Insertion Mutagenesis of the *lac* Repressor and Its Implications for Structure-Function Analysis

BRYN D. NELSON,¹ COLIN MANOIL,² AND BETH TRAXLER^{1*}

Department of Microbiology¹ and Department of Genetics,²
University of Washington, Seattle, Washington 98195

Received 20 December 1996/Accepted 26 March 1997

We recently developed a simple technique for the generation of relatively large (31-codon) insertion mutations in cloned genes. To test whether the analysis of such mutations could provide insight into structure-function relationships in proteins, we examined a set of insertion mutants of the *Escherichia coli* *lac* repressor (LacI). Representatives of several LacI mutant classes were recovered, including mutants which exhibit fully active, inducer-insensitive, or weak dominant-negative phenotypes. The various properties of the recovered mutants agree with previous biophysical, biochemical, and genetic data for the protein. In particular, the results support the prior designation of mutationally tolerant spacer regions of LacI as well as proposed differences in dimerization interactions among regions of the protein core domain. These findings suggest that the analysis of 31-codon insertion mutations may provide a simple approach for characterizing structure-function relationships in proteins for which high-resolution structures are not available.

A powerful combination of biochemistry, genetics, biophysics, and molecular biology has provided fundamental insights into the complex relationships which exist among protein sequence, structure, and function. Such a combination of approaches has produced a detailed view of the way in which the *Escherichia coli* *lac* repressor (LacI) structure gives rise to the multiple functions of this homotetrameric three-domain protein (17, 23, 25, 31, 39, 48, 51). The first of these is the N-terminal headpiece domain of LacI. Interactions between headpiece domains of the tetramer allow LacI to bind specifically to *lac* operator DNA sequences and function as a transcriptional repressor of the *lac* operon (19, 21, 22, 25, 41, 42, 43). At the C-terminal end of the headpiece domain is a linker segment which joins the headpiece of each monomer to the allosteric core domain of the protein (18). This linker segment is disordered in the absence of target DNA but becomes α -helical when the DNA headpiece domain binds DNA (25).

The large core domain of the protein is composed of N and C subdomains formed from noncontiguous segments of the amino acid chain as predicted by comparison to the effector domains of periplasmic binding proteins and PurR (37, 39, 46, 48). This core domain functions mainly in protein dimerization and in binding to *lac* inducers such as allolactose and isopropyl- β -D-thiogalactopyranoside (IPTG) (16, 42, 43). IPTG binding alters the interface between the core N subdomains within the dimer, and this structural change is likely transmitted through the hinge helices to the DNA-binding headpiece domains (25, 49). A resulting reduction in the overall affinity of the LacI DNA-binding domains for operator DNA sequences effects an increase in *lac* transcription (5, 21, 33). Another linker segment joins the core domain to the third domain of LacI, a C-terminal α -helix which is largely responsible for the formation of functional LacI tetramers (2, 3, 6, 14, 16, 40).

LacI mutants disrupting each of its activities have been characterized. Many mutations occurring within the first 60 codons of *lacI* result in a dominant-negative phenotype (LacI^{-d}) since

the mutant proteins retain oligomerization and inducer binding activity but are deficient in DNA binding (1, 6, 20, 34). Certain mutations in the core domain alter responses to the inducer (LacI^s, LacI^r) (11, 23, 36, 50, 54). Finally, mutations disrupting oligomerization have helped to identify dimer and tetramer interfaces in the core domain and C-terminal α -helix domain (2, 9, 10, 14, 15, 36, 47).

Previous genetic studies examining structure-function relationships in the *lac* repressor and other proteins have exploited a variety of techniques, including gene fusions, alanine-scanning mutagenesis, combinatorial cassette mutagenesis, and more comprehensive amino acid substitution mutageneses (7, 31, 36, 44, 45, 53). Since the *lac* repressor is relatively well characterized, we have used it here as a model to test the value for structure-function studies of a new, transposon-based method for generating 31-codon insertion mutations (29).

MATERIALS AND METHODS

Bacterial strains and plasmids. The following *E. coli* K-12 strains were used in this study: CC191 [F₁₂₈ *lacI*^r Δ (*lacZ*)m15/ Δ (*ara-leu*)7697 Δ (*lac*)X74 *phoA20 galE galK thi rpsE rpoB argE*(Am) *recA1* (24)]; CSH140 [F₁₂₈ *lacI*/*ara* Δ (*gpt-lac*)5 (35)]; BN29 [CSH140 *pcnB* *zad*::Tn10 (this study)]; DHB4 [F₁₂₈ *lacI*^r/ Δ (*ara-leu*)7697 Δ (*lac*)X74 Δ *phoA* *PvuII* Δ *malF3* *phoR galE galK thi rpsL* (8)]; DHB24 [DHB4 *pcnB* *zad*::Tn10 (8)]. The starting plasmids used for this study were all derived from pTrc99A (4), which carries *lacI*^r. The pTrc99A derivatives originally contained either the *malF*, *malG*, or *malK* gene cloned into the plasmid polylinker. Following the isolation of the insertion mutations in *lacI*, these plasmids were restricted with *NcoI* and *HindIII* at the polylinker to completely remove the *mal* gene and thus simplify characterization of the *lacI* mutations. The resulting plasmids were designated *lacI*-4535, e.g., to correspond with the respective mutant *lacI* allele (e.g., *lacI*-4535::i31). In the text and in Tables 2 through 5, each mutant LacI protein expressed from this set of plasmids is denoted by a number corresponding to the amino acid after which the 31-codon insertion occurs (e.g., *lacI*-4535 expresses LacI-7) (see Table 1).

Media and cell growth. The media used have been described previously (29, 32). The following medium supplements were used at the concentrations indicated: kanamycin, 30 μ g/ml; ampicillin, 100 μ g/ml; chloramphenicol, 30 μ g/ml; 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside (X-Gal), 40 μ g/ml; sucrose, 5% (wt/vol).

DNA manipulations. Standard DNA preparations and manipulations were used (28). Insertion site DNA was sequenced by the dideoxynucleotide termination method by use of Sequenase (United States Biochemical) with double-stranded plasmid DNA templates and the appropriate oligonucleotide primer (26).

