1) polyacrylamide gel electrophoresis (Fried & Crothers, 1981; Garner & Revzin, 1981) in the absence of loading dyes. The bound and free bands were excised from the gel, and the DNA was extracted from the gel slices with TE (10 mM Tris–1 mM EDTA, pH 8.0) buffer. The solution containing the DNA was collected and precipitated. Cleavage of the material was achieved by incubation with 20 μL of 10 mM phosphate buffer and 150 mM NaOH for 30 min at 90 °C. To determine if one-hit kinetics per DNA was achieved, the quantity of intact duplex DNA before and after hydrolysis was measured and found not to exceed 30% of the total operator, indicating that most of the material remains unaltered by the limited incubation with the chemical modifying agent. All DNA samples were adjusted to the same specific activity by the addition of standard dyes in 80% formamide, and each sample was heated to 90 °C for 5 min. The same volume for each of the operators (identical cpm) was applied to a 20% polyacrylamide gel containing 7 M urea and then exposed to X-ray film. All footprinting experiments were repeated at least twice, and the wild-type operator served as a control for all footprinting studies. Bands in the interference pattern were identified by sequencing reactions.

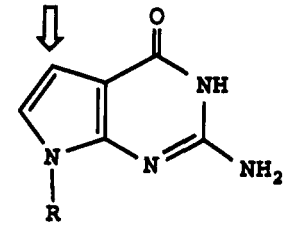


FIGURE 2: Chemical structure of the nucleoside analogue 7-deazadeoxyguanosine. The arrow indicates the position at which a carbon has been substituted for nitrogen. The letter R symbolizes the 2′-deoxyribose sugar.

## RESULTS

*Wild-Type and Modified Operators.* Operator sequences with or without the modified base were constructed from two complementary synthetic strands. Each strand was 47 bases, and after annealing, a 43-base-pair duplex was formed with a 4-base overhang at both ends. The sequence was identical to the segment of the *lac* operon called UV-5 and is shown in Figure 1. Embedded within this sequence is the wild-type operator, which is emphasized in bold type and defined by a numbering system to be used throughout this work. The axis of two-fold pseudosymmetry is indicated as well. Seven different substituted operator sequences were synthesized, in which a single guanine residue was replaced by 7-deazaguanine (Figure 2). These operators had a single substitution at position 5, 7, 9, 11, or 12 in the top strand or at position 10 or 17 in the bottom strand.

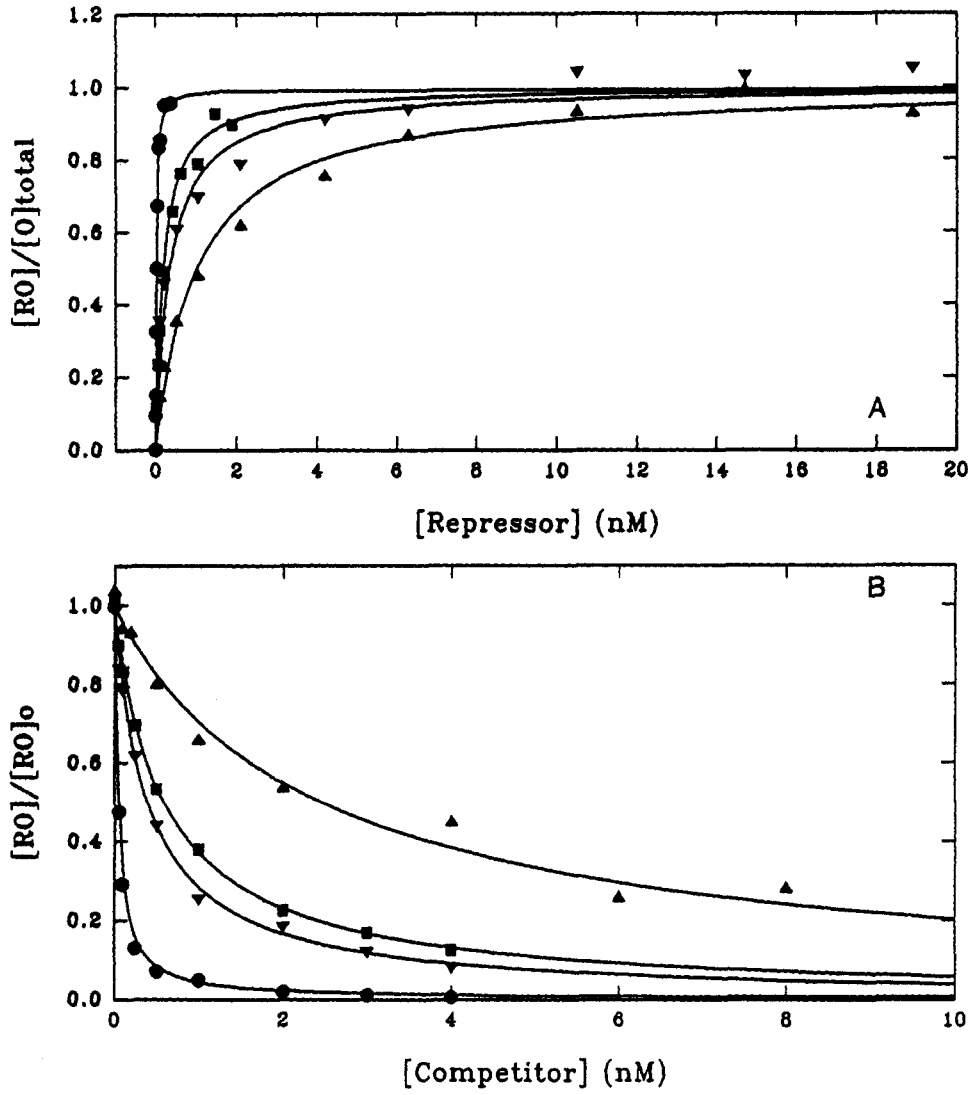


FIGURE 3: Measurements of the equilibrium binding constant for wild-type (●) and 5- (▲), 12- (▼), and 17-substituted operators (■) using (A) a standard equilibrium binding assay, which measures the amount of complex formed as a function of protein concentration, and (B) a competition assay, which measures the competition between wild-type and substituted operators for repressor protein. In each case, the solid line is the best fit to the data.

*Equilibrium Dissociation Constants.* The equilibrium dissociation constants were determined for the wild-type and analogue operators at 150 mM KCl, pH 7.6, and 25 °C. The observed equilibrium dissociation constants ($K_{obs}$) were derived from binding data depicted, for example, in Figure 3A. The results for all the operator sequences tested are summarized in Table I. The wild-type operator showed a $K_{obs} = 1.5 \times 10^{-11}$ M. These values are consistent with those previously published for this size fragment and this concentration of monovalent salt (Whitson et al., 1986; Whitson & Shive-Matthews, 1986). Only three of the operators having a single substitution demonstrated significant changes in affinity (greater than 2-fold). These were substitutions at positions 5 and 17 of the operator, which are related by a dyad axis of symmetry, and of the guanine residue at position 12 (Figure 1). The affinities with position 5 and 17 substitutions were reduced by approximately 73-fold and 15-fold, respectively, and the position 12 modified operator was reduced by 18-fold.

