BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

# Altering Residues N125 and D149 Impacts Sugar Effector Binding and Allosteric Parameters in *Escherichia coli* Lactose Repressor

Jia Xu,*[,†,⊥] Shirley Liu,[†] Mingzhi Chen,[‡] Jianpeng Ma,[‡,§,∥] and Kathleen S. Matthews[†]

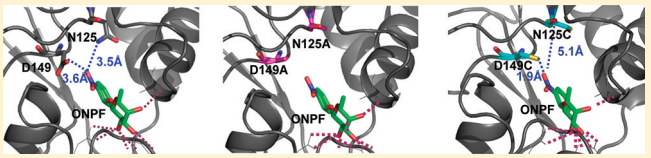
[†]Department of Biochemistry and Cell Biology, Rice University, Houston, Texas 77005, United States

[‡]Graduate Program of Structural and Computational Biology and Molecular Biophysics, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030, United States

[§]Department of Bioengineering, Rice University, Houston, Texas 77005, United States

[∥]Department of Biochemistry and Molecular Biology, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030, United States

**ABSTRACT:** Lactose repressor protein (LacI), a negative transcriptional regulator in *Escherichia coli,* relies on an allosteric conformational change for its function. The LacI effector isopropyl-*β*,D-thiogalactoside (IPTG) promotes this allosteric response and engages the side chains of residues N125 and D149 based on the crystallographic structure of LacI·IPTG. Targeted molecular dynamics (TMD) simulations have indicated involvement of these side chains during the protein structural changes in response to inducer binding. To examine this region further, we applied stochastic boundary molecular dynamics (SBMD) simulation and identified a transient interaction between residues N125 and D149. On the basis of these data, we introduced substitutions for either/both residues and analyzed their impact on protein function. The substitutions utilized were alanine to preclude hydrogen bonding or cysteine to allow disulfide bond formation, which was not observed for N125C/D149C. Minimal impacts were observed on operator affinity for all substitutions, but D149C, N125A/D149A, and N125C/D149C bound to IPTG with 5−8-fold lower affinity than wild-type LacI, and exhibited decreased allosteric amplitude ($K_{RI/O}/K_{R/O}$). Of interest, the double mutants did not exhibit an allosteric response to an alternate inducer, 2-phenylethyl-*β*,D-galactoside (PhEG), despite demonstration of PhEG binding. Further, the presence of the anti-inducer, *o*-nitrophenyl-*β*,D-fucoside (ONPF), enhanced operator affinity for wild-type LacI and all other mutant proteins examined, but behaved as an inducer for N125A/D149A, decreasing operator binding affinity. These results confirm the role of residues 125 and 149 in ligand binding and allosteric response and illustrate how readily the function of a regulatory protein can be altered.




The lactose repressor protein (LacI) is a negative transcription regulator in *E. coli* and has long served as a model system to study allosteric mechanisms.[1−4] In the presence of lactose, the metabolite allolactose is produced by *β*-galactosidase, binds to LacI, and elicits a conformational change that releases *lac* operon operator DNA, thereby allowing transcription of the structural genes required to transport and utilize lactose.[3−7] LacI can bind multiple sugars at the same effector binding site with varying outcomes on DNA-binding affinity.[5,7−10] Isopropyl-*β*,D-thiogalactoside (IPTG) (Figure 1A) is a highly effective inducer for LacI[1,5,10] and is widely used in experiments instead of the natural inducer allolactose. *o*-Nitrophenyl-*β*,D-fucoside (ONPF) serves as an anti-inducer, increasing the affinity of LacI for operator by ∼3-fold, whereas 2-phenylethyl-*β*,D-galactoside (PhEG) is an inducer with weaker affinity for LacI compared to IPTG (Figure 1A).[5,10]

The allosteric signal must be passed from the ligand binding site through the structure of the protein (∼40 Å) to reach the remote DNA binding site[3,4,11,12] (Figure 1B). Numerous residues are actively involved in transmitting the allosteric signal between these sites.[11−15] Crystallographic structures demonstrated that D149 directly contacts the inducer ligand IPTG, and targeted molecular dynamics (TMD) simulation to follow changes upon inducer binding identified residues in the allosteric pathway, including interactions between D149/S193 and D149/N125.[11−13] Mutation to allow disulfide formation between D149C and S193C demonstrated the importance of flexibility between these residues for allosteric response.[14] TMD simulation also demonstrated that N125 transiently moves closer to D149 during the conformational transition (Figure 1C). This simulated interaction indicates that N125 might function as a key residue in transmitting the signal from D149 along the allosteric pathway at the N-subdomain interface.

In high-resolution crystal structures of LacI in the presence of multiple sugar ligands, N125 and D149 form the N-terminal end of the effector binding site. In the IPTG-bound structure, direct contact is made with D149, and multiple waters and residues that include N125 are engaged in the binding site to form an intensive water-mediated hydrogen bond network (Figure 1D).[8] This network may function not only to secure

**Received:** June 9, 2011
**Revised:** September 15, 2011
**Published:** September 19, 2011

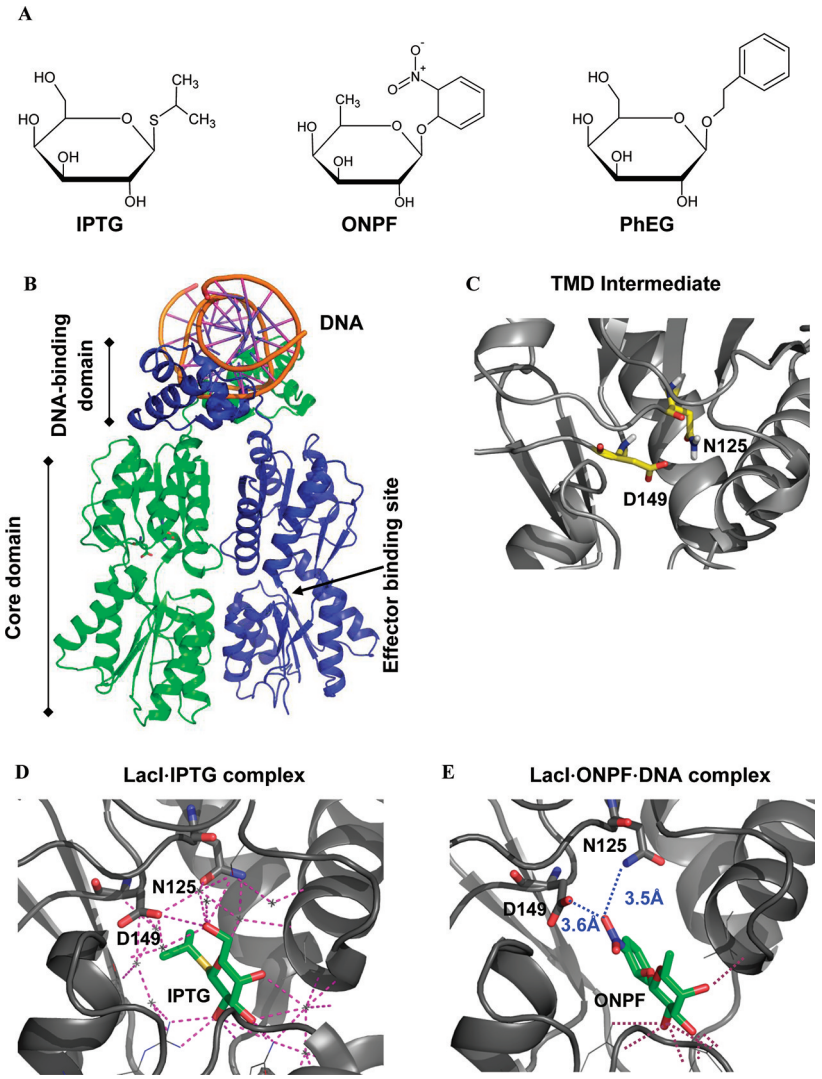



**Figure 1.** (A) Structures of IPTG, ONPF, and PhEG. (B) Structure of dimeric LacI bound to $O^{sym}$ DNA and ONPF (not shown). Structure is from PDB file 1EFA.[12] LacI is a tetramer of identical monomers, and the dimer (shown here without the C-terminal tetramerization domain) is the DNA-binding unit. The two monomers are designated with blue and green. The DNA-binding domain (amino acids ~1−50) is located at N-terminus of each dimer within a tetramer and contacts the DNA major groove; the hinge helix (amino acids ~51−60) inserts into the DNA minor groove, and the effector binding site is located between the N- and C-subdomains of the core domain (amino acids ~61−330).[12] Note that the C-terminal tetramerization domain (amino acids ~331−360) that comprises a leucine heptad repeat is not shown. (C) Close-up of effector binding site for an intermediate in the TMD simulation.[13] Note that side chains of D149 and N125 transiently move toward each other in this simulation (the closest distance is 2.6 Å). (D) Detailed view of IPTG binding region in LacI·IPTG structure (PDB file: 2P9H, ref 8). Note that both D149 and N125 are involved in forming an extensive water-mediated hydrogen-bonding network (dashed red lines) in the repressor·IPTG complex. (E) Detailed view of ONPF binding region in LacI·ONPF·DNA structure (coordinates from Dr. Mitchell Lewis, University of Pennsylvania). Note that the side chains of residues 125 and 149 are within distances (3.5 and 3.6 Å, respectively; dashed blue lines) of ONPF that could allow hydrogen bond formation with the nitro group (~3.5 Å). Hydrogen bonds with the sugar moiety are shown (dashed red lines).

