Published online 29 April 2011 *Nucleic Acids Research, 2011, Vol. 39, No. 15* ***6741–6752***
doi:10.1093/nar/gkr262

Solution structure and dynamic analysis of chicken MBD2 methyl binding domain bound to a target-methylated DNA sequence

J. Neel Scarsdale[1,2,3], Heather D. Webb[4], Gordon D. Ginder[3,5,6] and David C. Williams Jr[3,4,7,*]

[1]Institute of Structural Biology and Drug Design, [2]Center for the Study of Biological Complexity, [3]Massey Cancer Center, [4]Department of Pathology, [5]Department of Internal Medicine, [6]Department of Human and Molecular Genetics and [7]Department of Physiology and Biophysics, Virginia Commonwealth University, Richmond, VA 23298-0035, USA

Received August 9, 2010; Revised April 4, 2011; Accepted April 6, 2011

ABSTRACT

The epigenetic code of DNA methylation is interpreted chiefly by methyl cytosine binding domain (MBD) proteins which in turn recruit multiprotein co-repressor complexes. We previously isolated one such complex, MBD2-NuRD, from primary erythroid cells and have shown it contributes to embryonic/fetal β-type globin gene silencing during development. This complex has been implicated in silencing tumor suppressor genes in a variety of human tumor cell types. Here we present structural details of chicken MBD2 bound to a methylated DNA sequence from the *ρ-globin* promoter to which it binds *in vivo* and mediates developmental transcriptional silencing in normal erythroid cells. While previous studies have failed to show sequence specificity for MBD2 outside of the symmetric mCpG, we find that this domain binds in a single orientation on the *ρ-globin* target DNA sequence. Further, we show that the orientation and affinity depends on guanine immediately following the mCpG dinucleotide. Dynamic analyses show that DNA binding stabilizes the central β-sheet, while the N- and C-terminal regions of the protein maintain mobility. Taken together, these data lead to a model in which DNA binding stabilizes the MBD2 structure and that binding orientation and affinity is influenced by the DNA sequence surrounding the central mCpG.

INTRODUCTION

DNA methylation has been the focus of extensive research for the past several decades. This epigenetic modification involves the enzymatic addition of methyl groups at the C5 position of both symmetrically related cytosine bases in a CG dinucleotide sequence (CpG). Areas of increased CpG content (CpG islands) are often associated with gene promoters and when methylated are bound by regulatory complexes that downregulate transcription. Only a subset of CpG islands is methylated in adult tissues, which silence expression of the associated gene in a tissue-specific manner (1,2). Carcinogenesis has been associated with aberrant global DNA hypomethylation and hypermethylation of CpG islands associated with tumor suppressor genes (3–5).

The majority of methyl cytosine binding proteins specifically recognize the methylated CpG sequence through an ~60 amino acid methyl cytosine binding domain (MBD). There are five members of the MBD family in mammals: MeCP2, the first to be identified (6) and MBD1 through MBD4 (7). Outside of the methyl binding domain itself, the amino acid sequence of each protein is unique (with the exception of a high level of homology between MBD2 and MBD3). The regulatory complexes recruited and the promoter regions occupied by each appear to be at least partially non-overlapping and unique (8). Genetic knockouts of each MBD protein demonstrate unique phenotypes suggesting distinct functional roles (9). For example, mutations of MeCP2, many of which are within the MBD, are associated with Rett syndrome, a severe developmental neurological disorder (10) and MBD2 regulatory complexes have been implicated in silencing a small group of genes in normal tissues including chicken and human globin genes (11–14), the mouse *IL-4* gene (15,16) and genes in the gut of the developing mouse (15), as well as a large number of aberrantly methylated tumor suppressor genes in cancers such as *GSTP1* (5,17–19), *p14*/*p16* (20), *DAPK1* (21) and *KLK10* (22).

*To whom correspondence should be addressed. Tel: +1 (804) 628 2073; Fax: +1 (804) 828 2812; Email: dwilliams8@mcvh-vcu.edu

Recently Chatagnon *et al*. (23) investigated the role of DNA methylation and silencing of the estrogen regulated *pS2* gene. They showed that MBD2 down-regulated the expression of *pS2* when the TATA box region was methylated and that knockdown of MBD2 restored estrogen-dependent expression even though the DNA remained methylated. Therefore, other MBD proteins could not functionally substitute for MBD2 to silence expression of pS2. These results underscore the open question of how different MBD proteins selectively silence different methylated promoters.

In addressing why different MBD proteins silence distinct subsets of methylated promoters, studies have demonstrated that MeCP2 prefers A/T sequences adjacent to the mCpG (24) and that MBD1 preferentially binds TmCpGCA and TGmCpGCA sequences (25). In contrast, sequence specificity for bases outside of the mCpG has not been previously identified for MBD2. This latter observation raises the question of why MBD2 does not substitute for genes regulated by MBD1 and MeCP2. One hypothesis is that the regulatory complexes recruited by MBD2, which contain other DNA binding domains, contribute to promoter selectivity. For example, the MIZF protein binds to MBD2 and recognizes a specific DNA sequence, which could confer sequence specificity to the promoter targeted by MBD2. (26,27) Alternatively, the methyl binding domain itself could dictate which promoters are silenced. In support of the latter, Fraga *et al*. (28) demonstrated variable binding affinities between isolated MBD proteins that depends on the CpG density of the different promoters studied.

The structures of MBD1 (29) and MeCP2 (30) methyl binding domains bound to methylated DNA have been solved by nuclear magnetic resonance (NMR) spectroscopy and X-ray crystallography, respectively. These structures have shown that the MBD selectively binds methylated DNA through conserved arginine and tyrosine residues that make base-specific interactions with the mCpG sequence. The crystal structure of MeCP2 reveals that two arginine residues hydrogen bond with the symmetrically related guanine bases of the mCpG while a tyrosine residue makes water mediated hydrogen bonds to the methyl group of a methylated cytosine (30). This tyrosine has been directly implicated in the binding selectivity for methylated DNA by mutagenesis studies (28). Nonetheless, these studies have not provided clear structural evidence to explain sequence specificity for bases outside of the central mCpG.

Ginder and colleagues (11) previously identified a direct gene promoter target for MBD2 that contributes to silencing of the *ρ-globin* gene during normal avian erythroid development. These results gave us an opportunity to study the structural details of MBD2 bound to a *bona fide* target-methylated DNA sequence, important structural information that was previously unavailable. Among the methyl cytosine binding family of proteins, MBD2 is of particular interest since: (i) MBD2 represents the most phylogenetically ancient methyl cytosine binding protein, found across vertebrate, invertebrate and plant species (31–34); (ii) MBD2 shows the greatest degree of selectivity for methylated versus unmethylated CpG sequences (28) and (iii) MBD2 has been directly implicated in silencing tumor suppressor genes in cancer (5,17–22) and as such has been proposed as a therapeutic target for a wide-range of human malignancies (35).

In the structural and dynamic studies reported here, we show that the MBD of cMBD2 (96% identical to human MBD2), recognizes a target-methylated DNA sequence from the *ρ-globin* gene promoter in a similar manner to both MeCP2 and MBD1. Structural details reveal differences that likely contribute to greater DNA affinity and selectivity for methylated DNA. Surprisingly, MBD2 binds this target DNA sequence in a single orientation, which indicates previously unrecognized sequence specificity for bases outside of the symmetric mCpG. We show that binding orientation depends largely on the base pairs immediately flanking the mCpG dinucleotide and that reversing these base pairs largely, but not completely, reverses the binding orientation. Furthermore, changing the guanine that immediately follows the mCpG dinucleotide reduces binding affinity by an order of magnitude. NMR relaxation studies of MBD2 that show the DNA contacting region is well-structured and stable while both the N- and C-terminal regions of the domain undergo internal dynamic motions on fast and slow time scales, respectively. Thus our studies suggest a model in which the MBD2 methyl binding domain adopts a locally well-formed structure upon DNA binding and that binding is sensitive to the methylation status as well as the base pairs immediately flanking the mCpG. The latter finding suggests a basis for preferential binding of MBD2 to certain methylated and CpG rich promoters.

MATERIALS AND METHODS

Protein expression and purification

Amino acid residues 2–72 from chicken MBD2 were cloned and expressed as a fusion protein with thioredoxin and a hexahistidine N-terminal tag using a modified pET32a (Novagen) vector previously described (36). After affinity purification using a nickel sepharose column, the thioredoxin and hexahistidine tag were removed by cleavage with thrombin. The protein was further purified with sequential chromatographic isolation over (i) benzamidine sepharose (GEHealthcare), (ii) MonoS 10/100 GL (GEHeathcare) and (iii) Superdex 75 26/60 columns. The resulting MBD was >95% pure as estimated by SDS–PAGE analysis. Uniform double (^{13}C, ^{15}N) and triple (^{13}C, ^{15}N, ^{2}H) labeled protein samples were generated by standard techniques.

