Methyl Groups of Thymine Bases Are Important for Nucleic Acid Recognition by DtxR[†,‡]

C. Shuyan Chen,[§] André White,[||] John Love,[⊥] John R. Murphy,[⊥] and Dagmar Ringe*[,||,#]

Department of Biochemistry, Program in Bioorganic Chemistry, Departments of Biochemistry and Chemistry, The Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, Massachusetts 02454-9110, and Department of Medicine, Boston University School of Medicine, Boston, Massachusetts 02118-2392

Received April 24, 2000; Revised Manuscript Received June 27, 2000

ABSTRACT: The expression of diphtheria toxin is controlled by the diphtheria toxin repressor (DtxR). Under conditions of high iron concentration, DtxR binds the *tox* operator to inhibit transcription. To study how DNA binding specificity is achieved by this repressor, we solved the crystal structure of the nickel(II) activated DtxR(C102D) mutant complexed with a 43mer DNA duplex containing the DtxR consensus binding sequence. Structural analysis of this complex and comparison with a previously determined DtxR-(C102D)-Ni(II)-*tox* operator ternary complex revealed unusual van der Waals interactions between Ser37/Pro39 of the repressor helix-turn-helix (HTH) motif and the methyl groups of specific thymine bases in the consensus binding sequence. Gel mobility shift assays utilizing deoxyuridine modified duplex DNA probes proved the importance of these interactions: the four methyl groups shown to interact with Ser37/Pro39 in the crystal structure contribute a total of 3.4 kcal/mol to binding energy. Thus, in addition to making base-specific hydrogen-bonding interactions to the DNA through its Gln43 residue, DtxR also recognizes methyl groups at certain positions in the DNA sequence with its Ser37 and Pro39 side chains, to achieve binding specificity toward its cognate operator sequences.

Diphtheria toxin is the principle virulence factor secreted by toxigenic *Corynebacterium diphtheriae*, the causative agent of clinical diphtheria. Once secreted, the toxin is taken up by the blood stream and is distributed to and causes pathology in essentially all organ systems in sensitive mammals. It is well-known that following binding to its cell surface receptor, a heparin-binding epidermal growth factor-like precursor, the toxin is internalized into the cell by receptor-mediated endocytosis. Upon acidification of an early endosomal compartment, the transmembrane domain of the toxin spontaneously inserts into the membrane, forming a pore or channel, and facilitates the delivery of the catalytic domain into the cytosol. The catalytic domain of the toxin is an NAD^+-dependent adenosine diphosphate ribosyl (ADPR)[1] transferase. Once delivered to the eukaryotic cell cytosol, it specifically inactivates eukaryotic elongation factor 2 (EF-2) by irreversible covalent modification, thereby inhibiting protein synthesis (*1*).

Expression of diphtheria toxin in lysogenic toxigenic *C. diphtheriae* is tightly regulated by the concentration of iron in the external medium, with maximal synthesis occurring when iron becomes rate-limiting for growth. Regulation of *tox* gene expression is mediated by an iron-activated repressor, DtxR, which is encoded on the bacterial genome. In addition to controlling toxin production, DtxR has been shown to bind five iron-responsive promoters *irp1* to *5* and *hmuO* (*2*−*5*) within its own genome. Some of the downstream genes of the above promoters are unknown, but presumed to encode putative ferrisiderophore biosynthesis genes which are involved in acquiring and transporting the ferric ion from the environment into the bacterial cytosol. The downstream gene of promoter *hmuO* encodes a heme oxygenase that is involved in the degradation of heme and the release of iron (*4*). Thus, DtxR serves as a global regulatory protein to coordinately regulate the expression of diphtheria toxin, the high-affinity iron uptake system, and other iron sensitive genes of *C. diphtheriae* in response to changes in environmental iron concentrations.

The operator sequences that are recognized by the activated repressor comprise an interrupted pseudo-palindrome of 27 bps as determined by footprinting assays of the *tox* operator (*6*, *7*). To define the sequence specificity of the repressor, Tao and Murphy selected DNA sequences that bound DtxR by the in vitro genetic method CAST (cyclic amplification of selected targets) (*8*). The selected DNA sequences bind activated DtxR with similar affinity to that of the wild-type *tox* operator. Sequence analysis of this family of selected DNA sequences shows that the consensus sequence for DtxR binding consists of a 19 bp core, which is a palindrome

[†] This work was supported by a grant from the NIH (AI21628 to J.R.M. and D.R.) and, in part, by a grant from the Lucille P. Markey Charitable Trust.

[‡] The coordinates and structure factors for the crystal structure described in this paper are deposited in the Protein Data Bank (www.rcsb.org/pdb/) under accession code 1F5T.

* To whom correspondence should be addressed. Phone: (781) 736-4902. Fax: (781) 736-2405. E-mail: ringe@brandeis.edu.

[§] Department of Biochemistry.

[||] Department of Medicine, Boston University School of Medicine.

[⊥] The Rosenstiel Basic Medical Sciences Research Center.

[#] Departments of Biochemistry and Chemistry.

[1] Abbreviations: NAD^+, nicotinamide adenine dinucleotide; ADPR, adenosine diphosphate ribosyl; DtxR, diphtheria toxin repressor; bp, base pair; bps, base pairs; CAST, cyclic amplification of selected targets; SH3, src-homology domain 3; HTH, helix-turn-helix.

10.1021/bi0009284 CCC: $19.00
Published on Web 08/04/2000

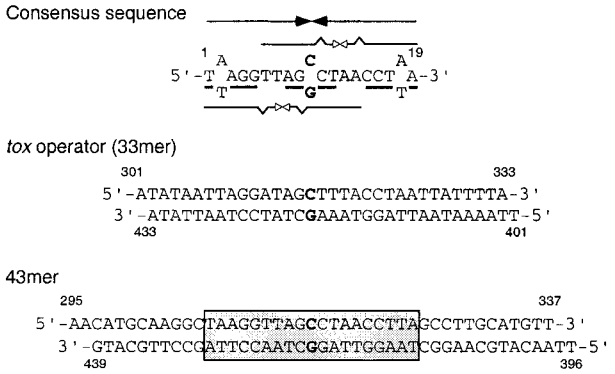

FIGURE 1: DtxR-recognized consensus sequence. (Top) Only the sense strand is shown for the consensus sequence core, numbered from 1 to 19. The DtxR consensus sequence is an interrupted palindrome of 19 bps in which a C·G forms the central base pair. This palindrome has symmetrical features that accommodate the binding of four protein subunits in the form of two dimers: the sequence includes two overlapping imperfect palindromes (14 bps, short arrows) that have conserved bases at symmetrical positions. This sequence will ultimately make a dyad symmetry of protein interaction both within the dimer and across the two dimers. (Middle) For comparison, the 33mer *tox* operator sequence, used in the DtxR(C102D)-Ni(II)-*tox* operator complex structure, is also shown. It contains an interrupted pseudo-palindrome. (Bottom) The structure reported here contains the 43mer shown with the 19 bp consensus core (shaded box).

separated by a single base pair (Figure 1). Analysis of 21 unique in vitro affinity-selected DtxR-responsive elements revealed strong base preferences at certain positions (underlined). Positions 15, 16, 17 and 19 were almost invariably found to be C, C, T, and A, which by palindromic symmetry, correspond to T, A, G, and G at positions 1, 3, 4, and 5. Positions 8 and 12 have a preference for A and T, respectively, whereas G and C are preferred at positions 9 and 11. The rest of the sequence shows less conservation. The DtxR consensus sequence was also deduced from in vivo studies of iron-responsive promoters from *C. diphtheriae* that bind DtxR (*2*). It is remarkable that the consensus sequence determined by in vitro selection is identical to that determined by sequence analysis of DtxR-sensitive operators.

Crystallographic studies have enabled us to understand the structural basis of DtxR metal ion induced activation and DNA binding. In solution in the absence of a transition metal ion, the apo-repressor exists mainly in the monomeric state in weak equilibrium with a dimer form. Upon addition of activating metal ions, a dimeric complex is stabilized (*9*). Each DtxR monomer of 226 amino acid residues consists of an N-terminal DNA binding domain, a dimerization domain, and a highly flexible C-terminal domain (*10*, *11*). A number of X-ray crystal structures of wild-type DtxR and a biologically active mutant have been determined in the presence or absence of transition metal ions (*10*−*15*). Both the apo- and holo-repressors adopt the dimeric forms in these structures. The C-terminal domain of approximately 90 residues is either disordered (*10*, *11*, *13*) or highly flexible (*12*, *14*); in the latter structures it adopts an SH3-like conformation. Recently, Wang et al. (1999) (*16*) solved the structure of the isolated C-terminal domain by NMR and confirmed its SH3-like conformation. The N-terminal and dimerization domains provide residues for metal ion binding, and two metal-ion binding sites have been identified from the holo-repressor structures.

