Mutant Trp Repressors with New DNA-Binding Specificities

STEVEN BASS,* VINCENZA SORRELLS,† PHILIP YOUDERIAN

Oligonucleotide-directed mutagenesis of the codons for glutamine-68 (Gln^{68}), lysine-72 (Lys^{72}), isoleucine-79 (Ile^{79}), alanine-80 (Ala^{80}), and threonine-81 (Thr^{81}) of the *Escherichia coli trpR* (tryptophan aporepressor) gene was used to make mutant repressors with each of 36 different amino acid changes. Mutant repressors were tested for binding to each member of a set of 28 different operators closely related to the consensus *trp* operator. Of the 36 mutant repressors, 11 bind a subset of the 28 operators; 5 of these have new binding specificities. These new specificities indicate that the hydroxyl group of Thr^{81} makes a specific contact with one of the four critical base pairs in a *trp* operator half-site, and the methyl group of Thr^{81} determines specificity at a second, critical base pair. The Trp repressor does not use the first two amino acids of its "recognition α-helix," Ile^{79} and Ala^{80}, to make sequence-specific DNA contacts, and interacts with its operator in vivo in a way fundamentally different from the way that phage lambda repressor, lambda Cro protein, and coliphage 434 repressor contact their respective binding sites.

THE RESULTS OF SEVERAL CRYSTALLOGRAPHIC INVESTIGATIONS have advanced our understanding of the nature of the specific, weak chemical interactions between DNA-binding proteins and their recognition sites. Detailed solutions of the structures of three protein-DNA cocrystals, endonuclease Eco RI bound to its site (*1*), coliphage 434 repressor bound to its operator (*2*), and the amino-terminal domain of λ repressor bound to its operator (*3*), indicate that the ability of these proteins to bind DNA selectively is determined by a small number of specific hydrogen bonds (H bonds) and van der Waals interactions between functional groups on amino acid side chains and nucleotide bases.

However, detailed crystal structures have not been determined for most DNA-binding proteins, in the absence or in the presence of DNA. Nonetheless, a number of crystal structures are known for several DNA-binding proteins in the absence of their sites, including Cro protein (*4, 5*), *Escherichia coli* catabolite activator protein (*6, 7*), and *E. coli* Trp (tryptophan) repressor (*8, 9*). Models of the specific interactions between these proteins and their sites have been inferred from these crystal structures as starting points (*5–14*). The

The authors are in the Department of Biological Sciences, University of Southern California, Los Angeles, CA 90089–1481.

*Present address: Genentech, Inc., South San Francisco, CA 94080.
†Present address: University of Southern California School of Medicine, Los Angeles, CA 90033.

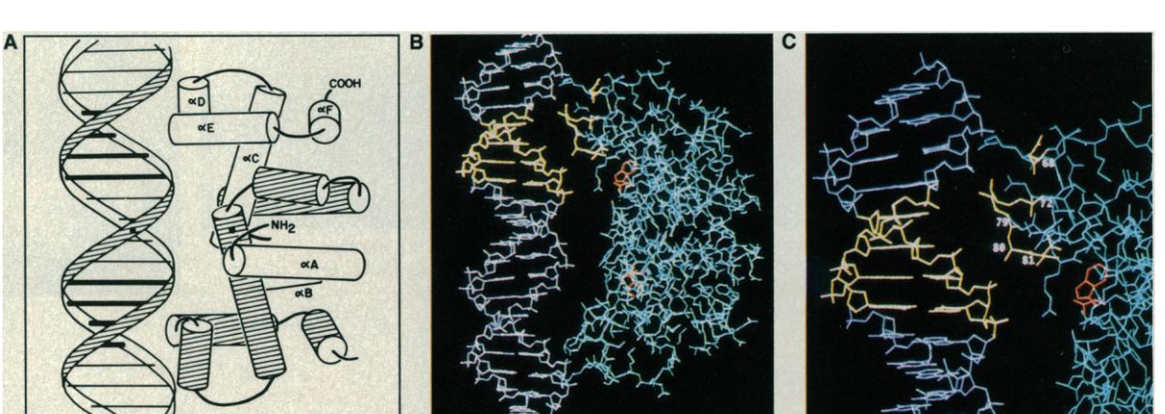

Fig. 1. Three views of a Trp repressor dimer docked with B-form DNA. (**A**) Schematic representation of the docked complex. The α-helices of Trp repressor are represented by cylinders, and the operator DNA is shown as a ladder with rungs representing base pairs. The four base pairs in each operator half-site critical for binding are indicated by thicker rungs; dots in the centers of each molecule designate dyad axes. (**B**) Representation of the complex from the trigonal crystal structure of Trp repressor (*8*). The four critical operator base pairs are highlighted in yellow, as are the amino acids changed by mutation. (**C**) Operator binding surface of the Trp repressor monomer. Residues Gln^{68}, Lys^{72}, Ile^{79}, Ala^{80}, and Thr^{81} are numbered. It is possible that in the bound repressor-operator complex, the DNA is unwound or kinked near the center of symmetry of the operator, and has a secondary structure significantly distinct from the B form, as is the case for the bound Eco RI and phage 434 repressor complexes (*1, 2*). Distortion of the central region of the operator might be favored by the hydrophobic stacking of the four central thymine methyl groups (*36*), and facilitate a more intimate alignment of the critical operator base pairs with the hydrophilic surfaces of α-helices D and E.

validity of these models, as well as those derived directly from cocrystal structures, residues in their ability to satisfy genetic tests for the presence in vivo of interactions between specific amino acids and base pairs.

