Engineered Disulfide Linking the Hinge Regions within Lactose Repressor Dimer Increases Operator Affinity, Decreases Sequence Selectivity, and Alters Allostery†

Catherine M. Falcon‡ and Kathleen S. Matthews*

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

Received July 6, 2001; Revised Manuscript Received October 17, 2001

ABSTRACT: The hinge domain encompasses amino acids 51−60 of lactose repressor (LacI) and plays an important role in its regulatory interaction with operator DNA. This segment makes both hinge−DNA and hinge−hinge′ contacts that are critical to DNA binding. Furthermore, this small region serves as a central element in communicating the allosteric response to inducer. Introducing a disulfide bond between partner hinges within a dimer via the mutation V52C results in a protein that has increased affinity for O^1 operator DNA compared to wild-type LacI and abolishes allosteric response to inducer [Falcon, C. M., Swint-Kruse, L., and Matthews, K. S. (1997) *J. Biol. Chem. 272*, 26818]. We have established that this high affinity is maintained for the disulfide-linked protein even when symmetry and half-site spacing within the operator region are altered, whereas binding by the reduced protein, as for wild-type LacI, is severely diminished by these alterations. Interestingly, the allosteric response to inducer for V52C-oxidized remains intact for a small group of operator variants. Temperature studies demonstrate that the presence of the disulfide alters the thermodynamics of the protein−DNA interaction, with a ΔC_p of significantly smaller magnitude compared to wild-type LacI. The results presented here establish the hinge region as an important element not only for LacI high-affinity operator binding but also for the essential communication between ligand binding domains. Moreover, the results confirm that DNA sequence/conformation can profoundly influence allostery for this prototypic regulatory protein.

Determining how transcription factors identify their cognate DNA sites and how these interactions can be regulated is important for understanding cellular processes. Both protein and DNA contribute features of import to the interaction, and the intertwined roles of these molecules in recognition and regulation are complex (*1*). One approach to define required elements is to introduce specific alterations in protein and/or DNA sequences. For DNA in particular, the effects of cumulative alterations provide a means to assess the role of selected regions in generating the overall stability of the complex and, where applicable, the allosteric response of the protein to ligand binding.

We have utilized the *lac* repressor (LacI),[1] a well-studied transcription regulator from *Escherichia coli* (*2*, *3*), to assess the role of both protein and DNA components in recognition and allostery. LacI is a 150 000 dalton homotetramer that binds to a specific DNA sequence within the *E. coli* genome to inhibit transcription of the lactose metabolic enzymes (*4*−*8*). Repression is relieved when an effector molecule (inducer) binds to a distant site within the protein and elicits a conformational change that precludes specific DNA binding (*7*−*9*).

The first ∼50 amino acids of each monomer comprise the helix−turn−helix DNA binding domain (*10*−*12*). The core domain (amino acids ∼61−340) encompasses the surface for monomer−monomer interaction as well as the inducer binding site (*13*, *14*). Although the dimeric unit is sufficient for DNA binding, the C-terminal 20 amino acids form a leucine heptad repeat dimer−dimer assembly motif that allows formation of looped DNA structures via the two DNA binding sites (reviewed in ref *3*). Finally, the hinge region (amino acids 51−60) links the DNA and inducer binding domains (*15*, *16*). This hinge segment plays a key role in binding to operator and is also presumed to provide a communicative link between the binding domains (*15*−*17*). The hinge folds into an α-helix (*16*) that appears only in the presence of DNA (*17*). Unfolding may be precipitated by disruption of the interface between the partner hinges within a dimer. When inducer binds the distant core domain, consequent reorientation of the N-subdomain of the core potentially disrupts the stabilizing interactions (*16*, *18*, *19*). Without the structure of the hinge helices, the protein is no longer able to make the appropriate contacts for high-affinity binding and binds to DNA only in a nonspecific manner.

Previously, we have reported a valine to cysteine mutation at position 52 (V52C) within the hinge region (Figure 1) (*20*), a site at which some substitutions generate a tight-binding mutant phenotype (*21*, *22*). Position 52 is located toward the N-terminus of the hinge segment near the linker

† Support for this project was provided by grants to K.S.M. from the 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; email: ksm@rice.edu.

‡ Current address: Tanox, Inc., 10301 Stella Link, Houston, TX 77025.

[1] Abbreviations: CytR, cytosine repressor protein; EDTA, ethylenediaminetetraacetic acid; IPTG, isopropyl-β,D-thiogalactoside; LacI, lactose repressor protein; ONPF, *o*-nitrophenyl-β,D-fucoside; PurR, purine repressor protein.

10.1021/bi0114067 CCC: $20.00
Published on Web 12/01/2001

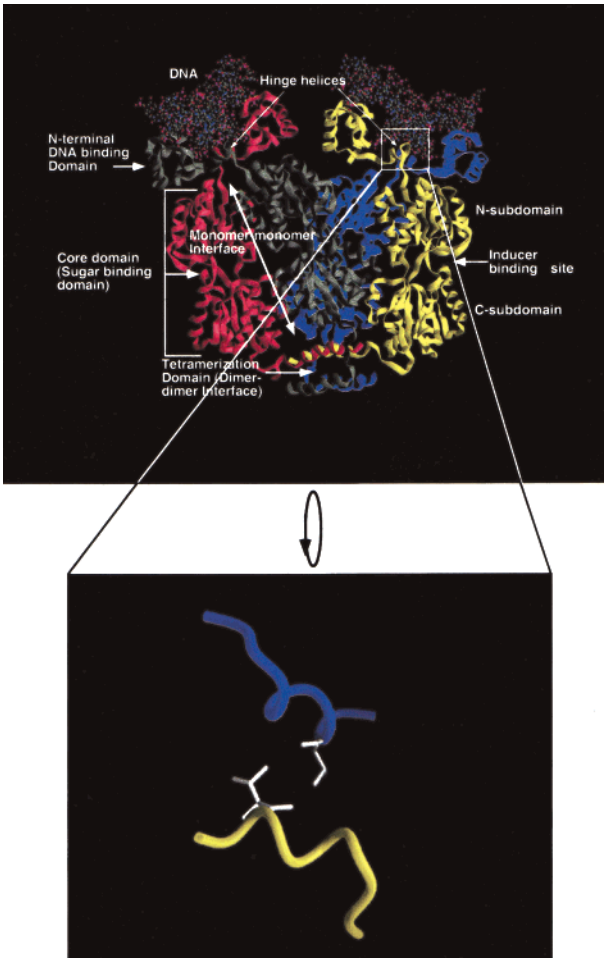

FIGURE 1: Structure of LacI tetramer bound to operator. The tetramer is shown bound to the symmetric operator, from pdb file 1LBG. The structure is displayed using Ribbons (*55*). The monomers are shown in different colors. The expansion shows the indicated region rotated by 90° out of the page in order to show the relationship between the α-helical, folded hinge segments within a dimer in the DNA-bound conformation. The amino acid side chains corresponding to position 52 are highlighted in white. Note the close apposition of these side chains that is required for disulfide bond formation.

to the N-terminal DNA binding domain, a region postulated by Bell and Lewis (*19*) to be critical for LacI allosteric response. The V52C change provides a covalent hinge−hinge′ linkage under conditions that favor disulfide formation (the prime designation denotes the partner subunit within a dimer). In its disulfide form, V52C protein exhibits higher affinity than wild-type LacI for the natural operator, O^1, but with an accompanying loss of allosteric response despite wild-type affinity for inducer (*20*). Indeed, introducing this mutation into the N-terminal 62 amino acid DNA binding fragment resulted in affinity comparable to full-length LacI for the disulfide-linked species (*23*). Under reducing conditions, the covalent bond between hinges is disrupted to restore behavior similar to wild-type LacI (*20*).