One of the metal ion binding sites, the ancillary metal ion-binding site (site 1), is observed in every holo-repressor structure determined so far. This site is composed of the side chains of His79, Glu83, His98, and a water molecule, forming a near tetrahedral geometry. The primary metal ion binding site (site 2) is not always observed, and when it is observed, the metal ion site is poorly occupied (*11*, *14*). The apparent reason for this anomaly is the oxidation of the sulfur atom in Cys102 that seems to occur readily. Analysis of point mutants at residue 102 had shown that only Asp can substitute for Cys to generate an active DtxR protein (*17*). The crystal structure of the mutant repressor DtxR(C102D) in the presence of nickel ion reveals that the primary site consists of the side chains of Met10, Asp102, Glu105, and His106, the mainchain oxygen of Asp102 and a water molecule, forming an octahedral coordination (*13*). Site-directed mutagenesis of DtxR has indicated that only the primary site is essential for repressor activity (*6*, *13*, *17*) as mutations of individual residues in the primary site result in the loss of repressor function. Upon transition metal ion binding, DtxR undergoes conformational changes that are essential in allowing the repressor to interact optimally with two adjacent major grooves of the operator DNA (*18*). These changes include a rearrangement of the N-terminal helix that is stabilized by a second sphere ligand to the activating metal ion and a shift of the N-terminal subdomain relative to the C-terminal subdomain in a caliper-like movement that moves the helix C, one from each monomer closer together by 2 Å.

The three-dimensional crystal structure of the complex between Ni(II)-bound DtxR(C102D) and a 33 base pair DNA segment containing the *tox* operator sequence (*18*) shows that two DtxR dimers bind to opposite sides of the operator at sites 5 bps apart, and the two dimers do not seem to interact directly with each other. Analysis of this ternary complex shows that DtxR binds to the major groove of the nucleic acid with its helix-turn-helix (HTH) motif, which is formed by the region spanning α-helices B and C (residues 27−50). The majority of the interactions between the operator and the repressor are nonspecific phosphate backbone−amino acid side chain interactions involving helices B and C (the first and second helices in the HTH motif, respectively) as well as one side chain of helix A and residues joining β-strands 1 and 2 between α-helices C and D.

Pohl et al. (1999) recently have reported a crystal structure of the Co(II)-activated DtxR-DNA complex. The 21 bp DNA duplex, which includes two T−C mismatches, was based on the consensus operator sequence. This structure also reveals that two DtxR dimers bind to the DNA. The authors reported that the SH3-like C-terminal domain contains two residues (Glu170 and Gln173) that coordinate to the metal ion at the ancillary site via their side chains.

The three-dimensional structures of both complexes show that only one amino acid residue, Gln43 of the second helix of the HTH motif from each subunit, makes specific hydrogen-bonding interactions with the nucleic acid bases. In the case of the DtxR(C102D)-Ni(II)-*tox* operator complex, the Gln43 side chains from subunits b and d interact with C321 (*C15*)[2] and C424 (*C15′*), respectively. The Gln43 side

[2] The numbering according to the 19bp consensus core will be given in parentheses for reference.

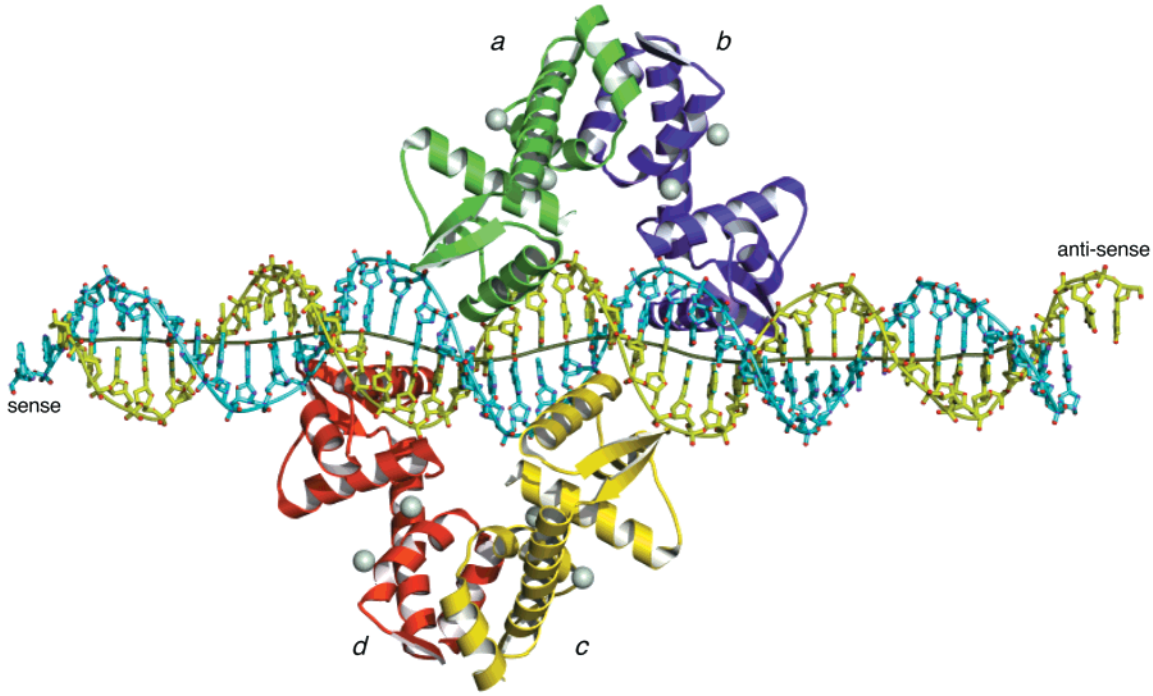

FIGURE 2: Overall structure of the Ni(II)-DtxR(C102D)-43mer DNA complex. The four protein subunits are colored green (*a*), purple (*b*), yellow (*c*), and red (*d*). The sense strand (bases 295−337) of the 43mer DNA is cyan, and the anti-sense strand (bases 396−438) is yellow. The axis of the DNA is shown in dark green. This figure and the following structure figures are prepared with MOLSCRIPT (*52*), POVSCRIPT (D. Peisach and E. Peisach, personal communication), and rendered with the public domain program POVRAY.

chain from subunit *a* interacts with two adjacent bases, G419 (*G10′*) and C420 (*C11′*), while Gln43 from subunit *c* interacts with C316 (*C10*) and T317 (*T11*) (*18*). In the structure reported by Pohl et al. (1999), the side chains of Gln43 in monomers 1 and 3 reach the C•G central base pair and possibly also the adjacent base pairs. Gln43 of monomers 2 and 4 approach two cytosine bases corresponding to the ones in the DtxR(C102D)-Ni(II)-*tox* operator structure.

The base-specific interaction between the single residue Gln43 of the repressor and the operator in the context of two DtxR dimers binding to DNA cannot account for all the specificity observed among DtxR-recognized sequences. The preferred interactions of Gln43 of each subunit for a C or a G base can only explain at most the presence of five conserved bases in the operator sequence. The in vitro selection experiment demonstrated that at least 12 bases are almost invariant among the 19 bp consensus core. Thus, other factors must contribute to recognition. The precise mechanism that controls the binding of DtxR to its cognitive sequences is not well understood. One possibility is that the DNA sequence itself may determine the conformation observed in the complex, and those nucleic acids with optimal conformations may become the target of DtxR binding. On the other hand, it could be that, in addition to the Gln43 interactions, DtxR utilizes other side chains for specific recognition of the operator. The subtle disruption of these interactions would reduce the strength of binding of DtxR to DNA. To study the DNA binding specificity of DtxR(C102D), we cocrystallized it with its experimentally determined consensus DNA sequence (*8*). The structure shows that the protein still utilizes extensive nonspecific phosphate backbone interactions in addition to limited base-specific interactions: for each subunit, Gln43 recognizes a G and/or a C, and two amino acid side chains, Ser37 and Pro39, make van der Waals interactions with an invariant thymine base. Unexpectedly, methyl groups on four thymines are shown by gel mobility-shift assay to be important in determining tight binding—with one methyl group contributing ~0.8 kcal/mol (average) to binding. The four thymine bases recognized by the Ser37/Pro39 pair together with the G or C base recognized by Gln43 thus seem to determine the specificity of binding. A comparison to other iron-regulated promoter sequences suggests these nucleotides are important in DtxR-operator recognition.

