

Influence of the disordered domain structure of MeCP2 on its structural stability and dsDNA interaction

David Ortega-Alarcon ^{a,1}, Rafael Claveria-Gimeno ^{a,b,c,1,2}, Sonia Vega ^a, Olga C. Jorge-Torres ^d, Manel Esteller ^{d,e,f,g}, Olga Abian ^{a,b,c,h,i,*}, Adrian Velazquez-Campoy ^{a,c,h,i,j,**}

^a Institute of Biocomputation and Physics of Complex Systems (BIFI), Joint Units IQFR-CSIC-BIFI, and GBSC-CSIC-BIFI, Universidad de Zaragoza, Zaragoza 50018, Spain

^b Instituto Aragonés de Ciencias de la Salud (IACS), 50009 Zaragoza, Spain

^c Instituto de Investigación Sanitaria Aragón (IIS Aragón), 50009 Zaragoza, Spain

^d Josep Carreras Leukaemia Research Institute (IJC), 08916 Badalona, Barcelona, Spain

^e Centro de Investigación Biomedica en Red Cancer (CIBERONC), 28029 Madrid, Spain

^f Institutio Catalana de Recerca i Estudis Avançats (ICREA), 08010 Barcelona, Spain

^g Physiological Sciences Department, School of Medicine and Health Sciences, University of Barcelona (UB), 08907, l'Hospitalet de Llobregat, Barcelona, Spain

^h Centro de Investigación Biomédica en Red en el Área Temática de Enfermedades Hepáticas y Digestivas (CIBERehd), 28029 Madrid, Spain

ⁱ Departamento de Bioquímica y Biología Molecular y Celular, Universidad de Zaragoza, 50009 Zaragoza, Spain

^j Fundación ARAID, Gobierno de Aragón, 50018 Zaragoza, Spain

ARTICLE INFO

Article history:

Received 9 November 2020

Received in revised form 17 January 2021

Accepted 29 January 2021

Available online 03 February 2021

Keywords:

Methyl-CpG-binding protein 2 (MeCP2)

Intrinsically disordered protein (IDP)

Protein stability and interactions

ABSTRACT

Methyl-CpG binding protein 2 (MeCP2) is a transcriptional regulator and a chromatin-associated structural protein. MeCP2 deregulation results in two neurodevelopmental disorders: MeCP2 dysfunction is associated with Rett syndrome, while excess of activity is associated with MeCP2 duplication syndrome. MeCP2 is an intrinsically disordered protein (IDP) constituted by six structural domains with variable, small percentage of well-defined secondary structure. Two domains, methyl-CpG binding domain (MBD) and transcription repressor domain (TRD), are the elements responsible for dsDNA binding ability and recruitment of the gene transcription/silencing machinery, respectively. Previously we studied the influence of the completely disordered, MBD-flanking domains (N-terminal domain, NTD, and intervening domain, ID) on the structural and functional features of the MBD (Claveria-Gimeno, R. et al. *Sci Rep.* 2017, 7, 41,635). Here we report the biophysical study of the influence of the remaining domains (transcriptional repressor domain, TRD, and C-terminal domains, CTD α and CTD β) on the structural stability of MBD and the dsDNA binding capabilities of MBD and ID. The influence of distant disordered domains on MBD properties makes it necessary to consider the NTD-MBD-ID variant as the minimal protein construct for studying dsDNA/chromatin binding properties, while the full-length protein should be considered for transcriptional regulation studies.

© 2021 Published by Elsevier B.V.

1. Introduction

Methyl-CpG binding protein 2 (MeCP2) is a transcriptional regulator involved in early stages of neuronal development, differentiation, maturation, and synaptic plasticity control [1]. Besides this promoter-specific dsDNA interaction required for finely tuning gene transcription, MeCP2 also binds massively to chromatin, thus acting as a chromatin

architecture remodeling factor by replacing histone H1 as a nucleosomal linker [2–4]. These different modes of interaction with DNA, as well as its ability to interact with many other different biological partners (RNA, structural and transcriptional proteins, nucleosomal elements) and its multifaceted cellular role regulated by post-translational modifications are made possible by its modular, dynamic and adaptive structure [5,6]. MeCP2 is an important hub within gene transcription regulation networks. However, recent evidences suggest a primary role consisting of recruiting co-repressor complexes to methylated sites in the genome, resulting in dampened gene expression [7].

MeCP2 deregulation leads to disease [8–10]. MeCP2 point mutations or deletions causing activity loss or deficiency are associated with Rett syndrome (RTT). Although a rare disease (1:10000 births), RTT is the main cause of mental retardation in females, characterized by a clinically varied expression and sharing features with other neurological

* Correspondence to: Olga Abian, Instituto Aragonés de Ciencias de la Salud (IACS), San Juan Bosco 13, 50009 Zaragoza, Spain.

** Correspondence to: Adrián Velázquez-Campoy, Institute BIFI - University of Zaragoza, Mariano Esquillor s/n, 50018 Zaragoza, Spain.

E-mail address: adrianvc@unizar.es (A. Velázquez-Campoy).

¹ Contributed equally

² Current address: Certest Biotec S.L., 50,840, Zaragoza, Spain

autistic diseases. On the other hand, duplication of *mecp2* causes over-expression of MeCP2 and leads to MeCP2 duplication syndrome (MDS), another much rarer disorder affecting males and sharing some phenotypic features with RTT (e.g., severe intellectual disability and impaired motor function).

MeCP2 is an intrinsically disordered protein (IDP). Most of its polypeptide chain (~60%) lacks well-defined secondary/tertiary structure. MeCP2 share some features with other IDPs, such as a high content in charged and polar residues (Table S1). The presence of flexible regions facilitates structural rearrangements necessary for allosteric regulation (in a broad sense, allosteric control consists on the modulation of the protein conformational landscape through ligand binding) and for exposing different motifs to interact with different partners. Interactions in IDPs are often characterized by a moderate-to-low binding affinity and a transient nature, because of the energetic penalty stemming from the conformational change coupled to the binding.

MeCP2 is organized into six domains: N-terminal domain (NTD), methyl binding domain (MBD), intervening domain (ID), transcriptional repression domain (TRD), C-terminal domain α (CTD α), and C-terminal domain β (CTD β) (Supplementary Fig. S1) [2,11]. The most important domains are MBD, initially associated with methylated CpG (mCpG) DNA binding, and TRD, associated with transcription repression activity [12,13]. Most of the mutations associated with RTT are located in these two domains [14]. However, it has been shown that other domains (ID, CTD α and CTD β) are also directly or indirectly involved in methyl-independent DNA interaction [11,15].

MBD is the best characterized domain in MeCP2. MBD structure basically consists of a wedge-shaped structured core containing a 3-stranded anti-parallel β -sheet with an α -helix on the C-terminal side, with two unstructured regions flanking this core [16,17]. MBD is considered to be directly involved in maintaining the global organization of the protein through interactions with other domains through inter-domain coupling [4]. Mutations in this domain would have an impact on the local and the global stability in MeCP2 [2,18].

There are many (completely or partially) unexplained issues related to MeCP2 function, such as: 1) the small difference in the binding affinities determined in vitro for methylated and unmethylated DNA, which is far from being biologically significant to explain its preferential distribution tracking the density of 5-methyl-cytosines heterochromatin foci [5,19,20]; 2) the mechanisms by which MeCP2 is able to both activate and deactivate the transcription of hundreds of genes, depending on the context (e.g., decreasing transcriptional noise and adapting gene expression pattern to different physiological or environmentally-induced conditions) [21,22]; 3) the inter-domain interactions and dependencies giving rise to cooperative phenomena regarding structural and functional features in MeCP2; 4) the role of the different DNA binding motifs along its sequence; are they functionally independent allowing the interaction with several DNA stretches simultaneously?; 5) the possibility for MeCP2 to bind to dsDNA in a cooperative fashion, where some domains would play a structural/functional role [4,23]; 6) the ability of MeCP2 for simultaneously showing methylation-dependent specific binding to certain target gene promoter locations and genome-wide methylation-independent unspecific binding to heterochromatin [24–26]; 7) the ability to recognize not only methyl-cytosine, but also other methylated nucleotides with high affinity [27]; 8) the large number of similar symptoms associated with RTT and MDS, two disorders caused by the down- and upregulation of MeCP2, respectively; and 9) the presence of two MeCP2 isoforms (E1 and E2) differing in just a few N-terminal residues and showing different expression patterns and non-redundant functions, being E2 the most widely studied and characterized isoform, but the one with lower expression level [28].