IPTG binding but also to stabilize the altered orientation of the N- and C-subdomains in the induced state. The structures observed with IPTG and ONPF effectors[a] demonstrate different orientations of binding and distinct water-mediated structures in the site that may account for their differential consequences on LacI binding to operator DNA.[8] In the ONPF-bound structure (Figure 1E), the extended hydrogen bond network is not present, although direct hydrogen-bond interactions with the nitro moiety by both N125 and D149 are possible based on distance.[8] From these structures, N125 and D149 participate in binding IPTG and ONPF in distinct ways.

To further understand the role of these residues in LacI allosteric response, we have explored the interactions between N125 and D149 in the IPTG-bound form using stochastic boundary molecular dynamics (SBMD) simulation[16,17] and were able to demonstrate a transient interaction with a distance characteristic of a hydrogen bond between these two side chains. We therefore altered these residues individually[b] and simultaneously to Ala to assess the impact of removing hydrogen bond capacity and to Cys to determine whether a disulfide bond could form between them. The results demonstrate that residues 125 and 149 together are critical in transmitting functional information within LacI.

## MATERIALS AND METHODS

**Computational Simulation.** The method of stochastic boundary molecular dynamics (SBMD) simulation[16,17] was


dx.doi.org/10.1021/bi200896t | *Biochemistry* 2011, 50, 9002−9013

used to examine the IPTG binding region of wild-type LacI. A 16 Å sphere from the LacI structure (PDF file 1LBH, ref 11) was defined and centered at the midpoint between residues D149 and S193, which are critical in the IPTG-induced conformation change of the protein,[14] and empty spaces were filled with TIP3P water molecules.[18] The macromolecular modeling software CHARMM was used to perform this SBMD simulation.[19] This entire system was set to be at thermal equilibrium at 300 K. A brief energy minimization was performed before a period of dynamical equilibrium for 50 ps, and the final production run was 5 ns. A total of five simulation trajectories were conducted for the structure with different initial parameter settings to ensure no bias was produced by the values selected.

**Site-Specific Mutagenesis.** Single or double mutations in LacI were introduced on plasmid pLS1[20] with a QuickChange PCR site-directed mutagenesis kit (Stratagene) using the corresponding oligonucleotides (Invitrogen). Plasmid DNA was grown in GC5 (Sigma) competent cells and purified with a Qiagen miniprep kit. Full sequencing (SeqWright) of the gene encoding LacI in each mutant plasmid confirmed the presence of only the designed mutation(s).

**Protein Expression, Purification, and Characterization.** Proteins were expressed in *E. coli* strain BLIM cells.[21] Cells with specific plasmids were grown in 2 × YT liquid media in a shaker at 37 °C for 20−24 h. The cells were harvested by centrifugation, and the supernatant was resuspended in breaking buffer (0.2 M Tris-HCl, pH 7.6, 0.2 M KCl, 0.01 M Mg acetate, 5% glucose, 0.3 mM DTT and 0.3 μM PMSF) containing 0.5 mg/mL lysozyme. The cells were kept frozen at −20 °C. For purification, cells were thawed slowly and treated with DNaseI (80 μL of 10 mg/mL for 40 mL cell pellet) and 10 mM $MgCl_2$. The supernatant following centrifugation was slowly mixed with ammonium sulfate to final 37% saturation. The precipitant was collected by centrifugation and resuspended in 0.09 M KP buffer (0.09 M potassium phosphate, pH 7.6, 0.3 mM DTT, and 5% glucose). The resuspended cell pellet was dialyzed against the 0.09 M KP buffer overnight at 4 °C and loaded on a phosphocellulose column. Proteins were eluted with a gradient of 0.12−0.3 M KP buffer.

The fractions containing repressor were assessed by SDS-PAGE, and the purity obtained was ≥90%. Protein molecular mass was determined by gel filtration chromatography. Samples were applied on Superdex 200 column equilibrated with 0.18 M KP buffer using an AKTA fast-protein liquid chromatography system (GE Healthcare). All of the mutants display elution volumes similar to wild-type LacI, consistent with formation of tetrameric structure.

**Operator Binding.** Nitrocellulose filter binding was used to determine the DNA affinities for purified repressors.[22,23] The 40 bp wild-type operator $O^1$ sequence (5′-TGTTGTGTGG-AATTGTGAGCGGATAACAATTTCACACAGG-3′) was utilized, and the assay was conducted at room temperature in FB buffer containing 0.01 M Tris-HCl, pH 7.4, 0.15 M KCl, 0.3 mM DTT, 0.1 mM EDTA, and 5% DMSO with 100 μg/mL of bovine serum albumin. The operator $O^1$ was radiolabeled with [$^{32}$P] and purified by a Nick column (Amersham Biosciences). The DNA concentration was kept at least 10-fold lower than the $K_d$ for operator binding, and protein concentrations were varied from $1 \times 10^{-13}$ to $5 \times 10^{-9}$ M. Protein was incubated with [$^{32}$P]-labeled DNA in a 96-well plate for ∼20 min before filtering through nitrocellulose. For assays with sugars present, IPTG concentration was set to 1 mM, ONPF to 3 mM, and PhEG to 10 mM (the concentrations of ONPF and PhEG are near their saturation values). After exposure to a Fuji phosphorimaging plate overnight, the retained radioactivity was quantified by a Fuji phosphorimager. The data were analyzed with Igor Pro (Wavemetrics, CA) to determine the DNA binding affinities with the following equation:

$$Y_{obs} = Y_{max}\left(\frac{[LacI]^n}{K_d{}^n + [LacI]^n}\right) + c \quad (1)$$

where $Y_{obs}$ is the retained radioactivity at a given protein concentration, $Y_{max}$ is the level of maximum radioactivity retained, $c$ is the background radioactivity detected, $K_d$ is the equilibrium dissociation constant, and $n$ is the value of the Hill coefficient, which is generally ∼1 for wild-type LacI. A similar assay with the concentration of operator ∼100-fold higher than the $K_d$ for LacI-operator binding was utilized to measure activity levels; all mutants exhibited percent binding activity similar to the wild-type repressor.

**IPTG Binding.** Inducer binding affinity for repressors was monitored by alterations in fluorescence emission with IPTG concentration variations.[24] Fluorescence spectra were obtained by excitation at 285 nm and recording the emission from 300 to 380 nm. IPTG concentration was varied from $1 \times 10^{-8}$ to $1 \times 10^{-3}$ M in fluorescence buffer (0.01 M Tris-HCl, 0.15 M KCl, pH 7.4), and protein concentration was set at $1.5 \times 10^{-7}$ M monomer. The total fluorescence emission change was determined by summing the values at wavelengths above 340 nm. The data were analyzed with Igor Pro, and IPTG binding affinity was determined using eq 2:

$$Y_{obs} = Y_{max} - \left(Y_{max}\frac{[IPTG]^n}{K_d{}^n + [IPTG]^n}\right) + c \quad (2)$$

where $Y_{obs}$ corresponds to the measured fluorescence at a specific IPTG concentration, $Y_{max}$ is the maximum change in fluorescence signal for a given protein in response to IPTG binding, $K_d$ is the equilibrium dissociation constant, $c$ is the background fluorescence detected, and $n$ is the Hill coefficient for IPTG binding.

**Operator Release and Capture Assays.** To assess the impact of sugar binding on DNA dissociation from the repressor−DNA complex, the property defined as "operator release" was measured using a variation of the filter binding assay for operator binding. Protein concentration was set to ∼80% of DNA binding saturation for each mutant, and DNA concentration was set at least 10-fold lower than $K_d$ for operator binding (∼$1.5 \times 10^{-12}$ M). The mixture of protein and DNA was incubated in FB buffer with 100 μg/mL of bovine serum albumin for ∼30 min, and a specific sugar was added at the concentration indicated. Retained radioactivity was determined by exposure to a phosphorimaging plate and quantified using a Fuji phosphorimager. The data obtained were analyzed with Igor Pro using eq 2 with $K_d$ substituted by $[Sugar]_{mid}$. $[Sugar]_{mid}$ corresponds to the concentration of inducer required to release 50% of the protein−DNA complex (or the concentration of anti-inducer required for 50% of the total increase in binding observed). These values do not correspond to a binding constant but do provide an upper limit for $K_d$ of sugar binding and are useful for comparison purposes.