MATERIALS AND METHODS

Materials. Recombinant DtxR(C102D) was overexpressed and purified as described (*17*). The DNA oligonucleotides used for crystallization were synthesized at Boston University Medical Center and their sequences are 5′-AAC ATG CAA GGC TAA GGT TAG CCT AAC CTT AGC CTT GCA TGT T-3′ and its complement 5′-TTA ACA TGC AAG GCT AAG GTT AGG CTA ACC TTA GCC TTG CAT G-3′. The two strands were annealed to produce a 43mer DNA duplex containing the consensus sequence for DtxR binding. This duplex DNA also contains 2 base overhangs on the 5′ ends of each strand (Figure 1).

The oligonucleotides used in gel shift assays were purchased from IDT (Coralville, IA) and purified by polyacrylamide gel electrophoresis. Their sequences are (numbered according to the consensus core) as follows: sense, 5′-GCG GAA TTC GCT AAG GTT AGC CTA ACC TTA GCG GAT CCG C-3′; antisense, 5′-GGC GGA TCC GCT AAG GTT AGG CTA ACC TTA GCG AAT TCC G-3′; sense *T12dU*, 5′-GCG GAA TTC GCT AAG GTT AGC CUA ACC TTA GCG GAT CCG C-3′; antisense *T17′dU*, 5′-GGC GGA TCC GCT AAG GTT AGG CTA ACC UTA GCG AAT TCC G-3′; sense *T12dU:T17dU*, 5′-GCG GAA

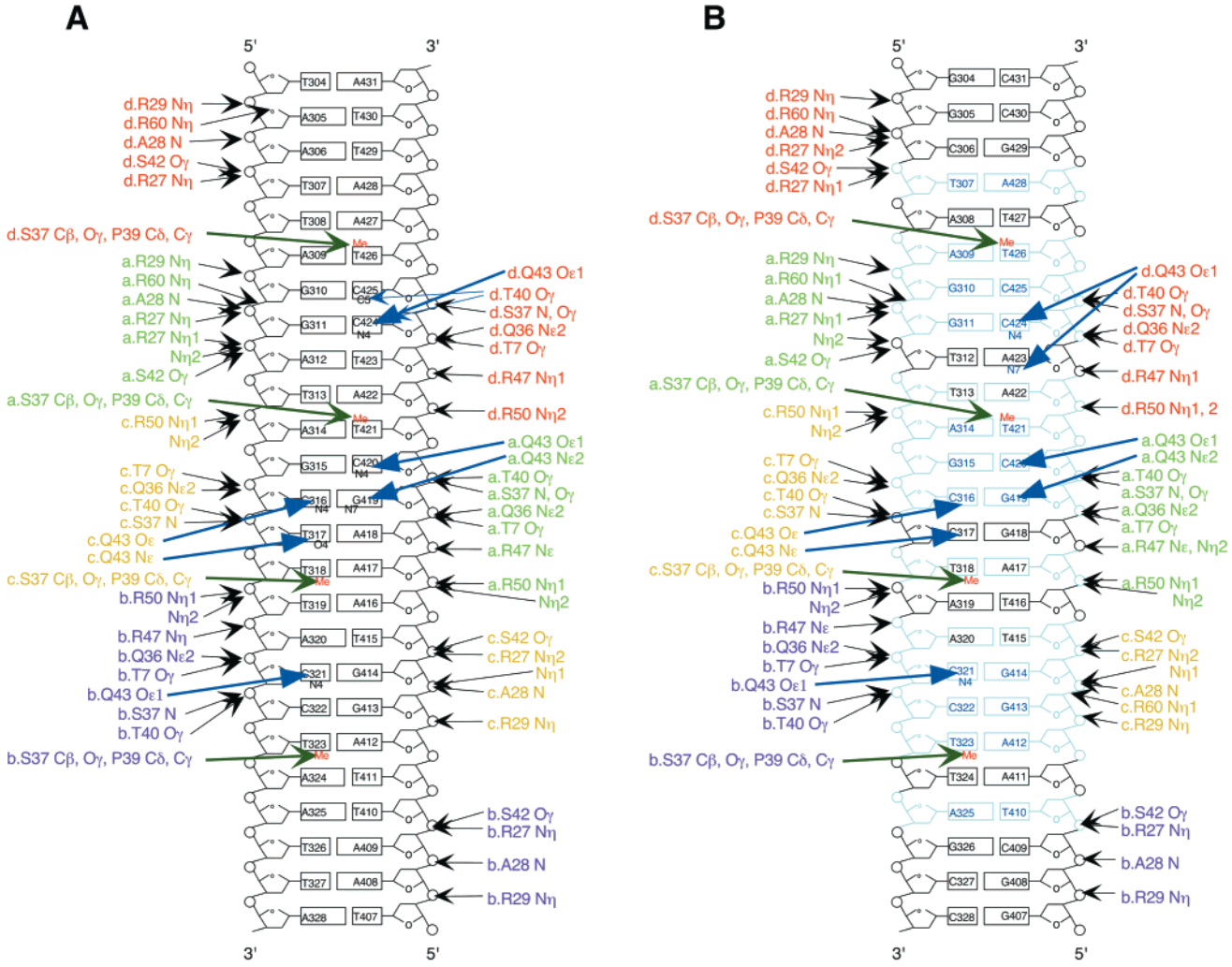

FIGURE 3: Protein−DNA interactions. Schematic representations of the protein−DNA contacts in (A**)** the DtxR(C102D)-Ni(II)-*tox* operator complex and (B**)** the DtxR(C102D)-Ni(II)-43mer consensus DNA complex. Only the section of DNA that interacts with the protein is shown. Protein-DNA backbone hydrogen bonds are shown as short black arrows. Base specific hydrogen bonds are shown in thick blue arrows. van der Waals interactions with specific bases are shown as thick green arrows. The methyl groups involved in van der Waals interactions are shown in red. Previously recognized invariant bases are shown in blue.

TTC GCT AAG GTT AGC CUA ACC UTA GCG GAT CCG C-3′; antisense *T12′dU:T17′dU*, 5′-GGC GGA TCC GCT AAG GTT AGG CUA ACC UTA GCG AAT TCC G-3′.

The DNA oligonucleotides (sense strand) were 5′ end labeled by T4 PNK in the presence of γ-^{32}P ATP, followed by annealing with unlabeled antisense strand. Six dsDNA probes were generated to harbor from 0 to 4 mutations in the four thymines (Figure 6).

Crystallization. Single crystals of the ternary Ni(II)-DtxR-(C102D)-consensus DNA complex were obtained by the hanging-drop vapor diffusion method at room temperature. DtxR(C102D) at a concentration of 16 mg/mL was activated by addition of 1.96 mM $NiCl_2$. The activated repressor was mixed with its consensus DNA (*8*) at a molar ratio of protein: DNA of 1:1.32. The protein−DNA mixture was at a concentration of 0.5 mM protein, 0.66 mM DNA, in 27 mM Tris pH 7.5, 320 mM NaCl, 0.74 mM DTT, 4.9 mM $MgCl_2$, 49 mM KCl, and 1.48 mM $NiCl_2$. A 2 μL drop of the protein−DNA solution was combined with a 1 μL reservoir solution containing 100 mM sodium cacodylate buffer, pH 7.3, 10 mM $MgCl_2$, and 2 mM spermine and equilibrated against 0.7 mL of the reservoir solution. Tetragonal rod-shaped crystals grew in 2 weeks. Crystals were transferred gradually into mother liquor with 30% glycerol as the cryoprotectant before freezing in liquid nitrogen.

Data Collection. X-ray diffraction data were measured at 110 K at beamline A1 of CHESS using the Quantum ADSC CCD detector and a wavelength of 0.92 Å. A total of 132 oscillation photographs collected with an oscillation range of 1° were obtained from one mono crystal. The data were integrated, scaled and merged with programs DENZO and SCALEPACK (*19*). The crystal has the symmetry of the tetragonal space group $P4_1$, with unit cell dimensions $a = b = 116.2$ Å, $c = 142.9$ Å.

