BIOCHEMISTRY
including biophysical chemistry & molecular biology

ARTICLE

pubs.acs.org/biochemistry

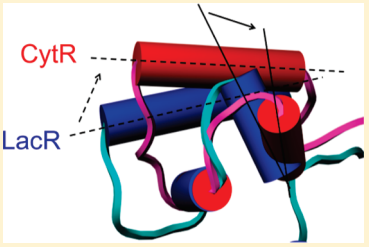
Multiple Conformations of the Cytidine Repressor DNA-Binding Domain Coalesce to One upon Recognition of a Specific DNA Surface

Colleen L. Moody,[†] Vira Tretyachenko-Ladokhina,[†] Thomas M. Laue,[‡] Donald F. Senear,[†] and Melanie J. Cocco*[,†]

[†]Department of Molecular Biology and Biochemistry, University of California, Irvine, California 92697, United States

[‡]Department of Biochemistry, University of New Hampshire, Durham, New Hampshire 03824, United States

ABSTRACT: The cytidine repressor (CytR) is a member of the LacR family of bacterial repressors with distinct functional features. The *Escherichia coli* CytR regulon comprises nine operons whose palindromic operators vary in both sequence and, most significantly, spacing between the recognition half-sites. This suggests a strong likelihood that protein folding would be coupled to DNA binding as a mechanism to accommodate the variety of different operator architectures to which CytR is targeted. Such coupling is a common feature of sequence-specific DNA-binding proteins, including the LacR family repressors; however, there are no significant structural rearrangements upon DNA binding within the three-helix DNA-binding domains (DBDs) studied to date. We used nuclear magnetic resonance (NMR) spectroscopy to characterize the CytR DBD free in solution and to determine the high-resolution structure of a CytR DBD monomer bound specifically to one DNA half-site of the uridine phosphorylase (*udp*) operator. We find that the free DBD populates multiple distinct conformations distinguished by up to four sets of NMR peaks per residue. This structural heterogeneity is previously unknown in the LacR family. These stable structures coalesce into a single, more stable *udp*-bound form that features a three-helix bundle containing a canonical helix−turn−helix motif. However, this structure differs from all other LacR family members whose structures are known with regard to the packing of the helices and consequently their relative orientations. Aspects of CytR activity are unique among repressors; we identify here structural properties that are also distinct and that might underlie the different functional properties.



In *Escherichia coli*, nine operons that encode transport proteins and enzymes involved in ribonucleotide salvage and degradation pathways are regulated by the cytidine repressor protein (CytR),[1,2] a member of the lactose repressor (LacR) family of bacterial repressors.[3] Members of this widespread family have been found in all bacterial genomes examined, with more than 1000 members identified.[4] Most of the 13 LacR family members in *E. coli* regulate genes for specific catabolic pathways in response to nutrient availability.[1,3] These regulators bind as homodimers to palindromic DNA operators.

LacR family DNA-binding domains (DBD) share a high degree of sequence similarity; e.g., the LacR and CytR DBDs are 36% identical and 68% similar. The DBDs contain canonical helix−turn−helix (HTH) motifs[5] with several shared structural features. In the four LacR family DBDs whose structures are known (LacR,[6] PurR,[7] FruR,[8] and CcpA[9]), the first helix contacts the DNA and positions the second helix in the major groove of the recognition site. The third helix docks onto helices 1 and 2 with minimal contact with the DNA. Mutational analysis has identified several amino acids within the CytR DBD and in particular within helix 2 that are essential for repressor function in vivo,[10] suggesting that CytR helix 2 recognizes DNA as it does in other LacR family DBDs.

CytR-regulated promoters are unique in that the central spacers that separate the recognition motifs vary significantly in length, ranging from 0 to 9 bp.[11] Consequently, the operators to which CytR binds differ in both the distance between and the register of the recognition site grooves. This raises the question of how the structure of CytR accommodates these variations. CytR's regulatory properties are also unique among LacR family repressors in that these rely on interactions with the cyclic AMP receptor protein (CRP). The requirement for long-range allosteric interactions raises the possibility that the DBD structure is modulated to communicate details of the operator sequence to CRP.

The structures of many other DNA-binding proteins have been defined in various bound states.[12,13] It is common for DNA-binding proteins to undergo some degree of protein folding coupled to recognition of a specific DNA sequence.[14,15] When free in solution, DNA-binding proteins show large variations in structure; for example, the LacR DBD is well-ordered, whereas DtxR[16] and LEF-1[17] are molten and relatively unstructured. In the case of LacR, the structure of the HTH DNA-binding domain remains largely intact whether it is bound either specifically or nonspecifically to DNA or is free.[18] A fourth helix in LacR, called the hinge helix, does fold completely only upon binding to a full operator. In many cases, a linker or capping helix is the element that folds upon binding DNA; this has been termed an "inducible

Received: February 9, 2011
Revised: June 16, 2011
Published: June 21, 2011

snap-lock" mechanism of DNA recognition.[15] Understanding how protein folding events are coupled to efficient binding to a specific DNA sequence is essential for building a complete model of their mechanism and details of their function in the cell. A reasonable starting point for addressing these questions is to compare the structures of bound and free states.

We have used nuclear magnetic resonance (NMR) spectroscopy to investigate the structure and properties of the CytR DNA-binding domain both free in solution and bound to a recognition site-containing oligonucleotide. In the bound state, the CytR DBD contains a HTH motif common among other proteins that interact with DNA; however, it is distinct when compared to the conformations of the other four LacR DBDs whose structures are known.[10] We find another notable difference in that the CytR DBD free state populates multiple stable conformations. Thus, operator−DNA binding is accompanied by a disorder to order transition in which the disordered state is an ensemble of distinct conformations.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification. The 67 N-terminal residues of CytR that comprise the DNA-binding domain (MKAKKQETAA TMKDVALKAK VSTATVSRAL MNPDKVSQAT RNRVEKAARE VGYLPQPMGR NVKRNES) were expressed in *E. coli* strain BL21(DE3)pLysS transformed with expression plasmid pSS584DBD. This contains the coding sequence cloned into the NdeI and BamHI sites of pSS584.[19] ^{15}N-labeled protein and ^{13}C- and ^{15}N-labeled protein were prepared by growing the cells in Neidhardt's minimal medium[20] with kanamycin and chloramphenicol containing [^{15}N]ammonium chloride as the sole nitrogen source and either unenriched glucose or [^{13}C]glucose as the sole carbon source with expression induced by isopropyl β-D-thiogalactopyranoside. Cells were lysed by sonication in 50 mM sodium phosphate, 1 mM EDTA (pH 7.7), and 1 mM Pefabloc, with Complete, mini EDTA protease inhibitor tablets (Roche Applied Science, Indianapolis, IN) added. Soluble extract was prepared by high-speed centrifugation (43000 rcf) for 45 min at 4 °C. The soluble fraction resulting from precipitation with 60% ammonium sulfate was dialyzed (molecular weight cutoff of 3500) into 50 mM sodium phosphate, 30 mM NaCl, and 1 mM EDTA (pH 7.6). CytR DBD was isolated by chromatography on a MonoS column (GE Healthcare) that was eluted using a NaCl gradient. The CytR DBD eluted as the major peak at ~0.5 M NaCl. Peak fractions were pooled, dialyzed into 50 mM sodium phosphate, 30 mM NaCl, and 1 mM EDTA (pH 6), and concentrated. The purity was above 95% as assessed by sodium dodecyl sulfate−polyacrylamide gel electrophoresis and MALDI mass spectrometry. The concentration was determined using an extinction coefficient (A_{280}) of 0.202 L g^{-1} cm^{-1} calculated from the amino acid composition.[21] All NMR experiments were performed with samples in a 90% H_2O/10% D_2O mixture in 50 mM sodium phosphate, 30 mM NaCl, and 1 mM EDTA (pH 6). Samples for structure determination contained 1 mM [^{15}N,^{13}C]CytR DBD with 1.2 mM natural abundance *udp* half-site DNA (5′-ATTTATGCAACGCA-3′) or 0.4 mM [^{15}N]CytR DBD and 0.6 mM natural abundance *udp* half-site DNA. The free DBD samples contained 1 mM [^{15}N]CytR DBD or 1 mM [^{15}N,^{13}C]CytR DBD.

Circular Dichroism (CD). CD spectra of CytR in the far-UV range (190−240 nm) were recorded at 1 nm wavelength intervals using a Jasco model 715 CD spectrometer. CytR DBD was dialyzed into 10 mM sodium phosphate buffer (pH 6.0), containing 0.03 M Na_2SO_4. The temperature was controlled with a Peltier device. Measurements in the far-UV range, from 260 to 190 nm, were taken using a DBD concentration of 0.26 g/L (3.4 μM) using a 2 mm path length cell. Spectra were recorded with 1 nm intervals as an average of 25 scans at a scan rate of 50 nm/min. Buffer spectra alone were subtracted using Spectra Manager. No smoothing of the spectra was used at any point. The background-subtracted far-UV CD spectrum of CytR was used for secondary structure analysis using CDSSTR as implemented in CDPro[22] optimized for 190−240 nm reference protein data sets SP43 and SDP48. Results of these six fits were averaged to yield the predicted composition reported.

