Crystal Structure of the Lactose Operon Repressor and Its Complexes with DNA and Inducer

Mitchell Lewis,* Geoffrey Chang, Nancy C. Horton,† Michele A. Kercher, Helen C. Pace, Maria A. Schumacher, Richard G. Brennan, Ponzy Lu

The *lac* operon of *Escherichia coli* is the paradigm for gene regulation. Its key component is the *lac* repressor, a product of the *lacI* gene. The three-dimensional structures of the intact *lac* repressor, the *lac* repressor bound to the gratuitous inducer isopropyl-β-D-1-thiogalactoside (IPTG) and the *lac* repressor complexed with a 21–base pair symmetric operator DNA have been determined. These three structures show the conformation of the molecule in both the induced and repressed states and provide a framework for understanding a wealth of biochemical and genetic information. The DNA sequence of the *lac* operon has three *lac* repressor recognition sites in a stretch of 500 base pairs. The crystallographic structure of the complex with DNA suggests that the tetrameric repressor functions synergistically with catabolite gene activator protein (CAP) and participates in the quaternary formation of repression loops in which one tetrameric repressor interacts simultaneously with two sites on the genomic DNA.

More than 30 years ago, Jacob and Monod (*1*) introduced the *E. coli* lactose operon as a model for gene regulation. The model persists as a cogent depiction of how a set of structural genes may be coordinately transcribed or repressed depending upon the concentration of metabolites in the growth medium. The product of the *i* gene, the repressor molecule, binds to a specific DNA sequence and functions as a molecular switch in response to inducer molecules. In the absence of extracellular lactose, the repressor binds tightly to the operator DNA and prevents transcription of *lacZ*, *lacY*, and *lacA*, which code for β-galactosidase, *lac* permease, and a transacetylase, respectively. Conversely, when lactose is present in the medium, the repressor dissociates from the operator, allowing transcription of the structural genes.

The *lac* repressor is a protein (*2*) of 360 amino acids that associates into a homotetramer of 154,520 dalton molecular mass (*3*, *4*). This tetramer can be cleaved by limited protease digestion into five distinct fragments: four NH_2-terminal fragments (about 60 residues) that bind specifically to operator DNA; and a COOH-terminal tetrameric core that binds inducers and their analogs (*5*). The last 30 amino acids of the core are responsible for the oligomeric state of the repressor and mediate tetramerization (*6*–*8*). The operator of the *lac* operon was originally identified from cis-acting constitutive (o^c) mutants (*1*). A 27–base pair (bp) DNA fragment, which encompasses these o^c mutations and is protected from deoxyribonuclease digestion by bound *lac* repressor, is centered 11-bp downstream from the start of transcription of the *lacZ* gene (*9*). This operator has an axis of approximate dyad symmetry, a feature common to many transcriptional regulatory sites. The *lac* repressor binds with 10 times greater affinity to a palindrome of the left half of the operator, which lacks the central base pair (*10*, *11*). The natural inducer of the repressor molecule is allolactose, an analog of lactose created by a side reaction of β-galactosidase (*12*). However, a gratuitous inducer, isopropyl-β-D-1-thiogalactoside (IPTG), that is not a substrate for β-galactosidase, can be used to turn on transcription of the lactose operon (*13*).

In their first paper on allostery, Monod, Changeaux, and Jacob (*14*) presented a unified view of conformational transitions in multimeric proteins. They considered the possibility that the affinity of the repressor for operator DNA could be regulated by allosteric effectors. An inducer or a co-repressor would cause a change in the conformation of the repressor and interfere with its binding to operator DNA. When crystals of the *lac* repressor–*lac* operator complex are exposed to the allosteric effector IPTG, they immediately shatter (*15*), reminiscent of the shattering of crystals of deoxyhemoglobin when exposed to air (*16*, *17*).

Control of the *lac* operon is complex and involves positive activation by a cyclic AMP–dependent catabolite gene activator protein (CAP). At high concentrations of cyclic AMP, the CAP protein raises the amount of transcription by binding a recognition site on the DNA adjacent to RNA polymerase and thereby the affinity of RNA polymerase for the promoter is increased (*18*). In addition to the principal operator O_1, the *lac* operon has two auxiliary operators, O_2 and O_3, which were discovered after the *lac* repressor was isolated and its operator DNA was sequenced (*19*). The center of the second operator, O_2, is located 401 bp downstream from the primary operator in the *lacZ* gene. The third operator, O_3, overlaps the COOH-terminal coding sequence for the *lacI* gene and is centered 93 bp upstream from O_1 (Fig. 1). All three operators are required for maximum repression in vivo (*20*).

Few proteins have had such a strong impact in molecular biology. For decades, a prodigious amount of genetic and biochemical information on *lac* repressor has awaited structural interpretation (*21*). A site-specific substitution analysis by Miller and co-workers (*22*–*24*) resulted in more than 4000 single amino acid replacements of known phenotype between codons 2 through 329 of the *lacI* gene. There are more known functional variants of the *lac*

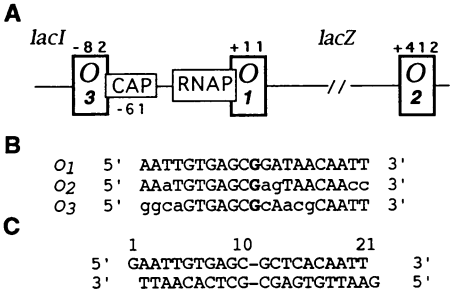

Fig. 1. (**A**) The *lac* operon of *E. coli* showing the locations of the primary operator O_1 and the two auxiliary operators O_2 and O_3 with respect to the CAP binding site and promoter region for RNA polymerase. The ends of the *lacI* and *lacZ* genes are also provided as a reference. (**B**) Sequence of the operators O_1, O_2, and O_3 aligned with the center of each operator in the *lac* operon in bold type. Bases listed with capital letters are identical to the bases in the primary operator O_1. (**C**) The sequence of the DNA duplex used in *lac* repressor–DNA co-crystallization aligned with operators in (B). This sequence lacks the central base and is a palindrome of the left half-site of O_1.

M. Lewis and G. Chang are in the Johnson Research Foundation and Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, PA 19104, USA. N. C. Horton, M. A. Kercher, H. C. Pace, and P. Lu are in the Department of Chemistry, University of Pennsylvania, Philadelphia, PA 19104, USA. M. A. Schumacher and R. G. Brennan are in the Department of Biochemistry and Molecular Biology, Oregon Health Sciences University, Portland, OR 97201, USA.

*To whom correspondence should be addressed.
†Present address: The Upjohn Company, Kalamazoo, MI 49001, USA.

repressor than of any other protein, including hemoglobin (*25*).

The *lac* repressor and its complexes with operator DNA and effector molecules have both contemporary and historical importance for the understanding of gene regulation. As a consequence, considerable effort has been devoted to determining the three-dimensional structures of the *lac* repressor and related proteins. The structures of proteolytic fragments of both the NH_2-terminal DNA binding domain (*26*) and the COOH-terminal tetrameric core bound to inducer (*27*) have been determined. The *lacI* family of proteins consists of about 20 transcriptional regulators that have sequence homology to the *lac* repressor (*28*). Within this family, only the dimeric purine repressor (PurR)–DNA complex structure has been determined (*29*). We present here the three-dimensional structures of the intact *lac* repressor, the *lac* repressor bound to the gratuitous inducer IPTG, and the *lac* repressor complexed with a 21-bp symmetric operator DNA. These structures provide not only the three-dimensional framework for interpreting a plethora of biochemical and genetic information but also suggest an explicit model for how the tetrameric form of the *lac* repressor interacts with DNA and functions synergistically with CAP to regulate the *lac* operon.