Each of the above operator sequences was tested by the competition method previously formulated by Lin and Riggs (1972) for protein–DNA interaction. We used relatively low concentrations of reactants; both protein and labeled operator were typically in the concentration range of $10^{-11}$ M. The assay tested the ability of the modified unlabeled operator sequence to compete with labeled wild-type operator for *lac* repressor protein under equilibrium conditions. This was accomplished by equilibrating varying concentrations of the modified DNA with a preset amount of wild-type protein–DNA complex. As the test duplex increased in concentration, less wild-type operator was retained on the filters, thereby allowing the competitive strength of the test operator to be assessed. For single substitutions that were tested, $K_{obs}$ values derived from the competition experiments (Figure 3B) were similar to those from the standard equilibrium assay described above. The competition binding experiment had a $\Delta\Delta G°$ of approximately 1.5 kcal/mol at substituted positions 12 and 17, while the position 5 substitution had a value of 2.4 kcal/mol.

Four additional operators were tested. The first was a *lac* operator sequence that contained two analogue substitutions, one at position 5 and the other at position 17, designated operator 5/17. The affinity of operator 5/17 was reduced by approximately 1770-fold ($\Delta\Delta G° = 4.4$ kcal/mol) in the standard equilibrium assay and 560-fold ($\Delta\Delta G° = 3.8$ kcal/mol) in the competition assay (Figure 4 and Table I). The second operator was a transition mutation at position 5 referred to as 5(A-T) and is a well characterized constitutive operator sequence (Mossing & Record, 1985). This operator had a

Table I: Equilibrium Dissociation Constants for the Wild-Type and Modified Operator Sequences

```
                1    5    10   15   20
5'              AATTGTGAGCGGATAACAATT        3'
3'              TTAACACTCGCCTATTGTTAA        5'
```

| operator sequence | equilibrium assay $K_{obs}$[a] ($\times 10^{-11}$ M) | ratio | equilibrium assay $\Delta\Delta G°$(25 °C)[b] (kcal/mol) | competition assay $K_{obs}$[a] ($\times 10^{-11}$ M) | ratio | competition assay $\Delta\Delta G°$(25 °C)[b] (kcal/mol) |
|---|---|---|---|---|---|---|
| wild type | 1.5 ± 0.1 | 1 | | 1.2 ± 0.1 | 1 | |
| 17 | 23.0 ± 6.0 | 15 | 1.6 ± 0.2 | 17.4 ± 3.6 | 14.3 | 1.6 ± 0.2 |
| 12 | 28.0 ± 14 | 18 | 1.7 ± 0.3 | 12.3 ± 1.0 | 10 | 1.4 ± 0.1 |
| 11 | 2.4 ± 0.4 | 1.6 | 0.3 ± 0.1 | 1.4 ± 0.6 | 1.2 | 0.1 ± 0.2 |
| 10 | 3.0 ± 1.2 | 2.0 | 0.4 ± 0.2 | 1.7 ± 0.3 | 1.4 | 0.2 ± 0.2 |
| 9 | 2.3 ± 0.5 | 1.5 | 0.2 ± 0.2 | 1.7 ± 0.4 | 1.4 | 0.2 ± 0.2 |
| 7 | 2.0 ± 0.6 | 1.3 | 0.2 ± 0.2 | 1.6 ● 1.0 | 1.3 | 0.2 ± 0.1 |
| 5 | 110 ± 20 | 73 | 2.5 ± 0.2 | 70 ± 22 | 58 | 2.4 ± 0.2 |
| 5/17 | 2650 ± 1150 | 1770 | 4.4 ± 0.3 | 677 ± 231 | 560 | 3.8 ± 0.2 |
| 5(A-T) | 1400 ± 330 | 930 | 4.0 ± 0.2 | 285 ± 53 | 236 | 3.2 ± 0.2 |
| $O^8$ | ND | | | 44 300 ± 11 400 | 36 900 | 6.2 ± 0.2 |
| $O^{IN}$ | ND | | | 35 100 ± 8 200 | 29 300 | 6.1 ± 0.2 |

[a] Binding was measured in 10 mM Tris and 150 mM KCl, pH 7.6, at 25 °C. Values tabulated above are the average of at least three independent determinations and the standard deviation. The equilibrium constants assume one binding site per DNA molecule. [b] $\Delta\Delta G°(25\ °C) = RT \ln[K_{obs}(\text{substitution})/K_{obs}(\text{wild type})]$ at 25 °C.

