THE PROTEIN SOCIETY

Lactose repressor hinge domain independently binds DNA

Joseph S. Xu, Madeleine N. Hewitt, Jaskeerat S. Gulati, Matthew A. Cruz, Hongli Zhan, Shirley Liu, and Kathleen S. Matthews*

Department of BioSciences, MS-140, Rice University, Houston, Texas 77251

Received 19 November 2017; Accepted 2 January 2018
DOI: 10.1002/pro.3372
Published online 10 January 2018 proteinscience.org

Abstract: The short 8–10 amino acid "hinge" sequence in lactose repressor (LacI), present in other LacI/GalR family members, links DNA and inducer-binding domains. Structural studies of full-length or truncated LacI-operator DNA complexes demonstrate insertion of the dimeric helical "hinge" structure at the center of the operator sequence. This association bends the DNA ~40° and aligns flanking semi-symmetric DNA sites for optimal contact by the N-terminal helix-turn-helix (HtH) sequences within each dimer. In contrast, the hinge region remains unfolded when bound to nonspecific DNA sequences. To determine ability of the hinge helix alone to mediate DNA binding, we examined (i) binding of LacI variants with deletion of residues 1–50 to remove the HtH DNA binding domain or residues 1–58 to remove both HtH and hinge domains and (ii) binding of a synthetic peptide corresponding to the hinge sequence with a Val52Cys substitution that allows reversible dimer formation via a disulfide linkage. Binding affinity for DNA is orders of magnitude lower in the absence of the helix-turn-helix domain with its highly positive charge. LacI missing residues 1–50 binds to DNA with ~4-fold greater affinity for operator than for nonspecific sequences with minimal impact of inducer presence; in contrast, LacI missing residues 1–58 exhibits no detectable affinity for DNA. In oxidized

Abbreviations: EMSA, electrophoretic mobility shift assay; His_6, sequence Leu-Glu-His-His-His-His-His-His added to the C-terminal end of the protein; HtH, helix-turn-helix of LacI N-terminal domain; IPTG, isopropyl-β,D-thiogalactoside; LacI, lactose repressor protein; LacI-51, LacI sequence missing the first 50 residues; LacI-59, LacI sequence missing the first 58 residues; O1, primary *lac* operator DNA sequence; O_{scram}, DNA sequence with similar nucleotide content to O1, but scrambled in order

Additional Supporting Information may be found in the online version of this article.

Joseph S. Xu, Madeleine N. Hewitt, Jaskeerat S. Gulati, and Matthew A. Cruz contributed equally to this work.

Statement of importance: The short "hinge" sequence in LacI, also found in other members of the LacI/GalR family, links the larger DNA-binding and inducer-binding domains and forms a dimeric helical structure only in complex with operator DNA. Insertion of the hinge helix dimer into the minor groove is essential for high affinity binding by adjacent DNA binding domains. Using truncated LacI proteins and a synthetic hinge peptide, we demonstrate the capacity for independent DNA binding by the LacI hinge domain.

Grant sponsor: National Institute of General Medical Sciences; Grant number: GM022441; Grant sponsor: Robert A. Welch Foundation; Grant number: C-576.

Joseph S. Xu's current address is 334 East 26th St., New York, NY 10010
Madeleine N. Hewitt's current address is Molecular & Cellular Biology Program, University of Washington, Box 357275, Seattle, WA 98195
Jaskeerat S. Gulati's current address is UT Southwestern Medical School, Dallas, TX 75390
Matthew A. Cruz's current address is Division of Biology and Biomedical Sciences, Washington University, St. Louis, MO 63110-1093.
Hongli Zhan's current address is Center for Molecular Biology and Pathology, LabCorp, 1912 T. W. Alexander Dr., Research Triangle Park, NC 27709
*Correspondence to: Kathleen S. Matthews, Department of BioSciences, MS-140, Rice University, Houston, TX 77251. E-mail: ksm@rice.edu

form, the dimeric hinge peptide alone binds to O1 and nonspecific DNA with similarly small difference in affinity; reduction to monomer diminished binding to both O1 and nonspecific targets. These results comport with recent reports regarding LacI hinge interaction with DNA sequences.

Keywords: allosteric regulation; DNA binding protein; DNA operator; DNA–protein interaction; hinge helix; lactose repressor protein; structure–function

Introduction

Lactose repressor (LacI) regulates the expression of the bacterial genes involved in lactose metabolism in response to environmental availability of this energy resource.[1,2] LacI, a member of the extended LacI/GalR family,[3] is unusual because of its tetrameric structure as a dimer of dimers.[1,2,4] Each LacI monomer comprises four key domains (Fig. 1): helix-turn-helix DNA binding domain (HtH), hinge, core inducer binding domain, and tetramerization domain. Both X-ray[4–8] and NMR studies[9–16] of the LacI•Lac operator complex demonstrate the interaction of the two helix-turn-helix DNA binding domains of each dimer with symmetric outer regions of the operator DNA sequence (O1). The conserved hinge region in multiple LacI/GalR family members provides the only covalent connection between the ligand-binding site formed by the core domain and the N-terminal HtH DNA binding site [LacI structure shown in Fig. 1(A)].[1–4] When folded into its helical form in the absence of inducer, this hinge region in LacI interacts with its partner hinge within a dimer as well as with the surface of the adjacent core domain. This pair intercalates into the minor groove to bend and rotate the target DNA, optimizing contacts in the major groove with HtH recognition sequences.[4–17]

