BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **2010,** *49,* 4395–4410 4395
DOI: 10.1021/bi9019753

Unique Physical Properties and Interactions of the Domains of Methylated DNA Binding Protein 2[†]

Rajarshi P. Ghosh,[§] Tatiana Nikitina,[‡] Rachel A. Horowitz-Scherer,[‡] Lila M. Gierasch,[§,∥] Vladimir N. Uversky,[⊥,@] Kristopher Hite,[#] Jeffrey C. Hansen,[#] and Christopher L. Woodcock*[,‡,§]

[‡]*Department of Biology, University of Massachusetts, Amherst, Massachusetts 01003,* [§]*Program in Molecular and Cellular Biology, University of Massachusetts, Amherst, Massachusetts 01003,* [∥]*Department of Biochemistry and Molecular Biology and Department of Chemistry, University of Massachusetts, Amherst, Massachusetts 01003,* [⊥]*Institute for Intrinsically Disordered Protein Research, Center for Computational Biology and Bioinformatics, Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, Indiana 46202,* [@]*Institute for Biological Instrumentation, Russian Academy of Sciences, 142290 Pushchino, Moscow Region, Russia, and* [#]*Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, Colorado 80523*

Received November 16, 2009; Revised Manuscript Received April 5, 2010

ABSTRACT: Methylated DNA binding protein 2 (MeCP2) is a methyl CpG binding protein whose key role is the recognition of epigenetic information encoded in DNA methylation patterns. Mutation or misregulation of MeCP2 function leads to Rett syndrome as well as a variety of other autism spectrum disorders. Here, we have analyzed in detail the properties of six individually expressed human MeCP2 domains spanning the entire protein with emphasis on their interactions with each other, with DNA, and with nucleosomal arrays. Each domain contributes uniquely to the structure and function of the full-length protein. MeCP2 is ~60% unstructured, with nine interspersed α-molecular recognition features (α-MoRFs), which are polypeptide segments predicted to acquire secondary structure upon forming complexes with binding partners. Large increases in secondary structure content are induced in some of the isolated MeCP2 domains and in the full-length protein by binding to DNA. Interactions between some MeCP2 domains in *cis* and *trans* seen in our assays likely contribute to the structure and function of the intact protein. We also show that MeCP2 has two functional halves. The N-terminal portion contains the methylated DNA binding domain (MBD) and two highly disordered flanking domains that modulate MBD-mediated DNA binding. One of these flanking domains is also capable of autonomous DNA binding. In contrast, the C-terminal portion of the protein that harbors at least two independent DNA binding domains and a chromatin-specific binding domain is largely responsible for mediating nucleosomal array compaction and oligomerization. These findings led to new mechanistic and biochemical insights regarding the conformational modulations of this intrinsically disordered protein, and its context-dependent in vivo roles.

Methylated DNA binding protein 2 (MeCP2)[1] belongs to a family of conserved vertebrate proteins that bind to symmetrically methylated CpG dinucleotides and, at least in some cases, transmit epigenetic signals encoded in DNA methylation (reviewed in refs *1* and *2*). The finding that mutations in human MeCP2 result in Rett syndrome (RTT), a debilitating neurodevelopmental disorder (*3*), and that its misregulation is common in other patients with autism spectrum disorders (*4*, *5*) stimulated a focused effort to determine its function(s) and mechanism(s) of action. It is now clear from work with both humans and mouse models that recapitulate many of the human RTT symptoms, that MeCP2 is required for the development and maintenance of neurons in some regions of the brain (*5*). However, it has been difficult to determine the molecular event(s) that is affected by MeCP2 deficiency or by RTT-causing mutations that lead to a functional deficit(s). A complicating issue is the fact that the function of MeCP2 appears to be largely context-dependent, varying with the species, tissue and cell type, and presence of binding partners. *Xenopus* MeCP2, for example, is important in neuronal fate decisions in early embryogenesis, a function not seen in mice (*6*). In humans, MeCP2 deficits have been linked to conditions other than RTT, including cancer (*7*−*10*).

Protease resistance and modeling studies have revealed that MeCP2 is a striking example of an intrinsically unstructured protein containing protease resistant domains having varying degrees of disorder (*11*). Early work identified a short (~90-residue) ordered region of MeCP2 (between residues 75 and 164) with the ability to bind methylated DNA (*12*). This region, named the methylated DNA binding domain (MBD), is highly conserved (only four amino acid differences between *Xenopus* and humans). Its structure has been determined by both NMR and X-ray diffraction (*13*−*15*). DNA methylation is typically associated with transcriptional repression, and indeed, MeCP2 has been shown to repress methylated genes in artificial systems in vitro (*16*). This transcriptional downregulation activity was

[†]Supported by National Institutes of Health (NIH) Grant GM070897 and the International Rett Syndrome Foundation (C.L.W.); NIH Grant GM066834 (J.C.H.); NIH Grants GM027616 and OD945 (L.M.G.); and NIH Grants LM007688 and GM071714, the program of the Russian Academy of Sciences for "Molecular and Cellular Biology", and the IUPUI Signatures Center Initiative (V.N.U.).

*To whom correspondence should be addressed: Department of Biology, University of Massachusetts, Amherst, MA 01003. Telephone: (413) 545-2825. Fax: (413) 455-3243. E-mail: chris@bio.umass.edu.

[1]Abbreviations: MeCP2, methylated DNA binding protein 2; RTT, Rett syndrome; PONDR, predictor of naturally disordered regions; FlAsH, fluorescein arsenical hairpin; NA, nucleosomal array; MoRF, molecular recognition feature.

Published on Web 04/20/2010

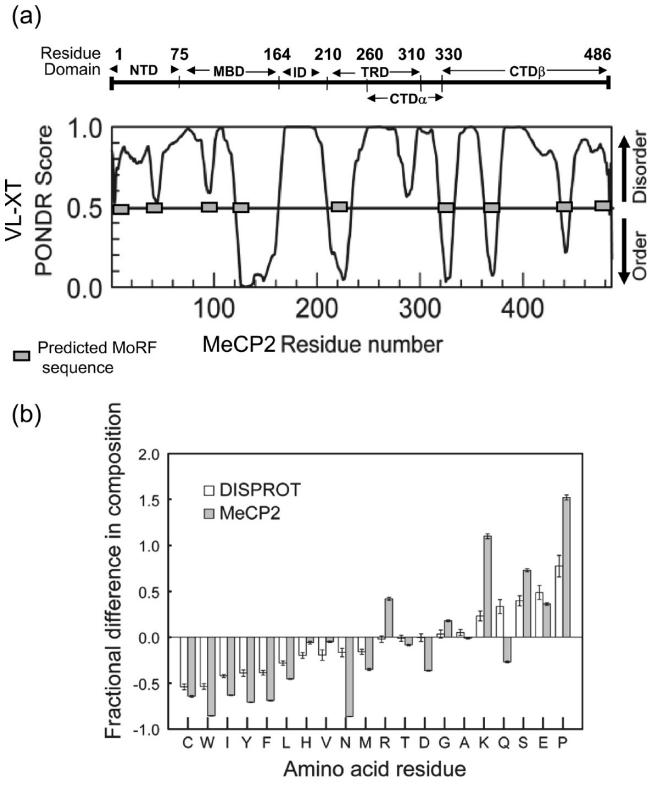

FIGURE 1: Organization of MeCP2 and relation to disorder predictions. (a) The top panel is a map of MeCP2 showing the six major domains identified by partial proteolysis (*11*). The graph shows the order−disorder score of MeCP2 predicted by PONDR VLXT, a neural network predictor of native disorder (*39*). Gray bars denote predicted molecular recognition features (MoRFs) (see Discussion). (b) The amino acid composition of MeCP2 is characteristic of a highly unstructured protein. The bar chart (gray bars) shows differences in amino acid composition between MeCP2 and the average composition of a set of ordered proteins for each amino acid. Positive values and negative values correspond to the higher and lower abundance of an amino acid, respectively, in MeCP2 compared to ordered proteins. White bars show the differences in average composition for each amino acid between disordered proteins from the DisProt database (*54*) and the same set of ordered proteins. The amino acid residues are arranged in an increasing order of disorder promoting potential (*54*). For an explanation regarding the calculation of the fractional difference in composition, see Materials and Methods.

mapped to amino acids 207−310, and the region was accordingly termed the transcriptional repression domain (TRD) (*16*). The MBD and TRD together comprise ~40% of the 486-residue hMeCP2 sequence (Figure 1a, top).

One early proposed mechanism of methylation-dependent repression involves an initial binding of MeCP2 to methylated DNA via the MBD, followed by TRD-mediated recruitment of the Sin3A corepressor and histone deacetylase complexes (HDACs). The subsequent deacetylation of histones in nucleosomes would render the local chromatin region more refractory to transcription (*17*). However, it is now clear that the mechanism of repression is often more complex, involving multiple MeCP2 binding partners. For example, in non-neuronal Rat-1 cells, repression of the neuron-specific *NaCh type II* gene by MeCP2 appears to involve at least three direct MeCP2 binding partners (methylated DNA, CoREST, and the histone H3 lysine methyl transferase SUV39H1) as well as indirect contributions from REST/NRSF and HP1 (*18*). In addition, the TRD of MeCP2 has been shown to be an important recruitment platform for several transcriptional modulators and epigenetic regulators in addition to mSin3A and HDACs. These include Ski, N-COR (*19*), DNMT1 (*20*), histone H3K9 methyltransferase (*21*), PU1 (*22*), splicing factors (*23*), BRM (*24*), RNA (*25*), and the RNA splicing machinery (*26*). The identification of numerous complexes that interact with MeCP2 suggests that additional modes of MeCP2 function remain to be discovered and calls further attention to its identity as an intrinsically unstructured protein (*11*), which characteristically has a large number of binding partners and multiple functions (*27*).

Further insight into the complexity of MeCP2 biology has come from recent genome-level studies. These revealed that MeCP2 binding is not confined to chromatin containing methylated DNA (*28*), and that MeCP2 binding can lead to both repression and upregulation depending on the gene context (*29*). Evidence that MeCP2 is involved in the maintenance of large-scale chromatin loops, perhaps by physically anchoring loop bases, has also been presented (*30*). This suggestion is consistent with the ability of MeCP2 to promote nucleosome−nucleosome interactions in vitro, a property that is enhanced by, but not dependent on, DNA methylation (*31*−*33*). These findings establish that MeCP2 is a multifunctional protein and suggest that the different functions are highly dependent on context.

With the exception of the MBD, very little is known about the structural properties of MeCP2 and how they contribute to the functional complexity of the intact protein. Studies of RTT-causing MeCP2 mutations show that the most prominent are a few missense mutations in the MBD that disrupt its structure and affect folding (*13*, *34*). There are, however, RTT-causing mutations throughout the entire molecule (see the Rett syndrome database at http://mecp2.chw.edu.au/mecp2/), indicating that regions of MeCP2 other than the MBD and TRD contribute to its multiple functions. Indeed, several reports have associated specific functions with individual regions. For example, the N-terminal domain of MeCP2 has been shown to mediate interactions with HP1 needed for transcriptional silencing during myogenic differentiation (*35*). The short (~45-residue) domain connecting the MBD and TRD has been recently shown to be instrumental in stable MeCP2 binding to chromatin in vivo (*36*), and a study of RTT patients showed that mutations tend to be located in this region (*37*). Finally, the C-terminal portion of MeCP2 required for chromatin interactions in vitro (*31*) also harbors the group II WW domain binding motif required for binding to splicing factors (*23*), and the SPxK DNA binding motif found in histone H1. The importance of the C-terminal region for MeCP2 function is underscored by the frequent occurrence of C-terminal deletions in RTT patients (*37*).

To improve our understanding of the interactions and functions of the different domains of human MeCP2, we have undertaken a systematic study of their properties, focusing on their structure, their interactions, and their DNA and chromatin binding abilities. Our work shows that the different domains are highly diverse in many respects, revealing novel properties and providing new mechanistic insights regarding the overall structure of the protein. The MBD and TRD, with their ability to bind methylated DNA and unmethylated DNA, respectively, are clearly key functional elements. Here we show that the N-terminal domain flanking the MBD modulates the affinity of MBD−DNA binding. Further, the intervening domain (ID) between the MBD and TRD possesses a strong, autonomous methylation-independent DNA binding activity and also facilitates MBD-dependent binding. We also report that some

domains exhibit a dramatic acquisition of secondary structure upon DNA binding, and while there are four autonomous DNA binding domains in MeCP2, considerable synergism exists in their mode of binding. Further, when bound to DNA, some domains increase the stability of MeCP2. Specific interdomain interactions are seen both *in cis* and *in trans*, suggesting that these physical couplings play an important role in MeCP2 structural organization and function. We have also improved our understanding of MeCP2 as an intrinsically unstructured protein and show that it has an unusually large number of interspersed molecular recognition features (MoRFs) (*38*, *39*), short regions predicted to acquire structure when in complexes with binding partners. The occurrence of several RTT-causing mutations within MoRFs further underscores their importance in MeCP2 function. Taken together, these studies significantly improve our understanding of the molecular basis of the unusual structure of MeCP2, and its relationships to DNA binding and the modulation of chromatin conformation.

