A permutational approach toward protein–DNA recognition

LIXUAN HUANG, TAKASHI SERA, AND PETER G. SCHULTZ

Department of Chemistry, University of California, Berkeley, CA 94720

Contributed by Peter G. Schultz, January 18, 1994

ABSTRACT The cI repressor of bacteriophage 434, known as 434 repressor, binds to 14-bp operator sequences by means of a helix–turn–helix motif. To probe the requirements for selective DNA recognition by this class of DNA binding proteins, as well as to generate new proteins with altered specificities, a library of \( \sim 3 \times 10^6 \) mutants was generated that contains all permutations of five residues in the recognition helix (helix 3) of the repressor. These mutants were then selected in vivo for their ability to bind both wild-type (WT) and operator mutant sequences. The results of the selection demonstrate that four of these residues—Gln29, Gin29, Ser30, and Gin33—play a critical role in recognition of the WT operator. A number of repressors with mutations at Thr27 showed altered DNA binding affinities and specificities. The approach described here may also prove useful in studies of DNA recognition by other classes of DNA binding proteins.

A number of classes of DNA binding proteins have been identified, including helix–turn–helix, zinc finger, leucine zipper, and \( \beta \)-ribbon proteins, in which specificity for a given sequence of DNA appears to be dictated to a large degree by hydrogen bonding interactions between amino acid side chains and the base functionality in the major and minor grooves of DNA (1–3). One such example is 434 repressor, a 209-amino acid protein that specifically binds as a dimer to several 14-bp operator sequences in the phage 434 genome by means of a helix–turn–helix motif (4).

The structure of the N-terminal, DNA-binding domain of the repressor complexed with the wild-type (WT) operator sequence O1 has been determined to 2.5 Å resolution by x-ray crystallography (5). The crystal structure reveals that repressor binding to O1 (5'-ACAAAGAGTGTGTTG-3') depends largely on residues in or near helices 2 and 3. Helix 3 from each monomer is inserted into the major groove of the DNA, while helix 2 lies across the groove. The N termini of helices 2 and 4 and the loop between helices 3 and 4 provide favorable interactions with the sugar phosphate backbone. The specificity of 434 repressor for operator DNA is thought to be determined to a large degree by hydrogen bonding and van der Waals interactions between the side chains of residues in helix 3 and the base pairs 5'-ACAATGGTTG-3' of the operator DNA. Binding of the repressor to DNA also results in significant distortion from canonical B-DNA including a 25° bend and overwinding of the minor groove at the center of the operator (5).

A number of mutagenesis studies have been carried out to probe the role of the residues in helix 3 of 434 repressor in determining DNA binding specificity. These studies include the generation and characterization of a chimeric 434 repressor containing the recognition helix of P22 repressor (6) as well as site-directed mutagenesis studies of individual residues (7). To gain further insight into the requirements for sequence-specific recognition of DNA by 434 repressor, we have generated a library of \( 3.2 \times 10^6 \) repressor mutants in which residues Thr27, Gin29, Gin29, Ser30, and Gin33 have been randomized. These mutants were then selected for their ability to bind WT operator and a number of mutant operators using an in vivo transcription interference assay (8). The DNA binding properties, both in vivo and in vitro, of a number of selected mutants have been characterized.

MATERIALS AND METHODS

Bacterial Strains, Plasmids, and Reagents. Cells were grown in Bacto-tryptone (20 g/liter), Bacto-yeast extract (5 g/liter), NaCl (0.5 g/liter), 2.5 mM KCl, 10 mM MgCl2, and 20 mM glucose (SOC medium). Escherichia coli MC1061 [hsdR, mcrI3, araD139, (araABC-leu)7679, ΔlacX74, galU, galK, rpsL, thi−] was used for plasmid library construction, E. coli HB101 (F'LacI) was used as the host strain for the selection experiments, and E. coli XA90 [Δ(lac-pro)XIII, ara−, nala, argE, am, thi−, rfr, (F', z−, y+, proA'B', LacI(39))] was used for expression of WT and mutant repressors. HB101 (F' lacI) containing the reporter plasmid was generated by crossing HB101 [supE4, 44 hsdS, 20(rgm)] with HB101 [supE4, 44 hsdS, 20(rgm)] containing the lac operon to form a viable strain. The lac operon was then cloned into the vector pLH17 and expressed in E. coli DH5α to generate the plasmid pDS61. This plasmid was then used to transform the E. coli strain to obtain the desired mutants.