Structure determination of the repressor and its complexes. Methods for protein isolation, purification, and initial crystallization of the intact *lac* repressor have been described (*15*). The *lac* repressor crystallized in a monoclinic space group C2 (*30*). The structure was solved by multiple isomorphous replacement (MIR) and anomalous scattering (Table 1). A mercury analog of the inducer IPTG, methylmercury(II)-β-D-1-thiogalactopyranoside (MMG) was synthesized by methods established for derivatives of sugars that bind pertussis toxin (*31*, *32*). Isomorphous crystals were obtained with MMG and the difference Patterson function was solved by a genetic algorithm search procedure (*33*). The mercury bound to Cys^{281} on the surface of the repressor and close to, but not in, the sugar binding site (Fig. 2). This was surprising since MMG competed for IPTG binding in vitro. A second derivative, prepared by soaking crystals with $K_2Pt(CN)_4$, was far more stable in the x-ray beam and the limit of diffraction was extended from 3.5 to 2.5 Å. This heavy atom derivative did not alter the inducer binding as measured by equilibrium dialysis. Anomalous scattering data were collected from a single crystal with x-rays at the platinum absorption edge (λ = 1.071 Å), which gave a Patterson function with a clear heavy atom signal. However, these data were not isomorphous with the native data, and a substitute "native" data set was collected from crystals soaked in $K_2Ni(CN)_4$. A single heavy atom, which was positioned on the noncrystallographic twofold axis that relates the two dimers, bound per tetramer and stabilized the tetramer by noncovalently cross-linking the dimers (Fig. 2).

Parallel attempts were made to solve the structure by molecular replacement methods. Self-rotation functions (*34*) revealed *222* symmetry in the diffraction data, indicating a noncrystallographic twofold axis perpendicular to the unique axis. The orientation of the molecule was determined from a cross-rotation function with the dimeric purine co-repressor binding domain (*29*). Conventional search procedures to translationally position this model were unsuccessful. Using a genetic algorithm (GA_RB), the molecule was positioned by simultaneously determining the orientation of the search molecule while searching for the position of the subdomains within the cell (*35*). The molecular replacement solution was verified by calculating the isomorphous and anomalous difference Fourier in each case, which produced large peaks at the expected heavy atom sites.

Table 1. Data collection and crystallographic analysis. Data were collected in our laboratory at the University of Pennsylvania (PENN), on the X-8-C beamline at the NSLS Brookhaven National Laboratory (BNL), and on the F2 beamline at Cornell High Energy Synchrotron Source (CHESS). All Pt and Ni data sets were collected at −165°C. For cryoprotection, *lac* repressor crystals, which had been previously exposed to $K_2Pt(CN)_4$ or $K_2Ni(CN)_4$, were soaked stepwise in a series of solutions of 2-methyl-2,4-pentandiol ranging from 5 to 25 percent over a 10-minute period, mounted in rayon loops, and then flash-cooled in a stream of nitrogen gas chilled to −165°C. PENN data were collected with a Siemens multiwire area detector mounted on a Siemens rotating anode generator operated at 40 kV, 80 mA. Diffraction images were processed and scaled with the XDS package (*72*). BNL data were collected with a CCD detector and diffraction images were processed with MADNES (*73*) and scaled with XSCALE (Kabsch). CHESS data were collected with the Princeton 1K CCD detector. Images collected at CHESS were integrated and scaled [DENZO and SCALEPACK (*74*)]. Anomalous data for platinum-soaked crystals as well as the isomorphous nickel crystals were collected at the absorption maximum of Pt LIII edge (λ = 1.071 Å). The Pt-absorption edge was determined by x-ray fluorescence at 90 degrees from the crystal. Data collection for the unsoaked, uncomplexed *lac* repressor, the *lac* repressor co-crystallized with MMG, the DNA–*lac* repressor and the IPTG–*lac* repressor crystals was done at room temperature (designated by RT). The following summarizes the space group and cell parameters of the crystals given in Table 1: Co-crystallization with MMG: Monoclinic, *C2*; a = 166.4, b = 75.4, c = 150.2 Å; β = 119.0°. $K_2Pt(CN)_4$ and $K_2Ni(CN)_4$ soak: Monoclinic, *C2*; a = 160.3, b = 73.3, c = 147.8 Å; β = 120.3°. Apo *lac* repressor: Monoclinic, C2; a = 164.5, b = 75.4, c = 148.7 Å; β = 118.8°. Co-crystallization with IPTG: Orthorhombic, $P2_12_12$; a = 149.2, b = 141.2, c = 75.1 Å. Co-crystallization with DNA: Monoclinic, *C2*; a = 104.3, b = 224.4, c = 112.1 Å; β = 95.7°.

Crystal	Data set		Data statistics					MIR and anomalous analysis				
	Source	No.	Resolution (Å)	Reflections		R_{sym}* (%)	Completeness (%)	Isomorphous† difference (%)	Phasing power‡		Figure of merit	
				Total	Unique				Acentric	Centric	Acentric	Centric
MMG	PENN-RT	8	15.0–3.50	89,018	18,234	7.4	87	18.0	1.6	1.1	0.42	0.61
$K_2Pt(CN)_4$	CHESS	1	15.0–3.15	54,876	17,781	7.1	65	17.5	1.4	1.1	0.38	0.63
	BNL	1	15.0–2.53	72,476	29,211	5.2	56	18.0	1.6	1.2	0.40	0.62
	PENN	1	15.0–2.70	71,187	37,143	4.4	89	17.8	1.6	1.2	0.40	0.60
$K_2Ni(CN)_4$	BNL	1	15.0–2.97	68,473	24,293	4.8	80					
	PENN	1	15.0–4.80	26,039	7,455	6.4	98					
Apo	PENN-RT	8	15.0–4.80	56,580	7,764	7.2	95					
IPTG	PENN-RT	1	15.0–3.20	110,526	26,094	9.1	97					
DNA	PENN-RT	5	15.0–4.80	43,976	11,283	5.8	90					

Model statistics			
Model	R (%)	rms bonds (Å)	rms angles (degrees)
LacR	25	0.005	2.1
IPTG	23	0.014	1.9
DNA	26	0.011	1.8

*$R_{sym} = \Sigma_{i,h}|I(i,h) - \langle I(h)\rangle|/\Sigma_{i,h}|I_{(i,h)}$, where $\langle I(h)\rangle$ is the mean of the i observations of the reflection h. †Isomorphous difference is $\Sigma_h||F_{PH} - F_P|/\Sigma F_{PH}$, where F_P and F_{PH} are the derivative and native structure factor amplitudes, respectively. $K_2Ni(CN)_4$ (BNL and PENN) data were used as native for the $K_2Pt(CN)_4$ (BNL and PENN) data, respectively. $K_2Ni(CN)_4$ (BNL) was used as a native for the $K_2Pt(CN)_4$ (CHESS) data. PENN-RT data were used as native for the PENN-RT MMG data set. ‡Phasing power, mean value of heavy atom structure factor amplitudes divided by the lack of closure.

An unbiased view of the *lac* repressor was obtained independently with the phases derived from the heavy atom derivatives. A preliminary set of phases was computed with PHASES (36) followed by solvent flattening. Electron density maps computed with these initial phases were sufficient to see the molecular boundary of the tetramer in the asymmetric unit. The experimentally derived phases were refined by separate fourfold averaging of the noncrystallographically related inducer binding domains and COOH-terminal oligomerization helices. A model built to the averaged electron density with CHAIN (37) revealed the core structure (residues 62 to 329) as well as the COOH-terminal oligomerization helices (residues 340 to 357). Residues 1 to 61 were disordered in the absence of DNA.

A model of the repressor, containing residues 62 to 357, was refined with the use of both the experimentally measured amplitudes and the symmetry averaged phases as implemented in X-PLOR (38). The fit of the model to the map is described by a vector residual of 39 percent, corresponding to a crystallographic *R* factor of 25 percent, for all data with *d* spacing greater than 2.7 Å. The stereochemistry of the model has a root-mean-square (rms) error in bond lengths and angles of less than 0.005 Å and 2.1°, respectively. As an independent check of the model, the averaging was repeated starting with phases derived from a polyalanine model of *lac* repressor. The resulting electron density from this "side chain omit map" (Fig. 3) allows the objective placement of the side chains in the model.

