

# A Comprehensive *In Silico* Perspective For Discovery of Novel Inhibitory Candidates Targeting Versatile Transcriptional Repressor MBD2

Zihni Onur Çalışkaner (✉ [zihnionur@gmail.com](mailto:zihnionur@gmail.com))

Üsküdar University

---

## Research Article

**Keywords:** DNA methylation, epigenetics, cancer, iPSC reprogramming, MD simulation

**Posted Date:** December 27th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-578382/v2>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

Genome methylation is a key epigenetic mechanism in various biological events such as development, cellular differentiation, cancer progression, aging, and iPSC reprogramming. Crosstalk between DNA methylation and regulation in gene expression is employed through MBD2, known as reader of DNA methylation and suggested as a drug target. Despite its magnitude of significance and rationale of nomination, a scarcely limited number of druggable ligands has been detected so far. Hence, we screened a comprehensive compound library, and then certain of them were subjected to computational docking analysis by targeting the methylated DNA-binding domain of human MBD2. We could detect reasonable binding energies and docking residues presumably located in druggable pockets. Docking results were also validated via MD simulation and per-residue energy decomposition calculation. Drug-likeness of tested ligands was assessed through ADMET prediction in order to foresee off-target side effects for future studies. Herein, on the basis of collaborating approaches such as molecular docking, MD simulation, energy decomposition, and ADMET prediction, notably two compounds named CID3100583 and 8,8-Ethylenebistheophylline, have become prominent as novel candidates, possibly disrupting MBD2<sub>MBD</sub>-DNA interaction. Hereby, these compounds exhibit a promising usage potential in a wide range of implementations from cancer treatment to somatic cell reprogramming protocols.

## Introduction

Epigenetic regulation is a key process in which involves various mechanisms, dynamically altering and specifying gene expression and cell fate (Aloia, 2021). As a predominant epigenetic mechanism, DNA methylation involves covalent attachment of methyl groups ( $-CH_3$ ) from S-adenosyl-L-methionine as the methyl donor to the carbon-5 positions (C5) of the cytosine residues within a context of cytosine-guanine (CpG) dinucleotides resulting in 5-methylcytosine (5-mC) analogue (Moore et al., 2013). Principally, three catalytically active DNA Methyltransferases in mammals, DNMT1, DNMT3A, and DNMT3B, maintain an overall DNA methylation pattern specific to either cell type, developmental period, and differentiation status (Gowher & Jeltsch, 2019; Jurkowska & Jeltsch, 2016). Methylation on CpG dinucleotides near and/or inside promoter sequences reveals a well-known hallmark for DNA-binding proteins, and thus gene transcription is repressed due to avoidance of transcription factors and machinery (Kribelbauer et al., 2020; H. Zhu et al., 2016). Besides its importance in normal development and cellular function (Vogt, 2017), methylation status of related genes also constitutes a significant diagnostic marker, causation, and therapeutic target in a variety of human diseases such as cancer (Locke et al., 2019), birth defects and syndromes (Dean et al., 2005), metabolic disorders (Samblas et al., 2019), autoimmune diseases (Brandt et al., 2019), muscular dystrophies (Rugowska et al., 2021), cardiovascular diseases (Zhong et al., 2016), aging (Bell et al., 2019), neurological dysfunctions (Stricker & Götz, 2018).

Correspondingly, miscellaneous epigenetic/genetic arrangements, mostly deacetylation of histones along related regions by Histone Deacetylases (HDACs), are accompanied and coordinated with DNA methylation in order to furtherly cement electrostatic interactions between histone core and DNA strands, resulting in compacted chromatin, which downregulates transcription (Cedar & Bergman, 2009; Irvine et

al., 2002). Herein, Methyl-CpG Binding Domain (MBD) family members, also known as “readers of DNA methylation” (H. Zhu et al., 2016), are outstanding transcriptional repressors, which facilitate crosstalk between DNA methylation and histone deacetylation (Baubec et al., 2013). All members of MBD-containing protein family in mammals share a highly conserved methylated DNA-binding domain, consists of 70-85 amino acids (Stirzaker et al., 2017; H. Zhu et al., 2016).

