Dynamic Consequences of Specificity within the Cytidine Repressor DNA-Binding Domain

Colleen L. Moody¹, Jenaro Soto², Vira Tretyachenko-Ladokhina¹, Donald F. Senear¹, and Melanie J. Cocco^{1,2*}

¹Department of Molecular Biology and Biochemistry, University of California, Irvine, Irvine, CA 92697, USA

²Department of Molecular Biology and Biochemistry, University of California, Irvine, Irvine, CA 92697, USA

*To whom correspondence should be addressed. Email: mcocco@uci.edu

ABSTRACT

The *E. coli* cytidine repressor (CytR) is a member of the LacR family of bacterial repressors that regulates nine operons with distinct spacing and orientations of recognition sites. Understanding the structural features of the CytR DNA-binding domain (DBD) when bound to DNA is critical to understanding differential mechanisms of gene regulation. We previously reported the structure of the CytR DBD monomer bound specifically to half-site DNA and found that the DBD exists as a three-helix bundle containing a canonical helix-turn-helix motif, similar to other proteins that interact with DNA [Moody, et al (2011), *Biochemistry* **50**:6622-32]. We also studied the free state of the monomer and found that since NMR spectra show it populates up to four distinct conformations, the free state exists as an intrinsically disordered protein (IDP). Here, we present further analysis of the DBD structure and dynamics in the context of full-site operator or nonspecific DNA. DBDs bound to full-site DNA show one set of NMR signals, consistent with fast exchange between the two binding sites. When bound to full-length DNA, we observed only slight changes in structure compared to the monomer structure and no folding of the hinge helix. Notably, the CytR DBD behaves quite differently when bound to nonspecific DNA compared to LacR. A dearth of NOEs and complete lack of protection from hydrogen exchange are consistent with the protein populating a flexible, molten state when associated with DNA nonspecifically, similar to fuzzy complexes. The CytR DBD structure is significantly more stable when bound specifically to the *udp* half-site substrate. For CytR, the transition from nonspecific association to specific recognition results in substantial changes in protein mobility that are coupled to structural rearrangements. These effects are more pronounced in the CytR DBD compared to other LacR family members.

INTRODUCTION

The cytidine repressor protein (CytR), a member of the LacR family of bacterial repressors [1] [2] regulates multiple operons whose genes encode transport proteins and enzymes involved in ribonucleotide salvage and catabolism. Recently, CytR has been shown to control the production of flagella in uropathogenic *E. coli* that cause urinary tract infections. In *V. cholerae*, CytR is involved in nutrient sensing and controls competence [3-6], adhesion [7] and biofilm production [8].

In contrast to other transcription regulators, CytR tolerates widely variable spacing between the inverted repeats that comprise its natural operators, ranging from zero base pairs at *deoP2* to eleven base pairs at *rpoH*, a variation that accounts for up to a 38Å translation and a complete rotation of the half-sites. We have shown that this accommodation requires distinctly different thermodynamic modes of CytR-DNA binding, which we interpret as ranging from a relatively rigid conformation of CytR bound to zero base pair spacing at *deoP2* to a flexible and dynamic conformation bound to the nine base pair spacing of *nupG* [9, 10]. The *udp* promoter is involved in the regulation of expression of uridine phosphorylase [11] and is an example of an intermediate spacing of three base pairs between CytR recognition sites.

To address the structural mechanism underlying these effects we have used NMR spectroscopy to investigate the structure and dynamics of isolated DBD and linker peptide bound to operator DNA and when associated with nonspecific DNA. Previously we determined the structure of the CytR DBD monomer bound specifically to one DNA half-site of the uridine phosphorylase (*udp*) operator [12]. The CytR DBD structure demonstrated that the conformation of specifically-bound protein is different from the common conformation of the other four LacR DBDs whose structures are known. If helix 2 is superimposed, both helices 1 and 3 move in the CytR DBD compared to LacR.

We also showed that operator-DNA binding is accompanied by a substantial disorder to order transition in going from the free state to the three-helix bundle that comprises the DBD core. In contrast, the three helices of the LacR DBD remain largely folded regardless of whether this domain is free in solution or bound to DNA. We found that the free CytR DBD is intrinsically disordered, existing in multiple conformations with only 20% helicity as measured by circular dichroism and very few NOEs. We found two sets of NMR signals in helices 1 and 2 and four sets of signals in helix 3 (figure 6 in [12]). For a residue in multiple conformations, the NMR timescale of exchange is defined by the separation between NMR resonance peaks in cycles/sec (Hz). Measuring the distance in Hz between

peaks sets an upper limit for the rate of exchange. We calculate that, based on the largest peak separation, the rate of exchange between species must be slower than 0.044 sec^{-1} when the DBD is free in solution. Thus, the alternate conformations of the DBD are stable and could represent unique conformations optimized for different operator structures. Thus, the CytR DBD can be characterized as an intrinsically disordered protein (IDP). Since our CytR DBD structure was published, it has been used as a paradigm in understanding IDPs in comparison to other folded LacR DBDs [7, 13-18].

