Reprint Series
11 November 1988, Volume 242, pp. 899–907

SCIENCE

Recognition of a DNA Operator by the Repressor of Phage 434: A View at High Resolution

ANEEL K. AGGARWAL, DAVID W. RODGERS, MARIE DROTTAR, MARK PTASHNE, STEPHEN C. HARRISON*

Recognition of a DNA Operator by the Repressor of Phage 434: A View at High Resolution

ANEEL K. AGGARWAL, DAVID W. RODGERS, MARIE DROTTAR, MARK PTASHNE, STEPHEN C. HARRISON*

The repressors of temperate bacteriophages such as 434 and lambda control transcription by binding to a set of DNA operator sites. The different affinity of repressor for each of these sites ensures efficient regulation. High-resolution x-ray crystallography was used to study the DNA-binding domain of phage 434 repressor in complex with a synthetic DNA operator. The structure shows recognition of the operator by direct interactions with base pairs in the major groove, combined with the sequence-dependent ability of DNA to adopt the required conformation on binding repressor. In particular, a network of three-centered bifurcated hydrogen bonds among base pairs in the operator helps explain why 434 repressor prefers certain sites over others. These bonds, which stabilize the conformation of the bound DNA, can form only with certain sequences.

PROTEINS THAT REGULATE GENE EXPRESSION GENERALLY recognize specific DNA sequences through the binding properties of a distinct domain. Several classes of such domains have now been identified (*1, 2*). Members of the best studied class contain a "helix-turn-helix" element—a 20-residue motif of nearly invariant geometry but considerable sequence variation (*1, 3*). A first crystallographic analysis of a complex of one of these, the DNA binding domain of phage 434 repressor, with a synthetic DNA operator revealed an intricate combination of direct protein-DNA interactions and DNA conformational effects (*4*). That structure was determined from crystals diffracting to spacings of 3.2 Å in some directions but only to about 4.5 Å in others. Taken

The authors are in the Department of Biochemistry and Molecular Biology, Harvard University, 7 Divinity Avenue, Cambridge, MA 02138. A. K. Aggarwal and S. C. Harrison are in the Howard Hughes Medical Institute.

*To whom correspondence should be addressed.

together with a series of systematic binding measurements, it showed that 434 repressor recognizes in DNA a defined array of base-pair and backbone contacts displayed in the context of the conformation adopted on binding (*5, 6*). A number of important questions about operator recognition remained unanswered: How do variations in the DNA sequence actually affect its conformation and hence modulate repressor binding? Are there significant structural asymmetries or additional direct contacts in a complex with a naturally occurring, incompletely symmetric operator? Do bound water molecules, or water-mediated hydrogen bonds, play any role? Is binding accompanied by any changes in protein conformation? To what extent is the observed DNA conformation imposed by the protein?

In this article, we describe a crystallographic analysis that answers many of the above questions. We have studied a complex between the DNA-binding domain of 434 repressor and a 20-bp DNA fragment. Following earlier terminology, we refer to the domain as R1-69. In temperate bacteriophages, such as lambda and 434, an efficient genetic switch regulates the choice between lysogeny and lytic growth (*7*). The switch requires differential affinity of two proteins, repressor and Cro, for six operator sites, designated O_R1-O_R3 and O_L1-O_L3 (Table 1). The sites have an approximate twofold symmetry, and the proteins bind as dimers. In phage 434, the conventionally assigned sites are 14 bp long, but both the earlier structural analysis and phosphate modification experiments show that repressor interactions also occur outside this consensus region (*4, 8*). The sequence of our 20-bp fragment contains the 14-bp O_R1, embedded in sequences related to the natural flanking sequences.

The complex of R1-69 with this 20-bp fragment closely resembles the complex with a 14-bp fragment studied previously (*4*). The high-resolution structure confirms assignments of interactions in conserved parts of the operator and strengthens principal conclusions from earlier work. The DNA is bent, with marked variation in twist, and the minor groove is compressed at the center. There is an extended protein-DNA interface. Amino acid side chains form an elaborate pattern of polar and nonpolar contacts to bases in the major groove. Main chain NH groups and neutral and charged side chains form hydrogen bonds with phosphates and sugars. There are modest but crucial differences between free and bound repressors. Contacts to base pairs and DNA backbone are linked by chains of hydrogen bonds. A network of three-centered bifurcated hydrogen bonds among base pairs in the center of the operator reinforces the conformation O_R1 adopts when it binds repressor. Because sequences in the center of the two half-sites in O_R1 are not the same, we can analyze two different ways in which operator base pairs adjust to conformational constraints. These observed DNA conformational features and the related protein interactions explain how 434 repressor distinguishes among different operator sequences. In an accompanying article, Jordan and Pabo (*9*) describe the structure of a complex between the lambda repressor DNA-binding domain and a synthetic lambda operator.

Structure determination and refinement. Following the approach of Jordan *et al.* (*10*), we attempted to optimize the quality of the cocrystals by using several different operator fragments. An overhanging 20-bp fragment, containing the exact O_R1 sequence and single unpaired bases at the 5′ end of each strand, gave the best cocrystals with R1-69 (Fig. 1). Two forms were obtained, both of which diffracted to spacings of 2.5 Å along the DNA axis and to at least 3.0 Å perpendicular to the DNA (Table 2). The monoclinic form occurred only once (two crystals), but the orthorhombic form was quite reproducible, and we were able to prepare a bromine derivative by substituting bromodeoxyuracil for thymine at positions −1L, −1R, 5′R, and 7R (see Fig. 1 for numbering scheme). Bromines at positions −1L and −1R are symmetrically related by the pseudo dyad through the middle of the operator, but bromines at 5′R and 7R are not. This asymmetry was important in helping us fix the orientation of the DNA in the crystal.

Using the structure of the refined complex with a 14-bp fragment as a starting model, we obtained molecular replacement solutions from a fast rotation function (*11*) and a translational *R*-factor search with the program TSEARCH (*12*). Two solutions, corresponding to two possible orientations about the pseudo dyad of the operator, clearly stood out from this search. When phases from one of these solutions were applied to $|F_{Br}| - |F_{nat}|$ in order to obtain a difference map, the four bromine positions appeared clearly, and the asymmetry of bromines at 5′R and 7R revealed the correct orientation of the complex in the crystal. We subjected this initial model (R = 44.4 percent) to rigid body refinement at 3.5 Å, using the program CORELS (*13*). Starting with the entire complex as a single rigid body and then gradually breaking it up into an increasing number of fragments resulted in a lowering of the R factor to 26.3 percent over 38 cycles. At this point, the resolution was increased to 3 Å, and 13 further cycles brought the R factor to 26.8 percent. Maps with $2|F_o| - |F_c|$ and $|F_o| - |F_c|$ coefficients were computed, and the model was rebuilt on an Evans and Sutherland PS390 with the aid of the program FRODO (*14*). The rebuilt complex was refined at 3.0 Å with the program TNT (*15*), which adjusts individual atomic positions subject to appropriate geometric restraints. This procedure reduced the R factor to 24.2 percent. The resolution was gradually increased to 2.5 Å and, after positional refinement had been completed, isotropic temperature factors, correlated for bonded atoms, were also refined in a series of separate cycles. The final R factor at this stage was 24.1 percent. Maps with $2|F_o| - |F_c|$ and $|F_o| - |F_c|$ coefficients were again calculated. At this point, 44 solvent molecules were assigned to peaks that had $|F_o| - |F_c|$ density > 3σ and were in positions consistent with reasonable stereochemistry. The model was again rebuilt, and side chains that refined to asymmetric positions in the two half-sites of the complex were checked by difference maps based on phases from models with the side chain (or side chains) in question deleted (omit maps). The DNA was regularized with CORELS, so that all the sugars were kept in C2′ endo conformation. The rebuilt complex and the water molecules were further refined at 2.5 Å until the R factor converged to 19.4 percent. At this stage, an anisotropy