This reversible covalent connection between the hinges provides the opportunity to dissect the role of hinge−hinge′ interaction in the function of LacI. We have recently identified a set of operator DNA variants that profoundly affect affinity and allosteric communication for a series of mutants with an increasing number of glycine residues between the hinge segment and core inducer binding domain (*24*, *25*). Using these variants of the wild-type operator DNA with the V52C LacI protein may allow us to parse various contributions to DNA binding and allosteric communication. We find that the presence of the disulfide has a profound influence on the affinity of LacI for operator DNAs. By stabilizing the proximity of the N-terminal region of the hinge segments within a dimer, the selectivity for sequence is diminished, with consequent high affinity for variant sequences that are not recognized by wild-type LacI. Interestingly, a subset of these altered DNA sequences results in effective inducer response by the oxidized form of the protein.

One possible source of this differential behavior is the coil to helix transition that appears to occur upon association with operator DNA (*19*, *23*). In situations in which specific binding of protein to DNA is coupled to a folding event, a large negative ΔC_p is anticipated (*26*). Using the oxidized form of V52C protein, wild-type LacI, and one of the hinge mutants, Q60G, in which additional flexibility was engineered at the C-terminal end of the hinge (*24*), we have examined the temperature dependence of DNA binding to assess the influence of hinge helix folding coupled with LacI−DNA binding. In contrast to wild-type LacI and Q60G proteins, binding of V52C-oxidized to DNA yields a very small, negative ΔC_p. This result suggests, as might be anticipated, that protein folding is not coupled to DNA binding in the presence of the disulfide bond.

The results of these studies establish clearly the significant role of the hinge region of *lac* repressor not only in high-affinity DNA binding but also in sequence selectivity and in the allosteric response. By stabilizing the juxtaposition of partner hinges within a dimer, LacI binding to DNA can be enhanced, but sequence stringency is diminished. Of particular importance, DNA sequence influence on the allosteric response to inducer is evident, an observation of considerable interest in understanding the ways in which DNA sequence may modulate the function of regulatory proteins.

EXPERIMENTAL PROCEDURES

DNA Construct. Plasmid pJC1 (*27*) contains the LacI gene and was used to mutate position 52 (valine) to a cysteine using the double-stranded mutagenesis method (Chameleon, Stratagene) as previously reported (*20*).

Protein Expression and Purification. The altered pJC1 plasmid containing the V52C mutation was transformed into BLIM cells for expression (*28*). The expressed protein was purified using a phosphocellulose column or a heparin column, with a final gradient of 0.12−0.3 M potassium phosphate, pH 7.6, 5% glucose used to elute the protein (*20*). The fractions containing V52C LacI were pooled and stored at −70 °C.

Oxidation and Reduction of V52C Protein. The purified protein was dialyzed overnight against 0.1 M Tris-HCl, pH 7.6, and 0.1 M NaCl without dithiothreitol. To further favor disulfide formation, 10 mM glutathione, oxidized form (Sigma), was added to the protein solution. To favor reduction of the cysteine, 10 mM glutathione, reduced form (Sigma), was added to the protein solution. Oxidation of the cysteine was most efficient when the protein was very dilute, <0.5 mg/mL.

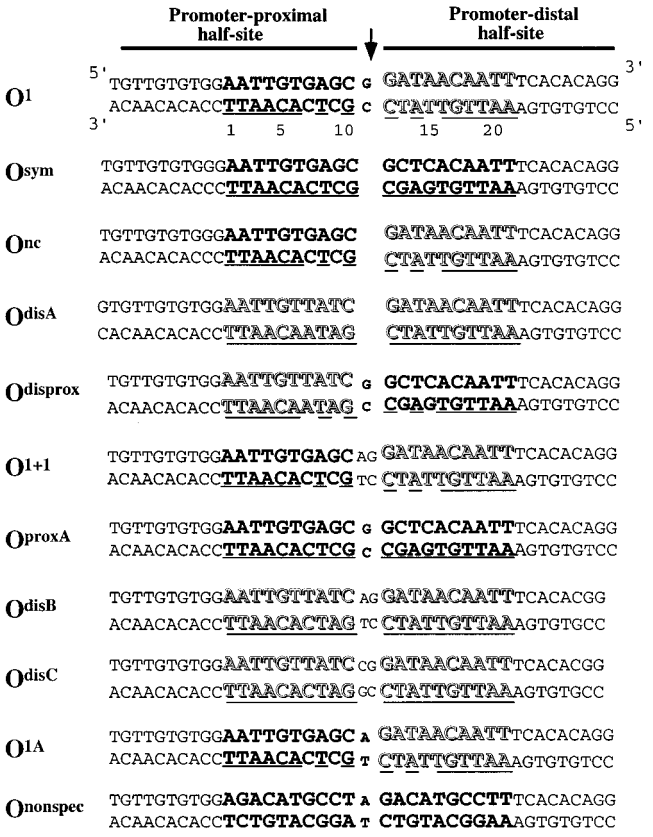

FIGURE 2: Sequence of the variant operators used for the DNA binding studies. The 40 base pair sequences are shown. The region of the operator that is protected by *lac* repressor from DNase footprinting is printed in boldface or outlined print, while flanking sequences are shown in smaller font. Symmetric sites within each operator are underlined. The two half-sites are labeled as proximal and distal based on the natural operator, O^1, with the point of symmetry for the sequences indicated by the arrow. The numbering of the bases is according to the transcription start site for the *lac* operon (*56*).

Operator Binding. DNA binding was measured using nitrocellulose filters in a 96-well plate, as previously reported (*20*), using a variation of the technique developed by Wong and Lohman (*29*). Each individual experiment contained duplicate points, and data from multiple experiments involving at least two separate operator and protein preparations were utilized in these studies. Assays were performed at room temperature in buffer containing 0.01 M Tris-HCl, pH 7.6, 0.15 M KCl, 5% dimethyl sulfoxide, 0.1 mM EDTA. Proteins were diluted in the buffer with the addition of 50 μg/mL bovine serum albumin. Where required, to ensure cysteine reduction, 10 mM DTT was added to the dilution buffer (no glutathione was added to the dilution buffers for this assay since it interfered with binding on the nitrocellulose filters, data not shown). The 40 base pair operators were synthesized commercially (Great American Gene Co. and Genosys) and heated to 80 °C for several minutes and then cooled slowly using annealing buffer (8 mM Tris-HCl, pH 7.5, 4 mM $MgCl_2$, and 10 mM NaCl). The complete 40 base pair sequences of the operators used in this study are shown in Figure 2. For these assays, the DNA concentration was either 1×10^{-13} or 1×10^{-12} M, and the protein concentration varied from 1×10^{-14} to 1×10^{-6} M, depending on the expected K_d value. Assay results were imaged using a Fugi phosphor imaging plate and quantitated using the MacBas program. Using IgorPro (Wavemetrics), these data were then fit to the equation:

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

where R is the fractional saturation, Y_m is a correction factor that allows the maximum value of R to float, K_d is the apparent equilibrium dissociation constant, and [P] is the free protein concentration in tetramer. The free protein concentration is equivalent to the total protein concentration under conditions at which the K_d value is >10-fold above the DNA concentration. However, for conditions where the DNA concentration is closer to the K_d value (i.e., where K_d values were only ~7-fold greater than operator concentration), the applicability of this method was evaluated by comparison to the analytical solution to the protein and DNA mass conservation equations for [P], an approach requiring no assumptions. The fits generated by this method were essentially identical and yielded very similar K_d values. Values reported were therefore derived using eq 1 and assumed that free and total protein concentrations were equivalent. Binding in the presence of saturating amounts of inducer (IPTG) was monitored and compared to binding of the protein without inducer.