**[$^{14}$C]-IPTG Competition Assay.** A [$^{14}$C]-IPTG competition assay was applied for the double mutants to assess approximate binding affinities for ONPF and PhEG.[25] Non-radiolabeled IPTG

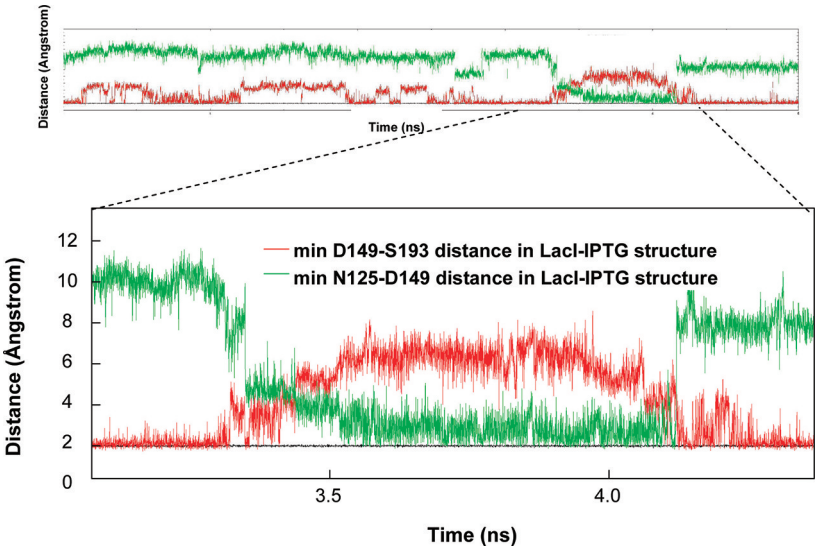



**Figure 2.** SBMD Results. Shown in this figure is the analysis of the distances between specific residues in IPTG·LacI structure during the course of a single SBMD simulation. The structure utilized for this simulation was PDB entry 1LBH.[11] The red line shows the minimal distance between the side chains of D149 and S193, whereas the green line corresponds to the minimal distance between the side chains of N125 and D149. Part of the trajectory is highlighted in the enlarged figure to show that the distance between N125 and D149 closes for a period to that consistent with forming a hydrogen bond.

was used as a control to ensure the accuracy of the assay. The assay used fluorescence buffer, and protein concentration was set to $2 \times 10^{-6}$ M monomer and to $\sim 2.5 \times 10^{-7}$ M for [$^{14}$C]-IPTG. The concentration of cold sugars, including IPTG, ONPF, or PhEG, was varied from $5 \times 10^{-8}$ to $5 \times 10^{-2}$ M where possible. The mixture was incubated for ~30 min before passing through the nitrocellulose membrane. The amount of radioactivity retained on the filter was quantified by a Fuji phosphorimager, and the data were analyzed by Igor Pro using a variation of eq 2. These experiments were also carried out in the presence of 40 bp $O^1$ operator DNA at a concentration of $4 \times 10^{-6}$ or $2 \times 10^{-5}$ M.

**Chemical Modification.** To measure the amount of free sulfhydryl in repressor, Ellman's reagent (DTNB, 5,5′-dithiobis-(2-nitrobenzoic acid)) was used.[26,27] The assay was carried out in reaction buffer containing 0.1 M sodium phosphate, pH 7.6, 1 mM EDTA. A volume of 100 $\mu$L of the denatured protein sample (with a final monomer concentration in the range ~1–10 $\mu$M in the presence of 8 M urea) was incubated with 20 $\mu$L of Ellman's reagent (4 mg of DTNB in 1 mL of reaction buffer) and 50 $\mu$L of reaction buffer at room temperature for ~15 min before measurement of absorbance at 412 nm. To measure the background, 100 $\mu$L of reaction buffer was substituted for protein solution in the mixture. The extinction coefficient ($\varepsilon$) used was 14 290 $M^{-1}$ $cm^{-1}$ in the presence of urea.

## RESULTS

**Stochastic Boundary Molecular Dynamics Simulation.** Both operator- and inducer-bound crystal structures show that side chains of N125 and D149 are well separated and not in direct hydrogen bond contact[8,12] (Figure 1D,E). These residues are found at the N-terminal end of the effector binding site and, in the presence of IPTG, interact with a complex water-mediated, hydrogen-bonding network that has been deduced to be essential to stabilize the induced conformation.[8] In TMD simulations, these side chains moved closer together transiently (Figure 1C).[13] These observations suggested the side chains of 125 and 149 might function as important points in the allosteric pathway.

To further examine 125/149 region, a molecular dynamics simulation was carried out utilizing stochastic boundary molecular dynamics (SBMD).[16,17] This method has been widely used to focus on a localized region using explicit water molecules and thereby to examine active site dynamics and interactions.[28,29] A sphere of 16 Å surrounding this region, hydrated with explicit solvent molecules, was utilized for this simulation. During SBMD, the most significant change observed in each run is that the distance between the side chains of N125 and D149 transiently closes to a distance that would allow formation of a hydrogen bond in wild-type LacI (~3.0 Å, Figure 2). By the end of each simulation, the side chain of D149 resumed its initial starting position and re-formed a hydrogen bond with S193 (a single simulation run shown in Figure 2). This result is concordant with results from TMD simulation.[13]

**Generation of Mutant Proteins.** On the basis of the results of SBMD and TMD simulations, we have explored this region experimentally by introducing alanine at positions 125 and/or 149.[b] These substitutions were designed to preclude hydrogen bond formation and inhibit interaction to determine if the transient association observed by simulation was critical to allosteric response. In addition, we substituted cysteine at these positions in an attempt to form a disulfide linkage. The mutant proteins were generated by site-specific methods, and the product proteins purified similarly to wild-type LacI, consistent with formation of tetrameric structure. Formation of this oligomer was confirmed by gel filtration chromatography.

**Disulfide Bond Formation in N125C/D149C.** The interaction between N125 and D149 illuminated by the crystal structures as well as the results of TMD and SBMD simulations suggested that these two residues might be sufficiently close to form a disulfide linkage under appropriate conditions. To determine whether such a bond could be formed between the side chains of the variant containing both N125C and D149C, this doubly substituted protein was exposed to conditions that

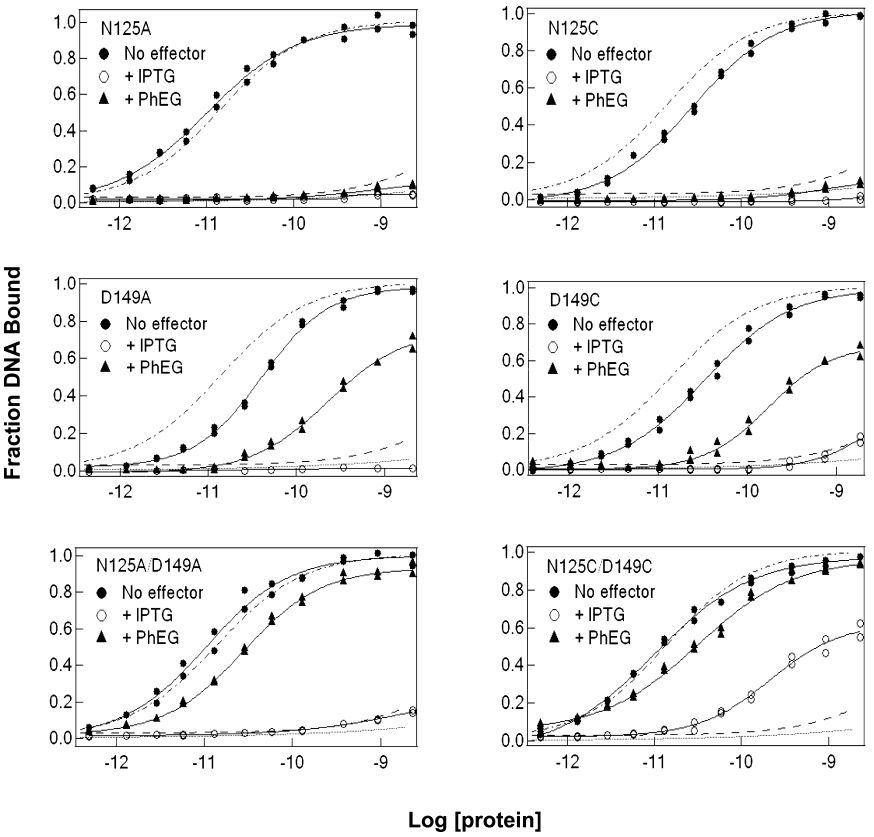



**Figure 3.** Operator binding curves and impact of IPTG and PhEG on operator binding. Experiments were conducted in FB buffer with 100 $\mu$g/mL bovine serum albumin as described in Materials and Methods. Binding curves in the absence of effector are shown as filled circles; empty circles represent binding in the presence of 1 mM IPTG and filled triangles in the presence of 3 mM PhEG. For comparison in each panel, wild-type LacI binding is represented by the dotted/dashed lines with no effector, dotted lines with IPTG present, and dashed lines for PhEG present. Results from multiple experiments are summarized in Tables 1 and 3.

promoted disulfide bond formation.[14] In monomeric wild-type LacI, there are three free cysteines located at positions 107, 140, and 281.[30] For the double mutant N125C/D149C, 4.5 ± 0.2 free cysteines were detected in the native monomeric state, suggesting that primarily the reduced form is present in solution. Following removal of DTT through a desalting column and dialysis in air-bubbled buffer to promote oxidation, 4.0 ± 0.4 free cysteines were detected, indicating only a small fraction of the protein is able to form a disulfide bond. These data suggest that these two side chains are not in ideal proximity and/or orientation to fully promote this covalent interaction.