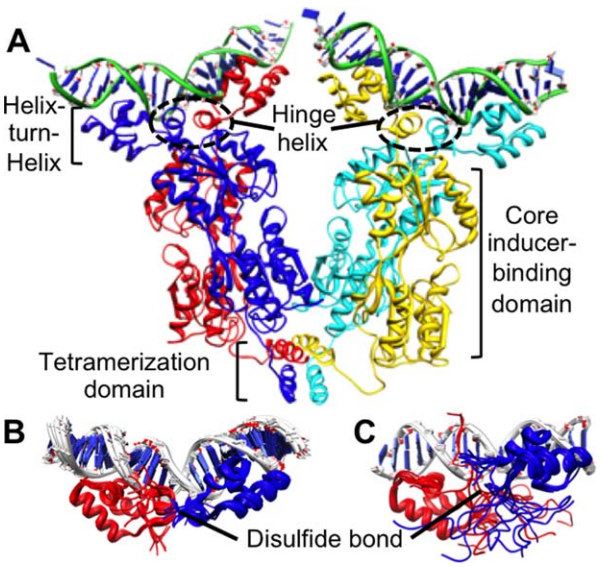

Figure 1. X-ray crystallographic and NMR structures of LacI and its domains. (A) X-ray structure for tetrameric LacI protein (PDB file 1LBG[4]). (B) and (C) show NMR structures of DNA complexes of truncated LacI containing the N-terminal HtH DNA binding domain and the hinge helix sequence with Val52Cys substitution. To enhance clarity, 10 of the 20 structures in the NMR-based PDB files were deleted in these NMR structures. (B) O1 operator DNA (PDB file 1L1M[14]). (C) Nonspecific DNA (PDB file 1OSL[16]). The black lines indicate the position of the disulfide linkage between the hinge regions in the oxidized Val52Cys variant. Note that the complex with nonspecific DNA is linear rather than bent with no hinge helix present.

The side chains at position 52 are positioned in sufficiently close proximity within each dimer of the LacI tetramer to form a disulfide linkage under oxidizing conditions when the native valine is substituted by cysteine.[18,19] Under these conditions, enhanced affinity for operator and loss of response to inducer binding are observed.[18,19] Although Val52 is the only site within the hinge helix that does not interact with DNA and/or the inducer-binding domain,[4–8] substitutions at this position impact operator binding affinity and specificity.[19,20] A subset of these mutations results in impaired functional communication with the inducer-binding domain,[20] pointing to the importance of the hinge region to the overall function of this protein.

Kalodimos and colleagues[12] utilized the Val52Cys mutation to generate a dimeric truncated LacI comprised of only the HtH-DNA binding domain and hinge helix region. NMR structural analysis demonstrated folding and insertion of the hinge helix into the minor groove of operator DNA, bending the DNA to align specific contacts by the helix-turn-helix in the adjacent major grooves [Fig. 1(B)].[12,14,16] Interestingly, helix formation by the hinge region was not observed in NMR analysis of this dimeric construct complexed with nonspecific DNA [Fig. 1(C)].[16] Note that this region is not observed in crystallographic structures in the absence of operator DNA nor in the presence of inducer.[2,4–8] However, NMR structural data[16] and recent theoretical/computational studies by Xu *et al.*[21] and Sun *et al.*[22] indicate that the hinge region contributes significantly to binding both operator and nonspecific DNA sequences, but utilizes different conformations and types of bonding for these two modes of binding.

To explore in more detail the capacity of the hinge helix to independently direct specific and nonspecific DNA binding, we have utilized (i) LacI missing the N-terminal HtH domain, (ii) LacI missing HtH domain *and* the hinge region, and (iii) synthetic hinge helix peptide with the Val52C alteration to allow stable