Temperature Studies. To study the dependence of K_d on temperature, DNA binding assays were performed over a range of temperatures from 0 to 41 °C. For these assays, the final DNA concentration was either 1×10^{-13}, 5×10^{-13}, 9×10^{-13}, or 6×10^{-11} M, depending on the affinity of the protein for the particular operator studied. The final protein concentration ranged from 1×10^{-13} to 1×10^{-8} M in the buffer described for DNA binding assays. For consistency, we used the same buffer for the temperature studies as that for DNA binding, even though the pH of Tris buffers is temperature-sensitive. Variation in pH over the temperature range utilized for these measurements was determined to be small, and the pH effect on operator binding in this buffer over this range is also minimal (*30*). For each protein−DNA assay, a stock solution of protein and DNA at the final concentrations mentioned was made and kept at 0 °C (*31*). To obtain binding at the designated temperature, a set volume was taken from the stock and equilibrated at the appropriate temperature for at least 30 min in a temperature-controlled water bath. After equilibration, samples were filtered onto presoaked nitrocellulose filters as rapidly as possible to minimize temperature fluctuations and then dried. Data were collected from at least three separate assays performed in duplicate and analyzed by MacBas version 2.0 (Fugi) to determine K_d values. Temperature dependence was analyzed using van't Hoff plots, using the following modified equation that assumes a large, negative ΔC_p (*31*):

$$-\log K_d = \frac{\Delta C_p}{2.303R}[T_H/T - \ln(T_S/T) - 1] \quad (2)$$

where K_d is the apparent equilibrium dissociation constant, ΔC_p is the heat capacity, T_H is the temperature at which ΔH is zero, and T_S is the temperature at which ΔS is zero. The data were analyzed using the program IgorPro (Wavemetrics).

RESULTS

Symmetric Operator. Previous work established that introducing a cysteine at position 52 within the LacI hinge

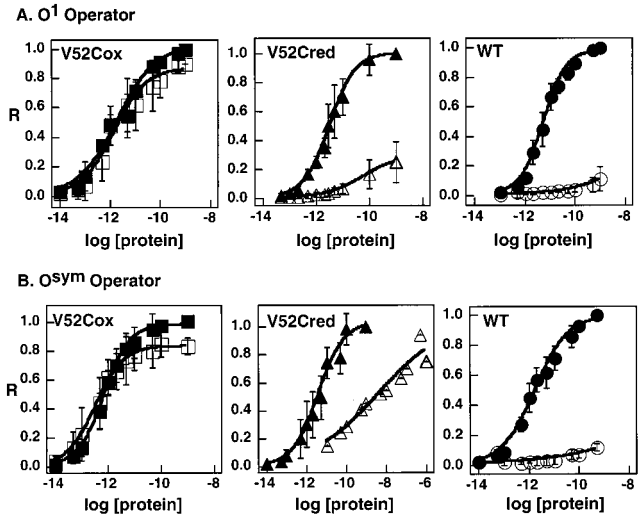

FIGURE 3: Binding isotherms of the proteins with operators O^1 and O^{sym}. V52C-oxidized (■), V52C-reduced (▲), and wild-type LacI (●) were each assayed with the 40 base pair operator under the conditions indicated under Experimental Procedures. The open symbols represent binding experiments in the presence of 1 mM IPTG. The lines drawn represent the fits generated using eq 1. Individual data points represent the average of at least 3 separate experiments, each performed in duplicate. (A) Wild-type, V52C-reduced, and V52C-oxidized proteins binding to O^1 operator; (B) wild-type, V52C-reduced, and V52C-oxidized proteins binding to O^{sym} operator.

produced a protein that binds O^1, the natural operator, with slightly greater affinity than the wild-type protein in its reduced form (*20*). With disulfide formation, the oligomeric state of the protein is not altered, but each monomer becomes covalently linked to its partner within a dimeric pair. V52C-oxidized binds O^1 operator with ~5-fold greater affinity compared to wild-type LacI and demonstrates a complete lack of response to inducer (Figure 3). Although V52C-reduced binds more tightly to O^1 compared to wild-type LacI, the allosteric communication remains intact so that the affinity in the presence of IPTG diminishes to near nonspecific levels (*20*). Both proteins bind inducer with the same affinity as wild-type LacI when no DNA is present (*20*). To assess whether the increased operator binding affinity of V52C results from overall higher affinity for any DNA sequence, as observed for some tight-binding LacI mutants (*32*), V52C was assayed with a series of variant operators (*25*).

The majority of structural work to date for LacI−DNA interactions has utilized the "ideal" operator sequence, O^{sym}, in which the target site is completely symmetric (*16*−*19*, *33*−*35*). V52C-oxidized binds O^{sym} with ~3-fold higher affinity compared to O^1 and demonstrates minimal response to inducer (Table 1, Figure 3). Indeed, V52C-oxidized may exhibit a very slightly greater affinity for O^{sym} in the presence of inducer. This result suggests that in this instance, IPTG may elicit change similar to that of an anti-inducer, stabilizing the LacI−LacO complex (*36*). V52C-reduced demonstrates comparable affinity for O^{sym} compared to O^1; however, the response to the presence of inducer, while still present, is diminished by ~100-fold (10 mM DTT is present in the buffer). This decreased allosteric response may result from very low residual levels of oxidized protein even though the conditions for this assay should strongly discourage oxidation. DNA may bring the V52 side chains into optimal alignment for disulfide formation despite the presence of reducing agent. The observation that introducing the V52C disulfide into the N-terminal 1−62 fragment generates wild-type LacI binding affinity suggests optimal alignment for oxidation in the complex with DNA (*23*). Additional variations in DNA sequence have been explored to determine the effect on affinity and allostery for V52C-oxidized and V52C-reduced LacI.