Here we find that although isolated DBD also binds nonspecific DNA with only moderately lower affinity, this binding is not accompanied by the same magnitude of folding of binding to specific site DNA. In this respect, the CytR DBD with nonspecific DNA is a fuzzy complex where the protein is only partially folded [19-21]. Further, even when DBD is bound to full-site operator-DNA, the CytR hinge helix does not fold and a dimer interface is not formed as it is in LacR. We thus conclude that specific DBD-DNA contacts are necessary for full stabilization of the bound structure with folding of the DBD being coupled to DNA binding.

MATERIALS AND METHODS

Protein expression and purification

The 67 residue CytR DNA binding domain (DBD) was expressed in *E. coli* strain BL21(DE3)pLysS transformed with expression plasmid pSS584DBD[22]. ^{15}N labeled protein was produced by growing the cells in Neidhardt's minimal media [23] containing ^{15}N -ammonium chloride as the sole nitrogen source and purified as described [24]. Purity was above 95% as assessed by SDS PAGE and MALDI mass spectrometry. Concentration was determined using extinction coefficient $A_{280} = 0.202 \text{ (g/L)}^{-1} \text{ cm}^{-1}$ calculated from the amino acid composition [25].

Oligonucleotide preparation

Deoxyribonucleotides were purchased as HPLC purified products from IDT Technologies, San Diego, CA.. Complimentary single strand oligonucleotides were annealed to double strands as described previously [24]. Each double strand sequence contained one strand conjugated to either Alexa 532 or Oregon Green 514 fluorescent dye through a 6-carbon linker to the 5'-phosphate.

DBD binding to DNA by steady-state fluorescence anisotropy

Steady-state fluorescence measurements were made using an SLM 8100 equipped with a double-grating excitation monochromator and a single-grating emission monochromator. Fluorescence was excited at 530 nm and emission measured at 580 nm, both with excitation 4 nm and emission 8 nm band pass. Fluorescence intensity was recorded for 2 minutes at each position of the excitation and the emission polarizers. Anisotropy was calculated from the average intensities according to

$$\langle r \rangle = (I_{vv} - G \times I_{vh}) / (I_{vv} + 2G \times I_{vh}) \quad \text{eq. 1}$$

I_{vv} and I_{vh} are the intensities at the excitation polarizer in the vertical position and emission in the vertical and horizontal positions of the emission polarizer, respectively. The G factor,

$$G = I_{hv} / I_{hh} \quad \text{eq. 2}$$

was calculated from the average of several measurements of free labeled DNA.

Binding titrations were conducted in 50 mM phosphate buffer, pH 6 containing 30 mM NaCl and 1 mM EDTA. Aliquots of DBD from a 5 mM concentrated stock were added to a solution of Alexa 532-labeled DNA (~100 nM) in 2x10 mm quartz cuvettes at a temperature of 20 ± 0.1 °C maintained by a Peltier device. After each addition of protein, the cuvette was incubated in the thermo-regulated cell holder for 10 minutes prior to taking intensity measurements to ensure that equilibrium had been attained. Anisotropy values were analyzed as a function of DBD concentration using non-linear least squares analysis in Origin 7 according to:

$$\langle r \rangle = (r_{\text{bound}} - r_{\text{free}}) \left(\exp(2.303 \ln[\text{DBD}]) - \Delta G^\circ / RT \right) / \left(1 + \exp(2.303 \ln[\text{DBD}]) - \Delta G^\circ / RT \right) + r_o \quad \text{eq. 3}$$

where r_o is a free labeled DNA anisotropy in the absence of protein, r_{max} is the anisotropy of DBD-DNA complex, ΔG° is a standard free energy of binding, and R and T are the gas constant and absolute temperature. Dissociation constants are calculated as $k_d = 1/K_a = \exp(\Delta G^\circ / RT)$.

Chemical shift analysis: Calculation of chemical shift changes, $\Delta\delta_{\text{amide}}$ was calculated using the changes in chemical shift of ^1H and ^{15}N according to the following equation from Williamson, 2013 [26]:

$$\Delta\delta_{\text{amide}} = \sqrt{\frac{1}{2} [\delta_H^2 + (\alpha \cdot \delta_N^2)]}$$

Where δ_H is the change in chemical shift of ^1H and δ_N is the change in chemical shift of ^{15}N between CytR DBD in the presences of specific and non-specific DNA fragments. α is the nitrogen scaling factor

of 0.14 [26].

