

# NMR Structure and Functional Studies of the Mu Repressor DNA-Binding Domain<sup>†,‡</sup>

Udayar Ilangovan, Jonathan M. Wojciak, Kevin M. Connolly, and Robert T. Clubb\*

Department of Chemistry and Biochemistry and UCLA-DOE Laboratory of Structural Biology and Genetics,  
University of California—Los Angeles, 405 Hilgard Avenue, Los Angeles, California 90095-1570

Received March 5, 1999; Revised Manuscript Received April 23, 1999

**ABSTRACT:** The repressor protein of bacteriophage Mu establishes and maintains lysogeny by shutting down transposition functions needed for phage DNA replication. It interacts with several repeated DNA sequences within the early operator, preventing transcription from two divergent promoters. It also directly represses transposition by competing with the MuA transposase for an internal activation sequence (IAS) that is coincident with the operator and required for efficient transposition. The transposase and repressor proteins compete for the operator/IAS region using homologous DNA-binding domains located at their amino termini. Here we present the solution structure of the amino-terminal DNA-binding domain from the repressor protein determined by heteronuclear multidimensional nuclear magnetic resonance spectroscopy. The structure of the repressor DNA-binding domain provides insights into the molecular basis of several temperature sensitive mutations and, in combination with complementary experiments using fluorescence anisotropy, surface plasmon resonance, and circular dichroism, defines the structural and biochemical differences between the transposase and repressor DNA-binding modules. We find that the repressor and enhancer domains possess similar three-dimensional structures, thermostabilities, and intrinsic affinities for DNA. This latter result suggests that the higher affinity of the full-length repressor relative to that of the MuA transposase protein originates from cooperative interactions between repressor protomers and not from intrinsic differences in their DNA-binding domains. In addition, we present the results of nucleotide and amino acid mutagenesis which delimits the minimal repressor DNA-binding module and coarsely defines the nucleotide dependence of repressor binding.

The Mu bacteriophage is a temperate transducing phage of *Escherichia coli* (1) that replicates its 37 kb<sup>1</sup> linear double-stranded genome by multiple rounds of DNA transposition (2–4). The choice between lysogenic or lytic growth is governed by the Mu repressor and Ner proteins. During lysogeny, the Mu prophage is maintained in the repressed state by the Mu repressor protein, which binds to three adjacent operators that control the synthesis of early genes required for lytic growth. The Ner protein binds to a separate operator that controls the synthesis of the Mu repressor and initiates the lytic cycle when sufficient quantities of the protein are produced to inhibit the production of the Mu repressor. In addition to this transcription-based regulatory cycle, the Mu repressor also maintains the lysogenic state by directly competing with the MuA transposase for DNA sites required for efficient transposition.

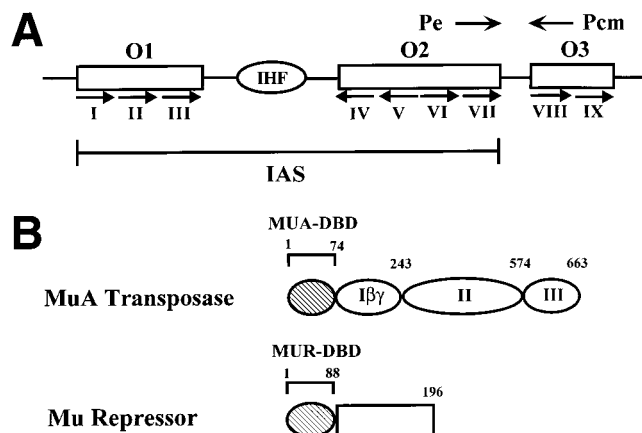
The transcription of early genes required for lytic growth is inhibited by repressor binding to an operator region positioned 1 kb from the left end of the phage genome. The operator is ~200 base pairs in length and contains three distinct operator sites (O1–O3), which function to modulate transcription from two divergent promoters, called Pe and Pcm (Figure 1A) (5, 6). Mu repressor binding to the operator region is highly cooperative and occurs through several distinct higher-order nucleoprotein complexes. At low concentrations, the repressor binds to sites O1 and O2 and prevents initiation of transcription at the early promoter Pe. Transcripts from Pe encode several proteins required for lytic growth, including the negative regulatory protein (ner), the MuA transposase protein, and the MuB protein which stimulates transposition and directs the transposase to sites of integration (7). At higher concentrations, the O3 operator is occupied by the repressor and transcription from the Pcm promoter is inhibited, preventing the synthesis of the c gene transcript encoding the repressor protein itself (6, 8, 9). The repressor also maintains the lysogenic state by directly inhibiting transposition. Efficient transposition of the Mu genome requires a transpositional DNA enhancer element called the internal activation sequence (IAS) (10, 11). The IAS element is coincident with the repressor operator (O1 and O2 sites) and is believed to be transiently bound by the MuA transposase protein during transposition. The repressor directly downregulates transposition in vivo and in vitro by competing with the MuA transposase for the IAS element.

<sup>†</sup> This work was supported by a grant from the U.S. Department of Energy (DE-FC-03-87ER60615).

<sup>‡</sup> The structure of the Mu repressor DNA-binding domain has been deposited in the Protein Data Bank (file name 1qPM).

\* To whom correspondence should be addressed. Telephone: (310) 206-2334. Fax: (310) 206-4749. E-mail: rclubb@mbi.ucla.edu.

<sup>1</sup> Abbreviations: HSQC, heteronuclear single-quantum coherence; IAS, internal activation sequence; kb, kilobase; MUA-DBD, amino-terminal DNA-binding domain from the MuA transposase protein; MUR-DBD, amino-terminal DNA-binding domain from the Mu bacteriophage repressor protein; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser enhancement; NOESY, nuclear Overhauser enhancement spectroscopy; R, purine; rms, root-mean-square; SPR, surface plasmon resonance; Y, pyrimidine.



**FIGURE 1:** Repressor and MuA transposase proteins contain homologous DNA-binding domains that compete for same DNA site within the Mu bacteriophage genome. (A) Schematic of the Mu repressor early operator. The three distinct regions within the operator that interact with the repressor are denoted as O1–O3. The internal activation sequence (IAS) that interacts with the MuA transposase protein is drawn below the operator. Promoters that initiate synthesis of the repressor protein and proteins involved in lytic growth are denoted by Pcm and Pe, respectively. The nine repressor and transposase binding sites that contain a conserved 14 bp consensus sequence are depicted as arrows and numbered I–IX. The binding site for the integration host factor protein is depicted as an oval between sites O1 and O2. (B) Schematic of the structures of the MuA transposase and repressor proteins. The conserved amino-terminal DNA-binding domain in both proteins is hatched and labeled. In the MuA transposase protein, ovals representing domains involved in end site DNA binding, catalysis, and interactions with MuB are labeled Iβγ, II, and III, respectively. The putative oligomerization domain in the Mu repressor protein is depicted as a rectangle.

Several host-encoded proteins modulate the activity of the repressor protein. Integration host factor (IHF) binds between the O1 and O2 sites and presumably stabilizes a looped complex in which sites O1 and O2 are occupied by the repressor and transcription from Pe is arrested (6, 12). A second cellular factor, the H-NS protein, has been shown *in vitro* to stabilize repressor binding to the operator, and *in vivo*, mutations in the H-NS protein increase the level of Mu-specific transcription and induce mini-Mu transposition (13). The repressor protein is selectively degraded by the ClpXP protease. Degradation alleviates inhibition of the Pe promoter, resulting in lytic growth, and may provide a mechanism in which environmental changes in the bacterial cell are propagated to the prophage to induce genome replication. Determinants for protease sensitivity are located at the carboxyl terminus of the repressor protein, and several virulent mutants contain frameshift mutations within this region and are rapidly degraded by the ClpXP (14, 15).

The full-length Mu repressor protein (21 850 Da, 196 amino acids) behaves as an oligomer of varying protomer concentrations in solution (16) and is comprised of two distinct domains: an amino-terminal DNA-binding domain (MUR-DBD) and a carboxyl-terminal region that is presumably involved in protein oligomerization (5). The sequence of MUR-DBD is homologous with the amino-terminal DNA-binding domain of the MuA transposase (MUA-DBD), and both proteins bind to the same DNA sequence within the operator/IAS element (Figure 1B) (17). Footprinting studies indicate that the DNA-binding modules interact with a nonsymmetric 14 bp consensus sequence CTTTTYARA/

TAANNA/T that is repeated nine times within the operator. The structure of the MUA-DBD has been determined, and this protein adopts an unusual winged helix–turn–helix motif (18). In this paper, we attempt to gain insights into repressor protein function and the molecular basis of its competition with the MuA transposase. Here we present (1) the NMR solution structure of the amino-terminal DNA-binding domain of the Mu repressor protein (MUR-DBD), (2) a comparison of the structures, DNA binding affinities, and thermostabilities of the isolated transposase and repressor domains, (3) the results of limited nucleotide mutagenesis of the repressor binding site, and (4) a map of the DNA binding surface on the MUR-DBD as determined by NMR studies of its complex with DNA.