Besides recent, spectacular advances in cellular and gene therapy (e.g., gene edition using CRISPR techniques), drug discovery and development remains as a useful strategy for efficient therapeutic tools. Although many efforts have been devoted to improve palliative care or

to restore intracellular signaling and metabolic routes altered because of abnormal MeCP2 activity, a possible strategy for RTT and MDS drug development consist in the search for small molecules able to bind to MeCP2 stabilizing and rescuing defective mutants associated with RTT, or able to bind to MeCP2 inhibiting its activity for MDS treatment. From a practical point of view, full-length MeCP2 is a protein presenting considerable difficulties because it is prone to degradation and precipitation. Thus, depending on the goal (i.e., improving or rescuing DNA binding or gene transcription), different protein constructs or variants can be specifically designed for laboratory work (e.g., biophysical structural and functional studies, cell-based assays, high-throughput molecular screening and drug discovery). In a previous biophysical study of three MeCP2 variants (MBD, and NTD-MBD, and NTD-MBD-ID), we established that the isolated MBD might not be the appropriate construct to study and assay its DNA binding features, because the presence of NTD and ID increased considerably the DNA binding affinity and the structural stability, besides adding a second, functionally independent and relevant DNA binding site [29].

Here we report a biophysical study of the structural stability and the DNA interaction of the remaining MeCP2 variants: NTD-MBD-ID-TRD, NTD-MBD-ID-TRD-CTD α , and NTD-MBD-ID-TRD-CTD α -CTD β (full-length MeCP2). The purpose was to find out if distant disordered domains would affect MBD structural and functional features, if a longer variant for MeCP2 would be more appropriate for biophysical studies and drug discovery programs, and if the presence of additional AT-hooks (in TRD and CTD α) would result in additional functionally independent DNA binding sites.

2. Materials & methods

2.1. Plasmid construction

MeCP2 (isoform E2) variants were inserted in a pET30b plasmid for protein expression. The different protein variants were obtained by inserting appropriate stop codons: NTD-MBD-ID-TRD, NTD-MBD-ID-TRD-CTD α , and NTD-MBD-ID-TRD-CTD α -CTD β (Fig. S1). The protein sequences contained an N-terminal polyhistidine-tag which was always removed after purification through an inserted PreScission Protease recognition cleavage site. All sequences were checked by sequencing analysis. The protein variants were checked and corroborated by Sanger sequencing using a BigDye Terminator v3.1 Cycle Sequencing Kit (Life Technologies, Carlsbad, CA, USA) in an Applied Biosystems 3730/DNA Analyzer (Thermo Fisher Scientific, Waltham, MA, USA). The data were analyzed with BioEdit Sequence Alignment Editor [30].

2.2. Protein expression and purification

All protein variants (NTD-MBD-ID-TRD, NTD-MBD-ID-TRD-CTD α , and NTD-MBD-ID-TRD-CTD α -CTD β /full-length MeCP2) were expressed and purified following the same procedure. Plasmids were transformed into BL21 (DE3) Star *E. coli* strain. Bacteria cultures were grown in 150 mL of LB/kanamycin (50 μ g/mL) media at 37 °C overnight. Then, 4 L of LB/kanamycin (25 μ g/mL) were inoculated (1:100 dilution) and incubated under the same conditions until reaching an OD (at a wavelength of 600 nm) of 0.6. Protein expression was induced with 1 mM isopropyl 1-thio- β -D-galactopyranoside (IPTG) at 18 °C overnight. Cells were ruptured by sonication in ice and benzonase (Merck-Millipore, Madrid, Spain) was added (20 U/mL) to remove nucleic acids. Proteins were purified using immobilized metal ion affinity chromatography (IMAC) in a HiTrap TALON column (GE-Healthcare Life Sciences, Barcelona, Spain) with two washing steps: buffer sodium phosphate 50 mM, pH 7, NaCl 300 mM, and in buffer sodium phosphate 50 mM, pH 7, NaCl 800 mM (to remove potential DNA contamination from the protein), before an imidazole 10–150 mM elution gradient. Purity was checked by SDS-PAGE.

Removal of the histidine-tag was performed by GST-tagged PreScission Protease processing in cleavage buffer (50 mM Tris-HCl,

150 mM NaCl, pH 7.5) at 4 °C for 4 h. Progress of the protease processing was checked by SDS-PAGE. Finally, proteins were further purified using a combination of two affinity chromatographic steps to remove the histidine-tag (HiTrap TALON column, from GE-Healthcare Life Sciences, Barcelona, Spain) and the GST-tagged PreScission Protease (GST TALON column, from GE-Healthcare Life Sciences, Barcelona, Spain). Purity and homogeneity were checked by SDS-PAGE and size-exclusion chromatography. The proteins were stored in buffer Tris 50 mM pH 7.0 at –80 °C. The identity of all proteins was checked by mass spectrometry (4800plus MALDI-TOF/MS, from Applied Biosystems - Thermo Fisher Scientific, Waltham, MA, USA). Potential DNA contamination was always checked determining the ratio of UV absorption at 260 nm vs absorption at 280 nm. An extinction coefficient of 11,460 M⁻¹ cm⁻¹ at 280 nm was employed for all variants (a single tryptophan is located in MBD), except for the full-length protein (13,075 M⁻¹ cm⁻¹).

Stability and binding assays were performed at different pH and buffer conditions (Tris 50 mM pH 7–9, NaCl 0–150 mM; Pipes 50 mM, pH 7; Phosphate 50 mM, pH 7). When needed, buffer exchange was performed using a 3 or 10 kDa-pore size ultrafiltration device (Amicon centrifugal filter, Merck-Millipore, Madrid, Spain) at 4000 rpm and 4 °C.

2.3. Double-stranded DNA

HPLC-purified methylated and unmethylated 45-bp single-stranded DNA (ssDNA) oligomers corresponding to the promoter IV of the mouse brain-derived neurotrophic factor (BDNF) gene [15,18] were obtained from Integrated DNA Technologies. Two complementary pairs of DNA were used for DNA binding assays: forward unmethylated: 5'- GCCA TGCCCTGGAACGGAAGTCTCTCTAATAAAAAG-ATGTATCATTT-3'; reverse unmethylated: 5'- AAATGATACATCTTTTATTAGGAGAGITCCGTTCC-AGGGCATGGC-3'; forward mCpG: 5'- GCCATGCCCTGGAA(5-Me)CGGA ACTCTCTAATAAAA-AGATGTATCATTT-3'; reverse mCpG: 5'- AAATGA TACATCTTTTATTAGGAGAGITTC(5-Me)CGTT-CCAGGGCATGGC-3'.

The DNA fragments were purchased as ssDNA oligonucleotides and they were subsequently annealed to obtain 45-bp double-stranded DNA (dsDNA) for the interaction experiments. Briefly, they were dissolved to obtain a 0.5 mM ssDNA solution for each oligonucleotide; then, they were mixed at an equimolar ratio and were annealed using a Stratagene Mx3005P qPCR real-time thermal cycler (Agilent Technologies, Santa Clara, CA, USA). The thermal annealing profile consisted of four steps: 1) equilibration at 25 °C for 30 s; 2) heating ramp up to 99 °C; 3) equilibration at 99 °C for 60 s; and 4) 3-h cooling process down to 25 °C at a rate of 1 °C/180 s.

2.4. Circular dichroism

Far-UV circular dichroism spectra were recorded in a Chirascan (Applied Photophysics, Leatherhead, UK) using a 0.1 cm path-length quartz cuvette (Hellma Analytics, Müllheim, Germany). The temperature was controlled by a Peltier unit and monitored using a temperature probe.

Circular dichroism (CD) spectra were recorded in a Chirascan spectropolarimeter (Applied Photophysics, Leatherhead, UK) at 25 °C. Far-UV spectrum was recorded at wavelengths between 200 and 260 nm in a 0.1 cm path-length cuvette (Hellma Analytics, Müllheim, Germany). Protein concentration was 10–20 µM. The temperature was controlled by a Peltier unit and monitored using a temperature probe. Molar residue ellipticity was calculated considering the concentration of protein and the number of residues for each variant.