One way to identify protein-DNA contacts is to isolate repressor mutations that are second-site suppressors of (constitutive) mutations that inactivate an operator. These mutant repressors bind some mutant operators better than wild-type repressor does, but they bind other operator sites less well than wild-type repressor; such mutant repressors are said to have "altered" or "new" specificities of binding. New specificity mutations must either create new contacts between mutant repressors and mutant operators, or eliminate unfavorable contacts between wild-type repressor and a mutant operator, and often pinpoint new weak chemical bonds between mutant amino acids and mutant base pairs. Repressors with new specificities have revealed specific contacts between *Salmonella* phage P22 Mnt repressor and its operator (*15, 16*) and between coliphage 434 repressor and its preferred site (*17*). In the second case, the specific interaction observed after mutation confirms an interaction predicted from the structure of the 434 repressor-DNA cocrystal (*2*).

Trp repressor. The *E. coli trpR* gene encodes a polypeptide 108 amino acids long that dimerizes to form Trp aporepressor. The aporepressor dimer binds two molecules of L-tryptophan, its corepressor, to form active repressor (*18–20*). Trp repressor is a global regulator; it binds at least three sites on the *E. coli* genome to inhibit transcription of the *trp* operon (*21*), the *aroH* gene (*22, 23*), and the *trpR* gene itself (*18*).

The crystal structures of Trp repressor show that the repressor monomer contains six α-helices joined by short surface turns, to which a disordered NH_2-terminal arm of 11 residues is attached (*8, 9*) (Fig. 1). Two α-helices, D and E, correspond to the conserved "helix-turn-helix" secondary substructure characteristic of many prokaryotic DNA-binding proteins (*13, 24*) and are positioned on the surface of Trp repressor to contact two successive major grooves of *trp* operator DNA. Models for the binding of other repressors suggest that the primary determinants of sequence specificity should reside in α-helix E of Trp repressor, the "recognition helix," and include the side chains of its first two amino acids (*13*). Unlike the case for many other repressors, however, the first two residues of α-helix E of Trp repressor, Ile^{79} and Ala^{80}, have a distinctly hydrophobic character. Few specific hydrophobic interactions are possible with functional groups in the major groove of DNA.

To begin a genetic analysis of the Trp repressor-operator interaction, we synthesized and cloned all variant operators differing by a symmetric pair of base pair substitutions from a consensus (reference-type) *trp* operator. The interactions of wild-type Trp repressor with each of these 28 operators in vivo reveal that four adjacent base pairs of a half-site are critical for binding (Fig. 2) (*14*). These results and others (*8, 9, 24, 25*) led us to propose a model for the specific interactions between Trp repressor and these four base pairs (*14*) (Fig. 3). This model predicts that, unlike the cases for λ repressor, Cro protein, and 434 repressor, specific contacts between Trp repressor and its operator are made with side chains from both the first α-helix (D) and the second α-helix (E) of the helix-turn-helix substructure. To assess this model, we have made a genetic analysis of *trpR* by constructing mutant Trp repressors with amino acid changes at the first three positions of its putative "recognition helix," to see how this region of the protein is involved in DNA binding.

Codon-directed mutagenesis of the *trpR* gene. To isolate mutant Trp repressors with amino acid substitutions at positions Ile^{79}, Ala^{80}, and Thr^{81}, we combined the site-directed mutagenesis of the corresponding codons on a plasmid-borne *trpR* gene with two genetic enrichments, one of which is for negative dominant mutant

Table 1. Changes of amino acids 68, 72, 79, 80, and 81 alter the DNA-binding properties of Trp repressor: Amino acid substitutions for residues Gln^{68}, Tyr^{72}, Ile^{79}, Ala^{80}, and Thr^{81} of Trp repressor obtained by codon-directed mutagenesis are listed, as are their mutant codons and frequencies of isolation. Mutant repressors have one of three DNA-binding phenotypes. Negative dominant mutants (−d) fail to bind any of the operators; mutants with restricted specificities (rs) bind a proper subset of operators bound by wild-type (wt) repressor; and mutants with new specificities (bs) bind some operators better than wild-type (wt) repressor. The efficiency of survival (EOS) is the percentage of cells that survive infection by a challenge phage with operator 11T (TCGAACTAGTT·AACTAGTTCGA), one of two reference types (*14*); many mutant repressors retain affinity for this operator. Values not shown are less than 10^{-5}. Repressors designated with asterisk (*) were isolated after spontaneous or ultraviolet mutagenesis. The dagger (†) indicates that the mutation was made by site-directed mutagenesis with the use of the following oligonucleotides: 5′-CGGCGCAGGCGCCGCGACGATTA-3′ (Ala^{79}) or 5′-CAGCGT-GAGTTGGCCAATGAACTCGGC-3′ (Ala^{72}).