**Operator Binding for LacI Variants.** Operator binding for the six LacI variants was measured by filter binding assay. Equilibrium dissociation constants ($K_{R/O}$) were derived from the binding curves (Figure 3; results of multiple determinations reported in Table 1). In the absence of IPTG, all of the mutants displayed DNA binding affinities comparable to wild-type LacI (<3-fold difference) (Note: D149A and D149C were examined previously with similar results.[14]) Wild-type operator DNA binding requires the formation of stable tetramer and precise arrangement of the DNA binding domains that in turn relies on features of the monomer·monomer interface (e.g., refs 11, 12, 15, and 31). Thus, the minimal difference in DNA binding properties from wild-type LacI for these proteins, even at the low concentrations required for DNA binding, confirms the tetrameric character (as demonstrated by gel filtration experiements). Further, the similarity of operator binding parameters in the absence of effector indicates that arrangement of the subunits within the variant proteins is similar to wild-type LacI.

In the presence of 1 mM IPTG, the equilibrium dissociation constants ($K_{RI/O}$) for N125A, N125C, and D149A were similar to that for wild-type LacI, whereas D149C and N125A/D149A were less sensitive to the presence of this inducer (Figure 3 and Table 1). Under same conditions, N125C/D149C retained significant affinity for operator in the presence of IPTG, with a $K_{RI/O}$ value within ~20-fold the binding constant in the absence of inducer (Figure 3 and Table 1); these differences are reflected in the allosteric amplitude (defined as $K_{RI/O}/K_{R/O}$). The latter value provides information regarding the ability of the protein to respond to IPTG binding and is diminished ~10–40-fold for D149C and N125A/D149A. The allosteric amplitude (and hence the impact of IPTG binding) is decreased by ~280-fold for N125C/D149C. This behavior is similar for protein exposed to oxidizing or reducing conditions, consistent with the minimal disulfide bond formation observed by DTNB reaction.

**IPTG Binding and Operator Release for LacI Mutants.** To explore the relationship between effector binding and allosteric response for each LacI variant, two values were measured and compared: (1) the equilibrium dissociation constant for IPTG binding ($K_{R/I}$) and (2) the amount of sugar required to make the allosteric switch, quantitated as the concentration of IPTG at which 50% of DNA is released from the repressor–operator complex; results of these "operator release" experiments are reported as $[IPTG]_{mid}$.


dx.doi.org/10.1021/bi200896t | *Biochemistry* 2011, 50, 9002–9013

**Table 1. Operator $O^1$ Binding Properties of N125/D149 Variants[a,b]**

| | $K_{R/O}$ ($M \times 10^{-11}$) | ratio $K_{R/O\ mutant}/K_{R/O\ WT}$ | $K_{RI/O}$ ($M \times 10^{-11}$) | allosteric amplitude $K_{RI/O}/K_{R/O}$ |
|---|---|---|---|---|
| WTLacI | 1.9 ± 0.2 | | >10000 | >5000 |
| N125A | 1.5 ± 0.2 | 0.8 | >10000 | >5000 |
| N125C | 2.8 ± 0.3 | 1.5 | >10000 | >4000 |
| D149A[c] | 5.0 ± 0.4 | 2.6 | >10000 | >2000 |
| D149C[c] | 4.0 ± 0.3 | 2.1 | **>500** | **>125** |
| N125A/D149A | 1.2 ± 0.1 | 0.6 | **>1000** | **>800** |
| N125C/D149C | 2.0 ± 0.2 | 1.1 | **35 ± 3.0** | **18** |

[a]Standard deviations shown represent a minimum of three measurements. Bold indicates values significantly different from wild-type LacI. [b]Operator binding experiments were performed in FB buffer containing 0.01 M Tris-HCl, pH 7.4, 0.15 M KCl, 0.3 mM DTT, 0.1 mM EDTA, 5% DMSO, and 100 μg/mL bovine serum albumin with $O^1$ concentration below $1.5 \times 10^{-12}$ M. IPTG, where present, was 1 mM. [c]Data for these mutants were also published in ref 14.

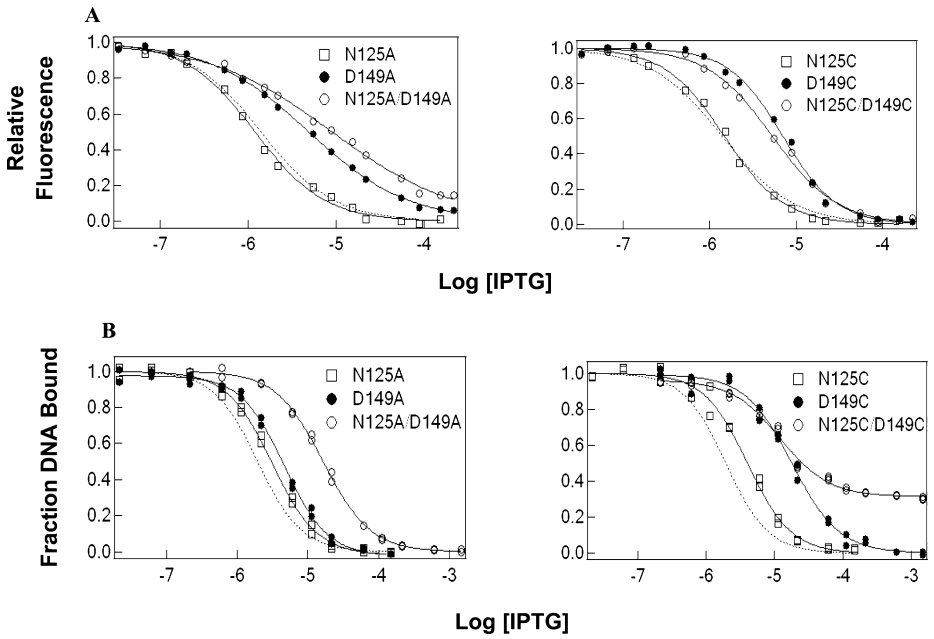



**Figure 4.** (A) Inducer binding and (B) operator release curves in response to IPTG binding for wild-type LacI and mutants. IPTG binding was conducted in fluorescence buffer and operator release assays in FB buffer with 100 μg/mL bovine serum albumin as described in Materials and Methods. Wild-type LacI behavior is represented by a dotted line in each panel for comparison. Results from multiple experiments are summarized in Table 2.

Using fluorescence titration to monitor IPTG binding, both N125A and N125C exhibit wild-type LacI affinities (Figure 4A and Table 2). D149A and N125C/D149C show ~3-fold weaker binding, whereas D149C and N125A/D149A display an even greater decrease (~5−6-fold) in binding affinities for IPTG (Figure 4A and Table 2). A consistent observation is that binding to IPTG by N125A/D149A (and to some degree D149A) displays an $n$ value, corresponding to the Hill coefficient of sugar binding, less than unity. The binding of one inducer within a dimer appears to impede sugar binding in the partner subunit for this double mutant, a situation that may also be reflected in the ~7-fold decreased allosteric amplitude ($K_{RI/O}/K_{R/O}$).

For operator release experiments, N125A, N125C, and D149A exhibit $[IPTG]_{mid}$ values within ~2-fold that for wild-type LacI (Figure 4B and Table 2). In contrast, for D149C, N125A/D149A, and N125C/D149C, the values are elevated ~6−8-fold over that for wild-type LacI (Figure 4B and Table 2), and higher IPTG concentrations are required to elicit DNA release. Complete release of DNA from N125C/D149C does not occur. The $[IPTG]_{mid}$ values comprise multiple components, including affinity for IPTG; thus, the impact on $K_{R/I}$ will be reflected in the release data. To better interpret these data, IPTG binding affinity is combined with operator release values to generate an induction ratio ($[IPTG]_{mid}/K_{R/I}$), which is useful for comparison of allosteric properties among LacI mutants (Table 2). All the mutant proteins exhibit induction ratios within 2-fold that for wild-type LacI. These results suggest that diminished response to IPTG, observed most dramatically for N125C/D149C but also for D149C and N125A/D125A, does not derive from impaired inducer binding affinity but from an impediment to assuming the fully induced conformation.