A different crystal form appeared when the repressor was co-crystallized in the presence of IPTG (30). The diffraction quality of these crystals was also improved by $K_2Pt(CN)_4$. These crystals are in a space group $P2_12_12$ with one tetramer in the asymmetric unit. The structure of the inducer complex was solved by conventional molecular replacement techniques (34) with amino acids 62 to 329 of the *lac* repressor structure. Initial phases were calculated on the basis of the uncomplexed structure and averaged from the fourfold noncrystallographic symmetry. The electron density map revealed the location and position of the inducer molecule as well as the platinum atom. The model of the repressor-inducer complex was also refined employing strict noncrystallographic symmetry and the phase constraints. The agreement between the observed data and that calculated from the model for the data to 3.2 Å has a vector residual of 38 percent and the conventional *R* factor of 23 percent. Again, there is no electron density for the headpieces and the hinge helices.

Several dozen synthetic double helical fragments of DNA that varied in length and sequence produced crystals. Only one synthetic operator produced crystals that diffracted sufficiently well. These crystals contain a 21-bp symmetric operator segment to which the repressor binds with greater affinity than the wild-type sequence (39) (Fig. 1). Crystals of the binary complex grew in space group C2 yielding diffraction

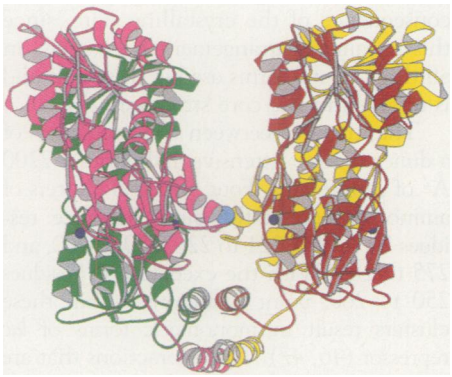

Fig. 2. View of the uncomplexed *lac* repressor tetramer with heavy atoms shown. Residues 62 to 357 of each monomer are illustrated with Molscript (76). Individual monomers are colored in green, violet, red, and yellow. The Pt^{2+} atom and the Hg^{2+} atoms are colored light blue and dark blue, respectively.

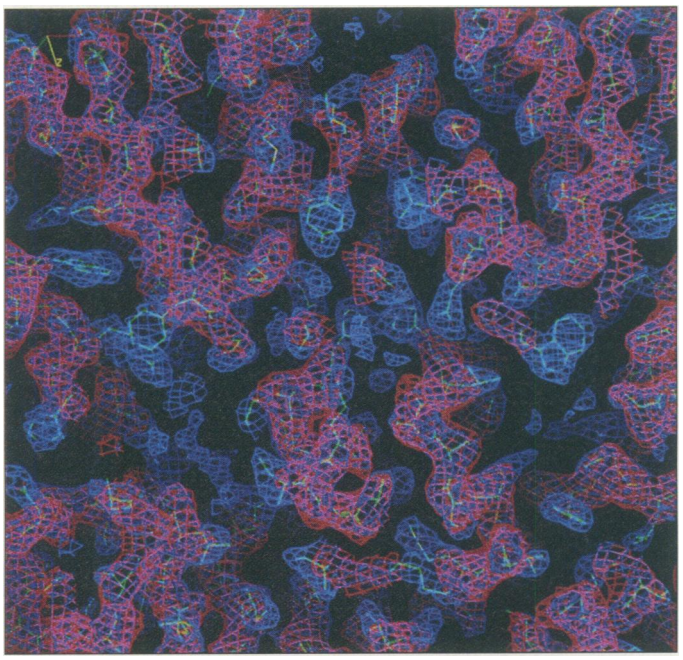

Fig. 3. Averaged electron density omit map of the uncomplexed *lac* repressor. The red density (contoured at 1.2 σ) is a map calculated with the observed structure amplitudes and phases derived from a polyalanine model. These preliminary phases were improved by solvent flattening and averaging. The "side chain omit map," also contoured at 1.2 σ, was calculated with the observed amplitudes and averaged phases. The information about the location of the side-chain residues is derived entirely from the amplitudes and is not biased by the model phases. The maps were viewed with CHAIN.

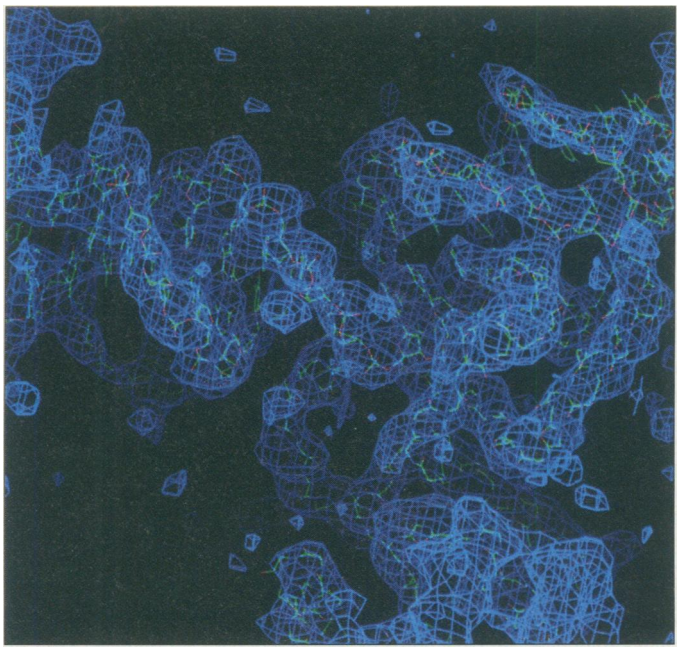

Fig. 4. Electron density of a portion of the *lac* repressor–DNA complex showing the density for the phosphate backbone of the DNA and the DNA binding domain of the repressor. Initial phases were obtained by molecular replacement with the use of the uncomplexed *lac* repressor structure (residues 62 to 329). After five cycles of symmetry averaging with F_o coefficients and subsequently with five cycles of $2F_o - F_c$ averaging, electron density maps were calculated with the use of the observed structure amplitudes and phases derived from the averaging. The electron density map was contoured at 0.8σ and displayed from CHAIN.

data with d spacings greater than 4.8 Å (Table 1). Molecular replacement methods were used to find the orientation of the *lac* dimers and their positions were determined with the use of GA_RB (*35*). Several iterations of fourfold noncrystallographic symmetry averaging improved the quality of the phases and produced unambiguous electron density for building the intact repressor as well as the 21-bp DNA. Again, the coordinates were improved by means of a phase-constrained refinement and maintaining strict noncrystallographic symmetry. The structure of the repressor-DNA complex agrees with the experimental data having a vector residual of 46 percent and a crystallographic R value of 26 percent. When the repressor was complexed to the DNA, the headpieces were ordered and were visible in the electron density (Fig. 4).

Architecture of the repressor. The repressor monomer has four functional units (Fig. 5, A and B), namely, an NH_2-terminal headpiece which binds specifically to the DNA, the hinge region, a sugar binding domain, and a COOH-terminal helix. The DNA binding domain, or headpiece of the *lac* repressor contains a helix-turn-helix (HTH) motif and is similar to other prokaryotic regulatory proteins in this class (*40*). A small, compact globular domain with a rich hydrophobic core is formed by two α helices in a HTH motif (residues 6 to 25) that is connected by a short stretch of polypeptide chain to the third helix (residues 32 to 45).

A linker (residues 46 to 62) connects the DNA binding domain to the core of the repressor. This segment of the polypeptide chain, referred to as a hinge, was thought to be devoid of secondary structure because it is most susceptible to proteolytic cleavage (*41*). Nuclear magnetic resonance (NMR) studies of the intact repressor in the presence and absence of DNA demonstrate that the headpiece of repressor moves independently of the core (*42*, *43*). In the presence of DNA, residues 50 to 58 are ordered and form an α helix that appears to (i) make specific interactions with the *lac* operator DNA and (ii) orient the headpieces. The coil-to-helix transition for residues 50 to 58 occurs only when the repressor associates with the operator. In the absence of the DNA, the hinge helices are disordered giving the headpieces a broad range of structural freedom.