As one of the family members, human *MBD2* gene encodes three isoforms, namely MBD2a, MBD2b, and MBD2c (or MBD2t) with 411 amino acids (~43 kDa), 262 amino acids (~29 kDa), and 301 amino acids (~32 kDa), respectively (Hendrich & Bird, 1998; Jin et al., 2005; Ng et al., 1999; Wood et al., 2016). A full-length isoform, MBD2a, is the most predominant variant that expressed ubiquitously in several tissues while expression of MBD2c, the second prevalent isoform, is especially restricted to testes, embryonic stem cells (ESCs), and human pluripotent stem cells (PSCs), suggesting a functional discrepancy in accordance with significant structural (domain) and expressional inequality (Baubec et al., 2013; Du et al., 2015; Wood et al., 2016). An approximately 80-amino acid Methyl-CpG Binding Domain (MBD) corresponding to residues between 145-225, exists in all isoforms (Buchmuller et al., 2020; Stirzaker et al., 2017; Wood et al., 2016). MBD is responsible to selectively bind methylated promoters via recognition of mCGG, mCAG, mCAT, mCC, mCT di-/tri-nucleotide patterns (K. Liu et al., 2018).

MBD2-mediated epigenetic regulation has been implicated in important biological functions such as neuronal activity and development, proliferation and maturation of various immune cells, hematopoiesis, pluripotency, and cancer initiation and progression (Ginder & Williams, 2018; Menafra & Stunnenberg, 2014; Wood & Zhou, 2016). For instance, it has been evidenced that MBD2 augments metastatic potential, cell proliferation, and growth by especially downregulating hypermethylated tumor suppressor genes in various tumor types (M. Y. Kim et al., 2019; Li et al., 2020; Martin et al., 2008; Mian et al., 2011; D. Zhu et al., 2011). Hereby, KCC-07, a brain-permeable small molecule inhibitor of the MBD2, was clearly demonstrated to inhibit medulloblastoma (MB) cell growth in vitro and reduce tumor volume in xenograft models through the reactivation of BAI1/p53 pathway (D. Zhu et al., 2018). Additionally, MBD2 depletion together with DNMT inhibitor (5-azaCdR) treatment in breast cancer cells resulted in synergistic anti-proliferative and anti-invasiveness effects through recovery of hypermethylation-mediated repression of apoptotic genes (Cheishvili et al., 2014). In other respects, MBD2 was shown to facilitate epigenetic suppression of pluripotency and self-renewal related genes such as Nanog, Oct3/4, and c-myc during the commitment of ESCs, consistent with its higher expression level in terminally differentiated cells. Besides, MBD2 is primarily downregulated by essential miRNA clusters at the early stage of somatic cell reprogramming and its depletion through specific siRNA increases the efficiency of iPSC production. Therefore, MBD2 stands out as a potential target in especially treatment of cancer (Ginder & Williams, 2018), and regulation of pluripotency-differentiation balance in iPSC reprogramming (Lu et al., 2014).

Although the structure of MBD2 and its function in several biological processes and disease progression have been well-characterized suggesting as key targets, the number of available inhibitors against that epigenetic regulator is still barely limited and therefore submission of novel potential ones maintains its importance yet. To date, NF449, Mitoxantrone, Idarubicin, Aurintricarboxylic acid have been proposed as

preventing MBD2 from methylated DNA binding (Wyhs et al., 2014), and KCC-07 (commercially available) and KCC-08 have been biologically tested *in vitro* and/or *in vivo* (Giovinazzo et al., 2018; D. Zhu et al., 2018). Hence, we aimed to evaluate novel inhibitor candidates targeting MBD2<sub>MBD</sub>-DNA interaction in order to reverse transcriptional silencing on crucial genes such as tumor suppressors, pluripotency factors, and etc. for multiple purposes. In this regard, we accomplished a comprehensive screening to reveal small molecules as potential inhibitors in this study. Our results have pointed these small molecules prominently docked significant residues such as Val177, Arg188, Lys190, Phe206, Phe208, and Arg209 in DNA-binding domain, predicting an interrupting impact on MBD2-methylated DNA recognition and interaction. Furthermore, Molecular Dynamics (MD) simulation validated MBD2<sub>MBD</sub>-ligand interaction by meaningful RMSD, RMSF, B-Factor values, and energetics. Besides, it has been noticed that predicted 3-dimensional binding pockets on MBD2<sub>MBD</sub> and residues in which ligands occupied, were considerably overlapped.