NMR dynamics measurements

All NMR experiments were performed with samples in 90% H_2O /10% D_2O in 50 mM sodium phosphate, 30 mM NaCl, and 1 mM EDTA at pH 6. Samples with *udp* full-site contained 0.411 mM ^{15}N CytR DBD and 0.889 mM natural abundance DNA. Samples with nonspecific DNA contained 0.4 mM ^{15}N CytR DBD and 0.6 mM natural abundance half-site nonspecific DNA. The free DBD samples contained 1 mM DBD ^{15}N CytR DBD.

NMR data were acquired on a Varian Inova 800MHz spectrometer. Data processing was performed using NMR Pipe (30) and spectra were analyzed using CCPNMR Analysis (31). Assignments were previously published[24]. ^{15}N T1 and T2 relaxation measurements were made using ^{15}N HSQC pulse sequences and varying relaxation times (46, 47) (T1 = 0, 50, 100, 200, 750, and 1500 ms; T2 = 0, 10, 30, 50, 70, 90, 110, 130, 150, 170, 190, 210, 230, and 250 ms). Peak intensities were fit to single exponential functions to determine the relaxation times (errors were calculated from the fit). Steady-state NOE values were calculated as the ratio between the peak intensities from the saturated and unsaturated spectra (48).

Hydrogen exchange measurements were performed by lyophilizing ^{15}N CytR DBD with natural abundance DNA in standard buffer and then resuspending the sample in D_2O . ^1H -1D and HSQC experiments were acquired immediately upon resuspension and transfer to Shigemi NMR tubes. ^1H -1Ds were acquired in between each HSQC (nt=16, ni=40, np=1024) for calibration purposes. HSQCs were processed without apodization and peak intensities were fit to single exponential curves to extract hydrogen-deuterium exchange rates for backbone amides in the CytR DBD *udp* half-site sample. Backbone amide protons of CytR DBD with nonspecific DNA exchanged too rapidly to obtain quantitative rate information. Protection factors were calculated for the CytR DBD with *udp* half-site sample using the following equation: Protection Factor = $k_{\text{intrinsic}}/k_{\text{measured}}$. The intrinsic rates were calculated based on studies by Bai et al. [27].

RESULTS We characterized the structure and dynamics of the CytR DBD in the context of three types of DNA substrates. These are: bound in a 1:1 complex to an operator half-site (*udp*), bound to a nonspecific DNA sequence (the same size as the *udp* half-site), and bound to full-length operator (*udp*), with the possibility that the DBD might dimerize on the DNA. CytR operators are comprised of inverted repeats of an eight base pair recognition motif [28] separated by variable spacing ranging from zero to nine base pairs [10]. These experiments used the natural operator (CytO) from *udp* (5'-GTAAATTTATGCAACGCAATTTGCGTCATGGTG-3'), an operator with three base pair spacing (underlined) hereafter referred to as full-site DNA. For the single recognition site sample, the operator left half-site of *udp* (5'-ATTTATGCAACGCA-3') is the best match to the consensus recognition motif defined by SELEX [28]. The nonspecific DNA substrate (5'-AGCATGCTGGGGAC-3') is the same length as the *udpP* half-site (14 bp) but contains the binding sequence of an unrelated transcription factor, i.e., the human p53 binding sequence found in the *gadd45* promoter [29]. As **Figure 1** indicates, the NMR

spectrum of the DBD bound to the full-site operator has many similarities to that of the half-site. When chemical shift changes between half- and full-site DNA spectra are plotted onto the monomer structure diagram in **Figure 1**, it appears that the largest changes in chemical environment are proximal to the first helix. In the presence of the nonspecific DNA, the DBD gives one set of signals at low temperatures (20 °C), but similar to the free protein [24], is mostly unfolded at 35 °C (**Figure 2**). Although the DBD is thermodynamically less stable when bound to nonspecific DNA than when bound to *udpP*, binding still produces a single structure. However, a comparison of HSQC spectra in Figure 2 and Figure 1B of [24] reveals that the chemical shifts of the protein/nonspecific DNA complex are different when compared to those of the protein/*udp* half-site (calculated in **Figure 3**). Comparison of spectra presented here and previously [24], shows that many DBD NMR peak positions are unique to the particular state of the protein (free, bound to specific DNA, or bound to nonspecific DNA).