Table 1: Binding of V52C to Variant Operators

	V52C protein ($K_d \times 10^{11}$ M)			
	oxidized		reduced	
operator	no IPTG	+IPTG	no IPTG	+IPTG
O^1	0.21 ± 0.05	0.15 ± 0.05	0.34 ± 0.02	>10000
O^{sym}	0.076 ± 0.005	0.036 ± 0.003	0.43 ± 0.12	470 ± 100
O^{nc}	0.080 ± 0.07	1.7 ± 1.0	300 ± 200	>10000
O^{disA}	200 ± 200	>10000	>10000	>10000
$O^{disprox}$	0.47 ± 0.11	20 ± 18	~10000	>10000
O^{1+1}	~10000	>10000	>10000	>10000
O^{proxA}	0.063 ± 0.02	0.63 ± 0.1	70 ± 30	200 ± 200
O^{disB}	2.2 ± 0.7	>10000	>10000	>10000
O^{disC}	0.074 ± 0.011	0.14 ± 0.03	3.5 ± 1.1	200 ± 100
O^{1A}	0.26 ± 0.15	7.1 ± 5.1	30 ± 3	1100 ± 400
$O^{nonspecific}$	1500 ± 500	>10000	>10000	>10000

Removal of the Central Base Pair. One of the differences between the O^{sym} and O^1 operators is that the central base pair (G/C) is not present in the symmetric sequence. Thus, the point of symmetry for the O^{sym} operator lies between the central C-G base sequence (Figure 2). The O^{nc} operator is identical to both O^1 half-sites but lacks its central G/C base pair. O^{nc} has similar half-site spacing to O^{sym}, but differs from the symmetric operator at positions 12 and 14, which disrupts the complete symmetry. When assayed with the O^{nc} operator, V52C-oxidized exhibits affinity comparable to that for O^{sym}; however, a significant response to inducer, on the order of ~20-fold decrease in affinity, is observed (Table 1, Figure 4). Therefore, this spacing change in the central region of the operator does not alter the affinity for the oxidized mutant but provides structural alterations that allow communication between the DNA and inducer binding sites. In contrast, V52C-reduced has a lower affinity, by ~1000-fold compared to O^1, and binding is reduced further by inducer presence.

Half-Site Symmetry. O^{sym} was constructed originally on the basis of the higher affinity of LacI for the promoter-proximal half-site (*34*). For comparison, we have designed an operator with symmetry based on the promoter-distal half-site, but with the spacing of O^{sym} (Figure 2). This operator, termed O^{disA} (distal half-sites with no central base), was assayed with both V52C-oxidized and V52C-reduced. In this instance, the affinity of the oxidized protein for this sequence is lower by ~1000-fold compared to O^1, and binding is further diminished to nonspecific levels in the presence of inducer (Table 1). Thus, despite stabilization of the hinge region, V52C-oxidized binds with low affinity to this sequence, and an inducer effect is observed. V52C-reduced demonstrates no detectable specific binding for O^{disA} and has comparably low affinity in the presence of inducer. This pattern is similar to that observed for wild-type LacI and this operator sequence (*25*, *37*).

Reordering of Half-Sites. The operator $O^{disprox}$ has the proximal and distal half-sites of O^1 inverted but maintains

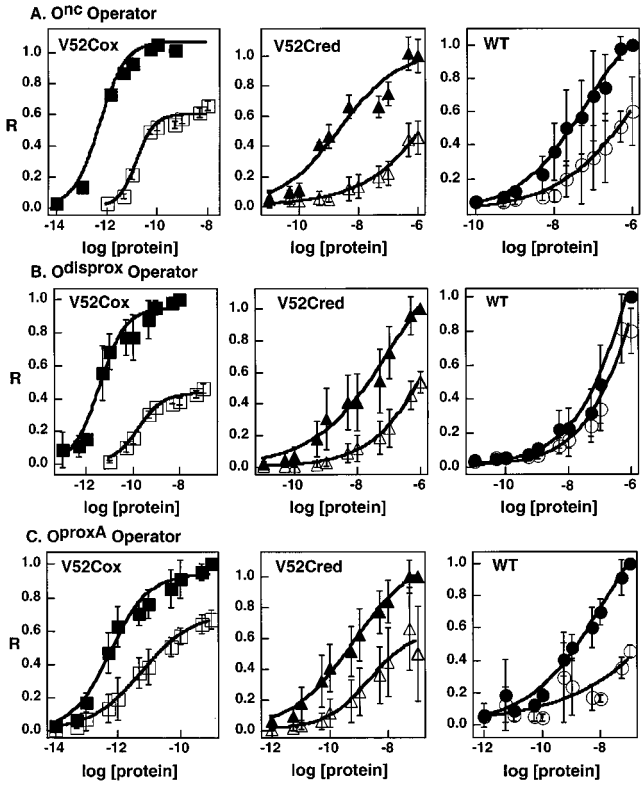

FIGURE 4: Binding isotherms of the proteins with operator variants. V52C-oxidized (■), V52C-reduced (▲), and wild-type LacI (●) were each assayed with the indicated operator under the conditions described under Experimental Procedures. These assays were also performed the presence of 1 mM IPTG (open symbols). The lines drawn represent the fits generated using eq 1. Individual data points represent the average of at least 3 separate experiments, each performed in duplicate. (A) Wild-type, V52C-reduced, and V52C-oxidized proteins binding to O^{nc} operator; (B) wild-type, V52C-reduced, and V52C-oxidized proteins binding to $O^{disprox}$ operator; (C) wild-type, V52C-reduced, and V52C-oxidized proteins binding to O^{proxA} operator. Note that the *x*-axes differ in order to show complete binding curves.

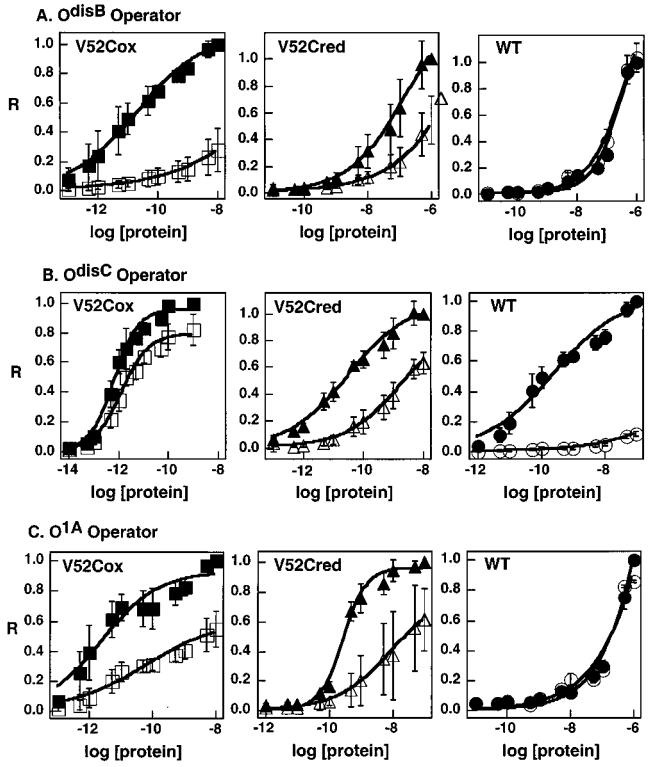

FIGURE 5: Binding isotherms of the proteins with operator variants. V52C-oxidized (■), V52C-reduced (▲), and wild-type LacI (●) were each assayed with the indicated operator under the conditions described under Experimental Procedures. These assays were also done in the presence of 1 mM IPTG (open symbols). The lines drawn represent the fits generated using eq 1. Individual data points represent the average of at least 3 separate experiments, each performed in duplicate. (A) Wild-type, V52C-reduced, and V52C-oxidized proteins binding to O^{disB} operator; (B) wild-type, V52C-reduced, and V52C-oxidized proteins binding to O^{disC} operator; (C) wild-type, V52C-reduced, and V52C-oxidized proteins binding to with O^{1A} operator. Note that the *x*-axes differ in order to show complete binding curves.

the central base, which essentially rotates the half-sites around the central base pair/axis of symmetry (Figure 2). Wild-type LacI binds this particular operator with low affinity and exhibits no alteration in the presence of inducer (*25*). V52C-reduced has very low affinity for this operator and demonstrates a marginal inducer response. However, V52C-oxidized maintains very high affinity, but shows significant response to inducer with a 2 order of magnitude decrease in affinity (Table 1, Figure 4).

