Operator DNA Sequence Variation Enhances High Affinity Binding by Hinge Helix Mutants of Lactose Repressor Protein†

Catherine M. Falcon and Kathleen S. Matthews*

Department of Biochemistry & Cell Biology, Rice University, Houston, Texas 77251

Received April 24, 2000; Revised Manuscript Received June 30, 2000

ABSTRACT: The mechanism by which genetic regulatory proteins discern specific target DNA sequences remains a major area of inquiry. To explore in more detail the interplay between DNA and protein sequence, we have examined binding of variant *lac* operator DNA sequences to a series of mutant lactose repressor proteins (LacI). These proteins were altered in the C-terminus of the hinge region that links the N-terminal DNA binding and core sugar binding domains. Variant operators differed from the wild-type operator, O^1, in spacing and/or symmetry of the half-sites that contact the LacI N-terminal DNA binding domain. Binding of wild-type and mutant proteins was affected differentially by variations in operator sequence and symmetry. While the mutant series exhibits a 10^4-fold range in binding affinity for O^1 operator, only a ∼20-fold difference in affinity is observed for a completely symmetric operator, O^{sym}, used widely in studies of the LacI protein. Further, DNA sequence influenced allosteric response for these proteins. Binding of this LacI mutant series to other variant operator DNA sequences indicated the importance of symmetry-related bases, spacing, and the central base pair sequence in high affinity complex formation. Conformational flexibility in the DNA and other aspects of the structure influenced by the sequence may establish the binding environment for protein and determine both affinity and potential for allostery.

Genetic regulatory proteins target specific sites within the genome and either enhance or repress transcriptional activity to elicit downstream cellular responses. In contrast to the precision and predictability of the genetic code, only general features of protein−DNA recognition processes have been identified (*1*−*3*). When complexed, the surfaces of the protein and DNA are complementary and match chemical, ionic, and hydrophobic interactions (*1*−*3*). Although common structural motifs have been identified in different families of regulatory proteins, the specific nature of the interaction can vary within a family (*4*). Moreover, structural flexibility for both protein and DNA is emerging in many cases as a requisite for specific binding (*5*−*10*). Repeating units/motifs within both the protein and the DNA (e.g., symmetric DNA sequence binding to a dimeric protein) may enhance the affinity and specificity of the complex (*3*, *11*). Solvent can also play a key role in protein−DNA binding (*12*−*14*), and effector molecules—metabolites, protein partners—can influence the protein−DNA interface indirectly by allosteric changes (*1*, *15*). These characteristics converge in the ability of a regulatory protein to locate its target DNA sequence in the substantial background of nontarget sites to elicit a specific transcriptional effect. Genetic regulation relies on these abilities of the protein to select a specific site and to modulate affinity in a manner responsive to the needs of the cell.

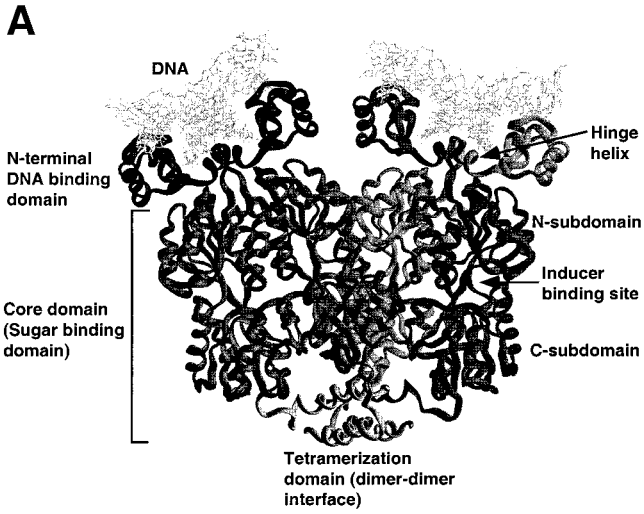
The *lac* operon in *Escherichia coli* is a well-characterized system of genetic regulation (*15*−*17*). The *lac* repressor protein (LacI) binds to its target site ($LacO^1$), thereby precluding transcription of the *lac* metabolic enzymes by RNA polymerase (*15*, *17*−*21*). Repression is relieved when inducer sugars bind to a distant site within LacI, thus eliciting a conformational change that diminishes affinity for the target operator DNA without effect on nonspecific DNA binding properties (*15*, *17*, *22*). LacI is a homotetramer of 360 amino acids per monomer with distinct functional domains (Figure 1) (*17*, *22*−*24*). The N-terminal ∼50 amino acids comprise the helix−turn−helix DNA binding domain (*25*−*29*). Each N-terminus makes contact with a half-operator so that two N-termini are required for binding to a single operator site, necessitating assembly to dimer for high affinity operator binding (*22*, *30*, *31*). The inducer binding site is located within the core domain, which encompasses amino acids ∼60−340 (*22*, *25*, *32*). The monomer−monomer subunit interface is also contained within this large core domain (*22*, *25*, *32*, *33*). A hinge domain (amino acids 51−59) links the N-terminal DNA binding and core inducer binding domains (*22*, *26*, *34*). This region forms a helix when LacI complexes with $LacO^1$ and may be involved in allosteric communication (*22*, *35*). A short segment at the C-terminus of the protein (amino acids 340−360) forms a four-helical coiled-coil structure that serves as the dimer−dimer interface to form the tetrameric structure that allows loop formation when two DNA binding sites within a single DNA molecule are occupied (*22*, *32*, *36*−*39*).

The *lac* operator, $LacO^1$, is a sequence of ∼24 base pairs that overlaps the *lac* operon promoter region (*40*). The O^1 operator exhibits a high degree of symmetry, a feature

† Support for this project was provided by grants to K.S.M. from NIH (GM 22441) and the Robert A. Welch Foundation (C-576). Spectroscopic facilities utilized were provided by the Keck Center for Computational Biology and the Lucille P. Markey Charitable Trust.

* To whom correspondence should be addressed. Phone: 713-348-4871; fax: 713-348-6149; e-mail: ksm@rice.edu.

10.1021/bi000924z CCC: $19.00
Published on Web 08/11/2000



B

Position →	51	52	53	54	55	56	57	58	59	60	61
Protein											
Wild type	R	V	A	Q	Q	L	A	G	K	Q	S
Gln60Gly	R	V	A	Q	Q	L	A	G	K	**G**	S
Gly60+1	R	V	A	Q	Q	L	A	G	K	**G**	**G**	S
Gly60+2	R	V	A	Q	Q	L	A	G	K	**G**	**G**	**G**	S
Gly60+3	R	V	A	Q	Q	L	A	G	K	**G**	**G**	**G**	**G**	S
Gln60Pro	R	V	A	Q	Q	L	A	G	K	**P**	S

FIGURE 1: Structure of LacI and sequences for variant proteins used in this study. (A) Wild-type LacI tetramer−O^{sym} complex is shown, with the functional domains labeled. The hinge helix, which links the functional domains within the protein, is designated. Likewise, the inducer binding site is indicated. This structure was derived from Protein Data Bank file 1LBG and displayed using Ribbons (*118*). (B) Amino acid sequence of the wild-type LacI hinge region is shown. Sequences for the mutant proteins that were engineered are shown with the specific mutation or insertion listed in larger font size.