Protein/DNA binding affinity measured by fluorescence anisotropy

The spectral differences of the DBD-DNA complexes formed with specific versus nonspecific DNA prompted an analysis of the affinity of DBD for binding to both half-site and full-site operator sequences as well as for binding to two nonspecific DNA oligonucleotides. One nonspecific DNA was the 14 bp oligonucleotide described above; the second was an extended version of this sequence (5'-GTACAGAACATGTCTAAGCATGCTGGGAC-3'), whose length of 30 bp was chosen to match the

udpP full-site oligonucleotide. We monitored fluorescence anisotropy of the dye-conjugated DNA substrates to yield an experimental observable for binding. However to analyze the titrations that result, it should be noted that while the fluorescence anisotropy method distinguishes bound from free DNA, it is insensitive to the position in which the DBD is bound to the DNA. It is also relatively insensitive to the stoichiometry and so does not distinguish between a single DBD bound versus multiple DBD's bound.

Figure 4 compares binding of CytR DBD to these four oligonucleotides. The data shown for specific binding to the full-length *udp* operator were reported previously [24]. Analysis according to eq. 3 yields a binding free energy change, $\Delta G^\circ = -7.14 \pm 0.04$ kcal/mol, which corresponds to the equilibrium constant for dissociation, k_D , equal to 4.7 ± 0.3 μM . DBD binding to the *udp* operator left half-site yielded $\Delta G^\circ = -6.73 \pm 0.10$ kcal/mol, or $k_D = 9.6 \pm 1.8$ μM . Eq. 3 reflects a simple 1:1 binding stoichiometry. The difference of two-fold in apparent affinity between full-site *udp* operator and the left operator half-site is exactly as expected given that the full operator offers two sites for DBD binding. Accounting for the stoichiometry, and assuming no interaction between DBD binding to the left- and right half-sites of the full-length operator,

these values are consistent with an intrinsic affinity of 10 μM .

Analysis of the nonspecific binding isotherms using eq. 3 yields $\Delta G^\circ = -7.70 \pm 0.07$ kcal/mol ($k_D = 1.81$ μM) for binding to the 30 bp oligonucleotide and $\Delta G^\circ = -7.03 \pm 0.05$ kcal/mol ($k_D = 5.7$ μM) for binding to the 14 bp oligonucleotide. The difference in apparent binding free energy $\Delta\Delta G^\circ = 0.67$ kcal/mol, corresponds to a three-fold difference in apparent affinity. Although the site size for nonspecific binding is uncertain, a value of approximately eight base pairs is expected, because this is the length of the inverted repeat sequences for specific binding. With this assumption, the 30 bp oligonucleotide offers three times as many binding sites as the 14 bp oligonucleotide, exactly accounting

for the apparent affinity difference observed. Accounting for the stoichiometry these values are consistent with an intrinsic affinity for nonspecific binding (41 μ M) that is only four-fold lower than for binding to the *udp* operator. These results indicate that the sequence specificity of the isolated monomeric DBD is modest. The results also show that the DBD is fully bound to both specific and nonspecific DNA under conditions for which NMR spectra are collected.

Structural Analysis of the CytR DBD with Nonspecific DNA

We assessed the structure of the DBD in the presence of nonspecific DNA using NMR spectroscopy. In contrast to both *udp* half and full-site oligonucleotides, many of the ^1H - ^{15}N HSQC DBD peak positions in the nonspecific DNA spectrum change dramatically (**Figure 3**). At positions where chemical shifts changes were small or moderate, assignments were extrapolated and confirmed by comparison of sequential ^{15}N NOESYHSQC data and deduction based on amino acid content. Analysis of NOE patterns showed strong similarities within the unfolded N- and C-termini of the construct; specifically, amino acids 3-9 and 54-67 (excluding prolines) show both similar NOE patterns regardless of DNA substrate. However, in the intervening regions (10 to 53), most amino acids show very few intermediate- or long-range NOEs in the DBD/nonspecific DNA ^{15}N NOESYHSQC spectrum, providing only 36% of the total NOEs compared to the DBD/*udp* half-site (e.g., **Figure S1**). Notably, the few NOE patterns present in the DBD/nonspecific DNA spectrum do correlate with peaks in the DBD/*udp* half-site spectrum. Unfortunately, these limited NOE data are insufficient to determine the structure of the DBD bound to the nonspecific DNA; however, since we have assigned the backbone amide signals, we can make a quantitative comparison of the dynamics between specific and nonspecific DNA.