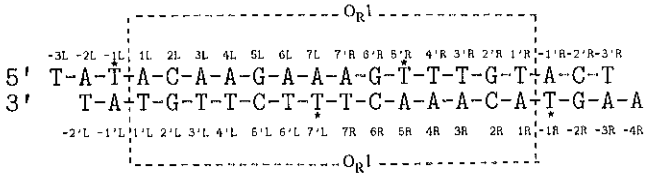

Fig. 1. Sequence of the 20-bp fragment and the numbering scheme used in the text. The approximately symmetric 14-bp operator site O_R1 is divided into right (R) and left (L) half-sites. Nucleotides are numbered 1 to 7 from either end in a 5′ to 3′ direction; nucleotides on the opposite strand are designated with primed numbers. Nucleotides outside the conventional operator sites are indicated by minus signs. Strands were synthesized (Applied Biosystems synthesizer) separately and purified by gel and column chromatography. The strands were then mixed at equimolar concentrations. Rectangular, platelike crystals were grown at 4°C by vapor diffusion from solutions containing 1 m*M* DNA, 2 m*M* protein, 100 m*M* NaCl, 120 m*M* $MgCl_2$, 2 m*M* spermine, and 12 to 14 percent polyethylene glycol 3K. Two forms were obtained: orthorhombic, space group $P2_12_12_1$, a = 148.3 Å, b = 27.7 Å, c = 65.0 Å; and monoclinic, space group $P2_1$, a = 70.5, b = 26.9, c = 64.7, β = 106.1°. The two crystal forms are indistinguishable in morphology and grow under identical conditions. They represent alternative crystal packings of the complexes. Both forms are sensitive to temperature fluctuations and survive for only a few weeks in the growth solution. The Br derivative was prepared by substituting 5-bromodeoxyuracil for T at positions designated by asterisks.

correction was determined for each crystallographic asymmetric unit (*16*). The improvement in overall *R* factor was small (from 19.4 to 19.2 percent), but the 3.0 to 2.5 Å shell showed an improvement of 0.6 percent. Further cycles were calculated, refining positional parameters and then isotropic B's, but this time the anisotropy correction was also applied in the calculations of $|F_c|$ prior to the gradient calculations. The *R* factor fell to 17.9 percent, with root-mean-square (rms) errors in bond length of 0.019 Å and in bond angle of 2.9°. Final $2|F_o| - |F_c|$ and $|F_o| - |F_c|$ maps (Fig. 2) were calculated, with the anisotropy correction applied to $|F_c|$. As before, asymmetric features of the structure were checked by $|F_o| - |F_c|$ omit maps.

DNA conformation. The DNA is a B-type helix, distorted by bending and by variations in twist and other helical parameters. The bases in the L half-site display significant noncoplanarity and form several bifurcated hydrogen bonds. Adjacent DNA fragments stack precisely in the crystal, with overlap base pairing as designed, forming a pseudo-continuous helix parallel to the *c* axis.

The structure of the 20-bp fragment seen in these crystals conforms very well to that of the 14-bp fragment complexed with R1-69 seen at lower resolution (*4*). By contrast, the 14-bp fragment complexed with 434 Cro has a somewhat different configuration (*17*). Thus, an operator, whether embedded covalently in a larger sequence or not, adopts one defined conformation on binding to R1-69, and another on binding to 434 Cro.

1) Interactions with repressor, described below, bend the operator. If the local helical axes for individual base steps are projected onto the mean plane of bending, they lie approximately along a circular arc of radius 65 Å. This radius may be compared to 43 Å for DNA bent around the histone octamer in a nucleosome core (*18*). As in the nucleosome, bending is not smooth. The DNA is relatively straight in the middle of the operator. It bends on each side of the operator by about 12° at two- to three-base steps out from the center (see Fig. 3A). The bent configuration permits contacts between the sugar-phosphate backbone and the NH_2-terminus of helix 2 (along with Arg^{10}), as well as the interactions of Gln^{28} and Gln^{29} with base pairs 1 and 2.

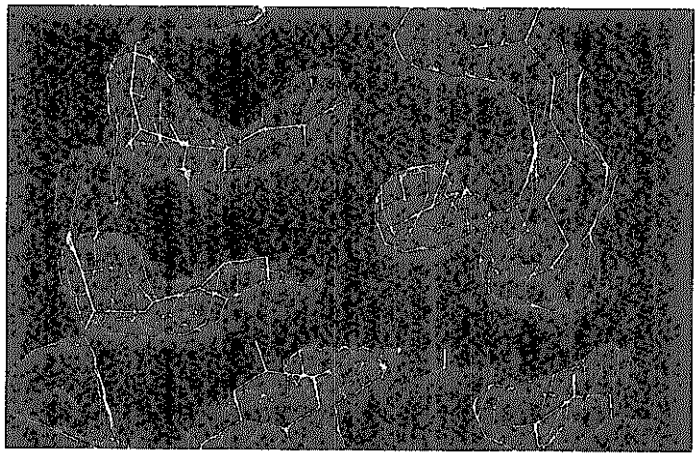

Fig. 2. Part of the electron density from a $2|F_o| - |F_c|$ map at 2.5 Å, showing residues Thr^{27}, Gln^{28}, and Gln^{29} from α3 and portions of base pairs −1 to 3 in the L half-site.

Table 1. Binding of 434 repressor and R1-69 to various operators. Sequences of the six naturally occurring 434 operator sites (*6*) and of the symmetrical 14-bp fragment (*4*). The conventional 14-bp sites are shown, together with flanking base pairs, and the conserved base pairs 1 to 4 are boxed. Binding data (*6*) are summarized as normalized dissociation constants for each protein. The values for repressor (repr.) binding to natural operators include the contribution from cooperativity between repressors bound on adjacent sites. The values for R1-69 reflect the intrinsic affinity for the particular site.

Site	Sequence																Repr.	R1-69
	−1L	1L	2L	3L	4L	5L	6L	7L	7′R	6′R	5′R	4′R	3′R	2′R	1′R	−1′R		
O_R1	T	A	C	A	A	G	A	A	A	G	T	T	T	G	T	T	2	1
	A	T	G	T	T	C	T	T	T	C	A	A	A	C	A	A		
	−1′L	1′L	2′L	3′L	4′L	5′L	6′L	7′L	7R	6R	5R	4R	3R	2R	1R	−1R		
O_R2	A	A	C	A	A	G	A	T	A	C	A	T	T	G	T	A	2	4
	T	T	G	T	T	C	T	A	T	G	T	A	A	C	A	T		
O_R3	C	A	C	A	A	G	A	A	A	A	A	C	T	G	T	A	12	4
	G	T	G	T	T	C	T	T	T	T	T	G	A	C	A	T		
O_L1	T	A	C	A	A	G	G	A	A	G	A	T	T	G	T	A	1	2.5
	A	T	G	T	T	C	C	T	T	C	T	A	A	C	A	T		
O_L2	A	A	C	A	A	T	A	A	A	T	A	T	T	G	T	A	1	2.5
	T	T	G	T	T	A	T	T	T	A	T	A	A	C	A	T		
O_L3	A	A	C	A	A	T	G	G	A	G	T	T	T	G	T	T	40	11
	T	T	G	T	T	A	C	C	T	C	A	A	A	C	A	A		
14 bp		A	C	A	A	T	A	T	A	T	A	T	T	G	T		24	9
		T	G	T	T	A	T	A	T	A	T	A	A	C	A			