Insertion of Base between Half-Sites. Presuming that the disulfide confines the spacing between the two hinge regions, we explored whether V52C-oxidized can bind specifically to an operator with half-sites spaced further than those found in O^1 and O^{sym} sequences. The O^{1+1} operator has an additional A/T base pair inserted between the proximal and distal half-sites, so that these sites are one base further apart in comparison to O^1 (Figure 2). The reduced form of the protein and wild-type LacI both show extremely low binding affinity for this sequence (*25*). V52C-oxidized also has very low affinity for this operator (Table 1). However, a small additional decrease in affinity is observed in the presence of inducer. Thus, there appears to be an inducer response, but the K_d value is sufficiently high to preclude a quantitative estimate of the difference between the two conditions.

Another operator was constructed in which symmetrically disposed half-sites were separated by a central G/C, so that the symmetry corresponds to O^{sym} but the spacing to O^1 (Figure 2). The reduced protein displays lower affinity for this operator, O^{proxA}, compared to O^1 ($K_d = 7 \times 10^{-10}$ M), although the presence of inducer elicits a ~3-fold decrease in affinity. For this DNA sequence, V52C-oxidized has a very high affinity ($K_d = 6 \times 10^{-13}$ M), and a 10-fold decrease in affinity is observed when inducer is present (Table 1, Figure 4).

Previous work demonstrated that wild-type LacI binds to an operator with symmetry based on the distal half-sites but with increased spacing between the sites (*18*, *37*). This operator, termed here O^{disB}, was constructed with an A-G central sequence (Figure 2). For V52C-reduced and wild-type protein (*24*), binding to this operator is of very low affinity, whereas V52C-oxidized bound with remarkably high affinity ($K_d = 2.2 \times 10^{-11}$ M) (Table 1, Figure 5). However, in the presence of inducer, the V52C-oxidized affinity for O^{disB} decreases to nonspecific levels, consistent with an effective allosteric transition despite the presence of the disulfide. When the central sequence is converted to C-G and combined with symmetric distal half-sites (O^{disC}), the reduced protein and wild-type LacI (*35*) exhibit high affinity with <100-fold decrease in the presence of saturating

amounts of inducer for V52C-reduced (Table 1, Figure 5). Indeed, the affinity of V52C-oxidized for the O^{disC} sequence is substantially higher ($K_d = 7 \times 10^{-13}$ M) than for O^{disB}, but, surprisingly, only a 2-fold decrease in affinity is observed when a saturating amount of inducer is present (Figure 5). This striking difference in affinity and allosteric communication derives solely from the variation of A to C in the central sequence—almost 2 orders of magnitude difference in the affinity of V52C-oxidized protein for DNA and $\sim 10^5$-fold difference in the protein−DNA interaction in the presence of inducer. Clearly, the central region of the operator exerts a significant influence on the binding interaction, as observed previously for the LacI variants (*25*) and the 454 and p22 repressors (*38*−*45*).

Alternate Central Base. The central base sequence of the operator appears to dramatically affect the binding interaction (*46*). To explore this phenomenon further, the central sequence was altered to A/T, while maintaining the half-site spacing and sequence of O^1. Converting the central base pair of O^1 to A/T in this O^{1A} operator profoundly affects binding by wild-type LacI and other mutant proteins with essentially nonspecific levels of binding (*25*, *47*). The reduced form of V52C has diminished affinity for O^{1A}, but maintains an allosteric response to inducer. However, V52C-oxidized binds with high affinity to this DNA sequence (Table 1, Figure 5), and a modest response (~30-fold) to inducer is observed.

Nonspecific Sequence. A subset of tight-binding LacI mutants exhibits increased affinity for nonoperator DNA (*32*, *48*). To determine if this were the case for V52C in either state, the oxidized and reduced forms of the protein were assayed with a nonspecific DNA sequence. In this instance, we used the recognition site of p53 protein (*49*). V52C-oxidized has a measurable binding affinity for this operator, with a K_d of $\sim 1 \times 10^{-8}$ M, and exhibits a decrease in affinity when assayed in the presence of inducer (Table 1). V52C-reduced exhibits low affinity, comparable to wild-type protein, with no change in affinity in the presence of inducer. Thus, the presence of the disulfide even enhances the affinity for nonoperator sequence by ~10-fold, and also allows for a moderate allosteric response. Whether this enhancement derives from hinge helix folding with nonspecific DNA or entropic effects of disulfide formation is not clear.

Temperature Studies. To assess the contribution of coupled hinge folding with DNA binding in the V52C-oxidized protein, we examined DNA binding as a function of temperature to determine ΔC_p for the interaction. A large negative ΔC_p reflects the burial of apolar residues concomitant with a conformational change that is coupled to protein binding with DNA (*26*). V52C-oxidized was compared with wild-type LacI as well as a mutant protein with an alteration at the C-terminal region of the hinge, Gln60Gly (*24*). We can thus compare wild-type protein to mutants—one with flexibility within the hinge (Gln60Gly) and one with restricted hinge motions (V52C-oxidized). For these studies, two different operators that exhibit high affinity for these three proteins were used to observe any differences in binding—the natural operator, O^1, and O^{sym}.

Substantial difference in binding to O^1 over the temperature range examined was observed among the three proteins (Figure 6A). Wild-type LacI generates a concave curve, which corresponds to a large negative ΔC_p (*26*, *31*). The

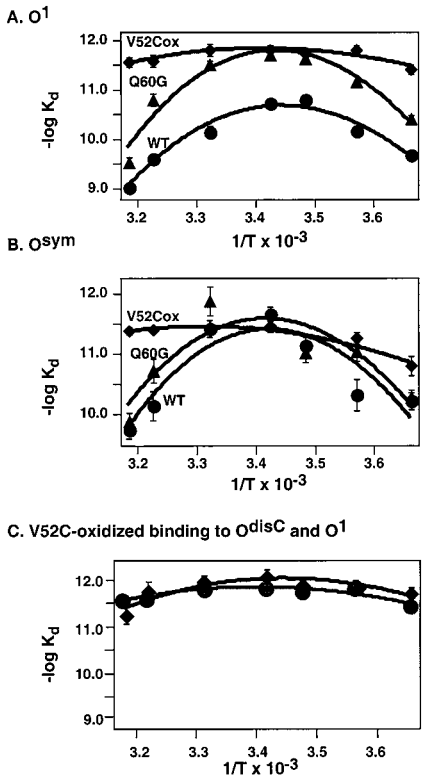

FIGURE 6: van't Hoff plots of V52C-oxidized, Q60G, and wild-type LacI binding to operators. V52C-oxidized (■), Q60G (▲), and wild-type LacI (●) were each assayed with the indicated operator at various temperatures from 0 to 41 °C. Individual data points represent the average of at least 3 separate experiments, each performed in duplicate. The lines drawn are the fits generated to the data using eq 2. (A) Wild-type, Q60G, and V52C-oxidized proteins binding to O^1 operator; (B) wild-type, Q60G, and V52C-oxidized proteins binding to O^{sym} operator; (C) V52C-oxidized binding to O^1 (●) and O^{disC} (◆) operators.