2.5. Fluorescence spectroscopy

Thermal unfolding studies were performed in a Cary Eclipse fluorescence spectrophotometer (Varian – Agilent, Santa Clara, CA, USA) using a 1 cm path-length quartz cuvette (Hellma Analytics, Müllheim, Germany). The temperature was controlled by a Peltier unit and monitored using a temperature probe. Fluorescence emission spectra were

recorded from 300 to 400 nm using an excitation wavelength of 290 nm and a bandwidth of 5 nm. Protein concentration was set at 5 µM.

Thermal stability assays were performed at a heating rate of 1 °C/min and at the emission wavelength for maximal spectral change (330 nm). Thermal unfolding of the protein was reversible and the experiments were analyzed considering a two-state unfolding model:

$$F = \frac{(A_N + B_N T) + (A_U + B_U T) \exp\left(-\frac{\Delta G(T)}{RT}\right)}{1 + \exp\left(-\frac{\Delta G(T)}{RT}\right)} \quad (1)$$

$$\Delta G(T) = \Delta H(T_m) \left(1 - \frac{T}{T_m}\right) + \Delta C_p \left(T - T_m - T \ln \frac{T}{T_m}\right)$$

where F is the fluorescence signal, T is the absolute temperature, T_m is the midtransition temperature, ΔH(T_m) is the unfolding enthalpy, ΔC_p is the unfolding heat capacity, and ΔG is the stabilization Gibbs energy. The adjustable parameters A_N, B_N, A_U, and B_U define the pre- (native) and post-transition (unfolded) regions in the temperature unfolding trace. The stabilizing effect upon dsDNA interaction was assessed by performing thermal denaturations of the different proteins (at 5 µM) in the presence of methylated and unmethylated DNA (at 10 µM) under the same conditions.

2.6. Isothermal titration calorimetry (ITC)

The interaction between the different proteins and dsDNA was characterized using an Auto-iTC200 microcalorimeter (MicroCal, Malvern-Panalytical, UK). Protein in the calorimetric cell at 3–5 µM was titrated with dsDNA at 50 µM. All solutions were degassed at 15 °C for 2 min before each assay. A sequence of 2 µL-injections of titrant solution every 150 s was programmed and the stirring speed was set to 750 rpm. The association constant, K_{B,obs}, and the enthalpy of binding, ΔH_{B,obs}, were estimated through non-linear regression of the experimental data employing a single ligand binding site model (1:1 protein:dsDNA stoichiometry) for ID interacting with dsDNA, or a two ligand binding sites model (1:2 protein:dsDNA stoichiometry) for NTD-MBD-ID and longer variants interacting with dsDNA, implemented in Origin (OriginLab, Northampton, MA). A detailed description of those models applied in ITC can be found elsewhere [31,32]. The dissociation constant K_d was calculated as the inverse of K_{B,obs}, and the binding Gibbs energy and entropy were calculated applying standard well-known relationships: ΔG = –RT lnK_B, ΔG = ΔH – TΔS.

The association constant K_{B,obs} is not affected by the buffer ionization as long as the pK_a of the buffer is close to the experimental pH. However, the observed binding enthalpy (and, therefore, the entropic contribution) will contain an additional contribution from buffer ionization. Those extrinsic contributions from the buffer can be removed. The buffer-independent binding enthalpy, ΔH, was determined according to [33–35]:

$$\Delta H_{B,obs} = \Delta H + \Delta n_H \Delta H_{buffer} \quad (2)$$

where Δn_H is the net number of protons exchanged between the protein–dsDNA complex and the bulk solution upon dsDNA binding, and ΔH_{buffer} is the ionization enthalpy of the buffer. Titrations were performed in buffers with different ionization enthalpies (Tris, 11.35 kcal/mol; Pipes, 2.67 kcal/mol; and phosphate, 0.86 kcal/mol) [36] in order to estimate the buffer-independent thermodynamic parameters (ΔH and Δn_H) from linear regression using eq. 2. Knowing the binding Gibbs energy and the buffer-independent binding enthalpy, the buffer-independent binding entropy can be readily calculated. The parameter Δn_H may be non-zero when ligand binding results in changes in the proton dissociation constant of certain ionizable residues (in the protein or the ligand) as a consequence of changes in their microenvironment upon complex formation. Interestingly,

according to Wyman's linkage relationships [37], Δn_H reports the change in binding affinity as a result of a change in pH:

$$n_H = - \left(\frac{\partial \log K_{B,obs}}{\partial pH} \right)_p \quad (3)$$

3. Results

3.1. Distant domains in MeCP2 exert a considerable stabilizing effect on MBD

The poor circular dichroism signal related to the low content in secondary structure of the protein variants (Fig. S2) and its small change during the thermal denaturation process within the 10–90 °C temperature range favored the use of fluorescence spectroscopy in the thermal unfolding assays. The single tryptophan residue in MeCP2, located in the MBD, allowed monitoring the thermal denaturation of the different variants by focusing specifically on the intrinsic stability of this domain. Importantly, the lack of tryptophan residues in all other domains was instrumental for directly observing the stabilizing effect of the other domains on the MBD. There is an additional tyrosine in full length MeCP2, but its contribution to the overall fluorescence intensity is expected to be small. No major differences were observed in the shape of the spectra of the three variants; all of them showed a maximum around 330 nm, which underwent a reduction in intensity and a red-shift to 335 nm when the temperature was increased. The fluorescence emission intensity was strongly affected by the presence of distant domains (Fig. S2), indicating those domains exert an influence on the tryptophan environment in MBD through inter-domain long-distance interactions. Major differences between the variants were the emission intensity temperature range at which the intensity decrease and the red-shift occur, and this is what the unfolding fluorescence traces captured (Fig. 1).

The three variants (NTD-MBD-ID-TRD, NTD-MBD-ID-TRD-CTD α , and NTD-MBD-ID-TRD-CTD α -CTD β /full length) showed a minor stability dependency on the pH, much smaller for the full-length protein (Fig. S1, and Table S2). This indicates the potential influence of ionizable amino acids in this pH range is negligible, and the selection of pH 7 for further experiments will not condition the reliability of the results. The unfolding temperature, T_m , showed an increasing trend with the length of the protein variant at pH 7, although that stabilizing effect already reaches the maximal extent when NTD and ID are present (Fig. 2 and Table S2). In addition, physiological high ionic strength (NaCl 150 mM) increased the stability for all variants at pH 7 (Figs. 1 and S2-S3, and Table S2), with regard to low ionic strength, which indicates that the unfolding process is coupled to the uptake of ions from the bulk solution, as previously observed with other smaller MeCP2 variants [29]. The experiments at high ionic strength also confirm the increasing trend for T_m with the length of the protein variant.

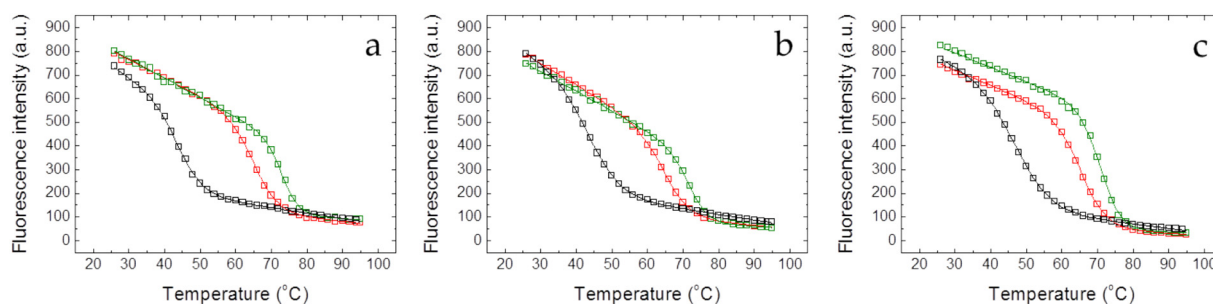


Fig. 1. Fluorescence thermal denaturations for the MeCP2 variants (a) NTD-MBD-ID-TRD, (b) NTD-MBD-ID-TRD-CTD α , and (c) NTD-MBD-ID-TRD-CTD α -CTD β , at pH 7 and low ionic strength. The influence of the presence of dsDNA on the thermal stability was assessed for the three variants (dsDNA-free, black squares; unmethylated CpG-dsDNA, red squares; and methylated mCpG-dsDNA, green squares). All unfolding traces could be fitted employing a two-state unfolding model (continuous lines) according to eq. 1. For clarity purposes, only one out two experimental data points are shown. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