## EXPERIMENTAL PROCEDURES

**Preparation of the Repressor and MuA Transposase DNA-Binding Domains.** Plasmids encoding the amino-terminal DNA-binding domain from the MuA transposase protein (MUA-DBD) were obtained as previously described (18). Expression plasmids encoding polypeptides corresponding to the DNA-binding domain from the Mu repressor protein (MUR-DBD<sup>1–88</sup>, residues 1–88, and MUR-DBD<sup>13–81</sup>, residues 13–81) were generated using standard techniques. Briefly, for each repressor construct, a *NdeI*–*Bam*HI fragment containing the coding sequence was generated by PCR from the full-length gene located on plasmid pMK282 (a gift from K. Mizuuchi). The fragments were then cloned into a pET11a expression vector, and used to transform *E. coli* host strain BL21(DE3). All proteins were overexpressed and purified using the following protocol. Eight liters of culture was grown at 37 °C and induced with an IPTG concentration of 1 mM when the absorbance at 600 nm reached ~0.6–0.7 optical density unit. The cells were then harvested, resuspended in 100 mM Tris-HCl (pH 7.2), 5 mM EDTA, 2.5 mM dithiothreitol (DTT), and 5 mM benzamidine, and mechanically lysed in a pressure cell. The lysate was centrifuged at 11 000 rpm for 45 min, and again at 30 000 rpm for 1 h. The supernatant was applied to a DEAE-Sepharose Fast Flow XK-50 column (Pharmacia) (500 mL bed volume) and the column eluted with a gradient of 1 M NaCl (0 to 100%) in 100 mM Tris-HCl (pH 7.2), 5 mM EDTA, and 2.5 mM DTT. Pooled fractions were dialyzed overnight against low-salt buffer containing 1 M urea and applied to a SP-Sepharose Fast Flow XK-50 column (Pharmacia) (250 mL bed volume), and the column was eluted with a gradient of 1 M NaCl (0 to 100%) in 100 mM Tris-HCl (pH 7.2), 5 mM EDTA, and 2.5 mM DTT. The repressor protein-containing fractions were concentrated to ~10 mL and loaded onto a Sephacryl S-100 column run in 50 mM phosphate (pH 6.2), 100 mM NaCl, and 2.5 mM DTT.

For structural studies, isotopic labeling of the repressor protein with <sup>15</sup>N, or <sup>15</sup>N and <sup>13</sup>C, was accomplished by growing the cells in minimal medium that contained <sup>15</sup>NH<sub>4</sub>-Cl and [<sup>13</sup>C<sub>6</sub>]glucose as the sole nitrogen and carbon sources, respectively. Samples for the NMR of the DNA free repressor protein were ~1.0 mM Mu repressor protein, 50 mM phosphate (pH 6.2), 100 mM NaCl, and 2.5 mM deuterated DTT. Three samples of the MUR-DBD<sup>1–88</sup> were made: [<sup>15</sup>N]-MUR-DBD<sup>1–88</sup> in 93:7 H<sub>2</sub>O/D<sub>2</sub>O, [<sup>15</sup>N,<sup>13</sup>C]MUR-DBD<sup>1–88</sup> in 93:7 H<sub>2</sub>O/D<sub>2</sub>O, and [<sup>15</sup>N,<sup>13</sup>C]MUR-DBD<sup>1–88</sup> in 100% D<sub>2</sub>O. The protein–DNA complex consisted of a 1:1 complex of

$^{15}\text{N}$ -labeled MUR-DBD<sup>13–81</sup> with its cognate binding site (dGCTTTTCAGTAATCTG•dCAGATTACTGAAAAGC). The corresponding NMR sample consisted of  $\sim 1.3$  mM MUR-DBD<sup>13–81</sup>–DNA complex, 25 mM phosphate (pH 6.0), 5 mM NaCl, and 25  $\mu\text{M}$  EDTA dissolved in 93:7  $\text{H}_2\text{O}/\text{D}_2\text{O}$ .

**NMR Spectroscopy.** NMR experiments were performed at 300 K on Bruker DRX 500 and 600 MHz spectrometers equipped with *xyz*-gradient triple-resonance probes. Protein resonances ( $^1\text{H}$ ,  $^{15}\text{N}$ , and  $^{13}\text{C}$ ) were assigned using three-dimensional (3D) HNCA, HNCO, and HNCOA (19); HNCACB (20); CBCA(CO)NH (21); C(CO)NH (22); HCCH total correlation spectroscopy (TOCSY) (23); HCCH COSY (24); and  $^{15}\text{N}$ -edited TOCSY (25) experiments.  $^3J_{\text{HN}\alpha}$  and  $^3J_{\text{HN}\beta}$  coupling constants were measured using two-dimensional and 3D HNHA (26) and HNHB (27) experiments. 3D  $^{15}\text{N}$ - and  $^{13}\text{C}$ -edited ROESY experiments aided in stereospecific methylene assignments (28). Distance restraints were obtained from 3D  $^{15}\text{N}$ - and  $^{13}\text{C}$ -edited NOESY (25, 29), and four-dimensional (4D) HCCH NOESY (30) experiments (mixing times of 70–120 ms). One-bond HN  $^1D_{\text{HN}}$  residual dipolar couplings were measured on a sample containing 0.8 mM  $^{15}\text{N}$ -labeled protein weakly aligned in 5% 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine/1,2-Dihexanol-*sn*-glycero-3-phosphocholine (3:1 DMPC/DHPC) bicelles (31, 32). Residual dipolar couplings were measured by recording two-dimensional F1- $^1\text{H}$ -coupled ( $^{15}\text{N}$ ,  $^1\text{H}$ )-HSQC spectra at 27 (isotropic) and 38 °C (partially aligned) and calculating the difference in the  $^1J_{\text{HN}}$  couplings. The magnitudes of the axial and rhombic components of the alignment tensor were determined by examining the distribution of dipolar couplings (33). This analysis yielded values for  $D_a^{\text{NH}}$  of 3.7 Hz and for  $R$  of 0.4, where  $D_a^{\text{NH}}$  is the axial component of the tensor and  $R$  is the rhombicity defined as the ratio of the rhombic and axial components of the tensor. Where appropriate, solvent suppression was achieved using pulse-field gradients (34). Spectra were processed using NMRPipe (35) and analyzed using the programs PIPP, CAPP, and STAPP (36).

**Structure Calculations.** Structures of the free MUR-DBD<sup>1–88</sup> were determined using the program XPLOR (37) modified to include terms for  $^3J_{\text{HN}\alpha}$  couplings (38),  $^{13}\text{C}\alpha$  and  $^{13}\text{C}\beta$  chemical shifts (39), conformational database refinement (40), and residual dipolar couplings in the target function (a gift from G. M. Clore). Distance restraints were grouped into four distance ranges: 1.8–2.7 Å (1.8–2.9 Å for distances involving  $^{15}\text{N}$ -bound protons), 1.8–3.3 Å (1.8–3.5 Å for distances involving  $^{15}\text{N}$ -bound protons), 1.8–5.0 Å, and 1.8–6.0 Å. A total of 0.5 Å was added to the upper distance limits of NOE distances involving methyl protons to account for the increased apparent intensities of methyl resonances. Distances involving methyl protons, aromatic ring protons, and non-stereospecifically assigned methylene protons were represented as a  $(\sum r_i^{-6})^{-1/6}$  sum. Hydrogen bond restraints were employed in areas of regular secondary structure and were introduced at the final stages of refinement. Two distance restraints were used for each hydrogen bond ( $r_{\text{NH}\cdots\text{O}} < 2.5$  Å and  $r_{\text{N}\cdots\text{O}} < 3.5$  Å).