Tn*lacZ*/in-mediated insertion mutagenesis. Our protocol was based on the two-step mutagenesis procedure of Manoil and Bailey (29; see also reference 26).

* Corresponding author. Mailing address: University of Washington, Department of Microbiology, Box 357242, Seattle, WA 98195-7242. Phone: (206) 543-5485. Fax: (206) 543-8297. E-mail: btraxler@u.washington.edu.

FIG. 1. *TnlacZ/in* insertion coding sequence. An insertion of 93 bp derived from the transposon *TnlacZ/in* disrupts the wild-type sequence of the target gene (represented by italic type). Numbers 1 to 4 represent the codons of the target gene directly preceding the insertion site. At the distal end of the insertion site, a 9-bp repeat made up of sequences from codons 1 to 4 occurs, followed by the wild-type target gene sequence. Shown below each codon is the corresponding amino acid in the protein sequence.

Cells containing the *lacI* plasmid were infected with a replication-deficient lambda phage carrying the *TnlacZ/in* transposon, divided into independent pools, plated at 30°C, and grown for 36 to 40 h on L agar plus chloramphenicol, ampicillin, and sucrose to select for the transposition of the *ISlacZ/in* element. Plasmid DNA was isolated by alkaline lysis and then transformed into CC191 cells. Cells with plasmids expressing LacI fused to β -galactosidase from the *ISlacZ/in* element were identified as blue colonies on plates containing X-Gal, chloramphenicol, and ampicillin. After restriction mapping of candidate plasmids to approximate the insert location, appropriate candidates were sequenced to determine the precise insert location. Plasmids containing the *ISlacZ/in* element in *lacI* were then restricted with *Bam*HI to remove the majority of the insertion sequences and religated to produce the 93-bp insert (Fig. 1).

β -Galactosidase assays. Assays were carried out as described by Kleina and Miller (23) with a few modifications: from fresh plates, mutant strains (CSH140 or BN29 background) were grown overnight at 32°C in minimal glycerol broth, diluted into fresh minimal glycerol broth, and grown at 37°C for approximately 105 min (optical density at 600 nm, between 0.3 and 0.6). Miller units (32) were used to calculate β -galactosidase activity levels. All assays were performed at least three times. For induction assays, 10^{-2} , 10^{-3} , or 10^{-4} M IPTG was added to each culture for 30 min prior to assaying β -galactosidase levels.

Western blot (immunoblot) analysis. Western blot analysis was carried out as described by Traxler and Beckwith (52). CSH140 cells were transformed by all 18 mutant *lacI* plasmids and pTrc99A as a negative control, since wild-type LacI does not contain the 31-codon insert. Protein levels were normalized by use of the Bio-Rad protein assay, and protein samples were heated at 65°C for 20 min prior to gel loading. A sodium dodecyl sulfate–12.5% polyacrylamide gel was used to separate proteins, which were then transferred to a 0.2- μ m-pore-size nitrocellulose filter. LacI mutant proteins were detected after incubation with a 1:1,000 dilution of polyclonal antipeptide antibody directed against the 31-codon insert (29).

LacI dominance testing assay. Selected LacI mutants were tested for potential dominant-negative phenotypes by use of the β -galactosidase assay as described above. Plasmids expressing LacI-7, LacI-11, and LacI-53 were tested in strains DHB4 and DHB24; LacI-261 and wild-type LacI (pTrc99A) were both included as negative controls. Results were expressed in β -galactosidase units (32) as the average of three or more assays.

RESULTS

Randomness of *lacI* mutagenesis. We wanted to characterize the effects of *TnlacZ/in*-mediated insertions of 31 amino acids (aa) (29) on a well characterized protein, the *lac* repressor. We selected for transposition events of the *TnlacZ/in* transposon involving only the *IS50_L* element (*ISlacZ/in*), which contains both a *cat* (chloramphenicol resistance) gene and a truncated *lacZ* gene. Approximately 75,000 colonies containing *ISlacZ/in* inserts were then screened to identify potential insertion mutants of LacI. Thirty candidates were selected for sequencing based on β -galactosidase activity of the protein resulting from fusion of LacI and LacZ and on restriction mapping data. From these candidates, 18 different in-frame *lacI* insertion mutants were obtained (Table 1). Five insertion mutants were

isolated twice, and insertions at one position, represented by LacI-258, were recovered independently at least five times (the number designation of the LacI mutant refers to the amino acid of LacI after which the 31-residue insertion occurs) (Table 1). Our insertion mutagenesis of the *lacI* gene involved no screen or selection for the activity of the LacI protein, thus minimizing potential bias towards the isolation of any particular class of mutant. Additionally, the 18 different insertions were distributed throughout the *lacI* gene and fell into all three of the previously described protein domains. The lack of major hot spots, with the possible exception of the site corresponding to mutant LacI-258, suggests that *TnlacZ/in* displayed no strong insertional bias for particular *lacI* sequences.

LacI insertion mutant classifications. Miller and colleagues classified an extensive collection of LacI mutants by their ability to repress expression of *lacZ* (23, 31). We have categorized the *in vivo* activity of our set of LacI insertion mutant proteins by use of a similar classification scheme, although our mutants were expressed from a high-copy-number plasmid instead of from an *Flac* factor. In our analysis, the β -galactosidase (LacZ) level of the *lacI lacZ⁺* strain CSH140 represents the baseline. The β -galactosidase levels of CSH140 transformed with plasmid pTrc99A expressing LacI⁺ and with pTrc99A-derived plasmids expressing the different LacI insertion mutants were all compared to the baseline to determine the activity classification of each mutant.