Sample preparation

Ten base complementary oligonucleotides with a central-methylated cytosine were purchased (Integrated DNA Technologies) and further purified over a MonoQ (GE Healthcare) ion exchange column before and after annealing. The DNA sequence was derived from the *ρ-globin* promoter known to be a native target sequence for MBD2 (GGAT(mC)GGCTC) (11). Purified MBD2 protein was combined with 10% excess double stranded

oligonucleotide, buffer exchanged into 10 mM $NaPO_4$, pH 6.5, 1 mM dithiothreitol, 10% ^2H_2O and 0.02% sodium azide and concentrated to ~1 mM.

For paramagnetic resonance enhancement measurements, EDTA-conjugated thymidine-modified oligonucleotides were purchased (Midland Certified Reagent Company, Inc). The modified thymidine was incorporated as (i) an additional base pair between positions 3 and 4 with EDTA–thymidine in the reverse strand and (ii) at base pair position 9. Purified EDTA conjugated DNA was first stripped of any divalent cations by adding 5 mM EDTA to the sample and then washed with 10 mM MES pH 6.5, 500 mM NaCl. Excess $CaCl_2$ or $MnCl_2$ were added and the DNA extensively washed by serial dilution and spin concentration with first a high salt buffer (10 mM MES pH 6.5, 500 mM NaCl) then low salt NMR buffer (10 mM MES pH 6.5). Then, $^2H^{13}C^{15}N$-cMBD2 was added to a slight excess of EDTA conjugated DNA and the sample concentrated to ~500 μM.

NMR data collection

Standard NMR experiments for resonance assignments, distance and torsional angle restraints were measured on a Varian 500 MHz Unity+ and Bruker Avance III™ 700 MHz NMR spectrometers at 25°C. Residual dipolar couplings were measured by adding ~12 mg/ml of pf1 bacteriophage to triple labeled MBD2:DNA samples and $^1D_{NH}$, $^1D_{NC'}$, $^1D_{HNC'}$ and $^1D_{C\alpha C'}$ couplings determined using standard inphase antiphase (IPAP)- and transverse relaxation optimized NMR spectroscopy (TROSY)-based experiments for both isotropic and partially aligned samples.

Paramagnetic relaxation enhancement (PRE) measurements were carried out as described by Iwahara *et al.* (37,38) and Iwahara and Clore (39). ^1H-^{15}N resonance peak intensities were measured at four ^1H_N T_1 (100, 300, 500 and 900 ms) inversion recovery delays and two ^1H_N T_2 (0 and 24 ms) relaxation delays for cMBD2 bound to Ca^{2+} and Mn^{2+} saturated EDTA-conjugated DNA. Relaxation rates were derived from fitting the intensities to an inversion recovery model (R_1) or calculated from the ratio of the intensities (R_2) [Equations (20) and (21), Iwahara *et al.* (38)] and ^1H_N-Γ_1 and ^1H_N-Γ_2 calculated as the difference in rates between Mn^{2+} and Ca^{2+} saturated samples. The ^1H_N-Γ_1 / ^1H_N-Γ_2 ratio was used to estimate τ_c^{app} [Equation (16), Iwahara *et al.* (38)] which was further optimized during simulated annealing (38).

Structure calculation

The structure of the complex was calculated by simulated annealing using the Xplor-NIH software package (40). The minimized target function included the experimental NMR restraints (nuclear Overhauser effect (NOE)-derived interproton distances, torsion angles, residual dipolar couplings and paramagnetic relaxation enhancement), a quartic van der Waals repulsion term for the non-bonded contacts (41), a torsion angle data base potential of mean force (42) and a radius of gyration restraint to ensure optimal packing (43). Backbone torsion angle restraints were based on chemical shifts as determined from TALOS (44,45) and a limited number of sidechain torsion angle restraints were derived from measured $^3J_{N\text{-}C\gamma}$ and $^3J_{CO\text{-}C\gamma}$ coupling constants. α-Helix and β-sheet hydrogen bond distance and angle restraints were incorporated based on the backbone torsion angle and characteristic NOE crosspeak patterns.

DNA assignments and NOE restraints were determined by double filtered (^{13}C,^{15}N) homonuclear NOE experiments collected at 4°C, 10°C and 25°C. Importantly, assignments of the key 5-methylcytosine H5 protons were confirmed by (i) the presence of strong NOEs between Thy104 H6 and both Thy104 H5 and mCyt105 H5 and (ii) comparison to the double filtered spectrum of a complex with Thy104Uri-modified methylated DNA. In addition to the NOE restraints, hydrogen bond distance and planarity restraints as well as B-form DNA backbone torsion angle restraints were incorporated into structure calculations. Furthermore, PRE restraints were incorporated using a PRE target function and a PRE Q-factor was calculated as described by Iwahara *et al.* (38). A hybrid DNA molecule incorporating an ensemble of three alternative EDTA conformations for each EDTA-conjugated thymidine was used in the simulated annealing calculations.

Initial simulated annealing calculations did not incorporate any intermolecular (protein:DNA) hydrogen bond restraints. Given that both R24 and R46 sidechain Hε showed strong NOE cross correlations with the mCyt105 and mCyt115 methyl protons, respectively, and that initial calculations consistently placed these two sidechains in close proximity to Gua106 and Gua116 bases, hydrogen bond distance and angular restraints were incorporated between R24/R26 NH_2 and Gua106/Gua116 O6 and N7 in the final simulated annealing calculations.

PRE and MBD orientation

In order to test for evidence of the alternative, symmetrically related binding orientation, an ensemble of 10 identical MBD bound to the methylated DNA were generated and identified by 10 different segment ids within XPLOR_NIH (40). The CpG bases from one DNA strand were aligned to the CpG bases on the opposite strand to generate the symmetrically rotated conformation of MBD. The PRE Q% was calculated with the experimental PRE data (EDTA–thymidine between base pairs 3 and 4) for this ensemble with 1–10 of the members rotated to the symmetrically related orientation. This same procedure was applied to PRE data collected with wild-type DNA (GGAT(mC)GGCTC) and with an inverted central 4 bp (GGAC(mC)GACTC).

Binding affinity

Wild-type and mutant 10 bp oligonucleotides (3′-biotinylated on the forward strand) were purchased (Integrated DNA Technologies), annealed and further purified by ion exchange chromatography on a MonoQ column (GE Healthcare). Wild-type and mutant MBD2 methyl-binding domain were expressed as previously and purified by nickel sepharose and size exclusion

chromatography (10 mM HEPES, 50 mM NaCl, 3 mM $MgCl_2$, 0.1 mM EDTA, 1 mM DTT, pH 7.4). The purified DNA was bound to a Sensor SA chip on a Biacore T100 (GE Healthcare) (10 ng/ul DNA, 10 ul/min flow rate, 100 s) until a final relative response of ~120 U. Kinetic and steady state binding analysis was carried out for varying concentrations of MBD2 proteins at a flow rate of 30 ul/min (10 mM HEPES, 50 mM NaCl, 3 mM $MgCl_2$, 0.1 mM EDTA, 1 mM DTT, 0.05% polysorbate 20, pH 7.4). The data were fit by steady state analysis using the manufacturer's software. At least one concentration of MBD2 was repeated in triplicate to determine the analytical uncertainty of the steady state response.

^{15}N relaxation measurements

^{15}N-R_1, $R_{1\rho}$ and heteronuclear NOEs were measured using standard pulse sequences on a single deuterated sample at 500 MHz. The crosspeak intensities were fit using scripts accompanying the NMRPipe software (46). A spin-lock field strength of 1.5 kHz was used for $R_{1\rho}$ measurements, which were converted to R_2 based on resonance offset and the observed R_1. Data were analyzed based on the extended model free formalism (47–49) using the Modelfree4 software and following the protocol described by Palmer and colleagues (50). This formalism incorporates from one to three motional parameters for each residue, choosing from an order parameter ($S^2 = S_f^2 S_s^2$) for both fast (S_f) and slow time scales (S_s), an internal time scale parameter (τ_e) representing either a fast (τ_f) or slow (τ_s) motions and a chemical exchange (R_{ex}) term that incorporates pseudo-first-order exchange processes. The rotational correlation time (τ_c) estimated from the trimmed R_1/R_2 ratio was used to determine the most appropriate model of internal motion for each residue [(i) S^2; (ii) S^2, $\tau_e = \tau_f$; (iii) S^2,R_{ex}; (iv) S^2, $\tau_e = \tau_f$, R_{ex} and (v) S_f^2,S^2, $\tau_e = \tau_s$] before fitting the rotational correlation time and internal motion parameters globally. The appropriate model for each residue is selected based on the sum-squared error (Γ_i) in the fit as compared to a critical value of a simulated Γ_i distribution and an *F*-statistic as described (50).

RESULTS

Solution structure of cMBD2 bound to methylated DNA

We determined the solution structure of the methyl binding domain from cMBD2 (residues 2–71) bound to a 10 bp fragment from the *ρ-globin* promoter containing a centrally located methylated CpG sequence using multidimensional NMR spectroscopy. Structures were determined based on 783 (20 intermolecular) NOE-derived distance constraints, 202 residual dipolar coupling constraints, 102 protein backbone torsion angle constraints and 136 B-form DNA torsion angle constraints (Table 1). In addition, EDTA conjugated thymidine was incorporated in one of the two different positions in the DNA and saturated with Mn^{2+} (or Ca^{2+} as a reference), as described by Iwahara *et al.* (38). Paramagnetic enhanced relaxation rates were measured for 73 backbone amide hydrogens (46 for EDTA–Thy119 and 27 for EDTA–Thy109) and incorporated into the structure calculations as described in the 'Materials and Methods' section (38). These data provide an independent measure of the relative orientation between protein and DNA and supplements the limited number of observed intermolecular NOEs.