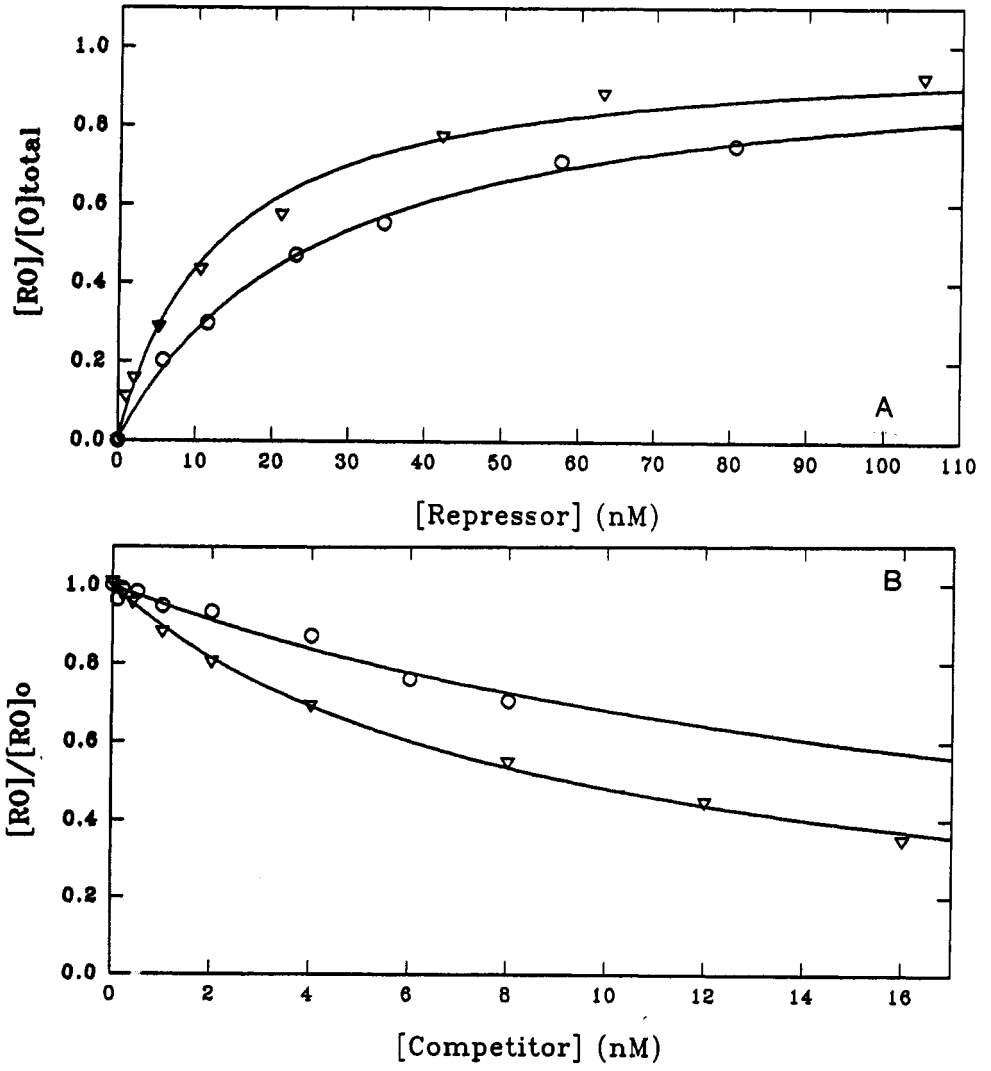


FIGURE 4: Measurements of the equilibrium binding constant for an operator with modifications at positions 5 and 17, operator 5/17 (○), and an operator with a transition mutation at position 5, called 5(A-T) (▽), using (A) a nitrocellulose binding assay which measures the amount of complex formed as a function of protein concentration and (B) a competition assay which measures the competition between wild-type and modified operators for repressor. In each case, the solid line is the best fit to the data.

930-fold weaker affinity ($\Delta\Delta G°$ = 4.0 kcal/mol) as measured in the standard equilibrium assay, while in the competition assay the relative affinity was 236-fold reduced ($\Delta\Delta G°$ = 3.2 kcal/mol) (Figure 4 and Table I). The third operator sequence

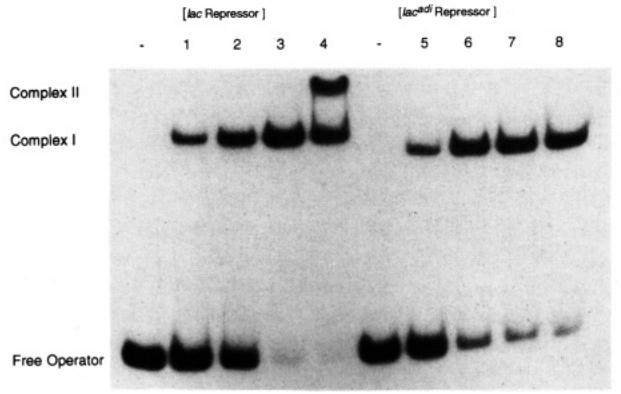


FIGURE 5: Separation of the bound and unbound operator sequences of previously alkylated DNA using polyacrylamide gel electrophoresis. DNA was labeled at the 5′ end of the top strand only, and DNA concentration of the duplex was $2.5 \times 10^{-8}$ M. Repressor concentrations used were (1) $5.5 \times 10^{-9}$, (2) $1.1 \times 10^{-8}$, (3) $2.75 \times 10^{-8}$, and (4) $1.1 \times 10^{-7}$ M. DNA in the absence of protein was used to identify free operator. The mutant *lac*$^{adi}$ repressor was also incubated with operator DNA at the following concentrations: (5) $9 \times 10^{-9}$, (6) $2.3 \times 10^{-8}$, (7) $4.5 \times 10^{-8}$, and (8) $1.4 \times 10^{-7}$ M. This allowed us to identify the dimer repressor bound to operator (complex I) and tetramer repressor bound to operator (complex II). Equivalent results were obtained with labeled DNA at the 5′ end of the bottom strand.

tested, $O^8$, contained all eight constitutive mutations, and the fourth, $O^{IN}$, was an inverted operator. These latter two operator sequences were designed to completely destroy specific binding within the operator sequence, in order to determine approximate binding constants for DNA of this size which has only nonspecific binding. Using the competition assay, the observed equilibrium constant was between $3.5 \times 10^{-7}$ and $4.4 \times 10^{-7}$ M with an approximate $\Delta\Delta G°$ value of 6.0 kcal/mol (see Table I). This is 2 kcal/mol more unfavorable than for operator 5/17, which has the weakest affinity of all the operators tested, suggesting that the contribution by nonspecific binding in assessing the specific binding of the operators was minimal.

*Phosphate Alkylation Interference Pattern.* Phosphate groups that interfered with complex formation, for either the wild-type or the modified operator, were identified using the method of phosphate alkylation. DNA operator duplexes containing randomly introduced alkylated sites were incubated with varying concentrations of *lac* repressor, and the bound and unbound operator sequences were separated using nondenaturing polyacrylamide electrophoresis (Figure 5). The samples were loaded in the absence of loading dyes. We have found that the addition of either bromphenol blue (0.01%) or xylene cyanol (0.01%) can significantly destabilize the repressor–operator complex (data not shown).

Figure 5 shows the gel retardation of the protein–DNA complex of alkylated DNA ($2.5 \times 10^{-8}$ M) with increasing concentrations of repressor protein ($5.5 \times 10^{-9}$ to $1.1 \times 10^{-7}$ M) or with a mutant form of repressor called *lac*$^{adi}$ ($9 \times 10^{-9}$ to $1.4 \times 10^{-7}$ M). The latter protein is incapable of forming tetrameric repressor and exists exclusively in a dimeric form at these concentrations of protein (Brenowitz et al., 1991; Chen & Matthews, 1992; Fickert & Müller-Hill, 1992). At the highest concentration of wild-type repressor used in this assay, two bound bands were evident (Figure 5) with no free DNA operator available. The faster of the two complexes (complex I) migrated to approximately the same position that was seen for the *lac*$^{adi}$ mutant protein, indicating that at this concentration of wild-type protein binding was exclusively in the form of dimer repressor. A second slower migrating band (called complex II) was presumed to be the tetrameric form of repressor binding to the operator sequence.