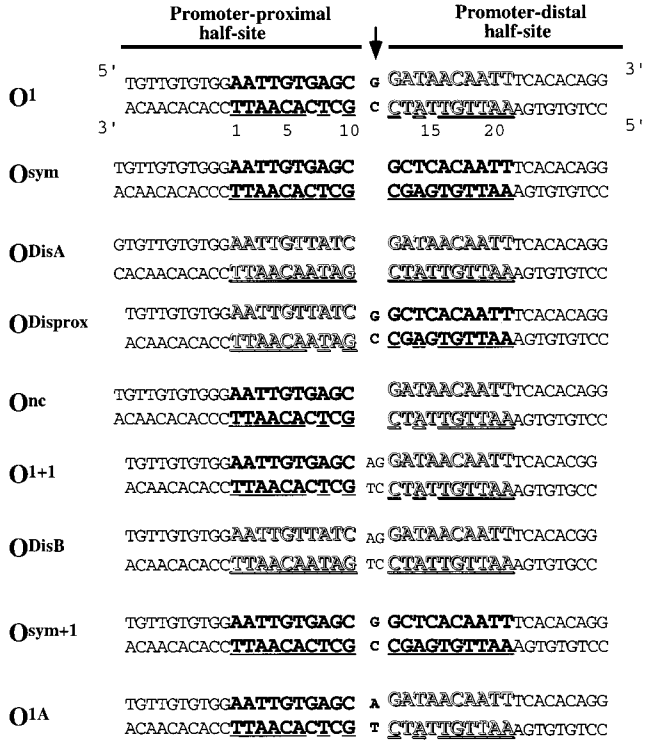


FIGURE 2: Variant operator constructs. The 40-base pair operator sequences used in the binding assays reported are shown. The natural operator, O^1, is listed first for comparison. The central sequence protected by LacI from DNase digestion is shown in larger font. The half-sites are labeled above the sequences, with the point of symmetry for the sequences indicated by the arrow. The sequences that are symmetric within the central region are underlined for each of the operators listed. Promoter-proximal sequences are in bold, and promoter-distal sequences are indicated by shadowed letters. Numbering is according to the start of transcription for the *lac* operon (*42*).

characteristic of many target sequences for regulatory proteins, presumably due to the required oligomeric structure of the cognate protein (Figure 2) (*3*, *11*, *41*, *42*). Two secondary operators within the *E. coli* genome have been located, O^2 and O^3, to which repressor binds with 10−100-fold lower affinity than O^1 (*43*−*45*). These sites are occupied in vivo and are required for maximal repression of *lac* operon expression, presumably due to formation of exceptionally stable looped structures with the higher affinity O^1 sequence (*46*−*49*). The interaction between LacI and O^1 has been assumed to be symmetric in nature, given the extensive bilateral symmetry of the two N-terminal contact regions and the DNA sequence (*40*, *42*, *50*, *51*). However, evidence has accumulated that the interaction is not entirely symmetric. Methylation and DNase protection studies (*31*, *52*, *53*), effects of sequence variation on binding (*44*, *54*, *55*), and analyses of repressor−operator complexes (*56*, *57*) have demonstrated asymmetric as well as symmetric contacts. Interestingly, however, a completely symmetric operator has been identified that binds to LacI with higher affinity than O^1 (*58*, *59*). This operator, O^{sym}, which is a symmetric sequence based on the promoter proximal half-site, differs from O^1 at three base positions; most notably O^{sym} is missing the central base pair (Figure 2) (*58*). The higher affinity for O^{sym} underscores the importance of symmetry in this interaction, although half-site spacing also plays a central role in effective complex formation (*60*, *61*).

To address more completely the nature of the LacI−DNA interaction, we have examined the binding of a series of LacI mutants with a range of operator variants. The mutant proteins (Figure 1) were designed with alterations at the C-terminus of the hinge domain (*62*), the short segment that links the DNA binding and inducer binding domains (*22*, *26*, *34*). These mutants exhibit progressive loss of O^1 affinity with insertion of Gly residues in this region, although inducer responsivity is maintained (*62*). Since increased flexibility between N-terminal and core domains was engineered into these proteins, we explored whether this variation would enable alternate binding specificities for DNA sequences. Multiple operators were constructed based on variations of O^1 and O^{sym} and assayed with the mutant LacI proteins. The results indicate that DNA sequence exerts a profound influence on the binding parameters for these proteins. Moreover, the DNA sequence influences the allosteric response to inducer, an effect suggested previously (*10*). The active participation of DNA in protein binding, rather than serving as a passive B-form target, is emerging from a variety of studies (*5*−*7*, *10*, *63*). From our results with LacI, the role of DNA sequence in establishing an interaction with a regulatory protein can be significant, and the active role of

Table 1: Operator Binding of LacI Mutants with O^{nc} (no central base pair)[a]

	5′										↓										3′
O^1	A	A	T	T	G	T	G	A	G	C	G	G	A	T	A	A	C	A	A	T	T
O^{sym}	A	A	T	T	G	T	G	A	G	C		G	C	T	C	A	C	A	A	T	T
O^{nc}	A	A	T	T	G	T	G	A	G	C		G	A	T	A	A	C	A	A	T	T

	O^1		O^{sym}		O^{nc}	
	K_d ($M \times 10^{11}$)					
protein	− IPTG	+ IPTG	− IPTG	+ IPTG	− IPTG	+ IPTG
wild-type	1.0 ± 0.2	>10000	0.18 ± 0.03	>10000	3000 ± 1000	3000 ± 1000
Gln60Gly	0.42 ± 0.05	>10000	0.20 ± 0.05	>10000	>10000	>10000
Gly60 + 1	2.5 ± 0.5	>10000	0.35 ± 0.1	>10000	>10000	>10000
Gly60 + 2	500 ± 200	>10000	1.4 ± 1.0	>1000	>10000	>10000
Gly60 + 3	>10000	>10000	0.41 ± 0.06	>100	>10000	>10000
Gln60Pro	20 ± 8	>10000	4.1 ± 2.0	>10000	>10000	>10000

[a] Sequences of the central region of the 40-bp operators are shown, top strand only. The center of symmetry is indicated by the arrow.

this partner in genetic regulation must be included in our assessment of these key cellular interactions.

EXPERIMENTAL PROCEDURES

Protein Purification. Plasmid DNA encoding either wild-type LacI or one of the mutant proteins (*62*) was transformed into BL26 cells [BL26Blue cells from Novagen, cured of the episome that carries the I^q promoter and the I gene (*64*)]. Wild-type and mutant proteins were purified according to protocols described previously (*39, 62*). Cells were frozen in lysing buffer (0.2 M Tris-HCl, pH 7.5, 0.2 M KCl, 0.01 M $Mg(OAc)_2$, 5% glucose, and 50 mg/L phenylmethylsulfonyl fluoride) and stored. After thawing in the presence of lysozyme (0.5 mg/mL), DNase was added, and the lysed cells were centrifuged followed by precipitation of the supernatant with 40% ammonium sulfate. The precipitate was centrifuged, resuspended, and dialyzed overnight against 0.09 M potassium phosphate, pH 7.6, 5% glucose, and 1 mM DTT.[1] The protein was loaded onto a phosphocellulose column equilibrated with the same buffer and eluted with a gradient from 0.12 to 0.3 M potassium phosphate, pH 7.6, 5% glucose, and 1 mM DTT. Fractions containing LacI activity were collected, and each protein was found to be >95% pure by SDS−PAGE (data not shown) (*62*).

Operator DNAs. Operator DNAs (Figure 2) were produced from single-stranded DNA sequences manufactured by Great American Gene Company or Genosys. The paired 40-mers were annealed using annealing buffer (20 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, and 25 mM NaCl) at 80 °C for 3−5 min and then cooled slowly for several hours. The absorbance at 260 nm for each operator was measured prior to the labeling reaction, and the double-stranded character was evaluated by polyacrylamide gel electrophoresis. DNA (0.5 μg) was labeled in the presence of 1 μL polynucleotide kinase, 10 U/μL (Promega), 1 × kinase buffer, 2.5 mM spermidine (Sigma), 10 μL of [γ-^{32}P]ATP, 10 mCi/mL (ICN) in a volume of 30 μL for 1−2 h. The reaction was terminated with the addition of 2 μL of 0.5 M EDTA plus 18 μL of H_2O. The reaction components were then purified using a Nick column (Pharmacia), and the dsDNA was stored in TE at a concentration of 4.5×10^{-8} M at −20 °C. All operator DNAs were synthesized at least twice, and results from at least two preparations of protein are included in the replicates reported in this manuscript.