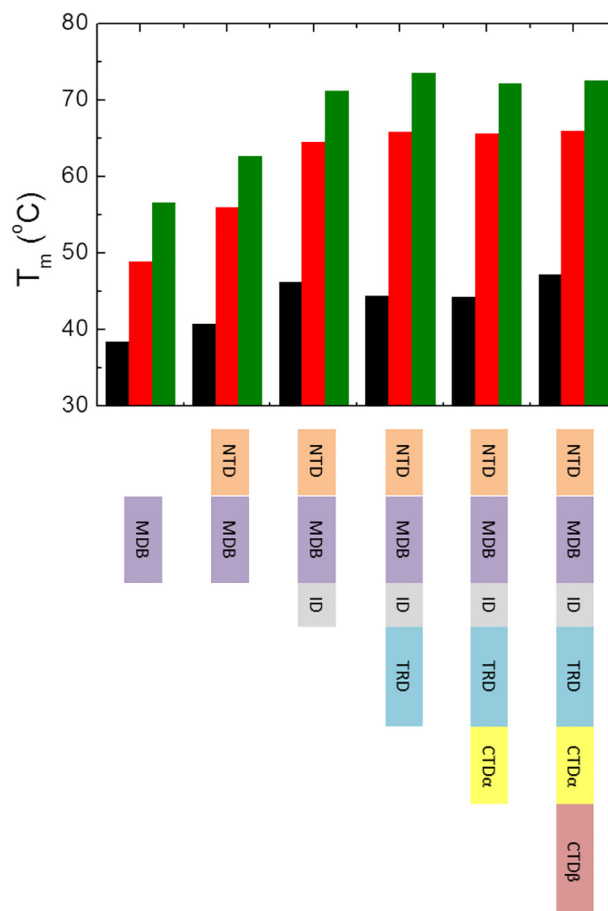


Fig. 2. Unfolding temperature of all MeCP2 protein variants in the absence of dsDNA (black bars), and the presence unmethylated CpG-dsDNA (red bars) and methylated mCpG-dsDNA (green bars), obtained from thermal denaturations followed by intrinsic tryptophan fluorescence.

The presence of dsDNA stabilized all protein variants (Figs. 1 and 2, and Table S3). Longer protein variants underwent a larger stabilization effect upon dsDNA interaction. Importantly, the stabilization effect was considerably larger for methylated mCpG-dsDNA, compared to unmethylated CpG-dsDNA, which would point to a higher binding affinity for methylated dsDNA (Figs. 1 and 2, and Table S3). The difference in the stabilization effect exerted by both dsDNA types was conserved for all variants (Fig. 2), pointing to an affinity difference conserved for all protein variants. At high ionic strength the stabilizing effect of dsDNA was smaller (Fig. S4, and Table S4), which indicates that the dsDNA binding affinity decreases with the ionic strength, and that the dsDNA binding must be coupled to the release of ions from the complex to

the bulk solution. Still, the stabilization effect was considerably larger for methylated mCpG-dsDNA, compared to unmethylated CpG-dsDNA. Somewhat different results, but showing a similar stabilization trend, have been reported before under different conditions [15].

3.2. Distant domains in MeCP2 exert a minor influence on dsDNA binding to MBD and ID

The ability of ID to interact with DNA was reported before [4]. In fact, some mutations in ID hinder MeCP2 chromatin association [4,38]. It was previously observed that the presence of ID (variant NTD-MBD-ID) has important consequences regarding the dsDNA binding capabilities of the MBD: 1) the ID provides another dsDNA binding site that is functionally independent from that located in MBD (i.e., the variant NTD-MBD-ID is able to interact with two dsDNA fragments); and 2) the ID dramatically increases the dsDNA binding affinity of the MBD [29]. However, a precise binding affinity determination for ID was lacking.

Therefore, we first assessed the ability of isolated ID to interact with dsDNA. ID interacted with submicromolar binding affinity with both methylated and unmethylated dsDNA (Fig. 3 and Table S5), therefore, confirming its ability to interact with dsDNA and its inability to discriminate between both dsDNA types. Surprisingly, although ID is a completely disordered domain, the dsDNA interaction is characterized by an unfavorable binding enthalpy and a largely favorable binding entropy, likely associated with a large desolvation of binding interfaces.

Then, we proceeded to assess the dsDNA binding characteristics of MeCP2 variants. The dsDNA binding parameters for NTD-MBD-ID-TRD and NTD-MBD-ID-TRD-CTD α variants are quite similar, and essentially identical to those observed for NTD-MBD-ID (Fig. 4 and Table S5), except for a less exothermic binding of dsDNA to the MBD high-affinity binding site. Therefore, we proceeded to evaluate the dsDNA binding for the NTD-MBD-ID-TRD-CTD α -CTD β variant (full-length MeCP2), which is the most physiologically relevant protein variant (except in the case of clinically relevant deletion mutations associated with RTT). Except a slightly higher affinity for the second low-affinity binding site and a more endothermic binding for the low-affinity binding site, the interaction of the three variants (NTD-MBD-ID-TRD and NTD-MBD-ID-TRD-CTD α , and NTD-MBD-ID-TRD-CTD α -CTD β) with dsDNA was similar (Fig. 5, and Table S5). The differences in the binding affinity between

methylated and unmethylated dsDNA (i.e., the ability to discriminate between both dsDNA forms) in both binding sites for these three variants seem to be similar to those observed for NTD-MBD-ID.

4. Discussion

The structural and functional role of disordered regions in proteins is controversial. These regions are characterized by a biased amino acid composition, where residues exhibiting considerable propensity to be exposed to the solvent predominate (polar and charged amino acids) [39]. However, they may exert a steric hindrance effect, or establish attractive or repulsive electrostatic interactions, or make key contacts with other structured regions and affect the global stability and the dynamics of the protein, as well as modulate the interaction with binding partners. Therefore, the impact of disordered regions on the global stability may be based on specific or unspecific effects. Specific effects may derive from long-lived or transient interactions between residues from disordered and structured regions, while unspecific effects may be due to reciprocal constrained flexibility/mobility of the polypeptide chain because of steric or electrostatic hindrance. Long-range electrostatic and dipolar interactions are extremely important in IDPs, especially at low ionic strength, because of the large fraction of charged and polar residues. Therefore, it may be possible that, even lacking well-defined structure, disordered regions can contribute to the overall stability of the protein [40]. In the case of MeCP2, we have recently provided evidence for that phenomenon: 1) NTD and ID, two completely disordered domains flanking the MBD, increase significantly the thermal stability of the MBD [29]; and 2) the two MeCP2 isoforms E1 and E2, which only differ in a few N-terminal aminoacids in the completely disordered domain NTD, show different thermal stability [28].

As indicated above, the purpose of this work was to determine: 1) if distant disordered domains in MeCP2 would affect MBD structural and functional features, 2) if a longer variant for MeCP2 would be more appropriate or compulsory for structural/functional assays, and 3) if the presence of additional AT-hooks (in TRD and CTD α) would result in additional functionally independent DNA binding sites.