2) A striking feature of the 434 operator when bound to repressor is compression of the minor groove in the center and gradual widening to either side. The minor groove width, defined by phosphate-phosphate distances, is 8.8 Å at the midpoint of the operator (compared with 11.5 Å for uniform B-DNA) (*19*); it increases to nearly 14 Å at the ends, opposite the positions where the major groove receives helix 3 (Fig. 4). The central minor groove compression appears to result from the bending of the DNA, together with the way in which sugar-phosphate backbone on either side of the minor groove interacts with the loops between repressor helices 3 and 4 (Fig. 4). The opening up of the minor groove at the ends of the operator is a consequence of bending toward the repressor, aided by underwinding. The width of the major groove shows much less variation, since the effect of bending compensates for the increase that might have been expected to occur on either side of the compressed minor groove. The average major groove width is 17.5 Å, compared with the 17.3 Å for ideal B-DNA (*19*).

3) Local helical parameters vary along the DNA, so that the average helical twist for the 18 steps within the covalently linked portion of the double helix is 35.6°, corresponding to 10.1-bp turn, and the average rise per base pair is 3.3 Å (Table 3). As observed in the complex with a 14-bp fragment (*4*), the center of the operator is overwound by several degrees, and its ends are underwound. The local helical twists are quite similar for corresponding steps in the two half-sites, which were refined independently, but values for the rise seem to be less well correlated across the operator center.

Propeller twist, defined as the relative rotation of bases about their long axis, varies from 2.4° to 29.5°. The central base pairs are generally more highly propeller twisted than those near the ends of the operator. In a double helix containing three or more A · T base pairs in a row, large propeller twist can displace the N6 of A and O4 of T toward the 3′ ends of their respective strands (*20*). As a result, a non–Watson-Crick hydrogen bond can form diagonally across the major groove, in addition to the normal Watson-Crick bonding. We observe such a bifurcated scheme at the center of the operator: O4 of T7R forms a hydrogen bond (2.68 Å) with N6 of A7L (Fig. 5). In a three-centered, bifurcated hydrogen bond, a single proton is shared between two acceptors.

We also observe two novel bifurcated hydrogen bonds—one in the minor groove and one near the helix axis (Fig. 5). Base pair 5L

has an unusually high propeller twist for G · C, which allows N2 of G5L to form a minor-groove hydrogen bond with O2 of T4′L, in addition to its lengthened Watson-Crick hydrogen bond with O4 of C5′L (Fig. 5). The bifurcated bond appears to compensate for weakening of the usual one. A similar minor-groove hydrogen bond between guanine and thymine has been observed in a G · A mismatch DNA structure (*21*). At A · T base pair 6L, the thymine is displaced along the helix axis so that its N3 forms a bifurcated hydrogen bond, involving N1 of A7L in addition to the normal Watson-Crick acceptor, N1 of A6L (Fig. 5). The displaced thymine makes room for the highly propeller twisted G · C 5L. The displacement is correlated with a strong contact from its phosphate to the α3–α4 loop (Fig. 4).

Buckle, the bending of a base pair about its short axis, is a second component of noncoplanarity in base pairs. The buckle of base pairs in the 20-bp fragment varies from −12.9° to 19.3°, values greater than in DNA structures alone. In the dodecamer studied by Dickerson and co-workers, for example, the angle of buckle ranges from −7.0° to 2.7° (*22*); and in the dodecamer studied by Nelson *et al.* (*20*), it ranges from −2.9° to 5.5°. The most highly buckled base pair is C · G 6R. The buckling results from a shift in sugar-phosphate backbone near 6′R, symmetrically related to the shift at 6′L just described. It is not possible to stabilize a displacement of guanine 6′R by a pattern of additional, non–Watson-Crick hydrogen bonds analogous to those stabilizing displacement of thymine 6′L. The guanine therefore tilts to preserve normal base-pairing.

Protein conformation. R1-69 is a bundle of five α helices linked by turns of varying lengths (*23*). The first four helices correspond to helices 1 to 4 of lambda repressor (*24*). The short fifth helix corresponds to a single helical turn in lambda repressor in the connection between its helices 4 and 5 (Table 4). This helix was not correctly built in the model of the complex of R1-69 and the 14-bp

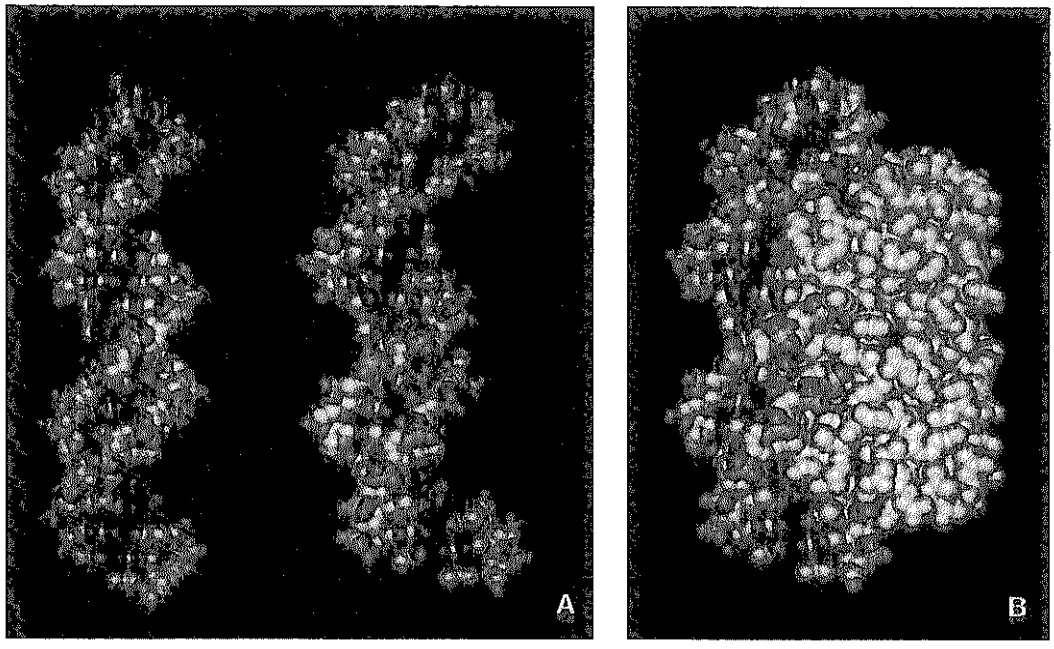

Fig. 3. (**A**) Space-filling representation of canonical straight B-DNA (10.0 bp per turn) and DNA from the complex, oriented identically. The DNA in the complex is bent, and its minor groove is compressed at the center, bringing the phosphate groups very close to each other. (**B**) Space-filling representation of the complex, showing the matched van der Waals surfaces of the protein and DNA.