Structure Determination. The structure was solved by molecular replacement using AMORE (*20*) and the DtxR-(C102D)-Ni(II)-*tox* operator complex (*18*) as a search model. A rotation function using all data between 10 and 4 Å yielded two top solutions with correlation coefficients of 37.2%. The translation function resulted in two related top solutions with correlation coefficients of 56.4 and 56.2% and R_{factor}s of 40.5 and 40.6%. The first solution resulted in a correlation coefficient of 60.3% and an R_{factor} of 38.6% after rigid body refinement. From this point on, all subsequent calculations were performed with CNS (*21*)]. The data were divided into

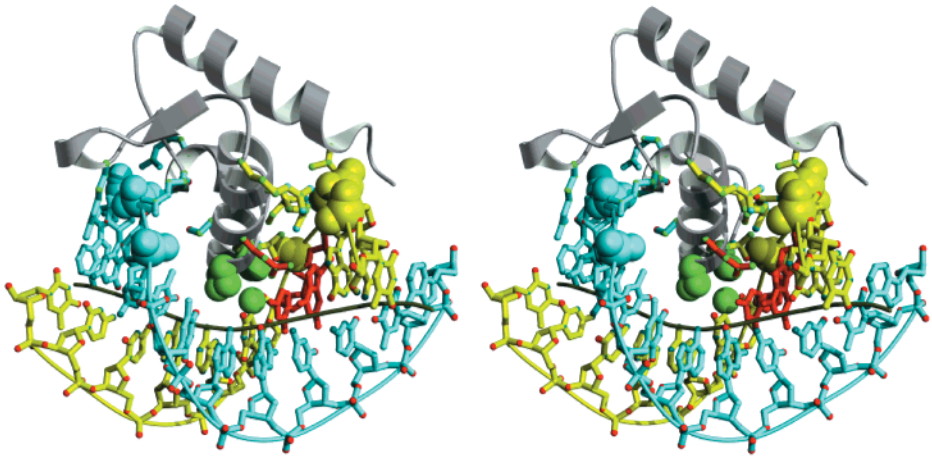

FIGURE 4: Stereoview of the interactions between the DNA and the amino terminal 60 residues of subunit *a*. The phosphates of the DNA backbone that interact with the protein are shown as space-filled balls labeled with the color of the DNA strand to which they belong (cyan for the sense strand, yellow for anti-sense). Amino acid side chains making hydrogen-bonding interactions to DNA phosphates are shown with the color corresponding to that of the phosphate with which they interact. The Gln43 side chain and the two bases with which it interacts are labeled in red. Ser37/Pro39 side chains and the methyl group with which they interact are space-filled (green).

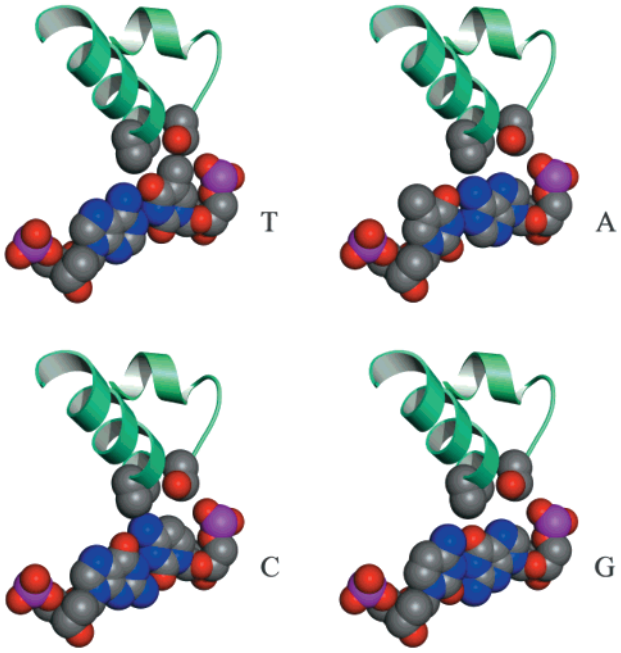

FIGURE 5: Ser37/Pro39 pair of the helix-turn-helix motif of subunit *a* makes van der Waals interactions with the methyl group of thymine 421 (*T12′*). This interaction was modeled to explore the nature of this interaction with bases that lack a methyl in the C5 position. Models of base substitutions were generated with the graphics program O assuming the rest of the structure remains unchanged. In all cases the fit between the amino acid side chains and the DNA is not as optimal as that observed for thymine.

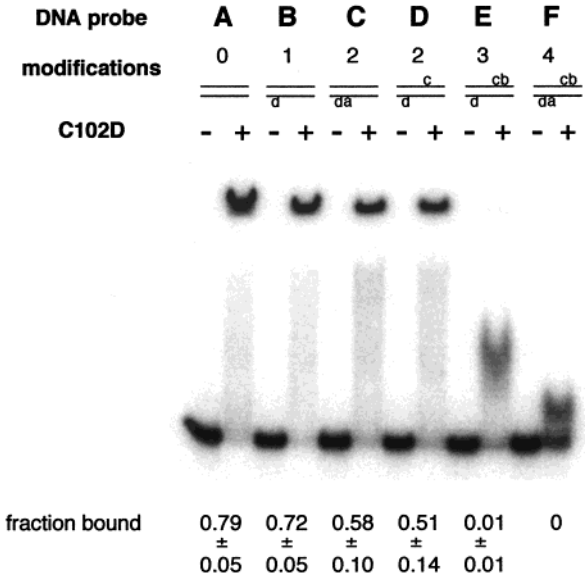

FIGURE 6: Representative gel mobility-shift assay demonstrating the importance of the thymine methyl group in repressor binding. DNA probes containing the consensus sequence were modified by substituting dU for T at those positions where the Ser37/Pro39 interaction takes place. The modifications on the probes are labeled with letters (a, b, c, or d) indicating which DtxR subunit(s) is(are) affected. For each probe (A to F), 20 fmol of probe was mixed with buffer (−) or 50 ng of DtxR(C102D) (+) in 16 μL as previously described (26). This experiment was repeated twice more, and analyzed as described in the materials and methods. The average fractions of DNA probe bound are given along with their standard deviations. The errors given for each probe were calculated from a total of three such experiments and are shown as error bars in Figure 7.

a working set (95%) and a test set (5%) which was used to monitor R_{free} (*22*). After the initial rigid body refinement of each protein subunit and DNA model using data from 50 to 3 Å the R_{factor} dropped from 43.2 to 40.1% (R_{free} from 47.3 to 44.7%). Rigid body refinement was also carried out using only the four protein subunits. The R_{factor} dropped from 42.3 to 39.1% (R_{free} from 42.2 to 39.3%). Difference Fourier unweighted $F_o - F_c$ and σ_A-weighted (*23*) $2F_o - F_c$ electron density maps were calculated. These maps clearly show the DNA double-helix. Subsequently, the correct DNA sequence was built into the electron density using the molecular graphics program O (*24*). Positional refinement and B group refinement were carried out. The above types of electron density maps were calculated at each step. Manual model-building and refinement were repeated for several cycles. A composite simulated annealing omit map was calculated for the final structure (Figure 2).

The model has a crystallographic R_{factor} of 0.249 and an R_{free} of 0.276 for all data from 10 to 3 Å resolution. Grouped temperature factor values were refined, giving an average of 48.5 Å^2 for the complex (Table 1). The quality of the final model was monitored by PROCHECK (*25*). The

Table 1: Statistics for DtxR(C102D)-Ni(II)-consensus Sequence Complex Data Collection and Refinement

data collection	
temperature	110 K
space group	$P4_1$
unit cell	
$a = b$ (Å)	116.2
c (Å)	142.9
resolution range (Å)	50.0−3.0
total no. of reflections	520 936
no. of unique reflections	37 719
overall $\langle I/\sigma(I)\rangle$	10.5
completeness of data (%; 3.0−3.11 Å shell)	99.2 (97.0)
R-merge[a](%)	14.9
refinement	
resolution range (Å)	10.0−3.0
reflections used (working/free)	33 634/1781
temperature factor model	restrained group
R_{factor}[b]/R_{free} (%)	24.9/27.6
model used in refinement	
no. of atoms in protein	3852
no. of atoms in DNA	1757
no. of waters	4
no. of Ni(II) ions	8
total non-hydrogen atoms	5621
deviations from ideal geometry[c]	
rms dev (bond lengths, Å)	0.007
rms dev (bond angles, deg)	1.2
rms dev (dihedral angles, deg)	19.8
rms dev (improper angles, deg)	1.1

[a] $R_{merge} = \sum|I_{obs} - I_{avg}|/\sum I_{avg}$, over all symmetry related observations. [b] $R_{factor} = \sum|F_o - F_c|/\sum F_o$, over all reflections. [c] Engh and Huber parameters were used in refinement (*51*).