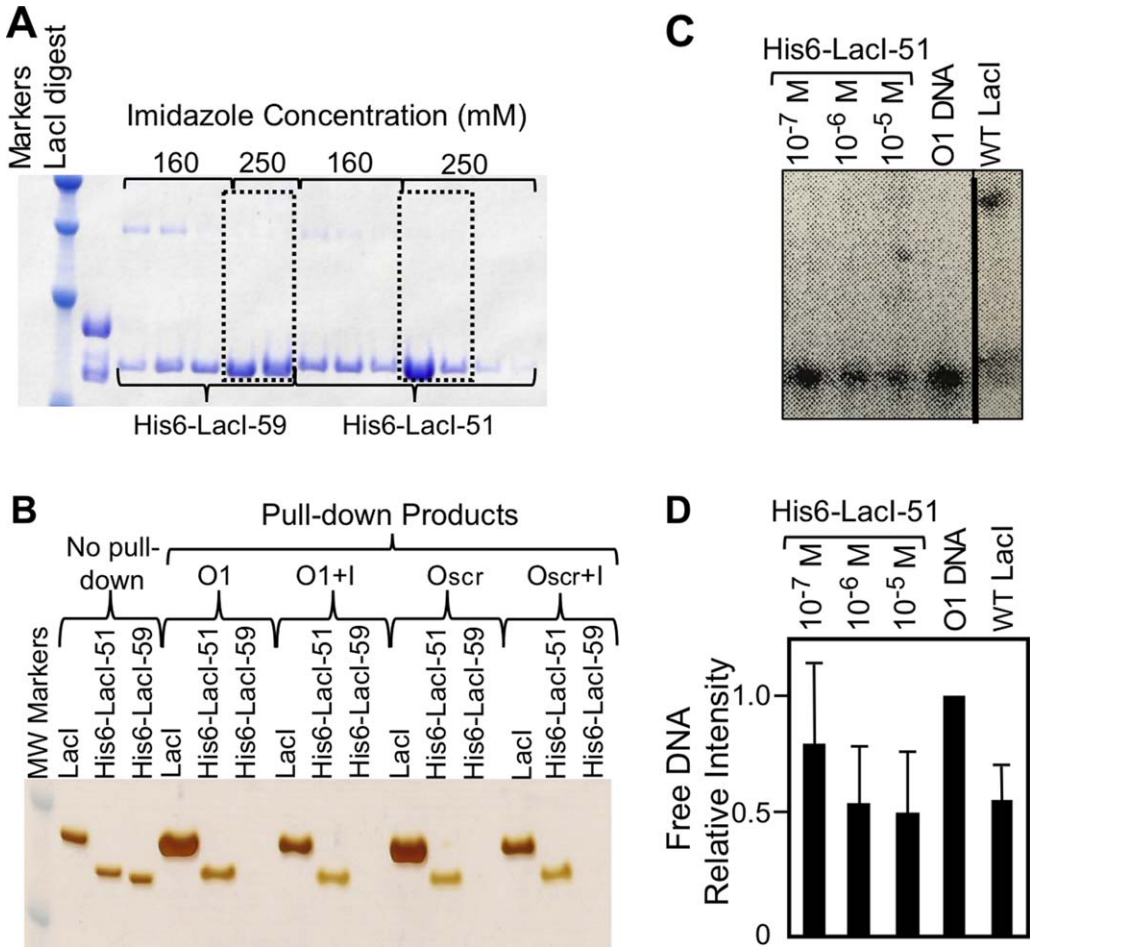

Figure 2. Purification and binding assays for His-tagged LacI variants. (A) Purification of His_6-LacI-51 and His_6-LacI-59. Products of purification were examined using SDS-PAGE to confirm purity (dashed boxes indicate samples used for binding assays). Lane 2 shows trypsin digestion products of wild-type LacI (no His-tag) for reference. (B) His-tagged deletion mutants were used in a pull-down assay with biotinylated O1 and O_{scram} DNA bound to streptavidin-coated beads with and without IPTG (I). Note that removal of the hinge helix in His_6-LacI-59 results in complete loss of DNA binding. (C) Electrophoretic mobility shift assays. Varying concentrations of His_6-LacI-51 and 10^{-7} *M* LacI were each mixed with [^{32}P]-labeled 40 bp O1 DNA ($\leq 10^{-11}$ *M*) and equilibrated, followed by rapid loading onto a polyacrylamide gel. Note the loss of free DNA and detection of bound bands for His_6-LacI-51 at 10^{-5} *M* protein; the lowest bound band exhibits slightly higher mobility compared to wild-type LacI, as expected for this smaller tetramer. Note that wild-type LacI was separated in this experiment from the His_6-LacI-51 and O1 DNA, and the dark line indicates removal of wells that contained other materials for analysis. (D) Band intensities for free DNA were derived from phosphorimaging of four separate experiments for His_6-LacI-51 and three separate experiments for wild-type LacI; the values were normalized to the free O1 DNA band intensity for each experiment. Standard deviations for free DNA in samples with bound species are indicated.

dimer formation. Results of DNA binding experiments for these constructs indicate that the hinge region mediates both specific and nonspecific binding, with only small differences in affinity for these targets and no impact of inducer binding. These data are consistent with NMR analysis[16] and with theoretical predictions that the hinge region contributes to both operator and nonspecific DNA binding.[21,22]

Results

Generation and purification of LacI N-terminal deletion mutants

To purify truncated LacI variants, three constructs were generated with a tag comprised of a Leu-Glu linker followed by six His residues added to the C-terminus of each protein. Plasmids were generated for expression of full-length LacI, LacI-51 missing the N-terminal 50 amino acids that encode the HtH domain, and LacI-59 missing the N-terminal 58 amino acids that encode both the HtH and hinge region. The proteins produced were purified using Ni-NTA columns [Fig. 2(A)]. No significant difference in DNA binding behavior was observed for His_6-LacI compared to untagged protein (data not shown).

DNA pull-down assays

To assess DNA binding, pull-down assays using streptavidin-coated magnetic beads and biotinylated

DNA sequences were performed [Fig. 2(B)]. Note that assessing affinity quantitatively for specific and nonspecific DNA is not possible for this assay given the high local concentration of DNA at the bead surface (>5 μ*M* minimum). In addition, both affinity and avidity (i.e., binding to multiple closely packed DNA sequences) would be detected by this process. Wild-type LacI bound similarly to operator (O1) in the absence of IPTG, O1 in the presence of IPTG, and nonspecific DNA (O_{scram}, a DNA sequence with similar nucleotide content to O1, but arranged randomly; see Supporting Information for sequence and details). Similarly, His_6-LacI-51 bound to O1 and to O_{scram} in both the presence and absence of IPTG. However, His_6-LacI-59, missing the hinge helix, was not pulled down by either O1 or nonspecific DNA sequences despite their high local concentration on the beads [Fig. 2(B)].

Gel electrophoretic mobility shift assays

The ability of a bound protein to retard DNA movement in gel electrophoresis has been widely applied to characterization of binding.[23] Rapid dissociation, found often for weak protein•DNA complexes ($K_d > 10^{-7}$), requires using loss of free DNA to assess binding.[23] For His_6-LacI-51 in the presence of O1 DNA at a concentration >100-fold below the K_d for wild-type LacI, levels of free DNA decrease with increasing concentration of protein, and a bound band is observed only at the highest concentration of His_6-LacI-51 [Fig. 2(C, D)]. No change in intensity of free DNA (or presence of lower mobility bands) was observed for His_6-LacI-59 at the same concentrations (data not shown).

Nitrocellulose DNA binding assays

Binding of His_6-LacI-51 was examined using retention of radiolabeled DNA by protein using nitrocellulose filters (Fig. 3, Supporting Information). The experiments were performed in the presence and absence of IPTG. His_6-LacI-51 bound with ~4-fold higher affinity to the O1 sequence than to nonspecific DNA, O_{scram} (Table I, Supporting Information). This differential is significantly smaller than that for wild-type LacI (>10^4 difference). Further, the presence of IPTG has no impact on binding by His_6-LacI-51 to either specific or nonspecific DNA sequences, consistent with computational studies indicating the importance of the hinge domain in stabilizing both types of binding, albeit with different types of bonds (i.e., apolar for O1 binding vs. ionic and hydrogen bonding for nonspecific binding).[21,22] The apparent cooperativity observed for O1 binding may reflect the requirement for hinge folding and association of two hinge helices within each dimer to form the bound complex (see Supporting Information for further details). However, at the concentrations and conditions required for these measurements, interpretations regarding cooperativity should be considered speculative. As would be anticipated, no DNA binding in similar experiments was observed for His_6-LacI-59.