The patterns generated from alkylated DNA after hydrolysis of the bound and unbound populations are shown in Figure 6. These experiments were done at high concentrations of DNA and protein and will therefore give detailed information about interference sites which affect the formation of complex at 2–3 orders of magnitude above the wild-type binding constant. In the top strand the percent of bound material for increasing protein concentration was 20%, 43%, and 100% (Figure 5, lanes 1–3, respectively). For the bottom strand, binding was 11%, 27%, and 100% (data not shown). Figure 6A shows the interference sites derived from each of these DNA populations. In the top strand, positions $T_3pT_4$, $T_4pG_5$, $G_5pT_6$, $G_{12}pA_{13}$, $G_{11}pG_{12}$, and $C_{10}pG_{11}$ showed interference with complex formation. The bottom strand had a series of interrference sites at $T_{19}pT_{18}$, $T_{18}pG_{17}$, $G_{17}pT_{16}$, $C_{11}pG_{10}$, $G_{10}pC_9$, $C_9pT_8$, $T_8pC_7$, and $C_7pA_6$. We were able to segregate these sites into arbitrary categories of "strong" and "weak" sites by lowering the concentration of DNA ($2 \times 10^{-10}$ M with protein concentrations of $2–20 \times 10^{-10}$ M (data not shown). The strong interference sites are $T_3pT_4$, $T_4pG_5$, and $G_{11}pG_{12}$ on the top strand and $T_{19}pT_{18}$, $T_{18}pG_{17}$, $G_{10}pC_9$, and $C_9pT_8$ on the bottom strand. These sites are identical to those found in previous studies [see Figure 4 of Barkley and Bourgeouis (1980)].

The category of strong sites can be further subdivided on the basis of ability of the alkylated operator to be saturated at a specific concentration of protein. For example, the site $G_{11}pG_{12}$ is no longer part of the interference pattern at the high concentration of protein (100% bound), while two phosphate sites located at the outer extremities ($T_3pT_4$ and $T_{19}pT_{18}$) of the operator site on each strand are still evident (Figure 6A, lane 3, and Figure 6B, lane 3). The mutant repressor used to identify the dimer form of repressor gave results similar to those with the wild-type operator (data not shown) but required higher concentrations to achieve the same effect, presumably due to its different binding constant.

All the singly substituted operator sequences had interference patterns similar to wild-type operator at the concentrations of protein–DNA tested, especially for the strong sites. The doubly substituted operator 5/17 and the constitutive mutant 5(A-T) were analyzed in the same manner (Figure 7). The first bound band of the doubly substituted operator (complex I) indicated that the interference pattern is similar to that of the wild type. These experiments indicated that the essential phosphate sites required for the formation of the wild-type complex were present in these modified operator sites as well. However, the analysis was complicated by the modified base, because treatment of this DNA with *N*-ethylnitrosourea followed by the alkali reaction enhanced the cleavage at the modified base along the phosphate chain. Since the interference assay relies on single-hit kinetics, hydrolysis of the DNA should be equally efficient for all DNA phosphate sites; the more rapid rate of hydrolysis was a potential source of error. We have measured this enhanced rate associated with the presence of the modified nucleotide and have determined that it represents a 3-fold increase compared to the natural guanine base. Given that no more than 30% of the molecules are cleaved in these reactions, each phosphate site will represent less than 1% of the entire population of operator sequences sampled. The enhanced rate represents approximately 3% of the entire population and consequently will not significantly enhance the population of molecules that undergo two cleavage reactions instead of one, and therefore

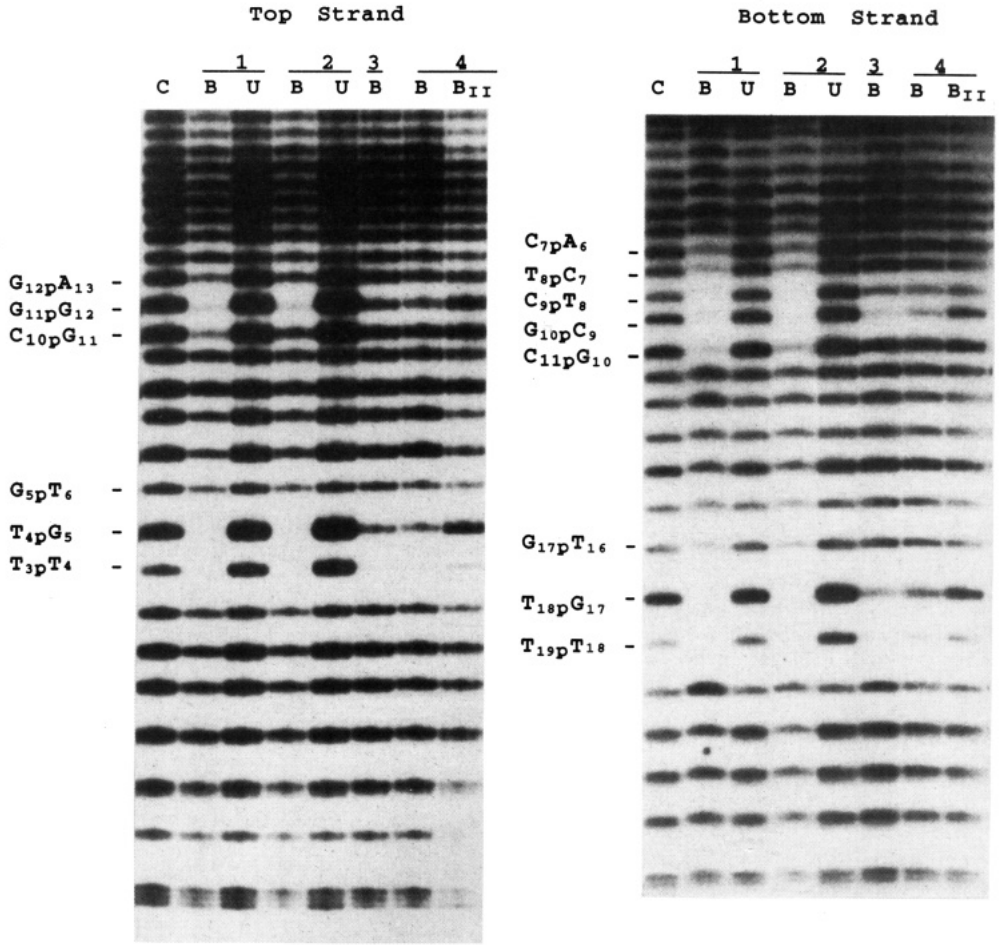


FIGURE 6: Bound and unbound DNA were hydrolyzed and applied to a sequencing gel. Where possible, bound (B) and unbound (U) products from a single protein concentration were positioned side-by-side. At saturating amounts of protein, a slower moving complex appeared ($B_{II}$) which was analyzed alongside complex I (B). The alkylated DNA alone is displayed in the control lane (C). The numbers above the lanes refer to the following concentrations of protein used: (1) $5.5 \times 10^{-9}$, (2) $1.1 \times 10^{-8}$, (3) $2.75 \times 10^{-8}$, and (4) $1.1 \times 10^{-7}$ M. Panel A (left) shows the results of hydrolysis of the top strand, and panel B (right) shows the results of hydrolysis of the bottom strand.

it does not affect the analysis. In addition, the modified nucleotide position appears to undergo some base breakdown, altering the migration of that DNA, shifting the band to a position in the gel that is slightly above the adjacent phosphate (see asterisk, Figure 7).