Thermal melting was monitored at 222 nm. Data were collected as the temperature was first increased from 5 to 70 °C at a rate of 1 °C/min and then decreased to 5 °C. Faster rates led to unacceptable hysteresis in heating and cooling phases. Averaged values from heating and cooling phases were plotted as mean residue ellipticity θ_N (degrees square centimeter per decimole per residue) versus absolute temperature (Figure 2A). Nonlinear least-squares analysis with Origin was used to fit these data to the equations

$$\theta_{obs} = [\theta_N + m_N(T - T_m)]f_N + [\theta_D + m_D(T - T_m)](1 - f_N) \quad (1a)$$

$$f_N = 1/\{1 + exp[(-\Delta H/R)(1/T - 1/T_m)]\} \quad (1b)$$

where θ_{obs} is the observed ellipticity, θ_N and θ_D are ellipticities of native and denatured protein, respectively, at the melting temperature, T_m, m_N and m_D are temperature-dependent slopes for native and denatured ellipticities, respectively, f_N is the fraction of native protein at temperature T, and ΔH is the enthalpy of unfolding.

Analytical Ultracentrifugation (AUC). High-speed sedimentation equilibrium and sedimentation velocity experiments were conducted in a Beckman XL-A analytical ultracentrifuge using interference optics. ^{15}N-labeled DBD prepared for NMR structural analysis was analyzed in NMR buffer [50 mM sodium phosphate, 30 mM NaCl, and 1 mM EDTA (pH 6.00)]. To mimic the high concentrations used in the NMR experiments, DBD was loaded into four-channel, sector-shaped cells at concentrations of 5, 2.5, 1.25, and 0.63 g/L. Interference data were acquired after sedimentation to equilibrium in an AN-60 Ti rotor at 20K, 30K, and 40K rpm and 20 °C. Water blanks were subtracted from the data. The multiple channels of data obtained at different speeds and loading concentrations were analyzed globally according to models accounting for only a single species, both with and without nonideality, or two species in a self-association equilibrium[23] using Nonlin.[24] Sedimentation velocity was determined at a DBD concentration of 0.63 g/L and a speed of 60K rpm. Data were analyzed using Sedfit[25] to determine the sedimentation coefficient distribution, $c(s)$. A value of the partial specific volume ($\bar{v}$ = 0.7343) was calculated from the amino acid composition[26] and the buffer density from the composition, both using SedNterp.[27]

Oligonucleotide Preparation. Deoxyribonucleotides were purchased from IDT Technologies (San Diego, CA). It was dissolved in 10 mM TRIS (pH 8) and 0.1 M NaCl at a concentration of 1 mM. Complementary single-strand oligonucleotides were annealed to double strands by mixing equimolar quantities, heating to 80−85 °C in a water bath for 10−15 min, and cooling overnight to room temperature by shutting off the water bath heating. Resulting double-stranded oligonucleotides were exchanged into

50 mM sodium phosphate (pH 6) containing 30 mM NaCl and 1 mM EDTA and concentrated to 2−5 mM using Centricon-YM3 concentrators (molecular weight cutoff of 3000 from Millipore). Oligonucleotides for binding experiments were purchased as products purified via high-performance liquid chromatography. Each double-strand sequence contained one strand with the 5′-phosphate conjugated to Alexa 532 or Oregon Green 514 fluorescent dye through a six-carbon linker. A 7% excess of unlabeled complementary oligonucleotide was used to ensure complete annealing of the dye-conjugated DNA to the double-strand sequence.

DBD−DNA Binding Affinity. Steady-state fluorescence measurements were taken using an SLM 8100 instrument 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 an 8 nm band-pass. Fluorescence intensity was recorded for 2 min at each position of the excitation and emission polarizers. Anisotropy was calculated from the average intensities according to

$$\langle r \rangle = (I_{vv} - GI_{vh})/(I_{vv} + 2GI_{vh}) \quad (2)$$

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

$$G = I_{hv}/I_{hh} \quad (3)$$

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 4 mm × 10 mm quartz cuvettes at a temperature of 20 ± 0.1 °C maintained with a Peltier device. After each addition of protein, we incubated the cuvette in the thermoregulated cell holder for 10 min prior to taking intensity measurements to ensure that equilibrium had been attained. Anisotropy values were analyzed as a function of DBD concentration using nonlinear least-squares analysis in Origin 7 according to

$$\langle r \rangle = (r_{bound} - r_{free})[exp(2.303\ ln[DBD]) - \Delta G^\circ/RT]/[1 + exp(2.303\ ln[DBD]) - \Delta G^\circ/RT] + r_o \quad (4)$$

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

Hydrogen−Deuterium Exchange. Samples (pH 6) were lyophilized and then resuspended in D_2O immediately prior to collection of ^{15}N HSQC spectra. Collection of NMR data was initiated 10 min following resuspension.

NMR Structure Determination. NMR data were acquired on a Varian Inova 800 MHz spectrometer. Data processing was performed using NMRPipe,[28] and spectra were analyzed using CCPNMR Analysis.[29] With the exception of residual dipolar coupling (RDC) measurements, all spectra used for structure determination were recorded at 35 °C. Backbone assignments were performed using ^{15}N HSQC,[30] HNCACB,[30,31] and CBCA-(CO)NH[30,31] spectra manually with the aid of MARS.[32,33] Mixing times (τ_m) in NOE experiments range from 50 to 200 ms. Side chain assignments and NOE (nuclear Overhauser effect) restraints were derived from ^{15}N NOESY-HSQC (τ_m = 150 ms),[34] ^{15}N TOCSY-HSQC (τ_m = 50 ms),[34] ^{13}C NOESY-HSQC (τ_m = 100 ms), and HCCH-TOCSY spectra.[35] Backbone dihedral angles were predicted using TALOS from NH, CA, and HA chemical shift information.[36] φ angles were also calculated using J couplings from the HNHA experiment.[37] RDC measurements for the backbone NH bonds were obtained using the IPAP (Inphase Antiphase) experiment.[38] Aligned data were collected with 0.4 mM DBD and 0.6 mM *udp* half-site DNA in 4% C12E5 PEG/hexanol (r = 0.85−1) bicelles at 25 °C as described previously.[39,40] The sample was allowed to incubate for 4 days at room temperature, at which point alignment was confirmed via a D_2O peak splitting of 14 Hz.

Structures were calculated using Xplor-NIH, version 2.20,[41,42] simulated annealing protocols incorporating NOEs, dihedral angles, and RDCs. Anneal scripts were similar to those described in ref 43 with the addition of an internal variable module for dynamics minimizations using torsion angle space for incorporating the RDC data. RDCs were incorporated with direct input of per residue changes in coupling (hertz). Initital Da and rhombicity values were calculated using Xplor-NIH.[44] The Da and rhombicity values were then allowed to optimize during the simulated annealing. NOE potential energy calculations were conducted using the sum averaging method. Initially, an extended structure generated with random φ and ψ angles that were energy minimized (Powell minimization, 1000 steps) was used as the starting structure. Simulated annealing consisted of three steps: 20 ps high-temperature dynamics at 3000 K with experimental restraints but with no van der Waals term, a 100-step cycle or 0.2 ps with van der Waals turned on (radius and force constants incrementally increased; force constants for angles and improper dihedral angles adjusted), and a gradual cooling to 25 K with a time step of 12.5 ps (all force constants increased to final values). The structure generated was subjected to 500 steps of energy minimization using the conjugate gradient Powell algorithm. Several rounds of calculations were performed, with each round generating 50−100 structures. After each cycle, the structure with the lowest energy was used at the beginning structure for the subsequent calculation. Refinement was performed similarly with a less drastic high-temperature dynamics step. The 10 lowest-energy structures of a run of 100 are presented here (Figure 3; statistics in Table 1). The structures and constrains have been deposited in the Protein Data Bank (PDB) (entry 2L8N). The chemical shift assignments have been deposited in the Biological Magnetic Resonance Data Bank (BMRB) (accession number 17419). An additional set of structures was further calculated using the same constraints and refinement script with the torsion angle database (RAMA) included. This ensemble of structures can be viewed as PDB entry 2LCV and BMRB entry 17634. Including the torsion angle database results in a slightly improved rmsd, but also a slight increase in the number of violations and an improvement in the relationship to experimental values (see Table 1).

Structures were visualized using UCSF Chimera.[45] Contact maps were generated using MOLMOL.[46] Where possible, sequential assignments of the CytR DBD free state were made on the basis of HNCACB/CBCA(CO)NH and NOESY (τ_m = 150 ms)/TOCSY (τ_m = 75 ms) data.