The 18 LacI insertion mutants fell into five classes based on their relative abilities to repress *lacZ* expression (Table 2). Three mutants, LacI-152, LacI-317, and LacI-338, possessed increased repressor activity in excess of 2,000-fold that of the LacI[−] baseline and were classified as highly active (+ +). The most active of these, LacI-317 (26,400-fold increase over the LacI[−] baseline), had repression levels similar to those of the LacI⁺ control (33,400-fold). Mutants LacI-94 and LacI-340 each retained increased repressor activity in excess of 200-fold compared to the LacI[−] baseline and were considered to be active (+). Mutants LacI-67 and LacI-72 exhibited partial repressor activity (between 20-fold and 200-fold greater than LacI[−] levels) and constituted a third mutant class (+/−). Mutant LacI-142 also retained significant levels of repressor activity, about 13-fold greater than the LacI[−] baseline, and was classified as weakly active (−/+). The 10 other LacI mutants each retained less than twofold-greater repression activity above LacI[−] levels and were considered to be inactive (−).

TABLE 1. *lacI* allele insertion junction sequences

Mutant <i>lacI</i> allele (no. of isolations) ^a	Site of insertion ^b	Sequence of insertion junction ^c
<i>lacI4535::i31</i> (1)	Tyr 7	<i>ACG TTA TAC GCT GAC...ACAGCGTTA TAC GAT</i> Thr Leu Tyr Ala Asp Thr Ala Leu Tyr Asp
<i>lacI4536::i31</i> (2)	Glu 11	<i>GTC GCA GAG TCT GAC...ACAGTC GCA GAGTAT</i> Val Ala Glu Ser Asp Thr Val Ala Glu Tyr
<i>lacI4537::i31</i> (2)	Ala 53	<i>CGC GTG GCA CCT GAC...ACAGGCGTG GCACAA</i> Arg Val Ala Pro Asp Thr Gly Val Ala Gln
<i>lacI4538::i31</i> (1)	Ala 67	<i>GGC GTT GCC ACT GAC...ACAGGCGTT GCCACC</i> Gly Val Ala Thr Asp Thr Gly Val Ala Thr
<i>lacI4539::i31</i> (1)	Ala 72	<i>AGT CTG GCC CCT GAC...ACAGGT CTG GCCCTG</i> Ser Leu Ala Pro Asp Thr Gly Leu Ala Leu
<i>lacI4540::i31</i> (1)	Val 94	<i>GCC AGC GTG GCT GAC...ACAGCCAGCGTG GTG</i> Ala Ser Val Ala Asp Thr Ala Ser Val Val
<i>lacI4541::i31</i> (2)	Gly 121	<i>GTC AGT GGG CCT GAC...ACAGTC AGT GGGCTG</i> Val Ser Gly Pro Asp Thr Val Ser Gly Leu
<i>lacI4542::i31</i> (1)	Asn 142	<i>TGC ACT AAT GCT GAC...ACAGGC ACT AATGTT</i> Cys Thr Asn Ala Asp Thr Gly Thr Asn Val
<i>lacI4543::i31</i> (2)	Asp 152	<i>GTC TCT GAC CCT GAC...ACAGTC TCT GACCAG</i> Val Ser Asp Pro Asp Thr Val Ser Asp Gln
<i>lacI4544::i31</i> (1)	Ile 160	<i>AGT ATT ATT TCT GAC...ACAGGT ATT ATT TTC</i> Ser Ile Ile Ser Asp Thr Gly Ile Ile Phe
<i>lacI4545::i31</i> (1)	Arg 197	<i>CGT CTG CGT CCT GAC...ACAGGT CTG CGT CTG</i> Arg Leu Arg Pro Asp Thr Gly Leu Arg Leu
<i>lacI4546::i31</i> (5)	Thr 258	<i>GCC ATT ACC GCT GAC...ACAGCCATT ACCGAG</i> Ala Ile Thr Ala Asp Thr Ala Ile Thr Glu
<i>lacI4547::i31</i> (1)	Gly 261	<i>GAG TCC GGG CCT GAC...ACAGAGTCC GGGCTG</i> Glu Ser Gly Pro Asp Thr Glu Ser Gly Leu
<i>lacI4548::i31</i> (2)	Cys 281	<i>AGC TCA TGT TCT GAC...ACAGGCTCA TGT TAT</i> Ser Ser Cys Ser Asp Thr Gly Ser Cys Tyr
<i>lacI4549::i31</i> (1)	Gly 297	<i>CTG CTG GGG CCT GAC...ACAGTG CTG GGGCAA</i> Leu Leu Gly Pro Asp Thr Val Leu Gly Gln
<i>lacI4550::i31</i> (1)	Gln 317	<i>GGC AAT CAG CCT GAC...ACAGGCAAT CAGCTG</i> Gly Asn Gln Pro Asp Thr Gly Asn Gln Leu
<i>lacI4551::i31</i> (1)	Ser 338	<i>ACC GCC TCT CCT GAC...ACAGCCGCCTCT CCC</i> Thr Ala Ser Pro Asp Thr Ala Ala Ser Pro
<i>lacI4552::i31</i> (1)	Arg 340	<i>TCT CCC CGC GCT GAC...ACAGCT CCCCGCGCG</i> Ser Pro Arg Ala Asp Thr Ala Pro Arg Ala

^a Mutant *lacI* alleles are listed in order of 31-codon insertion locations. Numbers in parentheses indicate the number of independent isolations of each mutant allele, based on restriction mapping and direct sequencing.

^b The amino acid of the LacI sequence directly preceding each 31-aa insertion site is given.

^c Nucleotides shown in italic type represent the wild-type *lacI* sequence, whereas nucleotides shown in boldface type represent the *Tn*lacZ/in-derived insertion sequence (see also Fig. 1). Shown below each codon is the corresponding amino acid.

Oehler and coworkers (40) reported that a mutant *lac* repressor protein capable of forming only dimers retained greater than 50% of wild-type LacI activity when assayed at high intracellular levels. At low protein levels, however, the same LacI mutant retained less than 2% of wild-type repressor activity. Based on these observations, we decided to assay the repression activities of mutants LacI-338 and LacI-340 at low intracellular concentrations by using a *pcnB* strain which reduces plasmid copy number approximately 10-fold (27). The insertions in mutants LacI-338 and LacI-340 both map near the C-terminal α -helix tetramerization domain of LacI; LacI-338 also retained greater than 50% of wild-type activity under high-plasmid-copy number conditions (Table 2). Under low-plasmid-copy number conditions, LacI-338 behaved similarly to the tetramerization-defective mutant in the Oehler et al. study (40), retaining only about 2% of wild-type LacI repression levels. Under these same conditions, LacI-340 retained

slightly less than 2% of wild-type activity or about twofold-greater activity than that of the LacI⁻ control (Table 2).