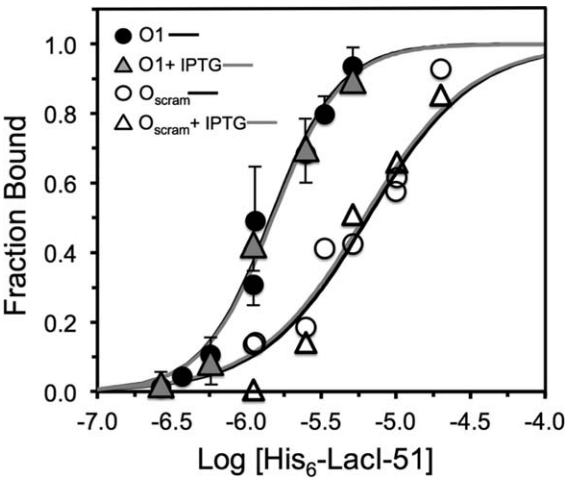

Figure 3. Nitrocellulose filter binding assays of His_6-LacI-51. Data were normalized to maximum binding observed to allow compilation of multiple experiments. Black-filled circles, His_6-LacI-51 and O1; gray-filled triangles, His_6-LacI-51, O1, and 2 m*M* IPTG. Open circles, His_6-LacI-51 and O_{scram}; open triangles, His_6-LacI-51, O_{scram} and 2 m*M* IPTG. Results are shown from four experiments for O1 binding and three experiments for O_{scram} binding with each experiment comprised of internal duplicate or triplicate samples. Error range corresponding to the standard deviation is shown either above or below each point to facilitate visualization (if there is no error bar, the error is smaller than the point symbol). The lines shown are fits to the data using the Hill equation (see Supporting Information for details). The protein concentration for 50% fractional saturation ($R_{50\%}$) derived from the fitted data is shown in Table I for each condition as estimated equilibrium dissociation constants.

Binding of hinge helix peptide to operator and nonspecific DNA

A synthetic peptide (Fig. 4) was obtained that corresponded to the hinge helix region and included the Val52Cys substitution and a 5'-fluorescein label at its C-terminus. The peptide was purified and oxidized to form dimer and then assayed for binding to operator and nonspecific DNA. Analysis by mass spectrometry indicated that ~70% of the oxidized peptide was in dimer form. Using biotinylated DNA sequences bound to streptavidin-coated magnetic beads, the fluorescent hinge helix peptide in its oxidized, disulfide-linked, dimeric state bound with only a small difference in affinity to O1 and O_{scram}. However, upon reduction, the level of binding to both O1 and O_{scram} decreased [Fig. 4(B)].

Nitrocellulose filter binding assays for peptides

After establishing that the peptide bound effectively to nitrocellulose by monitoring retained fluorescence,

Table I. *Estimated Equilibrium Binding Constants ($R_{50\%}$)*

Protein	Operator DNA (O1)	Operator DNA (O1) + IPTG	Nonspecific DNA (O_{scram})	Nonspecific DNA (O_{scram}) +IPTG
His_6-LacI-51	$\sim 1.5 \times 10^{-6}\ M$	$\sim 1.5 \times 10^{-6}\ M$	$\sim 6.5 \times 10^{-6}\ M$	$\sim 5 \times 10^{-6}\ M$
His_6-LacI-59	Not detectable	Not detectable	Not detectable	Not detectable
Monomeric peptide	$\sim 9 \times 10^{-6}\ M$	n/a	$\geq 4 \times 10^{-5}\ M$	n/a
Dimeric peptide	$\sim 3 \times 10^{-6}\ M$	n/a	$\sim 9 \times 10^{-6}\ M$	n/a

Estimated equilibrium constants ($R_{50\%}$) for His_6-LacI-51 and His_6-LacI-59 were derived from the 50% saturation value from nitrocellulose filter binding assays using operator DNA and nonspecific DNA (O_{scram}), with and without IPTG (Fig. 3 and Supporting Information). Estimated equilibrium constants ($R_{50\%}$) were determined for monomeric and dimeric peptide binding to operator and nonspecific DNA from the nitrocellulose filter binding data shown in Fig. 4 (no values for +IPTG are provided, since there is no core domain for inducer binding). Given the experimental limitations at the high protein/peptide concentrations required for these studies, these values are considered to be *approximations* to allow facile comparison. Analysis of the His_6-LacI-51 data using curve fitting to the Hill equation can be found in the Supporting Information.

filter binding assays were utilized to measure interaction of the hinge helix peptide with O1 and O_{scram} [Fig. 4(C)]. Although binding in all cases tested was weak, the highest extent of binding occurred for the oxidized hinge sequence with O1 operator. Dimeric hinge also bound to O_{scram}, but with lower affinity. Generation of the monomeric peptide by incubation with dithiothreitol resulted in diminished binding to both operator and nonspecific DNA under comparable conditions [Fig. 4(C), Table I].

Discussion

Early proteolytic experiments with LacI were used to separate the inducer-binding domain and the N-terminal helix-turn-helix DNA binding region.[24,25] Under conditions that targeted Arg 51, leaving the hinge region intact,[24,25] this truncated protein bound DNA with significantly lower affinity than wild-type LacI.[26–28] Later studies produced crystallographic and NMR structures of LacI and other members of the LacI/GalR family of regulatory proteins that illuminated the role of the hinge helix in bending the central region of the DNA target sequence to align the flanking sequences for HtH binding.[4–16,29–36] Within the LacI/GalR family, specific DNA binding and regulatory responses for each protein are generated by variations in shared structural motifs.[4–16,29–36] Structural homology is evident from detailed sequence analysis and crystallographic structures that include LacI, PurR, and CcpA.[4–8,29–36] Insertion of the hinge helix into the minor groove is a common feature of these structures, enabling high affinity binding by the HtH to flanking sequences.[4–8,33–36] Nonetheless, the DNA bend elicited by hinge insertion varies from ~50° for PurR, ~40° for LacI, to ~35° for CcpA.[36] The structural changes in DNA elicited by the hinge helix optimize the contacts that allow the major groove to