The single tryptophan residue located in the MBD facilitated the task of observing local stability changes in the MBD induced by other distant domains. The presence of additional partially disordered domains enhanced the stability of MBD, although the increase in stability was

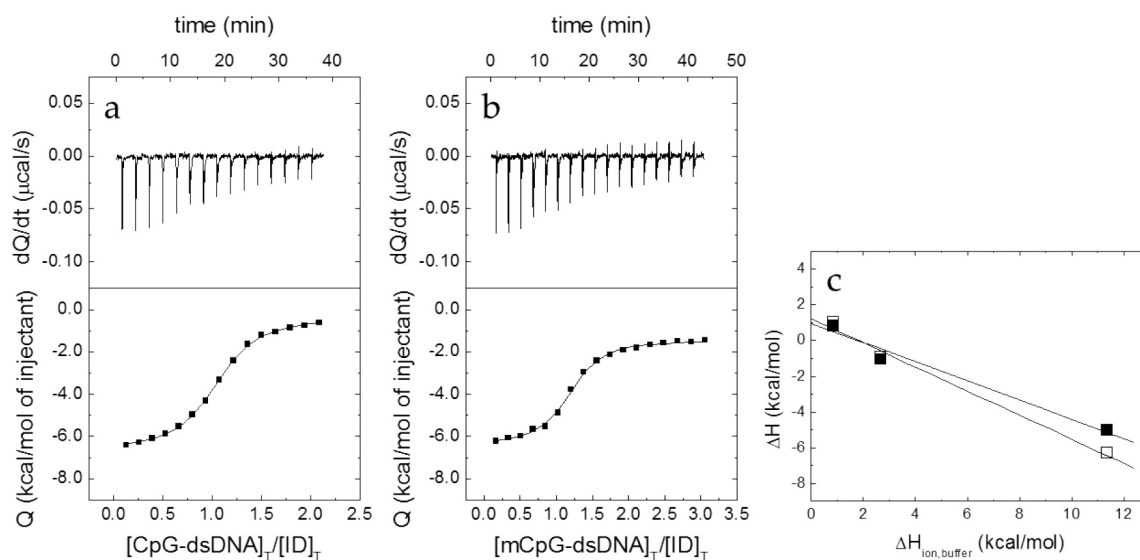


Fig. 3. Interaction of the ID with dsDNA by isothermal titration calorimetry (ITC) at pH 7. The interaction of unmethylated CpG-dsDNA (a) and methylated mCpG-dsDNA (b) was assessed at pH 7. Upper plots show the thermogram (raw thermal power as a function of time) and the lower plots show the binding isotherm (ligand-normalized heat effects as a function of the molar ratio). Non-linear least-squares analysis using a single binding site model allowed estimating the observed binding affinity and enthalpy (continuous lines). (c) Eq. 2 was employed for estimating the buffer-independent binding enthalpy: CpG-dsDNA (open squares) and mCpG-dsDNA (closed squares). Errors for the observed binding enthalpies are smaller than the size of the symbols.

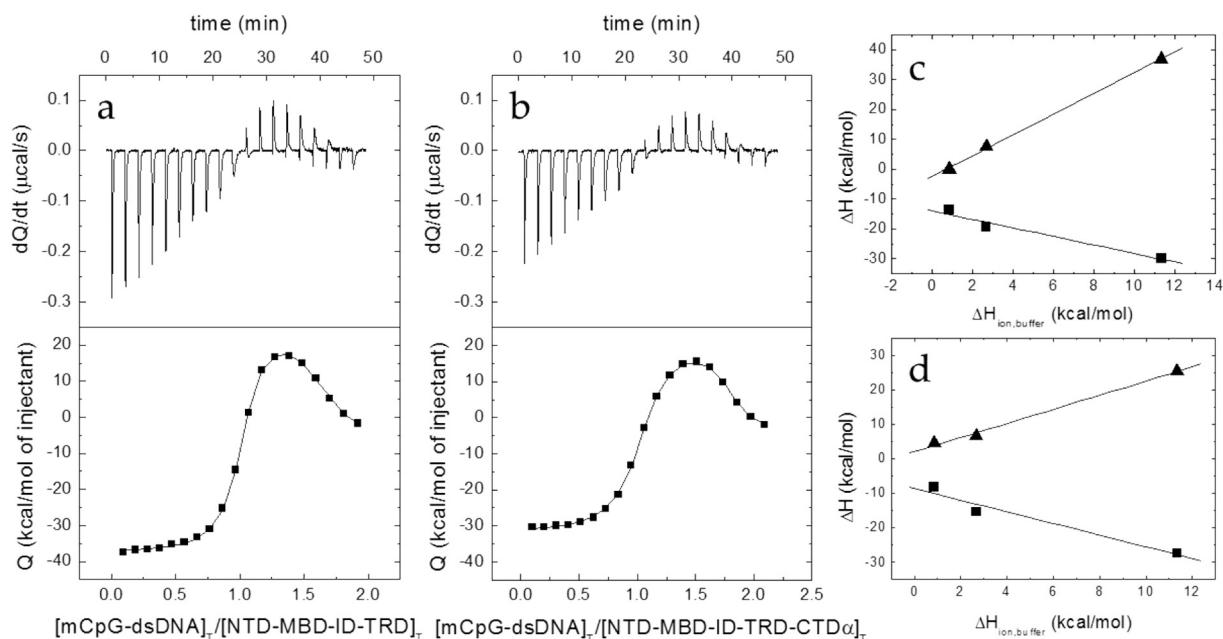


Fig. 4. Interaction of the NTD-MBD-ID-TRD (a) and NTD-MBD-ID-TRD-CTD α (b) with methylated mCpG-dsDNA by isothermal titration calorimetry (ITC) at pH 7. Upper plots show the thermogram (raw thermal power as a function of time) and the lower plots show the binding isotherm (ligand-normalized heat effects as a function of the molar ratio). Non-linear least-squares analysis using a two binding sites model allowed estimating the observed binding affinity and enthalpy (continuous lines). Eq. 2 was employed for estimating the buffer-independent binding enthalpy for the MBD high-affinity binding site (squares) and the ID low-affinity binding site (triangles) for NTD-MBD-ID-TRD (c) and NTD-MBD-ID-TRD-CTD α (d). Errors for the observed binding enthalpies are smaller than the size of the symbols.

smaller the more distant the domain. An increase in the unfolding temperature could be observed when increasing the length of the polypeptide chain (Fig. 2 and Table S2). Albeit the unfolding temperature may often be a quick, convenient index to judge or rank structural stability, it is an index reporting stability at high temperature, and very frequently a protein stability ranking based on the unfolding temperature does not correlate with a protein stability ranking based on the stabilization energy at room or physiological temperature. Thus, the stabilization Gibbs

energy at 20 °C for each variant was calculated employing Eq. (1) (Fig. 6). The calculation of the stabilization Gibbs energy at 20 °C involves a substantial extrapolation, in some cases over a range of more than 50 °C, and caution must be taken when performing these calculations. A similar trend to that of T_m can be observed for the stabilization Gibbs energy: the longer the variant, the larger the stability Gibbs energy. In addition, the binding of methylated and unmethylated dsDNA further increases the protein stability; in all protein variants, the stabilization effect