The inducer binding domain or core structure of the *lac* repressor is similar to PurR (*29*) and to the bacterial periplasmic binding proteins (*44*). The inducer binding domain is composed of two subdomains (Fig. 5) that are topologically similar, with a six-stranded parallel β sheet that is sandwiched between four α helices. These two compact subdomains are hinged together by three linkers (Fig. 5B). At the COOH-terminus of the core there is a short segment of 11 residues followed by a COOH-terminal α helix (residues 340 to 357) that contains two leucine heptad repeats. An oligomerization domain is formed when all four COOH-terminal helices of the tetramer associate. Mutations and deletions of the COOH-terminus destroy the tetramer and result in dimers with partial repressor function (*6*, *8*, *20*, *45*).

The quaternary structure of the *lac* repressor (Fig. 2) and the core fragment (*27*) are unusual and do not maintain the point group symmetry of other oligomeric proteins of known structure. Since virtually all homotetramers of known structure have *D2* (*222*) symmetry, it may be more useful to consider the *lac* repressor as a dimer of dimers. The twofold axis that relates the monomers within the dimer are skewed ($\Delta\phi \approx 20°$; $\Delta\psi \approx 10°$) with respect to the twofold axis that relates the dimers within the tetramer. This creates a repressor tetramer that is roughly V-shaped. It is unlikely this unusual quaternary structure is a consequence of the crystalline state, since this tetrameric arrangement is observed in all three crystal forms used in this study and is retained in the core structure (*27*).

The interface between two monomers of a dimer is quite extensive and buries ~2200 $Å^2$ of surface area. Four principal clusters of amino acids create this dimer interface: residues 70 to 100, 221 to 226, 250 to 260, and 275 to 285. With the exception of residues 250 to 260, point mutations within these clusters result in monomeric forms of *lac* repressor (*46*, *47*). The interactions that are responsible for the quaternary structure of the tetramer are almost entirely due to the association of the COOH-terminal helices, which is consistent with genetic analysis (*8*). Apart from the tetramerization domain, the surface area of a dimer that is buried when the tetramer forms is about 300 $Å^2$, which is

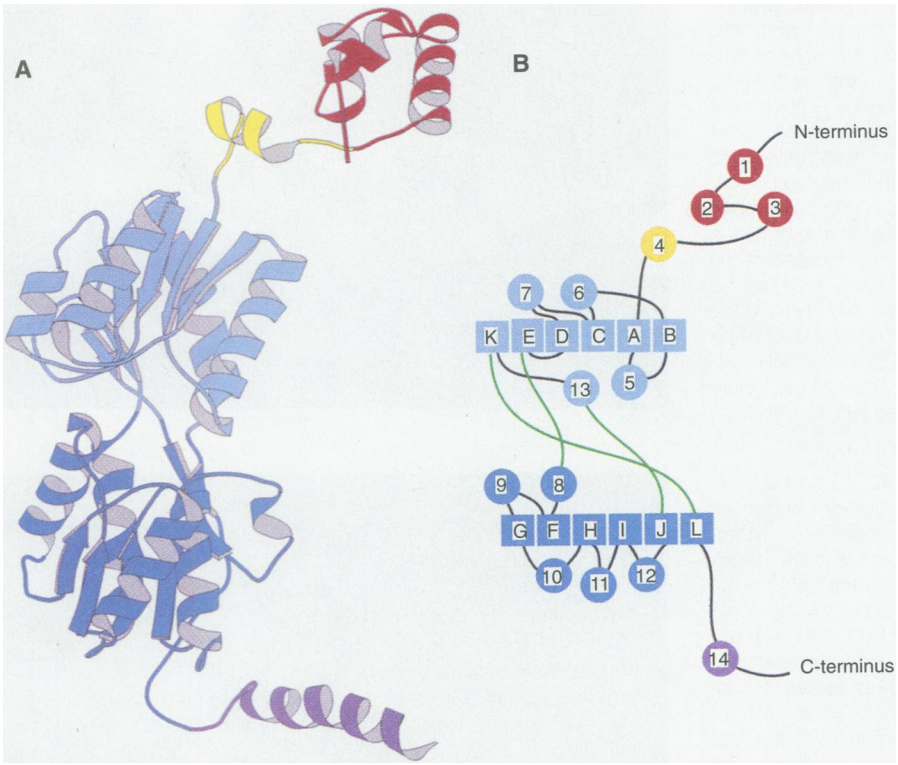

Fig. 5. (**A**) View of the *lac* repressor–DNA complex monomer where the domains have been colored separately. The DNA helix-turn-helix binding domain is colored in red, the DNA binding hinge helix is shown in yellow, the NH_2-terminal subdomain of the core is light blue, the COOH-terminal subdomain of the core is dark blue, and the helix of the tetramerization domain is purple. (**B**) Topology diagram of the *lac* repressor–DNA complex monomer. Helices are depicted as circles and strands as squares. The DNA binding domain consists of helix 1 (residues 6 to 12), helix 2 (17 to 25), helix 3 (32 to 45), and the hinge helix 4 (50 to 58) (red and yellow). The NH_2-terminal subdomain of the core (light blue) is comprised of strand A (63 to 68), helix 5 (74 to 90), strand B (92 to 98), helix 6 (104 to 116), strand C (121 to 124), helix 7 (131 to 137), strand D (145 to 149), strand E (158 to 161), helix 13 (293 to 309), and strand K (316 to 318). The COOH-terminal subdomain of the core (dark blue) consists of helix 8 (164 to 175), strand F (182 to 185), helix 9 (192 to 205), strand G (214 to 217), helix 10 (222 to 233), strand H (240 to 244), helix 11 (247 to 259), strand I (269 to 274), helix 12 (279 to 281), strand J (287 to 290), and strand L (322 to 324). The helix of the tetramerization domain (purple) is helix 14 (340 to 357) and residues 358 to 360 are disordered. The linkers between the NH_2-terminal and COOH-terminal core subdomains are highlighted in green. N-terminus, NH_2-terminus; C-terminus, COOH-terminus.

similar to the average surface area buried in a crystal lattice contact, about 250 Å^2 (*48*). The arrangement of strong dimer contacts and weak tetramer interactions suggests that the observed tetrameric structure of the repressor is essentially a tethered dimer. There are no obvious reasons for the two dimers to associate with this particular geometry and one might expect that the pair of tethered dimers could adopt a variety of conformations. Comparing the quaternary structures of repressor, in the various crystal forms, demonstrates that the orientation of the two dimers is not fixed precisely. Repressor dimers could be arranged with their DNA binding domains pointing in opposite directions.

DNA binding and structure. In the crystals of the repressor-DNA complex, each repressor tetramer is bound to two independent, symmetric operator DNA double helices (Figs. 6 and 7). The primary sites of interaction are concentrated in the HTH motif, which fits snugly in the major groove, and are consistent with the protection and synthetic DNA studies (*9*, *49*). The orientation of the HTH motif with respect to the DNA is essentially the same as that observed with PurR (*29*) but unlike other prokaryotic HTH proteins of known structures. Because of the limited resolution of the co-crystal data, atomic interactions between the protein and DNA cannot be assigned in detail. However, residues Leu^6, Tyr^{17}, Gln^{18}, Ser^{21}, Arg^{22}, and His^{29} are close to the DNA, and their side chains most likely form direct base pair interactions with the operator in the major groove of the DNA. Substitution of the amino acid side chains at these positions result in repressor molecules that are incapable of binding to DNA and have an I^- phenotype (*24*). From the position of the hinge helices, it is also likely that the pair of leucine residues at position 56 (one from each monomer) make direct contacts with the bases in the minor groove of the operator. These leucine residues are in close proximity to the center of the operator DNA and, as also seen in the PurR structure (*29*), work as a lever to pry open the minor groove. The wild-type *lac* operator is asymmetric with an additional G·C base pair at the center of the operator. If we assume that the repressor binds the natural operators in a similar fashion, these leucine residues would introduce an asymmetry in repressor binding that is observed in the methylation protection data as well as the distribution of constitutive mutations (*9*).