We also attempted to identify interference sites on the operators containing all eight constitutive mutants and the inverted site. In both cases, separation of the protein–DNA to isolate bound and unbound material produced two populations of DNA: one migrated as free DNA, and the other migrated as complex II. We were unable to detect any complex I with these mutant operator sequences, and hydrolysis of complex II did not provide a specific interference pattern. In general, once the protein concentration reached sufficiently high levels to allow the formation of the tetrameric form of the repressor, the footprinting pattern resembled the nonspecific binding pattern.

## DISCUSSION

*Considerations regarding DNA Fragment Size.* Thermodynamic measurements of *lac* repressor binding to its operator sequence are strongly influenced by the size of the DNA fragment. Shive-Matthews and co-workers have shown that a 40-base-pair fragment containing the intact operator has nearly all the thermodynamic properties associated with larger fragments of DNA that contain the *lac* operator (Whitson et al., 1986; Whitson & Shive-Matthews, 1986). Earlier work with *lac* operators of lengths between 21 and 26 base pairs did not display all the thermodynamic properties observed in larger fragments. Thus, measured counterion release upon formation of the *lac* operator–repressor complex was smaller in shorter fragments than observed in longer DNA fragments (Goeddel et al., 1977; Record et al., 1977; O'Gorman et al., 1980; Whitson et al., 1986). The limitation of fragment size was further demonstrated by thermodynamic and kinetic measurements of a transition mutation at position 8, which in a 21 base-pair duplex produced a smaller effect on the association of the complex than the same mutation in a larger DNA molecule (Mossing & Record, 1985; Fisher & Caruthers, 1979). On the basis of these observations, we used a base analogue incorporated into operator DNA that was embedded within a fragment of DNA greater than 40 base pairs.

*Comparison of Competition and Standard Equilibrium Dissociation Assays.* In general the values obtained for equilibrium and competition assays agree (Table I). However, significant differences were observed between these two assays for weakly interacting protein–operator complexes. For example, the affinity of the constitutive mutant at position 5 in the standard equilibrium assay was 0.8 kcal/mol greater than that observed in the competition assay. This difference in affinity could be due to problems inherent in the filter binding assay (Woodbury & von Hippel, 1983). Theoretically, the interpretation of data is not unambiguous when the complex is affected by the efficiency of the assay, the cooperativity of

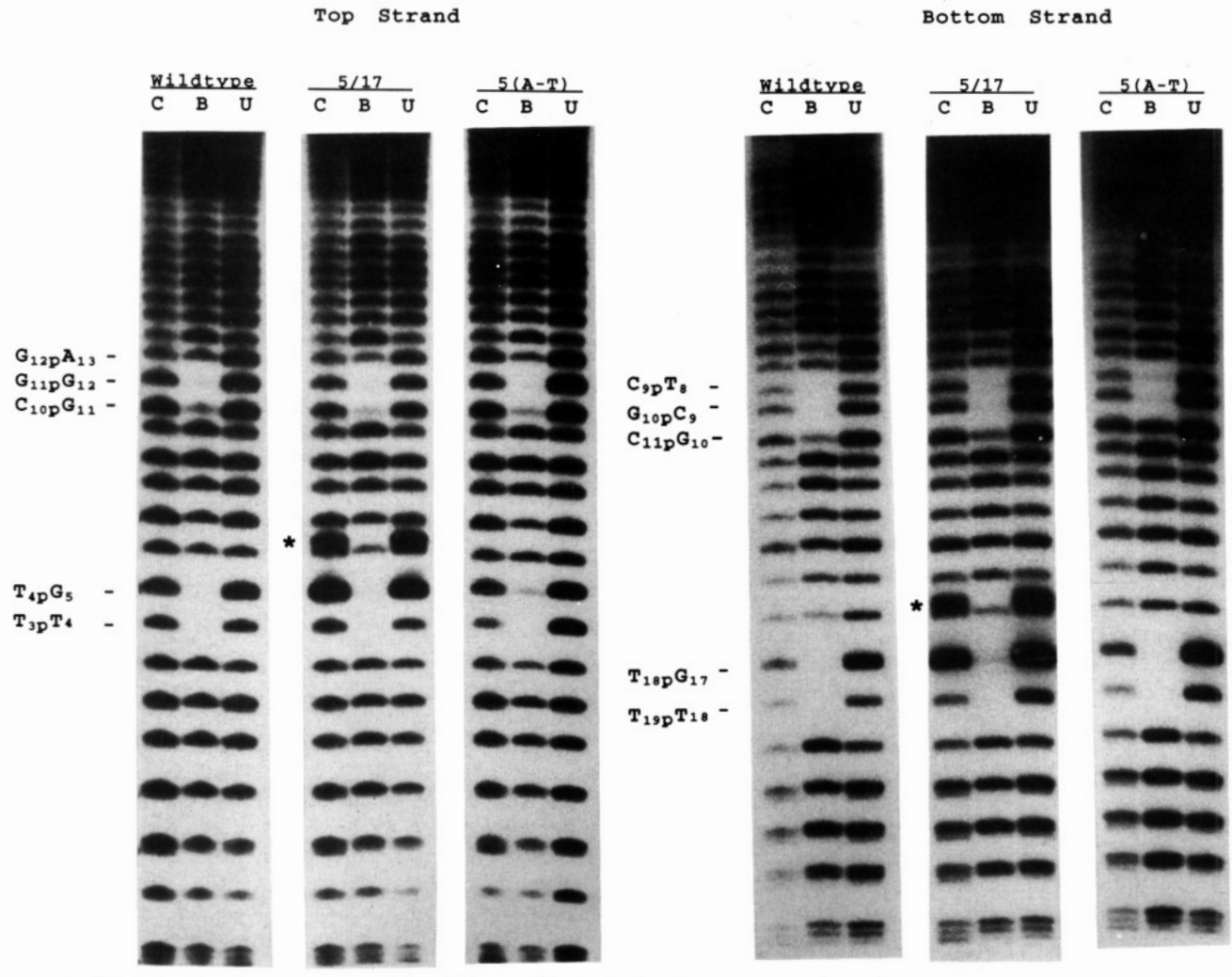


FIGURE 7: Alkylation interference pattern of the wild-type, 5/17, and 5(A-T) operators. Each operator has three lanes: C is the control lane, which is the alkylated DNA; B is the alkylated DNA isolated from the bound population, called complex I; and U is the alkylated DNA isolated from the unbound population. The asterisk to the left of the C lane for the 5/17 operator marks a band that migrates anomalously. The pattern of bands in panel A (left) is from the top strand of each operator, and the pattern in panel B (right) is from the bottom strand.

the protein, nonspecific binding, or the stoichiometry of binding. We have assessed the contribution of nonspecific DNA binding by using an equivalent-sized fragment of DNA which carries a sequence that is unrelated to the *lac* operator and have found that retention of the nonspecific complex occurs only at a much higher concentration of protein than that used for our measurements, indicating that flanking sequences did not play a role in binding (Table I). However, any of the other parameters mentioned may contribute to the differences observed in the binding assays.