Construction of 434 Repressor Library. Plasmid pLH17 was constructed by introducing Kpn I and Not I sites into pDS61 by site-directed mutagenesis (Amersham mutagenesis kit). A synthetic DNA linker (5'-CACAGTGATCGATGATG-3') containing Kpn I and Not I ends and an internal BamHI site was then added to the plasmid to generate pDS61KpnI. This plasmid was then used for library construction. A 65-nt oligonucleotide with the sequence 5'-GCTTTCGCTTGGTACC CTK NNN KNN NNK NNK ATG GAG NNN CTC GAA AAT GGT AAA ACT AAG CGG CGG CGG T-3' [where N is A, C, G, or T (equimolar) and K is G or C (1:1)] encoding the helix 3 library was then synthesized by using a standard on-line mixing protocol for K and a commercially available mixture for N (Applied Biosystems). The primer 5'-ACAC CGCA CGCG CGCG TTG ATG TTT ACC-3' (11.3 μg) was then annealed to 5.2 μg of the 65-mer by heating to 65°C for 5 min and cooling to room temperature for 10 min. The primer was extended by using Klenow fragment of DNA polymerase I and the resulting duplex DNA was digested with Kpn I and Not I and purified by 12.5% polyacrylamide gel electrophoresis. This fragment (13 μg) was ligated to 17 μg of Kpn I/Not I digested pLH21 (purified by agarose gel electrophoresis) using 400 units of T4 DNA ligase in a total vol of 300 μl of

Abbreviations: WT, wild type; Cm', chloramphenicol resistant; Ap', ampicillin resistant; Sp', spectinomycin resistant; IPTG, isopropyl β-D-thiogalactopyranoside.

*To whom reprint requests should be addressed.
ligase buffer (New England Biolabs). The ligation mixture was then digested with BamHI, divided into four batches, and transformed into freshly prepared E. coli MC1061 by electroporation (11). After 1 hr of nonselective growth, a library of 1.4 × 10^8 transformants per μg of DNA was obtained. The transformed cells were grown overnight in LB (4 × 1 liter) containing 50 μg of Ap per ml in order to amplify the library. Plasmids were isolated in pools of independent pools by the CsCl ultracentrifugation method (12). The library DNA was sequenced as a pool and revealed equal band intensities in each of the sequencing lanes at the randomized positions.

Construction of Reporter Plasmids and Operator-Containing Vectors. Reporter plasmid pConI-1, containing the WT operator, was constructed in two steps. A synthetic DNA fragment containing the operator with Kpn I and Pst I ends (5'-CTATACAGAGAAATGTTGACTTCGCA-3') was ligated into Kpn I/Pst I digested pNN396 (10) to yield plasmid pLH2. The Not I/HindIII fragment of pLH2 was then cloned into Not I/HindIII digested pNN388, which carries a Cam gene, to afford the plasmid pConI-0q. This plasmid contains the operator sequence inserted at the +4 position relative to the start of transcription from the ConII promoter.

Mutant reporter plasmids pConII-M1 and pConII-M2 corresponding to the operator sequences 5'-CTATACGA-GAAGTTGCTACCTTCGCA-3' and 5'-CTATATAGGAGAGTCTAGTTCTGCA-3', respectively, were similarly constructed. A reporter plasmid with the operator deleted was constructed by deleting the Kpn I/Not I sequence from pNN396 by restriction digestion and subsequent blunt-end ligation to give pConII-Op. The Not I/HindIII fragment of pConII-Op was cloned into pNN388 to yield pConII-ΔOR. Vectors containing operator sequences for gel-shift assays were constructed by ligating synthetic operator sequences with EcoRI and HindIII ends into EcoRI/HindIII digested pBluescript II KS' (Strategene). A control plasmid (pLHGS5) was constructed with the sequence 5'-CAGCTCTGTTCAAG-3' (Stratagene).