Ramachandran angle distribution indicates 90.2% of most favorable and 8.9% of allowed conformations, except for Val 5, which is involved in a water mediated interaction to the metal ion (*18*).

Gel Mobility-Shift Assay. The standard gel shift assay was done in 16 μL as previously described (*26*), each containing 20 fmol of probe and 50 ng of DtxR(C102D). The amounts of DtxR(C102D) were varied in the assays to determine binding affinity. The reaction was carried out at room temperature for 30 min and the mixture was then loaded onto a 6% acrylamide:bis-acrylamide (19:1) nondenaturing gel containing 20 mM BisTris borate buffer (pH 7.3). Electrophoresis was conducted at 200 V for 140 min. After drying, the gel was applied to a Molecular Dynamics phosphorimager. The DNA−protein complexes were detected by ^{32}P autoradiography. The signals were analyzed by the Image Quant software (Molecular Dynamics, Sunnyvale, CA) with the local background redefined as zero. The mobility shift for each protein concentration was determined three times and the results averaged. Data were presented by plotting the fraction of DNA bound versus the log of C102D dimer concentration using KaleidaGraph. The binding model used was

$$2\ \text{dimer} + \text{DNA} \leftrightarrow \text{dimer}_2\cdot\text{DNA}$$

The binding data were fitted to the equation

$$Y = 1/(1 + K_d/U^2)$$

where Y is the fraction of bound, ^{32}P-labeled DNA probe, K_d is the apparent equilibrium constant for dissociation of the protein−DNA complex into two free dimers and DNA, U is the free concentration of DtxR(C102D) dimer. Since in most experiments the protein concentration was about 100-fold greater than that of the DNA, the input protein concentration was taken as the free protein concentration.

RESULTS AND DISCUSSION

The active C102D mutant of DtxR was cocrystallized with the 43mer DNA duplex shown in Figure 1, that contains the 19 bp DtxR consensus core (shaded box). The X-ray crystal structure of this complex has been determined at 3 Å (Figure 2). The numbering of the DNA bases in the structure is from 295 to 337 for the sense strand (307 = no. 1 of consensus core), and from 396 to 438 for the anti-sense strand (410 = no. 1 of consensus core). These numbers directly correspond to those in the DtxR(C102D)-Ni(II)-*tox* operator complex (*18*). (When necessary, the numbering of the 19 bp consensus core will also be given.) Parts of the structure of this complex are quite similar to those of the DtxR(C102D)-Ni(II)-*tox* operator complex (*18*) and the DtxR-Co(II)-21mer DNA complex (*27*): two dimers bind the DNA 5 bps apart on almost opposite sides of the nucleic acid duplex, with no obvious dimer-dimer interactions. The dimers of the repressor bind to the consensus sequence on opposite sides of the DNA in such a way that the *a* and *c* subunits occur almost opposite each other. Analysis of the DNA sequence shows that it is composed of two overlapping imperfect palindromes within the interrupted palindrome consensus sequence (Figure 1). The two imperfect palindromes of 14 bps overlap with each other by 9 bps. This is correlated with the fact that the two dimers bind to DNA with an overlap of 5 bps. In this complex structure, each DtxR(C102D) monomer binds two nickel ions as seen in the DtxR(C102D)-Ni(II)-*tox* operator complex structure. However, a number of differences exist between the structure reported here and the structure of the DtxR-Co(II)-21mer DNA complex (*27*). The most obvious is that the C-terminal domain is still disordered in our structure. The absence of electron density for the C-terminal domain might be accounted for by a difference in space group: $P4_222$ (*27*) versus $P4_1$ (*18*; this report). Nevertheless, our new structure helps to better elucidate the determinants of DtxR specificity of binding and the conformation of its target nucleic acid in the macromolecular crystalline environment.

DNA Conformation. The conformation of the DNA in the complex was analyzed with the program CURVES52 (*28*, *29*) using the global parameters. The average helical twist of 32.6° (11 bps/turn) is slightly different from that of a typical canonical B-DNA (10 bps/turn). The DNA is in a nonlinear conformation (Figure 2), as shown by the curves in the helical axis: the DNA bends toward the helices C bound in its major groove, yielding four bends in the DNA. It is not clear whether the bends are present in the DNA alone, to which the protein adapts itself, or the bends are induced by the interaction with the protein. However, we have shown with two different DNA sequences that the same bends exist. This is consistent with the induction of DNA curvature with protein interaction. The central region of the DNA has an average helical twist of 31.8° (11.3 bps/turn). This slight unwinding also correlates with protein binding. Both ends of the double stranded nucleic acid are relatively linear (see Figure 2). Although these ends show higher mobility in the structure of the complex, as suggested by

their higher average B_{factor}s when compared with the remaining double stranded DNA, they are clearly visible in the electron density map. The overall shape of the part of the DNA interacting with DtxR is quite similar to that in the DtxR(C102D)-Ni(II)-*tox* operator complex (*18*).

DNA−DNA Crystal Contacts. The DNA helical axis lies roughly in the crystallographic *c*-direction. This direction is corroborated by scattering features observed near 3.4 Å resolution in the diffraction pattern that are reminiscent of the fiber diffraction of nucleic acids (data not shown). The length of the DNA strands allows the overhanging ends (see Figure 1) to interact with the neighboring DNA strands from the next asymmetric unit along the *c*-axis. Thus, the DNA in the crystal forms a contiguous double helix. The DNA was designed to form such a contiguous double helix in order to increase the interactions between molecules in the crystal, thereby producing better diffracting crystals. Unfortunately, the resolution limit was not improved, possibly due to high solvent content or the disordered C terminal domain.

HTH Motif Recognizes the DNA. The HTH motif (residues 27−50; helices B and C) is responsible for most of the DNA binding and recognition. Three residues, Gln43, Ser37, and Pro39, of the HTH motif make base specific interactions in the major groove. In addition, there are extensive nonspecific contacts between the phosphate groups of both DNA strands and each protein subunit. Figure 3 shows the interactions between protein and DNA in both the 33-mer *tox* operator complex structure (A) and the 43-mer consensus sequence structure (B). Despite the nucleotide differences between these sequences, such as G + C content, the 43-mer complex shows many similarities to the DtxR(C102D)-Ni(II)33-mer *tox* operator complex. Amino acid residues Arg27, Arg29, Ser37, Thr40, Ser42, Arg47, and Arg50 make hydrogen-bonding interactions with phosphate groups via their side chains, and the backbone amides of Ala28 and Ser37 form hydrogen bonds to phosphate groups (Figure 4). The second subunit of a dimer makes similar interactions with the DNA phosphate backbone and the interactions with the two subunits are related to each other by dyad symmetry across the smaller imperfect palindrome within the longer palindrome sequence. The two dimers bind 5 bps away from each other; thus, the interactions cover 23 bps, with the center located on the central C·G base pair.

In this structure, Gln43 still makes the base-specific interactions that had been observed earlier. In the *tox* operator complex, the two Gln43 residues from subunits *a* and *c* interact with the center of the palindrome as well as an adjacent base [G419C420 (*G10′C11′*) for *a* and C316T317 (*C10T11*) for *c*, respectively] (*18*). In the consensus sequence complex, the electron densities for the Gln43 side chains in subunits *a* and *c* are not as strong as those in the *tox* operator complex, indicating that this residue can adopt multiple conformations as opposed to the defined conformation in the *tox* operator complex. These two Gln43 residues prefer to interact with the central C·G base pair. The electron densities for the side chains of Gln43 in subunits *b* and *d* are well defined and interact with single bases: C321 (*C15*) for *b* and C424 (*C15′*) for *d*, observed similarly in the *tox* operator complex. In both complex structures, Gln43 is the only amino acid residue directly interacting with DNA bases. The importance of this residue to binding specificity is also demonstrated by the fact that DtxR carrying a Gln43 to Glu mutation fails to bind the *tox* operator in gel mobility-shift assays (*10*). The contribution of Gln43 to specificity is also demonstrated by a corynebacteriophage β mutant β^{201} which carries a single G → A mutation at position 311 (*G5A*) (*30*). This strain produces toxin even under high iron concentrations. This lack of activity can be explained by the crystal structures since C424 (*C15′*), the nucleotide that base pairs with G311 (*G5*), is shown in both structures to interact with Gln43 from subunit *d*. Thus activated DtxR fails to bind this mutated operator.