The structures were calculated by a hybrid distance geometry-simulated annealing (DGSA) protocol (41), followed by an additional simulated annealing step on each of the initial DGSA structures. The simulated annealing protocol was as follows. The initial phase comprised 10 ps of

dynamics (5000 integration time steps each lasting 2 fs) at 3000 K with force constants for the NOEs, dihedral angles, carbon chemical shifts, coupling constants, bonds, angles, impropers, residual dipolar couplings, conformational database, and van der Waals terms set to 2 kcal mol<sup>−1</sup> Å<sup>−2</sup>, 10 kcal mol<sup>−1</sup> rad<sup>−2</sup>, 0.5 kcal mol<sup>−1</sup> ppm<sup>−2</sup>, 1 kcal mol<sup>−1</sup> Hz<sup>−2</sup>, 1000 kcal mol<sup>−1</sup> Å<sup>−2</sup>, 200 kcal mol<sup>−1</sup> rad<sup>−2</sup>, 50 kcal mol<sup>−1</sup> rad<sup>−2</sup>, 0.01 kcal mol<sup>−1</sup> Hz<sup>−2</sup>, 0.002, and 1 kcal mol<sup>−1</sup> Å<sup>−4</sup> (with the van der Waals radius scale factor set to 0.9), respectively. In the second phase, the system was slowly cooled (annealed) from 3000 to 100 K over the course of 24 ps. There were 116 cycles of cooling, each lasting 0.206 ps (103 integration times of 2 fs each), with a decrease in temperature of 25 K per cycle. During this period, the force constants for the dihedral angle, carbon shift, coupling constant, and bond terms were held constant at 200 kcal mol<sup>−1</sup> rad<sup>−2</sup>, 0.5 kcal mol<sup>−1</sup> ppm<sup>−2</sup>, 1 kcal mol<sup>−1</sup> Hz<sup>−2</sup>, and 1000 kcal mol<sup>−1</sup> Å<sup>−2</sup>, respectively, while the other force constants were increased in each cycle by (final value/initial value)<sup>1/116</sup>. The force constants for NOEs, angles, impropers, conformational database, van der Waals terms, and dipolar couplings were increased from 2 to 30 kcal mol<sup>−1</sup> Å<sup>−2</sup>, 200 to 500 kcal mol<sup>−1</sup> rad<sup>−2</sup>, 50 to 500 kcal mol<sup>−1</sup> rad<sup>−2</sup>, 0.002 to 1, 0.01 to 2 kcal mol<sup>−1</sup> Hz<sup>−2</sup>, and 0.004 to 4 kcal mol<sup>−1</sup> Å<sup>−4</sup>, respectively; the van der Waals radius scale factor was decreased from 0.9 to 0.8. Finally, 250 cycles of Powell minimization were performed using the final values for the various force constants. The figures were prepared using the program MOLMOL (42).

**Fluorescence Anisotropy Measurements.** Fluorescence anisotropy titration measurements were performed using a Beacon 2000 fluorescence polarization system (PanVera Corp., Madison, WI). Protein–DNA binding was monitored by observing changes in fluorescence polarization of a fluorescein-tagged DNA duplex (reviewed in ref 43). The amount of fluorescence polarization was measured directly as a function of the amount of protein added using the appropriate fixed excitation (490 nm) and emission (530 nm) wavelengths. The DNA molecule used in the binding assay was constructed from two oligonucleotides synthesized by standard phosphoramidite chemistry and contained a single fluorescein molecule (label) attached to the 5′-end of one strand (5′-F-CGCCTTTTCAGTAAGCTGGGG-3′ and 5′-CCCCAGCTTACTGAAAAGGCG-3′, where F is the fluorescein molecule). The 21 bp double-stranded duplex was then annealed in buffer containing 1 M NaCl, 50 mM PO<sub>4</sub>, and 0.1 mM EDTA (pH 7.6). Mu repressor protein dissolved in binding buffer [50 mM PO<sub>4</sub> and 0.1 mM EDTA (pH 7.6)] was serially diluted from 150  $\mu\text{M}$  to 10 nM in disposable borosilicate glass tubes. Twenty-five microliters of the Mu DNA at a concentration of 50 nM was then added to each tube and the total volume adjusted to 250  $\mu\text{L}$  by addition of binding buffer. Five polarization measurements were made for each point in the assay, and the average was used for subsequent analysis. The average fluorescence polarization ( $P$ ) obtained for each sample was converted to fluorescence anisotropy ( $A$ ) using eq 1.

$$A = (2P)/(3 - P) \quad (1)$$

Dissociation constants were calculated by nonlinear least-squares fitting of the fluorescence anisotropy values to a



binding model that assumes 1:1 stoichiometry (eq 2).

$$A = (A_f + [(A_b - A_f)[L]) / (K_D + [L]) \quad (2)$$

where  $A$  is the measured anisotropy in a mixture of free and bound fluorescent molecules with anisotropies of  $A_f$  and  $A_b$ , respectively,  $K_D$  is the equilibrium dissociation constant, and  $[L]$  is the concentration of the protein. All proteins were assayed in an identical manner, except that in the MuA transposase studies the protein was serially diluted from 150  $\mu$ M to 100 nM. The reported dissociation constants are the average of at least two independent titration experiments, and the errors are the maximum observed deviation from this average. The binding data were fit using the program DeltaGraph (version 4.0.1, Deltapoint Inc., Monterey, CA).

**Surface Plasmon Resonance.** Binding of the repressor protein to wild-type and mutated DNA binding sites was monitored in real time by surface plasmon resonance (SPR) experiments performed on a BIAcore instrument (Pharmacia Biosensor, Uppsala, Sweden). For the SPR measurements, the following oligonucleotides were synthesized: 5'-XGGGGCCTTTTCAGTAAGCTGGGGG-3', 3'-CCCCGAAAA-GTCATTCGACCCCC-5', 5'-XGGGGCCGCGTCAGTAAGCTGGGGG-3', 3'-CCCCGGCGCAGTCATTCGACCCCC-5', 5'-XGGGGCCTTTTCAGGCGGCTGGGGG-3', 3'-CCCCGAAAAGTCCGCCGACCCCC-5', where X is biotin. After annealing of the appropriate oligonucleotides, the biotinylated DNA duplex was immobilized on the streptavidin-derivatized dextran matrix of the sensor chip according to the manufacturer's recommendations. DNA immobilization was monitored by observing the SPR response and continued until the DNA gave a response of  $\sim 500$  RU (resonance units).

The SPR response was measured in RU as the protein associated and dissociated from the duplex. The flow rate in the assay was 5  $\mu$ L/min, and the binding buffer consisted of 0.01 M HEPES [*N*-(2-hydroxyethyl)piperazine-*N'*-2-ethanesulfonic acid] (pH 7.4), 0.15 M NaCl, 3.0 mM EDTA, and 0.005% v/v surfactant P20. The response data were recorded as sensograms and analyzed using the BIAevaluation 3.0 software program. First, the dissociation rate ( $k_d$ ) of the complex was determined using eq 3:

$$\ln(dR/dt) = \ln(-k_d R_0) - k_d(t - t_0) \quad (3)$$

where  $R$  is the response at time  $t$  and  $R_0$  is the response at an arbitrary starting time  $t_0$  (not necessarily the beginning of the dissociation phase). The dissociation rate was extracted by plotting the value of  $\ln(dR/dt)$  versus  $t - t_0$ , which yields a straight line with the slope  $k_d$ . The next step in the analysis determined the association rate ( $k_a$ ) of the complex using eq 4:

$$\ln(dR/dt) = \ln(k_a C R_{\max}) - (k_a C + k_d)t \quad (4)$$

where  $C$  is the concentration of the protein,  $R_{\max}$  is the maximum protein binding capacity in RU, and  $R$  is the SPR signal in RU at time  $t$ . A plot of  $\ln(dR/dt)$  versus  $t$  yields a straight line with a slope equal to  $-(k_a C + k_d)$ . Because the concentration of the protein is known and  $k_d$  has been previously determined using eq 3, the value of  $k_a$  could be calculated directly. Finally, the equilibrium dissociation constant was then calculated from the ratios of the previously

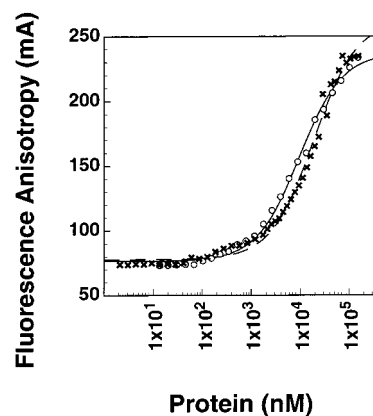


FIGURE 2: Representative fluorescence anisotropy measurements of the MuA transposase and repressor proteins. The figure shows a plot of the changes in fluorescence anisotropy of a fluorescein-tagged DNA duplex as a result of protein binding. Circles and crosses depict data for the MUA-DBD<sup>1-74</sup> and MUR-DBD<sup>1-88</sup>, respectively. The best fit of the data to the binding model assuming 1:1 stoichiometry is shown as solid and dashed lines for the MUA- and MUR-DBDs, respectively.

determined rate constants. The reported dissociation constants are the average of at least two independent titration experiments, and the errors are the maximum observed deviation from the average.