Operator Binding. Operator binding was assayed at 22 °C in buffer containing 0.01 M Tris-HCl, pH 7.6, 0.1 mM EDTA, 0.15 M KCl, and 5% dimethyl sulfoxide. The proteins were diluted into this buffer with the addition of 50 μg/mL bovine serum albumin. The labeled 40-base pair operator was used at a final concentration of either 1×10^{-13}, 1×10^{-12}, or 1×10^{-10} M, depending on the expected K_d. The protein concentration varied from 1×10^{-14} to 1×10^{-6} M, but each assay covered a range of concentrations of at least 4 log units. The final concentration of IPTG, when present, was 1 mM. The assay was a variation of that described by Wong and Lohman (*65*), using a dot-blot apparatus. Reactions were incubated for 20−30 min prior to filtering onto presoaked nitrocellulose filters. After filtration, the nitrocellulose filters were dried, and the bound DNA was detected using a Fugi Bioimaging System, with the amount of radiolabel quantitated using the program Mac-Bas (Fugi). All data were analyzed using the program Igor to fit the binding curves using nonlinear least-squares analysis to the following equation:

$$R = Y_m[P]/(K_d + [P]) \quad (1)$$

where R is the fraction of operator in complex, Y_m is the correction that allows the maximum value of R to float, [P] is the protein concentration in tetramer, and K_d is the apparent dissociation constant in tetramer concentration. Some binding curves did not reach saturation even at maximal protein concentrations, and therefore the K_d values derived from these curves are only approximate. Data in the presence of IPTG were compared to the saturation value for protein with no ligand to determine fractional saturation. Alterations in dissociation rate constants can affect the maximum value for DNA retention and thereby influence K_d determination. However, this method nonetheless appeared most effective in making comparisons between different DNAs.

RESULTS

Binding of Gly Mutant Series to O^1 and O^{sym}. LacI mutants (Figure 1) with conversion of Gln60 to Pro or Gly and with insertion of up to three additional Gly residues (Gly60 + 1,

[1] Abbreviations: DTT, dithiothreitol; TE, Tris-EDTA; IPTG, isopropyl-β,D-thiogalactoside.

Gly60 + 2, Gly60 + 3) following Gln60Gly were examined for affinity to O^{sym} and compared with O^1 (Table 1) (Figure 3) (*62*). For O^1, a loss in binding affinity was observed with increasing number of Gly insertions, with Gly60 + 3 yielding essentially nonspecific affinity for the DNA sequence (*62*). The difference in affinity between Gln60Gly and Gly60 + 3 for binding to O^1 was $\sim 10^4$. When inducer was added, the affinity for O^1 decreased to nonspecific levels for all of the mutant proteins. A strikingly different pattern was observed for binding of these mutant proteins to O^{sym}. Wild-type LacI bound with ~5-fold greater affinity to O^{sym} as compared to O^1, as anticipated from previous results (*58*, *59*, *66*) (Figure 3B). The mutant proteins all bound with high affinity to O^{sym}, with only 20-fold difference in affinity among this entire series.

Interestingly, the ability of inducer to elicit decreased affinity diminished with increased Gly insertions. All mutant proteins exhibited wild-type affinity for inducer (*62*) so that diminished affinity for this ligand cannot account for loss of allosteric response. Gln60Gly had approximately the same affinity for O^1 and O^{sym}, with a similar diminution of binding in the presence of inducer. With three Gly residues inserted (Gly60 + 3), substantial binding remained in the presence of 1 mM IPTG, a result that suggests loss of allosteric communication between the two functional domains in this protein. Even the Gln60Pro mutation exhibited high affinity for the symmetric operator sequence as compared to its affinity for O^1, and this protein also maintained inducibility with O^{sym}. The striking contrast between the binding behavior for the mutant proteins with O^1 versus O^{sym} motivated further exploration of the effects of sequence and spacing between the half-sites to assess key features of the LacI−LacO interaction.

Effects of Variation in Operator Symmetry. Variant operators were constructed with alterations in half-site symmetry. The promoter-distal (right) half-site of O^1 differs by two base pairs as compared to the corresponding promoter-proximal (left) half-site (Figure 2), which has been deduced to have higher affinity for LacI (*31*, *52*, *54*, *55*). The O^{DisA} operator was designed so that symmetry was generated based on the promoter-distal half-site, and the sequence was missing the central base pair; that is, this operator was similar to O^{sym} in its spacing pattern but contained sequence corresponding to the right rather than the left half-site (Figures 2 and 4). The binding of O^{DisA} to wild-type and the Gly series of mutant proteins was measured under conditions identical to those used for O^1 and O^{sym} binding. Wild-type LacI and all of the mutant proteins exhibited affinity approaching nonspecific levels. Thus, provision of the right half-site in a symmetric operator comparable to O^{sym} in spacing resulted in loss of specific wild-type LacI binding capacity, as previously reported (*60*) and confirmed in recent studies (*61*). Further, none of the mutant proteins bound this operator effectively. As a confirmation of the nonspecific character of this interaction, addition of inducer had no effect on binding of any of the proteins. From the binding pattern for O^{DisA} with wild-type and all the mutant proteins, flexibility between the N-terminal and core domains is not sufficient to generate the binding contacts required for this sequence. Kaptein and colleagues (*61*) have suggested recently that orientation of the N-terminal DNA binding domains on this sequence misposition the N-termini for hinge helix interaction and

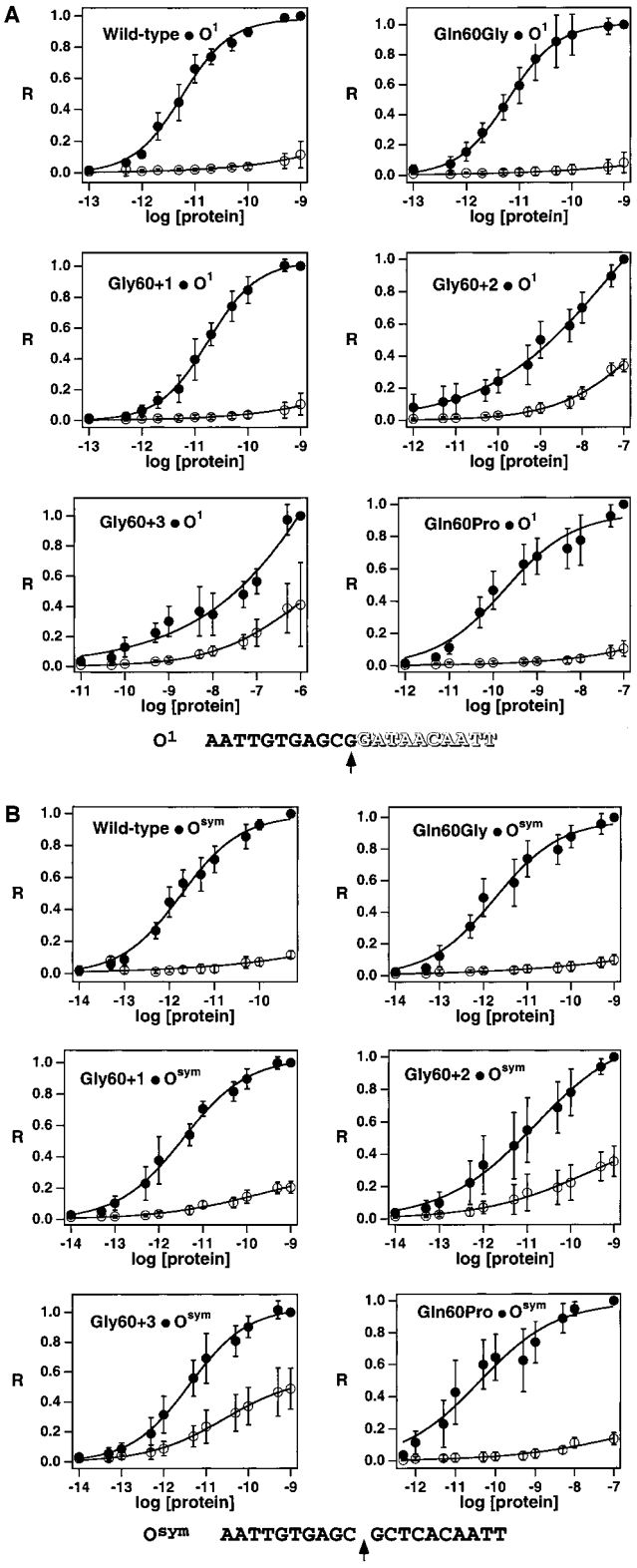


FIGURE 3: Binding of LacI and mutant proteins to O^1 and O^{sym}. (A) Binding isotherms for wild-type LacI and the mutant proteins with the natural operator, O^1. DNA concentration was 1×10^{-12} M for all but Gln60Gly, for which a DNA concentration of 1×10^{-13} M was used. (B) Binding isotherms for wild-type LacI and mutants with the symmetric operator, O^{sym}. DNA concentration was 1×10^{-13} M. (●) protein alone; (○) protein plus 10^{-3} M IPTG. The points represent the average of at least five experiments, and the fits (solid lines) were generated as described in Materials and Methods. The top strand of operator sequence is shown. The center of symmetry is indicated by the arrow. Note that the *x* axis differs for some panels due to differences in protein affinity for DNA.