Substitution	Codon	Isolates	Phenotype	EOS	Substitution	Codon	Isolates	Phenotype	EOS	Substitution	Codon	Isolates	Phenotype	EOS
	α-Helix D									*α-Helix E*				
Gln^{68}	CAG		wt	50	Ile^{79}	ATC		wt	50	Thr^{81}	ACG		wt	50
Cys^{68}	TGC	1	−d		Asp^{79}	GAC	13	−d		Ala^{81}	GCG	5*	−d	
Lys^{68}	CTC	1	−d		Glu^{79}	GAG	6	−d		Asp^{81}	GAC	3	−d	
	AAG	4			Lys^{79}	AAG	1	bs		Glu^{81}	GAG	1	−d	
Leu^{68}	CCC	1	−d		Phe^{79}	TTC	2	bs	40	Gly^{81}	GGG	1	−d	
	CTG	6			Pro^{79}	CCC	1	−d			GGC		3	
Met^{68}	ATG	1	−d		Arg^{79}	AGG	2	bs	10	Leu^{81}	CTC	1	−d	
Pro^{68}	CCG	5	−d		Tyr^{79}	TAC	1	rs	1	Pro^{81}	CCC	2*	−d	
	CCC	1			Trp^{79}	TGG	1	rs	0.1	Arg^{81}	CGC	1	−d	
Arg^{68}	TGC	1	−d		Ala^{79}	GCC	1†	rs	20	Ser^{81}	TCG	1	bs	50
	CGG	4			Ala^{80}	GCG		wt	50	Val^{81}	GTG	3	−d	
am^{68}	TAG	1*	−d		Cys^{80}	TGC	1	rs	0.0005		GTC	2		
Lys^{72}	AAA		wt	50	Lys^{80}	AAG	6	−d		am^{81}	TAG	2	−d	
Ala^{72}	GCC	1†	−d		Pro^{80}	CCG	1	−d						
Asp^{72}	GAC	2	−d		Arg^{80}	AGG	4	−d						
Glu^{72}	GAG	6	−d			CGG	7							
	GAA	3				CGC	2							
Asn^{72}	AAC	2	−d		Thr^{80}	ACG	1	bs	20					
Tyr^{72}	TAC	1	rs	10	Val^{80}	GTG	3	rs	2					
am^{72}	TAG	2	−d		Trp^{80}	TGG	1	−d						
oc^{72}	TAA	1	−d		am^{80}	TAG	3	−d						

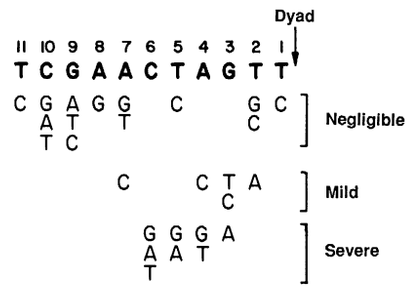

Fig. 2. Critical operator determinants of wild-type Trp repressor binding. The left half-site of an idealized Trp operator sequence is numbered by position from the axis of symmetry (arrow). Below the reference sequence are listed the changes that have negligible, mild, or severe effects on binding of wild-type repressor (Table 2). (*14*). Mutations that create operators bound tightly by *Salmonella* Trp repressor (8T, 8C, 1G, and 1A) are not shown.

repressors (that bind the natural *trp* operator with less affinity than the wild-type Trp repressor) (*26–30*). The sequence changes obtained for 84 independent mutants with amino acid changes at positions 79, 80, and 81, are summarized in Table 1.

Previously, we have used challenge phages to describe the interaction between wild-type Trp repressor and consensus and variant *trp* operators (*14, 31*). Trp challenge phages are derivatives of *Salmonella* phage P22 carrying one of 28 synthetic *trp* operators that differ from a reference-type operator by pairs of symmetric substitutions. The *trp* operators are positioned to regulate expression of the P22 antirepressor (*ant*) gene, which must be repressed for the phage to establish lysogeny. Phages with operators that retain the critical determinants of binding lysogenize a host that produces Trp repressor; phages with operators not bound by Trp repressor kill such hosts. The efficiency of lysogeny of a Trp challenge phage is a measure of how well its operator is bound.