The cMBD2–DNA complex structure is well defined with an overall root mean square deviation (RMSD) for all heavy atoms of 1.1 ± 0.1 Å^2 and for backbone atoms

Table 1. NMR and refinement statistics

Constraints and Statistics	Protein	Nucleic acid
NMR distance and dihedral constraints		
Distance restraints		
Total NOE	664	119
Intraresidue	194	72
Interresidue	470	47
Sequential ($\|i–j\| = 1$)	232	32
Non-sequential ($\|i–j\| > 1$)	238	15
Hydrogen bonds	16	42
Hydrogen bonds protein–nucleic acid	4	
Protein–nucleic acid intermolecular	20	
Total dihedral angle restraints		
Protein		
ϕ	56	
ψ	46	
Nucleic acid		
Backbone		114
Sugar pucker		22
RDC Q% (number of constraints)		
NH	6.1 ± 0.8 (56)	
H^NC'	30.6 ± 1.4 (49)	
NC'	36.3 ± 2.5 (49)	
$C'C\alpha$	44.4 ± 2.6 (49)	
PRE Q% (number of constraints)		
EDTA T119	23.9 ± 1.8 (46)	
EDTA T109	26.9 ± 1.4 (27)	
Structure statistics		
Violations (mean and SD for the complex)		
Distance constraints (Å)	0.026 ± 0.004	
Dihedral angle constraints (°)	0.62 ± 0.08	
Max. dihedral angle violation (°)	10.0	
Max. distance constraint violation (Å)	0.71	
Deviations from idealized geometry		
Bond lengths (Å)	0.0089 ± 0.0004	
Bond angles (°)	0.604 ± 0.004	
Impropers (°)	0.54 ± 0.06	
Average pairwise r.m.s. deviation[a] (Å)		
Protein		
Heavy	1.1 ± 0.2	
Backbone	0.7 ± 0.2	
DNA		
Heavy		1.1 ± 0.1
Backbone		0.2 ± 0.1
Complex		
Heavy	1.1 ± 0.1	
Backbone	0.8 ± 0.2	
Ramachandran plot summary (structured residues) (%)		
Most favored regions	86.1 (91.4)	
Additionally allowed regions	10.6 (8.6)	
Generously allowed regions	2.0 (0.0)	
Disallowed regions	1.3 (0.0)	

[a]Pairwise r.m.s. deviation was calculated among 20 refined structures for structured residues (amino acids 8–69 of MBD2 and base pairs 103–108 of DNA).

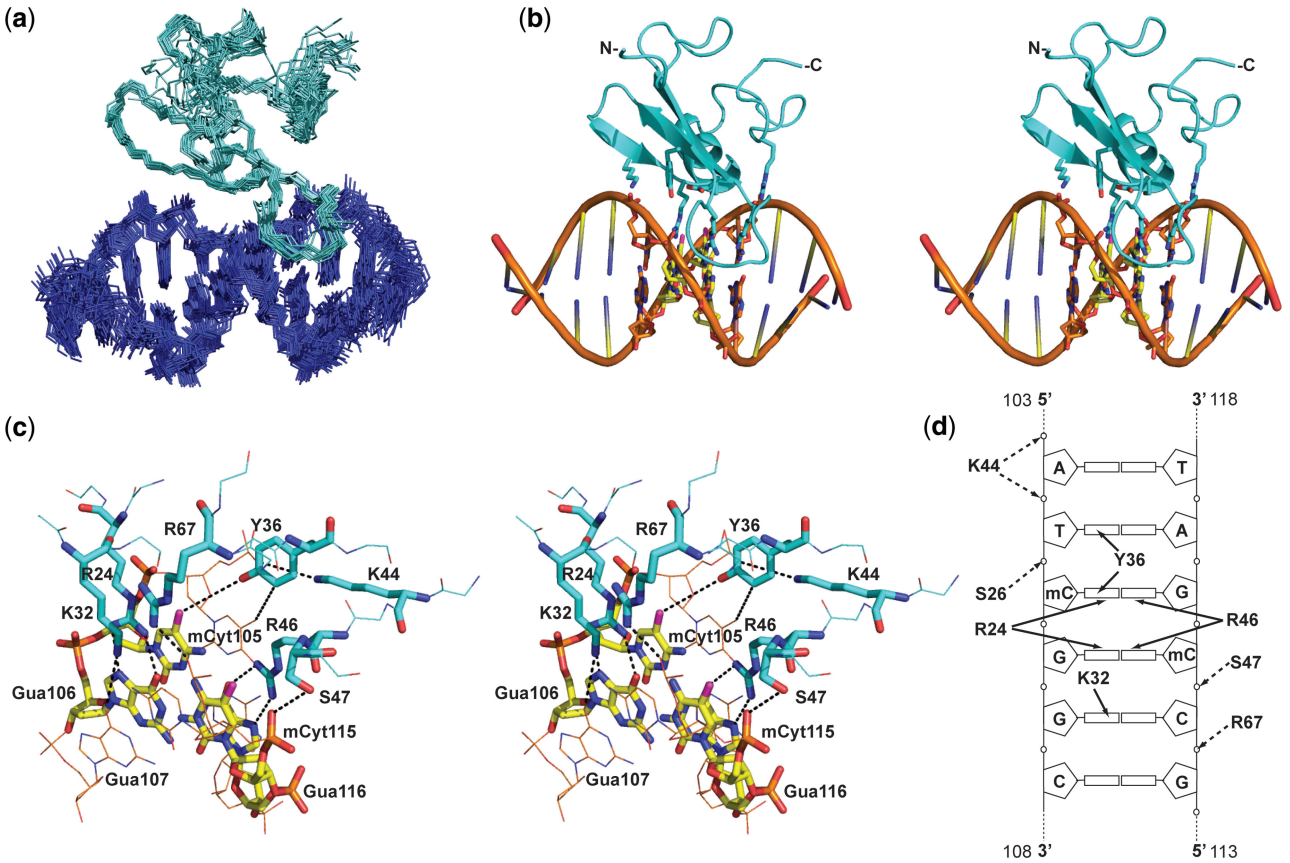

Figure 1. Solution structure of cMBD2 methyl binding domain bound to methylated DNA. (**a**) A best-fit superimposition licorice diagram of protein backbone (cyan) and DNA heavy atoms (blue) is shown for the ensemble of 20 calculated structures. (**b**) A stereo cartoon diagram of a representative MBD2 (cyan) and DNA (blue/orange) is shown. (**c**) A detailed line diagram of the protein:DNA interface is depicted with contacting protein residues (cyan) and mCpG DNA bases (yellow) shown as sticks. (**d**) A diagram depicting base-specific (solid lines) and phosphate backbone (dashed lines) contacts between MBD2 and DNA (for simplicity, only the central 6 bp are shown). Structure figures were generated with (a) VMD-XPLOR (65) and (b, c) PyMOL (Delano Scientific, LLC).

of 0.8 ± 0.2 Å^2 (Table 1 and Figure 1a) and is similar to other methyl binding domains. The cMBD2 structure consists of a long finger-like projection formed by a three strand β-sheet (residues 18–23, 32–38 and 43–45) with a fairly large loop between strands 1 and 2 and a tight turn between strands 2 and 3. Immediately following the last β-strand, the backbone turns back to form a short α-helix (residues 47–55). The N-terminal residue of this α-helix, S47, forms a classic N-cap through a sidechain hydroxyl hydrogen bond with the amide hydrogen of Q50. Neither the N- nor C-terminal regions of the MBD domain (residues 1–17 and 56–72) forms a regular secondary structure and both of these regions pack against the β-sheet opposite the DNA binding surface of the protein.

The finger-like projection of cMBD2 extends down into the major groove of DNA to make contact with the symmetrically methylated CpG sequence. Three residues make base-specific contacts with the methylated CpG: R24, Y36 and R46. The two arginines form hydrogen bonds with the two symmetrically opposed guanine bases (Gua106 and Gua116), which permits the aliphatic side chains of each arginine to pack against the neighboring methyl-cytosine methyl groups (mCyt105 and mCyt115). The aromatic side chain of Y36 interacts with the methyl groups of mCyt105 and the neighboring Thy104 (Figure 1c). D34 potentially stabilizes the conformation of R24 through a sidechain hydrogen bond and makes a direct hydrogen bond with the hydroxyl of Y36. The positively charged amino acids R67 and K44 form close ionic interactions with the phosphate backbone from Cyt114 and Ade103–Thy104, respectively. K32 adopts multiple conformations in the 20 simulated annealing structures, forming either base-specific contacts with Gua107 (in the majority of structures) or an ionic interaction with the backbone phosphate of Ade112.

