Trp Repressor−Operator Binding: NMR and Electrophoretic Mobility Shift Studies of the Effect of DNA Sequence and Corepressor Binding on Two Trp Repressor−Operator Complexes†

Mahesh Jaseja,‡ Mark Jeeves,§ and Eva I. Hyde*

School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, U.K.

Received January 23, 2002; Revised Manuscript Received September 16, 2002

ABSTRACT: In Trp repressor−DNA complexes, most interactions either occur with phosphate groups or are water-mediated hydrogen bonds to bases. To examine the factors involved in DNA selectivity, we have studied Trp repressor binding to two operator sequences, $trpR^S$ and $trpO^M$, with L-tryptophan or 5-methyltryptophan as corepressor. These operators contain all the consensus bases but differ at base pairs contacted by their phosphate groups. In electrophoretic mobility shift assays (EMSAs) the $trpR^S$ sequence gives solely 1:1 protein−DNA complexes with either corepressor. The $trpO^M$ sequence binds more weakly than $trpR^S$. It gives dissociating 2:1 complexes in EMSAs with L-tryptophan, but both 1:1 and 2:1 complexes are observed with 5-methyltryptophan or if glycerol is present in the gel. The backbone resonances of the TrpR−L-tryptophan−DNA complexes were assigned using triple-resonance experiments and selectively ^{15}N labeled protein. On changing the DNA sequence, the largest differences in the NMR spectra are at residues 78−81, at the turn of the helix−turn−helix motif and the tip of the recognition helix. I79 and A80 interact with the conserved bases of the operators, while G78 and T81 interact with phosphate groups at bases that differ between the two sequences. Changing the corepressor from L-tryptophan to 5-methyltryptophan causes effects at residues 52, 60, 61, and 85, which do not interact with the DNA. The spectra suggest that there is mutual induced fit between protein and DNA so that sequence changes at bases contacted only by the phosphate groups affect the environment of the protein at residues that bind to conserved bases elsewhere in the DNA.

The *E. coli* Trp repressor (TrpR) binds to at least five operators in the *E. coli* genome, repressing gene expression (reviewed in ref *1*). The operators at which it binds vary considerably in DNA sequence and are located at different positions within the promoters or leader sequence of the genes controlled. The operons controlled, namely, the *trpEDCBA* (*trpO*) operon and the genes for *trpR*, *aroH*, *mtr*, and *aroL* (*1*, *2*), are involved in the biosynthesis and uptake of the amino acid tryptophan. The repressor binds to the operators only in the presence of L-tryptophan, hence controlling the intracellular level of its effector. 5-Methyltryptophan acts as a stronger corepressor than tryptophan as it increases the affinity of Trp repressor to *trpO* (*3*).

As it is one of the smallest DNA-binding proteins that is regulated by binding of a ligand, the Trp repressor has been extensively studied by genetic and biophysical techniques, yet its mechanism of DNA binding and, in particular, its DNA selectivity are poorly understood (reviewed in ref *4*). The protein is a dimer, each subunit containing six helices (Figure 1A). Four of the helices (A, B, C, and F) form the core of the protein and intertwine with the corresponding helices from the other subunit. Helices D and E of each subunit form a helix−turn−helix motif (*5*, *6*). In the presence of L-tryptophan, the orientation of these helices relative to the core is altered so that they are correctly spaced to bind to DNA (*7*, *8*). Tryptophan also affects the oligomerization of the protein (*9*, *10*) and interacts directly with the DNA backbone (*11*).

Most studies of Trp repressor−DNA interactions have been with the *trpO* operator and symmetrical $trpO^S$ [1] variants containing TA at base pair 3, rather than AT. The $trpO^S$ operator contains three axes of symmetry, a central one (α) and two secondary ones (β), four base pairs apart. In this study we have numbered the base pairs 1−20 as in Chart 1.

The crystal structure of Trp repressor bound to the $trpO^S$ operator (*11*) shows one protein dimer per DNA duplex with only one direct hydrogen bond between the protein and the

† This work was partially funded by Grant 040373 from the Wellcome Trust. M. Jeeves was supported by a BBSRC studentship. The Biological NMR Centre at the University of Birmingham was funded by grants from the Wellcome Trust, the BBSRC, and the Arthritis and Rheumatism Campaign. The computing facilities at the University of Birmingham were funded by Grant G4600017 from the Medial Research Council.

‡ Present address: Department of Chemistry, North Dakota State University, Fargo, ND 58105.

§ Present address: Department of Biochemistry and Molecular Biology, University College London, London WC1E 6BT, U.K.

* To whom correspondence should be addressed. Phone: 00-44-121-414-5393. Fax: 00-44-121-414-5925. E-mail: e.i.hyde@bham.ac.uk.

[1] Abbreviations: EMSA, electrophoretic mobility shift assay; TBE, Tris borate, EDTA buffer; TPPI, time-proportional phase incrementation; $TrpO^S$, variant *trpO* operator, made symmetrical on the basis of the right half *trpO* operator sequence; $TrpO^M$, mutant $trpO^S$ operator, with mutation at base pairs 7 and 14; $TrpR^S$, variant *trpR* operator, made symmetrical on the basis of the right half *trpR* operator sequence.

10.1021/bi020072y CCC: $22.00
Published on Web 11/15/2002

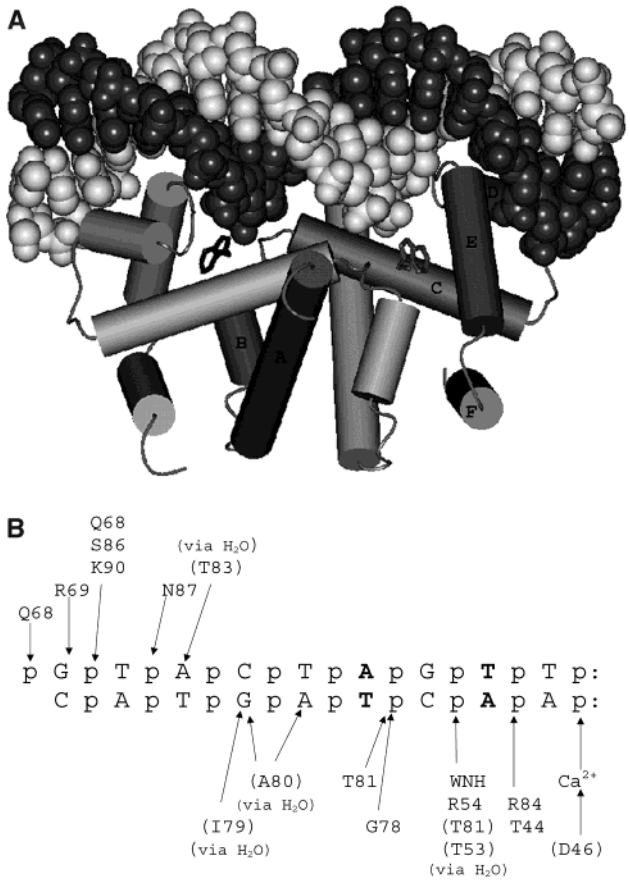

FIGURE 1: Schematic representation of the structure of the TrpR−*trpO*^{S}−L-tryptophan ternary complex. (A, top) Overall structure, from 1TRO (*11*). The DNA is shown as van de Waals spheres, the α-helices of the protein are shown as rods joined by ribbons, and the corepressor is shown in stick representation. The α-helices of one subunit are labeled A−F. (B, bottom) Protein−DNA contacts in one half-site of the complex. Indirect contacts via water or Ca^{2+} are shown in parentheses. The base pairs changed in *trpR*^{S} and *trpO*^{M} are shown in bold.

Chart 1: Sequence of the *TrpEDCBA* Operator Showing the Axes of Symmetry and the Numbering Used in This Paper[a]

trpO

β α β

1 10:11 20

C**GA**A**A**CT.AGTT:AACT.**AG**TACG

GCTT**GA**.TCAA:TTGA.TC**A**T**G**C

[a] The central α axis of symmetry is designated with a colon, the secondary *β* axes of symmetry are designated by periods. Bases contacted in the crystal structure of the TrpR−*trpO*^{S} complex *(11)* are shown in bold.

bases, to base G2. In mutagenesis studies this base was not found to be important for DNA binding (*12*). There are a number of protein/phosphate contacts, but the remaining protein/base contacts occur via water molecules to bases A4, G16, and A15 (in bold in Chart 1) (Figure 1B). An NMR study of the 1:1 TrpR−*trpO*^{S} complex showed contacts similar to those in the crystal structure, but could not resolve whether the hydrogen bonds to the base pairs were direct or water mediated (*13*). The crystal structure of Trp repressor bound to an alternative sequence, based on the *β*-axis of symmetry, shows tandem binding of the protein along the DNA (*14*), with two protein dimers per DNA duplex, and further water-mediated contacts to bases corresponding to A7, G8, and A12. Preliminary NMR studies of a tandem 2:1 complex have been reported (*15*, *16*). Other solution studies show that the stoichiometry of Trp repressor binding depends on the exact sequence and length of DNA studied (*17*−*22*). These findings suggest that sequence-specific recognition of operator DNA by Trp repressor may be due to factors other than direct hydrogen bonding. These could include water-mediated hydrogen bonds, the conformation of the DNA backbone (indirect readout), and cooperative binding of the protein to multiple sites. These factors have not often been considered in studies of DNA sequence selectivity but may be important for a number of protein−DNA complexes for discrimination between cognate and noncognate sequences.