Symmetric variant *trp* operators fall into three classes in terms of their ability to bind wild-type repressor (Fig. 2). Each variant operator is designated by the position of its symmetric change and the variant nucleotide in the top strand of the left operator half-site; positions in the half-site are numbered outward from the center of symmetry (for example, the top strand of operator 4T is 5′-CGAACTTGTTAACAAGTTCG-3′) (*14*). Many variant sites are bound as well as the reference types (11T and 11C), some are bound slightly less well, and eight variant sites are not bound. The eight operators that are not bound have changes that define the four base pairs critical for the binding of wild-type repressor (*14*).

Mutant repressors with changes of Ile^{79}, Ala^{80}, and Thr^{81} that we have isolated may be categorized by their abilities to bind the set of *trp* operators (Table 1). Fifteen of the 24 mutant repressors do not bind any operators. These mutants have a "negative dominant" phenotype; when plasmids that produce these mutant repressors are transformed into $trpR^+$ *E. coli* (or *Salmonella*) strains, the resulting meropolyploids are resistant to the tryptophan analogue, 5-methyltryptophan (*21, 24*). Presumably, the mutant proteins can fold, dimerize, and form heterodimers with wild-type repressor that cannot bind the natural *trp* operator (or corepressor). Amino acid changes that lead to a negative dominant phenotype include both changes that introduce negatively charged or large, bulky side chains onto the binding surface of the repressor, and changes that are likely to disrupt the secondary structure of α-helix E.

Nine mutant repressors recognize spectra of the set of variant operators different from the spectrum recognized by the wild type (Table 2). Four mutant repressors (Ile^{79}→Trp, Ile^{79}→Tyr, Ala^{80}→Cys, and Ala^{80}→Val) bind a proper subset of the operators bound by wild-type Trp repressor, and have restricted binding specificities. Changes that lead to restricted specificities of binding must perturb the repressor-operator interaction only slightly. Five mutant repressors (Ile^{79}→Arg, Ile^{79}→Lys, Ile^{79}→Phe, Ala^{80}→Thr, and Thr^{81}→Ser) have new binding specificities, and recognize some sites better than wild-type repressor, yet the binding of wild-type repressor is better at other sites. Mutations that create new specificities either create new contacts between the repressor and operator, or eliminate negative interactions between the wild-type amino acid and wild-type operator.

Changes of Ile^{79}, Ala^{80}, and Thr^{81}. Among eight different mutant repressors with changes of Ile^{79}, three (Ile^{79}→Phe, Ile^{79}→Lys, and Ile^{79}→Arg) have new and two (Ile^{79}→Trp and Ile^{79}→Tyr) have restricted specificities (Table 2). Four of these mutants (Ile^{79}→Arg, Ile^{79}→Phe, Ile^{79}→Trp, and Ile^{79}→Tyr) retain nearly normal abilities to bind the reference-type operators in vivo (Table 1) and the natural *trp* operator in vitro (*32*).

The mutant Ile^{79}→Phe repressor has a new specificity. It binds the reference operator as well as wild-type repressor, binds less well to most variant operators, yet recognizes the 4T operator better than wild type. The large, hydrophobic side chain of Phe^{79} may participate in a new van der Waals interaction with the thymine methyl group of the variant T·A base pair at position 4. Unlike the Ile^{79}→Phe repressor, the Ile^{79}→Tyr repressor does not bind operator 4T. Thus, the addition of a hydroxyl group to carbon-4 of the aromatic ring of Phe^{79} disrupts this new interaction.

The Ile^{79}→Lys repressor also has a new specificity, and shows dramatically reduced binding to most operators, yet binds operator 8G with high affinity, and operator 7C better than wild type. The basic ε-amino group of Lys^{79} may form new specific H bonds with the operator. The mutant Ile^{79}→Arg repressor binds operators 7C and 8G, yet retains the ability to bind the reference-type operators. Arginine has a long, complex side chain featuring a guanidino group that may donate one or more new H bonds to the operator in a way akin to the ε-amino group of lysine. Apparently, amino acid side chains of the Lys^{79} and Arg^{79} mutant repressors may extend to contact base pairs 7 and 8, which are not likely to be involved in specific bonds with wild-type repressor (*14*). The variant 7T operator has a change that creates a Dam (DNA adenine methylation) recognition site (5′ GATC 3′) in each *trp* operator half-site. The carbon-6 amino groups of both adenines in base pairs 7 and 8 are methylated in the major groove of this site (*33*), which is bound well by wild-type repressor. In contrast, neither the Ile^{79}→Lys nor the Ile^{79}→Arg repressor binds the 7T site.

Three of seven mutant repressors with changes of Ala^{80} have restricted or new specificities. The Ala^{80}→Cys and Ala^{80}→Val repressors show reduced abilities to bind most operators. The phenotypes of these changes suggest that Ala^{80} is positioned near the DNA, since substitutions of slightly larger side chains interfere with binding. Curiously, the Cys^{80} repressor has a new specificity; it retains affinity for the 5C operator, and prefers it to the consensus site.