Definition of exposed residues in the CytR DBD free state was determined using addition of gadodiamide (OmniScan, Nycomed, Melville, NY) followed by collection of the two-dimensional

Table 1. Structural Statistics for the Ensembles of the 10 Lowest-Energy CytR DBD Structures Calculated without or with the Xplor-NIH Torsion Angle Database

total no. of NOE restraints (for 47 amino acids)		555
intraresidue		232
sequential		134
medium-range		148
long-range		41
no. of other restraints		113
RDC		36
dihedrals		77
	28LN	2LCV[c]
restraint violations (average)		
NOE violations >0.3 Å	0	0
dihedral violations >5°	0	0.1 ± 0.3
mean rmsd from experimental restraints		
NOE (Å)	0.032 ± 0.001	0.043 ± 0.003
dihedral (deg)	0.645 ± 0.048	0.867 ± 0.166
RDC (Hz)	0.093 ± 0.012	0.138 ± 0.035
mean rmsd from idealized		
covalent geometry		
bonds (Å)	0.003 ± 0.000	0.002 ± 0.001
angles (deg)	0.440 ± 0.006	0.417 ± 0.019
impropers (deg)	0.301 ± 0.008	0.400 ± 0.031
RDC Statistics		
R-factor[a] (%)	0.89 ± 0.15	1.33 ± 0.39
Da[b] (Hz)	−7.45 ± 0.40/ 7.14 ± 0.53	−7.30 ± 0.29/ 7.25 ± 0.46
rhombicity	0.62 ± 0.06	0.63 ± 0.04
backbone rmsd (Å)		
relative to average coordinates [9−55 (N, Cα, C)] (Å)	1.07	0.96
Ramachandran plot (%)		
residues in the most favored region	83.5	84.2
residues in additional allowed region	14.4	14.2
residues in generously allowed region	2.1	1.6
residues in disallowed region	0.0	0.0

[a] Equation 3 from ref 57. [b] Because the rhombicity is close to 0.66 (2/3), the sign of Da is indeterminate. Within each group of 10 lowest-energy structures, approximately half (4 of 10 in 2L8N and 5 of 10 in 2LCV) of the values were negative and the remaining were similar in magnitude, but opposite in sign. Two Da values are shown for each structure averaged from either the positive or negative group. [c] Structure calculation included Xplor-NIH torsion angle database.

(2D) NOESY spectrum to determine which peaks were affected by the paramagnetic reagent.[47] A sample of free CytR DBD was lyophilized and resuspended in D_2O and 4.1 mM gadodiamide.

Maximal exchange rates for conformational changes were calculated using the following equation

$$1/k_r = \tau_A^c = \sqrt{2}/\pi(\nu_A - \nu_X) \quad (5)$$

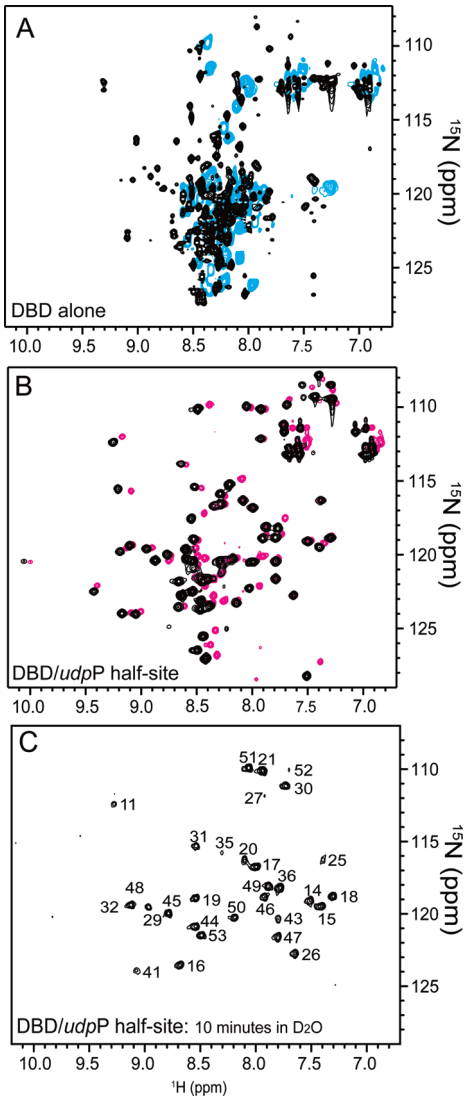


Figure 1. $^{15}N-^{1}H$ HSQC spectra of the CytR DBD (A) alone at 20 (black) and 35 °C (cyan) and (B) bound to *udp* half-site DNA at 20 (black) and 35 °C (magenta). Sixty-three peaks are expected from the backbone amide protons. The DNA:protein molar ratio was 1.5. (C) $^{15}N-^{1}H$ HSQC spectra of the CytR DBD bound to *udp* half-site DNA. Amide signals were protected from exchange after the lyophilized sample was resuspended in D_2O for 10 min. When a similar free-state DBD sample was resuspended in D_2O, no peaks were detected within 10 min.

where k_r is the rate of exchange at coalescence of the multiple peaks displayed, τ_A^c is the lifetime of the nucleus at coalescence, and ν_A and ν_X are the frequencies in hertz for the two states.[48]

RESULTS

We began our studies of the DBD free state and first assessed the structure and stability of the protein by collecting ^{15}N HSQC NMR spectra over a range of temperatures. As the temperature is increased, the dispersion of the peaks is reduced, indicating that the protein unfolds until only one set of random coil-like signals is obtained at 35 °C (Figure 1A). Notably, below the melting transition, the ^{15}N HSQC spectrum reveals more amide signals than expected from the 63 non-proline amino acids. Specifically, at least 87 distinct peaks are identified.

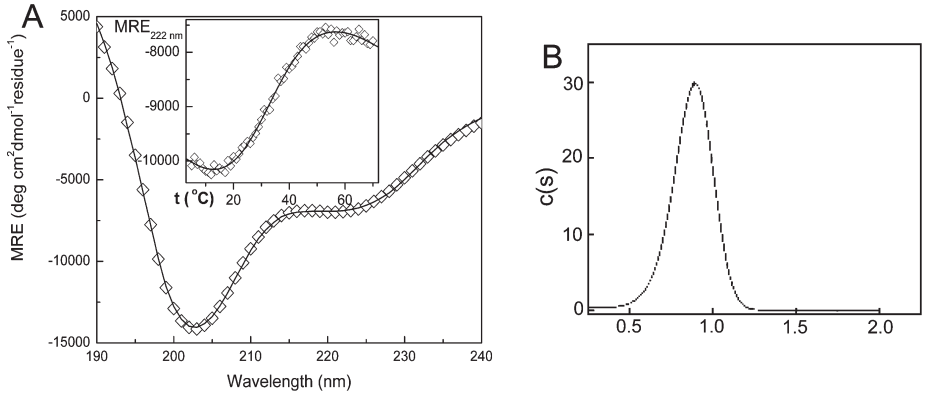


Figure 2. CD and AUC of free CytR DBD. (A) Circular dichroism spectra of free CytR DBD at 20 °C. The line represents the best fit using CDSSTR (CDPro).[22] The inset depicts the melting curve at 222 nm. (B) Sedimentation velocity indicates a single monomeric species with a sedimentation coefficient of 0.89 S.

The structure and stability of the free-state DBD were probed further using CD spectroscopy. The far-UV CD spectrum (Figure 2A) indicates primarily α-helix and random coil. The secondary structure composition was estimated using CDSSTR as implemented in CDPro.[22] The reconstructed spectrum that gave the best fit (Figure 2A) predicts a helical content of 22% at 20 °C, with the remainder predicted to be random coil. This is equivalent to only 15 residues in a helical conformation, slightly more than half the number of helical residues that we find in the bound form (see below). Thermal unfolding of this residual structure was induced when the solution was heated from 5 to 70 °C and monitored by the ellipticity at 222 nm (inset of Figure 2A). A clear transition was observed between sloping baselines, with a midpoint near 35 °C. The increase in ellipticity indicates the loss of helical structure. The transition was reversed and the initial ellipticity recovered upon cooling, suggesting a reversible two-state transition. Accordingly, the data were analyzed using the van't Hoff equation to obtain the melting temperature (T_m) and the enthalpy and entropy of unfolding. This yielded the following values: T_m = 33.3 ± 2.6 °C, $\Delta H°$ = 20.2 ± 4.7 kcal/mol, and $\Delta S°$ = 66 ± 16 cal mol^{-1} K^{-1}. The enthalpy and entropy changes are 2−3-fold lower than might be expected for a protein of this size.[49] This is consistent with the low degree of order suggested by the CD and NMR analyses.

Many of the extra peaks in the ^{15}N HSQC spectrum occur as doubled signals, suggesting that the DBD alone may populate multiple conformations at lower temperatures (i.e., 20 °C). MALDI mass spectrometry confirmed a single-length chain of the correct molecular weight; thus, multiple signals are not a consequence of degradation. One possible explanation for multiple signals in the free state would be multiple oligomeric states. We performed AUC (interference optics) on an NMR sample to assess the possibility of multiple species as this might provide an explanation for multiple spin systems. The sedimentation coefficient distribution obtained from sedimentation velocity analysis of DBD is shown in Figure 2B. This yields a sharp and slightly skewed peak consistent with a single, thermodynamically nonideal sedimenting component with a sedimentation coefficient of 0.89 S. Compared to the value for the equivalent sphere of the monomoeric protein (1.33 S), this suggests that the free DBD is monomeric with an extended structure.