In addition to the individual amino acid–DNA interactions, the positive end of the helical dipole points towards the negatively charged backbone phosphate of mCyt115, contributing to binding. S47 from this helix closely approaches the phosphate backbone of DNA and can form a sidechain hydrogen bond with the backbone phosphate of mCyt115 (in addition to the N-cap hydrogen bond with the backbone amide of Q50). Hence, the MBD spans the major groove to make base-specific hydrogen bond and aliphatic interactions with the central-methylated CpG, ionic, hydrogen bond and helical dipole interactions with the phosphate backbone on both sides of the major groove and

buttressed by a few base-specific interactions involving residues outside of the mCpG.

Comparison with other MBD proteins

The complex between cMBD2:DNA is similar to the previously reported structures of MeCP2 and MBD1 methyl binding domains bound to methylated DNA (29,30). The methyl binding domain from cMBD2 shares 60% identity (74% homology) with human MBD1. The structures of these two domains are very similar, with a backbone RMSD of 2.2 Å. The secondary structures align closely and the domains share many of the same features discussed for MBD2. The residues that directly interact with DNA are conserved between the two proteins and make very similar interactions with the DNA. However, MBD1 is rotated slightly and the binding surface more closely approximates the DNA than MBD2 (Supplementary Figure S1). This difference in global orientation alters some of the fine details and reflects subtle differences in how the two proteins bind methylated DNA. For example, the loop between β1 and β2 of MBD1 closely approaches the DNA such that the amide hydrogen of A26 forms a hydrogen bond with the backbone phosphate of Gua107. The hydroxyl of Y34 of MBD1 can form a direct hydrogen bond with the N4 of mCyt106, as opposed to an interaction with the methyl group of mCyt seen for MBD2 (potentially a water-mediated hydrogen bond as seen for MeCP2) (30). V47 sidechain methyl groups pack against the deoxyribose of Cyt117 and R18 of MBD1 closely interacts with the phosphate backbone of Thy104.

The MBD from cMBD2 shares 50% identity (56% homology) with human MeCP2. While the β-sheet region is very similar, with a backbone RMSD of 2.4 Å for alignment with cMBD2 residues 9–55, MeCP2 has a longer α-helix incorporating an additional turn (4 residues) and consequently an additional 5–6 residues in the C-terminal region that packs against this same helix. These additional residues cause the α-helix to adopt a more oblique angle with respect to the DNA phosphate backbone (Supplementary Figure S1). Despite these changes, MeCP2 and cMBD2 bind the methylated CpG sequence in a very similar manner. Both proteins contact the central mCpG with virtually identical interactions involving residues R24(111), R46(133) and Y36(123) in cMBD2(MeCP2).

The most notable differences between the cMBD2 and MeCP2 complexes involve residues outside of the mCpG binding region. K44 of cMBD2 (A31 of MeCP2) provides an additional ionic interaction while L28 (R114 of MeCP2) eliminates an ionic interaction with the phosphate backbone of DNA. The latter change helps explain why the loop between β-strands 1 and 2 of cMBD2 (residues 24–31) deviates away from the DNA relative to MeCP2. R67 replaces T158 to form a close ionic interaction with DNA phosphate backbone. The net result is an additional ionic interaction that is likely to contribute to the observed higher binding affinity and possibly greater specificity for methylated DNA by MBD2, as confirmed by the reduced binding affinity for the R67M mutant (Table 2; see below).

Table 2. Binding affinity

MBD2	mCpG	K_D (μM) ± SE	R_{max}	χ^2
WT	WT	2.1 ± 0.1	323	15.1
K32A	WT	291 ± 19	501	0.4
Y36F	WT	109 ± 3	497	0.69
R46C	WT	590 ± 71	678	1.9
R67M	WT	197 ± 17	464	8.2
K19W	WT	135 ± 17	334	17.5
WT	Thy104Gua	2.2 ± 0.1	247	2.8
WT	Gua107Thy	29 ± 2	402	10.6
WT	Inverted	2.3 ± 0.5	86	5.5

Orientation preference for cMBD2 binding to methylated DNA

Given the apparent lack of sequence selectivity previously reported (24) and that the mCpG DNA sequence is palindromic, we fully anticipated that cMBD2 would bind equally in either of two symmetrically related orientations. Surprisingly, the NOE and PRE data are most consistent with a single orientation of cMBD2 on this DNA. Among the 20 intermolecular protein:DNA NOEs, two particularly strong orientation specific intermolecular NOE peaks were identified between the Hε of R24 and R46 and the H5 methyl hydrogens of mCyt105 and mCyt115, respectively (Figure 2a). Likewise, the PRE data fit well with this orientation of cMBD2 (EDTA Thy119 PRE Q% = 23.9 ± 1.8, see Table 1). To explore how strongly the data favored one orientation, we reversed the orientation of cMBD2 on the DNA *in silico* and minimized the conformation of the EDTA-conjugated Thy119 ensemble. The fit to PRE data for the reverse orientation was significantly worse (Q% = 36.7) even though minimization allowed the conjugated EDTA ensemble to adopt disparate conformations.

This observation suggests that cMBD2 adopts predominantly one orientation on the methylated *ρ-globin* gene promoter sequence. However, the data do not exclude the possibility that the complex exists as a rapidly exchanging, albeit skewed, mixture between these two orientations. Paramagnetic relaxation enhancement is very sensitive to minor conformations and has been used to detect rarely populated states (51). To explore the possibility of rapid exchange between these two orientations, we generated 10 alternative copies of the cMBD2 domain *in silico* and evaluated whether the experimental PRE data fit better when averaged over an ensemble of the two possible orientations. As can be seen in Figure 2b, the data fit best (lowest PRE Q%) when all 10 members of the ensemble adopt the same orientation as that determined initially. Therefore, these experimental results are most consistent with a single orientation of cMBD2 on this methylated target sequence.

The propensity to bind in a single orientation implies unanticipated sequence selectivity for bases surrounding the mCpG. Since (i) previous SELEX experiments failed

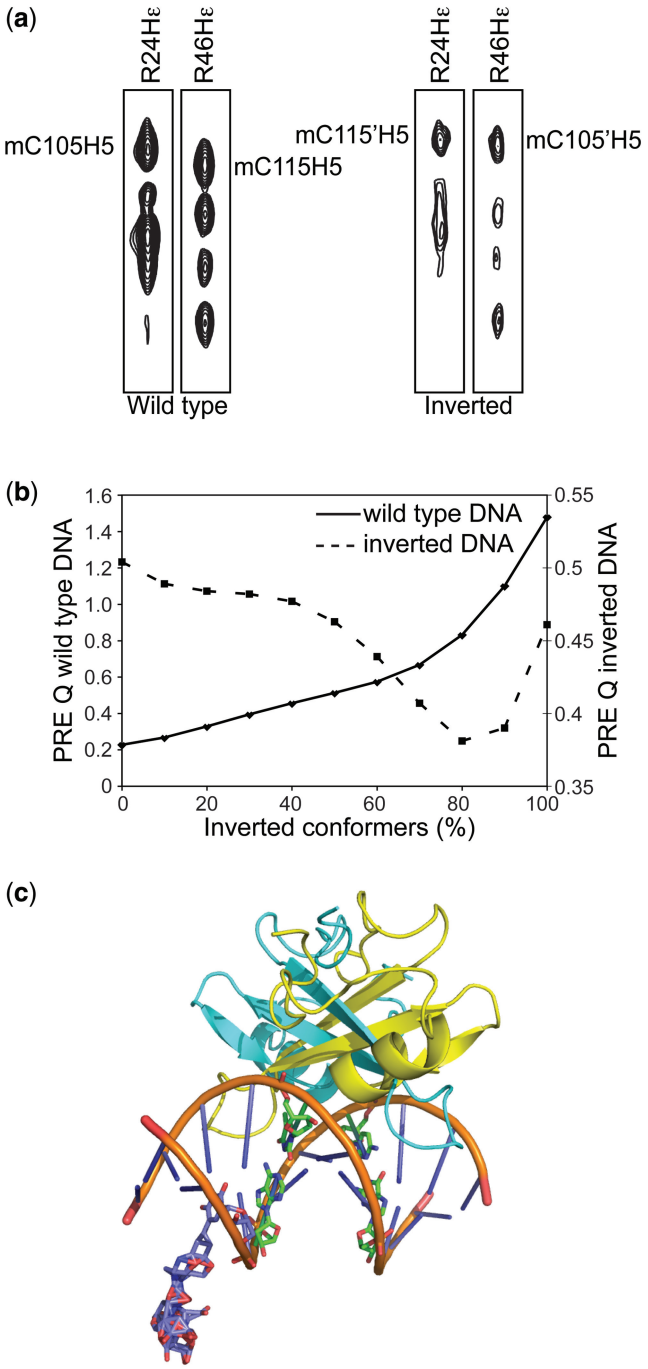

Figure 2. Orientation preference for MBD2 bound to methylated *ρ-globin* DNA sequence. (**a**) 3D ^{15}N-HSQC-NOESY slices corresponding to the N_{ε}-H_{ε} of Arg24 and Arg46 of MBD2 when bound to wild-type and inverted DNA sequences. NOE crosspeaks for mCyt105H5 and mCyt115H5 are labeled. (**b**) PRE Q values are calculated over an ensemble of MBD2 orientations for data obtained when bound to wild-type (solid line) and inverted (dashed line) DNA sequences. The ensemble consists of 10 copies of MBD2 with 0–10 of these having a reversed orientation with respect to the DNA. (**c**) A cartoon diagram shows the two alternative MBD2 orientations that make up the ensemble. The orientation as solved by NMR for wild-type DNA (cyan) and a symmetrically reversed orientation (yellow) are depicted.