MATERIALS AND METHODS

Cloning MeCp2 Domains and Linear Combinations of Domains. To construct the NTD (residues 1−90), an amplicon extending 100 bp 5′ of the NdeI site into the pTYB1 vector sequence and carrying a 3′ EcoRI linker (GACCGTGAATTC) was generated by PCR from full-length pTYB1-MeCP2 cDNA carrying MeCP2 cDNA between its NdeI and EcoRI sites, using the following primer pairs: NTD Forward (5′ CCGGTTTAAACCGGGGATCTCGATCC 3′) and NTD Reverse (5′ GTTAGAGAATTCGTCACGGATGATGGAGCGCCGCTG 3′). The forward primer used in this amplification reaction was complementary to a site ∼100 bp upstream of the MeCP2 start codon in the pTYB1 vector.

To construct the MBD (residues 75−164), ID (residues 165−210), TRD (residues 207−310), CTD-α (residues 261−330), and CTD-*β* (residues 335−486), we engineered amplicons with 5′ NdeI and 3′ EcoRI linkers and additional hexanucleotide overhangs at each end using the following primer pairs: MBD Forward (5′ CAATGACATATGGAAGCTTCTGCCTCCCCCAAACAGC 3′) and MBD Reverse (5′ GTTAGAGAATTCGCTCCCTCTCCCAGTTACCGTGAAG 3′), ID Forward (5′ CAATGACATATGCCCTCCCGGCGAGAGCAGAAACC 3′) and ID Reverse (5′ GTTAGAGAATTCCACCTGCACACCCTCTGACGTGGC 3′), TRD Forward (5′ CAATGACATATGGTGCAGGTGAAAAGGGTCCTGGAG 3′ and TRD Reverse (5′ GTTAGAGAATTCCTCCCGGGTCTTGCGCTTCTTGATG 3′), CTD-α Forward (5′ CAATGACATATGCCTCAGGCCATTCCCAAGAAACGGG 3′) and CTD-α Reverse (5′ GTTAGAGAATTCCTCACCGAGGGTGGACACCAGCAG 3′), and CTD-*β* Forward (5′ CAATGACATATGGGACTGAAGACCTGTAAGAGCCCTGG 3′) and CTD-*β* Reverse (5′ GTTAGAGAATTCGCTAACTCTCTCGGTCACGGGCGTC 3′).

The NTD−MBD (residues 1−164), MBD−ID (residues 75−210), and TRD−CTD (residues 207−486) domain combinations were engineered using the following primer pairs: NTD Forward and MBD Reverse for NTD−MBD, MBD Forward and ID Reverse for MBD−ID, TRD Forward and CTD-*β* Reverse for TRD−CTD-α−CTD-*β*. The NTD−MBD−ID triple domain was synthesized using primers 5′ CCCGGTTTAAACCGGGGATCTCGATCCCGC 3′ (forward) and 5′ TTTCAGAATTCCTGCACACCCTCTGACGTGGCCGC 3′ (reverse).

Following PCR amplification, the amplicons were doubly digested with EcoRI and NdeI and cloned into doubly digested vector pTYB1 following the standard ligation procedure.

Full-length wild-type MeCP2 and R294X were prepared as described previously (*31*).

To synthesize MBD-tetraCys, a modified pTYB1 expression vector was constructed in which the cDNA corresponding to the tetraCys sequence AEAAHRWCCPGCCKTF [GTTAGA-*GATTC*GCTGCTCATCGTTGGTGTTGTCCTGGTTGTTGTAAAACTTTT*CTCGAG*GATTGA (underlined bases constitute the hexanucleotide extension to facilitate restriction digestion, and italicized bases represent the EcoRI and XhoI site] was inserted in frame between the EcoRI and XhoI sites in the polylinker preceding the intein tag (the *Sce* VMA intein/chitin binding domain) such that insertion of the MBD amplicon between the NdeI and EcoRI sites generated the MBD-tetraCys−intein−CBD tag fusion. The tetraCys peptide with flanking HRW and KTF tripeptides was chosen because HRWCCPGCCKTF was shown to have a better quantum yield than the core CCPGCC peptide on binding to FlAsH reagent (*40*).

Protein Purification. Isoform 1 of human MeCP2 [wild type (WT) and R294X] as well as the individual domains and linked domain constructs were purified using the IMPACT system (New England Biolabs) as described previously (*31*, *34*). For the NTD and CTD-*β*, proteins were applied to heparin HP columns in 100 mM NaCl and eluted using salt steps from 0.1 to 1.0 M NaCl with increments of 0.1 M. Salt fractions containing the pure proteins were pooled and, if required, concentrated using Centricon concentrators (Amicon Inc.).

DNA and NA Preparation. Methylated and unmethylated 45 bp segments of promoter IV of the mouse brain-derived neurotrophic factor (BDNF) gene were prepared as described previously (*34*). 601-12 DNA was purified, methylated, and reconstituted into saturated/undersaturated nucleosomal arrays as described previously (*31*).

Electrophoretic Mobility Shift Assay (*EMSA*). To analyze the DNA/nucleosomal array (NA) binding efficiency and methylation specificity of the MBD, ID, MBD−ID, NTD−MBD, NTD−MBD−ID, and TRD−CTD-α−CTD-*β* constructs, we mixed and incubated unmethylated or methylated target DNA/NA (200 ng) with various amounts of polypeptide in binding buffer [100 mM NaCl, 10 mM Tris, 0.025% NP-40, and 0.25 mM EDTA (pH 7.4)] at room temperature for 30 min; 2-fold unmethylated competitor DNA or mononucleosome (400 ng) was included in experiments using DNA or NA as the substrate, respectively. Electrophoresis was performed on prechilled 1% agarose type IV gels, which were run at 85 V for 4 h at 4 °C in TAE [40 mM Tris, 24 mM acetic acid, and 0.5 mM EDTA (pH 8.3)] buffer.

To compare the DNA/NA binding efficiency of the rest of the MeCP2 fragments, methylated 601-12 DNA/NA was incubated with various amounts of protein (NTD, MBD, ID, TRD, CTD-α, CTD-*β*, or TRD−CTD-α−CTD-*β*) in the same binding buffer in the absence of competitor and electrophoresed as mentioned above. Gels were stained with ethidium bromide, photographed with the Kodak Gel 200 system, and analyzed using ImageJ. For each EMSA experiment, two to three trials were performed.

Electron Microscopy (*EM*). Sample fixation, grid preparation, and darkfield EM imaging were as described previously (*31*).

Circular Dichroism (*CD*). For each CD-related experiment, two to four trials were performed. CD spectra of domains and

their DNA complexes were acquired and analyzed as described previously (*43*). Estimates of secondary structure content were calculated using CONTINLL and reference set 7 on Dichroweb (*41*, *42*), which, in addition to a set of structured proteins (*43*, *44*), also contains five denatured proteins, aimed at moderating any possible structural bias of CONTINLL. Estimates of secondary structure content derived using CDSSTR closely resembled those from CONTINLL. We showed previously (*34*) that estimates of secondary structure in the MBD and full-length MeCP2 derived using CONTILL and LINCOMB are almost identical (*45*). To further probe the consistency and reproducibility of structural estimates derived using CONTINLL, two to four independent data acquisitions for each individual fragment were deconvolved.

To determine the nature of the spatial packing between different domains of MeCP2, we used a fragment complementation approach. For each pairwise comparison (NTD/MBD, MBD/ID, MBD/TRD, MBD/CTD-β, NTD/ID, ID/TRD, ID/CTD-β, and NTD/TRD), CD data were acquired separately for the two domains and also for their mixture, keeping the concentrations constant (*46*). Each mixture was incubated for 15 min at room temperature prior to the acquisition of data. The CD spectrum for each pair mix (A + B) was then subtracted from the spectrum obtained by addition of the individual spectra [(A) + (B)] of the constituent domains, and the difference at each wavelength was expressed as a percent of the sum of the individual spectra $\{[(A)+(B)] - (A + B)\}/[(A) + (B)]$.

Fluorescence Spectroscopy. For thermal unfolding studies, fluorescence emission spectra of MBD and polypeptides containing the MBD together with contiguous domains were collected and analyzed, and T_m values were derived as described previously (*34*). Experiments with domains alone included a recooling step to verify reversibility. DNA-containing samples used a 1:2 protein:DNA ratio. Thermal melting reversibility cannot be assessed for DNA protein complexes, and the T_m values are therefore denoted as "apparent". Data shown in Table 4 are averages of three independent sets of data.

Solvent Accessibility of Trp104 Using Acrylamide Quenching. Two or three independent fluorescence quenching measurements for each polypeptide were performed on a PTI QM1 spectrofluorometer over a 95 nm window from 305 to 400 nm using 2 nm emission and excitation slits with an integration time of 0.3 s and 0.5 nm steps. A 4 M acrylamide stock solution was prepared in buffer containing 10 mM Tris (pH 7.4), 100 mM NaCl, and 0.25 mM Na_2EDTA. Fluorescence excitation was conducted at 295 nm. At this wavelength, there is no inner filter effect due to acrylamide (*47*). In agreement with this, at 295 and 324 nm the maximum concentration of acrylamide used in the assay (310 μM) had negligible absorbance. For each acrylamide concentration, solvent-only spectra were subtracted from the data for the solvent with protein. Examination of the solvent-only spectra showed that the emission intensity was not affected by acrylamide concentration. Fluorescence quenching was assessed by the addition of varying amounts of a 4.0 M acrylamide stock solution to 2.5 μM protein in a final volume of 600 μL. For each acrylamide concentration, a separate reaction mixture was prepared and incubated for ~30 min prior to data acquisition to ensure attainment of equilibrium. There was no spontaneous quenching of Trp104 during the ~3 min data acquisition period. Overlays of spectra of the same reaction mixture after 30 min showed no visible change. Fluorescence quenching data were analyzed by using the Stern−Volmer equation:

$$F_0/F = 1 + K_{sv}Q \quad (1)$$

where F_0 and F are the initial (in the absence of quencher) and final (in the presence of quencher) fluorescence intensities, respectively, Q is the quencher concentration, and K_{sv} is the effective quenching constant. A fit of this equation to the raw data (F_0/F) yielded the K_{sv} and a y intercept of 1.

Fluorescence Labeling of MBD-tetraCys and Anisotropy. The MBD-tetraCys stock solution was incubated with 10 mM DTT for 6 h at 4 °C and then dialyzed extensively against 50 mM Tris-HCl (pH 7.8), 100 mM NaCl, 2.5 mM TCEP, 0.1 mM EDT, and 1 mM EDTA. MBD-tetraCys (~50 μM) was then incubated overnight at room temperature with 2 molar equiv of FlAsH EDT2 (LumioGreenTM labeling reagent, Invitrogen) and dialyzed extensively against the same buffer in the dark at 4 °C. A 2-fold molar excess of FlAsH label for purified tetraCys fusions has been shown to be sufficient for efficient labeling (*48*, *49*). FlAsH-EDT alone has negligible fluorescence, whereas the tetraCys peptide-bound FlAsH undergoes a boost in quantum yield in excess of 10^4-fold (*50*), indicating that the observed fluorescence originated from the bound reagent. Fluorescence anisotropy measurements were performed on a PTI QM1 spectrofluorometer equipped with an excitation and emission polarizer, using 8 nm emission and excitation slits with an integration time of 1 s. The excitation wavelength (λ_{ex}) was 500 nm, and the emission scan used a window of 10 nm (525−535 nm). A constant amount of labeled MBD (100 nM) was mixed with various amounts of other domains covering a range from 10 nM to 10 μM, and the mixture was incubated for 5 min prior to data acquisition. Fluorescence anisotropy was calculated using

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

where r is the fluorescence anisotropy of FlAsH-labeled MBD-tetraCys and I_{vv} and I_{vh} are the fluorescence intensities collected with a vertically oriented excitation polarizer and vertically (I_{vv}) and horizontally (I_{vh}) oriented emission polarizers, respectively (*47*). G is the correction factor for the difference in sensitivity of the detection system for vertically and horizontally polarized light and expressed as

$$G = I_{hv}/I_{hh}$$

where I_{hv} and I_{hh} are the fluorescence intensities collected with a horizontally oriented excitation polarizer and vertically (I_{hv}) and horizontally (I_{hh}) oriented emission polarizers, respectively. G factor correction was done for each data acquisition cycle, keeping the machine settings identical. The basal anisotropy of FlAsH-labeled MBD-tetraCys varied from ~0.1 to 0.12 between experiments. The anisotropies of complexes were normalized by dividing the anisotropies at each input concentration by the anisotropy of the MBD-tetraCys fusion in the respective experiment. Plots of normalized anisotropy versus increasing complementary unlabeled protein fragment (TRD and ID) concentration were fit to a four-parameter logistic binding model using Psi plot (*51*). The goodness of fit values for the TRD and ID were 0.998 (r^{TRD}) and 0.992 (r^{ID}), respectively. The dissociation constant $y = D + (A - D)/[1 + (X/C)^B]$, where Y is the normalized anisotropy, D is the anisotropy at an infinite concentration of the complementary MeCP2 domain, A is the anisotropy in the absence of the complementary MeCP2 domain, X is the

concentration of the different MeCp2 domains, and C is the inflection point on the fitting curve that is equivalent to the dissociation constant.

DNA Binding Affinities of MeCP2 Fragments. The blunt-ended fluorescein-labeled 22 bp duplex with a single symmetrically methylated CpG was synthesized by annealing complementary single strands of an HPLC-purified 22 bp DNA segment of mouse BDNF promoter IV: strand 1, 5′ /56-FAM/ CCCTATAA/Me-dC/GGAATTCATAATG 3′; and strand 2, 5′ CATTATGAATTC/Me-dC/GTTATAGGG.