## RESULTS

**Binding Studies of the Isolated Repressor and Transposase DNA-Binding Domains.** The repressor and MuA transposase proteins contain homologous amino-terminal DNA-binding domains that compete for the same DNA site binding within the early operator of the Mu bacteriophage. DNase I protection studies of the full-length repressor and MuA transposase proteins indicate that these domains exhibit substantially different affinities for the operator; the intact MuA transposase binds with a low affinity ( $K_D \sim 1-100$   $\mu$ M) (10, 11), while the full-length repressor protein binds with a high affinity ( $K_D \sim 3-5$  nM) (5). To determine whether the affinity difference between the repressor and MuA transposase proteins is an intrinsic property of each DNA-binding domain or a result of positive cooperativity between repressor protomers, we subcloned, overexpressed, and purified two polypeptide fragments that comprise the amino-terminal DNA-binding domains of the MuA transposase (MUA-DBD<sup>1-74</sup>, residues 1-74) and repressor (MUR-DBD<sup>1-88</sup>, residues 1-88) proteins. Fluorescence anisotropy measurements were then used to measure the binding affinities of the individual MUA- and MUR-DBDs for their cognate DNA site. The assay was performed using a fluorescein-labeled DNA molecule that contained the nucleotide sequence CTTTTCAGTAAGCT, which completely matches the consensus sequence (CTTTTYARA/TAANNA/T) found in all nine binding sites (Figure 1A). The fluorescence anisotropy of the DNA duplex was measured as a function of varying MUR- or MUA-DBD concentrations and was used to extract the dissociation constant for protein binding. The results of this assay are displayed in Figure 2A. At saturating concentrations, both proteins form complexes with similar total fluorescence anisotropy values ( $\sim 250$  mA), indicating that the final sizes (molecular correlation times) of the respective complexes are nearly

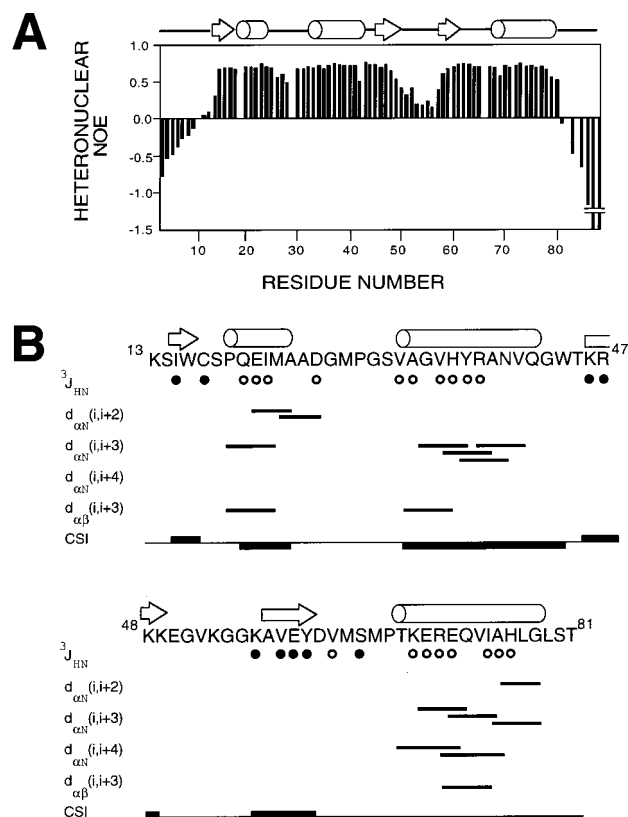


FIGURE 3: Summary of NMR data used to compute the solution structure of the Mu repressor DNA-binding domain. (A)  $^{15}\text{N}$ – $^1\text{H}$  heteronuclear NOE data recorded on a uniformly  $^{15}\text{N}$ -labeled sample of MUR-DBD<sup>1–88</sup>. The figure has been truncated for clarity, and the measured NOEs for residues Gln87 and Ala88 are  $-1.9$  and  $-3.4$ , respectively. (B) NMR data indicative of secondary structure. The amino acid sequence of the structured portion of the MUR-DBD<sup>1–88</sup> is shown (residues Lys13–Thr81). Lines are used to connect residues that exhibit medium-range NOEs.  $d_{\alpha N}(i, i+2)$ ,  $d_{\alpha N}(i, i+3)$ , and  $d_{\alpha N}(i, i+4)$  represent NOEs between the  $\alpha$  proton of residue  $i$  and the amide proton of residue  $i+2$ ,  $i+3$ , and  $i+4$ , respectively.  $d_{\alpha\beta}(i, i+3)$  represents an NOE between the  $\alpha$  proton of residue  $i$  and the  $\beta$  proton of residue  $i+3$ . The black and white circles correspond to  $^3J_{HN\alpha}$  coupling constants that are  $>8$  and  $<5.5$  Hz, respectively. The row marked CSI indicates the chemical shift index of the repressor and is a composite of the secondary chemical shifts of the  $\alpha$  and  $\beta$  carbons in the protein (48). Values above and below the line correspond to secondary chemical shifts that are consistent with  $\beta$ -sheet and  $\alpha$ -helical conformations, respectively. The CSI values were calculated using the program CSI (Department of Biochemistry, University of Alberta, Edmonton, AB).

equal. Since the polypeptides are of similar length, this suggests that the stoichiometries of the respective protein–DNA complexes are identical. A nonlinear least-squares analysis of the experimental data using a 1:1 protein–DNA binding model yielded reliable values for the equilibrium dissociation constants. The MUA- and MUR-DBDs were found to bind to the operator with dissociation constants ( $K_D$ ) of  $5.5 \pm 0.3$  and  $17.8 \pm 1.1 \mu\text{M}$ , respectively. These results suggest that contrary to the full-length proteins, the MUA- and MUR-DBDs have approximately the same intrinsic affinity for the consensus binding site within the operator. It thus appears that the higher affinity of the full-length repressor relative to that of the MuA transposase protein arises from cooperative interactions between repressor promoters.

*Structure of the Mu Repressor DNA-Binding Domain.* Having established that the DNA-binding modules within

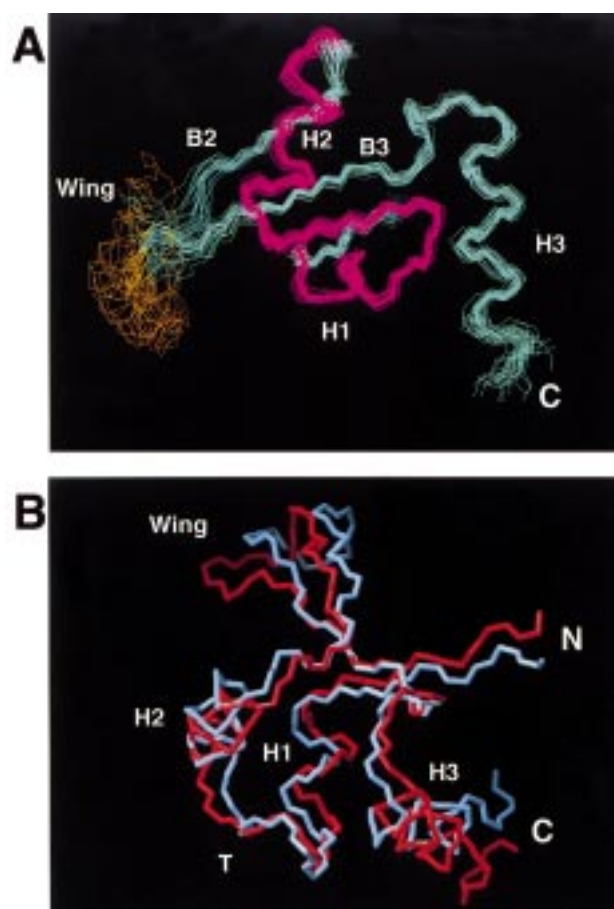


FIGURE 4: NMR solution structure of the amino-terminal domain of the Mu repressor protein. (A) Ensemble of 25 structures of the Mu repressor DNA-binding domain. The backbone atoms (C, C $\alpha$ , and N) of residues Lys13–Leu79 are shown: gold, wing structure; magenta, helix–turn–helix (H1–T–H2); and blue, the remainder of the polypeptide. (B) Comparison of the structures of the Mu repressor and MuA transposase DNA-binding domains. The energy-minimized average coordinates of the MUA-DBD (residues Met1–Glu67, Protein Data Bank file name 1tns) and the MUR-DBD (residues Lys13–Thr81) are colored blue and red, respectively. Residues Leu3–Ala38 and Ala45–Gly66 of the MUA-DBD were superimposed with residues Ile15–Glu50 and Ala57–Gly78 of the MUR-DBD.