The aim of our work is to examine the molecular basis for the selectivity of Trp repressor for operators over nonoperator sequences. We have previously examined Trp repressor binding to all of its five operators and to variant *trpR* operator sequences in the presence of tryptophan by electrophoretic mobility shift assay (EMSA) (*23*). We found that the repressor binds to DNA containing each of the five natural operator sequences with a stoichiometry of at least 2:1 protein dimer:DNA, but a change in a single base pair within the *trpR* sequence could change the stoichiometry to 1:1. We have also examined the effect on the DNA of Trp repressor binding to a symmetrized form of the *trpR* operator, *trpR*^{S}, using NMR spectroscopy (*24*). Here we compare Trp repressor binding to two symmetrized operator variants, *trpR*^{S} and *trpO*^{M}, by EMSA and NMR studies of the protein. Both of these operators contain all the bases that form hydrogen bonds to the protein in the 1:1 TrpR−*trpO*^{S} complex. The *trpR*^{S} operator differs from *trpO*^{S} at two of the ten base pairs in each half-site. It binds with similar affinity to Trp repressor as *trpO*^{S} but gives solely a 1:1 complex in EMSA (*23*). The *trpO*^{M} mutant differs from *trpO*^{S} at only one base pair per half-site, at a base also altered in *trpR*^{S}. Despite this, it acts as a partially constitutive operator *in vivo* (*25*), and in an *in vivo* phage challenge assay (*12*), this mutation showed a large effect on DNA binding. We show below that *trpO*^{M} binds to the Trp repressor more weakly than *trpR*^{S} and in EMSAs with L-tryptophan only a 2:1 complex is observed that dissociates in the gel. However, both 1:1 and 2:1 complexes are observed in EMSAs with the *trpO*^{M} operator in the presence of 5-methyltryptophan, or when glycerol is present in the gel. Under these conditions, the *trpR*^{S} operator still gives only a 1:1 complex. To examine the molecular basis of the differences observed in the binding of the two sequences to Trp repressor in the presence of the two corepressors, we examined the protein in each of the complexes by NMR spectroscopy. ^{1}H and ^{15}N chemical shift mapping was used to examine the effects of complex formation on the chemical environment of each amide group in the protein. The results are discussed with reference to the crystal structure of the TrpR−*trpO*^{S}−L-tryptophan complex (*11*).

EXPERIMENTAL PROCEDURES

Oligonucleotide. The self-complementary sequences 5′-dCGTACTCGCTAGCGAGTACG-3′ (*trpR*^{S}) and 5′-

Table 1: DNA Sequences Studied

trpO^S	
	1--------10---------20
	CGTACTAGTT.AACTAGTACG
	GCATGATCAA.TTGATCATGC
trpO^M	
	CGTACTGGTT.AACCAGTACG
	GCATGACCAA.TTGGTCATGC
trpR	
	CGTACTCTTT.AGCGAGTACA
	GCATGAGAAA.CTGCTCATGT
trpR^S	
	CGTACTCGCT.AGCGAGTACG
	GCATGAGCGA.TCGCTCATGC

dCGTACTGGTTAACCAGTACG-3′ (*trpO^M*) were synthesized by the phosphoramidite method. They were purified by reversed-phase HPLC before and after detritylation. For NMR experiments, 3−6 mg was dissolved in 0.4 mL in 25 mM sodium phosphate, pH 7.6, 150 mM NaCl, 0.05 mM EDTA and freeze-dried. The sample was heated to 70 °C in a water bath and allowed to cool overnight to anneal the strands. The amount of oligonucleotide was estimated by weight.

Isolation and Purification of the Protein. The Trp aporepressor was isolated from the overproducing strain of *E. coli* CY15070 carrying the plasmid pJPR2 (*26*). ^{15}N,^{13}C-labeled protein was made by growing the strain in M9 media containing 1 g/L [^{15}N]NH_4Cl and 2 g/L [$^{13}C_6$]glucose as the sole nitrogen and carbon sources, respectively. ^{15}N,^{14}N-Leu-labeled protein was made using M9 media containing 1 g/L [^{15}N]NH_4Cl and unlabeled leucine (0.26 g/L) and glutamic acid (0.46 g/L) as the nitrogen sources. Samples that were selectively labeled with [^{15}N]amino acids were made as by Muchmore et al. (*27*). Deuterated Trp repressor protein was made by acclimatizing the *E.coli* CY15070 containing pJPR2 to grow in D_2O by successively subculturing it in higher concentrations of D_2O in M9 media. It was then grown in M9 media containing 50% D_2O, [^{15}N]NH_4Cl, and [$^{13}C_6$]-glucose. In all cases the protein was purified according to the method of Paluh and Yanofsky (*26*). Traces of nuclease were removed using a Trisacryl Blue column (IBF) as described previously (*24*). The concentration of protein was estimated from its absorbance at 280 nm, using an extinction coefficient of 1.2 $dm^3 \cdot mg^{-1} \cdot cm^{-1}$ (*28*).

Cloning for EMSA. All cloning and DNA preparations were performed as by Sambrook et al. (*29*). The oligonucleotides were blunt-end ligated, after annealing, into the *Sma*I restriction site of plasmid pUC19. Transformants containing the desired plasmids were selected by lack of α-complementation in *E. coli* JM101. The length of the insert was determined. The polycloning region of the plasmids was sequenced to confirm the sequence of the insert and, for asymmetric sequences, to determine its orientation. The *trpR* operator was cloned with the *Eco*R1 site closest to bp 1 and the *Bam*H1 site closest to bp 20 in the orientation shown in Table 1. Large-scale preparations of plasmids were made by the alkaline lysis method. Contaminants were removed by precipitation with polyethylene glycol, or by isopycnic centrifugation through cesium chloride. The concentration of purified plasmid DNA was determined fluorimetrically using Hoescht 33258, with sheared calf thymus DNA as a standard.

EMSA. A 400−600 ng sample of plasmid DNA was restricted with *Eco*RI and *Hin*dIII to excise the fragment containing the operator sequences. The 5′-ends of the fragments were labeled using [α-^{32}P]dATP by the Klenow fragment of DNA polymerase I. These fragments were separated from the remaining plasmid DNA and the unincorporated label by polyacrylamide gel electrophoresis in TBE. The fragments were eluted from the gel by crushing the part of the gel of interest and soaking it in 10 mM Tris HCl, pH 8.0, 0.1 mM EDTA for a few hours followed by centrifugation.

EMSAs were performed on the basis of the method of Carey (*30*). DNA (∼50 pM) was incubated on ice with 0−20 nM protein in 10 mM sodium phosphate buffer (pH 6.0), 25 mM NaCl, 1 mM EDTA, 1 mg/mL BSA, 10% v/v glycerol (or 12% for gels containing glycerol), and either 0.4 mM L-tryptophan or 0.8 mM D,L-5-methyltryptophan. The samples were loaded under tension onto an 8% acrylamide, 1.6% bisacrylamide gel containing 10 mM sodium phosphate buffer (pH 6.0) and either 0.1 mM L-tryptophan or 0.2 mM D,L-5-methyltryptophan, plus or minus 3% glycerol. The electrophoresis buffer was the same as in the gel but plus or minus 2% glycerol. Gels were electrophoresed at 4 °C at 20 mA for about 5 h with recirculation of the electrophoresis buffer. The gels were dried and autoradiographed, and the intensity of each band was measured using a phosphorImager plate and the program ImageQuant 3.3 (Molecular Dynamics). The fraction of the total radioactivity found in each band was fitted to the following equations (*31*) using nonlinear regression in Sigmaplot (SPSS Inc.):

$$F_{\text{free}} = 1/Z$$

$$F_1 = K_1[\text{P}]/Z$$

$$F_2 = K_2[\text{P}]^2/Z$$

$$Z = 1 + K_1[\text{P}] + K_2[\text{P}]^2$$

where F_{free} is the fraction of DNA species in the free band,

F_1 is the fraction of DNA species in the first retarded band,

F_2 is the fraction of DNA species in the second, more retarded, band,

[P] is the protein concentration in subunits,

K_1 is the macroscopic association constant for a single protein molecule binding to the DNA, and

K_2 is the macroscopic association constant for two protein molecules binding simultaneously to free DNA.

Formation of Complexes for NMR Experiments.

The protein was added to an excess of oligonucleotide (ratio approximately 0.9:1.0), in the presence of 2 mM L-tryptophan or 4 mM D,L-5-methyltryptophan, and dialyzed into 5 mM sodium phosphate buffer, pH 6.0, 0.05 mM EDTA, 50 mM NaCl, 2 mM L-corepressor. The final concentration of protein in the NMR experiments of the ternary complexes was 1−2 mM in subunits in this buffer. Additional oligonucleotides were added to a final ratio of 1.5−1.8 equivalents of protein. The concentrations of L-tryptophan or 5-methyltryptophan were determined spectro-

photometrically. For the binary complexes, the protein was dialyzed into the same buffer, containing 500 mM NaCl.

NMR Spectroscopy. NMR spectroscopy was performed using a Bruker AMX500 spectrometer operating at 500.13 MHz or on a Varian Unity-plus spectrometer operating at 599.98 MHz. The spectra were recorded at 37 °C.

^{15}N−^{1}H HSQC spectra (*32*) were collected with 2K points with a ^{1}H spectral width of 13 ppm in the directly observed dimension, and 340−400 points in F1 with a 41 ppm sweep width in ^{15}N. Experiments were collected in States-TPPI (time-proportional phase incrementation) mode using the standard pulse sequences and phase cycling schemes and at least a 1 s delay after acquisition and before the first pulse of the program. The data were zero-filled in the indirect dimension to 1K points prior to Fourier transformation. The spectra were converted to Bruker format and processed using UXNMR software with shifted sine bell window functions in F1, Gaussian resolution enhancement in F2, and baseline correction in both dimensions.

HNCA (*33*), HNCOCA (*33*), HNCO (*34*), HNCACB (*35*), and CBCACONH (*36*) spectra were collected with 1K points and a spectral width of 13 ppm in the direct ^{1}H dimension, 32−48 points in the ^{15}N dimension (zero filled to 64−128 points) with a sweep width of 38 ppm, and 56−72 points in ^{13}C (zero filled to 128 points). For HNCO spectra, the ^{13}C spectral width was 10.5 ppm, for HNCA 30 ppm, and for HNCACB 63 ppm. 3D ^{15}N−^{1}H NOESY/HSQC spectra (*32*) were collected with a mixing time of 125−150 ms and 128 points in the ^{1}H NOE dimension (zero filled to 256 points), over a sweep width of 13 ppm. Proton chemical shifts are referenced to internal sodium 3-trimethylsilyl-(2,2,3,3-$^{2}H_4$)-propionate at 0 ppm. ^{15}N shifts and ^{13}C shifts are referenced indirectly to ^{1}H (*37*).