Fluorescence anisotropy measurements were performed on a PTI QM1 spectrofluorometer equipped with an excitation and emission polarizer, using 20 nm emission and 26 nm excitation slits with an integration time of 8 s. The λ_{ex} was 480 nm, and the emission was collected over a window of 4 nm (518−521 nm). A constant amount of labeled DNA (250 pM) was mixed with various amounts of the MBD and MBD−ID, NTD−MBD, and TRD−CTD-α−CTD-β constructs covering a range from 100 pM to 60 nM, and the same DNA at 1 nM was mixed with various amounts of NTD, ID, TRD, CTD-α, and CTD-β covering a range from 1 to 600 nM and incubated for 10 min prior to data acquisition. Anisotropy was calculated at 520 nm using eq 2. Anisotropy values were normalized using the equation $r_{norm} = (r_n - r_0)/(r_{max} - r_0)$, where r_0 is the raw anisotropy without protein input, r_{max} is the raw anisotropy at maximum protein input, r_n is the raw anisotropy at each protein concentration, and r_{norm} is the corresponding normalized anisotropy. The global dissociation constant of domain DNA interaction was obtained from least-squares fits of plots of normalized fluorescence anisotropy versus protein concentration. For comparison of the binding affinity of NTD and CTD-β to those of the other domains (e.g., ID), anisotropy values were normalized using the equation $r_{norm} = r_n/r_0$.

Sedimentation Velocity. Sedimentation velocity experiments were performed with a Beckman Optima XL-I analytical ultracentrifuge using absorbance optics. 208-12 nucleosomal arrays reconstituted on methylated DNA were mixed with the appropriate MeCP2 construct in 50 mM NaCl, 10 mM Hepes, and 0.25 mM EDTA and sedimented at a velocity of 18000 rpm and 20 ± 0.1 °C. The sample absorbance was 0.7 A_{260} unit, and the molar ratio of protein to nucleosomal array ranged from 1 to 4. The data were analyzed (*52*) using the Ultrascan data analysis program. Plots of the boundary fractions against their corresponding $S_{20,w}$ values yielded the integral distribution of sedimentation coefficients. Each experiment was repeated two or three times.

Compositional Profiling. To gain insight into the relationships between sequence and disorder, the amino acid composition of MeCP2 was analyzed using an approach recently developed for intrinsically disordered proteins (*53*). To this end, the fractional difference in composition between MeCP2 [or a set of disordered proteins from the DisProt database (*54*)] and a set of ordered proteins was calculated for each amino acid residue. The fractional difference was calculated as $(C_X - C_{order})/C_{order}$, where C_X is the content of a given amino acid in a given protein (or protein set) and C_{order} is the corresponding content in a set of ordered proteins and plotted for each amino acid. In corresponding plots, the amino acids were arranged from the most order-promoting to the most disorder-promoting according to the amino acid distribution in the DisProt database (*54*). For each amino acid residue, this analysis provides two important parameters, the sign of a corresponding bar and its height. The sign of a bar reflects enrichment or depletion of a given residue in a query sequence as compared to the content of this same residue in a set of ordered proteins. The height of the bar corresponds to the value of the fractional difference, where large bars correspond to the residues whose content in the query sequence is the most different from their content in the set of ordered proteins.

Disorder, α-MoRF Prediction, and Modeling. Disorder predictions for MeCP2 were made using PONDR VLXT (*53*). Potential interaction sites, molecular recognition features (MoRFs) that gain functionality upon a disorder-to-order transition induced by binding to a partner, were identified by the α-MoRF predictor which detects short (≤20 residue) stretches within long regions of disorder with the potential for helical structure acquisition upon binding (*38*, *39*). The algorithm utilizes a stacked architecture, where PONDR VLXT is used to identify short predictions of order within long predictions of disorder, and then a second level predictor determines whether the order prediction is likely to be a binding site on the basis of attributes of both the predicted ordered region and the predicted disordered region surrounding it.

The UCSF Chimera software (www.cgl.ucsf.edu) was used to visualize the MoRF residues within the structure of a MBD-methylated DNA complex (Protein Data Bank entry 3c2i) (*15*).

RESULTS

MeCP2 Domain Nomenclature. In this study, we establish the structural and functional properties of six domains of hMeCP2 isoform 1 both individually and as contiguous fusions. With the exception of the MBD, these regions do not constitute independently evolving structural units. However, in addition to their protease resistance, there is strong evidence, discussed below, that they have unique structural and functional properties and, in that context, are considered domains. Table 1 lists the salient features of the MeCP2 polypeptides we prepared, and Figure 1a (top panel) shows their locations within the parent protein. Flanking the MBD are the N-terminal and intervening domains, termed NTD and ID, respectively. The long C-terminal domain (CTD) includes a highly protease sensitive segment, which, when cleaved, results in two fragments (*11*) denoted CTD-α and CTD-β.

Most MeCP2 Domains Are Extensively Disordered. The predicted distribution of structured and unstructured domains in MeCP2 is well-illustrated by the output of PONDR (predictor of naturally disordered regions) (*53*, *55*) (Figure 1a, bottom). A portion of the MBD contains a region predicted to adopt a stable secondary structure and is the only region for which structure is known at the atomic level (*13*, *15*). Other short segments of predicted order occur throughout the protein and are found in all the domains except the short ID (Figure 1a). This alternating pattern of disorder and order is also predicted by the FoldIndex algorithm (*11*). Unstructured proteins tend to have an amino acid composition that favors structure-disrupting residues (*53*, *54*, *56*), and MeCP2 represents an extreme case of this skewing, the level exceeding that of the unstructured proteins in the DisProt (*54*) database (Figure 1b).

To assess the inherent secondary structure content of individual domains and determine whether the distribution of secondary structure agrees with the predicted disorder−order map of MeCP2, two to four independent circular dichroism (CD) spectra were recorded for each domain (Figure 2a). Of the six domains, the MBD was the only one showing a characteristic positive band

Table 1: Domains of Human MeCP2 Used in This Study

nomenclature used here	alternate name	polypeptides prepared for this study	no. of residues
N-terminal domain (NTD)	HMGD1 (*11*)	1−90	90
methylated DNA binding domain (MBD) (*34*)		75−164	90
intermediate domain (ID) (*36*)	HMGD2 (*11*)	165−210	46
transcriptional repression domain (TRD) (*16*)		207−310	104
C-terminal domain (CTD)		261−330	70
CTD-α			
CTD-*β*		335−486	156
NTD−MBD		1−164	164
MBD−ID		75−210	136
NTD−MBD−ID		1−208	208
TRD−CTD		207−486	280
$MeCP2^{1-294}$		1−294	294

in its CD spectrum at ~197 nm, indicative of significant ordered secondary structure (Figure 2a). The others had a negative band in this region, indicating extensive disorder, with the NTD and TRD being the most strongly disordered. Estimates of the different types of secondary structure were obtained using CONTINLL (*34*, *41*, *42*) deconvolution (see Materials and Methods for details). Deconvolution produced highly reproducible estimates of secondary structure for each domain (Table 2).

In agreement with previous findings (*11*, *13*, *34*), the MBD is ~60% structured, with ~45% *β*-sheet and turn and ~15% α-helix. The proportion of predicted unstructured sequence for the NTD, ID, and CTD ranged from ~62 to ~78% (Table 2), consistent with the net 60% unstructured sequence in intact MeCP2 (*11*, *34*). The amount of disorder in each of the domains determined by CD (Table 2) is close to that predicted by the PONDR (Figure 1a) and FoldIndex (*11*) algorithms.

MeCP2 Domains Differ in Their Ability To Bind DNA and Chromatin. Native electrophoretic mobility shift assays (EMSAs) provide a qualitative estimate of the DNA and chromatin interaction properties of MeCP2 (*31*, *32*), allowing the exploration of a wide range of DNA:protein ratios. The distribution of shifted species provides some indication of the nature of the interactions involved. We previously reported strikingly large mobility shifts when full-length MeCP2 interacts with DNA and nucleosomal arrays (NAs) (*31*). Here, we dissect the extent of these mobility shifts domain by domain. Substrates consisted of tandem (n = 12) arrays of a 207 bp sequence containing the "601" nucleosome positioning sequence (*57*), either as naked DNA or after reconstitution with core histones to yield 12-mer nucleosomal arrays (NAs). Input ratios are expressed as moles of protein per nucleosome or 207 bp DNA. Since each 207 bp fragment contains 18 methylatable CpG sequences, the highest molar ratio of peptides used here in the absence of competitor is approximately equivalent to one polypeptide per two methyl CpG sequences. The higher input ratios of peptides in EMSA and EM experiments were used to simulate a situation of local enrichment of peptides which may occur in cases such as MeCP2-regulated BDNF (brain-derived neurotropic factor) promoter III which contains a region of closely spaced CpG sequences (*58*). Also, we have found that closely spaced methylated CpG sequences favor cooperative DNA binding of MeCP2 (R. P. Ghosh and C. L. Woodcock, manuscript in preparation).

Full-length MeCP2 induces a pronounced shift with both DNA and NAs even at low molar inputs of the protein. These shifts are enhanced in the case of methylated substrates (Figure 2b, far right). Among the individual domains, the NTD and CTD-*β* stand out as inducing only very minor mobility shifts in DNA, whereas the ID, TRD, and CTD-α fragments all induce marked shifts (Figure 3a). The shifts induced by these three domains are methylation-independent (Figure 3c), accounting for the substantial methylation-independent binding observed with intact MeCP2 (Figure 3b) (*11*, *32*, *33*). Further support for DNA binding by the ID comes from the finding that a R188E MeCP2 mutant has a significantly lowered gel mobility compared to the wild type (data not shown). A slightly different fragment containing the TRD (residues 198−305) also has been shown to bind DNA (*11*). In general, the shifts with DNA and NAs are qualitatively similar. However, the CTD-*β* is a clear exception, inducing a moderate but reproducible shift with NAs but not with naked DNA (Figure 3a,c). This result is consistent with earlier findings that deletion of 192 residues of the C-terminal portion of MeCP2 results in deficient NA compaction and oligomerization (*31*) and suggests that the CTD-*β* contains a unique histone binding region(s).

The DNA and NA Binding Properties of the MBD and CTD Are Modulated by Their Flanking Domains. To determine if there were interdomain effects, we compared the DNA and chromatin interactions of several constructs comprising multiple contiguous MeCP2 domains. (Note that throughout, we use dashes to denote constructs encompassing adjacent MeCP2 domains yielding information about DNA and chromatin interactions in *cis* and the plus symbol to denote different MeCP2 domains combined in solution and providing information on *trans* interactions.) The TRD−CTD fusion comprising the 280 C-terminal residues of MeCP2 promotes pronounced shifts with both DNA (Figure 3a, far right) and NAs (Figure 3d, far right). This is consistent with DNA binding by the TRD and CTD-α, and chromatin binding by the CTD-*β* (Figure 3a,d). The gel shifts seen with the TRD−CTD construct reflect the additive binding effect of the constituent domains and are consistent with the concerted binding by these domains that would be required for chromatin condensation and/or oligomerization (also see Figures 4 and 5).

For the MBD-containing constructs, it was important to compare gel shifts obtained with methylated and unmethylated DNA or chromatin substrates. For these experiments, we used a 2-fold excess of unmethylated 207 bp DNA or mononucleosome competitor to enhance methylation-dependent effects (*31*, *34*). With this level of competitor, higher protein:DNA ratios are needed to observe significant mobility shifts. As a control, the MBD alone shows a reproducible methylation-dependent enhancement in mobility shift with methylated DNA (Figure 3b) and chromatin (Figure 3d, far left). The NTD−MBD construct

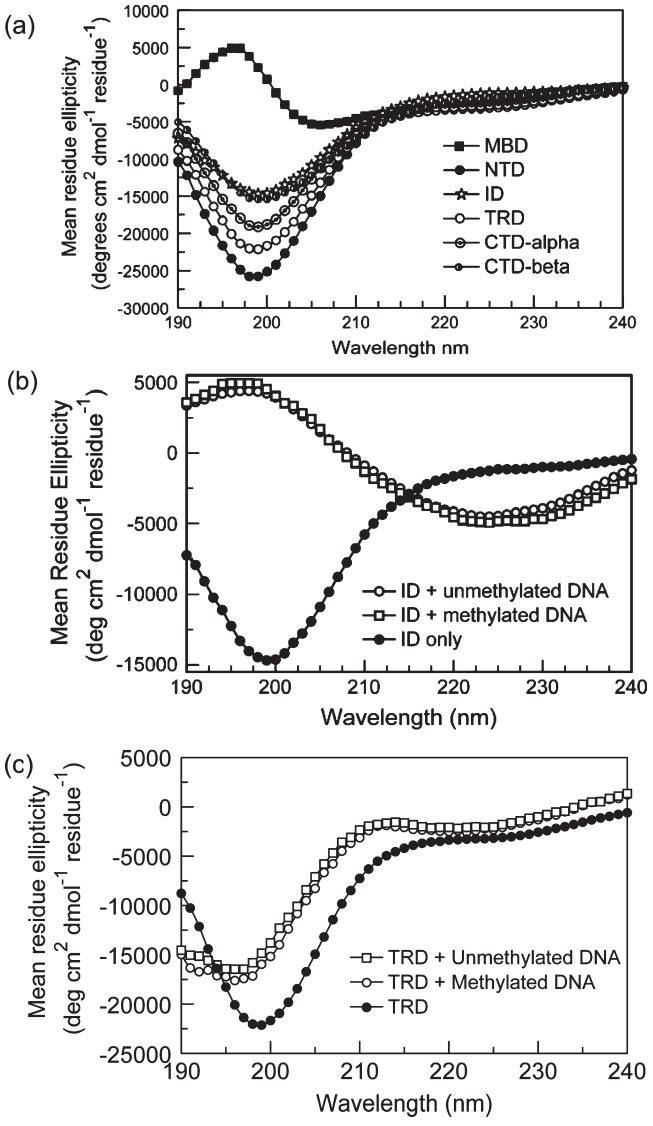

FIGURE 2: Circular dichroism spectra of MeCP2 domains reveal marked differences in secondary structure content. CD spectra are representative of two to four separate acquisitions. (a) Compared to the 195 nm peak indicative of β-sheet structure within the MBD (■), all the other domains show a negative band in the 195−198 nm region indicative of disorder. NTD (●), TRD (○), and CTD-α (⊙) have lower structure content than ID (☆) and CTD-β (half-filled circles) (see Table 1 for quantitation). (b) Addition of DNA (methylated as well as unmethylated) to the ID induces changes typical of the formation of β-structure, namely a marked increase in positive ellipticity at 195 nm and negative ellipticity in the 220−225 nm range. (c) Addition of DNA (methylated as well as unmethylated) to the TRD results in an increase in the level of order irrespective of the methylation status of the DNA.