It is particularly revealing to compare the DNA-binding properties of the Ala^{80} (wild-type), Thr^{80}, and Val^{80} repressors. The mutant Thr^{80} repressor shows decreased binding to most operators but has a new specificity; it binds operators 4G and 4T, which are not bound by wild-type repressor. The Thr^{80} repressor may make new contacts with variant G·C and T·A base pairs at position 4, and a variant A·T base pair at position 3 of the operator. The Thr^{80} hydroxyl group is a good H-bond donor. It may donate a proton to the congruent acceptor groups of G·C and T·A at position 4 (guanine O6 and thymine O4), and perhaps less well to the thymine O4 of A·T at position 3, positions that are proximate in three dimensions. When this Thr^{80} hydroxyl is replaced by a methyl group, operators 4G, 4T, and 3A are no longer bound better than by the wild-type repressor; the mutant Val^{80} repressor has a restricted specificity. Alternatively, the mutant Ala^{80}→Thr change may create a new nonspecific (phosphate backbone) contact that

compensates for negative interactions between wild-type repressor and these variant operators.

At position 81, all amino acid substitutions, except Thr^{81}→Ser, prevent binding of the repressor to the entire set of operators. These mutants include Thr^{81}→Ala, missing the hydroxyl and methyl groups of Thr^{81}, and Thr^{81}→Val, which replaces the Thr^{81} hydroxyl with a methyl group. These results indicate that the Thr^{81} hydroxyl group is critical for binding, and may make a phosphate backbone (nonspecific) contact with DNA, or may make a specific contact. We favor the specific contact, and propose that the Thr^{81} hydroxyl group forms a critical H bond with the thymine O4 group of the A·T base pair at position 4 (Fig. 3). This could account for our finding that wild-type repressor requires either an A·T or C·G at this operator position (*14*) (Fig. 2), because both base pairs have similarly positioned H-bond acceptor groups in the major groove (thymine O4 and guanine O6).

The mutant Ser^{81} repressor has reduced affinity for some variant operators. This result indicates that the methyl group of Thr^{81} may contribute to a modest hydrophobic contact in the interaction of the wild-type repressor and operator, that it may help position the hydroxyl group favorably, or that it may contribute in a more general way to the structure of the bound complex. The Ser^{81} repressor has a new specificity; it binds better to operator 3A than wild-type repressor, suggesting that the Thr^{81} methyl group is responsible for the specificity seen at position 3. The proposed interaction between the Thr^{81} hydroxyl group and thymine O4 at base pair 4 might position the Thr^{81} methyl group near the C-5 atom of cytosine at adjacent base pair 3, in a steric clash with the thymine methyl group of an A·T base pair at this position. When Thr^{81} is replaced by Ser, this clash is mitigated by the loss of the methyl group.

Ile^{79} and Ala^{80} are not involved in specific binding. In summary, we have tested each of 24 mutant Trp repressors with changes in the first three residues of α-helix E, Ile^{79}, Ala^{80}, and Thr^{81}, for its ability to bind each of a set of 28 *trp* operators in vivo, as judged from the P22 challenge phage assay. The analysis of these 672 interactions indicates that more subtle changes in Trp repressor at a position critical for DNA binding (Thr^{81}) have more drastic effects

Table 2. Interactions of mutant Trp repressors with variant *trp* operators in vivo. Symmetric operators carried by P22 *Kn9 arc-am*H1605 challenge phages are designated by operator position and variant nucleotide. Operator positions are numbered outward from the center of symmetry in the top strand of the left operator half-site (Fig. 2); reference type *trp* operators for each position are indicated by an asterisk. Variant operators differ by symmetric pairs of substitutions from the reference-types (for example, operator 6G is TCGAAGTAGTT·AACTACTTCGA). The reference type operator for positions 1 to 7 is shown in Fig. 2; the reference type for positions 8, 9, and 10 carries the 11C symmetric change (*14*). The binding strength of each repressor, wild-type (wt) or mutant, to a particular operator is given as the efficiency of survival (EOS), the percentage of repressor-producing cells (*Salmonella* strain MS1868/F′*lacI*^{Q} (*31*) carrying a derivative of plasmid pPY150) that survive infection with a challenge phage carrying that operator (*14*). The EOS was not determined for four operators (8T, 8C, 1G, and 1A) that are bound tightly by *Salmonella* Trp repressor. Results indicating new binding specificities are indicated in bold face and underlined. The data represent the average from at least three independent determinations; measurements varied less than fivefold from experiment to experiment. Values not shown are less than 10^{-5}.