All amide ^{1}H, ^{15}N, and ^{13}Cα resonances for the binary and ternary complexes with *trpR*^{S} have been assigned, apart from those of M11, E13−E18 inclusive, and E47 in the ternary complex, due to resonance overlaps. ^{13}Cβ resonances have been assigned for the majority of residues, excluding the ones above, the prolines, and an additional seven residues in the ternary complex. The corresponding amide ^{1}H, ^{15}N, and ^{13}Cα resonances for the *trpO*^{M} complex have been assigned, but no side chain resonances.

Comparison of Shifts to Those for trpO^{S} *Complexes Reported Previously (13, 38)*. The

NMR studies of the six complexes reported in this study were conducted at the same temperature and pH. The chemical shifts could therefore be compared directly. However, when comparisons were made with two previous NMR studies of TrpR−*trpO*^{S} complexes (*13*, *38*), small systematic differences in the chemical shifts were found among all the studies. To minimize the systematic differences, we determined the mean difference in chemical shifts between the resonances in each of the complexes. This mean difference was then subtracted from the actual difference in chemical shift observed at each residue. After this correction, the differences in chemical shifts in the N-terminal half of most of the complexes are very small. However, at a few residues there are large chemical shift differences of the TrpR−*trpO*^{S}−L-tryptophan spectra from those of all the other complexes. We therefore compare the chemical shifts of only the C-terminal three helices, from residue 70 onward, including the helix−turn−helix region. For these residues, after our correction is applied, the average difference in ^{15}N shifts between the TrpR−*trpO*^{S}−L-tryptophan complex (*13*) and the TrpR−*trpO*^{S}−5-methyltryptophan complex (*38*) is 0.15 ppm. The average difference in ^{15}N shifts between the TrpR−*trpO*^{S}−L-tryptophan complex and the TrpR−*trpR*^{S}−L-tryptophan and TrpR−*trpO*^{M}−L-tryptophan complexes, after correction, is 0.19 ppm. We have only considered shift differences greater than twice the average difference as significant.

RESULTS

EMSA. To examine the molecular basis of the selectivity of Trp repressor for its operator DNAs, we have compared the binding of the Trp repressor to two symmetrical operator variants, a constitutive mutant operator, *trpO*^{M}, and the symmetrized *trpR*^{S} operator (Table 1). Both contain all the bases recognized by a single TrpR dimer, but differ at bases near the center of the operator contacted only at phosphate groups. The first step in the comparison of the operators was to examine their binding to Trp repressor and the stoichiometry of the complexes formed, using EMSAs. These were compared to EMSAs with the symmetrized trpEDCBA operator (*trpO*^{S}), used for crystal studies, and with the *trpR* operator. For these studies 20 bp oligonucleotides containing the sequence of interest were cloned into the *Sma*I site of pUC19. From these plasmids we excised longer DNA fragments (approximately 70 bp) to be used in the assays, as in our previous study (*23*). By cloning the oligonucleotides into the same site, all the sequences studied have the same length and identical flanking sequences, so the only difference is the operator sequence inserted. This approach allows a direct comparison of the effects of the inserted sequence, while allowing the binding of the protein to extend into flanking sequences.

As shown in Figure 2A (lanes 1−8), on addition of Trp repressor to the *trpO*^{S} operator in the presence of L-tryptophan, two clear bands of mobility lower than the free operator are observed. Previous studies of this operator have shown that these correspond to protein−DNA complexes of stoichiometry 1:1 and 2:1 dimer:operator (*18*, *19*). In contrast, with the mutant *trpO*^{M} operator only one band is observed at relatively high protein concentrations (Figure 2A, lanes 9−16). This band has the same mobility as the 2:1 *trpO*^{S} complexes and dissociates in the gel, giving a smear between bound and free DNA.

5-Methyltryptophan binds more strongly to Trp repressor than tryptophan does and acts as a strong corepressor. It has been reported to tighten binding of wild-type Trp repressor (*3*) and Trp repressor mutants (*39*) to *trpO*^{S} DNA. To examine whether this occurs with other operators, we repeated the EMSA experiments in the presence of 5-methyltryptophan, rather than L-tryptophan, as corepressor. For the DNA fragment containing the *trpO*^{S} sequence, there was little difference in Trp repressor binding in the presence of the two corepressors (Figure 2, lanes 1−8). However, with DNA containing the *trpO*^{M} sequence a marked difference was seen in the presence of 5-methyltryptophan; both 2:1 and 1:1 TrpR−*trpO*^{M} complexes were observed, with no dissociation in the gel (Figure 2B, lanes 9−16). This suggests that 5-methyltryptophan stabilizes both complexes for *trpO*^{M} DNA.

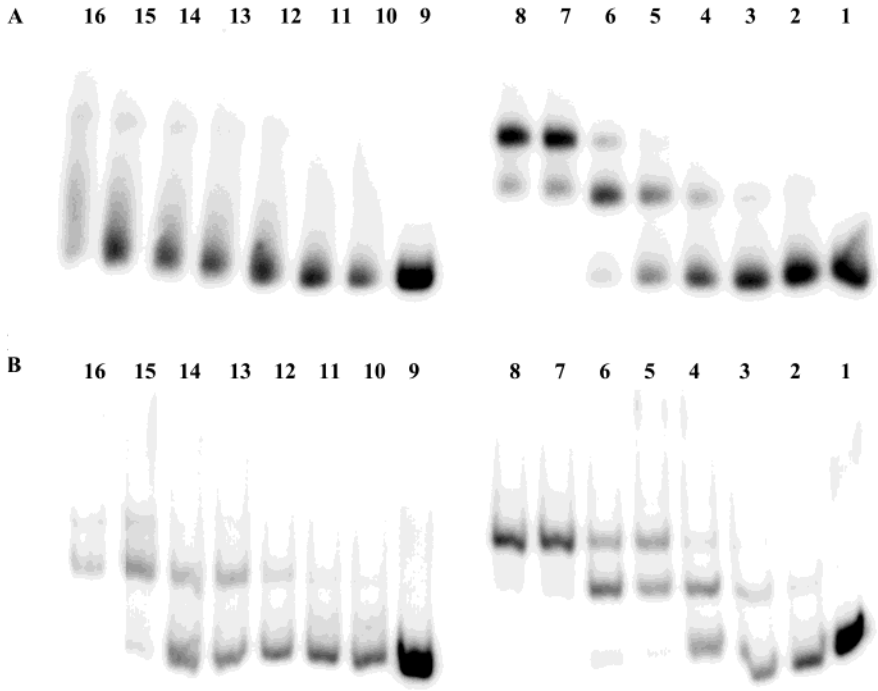

FIGURE 2: Electrophoretic mobility shift assays of Trp repressor binding to DNA fragments containing *trpO*^{M} and *trpO*^{S} sequences in buffer containing L-tryptophan (A) or 5-methyl-D,L-tryptophan (B) in gels electrophoresed in the absence of glycerol. (A) *trpO*^{M} (lanes 9−16) and *trpO*^{S} (lanes 1−8) with L-tryptophan. Trp repressor concentrations: lanes 1 and 9, 0; lane 2, 0.04 nM; lane 3, 0.08 nM; lane 4, 0.16 nM; lane 5, 0.32 nM; lane 6, 0.64 nM; lane 7, 1.25 nM; lane 8, 2.5 nM; lane 10, 0.32 nM; lane 11, 0.64 nM; lane 12, 1.25 nM; lane 13, 2.5 nM; lane 14, 5 nM; lane 15, 10 nM; lane 16, 20 nM. (B) *trpO*^{M} (lanes 9−16) and *trpO*^{S} (lanes 1−8) with 5-methyltryptophan. Trp repressor concentrations: lanes 1 and 9, 0; lanes 2 and 10, 0.02 nM; lanes 3 and 11, 0.04 nM; lanes 4 and 12, 0.08 nM; lanes 5 and 13, 0.12 nM; lanes 6 and 14, 0.16 nM; lanes 7 and 15, 0.24 nM; lanes 8 and 16, 0.32 nM.

To try to prevent dissociation of the *trpO*^{M} complexes with L-tryptophan, glycerol was added to the gel and the running buffer in the EMSA. Under these conditions, DNA fragments containing either the *trpO*^{S} sequence or the *trpO*^{M} sequence both gave 2:1 and 1:1 complexes with either corepressor (Figure 3A,B). The *trpO*^{M} sequence binds more weakly than the *trpO*^{S} sequence, requiring higher concentrations of Trp repressor for a complex to be formed. DNA fragments containing the *trpR* sequence show behavior similar to those containing the *trpO*^{S} sequence in EMSA, with two retarded bands of the same mobility as 1:1 and 2:1 complexes (Figure 3C,D, lanes 1−8). The symmetrized *trpR* sequence (*trpR*^{S}) however gives only one complex, with the same mobility as the 1:1 protein−complex even with 5-methyltryptophan and glycerol (Figure 3C,D, lanes 9−16). No 2:1 complex was observed with this sequence, even at high Trp repressor concentrations.

Table 2 shows the apparent association constants measured for the different fragments from the gels in Figure 3. The four fragments with the same corepressor were assayed together to allow comparisons between their relative affinities. When the gels were repeated, the relative affinities remained the same but the absolute values varied by a factor of 2. This is due partly to the relatively few points for each fragment; the accuracy of the association constant depends on the concentrations assayed. There is no significant difference in binding of the DNA fragments with tryptophan or 5-methyltryptophan as cofactor, except for the *trpO*^{M} sequence, which binds more tightly with the latter corepressor. The fragments containing *trpO*^{S} and *trpR* sequences have very similar association constants to form the 1:1 complex, while *trpO*^{M} binds 2−3-fold more weakly and *trpR*^{S} 1.5−2-fold more tightly. The binding of the second dimer is much tighter for *trpO*^{S} than for *trpR* or the other sequences, that of *trpO*^{M} is weak, and no second binding is seen for *trpR*^{S}.

NMR Spectroscopy.

To examine the molecular basis for the differences in binding, we have used NMR spectroscopy to examine Trp repressor in its complexes with the *trpR*^{S} and *trpO*^{M} oligonucleotides in the presence and absence of the two corepressors.