Dynamics of the CytR DBD

Relaxation in the laboratory frame is a useful means to determine motions in the ps-ns timescale range [30]. Having the backbone assignments for DBD bound both specifically and nonspecifically allows us to compare protein dynamics for these two types of interaction. We measured T1 and T2 ^1H - ^{15}N relaxation times, and ^1H - ^{15}N NOE intensities using ^{15}N HSQC methods. Relaxation and NOE values for the N- and C-terminal ends of the DBD are similar regardless of the DNA binding partner. However, differences occur in the central helical regions when the DBDs interacts with specific (**Figure S2A, C and E**) versus nonspecific (**Figure S2B, D and F**) DNA sequences. **Figure 5** compares the ^1H - ^{15}N NOEs when the DBD is bound nonspecifically vs binding to a specific site. Since the ^1H - ^{15}N NOE

intensity is a sensitive indicator of dynamics, these data show where the DBD becomes more flexible (lower NOE value) when bound to nonspecific DNA. Interestingly, changes in dynamics are not restricted to the loops. The start of helix 2 and middle of helix 3 show substantial changes in NOE intensities that indicate they become more flexible upon binding to nonspecific DNA.

Figure S2B shows that the DBD T_2 relaxation is too rapid to measure using the ^{15}N HSQC technique for most of the protein associate with nonspecific DNA. This is consistent with conformational exchange on the NMR timescale for this species. T_1 and T_2 values on their own are difficult to interpret, but the ratio and product of R_2 and R_1 (**Figure 6**) provide insight into deviations from the global correlation time. Relaxation rate ratios ($R_2/R_1 = [1/T_2]/[1/T_1]$) provide a useful qualitative analysis of dynamics as the average R_2/R_1 value for a backbone amide in ordered

parts of the protein provides a measure of the overall correlation time [31]. R_2/R_1 can also distinguish between fast (ps-ns) and slow (μ s-ms) dynamics [32, 33]. Residues with R_2/R_1 greater than the limit of one standard deviation from the average value for backbone positions indicate a region of slow (μ s-ms) motions and those less than the limit of one standard deviation correlate with fast (ps-ns) dynamics. The R_2/R_1 ratio for the DBD/nonspecific DNA complex shows that helix 3 appears to exhibit slow motions compared to helices 1 and 2. In addition, R_2^*/R_1 can be used to distinguish regions of fast exchange [34]. The R_2^*/R_1 product shows chemical exchange at the start of helix 2 and helix 3.

Hydrogen exchange is also a useful measure of protein dynamics and complements relaxation in the laboratory frame by sampling longer timescales of motion. Previously we found complete amide proton exchange occurred within ten minutes of transfer into D_2O for the free DBD [24]. Notably, we find here that binding to the nonspecific DNA substrate offered no additional protection with complete exchange of the DBD amide protons occurring in the first ten minutes. In contrast, with the DBD bound to the *udp* half-site, we find residues that are protected from exchange for up to three hours following transfer into D_2O and incubation at 20 °C (pH 6.0 prior to lyophilization). We measured exchange rates and calculated protection factors for all of the positions resistant to exchange. In contrast to the DBD/nonspecific DNA complex that exchanges immediately, all three DBD helices show protection from exchange when complexed with *udp* (Figure 7A and B).

DISCUSSION

We have now examined the CytR DBD in four states: alone, in the presence of *udp* full-site DNA, in the presence of *udp* half-site DNA, and in the presence of nonspecific DNA. The thermal melting properties of the CytR DBD/DNA systems are intriguing. Although both specific and nonspecific half-site DNA substrates influence the DBD structure by directing the protein into a distinct fold (one set of peaks), the DNA sequences have very different effects on protein thermal stabilities. Significantly, we find that the protein fold is strongly stabilized when bound to specific DNA, but with nonspecific DNA, the DBD is susceptible to thermal denaturation at much lower temperatures, in fact similar to the free state melting.

Our analysis of the DBD spectral features in the presence of the *udp* full-site indicates this species maintains a helix-turn-helix fold similar to the DBD bound to half-site DNA. We find one set of peaks for DBD/*udp* full-site DNA indicating that the protein is in fast exchange between the two recognition sites and consequently we observe an average bound state. Our analysis of DBD NOEs when bound to *udp* full-site DNA reveals that the hinge helix is not formed. We established binding properties of the CytR DBD to the *udp* half-site, as well as to the *udp* full-site. In addition, we studied both nonspecific half- and full-sites. The binding isotherms of CytR DBD to both half-site and full-site DNA are very similar (**Figure 4**); this is consistent with a lack of interaction between DBD monomers when they bind to *udp* full-site. There are two possible interpretations: a) that only a single monomer binds to both the half- and full-site DNA substrates, and b) that the monomers that bind the individual half-sites of the full-site sample do so in an equal and independent manner. Either interpretation supports the relevance of the monomer half-site complex structure presented here. In consideration of regulatory function, equivalent binding constants to half- and full-site DNA combined with the lack of hinge helix formation in full-site DNA suggests that cytidine binding domain (CBD) is necessary for dimer to form. This is either because only the CBD contributes to the subunit-subunit interface or because the hinge peptides require the bottom surface of the CBD dimer to form. The topological differences of CytR and LacR DBDs indicate that the interactions between the surface of the CytR DBD and the CBD should also be unique.