For example, the competition experiment was done at concentrations of protein in the range of $5 \times 10^{-11}$ M. Depending on the equilibrium constant used for the dimer to tetramer transition, the *lac* repressor protein is either totally in dimer form (Royer et al., 1990), partially in dimer form (Brenowitz et al., 1991), or virtually all in tetramer form (Fickert & Müller-Hill, 1992). To assess the contribution of the equilibrium constant, we measured the availability of multimeric forms of repressor by comparing a mutant repressor protein that is incapable of forming a tetramer with the wild-type repressor. We have shown that both alkylated and unmodified DNA bind protein primarily in the dimeric form at protein concentrations below $10^{-8}$ M in the conditions used for our assay (Figure 5). Since the competition method used protein concentrations below this value, our assay is measuring the dimer form of repressor only. In the standard equilibrium assay, the increase in protein concentration around the dimer to tetramer equilbrium constant would give rise to an apparent reduction in protein concentration which would contribute to a small error in the standard equilibrium assay.

Another potential source of error is the possibility that the increased size and surface area of the protein in the tetramer form would enhance the efficiency of retention on nitrocellulose filters. We have observed that for weaker binding operators efficiency of retenetion increases from 50% to 80%. This would compound the error in the standard equilibrium binding assay, since saturation is determined at high protein concentrations. This can account for the differences in the binding constant measured for some operators in the two assays, and it is consistent with the fact that both assays produced identical observable binding constants with high-affinity operator sites which require low concentrations of protein. The advantage of the competition assay is that the wild-type operator–protein complex is used as a standard, and the modified operator is tested against this standard in the same solution and under the same conditions. In addition, the low concentrations of reactants would prevent the formation of a complex with tetrameric forms of repressor, and there is always an excess of DNA to prevent aggregation of the protein. A problem that may be relevant to the competitor assay is the assumption that binding by the competitor DNA is by a single high-affinity site. As the operator deviates in composition from

the wild-type sequence and begins to resemble nonspecific DNA, the assumption of one high-affinity binding site will not be valid. However, footprinting results and nonspecific binding measurements indicate that the assumption of a single available site is valid for the sequences being used here, except in the case of the $O^8$ and $O^{IN}$ sequences.

*Singly Substituted Operators Identify Essential N7 Positions in the Operator.* Single substitutions of the analogue at three positions significantly increased the observed equilibrium dissociation constant compared to wild-type levels. The effect of substitutions at positions 5 and 17, which are symmetrically related along a pseudosymmetric dyad axis, were consistent with previous data which indicated that guanine residues at these positions are essential for forming the protein–DNA complex. Multidimensional NMR measurements for the *lac* headpiece (residues 1–56) in complex with half an operator sequence showed that the *lac* repressor headpiece has an orientation opposite that found in many known repressor–operator complexes (Boelens et al., 1987; Lamerichs et al., 1989). Modeling of the *lac* headpiece to a half-operator site indicated that the $Arg^{22}$ residue of the protein was in close proximity to the $GC^5$ base pair and presumably allowed the guanidinoyl group to form a bidentate hydrogen bond between the N7 and the O6 of the guanine residue. Müller-Hill and co-workers, using genetic experiments, extended these observations to the intact protein (Sartorius et al., 1989); substitutions that were made within the DNA at position 5 disrupted the *in vivo* binding and were reversed by substitutions made in the protein at position 22. Our data confirm the importance of the guanine residues at these positions and further imply that the N7 is directly involved in stabilizing the interaction. Recently, we have shown that the O6 position, when removed by the insertion of a 2-aminopurine analogue, perturbs the binding at these sites, supporting the proposed interaction between Arg and the GC base pair (X. Zhang and P. A. Gottlieb, unpublished results). Despite the symmetric positioning of residues at positions 5 and 17, binding is not affected to the same degree by substitution at these positions. The free energy change is approximately 1 kcal/mol greater when the analogue is substituted at position 5 (2.4 kcal/mol) than when the analogue substitution is at position 17 (1.6 kcal/mol). This dissimilarity is consistent with previous observations in which substitutions that are inserted in the wild-type operator at equivalent positions along the symmetry axis do not respond identically (Caruthers, 1980).

A third site which was affected by the substitution of the analogue was identified at position 12, where a change of 10–18-fold in affinity was observed, depending on the assay used. Previous experiments *in vivo* and *in vitro* have indicated that this position has no direct function for interaction with the protein, since substitutions incorporated at this site are without loss of function or complex stability (Betz et al., 1986). It was proposed that this site acts as a spacer to properly align major-groove sites in each half of the DNA operator. Since the position of the nucleotide is in the central region of the operator, it is conceivable that the insertion of a modified nucleotide affects structural features such as the flexibility of the DNA. This was observed for the 434 repressor binding to its operator; the insertion of a G·C base pair instead of an A·T base pair results in significant changes in the affinity (Koudelka et al., 1987, 1988). For *lac* repressor, such an effect would presume that the modified analogue affects the DNA more than any transition or transversion mutation would at this position, a fact which is not consistent with experiments that indicate that the DNA structure is influenced minimally by modified nucleotides. Our results predict that the repressor protein, when interacting with the wild-type operator, must form a hydrogen-bonding site with this guanine residue at the N7 position. The magnitude of the change in complex stability with this substitution is consistent with hydrogen-bonding perturbations found for other protein–DNA interactions (McLaughlin et al., 1987; Mazzarelli et al., 1992; Newmann et al., 1990a,b; Goeddel et al., 1977, 1978; Caruthers, 1980; Lesser et al., 1990).

*The Alkylation Interference Pattern for the Wild-Type Operator Is Protein Concentration Dependent.* We have investigated the phosphate requirements for complex formation using the alkylation interference assay. These phosphate groups appear to segregate into two arbitrary classes of strong and weak interference sites. These data are similar to those for concentration dependence interference sites demonstrated for the *Eco*RI restriction endonuclease sites (Becker et al., 1988). Two of the strong sites in the repressor–operator complex, even at a protein concentration of $1.1 \times 10^{-7}$ M, which results in virtually no unbound DNA, were never fully saturated in complex I or complex II. These two sites are the same ones that, when replaced with methyl phosphonates, resulted in major increases in the observable binding constant (Noble et al., 1984). They appear to be essential for repressor–operator interaction, since every population of alkylated operators tested had these phosphate interference sites. Recently, Fickert and Müller-Hill (1992) have argued that a specific series of contacts within the nucleobases must be formed for the repressor to locate the correct DNA sequence. Our results suggest that, for repressor to locate and bind operator sequences within a large fragment of DNA, these two phosphate sites must be available. Either these phosphate sites (and others) are simultaneously positioned with the nucleobase sites or perhaps these phosphates are the initial signal for proper alignment of the repressor protein. Currently, a more systematic study is underway to further clarify the role of these phosphates. These sites represent either electrostatic interactions, hydrogen-bonding sites, or regions where structural transitions occur in the DNA to facilitate the formation of the wild-type complex.