Operator	wt	Tyr^{72}	Arg^{79}	Lys^{79}	Phe^{79}	Trp^{79}	Tyr^{79}	Ala^{79}	Cys^{80}	Thr^{80}	Val^{80}	Ser^{81}
10G	30				0.0001	0.02		2		0.1	0.0007	0.1
10A	40	0.0003		0.0001	0.002	0.008	0.0005	20		0.1	0.007	0.1
10T	40	40	40	0.0002	30	0.1	40	20	5	30	20	20
10C*	50	10	40	0.0001	50	5	10	20	0.003	30	10	50
9G*	50	10	40	0.0001	50	5	10	20	0.003	30	10	50
9A	30	0.2			10	0.01	40	20		0.2	0.05	0.3
9T	20					0.006		0.0002		0.03		
9C	30	0.00002			0.04	0.002		0.6		0.005	0.0008	0.01
8G	40	0.2	30	40	0.00002	0.0002		30		20	0.1	0.2
8A*	50	10	40	0.0001	50	5	10	20	0.003	30	10	50
7G	40	0.00004						0.00004		0.02		
7A*	50	10	10		40	0.1	1	20	0.003	30	10	50
7T	30							0.00005		0.003		
7C	0.2	0.00003	**50**	**40**		0.01		0.2		0.02	0.00003	0.08
6G												
6A												
6T												
6C*	50	10	10		40	0.1	1	20	0.003	30	10	50
5G												
5A												
5T*	50	10	10		40	0.1	1	20	0.003	30	10	50
5C	30	0.00003	0.007		30	5	20	20	10	20	0.06	0.4
4G										**0.0004**		
4A*	50	10	10		40	0.1	1	20	0.003	30	10	50
4T					**0.02**					**0.001**		
4C	3				0.003	0.02		10		0.02	0.0003	0.04
3G*	50	10	10		40	0.1	1	20	0.003	30	10	50
3A	0.00005	0.00005						0.00009		**0.0005**		**0.06**
3T	6	0.00003						0.002		0.1	0.0006	0.3
3C	0.5	0.00007				0.00004		0.01		0.1	0.00005	0.2
2G	20							0.00005		0.0002		0.4
2A	0.7							0.00002				0.07
2T*	50	10	10		40	0.1	1	20	0.003	30	10	50
2C	60	0.0001	0.006			0.00005		0.1		0.06	0.0003	0.3
1T*	50	10	10		40	0.1	1	20	0.003	30	10	50
1C	50	20	0.0001	0.00003	0.006	0.0005	0.001	20	0.004	40	20	0.06

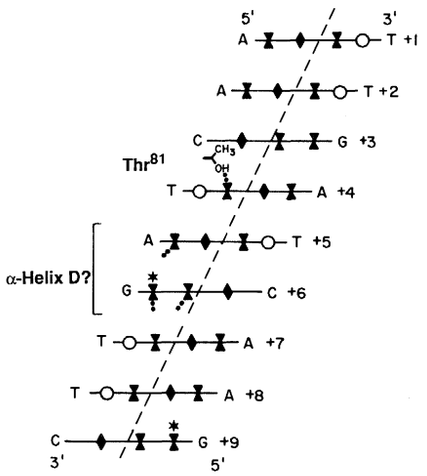

Fig. 3. Model for Trp repressor binding. The functional groups of the base pairs in the major groove of the reference-type *trp* operator are shown in two dimensions (*38*). The symbols indicate DNA functional groups as H-bond donors (♦), H-bond acceptors (⧗), and thymine methyl groups (○). Presumed H bonds between repressor side chains and the bases critical for binding (6, 5, 4, 3) are indicated as dotted lines. Guanines with N7 groups protected from methylation in vitro by dimethylsulfate upon binding of wild-type repressor are marked by stars (*26, 34*). The side chain of one residue from α-helix E, Thr^{81}, appears to make specific contacts with base pairs 4 and 3 of the operator. Residues in α-helix D may recognize base pairs 6 and 5 by making the three indicated H bonds, which suffice to account for the observed specificity at these positions. An important feature of this model is that a Trp repressor monomer specifically contacts one and only one strand of each operator half-site. This geometric constraint on repressor binding may enable the simultaneous recognition of the opposite strands of four critical base pairs in an operator half-site by two different repressor monomers belonging to two different dimers, and account for the ability of Trp repressor to bind multiple, tandem sites (*25*).

than more gross changes at positions not directly involved in binding (Ile^{79} and Ala^{80}). Although the first two amino acids of this helix, Ile^{79} and Ala^{80}, are predicted to be on the binding surface of Trp repressor by analogy with other repressors (*13*) and by modeling the repressor crystal structure to DNA (*8*) (Fig. 1), it is unlikely that their side chains make specific DNA contacts in the wild-type interaction. In contrast, the third amino acid of α-helix E, Thr^{81}, appears to make at least one specific contact. For Ile^{79} and Ala^{80}, we have obtained changes that lead to restricted specificities of binding and involve the substitution of side chains with larger specific volumes. Their changes have subtle but significant effects on repressor binding, indicating that Ile^{79} and Ala^{80} are not directly involved in specific operator contacts, but that lengthening the side chains at these positions may weaken binding by creating steric problems.