Binding of the repressor to the 21-bp symmetric operator distorts the conformation of the DNA so that it bends away from the repressor with an approximately 60 Å radius of curvature, slightly greater than the curvature for a nucleosome (*50*). Distortions on the DNA are localized to the center of the operator where there is a bend or kink of about 45° that results in the minor groove having a width of more than 11 Å and significantly reduces the depth of the groove to less than 1 Å. The central portion of the operator has a helical rise and a twist angle of 6.1 Å and 22°, respectively. The DNA was observed to have a local unwinding of more than 50°, which is consistent with solution measurements (*51*). These distortions accommodate the hinge helices in the minor groove. The average helical parameters that describe the conformation of the DNA in the environment of the HTH are consistent with the canonical B form. When a single headpiece associates with a half-site operator site, data from solution NMR studies indicated that the entire DNA fragment was undistorted from canonical B form (*26*). The presence of the full operator seems necessary for the transition of the hinge region into an α helix, which in turn is responsible for the local DNA deformation. The two fragments of DNA bound to a tetramer of the repressor are separated by approximately 25 Å and there are no interactions between these DNA fragments or with symmetry-related DNA fragments in the crystal structure (Fig. 7).

The inducer binding site. The inducer molecule, IPTG, binds to the repressor at the interface of the NH_2-terminal and COOH-terminal subdomains. Since the IPTG molecule is pseudo symmetric, the inducer molecule could be fit to the density in two ways. Although the same side chain residues were used in both instances, the contacts to the sugar differ. The repressor can make either three or four hydrogen bonds to the IPTG hydroxyl groups. The orientation with four hydrogen bonds involves the amino acid side chain Asn^{246} with the O2 of the galactoside, Arg^{197} with O3 and O4, and Asp^{149} with the O6 position. The sugar binding pocket also has a hydrophobic surface formed by Leu^{73}, Ala^{75}, Pro^{76}, Ile^{79}, Trp^{220}, and Phe^{293}, and the isopropyl group of the IPTG is within van der Waals contact of Trp^{220}. The alternative orientation is seen in the *lac* repressor core fragment (*27*).

There exists a collection of altered repressor molecules that bind to the operator DNA with wild-type affinity but are incapable of induction (*24*). These repressor molecules, which are classified by an I^S phenotype, are either defective in sugar binding or cannot transmit the allosteric signal to the DNA binding domain. The I^S point mutations are scattered throughout

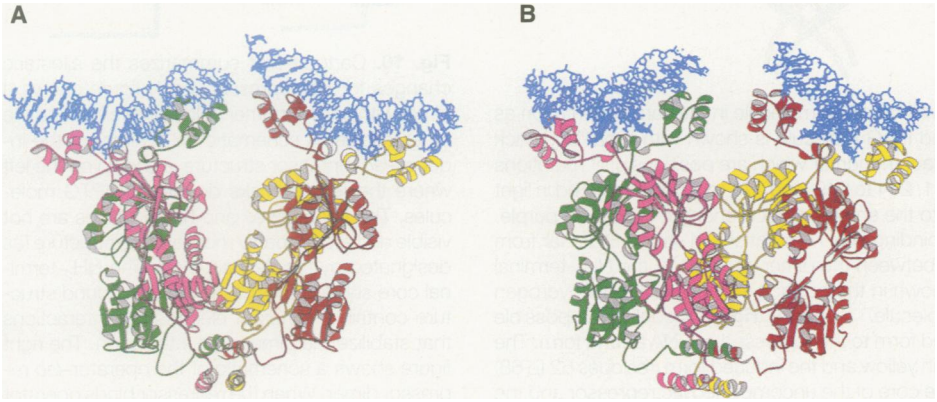

Fig. 6. (**A**) View of the *lac* repressor–DNA complex. Individual monomers (residues 1 to 357) are colored green, violet, red, and yellow. The two 21-bp deoxyoligonucleotides are in blue. The orientation of the molecule is similar to the orientation of the uncomplexed form in Fig. 5. (**B**) View of the tetramer which has been rotated around the vertical axis of the page by approximately 40° from the orientation shown in (A).

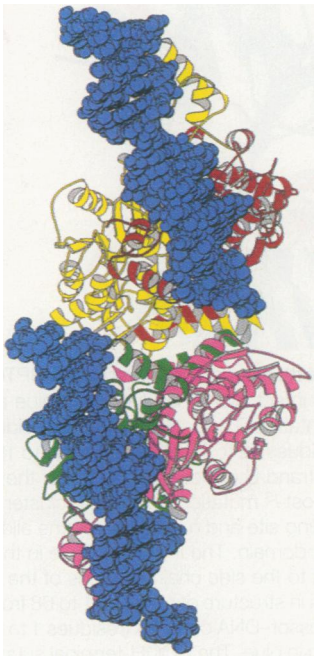

Fig. 7. View of the *lac* repressor–DNA complex illustrating that the two 21-bp fragments of deoxyoligonucleotide are separated in space and do not interact in the tetramer. The color scheme is the same as in Figs. 2 and 6.

the linear sequence of the protein but appear to cluster in five general locations that include residues 70 to 80, 90 to 100, 190 to 200, 245 to 250, and 272 to 277. In addition, residues Trp^{220} and Phe^{293} are also located within the binding pocket. Most of the I^S mutations are in close proximity to the sugar binding site, with the exception being mutations between residues 90 to 100 (Fig. 8). These residues form the first strand of the β sheet (strand B) within the NH_2-terminal subdomain and are at the dimer interface, suggesting that residues 90 to 100 transmit the allosteric signal.

The allosteric transition. Comparing the structures of the repressor in the three crystal forms suggests there are two distinct structural rearrangements of the monomer and the dimer, corresponding to the induced and repressed states. The conformation of the apo repressor monomer is extremely similar to the repressor in the presence of IPTG with an rms difference in the α-carbon coordinates of less than 0.4 Å. In contrast, the rms difference between the uncomplexed repressor monomer and that bound to DNA is greater than 1.9 Å for all α carbons. The repressor adopts a conformation in the presence of the operator that differs from both the unliganded and the inducer bound forms. When the NH_2-terminal and COOH-terminal subdomains in the induced conformation are superimposed independently on the corresponding subdomain of the repressor in the repressed conformation, the rms deviation in the α carbons is small (less than 1 Å). These subdomains are structurally invariant and do not significantly change conformation when switching between states. However, the relation of the NH_2- and COOH-terminal subdomains do change by what appears to be a small hinge motion.

This hingelike motion primarily alters the positional relations of the NH_2-terminal subdomains within the dimer changing the interface between the two NH_2-terminal subdomains. When the repressor binds to its operator DNA, these two subdomains rotate by about 10°, and their centers of mass move by about 2 Å preserving the dimer twofold axes. The hinge motion does not alter the relationship between COOH-terminal subdomains in the dimer. The structural change between the induced and repressed states is most likely propagated via the hinge helices to the headpieces (Fig. 9). The change in the dimer interface within the NH_2-terminal subdomain results in the displacement of the α carbon of residue 62, the first amino acid of the core. In the induced form of the repressor, this residue moves away from its twofold-related mate by about 3.5 Å which (i) disrupts the interaction of the hinge helices, (ii) frees the DNA binding HTH domains, and (iii) reduces the affinity of the repressor for the operator.

The allosteric transition appears remarkably similar to that observed in hemoglobin (*17*). The conformation of the repressor in the repressed state is analogous to the oxy form of hemoglobin or the R state while the repressor bound to inducer corresponds to the T state. In the induced state, the quaternary structure of the repressor, like deoxyhemoglobin, forms a number of electrostatic interactions across the dimer interface. Three electrostatic interactions between residues Lys^{84} and Glu^{100}, Gln^{117} and Arg^{118}, and His^{74} and Asp^{278} are observed in the induced (IPTG-bound) conformation of the repressor. In the repressed (DNA-bound) conformation, the α-carbon atoms of these residues move apart by over 3 Å and these interactions can no longer form across the dimer interface. The inducer molecule, IPTG, most likely performs the same role as the allosteric effector 2,3-bisphosphoglycerate, stabilizing the T conformation. Interestingly, a particular pair of residues that forms a salt bridge in the induced form of the repressor (His^{74} and Asp^{278}) is also used in hemoglobin and is responsible in part for the Bohr effect. It is possible that this interaction may be responsible for the pH dependence of repressor binding to operator DNA (*52*).