**Impact of Other Effectors on Operator Binding.** The effector ONPF is an "anti-inducer", exerting an impact on operator affinity opposite to that for IPTG and generating a small enhancement of DNA binding affinity when it binds to LacI.[5] For wild-type LacI, all of the single mutants, and the double mutant N125C/D149C, ONPF serves as an anti-inducer and elicits a small enhancement in affinity for operator compared to binding in its absence and therefore generates allosteric amplitude <1 (Table 3). In stark contrast, the operator


dx.doi.org/10.1021/bi200896t | Biochemistry 2011, 50, 9002−9013

Table 2. Inducer-Binding Properties[a]

| | $K_{R/I}$ (M × $10^6$)[b] | $[IPTG]_{mid}$ (M × $10^6$) operator release[c] | ratio[d] for $[IPTG]_{mid}$ mutant/WT | relative % operator release[e] | induction ratio[f] $[IPTG]_{mid}/K_{R/I}$ |
|---|---|---|---|---|---|
| WTLacI | 1.5 ± 0.2 | 2.2 ± 0.2 | | 100 | 1.5 ± 0.2 |
| N125A | 1.2 ± 0.2 | 2.9 ± 0.2 | 1.3 | 100 | 2.4 ± 0.2 |
| N125C | 1.4 ± 0.1 | 4.0 ± 0.3 | 1.8 | 100 | 2.9 ± 0.2 |
| D149A[g] | 4.2 ± 0.3 | 4.6 ± 0.5 | 2.1 | 100 | 1.1 ± 0.2 |
| D149C[g] | **6.0 ± 0.3** | **15 ± 0.5** | **6.8** | 100 | 2.5 ± 0.2 |
| N125A/D149A | **8.5 ± 0.2** | **18 ± 1.5** | **8.2** | 100 | 2.1 ± 0.2 |
| N125C/D149C | **5.2 ± 0.2** | **12 ± 0.4** | **5.5** | 70 | 2.3 ± 0.2 |

[a]Values shown represent a minimum of three measurements. **Bold** indicates values significantly different from wild-type LacI. [b]IPTG binding experiments were conducted in fluorescence buffer containing 0.01 M Tris-HCl, pH 7.4, 0.15 M KCl. [c]Operator release experiments were performed in FB buffer with 100 μg/mL bovine serum albumin. Protein concentration was as indicated in Materials and Methods. IPTG was present at varied concentrations. [d]This ratio is the value for $[IPTG]_{mid}$ for mutant divided by that for wild-type LacI. [e]This ratio is equal to amount of released operator over amount of total amount of operator in the complex. [f]The induction ratio is the IPTG concentration at the midpoint of operator release ($[IPTG]_{mid}$) divided by the $K_d$ for repressor–IPTG ($K_{R/I}$). [g]Data on these mutants were also published in ref 14.

binding affinity of N125A/D149A is decreased ~4-fold by ONPF, yielding behavior similar to a weak inducer with an allosteric amplitude of 4.3 (Table 3).

The effector PhEG serves as an inducer for wild-type LacI, with lower affinity than IPTG but comparable impact on operator binding (Figure 3).[5] Similar behavior to wild-type LacI is observed for both N125A and N125C single mutants, whereas PhEG acts as a weak inducer for D149A and D149C with only 4–5-fold decrease in DNA binding affinity observed (Table 3 and Figure 3). For both double mutants, N125A/D149A and N125C/D149C, PhEG has minimal impact on operator binding affinity. These differential effects are evident in the allosteric amplitudes for these effectors (Table 3).

**Impact of Varying ONPF or PhEG Concentrations on Operator DNA Binding.** The effects of varying the concentration ONPF or PhEG on the protein·operator complex were examined. All mutant proteins, with the exception of N125A/D149A, respond to ONPF in a manner similar to wild-type LacI with an increase in binding (Table 4). Unusual behavior is observed for N125A/D149A, with ONPF inhibiting the protein–DNA interaction (i.e., this anti-inducer acts as an inducer). The midpoint for ONPF effects on DNA binding is diminished by ~2-fold for N125C, D149A, D149C, and N125C/D149C. In response to PhEG, D149A and D149C release operator from the protein–operator complex, although the extent is lower than wild-type LacI at comparable concentrations and requires higher sugar concentrations. In contrast, N125A and N125C behave more similarly to wild-type LacI and release DNA at slightly lower sugar concentrations. For the double mutants N125A/D149A and N125C/D149C, significantly higher concentrations of PhEG than that observed for wild-type LacI are required to elicit a very low level of operator release (see Figure 3).

**ONPF and PhEG Competition Assays with IPTG.** The distinct responses of the double mutants to effector molecules raise the question of whether these sugar derivatives are binding to these LacI variants. To address this concern, competition of these effector molecules with protein-bound [$^{14}$C]-IPTG was conducted for the mutant proteins. Control experiments with unlabeled IPTG demonstrated the effectiveness of this assay. The results demonstrate that ONPF and PhEG can compete effectively with IPTG for binding to each of the double mutants in the concentration range examined in the operator release assay (Figure 5 and Table 5). Interestingly, the aromatic ONPF and PhEG sugars bind >4-fold more effectively to N125A/D149A compared to the corresponding double mutant with more hydrophilic cysteines present, suggesting potential formation of apolar interactions between these protein side chains and the aromatic effectors.

Because the presence of operator will alter binding of effector molecules, we pursued similar measurements in the presence of saturating operator DNA to determine (or estimate if necessary given solubility issues) a "lower bound" for the binding affinity under these conditions. The experiments are complicated by the high levels of DNA required that result in large experimental variations; nonetheless, the results obtained make clear that the effector ligands bind the proteins effectively in the presence of operator DNA (Figure 5 and Table 5). The data for these proteins indicate that ONPF binds to ~85% in the presence of saturating operator (Figure 5 and Table 5). For PhEG, competition with [$^{14}$C]-IPTG is significantly less effective for wild-type LacI in the presence of operator compared to the double mutants (Figure 5 and Table 5). These results for both ONPF and PhEG provide assurance that the altered impacts of effectors observed with these double mutants must derive primarily from changes in the character of the allosteric response rather than impaired sugar binding capacity.

## DISCUSSION

Structural studies of allosteric proteins have provided significant fodder for understanding the mechanisms by which ligand binding alters function. LacI has been a prototype for such studies of genetic regulatory proteins (e.g., refs 32 and 33): Computational approaches using structures in operator- and IPTG-bound states[13] have identified potential participants in allosteric pathways, some of which have been examined explicitly to confirm their role in this key process (e.g., refs 14 and 15). A recent crystallographic examination of LacI complexed with various effector ligands illuminated the requirements for binding of a sugar and indicated roles for N125 and D149.[8] These residues participate in an extended network of hydrogen bonds (including water-mediated interactions) in the effector site that is only observed in the structure of the IPTG-bound state.[8] This network has been deduced to stabilize the induced conformation by effectively cross-linking the N- and C-subdomains.[8]

The importance of D149 has been explored previously:[8,11−14] From phenotypic studies, almost all amino acid substitutions at position 149 diminish response to IPTG ($i^S$ phenotype),[34] a circumstance that reflects either compromised inducer affinity or impaired allosteric response. For seven

Table 3. Operator Binding Parameters in the Presence of ONPF and PhEG[a]

| | operator affinity[b] | | | | allosteric amplitude | | | |
|---|---|---|---|---|---|---|---|---|
| | no ligand $K_{R/O}$ ($M \times 10^{11}$) | +ONPF[c] $K_{RONPF/O}$ ($M \times 10^{11}$) | +PhEG[d] $K_{RPhEG/O}$ ($M \times 10^{11}$) | $K_{RONPF/O}/K_{R/O}$ ONPF | $K_{RPhEG/O}/K_{R/O}$ PhEG | for comparison, $K_{RIPTG/O}/K_{R/O}$ IPTG |
| WTLacI | 1.9 ± 0.2 | 0.9 ± 0.1 | >1000 | 0.5 | >500 | >5000 |
| N125A | 1.5 ± 0.2 | 1.0 ± 0.1 | >1000 | 0.7 | >600 | >5000 |
| N125C | 2.8 ± 0.3 | 1.5 ± 0.1 | >1000 | 0.5 | >350 | >4000 |
| D149A | 5.0 ± 0.4[e] | 2.5 ± 0.3 | **23 ± 3.0** | 0.5 | **4.6** | >2000 |
| D149C | 4.0 ± 0.3[e] | 1.6 ± 0.2 | **15 ± 1.6** | 0.4 | **3.8** | **>125** |
| N125A/D149A | 1.2 ± 0.1 | **5.2 ± 0.3** | **3.5 ± 0.2** | **4.3** | **2.9** | >800 |
| N125C/D149C | 2.0 ± 0.2 | 1.1 ± 0.1 | **3.0 ± 0.2** | 0.6 | **1.5** | **18** |

[a]Standard deviations shown represent a minimum of three measurements. Bold indicates values significantly different from those for wild-type LacI. [b]Operator binding experiments were performed in FB buffer with 100 μg/mL bovine serum albumin. Operator concentration was below $1.5 \times 10^{-12}$ M. [c]Operator binding affinities in the presence of 10 mM ONPF. [d]Operator binding affinities in the presence of 3 mM PhEG. [e]Data for these mutants were also published in ref 14.

purified variants at position 149, inducer binding affinity is decreased in most, but only for D149C are both decreased IPTG affinity and diminished allosteric response observed.[14] Further, forming a disulfide bond between D149C and S193C across the interface of N- and C-subdomains significantly impairs response to inducer IPTG (~15-fold increased induction ratio), presumably by impeding movement of D149 necessary for transmission of the structural changes required to release operator DNA and form the stabilized inducer-bound structure.[14]

This key role identified for D149 in the allosteric process was underscored in the present studies by SBMD simulation that showed that the side chains of this residue and N125 form a transient hydrogen bond, consistent with previous TMD simulation results.[13] Interestingly, almost all N125 substitutions also exhibit phenotypes ($I^s$ or $I^{ws}$) that indicate loss of inducer affinity and/or allosteric impediment.[34] We therefore utilized single and double substitutions at these two sites (a) using alanine to preclude hydrogen bonding or (b) using cysteine to promote disulfide bond formation if these residues come into appropriate proximity and orientation. In addition, two other effectors, the weak inducer PhEG and the anti-inducer ONPF, were used to examine impact on LacI allosteric communication in the variant proteins.