Taken together, extracted small molecules have emerged as potential inhibitor/drug candidates with binding affinity and inhibition constants in a range of -7.5-(-)11 kcal/mol and 4-1880 nM, respectively, as shown computationally. Hereby, the current study has provided substantial data, particularly for compounds CID3100583 and 8,8-Ethylenebistheophylline which might lead to *in vitro* and/or *in vivo* studies testing their applicability in cancer treatment and somatic reprogramming protocols.

## Results

### Docking Based Virtual Screening To Elicit Potential Binders

A screening library was prepared from ZINC15 database compounds. We regarded certain criteria such as logP value, MW, and charge for selection of molecule tranches. Wyhs et al. had demonstrated Mitoxantrone, Idarubicin, NF449, and Aurintricarboxylic Acid impaired MBD2<sub>MBD</sub> binding to methylated DNA in TR-FRET assay however mechanism of action remained unclear (Wyhs et al., 2014). Since these compounds predominantly have charges, logP, and MW ranging from -3 to +2, 0 to 3, 400 to 500 Da, respectively, we exploited related characteristics, emerging as a screening strategy to start out. Besides, logP, MW and charge were retained moderately in order to explore ligands with favorable solubility in aqueous vehicles, convenience for passing through the plasma membrane and intracellular delivery for potential use in prospective *in vitro*/*in vivo* studies. 15,893 different compounds yielded numerous docking poses including 120,310 available conformations with reasonable binding affinities. Among these, certain hit compounds with docking score over -5.0 kcal/mol which was pointed according to structure and/or biological activity (if available), were listed in Supplementary Table S1. Although most of the compounds have not been annotated, assayed and/or functionally identified, already identified/annotated hits were also recovered by the virtue of screening. Most of the hits have been attained to inhold heterocyclic groups, especially carboxamide, quinoxaline, imide, benzene, thiazole, piperazine, naphthalene, pyridazine, isoquinoline, pyrrole, pyrimidine/pyridine, furan, aniline, imidazole (see Supplementary Fig. S1). Moreover, ligands composed of xanthine, guanine, adenine or their

derivatives and/or methylated groups, were also quite notable in terms of implying a plausible compatibility between potential binders and methylated DNA-binding domain.

### Elaboration of Target Sites of Selected Leading Compounds Throughout MBD2<sub>MBD</sub>

The existence of significant structures and groups within hit compounds (see Supplementary Fig. S1) led us to determine binding regions more comprehensively by computational docking analysis. Thereby, selected ligands with higher screening scores were docked against refined receptor corresponding to human MBD2<sub>MBD</sub> (PDB: 7A08, Chain C). Among 10 of unidentified and 5 of annotated compounds, CID3100583 exhibited the best docking pose with -11.42 kcal/mol, 275.20 Å, and 4.22 nM as binding energy, RMSD, and estimated inhibition constant ( $K_i$ ), respectively (Table 1). It was followed by CID343482, CID136748749 as proximately. Nevertheless, estimated  $K_i$  of CID3100583 (4.22 nM) was ahead of other candidates as roundly 8-fold (Table 1), suggesting less amount could be sufficient for a half-maximal inhibition on MBD2<sub>MBD</sub>-DNA binding in actual fact. Correspondingly, each of CID3100583 and CID343482 constituted 9 H-bonds (Table 2) in company with 2 hydrophobic interactions and 3 hydrophobic interactions, respectively (Table 3). Albeit both ligands established the same number of H-bonds, more diverse amino acids contributed interaction with CID3100583 atoms.

8,8'-Ethylenebistheophylline (Ebis-Theophylline) committed the strongest binding energy (-10.32 kcal/mol) with the lowest  $K_i$  (27.30 nM) and the second rank among annotated drugs and all compounds, respectively (Table 1). Although Regadenoson and Glucosylhydroxymethylcytosine (GHMCytosine) formed 9 H-bonds (Table 2), their binding affinities and estimated  $K_i$  were relatively lower and besides less number of divergent amino acids participated in binding sites in a similar way with others (Table 2 and 3).