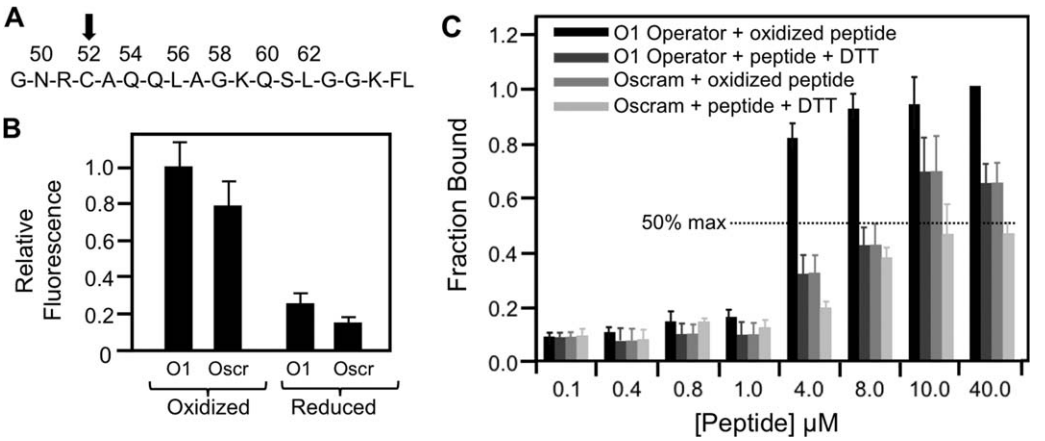

Figure 4. Hinge peptide sequence and binding data. (A) Sequence of the synthetic hinge peptide with Val52Cys substitution (arrow) to allow disulfide linkage. Amino acids in wild-type LacI are numbered 50–62. Note that fluorescein (FL) is attached to the side chain amino group of the C-terminal lysine residue. (B) Magnetic bead pull-down results for synthetic hinge peptide (10 μ*M*) in oxidized and reduced form using fluorescence to detect bound fluorescein. Data from three separate experiments with four replicates in each experiment were normalized to the average value for O1 DNA with oxidized peptide. (C), Nitrocellulose filter binding data with synthetic peptide. Values are normalized to the average value at 40 μ*M* oxidized peptide for O1 DNA. Black bars, oxidized peptide and O1; dark gray bars, reduced peptide and O1; medium gray bars, oxidized peptide and O_{scram}; light gray bars, reduced peptide and O_{scram}. The dotted line indicating 50% of the maximum binding observed for O1 operator can be used to determine approximate binding parameters and the relative stability of the complexes.

be "read" with specificity. Although mutations in PurR can result in changes in minor groove contacts, the hinge helix remains intact and bends the DNA to allow high affinity binding.[37] Thus, DNA binding engages similar features in these LacI/GalR proteins. In contrast, regulation of binding is mediated by distinct types of interactions with the core domain: LacI responds to allolactose and other galactosides to interfere with high affinity binding;[1] PurR responds to purines with increased DNA binding affinity;[33–35] and CcpA binding to DNA utilizes phosphorylated HPr as a co-repressor.[36] Each of these responses regulates the metabolic context within the bacterium.

Interestingly, comparison of LacI and PurR structures has indicated that these proteins utilize different contacts between the hinge region and flanking sequences and the core regulatory domain.[38] Within this small subset of LacI/GalR proteins with known structure, the variety of ways that similar folds can be utilized to regulate DNA is impressive, and the hinge domain that links the DNA- and effector-binding domains plays a critical functional role.[4–8,29–38] Although no structural data are available for the CytR protein, which interacts with cAMP receptor protein to regulate transcription, flexibility of the hinge domain and flanking linker region has been shown to be essential to its regulatory function.[39]

Unlike other members of this family, LacI is both tetrameric and capable of binding to a variety of semi-symmetric operator sites that differ in the length of the central spacer sequence.[13,16,19,40] Binding is achieved by an asymmetric arrangement of the two helix-turn-helix motifs within the dimeric DNA binding unit of the tetramer.[13,16,41,42] Two features of LacI are required for this flexible recognition process: (i) the hinge helix, which recognizes and kinks the central region of the operator in a symmetric target site,[4–7,12–16] and (ii) a YQ sequence at positions 17–18 in the N-terminal DNA binding domain that recognizes sequences in the flanking regions of the operator DNA sequence.[41,42] Not surprisingly, introducing the LacI hinge helix and YQ sequence into the purine repressor resulted in enhanced binding to alternative target DNA sequences.[42]

Xu and colleagues[21] have used molecular dynamics simulations and correlation network analysis to examine the recognition of DNA sequences by LacI. Electrostatic and hydrogen-bonding interactions were found to be prominent in mediating nonspecific interaction with DNA. In contrast, high-affinity operator binding relied on hydrophobic interactions by the hinge helices that introduce the bending required for specific interaction with operator DNA sequences. Xu *et al.*[21] concluded that the hinge region played an important role in modulating information flow between LacI and DNA sites. Sun *et al.*[22] utilized solvent molecular dynamics simulation and continuum electrostatic calculations to explore the hinge–DNA interaction for LacI and concluded that the hinge contributes significantly to stabilizing nonspecific DNA binding in addition to its key contribution to operator binding. Based on examination of solute impacts on kinetics for LacI–LacO binding, Record and colleagues[43] have recently proposed that nucleation of hinge helix formation in this complex provides the transition state to specific operator binding.

Our goal was to explore the independent role of the hinge domain in the absence of the helix-turn-helix domain. Early studies indicated that removal of this DNA binding domain by proteolysis lowered but did not destroy DNA binding capacity.[26–28] We were interested in the specific impact of the hinge domain in its native context of attachment to the inducer binding core domain of LacI and as an isolated peptide. We therefore generated His-tagged constructs of wild-type LacI and deletions of 50 (His_6-LacI-51) and 58 (His_6-LacI-59) N-terminal amino acids. Multiple assessments of DNA binding indicated that removing 58 amino acids resulted in no detectable capacity to bind DNA sequences. In particular, pull-down assays using magnetic beads are highly sensitive and, due to the very high local concentration of the DNA ligand, can detect very weak binding interactions; however, even this method did not demonstrate DNA binding by the purified His_6-LacI-59 protein.