Operator binding affinity is surprisingly robust to these substitutions, with <3-fold differences among all of the protein variants examined. However, D149C, as might be anticipated from prior studies,[14] and both double mutants exhibit compromised IPTG binding coupled with decreased allosteric response to IPTG, as reflected in their diminished allosteric amplitudes ($K_{RI/O}/K_{R/O}$). This impact is most dramatic for N125C/D149C, an effect that cannot be ascribed to the molar volume of the cysteine side chains since this property is similar to that for the parent asparagine and aspartic side chains.[35] Further, helical propensity also does not account for the distinction, as the differences for this parameter among these three amino acids are small,[36,37] and neither N125 nor D149 is in a helical segment of the protein. However, differences in hydropathy may contribute to the results observed, given the significant differences for this parameter between the parent Asn/Asp and the Ala/Cys substitutions,[38] particularly the latter. The alterations in hydropathy that accompany substitution may preclude (i) the direct interaction observed between N125 and D149 by molecular dynamics simulations and (ii) engagement in interactions that appear crucial to stabilizing the induced form of the sugar-bound protein. The dual alanine substitutions, with smaller volume, may allow conformations unavailable to the dual cysteine substitutions.

To further explore the impact of alterations in the region identified as crucial for stabilizing the IPTG-induced conformation, we examined two other effectors of wild-type LacI: the weak inducer PhEG and the anti-inducer ONPF. Response to PhEG is significantly compromised for all variants in which D149 is altered, consistent with observations for IPTG and the apparent role of this residue in transmitting the allosteric signal for this effector.[13] Indeed, behavior with PhEG serves a harbinger of the altered induction properties observed with IPTG. As apparent for wild-type LacI response to a range of sugars,[5] occupancy of the sugar site alone is insufficient to promote the formation of networks able to transfer allosteric information to the DNA binding site. The limited change for DNA binding affinity in response to PhEG in the subset of mutants with substitution at residue 149 recapitulates lack of

**Table 4. ONPF and PhEG Impact on Operator Binding**[a,b]

| | $[ONPF]_{mid}$ ($M \times 10^6$) | relative % increase[c] | $[PhEG]_{mid}$ ($M \times 10^6$) | relative % release |
|---|---|---|---|---|
| WTLacI | 110 ± 11 | 10–20 | 500 ± 35 | 100 |
| N125A | 200 ± 28 | 10–15 | 390 ± 15 | 100 |
| N125C | 40 ± 3.8 | 10–15 | 180 ± 12 | 100 |
| D149A | 45 ± 2.9 | 20–30 | >1000 | 60 |
| D149C | 30 ± 2.0 | 10–15 | >1000 | ~60 |
| N125A/D149A | >100 | (−) **30–40** | >1000 | **<30** |
| N125C/D149C | 55 ± 1.8 | 20–30 | >1000 | **<15** |

[a]Standard deviations shown represent a minimum of three measurements. Bold indicates values significantly different from wild-type LacI. [b]Operator release experiments were performed in FB buffer with 100 μg/mL bovine serum albumin. Protein concentration was at 80% saturation of operator binding. Operator concentration was below $1.5 \times 10^{-12}$ M. ONPF and PhEG were varied from $10^{-6}$ to $10^{-2}$ M. [c](−) indicates relative % decrease.

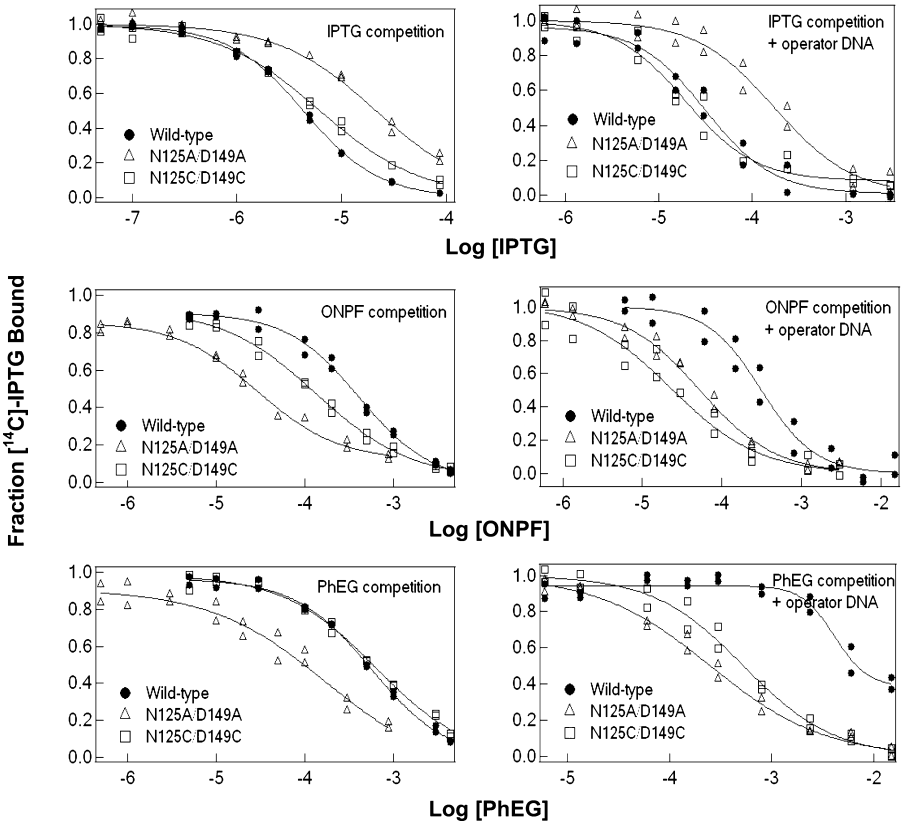



**Figure 5.** Sugar competition with [$^{14}$C]-IPTG binding to LacI. The assay was conducted in fluorescence buffer as described in Materials and Methods. The left panels show binding in the absence of operator DNA, and the right panels show binding in the presence of 40 bp operator DNA. Protein concentration was $2 \times 10^{-6}$ or $1 \times 10^{-5}$ M monomer with 2-fold excess of operator DNA, and [$^{14}$C]-IPTG was present at ~$2.5 \times 10^{-7}$ M. Results from multiple experiments are summarized in Table 5.

effect on DNA binding when the "neutral" sugar ONPG interacts with wild-type LacI.[5] The small chemical difference between ONPG and ONPF [the C6-hydroxyl on galactose is missing in ONPF] appears to preclude the interactions with side chains in the effector-binding site that stabilize distinct conformations and thereby alter DNA binding affinity. Similarly, introduction of changes in the side chain at position 149 preclude forming the interactions with PhEG necessary to stabilize the induced state.

What determines these distinctions in sugar impact is not evident, even from the structures available.[8,11,12] Although some commonality in behavior is found for the inducers IPTG and PhEG with the mutants examined, the response to ONPF differs. This fucoside acts as an anti-inducer to stabilize operator binding for wild-type LacI and all mutants except for N125A/D149A. For the latter protein, diminished hydrogen-bonding capacity, smaller side chain volume, and enhanced hydropathy at both positions 125 and 149 simultaneously improve binding affinity for ONPF and convert its allosteric effect to an inducer.

The observation of the extended hydrogen bond network only in IPTG-bound LacI has been interpreted as evidence for its importance in stabilizing the induced conformation by effectively cross-linking the N- and C-subdomains.[8] However, the double mutant N125A/D149A lacks the ability to promote this water-mediated hydrogen bond network with IPTG but still maintains a significant allosteric response to IPTG. These data indicate that formation of an apolar network in N125A/D149A might serve to stabilize the induced conformation of LacI. This apolar network may be less effective in the ONPF-bound


dx.doi.org/10.1021/bi200896t | *Biochemistry* 2011, 50, 9002–9013

Table 5. Competition of sugars with [$^{14}$C]IPTG[a]

| | $[IPTG]_{mid}$ ($\times 10^{-6}$ M) | relative % decrease | $[ONPF]_{mid}$ ($\times 10^{-5}$ M) | relative % decrease | $[PhEG]_{mid}$ ($\times 10^{-4}$ M) | relative % decrease |
|---|---|---|---|---|---|---|
| WTLacI | 4.4 | 100 | 41 | 90 | 5.6 | 90 |
| WTLacI + $O^1$ | 23 | 100 | 32 | 100 | 50 | 60 |
| N125A/D149A | 17 | 75 | 2.9 | 85 | 1.7 | 80 |
| N125A/D149A + $O^1$ | 180 | 90 | 7.2 | 95 | 2.8 | 90 |
| N125C/D149C | 6.2 | 85 | 13 | 80 | 6.2 | 85 |
| N125C/D149C + $O^1$ | 51 | 90 | 2.0 | 90 | 13 | 90 |

[a] [$^{14}$C]IPTG competition assays were conducted in fluorescence buffer. Protein concentration was $2 \times 10^{-6}$ M monomer, and [$^{14}$C]-IPTG was $\leq 2.5 \times 10^{-7}$ M. The concentration of the indicated competing effector sugar was varied from $5 \times 10^{-8}$ to $1 \times 10^{-2}$ M. Where indicated, $O^1$ operator DNA was present at $4 \times 10^{-6}$ or $2 \times 10^{-5}$ M. Values shown are averages from three experiments; however, the error ranges were broad, with midpoint values varying in some cases up to 3−5-fold in the presence of DNA. Single determinations are shown in Figure 5.

form of this protein but nonetheless partially stabilizes the induced conformation. In a similar vein, the alterations in N125C/D149C may impede formation of either type of network and thereby significantly impair response to IPTG and PhEG, whereas ONPF may be still able to stabilize the DNA-binding conformation. Thus, the specific impact of a sugar ligand bound to LacI derives from its ability to stabilize specific states of the protein, with a variety of consequent effects on the operator binding properties of the protein, from decreased to enhanced affinity.