Western blot analysis. We determined the relative protein expression levels of the 18 LacI insertion mutants by Western blot analysis (Fig. 2). Two prominent bands were observed in most samples, i.e., the LacI protein band migrating at about 49 kDa and a background band migrating at about 33.3 kDa. All but two of the inactive mutants (LacI-11 and LacI-197) accumulated protein to detectable levels, indicating that the loss of repressor activity was not due merely to an absence of protein in the cells. Inactive mutant LacI-53, whose insertion localizes to the linker helix between the DNA-binding domain and the core domain of LacI, produced two protein bands whose sizes are consistent with full-length and N-terminally truncated LacI (the latter likely resulting from proteolysis at the mutated linker region). Earlier reports have indicated that this linker is particularly susceptible to proteolysis (18, 42).

TABLE 2. Repression activity of LacI insertion mutants

Mutant ^a	Relative protein levels ^b	Repression activity ^c		Activity class ^d
		Avg	SEM	
LacI ⁺	NA	33,400	6,210	++
LacI ⁺ <i>pcnB</i>	NA	145	14.2	NA
LacI-7	++	1.01	0.013	—
LacI-11	—	0.943	0.010	—
LacI-53	++	0.904	0.036	—
LacI-67	+	67.4	1.65	+/-
LacI-72	+	33.8	0.812	+/-
LacI-94	+	506	37.1	+
LacI-121	+	0.925	0.034	—
LacI-142	++	13.3	0.286	-/+
LacI-152	+++	2,210	644	++
LacI-160	+	0.891	0.040	—
LacI-197	—	0.970	0.014	—
LacI-258	+	0.964	0.022	—
LacI-261	+	1.23	0.038	—
LacI-281	++	1.40	0.006	—
LacI-297	+	0.994	0.009	—
LacI-317	+++	26,400	4,600	++
LacI-338	+++	19,700	7,040	++
LacI-338 <i>pcnB</i>	NT	3.26	0.148	NA
LacI-340	+++	308	100	+
LacI-340 <i>pcnB</i>	NT	2.07	0.063	NA
LacI ⁻	NA	1	0.023	—
LacI ⁻ <i>pcnB</i>	NA	1	0.014	NA

^a All mutant LacI proteins were expressed from pTrc99A derivatives in the CSH140 background, and selected mutants were expressed in the BN29 (CSH140 *pcnB*) background, designated *pcnB* in the table. LacI⁺ was expressed from pTrc99A, and LacI⁻ represents untransformed CSH140 or BN29.

^b Relative protein level classifications are taken from Western blot analysis (Fig. 2). Symbols and abbreviations: —, no detectable protein; +++, most prominent protein band; NA, not applicable; NT, not tested. Relative protein levels were not seen for wild-type LacI since the 31-codon insertion antibody was used for protein detection.

^c Repression activity was calculated by dividing the β -galactosidase units (32) expressed from the CSH140 (or BN29) strain alone by the β -galactosidase units expressed by each mutant strain. Values obtained from multiple independent trials were then averaged to arrive at the final value for each mutant LacI protein. β -Galactosidase activity for CSH140 expressing LacI⁺ from pTrc99A ranged from 0.070 to 0.456 U, and β -galactosidase activity for CSH140 alone ranged from 4,600 to 8,230 u.

^d Activity classifications were based on repression levels (see Results for explanation of each activity class). NA, not applicable in a low-copy-number background since mutant activity classes were assigned based on repression levels in a high-plasmid-copy-number background.

Different mobilities among several of the mutant LacI proteins may reflect variations in the amount of residual structure present in the mutant proteins during sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Such mobility differences have also been observed among similar 31-residue insertion mutants of MalK and MalG (26, 37a). The mobility variations of the mutant proteins are not attributable to transposon-mediated rearrangements in *lacI* since restriction digest analysis confirmed the expected size and location of each *lacI* insertion (data not shown).

Inducer effects. Several of the active insertion mutants have insertion junctions in regions of LacI previously identified as important for proper inducer response. Mutant LacI-94 maps to a region (residues 95 to 97) which is believed to be responsible for transmitting the allosteric signal from the N-core subdomain to the DNA-binding head domain upon inducer binding (25, 51) (Fig. 3). Mutants LacI-67 and LacI-72 map to another region (residues 63 to 84) near the inducer binding site in which multiple LacI^s (or inducer-insensitive) mutants have been recovered (23, 25, 51) (Fig. 3).

We assayed the repressor activities of the potential inducer-

insensitive mutants LacI-67, LacI-72, and LacI-94 in the presence and absence of the inducer IPTG. Whereas the strain expressing mutant LacI-94 allowed a greater-than-100-fold increase in β -galactosidase levels when induced with 10^{-3} M IPTG, insertion mutant LacI-72 was unresponsive to three different concentrations of IPTG, as demonstrated by a relatively constant induction value near 1.00 (Table 3). LacI-67 also possessed an altered phenotype, as no increase in β -galactosidase levels was observed upon induction with three different IPTG concentrations; instead, a small but reproducible decrease in β -galactosidase expression occurred under all induction conditions tested (Table 3). When the induction period was extended to 105 min for 10^{-3} M IPTG, the induction value for LacI-67 decreased further to 0.384, whereas the LacI⁺ and LacI⁻ controls had induction values of 642 and 0.961, respectively. All other active insertion mutants responded positively to induction with 10^{-3} M IPTG for 30 min, as β -galactosidase units increased by a factor of 5-fold (for LacI-142) to about 350-fold (for LacI-317).

Dominance analysis. The inactive mutants LacI-7, LacI-11, and LacI-53 all map to the N-terminal headpiece domain of LacI. When tested at a high copy number for dominant-negative effects in a *lacI*⁺ background, LacI-7 and LacI-11 had almost no effect whereas LacI-53 was found to exhibit a weak dominant-negative phenotype (LacI^{-wd}; about a 10-fold increase in β -galactosidase expression compared to that of DHB4/pTrc99A) (Table 4). No dominant effect was evident, however, when LacI-53 was assayed in an isogenic *pcnB* strain (DHB24) in which the *lacI* plasmid copy number was reduced roughly 10-fold (27).