The sedimentation equilibrium data (not shown) are also consistent with DBD as a single, thermodynamically nonideal component. The highest concentrations observed at the base of the cell were approximately 20 g/L, yet even with such a high concentration, attempts to fit for self-association did not yield a lower limit to K_a for association, even to dimer. In contrast, when nonideality of a single species was considered, the rmsd decreased by a factor of 2 and a well-bound value of the second virial coefficient was obtained [B = (1.20 ± 0.21) × 10^{-3} mol mL g^{-2}]. Both the improvement of the fit and the tight confidence intervals suggest a reasonably accurate estimate of the nonideality, consistent with the velocity data given above.

AUC analysis yields an apparent molecular mass of 9000 ± 310 Da if the partial specific volume calculated on the basis of amino acid composition ($\bar{\nu}$ = 0.7434 mL/g) is used. The apparent molecular mass is significantly greater than the predicted molecular mass of the monomeric ^{15}N-containing protein, 7489.6 Da. However, the calculated partial specific volume does not account for either ^{15}N (which increases protein density) or the solvent electrostriction and the strong likelihood of territorially bound anions (Cl^- or PO_4^{2-}) due to the very high net charge (+9) of the protein. Overestimation of $\bar{\nu}$ by 0.03−0.04 mL/g would account for the difference between the fitted and monomeric molecular masses. The extra density contributed by ^{15}N would account for approximately one-third of this. Taken together, the data suggest that the free DBD is strictly monomeric and that self-interactions are dominated by repulsive nonideality.

We also characterized the structure and stability of the CytR DBD bound to an operator half-site. CytR operators are comprised of inverted repeats of an 8 bp recognition motif separated by variable spacing ranging from 0 to 9 bp.[11,50] The experiments described here use the natural operator left half-site of *udp* (5′-ATTTATGCAACGCA-3′); this was chosen because it is a better match to the consensus recognition motif defined by SELEX (Systematic Evolution of Ligands by EXponential enrichment)[11] than the right half-site. In the presence of the *udp* half-site DNA, we find a single set of well-dispersed DBD peaks in the NMR spectrum. This peak pattern exists at 20 °C and remains above 35 °C (Figure 1B). The structural consequences of DNA binding are reversible for DNA substrates; removal of DNA results in the return of the original DBD HSQC spectrum that reflects multiple protein conformations. We used the fluorescence anisotropy of a dye-conjugated (Alexa 532) DNA substrate to monitor the binding of the CytR DBD to DNA. Results of titration experiments (Figure 3) indicate that the CytR DBD binds the *udp* operator left half-site. We analyzed titration data assuming a 1:1

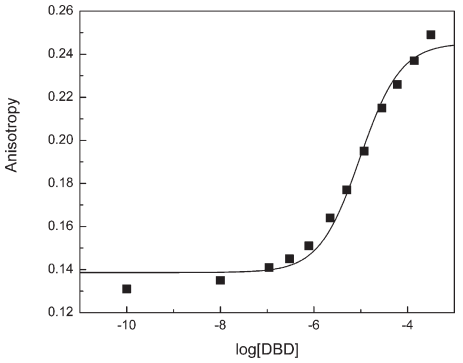


Figure 3. DBD−DNA binding affinity. Binding of CytR DBD to the *udp* left half-site at 20 °C and pH 6, as monitored by the steady-state anisotropy of the dye-conjugated oligonucleotide. The solid line indicates the fitted curve for *udp* half-site DNA with a k_D of 9.6 ± 4.0 μM. This has been transformed to a fraction bound scale on the ordinate using the fitted vales of r_{free} and r_{bound}.

stoichiometry between monomeric protein and DNA according to eq 4 (Experimental Procedures). These data fit well to our 1:1 binding model that yields a binding free energy change ($\Delta G°$) of −6.41 ± 0.12 kcal/mol and a corresponding equilibrium constant for dissociation (k_D) of 9.6 ± 4.0 μM. The results show that the DBD is fully bound to the *udp* half-site DNA under conditions under which NMR spectra are recorded where the concentrations exceed 0.4 mM.

Hydrogen exchange is another useful measure of protein stability. Rates of intrinsic exchange between amide protons and water are fast at pH 6.0, where we characterized the CytR DBD. Consequently, when the protein is transferred into D_2O, NMR amide signals corresponding to exposed positions will disappear entirely. In the case of the free-state DBD, no peaks were detectable in the NMR ^{15}N HSQC spectrum within 10 min of the transfer into D_2O, revealing complete amide proton exchange. In contrast, with the DBD bound to the *udp* half-site, we find that 30 peaks are present in the NMR spectrum following resuspension in D_2O (Figure 1C). These amide protons of the DBD−DNA complex exchange more slowly than those exposed to solvent, suggesting that they are involved in stable secondary structure or are buried within the protein interior or the protein−DNA interface.

Structural Analysis of the CytR DBD. Using standard ^1H, ^{15}N, and ^{13}C triple-resonance NMR experiments, we assigned the spectra of CytR DBD in the presence of the *udp* promoter half-site DNA substrate. We chose 35 °C as this temperature provided the best spectra. Chemical shift assignments were completed for amino acids 3−67 by assigning sequential connectivities derived from HNCACB and CBCA(CO)NH data. Side chains were assigned using HCCH-TOCSY data. NOEs were catalogued from ^{15}N and ^{13}C NOESY-HSQC spectra. The structure of the DBD−*udp* half-site complex was determined on the basis of 555 NOEs, 77 dihedral angles, and 36 dipolar couplings (Table 1). The family of 10 lowest-energy structures (rmsd = 1.07 Å) corresponding to amino acids 9−55 is presented here (Figure 4A). The N- and C-termini of the construct were not included in the structure calculation because of a lack of structural data. Our structure reveals a three-helix bundle consistent with other LacR family members. The CytR DBD helices vary in size [helix 1 having seven residues (12−18), helix 2 having six residues (23−28), and helix 3 having 13 residues (38−50)].

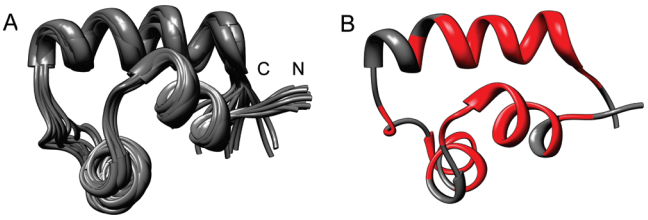


Figure 4. Ensemble of NMR structures of the CytR DBD in the presence of *udp* half-site DNA (positions resistant to exchange). (A) Solution structures of the CytR DBD in the presence of the *udp* half-site at 35 °C. An ensemble of the 10 lowest-energy structures of amino acids 9−55 is presented with an rmsd of 1.07 Å relative to the average structure (PDB entry 28LN). N and C denote the amino and carboxy directions, respectively. The structure adopts a helix−turn−helix fold as is common for DNA-binding domains. (B) Positions (colored red) stabilized by DNA binding that become resistant to hydrogen exchange mapped onto a single structure.

Figure 5. Conformational shifts in the free CytR DBD. In the free state, the CytR DBD exhibits regions of multiple distinct conformations. Colored purple are those amino acids with three or four conformations; residues colored orange show two conformations. Gray represents all other remaining residues that have been assigned where only one signal was clearly identified. Bold denotes the amino acids that are conserved throughout the LacR family. Prolines are colored red.

Helices 1 and 2 comprise a helix−turn−helix motif common to DNA-binding domains. Helix 3 docks against both helices 1 and 2, which also occurs in other LacR family DBD structures. The 30 positions resistant to exchange in the CytR DBD−DNA complex (Figure 1C) are distributed within the three helices (Figure 4B). Unfortunately, NOEs between the protein and DNA were not present at any mixing time tested. Our construct does include residues corresponding to the fourth (hinge) helix; however, this region is disordered on the basis of random coil chemical shifts, a lack of NOE data, and immediate backbone hydrogen exchange. Notably, LacR forms the hinge helix only when bound to a symmetric full-site DNA operator.[51]

Although the NMR data for the free state are complicated by the presence of more than one set of signals, we were able to characterize a substantial portion of the CytR DBD in the absence of DNA. Assignment of peaks in free-state NMR spectra was accomplished by a combination of three-dimensional NOESY and HNCACB/CBCA(CO)NH analyses and side chain identification based on TOCSY data. We were not able to make complete assignments, but sequential connectivities were sufficient to define peaks corresponding to 26 residues (Figure 5). These signals provide information about most of the protein; however, we are missing assignments for signals corresponding to residues of helix 2 in the bound state. We assigned at least one set of peaks for positions 8−11 at the N-terminus, positions 14−17 in helix 1, positions 21 and 22 in the first turn, positions 29−32 in the loop after helix 2, positions 38−41, 47, and 48 in helix 3, and positions 51−54, 59, and 67 at the C-terminus in the free state. In some cases, we were able to assign multiple sets of signals to a single position (highlighted in Figure 5).