Purification of WT and Mutant Repressors. WT and mutant repressors were expressed in E. coli XA90 transformed with pConII-1T and pConII-0q. One liter of cells was grown to OD_600 = 1.0 at 37°C, induced with 1 mM IPTG for 6 hr, and lysed in a French pressure cell at 8400 psi (1 psi = 6.9 kPa) in the presence of the protease inhibitors leupeptin, antipain, pepstatin, and chymostatin (3 μg/ml each) and phenylmethylsulfonfyl fluoride (0.2 mg/ml) in cold lysis buffer (100 mM Tris-HCl, pH 7.9/0.5 M NaCl/1 mM EDTA/2 mM CaCl_2/10 mM MgCl_2/1.4 mM 2-mercaptoethanol/5% (vol/vol) glycerol). After treatment with polyethyleneimine chloride (0.6%) and precipitation with 20% (NH_4)_2SO_4, the pellet was resuspended in 50 mM Tris-HCl, pH 7.0/0.1 mM EDTA/1.4 mM 2-mercaptoethanol/10% glycerol buffer and purified by chromatography on a Bio-Rex 70 column, eluting with a gradient from 0 to 500 mM NaCl in the above buffer. The yield of protein was normally ~2 mg/liter and protein was >95% homogeneous as judged by SDS/PAGE.

In Vivo Selection Against WT and Mutant Operator Sequences. The library DNA (4 × 10 μg) was transformed into 4 × 400 μl of freshly prepared electrocompetent HB101 or HB101 (F' lacD) containing the reporter plasmid. After 2 hr of nonselective growth in SOC medium and 2 hr of selective growth in the presence of Ap (50 μg/ml) and 0.1 mM IPTG, the transformed cells were plated onto LB agar containing Ap (80 μg/ml), Sp (80 μg/ml), and 0.1 mM IPTG and incubated at 37°C for 2 days. The cells from the selection plate were pooled and the plasmid was isolated (Magic Maxi Prep, Promega). The pooled DNA was digested with BamHI (restriction site present only on the reporter plasmid) in order to eliminate any reporter plasmid in the DNA preparation. The selection protocol was repeated two or three times until an ~10^4-fold enrichment was obtained. Individual clones were then isolated, sequenced, and tested for their in vivo DNA binding activity by retransforming into the selection strain and comparing the titers on LB agar plates containing Ap and Sp to the titers on agar plates containing Ap.

DNA Binding Assays. The gel-shift assay is essentially that described by Súch (9). A 32P-labeled operator-containing DNA fragment (5.6 fmol; 5 × 10^6 cpm) was incubated with WT and mutant repressors in 15 μl of binding buffer containing 10 mM Hepes (pH 7.0), 1 mM MgCl_2, 0.2 mM CaCl_2, 50 mM KCl, 5% glycerol, 1 μg of double-stranded poly(dI-dC) (nonspecific competitor DNA), and 4.5 μg of bovine serum albumin (carrier protein). Incubation was performed at 0°C for 1 hr. The mixture was then loaded onto an 8% nondenaturing polyacrylamide gel and electrophoresis was performed in 10 mM imidazole phosphate (pH 7.0) for 3 hr at 10 mA at 25°C. The gel was dried and subjected to autoradiography at ~80°C.