The CytR DBD has similar affinities for both the specific and nonspecific DNA sequences explored in this study. This is not a general feature of HTH DBDs; *e.g.*, HoxD9 has a 200-fold difference in K_D 's favoring specific site for binding [35]. Although it appears from fluorescence measurements that all of

these DNA substrates bind the DBD with similar affinities, clearly the NMR data reveal stark differences in the protein structure when bound to specific versus nonspecific DNA. *udp* half-site specific DNA can accommodate only one protein molecule when correctly bound. In contrast, protein molecules can bind in multiple positions on nonspecific DNA; collectively, multiple binding sites on a nonspecific substrate could give the appearance that the protein has a higher affinity than individual binding events. This type of binding-mode difference would be consistent with our finding that the structure is stable on specific DNA and much more flexible with nonspecific interactions. However, even in consideration of a lattice model for nonspecific binding, the binding to specific and nonspecific DNAs appear to be energetically similar. At most, specific binding would only be 4-fold higher (~1 kcal/mol) in the case of protein occupying all available sites on nonspecific DNA.

The energetic balances of the CytR system are quite interesting. Protein folding of the bound conformation must be unfavorable; if not, the DBD would populate the DNA-bound structure even when free in solution. Therefore, favorable interactions with DNA drive unfavorable protein folding for the CytR DBD. Either the two bound protein conformations (specific/nonspecific) are nearly isoenergetic and the intrinsic protein-DNA interactions are nearly isoenergetic (specific vs nonspecific) or the protein-DNA and protein folding are different in specific vs nonspecific cases, but such that they nearly compensate for one another. The DBD folds to a perhaps better defined and more rigid, but energetically less favorable structure in order to make better protein-DNA contacts associated with specific binding, whereas nonspecific binding is inherently weaker and consequently drives less protein folding. The net results are two energetically similar states. Regardless of the apparent similarities in binding affinities, the DNA sequence has a profound impact on the structural properties of the DBD in contrast to LacR.

The results of NMR dynamics experiments indicate significant alterations in the behavior of the protein when bound to a particular DNA substrate. ¹⁵N relaxation analysis indicates that residues throughout the range 12-50 of the DBD are more flexible in the presence of nonspecific DNA compared to specific DNA, potentially corresponding to movements within the secondary structure. Hydrogen-deuterium exchange reveals that the DBD/nonspecific DNA has greater overall flexibility compared to DBD/*udp* half-site. Despite evidence for a single structure of the DBD with nonspecific DNA, the protein takes on a considerably more dynamic state that allows for the solvent accessibility of all the protein amide groups within the timescale measured here. When bound to nonspecific DNA, larger domain motions could account for the lack of NOEs and consequently would predict that the average

structure is different depending on the DNA substrate. In the case of specific DNA, the amide protons of the DBD that were most resistant to exchange are highlighted on the structure (**Figure 7B**).

Although the structure of the CytR DBD resembles the LacR fold in the presence of a specific promoter target, there are differences in the structures induced by nonspecific DNA and the protein free in solution. In the case of LacR, the DBD adopts the same conformation regardless of state (bound or free) [36]. The positions of the first three helices are maintained in all of the LacR structures. Moreover, nonspecific binding does not significantly impact the conformation of the LacR DBD. The PurR DBD structure varies slightly between bound and unbound forms, however structures of the bound form also contain ligand binding domain [37, 38]. CytR is quite different from these structures, adopting a more varied set of conformations depending on the state and binding partner. Although the LacR DBD provided good NOE data to calculate the structure when bound nonspecifically [36], CytR with nonspecific DNA only provides a fraction (1/3) of the NOE constraints needed to define a high-resolution structure. However, having peak assignments allows us to define dynamics on a per-residue basis using NMR techniques. There are significant deviations in protein dynamics between LacR [39] and CytR DBDs. The result of the LacR H-exchange predicts stronger protection factors throughout the protein with helix 2 being most rigid. In contrast, CytR appears to be generally more flexible than LacR, even when bound to a target DNA site and the distribution is different with helices 1 and 3 show more extensive protection from exchange than helix 2.

The distinct structural properties of the CytR DBD free or bound to different DNAs are independent of the interactions with the CBD as this domain is not present in our sample. To generate any effect on differential gene regulation, the CBD interactions with CRP must be involved. Different spacing can affect the conformation of the hinge helices and packing between DBD and CBD; this was anticipated to influence coupling. What we show here is that even without these constraints, the DNA sequence dictates the DBD conformation. Hence, some effect on gene regulation might take place at the level of DBD structure. Therefore, a portion of the differential gene regulation of the full CytR in the cell may be attributed to an altered structure and dynamics of the DNA binding domain itself. Thus, intrinsic affinity for a particular DNA sequence may be overwritten by the structural and dynamical consequences of sequence recognition. In this way, the effects of DNA sequence on the structure and dynamics of the protein may be the basis for distinct levels of gene expression.