Together, *in silico* docking approach established that all subjected compounds occupied overlapping regions throughout MBD2<sub>MBD</sub>, containing mutual residues (see Supplementary Fig. S2). Hereby, residues Asp176, Val177, Arg188, Lys190, Val201, Phe206, Asp207, Phe208, and Arg209 were the main interactional contributors (Fig. 1), commonly encountered within docking sites of ligands. Therefore, we rendered intermolecular interactions specific to MBD2<sub>MBD</sub>-CID3100583 and MBD2<sub>MBD</sub>-Ebis-Theophylline complexes appeared as the most promising ones. As shown in Fig. 2, CID3100583 affiliated not only hydrogen bonds with amino acids Val177, Arg188, Lys190, Ser204, Phe206, Asp207, Phe208, and Arg209 (Table 2) and besides through van der Waals forces (Asp176, Tyr178, Ser189, and Phe208), pi-Alkyl (Val177), pi-Anion (Asp207) interactions (Table 3 and Fig. 2a-c). Furthermore, it's been noticed that interacted residues were aligned around by configuring a cavity-like surface, in which hypothetically facilitates settlement of CID3100583 in an easier and trapped manner (Fig. 2b). On the other part, Ebis-Theophylline built close and distinct interactions with the MBD2<sub>MBD</sub> receptor (Fig. 2d-e). Five H-bonds were bridged between Val177, Phe209 and Arg209 (Fig. 2d and Table 2) while amino acids Arg166, Asp176, Tyr178, Arg188, Ser189, Lys190, Pro191, and Asp207 fell within distinct interactive relations such as van der Waals,  $\pi$ -anion and  $\pi$ -alkyl forces (Fig. 2e and Table 3).

Since ligand docking may trigger conformational changes on protein, we aimed to better reflect the dynamics nature of protein-ligand interaction and local rearrangements of side-chains of active site amino acids via induced-fit docking (IFD) (Antunes et al., 2015; Huang & Zou, 2010) analysis. As a result of IFD (flexible receptor; flexible ligand), similar binding-pocket residues were appointed (Fig. 3c and 3d), consolidating our initial rigid docking results (rigid receptor; flexible ligand) again. We detected various transition states of amino acid side-chains, particularly surrounding CID3100583 and Ebis-Theophylline zones between apo (unbound) receptor and holo (ligand-bound) complex (Fig. 3a and 3b). It was elicited that Arg166, Asp176, Tyr178, Asp207, Phe208 became closer to CID3100583 atoms upon formation of the holo complex while Val177, Arg188 underwent a minor secession (Fig. 3a; see Supplementary Movie S1 for distances). A similar adjustment was also mapped for complex with Ebis-Theophylline according to unbound state (Fig 3c; see Supplementary Movie S2). Additionally, we re-calculated binding energies by using PRODIGY-Ligand interface (Vangone et al., 2019) in order to eliminate differences outputted by different software and to compare whether affinity could establish any alteration in IFD. Binding energies (kcal/mol) were listed as -5.21, -7.52, -5.28, -6.49 for CID3100583-MBD2<sub>MBD</sub> (rigid receptor), CID3100583-MBD2<sub>MBD</sub> (flexible receptor), Ebis-Theophylline-MBD2<sub>MBD</sub> (rigid receptor) and Ebis-Theophylline-MBD2<sub>MBD</sub> (flexible receptor), respectively.