DISCUSSION

A previous mutagenesis strategy involving relatively large insertions (average net insertion of 16 to 17 residues) has been used successfully to study the *E. coli* integral outer membrane protein LamB (12, 13, 38). We used a simple new genetic approach to randomly generate 18 in-frame insertions of 31 residues in the *lac* repressor protein and analyzed the resulting mutant phenotypes. We reasoned that this characterization could be used to correlate genetic, biochemical, and structural data for the protein. Active and inactive mutants of the protein were recovered and characterized for their abilities to fold into a stable protein, repress *lac* expression, oligomerize, and bind inducer. We obtained representatives of several previously characterized mutant classes, including LacI^{-d} (dominant-negative) and LacI^s (inducer-insensitive) mutants. In addition, we assigned our LacI mutant proteins to five different classes based on repression activity phenotype (summarized in Table 5). Below, we discuss several of the most interesting correlations that can be made between these LacI mutants and the recent biochemical, genetic, and crystal structure analyses (summarized in references 17, 23, 25, 31, 39, and 51).

FIG. 2. Western blot analysis of LacI mutant proteins. Proteins were detected by incubation with antibody directed against the 31-aa insert, with wild-type LacI expressed from pTrc99A as the negative control. The background band of approximately 33 kDa appeared in all lanes.

FIG. 3. Stereodigram representation of LacI showing the dimer interface between two protein monomers. Residues 61 to 356 of each monomer (excluding the N-terminal DNA-binding domain) were illustrated with Insight II (1995 version) based on the crystallographic data of Friedman et al. (17). The darkened residue side chains on the monomer to the right represent the amino acid directly preceding each insertion event and are labeled as such with the single-letter amino acid designation (e.g., A67 represents the insertion location of mutant LacI-67; see Table 1 for sequential context). Boldface underlined labels indicate active repressor mutants, whereas lightface labels without underlining indicate inactive mutants. Dashed lines indicate that the mutants were active in high intracellular concentrations but only very weakly active in low intracellular concentrations.

Insertion mutant LacI-53 has its 31-residue insert located in the hinge helix of the DNA-binding (headpiece) domain which links the headpiece to the core domain of LacI. The weak dominant-negative effect of LacI-53 at high intracellular concentrations agrees with earlier mapping of LacI^{-d} mutants which places the mutations within the first 60 aa of the protein (1, 6, 20, 34). The lack of dominance by this mutant at low intracellular concentrations may reflect partial folding or oligomerization interference of the core domain, thus favoring oligomerization among wild-type monomers.

Mutants LacI-67, LacI-72, and LacI-94 have insert junctions mapping to a region of the N subdomain of the LacI core implicated in both dimer formation and inducer binding (17, 25, 51) (Fig. 3). The 31-residue inserts in these three mutant proteins evidently interfere with oligomerization only partially, since each mutant retained significant levels of repression activity, especially LacI-94. In addition, mutants LacI-67 and LacI-72 possessed altered phenotypes in the presence of inducer. The insertion mutation of LacI-72 is positioned at the lip of the inducer-binding region of the LacI core domain and may directly interfere with this binding process by disrupting a hydrophobic contact normally mediated by Leu-73 (Table 1) (17). The repression capacity of LacI-67 not only was maintained upon induction with IPTG but also increased slightly. This aberrant behavior may be used to classify the mutant as inducer-insensitive (LacI^s) or even as a weak LacI^r mutant (reverse-curve) (11, 23; reviewed in reference 33).

The observations (i) that all three of these mutants possess significant repressor activity, (ii) that LacI-72 and LacI-67 display altered inducer binding phenotypes, and (iii) that inactive mutants LacI-258, LacI-261, and LacI-281 all map to the dimerization interface of the core C subdomain (see below) agree with the modified Venus's-flytrap model (30) suggested by Suckow et al. (51). This model proposes that specific dimer-

TABLE 3. Effect of IPTG induction on selected LacI mutants

Mutant ^a	IPTG concn (M)	Induction ^b	
		Avg	SEM
LacI ⁺	10 ⁻⁴	23.4	9.67
	10 ⁻³	410	62.9
	10 ⁻²	82.6	36.9
LacI-53	10 ⁻³	0.944	0.091
LacI-67	10 ⁻⁴	0.820	0.014
	10 ⁻³	0.654	0.018
	10 ⁻²	0.738	0.033
LacI-72	10 ⁻⁴	1.26	0.107
	10 ⁻³	1.11	0.053
	10 ⁻²	1.10	0.091
LacI-94	10 ⁻³	113	7.86
LacI-142	10 ⁻³	5.05	0.209
LacI-152	10 ⁻³	17.8	1.85
LacI-317	10 ⁻³	355	72.3
LacI-338	10 ⁻³	219	27.2
LacI-340	10 ⁻³	10.2	4.06
LacI ⁻	10 ⁻⁴	1.00	0.025
	10 ⁻³	0.971	0.020
	10 ⁻²	1.00	0.031

^a See Table 1, footnote a.

^b Cultures were induced for 30 min in different concentrations of IPTG at 37°C prior to termination of growth. Induction values were calculated by dividing the β-galactosidase units for an induced sample by the β-galactosidase units for the corresponding uninduced sample. All values represent the average of multiple independent assays. LacI-53 (LacI⁻) and CSH140 alone (LacI⁻) were both included as controls since β-galactosidase units would be expected to be maximal regardless of IPTG induction; the expected fold induction value in each case should be near 1.00. Competitive inhibition of the β-galactosidase assay by IPTG concentrations of 10⁻² M was corrected for by setting the 10⁻² M LacI⁻ induction value at 1.00 and dividing the other relevant induction values by this number.