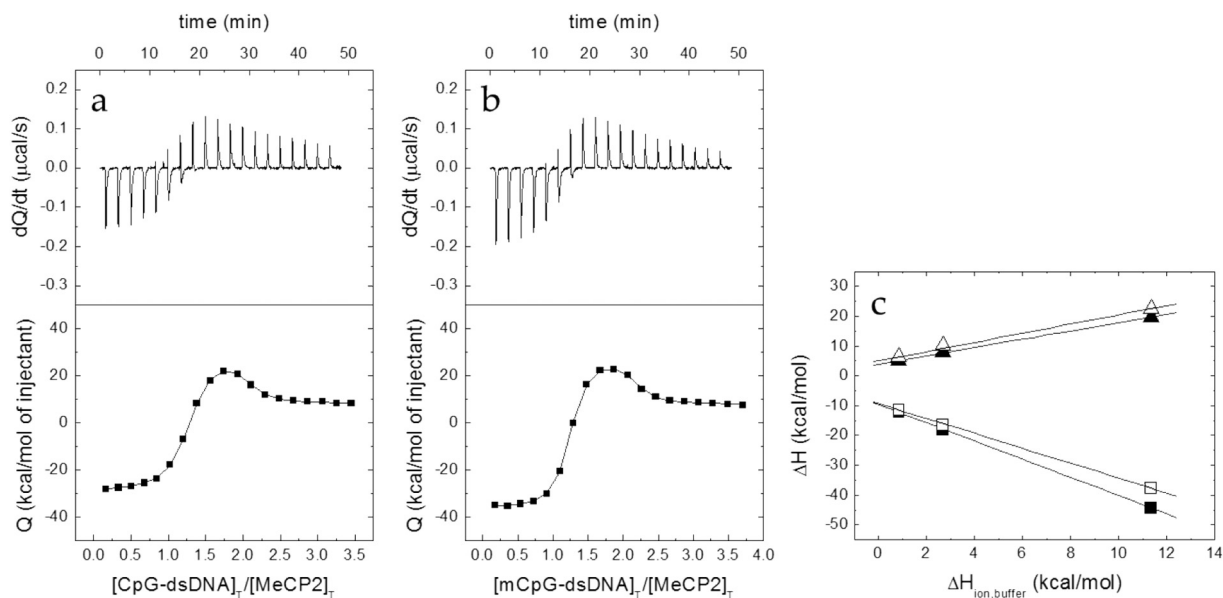


Fig. 5. Interaction of the NTD-MBD-ID-TRD-CTD α -CTD β (full-length MeCP2) with unmethylated CpG-dsDNA (a) and unmethylated CpG-dsDNA (b) by isothermal titration calorimetry (ITC) at pH 7. Upper plots show the thermogram (raw thermal power as a function of time) and the lower plots show the binding isotherm (ligand-normalized heat effects as a function of the molar ratio). Non-linear least-squares analysis using a two binding sites model allowed estimating the observed binding affinity and enthalpy (continuous lines). (c) Eq. 2 was employed for estimating the buffer-independent binding enthalpy for the MBD high-affinity binding site (squares) and the ID low-affinity binding site (triangles) CpG-dsDNA (open symbols) and mCpG-dsDNA (closed symbols). Errors for the observed binding enthalpies are smaller than the size of the symbols.

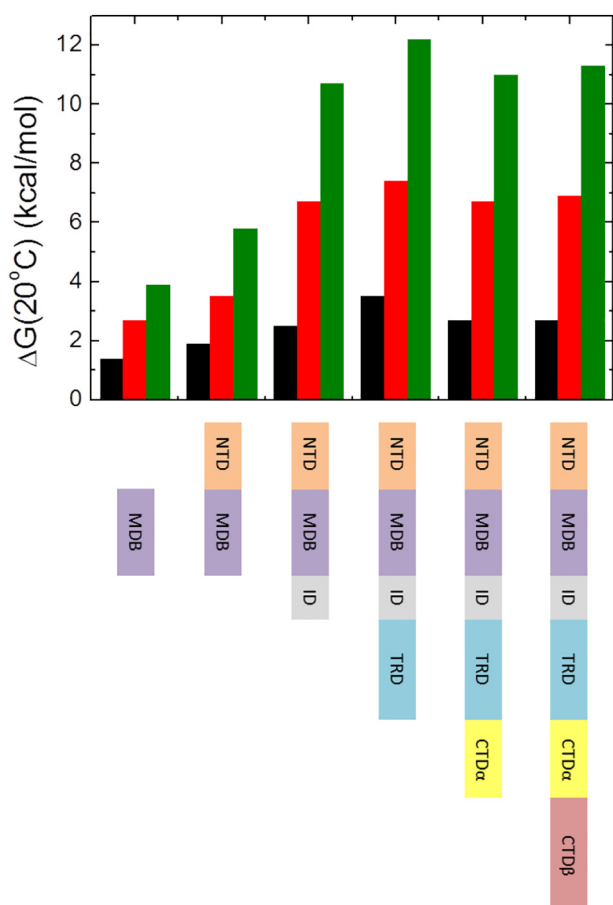


Fig. 6. Stability Gibbs energy at 20 °C of all MeCP2 protein variants in the absence of dsDNA (black bars), and the presence unmethylated CpG-dsDNA (red bars) and methylated mCpG-dsDNA (green bars). The stability Gibbs energy was estimated considering a two-state unfolding model (eq. 1).

of methylated dsDNA is considerably larger than that of unmethylated dsDNA. The overall stabilization Gibbs energy, ΔG , in the presence of an interacting ligand is equal to the intrinsic stabilization Gibbs energy, ΔG_0 , plus the excess average binding Gibbs energy, $\langle \Delta G_B \rangle$, which is the contribution to the stability from ligand binding. Focusing on full-length MeCP2, the intrinsic stabilization Gibbs energy is 2.7 kcal/mol, and the binding of unmethylated CpG-dsDNA and methylated mCpG-dsDNA increases the stabilization energy to 6.9 kcal/mol and 11.3 kcal/mol, respectively (Fig. 6). Then, the binding of dsDNA contributes 60% and 75% of the global stabilization energy of the complex (at the 1:2 protein:dsDNA concentration ratio employed in the thermal unfolding experiments, protein 5 μM and DNA 10 μM). From those data, it could be reasonable to hypothesize a higher binding affinity for methylated mCpG-dsDNA compared to unmethylated CpG-dsDNA.

ITC is currently considered the gold-standard for determining binding affinities in biomolecular interactions. This technique allows the simultaneous determination of the binding affinity (association constant, dissociation constant, or Gibbs energy of binding) and the binding enthalpy, as well as the stoichiometry. Very importantly, the binding affinity will not be affected by ionization properties of the buffer, as long as the experimental pH is close to the pKa of the buffer. On the contrary, the binding enthalpy (and, therefore, the binding entropy also) might contain considerable contributions from the ionization enthalpy and entropy of the buffer. Fortunately, there is a straightforward procedure that allows removing the buffer contributions, thus obtaining the buffer-independent binding parameters, and at the same time estimating the net number of protons that are exchanged between complex and bulk solution upon complex formation. This quantity may be explored in

detail in order to get insight into the ionizable functional groups involved in the binding process [41,42], and, from the practical point of view, it provides information, although limited, on the dependency of the binding affinity with pH (eq. 3). In addition, ITC is very well suited for studying biological interactions with more than one binding site, as it happens with NTD-MBD-ID and longer MeCP2 variants, since the interplay between binding affinity and enthalpy makes easy to observe different binding process occurring at different locations on a macromolecule through a multiphasic binding isotherm.

Overall, except some differences in the binding enthalpies, the three variants considered in this work do not behave differently compared to NTD-MBD-ID regarding dsDNA binding affinity, difference in binding affinity between the high- and low-affinity sites, ability to discriminate between unmethylated and methylated dsDNA in both sites, and binding stoichiometry. Focusing on the full-length protein, the interaction with methylated mCpG-dsDNA shows a more favorable enthalpy in the high-affinity site and a less unfavorable enthalpy in the low-affinity site. For both dsDNAs, the binding to the high-affinity site is enthalpically driven and the binding to the low-affinity site is entropically driven. Still, the entropic contribution to the binding of both sites is favorable in both cases. In addition, the dsDNA binding to the high-affinity site is coupled to a net deprotonation ($\Delta n_H < 0$) in the complex, while the dsDNA binding to the low-affinity site is coupled to a net protonation ($\Delta n_H > 0$).

There are several intriguing facts derived from the experimental results reported here. First, the large differences in stabilization effects induced by unmethylated CpG-dsDNA and methylated mCpG-dsDNA (in terms of increases in T_m or $\Delta G(20^\circ\text{C})$) do not correlate with the small differences in binding affinities estimated from ITC. This was previously observed, and it is under further scrutiny in our group. One explanation could be that in experimental assays where the steady-state effect of the interaction between protein and ligand is observed after long incubations (e.g., steady-state spectroscopy, chemical and thermal denaturations, inhibition assays) there is sufficient time for the complex to achieve the optimal binding configuration, whereas in experimental assays where the transient effect of the interaction in the first minutes after instantaneous mixing (e.g., ITC) there might not be sufficient time to achieve optimal final binding configuration and lower binding affinities may be estimated. In this sense, differences in binding affinities as determined by different techniques or a slow adaptation of the ligand within the binding site to achieve an optimal binding affinity have been observed before for biological systems [43–45]. However, steady-state fluorescence spectroscopy experiments provided binding affinities for methylated and unmethylated dsDNA similar to those determined by ITC (Fig. S5). MeCP2 was initially identified as a methyl-specific dsDNA binder, as confirmed by cellular and in vivo assays. Would the absence of discrimination between unmethylated CpG-dsDNA and methylated mCpG-dsDNA in both binding sites in in vitro assays be an indication for the need of an additional factor (a protein cofactor?) allowing such effective methyl-dependent discrimination? Maybe this is something requiring further attention.