In contrast, removal of only 50 amino acids, leaving the hinge region intact, resulted in tetrameric protein with DNA binding capacity. Interestingly, and perhaps not surprisingly in the context of NMR studies,[16] recent computational and theoretical analyses,[21,22] and examination of solute effects on binding,[43] His_6-LacI-51 bound to DNA with a K_d of $\sim 1.5 \times 10^{-6}$ M, only slightly higher affinity than the isolated hinge domain (Table I). In further concert with theoretical studies,[21,22] binding was not inducer-sensitive. This latter observation is presumably related to the dual capacities of this region: (1) two hinge helices can insert as a pair into the minor groove at a specific sequence and thereby bend the DNA and (2) the unfolded hinge can bind to nonspecific DNA with only slightly lower affinity. Not surprisingly, these two binding modes have been deduced to utilize different types of noncovalent bonding.[21,22] These observations are consistent with hinge binding to linear nonspecific DNA shown by NMR studies of the N-terminal region containing the HtH and hinge domains.[16]

Conclusion

A significant body of work on LacI structure and function demonstrates that multiple regions are crucial to its key functions: binding inducer, operator

DNA, and nonspecific DNA. However, the hinge domain—a sequence of <10 amino acids—stands out. This region is essential for recognition of specific DNA sequences in LacI and other members of the extended LacI/GalR family by altering the path of the DNA backbone.[4–7,12–16,33–36] For LacI, hinge helix insertion into the minor groove using primarily apolar contacts bends the DNA and positions the flanking major groove sequences in O1 DNA to make specific contacts with the helix-turn-helix motifs.[16,21,22] In contrast, nonspecific binding employs hydrogen bonding and ionic interactions to stabilize binding to the DNA backbone and maintain high local concentration of LacI.[16,21,22] The experimental studies presented, in concert with previous theoretical studies[21,22] and recent examination of the mechanism using solute effects on rate constants,[43] confirm that the hinge sequence contributes independently and significantly to binding affinity for both specific and nonspecific DNA and plays a key role in the transition from nonspecific to specific binding.[43] This region illustrates the varied, often essential, roles that even short sequences can play in the functional properties of a protein.

Experimental Procedures

Materials and methods

Information on sources of chemicals, peptide, and DNA sequences and details on common laboratory procedures are in Supporting Information.

Generation of plasmids encoding LacI-51, LacI-59, and wild-type LacI with His-tags

LacI-51, LacI-59, and wild-type LacI coding sequences were individually cloned into pET-20b(+) (Addgene) between NdeI and XhoI restriction sites under control of a T7 polymerase promoter. A Met codon was introduced just prior to the first amino acid in each version. Codons for the C-terminal sequence, Leu-Glu-His-His-His-His-His-His-Term, were placed at the end of the LacI coding sequence well separated from the DNA binding regions. Sequences of the constructs were confirmed by DNA sequencing (GeneWiz, South Plainfield, NJ).

Expression and purification of His-tagged LacI

Plasmids containing the desired coding sequences were cotransformed with the pTARA plasmid, which encodes T7 polymerase under arabinose control, into *E. coli* BLIM cells, a LacI-negative strain.[44] Cells were grown for 24 h at 37°C in 1 L of 2xYT media (15 g Bacto tryptone, 10 g yeast extract, 5 g NaCl, pH 7.4). Two hours prior to harvest, solution containing 2.5 g of arabinose was added to each 1 L liquid culture. Cells harvested from each liter of liquid culture by centrifugation were resuspended in 12.5 mL of lysis buffer (200 m*M* Tris-HCl, 200 m*M* KCl, 14 m*M* $Mg(OAc)_2$, 5% glucose, ~25 μg/mL DNase, 0.3 m*M* DTT, 0.4 m*M* AEBSF, one complete EDTA-free protease inhibitor cocktail tablet, 0.5 mg/mL lysozyme at a final pH of 7.8) and stored at −80°C. Harvested cells were thawed, mixed with 20 mL breaking buffer (200 m*M* Tris-HCl, pH 7.5, ~25 μg/mL DNase, 200 m*M* $MgCl_2$, 0.1 m*M* DTT, 5% glucose, one complete EDTA-free protease inhibitor cocktail tablet, 1 m*M* AEBSF), and centrifuged at 40,000*g* for 1 h at 4°C. Supernatant was applied to a Ni-NTA (Qiagen) column equilibrated with buffer containing 90 m*M* Tris-HCl, pH 7.5, 5% glucose, 10 m*M* imidazole, 0.1 m*M* DTT. Protein was eluted stepwise in the same buffer using increasing concentrations of imidazole (20, 40, 80, 160, and 250 m*M*). SDS-PAGE was used to identify fractions with the desired protein. Fluorescence shift in the presence of IPTG (100 μ*M*) demonstrated folded core domain for the purified His_6-tagged proteins.

Magnetic bead DNA pull-down assay

Streptavidin magnetic beads (New England BioLabs) were placed in a magnetic tube rack and the supernatant removed. Beads (25 μL of original suspension per sample) were washed three times in an equal volume of Buffer 1 containing 10 m*M* Tris-HCl, pH 7.5, 1 m*M* EDTA, 0.5 *M* NaCl. The beads were resuspended in one equivalent each of Buffer 1 and biotinylated operator DNA (O1, 2 μ*M*) or biotinylated scrambled DNA (O_{scram}, 2 μ*M*) in DNA-grade water and then washed three times with two equivalents of Buffer 2 (10 m*M* Tris-HCl, pH 7.5, 1 m*M* EDTA, 0.25 *M* NaCl). Buffer 2 was removed and beads were resuspended in two equivalents of Buffer 3 (10 m*M* Tris-HCl, pH 7.4, 0.1 m*M* EDTA, 150 m*M* KCl, 5% DMSO). For proteins, 30 μL of a purified sample at ~1.6 μ*M* in the same buffer was added to the beads and incubated at room temperature for 10 min with occasional mixing. The beads and bound protein/peptide were separated from the supernatant using a magnetic rack and washed three times in 50 μL Buffer 3. To dissociate bound protein, each sample was resuspended in 15 μL 13.3% SDS, 0.33 *M* DTT, 3.3 m*M* Tris-HCl, pH 7.4, 0.03 m*M* EDTA, and 50 m*M* KCl, and the solutions were heated at 90°C for 10 min. Supernatant was removed using a magnetic rack followed by SDS-PAGE separation and quantitation by silver staining.