Figure 4 shows the ^{15}N−^{1}H HSQC spectra of the fully ^{15}N labeled TrpR−L-tryptophan binary complex (A), the TrpR−*trpR*^{S}−L-tryptophan ternary complex (B), and the TrpR−*trpO*^{M}−L-tryptophan ternary complex (C). In all cases L-tryptophan, and DNA where present, was in excess over protein. The backbone ^{1}H, ^{15}N, and ^{13}C assignments of the TrpR−L-tryptophan binary complex and the TrpR−*trpR*^{S}−L-tryptophan ternary complex were determined by 3D triple-resonance HNCA, HNCOCA (*33*), and HNCO experiments (*34*), in combination with ^{15}N−^{1}H NOESY−HSQC (*32*). The side chain resonances of these two complexes were partially assigned by a combination of 3D triple-resonance and double-resonance experiments. The amide ^{1}H, ^{15}N, ^{13}Cα, and ^{13}C′ resonances of the TrpR−*trpO*^{M}−L-tryptophan ternary complex were assigned by comparison of the HNCA and HNCO spectra with those of the TrpR−*trpR*^{S}−L-tryptophan complex, but the side chain resonances were not assigned. The resonance assignments for the two ternary complexes are given as Supporting Information.

Because of the large mass of the ternary complexes (37 kDa) and the α-helical structure of the protein, the sensitivity of some of the triple-resonance experiments was low and there was much resonance overlap. Partial deuteration

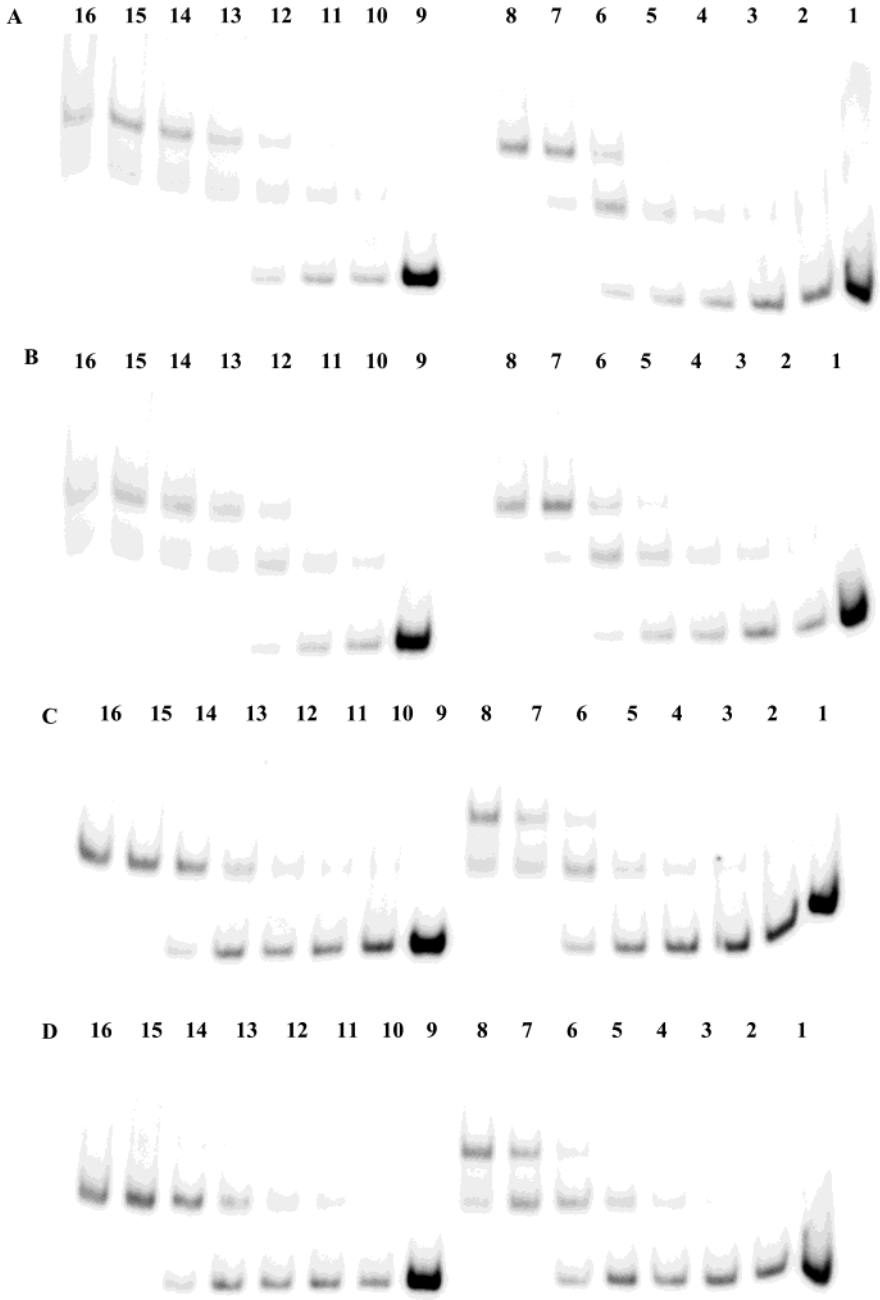

FIGURE 3: Electrophoretic mobility shift assays of Trp repressor binding to (A, B) *trpO*^{M} and *trpO*^{S} and (C, D) *trpR*^{S} and *trpR* in buffer containing L-tryptophan (A, C) or 5-methyl-D,L-tryptophan (B, D) in gels electrophoresed in the presence of glycerol. (A, B) *trpO*^{M} (lanes 16−9) and *trpO*^{S} (lanes 1−9) with (A) L-tryptophan and (B) 5-methyl-D,L-tryptophan. Trp repressor concentrations: lanes 1 and 9, 0; lane 2, 0.02 nM; lane 3, 0.04 nM; lane 4, 0.08 nM; lane 5, 0.16 nM; lane 6, 0.32 nM; lane 7, 0.64 nM; lane 8, 1.25 nM; lane 10, 0.16 nM; lane 11, 0.32 nM; lane 12, 0.64 nM; lane 13, 1.24 nM; lane 14, 2.5 nM; lane 15, 5.0 nM; lane 16, 10 nM. (C, D) *trpR*^{S} (lanes 9−19) and *trpR* (lanes 1−8) with (C) L-tryptophan and (D) 5-methyl-D,L-tryptophan. Trp repressor concentrations: lanes 1 and 9, 0; lanes 2 and 10, 0.02 nM; lanes 3 and 11, 0.04 nM; lanes 4 and 12, 0.08 nM; lanes 5 and 13, 0.16 nM; lanes 6 and 14, 0.32 nM; lanes 7 and 15, 0.64 nM; lanes 8 and 16, 1.25 nM.

(approximately 50%) of the protein was used to increase the sensitivity of some of the triple-resonance experiments. To help remove ambiguities, ^{15}N HSQC and NOESY− HSQC experiments were performed with protein samples that were selectively ^{15}N labeled. A, I, and T residues occur at the beginning of the recognition helix of the DNA-binding motif and contact the DNA, so they were chosen for the selectively labeled amino acids in the samples [IT-^{15}N]- and [AI-^{15}N]Trp repressor. Figure 5 shows the HSQC spectra of the [IT-^{15}N]-labeled TrpR−*trpR*^{S}−L-tryptophan and TrpR−*trpO*^{M}−L-tryptophan complexes, showing the selective shift differences of T81 and I79 between the two ternary complexes. Because of the importance of glycine in turns, we also used an [AGS-^{15}N]-Trp repressor sample to confirm the assignment of glycine residues and to examine the serine residues, which are labeled by metabolism from glycine. Leucine and glutamate are particularly abundant in Trp repressor (19 residues each of 107) and so were chosen as ^{14}N-labeled residues in an otherwise ^{15}N labeled protein sample. In this sample, in ^{15}N-filtered spectra there was a general background of ^{14}N-labeled protein from metabolism of glutamate. In the ^{15}N-observed spectra, however, the Leu residues were absent while some of the Ile and Val residues had lower intensity relative to other residues. A few leucine resonances

Table 2: Apparent Association Constants for the Four DNA Fragments Binding to Trp Repressor Obtained from EMSA in the Presence of Glycerol[a]

DNA fragment	$K_1 \times 10^{-9}$ (M^{-1})	$K_2 \times 10^{-18}$ (M^{-2})	DNA fragment	$K_1 \times 10^{-9}$ (M^{-1})	$K_2 \times 10^{-18}$ (M^{-2})
A. With Tryptophan as Corepressor					
trpO^S	3.3 ± 1.0	15.5 ± 5.1	*trpR*	2.9 ± 0.5	5.0 ± 1.1
trpO^M	1.2 ± 0.2	1.4 ± 0.3	*trpR*^{S}	6.3 ± 0.9	0
B. With 5-Methyltryptophan as Corepressor					
trpO^S	4.6 ± 1.1	21.8 ± 6.0	*trpR*	2.9 ± 0.5	5.0 ± 1.1
trpO^M	2.5 ± 0.5	2.3 ± 0.6	*trpR*^{S}	5.9 ± 0.9	0

[a] The errors shown are those from the fit for a single gel. When the gels were repeated, the relative affinities of the fragments remained the same as did the ratios of K_1/K_2, but the absolute values differed.

are clearly separated from others in the HSQC spectra, and so could be assigned directly from the 2D spectra. The majority of leucine peaks were assigned by their absence in 3D NOESY−HSQC spectra in this sample. Comparison of all these spectra allowed unambiguous assignment of resonances by amino acid type from most of the ten Ala, three Ile, four Thr, five Gly, six Ser, and nineteen Leu residues, comprising 44% of the residues of the protein.

Figure 6 shows histograms of the differences in chemical shifts of amide proton and ^{15}N resonances between the complexes. Parts A and B of Figure 6 show the differences in chemical shifts that occur on binding *trpR*^{S} DNA to the TrpR−L-tryptophan complex. On DNA binding, shift changes occur throughout the protein, with most large changes at the helix−turn−helix, DNA-binding region (residues 68−89). The largest ^{15}N shift change is 7.1 ppm for I79 in the recognition helix, while T81 and T83, in this helix, change by 3.2 ppm (Figure 6A). Outside this region, ^{15}N shift changes greater than 1 ppm on DNA binding occur at D46, G52, and R56 in the C helix, at S8 and A9 at the N-terminus of the protein, and at L38 in helix B. Amide ^{1}H shift changes on DNA binding occur in regions similar to those of the ^{15}N shifts (Figure 6B).