Second, the presence of additional domains with potential dsDNA binding motifs (AT-hooks in TRD and CTD α , besides the binding sites in MBD and ID) did not increase the dsDNA binding stoichiometry in MeCP2. The maximum number of dsDNA molecules MeCP2 is able to interact with simultaneously is two (1:2 MeCP2:dsDNA). Other stoichiometries or binding configurations (e.g., 2:1 MeCP2:dsDNA, or 1:1 MeCP2:dsDNA) are not compatible with the results presented here. This maximal stoichiometry already arises in the NTD-MBD-ID variant. Therefore, no additional, functionally independent DNA binding sites could be observed when TRD and both CTDs were present. Still, this may be sufficient for its genome-wide chromatin remodeling/compacting function. From the results presented here, there is no possibility to conclude unambiguously where the main DNA-attachment spot for the second, low-affinity binding site is located in the three variants. That site could be located in ID, TRD or CTD α . There is a need for structural data

(e.g., nuclear magnetic resonance) shedding light on this matter, but an independent binding site has already been identified and characterized in ID. It would be possible that those additional dsDNA binding motifs from TRD and CTD α , different from the main one located in the MBD and the secondary one in ID, would wrap around the same dsDNA fragments, thus cooperating to achieve a high binding affinity and a high kinetic barrier towards dissociation.

And third, the 45-bp dsDNA employed in the experiments would be long enough to accommodate several MeCP2 molecules bound, at least for the smaller variants (e.g., MBD or NTD-MBD-ID). However, according to the stoichiometry observed in the ITC experiments (this work and [29]) and ultracentrifugation experiments [29], no simultaneous binding of more than one MeCP2 molecule on the dsDNA molecule could be observed, thus, ruling out cooperativity for MeCP2 binding to dsDNA *in vitro*.

5. Conclusions

MeCP2 is a potential pharmacological protein target associated with RTT (caused by lack of MeCP2 activity) and MDS (caused by excess of MeCP2 activity), two neurological disorders with similar phenotypic features. MeCP2 is mainly involved in neuronal development and maturation, and synaptic plasticity. MeCP2 is a partially disordered, multiple-domain protein exhibiting multiple functions: activation/repression of transcription at specific promoter locations, genome-wide chromatin remodeling and nucleosomal compaction, and pre-mRNA maturation and splicing, among the most important ones. The structural properties (modularity and plasticity) in MeCP2 are responsible for its complex conformational/functional landscape and its many biological functions.

No drugs acting specifically on MeCP2 have been developed so far. Therefore, high-throughput screening programs for identifying bioactive molecules capable of modulating MeCP2 function would be desirable. However, heterologous expression of MeCP2 for *in vitro* assays exhibits some problems undermining its tractability: low expression yield, high propensity to degradation, and high propensity to precipitation. Although some appropriate MeCP2 constructs or variants can be envisaged considering MeCP2 modularity and the inferred functions for each domain, as well as the purpose of the assay (e.g., molecular screening, *in vitro* function, cell-based assay...), the assessment of the stability and the functional features of the different variants would be necessary. In this work we have evaluated the structural stability (stability against thermal denaturation, the stabilization effect of unmethylated and methylated dsDNA), and the ability to interact with unmethylated and methylated dsDNA (determination of binding affinity and enthalpy) of different MeCP2 variants with variable length. In particular, the NTD-MBD-ID variant would be appropriate for identifying small molecules able to recover or inhibit MeCP2 interaction with dsDNA (intended for therapeutic treatment of RTT or MDS, respectively), since it has sufficient structural stability and its functional behavior related to dsDNA interaction is similar to other longer variants, including full length MeCP2. Nevertheless, if the compound screening is aimed at identifying small molecules modulating MeCP2 protein-protein interactions, then, longer variants including TRD and/or CTDs would be required.

CRedit authorship contribution statement

Conceptualization, M.E., O.A. and A.V.C.; methodology, O.A. and A.V.C.; software, O.A. and A.V.C.; validation, D.O.A., R.C.G., S.V., O.C.J.T., M.E., O.A. and A.V.C.; formal analysis, D.O.A., R.C.G., S.V., O.C.J.T., M.E., O.A. and A.V.C.; investigation, D.O.A., R.C.G., S.V. and O.C.J.T.; resources, M.E., O.A. and A.V.C.; data curation, D.O.A. and R.C.G.; writing—original draft preparation, O.A. and A.V.C.; writing—review and editing, D.O.A., R.C.G., S.V., O.C.J.T., M.E., O.A. and A.V.C.; visualization, D.O.A., R.C.G., S.V., O.C.J.T., M.E., O.A. and A.V.C.; supervision, M.E., O.A. and A.V.C.; project administration, O.A. and A.V.C.; funding acquisition, O.A. and A.V.C. All authors have read and agreed to the published version of the manuscript.

Declaration of competing interest

The authors declare no conflict of interest.

Acknowledgments

This research was funded by the Spanish Ministry of Economy and Competitiveness and European ERDF Funds (MCIU/AEI/FEDER, EU) (BFU2016-78232-P to A.V.C.; BES-2017-080739 to D.O.A.); Miguel Servet Program from Instituto de Salud Carlos III (CPII13/00017 to OA); Fondo de Investigaciones Sanitarias from Instituto de Salud Carlos III and European Union (ERDF/ESF, “Investing in your future”) (PI15/00663 and PI18/00349 to O.A.); Diputación General de Aragón (Protein Targets and Bioactive Compounds Group E45_17R to A.V.C. and Digestive Pathology Group B25_17R to O.A.); and the Centro de Investigación Biomédica en Red en Enfermedades Hepáticas y Digestivas (CIBERehd).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ijbiomac.2021.01.206>.

References

- [1] J. Ausio, A. Martínez de Paz, M. Esteller, MeCP2: the long trip from a chromatin protein to neurological disorders, *Trends Mol. Med.* 20 (2014) 487–498.
- [2] N.L. Adkins, P.T. Georgel, MeCP2: structure and function, *Biochem. Cell Biol.* 89 (2011) 1–11.
- [3] X. Nan, F.J. Campoy, A. Bird, MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin, *Cell* 88 (1997) 471–481.
- [4] J.C. Hansen, R.P. Ghosh, C.L. Woodcock, Binding of the Rett syndrome protein, MeCP2, to methylated and unmethylated DNA and chromatin, *IUBMB Life* 62 (2010) 732–738.
- [5] J. Guy, H. Cheval, J. Selfridge, A. Bird, The role of MeCP2 in the brain, *Annu. Rev. Cell Dev. Biol.* 27 (2011) 631–652.
- [6] R. Claveria-Gimeno, O. Abian, A. Velazquez-Campoy, J. Ausio, MeCP2... Nature's wonder protein or medicine's most feared one? *Curr Genet Med Rep* 4 (2016) 180–194.
- [7] R. Tillotson, A. Bird, The molecular basis of MeCP2 function in the brain, *J. Mol. Biol.* 432 (2020) 1602–1623.
- [8] R.E. Amir, I.B. Van den Veyver, M. Wan, C.Q. Tran, U. Francke, H.Y. Zoghbi, Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2, *Nat. Genet.* 23 (1999) 185–188.
- [9] J. Lorenz, M.D. Neul, The relationship of Rett syndrome and MeCP2 disorders to autism, *Dialogues Clin. Neurosci.* 14 (2014) 253–262.
- [10] A.J. Sandweiss, V.L. Brandt, H.Y. Zoghbi, Advances in understanding of Rett syndrome and MECP2 duplication syndrome: prospects for future therapies, *Lancet Neurol.* 19 (2020) 689–698.
- [11] K.C. Hite, V.H. Adams, J.C. Hansen, Recent advances in MeCP2 structure and function, *Biochem. Cell Biol.* 87 (2009) 219–227.
- [12] X. Nan, R.R. Meehan, A. Bird, Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2, *Nucleic Acids Res.* 21 (1993) 4886–4892.
- [13] H. Adam, P.T. Georgel, CHD proteins: a diverse family with strong ties, *Biochem. Cell Biol.* 85 (2007) 463–476.
- [14] R.R. Shah, A.P. Bird, MeCP2 mutations: Progress towards understanding and treating Rett syndrome, *Genome Med* 9 (2017) 17.
- [15] R.P. Ghosh, T. Nikitina, R.A. Horowitz-Scherer, L.M. Gierasch, V.N. Uversky, K. Hite, J.C. Hansen, C.L. Woodcock, Unique physical properties and interactions of the domains of methylated DNA binding protein 2, *Biochemistry* 49 (2010) 4395–4410.
- [16] R.I.D. Wakefield, B.O. Smith, X. Nan, A. Free, A. Soteriou, D. Uhrin, A.P. Bird, P.N. Barlow, The solution structure of the domain from MeCP2 that binds to methylated DNA, *J. Mol. Biol.* 291 (1999) 1055–1065.
- [17] K.L. Ho, I.W. McNae, L. Schmiedeberg, R.J. Klose, A.P. Bird, M.D. Walkinshaw, MeCP2 binding to DNA depends upon hydration at methyl-CpG, *Mol. Cell* 29 (2008) 525–531.
- [18] R.P. Ghosh, R.A. Horowitz-Scherer, T. Nikitina, L.M. Gierasch, C.L. Woodcock, Rett syndrome-causing mutations in human MeCP2 result in diverse structural changes that impact folding and DNA interactions, *J. Biol. Chem.* 283 (2008) 20523–20534.
- [19] E. Ballestar, T.M. Yusufzai, A.P. Wolffe, Effects of Rett syndrome mutations of the methyl-CpG binding domain of the transcriptional repressor MeCP2 on selectivity for association with methylated DNA, *Biochemistry* 39 (2000) 7100–7106.
- [20] M.F. Fraga, E. Ballestar, G. Montoya, P. Taysavang, P.A. Wade, M. Esteller, The affinity of different MBD proteins for a specific methylated locus depends on their intrinsic binding properties, *Nucleic Acids Res.* 31 (2003) 1765–1774.
- [21] J. Singh, A. Saxena, J. Christodoulou, D. Ravine, MeCP2 genomic structure and function, *Nucleic Acids Res.* 36 (2008) 6035–6047.
- [22] R.J. Klose, A.P. Bird, Genomic DNA methylation: the mark and its mediators, *Trends Biochem. Sci.* 32 (2006) 89–97.