*Phosphate Interference Sites in Variant Complexes Are Similar.* The effects of the nucleotide analogue on the *lac* operator interaction with repressor can affect a number of "invisible" forces required for stability; the formation of a protein–DNA complex is favored by counterion and water release while paying energetic penalties for structural reorganization (Record et al., 1977; Record, 1988, 1990; Ha et al., 1989; Lesser et al., 1990). The phosphate interference assay performed on the *lac* repressor with the singly substituted operators is diagnostic for how the two macromolecules are arranged along the interface. The fact that no apparent major alterations occurred in the interference pattern in the presence of a modified base at this concentration of protein and DNA implies that the variant complex is similar to the wild-type complex and that the overall phosphate requirements in the variant complex remain unchanged. Salt-dependency measurements of the singly substituted operator–protein complexes showed that there were no significant changes in counterions released compared to the wild-type complex (X. Zhang and P. A. Gottlieb, unpublished results). The source of the weaker stability observed in the *lac* repressor–modified operator complex is most likely a result of the loss of a hydrogen-bonding site along with either the apposition of sterically unfavorable groups along the interface of the protein–DNA

complex or perturbations in the local structure that require additional free energy to drive the acquisition of the correct conformation during the association process, or a combination of both. This is similar to observations made in the *Eco*RI restriction endonuclease system in which the isosteric analogues did not significantly alter the phosphate interference pattern (Lesser et al., 1990). In that case, the insertion of a single isosteric analogue, 7-deazadeoxyguanosine, into the restriction site resulted in an overall free energy change of 2.4 kcal/mol. This was partitioned between the energetics of binding (1.1 kcal/mol) and the local structural perturbations required to achieve the transition state (1.3 kcal/mol).

One possible mechanism whereby the complex's local structure plays a role in facilitating protein–DNA interaction is based on observations made by Lu et al. (1985). They have shown that embedded within the sequence GTG is an imino proton from the thymine residue that has an anomalous rate of exchange with solvent, in the absence of proteins. The structural basis for this faster than expected proton exchange is a matter of speculation, but it may be linked to an unusual structure of DNA similar to the one found in the crystal structure of the CAP–DNA complex (Schultz et al., 1991). The analogue substitution that had the most effect on the *lac* repressor–operator was in fact the first G of this motif, although it would not be directly involved in the site that is bent. Thus, it is conceivable that the effect exerted by this analogue is due in part to the influence of structural flexibility on the stabilization of the complex. We emphasize that this can explain only part of the effects observed. The genetic analysis in which both protein and DNA are mutated to produce pseudorevertants supports some form of direct readout between *lac* repressor and operator (Lehming et al., 1987, 1990; Sartorius et al., 1989). This may also explain the differential response observed when the analogue is substituted at either position 5 or 17. The loss of a hydrogen bond would presumably be equivalent, but the local structural effects may influence the magnitude of the energetics of interaction.

*Complexes with Double Substitutions Are Additive in Free Energy of Binding.* The one double mutant tested (5/17) altered the stability of the protein–DNA complex by 1770-fold in the standard equilibrium assay and by almost 560-fold in the competition binding assay. When operator 5/17 was subjected to alkylation interference footprinting, we identified phosphate groups that inhibited complex formation and whose pattern was similar to those of the singly substituted and wild-type operators. Thus the complex appears to be identical to the wild-type operator. The introduction of the double mutant therefore represents a simple combination of each of the single mutants, since the sum of $\Delta\Delta G°$ for positions 5 (2.4 kcal/mol) and 17 (1.6 kcal/mol) accurately reflects the $\Delta\Delta G°$ for the doubly substituted operator (4.8 kcal/mol) within experimental uncertainty.

*Constitutive Mutant Complex.* Adaptability of the *lac* protein–operator to constitutive mutations was postulated by Mossing and Record (1985), who showed that introduction of this constitutive mutation in the *lac* operator not only increased the $K_{obs}$ for the interaction but also increased counterion released for protein–DNA association as measured by an increase in salt dependency. The additional dependency on ions for the variant complex suggested that more phosphate groups are sequestered to stabilize the interaction which was perturbed by the loss of specific hydrogen bonds between the protein and the DNA. The complex in that state is not formed by nonspecific binding, which requires approximately 12 counterions, nor is it the canonical wild-type sequence requiring 7 counterions released, and they argue for an intermediate state between specific and nonspecific binding. We have looked at the alkylation interference over a limited range of concentrations of repressor protein (Figure 7). Our results showed that the interference pattern for the constitutive operator complex was similar to that of the wild type. This indicated that, in this case, identification of the phosphates required for complex formation using chemical methods either is too insensitive a method to identify all important sites or does not neccessarily have to correlate with energetic measurements.

*Conclusions.* The combined thermodynamic and footprinting data showed that the wild-type operator, when bound to repressor, supports asymmetric interactions on the left and right halves of the operator. Assuming that the alkylation interference assay is representative of the protein–DNA complex along the phosphate backbone, then the lack of a substantive change in footprinting, along with the thermodynamic stability, indicates that these complexes are similar in their global phosphate requirements for complex formation. The loss in stability must be associated with sterically incompatible groups possibly accompanied by changes in the local properties of the DNA.

## ACKNOWLEDGMENT

We thank Gail Dinter-Gottlieb and Gregory Moe for critical reading of the manuscript; Ponzy Lu, Mitch Lewis, and Helen Pace for the gift of *lac* repressor protein; Michael Brenowitz for the gift of the mutant *lac* repressor; and Linda Jen-Jacobson for many valuable discussions.

## REFERENCES

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

Aikens, C. R., & Gumport, R. I. (1991) *Methods Enzymol. 208*, 433–457.

Barkley, M. D., & Bourgeouis, S. (1980) in *The Operon* (Miller, J. H., & Reznikoff, W. S., Eds.) p 192, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Becker, M. M., Lesser, D., Kurpiewski, M., Baranger, A., & Jen-Jacobson, L. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6247–6251.

Betz, J. L., Sasmor, H. M., Buck, H., Insley, M. Y., & Caruthers, M. H. (1986) *Gene 50*, 123–132.

Boelens, R., Scheek, R. M., van Boom, J. H., & Kaptein, R. (1987) *J. Mol. Biol. 193*, 213–216.

Brennan, C. A., Van Cleve, M. D., & Gumport, R. I. (1987a) *J. Biol. Chem 261*, 7270–7278.

Brennan, C. A., Van Cleve, M. D., & Gumport, R. I. (1987b) *J. Biol. Chem 261*, 7279–7286.

Brenowitz, M., Pickar, A., & Jamison, E. (1991) *Biochemistry 30*, 5986–5998.