RESULTS

Library Design and Construction. The binding of 434 repressor to its operator sequence involves contacts between the repressor and the sugar phosphate backbone of DNA as well as interactions between the side chains of the residues in helix 3 and the bases in the major groove (5). These latter residues include Thr_27, Gln_29, Gln_33, Ser_30, and Gln_33. The structure of the DNA–protein complex suggests that the carboxamide group of Gln_29 makes bidentate hydrogen bonds with adenine 1, and the side chain of Gln_29 makes a hydrophobic contact to thymine 1 (Fig. 1). Bidentate hydrogen bonding also links the N_7 and O_6 of guanine 2' to the carboxamide NH_2 group of Gln_29. The side chains of Thr_27 and Gln_29 form a hydrophobic pocket that makes van der Waals contact with the methyl group of thymine 3'. The carboxamide NH_2 group of Gln_33 donates a hydrogen bond to O_4 of thymine 4', while the methyl group of the same thymine makes van der Waals contact with the side chains of Gln_29 and

![Fig. 1](image-url)
Ser\(^{30}\). The carbonyl oxygen of the Gln\(^{33}\) side chain is hydrogen bonded to a water molecule that makes bridging contacts to phosphate 5' and Ser\(^{30}\).

To gain increased insight into the mechanism whereby 434 repressor binds its operator sequence, as well as to determine whether mutants of 434 can be isolated with altered binding specificities, a large library containing all possible permutations of the 20 amino acids at residues Thr\(^{27}\), Gin\(^{28}\), Gin\(^{29}\), Ser\(^{30}\), and Gln\(^{33}\) was generated. The codons encoding these amino acids were substituted with NNK (N = A, G, C, or T and K = G or C) to afford a library of \(3 \times 10^6\) gene sequences (3 \(\times 10^6\) protein sequences) (11). The library was constructed by introducing Kpn I and Not I restriction sites directly to the 5' and 3' side, respectively, of the sequence encoding recognition helix 3 (Fig. 2). The Kpn I/Not I fragment of WT repressor was then replaced with a 26-bp synthetic linker, which introduced both a frameshift mutation in the 434 repressor gene as well as a BamHI restriction site. The resulting vector, pLH21, contains a nonfunctional repressor structural gene behind the tac promoter in an Ap\(^{r}\) plasmid derived from pUC18. A synthetic Kpn I/Not I fragment encoding the randomized library was then substituted into pLH21. Subsequent digestion with BamHI and transformation into competent E. coli MC1061 afforded \(1 \times 10^8\) independent recombinants, encoding the entire library of mutant repressor sequences.

**In Vivo Selection Against WT Operator.** An in vivo selection system based on a previously developed transcriptional interference assay (8) was used to select individual mutants from the library that bound WT or mutant operator sequences. In this assay, the strong ConII bacterial polymerase promoter interferes with the expression of a drug resistance gene, aadA (encoding aminoglycoside 3'-adenyltransferase), which is transcribed in the opposite direction. Expression of functional repressor in a cell containing the reporter plasmid disrupts transcription from the ConII promoter by binding of the repressor to its operator site and confers the Sp\(^{r}\) phenotype on the host. By selecting for Sp\(^{r}\) on the transformation plates, clones encoding a repressor that binds a particular operator sequence can be selectively enriched. The dimeric nature of both the operator and the repressor should strongly bias binding by WT and mutant repressors at the operator site versus other sites in the reporter plasmid (nonspecific binding may be lethal to the cell). Unique restriction sites in the reporter plasmid allow this vector to be linearized after each round of selection, preventing selection for reporter mutants.

The helix 3 library was initially selected by using the reporter plasmid pConII-Op1, which contains the WT operator sequence. The ratio of colonies on Sp-containing plates vs. Cm-containing plates was in the range \(10^{-4}\) to \(10^{-5}\) for E. coli HB101 transformed with the reporter plasmid alone. This result demonstrates that the selection has a relatively low background. The helix 3 library was then transformed into this host and plated onto Ap plates and ApSp plates affording an Ap\(^{r}\)Sp\(^{r}\)/Ap\(^{r}\) ratio of \(10^{-5}\). Subsequent rounds of selection were carried out by isolating plasmid from Ap\(^{r}\)Sp\(^{r}\) colonies, digesting the pooled plasmids with BamHI and transforming into HB101 bearing the appropriate reporter plasmid.