Analysis of the structure and dynamics of the CytR DBD allows us to further hypothesize about the link between protein flexibility and DNA recognition. The intrinsic disorder of the free CytR DBD has

been confirmed and compared to LacR [7, 13-18]. Our characterization of the nonspecifically bound state classify CytR as forming a fuzzy complex with DNA where the bound protein is only partially structured [19-21]. The dynamics of the DBD may play an integral role in influencing the dynamic state of the rest of the protein. Therefore, the flexibility of the DBD could be transmitting information about the DNA sequence to the CBD and consequently directing the strength of interactions between CytR and CRP. The conformation of the DBD is coupled to DNA binding; however, hinge helix folding and dimer formation requires DBD/CBD interactions, consistent with our hypothesis that differences in spacing are accommodated by changes in protein conformation.

FUNDING

This work was supported by the National Science Foundation [MCB 02115769, MCB 06652875] and an NIH fellowship to JS [GM055246].

References

1. Keseler, I.M., et al., *EcoCyc: a comprehensive view of Escherichia coli biology*. Nucleic Acids Res, 2009. **37**(Database issue): p. D464-70.
2. Kristensen, H.H., P. Valentin-Hansen, and L. Sogaard-Andersen, *Design of CytR regulated, cAMP-CRP dependent class II promoters in Escherichia coli: RNA polymerase-promoter interactions modulate the efficiency of CytR repression*. J Mol Biol, 1997. **266**(5): p. 866-76.
3. Sun, Y., et al., *Competence and natural transformation in vibrios*. Mol Microbiol, 2013. **89**(4): p. 583-95.
4. Antonova, E.S., E.E. Bernardy, and B.K. Hammer, *Natural competence in Vibrio cholerae is controlled by a nucleoside scavenging response that requires CytR-dependent anti-activation*. Mol Microbiol, 2015. **97**(3): p. 605.
5. Watve, S.S., J. Thomas, and B.K. Hammer, *CytR Is a Global Positive Regulator of Competence, Type VI Secretion, and Chitinases in Vibrio cholerae*. PLoS One, 2015. **10**(9): p. e0138834.
6. Cohen, J.J., et al., *Control of competence in Vibrio fischeri*. Appl Environ Microbiol, 2021.
7. Das, S., et al., *Multifunctional transcription factor CytR of Vibrio cholerae is important for pathogenesis*. Microbiology (Reading), 2020. **166**(12): p. 1136-1148.
8. Haugo, A.J. and P.I. Watnick, *Vibrio cholerae CytR is a repressor of biofilm development*. Mol Microbiol, 2002. **45**(2): p. 471-83.
9. Tretyachenko-Ladokhina, V., J.B. Ross, and D.F. Senear, *Thermodynamics of E. coli cytidine repressor interactions with DNA: distinct modes of binding to different operators suggests a role in differential gene regulation*. J Mol Biol, 2002. **316**(3): p. 531-46.
10. Tretyachenko-Ladokhina, V., M.J. Cocco, and D.F. Senear, *Flexibility and adaptability in binding of E. coli cytidine repressor to different operators suggests a role in differential gene regulation*. J Mol Biol, 2006. **362**(2): p. 271-86.
11. Barbier, C.S. and S.A. Short, *Amino acid substitutions in the CytR repressor which alter its capacity to regulate gene expression*. J Bacteriol, 1992. **174**(9): p. 2881-90.
12. Moody, C.L., et al., *Multiple conformations of the cytidine repressor DNA-binding domain coalesce to one upon recognition of a specific DNA surface*. Biochemistry, 2011. **50**(31): p. 6622-32.
13. Munshi, S., et al., *Tunable order-disorder continuum in protein-DNA interactions*. Nucleic Acids Res, 2018. **46**(17): p. 8700-8709.
14. Munshi, S., et al., *Protein plasticity driven by disorder and collapse governs the heterogeneous binding of CytR to DNA*. Nucleic Acids Res, 2018. **46**(8): p. 4044-4053.
15. Munshi, S., D. Rajendran, and A.N. Naganathan, *Entropic Control of an Excited Folded-Like Conformation in a Disordered Protein Ensemble*. J Mol Biol, 2018. **430**(17): p. 2688-2694.
16. Munshi, S., et al., *Controlling Structure and Dimensions of a Disordered Protein via Mutations*. Biochemistry, 2020. **59**(2): p. 171-174.
17. Munshi, S., et al., *Engineering Order and Cooperativity in a Disordered Protein*. Biochemistry, 2019. **58**(19): p. 2389-2397.
18. Naganathan, A.N. and M. Orozco, *The conformational landscape of an intrinsically disordered DNA-binding domain of a transcription regulator*. J Phys Chem B, 2013. **117**(44): p. 13842-50.
19. Fuxreiter, M., *Fuzziness: linking regulation to protein dynamics*. Mol Biosyst, 2012. **8**(1): p. 168-77.
20. Sharma, R., et al., *Fuzzy complexes: Specific binding without complete folding*. FEBS Lett, 2015. **589**(19 Pt A): p. 2533-42.