The structural basis for the allosteric transition of the *lac* repressor can be formulated in view of the observation that the induced conformation and the unliganded form of the repressor are virtually identical and differ from the repressed conformation (Fig. 10). In the absence of inducer, the unliganded form of the repressor can adopt

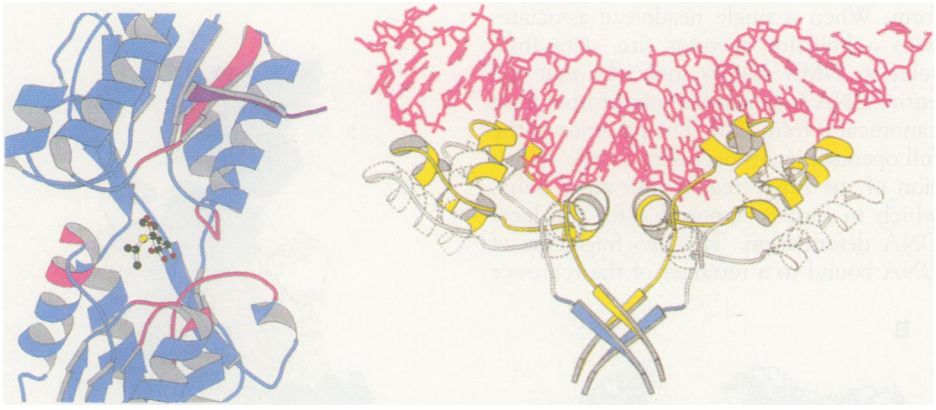

Fig. 8 (left). View of the *lac* repressor–IPTG monomer with an IPTG molecule in the same orientation as 4A. The intact monomer is colored blue and the bound IPTG molecule is shown with a ball and stick representation. The backbone of the residues on the intact monomer which are positions of I^S mutations (*24*) (residues 72 to 76, 92 to 100, 191 to 197, 219 to 221, 244 to 249, and 271 to 278) are colored in light violet. Strand B (residues 92 to 98) of the dimer mate to the shown intact monomer, is colored purple. While most I^S mutations appear to cluster in the IPTG binding site, I^S mutations in strand B are far from the binding site and must transmit the allosteric signal between the dimer interface at the NH_2-terminal core subdomain. The IPTG molecule in this figure is shown in the position which yields most hydrogen bonding to the side chain residues of the repressor molecule. **Fig. 9 (right).** Model of the possible changes in structure of residues 1 to 68 from the induced form to the repressed or DNA bound form. The *lac* repressor–DNA complex (residues 1 to 68) is shown in yellow and the induced form (residues 62 to 68) is shown in blue. The COOH-terminal subdomains of the core of the uncomplexed *lac* repressor and the *lac* repressor–DNA were superimposed to predict the relative positions of residues 1 to 61. The dashed lines show where residues 1 to 61 must go in the uncomplexed *lac* repressor form as a consequence of the change in position of residue 62. Residues 1 to 61 are not seen in the current electron density in the absence of DNA. In the uncomplexed form, the hinge helices would appear to move 3.5 Å apart in space and would no longer interact with the DNA.

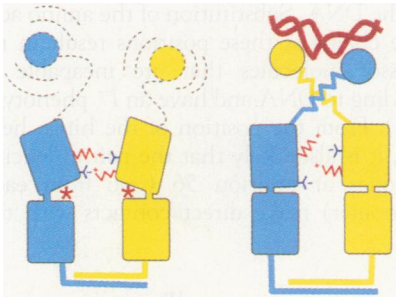

Fig. 10. Cartoon that summarizes the allosteric changes for *lac* repressor. Each figure shows a dimer with one monomer colored in yellow and the other in blue. A schematic for a dimer of the inducer–*lac* repressor structure is shown on the left where the red asterisks denote the IPTG molecules. The headpieces and hinge helices are not visible and are probably mobile in the structure (as designated by the dashed lines). The NH_2-terminal core subdomains of the inducer bound structure contain a series of electrostatic interactions that stabilize the dimer core interaction. The right figure shows a schematic of the operator–*lac* repressor dimer. When the repressor binds operator DNA, the hinge helices form, the NH_2-terminal core subdomain translates and rotates, the interactions between the NH_2-terminal subdomains are broken, and the COOH-terminal oligomerization helices change conformation.

an alternative conformation when presented with operator DNA. In this conformation, the electrostatic interactions that hold the dimer together are broken, and the NH_2-terminal subdomains move apart. The stability of this complex therefore results from the interaction of the two headpieces with the DNA and the interaction of the two hinge helices. As a consequence, the repressed conformation of the repressor may locally destabilize the quaternary structure of the dimeric core. Binding of inducer to the operator-repressor complex causes a change in the quaternary structure of the core that propagates through the hinge helices, which destabilizes the repressor-operator complex. This conformational change allows the electrostatic interactions to reform between the NH_2-terminal subdomains of the core structure.

The role of the tetramer. The allosteric transition is observed within a dimer, yet the fully functional *lac* repressor is a tetramer. A comparison of the different states of the repressor tetramer show that the tethered dimers move apart, broadening the V-shaped structure by approximately 12° and the dimers twist by approximately 8° about their twofold axes on binding the operator. The relative orientation of the COOH-terminal tetramerization helices with respect to the inducer binding domains changes with the entire four-helix bundle translated about 2 Å away from the core.

Removing the COOH-terminal helix of the repressor results in functional dimers, which bind a single operator with virtually the same affinity as the tetramer (*53*) but with diminished induction ratios (*20*). To achieve maximal repression of the *lac* operon requires three operators suggesting that the tetrameric form of the repressor is required for maintaining the physiology of the system. Removing O_1, O_2, or O_3 reduces the repression ratios by varying degrees (*20*). A single *lac* repressor tetramer binds two operators that are separated by 93 or 401 bp. Since the repressor tetramer binds two independent pieces of operator DNA, a continuous DNA must bend to form repression loops (*20*, *54*, *55*).

There are two plausible mechanisms for creating these looped structures which are consistent with the architecture of the *lac* repressor tetramer. Two subunits of the tetramer can bind to the primary operator site and the other dimer subsequently can associate with an ancillary operator. Alternatively, free repressor dimers bind to separate operators and a loop occurs when the dimeric repressors associate into a tetramer. Like the hinge region (helix 4), the COOH-terminal helices (helix 14) would undergo a coil-to-helix transition upon tetramer formation by associating into a four-helix bundle. Synthetic peptides made with the sequence of the COOH-terminal tetramerization domain associate cooperatively, forming four-helix bundles with no detectable intermediates (*56*). By either mechanism, the repressor acts like a double clamp, bringing two operators that are separated in linear sequence close together in space. The order of events is dictated by the concentration of the repressor in the cell, the dimer-tetramer association (*57*, *58*), and the binding affinity of dimeric repressor molecules to the operators (*59*–*61*). The reported values for the several dissociation constants vary widely. Both mechanisms are plausible and may depend on the precise physiological conditions.

For the *lac* repressor to bind two distinct operators and form repression loops depends on the physical properties of DNA as well as the length of the intervening loop (*62*, *63*). Since the two ancillary operators of the *lac* operon are separated by a different number of base pairs from the primary operator, the mechanism of loop formation in each case should be rather different. Forming stable looped complexes, particularly for relatively short stretches of DNA, may require additional DNA binding proteins which can dramatically change the physical properties of the nucleic acid. For example, CAP interacts with a DNA sequence between the primary and a secondary operator and kinks the DNA, inducing an approximately 90° bend over about 30 bp of DNA (*64*). This CAP-induced bending facilitates a single *lac* repressor tetramer to bind at O_1 and O_3, resulting in greater stability of the ternary complex (*65*).

It may seem paradoxical that a transcriptional activator (CAP) would stabilize or promote the repressed state. Monod extensively examined the diauxic behavior of bacterial cultures when glucose and lactose are present in the growth medium (*66*). Glucose is the primary energy and carbon source whereas lactose is an ancillary fuel that is used when the glucose supply becomes depleted. Repression loops between O_1 and O_3 most likely form when there are low concentrations of glucose (that is, an increase of cyclic AMP, promoting CAP binding) and lactose such that the bacteria need to reduce the baseline transcription of the *lac* operon to conserve energy. In this way the *lac* repressor and CAP proteins operate as an integrated switch that responds to the relative concentration of these metabolites.