For peptide experiments, oxidized peptide was generated by overnight exposure to air and reduced peptide by addition of 5 μL of 100 m*M* DTT in peptide buffer comprised of 10 m*M* Tris-HCl, pH 7.4, 150 m*M* KCl. Oxidized/reduced samples (30 μL each) at varying concentrations were mixed with an equal volume of prepared beads in peptide buffer without DTT and incubated at room temperature for 10 min with occasional mixing. The beads and bound peptide were separated and washed three times in 50

μL of peptide buffer without DTT. The final supernatant was removed using a magnetic rack, and beads with bound DNA–peptide were resuspended in peptide buffer. The samples were placed in a 96-well plate, and fluorescence was measured on a Tecan Infinite M1000 instrument.

Electrophoretic mobility shift assay (EMSA)

EMSA experiments[23] were performed with [^{32}P]-labeled 40 bp DNA encoding the *lac* O1 sequence (Supporting Information Table I). Radiolabeled DNA was added to a final concentration of 3.5×10^{-11} *M* and incubated at room temperature with purified His_6-LacI-51, His_6-LacI-59 or wild-type LacI for 10 min in binding buffer (10 m*M* Tris-HCl, 10 m*M* KCl, 0.1 *M* EDTA, 0.3 m*M* DTT, 1 mg/mL bovine serum albumin, pH 7.4). Following equilibration, glycerol was added to a final concentration of 10%. Samples were electrophoresed at 250 V on 4% native polyacrylamide gels in TBE buffer (45 m*M* Tris-borate, 0.75 m*M* EDTA, pH 8.3). The gels were dried and results analyzed after exposure of a Fuji phorphorimaging plate.

Nitrocellulose filter binding assay

Nitrocellulose filter binding was used to determine binding affinities for radiolabeled O1 and O_{scram} to His_6-LacI, His_6-LacI-51, and His_6-LacI-59 and monomeric and disulfide-linked peptides.[45] The assay was conducted at room temperature in buffer containing 10 m*M* Tris-HCl, pH 7.4, 10 m*M* KCl, 0.3 m*M* DTT, 0.1 m*M* EDTA. Samples with inducer contained 2 m*M* IPTG. For peptide samples, the buffer was 10 m*M* Tris-HCl, pH 7.4, 0.15 *M* KCl. DTT was not present in the assays for oxidized peptide. The DNA concentration was $\geq$100-fold lower than the K_d for wild-type LacI binding to O1, and peptide/protein concentrations were varied over several orders of magnitude. Protein or peptide was incubated with [^{32}P]-labeled O1 or O_{scram} DNA in a 96-well plate for $\sim$20 min before filtering through nitrocellulose. Retained radioactivity was quantified by a Fuji phosphorimager for further analysis.

Acknowledgments

We are deeply grateful to John S. Olson for his expertise in fitting the DNA binding data and for his interest in and long-held support for this project. We offer enthusiastic thanks to Ryan Fleming and Brett Baggett for performing early experiments that suggested that this project was worthy of deeper pursuit. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of Interest

The authors declare that they have no conflicts of interest with the contents of this article.

Author Contributions

JSX, MNH, JSG, and MAC conducted the experiments. JSX and MNH focused on the truncated LacI purification and characterization, and JSG and MAC focused on the "hinge helix" peptide experiments. SL provided training for methodologies utilized and supervised the experiments. HZ constructed the plasmid encoding the His-tagged proteins. The entire team analyzed the results and contributed to writing the paper. KSM conceived the idea for this project, performed a few experiments, analyzed data, and was responsible for drafting the paper with assistance and feedback from the entire team. All authors reviewed the results and approved the final version of the manuscript.

References

1. Matthews KS, Nichols JC (1998) Lactose repressor protein: functional properties and structure. Prog Nucl Acid Res Mol Biol 58:127–164.
2. Lewis M (2005) The *lac* repressor. C R Biol 328:521–548.
3. Weickert MJ, Adhya S (1992) A family of bacterial regulators homologous to Gal and Lac repressors. J Biol Chem 267:15869–15874.
4. Lewis M, Chang G, Horton NC, Kercher MA, Pace HC, Schumacher MA, Brennan RG, Lu P (1996) Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. Science 271:1247–1254.
5. Bell CE, Lewis M (2000) A closer view of the conformation of the Lac repressor bound to operator. Nat Struct Biol 7:209–214.
6. Bell CE, Lewis M (2001) Crystallographic analysis of Lac repressor bound to natural operator *O1*. J Mol Biol 312:912–926.
7. Bell CE, Lewis M (2001) The Lac repressor: a second generation of structural and functional studies. Curr Opin Struct Biol 11:19–25.
8. Daber R, Stayrook S, Rosenberg A, Lewis M (2007) Structural analysis of Lac repressor bound to allosteric effectors. J Mol Biol 370:609–619.
9. Spronk CAEM, Folkers GE, Noordman A-MGW, Wechselberger R, van den Brink N, Boelens R, Kaptein R (1999) Hinge-helix formation and DNA bending in various *lac* repressor-operator complexes. EMBO J 18: 6472–6480.
10. Spronk CAEM, Slijper M, van Boom JH, Kaptein R, Boelens R (1996) Formation of the hinge helix in the *lac* repressor is induced upon binding to the *lac* operator. Nat Struct Biol 3:916–919.
11. Spronk CAEM, Bonvin AMJJ, Radha PK, Melacini G, Boelens R, Kaptein R (1999) The solution structure of Lac repressor headpiece 62 complexed to a symmetrical *lac* operator. Structure 7:1483–1492.
12. Kalodimos CG, Folkers GE, Boelens R, Kaptein R (2001) Strong DNA binding by covalently linked dimeric Lac headpiece: Evidence for the crucial role of the hinge helices. Proc Natl Acad Sci USA 98:6039–6044.
13. Kalodimos CG, Boelens R, Kaptein R (2002) A residue-specific view of the association and dissociation pathway in protein–DNA recognition. Nat Struct Biol 9: 193–197.
14. Kalodimos CG, Bonvin AMJJ, Salinas RK, Wechselberger R, Boelens R, Kaptein R (2002) Plasticity in protein-DNA recognition: *lac* repressor interacts with