to detect selectivity of MBD2 for bases surrounding an invariant C(mC)GG central 4 bp (in contrast to the clear preferential binding of MeCP2 to sequences containing A/T rich stretches adjacent to the mCpG) (24) and (ii) the majority of DNA contacts involves these central 4 bp, we hypothesized that the binding orientation depends solely on the central four base T(mC)GG sequence. To test this hypothesis, we analyzed binding to a modified sequence in which the central T(mC)GG was reversed to C(mC)GA. As can be seen in Figure 2a, strong intermolecular NOEs were identified between the Hε of R24 and R46 and the methyl hydrogen of mCyt115′ and mCyt105′, respectively, which is the reverse of the pattern seen with the wild-type sequence. The chemical shift difference between these two methyl groups on the DNA is much smaller than for the wild-type sequence (0.013 ppm versus 0.06 ppm, respectively) and this chemical shift difference increases at lower temperature (0.02 ppm at 10°C). Together, these observations indicate that reversing the sequence of the central four bases has reversed the DNA binding orientation but also raise the possibility that binding now involves a rapidly exchanging mixture of orientations due to the decrease in chemical shift differences between the mCyt methyl groups and the line-broadening seen in Figure 2a. To further evaluate this possibility, PRE data were collected using EDTA-conjugated Thy119 in the modified DNA sequence. If we assume the PRE data pertain to a single orientation, the PRE data fit best to the reverse orientation of cMBD2; however, the overall best fit of PRE data occurs using a mixed ensemble of cMBD2 orientations with the predominant orientation in the reverse direction (~80%, Figure 2b).

Binding affinity

Wild-type MBD2 binds the methylated DNA sequence with very fast on- and off-rates and an overall K_D ~2.1 μM (Table 2, Figure 3a and b). Additional binding analyses were performed with select MBD2 mutants (K32A, Y36F, R46C, R67M, K19W), the latter three of which are homologous to the more common missense mutations of MeCP2 associated with Rett syndrome (R133C, T158M, R106W). Mutations that affect direct interaction with DNA (K32A, Y36F, R46C and R67M) reduce binding by at least 50-fold (Table 2). The R46C mutation in particular markedly decreases binding, affirming the central role of the R46-Gua116 hydrogen bonding interaction. The K32A mutation removes a residue that can form a base-specific interaction with Gua107 and an ionic interaction with the phosphate backbone. The Y36F mutation reduces binding affinity by removing a single hydroxyl group that interacts with the methyl group of the mCyt105 (28,52). R67M, which decreases binding by nearly 100-fold, is homologous to one of the most common Rett syndrome missense mutations (T157M); however, in MBD2 the sidechain of R67 interacts with the phosphate backbone of DNA while in MeCP2 T157 plays a role stabilizing two consecutive turns of the protein backbone. As discussed previously, this additional ionic interaction with DNA is likely to contribute to an

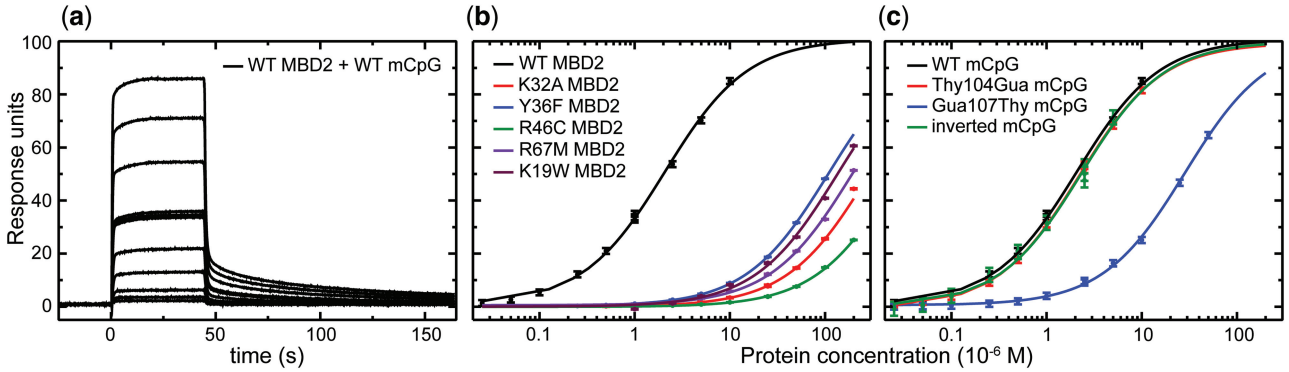

Figure 3. Binding affinity of cMBD2 to methylated DNA. (**a**) Surface plasmon resonance analysis for varying concentrations of wild-type cMBD2 binding to a 3′-biotinylated and methylated 10 bp target sequence coupled to a Sensor Chip SA on a Biacore T100 (GE Healthcare). Steady state binding response was analyzed for varying concentrations of cMBD2 with (b) specific point mutations or (c) binding to modified DNA sequences. The data were fit using the Biacore T100 evaluation software. For comparison, the response units for each were normalized to an R_{max} = 100 (Table 2).

overall higher binding affinity for MBD2 as compared to MeCP2 and MBD1.

The sidechain of K19 does not directly interact with DNA; instead the lysine sidechain forms a partially buried interaction with the backbone of G69—linking the first β-strand with the C-terminus of the MBD. The K19W mutation decreases binding affinity by nearly 70-fold suggesting that destabilizing this interaction indirectly affects DNA binding.

As discussed previously, Y36 and K32 interact with the bases of Thy104 and Gua107, respectively, forming the only base-specific interactions outside of the central mCpG. To test whether these base-specific interactions help dictate binding orientation, we determined the affinity of MBD2 for Thy104Gua and Gua107Thy substitutions. Although Thy104Gua removes a methyl group that contacts Y36, this change does not appreciably alter binding affinity (Table 2 and Figure 3c). In contrast, the Gua107Thy substitution reduced binding affinity by at least 10-fold, indicating that the K32–Gua107 interaction strongly favors the observed orientation of MBD2 on the wild-type DNA. Furthermore, reversing the central four bases to C(mC)GA did not affect the binding affinity ('inverted' DNA in Table 2 and Figure 3c). This latter observation is consistent with MBD2 binding in the reverse orientation (as shown above), which preserves the K32–Gua interaction.

Internal dynamics of cMBD2

^{15}N relaxation and heteronuclear NOE data were analyzed using the modified model-free formalism with the Modelfree4 software and following the protocol described by Palmer and colleagues (50). The overall rotational correlation time (τ_c) for the final model was 8.0 ns. The overall internal order parameters ($S^2 = S_f^2 S_s^2$), internal fast or slow rotational correlation times (τ_e) and slow exchange terms (R_{ex}) for each residue are shown in Figure 4a and Supplementary Table S1. Out of a total of 57 backbone ^{15}N analyzed, 23 were fit by model 1 ($<S^2> = 0.90 \pm 0.05$); 7 by model 2 ($<S^2> = 0.75 \pm 0.11$); 12 by model 3 ($<S^2> = 0.84 \pm 0.07$); 10 by model 4 ($<S^2> = 0.76 \pm 0.11$); 5 by model 5 ($<S^2> = 0.41 \pm 0.16$). Both the N- and C-terminal regions show evidence of internal motions, best characterized as fast time scale internal motions for the N-terminus (Model 5 with a large τ_e for residues 4–7 and 10) and a slow exchange process (R_{ex}) most pronounced for residues after the α-helix (residues 57–72).

The order parameters were mapped onto the solution structure and color coded with red the most dynamic (lowest S^2) and blue, most structured (highest S^2) residues. As can bee seen in Figure 4b, the β-sheet and DNA contacting regions are well structured while the N- and C-terminal regions form a structurally dynamic 'lid' sitting down on this stable β-sheet platform. In fact, part of the motivation for these measurements came from the observation that residues from both regions (i.e. T8, D65, R66 and T68) were broadened or completely absent from the ^{15}N-HSQC spectrum. These results show that DNA contact stabilizes the β-sheet region while regions remote from the protein:DNA interface show internal mobility.