A model of a 93-bp repression loop that corresponds to the *lac* operon DNA sequence −82 to +11 was constructed with a shape best described by the Folium of Descartes (Fig. 11). The model for this repression loop positions the contact surface for RNA polymerase on the inside surface, providing an additional mechanism for preventing polymerase access to the promoter (*67*). This higher ordered structure efficiently prevents polymerase from fully engaging the *lac* promoter and thus maximizes repression (*68*). In addition, a closed DNA loop between O_1 and O_3 could facilitate the interaction of CAP with RNA polymerase, consistent with the observation that *lac* repressor accelerates messenger RNA synthesis upon induction (*69*).

Looping of DNA over long distances is a recurring feature of transcriptional complexes in both prokaryotic (*70*) and eukaryotic systems (*71*) and the quaternary structure of *lac* repressor may serve as a model for some eukaryotic regulators. It is likely that DNA binding proteins that regulate transcription would bind to the outside of the looped structures when the DNA is supercoiled and associated with histones or their analogues. The structural model for a repression loop, as described here, provides a physical depiction of how higher order complexes may regulate complex events within the cell.

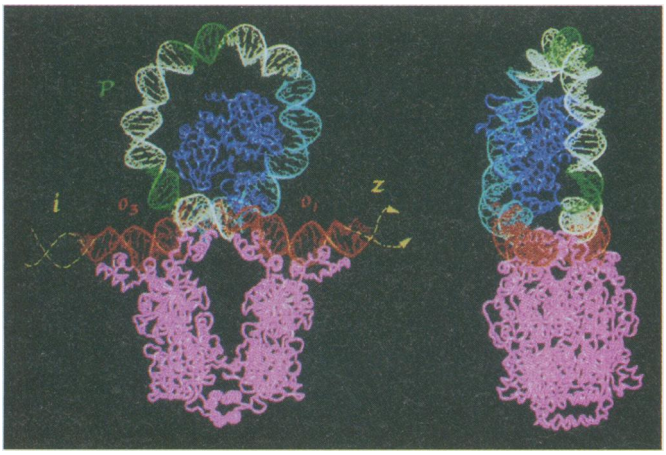

Fig. 11. A model of the 93-bp repression loop that corresponds to the *lac* operon −82 to +11. The ends of the loop contain operators O_1 and O_3 (red) to which the *lac* repressor tetramer is shown bound (violet). Inserted in the loop is the CAP protein and 30-bp DNA complex (blue) taken from the PDB coordinates (*64*). The grey DNA was created by applying a smooth curvature to B-DNA. The curvature of the modeled portion of the loop is about 40 Å and consistent with the curvature of the DNA observed in the CAP structure and that observed for the repressor. The handedness of the superhelix is arbitrary. Highlighted in green are the −10 and −35 sites of the *lac* promoter. (Figure produced by means of Biosym Insight II.)