The dispersion of the peaks in the ^{15}N HSQC spectrum and NOE patterns indicate that the domain alone does have

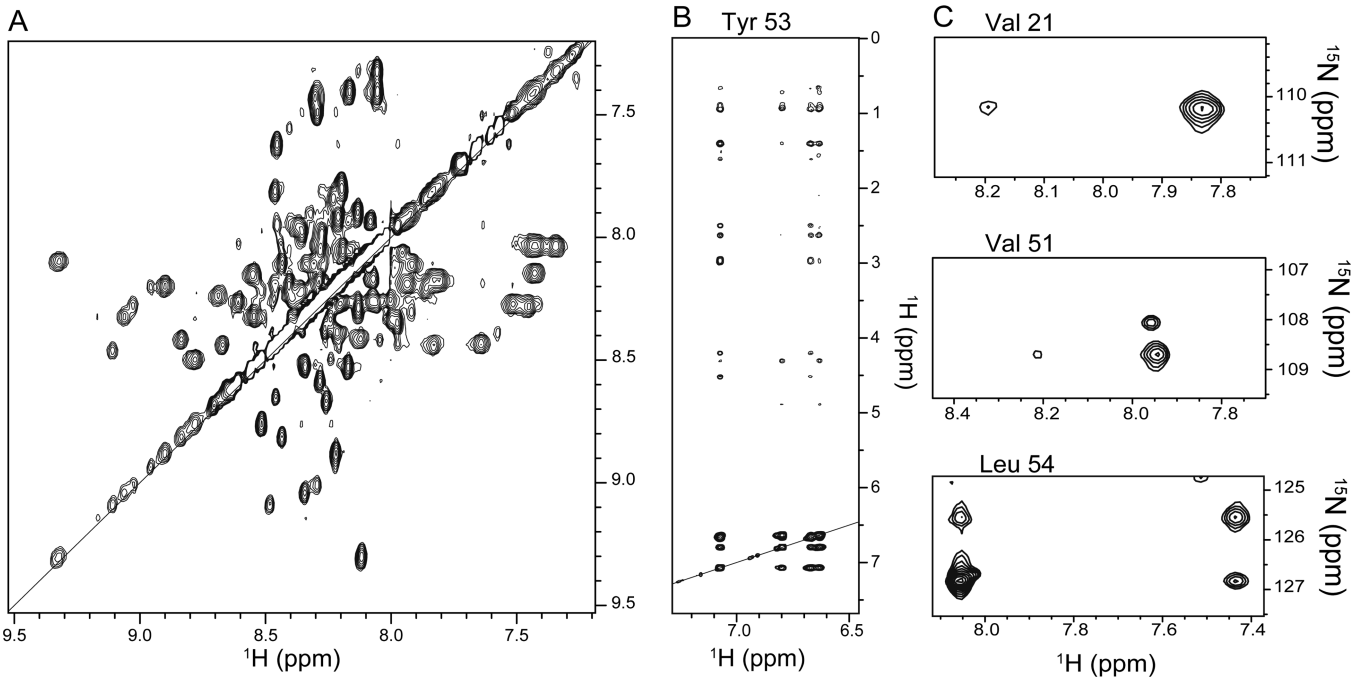


Figure 6. Spectra of free CytR DBD. (A) 2D NOESY spectrum, amide region, depicting the large number of backbone amide−amide contacts consistent with helical conformations. (B) 2D NOESY spectrum in D_2O, aromatic region, showing two distinct sets of resonances for Tyr53, the only aromatic side chain in the construct. (C) Regions of the ^{15}N HSQC spectrum demonstrating multiple signals for specific residues, Val 21, Val 51, and Leu 54.

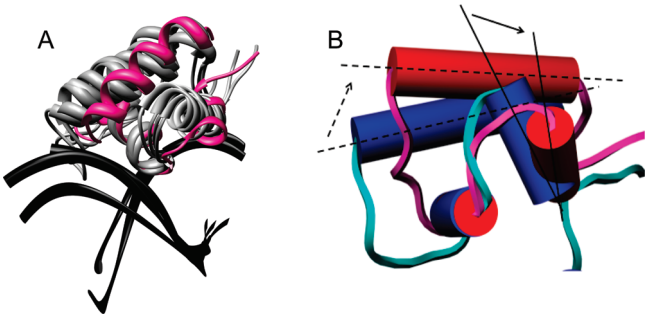


Figure 7. Model of CytR DBD-bound DNA and comparison to other LacR family members. (A) Representative structure of CytR DBD (PDB entry 28LN) (magenta) aligned along helix 2 with DBDs of other members of the LacR family: LacR (PDB entry 1CJG), PurR (PDB entry 1PNR), FruR (PDB entry 1UXC), and CcpA (PDB entry 1RZR) (gray). DNA from the LacR, PurR, and CcpA structures is present as well (black). The CytR DBD structure is different from that of the family in that a relative shift occurs between helix 2 and helices 1 and 3. (B) Helical shift in CytR DBD compared to LacR DBD. As seen in panel A, the other LacR family DBD structures maintain similar helical orientations. Here the LacR DBD (blue and cyan) helices are depicted as cylinders for comparison of the CytR DBD structure (red and magenta). With the structures aligned at helix 2, it is evident that the flanking ends of both helices 1 and 3 move outward.

structural features; this is supported by 59 backbone NOE patterns consistent with turns and helical conformations (for example, amide−amide NOEs in Figure 6A). CD data suggest that roughly 15 residues contribute to the helical signal. Of the NMR peaks assigned, NOE patterns provide evidence of helical conformations within both helices 1 and 3. Although we are lacking assigned signals for helix 2, unfolding of this helix could account for the decrease in helical content predicted by CD when the protein is not bound to DNA. Alternatively, the free-state structure we see by NMR could represent one folded population, whereas the CD signal represents the average ensemble, including unfolded populations whose signals are unassigned in our NMR data.

As revealed by the number of peaks the ^{15}N HSQC spectrum, there are multiple signals for individual positions throughout the sequence, implying multiple conformations. For example, a 2D TOCSY spectrum in D_2O of the aromatic region reveals two sets of peaks for the single tyrosine[53] at the end of helix 3 (Figure 6B). As the tyrosine side chain is not buried or restricted by packing, it is unlikely that the ring is in slow exchange. A 2D NOESY spectrum in the presence of the paramagnetic gadodiamide reagent showed that the ring is exposed to solvent under the experimental conditions. Interestingly, these sets of ring peaks have distinct NOE patterns consistent with multiple side chain rotomer conformations. Analysis of the HNCACB spectrum indicates that there are at least 20 spin systems with C_β peaks in the alanine range, double the number of alanine signals expected from the DBD sequence. Of the seven alanines we assigned, we found one set of signals for each spin system. The 13 remaining unassigned alanine spin systems could include (1) a second conformation of those we assigned, (2) multiple conformations for the three unassigned alanines, or (3) a combination of systems 1 and 2. Finally, the 2D TOCSY spectrum indicates only nine sets of peaks in the range for the seven glutamine and asparagine side chain signals. Thus, when the DBD is free in solution, it does not conform to a single structure. Our analysis reveals that the multiple conformations are most apparent when considering backbone signals. In contrast to the backbone amide signals, as the side chain length increases, the side chain signals become degenerate and appear to represent one species. ^{15}N HSQC peaks assigned to the same backbone amide proton vary as much as 0.617 ppm in ^{1}H and 1.25 ppm in ^{15}N [Leu 54 (Figure 6C)], indicating a rate of