The effect of ONPF on N125A/D149A function reflects an inversion from a "repressible" system to an "inducible" system. Recent *in vivo* and *in vitro* studies of LacI mutants demonstrated even a single change in amino acid sequence could result in an inverted response pattern from inducible to corepressible.[39−41] Elucidating the molecular mechanism by which such changes alter the conformational states within this delicate genetic switch is crucial to advance our understanding the allosteric communication. LacI has the ability to bind different effectors within the same region, but the product of this binding process results in distinct function, reflecting multiple allosteric modes. All the effectors (IPTG, ONPF, and PhEG) examined in this study share a common core structure: O2, O3, and O4 hydroxyls are responsible for binding to repressor, whereas C1 and C6 are responsible for generating specificity in sugar−repressor interaction.[8] Based on our results, when ONPF occupies the inducer binding pocket in N125A/D149A, the allosteric signal transduction pathway diverges from that for wild-type LacI.

To better understand the source of the observed differences, the double mutants were modeled in the presence of IPTG and ONPF, for which wild-type structures are available,[c] by using the PyMOL structure visualization and modification program. Desired substitutions can be introduced into the corresponding positions in available crystal structures and the conditions optimized to generate likely conformations. For the IPTG-bound N125A/D149A, modeling indicates that the extensive water-mediated hydrogen bond network characteristic of wild-type LacI (see Figure 1D) is disrupted in the N-subdomain segment of the binding pocket without the key polar residues at positions 125 and 149 (Figure 6A). Nonetheless, this LacI variant responds to IPTG, albeit less strongly than wild-type LacI. For the IPTG-bound N125C/D149C structure, the hydrogen bond network at the N-subdomain end of the site is less robust due to rotation of the side chain at both residues 125 and 149 away from the effector site (>5 Å, see Figure 6B), an effect that may account for the compromised induction response.

The intensive hydrogen-bond network observed for IPTG-bound LacI appears to stabilize the induced conformation.[8] However, as indicated above, alternate stabilization schemes through hydrophobic interactions appear feasible to support the induction process for N125A/D149A that are not feasible for D149C or, even more dramatically, for N125C/D149C. For PhEG, for which no structural information is available, without the polar residues at positions 125 and 149, other side chains (e.g., L148 that lies between N125 and D149 at the N-terminal end of the inducer binding pocket) may interact less effectively to stabilize the subdomain interface in the DNA-bound form, accounting for the compromised induction response to PhEG in the double mutants.

In the wild-type LacI structure, the polar side chains at N125 and D149 are in hydrogen bond distance from the nitro substituent of the phenyl ring of ONPF in the presence of DNA (Figure 1E). Based on the binding data for the variants at a single position, only one of these polar side chains appears necessary for stabilizing the DNA-bound form of the protein. For ONPF-bound N125C/D149C (Figure 6D), where a modestly polar side chain is present at both positions 125 and 149, predicted interaction with ONPF appears to stabilize the operator-binding conformation as observed for wild-type LacI (Figure 1E), resulting in operator binding affinities comparable to wild-type LacI (Table 3). Apolar substitution at both positions, however, results in conversion of ONPF from an anti-inducer to a weak inducer. The modeled structure for ONPF-bound N125A/D149A shows movement of the side chains away from the effector (Figure 6C). Although the extensive water-mediated network characteristic of the inducer-bound state is not present, the presence of ONPF may result in an arrangement of alternate stabilizing interactions sufficient to generate the limited decrease in operator affinity (Table 3).

Interestingly, LacI can be converted to a "repressible switch" in the LacI variant L148W (adjacent to D149), an arrangement where repression in the absence of effector is diminished and ONPF results in enhanced repression.[40] This result presumably derives from distortion in the effector binding site that is ameliorated by the presence of ONPF to generate the DNA-binding conformation. More recently, *in vivo* screens were utilized to identify S97P (in the monomer·monomer interface) as a LacI variant that yields an "inverse" response confirmed by *in vitro* measurements; this protein exhibits diminished operator binding in the unliganded state, enhanced affinity for operator in the presence of IPTG, and reduced affinity with ONPF.[41]


dx.doi.org/10.1021/bi200896t | *Biochemistry* 2011, 50, 9002−9013

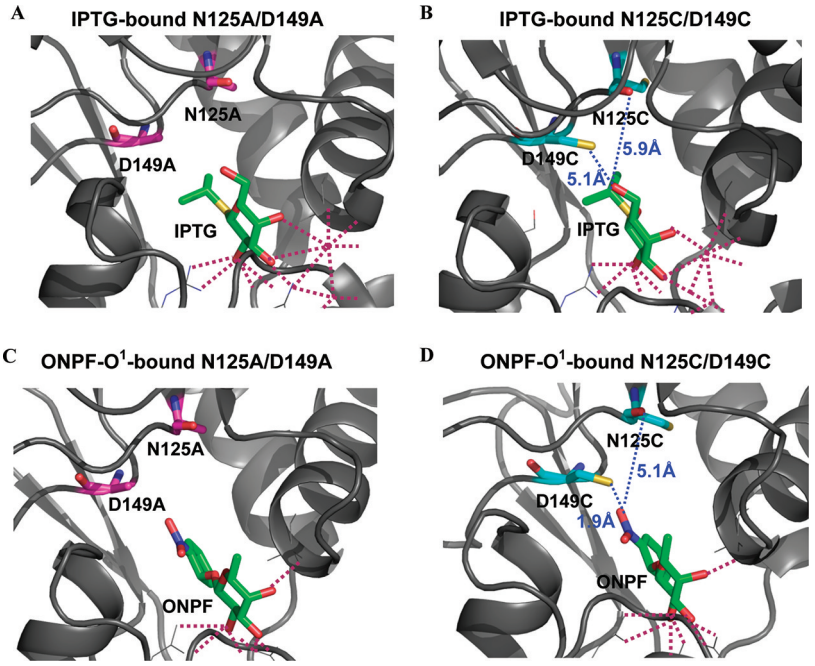



**Figure 6.** PyMOL-generated variant structures: (A) N125A/D149A and (B) N125C/D149C in the IPTG-bound state (PDB entry 2P9H, ref 8); (C) N125A/D149A and (D) N125C/D149C in ONPF-DNA-bound state (coordinates obtained from Dr. Mitchell Lewis, University of Pennsylvania). The substituted side chains are positioned at the optimal orientation after energy minimization. For the N125C/D149C structures, the sulfhydryl group is consistently rotated away from the plane of the figure and therefore difficult to discern. Note that the extensive water-mediated hydrogen-bonding network in the IPTG-bound state is not present for N125A/D149A (panel A), perhaps allowing reorientation of ONPF not detected by PyMOL analysis (panel C) for N125A/D149A. The minimal distances between D149C/N125C and IPTG/ONPF are presented in dashed blue lines (panels B and D). Note that 1.9 Å is measured between D149C-SH and ONPF-$O^6$ in panel D, suggesting potential for hydrogen-bond formation.

## CONCLUSION

The LacI response to effectors is initiated in the central ligand binding site, which accommodates many different sugars. However, only a subset of these sugars binds with high affinity and a smaller subset generates the allosteric response of decreased operator binding. All mutations at D149 impede the allosteric response to PhEG, despite direct evidence of binding, illustrating that small changes can convert an inducer to a neutral effector. Converting N125 and D149 to apolar residues removes the polar side chains that interact with LacI effector molecules[8] and with each other during the intermediate stages of the allosteric process[13] and diminishes response to even the strong inducer IPTG. The N125A/D149A double mutation not only impacts the extent of allosteric release but also converts ONPF from a weak anti-inducer to a weak inducer. This sensitivity underscores the malleability of protein function in response to minimal sequence alteration.[39−41] Not only can both random and designed screens generate variant regulatory properties, but amino acid changes that occur naturally may be selected in particular environments in an evolutionary process to provide differential gene regulatory response across a range of effectors that bind to a protein. The alterations presented here and in recent reports on other LacI variants with altered allosteric properties to sugars[40,41] demonstrate that strategies can be readily constructed to generate and selected for altered effector specificity or for conversion of a repressible to an inducible genetic regulatory protein (or *vice versa*). These data suggest mechanisms in this family of genetic regulatory proteins that can be utilized to evolve new functions, both in nature and by design.

## AUTHOR INFORMATION

### Corresponding Author

*Tel: 713-348-4936. E-mail: jiaxu@rice.edu. Fax: 713-522-6782.

### Present Address

[⊥]Department of Systems Medicine and Bioengineering, The Methodist Hospital Research Institute, Weill Cornell Medical College, Houston, TX 77030.

### Funding

This work was supported by grants from the Robert A. Welch Foundation to K.S.M. (C-576) and to J.M. (Q-1512) and from NIH (GM 22441) to K.S.M.