REFERENCES AND NOTES

1. F. Jacob and J. Monod, *J. Mol. Biol.* **3**, 318 (1961).
2. W. Gilbert and B. Müller-Hill, *Proc. Natl. Acad. Sci. U.S.A.* **56**, 1891 (1966).
3. S. Bourgeois and M. Pfahl, *Adv. Prot. Chem.* **30**, 1 (1976).
4. P. J. Farabaugh, *Nature* **274**, 765 (1978); K. Beyreuther, K. Alder, N. Geisler, A. Klemin, *Proc. Natl. Acad. Sci. U.S.A.* **70**, 3576 (1973).
5. T. Platt, J. G. Files, K. Weber, *J. Biol. Chem.* **248**, 110 (1973).
6. J. H. Miller, T. Platt, K. Weber, "Strains with the promoter deletion L1 synthesize an altered *lac* repressor," in *The Lactose Operon*, J. R. Beckwith and D. Zipser, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1970).
7. A. E. Chakerian *et al.*, *J. Biol. Chem.* **266**, 1371 (1991).
8. S. Alberti, S. Oehler, B. von Wilcken-Bergmann, B. Müller-Hill, *New Biol.* **3**, 57 (1991).
9. W. Gilbert and A. Maxam, *Proc. Natl. Acad. Sci. U.S.A.* **70**, 3581 (1973).
10. J. R. Sadler, H. Sasmor, J. L. Betz, *ibid.* **80**, 6785 (1983).
11. A. Simons, D. Tils, B. von Wilcken-Bergmann, B. Müller-Hill, *ibid.* **81**, 1624 (1984).
12. A. Jobe and S. Bourgeois, *J. Mol. Biol.* **69**, 397 (1973); C. Burstein, M. Cohn, A. Kepes, J. Monod, *Biochim. Biophys. Acta* **95**, 634 (1965); B. Müller-Hill, H.v. Rickenberg, K. Wallenfels, *J. Mol. Biol.* **10**, 303 (1964).
13. F. Jacob and J. Monod, *Cold Spring Harbor Symp. Quant. Biol.* **26**, 193 (1961).
14. J. Monod, J.-P. Changeux, F. Jacob, *J. Mol. Biol.* **6**, 306 (1963).
15. H. C. Pace, P. Lu, M. Lewis, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 1870 (1990).
16. F. Haurowitz, *Hoppe-Seyler's Z. Physiol. Chem.* **254**, 268 (1938).
17. M. F. Perutz, W. Bolton, R. Diamond, H. Muirhead, H. C. Watson, *Nature* **203**, 687 (1964).
18. G. Zubay, D. Schwartz, J. Beckwith, *Proc. Natl. Acad. Sci. U.S.A.* **66**, 104 (1970).
19. W. Gilbert, in *Protein-Ligand Interactions*, H. Sund and G. Blauer, Eds. (de Gruyter, Berlin, 1975), p. 193.
20. S. Oehler, E. R. Eismann, H. Krämer, B. Müller-Hill, *EMBO J.* **9**, 973 (1990).
21. J. H. Miller and W. S. Reznikoff, Eds., *The Operon* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1980).
22. J. H. Miller *et al.*, *J. Mol. Biol.* **131**, 191 (1979).
23. J. H. Miller, *ibid.* **180**, 205 (1984).
24. P. Markiewicz, L. Kleina, C. Cruz, S. Ehret, J. H. Miller, *ibid.* **240**, 421 (1994).
25. D. Voet and J. G. Voet, *Biochemistry* (Wiley, New York, 1990).
26. V. P. Chuprina *et al.*, *J. Mol. Biol.* **234**, 446 (1993).
27. A. M. Friedman, T. O. Fischmann, T. A. Steitz, *Science* **268**, 1721 (1995).
28. M. J. Weickert and S. Adhya, *J. Biol. Chem.* **267**, 15869 (1992).
29. M. A. Schumacher, K. Y. Choi, H. Zalkin, R. G. Brennan, *Science* **266**, 763 (1994).
30. Monoclinic *lac* repressor crystals were obtained from drops made in microbridges from equal volumes of 5.3×10^{-5} M protein in tris buffer (pH 7.4) and 11 to 14 percent 0.4 M sodium potassium phosphate buffer at pH 5.59 over a reservoir containing precipitant. Crystals, suitable for x-ray analysis, formed after 1 to 3 months at room temperature with average dimensions 0.5 mm by 0.3 mm by 0.2 mm. IPTG–*lac* repressor crystals (average dimensions 0.3 mm by 0.2 mm by 0.2 mm) were obtained by co-crystallizations with slightly lower percentages of 0.4 M sodium potassium phosphate buffer (pH 5.59) and an initial protein solution containing a mixture (1:8 molar ratio) of *lac* repressor and IPTG. Platinum- and nickel-containing *lac* and IPTG–*lac* repressor crystals were obtained by a slow exchange into 4.9 mM solution of either $K_2Pt(CN)_4$ or $K_2Ni(CN)_4$ followed by a 50-minute soaking in the final solution. The sequence of the plasmid carrying the gene coding for the repressor used in these studies was verified to be that reported by Betz (*75*). By all extensive biochemical criteria, the protein in these studies behaves exactly like that described (*3*). In addition, mass spectroscopy indicated that the purified protein was full length.
31. R. Shigeta, K. Forest, L. Yan, D. Kahne, C. E. Schutt, *Actn Crystallogr.* **D50**, 71 (1994).
32. MMG or methylmercury(II)-β-D-1-thiogalactoside was synthesized by reacting the sodium salt of β-D-1-thiogalactoside (Sigma Chemical Company) with methylmercury(II) chloride in methanol. The product was purified by preparative silica thin-layer chromatography. A mimic of *o*-nitrophenyl-β-D-fucoside (ONPF), methylmercury(II)-β-D-1-thiofucoside, was synthesized in a slightly longer procedure starting with (+) D-fucose (Janseen Chimica). MMG–*lac* repressor crystals were obtained as described for the protein alone. The initial protein solution contained a mixture (1:8 molar ratio) of *lac* repressor and MMG.
33. G. Chang and M. Lewis, *Acta Crystallogr.* **D50**, 667 (1994).
34. J. Navaza, *ibid.*, p. 157.
35. GA_RB is a genetic algorithm that is used for rigid-body refinement. The search object was divided into pieces corresponding to the structural subdomains of the core. The NH_2-terminal and COOH-terminal subdomains were allowed to move independently, but were constrained by the noncrystallographic symmetry as was the intact tetramer. A subset of rotational and all of translational space were explored with a continuous transform to interpolate the rotational component of the molecular scattering, and a genetic algorithm to optimize the correlation of the structure amplitudes. The orientation and position of the model pieces were further refined with a modified version of GA_RB that maximized the peak height in the difference Fouriers at the heavy atom sites.
36. W. Furey and S. Swaminathan, *Acta Crystallogr.* **18**, 73 (1990).
37. J. S. Sack, *J. Mol. Graphics* **6**, 224 (1988).
38. A. T. Brünger, J. Kuriyan, M. Karplus, *Science* **235**, 458 (1987).
39. The deoxyoligonucleotide used in this study was synthesized by the standard phosphoramidite method and purified by anion exchange chromatography (Pharmacia Mono Q column) at pH 10.0 at room temperature. The DNA was desalted and then annealed (10.0 mg/ml) by heating at 70° to 80°C for about 10 minutes and slowly cooling overnight. Initial co-crystals of DNA and *lac* repressor (0.1 mm by 0.05 mm by 0.05 mm) grew in 1 to 5 days at 22°C in a mixture of equal volumes of 2.88×10^{-5} M protein and 2.30×10^{-4} M DNA in tris buffer (pH 7.4) and 31 percent PEG 3K with 0.05 M potassium phosphate (pH 8.5) and 0.075 M sodium potassium tartrate. Macroseeding by addition of a 7-μl drop of 32 percent PEG 3K, 0.05 M potassium phosphate (pH 8.5), and 0.075 M sodium potassium tartrate with a seed crystal to a 5.5-μl drop of complex (5.50×10^{-5} M protein and 1.94×10^{-4} M DNA) and placement over a reservoir solution containing 29 to 31 percent PEG 3K in 0.05 M potassium phosphate (pH 8.5) and 0.075 M sodium potassium tartrate yielded larger crystals (1.0 mm by 0.5 mm by 0.5 mm) suitable for data collection.
40. S. C. Harrison, *Nature* **353**, 715 (1991).
41. N. Geisler and K. Weber, *FEBS Lett.* **87**, 215 (1978).
42. N. Wade-Jardetzky *et al.*, *J. Mol. Biol.* **128**, 259 (1978).
43. M. A. Jarema, P. Lu, J. H. Miller, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 2707 (1981).
44. S. L. Mowbray and L. B. Cole, *J. Mol. Biol.* **225**, 155 (1992).
45. A. E. Chakerian and K. S. Matthews, *J. Biol. Chem.* **266**, 22206 (1991).
46. A. Schmitz, U. Schmeissner, J. H. Miller, P. Lu, *ibid.* **251**, 3359 (1976).
47. W.-I. Chang, J. S. Olson, K. S. Matthews, *ibid.* **268**, 17613 (1993).
48. M. Lewis, unpublished data.
49. M. H. Caruthers, *Acc. Chem. Res.* **13**, 155 (1980).
50. R. D. Kornberg and A. Klug, *Sci. Amer.* **244**, 48 (February 1981).
51. R. Kim and S.-H. Kim, *Cold Spring Harbor Symp. Quant. Biol.* **47**, 451 (1982).
52. A. D. Riggs, S. Bourgeois, M. Cohn, *J. Mol. Biol.* **53**, 401 (1970).
53. M. Brenowitz, N. Mandal, A. Pickar, E. Jamison, S. Adhya, *J. Biol. Chem.* **286**, 1281 (1991).
54. Y. Flashner and J. D. Gralla, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8968 (1988).
55. E. R. Eismann and B. Müller-Hill, *J. Mol. Biol.* **213**, 763 (1990).
56. R. Fairman *et al.*, *Protein Sci.* **4**, 1457 (1995).
57. C. A. Royer, A. E. Chakerian, K. S. Matthews, *Biochemistry* **29**, 4959 (1990).
58. R. Fickert and B. Müller-Hill, *J. Mol. Biol.* **226**, 59 (1992).
59. M. Brenowitz, A. Pickar, E. Jamison, *Biochemistry* **30**, 5986 (1991).
60. J. Chen and K. S. Matthews, *ibid.* **33**, 8728 (1994).
61. J. Chen, S. Alberti, K. S. Matthews, *J. Biol. Chem.* **269**, 12482 (1994).
62. G. R. Bellomy, M. C. Mossing, M. T. Record, *Biochemistry* **27**, 3900 (1988).
63. M. C. Mossing and M. T. Record Jr., *Science* **233**, 889 (1986).
64. S. C. Schultz, G. C. Shields, T. A. Steitz, *ibid.* **253**, 1001 (1991).
65. M. J. Hudson and M. G. Fried, *J. Mol. Biol.* **214**, 381 (1990).
66. J. Monod, thesis, Université de Paris (1942).
67. U. Siebenlist, R. B. Simpson, W. Gilbert, *Cell* **20**, 269 (1980).
68. A. Schmitz and D. J. Galas, *Nucleic Acids Res.* **6**, 111 (1979).
69. S. B. Straney and D. M. Crothers, *Cell* **51**, 699 (1987).
70. R. Schleif, *Science* **240**, 127 (1988).
71. S. Adhya, *Annu. Rev. Genet.* **23**, 207 (1989).
72. W. Kabsch, *J. Appl. Crystallogr.* **21**, 916 (1988).
73. A. Messerschmidt and J. W. Pflugrath, *ibid.* **20**, 306 (1987).
74. Z. Otwinowski and W. Minor, *The HKL Program Suite*, in preparation.
75. J. L. Betz, *Gene* **42**, 283 (1986).
76. J. Kraulis, *J. Appl. Crystallogr.* **24**, 946 (1991).

77. We thank S. L. Ginell, L. J. Keefe, C. E. Schutt, R. Shigeta Jr., and A. Luptak for technical assistance and D. C. Rees and M. J. Bennett for discussions, and the Argonne National Laboratory Structural Biology Center and CHESS and MacCHESS for access to synchrotron x-ray facilities. Supported by National Institutes of Health (NIH) grant GM-44617 and an Army Research Office grant DAAL03-92-G-0713, NIH Cell and Molecular Biology training grants (N.C.H. and M.A.K.), and NIH Molecular Biophysics training grant 2-T32-GM082745 (G.C.). The Argonne National Laboratory Structural Biology Center beamline, X-8-C of the National Synchrotron Light Source, is supported by the Department of Energy, Office of Health and Environmental Research, and by the NIH-NCRR grant P41-RR06017. The Cornell High Energy Synchrotron Source is supported by the National Science Foundation (DMR-9311772). MacCHESS is supported by a NIH grant RR-016416. Atomic coordinates have been deposited in the Protein Data Bank with accession numbers 1LBG (*lac* DNA), 1LBH (*lac* IPTG), and 1LBI (*lac* Apo).

16 August 1995; accepted 5 January 1996