After three rounds of selection, 14 clones were isolated and the DNA was sequenced (Fig. 3). In the 14 clones, there were eight unique DNA sequences. In virtually every case Gin\(^{28}\), Gin\(^{29}\), and Ser\(^{30}\) were conserved. Thr\(^{27}\), on the other hand, was substituted with both hydrophilic and large hydrophobic residues, which include Ser, His, Arg, Phe, and Trp. Surprisingly, for each clone the codon for Gln\(^{33}\) was replaced with an amber stop codon. Because HB101 is a supE strain (13), it is likely that rather than encoding truncated protein (which cannot dimerize), these genes all yield full-length repressor containing Gln\(^{33}\) by suppression of the amber stop codon with glutamine. In fact, full-length protein was detectable by Western blot analysis (14) and attempts to express truncated protein were unsuccessful. Selection for the TAG vs. Gin codons at this position may result from altered translational efficiency of the RNA transcripts.

Transformation of the isolated clones into HB101 containing the WT reporter plasmid resulted in Ap\(^{r}\)Sp\(^{r}\)/Ap\(^{r}\) ratios ranging from 0.16 to 0.99 (Table 1). Transformation of the clones into HB101 containing a reporter plasmid (pConII-MT1) with the mutant operator sequence (ACGA-TCGT) afforded Ap\(^{r}\)Sp\(^{r}\)/Ap\(^{r}\) ratios of \(<10^{-4}\), indicating that the mutants bind selectively to Op1 in vivo. Four identical clones were also isolated from the library, which contained a multiple insert and premature stop codon. Retransformation of this clone into HB101 with reporter plasmids containing either the WT operator, mutant operator (MT1), or an operator deletion (AO) afforded Ap\(^{r}\)Sp\(^{r}\)/Ap\(^{r}\) ratios all \(>0.20\), suggesting that factors other than specific DNA binding were involved in the selection of this clone.

Multiple repressors were expressed behind the tac promoter (the amber stop codon was converted to the Gln codon CAG) and the resulting proteins were purified to homogeneity. Their binding affinities were determined by a gel-shift assay using both Op1 and the mutant operators 5'-ACGA-TCGT-3', 5'-ACTA-TAGT-3', and 5'-ACCA-TGGT-3', in which the proposed Thr\(^{27}\) operator contact thymine 3 is mutated (Table 2). In addition, DNase I footprinting assays were performed on the WT 434 repressor, Phe\(^{27}\) mutant, and Arg\(^{27}\) mutant using \(^{32}\)P-labeled 150-bp fragments containing Op1. Mutants show the same footprinting pattern as that of WT repressor (data not shown).

**In Vivo Selection Against Mutant Operators.** The helix 3 library was also subjected to four rounds of selection by using operators containing single (MT1, 5'-ACGAGAAGTTCC-
GT-3') and double (MT2, 5'-ATAGGAAGTCATCTAT-3') mutations. The first mutant operator contains an A → G transition at a base pair not involved in hydrogen bonding interactions with WT repressor. The second mutant involves two transitions (C → T and A → G), resulting in the substitution of two carboxyl groups with two NH₂ groups, which can also participate in hydrogen bonding interactions.

Four rounds of selection were carried out for each mutant operator; the sequences of selected clones are listed in Fig. 3. The clones varied considerably in their sequences and in the hydrophobicity and size of amino acids at a given site, ranging from a clone with 2 Arg and 1 Lys (CL10) to a clone with 2 Trp and 1 Leu (CL8). It is somewhat surprising that almost identical sequences, Gly27-Trp28-Trp29-Ser30-Xaa31-Xaa32-Val33 (MT2)/Gly27-Trp28-Trp29-Ser30-Xaa31-Xaa32-Leu33 (MT1) and Arg27-Trp28-Leu29-Ala30-Xaa31-Xaa32-Leu33 (MT1)/Arg27-Met28-Leu29-Ala30-Xaa31-Xaa32-Leu33 (MT2), were found in the two selections.