Parts D and E of Figure 6 show the differences in chemical shifts between the TrpR−*trpR*^{S}−L-tryptophan and TrpR−*trpO*^{M}−L-tryptophan complexes. The differences in chemical shifts between the two ternary complexes are much smaller than the changes observed on adding DNA to the binary TrpR−L-tryptophan complex. The differences in shifts between the two ternary complexes occur primarily in the DNA-binding region. The largest ^{15}N shift differences between the ternary complexes are 1.4 ppm for A80 and 1.1 ppm for G78 (Figure 6D). Amide ^{1}H shift differences greater than 0.15 ppm between the two ternary complexes occur only at residues A77−T81 (Figure 6e).

We also compared the ^{15}N HSQC spectra of Trp repressor complexes containing 5-methyltryptophan to those with L-tryptophan. Fully ^{15}N labeled Trp repressor was used for all three complexes. The resolved peaks in the HSQC spectra were assigned on the basis of comparisons with the corresponding L-tryptophan complex. In additional, spectra of selectively labeled [AI-^{15}N]TrpR−*trpO*^{M}−5-methyltryptophan and [AGS-^{15}N]TrpR−*trpR*^{S}−5-methyltryptophan complexes were taken to confirm the assignments of the corresponding residues in these complexes.

Figure 7 shows the differences in ^{15}N shifts of the resolved peaks between the L-tryptophan and 5-methyltryptophan

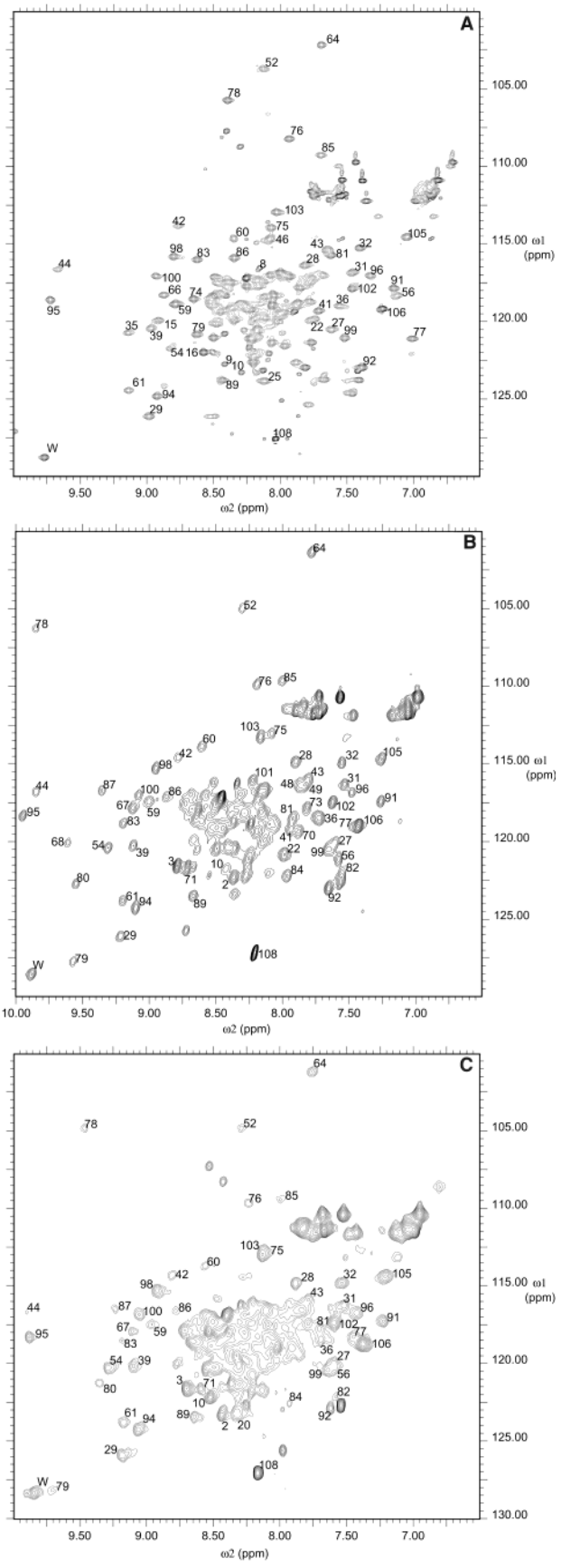

FIGURE 4: ^{15}N−^{1}H HSQC spectra of Trp repressor complexes: (A, top) TrpR−L-tryptophan binary complex, (B, middle) TrpR−*trpR*^{S}−L-tryptophan ternary complex, and (C, bottom) TrpR−*trpO*^{M}−L-tryptophan ternary complex. Selected isolated peaks are labeled with their assignments.

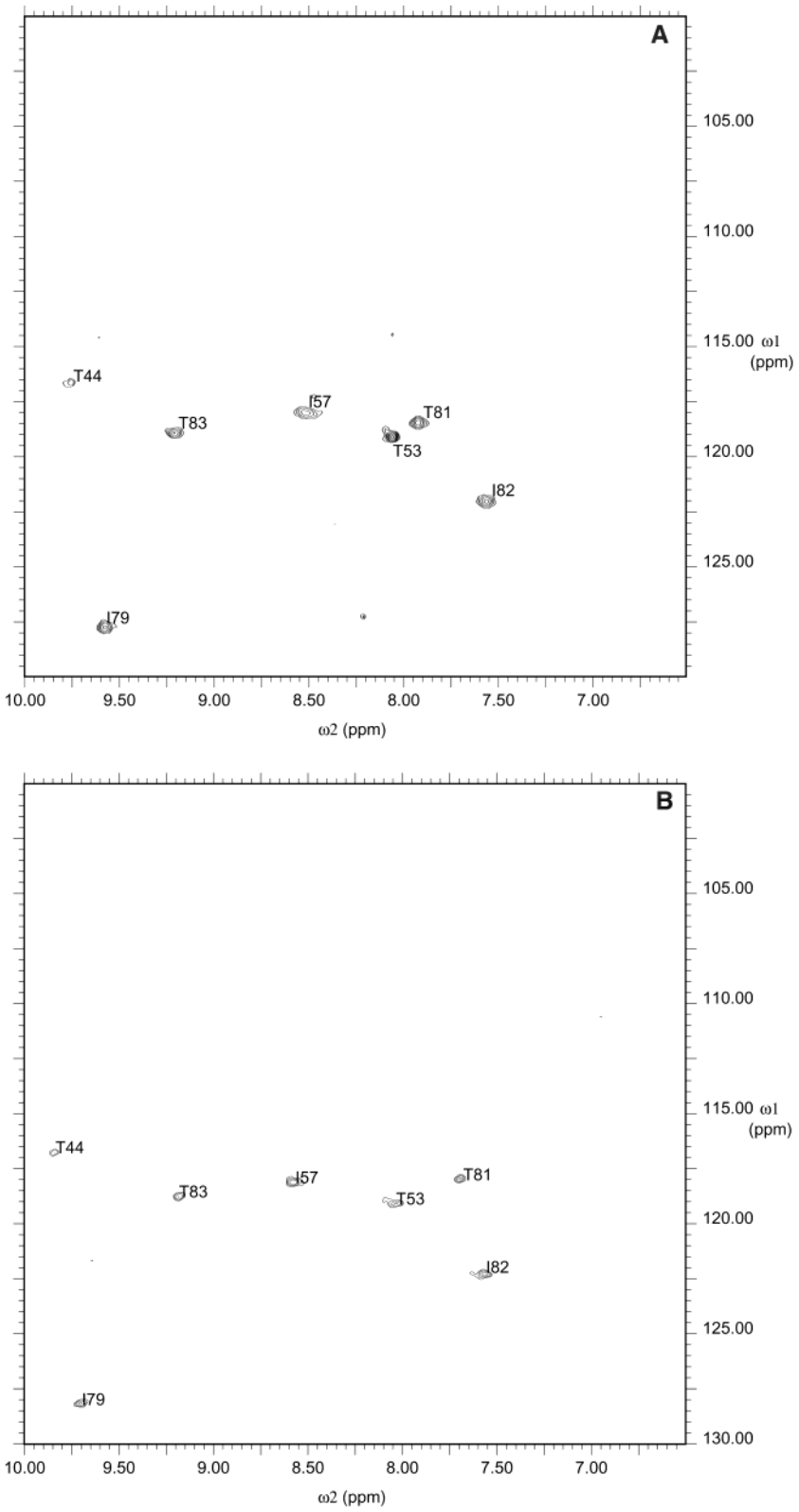

FIGURE 5: $^{15}N-^{1}H$ HSQC spectra of Trp repressor selectively ^{15}N labeled at Ile and Thr in complex with (A, top) *trpR*^{S} DNA and L-tryptophan and (B, bottom) *trpO*^{M} DNA and L-tryptophan. Peaks are labeled with their assignments.

complexes. In the HSQC spectrum of the TrpR−5-methyltryptophan binary complex 58 peaks were assigned. Figure 7A shows the differences in ^{15}N shifts for these peaks between the TrpR−5-methyltryptophan complex and the TrpR−L-tryptophan complex. The largest observed differences in ^{15}N shift between the two binary complexes is 0.63 ppm for L61 and G85. The largest ^{1}H shift difference observed between the two binary complexes is 0.13 ppm for G85 and S86. In addition, T44 gave a broad peak in the HSQC spectrum with 5-methyltryptophan, suggesting exchange broadening between two conformers. Figure 7C shows the corresponding differences in ^{15}N chemical shifts between the two TrpR−*trpO*^{M} ternary complexes with different corepressors, TrpR−*trpO*^{M}−5-methyltryptophan