dx.doi.org/10.1021/bi200205v |*Biochemistry* 2011, 50, 6622–6632

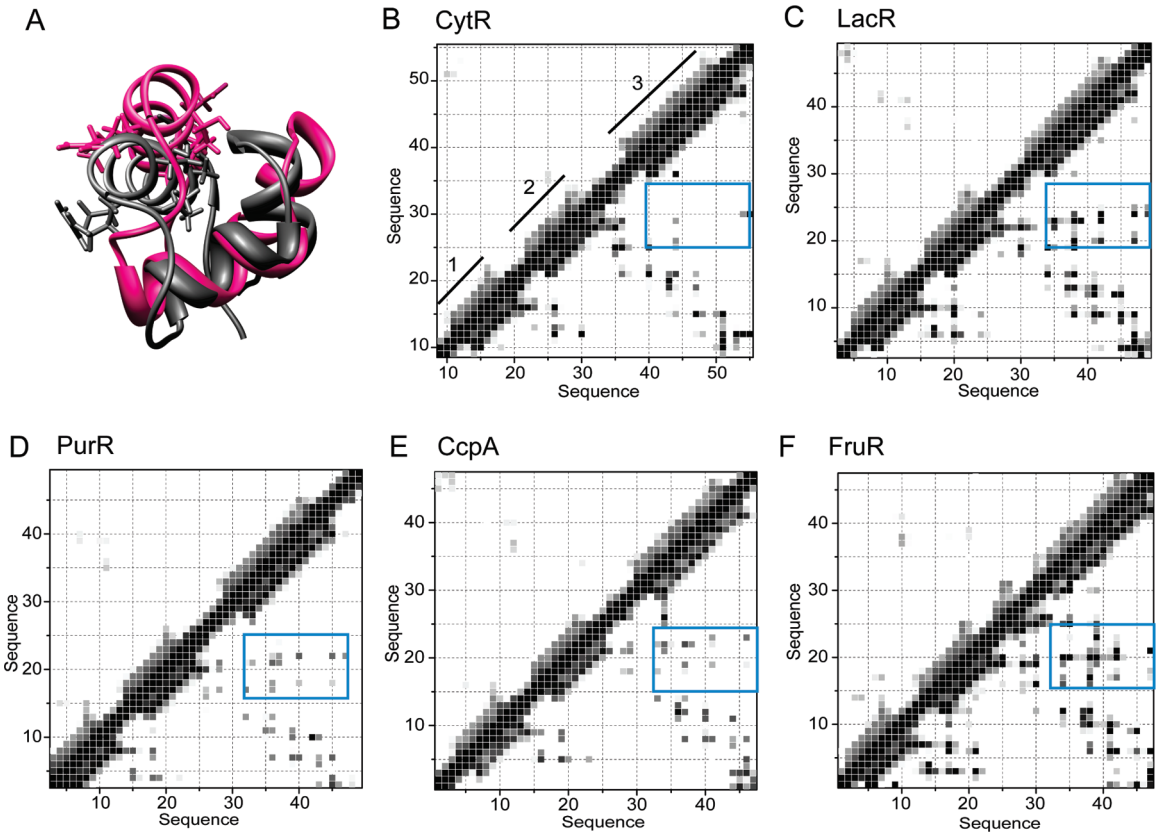


Figure 8. Comparison of the CytR DBD to other LacR family DBD structures. (A) Representative structure of CytR DBD (PDB entry 28LN) (magenta) aligned along helix 2 with the LacR DBD. Side chains of conserved residues are depicted for helix 3 to indicate a translation, but not a rotation, of the helix occurs. (B−F) Contact maps of CytR (PDB entry 2L8N, lowest energy), LacR (PDB entry 1CJG, lowest energy), PurR (PDB entry 1PNR), CcpA (PDB entry 1RZR), and FruR (PDB entry 1UXC). Contacts between 2 and 6 Å are indicated (fading from black at <2 Å to white at >6 Å). Above the diagonal, backbone−side chain contacts are indicated. Below the diagonal, side chain−side chain contacts are shown. Helices 1−3 are indicated in panel B. The region of protein depicted in each was chosen on the basis of alignment of the three helices. The blue boxes highlight the interactions between helix 3 and the end of helix 2; there is a notable lack of these contacts in the CytR DBD structure as compared to those of the other LacR family members.

conformational exchange slower than 0.0045 and 0.022 s^{-1} (eq 5). ^{15}N HSQC spectra were recorded as a function of temperature. We find that the protein unfolds before coalescence can be achieved, indicating that the conformational exchange is slow compared to the rate of unfolding.

DISCUSSION

Although no X-ray structure of CytR in any form has been produced to date, we have been able to use NMR spectroscopy to determine the structure of the CytR DBD bound to the *udp* half-site and characterize the free state. When associated with DNA, the DBD is a three-helix bundle containing a canonical helix−turn−helix motif. The three helices formed in the bound state (two short with six and eight amino acids each and one long with 13 amino acids) are arranged such that the middle helix is almost perpendicular to the other two that are antiparallel. Dipolar couplings define the relative angles between helices making the helix orientations more accurate in our CytR DBD structure than what would be calculated with NOE restraints alone. Hydrophobic groups such as valine and alanine form a hydrophobic core, while 15 positively charged arginine and lysines cover the surface of the DBD providing an electrostatic attraction to the negatively charged DNA backbone.

Structurally, the DNA-bound CytR DBD has features that are similar to those of other DBDs in the LacR family. When the sequences of CytR, LacR, PurR, FruR, and CcpA are aligned by conservation, the helices are all the same within two residues in length and comprise the same regions in sequence. We hypothesize that the CytR DBD makes contact with the DNA via the second helix as in other members of the LacR family and as is typical for HTH motifs (Figure 7A). These four are the only other family members for which the structures of the DBDs are known. For LacR, PurR, and CcpA, the relative positions of the DNA remain similar, indicating a common theme. However, comparison of the structures reveals that the CytR DBD differs from the DBDs of the other LacR family members in the relative orientation between helix 2 and helices 1 and 3. Helix 1 positions in LacR and CytR differ by an angle of 33°; helix 3 positions differ by an angle of 15° (Figure 7B). The average rmsd between LacR family DBDs is 2.67 Å; however, the rmsd between the LacR DBD and CytR DBD structured regions is considerably larger, 4.33 Å.

We generated contact maps for CytR and LacR family DBDs (Figure 8B−F) to illustrate differences in packing between these proteins. Notably, CytR DBD is more loosely packed than LacR; specifically, there are fewer interactions between the end of helix 2 and helix 3 in CytR. Helix 2 is shifted in the CytR DBD, opening the structure so that the end of the helix does not interact with helix 3. Residues 29 and 30 in CytR are more

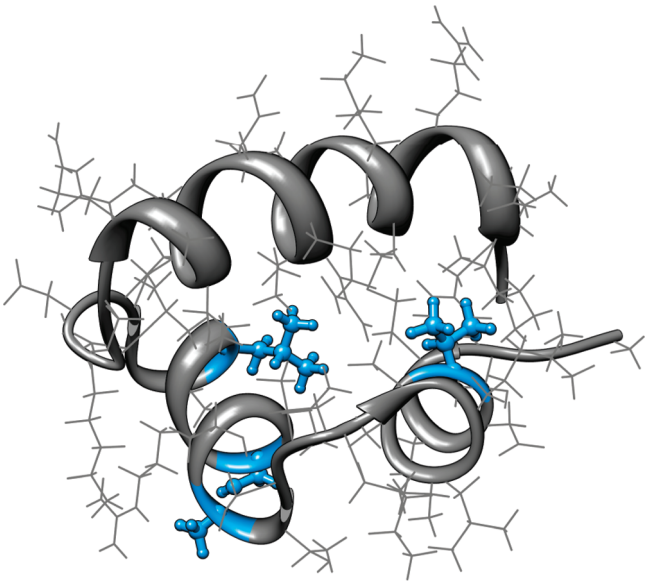


Figure 9. Activity mutants highlighted in the CytR DBD structure. The mutations (V15A, A24V, S27F, and L30S) found by Barbier et al.[10] of partially inactive CytR in vivo are depicted here in cyan ball-and-stick representation on the structure of CytR DBD bound to the *udp* half-site (PDB entry 28LN). The lowest-energy structure of the ensemble was chosen for this display. Two of the positions, 15 and 30, are involved in protein packing. The other two positions, 24 and 27, are on a solvent-exposed face of the protein and would therefore be available to interact with DNA.

exposed to solvent compared to the corresponding residues in the remainder of the LacR family structures. In addition, both the relative angles of the helices (described above) and the docking of helices within the CytR DBD are distinct when compared to those of other LacR family members in their DNA-bound forms. Helix 3 displays a significant translation, keeping the same side chains exposed albeit in different locations (Figure 8A). Thus, threading the CytR sequence through an existing LacR family coordinate file would not have produced an accurate model.

In the other LacR family structures, many of the amino acids that interact with the DNA backbone are conserved residues. Our model of the CytR DBD in which helix 2 binds in the major groove (Figure 7A) results in most of the same conserved residues interacting with the DNA backbone (K13, S22, T23, A24, R28, D34, K35, V36, and S37) despite changes in the orientation of the first and third helices. In LacR, two nonconserved residues, 17 and 18, contact the promoter sequence specifically at bases T7 and G8. Our model recapitulates this recognition feature seen in LacR with corresponding residues 23 and 24 in CytR. As mentioned earlier, mutational analysis located four DBD positions (Figure 9) that resulted in partially active forms of CytR (V15A, A24V, S27F, and L30S).[10] The first of these mutants, CytR-V15A, does not contact DNA but is buried at the interface of helices 1 and 3; mutation here could disrupt the folding of the DBD by destabilizing this interaction. The next two activity mutants, CytR-A24V and -S27F, are both located on helix 2 and are predicted to contact the DNA, although only A24V is in a position in our model to interfere directly with base-specific DNA binding. The mutation equivalent to CytR-S27F is LacR-S21F; mutation of LacR-S21F results in a more severe effect with complete inactivation of the protein. This amino acid is closer to the DNA backbone in LacR compared to our model of CytR. The last of the mutants, CytR-L30S, is the most exposed, located in the loop immediately following helix 2. However, it is oriented in such a way that it contacts the start of helix 1 through the M12 side chain and thus disruption may impact folding interactions as in CytR-V15A. Thus, the structure reveals that two of the mutations known to affect DNA binding are actually expected to affect the protein structure and are not predicted to contact DNA.