Favorably, amino acids interacted with ligands have been detected highly conserved between MBD1, MBD3, MBD4, and MeCP2 in human (see Supplementary Fig. S3) and MBD2 in different organisms via multiple sequence alignment. Residual counterparts in both MBD1 (Val20, Arg22, Tyr34, Arg44, Ser45) (Ohki et al., 2001) and MeCP2 (Arg106, Arg133, Trp155, and Tyr158) (Ballestar et al., 2000; K. Liu et al., 2018) construct a methylated-DNA binding site as reported, designating a highly conserved DNA binding tendency for MBDs. To test whether ligands exhibited an incidental complementarity with docking site on MBD2<sub>MBD</sub> or compounds were detained due to their DNA-like preferability for functional residues, we docked CID3100583 and Ebis-Theophylline against MeCP2<sub>MBD</sub> and MBD1<sub>MBD</sub>. Among CID3100583 docking residues, Asp121, Val122, Arg133, Thr158 was worthy of note for MeCP2<sub>MBD</sub> that these were previously reported as harboring to the backbone or aromatic bases of DNA duplex (Ghosh et al., 2010; K. Liu et al., 2018; Nikitina et al., 2007) while Arg44, Ser45, Lys46 and Lys 65, which mediates the interaction between MBD1<sub>MBD</sub> and negatively charged DNA backbone (Ohki et al., 2001; Zou et al., 2012) (See Supplementary Fig. S4a-b). Besides, Ebis-Theophylline was also shown to dock the similar and/or adjacent residues on MBD1<sub>MBD</sub> and MeCP2<sub>MBD</sub> (See Supplementary Fig. S4c-d). For instance, Arg133 and Thr158 were appeared as two common residues on MeCP2<sub>MBD</sub> which contributed to both CID3100583 and Ebis-Theophylline docking while Ser134 was shown to involve in van der Waals forces between only Ebis-Theophylline (See Supplementary Fig. S4c). Binding energies (kcal/mol) were calculated as -11.05 ( $K_i = 7.91$  nM), -9.52 ( $K_i = 104.35$  nM), -10.99 ( $K_i = 8.82$  nM), -10.67 ( $K_i = 15.21$  nM) for MBD1<sub>MBD</sub>-CID3100583, MBD1<sub>MBD</sub>-Ebis-Theophylline, MeCP2<sub>MBD</sub>-CID3100583, MeCP2<sub>MBD</sub>-Ebis-Theophylline complexes, respectively.

Because the mechanism of action remained unclear, it led us to survey the docking residues of the most effective inhibitors KCC-07 and CID3136570 patented by Nelson et al. (Nelson et al., 2010), in a similar

manner (see Supplementary Fig. 5). As a reference, binding energies (kcal/mol) were -7.89 (KCC-07) and -7.61 (CID3136570) in our docking analysis, even lower than assumptions for novel inhibitor candidates in Table 1. Surprisingly, we detected Val177, Arg188, Arg209 in H-bonding while Arg166, Tyr178, Lys190, Phe208, Arg209 set hydrophobic interactions with CID3136570. Nonetheless, KCC-07 was shown to interact with the residues between Asp151 and Lys160 (see Supplementary Fig. S5a).

### **Crosscheck of Dynamics and Flexibility of MBD2<sub>MBD</sub>-CID3100583 and MBD2<sub>MBD</sub>-Ebis-Theophylline Complexes via MD Simulation**

Because CID3100583 has emerged as the most promising druggable ligand in consequence of either virtual screening and docking analysis, we facilitated MD simulation in order to pursue dynamic behavior and to prove flexibility/rigidity of protein-ligand complex binding along given time. Along with CID3100583-MBD2<sub>MBD</sub> complex fitted as superior (Table 1), MD simulations of other ligand-protein complexes with lower docking scores were also conducted as a comparison (Fig. 4a). Apparently, MBD2<sub>MBD</sub> conformations with ligands CID3100583, CID343482, CID136748749, CID46959745, CID126001284, CID125615373, CID25826758, CID323153, CID24241 established mean RMSD values 0.89 Å, 0.87 Å, 0.94 Å, 1.04 Å, 1.14 Å, 1.14 Å, 1.27 Å, 0.97 Å, and 0.94 Å, respectively, against the average structure along the trajectory. Likewise, CID3100583-protein complex fluctuated slightly through snapshots and followed by CID343482- and surprisingly by CID24241-protein complexes. Despite these complexes were fluctuating proximately, CID3100583 could be distinguished owing to an earlier equilibrium state and a lesser chequered progress during simulation. The RMSD fluctuation suggested that the MD trajectories except for CID12561537, have been overall stable for ligand-protein structures during simulation (Fig. 4a). It was noted that both protein backbone and CID3100583 ligand in the complex have been reached equilibrium in a shorter time (in approximately 200 frames) and remained stable except for negligible wobbles in a wider time span (Fig. 4b). Afterward, RMSF and B-factors were interpreted in order to detect atomic fluctuations and x-ray scattering caused by thermal motion in specific to per residue for C<sub>α</sub> atoms. Importantly, we apparently detected sharp declines in fluctuation especially over three regions built by Val177, Arg188, Lys190, Ser204, Phe206, Asp207, Phe208, Arg209 and their adjacencies (Fig. 4c, orange line), demonstrating residues have rigidly docked to CID3100583 atoms and therefore been constrained less flexible. Correspondingly, the regions in which B-factor was lower, account for regular secondary structures (Fig. 4c, blue line). For instance, Val177 and Arg188 accommodate in strands and Lys190 is included in helix while residues Ser204-Arg209 locates in a loop (see Supplementary Fig. S3). In other respects, RoG has been moderately altering over time in parallel with the RMSD chart. As plotted in Fig. 4d (blue line) during the initial 200 frames, here detected an unsteadiness in which also corresponded to the protein-ligand equilibrium phase (Fig. 4b). However beyond this point, RoG decreased upon probably binding with CID3100583, in concordance with the equilibrium of the system (Fig. 4b and 4d). Fraction of native contacts (Q<sub>x</sub>) increases based upon more binding residues by determining transition in folding state, that is, the more the folding (lesser RoG), the higher the Q(x) is assumed (Best et al., 2013; Brylinski & Skolnick, 2008). Notably, Q(x) has been conducted verifying RoG data (Fig. 4d).