To demonstrate decisively that Ile^{79} plays little role in specific binding, we constructed the mutant Ile^{79}→Ala repressor. Of all of the mutant repressors with changes of residue 79, Ala^{79} has the specificity closest to wild type (Table 2). Also, Ala^{80} does not appear to make specific contacts with the *trp* operator. If it did, we would expect that the Ala^{80} methyl would participate in a van der Waals interaction with a thymine methyl and that substitutions of larger side chains for the small alanine side chain would disrupt such an interaction. The Ala^{28} methyl of a mutant 434 repressor makes just such a contact; substitutions of Ser or Leu for Ala^{28} disrupt this contact, as predicted (*17*). In contrast, we find that mutant Trp repressors with changes that replace the Ala^{80} side chain with slightly larger side chains (for example, Ala^{80}→Cys, Ala^{80}→Thr, or Ala^{80}→Val) still bind many variant *trp* operators. It takes more drastic changes at this position (for example, Ala^{80}→Lys, Ala^{80}→Arg, or Ala^{80}→Trp) to disrupt binding.

For Ile^{79} and Ala^{80}, we have also obtained changes that confer new binding specificities. A simple interpretation of these mutant repressors is that Ile^{79} and Ala^{80}, although not in direct contact with the operator, are close enough so that residues with extended side chains at these positions may make new bonds with the operator.

Trp repressor uses its "recognition helix" in a novel way. Trp repressor must recognize DNA in a way fundamentally different than that for λ repressor, Cro protein, and 434 repressor. For both λ repressor and Cro protein, the first two residues of their "recognition α helices" are Gln and Ser, which share in H bonds with base pairs 2 and 4, and base pairs 2 and 3, of the λ operator, respectively (*3–5, 13, 34, 35*). The first two amino acids of this helix of 434 repressor, Gln^{28} and Gln^{29}, make specific contacts with the 434 operator (*2, 17*). The Trp repressor, unlike these phage repressors, does not appear to use the first two hydrophobic residues of its so-called "recognition helix" to make specific DNA contacts.

Mutant Trp repressors with changes in residues Ile^{79}, Ala^{80}, and Thr^{81} that have new binding specificities retain the same specificity as the wild-type repressor at operator positions 5 and 6, consistent with the idea that these amino acids are not involved in recognition of these two critical base pairs. Two other regions of repressor, its 11-residue NH_2-terminal arm and α-helix D, might recognize these base pairs. Since a deletion that removes the NH_2-terminal arm of Trp repressor results in a protein that behaves like wild type in its ability to protect the *trp* operator from methylation (*32*), the first alternative appears unlikely.

These results lead us to suggest that side chains of amino acids on the hydrophilic surface of α-helix D, including Gln^{68}, Arg^{69}, Lys^{72}, or Asn^{73} (or any combination of these), may interact with these base pairs (Fig. 3). Alternatively, other residues such as Arg^{54} and Arg^{84}, which form part of the corepressor binding pocket, may contribute to specificity. To test these hypotheses, we have isolated mutations that change these amino acids and have characterized their effects on DNA binding. Residues Gln^{68} and Lys^{72} are critically involved in binding, since conservative changes of these residues either prevent binding (such as Gln^{68}→Met, Gln^{68}→Leu, Lys^{72}→Asn, Lys^{72}→Ala), or restrict specificity (Lys^{72}→Tyr) (Tables 1 and 2). However, whether the contributions of these amino acids to binding are specific or nonspecific remains to be determined.

About one-tenth of mapped *Escherichia coli* and coliphage genes encode specific DNA-binding proteins; thus, the ability to bind specific DNA sequences is a shared function of a very large group of proteins. It is not surprising that individual members of this large group would exhibit a wide range of both structural and functional diversity.

REFERENCES AND NOTES

1. J. A. McClarin *et al.*, *Science* **234**, 1526 (1986).
2. J. E. Anderson, M. Ptashne, S. C. Harrison, *Nature* **326**, 746 (1987).
3. S. R. Jordan and C. O. Pabo, *Science*, in press.
4. D. H. Ohlendorf, W. F. Anderson, R. G. Fisher, Y. Takeda, B. W. Matthews, *Nature* **298**, 718 (1982).
5. W. F. Anderson, D. H. Ohlendorf, Y. Takeda, B. W. Matthews, *ibid.* **290**, 754 (1981).
6. D. B. McKay and T. A. Steitz, *ibid.*, p. 744.
7. D. B. McKay, I. T. Weber, T. A. Steitz, *J. Biol. Chem.* **257**, 9518 (1982).
8. R. W. Schevitz, Z. Otwinowski, A. Joachimiak, C. L. Lawson, P. B. Sigler, *Nature* **317**, 782 (1985).
9. C. L. Lawson *et al.*, *Proteins* **3**, 18 (1988).
10. D. H. Ohlendorf, W. F. Anderson, M. Lewis, C. O. Pabo, B. W. Matthews, *J. Mol. Biol.* **169**, 757 (1983).
11. R. H. Ebright and J. R. Wong, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 4011 (1981).
12. F. R. Salemme, *ibid.* **79**, 5263 (1982).
13. C. O. Pabo and R. T. Sauer, *Annu. Rev. Biochem.* **53**, 293 (1984).
14. S. Bass, P. Sugiono, D. N. Arvidson, R. P. Gunsalus, P. Youderian, *Genes Dev.* **1**, 565 (1987).
15. P. Youderian, A. Vershon, S. Bouvier, R. T. Sauer, M. M. Susskind, *Cell* **35**, 777 (1983).
16. A. K. Vershon, P. Youderian, M. A. Weiss, M. M. Susskind, R. T. Sauer, in *Sequence Specificity in Transcription and Translation*, R. Calendar and L. Gold, Eds. (Liss, New York, 1985), p. 209.
17. R. P. Wharton and M. Ptashne, *Nature* **326**, 888 (1987).
18. R. P. Gunsalus and C. Yanofsky, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 7117 (1980).
19. D. N. Arvidson, C. Bruce, R. P. Gunsalus, *J. Biol. Chem.* **261**, 238 (1985).
20. A. Joachimiak, R. L. Kelley, R. P. Gunsalus, C. Yanofsky, P. B. Sigler, *Proc. Natl.*