DISCUSSION

The methyl binding domain demonstrates a remarkable ability to distinguish symmetrically methylated from unmethylated CpG sequences. This domain is highly conserved and can be found across vertebrate, invertebrate and plant species (32–34). In mammals, five MBD containing proteins have been identified. The MBD2 protein has been associated with silencing of embryonic/fetal hemoglobin expression (11–14), the mouse *IL-4* gene (15,16) and genes in the gut of the developing mouse (15) and is frequently associated with silencing of a subset of aberrantly methylated tumor suppressor genes in cancer (5,17–22,35). *In vitro* binding studies of an MBD2 containing complex from primary chicken red cells suggested a sequence preference for the *ρ-globin* promoter region over the generic CpG-rich sequence CG11 (14). These

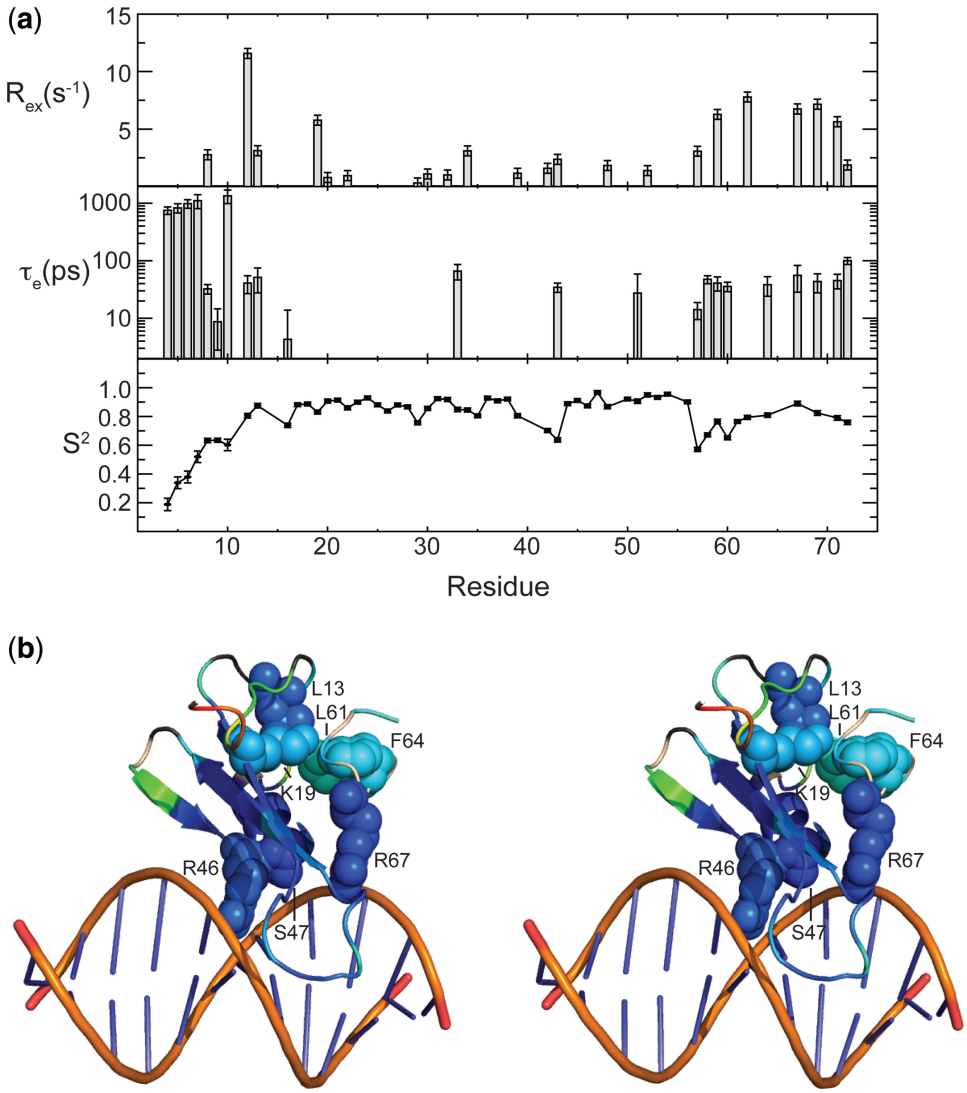

Figure 4. ^{15}N-relaxation dynamic analysis of cMBD2 bound to DNA. (**a**) Model free parameters (S^2, τ_e and R_{ex}) derived from ^{15}N-relaxation data are plotted against residue number. (**b**) A stereo cartoon diagram of the cMBD2-DNA structure is shown and colored according to generalized order parameter from high (blue) to low (red), with proline residues (not included in the analysis) colored dark gray and residues with broadened amide resonances (not included, but likely to undergo slow exchange motions) colored peach. Structurally equivalent residues for some of the most common Rett syndrome missense mutations are depicted as spheres.

observations raise the question of whether the methyl binding domain itself contributes to promoter-specific binding and functional divergence.

Binding affinity and orientation preference

In these structural studies, we use a combination of PRE and intermolecular NOEs to accurately determine the solution structure of MBD2 bound to DNA. PRE represents an independent measure to augment limited NOE data and allows one to investigate minor binding modes. (37,39,53) Together, the data strongly support a model in which cMBD2 binds almost entirely in one orientation on the wild-type-methylated target sequence from the chicken *ρ-globin* promoter. This orientation preference depends primarily, but not solely, on the bases immediately adjacent to the mCpG dinucleotide. Reversing the direction of the central four bases, which is equivalent to simply exchanging the base pairs on either side of the mCpG, reverses the predominant binding orientation of cMBD2. However, ~20% of the cMBD2 population binds in the original orientation on this reverse sequence, indicating that the sequence outside of the central four bases influences the preferred orientation.

We determined the binding affinity for wild-type and mutated MBD2 as well as variations of the target binding sequence to confirm the importance of individual protein–DNA interactions. Mutating residues that are involved in direct DNA contact significantly decreases binding affinity. Importantly, we find that a base-specific interaction between K32 and Gua107 contributes to high-affinity binding. This observation is consistent with prior studies showing that lysine residues preferentially form bidentate and complex base-specific hydrogen bonds with guanine bases (54). Therefore, the data

support a model in which the base-specific K32–Gua107 interaction increases binding affinity and leads to the observed preferred orientation on the target DNA sequence.

Bird and colleagues (24) tested for evidence of DNA sequence selectivity for both MBD2 and MeCP2 using a SELEX based experiment. In those studies, sequence preferences were identified by sequential enrichment from a pool DNA oligonucleotides with random bases surrounding a central 4 bp C(mC)GG sequence. The results of their SELEX experiments demonstrated that MeCP2 preferentially binds sequences with a short run ($\geq$4) of [A/T] base pairs while MBD2 failed to show any sequence selectivity for residues flanking the C(mC)GG bases (24).

Our solution NMR studies of MBD2, which involve only the MBD from MBD2, show strong evidence of a single orientation on this methylated target DNA sequence and a binding affinity preference for the guanine immediately following the mCpG. The wild-type DNA sequence was derived from a known *in vivo* target sequence for cMBD2 binding during normal erythroid development (11), which suggests potential functional significance for the orientation of the MBD. One possibility is that the MBD2 dictates the orientation of the associated coregulatory complex. MBD2 recruits and tightly interacts with the multi-protein nucleosome remodeling and deacetylation (NuRD) complex (55,56). The NuRD complex is one of many chromatin remodeling complexes that alter nucleosome position and chromatin structure to modify gene expression (57–59). One component of NuRD, the Mi2 protein, contains a helicase-like ATPase domain implicated in energy dependent repositioning of nucleosomes (55). Most models of nucleosome remodeling implicate sliding of chromatin remodeling complexes (either step-wise or continuous) along the DNA to reposition the nucleosome (60). The orientation of MBD2 on DNA could orient the NuRD complex, which would then direct the final localization of the nucleosome.

In addition, we have shown a sequence preference for a guanine residue immediately following the mCpG. To explore whether this sequence preference could explain MBD2 promoter selectivity, we examined the CpG islands from several known target promoters for MBD2 (*DAPK1, GSTP1* and *BRCA1*). We find an overrepresentation of either CGG or CGC sequences in these CpG islands (i.e. for the *DAPK1* CpG island, 62% of CpG sequences contain a CGG on either strand while only 16% contain CGT). However, it is unclear whether this bias reflects the CG rich nature of CpG islands in general or a feature that is selectively targeted by MBD2. To confidently assign functional significance to the mCpGG sequence preference of MBD2 will require more extensive *in vivo* analyses of methylated target promoters.

Hence, the orientation preference, we have detected could contribute to MBD2 function by (i) reflecting an affinity preference that dictates, which promoters are preferentially bound by MBD2 and/or (ii) determining the orientation of NuRD on the DNA to direct reposition of the nucleosome, thereby influencing which promoters are silenced by MBD2. The former possibility might in part explain the overlapping but distinct associations between MBD2 versus MeCP2 and specific methylated gene promoters.