slightly greater curvature for Gln60Gly suggests that this mutation results in enhanced folding upon DNA binding, which might be ascribed to the flexibility introduced by the glycine. In contrast, V52C-oxidized binding exhibits minimal differences over the range of temperatures studied. The consequent small negative ΔC_p suggests that coupled folding and binding does not occur for this mutant. When assaying the same series of proteins with O^{sym}, only slight differences from O^1 binding were observed (Figure 6B). V52C-oxidized again exhibits little change in binding affinity over this broad temperature range. Because V52C-oxidized exhibits binding behavior for O^{disC} comparable to wild-type·O^1 binding, we examined the temperature dependence for V52C-oxidized binding this operator (Figure 6C). Similar to behavior with O^1 and O^{sym}, the affinity of V52C-oxidized for O^{disC} exhibits very little change over the entire temperature range, corresponding to a small, negative ΔC_p. Thus, ΔC_p for V52C-oxidized is comparable for each of the operator sequences examined. These temperature studies reveal a substantial difference in the thermodynamics that describe the interactions that occur between variant proteins and DNAs in this system even when the binding affinities are similar. These results suggest that a careful analysis of ΔC_p for a larger family of related protein and DNA sequences may illuminate structural interpretation of this thermodynamic parameter.

DISCUSSION

Protein−DNA interactions are central to cellular regulatory processes. One fundamental approach to dissecting the key features of these interactions is variation of either protein or DNA sequence with subsequent assessment in the effects on function. This approach has been applied to the paradigmatic repressor protein, LacI, and its natural operator. The functional unit of LacI is the dimer (*16*, *50*), and the hinge region (amino acids 51−60) has been deduced to mediate protein−protein contact that is essential for aligning the N-terminal DNA binding helix−turn−helix motifs (*11*, *16*, *17*, *19*). We have focused the present studies on the hinge region of LacI because this region appears to play an important role not only in DNA recognition but also in the allosteric response.

Induced folding of the hinge sequence and hinge−hinge′ contact between LacI monomers within a dimer appear to be crucial for establishing the complex with DNA (*16*−*19*, *23*). Within the hinge sequence, position 52 has no direct contact with the operator (*16*) and is one of a few positions in this region that is tolerant to some substitutions (*22*). By substituting cysteine at position 52, a disulfide bond can be introduced to "fix" the partner hinges within a dimer in close proximity (*20*). This covalent linkage between hinge domains results in an increase in affinity for the natural operator, O^1, with loss in allosteric response to inducer (*20*). The importance of fixing the relationship between the hinge sequences is reflected in the enhanced affinity observed, not only for O^1 but also for almost all variants of the operator sequence, including a nonspecific sequence. Indeed, when the binding ability of the oxidized protein is compared to either the reduced form of V52C or wild-type LacI, these differences are profound. Thus, V52C-oxidized has greatly expanded recognition for various operator sequences. Interestingly, the spacing between half-sites within the operator has minimal effect on V52C-oxidized affinity, in sharp contrast to results for V52C-reduced, wild-type LacI, and glycine mutants in the C-terminus of the hinge region (*25*). Symmetry is also not a primary factor for V52C-oxidized, since this protein can bind with similar affinity to operators of varied symmetry arrangements.

The key influence on DNA binding affinity for V52C-oxidized is sequence within the central region of the operator. The recent crystallographic structure solved for LacI·ONPF·O^{sym} provides a background for interpreting these results (*19*). In composite, the *B*-factors for the N-terminal 60 amino acids are high in relation to the complete protein; however, within the protein−DNA interface, these factors are lowest in the central region occupied by the hinge helices (*19*). The central operator region in contact with the hinge segments of LacI also exhibits the highest resolution within the protein−DNA interface (*19*). Variations in the central base sequence result in profoundly decreased binding by V52C-reduced and wild-type repressor with dramatically smaller effects on the binding of V52C-oxidized. Only in the case of expanding the spacing by addition of an A/T base pair to the central sequence is high-affinity binding by V52C-oxidized compromised. The central base sequence also determines the allosteric response of the V52C-oxidized protein, with a 10^5 difference in the affinity observed in the presence of inducer for sequences with *only* a single base pair variation (O^{disB} vs O^{disC}, Table 1). The influence of this central region of the operator is further apparent in the differences found in the binding of V52C-oxidized, wild-type LacI, and V52C-reduced to O^1, $O^{disprox}$ (which simply inverts the proximal and distal sites around the central base pair), or O^{proxA} (which makes the distal site symmetric to the proximal site of O^1). LacI appears to be able detect the "orientation" of the sequence across this central region.

Similar influence on both DNA affinity and protein function has been noted for the central, noncontacted bases within symmetric operator sequences for the phage 434 and p22 repressors (*38*−*45*). Changes in these central bases may modulate the DNA twist and bending to affect the geometry of the minor groove in this central region, and structural effects are propagated to the entire protein−DNA interface (*40*−*43*). Further, alterations in noncontacted bases affect not only the affinity but also the functional capacity of the 434 protein−DNA complex to activate transcription (*44*). Noncontacted bases in multiple systems can alter the structure of the protein−DNA complex and modulate functional capacity (*45*). This effect is observed in the present studies for both affinity and allosteric response of the V52C-oxidized protein, which depend strongly on the sequence of the central region of the operator DNA.

The association of the protein and DNA at this central region may stabilize the remaining protein−DNA interface. If the hinge helices cannot form or if the two hinges separate, reduced affinity for DNA ensues (*18*, *23*). This interpretation is supported by results reported for CytR, another member the LacI family of bacterial repressor proteins. CytR contains a helix−turn−helix motif that contacts DNA, but the interdomain linker corresponding to the hinge differs in sequence from LacI and another family member, PurR, and would not be predicted to form a helix (*51*, *52*). This region appears to be highly flexible in a manner that allows the protein to bind to operator constructs with substantial variation in half-site spacing but with substantially lower affinity compared to either LacI or PurR (*51*, *52*). This decreased affinity presumably results from the elimination of hinge helix formation and the hinge recognition elements. High-affinity binding and regulation require CytR interaction with the catabolite repressor protein (CRP); thus, interprotein interactions compensate for intraprotein contacts and are important for CytR regulation as an alternate means to enhance DNA binding affinity (*52*). Covalent linkage of the hinge regions in LacI V52C-oxidized generates high-affinity binding, an effect opposite to the deleterious effect that precluding hinge helix formation and association have on affinity in CytR. Interestingly, this covalent connection also generates relaxed sequence specificity for V52C-oxidized. The formation of this linkage may allow substantial local rearrangement of both protein and DNA elements at the minor groove interface to optimize binding to variant sequences without disrupting the spacing required for effective major groove interactions by the helix−turn−helix motifs.