exhibited methylation-independent and -dependent binding to both DNA and chromatin, but surprisingly, the observed mobility shifts were much larger than expected from a mere sum of the moderate shift induced by the MBD and zero shift induced by the NTD (Figure 3a,b,d,e). This synergistic shift enhancement suggests a strong structural and/or functional coupling between the NTD and MBD eliciting a binding mode uncharacteristic of either of the individual domains. This phenomenon cannot be explained by a simple increase in the level of electrostatic shielding since the net charge of the MBD (5 at pH 7.4) is higher than that of the NTD−MBD construct (4 at pH 7.4) in the reaction buffer. Thus, while the MBD may be the minimal

Table 2: Secondary Structure Content of MeCP2 Domains Based on CONTINLL Deconvolution of CD Data

	% ordered secondary structure (standard error)				
domain	α-helix	β-strand	β-turn	total	% unstructured (standard error)
NTD	9 (0.3)	4 (0.3)	9 (0)	22	78 (0.6)
MBD (*34*)	15 (0.3)	27 (0.3)	18 (0.3)	60	40 (0.3)
ID	9 (0.9)	16 (0.33)	13 (0.6)	38	62 (0.6)
TRD	8 (0.3)	3 (0.8)	11 (0.4)	22	78 (0.5)
CTD-α	7 (0.5)	12 (1.0)	12 (0)	31	69 (0.5)
CTD-β	9 (0.5)	16 (0)	15 (0)	40	60 (0.5)

domain necessary for the recognition of methylated DNA, MBD function is clearly impacted by the adjacent non-DNA binding NTD in a way that is likely to enhance its DNA binding affinity (see later). The EMSA patterns obtained with the MBD−ID construct were also different from those with the MBD, with significant smearing (Figure 3b). This likely reflects formation of nonspecific higher-order complexes via cross-linking by the MBD−ID fragment, with its two independent DNA-binding regions. The NTD−MBD−ID construct interacted with DNA and chromatin much like the NTD−MBD construct, although the gel shifts were slightly more pronounced with the longer construct (Figure 3b,e). It should be stressed that methylation-enhanced shifts occur only with MBD-containing polypeptides. Examples of this are shown in Figure 3c, where, for the ID and the TRD−CTD fusion (TRD−CTD-α−CTD-β), the methylation state of the DNA has no effect on the induced shift.

Contiguous Fusions of Certain MeCP2 Domains Induce Condensation of NAs. Full-length MeCP2 is a potent chromatin architectural protein, inducing extensive compaction and self-association of nucleosomal arrays (*31*, *32*). To determine which domains of MeCP2 were important for this phenomenon, we first investigated changes in sedimentation velocity of defined nucleosomal arrays (NAs), which provide a sensitive and quantitative assessment of their state of compaction (*59*). We prepared methylated 601-12 NAs and measured the influence of individual domains and multidomain constructs on their sedimentation properties. MeCP2 domains and constructs comprising multiple domains, shown by an EMSA to interact with chromatin, were mixed with methylated 601-12 NAs, and diffusion-corrected sedimentation coefficient distributions were obtained by analysis of sedimentation boundaries using the van Holde−Weischet method that is particularly well suited for polydisperse systems such as population of nucleosomal arrays with multiple compaction states (*52*). Results from two or three independent experiments show that at the ionic strength used in these experiments (50 mM NaCl), NAs alone give a sedimentation profile typical of a slightly folded conformation, with a nearly homogeneous population between boundary fractions of ~20 and ~70%, and an average $S_{20,w}$ value of ~32 ± 2 S (Figure 4). Full-length MeCP2 at a 1:1 input ratio induced a dramatic increase to ~61 ± 2 S (Figure 4), similar to our earlier observation in the 5S 208-12 NA system (*32*). Higher input ratios of the full-length protein result in the formation of rapidly sedimenting complexes due to MeCP2-mediated self-association of NAs (data not shown).

For the MBD and constructs that included its flanking domains, the sedimentation coefficient distributions obtained at an input ratio of two polypeptides per nucleosome are presented in Figure 4. The MBD alone increased the sedimentation coefficient by only ~2 ± 1 S, indicating binding but little or no

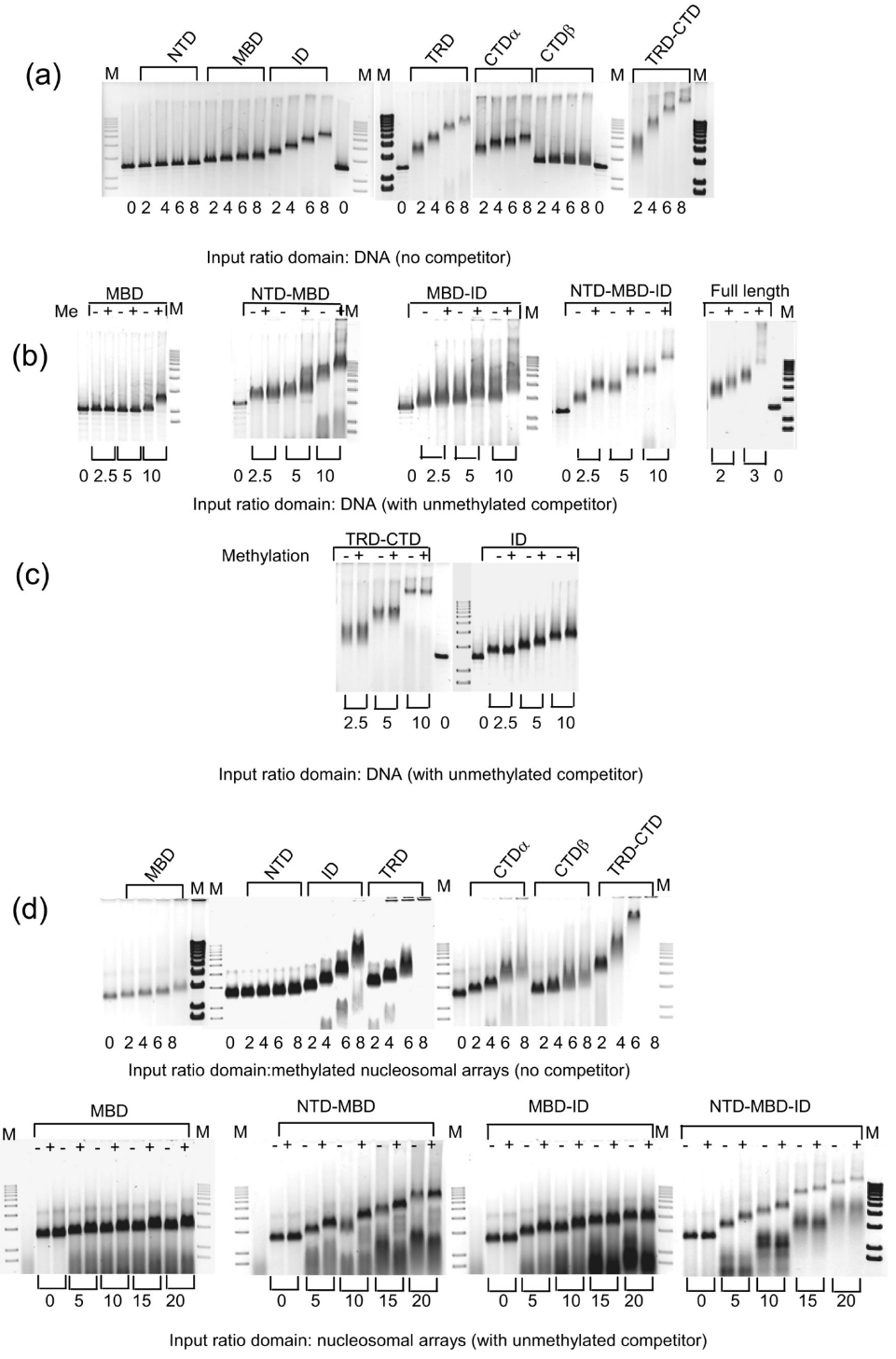

FIGURE 3: MeCP2 domains induce electrophoretic mobility shifts upon addition to DNA or chromatin. Gel images for each experiment are representative of two to three separate trials. (a) Interaction between individual MeCP2 domains and DNA. Domains (NTD, MBD, ID, TRD, CTD-α, CTD-*β*, and TRD−CTD-α−CTD-*β*) were incubated with methylated 601-12 DNA at molar input ratios of 0−8, and the products are displayed on 1% agarose gels. The ID, TRD, and CTD-α induce substantial retardation of the DNA. In contrast, the MBD shows only minor shifts and the NTD appears to have virtually no interaction with DNA. (b) To examine methylation specificity, the MBD and constructs that include its flanking domains were incubated with unmethylated (−) or methylated (+) DNA in the presence of a 2-fold excess of 208-1 DNA competitor at molar input ratios of 0−10. A distinct methylation-dependent enhancement of the gel shifts is seen in all constructs containing the MBD. Of particular interest is the large shift shown by the NTD−MBD construct, which suggests a synergism between these two domains. Full-length MeCP2 produces a pronounced gel shift at a much lower input than the MBD-containing contiguous domain fusions. (c) The ID and TRD−CTD polypeptides produce strong shifts, but there is no methylation-dependent enhancement. (d and e) Same as panels a and b, respectively, but with 601-12 nucleosomal arrays (NAs) as the substrate and 208-1 mononucleosomes as the competitor. With the exception of the CTD-*β* which induces a moderate but consistent mobility shift with chromatin but not with naked DNA, the patterns of electrophoretic shift with DNA and NAs are similar. M denotes molecular weight marker lanes.

array compaction (binding of two MBD molecules to each nucleosomal unit in an array results in an ~10% increase in mass for the complex and would be expected to increase the array sedimentation coefficient by a few siemens). The NTD−MBD construct, which produced a prominent electrophoretic mobility shift (Figure 3d), caused an increase in sedimentation coefficient

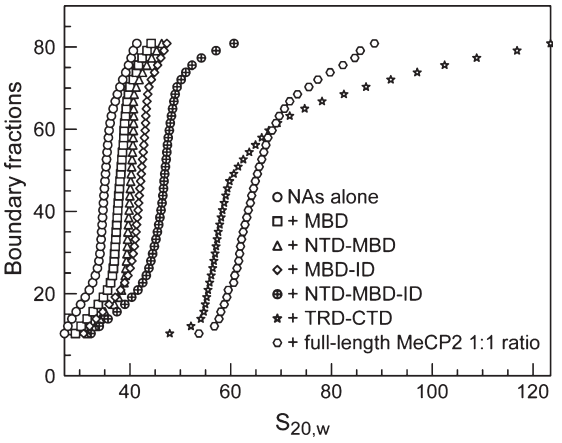

FIGURE 4: Sedimentation velocity reveals differences in the ability of MeCP2 domains to compact nucleosomal arrays. Methylated 601-12 nucleosomal arrays were incubated with a 2-fold molar input of MeCP2 constructs in 50 mM NaCl, 10 mM Hepes, and 0.25 mM EDTA and analyzed by sedimentation velocity: NAs alone (○), MBD (□), NTD−MBD (△), MBD−ID (◇), NTD−MBD−ID (⊕), TRD−CTD (☆), and full-length MeCP2 (⬡). For characterization of full-length MeCP2, an equimolar input of protein was used since a 2-fold input causes extensive self-association and oligomerization. Data were consistent over two or three separate trials for each experiment.

of ~4.5 ± 0.5 S, whereas the MBD−ID construct resulted in a sedimentation coefficient increase of ~6 ± 1 S. These results show that each of these fragments binds to nucleosomal arrays but induces only small increases in array compaction. A more substantial increase in compaction of ~11 ± 1 S was seen with the NTD−MBD−ID construct (Figure 4), and the most striking result was obtained with the TRD−CTD construct. Binding of the latter caused an increase of ~26 ± 3 S in the homogeneous segment of the population, indicating a level of NA compaction similar to that caused by full-length MeCP2, albeit at double the protein input. Binding of the TRD−CTD fragment also produced a significant fraction of heterogeneous self-associated arrays (Figure 4, boundary fraction of > 60%), as did full-length MeCP2. These results suggest that the TRD−CTD fragment may be able to recapitulate the chromatin condensing functions of the full-length protein. At a 4-fold molar input of domains and domain fusions, the relative differences in levels of compaction of NAs remained largely the same, although as expected the absolute sedimentation values increased (data not shown).