Acad. Sci. U.S.A. **80**, 668 (1983).
21. G. N. Bennett and C. Yanofsky, *J. Mol. Biol.* **121**, 179 (1978).
22. G. Zurawski, R. P. Gunsalus, K. D. Brown, C. Yanofsky, *ibid.* **145**, 47 (1980).
23. C. L. Grove and R. P. Gunsalus, *J. Bacteriol.* **169**, 2158 (1987).
24. R. L. Kelley and C. Yanofsky, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 483 (1986).
25. A. A. Kumamoto, W. Miller, R. P. Gunsalus, *Genes Dev.* **1**, 556 (1987).
26. To mutagenize the *trpR* gene, single-stranded plasmid pPY150 (*14*) DNA, prepared by M13 infection of *E. coli* strain BW313 (*27*) carrying pPY150, was annealed with Bam HI–cleaved plasmid pPY158 DNA to form a template with a single-stranded gap missing the sense strand of *trpR*. To confine the mutagenesis to a target codon, we synthesized mixed primers 23 bases long. These primers are complementary to two ten-base stretches of DNA preceding and following a target codon, and have the mixed sequence, 5′ NNS 3′, in place of the target codon (N represents a mixture of all four bases, and S a mixture of G and C). Because the genetic code is degenerate, the mixture NNS represents all 20 amino acids by 31 codons (the 32nd is TAG). Primed template DNA's were filled in, treated with T4 DNA ligase, and used to transform X90/F′*lacI*^{Q1} (*28*) to an ampicillin-resistant phenotype. Inheritance of the repaired strand was enriched by transformation into this restricting (*dut*$^{+}$ *ung*$^{+}$) host (*27*). Transformants carrying plasmids with negative dominant *trpR*$^{-}$ mutations were enriched for their ability to form larger colonies on media containing the corepressor analogue, 5-methyltryptophan (*21, 24*). Single-stranded DNA's from mutant plasmids (*29*) were sequenced (*30*) with [α^{32}P]dATP (ICN) as label. Synthetic oligonucleotides were made on an Applied Biosystems automated DNA synthesizer with phosphoramidite substrates.
27. D. Kunkel, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 488 (1985).
28. E. Amann, J. Brosius, M. Ptashne, *Gene* **25**, 167 (1983).
29. J. R. Zagursky and M. L. Berman, *ibid.* **27**, 183 (1984).
30. F. Sanger, S. Nicklen, A. Coulson, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463 (1977).
31. N. Benson, P. Sugiono, S. Bass, L. V. Mendelman, P. Youderian, *Genetics* **114**, 1 (1986).
32. A. A. Kumamoto and R. P. Gunsalus, unpublished results.
33. M. C. Gomez-Eichelmann, *J. Bacteriol.* **140**, 574 (1979).
34. N. Benson and P. Youderian, *Genetics*, in press.
35. Y. Takeda, A. Sarai, F. H. Hausheer, *Proc. Natl. Acad. Sci. U.S.A.*, in press.
36. These graphics were produced with the USC molecular graphics program by F. Lee, J. Hwang, and A. Warshel. Coordinates were provided by Z. Otwinowski and P. B. Sigler.
37. R. L. Jernigan, A. Sarai, A., K.-L. Ting, R. Nussinov, *J. Biomol. Str. Dyn.* **4**, 41 (1986).
38. C. P. Woodbury, Jr., O. Hagenbuckle, P. H. von Hippel, *J. Biol. Chem.* **255**, 11534 (1980).
39. We thank R. Gunsalus, G. Mosig, and M. Susskind for critical comments on the manuscript. Supported by NIH grant GM34150 (P.Y.)

13 June 1988; accepted 13 September 1988