the repressor and transposase have similar intrinsic affinities for the operator/IAS site, we next determined the structure of the repressor domain to enable a detailed comparison of their structures. Double- and triple-resonance 3D and 4D NMR spectroscopy was used to determine the structure of the amino-terminal DNA-binding domain from the Mu repressor protein. NMR experiments were performed on a polypeptide fragment corresponding to residues 1–88 of the Mu repressor protein (MUR-DBD<sup>1–88</sup>) that was uniformly labeled with either  $^{15}\text{N}$  or  $^{15}\text{N}$  and  $^{13}\text{C}$ . Nearly complete assignments of the  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{15}\text{N}$  nuclei and an analysis of the NMR data suggested that only residues Lys13–Ser80 in the MUR-DBD<sup>1–88</sup> were structured in solution. Figure 3A shows a plot of the  $^1\text{H}$ – $^{15}\text{N}$  heteronuclear NOE values for the full fragment (44). Residues Lys2–Asn10 at the amino terminus and residues Thr81–Ala88 at the carboxyl terminus exhibit negative heteronuclear NOE values, suggesting that they are disordered in solution. Extensive analysis of the NMR data supports this conclusion, since only residues Lys13–Ser80 exhibited features characteristic of a folded domain. Figure 3B summarizes this analysis and displays

Table 1: Structural Statistics<sup>a</sup>

	$\langle SA \rangle$	$\overline{(SA)}_r$
rms deviations from NOE interproton distance restraints ( $\text{\AA}$ ) <sup>b</sup>		
all (648)	$0.028 \pm 0.004$	0.032
interresidue sequential ( $ i - j  = 1$ ) (186)	$0.020 \pm 0.005$	0.035
interresidue short-range ( $1 <  i - j  < 5$ ) (96)	$0.031 \pm 0.007$	0.050
interresidue long-range ( $ i - j  > 5$ ) (183)	$0.035 \pm 0.005$	0.031
intraresidue (183)	$0.020 \pm 0.012$	0.017
rms deviations from hydrogen bonding restraints ( $\text{\AA}$ ) (46)	$0.041 \pm 0.013$	0.013
rms deviations from experimental dihedral angle restraints (deg) (109) <sup>c</sup>	$0.078 \pm 0.085$	0.001
rms deviations from experimental $^3J_{\text{HN}\alpha}$ coupling constants (Hz) (43)	$0.618 \pm 0.038$	0.704
rms deviations from experimental $^1D_{\text{HN}}$ residual dipolar coupling constants (Hz) (43)	$0.669 \pm 0.033$	0.698
rms deviations from experimental $^{13}\text{C}$ shifts		
$^{13}\text{C}\alpha$ (ppm) (47)	$1.115 \pm 0.053$	1.170
$^{13}\text{C}\beta$ (ppm) (47)	$1.154 \pm 0.046$	1.279
deviations from idealized covalent geometry		
bonds ( $\text{\AA}$ )	$0.003 \pm 0.0002$	0.005
angles (deg)	$0.431 \pm 0.031$	0.666
impropers (deg)	$0.451 \pm 0.053$	0.536
coordinate precision ( $\text{\AA}$ ) <sup>d</sup>		
protein backbone	$0.61 \pm 0.16$	
heavy atoms	$1.10 \pm 0.15$	

<sup>a</sup> The notation of the NMR structures is as follows.  $\langle SA \rangle$  are the final 25 simulated annealing structures;  $\overline{(SA)}_r$  is the restrained regularized mean structure obtained by averaging the coordinates of the individual SA structures best fitted to one another with respect to residues Lys6–Glu50 and Lys56–Leu79, and restrained regularization of this mean structure. The number of terms for each restraint is given in parentheses. Residues Met1–Glu12 and Pro82–Ala88 exhibited no long-range NOE cross-peaks in the data and were omitted from the final simulated annealing calculations.

<sup>b</sup> None of the structures exhibited distance violations of  $>0.5$   $\text{\AA}$ , dihedral angle violations of  $>5^\circ$ , or  $^3J_{\text{HN}\alpha}$  scalar coupling and  $^1D_{\text{HN}}$  residual dipolar coupling violations of  $>2$  Hz. <sup>c</sup> The experimental dihedral angle restraints comprised 49  $\phi$ , 38  $\psi$ , 16  $\chi_1$ , and 6  $\chi_2$  angle restraints. <sup>d</sup> The coordinate precision is defined as the average atomic rms deviation between the 25 individual SA structures and the mean coordinates. The reported values are for residues 13–50 and 56–79. The backbone value refers to the N, C, and CO atoms.

the observed medium-range NOEs, the magnitudes of  $^3J_{\text{HN}\alpha}$  couplings, and the consensus secondary  $^{13}\text{C}$  chemical shifts of the  $\alpha$  and  $\beta$  carbon atoms. Analysis of the data in terms of secondary structure indicates that all of these parameters are in good agreement in predicting the types of secondary structure expected for a particular amino acid. For instance, regions of the polypeptide that adopt an  $\alpha$ -helical conformation (Pro19–Met23, Val32–Val41, and Thr67–Gly78) exhibit medium-range NOEs, small  $^3J_{\text{HN}\alpha}$  coupling constants, and the appropriate secondary carbon chemical shifts.

The solution structure of MUR-DBD<sup>1–88</sup> was calculated using hybrid distance geometry and simulated annealing methods for residues Lys13–Thr81 and employed 983 experimental restraints: 648 NOE distance restraints, 46 hydrogen bond restraints for 23 hydrogen bonds (which were introduced in the final stages of the refinement), 109 dihedral angle restraints, 43  $^3J_{\text{HN}\alpha}$  coupling constant restraints, 106 carbon chemical shift restraints, and 31  $^1D_{\text{HN}}$  one-bond residual dipolar couplings. The latter values were measured from weakly aligned uniformly  $^{15}\text{N}$ -enriched MUR-DBD<sup>1–88</sup> samples in DMPC/DHPC bicelles and ranged from  $-7.2$  to  $5.7$  Hz. Stereospecific assignments were obtained for 10  $\beta$ -methylene groups and the methyl groups of four of the seven valine residues. Figure 4A shows the ensemble of 25 solution structures calculated from these data, and the complete structural and restraint statistics are presented in Table 1. All structures exhibit good covalent geometry and have no NOE, dihedral angle, or coupling (residual one-bond dipolar or three-bond scalar) violations of greater than  $0.5$   $\text{\AA}$ ,  $5^\circ$ , or  $2$  Hz, respectively. The MUR-DBD<sup>1–88</sup> is structured from residues Lys13–Glu50 and Lys56–Leu79, and the root-mean-square (rms) deviation between the atomic coordinates of the backbone atoms of these residues and the average coordinates is  $0.61 \pm 0.16$   $\text{\AA}$ . Residues Gly51–Gly55 are largely unstructured and flexible as judged by the

small magnitude of their  $^1\text{H}$ – $^{15}\text{N}$  heteronuclear NOEs (the NOEs range from  $0.15$  to  $0.42$ ) (Figure 3A).