its natural operator *O1* through alternative conformations of its DNA binding domain. EMBO J 21:2866–2876.
15. Kalodimos CG, Boelens R, Kaptein R (2004) Toward an integrated model of protein–DNA recognition as inferred from NMR studies on the *Lac* repressor system. Chem Rev 104:3567–3586.
16. Kalodimos CG, Biris N, Bonvin AMJJ, Levandoski MM, Guennuegues M, Boelens R, Kaptein R (2004) Structure and flexibility adaptation in nonspecific and specific protein–DNA complexes. Science 305:386–389.
17. Flynn TC, Swint-Kruse L, Kong Y, Booth C, Matthews KS, Ma J (2003) Allosteric transition pathways in the lactose repressor protein core domains: asymmetric motions in a homodimer. Protein Sci 12:2523–2541.
18. Falcon CM, Swint-Kruse L, Matthews KS (1997) Designed disulfide between N-terminal domains of lactose repressor disrupts allosteric linkage. J Biol Chem 272:26818–26821.
19. Falcon CM, Matthews KS (2001) Engineered disulfide linking the hinge regions within lactose repressor dimer increases operator affinity, decreases sequence selectivity, and alters allostery. Biochemistry 40:15650–15659.
20. Zhan H, Swint-Kruse L, Matthews KS (2006) Extrinsic interactions dominate helical propensity in coupled binding and folding of the lactose repressor protein hinge helix. Biochemistry 45:5896–5906.
21. Xu L, Ye W, Jiang C, Yang J, Zhang J, Feng Y, Luo R, Chen H-F (2015) Recognition mechanism between Lac repressor and DNA with correlation network analysis. J Phys Chem B 119:2844–2856.
22. Sun L, Tabaka M, Hou S, Li L, Burdzy K, Aksimentiev A, Maffeo C, Zhang X, Holyst R (2016) The hinge region strengthens the nonspecific interaction between Lac-repressor and DNA: a computer simulation study. PLoS One 11:e0152002.
23. Hellman LM, Fried MG (2007) Electrophoretic mobility shift assay (EMSA) for detecting protein–nucleic acid interactions. Nat Protoc 2:1849–1861.
24. Geisler N, Weber K (1978) Isolation of two different homogeneous headpieces and the existence of a hinge region between residues 50 and 60 in the repressor molecule. FEBS Lett 87:215–218.
25. Jovin TM, Geisler N, Weber K (1977) Amino-terminal fragments of *Escherichia coli lac* repressor bind to DNA. Nature 269:668–672.
26. Matthews KS (1979) Tryptic core protein of lactose repressor binds operator DNA. J Biol Chem 254:3348–3353.
27. O'Gorman RB, Dunaway M, Matthews KS (1980) DNA binding characteristics of lactose repressor and the trypsin-resistant core repressor. J Biol Chem 255: 10100–10106.
28. Dunaway M, Matthews KS (1980) Hybrid tetramers of native and core lactose repressor protein: assessment of operator and nonspecific DNA binding parameters and their relationship. J Biol Chem 255:10120–10127.
29. Swint-Kruse L, Matthews KS (2009) Allostery in the LacI/GalR family: variations on a theme. Curr Opin Microbiol 12:129–137.
30. Tungtur S, Parente DJ, Swint-Kruse L (2011) Functionally important positions can comprise the majority of a protein's architecture. Proteins 79:1589–1608.
31. Tungtur S, Egan SM, Swint-Kruse L (2007) Functional consequences of exchanging domains between LacI and PurR are mediated by the intervening linker sequence. Proteins 68:375–388.
32. Bondos SE, Swint-Kruse L, Matthews KS (2015) Flexibility and disorder in gene regulation: LacI/GalR and Hox proteins. J Biol Chem 290:24669–24677.
33. Schumacher MA, Choi KY, Zalkin H, Brennan RG (1994) Crystal structure of LacI member, PurR, bound to DNA: minor groove binding by alpha helices. Science 266:763–770.
34. Schumacher MA, Choi KY, Lu F, Zalkin H, Brennan RG (1995) Mechanism of corepressor-mediated specific DNA binding by the purine repressor. Cell 83:147–155.
35. Schumacher MA, Glasfeld A, Zalkin H, Brennan RG (1997) The X-ray structure of the PurR-guanine-*purF* operator complex reveals contributions of complementary electrostatic surfaces and a water-mediated hydrogen bond to corepressor specificity and binding affinity. J Biol Chem 272:22648–22653.
36. Schumacher MA, Allen GS, Diel M, Seidel G, Hillen W, Brennan RG (2004) Structural basis for allosteric control of the transcription regulator CcpA by the phosphoprotein HPr-Ser56-P. Cell 118:731–741.
37. Arvidson DN, Lu F, Faber C, Zalkin H, Brennan RG (1998) The structure of PurR mutant L54M shows an alternative route to DNA kinking. Nat Struct Biol 5: 436–441.
38. Swint-Kruse L, Larson C, Pettitt BM, Matthews KS (2002) Fine-tuning function: correlation of hinge domain interactions with functional distinctions between LacI and PurR. Protein Sci 11:778–794.
39. Kallipolitis BH, Valentin-Hansen P (2004) A role for the interdomain linker region of the *Escherichia coli* CytR regulator in repression complex formation. J Mol Biol 342:1–7.
40. Falcon CM, Matthews KS (2000) Operator DNA sequence variation enhances high affinity binding by hinge helix mutants of lactose repressor protein. Biochemistry 39:11074–11083.
41. Zuo Z, Stormo GD (2014) High-resolution specificity from DNA sequencing highlights alternative modes of Lac repressor binding. Genetics 198:1329–1343.
42. Zuo Z, Chang Y, Stormo GD (2015) A quantitative understanding of lac repressor's binding specificity and flexibility. Quant Biol 3:69–80.
43. Sengupta R, Capp MW, Shkel IA, Record MT Jr (2017) The mechanism and high-free-energy transition state of lac repressor-lac operator interaction. Nucl Acids Res 45:12671–12680.
44. Wycuff DR, Matthews KS (2000) Generation of an AraC-*ara*BAD promoter-regulated T7 expression system. Anal Biochem 277:67–73.
45. Xu J, Liu S, Chen M, Ma J, Matthews KS (2011) Altering residues N125 and D149 impacts sugar effector binding and allosteric parameters in *Escherichia coli* lactose repressor. Biochemistry 50:9002–9013.