Different Domains of MeCP2 Induce Distinctive Changes in NA Morphology. Observation by electron microscopy (EM) of the compaction state of individual NAs provides a direct measure of the impact of protein binding on array morphology and complements the EMSA and sedimentation studies. From the images, it is possible to compare compaction effects quantitatively and examine the initial changes at the nucleosome and linker DNA level that lead to overall increases in the level of condensation. For these purposes, it is useful to employ "subsaturated" NAs in which the linker remains visible during the initial stages of compaction (*31*). Hence, in this study, we reconstituted methylated 601 DNA templates with six to eight (rather than 12) nucleosomes, exposed them to the defined fragments of MeCP2 at various input ratios, and imaged the resulting complexes using darkfield EM.

In the absence of protein, the NAs were fully extended (Figure 5a). Little change in conformation was seen with either the MBD or the NTD alone (Figure 5b,c), consistent with the

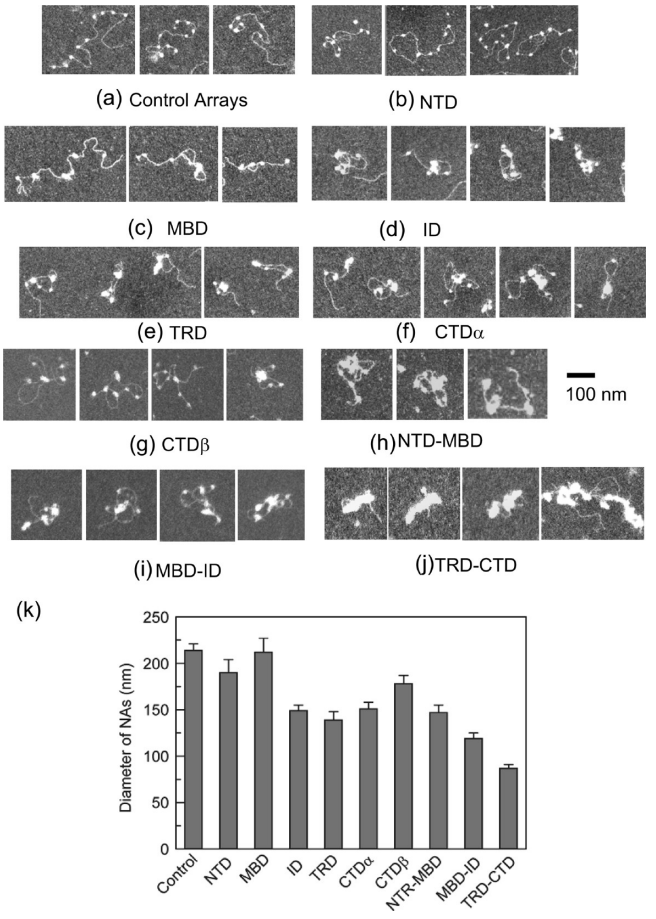

FIGURE 5: Direct EM observation reveals differences in conformational changes induced in undersaturated nucleosomal arrays by MeCP2 domains. Subsaturated NAs were mixed with different MeCP2 fragments at input ratios of eight molecules of protein per 208 bp DNA, fixed, and imaged using darkfield EM. (a−i) Representative images of NAs showing the range of conformational changes from none for the MBD and NTD to extensive compaction and self-association for the TRD−CTD fusion. (k) Mean array diameters with standard errors.

EMSA data on saturated NAs (Figure 3). All the other individual domains induced partial clustering of nucleosomes within arrays (Figure 5d−g), leading to significant ($p < 0.001$) reductions in array diameter [diameter of the smallest circle that fully encloses the array (Figure 5k)]. The three contiguous constructs examined (MBD−ID, NTD−MBD, and TRD−CTD) also exhibited nucleosome clustering and array compaction (Figure 5h−j), with decreased array diameters (Figure 5k) that parallel the increased sedimentation coefficients seen in Figure 4. Consistent with its potency in inducing gel shifts and increasing sedimentation velocity, the TRD−CTD construct was the most effective at compaction, with a mean diameter significantly ($p < 0.0001$) smaller than those of all of the other MeCP2 fragments. In addition to compacted individual arrays, the TRD−CTD construct induced self-association of arrays (Figure 5j, far right), consistent with our observation of a large fraction of heterogeneous rapidly sedimenting material (Figure 4). The full-length protein induces extensive array oligomerization under these conditions (*32*), precluding accurate measurements of array compaction.

The MBD Is Structurally Coupled to Other MeCP2 Domains. In a previous study focusing on the properties of the MBD, we used fluorescence spectroscopy to monitor the solvent

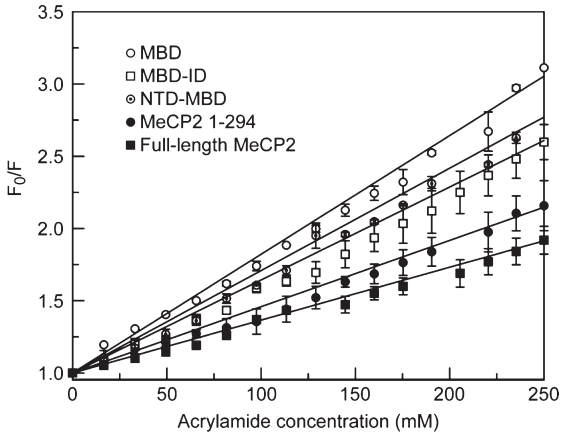

FIGURE 6: Interactions between the MBD and flanking domains revealed by tryptophan accessibility. Fluorescence quenching by acrylamide analyzed using Stern−Volmer plots shows that the fluorescence of Trp104 in the MBD is differentially accessible depending on the flanking domains present: MBD only (○). NTD−MBD (⊙), MBD−ID (□), MeCP2 residues 1−294 (●), and full-length MeCP2 (■). Plots are linear up to ~250 mM acrylamide. Error bars represent standard errors of mean.

exposure (*47*) of the single tryptophan at position 104 and reported that domains flanking the MBD provided solvent protection to W104 (*34*). Here, we examine this effect in more detail, using the collisional quenching agent acrylamide to avoid complications due to the differing fragment sizes, and hence tumbling rates. The extent of quenching of W104 by acrylamide provides a direct measurement of its solvent accessibility and allows the identification of flanking domains that provide protection from quenching *in cis*. The results of two or three independent measurements for each fragment (Figure 6) show that acrylamide most effectively quenches the fluorescence signal of tryptophan in the MBD alone (quenching constant K_{sv} = 8.4 M^{-1}) and least effectively in full-length MeCP2 (K_{sv} = 3.9 M^{-1}). Intermediate levels of quenching (K_{sv} = 7.3 and 6.8 M^{-1}) were obtained for the NTD−MBD and MBD−ID constructs, respectively. Tryptophan fluorescence from the truncation RTT mutant R294X, which is approximately equivalent to the NTD−MBD−ID−TRD construct, exhibited greater protection from quenching than other constructs (K_{sv} = 4.9 M^{-1}) but was less effective than the full-length protein. These results indicate that within the full-length protein, the structure of the MBD is influenced by associations with other domains that lead to the shielding of W104 from solvent exposure. All the MeCP2 domains contribute to this shielding (particularly NTD, ID, and TRD), suggesting that interdomain coupling occurs within the overall structure of the full-length protein. In this respect, it is interesting to note that an NMR study of a region of the chicken homologue of MeCP2 approximately equivalent to the human MBD−ID construct suggested that it may act as a platform for interaction with other regions of MeCP2, or the binding of other proteins (*60*).

Interdomain Coupling Occurs in trans and Affects Secondary Structure. Interdomain interactions that are crucial to the stability of the native state of the protein are often strong enough to be sustained *in trans* in mixtures of protein fragments (*61*). Fragment complementation approaches to defining tertiary organization of domains are particularly helpful for proteins refractory to crystallization as is the case with MeCP2. To test for such associations between MeCP2 domains, we prepared a construct in which a tetracysteine motif was fused to

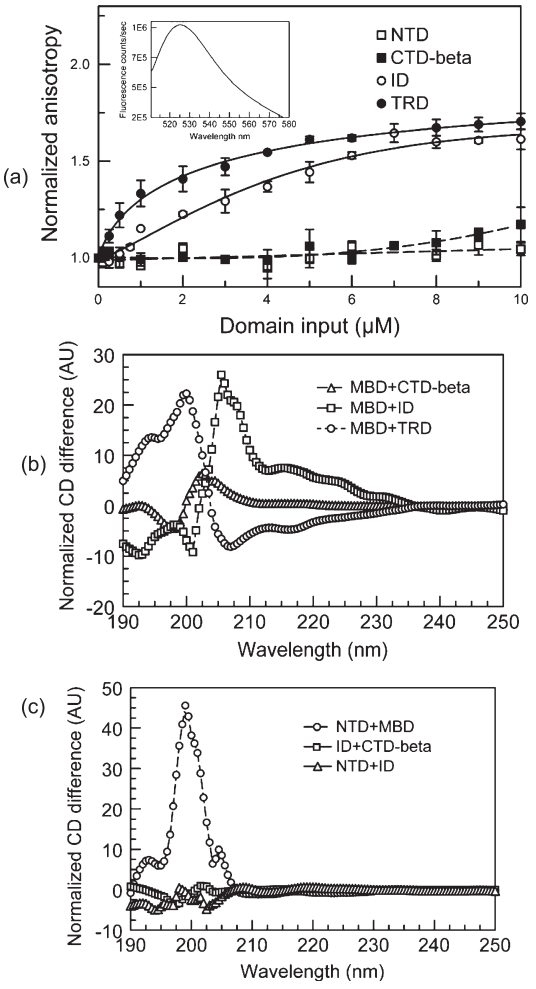

FIGURE 7: *In trans* interactions between MeCP2 domains revealed by fluorescence anisotropy and CD. (a) Fluorescence anisotropy of the fluorescently labeled MBD upon mixing with other MeCP2 domain constructs. In the presence of the ID (○) and TRD (●), the MBD shows a marked increase in anisotropy, whereas addition of the NTD (□) caused no change in anisotropy. The CTD-β (■) gave a small increase at higher input ratios. Error bars denote standard errors of mean. The inset shows the robust emission spectrum of the MBD-tetraCys-bound FlAsH complex. (b and c) Fragment complementation was also detected by using CD to monitor interactions between domain pairs caused by changes in secondary structure. For each pairwise comparison (A:B), data were acquired separately for the two different domains (A and B) and also for their mixture (A + B). Plots show the difference spectra at each wavelength expressed as a percent of the spectrum obtained by addition of the individual spectra {[(A) + (B)] − (A + B)}/[(A) + (B)]. (b) MBD−ID (□) and MBD−TRD (○) spectra show strong differences from the composite spectrum of the individual domains, while the MBD+CTD-β pair shows only a minor change. (c) The NTD+MBD pair (○) shows a distinct difference at ~198 nm, while NTD+ID (△) and ID+CTD-β (□) pairs show no differences.

the MBD and labeled with the FlAsH reagent (*50*, *62*) at a level that produced a robust emission (Figure 7A, inset). The labeled MBD was titrated with a second domain and evaluated using fluorescence anisotropy for interdomain interactions that cause it to tumble more slowly. The results indicated that the ID and TRD were able to bind the MBD in solution, with dissociation constants of ~4 and ~2.5 μM, respectively (Figure 7a). In contrast, no interactions were observed between the MBD and NTD. A weak effect was seen with the CTD-β, but only at high molar inputs.

We also investigated interdomain associations *in trans* using CD, which can detect changes in secondary structure of one or

Table 3: Changes in Secondary Structure of MeCP2 Domains upon DNA Binding

domain	% ordered secondary structure without DNA	with DNA
full-length MeCP2	35[a]	42[a]
NTD	22	23
MBD	60[a]	66[a]
ID	38	59
TRD	22	30
CTD-α	31	30

[a]From ref *34*.

both components in a mixture of the two (*63*). For each pairwise comparison, data were acquired separately for the individual domains and also for their mixture. With the NTD+MBD, MBD+ID, and MBD+TRD mixtures, we observed a clear difference between the summed individual CD spectra and the spectrum of an equimolar mixture (Figure 7b,c), indicating that an interaction between domains led to change(s) in secondary structure. The failure to observe an interaction between the MBD and NTD with anisotropy may be due to the location of the tetra-Cys moiety at the largely unstructured C-terminus of the MBD. Also, binding by the NTD distal to the C-terminus may not result in changes in the rotational freedom of the label and thus not affect anisotropy.

The CD approach also allows domain interactions not involving the MBD to be investigated. None of the domain combinations that lacked the MBD gave any evidence of interaction. Examples are shown for the NTD+ID and ID+CTD-β pairs (Figure 7c). Thus, it appears that the role of the MBD is not solely as a methylation-dependent DNA binding domain; it also acts as a structural "core" of the protein, participating in multiple interdomain interactions (see Discussion).