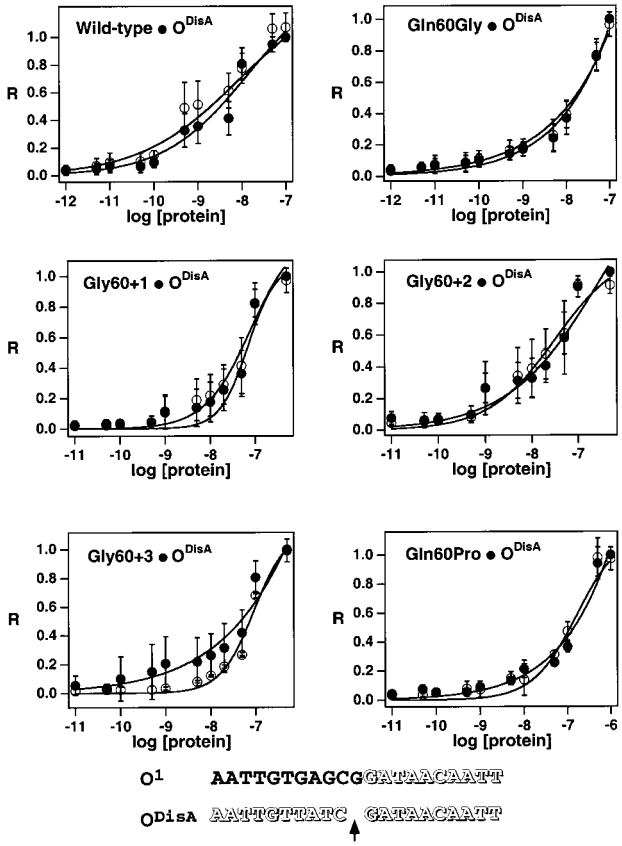


FIGURE 4: Binding isotherms of wild-type LacI and mutant proteins to the O^{DisA} operator. Data are from at least three experiments, and the fits (solid lines) were generated from eq 1 in Materials and Methods. DNA concentration was 1×10^{-12} M. (●) protein alone; (○) protein plus 10^{-3} M IPTG. O^{DisA} sequence is shown as compared to O^1. The center of symmetry is indicated by the arrow.

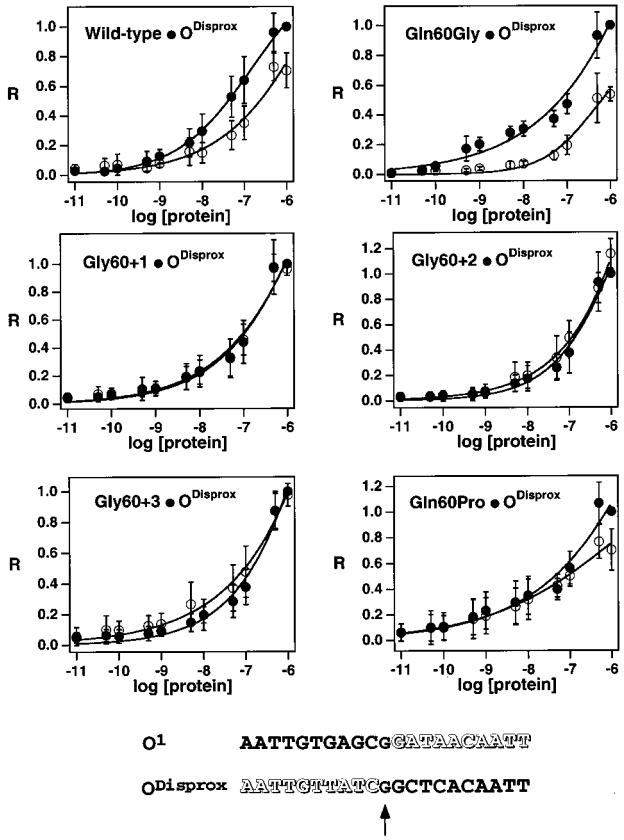


FIGURE 5: Binding isotherms for wild-type LacI and the mutant proteins with $O^{Disprox}$. Data are from at least three experiments, and the fits (solid lines) were generated using eq 1 in Materials and Methods. DNA concentration was 1×10^{-12} M. (●) protein alone; (○) protein plus 10^{-3} M IPTG. $O^{Disprox}$ sequence is shown as compared to O^1. The center of symmetry is indicated by the arrow.

thereby preclude high affinity binding.

To examine the effect of "order" of the half-sites, an operator was constructed in which the half-sites were "inverted", keeping the central base pair intact (Figures 2 and 5). This operator, $O^{Disprox}$ (promoter distal half-site rotated through the point of symmetry with the proximal half-site), also resulted in nonspecific levels of binding. Although the half-site sequences were maintained, the arrangement of these sites around the central base was altered, changing the order of bases in the central operator sequence reading from the promoter-proximal site. The wild-type and mutant proteins demonstrated no measurable affinity for this DNA sequence. Interestingly, Gln60Gly appeared to have a very small inducer response, although the affinity in the absence of inducer was already quite low.

Effects of Variation in Spacing. The central base pair occupies the axis of rotation for the O^1 operator, and this base pair is eliminated in O^{sym} for which the axis of rotation is between symmetric base pairs. With removal of the central base pair, the half-sites rotate 36° closer to the same plane (presuming the DNA is in the standard B-form) and move closer by 3.3 Å (*67*). Alternate operators were constructed to examine the influence of spacing on the binding interaction. In the first of this set of operator spacing variants, the central base pair in O^1 was deleted, and the half-site sequences were maintained. This operator, O^{nc} (for no central base pair), was examined with wild-type and the mutant Gly series (Table 1). Wild-type LacI exhibited a >1000-fold reduction in affinity for this operator, and binding did not appear to be influenced by the presence of inducer. Furthermore, the entire Gly series of mutant proteins could not recognize this sequence specifically. A similar result was observed recently for wild-type protein and N-terminal fragments (*61*). Moreover, in this study these proteins showed no effect of inducer presence on binding when assayed with this operator sequence. Despite the difference of only two base pairs between O^{sym} and O^{nc}, the binding differential was greater than 4 orders of magnitude for all proteins examined.

To explore whether additional separation of half-sites may be accommodated, a variant operator, O^{1+1}, was constructed with addition of an A/T base pair in the central region, separating the half-sites further. Wild-type LacI bound to O^{1+1} with >1000-fold reduced affinity, and binding was not altered by inducer presence (Table 2). Similar behavior was observed for the mutant series of proteins. Despite the additional flexibility and the presumed ability to span the added separation of half-sites, the Gly insertion mutants bound to this operator with nonspecific affinity. Diminished binding may derive from the inability of the hinge helices to form when separated by the additional base pairs. Another operator was constructed in which the half-sites were symmetric to the right (distal) half-site, and these sites were separated by two central bases, O^{DisB}. Previous reports have postulated that further separation of the distal half-sites beyond wild-type spacing would enhance binding for wild-type LacI (*60*, *61*). However, when this series of proteins

Table 2: Operator Binding of LacI Mutants with O^{1+1} (insertion of one extra base to central region) and with O^{DisB} (insertion of two central bases between promoter distal half-sites)[a]

	5′										↓										3′
O^1	A	A	T	T	G	T	G	A	G	C	G	G	A	T	A	A	C	A	A	T	T
O^{1+1}	A	A	T	T	G	T	G	A	G	C	A G	G	A	T	A	A	C	A	A	T	T
O^{DisB}	A	A	T	T	G	T	T	A	T	C	A G	G	A	T	A	A	C	A	A	T	T

	O^{1+1}		O^{DisB}	
	K_d ($M \times 10^8$)			
protein	− IPTG	+ IPTG	− IPTG	+ IPTG
wild-type	2 ± 1	2 ± 1	>10	>10
Gln60Gly	8 ± 5	4 ± 3	>10	>10
Gly60 + 1	0.8 ± 0.5	2 ± 2	>10	>10
Gly60 + 2	>10	>10	>10	>10
Gly60 + 3	>10	>10	>10	>10
Gln60Pro	>10	>10	>10	>10

[a] Sequences of the central region of the 40-bp operators are shown, top strand only. The center of symmetry is indicated by the arrow.