A number of selected clones were then transformed into HB101 harboring reporter plasmids containing WT, mutant, and deleted operator sequences (Table 1). The Ap³Sp/Ap⁺ ratios suggest that these clones do in fact bind the mutant operator sequences MT1 and MT2 in vivo relative to the control containing the operator deletion (pConII-ΔOR). However, none of the clones showed significant selectivity in vivo for the mutant operator sequences relative to O31. Three of the mutant repressors were then expressed and purified to homogeneity. Gel-shift assays indicated that all three clones bound WT operator with higher affinities than the corresponding mutant operators.

### Table 1. In vivo binding assays of mutants

<table>
<thead>
<tr>
<th>Repressor</th>
<th>pConII-O31 (ACAA-TTGT)</th>
<th>pConII-MT1 (ACGA-TGCT)</th>
<th>pConII-MT2 (ATAG-CTAT)</th>
<th>pConII-ΔOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT repressor</td>
<td>0.55</td>
<td>&lt;10⁻⁴</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Thr27 → Phe27</td>
<td>0.99</td>
<td>&lt;10⁻⁴</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Thr27 → Trp27</td>
<td>0.65</td>
<td>&lt;10⁻⁴</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Thr27 → Arg27</td>
<td>0.64</td>
<td>&lt;10⁻⁴</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Thr27 → His27</td>
<td>0.38</td>
<td>&lt;10⁻⁴</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Thr27 → Ser27</td>
<td>0.16</td>
<td>&lt;10⁻⁴</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>CL7</td>
<td>1.6</td>
<td>1.0</td>
<td>ND</td>
<td>0.20</td>
</tr>
<tr>
<td>CL8 (GWWS...L)</td>
<td>0.20</td>
<td>0.72</td>
<td>ND</td>
<td>7.5 x 10⁻³</td>
</tr>
<tr>
<td>CL10 (RARK...S)</td>
<td>2.1</td>
<td>0.68</td>
<td>ND</td>
<td>8.7 x 10⁻³</td>
</tr>
<tr>
<td>CL15 (GWWS...V)</td>
<td>0.25</td>
<td>0.50</td>
<td>4 x 10⁻⁴</td>
<td></td>
</tr>
<tr>
<td>CL16 (RMLA...L)</td>
<td>0.38</td>
<td>ND</td>
<td>0.73</td>
<td>1.3 x 10⁻³</td>
</tr>
</tbody>
</table>

ND, not determined.
Table 2. In vitro binding of repressors to WT and mutant operator sequences

<table>
<thead>
<tr>
<th>Operator sequence</th>
<th>pLHGS1 (ACAA- TTGT)</th>
<th>pLHGS2 (ACGA- TTGT)</th>
<th>pLHGS3 (ACAA- TAGT)</th>
<th>pLHGS4 (ACTA- TAGT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>6.0 × 10^{-9}</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>F27 pLHGS2</td>
<td>1.9 × 10^{-8}</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Trp27</td>
<td>7.5 × 10^{-8}</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Arg27</td>
<td>3.3 × 10^{-9}</td>
<td>ND</td>
<td>5.6 × 10^{-8}</td>
<td>1.7 × 10^{-7}</td>
</tr>
<tr>
<td>His27</td>
<td>1.7 × 10^{-9}</td>
<td>ND</td>
<td>2.5 × 10^{-7}</td>
<td>ND</td>
</tr>
<tr>
<td>Ser27</td>
<td>1.3 × 10^{-8}</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
</tbody>
</table>

Assays were carried out using a 32P-end-labeled 139-bp fragment generated by digestion of operator-containing plasmids (pLHGS1-pLHGS4) with Xho I and BstEII. Values shown (M) correspond to repressor concentration at half-maximal binding. ND, not detected (no specific DNA binding observed at 5 μM protein concentration). Mutant bases are in boldface type.

*Measured at 4°C.

DNA insert (pLHGS5). All attempts to footprint the mutant repressors using DNase I failed. Addition of E. coli cytoplasmic extracts to the assays had no effect.