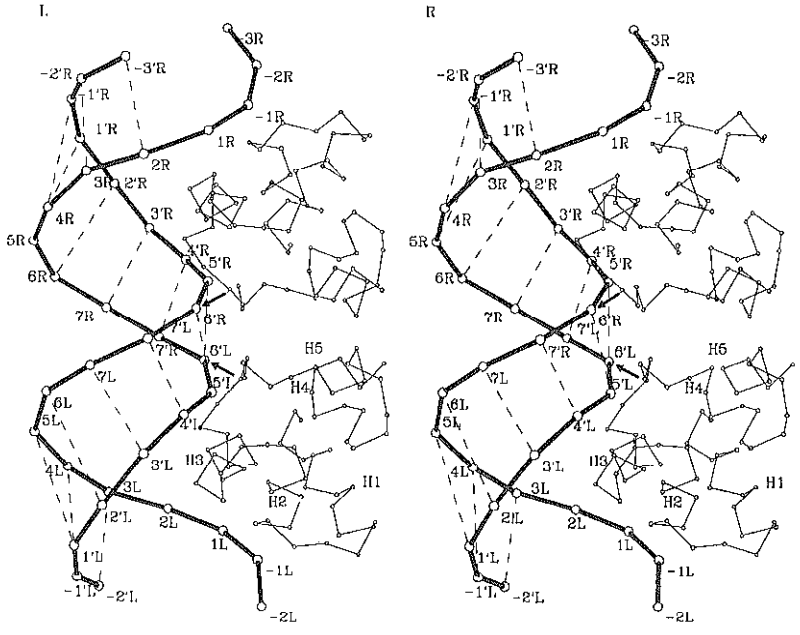

Fig. 4. Stereo view of the phosphate backbone of the 20-bp fragment and the α-carbon trace of dimer R1-69. The dashed lines connect phosphates across the minor groove. The minor groove is compressed in the middle of the operator (8.8 Å and 9.2 Å) and then gradually widens at the ends of the operator (13.5 Å to 14.0 Å). The turn from helix 3 to helix 4 of R1-69 is close to the phosphate backbone at the center of the operator. The arrow indicates close approach to phosphate 6′. H1 to H5 denote helices 1 to 5.

fragment (*4*) because of ambiguities in the electron density map in this region, but it was clearly seen in the free R1-69 structure (*23*). Thus, the DNA-binding domains of 434 and lambda repressors share a very similar, all-helical core. Augmenting this core, the lambda DNA-binding domain has an NH_2-terminal extension (the arm) and a COOH-terminal extension (its helix 5). Neither extension has a counterpart in R1-69.

1) Comparison of bound monomers with the structure of free R1-69 (*23*) shows that no large-scale conformational changes occur on binding, but it does reveal significant local adjustments of side chain configuration and a small shift in the turn from α2 to α3. Side chains that rearrange on binding include Gln^{29}, Glu^{32}, and Arg^{43}, all of which interact with DNA (Fig. 6). Other residues contacting DNA do not rearrange. In particular, Arg^{10}, Gln^{17}, and Gln^{28} appear to be held in position for specific binding by a network of hydrogen bonds, linking helices 1, 2, and 3 (*23*).

2) Contacts between R1-69 monomers involve a patch of hydrophobic residues (Leu^{45}, Pro^{46}, Val^{56}, and Leu^{60} from both monomers) as well as a salt bridge from Arg^{41} to Glu^{47} (*25*). A hydrogen bond between Arg^{41}L and the carbonyl of Arg^{43}R also contributes. The geometry of the protein dimer established by these interactions restricts the conformation of the central base pairs. R1-69 is a monomer in solution. That is, the contacts are strong enough to define the structure of the bound dimer but insufficient to maintain its integrity when not associated with DNA. We believe that dimer contacts formed by R1-69 are the same as those it forms when part of intact repressor, since similar rank orders of operator site affinity are observed with both species (*5*).

Table 2. Statistics for diffraction data. Data for crystals with the native operator fragment were collected at 4°C with CEA-25 x-ray film and oscillation photography (Elliot GX-6 x-ray generator). Three crystals were used for the data set from the orthorhombic form. A data set from the monoclinic form (two crystals) was also collected. The integrated intensities were obtained using the SCAN12 program package (*28*). Data for the bromine derivative were collected at 4°C on a Nicolet area detector (*29*). The crystal to detector distance was 15.0 cm, and the detector was swung out to a 2θ angle of 15°. The reflections were autoindexed and integrated with the BUDDHA program package (*30*). Scaling and merging of data of both native and derivative data were carried out by programs ROTAVATA and AGROVATA (*12*). Merging statistics in the table come from an average multiplicity of 2 to 3 for each reflection.

$$R_{sym} = \sum_i \sum_j |I_{ij} - \bar{I}_i| / \sum_i \sum_j |I_{ij}|$$

N = number of reflections after rejections

N_{poss} = maximum number of reflections possible in the particular resolution shell

% = Percentage of possible data in the particular resolution shell. The drop in this value beyond 3 Å reflects anisotropic falloff of intensities.

Native (25.0 to 2.5 Å)

	d_{min} (Å)	R_{sym}	N	N_{poss}	%
	7.57	0.089	284	417	68.1
	5.48	0.102	494	615	80.3
	4.51	0.109	641	787	81.3
	3.92	0.120	731	887	82.3
	3.52	0.131	819	982	83.4
	3.22	0.159	819	1091	75.1
	2.98	0.185	827	1195	69.2
	2.79	0.194	708	1247	56.8
	2.63	0.195	443	1356	32.7
	2.50	0.183	253	1337	18.9
Overall		0.123	6019	9914	60.7

Protein-DNA contacts. The R1-69 dimer binds to DNA with helix 3 lying in the major groove and with the NH_2-termini of helices 2 and 4 close to the sugar phosphate backbone. The loop from helix 3 to helix 4 follows the sugar phosphate backbone at the center of the operator. Many of the contacts described below were observed in the complex of R1-69 with the 14-bp fragment. The precision of our structure allows us to give a more detailed account of these interactions, to describe others missed previously at lower resolution, and to identify significant bound solvent molecules (Fig. 6).

Direct contacts to conserved base pairs 1 to 4 in the major groove,

Table 3. Helical parameters for the 20-bp fragment. Local twist and rise were calculated for individual base steps with program HELIX; propeller twist (Pr tw) and buckle were calculated by program BROLL (*22*). Reliable calculation of local twist and rise is difficult, since values for the whole base-pair step depend on a few vectors, which can be influenced by local conformational variations such as propeller twist and sugar pucker. We sought to minimize the effect of these variations by idealizing the helix using CORELS (*13*), imposing a C2′ endo conformation on the sugars. This idealized DNA superimposes on the refined DNA backbone with a 0.2 Å rms error and gives local helical twists with standard deviations between 0.2° and 2.0°. The refined DNA coordinates give the same pattern of twists, but individual values vary much more from the mean twist and rise and have much larger standard deviations. Propeller twist and buckle are defined in the text. They were calculated directly from refined DNA coordinates. The average torsion angles (α, β, γ, δ, ε, and ζ) in the sugar-phosphate backbone are −56° ± 37°, 161° ± 20°, 45° ± 35°, 139° ± 13°, −155° ± 24°, and −122° ± 49°. Most correspond to $g^-tg^+tt(g^-/t)$, characteristic of B-DNA (*19*). All the bases are in the *anti* conformation, with an average glycosidic torsion angle (χ) of −108° ± 20°.