Table 2: Distances between Selected Atoms in the 43mer Consensus (*tox* Operator) Complex[a]

	(a) Thy421 C5A	(b) Thy323 C5A	(c) Thy318 C5A	(d) Thy426 C5A
Ser37 Cβ	3.23 (3.94)	3.74 (3.73)	3.45 (3.87)	3.56 (3.54)
Ser37 Oγ	3.24 (3.76)	3.31 (3.47)	2.99 (3.67)	3.57 (3.37)
Pro39 Cδ	4.14 (3.99)	3.88 (4.19)	4.20 (3.92)	3.87 (4.09)
Pro39 Cγ	4.09 (3.73)	3.39 (3.42)	4.29 (3.58)	3.77 (4.08)

[a] The methyl group of thymine at positions 318, 323, 421, and 426 (12, 17, 12′, and 17′ in the consensus core) is involved in van der Waals interactions with Ser37 and Pro39 in DtxR(C102D). The distances between the methyl group and the Cβ, Oγ atoms in Ser37, the Cδ, Cγ atoms in Pro39 from the each subunit are listed for the complex structure reported here. The corresponding distances for the DtxR(C102D)-Ni(II)-*tox* operator complex are listed in parentheses.

Ser37/Pro39 Make van der Waals Interactions with the Methyl Group of a Thymine. For each of the four DtxR subunits, Ser37 and Pro39 are shown to participate in van der Waals interactions with an invariant thymine base (Figure 4). The distances between the methyl group and the atoms in Ser37 and Pro39 are listed in Table 2 for each subunit. Pro39 is located at the N-terminal end of helix C, the recognition helix, and Ser37 is on the preceding turn. Thus, DtxR serves as an example for DNA recognition by residues on the turn in a HTH motif. This study corroborates genetic studies on DtxR in which mutation of Pro39 to Lys abolished its binding to the *tox* operator (*31*). For instance, Pro39 of subunit *a* docks perfectly into the major groove between the two bases T421A422 (*T12′A13′*) of the antisense strand, and forms extensive interactions with nucleotide functional groups, especially the methyl group of thymine 421 (*T12′*), while the side chain of Ser37 is close enough to make van der Waals interactions with the same thymine methyl group. Of the two adjacent bases, T421(*T12′*) is shown to be more important than A422 (*A13′*) by in vitro selection experiments (*8*): at position 13′ (i.e., position 7 on the sense strand), other bases (A, C, or G) are tolerated for DtxR binding. Basepair substitutions were modeled using the graphics program O, assuming other parts of the structure remain unchanged (Figure 5). Shown in Figure 5, base substitution either results in interactions that are unfavorably too close between the DNA and the repressor or in fewer interactions. This hydrophobic interaction between S37/P39 and the thymine methyl group is seen in all subunits of the complex reported here, as well as in the *tox* operator complex.

In the Ser37/Pro39:thymine base interaction, the methyl group of T is of particular importance. It has been shown that the methyl group can serve as a functional group for DNA sequence recognition by a number of proteins. In these structures, the methyl group of a thymine was shown to be within van der Waals interaction distance to Thr [trp-repressor (*32*)], Thr, Ala, Asn [cro (*33*)], Glu, Gly [λ

repressor (*33*)], Thr [plasmid-encoded transcriptional repressor CopG (*34*)], Ile [HoxB1−Pbx1 heterodimer (*35*)], Ile [ultrabithorax-extradenticle homeodomain (*36*)], or Arg residues [Pax6 paired domain (*37*)]. In a number of protein−DNA complexes, the C5-methyl group of thymine was also suggested to make relatively weak CH···O interactions with Asp, Asn, Glu, Gln, Ser, and Thr, contributing to the specificity of recognition (*38*). Since the hydroxyl group of Ser37 is in close contact with the methyl group of the thymine, it might also be involved in such an interaction. This is the first time that a van der Waals interaction between the Ser37/Pro39 pair and the methyl group of a thymine base has been observed. A simple prediction from the Ser37/Pro39:thymine methyl group interaction is that the removal of the methyl group would weaken the DtxR-DNA interaction.

Ser37/Pro39:Methyl Group Interaction Contributes to Binding Energy. To test this hypothesis, we synthesized DNA oligonucleotides that contain either deoxythymine or deoxyuridine at the four sites (Figure 6) shown to interact with Ser37/Pro39 in the crystal structure (see Materials and Methods). The differences among these oligonucleotides are the presence (thymine) or absence (uridine) of the C5 methyl group on specific bases. Six 40 mer DNA duplex probes were generated: the wild-type probe A (wt/wt) containing the consensus core sequence, probe B (wt/*T17′dU*) with modification of one thymine on the antisense strand, probe C (wt/*T12′1dU:T17′dU*) with modifications of two thymines on the antisense strand which should interfere with one subunit binding in both dimers, probe D (*T12dU/T17′dU*) with two modifications on thymines that should affect one dimer binding but not the other, probe E (*T12dU/T12′dU:T17′dU*) containing three modifications, and probe F (*T12dU:T17dU/T12′dU:T17′dU*) with all four bases mutated. Gel mobility-shift assays were used to assess the ability of DtxR(C102D) to bind to these probes (Figure 6). Under the standard assay condition (see Materials and Methods), the wild-type probe was readily shifted by C102D in the presence of Mn^{2+}. It became more and more difficult to shift the DNA probe with increasing number of modifications. For probe E, the shifted band could barely be seen. For probe F, binding was abolished. For both of these probes, the smear underneath the shifted complex band was more prominent than for the first four probes. Such a smear indicates the presence of an unstable complex that can equilibrate between associated and dissociated DNA during the gel running time. The fraction of DNA bound for each probe was determined from at least three experiments. It is interesting to note that probes C and D, which both contain a total of two modifications, seem to behave slightly differently, implying that modifications impairing two DNA interactions in one dimer (D) hamper DNA binding more strongly than modifications that impair one DNA interaction in each dimer (C). However, the effect is modest and may be within the limits of resolution of this experiment.

We determined the affinity of DtxR(C102D) for the six DNA probes above (Figure 7) using data from the averages of at least three experiments each. The fraction of DNA probe bound was then plotted against the repressor dimer concentration. The data indicate that two DtxR dimers bind one probe based on the following argument. Suppose it only takes one dimer binding the probe to cause a mobility shift. When

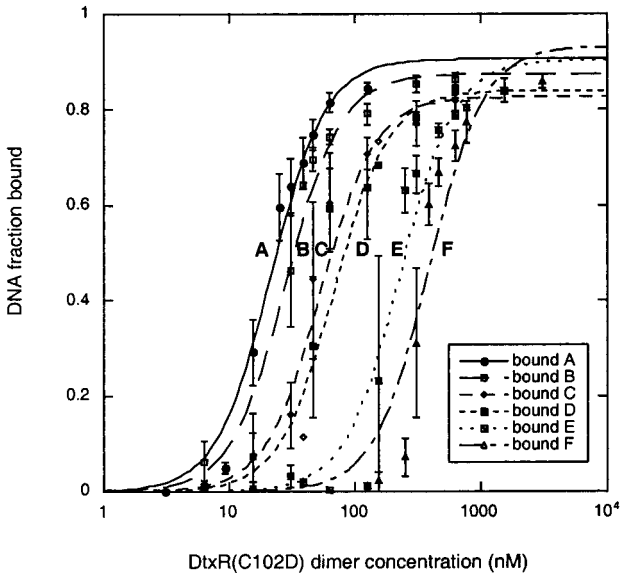

FIGURE 7: DNA binding curves determined by gel mobility-shift experiments. The data have been fitted to a simple one-step two-dimer binding model (see Materials and Methods). Each point is the result of three experiments. The error bars represent the standard deviation of the mean.

considering probes A and D, the former would have two identical binding sites for a dimer, while the latter has one wild-type dimer binding site and a defective dimer binding site. If we assume that the defective binding site is so changed that it no longer binds a dimer, then probe D only has one wild-type binding site. We can assume that in probes A and D the binding sites have identical intrinsic microscopic dissociation constants. Since A has two sites, its macroscopic dissociation constant would be one-half of that of probe D. This suggests that the half-saturation DtxR dimer concentration for probe D would be twice that for probe A. However, the experimental data show that it takes at least 5 times more DtxR dimer to half-saturate probe D than to half-saturate probe A. So the assumption above is not valid. It takes at least two dimers to bind the probe.