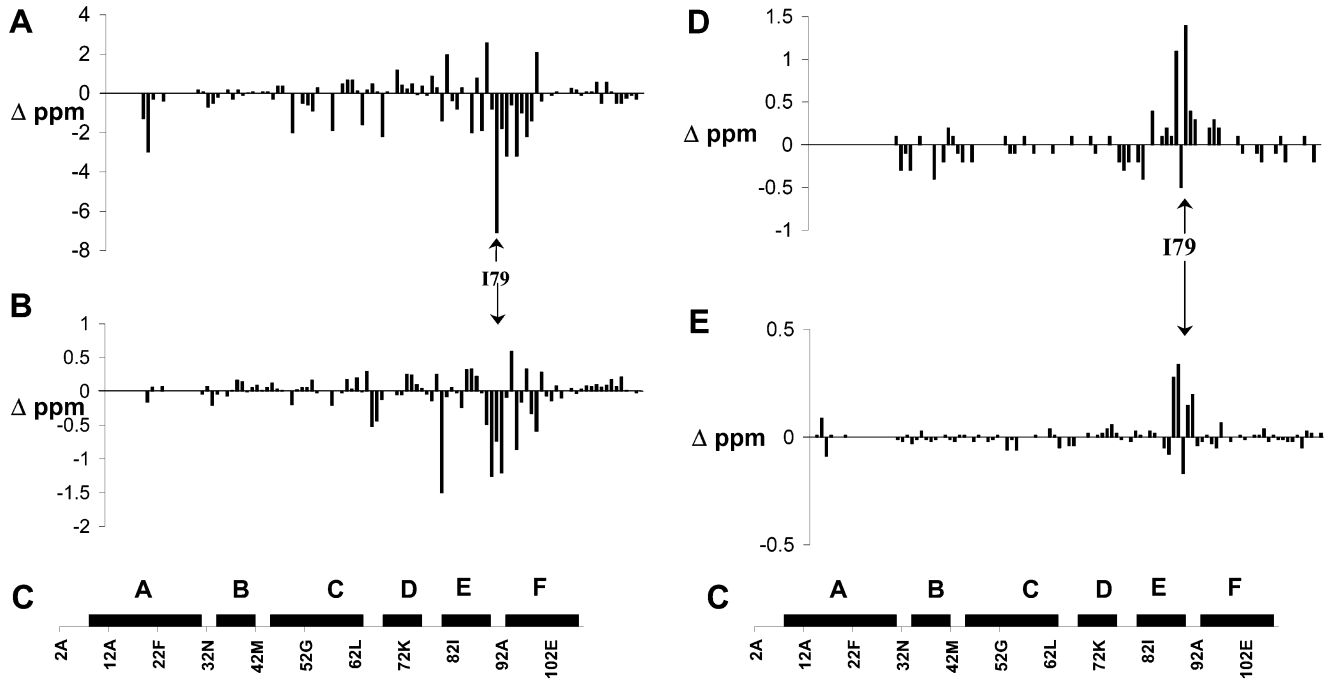

FIGURE 6: Histogram showing the changes in chemical shifts of TrpR−L-tryptophan residues on binding *trpR*^{S} DNA (A, B) and between the TrpR−*trpR*^{S}−L-tryptophan ternary complex and the TrpR−*trpO*^{M}−L-tryptophan ternary complex (D, E). (A, D) Differences in ^{15}N shift. (B, E) Differences in ^{1}H shifts. The chemical shift differences are given in parts per million. Chemical shifts are accurate to ±0.02 ppm in ^{1}H and ±0.05 ppm in ^{15}N. (C) shows the residue numbers, with the positions of α-helices as black rods labeled as in Figure 1. The position of I79 in each histogram is marked.

and TrpR−*trpO*^{M}−L-tryptophan. Several of the resonances affected on a change of corepressor in the binary complex are also affected in the *trpO*^{M} ternary complexes. Large effects are observed at E60, L61, and L105, with the largest effect at G52 (−1.0 ppm). For the two TrpR−*trpR*^{S} ternary complexes with different corepressors, TrpR−*trpR*^{S}−5-methyltryptophan and TrpR−*trpR*^{S}−L-tryptophan (Figure 7B), the same residues are affected on a change in corepressor as in the *trpO*^{M} ternary complexes. However, the pattern of shift changes is different, and for the *trpR*^{S} ternary complexes, the largest effect is at E60 (0.94 ppm).

DISCUSSION

To examine the molecular basis of the specificity of Trp repressor for its operators, we have compared its binding to two symmetrical operator variants. Each variant contains all the consensus bases recognized by a single TrpR dimer, contacted *via* hydrogen bonds to bases. They differ only at bases near the center of the operator, contacted at the phosphate groups. The *trpO*^{M} sequence is a variant of the *trpO*^{S} operator, differing only at base pair 14−7. The *trpR*^{S} sequence is a symmetrical version of the *trpR* operator. It differs from both the *trpO*^{S} and the *trpO*^{M} sequences at two base pairs, namely, 12−9 and 14−7 (Table 1, Figure 1B). The EMSA studies of Trp repressor binding to DNA containing the *trpO*^{S} and *trpR* sequences shows bands corresponding to both 2:1 and 1:1 TrpR−DNA complexes. Trp repressor binding to the *trpO*^{M} operator in the presence of L-tryptophan shows only a 2:1 TrpR−*trpO*^{M} complex, which dissociates in the gel. However, in the presence of 5-methyltryptophan, or when glycerol is present in the gel containing tryptophan, both 2:1 and 1:1 TrpR−*trpO*^{M} complexes are seen. We conclude that the TrpR−*trpO*^{M} complexes have fast dissociation rates and so are not observed in EMSA unless stabilized either by 5-methyltryptophan or by reducing the dissociation rate with glycerol. The association constant of the *trpO*^{M} fragment to give a 1:1 complex is 2−3-fold less than that of *trpO*^{S}, and that of the 2:1 complex is 15−20-fold less. In contrast, while the 1:1 TrpR−*trpR*^{S} complex is formed at concentrations of Trp repressor similar to those required for formation of 1:1 TrpR−*trpO*^{S} complex, the 2:1 complex is not observed, even in the presence of 5-methyltryptophan. These differences in Trp repressor complexes with *trpR*^{S} and *trpO*^{M} DNA led us to examine them using NMR spectroscopy.

Despite the EMSAs, evidence for both 1:1 and 2:1 complexes in the presence of L-tryptophan for TrpR−*trpR*^{S} as well as for TrpR−*trpO*^{M} is observed in NMR studies. On addition of only 1 equivalents of *trpR*^{S} DNA to the TrpR−L-tryptophan holoprotein, the line widths of the peaks in the NMR spectra of the protein are greater than expected for the size of the 1:1 complex (data not shown). As the amount of added DNA is increased above 1 equivalents, the line widths of the peaks decrease to those expected for a 37 kDa complex. This shows that there are two (or more) protein complexes in equilibrium and that the position of the equilibrium is affected by DNA concentration. These species are likely to be the 2:1 protein dimer−DNA complex and 1:1 protein dimer−DNA complex. At low DNA concentrations, these are in fast to intermediate exchange on the NMR time scale, leading to line broadening of the resonances. On adding excess DNA, more of the 1:1 protein−DNA complex is formed, so the signals from this dominate the appearance of the spectrum. The same effect is observed with *trpO*^{M} DNA, but more DNA needs to be added to obtain a sharp spectrum corresponding to the 1:1 protein−DNA complex. This is in agreement with the weaker binding of *trpO*^{M} to Trp repressor seen in the EMSA. Calorimetric and fluores-

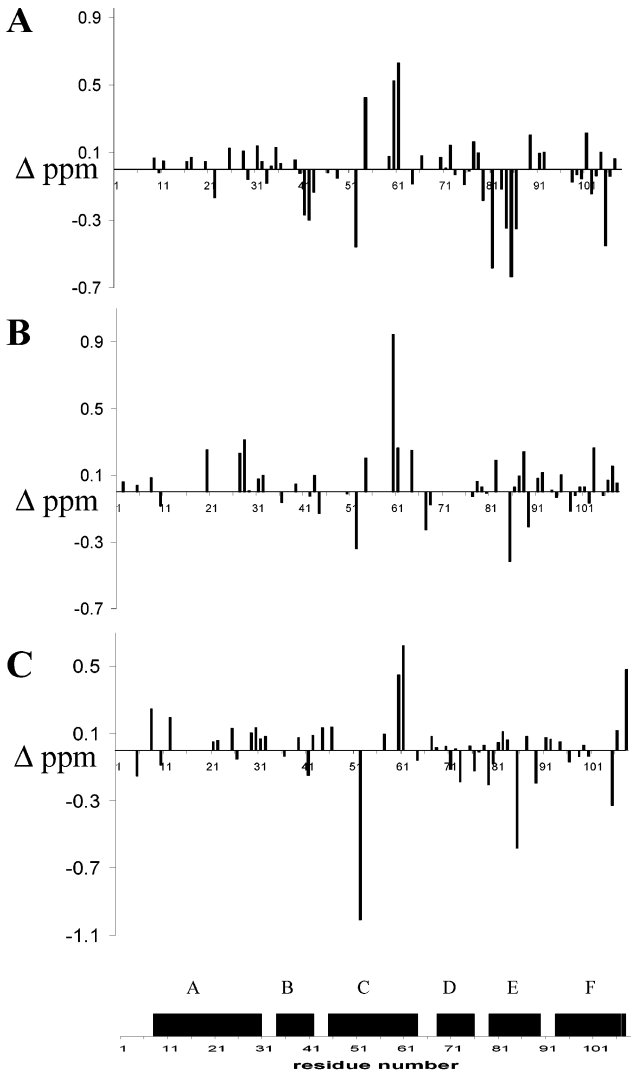

FIGURE 7: Histogram showing the difference in ^{15}N chemical shifts between Trp repressor complexes with L-tryptophan and with 5-methyltryptophan as corepressor. Chemical shift differences (A) between the two binary Trp repressor complexes TrpR−L-tryptophan and TrpR−5-methyltryptophan, (B) between the TrpR−*trpR*^{S} operator complexes TrpR−*trpR*^{S}−L-tryptophan and TrpR−*trpR*^{S}−5-methyl tryptophan, and (C) between the TrpR−*trpO*^{M} complexes TrpR−*trpO*^{M}−L-tryptophan and TrpR−*trpO*^{M}−5-methyltryptophan. The lowest histogram shows the residue numbers, with the positions of α-helices as black rods labeled as in Figure 1.

cence anisotropy studies with short (20−25 bp) oligonucleotides based on the *trpO* sequence also show formation of 2:1 TrpR−DNA complexes (*40*, *41*). These occur at Trp repressor concentrations lower than that of binding at nonspecific sequences and have been interpreted as tandem binding along the DNA, despite its short length. Our results suggest that Trp repressor forms 2:1 and 1:1 protein−DNA complexes in solution with all four DNA species studied, but EMSA selects only those that are kinetically stable. Thus, the use of EMSA to determine equilibrium constants and stoichiometry may be misleading for fast-exchanging DNA complexes.