The structure of the MUR-DBD consists of a three-stranded antiparallel  $\beta$ -sheet packed against three  $\alpha$ -helices with its secondary structural elements arranged in a B1–H1–T–H2–B2–W–B3–H3 topology (B, strand; H, helix; T, turn; and W, wing) (Figure 4A). At the amino terminus, the first strand of the sheet (B1, residues Trp16 and Cys17) pairs with strand B3 (Ala57–Tyr60). The chain then rises above the sheet, forming helix H1 (Pro19–Met23), which is followed by a seven-residue turn (Ala24–Gly30) that connects helix H1 to helix H2 (Ser31–Val41). The final two strands of the sheet follow helix H2 and are formed by strands B2 (Lys46–Lys49) and B3 (Ala57–Tyr60) which are connected by a flexible loop structure we describe as a “wing” (residues Glu50–Lys56). Strand B3 is the central strand in the sheet and is followed by helix H3 (residues Thr67–Gly78) which packs against helix H2 and strand B1. The hydrophobic core of the protein is comprised of the side chains of residues Cys17 from B1, Pro19 and Ile24 from H1, Ala25 and Met28 from the turn between H1 and H2, Val32, Val35, His36, and Ala39 from H2, Trp44 from the turn connecting H2 and B2, Val58 and Tyr60 from B3, Val62 and Met65 from the turn connecting B3 and H3, and Glu69 and Val73 from H3. The side chain of Glu69 is positioned at the edge of the hydrophobic core and in the majority of structures in the ensemble forms a salt bridge with the guanidino group of Arg38. The solution structures of the MUR- and MUA-DBDs adopt a conserved protein fold, and a detailed comparison will be presented in the Discussion.

*MUR-DBD Complex and Its Investigation by NMR and Surface Plasmon Resonance.* We next set out to study the solution properties of the MUR-DBD–DNA complex. First, we used NMR spectroscopy to map the DNA binding surface on the MUR-DBD protein. The 3D solution structure of the



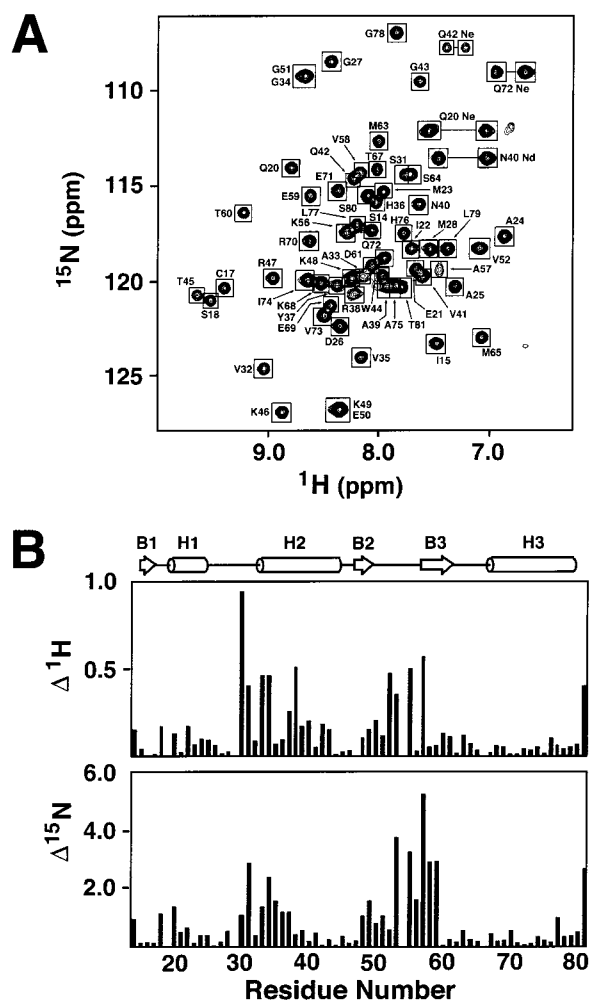


FIGURE 5: NMR spectra of the MUR-DBD<sup>13–81</sup>–DNA complex and a comparison with the free MUR-DBD<sup>1–88</sup>. (A)  $^{15}\text{N}$ – $^1\text{H}$  HSQC spectrum of the 1:1 MUR-DBD<sup>13–81</sup>–DNA complex. The identities of the backbone amide and side chain amine resonances are denoted. In general, the 17.2 kDa protein–DNA complex exhibits excellent spectral resolution and narrow line widths. However, the backbone amide groups of some residues are broadened. Ala33, Arg38, and Ala57 exhibit moderate broadening, and Gly54 and Gly55 are broadened beyond detection in this plot. (B) Comparison of the amide  $^1\text{H}$  and  $^{15}\text{N}$  chemical shifts in the DNA free MUR-DBD<sup>1–88</sup> and the MUR-DBD<sup>13–81</sup>–DNA complex. The observed secondary structural elements in the structure of the MUR-DBD<sup>1–88</sup> are displayed above the figure. The large magnitude changes in the shifts of Thr81 result from the different lengths of the polypeptide fragments (residues 1–88 vs 13–81).

MUR-DBD in the DNA free state revealed that residues Met1–Glu12 and Pro82–Ala88 were unstructured in solution. To simplify and improve the NMR spectra of the complex, we generated a truncated repressor polypeptide which consisted of residues Lys13–Thr81 (MUR-DBD<sup>13–81</sup>). The MUR-DBD<sup>13–81</sup> polypeptide retains affinity for the Mu operator DNA as judged by the fluorescence anisotropy DNA binding assay ( $K_D = 57.1 \pm 9.7 \mu\text{M}$ ). A 1:1 complex of  $^{15}\text{N}$ -labeled MUR-DBD<sup>13–81</sup> with duplex DNA containing the sequence 5'-d(GCTTTTCAGTAATCTG)•5'-d(CAGAT-TACTGAAAAGC) was studied by NMR. Figure 5A shows the quality of the  $^{15}\text{N}$ – $^1\text{H}$  HSQC spectrum of the MUR-DBD<sup>13–81</sup>–DNA complex. The amide  $^1\text{H}$  and  $^{15}\text{N}$  chemical shifts of MUR-DBD<sup>13–81</sup> in the complex were assigned by analysis of the 3D  $^1\text{H}$ ,  $^{15}\text{N}$  NOESY- and TOCSY-HSQC spectra of the complex and by reference to the assignments

of the free protein. In contrast to the NMR spectra of the transposase domain which exhibited severe line broadening (18), the spectrum of the MUR-DBD–DNA complex is well-dispersed and exhibits sharp resonance lines. The difference in the spectral properties of the MUR-DBD–DNA and MUA-DBD–DNA complexes most likely results from the different solution conditions used in the respective NMR studies. Previous studies of the MUA-DBD–DNA complex were performed using high concentrations of salt (250 mM NaCl), while the NMR spectra of the MUR-DBD–DNA complex were recorded at low salt concentrations. In the presence of high concentrations of NaCl, the MUA-DBD exhibits a substantially reduced affinity for its cognate site. The reduced affinity presumably results from the masking of important electrostatic interactions and would be expected to decrease the lifetime of the specific protein–DNA complex, resulting in line broadening. A comparison of the  $^{15}\text{N}$  and  $^1\text{H}$  shifts of the MUR-DBD<sup>1–88</sup> in the absence of DNA to the those of the bound MUR-DBD<sup>13–81</sup> protein is displayed in Figure 5B. The largest deviations in the  $^1\text{H}$  and  $^{15}\text{N}$  shifts occur in the helix–turn–helix unit (H1–T–H2) and the flexible loop between strands B2 and B3. The most drastic  $^1\text{H}$  and  $^{15}\text{N}$  chemical shift changes occur in the amide proton of Gly30 (from turn T) and the amide nitrogen of Ala57 (from the wing connecting strands B2 and B3), respectively.