TABLE 4. Dominant-negative phenotype analysis

Mutant tested	Strain background ^a	β-Gal level ^b	
		Avg	SEM
LacI ⁺ (pTrc99A)	DHB4 (<i>lacI</i> ^q)	7.18	0.123
LacI-7	DHB4	12.7	0.318
LacI-11	DHB4	11.8	0.536
LacI-53	DHB4	70.9	14.8
LacI-261	DHB4	8.88	0.192
None	DHB4	8.23	0.082
LacI ⁺ (pTrc99A)	DHB24 (<i>lacI</i> ^q <i>pcnB</i>)	11.6	0.608
LacI-53	DHB24	12.7	0.231
None	DHB24	12.0	0.285

^a DHB4/pTrc99A possesses the genotype *lacI*^q/*lacI*^q, and DHB24/pTrc99A possesses the genotype *lacI*^q *pcnB*/*lacI*^q. Mutant LacI-261 was expressed in the DHB4 background as an additional negative control.

^b β-Gal, β-galactosidase. Miller (32) β-galactosidase units were used. All values represent the average of multiple independent assays.

ization of the LacI monomers is mediated mostly by C-core subdomain sequences independent of inducer binding, whereas N-core subdomains are involved in structural rearrangements primarily during induction, similar to that seen for PurR (49).

An extensive series of amino acid substitutions through residue 329 of LacI and a more directed alanine replacement mutagenesis of the protein identified six stretches of mutationally tolerant amino acids, each 5 to 8 residues in length (31). X-ray crystallography data position four of these tolerant regions at or near the surface of the N-core subdomain and two at relatively surface-exposed regions in the C-core subdomain of LacI (17, 25). These regions are thought to serve as spacers which may connect more conserved hydrophobic residues important for protein folding and oligomerization (31, 51). Active 31-residue insertion mutants LacI-142 (in the 140- to 144-aa region), LacI-152 (in the 151- to 158-aa region), and LacI-317 (in the 311- to 318-aa region) each map to spacer regions in the N-terminal core subdomain (Fig. 3). Our results suggest that such surface-exposed spacer regions may be identified, in at least some cases, by their ability to accommodate large insertions.

Inactive mutants LacI-258, LacI-261, and LacI-281 all have insert junctions at relatively surface-exposed regions of the monomer. These three insertions, however, are contained in a region which is particularly intolerant of amino acid substitutions (31) and which has been implicated in dimer formation (specifically, residues 250 to 260 and 275 to 285) (9, 15, 25, 47, 51) (Fig. 3). The insertion junctions of two other inactive LacI mutants, LacI-197 and LacI-297, both interrupt relatively surface-exposed α-helices of the core domain. In each case, the inserted sequence begins with a proline residue (Table 1) which likely prevents the completion of the helical structure by ensuing residues of the insertion. In the case of LacI-197, the resulting mutant protein was destabilized to the extent that no protein was detectable on a Western blot (Fig. 2).

The insertion junction of LacI-338 lies in a linker region between the C-core subdomain and the C-terminal tetramerization α-helix of LacI (Fig. 3). The insertion junction of mutant LacI-340 is C-terminal to that of LacI-338 by only two residues and directly precedes a series of leucine heptad repeats which are critical for LacI tetramerization (2, 3, 14). Both mutants displayed significant repression activity in high intracellular concentrations, especially LacI-338. When assayed for activity in low intracellular concentrations, however, each mutant retained only about 2% of wild-type repression levels and

only twofold to threefold-higher activity than the LacI⁺ background. This dramatic decrease in repression activity agrees with earlier studies on tetramerization-defective mutants (40) and suggests that both LacI-338 and LacI-340 are severely impaired in their ability to form tetramers, resulting in a cellular population of low-activity dimers. LacI tetramerization in these mutants may be hampered by instability or misalignment of the C-terminal α-helix; the slight but reproducible difference in the repression activities of LacI-338 and LacI-340 perhaps relates to the respective proximities of their insertion junctions to the α-helix.

Use of insertion mutagenesis in analyzing protein structure and function. Transposon-mediated insertion mutagenesis has demonstrated several key features which make it an attractive method for protein analysis. This mutagenesis has provided a unique in vivo view of the contributions of different protein regions to stability, oligomerization, and function for a well-characterized protein and thus presents an alternative to other genetic techniques for complementing structural data. This approach has independently verified several important structural and functional aspects of the well-characterized *lac* repressor and should be useful for examining similar aspects of poorly characterized proteins and proteins not amenable to structural analysis by other methods. As a tool, it can highlight protein regions where more precise mutagenic analyses may be used to refine key elements of protein structure and function.

TABLE 5. Summary of mutant LacI phenotypes

LacI mutant	LacI phenotype ^a	Protein region and functions affected ^b
LacI-7	—	Headpiece α-helix; HTH motif stability
LacI-11	—	Headpiece α-helix; HTH motif stability
LacI-53	— ^{wd}	Hinge α-helix involved in DNA contacts, joining headpiece to core domain
LacI-67	+/- ^s or +/- ^{wr}	N-subdomain dimer interface (63–84); inducer binding site
LacI-72	+/- ^s	N-subdomain dimer interface (63–84); inducer binding site—Leu-73 important
LacI-94	+	N-subdomain dimer interface (95–97); allosteric signal transfer during induction
LacI-121	—	Buried N-subdomain β-sheet
LacI-142	-/+	N-subdomain spacer (140–144)
LacI-152	++	N-subdomain spacer (151–158)
LacI-160	—	Linker between N- and C-core subdomains
LacI-197	—	Inducer binding site—Arg-197 important; C-subdomain α-helix
LacI-258	—	C-subdomain dimer interface (250–260)
LacI-261	—	Near C-subdomain dimer interface (250–260)
LacI-281	—	C-subdomain dimer interface (275–285)—Tyr-282 important
LacI-297	—	N-subdomain α-helix
LacI-317	++	N-subdomain spacer (311–318)
LacI-338 ^c	++	Linker/COOH terminus tetramerization domain
LacI-340 ^c	+	COOH terminus tetramerization domain

^a Symbols: ++, repression activity over 2,000-fold above baseline levels (see Table 2 and Results); +, 200- to 2,000-fold increase; +/-, 20- to 200-fold increase; -/+, 2- to 20-fold increase; —, less than 2-fold increase. Superscript s indicates an inducer-resistant phenotype, superscript wd indicates a weak dominant effect, and superscript wr indicates weak reverse-curve phenotype.