Despite the fact that Ebis-Theophylline, as the most probable repositioning drug (Table 1), showed the closest RMSD profile among the annotated drugs, Ebis-Theophylline-MBD2<sub>MBD</sub> system could catch the equilibrium a bit later and also fluctuated much more during simulation trajectory compared to CID3100583-MBD2<sub>MBD</sub> system (Fig. 5a). Average RMSD along MD trajectory was 0.89 Å, 0.90 Å, 0.98 Å, 0.94 Å, and 1.05 Å for CID3100583, Ebis-Theophylline, Golgicide A Sanguinarine, and GHMCytosine respectively. It's been noticed that Golgicide A-MBD2<sub>MBD</sub> complex trajected more stably although it was the fifth rank in the sense of docking (Table 1) and reached steady state at the latest, flashing as another alternative repurposing drug candidate (Fig. 5a). In more detail, both protein backbone and Ebis-Theophylline in the complex could get the equilibrium by the 450 frames but later on, they were seemed to considerably fit in parallel until the end of time course (Fig. 5b). Similarly, atomic fluctuations and B-factor was diminished in line with interacted residues such as Asp176, Val177, Arg188, Lys190, Asp207, Phe208, Arg209 (Fig. 5c). In case of RoG and Q(x), it could be estimated that there has been a transition in folding state at around 300 frames and continued with almost none of remarkable alteration in contrast to CID3100583 system, implying Ebis-Theophylline could detain receptor in more stable as soon as it has interacted with MBD2<sub>MBD</sub> (Fig. 5d). In conclusion, dynamic behaviors and flexibilities of crystal structure of both MBD2<sub>MBD</sub>-CID3100583 and MBD2<sub>MBD</sub>-Ebis-Theophylline complexes during MD trajectory have been consistent, proving strict interactions constructed in computational docking approach.

### Calculation of MM-PB(GB)SA Binding Free Energy and E<sub>Dec</sub> Per Residue

After MD simulation, we employed MM-PBSA calculation and per-residue energy decomposition, aiming to elucidate putative binding free energy ( $\Delta G_{\text{Bind}}$ ) between MBD2<sub>MBD</sub> and CID3100583 or Ebis-Theophylline, and the energetic contribution of each amino acid involved in ligand docking, respectively. As expected, compound-ligand complexes with lower docking scores and unfine fluctuations established relatively lower binding free energy. As expected, 8,8-Ethylenebistheophylline had the strongest binding affinity (-15.03 kcal/mol) among annotated library and second rank within all ligands. However, CID3100583 was ranked as the third (-12.26 kcal/mol) in the list, and interestingly CID126001284 (-18.23 kcal/mol) and CID136748749 (-13.88 kcal/mol) demonstrated the strongest binding affinity (see Supplementary Table S2). In MBD2<sub>MBD</sub>-CID3100583 trajectory, van der Waals interactions (-34.49 kcal/mol) and polar solvation energy (-26.10 kcal/mol) mainly contributed to ligand binding while non-polar solvation energy (12.26 kcal/mol) conduced to unfavorable contributions (see Supplementary Table S2). From this point of view, we concluded van der Waals forces and polar interactions should have an indispensable role in CID3100583 docking on MBD2<sub>MBD</sub>.