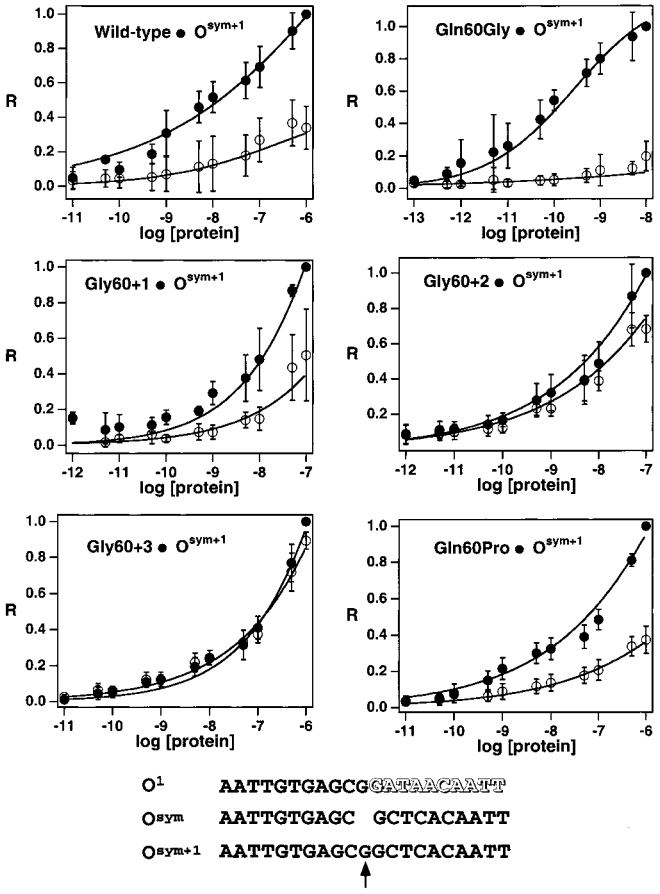


FIGURE 6: Binding isotherms for wild-type LacI and the mutants proteins with O^{sym+1}. Data are from at least three experiments, and the fits (solid lines) were generated using eq 1 in Materials and Methods. DNA concentration was 1×10^{-12} M. (●) protein alone; (○) protein plus 10^{-3} M IPTG. O^{sym+1} sequence is shown as compared to O^1 and O^{sym}. The center of symmetry is indicated by the arrow.

was assayed with the O^{DisB} operator, no high affinity binding was detected (Table 2).

A final spacing variant was constructed to insert the central G/C base pair characteristic of O^1 into the O^{sym} sequence (Figure 6). This operator, termed O^{sym+1}, had been examined in previous studies (*60*, *61*), and we demonstrate here that wild-type LacI bound with reduced affinity to this sequence by ~100-fold as compared to O^1 and ~1000-fold as compared to O^{sym}. IPTG elicited a decrease in wild-type LacI affinity for O^{sym+1}, but the differential was not as pronounced as that observed for the O^1 or O^{sym} sequences. Gln60Gly bound with higher affinity than wild-type ($K_d \sim 3 \times 10^{-10}$ M), although affinity was reduced as compared to either O^1 or O^{sym}. Moreover, inducer elicited a decrease to nonspecific binding levels for this mutant protein. Gln60Pro also exhibited an effect of inducer on DNA binding, although the affinity in the absence of inducer was significantly diminished as compared to wild-type or Gln60Gly. The remainder of the mutants exhibited nonspecific affinity and did not respond significantly to the presence of inducer. From these experiments, affinity and allostery for wild-type and mutant LacI proteins are influenced significantly by single base pair alterations in half-site spacing.

Alternate Central Base Pair. The effect of altering the central base pair of O^1 from G/C to A/T was examined. This operator, O^{1A}, has been shown to exhibit behavior in vivo, suggesting diminished binding affinity (*68*). Wild-type and all of the mutant proteins (Table 3) exhibited very low affinity for this operator variant. No effect of inducer was observed on the binding properties. These results indicate the key role of the central base pair in establishing the potential for interaction between LacI and LacO. Even a purine to purine change in this central sequence abolishes high-affinity complex formation.

Nonspecific DNA Binding. To confirm that wild-type and the Gly series of mutants bound similarly to nonspecific DNA, a 40-bp sequence was generated (Table 3). The sequence contained the same flanking ends as the variant operator sequences, but differed in the central region corresponding to the half-sites and central base pair. This sequence is the consensus target DNA for the human p53 protein (*69*). Wild-type and mutant proteins exhibited the anticipated nonspecific binding affinity for this sequence.

DISCUSSION

Protein−DNA interactions have been the subject of significant study over the past several decades (*1*, *41*, *70*, *71*). The formation of a high-affinity complex has been generally ascribed to a complementary fit between the surfaces of B-form DNA and the protein with the juxtaposition of functional groups that generate specific ionic, van der Waals, and hydrophobic contacts (*1*, *3*, *66*, *72*). Variations in the functional groups via protein mutation or

Table 3: Operator Binding of LacI Mutants with O^{1A} (the central base pair was altered from G/C to A/T) and $O^{nonspecific}$ (non-*lac* Operator Sequence)[a]

	5′										↓										3′
O^1	A	A	T	T	G	T	G	A	G	C	G	G	A	T	A	A	C	A	A	T	T
O^{1A}	A	A	T	T	G	T	G	A	G	C	A	G	A	T	A	A	C	A	A	T	T
O^{non}	A	G	A	C	A	T	G	C	C	T	A	G	A	C	A	T	G	C	C	T	T

	$K_{d\,(M \times 10^{8})}$			
	O^{1A}		$O^{nonspecific}$	
protein	− IPTG	+ IPTG	− IPTG	+ IPTG
wild-type	>10	>10	>10	>10
Gln60Gly	>10	>10	>10	>10
Gly60 + 1	>10	>10	>10	>10
Gly60 + 2	>10	>10	>10	>10
Gly60 + 3	>10	>10	>10	>10
Gln60Pro	>10	>10	>10	>10
wild-type	>10	>10	>10	>10

[a] Sequences of the central region of the 40-bp operators are shown, top strand only. The center of symmetry is indicated by the arrow.

chemical synthesis of DNA sequences have been shown to affect affinity and specificity of binding in a way that can be rationalized based on alterations in the complementarity of the binding species (*64*, *73*−*80*). Attention has focused largely on the protein as the "active" partner in this interaction, while DNA has been viewed in a more "passive" role, with sequence alterations misaligning the functional groups presented by the DNA. Indeed, the recognition helix of the helix−turn−helix motif in the N-terminal DNA binding domains of LacI appears to interact with moieties in the major groove of a predominantly B-form DNA region (*22*).

Protein conformational changes, adaptation to ligand, and concerted protein folding have been established as key aspects for protein recognition of DNA sequences (*8*, *68*, *80*). In contrast, until recently the structure of the DNA was presumed by many to be relatively rigid and largely B-form; alterations in DNA structure were not considered in detail in understanding the mode of interaction. X-ray crystallographic structures of multiple protein−DNA complexes have elicited renewed interest in the DNA as an equally malleable partner in complex formation. The identification of DNA bends (*81*−*83*), unwinding (*84*−*86*), and other conformational alterations (*87*−*92*) suggest a more active role for DNA structure (*10*) than merely presenting functional groups in the major and minor grooves within an otherwise uniform DNA background. Bending is observed in a number of protein−DNA complexes associated with transcription regulation, including LacI, and may be used in part to signal complex formation (*22*, *61*, *81*−*83*, *93*, *94*). The flexibility of the sugar−phosphate backbone (*95*, *96*) allows rearrangements of the grooves in protein complexes with minimal energy loss and without significant alterations of base stacking (*5*, *97*). Some forms of DNA bending may require a Py−Pu step, a consequence of the ability of this sequence to adopt a broader range of conformations (*82*, *98*). Py−Pu, and in particular C−G, steps have higher than average roll angles, resulting in enhanced flexibility and the ability to adopt greater positive rolls (*6*, *94*).