Brown, T., & Brown, D. J. S. (1991) *Oligonucleotides and Analogues: A practical approach* (Eckstein, F., Ed.) pp 1–24, IRL Press, Oxford.

Carey, J., Lewis, D. E. A., Lavoie, T. W., & Yang, J. (1991) *J. Biol. Chem. 266*, 24509–24513.

Caruthers, M. H. (1980) *Acc. Chem. Res. 13*, 155–160.

Chen, J. C., & Matthews, K. S. (1992) *J. Biol. Chem. 267*, 13843–13852.

Dubbendorf, J. W., deHaseth, P. L., Rosehdahl, M. S., & Caruthers, M. H. (1987) *J. Biol. Chem. 262*, 892–898.

Fickert, R., & Müller-Hill, B. (1992) *J. Mol. Biol. 226*, 59–68.

Fisher, E. F., & Caruthers, M. H. (1979) *Nucleic Acids Res. 7*, 401–407.

Fried, M., & Crothers, D. M. (1981) *Nucleic Acids Res. 9*, 248–254.
Garner, M. M., & Revzin, A. (1981) *Nucleic Acids Res. 9*, 3047–3060.
Goeddel, D. V., Yansura, D. G., & Caruthers, M. H. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 3292–3297.
Goeddel, D. V., Yansura, D. G., & Caruthers, M. H. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 3578–3582.
Ha, J.-H., Spolar, R. S., & Record, M. T., Jr. (1989) *J. Mol. Biol. 209*, 801–816.
Harrison, S. C., & Aggarwal, A. K. (1990) *Annu. Rev. Biochem. 59*, 933.
Jordan, S. R., & Pabo, C. O. (1988) *Science 242*, 893–899.
Koudelka, G. B., Harrison, S. C., & Ptashne, M. (1987) *Nature 326*, 886–889.
Koudelka, G. B., Harbury, P., Harrison, S. C., & Ptashne, M. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 4633–4637.
Lamerichs, R. M. J. N., Boelens, R. A., van der Marel, G. A., van Boom, J. H., Kaptein, R., Buck, F., Fera, B., & Ruterjans, H. (1989) *Biochemistry 28*, 2985–2991.
Lehming, N., Sarotrius, J., Neimoller, M., Genenger, D. W., & Müller-Hill, B. (1987) *EMBO J. 6*, 3145–3153.
Lehming, N., Sartorius, J., Kisters-Woike, B., von Wilcken-Bergmann, B., & Müller-Hill, B. (1990) *EMBO J. 9*, 615–621.
Lesser, D. R., Kurpiewski, M. R., & Jen-Jacobson, L. (1990) *Science 250*, 776–785.
Lewis, M., Lu, P., & Pace, H. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 1870–1875.
Lin, S.-Y., & Riggs, A. D. (1972) *J. Mol. Biol. 72*, 671–690.
Lin, S.-Y., & Riggs, A. D. (1975) *Cell 44*, 925–932.
Lu, P., Cheung, S., & Donlan, M. (1985) *Adv. Biophys. 20*, 153–175.
Maxam, A., & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 560–564.
Mazzarelli, J. M., Sharanbasva, B. R., Patricia, L. I., & McLaughlin, L. W. (1992) *Biochemistry 31*, 5925–5936.
McLaughlin, L. W., Benseler, F., Graeser, E., Piel, N., & Scholtissek. S. (1987) *Biochemistry 26*, 7238–7245.
Mossing, M. C., & Record, M. T., Jr. (1985) *J. Mol. Biol. 186*, 295–305.
Newmann, P. C., Nwosku, V. U., Williams, D. M., Cosstick, R., & Connolly, B. A. (1990a) *Biochemistry 29*, 9891–9901.
Newmann, P. C., Williams, D. M., Cosstick, R., Seela, F., & Connoly, B. A. (1990b) *Biochemistry 29*, 9902–9910.
Noble, S. A., Fisher, E. F., & Caruthers, M. H. (1984) *Nucleic Acids Res. 12*, 3387–3404.
O'Gorman, R. B., Dunaway, M., & Matthews, K. S. (1980) *J. Biol. Chem. 225*, 10100–10106.
Otwinowski, Z., Schevitz, R. W., Zhang, R.-G., Lawson, C. L., Joachimiak, A., Mormorstein, R. Q., Luis, B. F., & Sigler, P. B. (1988) *Nature 335*, 321–329.
Record, M. T., Jr. (1988) in *Unusual DNA Structures* (Wells, R. D., & Harvey, S. C., Eds) pp 237–251, Springer-Verlag, New York.
Record, M. T., Jr. & Spolar, R. (1990) *CRC-The biology of nonspecific DNA-protein interactions* (Revzin, A., Ed.) pp 33–71, CRC Press Inc., Roca Baton FL.
Record, M. T., Jr., deHaseth, P. L., & Lohman, T. M. (1977) *Biochemistry 16*, 4791–4796.
Riggs, A. D., Suzuki, H., & Bourgeois, S. (1970) *J. Mol. Biol. 48*, 67–83.
Royer, C. A., Chakerian, A. E., & Matthews, K. S. (1990) *Biochemistry 29*, 4959–4966.
Sartorius, J., Lehming, N., Kisters, B., Wilken-Bergmann, V., & Müller-Hill, B. (1989) *EMBO J. 8*, 1265–1270.
Schultz, S. C., Shields, G. C., & Steitz, T. A. (1991) *Science 253*, 1001–1007.
Seela, F., & Driller, H. (1986) *Nucleic Acids Res. 14*, 2319–2333.
Seela, F., & Kehne, A. (1987) *Biochemistry 26*, 2232–2238.
Seela, F., Kehne, A., & Winkeler, H. D. (1983a) *Liebigs Ann. Chem.*, 137–146.
Seela, F., Winkeler, H.-D., Ott, J., Tran-Thi, Q.-H., Hasselman, D., & Bussmann, W. (1983b) in *Nucleotides, Nucleosides and their Biological Applications* (Rideout, J. L., Henry, D. W., & Beacham III, L. M., Eds.) pp 181–208, Academic Press, New York.
Seela, F., Binding, U., Driller, H., Herdering, W., Kaiser, K., Kehne, A., Rosemyer, H., & Steker, H. (1987) *Nucleosides and Nucleotides 6*, 11–23.
Seeman, N. C., Rosenberg, J. M., & Rich, A. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 804–808.
Siebenlist, U., & Gilbert, W. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 122–126.
Staacke, D., Walter, B., Kisters-Woike, B., Wilcken-Bergmann, B., & Müller-Hill, B. (1990) *EMBO J. 9*, 1963–1967.
Whitson, P. A., & Shive-Matthews, K. (1986) *Biochemistry 25*, 3845–3851.
Whitson, P. A., Olson, J. S., & Shive-Matthews, K. (1986) *Biochemistry 25*, 3852-3858.
Woodbury, C. P., Jr., & von Hippel, P. H. (1983) *Biochemistry 22*, 4730–4737.