The DNA binding observed in our NMR studies is dependent on the presence of the corepressor. In the absence of corepressor, no changes in shift are observed in the NMR spectra of either *trpO*^{M} DNA or Trp aporepressor when they are added to each other (data not shown). The *trpR*^{S} DNA binds slightly to the aporepressor at low ionic strength, as some line broadening and shifts are seen in the NMR spectra of both DNA and protein on their addition. However, no changes are observed in the NMR spectra at high ionic strength. This suggests that the binding of Trp apoprotein to *trpR*^{S} DNA is weak and mainly electrostatic. The difference in Trp aporepressor binding to *trpR*^{S} DNA and to *trpO*^{M} DNA confirms that the apoprotein shows some DNA selectivity. This selectivity has also been observed in fluorescence polarization studies (*41*). The differences in effects observed on DNA binding between the holorepressor and aporepressor confirm that the ternary complexes studied here, containing Trp repressor, DNA, and corepressor, are specific.

In the presence of L-tryptophan, large changes are observed in the chemical shifts of Trp repressor on DNA binding. The largest changes in chemical shifts occur at the helix−turn−helix region of the protein, residues 68−89. The changes observed on binding *trpR*^{S} DNA are similar to those previously observed on forming the TrpR−*trpO*^{S}−L-tryptophan complex (*13*) and can be rationalized on the basis of the structure of this complex. Figure 1B shows a schematic view of the DNA−protein contacts in the crystal structure of the 1:1 TrpR−*trpO*^{S}−L-tryptophan complex (*11*). In the current study, the largest changes in chemical shifts of Trp repressor on binding *trpR*^{S} DNA are at I79, A80, T81, and T83. In the crystal structure of the *trpO*^{S} complex, I79, A80, and T83 hydrogen bond to the DNA bases *via* water-mediated contacts, while T81 hydrogen bonds to two phosphate groups. D46, Q68, R84, S86, and N87, which bond to the DNA backbone in the *trpO*^{S} complex, also show large changes in chemical shift on adding *trpR*^{S} DNA in this study. In addition, R69, which is the only residue that forms a direct hydrogen bond to a DNA base in the crystal structure, shows a large change in shift on binding *trpR*^{S} DNA. Changes are also observed in the C helix at G52 and R56, which lie close to the phosphate backbone and the corepressor, and at residues G76 and A77 in the turn before the recognition helix. These effects suggest that overall the orientations of the protein and DNA in the TrpR−*trpO*^{S} complex and the TrpR−*trpR*^{S} complex are similar.

The differences in chemical shifts between the two TrpR−DNA−L-tryptophan ternary complexes studied here, TrpR−*trpR*^{S}−L-tryptophan and TrpR−*trpO*^{M}−L-tryptophan, are much smaller than the effects observed on DNA binding to the binary complex. The chemical shift differences occur primarily at residues A77−T81, which are in the turn of the helix−turn−helix motif and in the first turn of the recognition helix. The changes in shifts for these residues on binding the two oligonucleotides vary in magnitude and direction. Both the ^{15}N and ^{1}H shifts of I79 are further downfield in the *trpO*^{M} complex than in the *trpR*^{S} complex, whereas the opposite is the case for A80. All of these shifts are further downfield in the two ternary complexes with DNA than in the binary TrpR−L-tryptophan complex. In contrast, the ^{15}N shift of G78 in the binary TrpR−L-tryptophan complex lies between the corresponding shifts in the two ternary DNA complexes TrpR−*trpR*^{S}−L-tryptophan and TrpR−*trpO*^{M}−L-tryptophan. The differences in effects observed at these three positions show that the shift differences are not due simply to a faster off rate for the DNA in the *trpO*^{M} complex. They reflect a difference in the conformation or environment of the protein in the two ternary complexes.

In the crystal structure of the TrpR−*trpO*^{S}−L-tryptophan complex, I79 and A80 interact with DNA base pairs that are common in the operators studied, while G78 and T81 interact with phosphate groups. G78 interacts with p14, while T81 interacts with p14 and p13. The sequences of the *trpR*^{S} and *trpO*^{M} operators differ at base pairs 12−9 and 14−7. These differences are likely to change the conformation of the backbone of the DNA, affecting its interactions with the protein. The indole proton of the bound corepressor also hydrogen bonds to p13. Its resonance is shifted downfield by 1.4 ppm on binding *trpR*^{S} DNA (*24*). However, its shift is the same in the *trpO*^{M} complex as in the *trpR*^{S} complex (data not shown), showing that its environment is unchanged. This suggests that p13 is not affected by the change in sequence. The effects observed at G78 and T81 are therefore probably due to changes at p14, rather than at p13. Little change is seen in the ^{15}N shifts of other residues that interact with central phosphate groups, but ^{15}N shift differences are seen at I79 and A80, which interact with the bases that are conserved in all these species. These data show that the changes to nonconserved bases, contacted only at the phosphate groups, affect the environment of the protein at residues that contact conserved bases elsewhere in the DNA.

We also examined the effect of a change in corepressor from L-tryptophan to 5-methyltryptophan in the binary complex and the ternary complexes with *trpR*^{S} and *trpO*^{M}. The tryptophan contacts in the crystal structures of the binary TrpR−L-tryptophan complex (*5*) and the ternary TrpR−*trpO*^{S}−L-tryptophan (*11*) are shown in Figure 8. Previous NMR studies and molecular dynamics calculations of corepressor binding to TrpR (*42*) suggest that the indole ring of 5-methyltryptophan is moved slightly relative to that of tryptophan. This allows the methyl group to be in a hydrophobic pocket. Comparison of the spectra of the two binary complexes TrpR−5-methyltryptophan and TrpR−L-tryptophan shows chemical shift differences at R54, G85, T81, and R84 and a broadening of the resonance of T44 in the former complex. All of these residues bond directly to the corepressor. Large shift changes between these two binary complexes also occur at G52, E60, L61, and S86. These residues are further from the corepressor binding site and are too far for these shift effects to be due to ring currents, but they may be influenced by changes in the hydrophobic pocket at V58 and L89.

Comparison of the spectra of the two TrpR−*trpO*^{M} ternary complexes, with 5-methyltryptophan and L-tryptophan, shows differences in chemical shifts at E60, L61, G85, and G52. The effects at E60, L61, and G85 are similar in magnitude to those in the two binary complexes, while the effect at G52 is larger in the ternary complex. R54 and R84 are not resolved in the TrpR−*trpO*^{M} complexes, so the effects here are unknown, but there is no effect at T81 or S86 on a change of corepressor from L-tryptophan to 5-methyltryptophan. In the two TrpR−*trpR*^{S} ternary complexes with 5-methyltryptophan and L-tryptophan, the same residues are affected as in the TrpR−*trpO*^{M} ternary complexes, but to different extents. This suggests that 5-methyltryptophan may be in a slightly different orientation in the *trpR*^{S} complex than in the *trpO*^{M} complex, leading to different longer range effects at G52, E60, and L61. This may explain the differential effect of the corepressor in the two complexes.

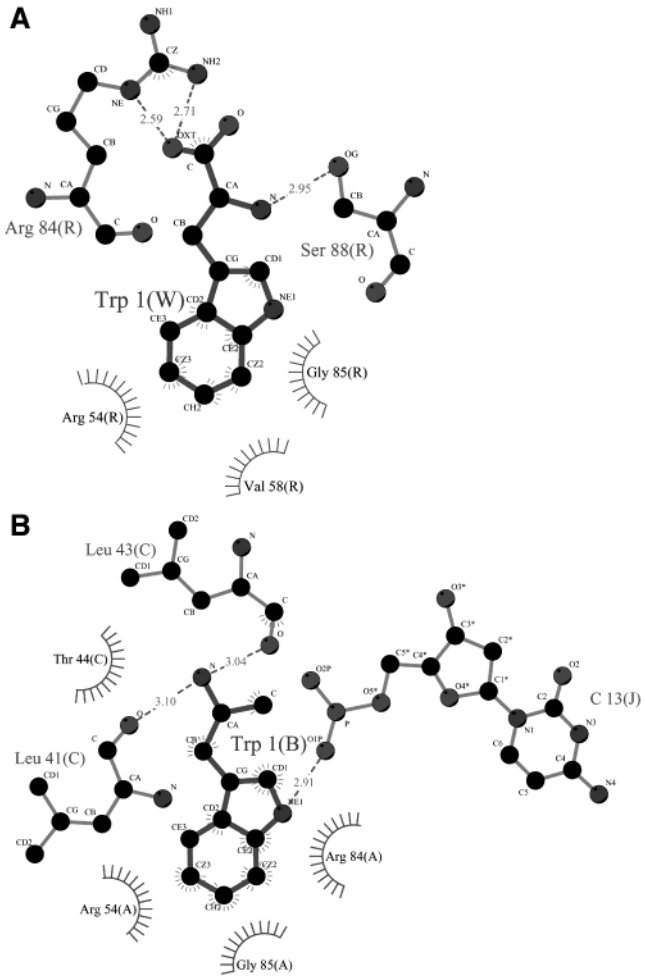

FIGURE 8: Schematic representation of the protein−L-tryptophan contacts in (A, top) the TrpR−L-tryptophan binary complex, 1WRP (*5*), and (B, bottom) the TrpR−*trpO*^{S}−L-tryptophan ternary complex, 1TRO (*11*). The diagrams were drawn using the program Ligplot (*44*). Residues forming ionic or hydrogen bond interactions with the corepressor are shown as ball-and-stick models, with the hydrogen bonds as lines with distances between the appropriate atoms. Residues with van der Waals contacts to the corepressor are shown as semicircles. In (A), (R) denotes the repressor and (W) the corepressor. In (B), (A) and (C) denote the two subunits of the repressor while (B) denotes the corepressor and (J) a strand of DNA.