Base step	Twist (deg)	Rise (Å)	Pr tw (deg)	Buckle (deg)
−4R A–				
−3R A–T −3′R	35.2	3.9	16.1	14.5
−2R G–C −2′R	34.4	2.8	8.1	9.7
−1R T–A −1′R	34.8	3.5	11.4	−2.0
1R A–T 1′R	34.9	3.4	9.2	3.6
2R C–G 2′R	32.3	3.3	2.4	6.1
3R A–T 3′R	34.2	3.3	4.2	7.6
4R A–T 4′R	38.2	3.7	6.3	−4.7
5R A–T 5′R	37.2	2.9	11.8	13.4
6R C–G 6′R	36.1	3.4	13.6	19.3
7R T–A 7′R	38.1	3.1	28.35	−12.9
7′L T–A 7L	39.6	3.4	29.5	4.9
6′L T–A 6L	36.2	3.1	9.1	11.4
5′L C–G 5L	38.0	3.0	29.1	−1.1
4′L T–A 4L	33.4	3.7	12.7	−11.7
3′L T–A 3L	32.4	2.9	10.6	8.8
2′L G–C 2L	34.8	3.9	3.1	−0.9
1′L T–A 1L	34.2	3.2	12.9	−12.5
−1′L A–T −1L	36.6	3.1	5.15	−10.3
−2′L T–A −2L			7.7	10.1
T −3L				

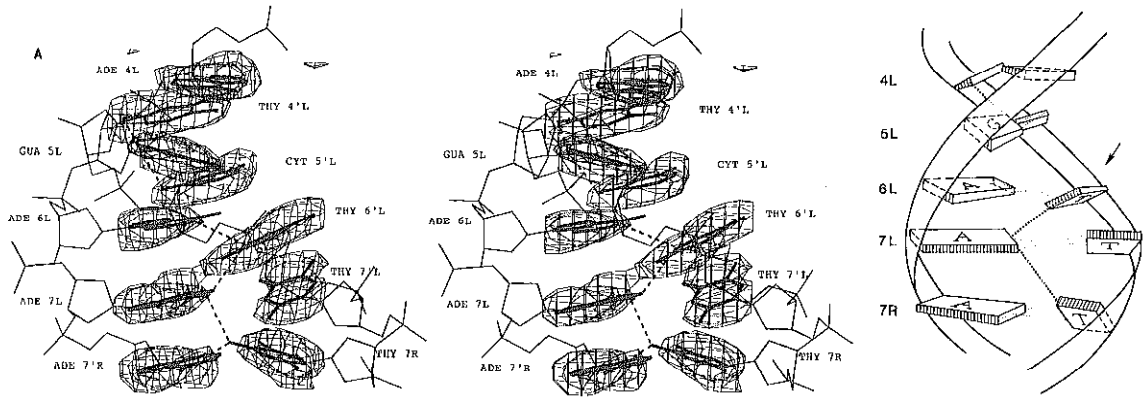

Fig. 5. (**A**) Stereo view of base pairs 4L, 5L, 6L, 7L, and 7′R in the final refined structure. The map is an "omit map," obtained by leaving out these five base pairs from the calculation of structure factors and then using the computed phases to calculate a difference $|F_o| - |F_c|$ map. Such a map is unbiased by the model in this region. The density is contoured at 3σ. Only the three sets of bifurcated hydrogen bonds are shown for clarity. The high propeller twist of A · T base pairs 7′R and 7L makes the first three-centered interaction in the major groove possible; the displacement of thymine 6′L from its Watson-Crick adenine base 6L makes the second three-centered interaction possible; and the high propeller twist of G · C base pair 5L makes the third three-centered interaction in the minor groove possible. (**B**) Diagram of the base pairs 4L, 5L, 6L, 7L, and 7′R, corresponding approximately to the view in (A). Non–Watson-Crick hydrogen bonds are shown by heavy dotted lines, and normal Watson-Crick bonds, by light dotted lines. Arrow shows position of strong interaction with the α3-α4 loop of repressor.

as well as to base pair −1, are similar for both half-sites of the operator. At base pair 1 (A · T), bidentate hydrogen bonding links adenine 1 and Gln^{28}. This is the only contacted base pair that interacts with a single amino acid side chain. Moreover, that side chain (Gln^{28}) also makes extensive van der Waals contact with T-1. This contact, outside the canonical 14-bp operator, may help explain the relative binding constants of 434 repressor for O_R2 and O_L1 (see below). At base pair 2 (C · G), bidentate hydrogen bonding joins guanine 2′ and the terminal -NH_2 of Gln^{29}. In addition, Glu^{32}, directed away from base pairs in the 14-bp fragment complex, points toward cytosine 2 in the structure presented. The distance between them is too large for a hydrogen bond, but some dipolar interaction may occur. At base pair 3 (A · T), no hydrogen bonding to protein occurs, and direct contacts are all nonpolar. The side chains of Thr^{27} and Gln^{29} form a van der Waals pocket to receive the methyl group of thymine 3′. At base pair 4 (A · T), a single hydrogen bond links O4 of thymine 4′ with Gln^{33}. In addition, the thymine methyl group has van der Waals contacts with Gln^{29} and Ser^{30}. The electron density for contacts to base pairs −1, 1, 2, and 3 is shown in Fig. 2. No direct hydrogen bonds link the repressor with base pair 5. Gln^{33} forms a hydrogen bond in both half-sites with T of base pair 4. In the L half-site, its Oε atom comes to within 3.6 Å of the C5 of cytosine 5′L. In the R site, this Oε is displaced by a thymine methyl group. In rotating out of the way of thymine 5′R, the terminal group of Gln^{33} retains its hydrogen bond to thymine 4′, but lengthens its hydrogen bond to the bound solvent molecule bridging to phosphate 5′R and to Ser^{30} (Fig. 6). A concerted set of additional changes involving residues 26 to 30 and at least one other bound solvent molecule is also observed. There is a small but significant change in the α2-α3 turn. We discuss below the significance of these rearrangements for specificity at base pair 5.

Extensive interactions with repressor occur along two segments of the sugar phosphate chain. In segment 1 (nucleotides −2 to 1), the DNA backbone lies against the NH_2-terminus of helix 2. In segment 2 (nucleotides 5′L to 7′L in one half-site and 4′R to 7′R in the other), the backbone contacts the loop between helices 3 and 4 and the NH_2-terminus of helix 4.

Segment 1 has five hydrogen bonds linking the protein to the DNA backbone. The phosphate of nucleotide −1 lies in a small depression in the protein surface. It accepts hydrogen bonds from Arg^{10} and from the main-chain -NH of Gln^{17}. Since Gln^{17} is at the NH_2-terminus of helix 2, the dipole moment of this helix may contribute to the strength of the latter interaction (*26*). Ethylation of this phosphate strongly interferes with repressor binding (*8*). Hydrogen bonds are also observed to the adjacent phosphates: P1 accepts bonds from Gln^{17} and Asn^{36}, and P-2 accepts a bond from Asn^{16}. These interactions were not detected in the ethylation interference experiment. Phosphates and sugars of nucleotide −1 also make extensive van der Waals contacts with main-chain and side-chain atoms of Asn^{16} and Gln^{17}. The sugar of nucleotide 1 is in van der Waals contact with Glu^{32}. Formation of all the protein-DNA contacts in segment 1 requires bending of the operator toward the NH_2-terminus of helix 2.