Central sequences in pseudosymmetric operators may play a key role in bending and structural flexibility to generate the DNA binding conformation necessary for high-affinity protein binding. The 434 and P22 repressors provide well-studied examples of proteins that rely on noncontacted base pairs that lie between the contacted base pairs of their respective operators to form the high-affinity complex (*99*−*101*). These central bases are involved in modulating DNA twist and the geometry of the minor groove in the center of the binding site for these proteins. In fact, the conformation of the DNA phosphate backbone has been shown to be altered throughout the 434 operator by a single base pair substitution in the central region (*100*). Subtle changes in DNA sequence therefore lead to structural changes that are detected by the protein and are reflected in large alterations in affinity and/or allostery (*102*, *103*). Binding affinity can be decreased if the binding surface of the protein does not complement the DNA surface or if protein−DNA binding cannot elicit the requisite conformational shifts to generate that surface.

To examine the importance of DNA sequence/structure in the formation of the high-affinity complex with LacI, we have examined the interaction of designed operator variants with wild-type LacI and derivatives with alterations in the C-terminal region of the hinge that links the DNA and inducer binding domains. The results indicate that O^{sym}, an operator made completely symmetric to the promoter-proximal half-site and missing the central base pair, can substantially overcome deleterious effects of hinge region alterations in this protein. The bending observed in the central region of O^{sym} (*22*, *61*, *104*, *105*) and O^1 (*61*) in complex with LacI derives from widening of the minor groove by insertion of Leu56 side chains from the two N-terminal hinge helices within a dimer, an arrangement also observed for the purine repressor in the PurR−PurF complex (*106*). Despite the widening of the minor groove, the central bases of O^{sym} do not appear to be contacted significantly by the protein (*22*, *105*), although hydrophobic interactions have been reported between residues of the hinge helix and the central base pair (*104*, *105*). The distinction between O^1 and O^{sym} is further emphasized when these LacI mutant proteins are assayed in the presence of IPTG. As the number of glycines inserted at the end of the hinge region increases, the allosteric communication between the domains decreases in the presence of O^{sym}. This communication may be mediated by hinge helix folding, which has been identified as a key component for high-affinity binding (*61*, *104*, *105*); DNA sequence may influence the establishment of requisite protein−protein interactions for hinge helix formation (*61*).

The importance of the central operator sequence is illuminated by the binding behavior of the $O^{Disprox}$, O^{1+1}, and O^{1A} operator sequences to wild-type and mutant LacI proteins. These sequences alter the CGG central sequence of O^1 (reading from the promoter-proximal sequence) to CCG, CAGG, and CAG, respectively, and may limit DNA conformational states. O^{sym} contains two GC sequences in tandem, an arrangement unique among the sequences examined. Thus, flexibility within the central operator region may be a primary determinant of overall DNA binding affinity. The difference in central sequence may also explain the difference in binding affinity between the O^{DisB} operator reported here and a similar sequence used by Spronk et al. (*61*), where they maintain the central CG sequence. Interestingly, the O^{nc} operator reported here contains a central CG sequence, yet none of the proteins appear to bind with measurable affinity. Therefore, the sequence of the central bases may be necessary but not sufficient for the binding interaction.

Symmetry appears important to generate high-affinity binding for LacI and other multimeric proteins (e.g., refs *6*, *58*, and *107*). However, when symmetry is produced to the promoter-distal half-site of O^1, high-affinity binding is precluded. Thus, symmetry alone is not sufficient for effective sequence recognition. In fact, LacI−$LacO^1$ binding has evidenced asymmetric character from the identification of the first operator-constitutive mutants, where base pair alterations deleterious to repression in vivo were found to be concentrated on the promoter-proximal side of the O^1 sequence (*42*, *55*, *74*). Further, methylation protection of purine residues in the LacI−$LacO^1$ complex (*31*, *108*), protection of BrdU-substituted oligonucleotides from UV-induced strand scission by LacI presence (*109*, *110*), and effects of base substitution on binding by synthetic oligonucleotides generate an underlying pattern of symmetric contacts overlaid by asymmetric contacts by LacI (*72*, *76*, *77*).

This mix of symmetry/asymmetry has been observed in NMR studies of the intact protein and the isolated N-terminal domain complexed with O^{sym} and O^1 sequences. Differences between nonspecific and specific complexes are observed for the chemical shifts of Tyr7 and Tyr17, amino acids that are essential for DNA binding (*111*, *112*). Using ^{19}F NMR and 5-fluoro-deoxyuridine-substituted O^1 DNA, distinct spectral differences were observed for binding of two independent N-terminal domains to the two-half-site sequences. The results corresponded closely to the data obtained in the presence of the intact LacI protein (*113*), suggesting that inherent differences in these DNA sequences influence the mode by which contact is made with the identical N-terminal DNA binding domains. Symmetric NMR signals were observed only when the half-sites were completely symmetric, and the key effect appears to be related to the +13 A/T base pair (see Figure 2 for position), which influences interactions up to three base pairs removed (*113*). The symmetric presence of this +13 A/T base pair in O^{DisA} may account for the lower affinity of this operator sequence for wild-type and mutant LacI proteins. Distortions in the backbone and base stacking that appear to be propagated from A/T at +13, and presumably from its symmetry partner at +9 in O^{DisA}, may generate a conformation that is unfavorable for LacI recognition and binding (*113*, *114*).

Spronk et al. (*61*) have recently shown that NMR chemical shifts for hinge helix resonances are different for complexes with O^1 and O^{sym}. The resonances assigned to Gly58 appear to be in nonsymmetric environments for O^1 but similar for O^{sym}. This position abuts the site for the mutant series examined, and the additional flexibility introduced by glycine insertion appears to compromise O^1 binding but not O^{sym} binding. An effect similar to Gly58 was observed for Gln18, which is located in the recognition helix. Whether these effects are correlated remains unknown, but the influence of hinge helix formation on DNA bending and hence minor/major groove structure provides the potential for the protein to serve as an allosteric effector of DNA structure, thereby determining the affinity of the complex. Spacing of the N-terminal domains on the DNA appears to influence the relative capacity to form the hinge helices essential for minor groove insertion, DNA bending, and optimal contacts with the major groove recognition helix (*61*).

Indeed, folding of the hinge region to a helix may be a key element for specific LacI binding and for allostery (*22*, *35*, *61*, *104*, *105*), as the hinge helix does not appear to be stable in the absence of specific operator binding (*35*, *61*, *104*). Specific contacts/arrangements of the minor groove may be required for formation of this secondary structure. On the basis of a high-resolution crystal structure of dimeric LacI complexed with O^{sym}, protein−protein contacts are suggested to be essential for both hinge helix folding and for allosteric response to IPTG (*105*). Further, base substitution studies of LacO sequences have established that substitution of either 2-amino-purine or purine paired with either C or U at position 12 in $LacO^1$ did not affect LacI binding affinity, while substitution at other sites, including the central position 11, diminished binding affinity (*115*). Thus, a specific DNA sequence in the central region of the operator may be required for formation of the hinge helix and hence complex formation.

Whatever the specific role of the central sequence, hinge helix formation is clearly required to generate the DNA bending necessary for specific contacts between the two helix−turn−helix motifs and the half-site sequences (*61*). If the hinge helix does not form and the Leu56 residues are not positioned in the minor groove, high-affinity specific binding will be prevented. Another member of the LacI family of repressors, CytR, illustrates this effect. CytR contains a flexible, unstructured interdomain linker that connects the DNA binding domain with the core domain. This arrangement allows CytR to adopt various conformations that enable it to interact with widely spaced half-sites and with CRP at different promoters but with significantly lower affinity as compared to LacI/LacO (*116*, *117*). This flexible linker frees up the protein to interact with various DNA sequences but at the cost of high affinity binding. One possible explanation for the low affinity for O^1 observed with Gly60 + 1, Gly60 + 2, and Gly60 + 3 is that the additional glycines diminish the capacity for α-helix formation by the hinge. However, the high affinity of the mutant proteins for O^{sym} argues against loss of helix formation as an exclusive explanation. A highly intriguing result is that insertion of a central G:C base pair (which yields the spacing of O^1) into O^{sym} to generate O^{sym+1} results in significantly decreased

binding for wild-type and Gln60Gly and nonspecific binding for the remaining proteins. Thus, promoter-distal sequences in O^1 must compensate for binding loss generated by the presence of the central G:C base pair. These sequences do not, however, support effective binding on their own, even with spacing similar to O^{sym} (O^{DisA} results).