21. Tompa, P. and M. Fuxreiter, *Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions*. Trends Biochem Sci, 2008. **33**(1): p. 2-8.
22. Barbier, C.S., S.A. Short, and D.F. Seneviratne, *Allosteric mechanism of induction of CytR-regulated gene expression. CytR repressor-cytidine interaction*. J Biol Chem, 1997. **272**(27): p. 16962-71.
23. Neidhardt, F.C., P.L. Bloch, and D.F. Smith, *Culture medium for enterobacteria*. J Bacteriol, 1974. **119**(3): p. 736-47.
24. Moody, C.L., et al., *Multiple Conformations of the Cytidine Repressor DNA-Binding Domain Coalesce to One upon Recognition of a Specific DNA Surface*. Biochemistry, 2011.
25. Gasteiger, E., et al., *The Proteomics Protocols Handbook: Protein Identification and Analysis Tools on the ExPASy Server*. 2005, Humana Press: Totowa, NJ. p. 571-609.
26. Williamson, M.P., *Using chemical shift perturbation to characterise ligand binding*. Prog Nucl Magn Reson Spectrosc, 2013. **73**: p. 1-16.
27. Bai, Y., et al., *Primary structure effects on peptide group hydrogen exchange*. Proteins, 1993. **17**(1): p. 75-86.
28. Pedersen, H. and P. Valentin-Hansen, *Protein-induced fit: the CRP activator protein changes sequence-specific DNA recognition by the CytR repressor, a highly flexible LacI member*. EMBO J, 1997. **16**(8): p. 2108-18.
29. Kastan, M.B., et al., *A mammalian cell cycle checkpoint pathway utilizing p53 and GADD45 is defective in ataxia-telangiectasia*. Cell, 1992. **71**(4): p. 587-597.
30. Kordel, J., et al., *Backbone dynamics of calcium-loaded calbindin D9k studied by two-dimensional proton-detected ¹⁵N NMR spectroscopy*. Biochemistry, 1992. **31**(20): p. 4856-4866.
31. Kay, L.E., D.A. Torchia, and A. Bax, *Backbone dynamics of proteins as studied by ¹⁵N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease*. Biochemistry, 1989. **28**(23): p. 8972-9.
32. Friedler, A., et al., *Structural distortion of p53 by the mutation R249S and its rescue by a designed peptide: implications for "mutant conformation"*. J Mol Biol, 2004. **336**(1): p. 187-96.
33. Rasquinha, J.A., et al., *Intrinsic Differences in Backbone Dynamics between Wild Type and DNA-Contact Mutants of the p53 DNA Binding Domain Revealed by Nuclear Magnetic Resonance Spectroscopy*. Biochemistry, 2017. **56**(37): p. 4962-4971.
34. Kneller, J.M., M. Lu, and C. Bracken, *An effective method for the discrimination of motional anisotropy and chemical exchange*. J Am Chem Soc, 2002. **124**(9): p. 1852-3.
35. Iwahara, J., M. Zweckstetter, and G.M. Clore, *NMR structural and kinetic characterization of a homeodomain diffusing and hopping on nonspecific DNA*. Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(41): p. 15062-15067.
36. Kalodimos, C.G., et al., *Structure and flexibility adaptation in nonspecific and specific protein-DNA complexes*. Science (New York, N.Y.), 2004. **305**(5682): p. 386-389.
37. Nagadai, A., et al., *Structural comparison of the free and DNA-bound forms of the purine repressor DNA-binding domain*. Structure, 1995. **3**(11): p. 1217-24.
38. Schumacher, M.A., et al., *Crystal structure of LacI member, PurR, bound to DNA: minor groove binding by alpha helices*. Science (New York, N.Y.), 1994. **266**(5186): p. 763-770.
39. Kalodimos, C.G., R. Boelens, and R. Kaptein, *Toward an integrated model of protein-DNA recognition as inferred from NMR studies on the Lac repressor system*. Chemical reviews, 2004. **104**(8): p. 3567.