Segment 2 faces the main chain of the loop between helices 3 and 4 (Fig. 4). The position of the loop appears to contribute to compression of the minor groove at the center. The close approach allows three main chain -NH groups to form hydrogen bonds with two phosphates: from Lys^{40} and Arg^{41} to phosphate 6′ and from Arg^{43} to phosphate 5′. The strength of the NH-phosphate 5′ link at Arg^{43} may be augmented by the dipole of helix 4. The side chains of Arg^{43} project into the minor groove, but their conformations are not identical on the two monomers. Both groups form hydrogen bonds with water, which in turn bond (in different ways) to base pairs 7R and 7L. The side chain of Arg^{43}L makes two direct hydrogen bonds to O3′ and O4′ in the sugar of nucleotide G7′R. The side chain of Arg^{43}R makes a direct hydrogen bond to the phosphate 6′L. These interactions are the only examples of crossover contacts from the L repressor monomer to the R half-site, or conversely. The presence of these two positively charged side chains in the minor groove probably stabilizes its compression, which brings negatively charged phosphates close together. Distances appear to be favorable for Coulombic interactions between Arg^{43}L and phosphates 5′L and 6′R and between Arg^{43}R and phosphates 5′R and 6′L, contributing to the overall stability of the complex. Changing Arg^{43} to Ala decreases binding by a factor of more than 200 (*5*).

In addition to the polar interactions just described, there is extensive van der Waals complementarity of the protein and DNA surfaces in segment 2. Nucleotide 5′ contacts parts of Pro^{42}, Arg^{43}, and Phe^{44}, and nucleotide 6′ contacts main and side chain atoms of Lys^{40}, Arg^{41}, and Pro^{42}, as well as Arg^{43} from both monomers. The sugar of 7′L contacts the side chain of Arg^{43}R. The conformation of the helix 2 to helix 3 turn in the R monomer brings side chains of Thr^{26} and Thr^{27} into van der Waals contact with phosphate 4′R.

Bound solvent molecules. Forty-four solvent molecules have been assigned. Most are probably water, but a few may be cations, since the crystallization is carried out in the presence of high $MgCl_2$

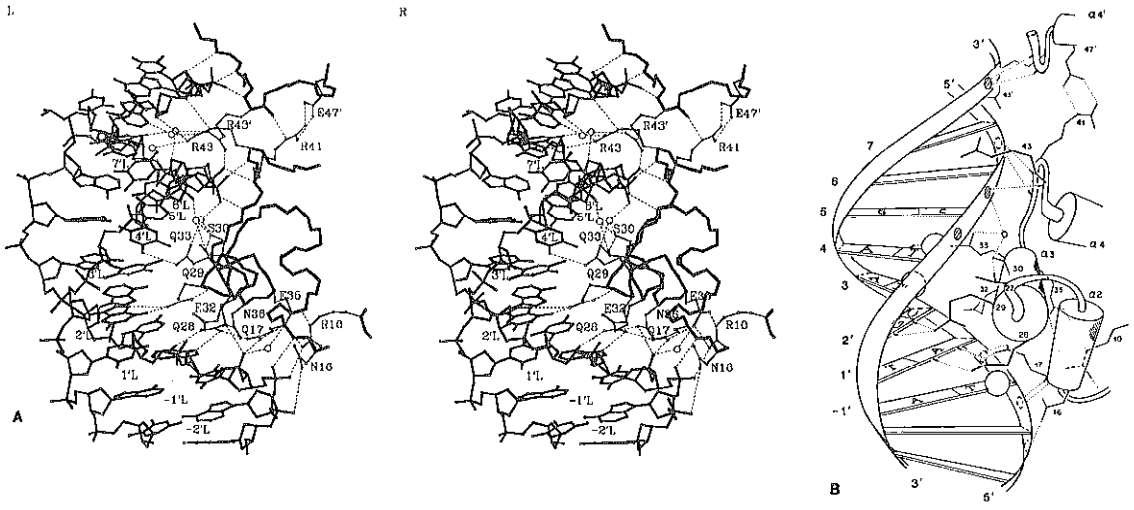

Fig. 6. (**A**) Stereo view of the contacts between protein and DNA. The main chain of the L monomer from residue 16 to 44 is displayed along with residues Arg^{10}, Asn^{16}, Gln^{17}, Thr^{27}, Gln^{28}, Gln^{29}, Ser^{30}, Glu^{32}, Gln^{33}, Glu^{35}, Arg^{41}, and Arg^{43}. The main chain of the R monomer from residue 39 to 47 is displayed along with residues Arg^{43} and Glu^{47}. Except for residues 29 on the L monomer and 43 on the R monomer, the carbonyls have been omitted for clarity. For the DNA, the following nucleotides are shown:

–3L	–2L	–1L	1L	2L	3L	4L	5L	6L	7L	7'R	6'R	5'R
T	A	T	A	C	A	A	G	A	A	A	G	T
	T	A	T	G	T	T	C	T	T	T	C	
	–2'L	–1'L	1'L	2'L	3'L	4'L	5'L	6'L	7'L	7R	6R	

Labels are present on nucleotides −2′L to 7′L and on all protein residues except Thr^{27}. Six water molecules involved in bridging the protein and DNA are shown as circles. The potential hydrogen bonds (up to 3.5 Å in length) are shown as dotted lines. Residues Thr^{27} and Ser^{30} make only van der Waals contacts to DNA. Glu^{35} does not make a direct hydrogen bond to DNA, but it is involved in bridging those residues that do. Also included is the salt link from Arg^{41} in the L monomer to Glu^{47} in the R monomer. Contacts from the two Arg^{43} side chains are asymmetric. (**B**) Diagram of the protein-DNA interface, corresponding approximately to the view in (A).

concentration. Many of the bound solvent molecules lie in the crevices at the edges of the protein-DNA interface. Six are in the minor groove at the center, of which three bridge Arg^{43} residues to base pairs and sugar phosphates (Fig. 6). In the major groove two solvent molecules are present at very similar positions in both half-sites. One of these bridges Gln^{33} to phosphate 5′ as well as to Ser^{30} (Fig. 6). It may play a role in the preference of 434 repressor for sites with G at position 5 (see below). The other solvent molecule common to both half-sites forms a hydrogen bond with Oε of Gln^{29}. Several solvent molecules are also seen in the major groove of one half-site, but not at corresponding positions in the other. (The absence of a peak can indicate either a relatively mobile and disordered molecule or none at all.) A solvent molecule bridges Glu^{35}L to phosphate 1L (Fig. 6), but it is not present at the corresponding R position. Conversely, a water bridging Gln^{33}R to base pair 3R is not seen in the L half-site. Finally, a solvent molecule that bridges Ser^{30}L to phosphate 4′L (Fig. 6) is not present in the R half-site, possibly because it is excluded by the close approach of Thr^{26}R and Thr^{27}R to phosphate 4′R.

Comparison with lambda complex. The complex of the DNA-binding domain of lambda repressor with a synthetic lambda operator (*9*) has a rather different DNA configuration and a very different protein dimer interaction from that in the 434 R1-69 complex that we describe. Nonetheless, it is remarkable that the character of many local interactions at homologous positions is conserved between the two structures. In particular, homologous interactions with DNA backbone appear to anchor each monomer to a half-site.

The structures show that the correct operator alignment matches 434 base pairs −1 to 7 in each half-site with lambda base pairs 1 to 8. The 17-bp lambda operator has an additional base pair in the center; its presence requires a very different dimer configuration. In both complexes, the first residue of helix 3 (Gln^{28} in 434; Gln^{44} in lambda) forms hydrogen bonds with an analogous adenine (position 1 in 434; position 2 in lambda) as well as with the glutamine side chain at the NH_2-terminus of helix 2 (17 in 434; 33 in lambda). This glutamine in turn interacts with the corresponding phosphate (1 in 434; 2 in lambda). Moreover, the same phosphate interacts with a conserved asparagine at the COOH-terminus of helix 3 (36 in 434; 52 in lambda), while the phosphate 5′ to it (−1 in 434; 1 in lambda) forms a hydrogen bond with the main chain -NH of the helix-2 glutamine (17 in 434; 33 in lambda).