We next examined the contribution per residue to binding free energy. Amino acids with energy decompose (E<sub>Dec</sub>) contribution ranked in top ten, were Asp207, Arg209, Phe208, Val177, Phe206, Lys190, Asp176, Ser205, Tyr178, Thr210 in the protein-ligand system (Fig. 6a, top panel). It should be noticed that top energetic contributors to binding free energy have consistently been amino acids that were diversely interacted with CID3100583 within the binding pocket (Fig. 2c). As illustrated in Fig. 6a (bottom panels),

the other residues except the top ten contributors had a less impact on binding free energy but certain neighboring residues around top ten could faintly contribute, as well (Fig. 6a, bottom panels). Herein, van der Waals forces have been detected as the main  $E_{Dec}$  contribution however electrostatic energetics have constitutively been provided by Val177, Arg188, Ser204, Asp207 and Phe208 (Fig. 6a, bottom). In other respects, Ebis-Theophylline showed the top van der Waals (-37.55 kcal/mol) and polar solvation (-28.21 kcal/mol) contributions among all (see Supplementary Table S2). Accordingly, Arg188, Arg209, Phe208, Tyr178, Asp176, Val177, Val164, Asp207, Lys186, Lys190 were orderly top ten energetic contributors resulted from MD simulation (Fig. 6b, top panel), demonstrating a harmony with the amino acids achieved in docking analysis (Fig. 2d–e). As plotted in Fig. 6b (bottom) Asp176, Val177, Tyr178, Arg188, Phe208, and Arg209 have been essential for van der Waals contribution and also principle contributors to overall binding free energy. Though Val177, Tyr178, and Lys186 provided electrostatic forces (Fig. 6b), such forces have been accounted weakly in total energy decompose (see Supplementary Table S2). Thus, these results demonstrated docking might have been enabled by favor of such energetics contributed by especially amino acids enclosing CID3100583 or 8,8-Ethylenebistheophylline binding pockets.

## Discussion

Since MBD2 is a miscellaneous participant in various biological phenomena and disorders, that links epigenetic methylation marks and gene expression regulation, it has arisen as a potent target for multi-purpose. In the current study, we aimed to inspect novel drug candidates targeting  $MBD2_{MBD}$  through docking-based virtual screening and drug repurposing approaches. In this respect, we generated a comprehensive library including over 15,000 non-annotated compounds and biologically identified drugs, from ZINC15 database. Certain ligands have been listed in Supplementary Table S1 as hitter ligands by filtering according to resulted binding affinity and physicochemical properties such as complexity, molecular weight, tPSA, and also convenience to be purchasable or synthesizable for *de facto* attempts. Here, we have realized that numerous hit compounds (see Supplementary Table S1) and also unlisted ones consist of remarkable groups such as pyridine, quinoxaline, imidazole, isoquinoline, which are incorporated either in nucleosides and nucleotide analogs, or at least related structures to directly nucleotides or DNA-related constructs (Ali et al., 2008; García et al., 2008; Röthlisberger et al., 2017). For instance, quinoxaline and isoquinoline derivatives have been demonstrated to bind and act as antagonists of cyclic nucleotide-related enzymes such as cyclic nucleotide phosphodiesterase (Parra et al., 2001) and P2X7 nucleotide receptor (Humphreys et al., 1998; Watano et al., 2002), suggesting these groups might be attractive for DNA- or nucleotide-recognizing domains. Interestingly, NF449, an inhibitor for P2X1 and P2X2 receptors, was shown to inhibit  $MBD2_{MBD}$  interaction with methylated DNA, suggesting that binding with negative DNA backbone could be disrupted due to interaction between negatively charged sulfonic acids on NF449 and positively charged basic residues on the MBD2 surface (Wyhs et al., 2014). Likewise, we also found that ring structures in our ligands such as benzene, thiazole, quinone, pyridine, have already been built in NF449, Mitoxantrone, Idarubicin, Aurintricarboxylic Acid, KCC-07 (Wyhs et al., 2014; D. Zhu et al., 2018). Furthermore, some of screened chemicals were directly purine- or pyrimidine-containing molecules (ex. CID3100583, CID136707272, CID135623722,