The data also indicate that repressor binding is cooperative. Let us assume that the binding is a two-step process:

$$\text{dimer} + \text{DNA} \leftrightarrow \text{dimer}\cdot\text{DNA}$$

$$\text{dimer}\cdot\text{DNA} + \text{dimer} \leftrightarrow \text{dimer}_2\cdot\text{DNA}$$

Then for probes A and D, the first step is the same (at most there is a factor of 2 difference), but probe D has a weaker binding site for the second step. If the two steps are not coupled, we would expect to see a binding intermediate for D in which a smaller complex (dimer·DNA) is detectable. However, such an intermediate is never stable enough to be detected by mobility shift assay. Thus, such an intermediate is very short-lived and must have a high affinity toward the binding of the second dimer. So the two binding steps must be tightly coupled. The difference between probes A and D is the one weaker dimer-binding site in D. However, we see an overall decrease in DtxR binding. This indicates that the two dimer-binding sites on the probe can affect each other's binding affinity. The model used to analyze the gel-shift data assumes independent binding of two dimers per probe. However, the above discussion and the systematic shift of the data points from the calculated curve at lower protein

		Helix	Turn	Helix
C. diphtheriae DtxR	27	RARIAERLEQ**S**G**P**TVS**Q**TVARMERDGLV		
C. glutamicum DtxR	27	RARIAERLEQ**S**G**P**TVS**Q**TVARMERDGLV		
M. tuberculosis IdeR	27	RARIAERLDQ**S**G**P**TVS**Q**TVSRMERDGLL		
M. smegmatis IdeR	27	RARIAERLDQ**S**G**P**TVS**Q**TVSRMERDGLL		
S. lividans DesR	27	RARIAERLDQ**S**G**P**TVS**Q**TVARMERDGLV		
S. pilosus DesR	27	RARIAERLDQ**S**G**P**TVS**Q**TVARMERDGLV		

FIGURE 8: Sequence alignment of DtxR homologues. The numbering is based on each sequence. The residues Ser37, Pro39, and Gln43 (numbering in *C. diphtheriae* DtxR) involved in DNA base recognition are bold.

Table 3: Dissociation Constants and Change in Binding Energy for the Six DNA Probes

probe	K_d (nM^2)	$\Delta\Delta G$ (kcal/mol)
A	$(4.6 \pm 0.8) \times 10^2$	
B	$(8.0 \pm 1.7) \times 10^2$	0.33
C	$(2.8 \pm 0.7) \times 10^3$	1.1
D	$(4.6 \pm 1.7) \times 10^3$	1.4
E	$(5.2 \pm 1.7) \times 10^4$	2.8
F	$(1.5 \pm 0.5) \times 10^5$	3.4

[a] K_d = [DNA][dimer]^2/[DNA·dimer$_2$]. The error in each case is determined from the fit of the points to the calculated curve. In these experiments, the concentration of DtxR(C102D) dimer ([dimer]) is far greater than the concentration of DNA ([DNA]). $\Delta\Delta G_{B\rightarrow A} = RT \ln(K_{d,B}/K_{d,A})$. An error of approximately 20% in $\Delta\Delta G$ arrives from the errors in K_d.

concentrations (for each probe) indicates that binding might be cooperative.

The dissociation constants and the corresponding ΔΔG values are listed in Table 3. Modification of one thymine causes a 0.33 kcal/mol decrease in binding energy with probe B compared to the wild-type probe A. However, the binding energy contribution is not additive: two modifications on each strand (probe F) contribute to a total of 3.4 kcal/mol, resulting in a 25-fold increase in the half-saturation concentration of DtxR(C102D) required by the mutated probe. The energy contribution of the thymine methyl group is in agreement with mutation studies on Cro and λ-repressor: operator interactions (*33*, *39*, *40*). In their studies, the authors saw a contribution of 0.3−2 kcal/mol in free energy change per methyl group.

The six probes all show sigmoidal behavior upon DtxR binding, indicating cooperativity between the two DtxR dimers. This is also consistent with the fact that mutations in one binding site affect the overall binding of the two repressor dimers to the DNA. The sigmoidal binding curve of DtxR on wild-type consensus DNA (probe A) also demonstrates that the repressor can work as a sensitive switch to turn on and off toxin production: physiological changes that can produce small changes in activated DtxR dimer concentration will have profound effects on toxin production. We do not understand how the cooperativity in DNA binding is achieved since in the crystal structure the two dimers show no obvious interactions. Maybe binding of the first dimer causes conformational changes in the DNA, such as unwinding of the DNA, to facilitate the binding of the second dimer, and this in turn stabilizes the first dimer on the DNA. In addition, there is the possibility that the C terminal domain participates in mediating dimer−dimer interactions.

This work and the previous structural study of the DtxR(C102D)-*tox* operator complex show that protein interactions with the DNA backbones contribute strongly to binding. Binding specificity, however, depends predominantly on specific hydrogen-bonding interactions. Our results also demonstrate that van der Waals contacts also play an important role in tight binding and thus, base specificity.

DtxR Homologues. As a prototype iron-dependent regulatory protein, DtxR is a member of a family of homologous proteins from, among others, *Corynebacterium glutamicum* (*41*), *Mycobacterium tuberculosis*, and *Mycobacterium smegmatis* (*42*−*44*), *Streptomyces lividans*, and *Streptomyces pilosus* (*45*), identified by sequence alignment using Clustal W (*46*, *47*). These sequences are compared based on a conserved cysteine residue at the position analogous to C102 in DtxR. A second subfamily of DtxR homologues can also be identified that have a glutamate at this position, but are not included in this comparison. Shown in Figure 8 are the helix-turn-helix motifs from these proteins. Even though the native promoter/operator sequences in some organisms have not been identified to date, based on the conserved repressor sequences we predict that at least in *C. glutamicum*, *M. tuberculosis*, *M. smegmatis*, *S. lividans*, and *S. pilosus*, they will have consensus sequences similar to the DtxR consensus DNA sequence.

DtxR-Binding Sequences. DtxR not only controls the production of diphtheria toxin, it has also been shown to bind to operators *irp1* to *5*, and an operator controlling heme oxygenase expression (*hmuO*) in *C. diphtheriae* (*2*−*5*), as well as operator regions of five putative "iron boxes" in *Mycobacterium tuberculosis* (*48*). Figure 9 shows the alignment of all the operators that have been shown to be under the control of DtxR. The triangles indicate the bases interacting with Gln43, and circles with Ser37/Pro39. The interactions with each DtxR dimer are labeled above and below the sequence of the *tox* operator, respectively. These interactions show that with the spacing of 5 bps between the two dimers and the symmetry within a bound dimer, our structures are able to explain the recognition of 9 bps out of the 19 bp consensus sequence. The DNA sequence itself may also determine how easily it can unwind and bend upon DtxR binding. However, without a structure of the DNA alone, we cannot make any reasonable assumptions.