## ACKNOWLEDGMENTS

The authors thank Dr. Liskin Swint-Kruse (Kansas University Medical Center) for stimulating discussions and insightful feedback on this manuscript and Dr. Mitchell Lewis (University of Pennsylvania) for the coordinates for ONPF·wild-type LacI and ONPF·operator·wild-type LacI.

## ABBREVIATIONS

DTNB, 5,5′-dithiobis(2-nitrobenzoic acid); DTT, dithiothreitol; IPTG, isopropyl-*β*,D-thiogalactoside; LacI, lactose repressor protein; ONPF, *o*-nitrophenyl-*β*,D-fucoside; ONPG, *o*-nitrophenyl-*β*,D-galactoside; PhEG, 2-phenylethyl-*β*,D-galactoside; SBMD, stochastic boundary molecular dynamics; TMD, targeted molecular dynamics; WT, wild-type.

## ADDITIONAL NOTES

[a]No crystallographic structure for LacI complexed with PhEG is available.

[b]D149A and D149C were originally produced and characterized partially in ref 14.
[c]The results observed for PhEG cannot be visualized, since a crystallographic structure for this inducer is not available.

## REFERENCES


(1) Jacob, F., and Monod, J. (1961) Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol. 3*, 318−356.
(2) Monod, J., Wyman, J., and Changeux, J. P. (1965) On the nature of allosteric transitions: A plausible model. *J. Mol. Biol. 12*, 88−118.
(3) Matthews, K. S., and Nichols, J. C. (1998) Lactose repressor protein: Functional properties and structure. *Prog. Nucleic Acid Res. Mol. Biol. 58*, 127−164.
(4) Müller-Hill, B. (1996) The *lac* operon: A short history of a genetic paradigm, Walter de Gruyter, Berlin, NY.
(5) Barkley, M. D., Riggs, A. D., Jobe, A., and Bourgeois, S. (1975) Interaction of effecting ligands with *Lac* repressor and repressor-operator complex. *Biochemistry 14*, 1700−1712.
(6) Riggs, A. D., Bourgeois, S., Newby, R. F., and Cohn, M. (1968) DNA binding of the *lac* repressor. *J. Mol. Biol. 34*, 365−368.
(7) Jobe, A., and Bourgeois, S. (1972) *lac* repressor-operator interaction. VI. The natural inducer of the *lac* operon. *J. Mol. Biol. 69*, 397−408.
(8) Daber, R., Stayrook, S., Rosenberg, A., and Lewis, M. (2007) Structural analysis of Lac repressor bound to allosteric effectors. *J. Mol. Biol. 370*, 609−619.
(9) Jobe, A., and Bourgeois, S. (1973) Lac repressor-operator interaction.VIII. Lactose is an anti-inducer of the *lac* operon. *J. Mol. Biol. 75*, 303−313.
(10) Riggs, A. D., Newby, R. F., and Bourgeois, S. (1970) *lac* repressor-operator interaction. II. Effect of galactosides and other ligands. *J. Mol. Biol. 51*, 303−314.
(11) Lewis, M., Chang, G., Horton, N. C., Kercher, M. A., Pace, H. C., Schumacher, M. A., Brennan, R. G., and Lu, P. (1996) Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science 271*, 1247−1254.
(12) Bell, C. E., and Lewis, M. (2000) A closer view of the conformation of the Lac repressor bound to operator. *Nat. Struct. Biol. 7*, 209−214.
(13) Flynn, T. C., Swint-Kruse, L., Kong, Y., Booth, C., Matthews, K. S., and Ma, J. (2003) Allosteric transition pathways in the lactose repressor protein core domains: Asymmetric motions in a homodimer. *Protein Sci. 12*, 2523−2541.
(14) Xu, J., and Matthews, K. S. (2009) Flexibility in the inducer binding region is crucial for allostery in the *Escherichia coli* lactose repressor. *Biochemistry 48*, 4988−4998.
(15) Zhan, H., Camargo, M., and Matthews, K. S. (2010) Positions 94−98 of the lactose repressor N-subdomain monomer-monomer interface are critical for allosteric communication. *Biochemistry 49*, 8636−8645.
(16) Brooks, C. L. 3rd, and Karplus, M. (1989) Solvent effects on protein motion and protein effects on solvent motion. Dynamics of the active site region of lysozyme. *J. Mol. Biol. 208*, 159−181.
(17) Kong, Y., and Ma, J. (2001) Dynamic mechanisms of the membrane water channel aquaporin-1 (AQP1). *Proc. Natl. Acad. Sci. U. S. A. 98*, 14345−14349.
(18) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys. 79*, 926−935.
(19) Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem. 4*, 187−217.
(20) Swint-Kruse, L., Elam, C. R., Lin, J. W., Wycuff, D. R., and Matthews, K. S. (2001) Plasticity of quaternary structures: Twenty-two ways to form a LacI dimer. *Protein Sci. 10*, 262−276.
(21) Wycuff, D. R., and Matthews, K. S. (2000) Generation of an AraC-*ara*BAD promoter-regulated T7 expression system. *Anal. Biochem. 277*, 67−73.
(22) O'Gorman, R. B., Rosenberg, J. M., Kallai, O. B., Dickerson, R. E., Itakura, K., Riggs, A. D., and Matthews, K. S. (1980) Equilibrium binding of inducer to *lac* repressor·operator DNA complex. *J. Biol. Chem. 255*, 10107−10114.
(23) Wong, I., and Lohman, T. M. (1993) A double-filter method for nitrocellulose-filter binding: Application to protein-nucleic acid interaction. *Proc. Natl. Acad. Sci. U. S. A. 90*, 5428−5432.
(24) Laiken, S. L., Gross, C. A., and von Hippel, P. H. (1972) Equilibrium and kinetic studies of *Escherichia coli lac* repressor-inducer interactions. *J. Mol. Biol. 66*, 143−155.
(25) Gardner, J. A., and Matthews, K. S. (1990) Characterization of two mutant lactose repressor proteins containing single tryptophans. *J. Biol. Chem. 265*, 21061−21067.
(26) Ellman, G. L. (1959) Tissue sulfhydryl groups. *Arch. Biochem. Biophys. 82*, 70−77.
(27) Riddles, P. W., Blakeley, R. L., and Zerner, B. (1983) Reassessment of Ellman's reagent. *Methods Enzymol. 91*, 49−60.
(28) Brünger, A. T., Brooks, C. L. III, and Karplus, M. (1985) Active site dynamics of ribonuclease. *Proc. Natl. Acad. Sci. U. S. A. 82*, 8458−8462.
(29) Cox, K., Watson, T., Soultanas, P., and Hirst, J. D. (2003) Molecular dynamics simulations of a helicase. *Proteins 52*, 254−262.
(30) Yang, D. S., and Matthews, K. S. (1976) Lactose repressor protein reaction with 2-chloromercuri-4-nitrophenol. *J. Mol. Biol. 103*, 433−437.
(31) Chen, J., and Matthews, K. S. (1994) Subunit dissociation affects DNA binding in a dimeric *Lac* repressor produced by C-terminal deletion. *Biochemistry 33*, 8728−8735.
(32) Wilson, C. J., Zhan, H., Swint-Kruse, L., and Matthews, K. S. (2006) The lactose repressor system: Paradigms for regulation, allosteric behavior and protein folding. *Cell. Mol. Life Sci. 2007*, 3−16.
(33) Swint-Kruse, L., and Matthews, K. S. (2004) Thermodynamics, protein modification, and molecular dynamics in characterizing lactose repressor protein: Strategies for complex analysis of protein structure-function. *Methods Enzymol. 379*, 188−209.
(34) Suckow, J., Markiewicz, P., Kleina, L. G., Miller, J., Kisters-Woike, B., and Müller-Hill, B. (1996) Genetic studies of the Lac repressor XV: 4000 single amino acid substitutions and analysis of the resulting phenotypes on the basis of the protein structure. *J. Mol. Biol. 261*, 509−523.
(35) Zamyatnin, A. A. (1984) Amino acid, peptide, and protein volume in solution. *Annu. Rev. Biophys. Bioeng. 13*, 145−165.
(36) Koehl, P., and Levitt, M. (1999) Structure-based conformational preferences of amino acids. *Proc. Natl. Acad. Sci. U. S. A. 96*, 12524−12529.
(37) Luque, I., Mayorga, O. L., and Freire, E. (1996) Structure-based thermodynamic scale of $\alpha$-helix propensities in amino acids. *Biochemistry 35*, 13681−13688.
(38) Kyte, J., and Doolittle, R. F. (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol. 157*, 105−132.
(39) Tungtur, S., Meinhardt, S., and Swint-Kruse, L. (2010) Comparing the functional roles of nonconserved sequence positions in homologous transcription repressors: Implications for sequence/function analyses. *J. Mol. Biol. 395*, 785−802.
(40) Daber, R., Sochor, M. A., and Lewis, M. (2011) Thermodynamic analysis of mutant *lac* repressors. *J. Mol. Biol. 409*, 76−87.
(41) Poelwijk, F. J., de Vos, M. G. J., and Tans, S. J. (2011) Tradeoffs and optimality in the evolution of gene regulation. *Cell 146*, 462−470.