^b HTH, helix-turn-helix. Numbers in parentheses refer to relevant stretches of amino acids; positional information was gathered from references 17, 25, 31, 39, and 51.

^c LacI-338 and LacI-340 are both classified as active at high intracellular concentrations, although they are only very weakly active at low intracellular concentrations (see Discussion).

ACKNOWLEDGMENTS

B.D.N. was supported in part by Molecular and Cellular Biology Training grant PHS NRSA T32 GM07270 from NIGMS. This work was additionally supported in part by NSF grant MCB 9306752 to B.T. and by NSF grant MCB 9506989 and NIH grant GM46493 to C.M.

We acknowledge Lawrence Schaufler for generous assistance with the structural analysis. In addition, we thank Eli Gachelet and Fred Taub for technical assistance and Jim Champoux and Peter Schatz for encouragement and helpful comments.

REFERENCES

- Adler, K., K. Beyreuther, E. Fanning, N. Geisler, B. Gronenborn, A. Klemm, B. Müller-Hill, M. Pfahl, and A. Schmitz. 1972. How *lac* repressor binds to DNA. *Nature* (London) **237**:322–327.
- Alberti, S., S. Oehler, B. von Wilcken-Bergmann, H. Krämer, and B. Müller-Hill. 1991. Dimer to tetramer assembly of *lac* repressor involves a leucine heptad repeat. *New Biol.* **3**:57–62.
- Alberti, S., S. Oehler, B. von Wilcken-Bergmann, and B. Müller-Hill. 1993. Genetic analysis of the leucine heptad repeats of *lac* repressor: evidence for a 4-helical bundle. *EMBO J.* **12**:3227–3236.
- Amann, E., B. Ochs, and K. J. Abel. 1988. Tightly regulated *tac* promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli*. *Gene* **69**:301–315.
- Barkley, M. D., A. D. Riggs, A. Jobe, and S. Bourgeois. 1975. Interaction of effecting ligands with *lac* repressor and repressor-operating complex. *Biochemistry* **14**:1700–1712.
- Betz, J. L. 1986. Cloning and characterization of several dominant-negative and tight-binding mutants of *lac* repressor. *Gene* **42**:283–292.
- Blaber, M., W. A. Baase, N. Gassner, and B. W. Matthews. 1995. Alanine scanning mutagenesis of the α -helix 115–123 of phage T4 lysozyme: effects on structure, stability and the binding of solvent. *J. Mol. Biol.* **246**:317–330.
- Boyd, D., C. Manoel, and J. Beckwith. 1987. Determinants of membrane protein topology. *Proc. Natl. Acad. Sci. USA* **84**:8525–8529.
- Chakerian, A., and K. S. Matthews. 1991. Characterization of mutations in oligomerization domain of *lac* repressor protein. *J. Biol. Chem.* **266**:22206–22214.
- Chakerian, A. E., V. M. Tesmer, S. P. Manly, J. K. Brackett, M. J. Lynch, J. T. Hoh, and K. S. Matthews. 1991. Evidence for leucine zipper motif in lactose repressor protein. *J. Biol. Chem.* **266**:1371–1374.
- Chamness, G. C., and C. D. Wilson. 1970. An unusual *lac* repressor mutant. *J. Mol. Biol.* **53**:561–565.
- Charbit, A., J. C. Boulain, A. Ryter, and M. Hofnung. 1986. Probing the topology of a bacterial membrane protein by genetic insertion of a foreign epitope: expression at the cell surface. *EMBO J.* **5**:3029–3037.
- Charbit, A., J. Ronco, V. Michel, C. Werts, and M. Hofnung. 1991. Permissive sites and topology of an outer membrane protein with a reporter epitope. *J. Bacteriol.* **173**:262–275.
- Chen, J., and K. S. Matthews. 1992. Deletion of lactose repressor carboxyl-terminal domain affects tetramer formation. *J. Biol. Chem.* **267**:13843–13850.
- Daly, T. J., and K. S. Matthews. 1986. Characterization and modification of a monomeric mutant of the lactose repressor protein. *Biochemistry* **25**:5474–5478.
- Files, J. G., and K. Weber. 1976. Limited proteolytic digestion of *lac* repressor by trypsin. Chemical nature of the resulting trypsin-resistant core. *J. Biol. Chem.* **251**:3386–3391.
- Friedman, A. M., T. O. Fischmann, and T. A. Steitz. 1995. Crystal structure of *lac* repressor core tetramer and its implications for DNA looping. *Science* **268**:1721–1727.
- Geisler, N., and K. Weber. 1978. *Escherichia coli* lactose repressor. 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.
- Gilbert, W., and B. Müller-Hill. 1967. The *lac* operator is DNA. *Proc. Natl. Acad. Sci. USA* **58**:2415–2421.
- Gordon, A. J. E., P. A. Burns, D. F. Fix, F. Yatagai, F. L. Allen, M. J. Horsfall, J. A. Halliday, J. Gray, C. Bernelot-Moens, and B. W. Glickman. 1988. Missense mutation in the *lacI* gene of *Escherichia coli*. Inferences on the structure of the repressor protein. *J. Mol. Biol.* **200**:239–251.
- Jacob, F., and J. Monod. 1961. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **3**:318–356.
- Kaptein, R., R. M. J. N. Lamerichs, R. Boelens, J. Anton, and C. Rullman. 1990. Two-dimensional NMR study of a protein-DNA complex: *lac* repressor headpiece-operator interaction. *Biochem. Pharmacol.* **40**:89–96.
- Kleina, L. G., and J. H. Miller. 1990. Genetic studies of the *lac* repressor. XIII. Extensive amino acid replacements generated by the use of natural and synthetic nonsense suppressors. *J. Mol. Biol.* **212**:295–318.
- Lee, E., and C. Manoel. 1994. Mutations eliminating the protein export function of a membrane-spanning sequence. *J. Biol. Chem.* **269**:28822–28828.