Impact of DNA Binding on Secondary Structure. Intrinsically disordered proteins often undergo binding-induced increases in secondary structure content. Given the multiple DNA-binding domains and the large degree of intrinsic disorder spread throughout MeCP2, we were interested in determining whether DNA binding leads to structural alterations. For DNA binding experiments, we selected a 45 bp segment of the brain-derived neurotropic factor (BDNF) III promoter that has a single CpG unit and is known to be a target for in vivo MeCP2 binding (*58*, *64*). A similar stretch of BDNF DNA was used in the MeCP2 complex for which the X-ray structure has been determined (*15*). The ID and TRD, when mixed with an equimolar amount of the 45 bp DNA substrate, resulted in striking changes in far-UV CD profiles, independent of the DNA methylation state (Figure 2b,c). These changes result from significant increases in secondary structure content [from ~38 to ~59% for the ID and from ~22 to ~30% for the TRD (Table 3)]. For the ID, the acquired structure is approximately equally divided between α-helix and β-strand, while for the TRD, the increase is in the β-strand component. As previously reported, in the presence of DNA, the MBD shows no change in secondary structure with unmethylated DNA but acquires a methylation-dependent ~6% increase in α-helix content (*34*).

Domains of MeCP2 Differ in Their Affinity for DNA and Contribution to Thermal Stabilization upon DNA Binding. To measure the affinity between DNA and different MeCP2 domains, we used a 22 bp segment of the BDNF promoter containing a single centrally located methylated CpG dinucleotide and a fluorescein label at one end. DNA binding by a given protein fragment reduces the DNA tumbling rate in solution and can be measured by following changes in the steady-state anisotropy of the fluorescein label. Figure 8a–c shows the changes in anisotropy as a function of concentration of MeCP2 domains, and constructs with linked domains. Among the MBD and MBD-containing contiguous domain fusions, the NTD–MBD construct exhibited the highest DNA binding affinity (K_d = 0.8 nM), 10-fold higher than that of the MBD

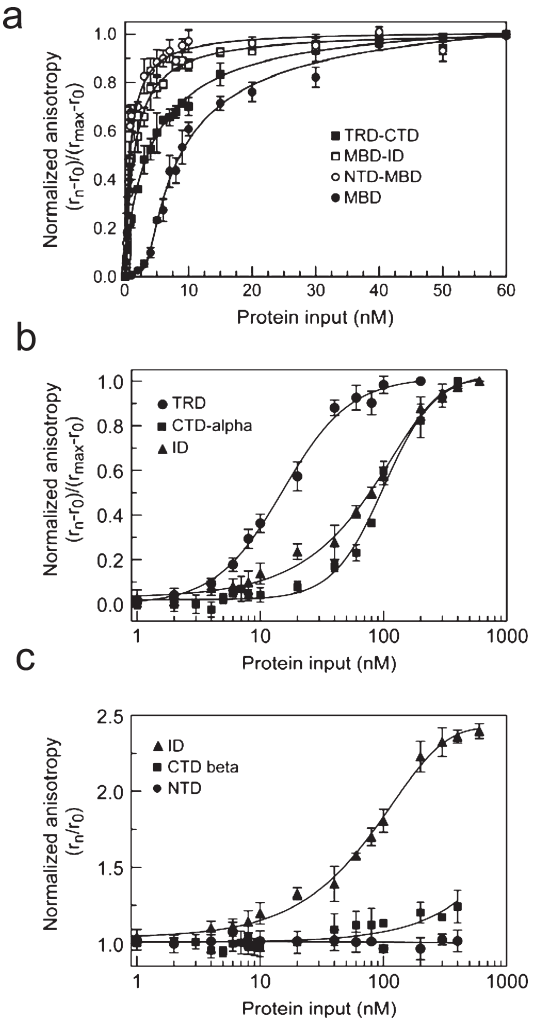

FIGURE 8: Quantitation of DNA binding affinity of MeCP2 fragments. (a) The normalized fluorescence anisotropy [$r_{norm} = (r_n - r_0)/(r_{max} - r_0)$] of a 5′-fluorescein-labeled 22 bp fragment of BDNF promoter DNA with a single methylated CpG was measured in the presence of increasing concentrations of MeCP2 fragments: MBD (■), NTD−MBD (○), MBD−ID (●), and TRD−CTD (□). Error bars denote standard errors of the mean. The *X* axis (protein concentration) and *Y* axis (normalized fluorescence anisotropy) are linear normal. (b) Same as panel a but with different MeCP2 domains: ID (▲), TRD (●), and CTD-α (■). Error bars denote standard errors of the mean. The *X* axis (protein concentration) is log decimal, and the *Y* axis (normalized fluorescence anisotropy) is linear normal. (c) Normalized fluorescence anisotropy ($r_{norm} = r_n/r_0$) of the same DNA substrate as in panels a and b in the presence of increasing concentrations of the NTD (●), ID (▲), and CTD-β (■). The ID is included in both panel b and panel c to provide a reference scale for the two different types of normalizations used in panels b and c. Error bars denote standard errors of the mean. r_0 is the raw anisotropy with no protein input. r_{max} is the raw anisotropy at maximum protein input. r_n is the raw anisotropy at each protein concentration. r_{norm} is the corresponding normalized anisotropy.

Table 4: Thermal Stability of MeCP2[a]

MeCP2 polypeptide	T_m alone (°C)	T_m(apparent) with unmethylated DNA (°C)	$\Delta T_{m(unmet)}$	T_m(apparent) with methylated DNA (°C)	$\Delta T_{m(met)}$
MBD^{75-164}	44.9 ± 0.1	46.6 ± 0.6	1.7	54.3 ± 0.3	9.4
$NTD-MBD^{1-164}$	40.4 ± 0.6	46.3 ± 0.1	5.9	54.5 ± 0.3	14
$MBD-ID^{75-210}$	46.7 ± 0.2	55.7 ± 0.3	9	63.8 ± 0.2	17.1
$NTD-MBD-ID^{1-208}$	43.1 ± 0.1	58.6 ± 0.4	15.5	67.1 ± 0.2	24
$MeCP2^{1-294}$	45.3 ± 0.1	56.3 ± 0.3	11	64.6 ± 0.2	19.3
full-length $MeCP2^{1-486}$	44.5 ± 0.2	55.7 ± 0.5	11.2	63.1 ± 0.4	18.6

[a] ΔT_m is the increase in the melting temperature upon DNA binding. T_m values in the presence of DNA are denoted apparent since reversibility cannot be tested.

alone which exhibited the weakest binding (K_d = 8.5 nM) (Figure 8a,c). The MBD−ID construct (K_d = 1.4 nM) bound to DNA with a 6-fold higher affinity than the MBD and a ~50-fold higher affinity than the ID alone (K_d = 75 nM) (Figure 8a,b). This clearly shows that while the ID contains an autonomous DNA binding domain, it also facilitates MBD-mediated binding. Unlike the ID, the NTD does not bind to DNA as an isolated domain (Figure 8c) but when coupled to the MBD markedly enhances its binding affinity, consistent with our EMSA data.

The constituent domains of TRD−CTD fusions showed considerable variability in their DNA binding affinity. Unlike the TRD and CTD-α, which bound with affinities of 20 and 96 nM, respectively (Figure 8b), the CTD-β induced small increases in fluorescence anisotropy and only at very high protein concentrations (~600 nM) (Figure 8c). Interestingly, the TRD−CTD fusion construct bound DNA more strongly (K_d = 3.6 nM) than its constituent domains, suggesting that, in addition to their autonomous binding capacities, coupling between the constituent domains results in emergent binding properties. In summary, in increasing order of DNA binding affinity, the domains of MeCP2 can be arranged as NTD, CTD-β, CTD-α, ID, TRD, and MBD.

Since binding of DNA to MeCP2 also confers thermal stability to the protein (*34*), it was of interest to determine the contributions of the different domains to the overall stability. To assess stability, we monitored the fluorescence emission of the single tryptophan in MeCP2 (W104 in the MBD) over the temperature range from 10 to 85 °C. We first compared the melting profiles of the MBD alone with those of the longer constructs containing the MBD. As previously reported, the MBD has an apparent T_m (50% melt temperature) of ~45 °C (*34*), and similar values were obtained with the MBD−ID construct and $MeCP2^{1-294}$, which includes most of the TRD. Unexpectedly, the T_m of NTD−MBD domain fusion was lower than that of the MBD (Table 4), suggesting that the NTD holds the MBD in a relatively destabilized state (thus lowering the T_m).

Each protein fragment was then mixed with methylated or unmethylated DNA consisting of the 45 bp segment of the BDNF promoter with a single centrally located methylatable site, and thermal melting profiles were obtained and analyzed (Table 4). Two important patterns emerge. First, unmethylated DNA acts as a stabilizing agent, consistently inducing increases in T_m. Here, the MBD alone stands out in inducing only a small stabilization of only ~2 °C, whereas additional domains generate a much greater enhancement in thermal stability. Second, when the DNA is methylated, there is a consistent increase of ~8 °C in the T_m of all constructs. Interestingly, the stabilizing effect of methylated DNA on the NTD−MBD−ID fragment exceeds that of the full-length protein. The very similar apparent thermal stabilities of $MeCP2^{1-294}$ and full-length $MeCP2^{1-486}$ indicate that the C-terminal domain of the molecule does not contribute to the overall thermal stability of the protein when it is bound to DNA.

DISCUSSION

Because of its high degree of disorder, the overall structure of full-length MeCP2 is not readily amenable to structure determination by X-ray crystallography or NMR. However, our domain-by-domain dissection of MeCP2 has provided important insights into the physical and functional properties of this unique unstructured protein. An overarching aspect of MeCP2 biology is the role that disorder plays in supporting its multiple functions. Importantly, the structure content of the full-length protein is largely similar to the sum of the weighted average structure of its domains, and the small difference can be attributed to changes in secondary structure resulting from interdomain associations largely involving the MBD in the context of the full-length protein. As further discussed below, the MBD appears to be the central hub for MeCP2 tertiary structure, forming contacts with the NTD, ID, and TRD.

The Large Number of MoRFs Might Account for the Functional and Structural Versatility of MeCP2. The functions of intrinsically disordered proteins are often coupled to the acquisition of structure upon binding to a partner (*65*−*68*). In this respect, it is significant that MeCP2 gains secondary structure and undergoes striking thermal stabilization upon binding to DNA (Table 4). α-MoRF predictors (*38*, *39*) (see Materials and Methods) predict nine α-MoRFs in MeCP2. This is an unusually large number for a protein of this size (*68*) but is consistent with its predicted distribution of order and disorder. There is at least one predicted MoRF in each MeCP2 domain except the ID. Some of these sites may be involved in long-range intraprotein contacts within MeCP2, promoting the *cis* and *trans* domain interactions documented above.

Two MeCP2 point mutations, R133C and A140V, located in the MBD resident α-MoRF 133−150, have been shown to impair MeCP2's interactions with the ATRX protein, leading to improper nuclear localization of ATRX, a phenomenon implicated in ATRX (α-thalassemia/mental retardation X-linked) syndrome (*69*). In the X-ray structure of the MBD−DNA complex (*15*), the two discontinuous MoRFs, 87−104 and 133−150 (Figure 1), form a continuous surface (Figure 9). Both α-MoRFs are thus maximally accessible and constitute potential interaction surfaces for MeCP2 domains or protein partners. Strikingly, several RTT-causing mutations in the MBD are proximal to these candidate interaction surfaces, explaining the deleterious effects of these mutations on MeCP2 function (Figure 9).

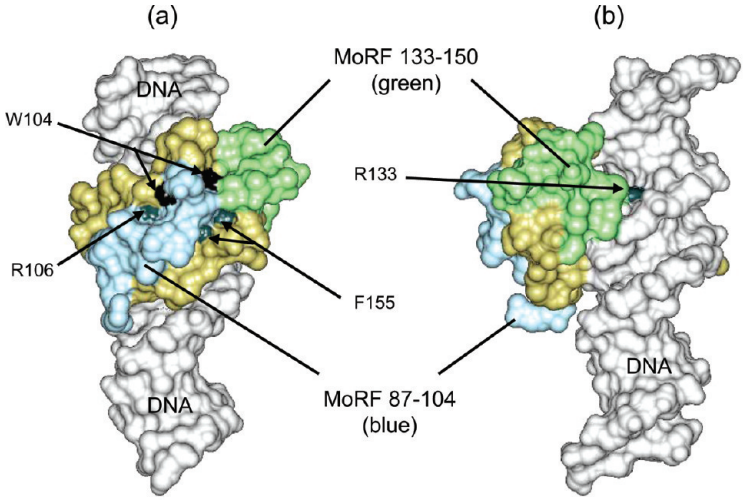

FIGURE 9: Structure of MBD bound to DNA which suggests that MoRFs flank interaction surfaces. Model of the MBD (tan, light blue, and green) of hMeCP2 bound to 20 bp of BDNF promoter DNA (gray) [Protein Data Bank entry 3c2i (*6*)]. The MBD α-MoRFs are located in residues 87−104 (light blue) and 133−150 (green). The two MoRFs form a contiguous surface that is predominantly hydrophilic, winding across the MBD opposite the DNA interaction surface. Arrows point to the solvent accessible surface area of Trp104 (black) and the surface-exposed regions (blue) of Arg106, Arg133, and Phe155, where Rett syndrome-causing point mutations result in significant changes in the local surface properties (*13*). These all contribute to a MoRF surface, suggesting a role in inter- and intraprotein interactions related to MoRF disorder-to-order transitions.