We demonstrate in this work that subtle alterations in DNA sequence have a profound effect on binding affinity and allosteric response for wild-type LacI and mutant proteins with alterations at the C-terminus of the hinge helix. Altering the central sequence, spacing between half-sites, or the symmetry within the operator results in significant effects on affinity and allostery. By engineering these particular LacI mutations to introduce flexibility between the N-terminal and core domains, the proteins were designed to have the conformational potential to extend their recognition of DNA sequences. However, the results demonstrate that this potential is highly dependent on the sequence, and presumably conformation, of the DNA. This work contributes to an evolving view of the importance of DNA sequence and associated conformational flexibility on protein recognition and affinity. The influence of DNA on complex formation and on allosteric response should be anticipated from basic thermodynamic principles, but the subtlety and range of this influence are only beginning to be appreciated. Base analogue substitution, temperature studies to examine coupled folding, and structure determinations for LacI and its derivatives complexed to variant operators may provide additional insight into the detailed mechanisms by which protein−DNA binding affinity and allosteric response are determined.

ACKNOWLEDGMENT

We thank Emilia Mrowczynski for technical support and the members of the Matthews lab for helpful discussions.

REFERENCES

1. Steitz, T. A. (1990) *Q. Rev. Biophys. 23*, 205−280.
2. Suzuki, M. (1994) *Structure 2*, 317−326.
3. Rhodes, D., Schwabe, J. W. R., Chapman, L., and Fairall, L. (1996) *Philos. Trans. R. Soc. London B 351*, 501−509.
4. Harrison, S. C. (1991) *Nature 353*, 715−719.
5. Laughton, C., and Luisi, B. (1998) *J. Mol. Biol. 288*, 953−963.
6. Juo, Z. S., Chiu, T. K., Leiberman, P. M., Baikalov, I., Berk, A. J., and Dickerson, R. E. (1996) *J. Mol. Biol. 261*, 239−254.
7. Nagaich, A. K., Appella, E., and Harrington, R. E. (1997) *J. Biol. Chem. 272*, 14842−14849.
8. Spolar, R. S., and Record, M. T., Jr. (1994) *Science 263*, 777−784.
9. Cleary, M. A., and Herr, W. (1995) *Mol. Cell. Biol. 15*, 2090−2100.
10. Lefstin, J. A., and Yamamoto, K. R. (1998) *Nature 392*, 885−888.
11. Schleif, R. (1988) *Science 241*, 1182−1187.
12. Li, L., and Matthews, K. S. (1997) *Biochemistry 36*, 7003−7011.
13. Otwinowski, Z., Schevitz, R. W., Zhang, R.-G., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q., Luisi, B. F., and Sigler, P. B. (1988) *Nature 335*, 321−329.
14. Arbuckle, N. D., and Luisi, B. (1995) *Struct. Biol. 2*, 341−346.
15. Miller, J. H., and Reznikoff, W. S., Eds. (1980) *The Operon*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
16. Jacob, F., and Monod, J. (1961) *J. Mol. Biol. 3*, 318−356.
17. Matthews, K. S., and Nichols, J. C. (1998) *Prog. Nucl. Acid Res. Mol. Biol. 58*, 127−164.
18. Deuschle, U., Gentz, R., and Bujard, H. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 4134−4137.
19. Straney, S. B., and Crothers, D. M. (1987) *Cell 51*, 699−707.
20. Lee, J., and Goldfarb, A. (1991) *Cell 66*, 793−798.
21. Schlax, P. J., Capp, M. W., and Record, M. T., Jr. (1995) *J. Mol. Biol. 245*, 331−350.
22. Lewis, M., Chang, G., Horton, N. C., Kercher, M. A., Pace, H. C., Schumacher, M. A., Brennan, R. G., and Lu, P. (1996) *Science 271*, 1247−1254.
23. Beyreuther, K., Adler, K., Geisler, N., and Klemm, A. (1973) *Proc. Natl. Acad. Sci. U.S.A. 70*, 3576−3580.
24. Farabaugh, P. J. (1978) *Nature 274*, 765−769.
25. Platt, T., Files, J. G., and Weber, K. (1973) *J. Biol. Chem. 248*, 110−121.
26. Geisler, N., and Weber, K. (1977) *Biochemistry 16*, 938−943.
27. Jovin, T. M., Geisler, N., and Weber, K. (1977) *Nature 269*, 668−672.
28. Matthews, B. W., Ohlendorf, D. H., Anderson, W. F., and Takeda, Y. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 1428−1432.
29. Kaptein, R., Zuiderweg, E. R. P., Scheek, R. M., Boelens, R., and van Gunsteren, W. F. (1985) *J. Mol. Biol. 182*, 179−182.
30. Kania, J., and Brown, D. T. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 3529−3533.
31. Ogata, R. T., and Gilbert, W. (1979) *J. Mol. Biol. 132*, 709−728.
32. Friedman, A. M., Fischmann, T. O., and Steitz, T. A. (1995) *Science 268*, 1721−1727.
33. Chakerian, A. E., and Matthews, K. S. (1991) *J. Biol. Chem. 266*, 22206−22214.
34. Müller-Hill, B. (1975) *Prog. Biophys. Mol. Biol. 30*, 227−252.
35. Spronk, C. A. E. M., Slijper, M., van Boom, J. H., Kaptein, R., and Boelens, R. (1996) *Nature Struct. Biol. 3*, 916−919.
36. Chakerian, A. E., Tesmer, V. M., Manly, S. P., Brackett, J. K., Lynch, M. J., Hoh, J. T., and Matthews, K. S. (1991) *J. Biol. Chem. 266*, 1371−1374.
37. Alberti, S., Oehler, S., von Wilcken-Bergmann, B., Krämer, H., and Müller-Hill, B. (1991) *New Biol. 3*, 57−62.
38. Alberti, S., Oehler, S., von Wilcken-Bergmann, B., and Müller-Hill, B. (1993) *EMBO J. 12*, 3227−3236.
39. Chen, J., and Matthews, K. S. (1992) *J. Biol. Chem. 267*, 13843−13850.
40. Gilbert, W., and Maxam, A. (1973) *Proc. Natl. Acad. Sci. U.S.A. 70*, 3581−3584.
41. Ohlendorf, D. H., and Matthews, B. W. (1983) *Annu. Rev. Biophys. Bioeng. 12*, 259−284.
42. Gilbert, W., Maizels, N., and Maxam, A. (1973) *Cold Spring Harbor Symp. 38*, 845−855.
43. Reznikoff, W. S., Winter, R. B., and Hurley, C. K. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 2314−2318.
44. Gilbert, W., Gralla, F., Majors, J., and Maxam, A. (1975) In *Symposium on Protein−Ligand Interactions* (Sund, H., and Blauer, G., Eds), pp 193−206, Walter de Gruyter, Berlin.
45. Pfahl, M., Gulde, V., and Bourgeois, S. (1979) *J. Mol. Biol. 127*, 339−344.
46. Sasse-Dwight, S., and Gralla, J. D. (1988) *J. Mol. Biol. 202*, 107−119.
47. Oehler, S., Eismann, E. R., Krämer, H., and Müller-Hill, B. (1990) *EMBO J. 9*, 973−979.
48. Krämer, H., Niemöller, M., Amouyal, M., Revet, B., von Wilcken-Bergmann, B., and Müller-Hill, B. (1987) *EMBO J. 6*, 1481−1491.
49. Whitson, P. A., Hsieh, W.-T., Wells, R. D., and Matthews, K. S. (1987) *J. Biol. Chem. 262*, 4943−4946.
50. O'Gorman, R. B., Rosenberg, J. M., Kallai, O. B., Dickerson, R. E., Itakura, K., Riggs, A. D., and Matthews, K. S. (1980) *J. Biol. Chem. 255*, 10107−10114.
51. Dickson, R. C., Abelson, J., Barnes, W. M., and Reznikoff, W. S. (1975) *Science 187*, 27−35.
52. Schmitz, A., and Galas, D. J. (1979) *Nucleic Acids Res. 6*, 111−137.