Thus, a similar network of hydrogen bonds and van der Waals contacts "clamps" each repressor onto one side of the major groove and links this attachment to its interaction with a base pair. On the other side of the major groove, phosphate 6′ of 434 and PE of lambda accept hydrogen bonds from homologous main chain -NH groups (40 in 434; 56 in lambda). Corresponding phosphates 5′ in 434 and PD in lambda are also both contacted, but differently, since the loops between helices 3 and 4 diverge beyond residue 40 when the two structures are superposed.

The DNA in the lambda complex is more regular than the DNA in the 434 complex. In particular, it is much straighter, the minor groove compression is absent, and central base pair noncoplanarity is less marked (*9*). The lambda arm contacts central base pairs directly, eliminating the functional need for conformational variability, on which 434 repressor relies for specificity at these positions.

A model for the structure of the 434 repressor–DNA complex would have been more difficult to predict from the structure of the free protein than a model for the lambda complex. The rearrangement of side chains and the irregularity of the DNA helix in the 434 structure could not have been anticipated.

Operator recognition. The DNA-protein interface has matched molecular surfaces over an extended area (Fig. 3B). The complementarity includes protein contacts to sugar-phosphate backbone as well as to base pairs in the major groove. Therefore, specificity depends on all these correlated interactions taken together, and changing any one may affect others.

Backbone contacts involve hydrogen bonds from main chain -NH groups and neutral side chains more prominently than close Coulombic interactions. There are also van der Waals contacts to deoxyribose sugars. Only two positively charged side chains (Arg^{10} and Arg^{43}) form hydrogen bonds to phosphates. Several lysine residues near the DNA backbone probably contribute longer range

Table 4. Alignments of 434 and lambda repressors. The residue alignment with lambda repressor was obtained (A. Mondragon) from the superposition of coordinates of free R1-69 with those of lambda repressor (*9*). Residues 63 to 69 were not seen in the map presented or in the map of free R1-69, suggesting inherent flexibility.

Helix	Repressor		Comments
	434	Lambda	
α1	1–12	13–24	434 lacks the NH_2-terminal arm of lambda
α2	16–22	32–38	The α1-α2 turn of 434 is four residues shorter than lambda
α3	28–36	44–52	
α4	45–51	62–68	The α3-α4 loop of 434 is one residue shorter than lambda
α5	56–61	73–76	These residues correspond to the short helical turn from α4-α5 in lambda. 434 lacks the residues described as helix 5 in lambda.

Coulombic interactions, but they are not anchored by direct hydrogen bonds.

In the major groove, nonpolar contacts appear to be as important as hydrogen bonds. Methyl groups of two of the three invariant thymines in each half-site (positions 3′ and 4′) are essential for operator recognition, and two others (positions −1 and 5′R), for differential operator affinity. Three glutamine side chains make hydrogen bonds with base pairs, but each with quite different character and with concomitant nonpolar interactions. A particular side chain can contact several base pairs, and a particular base pair can contact more than one side chain. Moreover, networks of additional hydrogen bonds link the glutamines to other groups on the protein and to phosphates.

Despite this complexity, we can arrive at a systematic understanding of 434 operator recognition by analyzing the structure in the context of known binding data. With the exception of Arg^{43}, each repressor monomer makes contact with just one half-site. Therefore, it is convenient for discussion to divide a half-site into conserved positions 1 to 4, variable position 5, and variable central base pairs 6 and 7. Because of DNA conformational effects, this useful division is probably not strictly observed. For example, the patterns of bifurcated hydrogen bonds we find suggest that effects of sequence changes at positions 5 and 6 cannot be independent. Moreover, since the central base pairs affect the propensity of the operator to adopt its required overwound configuration, contributions from the two half-sites may likewise be correlated.

Positions 1 to 4. All twelve 434 operator half-sites have ACA as the first three bases, and eleven have A as the fourth (see Table 1). The conformations of Thr^{27}, Gln^{28}, and Gln^{29} specify the ACA sequence, in the sense that any change in base pairs 1 to 3 eliminates favorable interactions with these three side chains without restoring compensating ones. At position 4, Gln^{29}, Ser^{30}, and Gln^{33} all interact with the A · T base pair—the first two side chains through van der Waals contact with the thymine methyl group, the last by a hydrogen bond to the thymine. Again, no base pair substitution can conserve the complementarity. In O_R3, there is a G · C at this position in one half-site. The anticipated large decrease in binding constant is relieved in part by a particularly favorable sequence in the center of O_R3 (see below). In their systematic analysis of 434 repressor binding specificity, Koudelka *et al.* (5) and Wharton (*6*) have examined all possible symmetric base pair substitutions in the 14-bp fragment sequence. Any substitution for ACAA at base pairs 1 to 4 leads to a reduction in the repressor-operator binding constant by more than two orders of magnitude. These results are fully consistent with the tightly specific structure we describe.

Position 5. The 12 operator half-sites vary substantially in sequence at positions 5 to 7. There are two versions of position 5 in our structure. We have described above how their difference affects the configuration of Gln^{33} and an associated network of hydrogen bonds including a bound solvent molecule. We have also described a shift in the base at position 6′, in response to positioning of its sugar-phosphate backbone by contacts from the α3-α4 loop, and the correlation of this shift with the configuration of base pair 5. Changes at position 5, made in the context of the 14-bp fragment sequence, show repressor preference in the order G > A, T > C. These affinity differences (approximately five- to tenfold) could arise either from the propagated protein conformational effects of repositioning Gln^{33} or from the way in which other base pairs can adjust to the configuration of base pair 5. Both probably contribute. A mutant repressor with Ala^{33} shows a different order of base-pair preference at position 5 (*6*); Ala^{33} can neither touch base pair 5 nor participate in the hydrogen-bond network. The DNA conformational effects described above could be responsible for the new specificity.

Central base pairs. Variation in binding of the 434 repressor with sequence at the central four base pairs has been analyzed thoroughly (*5*). Operators with A · T or T · A at these positions bind repressor more tightly than those with G · C or C · G. The effect is stronger at position 7 than at position 6. Arg^{43} makes water-mediated hydrogen bonds to base pair 7 in the minor groove. However, these bonds could be made in similar ways with other sequences, and indeed, the influence of base pairs 6 and 7 is independent of Arg^{43} (*5*). No other residues contact these base pairs. The effects are explained by conformational requirements (*5*). Repressor monomers bind tightly to each half-site, and interactions between monomers in turn constrain DNA conformation at the operator center. Relaxing these constraints, by introducing a single-strand nick or by appropriately mutating the protein, diminishes the selectivity (*5*). The structure presented here demonstrates marked noncoplanarity of all central base pairs—expressed as both propeller twist and buckle–a consequence of overwinding and minor-groove compression. The central AAAG sequence (6L to 6′R) has bifurcated hydrogen bonds linking its first three base pairs. AT-rich sequences accommodate such distortions from idealized Watson-Crick geometry more readily than GC-rich sequences. Moreover, the particular network of bifurcated hydrogen bonds that we observe occurs only in the sequence AAA (6L to 7′R). We have described above tight protein contacts to the backbone of nucleotide 6′. Repressor binding should therefore be favored by those sequences for which the conformation constrained by these contacts is energetically least costly. Koudelka *et al.* (*5*) have indeed found that the AAA sequence (6L to 7′R) is especially favored—by a factor of 3 over ATA, for example. Our structure accounts for this preference. A significant difference between the sequences of O_R1 and O_R2 is also this AAA in O_R1 (Table 1), and we propose that it determines the fourfold difference in affinity for R1-69. In O_R3, there is a five-base oligo(dA) sequence in the center of the site. We suggest that its properties can explain the smaller than expected effect of G at position 4R. Several patterns of bifurcated hydrogen bonding are possible, either by displacement of the 6′ bases or by formation of an extended network, as in the structure described by Nelson *et al.* (*20*). The overall DNA helical axis is relatively straight from 5L to 5R, consistent with the conclusion (*5*) that bending (determined by gel mobility) is not correlated with central base-pair selectivity.