Structural dynamics of MBD2

The results of the dynamic studies indicate that the β-sheet region forms a stable platform interacting with DNA. In contrast, the N- and C-terminal regions undergo internal dynamic fluctuations forming a dynamic lid that packs against the β-sheet opposite the DNA binding surface (Figure 3c). While not surprising that solvent-exposed N- and C-terminal residues are more mobile, the dynamic regions identified in MBD2 are fairly extensive and include residues involved in hydrophobic packing (i.e. L61 and F64). These observations, in conjunction with marked line-broadening seen in the ^{15}N-HSQC spectra of free MBD2, suggest that DNA contact stabilizes the core structure of the MBD. While as yet no known human disease has been associated with mutations of MBD2, some of the more common MeCP2 missense mutations associated with Rett syndrome are found in the MBD. In Figure 3b, residues in cMBD2 that are the structural equivalent of the more common MeCP2 missense mutations associated with Rett (L100, R106, R133, S134, P152, F155 and T158 of MeCP2) are depicted. Several of the mutations occur at the protein DNA interface and impact binding directly (R46, S47 and R67), as demonstrated by binding affinity analysis (Table 2). Many of these mutations, however, occur at the interface between the β-sheet and the dynamic N- and C-terminal regions (L13, K19, L61 and F64) indicating that these changes are likely to affect function by further destabilizing the more dynamic regions of the protein. We showed that the K19W mutation, which is remote from the DNA interface, does reduce binding affinity by 70-fold. These observations are consistent with recent work by Ghosh *et al.* (61), which showed that the Rett-associated mutations led to thermal destabilization and reduced DNA binding affinity of MeCP2. Hence, our data support a model in which destabilizing the packing of the N- and C-terminal regions against the β-sheet modifies interaction with DNA and disrupts function. Consistent with this model, the regions of increased internal dynamics we report for cMBD2 correlate well with the B-factors reported for the crystal structure of MeCP2-bound to methylated DNA (30). As has been suggested for destabilizing mutations of p53 (62–64), a potential molecular therapeutic approach could involve agents that bind and stabilize both the N- and C-terminal regions to overcome the functional deficits caused by these missense mutations.

Summary

The solution structure of cMBD2 MBD bound to a target-methylated sequence reveals common and unique features of how this domain recognizes the mCpG dinucleotide. This MBD has the highest affinity and greatest selectivity for methylated MBD proteins, yet a sequence preference has not been previously

demonstrated. In these studies, we show that cMBD2 MBD adopts a single orientation on a methyated *ρ-globin* promoter target sequence despite the symmetry of the mCpG. This orientation preference indicates sequence specificity primarily dependent on the bases immediately flanking the mCpG. Furthermore, binding to DNA leads to a well-structured core β-sheet packing against more dynamic N- and C-terminal regions. Both the orientation preference and internal dynamic suggest functional connections between DNA binding by this domain and recruitment of the NuRD remodeling complex in a preferential manner to specific methylated CpG-rich promoters to silence expression of the associated gene.

ACCESSION NUMBERS

The coordinates and NMR restraints for the cMBD2–dsDNA complex have been deposited in the Protein Data Bank (PDB ID: 2ky8); the NMR assignments have been deposited in the Biological Magnetic Resonance Bank (BMRB accession: 16936).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We would like to thank Dan Conrad and Joel Mathews for assistance with surface plasmon resonance analysis. NMR data were acquired using NMR instrumentation in the VCU Structural Biology Core and surface plasmon resonance data were acquired in the VCU Flow Cytometry Core.

FUNDING

Junior Faculty Scholar Award from the American Society of Hematology (to D.C.W.); NIH R01 DK29902 (to G.D.G.); NCI Cancer Center Core Support (Grant P30 CA016059). Funding for open access charge: Massey Cancer Center.

Conflict of interest statement. None declared.

REFERENCES

1. Straussman,R., Nejman,D., Roberts,D., Steinfeld,I., Blum,B., Benvenisty,N., Simon,I., Yakhini,Z. and Cedar,H. (2009) Developmental programming of CpG island methylation profiles in the human genome. *Nat. Struct. Mol. Biol.*, **16**, 564–571.
2. Illingworth,R.S. and Bird,A.P. (2009) CpG islands—'a rough guide'. *FEBS Lett.*, **583**, 1713–1720.
3. Brock,M.V., Herman,J.G. and Baylin,S.B. (2007) Cancer as a manifestation of aberrant chromatin structure. *Cancer J.*, **13**, 3–8.
4. Jiang,Y., Dunbar,A., Gondek,L.P., Mohan,S., Rataul,M., O'keefe,C., Sekeres,M., Saunthararajah,Y. and Maciejewski,J.P. (2009) Aberrant DNA methylation is a dominant mechanism in MDS progression to AML. *Blood*, **113**, 1315–1325.
5. Jones,P.A. and Baylin,S.B. (2002) The fundamental role of epigenetic events in cancer. *Nat. Rev. Genet.*, **3**, 415–428.
6. Meehan,R.R., Lewis,J.D., McKay,S., Kleiner,E.L. and Bird,A.P. (1989) Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell*, **58**, 499–507.
7. Hendrich,B. and Bird,A. (1998) Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol. Cell. Biol.*, **18**, 6538–6547.
8. Lopez-Serra,L., Ballestar,E., Fraga,M.F., Alaminos,M., Setien,F. and Esteller,M. (2006) A profile of methyl-CpG binding domain protein occupancy of hypermethylated promoter CpG islands of tumor suppressor genes in human cancer. *Cancer Res.*, **66**, 8342–8346.
9. Fatemi,M. and Wade,P.A. (2006) MBD family proteins: reading the epigenetic code. *J. Cell. Sci.*, **119**, 3033–3037.
10. Chadwick,L.H. and Wade,P.A. (2007) MeCP2 in Rett syndrome: transcriptional repressor or chromatin architectural protein? *Curr. Opin. Genet. Dev.*, **17**, 121–125.
11. Kransdorf,E.P., Wang,S.Z., Zhu,S.Z., Langston,T.B., Rupon,J.W. and Ginder,G.D. (2006) MBD2 is a critical component of a methyl cytosine-binding protein complex isolated from primary erythroid cells. *Blood*, **108**, 2836–2845.
12. McGhee,J.D. and Ginder,G.D. (1979) Specific DNA methylation sites in the vicinity of the chicken β-globin genes. *Nature*, **280**, 419–420.
13. Rupon,J.W., Wang,S.Z., Gaensler,K., Lloyd,J. and Ginder,G.D. (2006) Methyl binding domain protein 2 mediates γ-globin gene silencing in adult human βYAC transgenic mice. *Proc. Natl Acad. Sci. USA*, **103**, 6617–6622.
14. Singal,R., Wang,S.Z., Sargent,T., Zhu,S.Z. and Ginder,G.D. (2002) Methylation of promoter proximal-transcribed sequences of an embryonic globin gene inhibits transcription in primary erythroid cells and promotes formation of a cell type-specific methyl cytosine binding complex. *J. Biol. Chem.*, **277**, 1897–1905.
15. Berger,J., Sansom,O., Clarke,A. and Bird,A. (2007) MBD2 is required for correct spatial gene expression in the gut. *Mol. Cell. Biol.*, **27**, 4049–4057.
16. Hutchins,A.S., Mullen,A.C., Lee,H.W., Sykes,K.J., High,F.A., Hendrich,B.D., Bird,A.P. and Reiner,S.L. (2002) Gene silencing quantitatively controls the function of a developmental trans-activator. *Mol. Cell*, **10**, 81–91.
17. Lin,X. and Nelson,W.G. (2003) Methyl-CpG-binding domain protein-2 mediates transcriptional repression associated with hypermethylated GSTP1 CpG islands in MCF-7 breast cancer cells. *Cancer Res.*, **63**, 498–504.
18. Bakker,J., Lin,X. and Nelson,W.G. (2002) Methyl-CpG binding domain protein 2 represses transcription from hypermethylated pi-class glutathione S-transferase gene promoters in hepatocellular carcinoma cells. *J. Biol. Chem.*, **277**, 22573–22580.
19. Singal,R., van Wert,J. and Bashambu,M. (2001) Cytosine methylation represses glutathione S-transferase P1 (GSTP1) gene expression in human prostate cancer cells. *Cancer Res.*, **61**, 4820–4826.
20. Magdinier,F. and Wolffe,A.P. (2001) Selective association of the methyl-CpG binding protein MBD2 with the silent p14/p16 locus in human neoplasia. *Proc. Natl Acad. Sci. USA*, **98**, 4990–4995.
21. Qian,J., Wang,Y.L., Lin,J., Yao,D.M., Xu,W.R. and Wu,C.Y. (2009) Aberrant methylation of the death-associated protein kinase 1 (DAPK1) CpG island in chronic myeloid leukemia. *Eur. J. Haematol.*, **82**, 119–123.
22. Sidiropoulos,M., Pampalakis,G., Sotiropoulou,G., Katsaros,D. and Diamandis,E.P. (2005) Downregulation of human kallikrein 10 (KLK10/NES1) by CpG island hypermethylation in breast, ovarian and prostate cancers. *Tumour Biol.*, **26**, 324–336.
23. Chatagnon,A., Ballestar,E., Esteller,M. and Dante,R. (2010) A role for methyl-CpG binding domain protein 2 in the modulation of the estrogen response of pS2/TFF1 gene. *PLoS ONE*, **5**, e9665.
24. Klose,R.J., Sarraf,S.A., Schmiedeberg,L., McDermott,S.M., Stancheva,I. and Bird,A.P. (2005) DNA binding selectivity of MeCP2 due to a requirement for A/T sequences adjacent to methyl-CpG. *Mol. Cell*, **19**, 667–678.
25. Clouaire,T., de Las Heras,J.I., Merusi,C. and Stancheva,I. (2010) Recruitment of MBD1 to target genes requires sequence-specific