53. Becker, M. M., and Wang, J. C. (1984) *Nature 309*, 682−687.
54. Caruthers, M. H., Beaucage, S. L., Efcavitch, J. W., Fisher, E. F., Goldman, R. A., deHaseth, P. L., Mandecki, W., Matteucci, M. D., Rosendahl, M. S., and Stabinsky, Y. (1983) *Cold Spring Harbor Symp. Quantum Biol. 47*, 411−418.
55. Betz, J. L., Sasmor, H. M., Buck, F., Insley, M. Y., and Caruthers, M. H. (1986) *Gene 50*, 123−132.
56. Boelens, R., Scheek, R. M., van Boom, J. H., and Kaptein, R. (1987) *J. Mol. Biol. 193*, 213−216.
57. Lehming, N., Sartorius, J., Niemöller, M., Genenger, G., von Wilcken-Bergmann, B., and Müller-Hill, B. (1987) *EMBO J. 6*, 3145−3153.
58. Sadler, J. R., Sasmor, H., and Betz, J. L. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 6785−6789.
59. Simons, A., Tils, D., von Wilcken-Bergmann, B., and Müller-Hill, B. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 1624−1628.
60. Sasmor, H. M., and Betz, J. L. (1990) *Gene 89*, 1−6.
61. Spronk, C. A. E. M., Folkers, G. E., Noordman, A. G. W., Wechselberger, R., van den Brink, N., Boelens, R., and Kaptein, R. (1999) *EMBO J. 18*, 6472−6480.
62. Falcon, C. M., and Matthews, K. S. (1999) *J. Biol. Chem. 274*, 30849−30857.
63. Koudelka, G. B. (1998) *Nucleic Acids Res. 26*, 669−675.
64. Falcon, C. M., Swint-Kruse, L., and Matthews, K. S. (1997) *J. Biol. Chem. 272*, 26818−26821.
65. Wong, I., and Lohman, T. M. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 5428−5432.
66. Frank, D. E., Saecker, R. M., Bond, J. P., Capp, M. W., Tsodikov, O. V., Melcher, S. E., Levandoski, M. M., and Record, M. T., Jr. (1997) *J. Mol. Biol. 267*, 1186−1206.
67. Calladine, C. R., and Drew, H. R. (1997) *Understanding DNA*, 2nd ed., pp. 18−64, Academic Press, San Diego.
68. Mossing, M. C., and Record, M. T., Jr. (1985) *J. Mol. Biol. 186*, 295−305.
69. El-Deiry, W. S., Kern, S. E., Pietenpol, J. A., Kinzler, K. W., and Vogelstein, B. (1992) *Nature Genet. 1*, 45−49.
70. von Hippel, P. H., and McGhee, J. D. (1972) *Annu. Rev. Biochem. 41*, 231−300.
71. Takeda, Y., Ohlendorf, D. H., Anderson, W. F., and Matthews, B. W. (1983) *Science 221*, 1020−1025.
72. Caruthers, M. H. (1980) *Acc. Chem. Res. 13*, 155−160.
73. Capp, M. W., Cayley, D. S., Zhang, W., Guttman, H. J., Melcher, S. E., Saecker, R. M., Anderson, C. F., and Record, M. T., Jr. (1996) *J. Mol. Biol. 258*, 25−36.
74. Jobe, A., Sadler, J. R., and Bourgeois, S. (1974) *J. Mol. Biol. 85*, 231−248.
75. Ogata, R., and Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 4973−4976.
76. Goeddel, D. V., Yansura, D. G., and Caruthers, M. H. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 3292−3296.
77. Goeddel, D. V., Yansura, D. G., and Caruthers, M. H. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 3578−3582.
78. Sadler, J. R., Tecklenburg, M., Betz, J. L., Goeddel, D. V., Yansura, D. G., and Caruthers, M. H. (1977) *Gene 1*, 305−321.
79. Betz, J. L., and Fall, M. Z. (1988) *Gene 67*, 147−158.
80. Härd, T., and Lundbäck, T. (1996) *Biophys. Chem. 62*, 121−139.
81. Gartenberg, M. R., and Crothers, D. M. (1988) *Nature 333*, 824−829.
82. Suzuki, M., and Yagi, N. (1995) *Nucleic Acids Res. 23*, 2083−2091.
83. Dickerson, R. E., and Chiu, T. K. (1997) *Biopolymers 44*, 361−403.
84. Davis, N. A., Majee, S. S., and Kahn, J. D. (1999) *J. Mol. Biol. 291*, 249−265.
85. Giese, K., Pagel, J., and Grosschedl, R. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 12845−12850.
86. Wang, J. C., Barkley, M. D., and Bourgeois, S. (1974) *Nature 251*, 247−249.
87. Koudelka, G. B., and Carlson, P. (1992) *Nature 355*, 89−91.
88. Gorin, A. A., Zhurkin, V. B., and Olson, W. K. (1995) *J. Mol. Biol. 247*, 34−48.
89. Lahm, A., and Suck, D. (1991) *J. Mol. Biol. 221*, 645−667.
90. Harrington, R. E. (1992) *Mol. Microbiol. 6*, 2549−2555.
91. Lesser, D. R., Kurpiewski, M. R., and Jen-Jacobson, L. (1990) *Science 250*, 776−786.
92. Werner, M. H., Gronenborn, A. M., and Clore, G. M. (1996) *Science 271*, 778−784.
93. Pérez-Martín, J., and de Lorenzo, V. (1997) *Annu. Rev. Microbiol. 51*, 593−628.
94. Dickerson, R. E. (1998) *Nucleic Acids Res. 26*, 1906−1926.
95. Karslake, C., Botuyan, M. V., and Gorenstein, D. G. (1992) *Biochemistry 31*, 1849−1858.
96. Botuyan, M. V., Keire, D. A., Kroen, C., and Gorenstein, D. G. (1993) *Biochemistry 32*, 6863−6874.
97. Packer, M. J., and Hunter, C. A. (1998) *J. Mol. Biol. 280*, 407−420.
98. Subirana, J. A., and Faria, T. (1997) *Biophys. J. 73*, 333−338.
99. Koudelka, G. B., Harrison, S. C., and Ptashne, M. (1987) *Nature 326*, 886−888.
100. Koudelka, G. B., Harbury, P., Harrison, S. C., and Ptashne, M. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 4633−4637.
101. Wu, L., Vertino, A., and Koudelka, G. B. (1992) *J. Biol. Chem. 267*, 9134−9139.
102. Hilchey, S. P., and Koudelka, G. B. (1997) *J. Biol. Chem. 272*, 1646−1653.
103. Bell, A. C., and Koudelka, G. B. (1993) *J. Mol. Biol. 234*, 542−553.
104. Spronk, C. A. E. M., Bonvin, A. M. J. J., Radha, P. K., Melacini, G., Boelens, R., and Kaptein, R. (1999) *Structure 7*, 1483−1492.
105. Bell, C. E., and Lewis, M. (2000) *Nature Struct. Biol. 7*, 209−214.
106. Schumacher, M. A., Choi, K. Y., Zalkin, H., and Brennan, R. G. (1994) *Science 266*, 763−770.
107. Pabo, C. O., and Sauer, R. T. (1984) *Annu. Rev. Biochem. 53*, 293−321.
108. Gilbert, W., Maxam, A., and Mirzabekov, A. (1976) *Control of Ribosome Synthesis, Alfred Benzon Symposium IX*, 139−143.
109. Allen, T. D., Wick, K. L., and Matthews, K. S. (1991) *J. Biol. Chem. 266*, 6113−6119.
110. Wick, K. L., and Matthews, K. S. (1991) *J. Biol. Chem. 266*, 6106−6112.
111. Nick, H., Arndt, K., Boschelli, F., Jarema, M. A. C., Lillis, M., Sadler, J., Caruthers, M., and Lu, P. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 218−222.
112. Nick, H., Arndt, K., Boschelli, F., Jarema, M. A., Lillis, M., Sommer, H., Lu, P., and Sadler, J. (1982) *J. Mol. Biol. 161*, 417−438.
113. Rastinejad, F., Artz, P., and Lu, P. (1993) *J. Mol Biol. 233*, 389−399.
114. Schroeder, S. A., Roongta, V., Fu, J. M., Jones, C. R., and Gorenstein, D. G. (1989) *Biochemistry 28*, 8292−8303.
115. Zhang, X., and Gottlieb, P. A. (1995) *Nucleic Acids Res. 23*, 1502−1511.
116. Jørgensen, C. I., Kallipolitis, B. H., and Valentin-Hansen, P. (1998) *Mol. Microbiol. 27*, 41−50.
117. Pedersen, H., and Valentin-Hansen, P. (1997) *EMBO J. 16*, 2108−2118.
118. Carson, M. (1991) *J. Appl. Crys. 24*, 985−961.

BI000924Z