Data for determining the structure were collected at 35 °C. We found the protein alone melted at 33 °C; thus, the CytR DBD is stabilized by binding to the *udp* half-site. Specific residues stabilized by DNA binding are revealed as those amides found to be resistant to hydrogen exchange (Figure 4B). Specifically, DNA binding stabilizes all three helices and some loop structure compared to the free state, which exchanges immediately despite the helical implications of the CD and NMR data.

The free-state CytR DBD structure is strikingly different from that of the bound state, including regions with as many as four distinct conformations that are not a consequence of oligomerization. On the basis of a count of NMR peaks, two conformers are present in the free state in the regions containing helices 1 and 2 in the bound state; toward the end of helix 3, three to four conformations are seen. These folds are moderately stable at 20 °C but do not confer protection from exchange with D_2O. The structures of the unbound forms are significantly different when the CytR DBD is compared to others in the LacR family. In the case of both LacR and CcpA DBDs, the helices and their relative positions are maintained regardless of state (bound or free) with only slight changes in side chain conformations.[52,53] The PurR DBD structure varies slightly between bound and unbound forms; however, structures of the bound form also contain the ligand-binding domain.[7,54] The CytR DBD is quite different from these structures, adopting a more varied set of conformations when free in solution. Although the CytR DBD is identified as a member of the LacR family by homology, the consensus amino acids are insufficient to fold the free form of the protein. It is possible that the other members of the LacR family possess nonconserved residues that enhance protein folding, and these are missing in CytR. Another explanation for the structural difference of the free states could be that the CytR sequence contains antifolding residues that imbue additional flexibility for functional reasons. For example, CytR also contains two additional proline residues (at positions 33 and 57) compared to other LacR family proteins (which have one conserved proline at position 55). Proline cis−trans isomers can be responsible for the multiple stable conformations found in the CytR DBD free state. Similar occurrences are described by Whittaker et al.[55] Although two additional prolines unique to the CytR DBD are not DNA-contact groups, they may nevertheless play a functional role by modulating the populations of conformations present in the free state. In addition, positively charged groups may also play a role in destabilizing a single fold in the free state. The CytR DBD has a net charge of +9, whereas the LacR DBD has a net charge of +3 (however, the net charges relative to those of other LacR family members do not vary as drastically; PurR DBD has a net charge of +5 and CcpA a net charge of +7). The high charge density within the CytR DBD should oppose folding until these charges are neutralized by DNA binding.

In contrast to members of the LacR family, the CytR DBD free state is actually more like that of the DNA binding protein, LEF-1, where proline 68 has been attributed to structural heterogeneity. Members of the sequence-specific subclass of high-mobility group domains (including LEF-1) have flexible free states that

coalesce into a single structure upon target interaction.[17] Similar to the findings for LEF-1, the CytR DBD exhibits a significant increase in order and stability upon binding with target DNA. In both cases, it is arguable that the free protein retains some of its secondary structure but that a single tertiary structure is a product of sequence-specific DNA interactions.

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. Spolar and Record have previously considered the balance of energies between protein folding and DNA binding and concluded that an entropic contribution drives association.[14] Härd addressed the function for the mechanism of coupled folding in recognition and suggested that in general, a lack of a preformed interface allows for rapid scanning of target surfaces in a noncommittal manner with a rapid off rate.[56] Thus, although the CytR DBD has unique structural features within the LacR family, including a disordered free state, it displays a protein folding-coupled binding that is common to other DNA-binding proteins.

CytR operons are distinct in the LacR family because the 8 bp inverted repeats are separated by 0−9 bp. This corresponds to a difference in spacing of 31 Å and a 310° rotation along B-form DNA between recognition sites. Consequently, CytR must undergo dramatic structural rearrangements to recognize these different sequences. It is possible that the CytR DBD prolines stabilize four conformations in the free state, each unique for a specific architectural target; however, this would depend on the DBD accessing these conformations immediately before target binding potentially after nonspecific binding to the DNA had already occurred. Changes in the DBD structure may not be sufficiently dramatic to completely accommodate the changes in spacing; however, they could occur in conjunction with changes in the ligand-binding domain and ultimately communicate to CRP.

Accession Codes

Protein Data Bank entries 2L8N and 2LCV. Biological Magnetic Resonance Data Bank entries 17419 and 17634.

AUTHOR INFORMATION

Corresponding Author

*Phone: (949) 824-4487. Fax: (949) 824-8551. E-mail: mcocco@uci.edu.

Funding Sources

This work was supported by the funds from the National Science Foundation (MCB 02115769 and MCB 06652875).

ACKNOWLEDGMENT

We thank Evgeny Fadeev and Sheeja Vasudevan for help with the NMR spectroscopy and implementation of the MARS program, as well as for insightful discussions regarding structure determination.

ABBREVIATIONS

CytR, cytidine repressor; LacR, lactose repressor protein; DBD, DNA-binding domain; *udp*, uridine phosphorylase operator; HTH, helix−turn−helix; CRP, cyclic AMP receptor protein; HSQC, heteronuclear single-quantum coherence; CD, circular dichroism; T_m, melting temperature; MALDI, matrix-assisted laser desorption ionization; AUC, analytical ultracentrifugation; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; rmsd, root-mean-square deviation; TOCSY, total correlation spectroscopy; RDC, residual dipolar coupling; IPAP, Inphase Antiphase; SELEX, systematic evolution of ligands by exponential enrichment.

REFERENCES

(1) Keseler, I. M., Bonavides-Martinez, C., Collado-Vides, J., Gama-Castro, S., Gunsalus, R. P., Johnson, D. A., Krummenacker, M., Nolan, L. M., Paley, S., Paulsen, I. T., Peralta-Gil, M., Santos-Zavaleta, A., Shearer, A. G., and Karp, P. D. (2009) EcoCyc: A comprehensive view of *Escherichia coli* biology. *Nucleic Acids Res. 37*, D464–D470.

(2) Kristensen, H. H., Valentin-Hansen, P., and Sogaard-Andersen, L. (1997) 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. 266*, 866–876.

(3) Weickert, M. J., and Adhya, S. (1992) A family of bacterial regulators homologous to Gal and Lac repressors. *J. Biol. Chem. 267*, 15869–15874.

(4) Marchler-Bauer, A., Anderson, J. B., Chitsaz, F., Derbyshire, M. K., DeWeese-Scott, C., Fong, J. H., Geer, L. Y., Geer, R. C., Gonzales, N. R., Gwadz, M., He, S., Hurwitz, D. I., Jackson, J. D., Ke, Z., Lanczycki, C. J., Liebert, C. A., Liu, C., Lu, F., Lu, S., Marchler, G. H., Mullokandov, M., Song, J. S., Tasneem, A., Thanki, N., Yamashita, R. A., Zhang, D., Zhang, N., and Bryant, S. H. (2009) CDD: Specific functional annotation with the Conserved Domain Database. *Nucleic Acids Res. 37*, D205–D210.

(5) Aravind, L., Anantharaman, V., Balaji, S., Babu, M. M., and Iyer, L. M. (2005) The many faces of the helix-turn-helix domain: Transcription regulation and beyond. *FEMS Microbiol. Rev. 29*, 231–262.

(6) Chuprina, V. P., Rullmann, J. A., Lamerichs, R. M., van Boom, J. H., Boelens, R., and Kaptein, R. (1993) Structure of the complex of lac repressor headpiece and an 11 base-pair half-operator determined by nuclear magnetic resonance spectroscopy and restrained molecular dynamics. *J. Mol. Biol. 234*, 446–462.

(7) Schumacher, M. A., Choi, K. Y., Zalkin, H., and Brennan, R. G. (1994) Crystal structure of LacI member, PurR, bound to DNA: Minor groove binding by α helices. *Science 266*, 763–770.

(8) Penin, F., Geourjon, C., Montserret, R., Bockmann, A., Lesage, A., Yang, Y. S., Bonod-Bidaud, C., Cortay, J. C., Negre, D., Cozzone, A. J., and Deleage, G. (1997) Three-dimensional structure of the DNA-binding domain of the fructose repressor from *Escherichia coli* by ^{1}H and ^{15}N NMR. *J. Mol. Biol. 270*, 496–510.

(9) Schumacher, M. A., Allen, G. S., Diel, M., Seidel, G., Hillen, W., and Brennan, R. G. (2004) Structural basis for allosteric control of the transcription regulator CcpA by the phosphoprotein HPr-Ser46-P. *Cell 118*, 731–741.

(10) Barbier, C. S., and Short, S. A. (1992) Amino acid substitutions in the CytR repressor which alter its capacity to regulate gene expression. *J. Bacteriol. 174*, 2881–2890.