Another striking feature of V52C disulfide bond formation is its effect on the LacI allosteric response. Interestingly, the ability of V52C-oxidized to respond to inducer is dependent on DNA sequence, with significant variability—from no response to a 10^4 differential in the presence vs absence of inducer. A priori, one might predict that fixing the hinge−

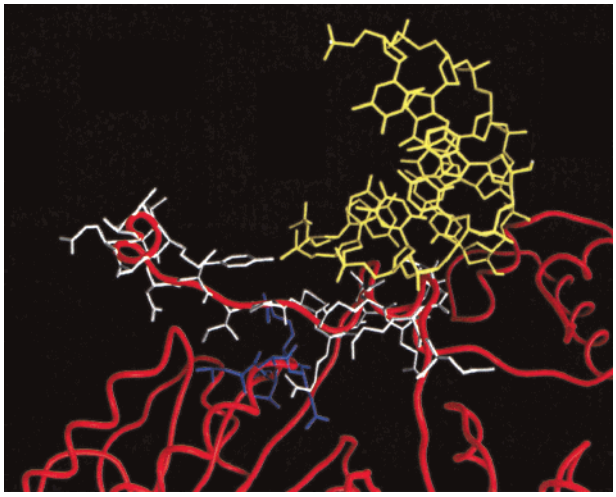

FIGURE 7: Portion of the DNA−LacI interface. The DNA backbone of the central region of O^{sym} is shown in relation to the core domain of LacI. The hinge region is in the center of the figure. DNA is shown in yellow, and the protein backbone is shown in red. Side chains from one monomer are in white. Blue side chains are from the core domain of the partner monomer and delineate the regions of the core domain that have been postulated to be important in the conformational response to inducer (*19*). Taken from pdb file 1EFA and shown using Grasp (*57*).

hinge′ interaction with the introduction of the disulfide would diminish or preclude altogether the allosteric response (*16*). Indeed, this behavior is observed for both O^1 and O^{sym} sequences. However, the data demonstrate clearly that the response of V52C-oxidized to inducer is DNA-sequence-dependent. Thus, the DNA sequence and presumably the consequent overall conformation of the protein·DNA complex profoundly influence the specific interactions and direct the response of the protein. The fixed proximity of the DNA recognition elements via the disulfide bond may allow greater structural flexibility and accommodate rearrangements in this central region related to sequence-specific properties of the base pairs (detailed discussion in *25*). This structural communication between DNA and protein is, of necessity, coupled with the intratetramer communication network to modulate the allosteric response.

Detailed inspection of the crystal structure reveals that the backbone of the bases flanking the central operator sequence comes in close proximity to an "unstructured" linker connecting the helix−turn−helix to the hinge segment. This linker has been proposed to be integral to the inducer response given the number of interactions made with the core N-subdomain of the partner subunit within a dimer (Figure 7, *19*). An interesting possibility is that variation of the DNA backbone conformation in this closely packed region may influence the orientation of the linker, thereby impacting its interface with the partner core N-subdomain and the ability of the protein to respond effectively to inducer. These N-subdomain−linker′ contacts between the dimer subunits of V52C-oxidized may be stronger for DNAs that enable effective inducer response, while these interactions may be diminished for DNAs that are insensitive to inducer.

Whatever the specific mechanism, the results for V52C-oxidized reflect the importance of hinge−hinge′ proximity in generating the conformation of the protein required for high-affinity binding. These studies also demonstrate clearly the influence of DNA sequence/conformation in contributing to the allosteric response elicited by inducer binding to the core domain. Further studies will be required to assess whether helix formation necessarily accompanies disulfide bond formation and the nature of the allosteric transition with different operator DNA sequences. Interestingly, binding parameters comparable to the intact protein have been observed for the N-terminal domain (amino acids 1−62) linked by a V52C disulfide bond (*23*). Without the core domain, the complex does not respond to inducer, but is nonetheless sensitive to reduction of the disulfide (*23*). Thus, this linkage between the partners within a dimer—in this case stabilized by a covalent linkage—is an essential element in generating high-affinity binding. Interestingly, under oxidizing conditions, hinge helix formation is not observed for the V52C(1−62) dimer alone (*23*). However, this helix forms in the complex with operators, and all DNAs examined (corresponding to the central 22 bp sequence of O^1, O^{sym}, and O^{disC}) resulted in DNA bending associated with minor groove insertion of these helices. Further, NMR spectroscopy results demonstrate that contacts differ on the two sides of the asymmetric O^1 sequence for the N-terminal V52C(1−62) disulfide-linked dimer (*23*), providing evidence for additional effects of DNA sequence on protein interaction.

The large negative ΔC_p observed for LacI binding to O^1 and O^{sym} operators has been interpreted to reflect a folding transition that occurs within the hinge region of LacI on binding to DNA (*26*, *31*). The burial of apolar surface area concomitant with complex formation has been presumed to reflect coupled folding and binding processes (*26*, *31*, *53*). However, a large negative ΔC_p has also been ascribed to the formation of a specific protein−DNA interface (*31*). The high affinity of V52C-oxidized for almost all DNAs examined, however, corresponds with a quite small negative ΔC_p for the three high-affinity operator sequences examined. Several alternative explanations emerge: The first is that the disulfide positions the N-terminal helix−turn−helix domains for high-affinity DNA binding independently of hinge helix formation. Another possibility is that helix formation is facilitated by the formation of the disulfide bond and is no longer linked with the DNA binding process. Finally, presuming that a large negative ΔC_p corresponds to a specific interface (*31*), the promiscuity in sequence selectivity by V52C-oxidized may reflect loss of specificity despite high affinity. As mentioned previously, N-terminal fragments encoding amino acids 1−62 with the V52C mutation form a disulfide-linked dimer that exhibits high affinity for operator DNAs similar to the intact protein (*23*). Hinge helix formation does not occur in the free V52C(1−62) dimer, but these helices are folded in complexes with operator DNA sequences (*23*). Whether the presence of the core domain would alter this behavior is not apparent. At least for the N-terminal fragments, hinge helix folding occurs only in the presence of operator sequences. However, in preliminary experiments with a different set of N-terminal fragments (encompassing residues 1−59) terminating with GlyCysGly sequences, under oxidizing conditions we find enhanced signal corresponding to α-helix by circular dichroism measurements (Swint-Kruse, Xu, and Matthews, unpublished data). Further experiments are required to determine whether hinge helix folding occurs in the full-length V52C-oxidized protein in the absence of DNA and to decipher the mecha-

nism by which sequence specificity is relaxed while maintaining high affinity for variant operator sequences.

In summary, introduction of a disulfide bond at the N-terminus of the hinge region of LacI results in high-affinity binding to a wide range of operator sequence variants. Response to inducer is highly dependent on the DNA sequence. The structural basis for varied inducer responsitivity may be the close apposition of the central operator sequence and residues from both the DNA binding domain and its partner core domain within the dimeric DNA binding unit. Specific orientations/contacts or sequence-dependent structural features (*54*) may be requisite for high-affinity binding and for effective allosteric communication to release the operator sequence for transcription in vivo. Differences in alignment in this closely packed region may profoundly affect the outcome of protein binding to inducer. Interestingly, introduction of the V52C disulfide bond to link the hinge helices significantly alters the thermodynamic character of the binding process, yielding an uncharacteristically small, negative ΔC_p for interaction with operator DNAs that bind with high affinity. Structural and biophysical studies of protein complexed with variant operator sequences and inducer will be required to establish firmly the basis for V52C-oxidized high affinity for multiple operator sequences, the reason for the diminished magnitude of ΔC_p for DNA binding by this protein, and the mechanism by which DNA sequence influences allosteric response in this extraordinarily interesting mutant. Studies of well-characterized regulatory proteins provide a firm foundation on which to build our understanding of DNA sequence recognition and its modulation by protein−protein and protein−ligand interactions. This information is crucial in our quest to understand the transcriptional control mechanisms central to cellular growth and differentiation.