MeCP2 Harbors Multiple Autonomous Binding Sites That Affect the Overall Interactions of the Protein with DNA and Chromatin. Although MeCP2 was originally identified as a protein that binds specifically to methylated DNA, it has been shown since that MeCP2 can also bind to unmethylated DNA, albeit less efficiently (*11*, *28*, *31*, *32*). In this study, we have shown that of the four autonomous DNA binding domains, the ID and TRD acquire significant secondary structure upon binding to DNA, a phenomenon that has also been reported for key proteins such as Jun, Fos, GCN4, and histone H1 (*70*−*72*). MBD is solely responsible for methylation-specific binding (Figure 3b). The autonomous nonspecific DNA binding capabilities of the ID, TRD, and CTD-α suggest that full-length MeCP2 is poised to make unusually extensive contacts with DNA which, coupled with their highly disordered nature and ability to undergo binding-induced structural changes, suggests that the simultaneous or selective engagement of these domains will promote considerable functional variability.

The slight but consistent enhancement in shift seen with nucleosomal arrays when incubated with the CTD-*β*, versus DNA with which there was virtually no shift (Figure 3d), can be attributed to a putative histone-binding region(s) in the C-terminal half of MeCP2. This is in agreement with earlier work suggesting a role of the C-terminal domain in specific binding to chromatin, most likely to histone H3 (*32*, *73*, *74*).

Properties of Multidomain Fragments of MeCP2 Reveal Structural and Functional Synergism between Domains. Upon incubation with the NTD−MBD construct, both DNA and NA undergo striking enhancement in their electrophoretic shifts compared to the minor shift with the MBD and no shift with the NTD. Furthermore, a 10-fold higher binding affinity of the NTD−MBD construct for methylated DNA compared to that of the MBD alone (Figures 3b,e and 8a,b) suggests that conformational coupling between these domains possibly through their MoRFs results in a synergistic increase in DNA binding efficiency, and/or methylation specificity. Furthermore, thermal unfolding suggests that the NTD holds the MBD in a relatively destabilized state which may be more potent in DNA binding which is evident from the higher affinity of the NTD−MBD construct for methylated DNA than of the MBD. This is also evident from the fact that upon binding to DNA, the NTD−MBD construct undergoes a more pronounced thermal stabilization than the MBD.

Two distinct mechanisms for coupled binding and folding of unstructured proteins have recently been proposed: folding upon binding and conformational selection (*75*). For MeCP2, both mechanisms may be operating: while the acquisition of structure by the ID and TRD upon binding to DNA is definitely a case of binding-induced folding, interdomain interactions within MeCP2 could select for conformations favorable for DNA binding, as in the NTD-induced enhancement of the DNA binding affinity of the MBD. In addition, the unstructured regions of MeCP2 will populate an ensemble of conformations from which those that favor binding by specific partner proteins may be selected.

Another case of synergistic binding is seen with the TRD−CTD construct, which binds to DNA with 6-fold higher affinity than the TRD and 30-fold higher affinity than the CTD-α. It is also the most potent in terms of inducing mobility shifts in DNA and NAs (Figure 3a,c,d). Furthermore, the remarkable compaction and self-association of NAs induced by the TRD−CTD fusion suggest contributions from both the DNA and histone binding sites in this fragment.

The apparent synergism between certain domains in DNA and chromatin binding is consistent with the conformational coupling between the MBD and other domains of MeCP2 detected *in cis* and *trans*. Long-range interactions between interspersed structured segments may result in a loose folded structure with intrinsically disordered domains extending from one or two structural "hubs". Such a structure is likely to promote simultaneous interaction with multiple partners conferring considerable functional flexibility to MeCP2.

Two Halves of MeCP2 Involved in DNA and Chromatin Binding. The DNA binding properties of the MBD are strongly modulated by the flanking domains, suggesting that the NTD−MBD−ID region constitutes a functional entity with

both methylation-dependent and -independent DNA binding abilities. The TRD−CTD construct also functions as an independent unit with chromatin compacting and oligomerizing properties. This suggests that MeCP2 is effectively organized into an N-terminal functional unit composed of the NTD, MBD, and ID and a C-terminal unit composed of the TRD, CTD-α, and CTD-β.

While both the TRD (residues 207−310) and CTD-α (residues 261−330) fragments bind DNA efficiently, DNA-induced structure acquisition is limited to the TRD polypeptide. This suggests that residues 207−260 within the TRD house DNA binding activity while residues 310−335 harbor the DNA binding activity seen in the CTD-α. In support of the former, deletion mutagenesis of the isolated TRD suggests its DNA binding region lies between residues 245 and 270 (data not shown). While the extent of NA compaction induced by the TRD−CTD construct approaches that seen with the wild-type protein, a higher ratio of the TRD−CTD polypeptide is needed to see an equivalent effect. This would be expected if NTD−MBD−ID construct-mediated DNA binding acts in concert with TRD−CTD construct-mediated DNA binding and also contributes to inter- and intranucleosomal associations of NAs.

In Vitro Functions of MeCP2 Domains Strongly Correlate with Their Function in Vivo. Our results indicating the importance of MeCP2 domains other than the MBD and TRD are supported by in vivo data. For example, the heterochromatin-associated HP1 protein has been shown to interact with the NTD (*35*), perhaps contributing to the colocalization of MeCP2 and HP1 in pericentromeric heterochromatin. However, while the NTD appears to be necessary for the pericentromeric localization of MeCP2, it is evidently not sufficient, also requiring the MBD (*76*). This is in agreement with our data that show that NTD largely plays a role in regulating MBD-mediated binding. Important roles for the ID have also been established. The ID has been shown to be a universal component of the MeCP2 fragments required for interaction with the corepressors mSin3A (*77*), N-CoR and Ski (*18*), H3 methyl transferase (*20*), p20, a putative *Xenopus* protease (*78*), and the DNA methyltransferase Dnmt1 (*19*). A fluorescence recovery after photobleaching study (*36*) has recently demonstrated that deletion of either the ID or the TRD markedly increases MeCP2 intranuclear mobility. This suggests that the DNA-induced disorder-to-order transition of the ID and the TRD reported here significantly stabilizes MeCP2−chromatin complexes, perhaps providing a larger window for downstream repressor recruitment or secondary interactions required for the structural modulation of chromatin.

Our finding that full-length MeCP2 and the TRD−CTD construct induce clustering of nucleosomes with looping out of non-nucleosomal DNA (Figure 5) (*31*) is consistent with the growing in vivo evidence of a role of MeCP2 in stabilizing large chromatin loops (*30*, *79*, *80*). The presence of independent DNA and chromatin-binding domains in multiple regions of MeCP2, allowing a single MeCP2 molecule to bind two or more regions of chromatin, would contribute to the stabilization of a loop base. Interestingly, in this regard, the MBD−ID fragment has been reported to contain a MAR binding site (*23*). A role in loop maintenance would not be possible with the MBD alone with its very limited compaction ability but may require the additional DNA binding properties of the ID and the chromatin binding properties of the CTD. The pathological effects of C-terminal truncations of MeCP2 in both RTT patients and a mouse model (*17*) may be understood in this context.

On the basis of the work presented here, a new picture of MeCP2 biology is emerging in which its intrinsically disordered nature is a key property. The novel properties of MeCP2 elucidated here support a global structure for MeCP2 that can engage in a wide range of potential binding events and conformational changes, promoting functional outcomes that will vary according to the specific context of the gene locus, DNA methylation density, and availability of binding partners.

ACKNOWLEDGMENT

We thank Dr. Stephen Eyles for assistance with analytical ultracentrifugation.

REFERENCES

1. Klose, R. J., and Bird, A. P. (2006) Genomic DNA methylation: The mark and its mediators. *Trends Biochem. Sci. 31*, 91–97.
2. Dhasarthy, J., and Wade, P. A. (2008) The MBD protein family: Reading an epigenetic mark? *Mutat. Res. 647*, 39–43.
3. Amir, R. E., den Veyber, I. B., Wan, M., Tran, C. Q., Francke, U., and Zoghbi, H. Y. (1999) Rett syndrome is caused by mutations in X-linked MeCP2, encoding methyl-CpG-binding protein 2. *Nat. Genet. 23*, 185–188.
4. LaSalle, J. M., Hogart, A., and Thatcher, K. N. (2005) Rett syndrome: A Rosetta stone for understanding the molecular pathogenesis of autism. *Int. Rev. Neurobiol. 71*, 131–165.
5. Moretti, P., and Zoghbi, H. Y. (2006) MeCP2 dysfunction in Rett syndrome and related disorders. *Curr. Opin. Genet. Dev. 16*, 276–281.
6. Kishi, N., and Macklis, J. D. (2004) MeCP2 is progressively expressed in post-migratory neurons and is involved in neuronal maturation rather than cell fate decisions. *Mol. Cell. Neurosci. 27*, 306–321.
7. Bernard, D., Gil, J., Dumont, P., Rizzo, S., Monte, D., Quatannens, B., Hudson, D., Visakorpi, T., Fuks, F., and de Launoit, Y. (2006) The methyl-CpG-binding protein MeCP2 is required for prostate cancer cell growth. *Oncogene 25*, 1358–1366.
8. Chatagnon, A., Bougel, S., Perriaud, L., Lachuer, J., Benhattar, J., and Dante, R. (2009) Specific association between the methyl-CpG-binding domain protein 2 and the hypermethylated region of the human telomerase reverse transcriptase promoter in cancer cells. *Carcinogenesis 30*, 28–34.
9. Ballestar, E., and Esteller, M. (2005) Methyl-CpG-binding proteins in cancer: Blaming the DNA methylation messenger. *Biochem. Cell Biol. 83*, 374–384.
10. Ballestar, E., Paz, M. F., Valle, L., Wei, S., Fraga, M. F., Espada, J., Cigudosa, J. C., Huang, T. H., and Esteller, M. (2003) Methyl-CpG binding proteins identify novel sites of epigenetic inactivation in human cancer. *EMBO J. 22*, 6335–6345.
11. Adams, V. H., McBryant, S. J., Wade, P. A., Woodcock, C. L., and Hansen, J. C. (2007) Intrinsic disorder and autonomous region function in the multifunctional nuclear protein, MeCP2. *J. Biol. Chem. 282*, 15057–15064.
12. Nan, X., Meehan, R. R., and Bird, A. (1993) Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. *Nucleic Acids Res. 21*, 4886–4892.
13. Wakefield, R. I. D., Smith, B. O., Nan, X., Free, A., Soteriou, A., Uhrin, D., Bird, A. P., and Barlow, P. N. (1999) The solution structure of the region from MeCP2 that binds to methylated DNA. *J. Mol. Biol. 291*, 1055–1065.
14. Ohki, I., Shimotake, N., Fujita, N., Jee, J.-G., Ikegami, T., Nakao, M., and Shirakawa, M. (2001) Solution structure of the methyl-CpG binding region of human MBD1 in complex with methylated DNA. *Cell 105*, 487–497.
15. Ho, K. L., McNae, I. W., Schmiedeberg, L., Klose, R. J., Bird, A. P., and Walkinshaw, M. W. (2008) MeCP2 binding to DNA depends upon hydration at methyl-CpG. *Mol. Cell 29*, 525–531.
16. Nan, X., Campoy, F. J., and Bird, A. (1997) MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell 88*, 471–481.
17. Shahbazian, M., Young, J., Yuva-Paylor, L., Spencer, C., Antalffy, B., Noebels, J., Armstrong, D., Paylor, R., and Zoghbi, H. (2002) Mice with truncated MeCP2 recapitulate many Rett syndrome features and display hyperacetylation of histone H3. *Neuron 35*, 243–254.
18. Lunyak, V. V., Burgess, R., Prefontaine, G. G., Nelson, C., Sze, S. H., Chenoweth, J., Schwartz, P., Pevzner, P. A., Glass, C., Mandel, G., and Rosenfeld, M. G. (2002) Corepressor-dependent silencing of