**DISCUSSION**

The results of library selection against O21 suggest that the helical motif Gin28-Gln29-Arg30-Xaa31-Xaa32-Gln33 is not only sufficient for specific binding to O21 by 434 repressor, it may also be unique for this sequence. The strong preference for Gin may reflect constraints imposed on the protein-DNA interface by other residues not in helix 3. The side chains of Asn16, Gin17, Asn26, and Arg30 anchor the protein along the upper strand of the helix binding site, while main-chain NH groups of residues in the loop between helices 3 and 4 anchor the protein to the bottom strand (5). Consequently, only Gin may be capable of satisfying both hydrogen bonding and van der Waals packing interactions between helix 3 of the repressor and the major groove of DNA. The strong bias for Ser30 is somewhat surprising given that this residue does not directly contact the bases (5). However, this residue is involved in hydrogen bonding to two water molecules that are bound to the phosphate backbone. Substitution of a less polar group at this site (e.g., Cys, Thr, or Ala) might disrupt hydration of the phosphate groups.

Both polar and nonpolar side chains of various sizes can substitute for Thr27. It has been suggested that this residue may be involved in either hydrogen bonding interactions with Ser30 and a phosphate-bound water molecule or in hydrophobic interactions with the methyl group of thymine 3'. Both the methyl (Thr27 → Ser) and hydroxyl (Thr27 → Phe) groups of Thr could be removed without a significant loss in DNA binding affinity. Substitution of Arg or His for Thr27 led to mutants with comparable or even higher binding affinities for O21 relative to WT repressor, possibly as a result of favorable electrostatic interactions with the DNA backbone. Even the bulky residue Trp was tolerated reasonably well at position 27. These results indicate that significant structural perturbation can be accommodated at this site, in contrast to sites 28, 29, 30, and 33.

Interestingly, the Thr27 → Arg mutant also displayed significant binding affinity to the mutant operators 5'-ACCA-TGGT-3' and 5'-ACTA-TAGT-3', but not to the sequence 5'-ACGA-TCGT-3' (mutant bases are in boldface type). The Thr27 → His mutant also bound the mutant operator ACCT-GTG. There was no detectable binding affinity of WT repressor to either mutant operator at 5 μM repressor. This increased interaction of the Thr27 mutants with the mutant operators could be the result of interactions between the side chain of residue 27 and bp 3 of the operator. The crystal structure (5) suggests that substitution of thymine 3' with a purine creates sufficient space to allow hydrogen bonding between the side chain of Arg27 and N' of the purine.

The fact that selection of the library against either of the two mutant operators did not afford a mutant repressor with high affinity for either the mutant or WT operator sequences suggests that the hydrogen bonding and packing and van der Waals requirements at these repressor operator interfaces are again quite severe. In addition, there may be significant conformational differences between the mutant and WT operators that affect repressor binding. In contrast to these results, it has been previously shown that a Gin28 → Ala mutant 434 repressor can bind a mutant operator in which the A-T base pair at position 1 is substituted with T-A (7). Consequently, modification of the hydrogen bonding, van der Waals and packing interactions at the DNA-protein interface, in order to generate new proteins with altered specificities, may require a series of interdependent mutations in both the DNA and protein. This has been shown to be the case with mutations in residues in the hydrophobic core of a repressor (15). Significant changes in the sequence specificity of 434 repressor may also require the mutation of additional residues surrounding helix 3.

Interestingly, the sequences selected against the mutant operators did result in Sp in vivo and homologous sequences were obtained in independent selections against the two mutant operators. Moreover, in vitro assays of these sequences against the reporter plasmid with the deleted operator sequence (ΔOR) resulted in APSp/Ap ratios of ~10^{-1}. Although the selective enrichment of repressor library in vitro appears to depend on the operator sequence, other factors not yet apparent from the in vitro assays are clearly playing a role. Attempts to screen libraries of 434 repressor in vitro using phage display systems were unsuccessful.

We are grateful for financial support for this work from the National Institutes of Health (Grant GM41679) and the Lucille P. Markey Charitable Trust.