ACKNOWLEDGMENT

We thank Drs. Tod D. Romo and Liskin Swint-Kruse for assistance with the figures derived from X-ray crytallographic coordinates. We also thank Drs. Mitchell Lewis and Charles Bell for providing the LacI dimer/O^{sym} structural data prior to publication. Discussion and suggestions from the members of the Matthews laboratory and feedback on the manuscript and assistance from Dr. Liskin Swint-Kruse and Markos Moraitis are greatly appreciated.

REFERENCES

1. Pabo, C. O., and Nekludova, L. (2000) *J. Mol. Biol. 301*, 597−624.
2. Miller, J. H., and Reznikoff, W. S., Eds. (1980) *The Operon*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
3. Matthews, K. S., and Nichols, J. C. (1998) *Prog. Nucleic Acid Res. Mol. Biol. 58*, 127−164.
4. Beyreuther, K., Adler, K., Geisler, N., and Klemm, A. (1973) *Proc. Natl. Acad. Sci. U.S.A. 70*, 3576−3580.
5. Farabaugh, P. J. (1978) *Nature 274*, 765−769.
6. Jacob, F., and Monod, J. (1961) *J. Mol. Biol. 3*, 318−356.
7. Riggs, A. D., Bourgeois, S., Newby, R. F., and Cohn, M. (1968) *J. Mol. Biol. 34*, 365−368.
8. Gilbert, W., and Müller-Hill, B. (1967) *Proc. Natl. Acad. Sci. U.S.A. 58*, 2415−2421.
9. Barkley, M. D., Riggs, A. D., Jobe, A., and Bourgeois, S. (1975) *Biochemistry 14*, 1700−1708.
10. Ogata, R. T., and Gilbert, W. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 5851−5854.
11. Kaptein, R., Zuiderweg, E. R. P., Scheek, R. M., Boelens, R., and van Gunsteren, W. F. (1985) *J. Mol. Biol. 182*, 179−182.
12. de Vlieg, J., Berendsen, H. J. C., and van Gunsteren, W. F. (1989) *Proteins: Struct., Funct., Genet. 6*, 104−127.
13. Platt, T., Files, J. G., and Weber, K. (1973) *J. Biol. Chem. 248*, 110−121.
14. Files, J. G., and Weber, K. (1976) *J. Biol. Chem. 251*, 3386−3391.
15. Müller-Hill, B. (1975) *Prog. Biophys. Mol. Biol. 30*, 227−252.
16. 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.
17. Spronk, C. A. E. M., Slijper, M., van Boom, J. H., Kaptein, R., and Boelens, R. (1996) *Nat. Struct. Biol. 3*, 916−919.
18. Spronk, C. A. E. M., Folkers, G. E., Noordman, A.-M. G. W., Wechselberger, R., van den Brink, N., Boelens, R., and Kaptein, R. (1999) *EMBO J. 18*, 6472−6480.
19. Bell, C. E., and Lewis, M. (2000) *Nat. Struct. Biol. 7*, 209−214.
20. Falcon, C. M., Swint-Kruse, L., and Matthews, K. S. (1997) *J. Biol. Chem. 272*, 26818−26821.
21. Betz, J. L. (1986) *Gene 42*, 283−292.
22. Markiewicz, P., Kleina, L. G., Cruz, C., Ehret, S., and Miller, J. H. (1994) *J. Mol. Biol. 240*, 421−433.
23. Kalodimos, C. G., Folkers, G. E., Boelens, R., and Kaptein, R. (2001) *Proc. Natl. Acad. Sci. U.S.A. 98*, 6039−6044.
24. Falcon, C. M., and Matthews, K. S. (1999) *J. Biol. Chem. 274*, 30849−30857.
25. Falcon, C. M., and Matthews, K. S. (2000) *Biochemistry 39*, 11074−11083.
26. Spolar, R. S., and Record, M. T., Jr. (1994) *Science 263*, 777−784.
27. Chen, J., and Matthews, K. S. (1992) *J. Biol. Chem. 267*, 13843−13850.
28. Wycuff, D. R., and Matthews, K. S. (2000) *Anal. Biochem. 277*, 67−73.
29. Wong, I., and Lohman, T. M. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 5428−5432.
30. Chang, W.-I., Olson, J. S., and Matthews, K. S. (1993) *J. Biol. Chem. 268*, 17613−17622.
31. Ha, J.-H., Spolar, R. S., and Record, M. T., Jr. (1989) *J. Mol. Biol. 209*, 801−816.
32. Betz, J. L. (1987) *J. Mol. Biol. 195*, 495−504.
33. 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.
34. Sadler, J. R., Sasmor, H., and Betz, J. L. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 6785−6789.
35. Simons, A., Tils, D., von Wilcken-Bergmann, B., and Müller-Hill, B. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 1624−1628.
36. Riggs, A. D., Newby, R. F., and Bourgeois, S. (1970) *J. Mol. Biol. 51*, 303−314.
37. Sasmor, H. M., and Betz, J. L. (1990) *Gene 89*, 1−6.
38. Koudelka, G. B., Harrison, S. C., and Ptashne, M. (1987) *Nature 326*, 886−888.
39. Koudelka, G. B., Harbury, P., Harrison, S. C., and Ptashne, M. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 4633−4637.
40. Wu, L., Vertino, A., and Koudelka, G. B. (1992) *J. Biol. Chem. 267*, 9134−9139.
41. Koudelka, G. B., and Carlson, P. (1992) *Nature 355*, 89−91.
42. Bell, A. C., and Koudelka, G. B. (1993) *J. Mol. Biol. 234*, 542−553.
43. Hilchey, S. P., and Koudelka, G. B. (1997) *J. Biol. Chem. 272*, 1646−1653.
44. Xu, J., and Koudelka, G. B. (1998) *J. Biol. Chem. 273*, 24165−24172.
45. Koudelka, G. B., Donner, A. L., and Ciubotaru, M. (2000) *J. Biomol. Struct. Dyn. 11*, 135−139.
46. Zhang, X., and Gottlieb, P. A. (1995) *Nucleic Acids Res. 23*, 1502−1511.

47. Mossing, M. C., and Record, M. T., Jr. (1985) *J. Mol. Biol. 186*, 295−305.
48. Pfahl, M. (1981) *J. Mol. Biol. 147*, 1−10.
49. El-Deiry, W. S., Kern, S. E., Pietenpol, J. A., Kinzler, K. W., and Vogelstein, B. (1992) *Nat. Genet. 1*, 45−49.
50. Kania, J., and Brown, D. T. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 3529−3533.
51. Jørgensen, C. I., Kallipolitis, B. H., and Valentin-Hansen, P. (1998) *Mol. Microbiol. 27*, 41−50.
52. Pedersen, H., and Valentin-Hansen, P. (1997) *EMBO J. 16*, 2108−2118.
53. Ladbury, J. E., Wright, J. G., Sturtevant, J. M., and Sigler, P. B. (1994) *J. Mol. Biol. 238*, 669−681.
54. Jen-Jacobson, L. (1997) *Biopolymers 44*, 153−180.
55. Carson, M. (1991) *J. Appl. Crystallogr. 24*, 985−961.
56. Gilbert, W., Maizels, N., and Maxam, A. (1973) *Cold Spring Harbor Symp. Quant. Biol. 38*, 845−855.
57. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins: Struct., Funct., Genet. 11*, 281−296.

BI0114067