chromosomal regions encoding neuronal gene. *Science 298*, 1747–1752.
19. Kokura, K., Kaul, S. C., Wadhwa, R., Nomura, T., Khan, M. M., Shinagawa, T., Yasukawa, T., Colmenares, C., and Ishii, S. (2001) The Ski protein family is required for MeCP2-mediated transcriptional repression. *J. Biol. Chem. 276*, 34115–34121.
20. Kimura, H., and Shiota, K. (2003) Methyl-CpG-binding protein, MeCP2, is a target molecule for maintenance DNA methyl transferase, Dnmt1. *J. Biol. Chem. 288*, 4806–4812.
21. Fuks, F., Hurd, P. J., Wolf, D., Nan, X., Bird, A. P., and Bird, A. P. (2003) The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. *J. Biol. Chem. 278*, 4035–4040.
22. Suzuki, M., Yamada, T., Kihara-Negishi, F., Sakurai, T., and Oikawa, T. (2003) Direct association between PU.1 and MeCP2 that recruits mSin3A-HDAC complex for PU.1-mediated transcriptional repression. *Oncogene 22*, 8688–8698.
23. Buschdorf, J. P., and Stratling, W. H. (2004) A WW region binding region in methyl-CpG-binding protein MeCP2: Impact on Rett syndrome. *J. Mol. Med. 82*, 135–143.
24. Harikrishnan, K. N., Chow, M. Z., Baker, E. K., Pal, S., Bassal, S., Brasacchio, D., Wang, L., Craig, J. M., Jones, P. L., Sif, S., and El-Osta, A. (2005) Brahma links the SWI/SNF chromatin-remodeling complex with MeCP2-dependent transcriptional silencing. *Nat. Genet. 37*, 254–264.
25. Jeffery, L., and Nakielny, S. (2004) Components of the DNA methylation system of chromatin control are RNA-binding proteins. *J. Biol. Chem. 279*, 49479–49487.
26. Young, J. I., Hong, E. P., Castle, J. C., Crespo-Barreto, J., Bowman, A. B., Rose, M. F., Kang, D., Richman, R., Johnson, J. M., Berget, S., and Zoghbi, H. Y. (2005) Regulation of RNA splicing by the methylation-dependent transcriptional repressor methyl-CpG binding protein 2. *Proc. Natl. Acad. Sci. U.S.A. 102*, 17551–17558.
27. Dunker, A. K., Silman, I., Uversky, V. N., and Sussman, J. L. (2008) Function and structure of inherently disordered proteins. *Curr. Opin. Struct. Biol. 18*, 756–764.
28. Yasui, D. H., Peddada, S., Bieda, M. C., Vallero, R. O., Hogart, A., Nagarajan, R. P., Thatcher, K. N., Farnham, P. J., and LaSalle, J. M. (2007) Integrated epigenomic analyses of neuronal MeCP2 reveal a role for long-range interaction with active genes. *Proc. Natl. Acad. Sci. U.S.A. 104*, 19416–19421.
29. Chahrour, M., Jung, S. J., Shaw, C., Zhou, X., Wong, S. T. C., Qin, J., and Zoghbi, H. Y. (2008) MeCP2, a key contributor to neurological disease, activates and represses transcription. *Science 320*, 1224–1229.
30. Horike, S., Cai, S., Miyano, M., Cheng, J. F., and Kohwi-Shigematsu, T. (2005) Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. *Nat. Genet. 37*, 31–40.
31. Nikitina, T. N., Shi, X., Ghosh, R. P., Horowitz-Scherer, R. A., Hansen, J. C., and Woodcock, C. L. (2007) Multiple interactions between MeCP and chromatin. *Mol. Cell. Biol. 27*, 864–877.
32. Georgel, P. T., Horowitz-Scherer, R. A., Adkins, N., Woodcock, C. L., Wade, P. A., and Hansen, J. C. (2003) Chromatin compaction by human MeCP2: Assembly of novel secondary structures in the absence of DNA methylation. *J. Biol. Chem. 278*, 32181–32188.
33. Chadwick, L. H., and Wade, P. A. (2007) MeCP2 in Rett syndrome: Transcriptional repressor or chromatin architectural protein. *Curr. Opin. Genet. Dev. 17*, 1–5.
34. 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.
35. Agarwal, N., Hardt, T., Brero, A., Nowak, D., Rothbauer, U., Becker, A., Leonhardt, H., and Cardoso, M. C. (2007) MeCP2 interacts with HP1 and modulates its heterochromatin association during myogenic differentiation. *Nucleic Acids Res. 35*, 5402–5408.
36. Kumar, A., Kamboj, S., Malone, B. M., Kudo, S., Twiss, J. L., Czymmek, K., LaSalle, J. M., and Schanen, N. C. (2008) Analysis of protein regions and Rett syndrome mutations indicate that multiple regions influence chromatin-binding dynamics of the chromatin-associated protein MECP2 in vivo. *J. Cell Sci. 121*, 1128–1137.
37. Laccone, F., Huppke, P., Hanefeld, F., and Meins, M. (2001) Mutation spectrum in patients with Rett syndrome in the German population: Evidence of hot spot regions. *Hum. Mutat. 17*, 183–190.
38. Oldfield, C. J., Cheng, Y., Cortese, M. S., Romero, P., Uversky, V. N., and Dunker, A. K. (2005) Coupled folding and binding with α-helix-forming molecular recognition elements. *Biochemistry 44*, 12454–12470.
39. Mohan, A., Oldfield, C. J., Radivojac, P., Vacic, V., Cortese, M. S., Dunker, A. K., and Uversky, V. N. (2006) Analysis of molecular recognition features (MoRFs). *J. Mol. Biol. 362*, 1043–1059.
40. Martin, B. R., Giepmans, B. N., Adams, S. R., and Tsien, R. Y. (2005) Mammalian cell-based optimization of the biarsenical-binding tetracysteine motif for improved fluorescence and affinity. *Nat. Biotechnol. 10*, 1308–1314.
41. Lobley, A., Whitmore, L., and Wallace, B. A. (2002) DICHROWEB: An interactive website for the analysis of protein secondary structure from circular dichroism spectra. *Bioinformatics 18*, 211–212.
42. Whitmore, L., and Wallace, B. A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res. 32*, W668–W673.
43. Provencher, S. W., and Glöckner, J. (1981) Estimation of globular protein secondary structure from circular dichroism. *Biochemistry 20*, 33–37.
44. Pancoska, P., Bitto, E., Janota, V., Urbanova, M., Gupta, V. P., and Keiderling, T. A. (1995) Comparison of and limits of accuracy for statistical analyses of vibrational and electronic circular dichroism spectra in terms of correlations to and predictions of protein secondary structure. *Protein Sci. 4*, 1384–1401.
45. Perczel, A., Park, K., and Fasman, G. D. (1992) Analysis of the circular dichroism spectrum of proteins using the convex constraint algorithm: A practical guide. *Anal. Biochem. 203*, 83–91.
46. Greenfield, N. J. (2004) Circular dichroism analysis for protein-protein interactions. *Methods Mol. Biol. 261*, 55–78.
47. Lakowicz, J. R. (2006) in Principles of fluorescence spectroscopy, 3rd ed., Springer, New York.
48. Luedtke, N. W., Dexter, R. J., Fried, D. B., and Schepartz, A. (2007) Surveying polypeptide and protein domain conformation and association with FlAsH and ReAsH. *Nat. Chem. Biol. 3*, 779–784.
49. Krishnan, B., and Gierasch, L. M. (2008) Cross-strand split tetra-Cys motifs as structure sensors in a β-sheet protein. *Chem. Biol. 15*, 1104–1115.
50. Griffin, B. A., Adams, R. S., and Tsien, Y. R. (1998) Specific covalent labeling of recombinant protein molecules inside living cells. *Science 281*, 269–272.
51. Findlay, J. W., and Dillard, R. E. (2007) Appropriate calibration curve fitting in ligand binding assays. *AAPS J. 9*, E260–E267.
52. Van Holde, K. E., and Weischet, W. O. (1978) Boundary analysis of sedimentation velocity experiments with monodisperse and paucidisperse solutes. *Biopolymers 17*, 1387–1403.
53. Romero, P., Obradovic, Z., Li, X., Garner, E. C., Brown, C. J., and Dunker, A. K. (2001) Sequence complexity of disordered protein. *Proteins 42*, 38–48.
54. Sickmeier, M., Hamilton, J. A., LeGall, T., Vacic, V., Cortese, M. S., Tantos, A., Szabo, B., Tompa, P., Chen, J., Uversky, V. N., Obradovic, Z., and Dunker, A. K. (2007) DisProt: The database of disordered proteins. *Nucleic Acids Res. 35*, D786–D793.
55. Iakoucheva, L. M., and Dunker, K. A. (2003) Order, disorder, and flexibility: Prediction from protein sequence. *Structure 11*, 1316–1317.
56. Vacic, V., Uversky, V. N., Dunker, A. K., and Lonardi, S. (2007) Composition Profiler: A tool for discovery and visualization of amino acid composition differences. *BMC Bioinf. 19*, 211.
57. Lowary, P. T., and Widom, J. (1998) New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol. 276*, 19–42.
58. Chen, W. G., Chang, Q., Lin, Y., Meissner, A., West, A. E., Griffith, E. C., Jaenisch, R., and Greenberg, M. E. (2003) Derepression of BDNF transcription involves calcium-dependent phosphorylation of MeCP2. *Science 302*, 885–889.
59. Hansen, J. C., Lebowitz, J., and Demeler, B. (1994) Analytical ultracentrifugation of complex macromolecular systems. *Biochemistry 33*, 13155–13163.
60. Heitmann, B., Maurer, T., Weitzel, J. M., Stratling, W., and Kalbitzer, H. R. (2003) Solution structure of the matrix attachment region-binding domain of chicken MeCP2. *Eur. J. Biochem. 270*, 3263–3270.
61. Berggard, T., Linse, S., and James, P. (2007) Methods for the detection and analysis of protein-protein interactions. *Proteomics 7*, 2833–2842.
62. Adams, S. R., Campbell, R. E., Gross, L. A., Martin, B. R., Walkup, G. K., Yao, Y., Llopis, J., and Tsien, R. Y. (2002) New biarsenical ligands and tetracysteine motifs for protein labeling in vitro and in vivo: Synthesis and biological applications. *J. Am. Chem. Soc. 124*, 6063–6076.
63. Shuman, C. F., Jiji, R., Kerfeldt, K. S., and Linse, S. (2006) Reconstitution of calmodulin from domains and subdomains: Influence of target peptide. *J. Mol. Biol. 358*, 870–881.
64. Martinowich, K., Hattori, D., Wu, H., Fouse, S., He, F., Hu, Y., Fan, G., and Sun, Y. E. (2003) DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. *Science 302*, 890–893.

65. Tompa, P. (2005) The interplay between structure and function in intrinsically unstructured proteins. *FEBS Lett. 579*, 3346–3354.
66. Cortese, M. S., Uversky, V. N., and Dunker, A. K. (2008) Intrinsic disorder in scaffold proteins: Getting more from less. *Prog. Biophys. Mol. Biol. 98*, 85–106.
67. Oldfield, C. J., Meng, J., Yang, J. Y., Yang, M. Q., Uversky, V. N., and Dunker, A. K. (2008) Flexible nets: Disorder and induced fit in the associations of p53 and 14-3-3 with their partners. *BMC Genomics 9* (Suppl. 1), S1.
68. Mészáros, B., Tompa, P., Simon, I., and Dosztányi, Z. (2007) Molecular principles of the interactions of disordered proteins. *J. Mol. Biol. 372*, 549–561.
69. Nan, X., Hou, J., Maclean, A., Nasir, J., Lafuente, M. J., Shu, X., Kriaucionis, S., and Bird, A. (2007) Interaction between chromatin proteins MeCP2 and ATRX is disrupted by mutations that cause inherited mental retardation. *Proc. Natl. Acad. Sci. U.S.A. 104*, 2709–2714.
70. Patel, L., Abate, C., and Curran, T. (1990) Altered protein conformation on DNA binding by Fos and Jun. *Nature 347*, 572–575.
71. Weiss, M. A., Ellenberger, T., Wobbe, C. R., Lee, J. P., Harrison, S. C., and Struhl, K. (1990) Folding transition in the DNA-binding region of GCN4 on specific binding to DNA. *Nature 347*, 575–578.
72. Roque, A., Iloro, I., Ponte, I., Luis, J., Arrondo, R., and Suau, P. (2005) DNA induced secondary structure of the carboxy terminal region histone H1. *J. Biol. Chem. 280*, 32141–32147.
73. Chandler, S. P., Guschin, D., Landsberger, N., and Wolffe, A. P. (1999) The methyl-CpG binding transcriptional repressor MeCP2 stably associates with nucleosomal DNA. *Biochemistry 38*, 7008–7018.
74. Nikitina, T. N., Ghosh, R. P., Horowitz-Scherer, R. A., Hansen, J. C., Grigoryev, S. A., and Woodcock, C. L. (2007) MeCP2-chromatin interactions include the formation of chromatosome-like structures and are altered in mutations causing Rett syndrome. *J. Biol. Chem. 282*, 28237–28245.
75. Wright, P. E., and Dyson, H. J. (2009) Linking folding and binding. *Curr. Opin. Struct. Biol. 19*, 31–38.
76. Marchi, M., Guarda, A., Bergo, A., Landsberger, N., Kilstrup-Nielsen, C., Ratto, G. M., and Costa, M. (2007) Spatio-temporal dynamics and localization of MeCP2 and pathological mutants in living cells. *Epigenetics 2*, 187–197.
77. Nan, X., Ng, H.-H., Johnson, C. A., Laherty, C. D., Turner, B. M., Eisenman, R. N., and Bird, A. (1998) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature 393*, 386–389.
78. Carro, S., Bergo, A., Mengoni, M., Bachi, A., Badaracco, G., Kilstrup-Nielsen, C., and Landsberger, N. (2004) A novel protein, *Xenopus* p20, influences the stability of MeCP2 through direct interaction. *J. Biol. Chem. 279*, 25623–25631.
79. Weitzel, J. M., Buhrmester, H., and Strätling, W. H. (2007) Chicken MAR-binding protein ARBP is homologous to rat methyl-CpG-binding protein MeCP2. *Mol. Cell. Biol. 17*, 5656–5666.
80. Eivazova, E. R., Gavrilov, A., Pirozhkova, I., Petrov, A., Iarovaia, O. V., Razin, S. V., Lipinski, M., and Vassetzky, Y. S. (2008) Interaction in vivo between the two Matrix attachment regions flanking a single chromatin loop. *J. Mol. Biol. 386*, 929–937.