(11) Pedersen, H., and Valentin-Hansen, P. (1997) Protein-induced fit: The CRP activator protein changes sequence-specific DNA recognition by the CytR repressor, a highly flexible LacI member. *EMBO J. 16*, 2108–2118.

(12) Rohs, R., Jin, X., West, S. M., Joshi, R., Honig, B., and Mann, R. S. (2010) Origins of specificity in protein-DNA recognition. *Annu. Rev. Biochem. 79*, 233–269.

(13) von Hippel, P. H. (2007) From "simple" DNA-protein interactions to the macromolecular machines of gene expression. *Annu. Rev. Biophys. Biomol. Struct. 36*, 79–105.

(14) Spolar, R. S., and Record, M. T., Jr. (1994) Coupling of local folding to site-specific binding of proteins to DNA. *Science 263*, 777–784.

(15) Dyson, H. J., and Wright, P. E. (2002) Coupling of folding and binding for unstructured proteins. *Curr. Opin. Struct. Biol. 12*, 54–60.

(16) Twigg, P. D., Parthasarathy, G., Guerrero, L., Logan, T. M., and Caspar, D. L. (2001) Disordered to ordered folding in the regulation of diphtheria toxin repressor activity. *Proc. Natl. Acad. Sci. U.S.A. 98*, 11259–11264.

(17) Love, J. J., Li, X., Chung, J., Dyson, H. J., and Wright, P. E. (2004) The LEF-1 high-mobility group domain undergoes a disorder-to-order transition upon formation of a complex with cognate DNA. *Biochemistry 43*, 8725–8734.

(18) Kalodimos, C. G., Boelens, R., and Kaptein, R. (2004) Toward an integrated model of protein-DNA recognition as inferred from NMR studies on the Lac repressor system. *Chem. Rev. 104*, 3567.

(19) Barbier, C. S., Short, S. A., and Senear, D. F. (1997) Allosteric mechanism of induction of CytR-regulated gene expression. Cytr repressor-cytidine interaction. *J. Biol. Chem. 272*, 16962–16971.

(20) Neidhardt, F. C., Bloch, P. L., and Smith, D. F. (1974) Culture medium for enterobacteria. *J. Bacteriol. 119*, 736–747.

(21) Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., Bairoch, A., and Walker, J. M. (2005) Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*, pp 571−607, Humana Press, Totowa, NJ.

(22) Sreerama, N., and Woody, R. W. (2000) Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem. 287*, 252–260.

(23) Laue, T. M. (1995) Sedimentation equilibrium as thermodynamic tool. *Methods Enzymol. 259*, 427–452.

(24) Johnson, M. L., Correia, J. J., Yphantis, D. A., and Halvorson, H. R. (1981) Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques. *Biophys. J. 36*, 575–588.

(25) Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J. 78*, 1606–1619.

(26) Cohn, E. J., Edsall, J. T., et al. (1943) *Proteins, Amino Acids and Peptides as Ions and Dipolar Ions*, Reinhold, New York.

(27) Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) *Computer-aided interpretation of analytical sedimentation data for proteins*, The Royal Society of Chemistry, Cambridge, U.K.

(28) NMRPipe System (C) (1995−2008) F. Delaglio and NMR Science Inc.

(29) Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D. (2005) The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins 59*, 687–696.

(30) Kay, L. E., Keifer, P., and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc. 114*, 10663.

(31) Muhandiram, D. R., and Kay, L. E. (1994) Gradient-enhanced triple-resonance three-dimensional NMR experiments with improved sensitivity. *J. Magn. Reson., Ser. B 103*, 203–216.

(32) Jung, Y.-S., and Zweckstetter, M. (2004) Backbone assignment of proteins with known structure using residual dipolar couplings. *J. Biomol. NMR 30*, 25–35.

(33) Jung, Y.-S., and Zweckstetter, M. (2004) Mars: Robust automatic backbone assignment of proteins. *J. Biomol. NMR 30*, 11–23.

(34) Zhang, O., Kay, L. E., Olivier, J. P., and Forman-Kay, J. D. (1994) Backbone ^{1}H and ^{15}N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. *J. Biomol. NMR 4*, 845.

(35) Kay, L. E., Xu, G. Y., Singe, A. U., Muhandiram, D. R., and Forman-Kay, J. D. (1993) A gradient-enhanced HCCH-TOCSY experiment for recording side-chain ^{1}H and ^{13}C correlations in H_2O samples for proteins. *J. Magn. Reson., Ser. A 101*, 333.

(36) Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR 13*, 289–302.

(37) Vuister, G. W., and Bax, A. (1993) Quantitative J correlation: A new approach for measuring homonuclear three-bond $J_{HNH\alpha}$ coupling constants in ^{15}N-enriched proteins. *J. Am. Chem. Soc. 115*, 7772.

(38) Ottiger, M., Delaglio, F., and Bax, A. (1998) Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J. Magn. Reson., Ser. A 131*, 373–378.

(39) Ruckert, M., and Otting, G. (2000) Alignment of biological macromolecules in novel nonionic liquid crystalline media for NMR experiments. *J. Am. Chem. Soc. 122*, 7793.

(40) Freyssingeas, E., Nallet, F., and Roux, D. (1996) Measurement of the membrane flexibility in lamellar and "sponge" phases of the $C_{12}E_5$/hexanol/water system. *Langmuir 12*, 6028–6035.

(41) Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Clore, G. M. (2003) The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson. 160*, 65–73.

(42) Schwieters, C. D., Kuszewski, J. J., and Clore, G. M. (2006) Using Xplor-NIH for NMR molecular structure determination. *Prog. Nucl. Magn. Reson. Spectrosc. 48*, 47.

(43) Vasudevan, S. V., Schulz, J., Zhou, C., and Cocco, M. J. (2010) Protein folding at the membrane interface, the structure of Nogo-66 requires interactions with a phosphocholine surface. *Proc. Natl. Acad. Sci. U.S.A. 107*, 6847–6851.

(44) Clore, G. M., Gronenborn, A. M., and Bax, A. (1998) A robust method for determining the magnitude of the fully asymmetric alignment tensor of oriented macromolecules in the absence of structural information. *J. Magn. Reson. 133*, 216–221.

(45) Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera: A visualization system for exploratory research and analysis. *J. Comput. Chem. 25*, 1605–1612.

(46) Koradi, R., Billeter, M., and Wuthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graphics 14*, 29–32, 51−55.

(47) Pintacuda, G., and Otting, G. (2002) Identification of protein surfaces by NMR measurements with a paramagnetic Gd(III) chelate. *J. Am. Chem. Soc. 124*, 372–373.

(48) Harris, R. K. (1986) *Nuclear magnetic resonance spectroscopy: A physicochemical view*, Longman Scientific & Technical, Harlow, Essex, England.

(49) Bava, K. A., Gromiha, M. M., Uedaira, H., Kitajima, K., and Sarai, A. (2004) ProTherm, version 4.0: Thermodynamic database for proteins and mutants. *Nucleic Acids Res. 32*, D120–D121.

(50) Tretyachenko-Ladokhina, V., Cocco, M. J., and Senear, D. F. (2006) Flexibility and adaptability in binding of *E. coli* cytidine repressor to different operators suggests a role in differential gene regulation. *J. Mol. Biol. 362*, 271.

(51) Spronk, C. A., Bonvin, A. M., Radha, P. K., Melacini, G., Boelens, R., and Kaptein, R. (1999) The solution structure of Lac repressor headpiece 62 complexed to a symmetrical lac operator. *Structure 7*, 1483.

(52) Kalodimos, C. G., Biris, N., Bonvin, A. M., Levandoski, M. M., Guennuegues, M., Boelens, R., and Kaptein, R. (2004) Structure and flexibility adaptation in nonspecific and specific protein-DNA complexes. *Science 305*, 386–389.

(53) Loll, B., Saenger, W., and Biesiadka, J. (2007) Structure of full-length transcription regulator CcpA in the apo form. *Biochim. Biophys. Acta 1774*, 732–736.

(54) Nagadoi, A., Morikawa, S., Nakamura, H., Enari, M., Kobayashi, K., Yamamoto, H., Sampei, G., Mizobuchi, K., Schumacher, M. A., and Brennan, R. G. (1995) Structural comparison of the free and DNA-bound forms of the purine repressor DNA-binding domain. *Structure 3*, 1217–1224.

(55) Whittaker, S. B., Boetzel, R., MacDonald, C., Lian, L. Y., Pommer, A. J., Reilly, A., James, R., Kleanthous, C., and Moore, G. R. (1998) NMR detection of slow conformational dynamics in an endonuclease toxin. *J. Biomol. NMR 12*, 145–159.

(56) Härd, T. (1999) NMR studies of protein−nucleic acid complexes: Structures, solvation, dynamics and coupled protein folding. *Q. Rev. Biophys. 32*, 57–98.

(57) Clore, G. M., and Garrett, D. S. (1999) R-factor, Free-R and complete cross-validation for dipolar coupling refinement of NMR structures. *J. Am. Chem. Soc. 121*, 9008–9012.