Within the five iron responsive promoters, *irp3* was shown to have the lowest affinity for DtxR (*5*): a 10-fold greater concentration of DtxR was required to obtain an optimal gel mobility shift with *irp3* compared to that with *tox*, *irp4*, and *5*. Nucleotide substitution experiments show that the change of C17 in *irp3* to the consensus T causes a dramatic increase in the binding of DtxR, while G or A had only small effects (*49*). This observation is readily explained by our results:

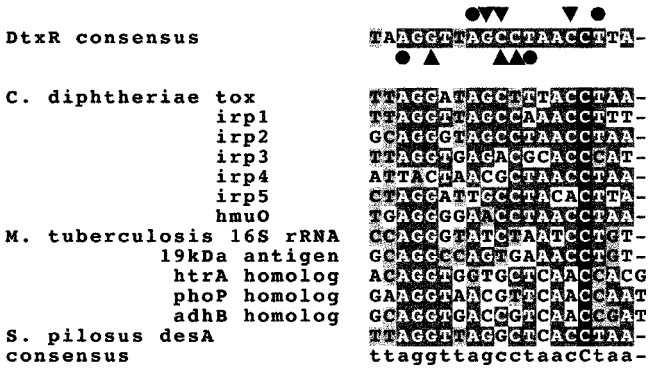

FIGURE 9: Sequence alignment of known DtxR binding sequences using ClustalW. The base specific interactions with DtxR are indicated on the consensus sequence: triangles indicate interaction with Gln43, circles with Ser37/Pro39 pair. The interactions with one DtxR dimer are shown above the consensus sequence, those for the other, below. Only one strand is shown for clarity, although the indicated interactions may be with the base pair partner. The most conserved bases from the overall alignment are indicated at the bottom.

if the DtxR-*irp3* complex is similar to the structure we observed here, then based on our complex structures, C17 and G12 in *irp3* are in positions to interact with the Ser37/Pro39 pair of subunits *b* and *c*, respectively. Since neither of these bases is the conserved T, this sequence is expected to have a lower affinity for DtxR, similar to that of the probe C used in this study. Restoration of the consensus nucleotide at base position 17 should make it a better binding substrate for the repressor. Further restoration at position 12 should make it even better.

In addition to the sequences known to interact with DtxR, the *desA* operator from *S. pilosus* has been found to be under the control of iron (*50*). The two genes controlled by the *desA* operator encode proteins involved in desferrioxamine production, the main siderophore of *S. pilosus*. The DNA sequence of this operator is also aligned in Figure 7. Interestingly, it bears significant similarity to the *tox* operator sequence. This operator sequence is sufficient for iron regulation in the original host, *S. pilosus*, and also in *S. lividans*, consistent with the fact that the putative iron repressor genes in these two organisms are well conserved and show high homology to DtxR (Figure 8) (*45*).

ACKNOWLEDGMENT

The authors thank Dr. Gregory A. Petsko for helpful discussions and comments. We thank Yi Wang, Vandana Arora, and Gloria E Hofileña for help in data collection. We also thank staff at CHESS beamline A1, BNL beamline X12B, and APS beamline 14-BMD for their technical support.

REFERENCES

1. Tao, X., Schiering, N., Zeng, H. Y., Ringe, D., and Murphy, J. R. (1994) *Mol. Microbiol. 14*, 191−7.
2. Schmitt, M. P., and Holmes, R. K. (1994) *J. Bacteriol. 176*, 1141−9.
3. Schmitt, M. P., Talley, B. G., and Holmes, R. K. (1997) *Infect. Immun. 65*, 5364−7.
4. Schmitt, M. P. (1997) *Infect. Immun. 65*, 4634−41.
5. Lee, J. H., Wang, T., Ault, K., Liu, J., Schmitt, M. P., and Holmes, R. K. (1997) *Infect. Immun. 65*, 4273−80.
6. Tao, X., and Murphy, J. R. (1992) *J. Biol. Chem. 267*, 21761−4.
7. Schmitt, M. P., and Holmes, R. K. (1993) *Mol. Microbiol. 9*, 173−81.
8. Tao, X., and Murphy, J. R. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 9646−50.
9. Tao, X., Zeng, H. Y., and Murphy, J. R. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 6803−7.
10. Schiering, N., Tao, X., Zeng, H., Murphy, J. R., Petsko, G. A., and Ringe, D. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 9843−50.
11. Qiu, X., Verlinde, C. L., Zhang, S., Schmitt, M. P., Holmes, R. K., and Hol, W. G. (1995) *Structure 3*, 87−100.
12. Qiu, X., Pohl, E., Holmes, R. K., and Hol, W. G. (1996) *Biochemistry 35*, 12292−302.
13. Ding, X., Zeng, H., Schiering, N., Ringe, D., and Murphy, J. R. (1996) *Nat. Struct. Biol. 3*, 382−7.
14. Pohl, E., Qui, X., Must, L. M., Holmes, R. K., and Hol, W. G. (1997) *Protein Sci. 6*, 1114−8.
15. Pohl, E., Holmes, R. K., and Hol, W. G. (1998) *J. Biol. Chem. 273*, 22420−7.
16. Wang, G., Wylie, G. P., Twigg, P. D., Caspar, D. L., Murphy, J. R., and Logan, T. M. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 6119−24.
17. Tao, X., and Murphy, J. R. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 8524−8.
18. White, A., Ding, X., vanderSpek, J. C., Murphy, J. R., and Ringe, D. (1998) *Nature 394*, 502−6.
19. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−26.
20. Navaza, J. (1994) *Acta Crystallogr., Sect. A 50*, 157−63.
21. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr., Sect. D 54*, 905−21.
22. Brünger, A. T. (1992) *Nature 355*, 472−5.
23. Read, R. J. (1986) *Acta Crystallogr., Sect. A 42*, 140−9.
24. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard. (1991) *Acta Crystallogr., Sect. A 47*, 110−9.
25. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−91.
26. Tao, X., Boyd, J., and Murphy, J. R. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 5897−901.
27. Pohl, E., Holmes, R. K., and Hol, W. G. (1999) *J. Mol. Biol. 292*, 653−67.
28. Ravishankar, G., Swaminathan, S., Beveridge, D. L., Lavery, R., and Sklenar, H. (1989) *J. Biomol. Struct. Dynam. 6*, 669−99.
29. Stofer, E., and Lavery, R. (1994) *Biopolymers 34*, 337−46.
30. Welkos, S. L., and Holmes, R. K. (1981) *J. Virol. 37*, 936−45.
31. Wang, Z., Schmitt, M. P., and Holmes, R. K. (1994) *Infect. Immun. 62*, 1600−8.
32. 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 335*, 321−9.
33. Albright, R. A., and Matthews, B. W. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 3431−6.
34. Gomis-Ruth, F. X., Sola, M., Acebo, P., Parraga, A., Guasch, A., Eritja, R., Gonzalez, A., Espinosa, M., del Solar, G., and Coll, M. (1998) *EMBO J. 17*, 7404−15.
35. Piper, D. E., Batchelor, A. H., Chang, C. P., Cleary, M. L., and Wolberger, C. (1999) *Cell 96*, 587−97.
36. Passner, J. M., Ryoo, H. D., Shen, L., Mann, R. S., and Aggarwal, A. K. (1999) *Nature 397*, 714−9.
37. Xu, H. E., Rould, M. A., Xu, W., Epstein, J. A., Maas, R. L., and Pabo, C. O. (1999) *Genes Dev. 13*, 1263−75.
38. Mandel-Gutfreund, Y., Margalit, H., Jernigan, R. L., and Zhurkin, V. B. (1998) *J. Mol. Biol. 277*, 1129−40.
39. Sarai, A., and Takeda, Y. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 6513−7.
40. Takeda, Y., Sarai, A., and Rivera, V. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 439−43.

41. Oguiza, J. A., Tao, X., Marcos, A. T., Martin, J. F., and Murphy, J. R. (1995) *J. Bacteriol. 177*, 465−7.
42. Schmitt, M. P., Predich, M., Doukhan, L., Smith, I., and Holmes, R. K. (1995) *Infect. Immun. 63*, 4284−9.
43. Doukhan, L., Predich, M., Nair, G., Dussurget, O., Mandic-Mulec, I., Cole, S. T., Smith, D. R., and Smith, I. (1995) *Gene 165*, 67−70.
44. Dussurget, O., Rodriguez, M., and Smith, I. (1996) *Mol. Microbiol. 22*, 535−44.
45. Gunter-Seeboth, K., and Schupp, T. (1995) *Gene 166*, 117−9.
46. Higgins, D. G., Bleasby, A. J., and Fuchs, R. (1992) *Comput. Appl. Biosci. 8*, 189−91.
47. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res. 22*, 4673−80.
48. Manabe, Y. C., Saviola, B. J., Sun, L., Murphy, J. R., and Bishai, W. R. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 12844−8.
49. Lee, J. H., and Holmes, R. K. (2000) *J. Bacteriol. 182*, 432−8.
50. Gunter, K., Toupet, C., and Schupp, T. (1993) *J. Bacteriol. 175*, 3295−302.
51. Engh, R. A., and Huber, R. (1991) *Acta Crystallogr. A 47*, 392−400.
52. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−50.

BI0009284