- [23] R.P. Ghosh, R.A. Horowitz-Scherer, T. Nikitina, L.S. Shlyakhtenko, C.L. Woodcock, MeCP2 binds cooperatively to its substrate and competes with histone H1 for chromatin binding sites, *Mol. Cell. Biol.* 30 (2010) 4656–4670.
- [24] P.T. Georgel, R.A. Horowitz-Scherer, N. Adkins, C.L. Woodcock, P.A. Wade, J.C. Hansen, Chromatin compaction by human MeCP2, Assembly of novel secondary chromatin structures in the absence of DNA methylation, *J Biol Chem* 278 (2003) 32181–32188.
- [25] P.J. Skene, R.S. Illingworth, S. Webb, A.R. Kerr, K.D. James, D.J. Turner, R. Andrews, A.P. Bird, Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state, *Mol. Cell* 37 (2010) 457–468.
- [26] S. Cohen, H.W. Gabel, M. Hemberg, A.N. Hutchinson, L.A. Sadacca, D.H. Ebert, D.A. Harmin, R.S. Greenberg, V.K. Verdine, Z. Zhou, W.C. Wetsel, A.E. West, M.E. Greenberg, Genome-wide activity-dependent MeCP2 phosphorylation regulates nervous system development and function, *Neuron* 72 (2011) 72–85.
- [27] B. Kinde, H.W. Gabel, C.S. Gilbert, E.C. Griffith, M.E. Greenberg, Reading the unique DNA methylation landscape of the brain: non-CpG methylation, hydroxymethylation, and MeCP2, *Proc. Natl. Acad. Sci. U. S. A.* 112 (2015) 6800–6806.
- [28] A. Martinez de Paz, L. Khajavi, H. Martin, R. Claveria-Gimeno, S. Tom Dieck, M.S. Cheema, J.V. Sanchez-Mut, M.M. Moksá, A. Carles, N.I. Brodie, T.I. Sheikh, M.E. Freeman, E.V. Petrotchenko, C.H. Borchers, E.M. Schuman, M. Zytznicki, A. Velazquez-Campoy, O. Abian, M. Hirst, M. Esteller, J.B. Vincent, C.E. Malnou, J. Ausio, MeCP2-E1 isoform is a dynamically expressed, weakly DNA-bound protein with different protein and DNA interactions compared to MeCP2-E2, *Epigenetics Chromatin* 12 (2019) 63.
- [29] R. Claveria-Gimeno, P. Lanuza, I. Morales-Chueca, O.C. Jorge-Torres, S. Vega, O. Abian, M. Esteller, A. Velazquez-Campoy, The intervening domain from MeCP2 enhances the DNA affinity of the methyl binding domain and provides an independent DNA interaction site, *Sci. Rep.* 7 (2017) 41635.
- [30] T.A. Hall, BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT, *Nucleic Acids Symp. Ser.* 41 (1999) 95–98.
- [31] S. Vega, O. Abian, A. Velazquez-Campoy, A unified framework based on the binding polynomial for characterizing biological systems by isothermal titration calorimetry, *Methods* 76 (2015) 99–115.
- [32] E. Freire, A. Schön, A. Velazquez-Campoy, Isothermal titration calorimetry: general formalism using binding polynomials, *Methods Enzymol.* 455 (2009) 127–155.
- [33] M. Eftink, R.L. Biltonen, Thermodynamics of interacting biological systems. In *Biological Microcalorimetry*; A.E. Beezer, Ed.; Academic Press: London, UK, 1980; pp. 343–412.
- [34] H.J. Hinz, D.D.F. Shiao, J.M. Sturtevant, Calorimetric investigation of inhibitor binding to rabbit muscle aldolase, *Biochemistry* 10 (1971) 1347–1352.
- [35] J. Gomez, E. Freire, Thermodynamic mapping of the inhibitor binding site of the aspartic protease Endothiapepsin, *J. Mol. Biol.* 252 (1995) 337–350.
- [36] R.N. Goldberg, N. Kishore, R.M. Lennen, Thermodynamic quantities for the ionization reactions of buffers, *J. Phys. Chem. Ref. Data* 31 (2002) 231–370.
- [37] J. Wyman, Linked functions and reciprocal effects in hemoglobin: a second look, *Adv. Protein Chem.* 19 (1964) 223–286.
- [38] T.I. Sheikh, R. Harripaul, M. Ayub, J.B. Vincent, MeCP2 AT-Hook1 mutations in patients with intellectual disability and/or schizophrenia disrupt DNA binding and chromatin compaction in vitro, *Hum. Mutat.* 39 (2018) 717–728.
- [39] V.N. Uversky, Unusual biophysics of intrinsically disordered proteins, *Biochim. Biophys. Acta, Proteins Proteomics* 1834 (2013) 932–951.
- [40] S.A. Leavitt, A. Schön, J.C. Klein, U. Manjappara, I.M. Chaiken, E. Freire, Interactions of HIV-1 proteins gp120 and Nef with cellular partners define a novel allosteric paradigm, *Curr. Protein Pept. Sci.* 5 (2004) 1–8.
- [41] A. Velazquez-Campoy, I. Luque, M.J. Todd, M. Milutinovich, Y. Kiso, E. Freire, Thermodynamic dissection of the binding energetics of KNI-272, a potent HIV-1 protease inhibitor, *Protein Sci.* 9 (2000) 1801–2000.
- [42] D. Xie, S. Gulnik, L. Collins, E. Gustchina, L. Suvorov, J.W. Erickson, Dissection of the pH dependence of inhibitor binding energetics for an aspartic protease: direct measurement of the protonation states of the catalytic aspartic acid residues, *Biochemistry* 36 (1997) 16166–16172.
- [43] N. Cremades, A. Velazquez-Campoy, E. Freire, J. Sancho, The flavodoxin from *Helicobacter pylori*: structural determinants of thermostability and FMN cofactor binding, *Biochemistry* 47 (2008) 627–639.
- [44] J.L. Neira, B. Rizzuti, A. Jimenez-Alesanco, O. Abian, A. Velazquez-Campoy, J.L. Iovanna, The paralogue of the intrinsically disordered nuclear protein 1 has a nuclear localization sequence that binds to human importin $\alpha 3$, *Int J Mol Sci* 21 (2020) 7428.
- [45] Y.J. Bollen, A.H. Westphal, S. Lindhoud, W.J. van Berkel, C.P. van Mierlo, Distant residues mediate picomolar binding affinity of a protein cofactor, *Nat. Commun.* 3 (2012) 1010.