The binding of the MUR-DBD<sup>1–88</sup> to wild-type and mutated forms of its DNA binding site was studied by surface plasmon resonance (SPR) to probe the nucleotide specificity of the interaction. In this experiment, a DNA duplex containing the protein binding site was immobilized on a surface via a streptavidin–biotin linker and binding of protein was then monitored directly as changes in surface plasmon resonance (a representative sensogram is shown in Figure 6A). The MUR-DBD<sup>1–88</sup> was found to bind to the tethered DNA molecule with an affinity slightly higher than that measured by fluorescence anisotropy ( $K_D = 0.58 \pm 0.05 \mu\text{M}$ ), and the lifetime of the MUR-DBD<sup>1–88</sup>–DNA complex was calculated to be approximately 2 s. As described above, chemical shift perturbation data and modeling studies of the MUR- and MUA-DBDs predicted that binding specificity would be mediated by insertion of helix H2 (the putative recognition helix) into the major groove. Reasoning that these contacts would occur over at most three or four nucleotides and be to the most highly conserved nucleotides within the binding site, we attempted to define the orientation of repressor binding by nucleotide mutagenesis. Two oligonucleotides containing mutations in the binding site were assayed by SPR for binding to the MUR-DBD<sup>1–88</sup>. The first mutant duplex altered three nucleotides at positions 2–4 (see Figure 6B for sequences and numbering). These nucleotides are completely conserved in all nine sites and are expected to be important determinants of binding affinity. A second mutant oligonucleotide altered the identity of base pairs 10–12, which are poorly conserved and presumably not recognized by helix H2 in a sequence-specific manner. Surprisingly, both mutant binding sites exhibited an approximately 10-fold reduction in binding affinity for the MUR-DBD<sup>1–88</sup> relative to the wild-type sequence. Although the orientation of binding cannot be deduced from these results, they do coarsely define the nucleotide dependence of repressor binding. This result suggests that important base-specific

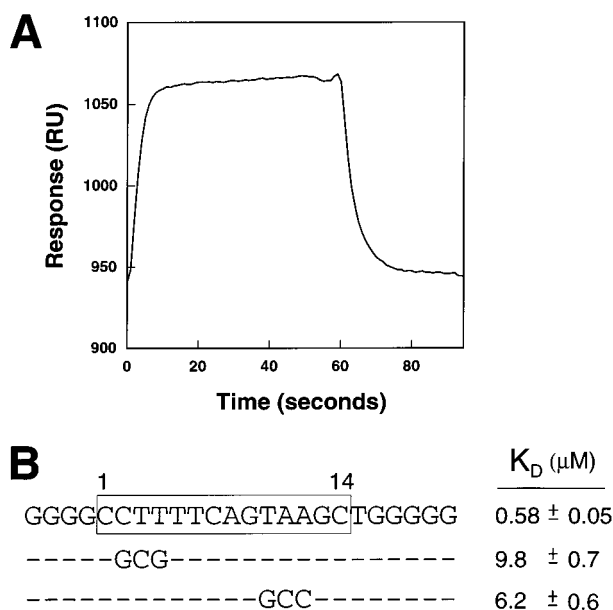


FIGURE 6: Surface plasmon resonance measurements of Mu repressor binding to wild-type and mutant DNA oligonucleotides. (A) Representative sensogram used to calculate the dissociation constants and kinetic parameters of the repressor–DNA complex. The curve shows the association and dissociation of the MUR-DBD<sup>1–88</sup> with a DNA fragment containing its consensus binding site. A plot of response vs time is shown for the MUR-DBD<sup>1–88</sup> binding to a DNA duplex immobilized on the surface of a chip. The DNA duplex contained a single biotin molecule and was tethered to a chip containing conjugated streptavidin. (B) The three oligonucleotides that were studied by surface plasmon resonance and their measured affinity for MUR-DBD<sup>1–88</sup>. The 14 bp consensus sequence for the repressor and transposase proteins is enclosed within a box and numbered. The dashed lines correspond to nucleotides that are conserved between the three oligonucleotides.

contacts are made between the repressor and one or more bases in the triplets formed by base pairs 2–4 and 10–12.

## DISCUSSION

The repressor and MuA transposase proteins compete for the same DNA binding site within the early operator of the Mu bacteriophage. This competition may provide an added level of control over the lytic–lysogenic growth choice, enabling the repressor to inhibit both transcription and transposition reactions. At a molecular level, the transposase versus repressor competition is mediated by homologous DNA-binding domains (42% sequence identity over 64 residues). The results of fluorescence anisotropy measurements with the isolated MUA- and MUR-DBDs indicate that they bind to the operator with weak affinity. In this assay, the total anisotropy is the sum of contributions from both specifically and nonspecifically bound protein–DNA complexes. For high-affinity protein–DNA interactions, nonspecific binding at higher protein concentrations results in a biphasic binding curve and the two modes of interaction can readily be separated. In this case, the specific binding is weak, and it is not possible to distinguish between the two effects. A direct comparison of the magnitudes of the binding constants measured for the isolated domains and the full-length proteins is not appropriate, since they were determined using different techniques in different research laboratories. The results presented here clearly indicate that the MUR- and MUA-DBDs exhibit similar weak binding for their

cognate site, while previous studies have shown that in the context of the full-length proteins, these domains bind to the operator/IAS with substantially different affinities; the transposase binds with a much lower affinity than the repressor [MuA transposase,  $K_D \sim 1$ – $100 \mu$ M (10, 11); Mu repressor protein,  $K_D \sim 3$ – $5$  nM (5)]. The most plausible explanation for this difference is that cooperative interactions between repressor protomers increase binding affinity, a conclusion consistent with the multimeric structure of the full-length repressor and previous footprinting studies (16).

The solution structure of the repressor DNA-binding domain was determined by NMR, enabling a structural comparison with its competitor domain from the MuA transposase. As expected, the overall protein folds of the MUR- and MUA-DBDs are similar. Both proteins define a distinct structural class of the winged helix–turn–helix family in which the secondary structural elements are permuted (45, 46). This class consists of at least six members that share sequence homology, the transposase and repressor proteins from bacteriophage Mu and the corresponding proteins from the related bacteriophages D108 and D3112. Figure 4B shows a comparison of the 3D structures of the MUR- and MUA-DBDs. The structured regions of the two proteins can superimposed with a rms deviation of 1.8 Å for the backbone atoms. The largest structural differences occur in the turn (T) within the helix–turn–helix unit and in the region connecting strand B3 and helix H3. In both of these regions, residues that pack into the core of the protein are different, causing slight changes in the path of the polypeptide backbone. The carboxyl-terminal helix in the repressor domain is also slightly longer. This structural difference occurs because several hydrophilic amino acids at the end of helix H3 in the transposase (Arg64 and Gln65) are replaced by more nonpolar amino acids in the repressor (His76–Leu79). In particular, the side chains of His76 and Leu77 from H3 of the repressor exhibit van der Waals contacts to residues in turn T, stabilizing the end of the helix. Although there are clear differences in both the compositions and packing arrangements of residues involved in the hydrophobic cores of each protein, these differences do not significantly affect thermostability. Temperature denaturation studies using circular dichroism indicate that the midpoint of unfolding in the MUR- and MUA-DBDs is  $52.7 \pm 0.7^\circ$  and  $51.2 \pm 1.1^\circ$  C, respectively (data not shown).

Chemical shift mapping experiments and structural homology to other winged helix–turn–helix proteins suggest a plausible mode of DNA binding. In this model, helix H2 would be inserted into the major groove for base-specific contacts and the adjacent phosphodiester backbone would be contacted by the wing structure which rests in the minor groove. Eleven positions within the MUR- or MUA-DBD fold may interact with DNA according to this model. Given their similar affinities for DNA, it is tempting to speculate that conserved residues at the protein–DNA interface will be important for binding. Three of the 11 residues are completely conserved and include Ala21 and Tyr37 (repressor numbering) within the putative recognition helix H2 and Lys56 within the wing structure. Six of the remaining eight potential contact residues are chemically conserved and include Gln20R and Lys8T, Ser31R and Thr19T, Asn40R and Lys28T, and Val58R and Ile46T, where “R” and “T” represent repressor and transposase amino acids, respectively.



The side chains of three residues are apparently not important for binding affinity; Lys18, Ile24, and Lys29 in the MUA-DBD are replaced in the MUR-DBD by Gly30, His36, and Val41, respectively. Limited mutagenesis of the enhancer domain is consistent with the apparent functional conservation of residues at the protein–DNA interface in the MUR- or MUA-DBD fold (47). The notable exception however is Lys29 in the MUA-DBD which is important for enhancer-dependent transposition in the full-length transposase, but not conserved in the repressor domain (Val41). Limited nucleotide mutagenesis data presented here suggest that site-specific interactions from this protein interface are extensive. In particular, the results of the SPR measurements suggest that the affinity of the complex is dependent on the base identities of nucleotides spanning a region up to nine base pairs in length that includes base pairs 3–12 (Figure 6B).