Comparisons of our data with previous NMR studies of TrpR− *trpO*^{S} complexes with L-tryptophan, and with 5-methyltryptophan (*13*, *38*), suggest that the complexes with all three operators differ at residues at the tip of the recognition helix. As discussed above, the changes in shift observed on adding *trpR*^{S} DNA to Trp holorepressor in this study are similar to those on adding *trpO*^{S} DNA. This suggests that TrpR−*trpR*^{S} complexes are very close in structure to the TrpR−*trpO*^{S} complexes. Detailed comparisons of the shifts are difficult, as there are small, systematic, chemical shift differences between all the studies. We have numerically minimized the average shift differences, as described in the Experimental Procedures. Here we compare the chemical shifts of the C-terminal 38 residues, from residue 70 onward, after this minimization. Comparisons of the shifts of the two *trpO*^{S} ternary complexes studied previously, TrpR−*trpO*^{S}−L-tryptophan and TrpR−*trpO*^{S}−5-methyltryptophan, in this region show only seven residues

with differences in ^{15}N shift greater than 0.3 ppm. These are at K72, A77, T81, R84, G85, L89, and L105. Many of these residues have been shown in this study to be affected by changes in corepressor from L-tryptophan to 5-methyltryptophan. This suggests that the changes in corepressor affect the TrpR−*trpO*^{S} complexes in the same way as the TrpR−*trpR*^{S} and TrpR−*trpO*^{M} complexes. The *trpO*^{M} complex differs from the *trpO*^{S} complex at only one base pair, 14. The largest difference in ^{15}N shift between the TrpR−*trpO*^{M}−L-tryptophan complex and the TrpR−*trpO*^{S}−L-tryptophan complex is only 0.5 ppm. This occurs at two residues, at G78, which interacts with p13, and also at A80, which interacts with the conserved bases 15 and 16. There is also a difference in shift of 0.4 ppm at T83, which interacts with the conserved base 4, further from the site of mutation. The *trpR*^{S} and *trpO*^{S} complexes differ at two base pairs, 12 and 14. The TrpR−*trpR*^{S}−L-tryptophan and TrpR−*trpO*^{S}−L-tryptophan complexes show differences in ^{15}N shift greater than 0.4 ppm at residues 77−80, T83, and S88. The largest differences are at A80, 0.9 ppm, and at I79, 0.74 ppm, both of which interact with the conserved bases.

Thus, in complexes with all three DNA sequences, which differ at positions contacted only *via* phosphate groups, changes in environment are observed at residues that hydrogen bond to conserved bases. Changing the corepressor from L-tryptophan to 5-methyltryptophan causes shift changes at different positions compared to the effects of changes in DNA sequence. The subtle differences observed between all the ternary complexes suggest that the DNA and the protein both undergo mutual "induced fit" on binding; i.e., the entire protein−DNA interface is affected by a change in a single base pair or protein (or corepressor) residue. The changes observed at the protein residues that contact the conserved bases may explain why the *trpO*^{M} sequence binds sufficiently poorly to Trp repressor to act as a constitutive operator *in vivo*, despite the presence of the conserved base sequence. In contrast, the *trpR*^{S} sequence differs from *trpO*^{S} at the same position as *trpO*^{M} and at an additional base pair, yet binds equally well to the first TrpR dimer. Coupling between phosphate contacts (indirect read-out, as proposed in ref *11*) and contacts to bases, either water-mediated or direct, has been seen in a recent study of a mutant MetJ protein binding to three DNA sequences (*43*). Sequence selectivity in TrpR, MetJ, and other DNA-binding proteins appears to be due to a complex interplay among the energy required to deform DNA by the protein, the energy obtained from release of water at the interface between the molecules, and both direct and indirect interactions between the protein and the DNA. The effects of mutations of either DNA or protein (or, in this case, corepressor) cannot be explained as simple effects at a single point.

ACKNOWLEDGMENT

We thank K. J. Smith for help with the triple-resonance experiments, R. Parslow for skillful preparation of the protein, N. Banos for preliminary EMSA experiments, V. J. Howard for quantitation of the EMSA experiments presented, and Y. Gao and A. J. Pemberton for maintenance of the NMR equipment.

SUPPORTING INFORMATION AVAILABLE

Chemical shifts for the TrpR−*trpR*^{S}−L-tryptophan and TrpR−*trpO*^{M}−L-tryptophan complexes. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Somerville, R. (1992) *Prog. Nucleic Acid Res. Mol. Biol. 42*, 1−38.
2. Lawley, B., and Pittard, A. J. (1994) *J. Bacteriol. 176*, 6921−6930.
3. Marmorstein, R. Q., and Sigler, P. B. (1989) *J. Biol. Chem. 264*, 9149−9154.
4. Youderian, Y., and Arvidson, D. N.(1994) *Gene 150*, 1−8.
5. Schevitz, R. W., Otwinowski, Z., Joachimiak, A., Lawson, C. L., and Sigler, P. B. (1985) *Nature 317*, 782−786.
6. Zhao, D., Arrowsmith, C. H., Jia, X., and Jardetzky, O. (1993) *J. Mol. Biol. 229*, 735−746.
7. Zhang, R.-G., Joachimiak, A., Lawson, C. L., Schevitz, R. W., Otwinowski, Z., and Sigler, P. B. (1987) *Nature 327*, 591−597.
8. Lawson, C. L., Zhang, R.-G., Schevitz, R. W., Otwinowski, Z., Joachimiak, A., and Sigler, P. B. (1988) *Proteins 3*, 18−31.
9. Fernando, T., and Royer, C. A. (1992) *Biochemistry 31*, 3429−3441.
10. Martin, K. S., Royer, C. A., Howard, K. P., Carey, J., Liu, Y. C., Matthews, K., Heyduk, E., and Lee, J. C. (1994) *Biophys. J. 66*, 1167−1172.
11. Otwinowski, Z., Schevitz, R. W., Zhang, R.-G., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q., Luisi, B. F., and Sigler, P. B. (1988) *Nature 355*, 321−329.
12. Bass, S., Sugiono, P., Arvidson, D. N., Gunsalus, R. P., and Youderian, P. (1987) *Genes Dev. 1*, 565−572.
13. Zhang, H., Zhao, D., Revington, M., Lee, W., Jia, X., Arrowsmith, C., and Jardetzky, O. (1994) *J. Mol. Biol. 238*, 592−614.
14. Lawson, C. L., and Carey, J. (1993) *Nature 366*, 178−182.
15. Shan, X., Gardner, K. H., Muhandiram, D. R., Rao, N. S., Arrowsmith, C. H., and Kay, L. E. (1996) *J. Am. Chem. Soc. 118*, 6570−6579.
16. Shan, X., Gardner, K. H., Muhandiram, D. R., Kay, L. E., and Arrowsmith, C. H. (1998) *J. Biomol. NMR 11*, 307−318.
17. Staake, D., Walter, B., Kisters-Woike, B., Wilcken-Bergmann, B., and Muller-Hill, B. (1990) *EMBO J. 9*, 1963−1967.
18. Haran, T. E., Joachimiak, A., and Sigler, P. B. (1992) *EMBO J. 11*, 3021−3030.
19. Liu, Y.-C., and Matthews, K. S. (1993) *J. Biol. Chem. 268*, 23239−23249.
20. Gunes, C., Staake, D., von Wilken-Bergman, B., and Muller-Hill, B. (1996) *Mol. Microbiol. 20*, 375−384.
21. Bareket-Samish, A., Cohen, I., and Haran, T. E. (1997) *J. Mol. Biol. 267*, 103−117.
22. Bareket-Samish, A., Cohen, I., and Haran, T. E. (1998) *J. Mol. Biol. 277*, 1071−1080.
23. Jeeves, M., Evans, P. D., Parslow, R. A., Jaseja, M., and Hyde, E. I. (1999) *Eur. J. Biochem. 265*, 919−928.
24. Evans, P. D., Jaseja, M., Jeeves, M., and Hyde, E. I. (1996) *Eur. J. Biochem. 242*, 567−575.
25. Bennett, G. N., and Yanofsky, C. (1978) *J. Mol. Biol. 121*, 179−192.
26. Paluh, J. L., and Yanofsky, C. (1986) *Nucleic Acids Res. 14*, 7851−7860.
27. Muchmore, D. C., McIntosh, L. P., Russell, C. B., Anderson, D. E., and Dahlquist, F. W. (1989) *Methods Enzymol. 177*, 44−73.
28. Joachimiak, A., Kelly, R. L., Gunsalus, R. P., Yanofsky, C., and Sigler, P. B. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 668−672.
29. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular cloning, a laboratory manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
30. Carey, J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 975−979.
31. Fried, M., and Crothers, D. M. (1981) *Nucleic Acids Res. 9*, 6505−6525.
32. Zhang, O. W., Kay, L. E., Olivier, J. P., and Forman-Kay, J. D. (1994) *J. Biomol. NMR 4*, 845−856.
33. Grzesiek, S., and Bax, A. (1992) *J. Magn. Reson. 96*, 432−440.
34. Muhandiram, D. R., and Kay, L. E. (1994) *J. Magn. Reson. B 103*, 203−216.

35. Grzesiek, S., and Bax, A. (1992) *J. Magn. Reson. 99*, 201−207.
36. Grzesiek, S., and Bax, A. (1993) *J. Biomol. NMR 3*, 185−204.
37. Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) *J. Biomol. NMR 6*, 135−140.
38. Yamazaki, T., Lee, W., Arrowsmith, C. H., Muhandiram, D. R., and Kay, L. E. (1994) *J. Am. Chem. Soc. 116*, 11,655−11,666.
39. Liu, Y.-C., and Matthews, K. S. (1994) *J. Biol. Chem. 269*, 1692−1698.
40. Ladbury, J. E., Wright, J. G., Sturtevant, J. M., and Sigler, P. B. (1994) *J. Mol. Biol. 238*, 669−681.
41. Grillo, A. O., Brown, M. P., and Royer, C. A. (1999) *J. Mol. Biol. 287*, 539−554.
42. Ramesh, V., Syed, S. E. H., Frederick, R. O., Sutcliffe, M. J., Barnes, M., and Roberts, G. C. K. (1996) *Eur. J. Biochem. 235*, 804−813.
43. Garvie, C. W., and Phillips, S. E. V. (2000) *Structure 8*, 905−914.
44. Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) *Protein Eng. 8*, 27−134.

BI020072Y