- Lewis, M., G. Chang, N. C. Horton, M. A. Kercher, H. C. Pace, M. Schumacher, R. G. Brennan, and P. Lu. 1996. Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science* **271**:1247–1254.
- Lippincott, J., and B. Traxler. 1997. MalFGK complex assembly and transport and regulatory characteristics of MalK insertion mutations. *J. Bacteriol.* **179**:1337–1343.
- Lopilato, J., S. Bortner, and J. Beckwith. 1986. Mutations in a new chromosomal gene of *Escherichia coli* K-12, *pcnB*, reduce plasmid copy number of pBR322 and its derivatives. *Mol. Gen. Genet.* **205**:285–290.
- Maniatis, T., E. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Manoil, C., and J. Bailey. 1997. A simple screen for permissive sites in proteins. Analysis of *Escherichia coli lac* permease. *J. Mol. Biol.* **267**:250–263.
- Mao, B., M. R. Pear, J. A. McCammon, and F. A. Quiocho. 1982. Hinge-bending in L-arabinose-binding protein. The Venus-flytrap model. *J. Biol. Chem.* **257**:1131–1133.
- Markiewicz, P., L. G. Kleina, C. Cruz, S. Ehret, and J. H. Miller. 1994. Genetic studies of the *lac* repressor. XIV. Analysis of 4000 altered *Escherichia coli lac* repressors reveals essential and non-essential residues, as well as 'spacers' which do not require a specific sequence. *J. Mol. Biol.* **240**:421–433.
- Miller, J. H. 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Miller, J. H. 1980. A review, p. 31–88. In J. H. Miller and W. S. Reznikoff (ed.), The operon. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Miller, J. H. 1984. Genetic studies of the *lac* repressor. XII. Amino acid replacements in the DNA binding domain of the *Escherichia coli lac* repressor. *J. Mol. Biol.* **180**:205–212.
- Miller, J. H. 1992. A short course in bacterial genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Miller, J. H., C. Coulondre, M. Hofer, U. Schmeissner, H. Sommer, A. Schmitz, and P. Lu. 1979. Genetic studies of the *lac* repressor. IX. Generation of altered proteins by the suppression of nonsense mutations. *J. Mol. Biol.* **131**:191–222.
- Mowbray, S. L., and L. B. Cole. 1992. 1.7 Å X-ray structure of the periplasmic ribose receptor from *Escherichia coli*. *J. Mol. Biol.* **225**:155–175.
- Nelson, B. D., and B. Traxler. Unpublished data.
- Newton, S. M. C., P. E. Klebba, V. Michel, M. Hofnung, and A. Charbit. 1996. Topology of the membrane protein LamB by epitope tagging and a comparison with the X-ray model. *J. Bacteriol.* **178**:3447–3456.
- Nichols, J. C., N. K. Vyas, F. A. Quiocho, and K. S. Matthews. 1993. Model of lactose repressor core based on alignment with sugar-binding proteins is concordant with genetic and chemical data. *J. Biol. Chem.* **268**:17602–17612.
- Oehler, S., E. R. Eismann, H. Krämer, and B. Müller-Hill. 1990. The three operators of the *lac* operon cooperate in repression. *EMBO J.* **9**:973–979.
- Ogata, R. T., and W. Gilbert. 1978. An amino-terminal fragment of *lac* repressor binds specifically to *lac* operator. *Proc. Natl. Acad. Sci. USA* **75**:5851–5854.
- Platt, T., J. G. Files, and K. Weber. 1973. *lac* repressor: specific proteolytic destruction of the NH₂-terminal region and loss of the deoxyribonucleic acid-binding activity. *J. Biol. Chem.* **248**:110–121.
- Platt, T., K. Weber, D. Ganem, and J. H. Miller. 1972. Translational restarts: AUG reinitiation of a *lac* repressor fragment. *Proc. Natl. Acad. Sci. USA* **69**:897–901.
- Reidhaar-Olson, J. F., and R. T. Sauer. 1988. Combinatorial cassette mutagenesis as a probe of the informational content of protein sequences. *Science* **241**:53–57.
- Rennell, D., S. E. Bouvier, L. W. Hardy, and A. R. Poteete. 1991. Systematic mutation of bacteriophage T4 lysozyme. *J. Mol. Biol.* **222**:67–88.
- Sams, C. F., N. K. Vyas, F. A. Quiocho, and K. S. Matthews. 1984. Predicted structure of the sugar-binding site of the *lac* repressor. *Nature* **310**:429–430.
- Schmitz, A., V. Schmeissner, and J. H. Miller. 1976. Mutations affecting the quaternary structure of the *lac* repressor. *J. Biol. Chem.* **251**:3359–3366.
- Schumacher, M. A., K. Y. Choi, H. Zalkin, and R. G. Brennan. 1994. Crystal structure of LacI member, PurR, bound to DNA: minor groove binding by α helices. *Science* **266**:763–770.
- Schumacher, M. A., K. Y. Choi, F. Lu, H. Zalkin, and R. G. Brennan. 1995. Mechanism of corepressor-mediated specific DNA binding by the purine repressor. *Cell* **83**:147–155.
- Spotts, R. O., A. E. Chakerian, and K. S. Matthews. 1991. Arginine 197 of *lac* repressor contributes significant energy to inducer binding. *J. Biol. Chem.* **266**:22998–23002.

51. **Suckow, J., P. Markiewicz, L. G. Kleina, J. Miller, B. Kisters-Woike, and B. Müller-Hill.** 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.
52. **Traxler, B., and J. Beckwith.** 1992. Assembly of a hetero-oligomeric membrane protein complex. *Proc. Natl. Acad. Sci. USA* **89**:10852–10856.
53. **Wales, M. E., and J. R. Wild.** 1991. Analysis of structure-function relationships by formation of chimeric enzymes produced by gene fusion. *Methods Enzymol.* **202**:687–706.
54. **Willson, C., D. Perrin, M. Cohn, F. Jacob, and J. Monod.** 1964. Non-inducible mutants of the regulator gene in the lactose system of *Escherichia coli*. *J. Mol. Biol.* **8**:582–592.