The structure of the MUR-DBD provides insights into the molecular basis of several temperature sensitive mutations. Four mutations within the repressor DNA-binding domain cause the Mu bacteriophage to undergo lytic development when the temperature is shifted from 30 to 42 °C (5). These mutations include Arg47Gln, Gly43Asp, Ser18Leu, and Met28Ile, all of which presumably affect DNA binding by reducing thermostability. Met28 is located in the turn of the helix–turn–helix unit and contributes its side chain to the hydrophobic core. Substitution with an isoleucine residue may destabilize the structure because its side chain may not pack well in the core of the protein. Ser18 is located just before helix H1, and its side chain hydroxyl may accept a hydrogen bond from the amide proton of Gln20, capping the amino-terminal end of the helix. Modification to a leucine residue would prevent this capping interaction and may lead to reduced thermostability. Gly43 is located in the turn that connects H2 to strand B2. Although the side chain of an aspartic acid is readily accommodated at this position, its backbone adopts a conformation that is somewhat unfavorable for  $\beta$  carbon-containing amino acids ( $\phi = -95.8 \pm 9.6^\circ$  and  $\psi = 63.8 \pm 5.3^\circ$  in the ensemble of 25 structures). The role of Arg47 in thermostability is ambiguous. It may stabilize the base of the protruding wing structure by forming a salt bridge to the side chain of Glu59 and by packing against the indole ring of Trp16 on the protein surface.

The structure of the repressor DNA-binding domain and biochemical studies have enabled a detailed comparison of the transposase and repressor DNA-binding modules. The ultimate goal of these studies is to define the molecular basis of operator site binding by the transposase and repressor proteins in three dimensions. Toward this objective, we have shown here that a minimal polypeptide fragment of the repressor protein consisting of residues Lys13–Thr81 (MUR-DBD<sup>13–81</sup>) is structured in solution and retains affinity for the operator. Furthermore, we have demonstrated that the MUR-DBD<sup>13–81</sup>–DNA complex is in slow exchange on the chemical shift time scale and yields high-quality NMR spectra. At present, we are attempting to determine the structure of the MUR-DBD<sup>13–81</sup>–DNA complex by NMR spectroscopy.

## ACKNOWLEDGMENT

We thank R. Peterson, M. Phillips, and T. Sutherland for technical support. We also thank Dr. G. M. Clore for generously providing his modified version of the program

XPLOR, Dr. K. Mizuuchi for providing the plasmid encoding the Mu repressor protein, and Dr. A. Berk for the use of his fluorescence spectrophotometer.

## REFERENCES

1. Taylor, A. L. (1963) *Proc. Natl. Acad. Sci. U.S.A.* 50, 1043–1051.
2. Chaconas, G., Lavoie, B. D., and Watson, M. A. (1996) *Curr. Biol.* 6, 817–820.
3. Lavoie, B. D., and Chaconas, G. (1996) *Curr. Top. Microbiol. Immunol.* 204, 83–102.
4. Mizuuchi, K. (1992) *Annu. Rev. Biochem.* 61, 1011–1051.
5. Vogel, J. L., Li, Z. J., Howe, M. M., Toussaint, A., and Higgins, N. P. (1991) *J. Bacteriol.* 173, 6568–6577.
6. Krause, H. M., and Higgins, N. P. (1986) *J. Biol. Chem.* 261, 3744–3752.
7. Krause, H. M., Rothwell, M. R., and Higgins, N. P. (1983) *Nucleic Acids Res.* 11, 5843–5495.
8. Krause, H. M., and Higgins, N. P. (1984) *Cold Spring Harbor Symp. Quant. Biol.* 49, 827–834.
9. Goosen, N., van Heuvel, M., Moolenaar, G. F., and van de Putte, P. (1984) *Gene* 32, 419–426.
10. Leung, P. C., Teplow, D. B., and Harshey, R. M. (1989) *Nature* 338, 656–658.
11. Mizuuchi, M., and Mizuuchi, K. (1989) *Cell* 58, 399–408.
12. Gama, M., Toussaint, A., and Higgins, N. P. (1992) *Mol. Microbiol.* 6, 1715–1722.
13. Falconi, M., McGovern, V., Gualerzi, C., Hillyard, D., and Higgins, N. P. (1991) *New Biol.* 3, 615–625.
14. Geuskens, V., Vogel, J. L., Grimaud, R., Desmet, L., Higgins, N. P., and Toussaint, A. (1991) *J. Bacteriol.* 173, 6578–6585.
15. Geuskens, V., Mhammedi-Alaoui, A., Desmet, L., and Toussaint, A. (1992) *EMBO J.* 11, 5121–5127.
16. Rousseau, P., Bétermier, M., Chandler, M., and Alazard, R. (1996) *J. Biol. Chem.* 271, 9739–9745.
17. Harshey, R. M., Getzoff, E. D., Baldwin, D. L., Miller, J. L., and Chaconas, G. (1985) *Proc. Natl. Acad. Sci. U.S.A.* 82, 7676–7680.
18. Clubb, R. T., Omichinski, J. G., Savilahti, H., Mizuuchi, K., Gronenborn, A. M., and Clore, G. M. (1994) *Structure* 2, 1041–1048.
19. Grzesiek, S., and Bax, A. (1992) *J. Magn. Reson.* 96, 432–440.
20. Wittekind, M., and Mueller, L. (1993) *J. Magn. Reson.* 101, 201–205.
21. Grzesiek, S., and Bax, A. (1992) *J. Am. Chem. Soc.* 114, 6291–6293.
22. Grzesiek, S., Anglister, J., and Bax, A. (1993) *J. Magn. Reson., Ser. B* 101, 114–119.
23. Bax, A., Clore, G. M., and Gronenborn, A. M. (1990) *J. Magn. Reson.* 88, 425–431.
24. Ikura, I., Kay, L. E., and Bax, A. (1991) *J. Biomol. NMR* 1, 299–304.
25. Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., and Clore, G. M. (1989) *Biochemistry* 28, 6150–6156.
26. Vuister, G. W., and Bax, A. (1993) *J. Am. Chem. Soc.* 115, 7772–7777.
27. Archer, S. J., Ikura, M., Torchia, D. A., and Bax, A. (1991) *J. Magn. Reson.* 95, 636–641.
28. Clore, G. M., Bax, A., and Gronenborn, G. M. (1991) *J. Biomol. NMR* 1, 13–22.
29. Fesik, S. W., and Zuiderweg, E. R. P. (1988) *J. Magn. Reson.* 78, 588–593.
30. Vuister, G. W., Clore, G. M., Gronenborn, A. M., Powers, R., Garrett, D. S., Tschudin, R., and Bax, A. (1993) *J. Magn. Reson., Ser. B* 101, 210–213.
31. Ottiger, M., and Bax, A. (1998) *J. Biomol. NMR* 12, 361–372.
32. Tjandra, N., and Bax, A. (1997) *Science* 278, 1111–1114.
33. Clore, G. M., Gronenborn, A. M., and Bax, A. (1998) *J. Magn. Reson.* 133, 216–221.
34. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR* 2, 661–665.

35. Delaglio, F. (1995) *J. Biomol. NMR* 6, 277–293.
36. Garrett, D. S., Powers, R., Gronenborn, A. M., and Clore, G. M. (1991) *J. Magn. Reson.* 95, 214–220.
37. Brünger, A. T. (1993) *X-PLOR (Version 3.1): A system for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT.
38. Garrett, D. S. (1994) *J. Magn. Reson., Ser. B* 104, 99–103.
39. Kuszewski, J., Qin, J., Gronenborn, A. M., and Clore, G. M. (1996) *J. Magn. Reson., Ser. B* 106, 92–96.
40. Kuszewski, J., Gronenborn, A. M., and Clore, G. M. (1996) *Protein Sci.* 5, 1067–1080.
41. Nilges, M., Clore, G. M., and Gronenborn, A. M. (1988) *FEBS Lett.* 229, 129–136.
42. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J. Mol. Graphics* 14, 51–55.
43. Lundblad, J. R., Lawrance, M., and Goodman, R. H. (1996) *Mol. Endocrinol.* 10, 607–612.
44. Kay, L. E., Torchia, D. A., and Bax, A. (1989) *Biochemistry* 28, 8972–8979.
45. Brennan, R. G. (1993) *Cell* 74, 773–776.
46. Wintjens, R., and Rooman, M. (1996) *J. Mol. Biol.* 262, 294–313.
47. Clubb, R. T., Mizuuchi, M., Huth, J. R., Omichinski, J. G., Savilahti, H., Mizuuchi, K., Clore, G. M., and Gronenborn, A. M. (1996) *Proc. Natl. Acad. Sci. U.S.A.* 93, 1146–1150.
48. Wishart, D. S., and Sykes, B. D. (1994) *J. Biomol. NMR* 4, 171–180.

BI990530B