Position −1. The conventional 14-bp designation of the 434 operator comes originally from sequence consensus. The contact between Gln^{28} and T-1 suggests that the proper definition of a 434 operator should also include this base pair as well. Indeed, backbone contacts extend to phosphate −2. We suggest that the T-1 interac-

tion can explain the observed ratio of repressor affinities of O_R2 and O_L1. These operators differ only at base pairs 6L to 6R, from which the rules worked out by Koudelka *et al.* (*5*) predict that O_R2 should be the stronger site (by a factor of 2 to 3). The reverse is in fact the case. O_L1 has thymine at position −1 on both sides; O_R2, on only one. The additional contact to a thymine methyl group can account for the actual affinity ratio. We also suggest that the low affinity of O_L3 is due not only to the presence of three G · C–type base pairs in the center but also to the absence of thymine at both −1 positions.

In conclusion, when 434 repressor binds to an operator, extended contacts between the protein and the DNA backbone constrain a bent and twisted structure. These interactions determine the position of α3 in the major groove as well as the shape of the operator DNA. Direct contacts from α3 in the major groove establish specificity for conserved base pairs 1 through 4 and variable base pairs −1 and 5. The irregularities in sugar-phosphate backbone are essentially the same on both sides of O_R1, but the ways in which base pairs in the two sequences adjust to these distortions are different. That is, when fixing a particular conformation for the DNA, the protein appears to meet less resistance from one sequence than from the other. Three-centered bifurcated hydrogen bonding is an important component of this adjustment. It has been suggested that proteins might recognize sequence-dependent nonuniformities in DNA (*27*). In the interaction between 434 repressor and its operator, the protein appears to position the backbone, with a cost in free energy that depends on base sequence. This effect explains observed specificities at base pairs 6 and 7 (*5*).

REFERENCES AND NOTES

1. C. O. Pabo and R. T. Sauer, *Annu. Rev. Biochem.* **53**, 293 (1984).
2. J. Miller, A. D. McLachlan, A. Klug, *EMBO J.* **4**, 1609 (1985); R. M. Evans and S. M. Hollenberg, *Cell* **52**, 1 (1988).
3. W. F. Anderson, D. H. Ohlendorf, Y. Takeda, B. W. Matthews, *Nature* **290**, 75 (1981); D. B. McKay and T. A. Steitz, *ibid.*, p. 744.
4. J. E. Anderson, M. Ptashne, S. C. Harrison, *ibid.* **326**, 846 (1987).
5. G. B. Koudelka, S. C. Harrison, M. Ptashne, *ibid.*, p. 886; G. B. Koudelka, P. Harbury, S. C. Harrison, M. Ptashne, *Proc. Natl. Acad. Sci. U.S.A.*, **85**, 4633 (1988); G. B. Koudelka, S. C. Harrison, M. Ptashne, unpublished results.
6. R. Wharton, thesis, Harvard University, Cambridge, MA (1985).
7. M. Ptashne, *A Genetic Switch* (Cell Press, Cambridge, MA, 1986).
8. R. Bushman, J. E. Anderson, S. C. Harrison, M. Ptashne, *Nature* **316**, 651 (1985).
9. S. Jordan and C. O. Pabo, *Science* **242**, 895 (1988).
10. S. R. Jordan, T. V. Whitcombe, J. M. Berg, C. O. Pabo, *ibid.* **230**, 1383 (1985).
11. R. A. Crowther, in *The Molecular Replacement Method*, M. G. Rossman, Ed. (Gordon & Breach, New York, 1972), p. 173.
12. CCP4 Suite for Protein Crystallography (provided by P. Evans).
13. J. L. Sussman, S. R. Holbrook, G. M. Church, S. Kim, *Acta Cryst.* **A3**, 800 (1977).
14. T. A. Jones, *J. Appl. Cryst.* **11**, 268 (1978).
15. D. E. Tronrud, L. F. Ten Eyck, B. W. Matthews, *Acta Cryst.* **A43**, 489 (1987).
16. S. Sheriff and W. A. Hendrickson, *ibid.* **A43**, 118 (1987).
17. C. Wolberger, Y.-C. Dong, M. Ptashne, S. C. Harrison, *Nature*, in press.
18. T. J. Richmond, J. T. Finch, B. Rushion, D. Rhodes, A. Klug, *ibid.* **311**, 532 (1984).
19. S. Arnott, P. J. Campbell Smith, R. Chandrasekharan, *CRC Handbook of Biochemistry and Molecular Biology, Nucleic Acids* (CRC, Cleveland, OH, 1975), vol. 2, p. 411. W. Saenger, *Principles of Nucleic Acid Structure* (Springer-Verlag, New York, 1984).
20. H. C. Nelson, J. T. Finch, B. F. Luisi, A. Klug, *Nature* **330**, 221 (1987); M. Coll, C. A. Frederick, H.-J. Wang, A. Rich, *Proc. Nat. Acad. Sci. U.S.A.* **84**, 8385 (1987).
21. G. G. Prive, U. Heinemann, S. Chandrasegaran, L.-S. Kan, M. L. Kopka, R. E. Dickerson, *Science* **238**, 498 (1987).
22. R. E. Dickerson and H. R. Drew, *J. Mol. Biol.* **149**, 761 (1981); DNA Helix Analysis Programs (provided by J. M. Rosenberg and R. E. Dickerson).
23. A. Mondragon, S. Subbiah, S. C. Almo, M. Drottar, S. C. Harrison, *J. Mol. Biol.*, in press.
24. C. O. Pabo and M. Lewis, *Nature* **298**, 443 (1982).
25. A similar intersubunit salt bridge has now been found in the refined R1-69 14-bp structure (S. Subbiah, J. E. Anderson, H. Holley, S. C. Harrison, in preparation).
26. W. G. J. Hol, P. T. Van Duijnen, H. J. C. Berendsen, *Nature* **273**, 443 (1978).
27. R. E. Dickerson, *Sci. Am.* **249** (no. 6), 94 (1983).
28. J. Crawford, thesis, Harvard University, Cambridge, MA (1986).
29. R. Durbin *et al.*, *Science* **232**, 1127 (1986).
30. M. Blum, P. Metcalf, S. C. Harrison, D. C. Wiley, *J. Appl. Cryst.* **20**, 235 (1987).
31. Coordinates are being deposited in the Brookhaven Protein Structure Database.

We thank J. Anderson, G. Koudelka, A. Mondragon, R. Wharton, C. Wolberger, F. Bushman, S. Subbiah, and D. C. Wiley for helpful discussions; A. Mondragon for program advice; M. Blum for detector assistance; and M. Saper for help with computer graphics. Supported by a NATO-SERC fellowship (A.A.), an American Cancer Society fellowship (D.R.), and by NIH grant GMS-29109 (M.P. and S.C.H.).

8 July 1988; accepted 30 September 1988