interaction of the MBD domain with methylated DNA. *Nucleic Acids Res.*, **38**, 4620–4634.
26. Sekimata,M. and Homma,Y. (2004) Sequence-specific transcriptional repression by an MBD2-interacting zinc finger protein MIZF. *Nucleic Acids Res.*, **32**, 590–597.
27. Sekimata,M., Takahashi,A., Murakami-Sekimata,A. and Homma,Y. (2001) Involvement of a novel zinc finger protein, MIZF, in transcriptional repression by interacting with a methyl-CpG-binding protein, MBD2. *J. Biol. Chem.*, **276**, 42632–42638.
28. Fraga,M.F., Ballestar,E., Montoya,G., Taysavang,P., Wade,P.A. and Esteller,M. (2003) The affinity of different MBD proteins for a specific methylated locus depends on their intrinsic binding properties. *Nucleic Acids Res.*, **31**, 1765–1774.
29. Ohki,I., Shimotake,N., Fujita,N., Jee,J., Ikegami,T., Nakao,M. and Shirakawa,M. (2001) Solution structure of the methyl-CpG binding domain of human MBD1 in complex with methylated DNA. *Cell*, **105**, 487–497.
30. Ho,K.L., McNae,I.W., Schmiedeberg,L., Klose,R.J., Bird,A.P. and Walkinshaw,M.D. (2008) MeCP2 binding to DNA depends upon hydration at methyl-CpG. *Mol. Cell*, **29**, 525–531.
31. Hendrich,B. and Tweedie,S. (2003) The methyl-CpG binding domain and the evolving role of DNA methylation in animals. *Trends Genet.*, **19**, 269–277.
32. Albalat,R. (2008) Evolution of DNA-methylation machinery: DNA methyltransferases and methyl-DNA binding proteins in the amphioxus branchiostoma floridae. *Dev. Genes Evol.*, **218**, 691–701.
33. Matsumoto,M. and Toraya,T. (2008) cDNA cloning, expression and characterization of methyl-CpG-binding domain type 2/3 proteins from starfish and sea urchin. *Gene*, **420**, 125–134.
34. Springer,N.M. and Kaeppler,S.M. (2005) Evolutionary divergence of monocot and dicot methyl-CpG-binding domain proteins. *Plant Physiol.*, **138**, 92–104.
35. Sansom,O.J., Maddison,K. and Clarke,A.R. (2007) Mechanisms of disease: Methyl-binding domain proteins as potential therapeutic targets in cancer. *Nat. Clin. Prac. Oncol.*, **4**, 305–315.
36. Cai,M., Williams,D.C. Jr, Wang,G., Lee,B.R., Peterkofsky,A. and Clore,G.M. (2003) Solution structure of the phosphoryl transfer complex between the signal-transducing protein $IIA^{Glucose}$ and the cytoplasmic domain of the glucose transporter $IICB^{Glucose}$ of the *Escherichia coli* glucose phosphotransferase system. *J. Biol. Chem.*, **278**, 25191–25206.
37. Iwahara,J., Anderson,D.E., Murphy,E.C. and Clore,G.M. (2003) EDTA-derivatized deoxythymidine as a tool for rapid determination of protein binding polarity to DNA by intermolecular paramagnetic relaxation enhancement. *J. Am. Chem. Soc.*, **125**, 6634–6635.
38. Iwahara,J., Schwieters,C.D. and Clore,G.M. (2004) Ensemble approach for NMR structure refinement against (1)H paramagnetic relaxation enhancement data arising from a flexible paramagnetic group attached to a macromolecule. *J. Am. Chem. Soc.*, **126**, 5879–5896.
39. Iwahara,J. and Clore,G.M. (2006) Detecting transient intermediates in macromolecular binding by paramagnetic NMR. *Nature*, **440**, 1227–1230.
40. 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.
41. Clore,G.M. and Gronenborn,A.M. (1998) New methods of structure refinement for macromolecular structure determination by NMR. *Proc. Natl Acad. Sci. USA*, **95**, 5891–5898.
42. Clore,G.M. and Kuszewski,J. (2002) Chi(1) rotamer populations and angles of mobile surface side chains are accurately predicted by a torsion angle database potential of mean force. *J. Am. Chem. Soc.*, **124**, 2866–2867.
43. Kuszewski,J., Gronenborn,A.M. and Clore,G.M. (1999) Improving the packing and accuracy of NMR structures with a pseudopotential for the radius of gyration. *J. Am. Chem. Soc.*, **121**, 2337–2338.
44. 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.
45. Shen,Y., Delaglio,F., Cornilescu,G. and Bax,A. (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR*, **44**, 213–223.
46. Delaglio,F., Grzesiek,S., Vuister,G.W., Zhu,G., Pfeifer,J. and Bax,A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
47. Lipari,G. and Szabo,A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules 2. analysis of experimental results. *J. Am. Chem. Soc.*, **104**, 4559–4570.
48. Lipari,G. and Szabo,A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules 1. theory and range of validity. *J. Am. Chem. Soc.*, **104**, 4546–4559.
49. Clore,G.M., Szabo,A., Bax,A., Kay,L.E., Driscoll,P.C. and Gronenborn,A.M. (1990) Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. *J. Am. Chem. Soc.*, **112**, 4989–4991.
50. Mandel,A.M., Akke,M. and Palmer,A.G. 3rd (1995) Backbone dynamics of *Escherichia coli* ribonuclease HI: correlations with structure and function in an active enzyme. *J. Mol. Biol.*, **246**, 144–163.
51. Clore,G.M., Tang,C. and Iwahara,J. (2007) Elucidating transient macromolecular interactions using paramagnetic relaxation enhancement. *Curr. Opin. Struct. Biol.*, **17**, 603–616.
52. Saito,M. and Ishikawa,F. (2002) The mCpG-binding domain of human MBD3 does not bind to mCpG but interacts with NuRD/Mi2 components HDAC1 and MTA2. *J. Biol. Chem.*, **277**, 35434–35439.
53. Clore,G.M., Tang,C. and Iwahara,J. (2007) Elucidating transient macromolecular interactions using paramagnetic relaxation enhancement. *Curr. Opin. Struct. Biol.*, **17**, 603–616.
54. Luscombe,N.M., Laskowski,R.A. and Thornton,J.M. (2001) Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. *Nucleic Acids Res.*, **29**, 2860–2874.
55. Denslow,S.A. and Wade,P.A. (2007) The human Mi-2/NuRD complex and gene regulation. *Oncogene*, **26**, 5433–5438.
56. Le Guezennec,X., Vermeulen,M., Brinkman,A.B., Hoeijmakers,W.A., Cohen,A., Lasonder,E. and Stunnenberg,H.G. (2006) MBD2/NuRD and MBD3/NuRD, two distinct complexes with different biochemical and functional properties. *Mol. Cell. Biol.*, **26**, 843–851.
57. Brehm,A., Längst,G., Kehle,J., Clapier,C.R., Imhof,A., Eberharter,A., Müller,J. and Becker,P.B. (2000) dMi-2 and ISWI chromatin remodelling factors have distinct nucleosome binding and mobilization properties. *EMBO J.*, **19**, 4332–4341.
58. Schnitzler,G.R. (2008) Control of nucleosome positions by DNA sequence and remodeling machines. *Cell Biochem. Biophys.*, **51**, 67–80.
59. Clapier,C.R. and Cairns,B.R. (2009) The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.*, **78**, 273–304.
60. Bowman,G.D. (2010) Mechanisms of ATP-dependent nucleosome sliding. *Curr. Opin. Struct. Biol.*, **20**, 73–81.
61. Ghosh,R.P., Horowitz-Scherer,R.A., Nikitina,T., Gierasch,L.M. and Woodcock,C.L. (2008) Rett syndrome-causing mutations in human MeCP2 result in diverse structural changes that impact folding and DNA interactions. *J. Biol. Chem.*, **283**, 20523–20534.
62. Basse,N., Kaar,J.L., Settanni,G., Joerger,A.C., Rutherford,T.J. and Fersht,A.R. (2010) Toward the rational design of p53-stabilizing drugs: probing the surface of the oncogenic Y220C mutant. *Chem. Biol.*, **17**, 46–56.
63. Boeckler,F.M., Joerger,A.C., Jaggi,G., Rutherford,T.J., Veprintsev,D.B. and Fersht,A.R. (2008) Targeted rescue of a destabilized mutant of p53 by an in silico screened drug. *Proc. Natl Acad. Sci. USA*, **105**, 10360–10365.
64. Joerger,A.C. and Fersht,A.R. (2007) Structure-function-rescue: the diverse nature of common p53 cancer mutants. *Oncogene*, **26**, 2226–2242.
65. Schwieters,C.D. and Clore,G.M. (2001) The VMD-XPLOR visualization package for NMR structure refinement. *J. Magn. Reson.*, **149**, 239–244.