CID124903520, CID136732748, TB6:RiboPyrimidoP) or straightly nucleotide/nucleoside derivatives (ex. Regadenoson, Wyosine, 8-Methylamino-Guanosine, Thiosangivamycin, Adenosine 5'-monophosphoramidate, 5,6-Trimethylenuridine, Fludarabine Base, 8-Aminoguanosine, Tubercidin, Hypoxanthosine, GHMcytosine, 6-Mercapto-7-Methylguanosine, 8-Chloroguanosine, Isovaleriansaures Coffein, 2-Methylformycin, 5'-N-Methylcarboxamidoadenosine, etc.), documented in Supplementary Table S1. Additionally, these compounds highly contain methyl groups. Indeed, it has been quite interesting that a considerable number of particularly adenosine and guanosine derivatives have become prominent as a consequence of drug screening. It was reported that residues in MBDs of both MeCP2 and MBD2 exhibited binding affinity to mCAT, mCAC, mCAT, mCGG patterns, and surprisingly unmethylated TG-dinucleotides (Fraga et al., 2003; K. Liu et al., 2018). Thereby, such interplay preferences of MBD might reflect why these compounds were recorded as hit molecules amongst the whole library. Then top ten of non-annotated and top five annotated molecules were proceeded for further docking analysis. Principally, CID3100583 (ZINC3109386) and 8,8-Ethylenebistheophylline (ZINC8612354) urged us for the exhaustive investigation due to the highest binding energy and the lowest  $K_i$  (Table 1). At this point, chemical contexts of both substances have been intriguing since CID3100583 contains two pyrimidine-based rings, and Ebis-Theophylline is composed of two interconnected druggable Theophylline (dimethylxanthine) molecules, a methylated xanthine analog which antagonizes human A1, A2a, A2b Adenosine Receptors as in the treatment of airways diseases and vasodilation (Cheng et al., 2017; Fieger & Wong, 2010; Landells et al., 2000). Together, similarities between the actual recognition target of MBD2<sub>MBD</sub> (various nucleotide patterns above) and top hitters in the context of chemical and structural features have been speculated as possible reasons underlying such complementarity and fine docking.

Actually, all ligands have been demonstrated interacting with similar residues (Fig. 1). Particularly, hydrogen bonds, van der Waals forces, pi-Alkyl, and pi-Anion interactions between Val177, Arg188, Ser189, Lys190, Phe208, and Arg209 were the most prominent common residues, signifying both ligands have been stabilized by cementing highly reactive surroundings. It was previously reviewed that arginine, lysine, serine, threonine was orderly top interacting amino acids in protein-DNA complexes and arginine was featured prominently through hydrogen bonding/cation- $\pi$  interactions with pyrimidine-guanine dinucleotides in DNA backbone (Luscombe et al., 2001; Zou et al., 2012). As shown, such kinds of amino acids were also favorably frequent in our analysis. In this respect, Liu et al. also reported MBD2 bound to DNA via constitution of H-bonding between Arg166, Asp176, Tyr178, Arg188 within MBD2<sub>MBD</sub> and methylated-Cytosine (mC) along with adjacent Thymine or Guanine bases in the context of mCGG, mCAG, mCAT, mCC, mCT di-/tri-nucleotide patterns (K. Liu et al., 2018). Along with pi-Anion interactions between referred amino acids and aromatic bases above, alternate residues such as Lys167, Ser168, Ser171, Lys186, Ser189, Lys190, Arg209, interacted with the sugar-phosphate backbone of DNA. Buchmuller et al. also attributed regions 164-177 and 186-195 in the vicinity to DNA duplex (Buchmuller et al., 2020). In other studies, Lys174 and Tyr178 have been reported as critical amino acids to selectively recognize and bind methylated CpG probes (Cramer et al., 2014; Fraga et al., 2003). Here, we demonstrated our compounds firmly settled in this functional mCpG-binding/recognition sphere.

As shown in Supplementary Fig. S4, CID3100583 and Ebis-Theophylline also fitted into MeCP2<sub>MBD</sub> and MBD1<sub>MBD</sub>, supporting these inhibitor candidates have potentially mimicked related nucleotide patterns and favorably invaded conserved methylated DNA recognition residues on MBDs. Importantly, Arg188 (MeCP2; R133, MBD1; R44), Ser189 (MeCP2; S134, MBD1; S45) and Phe208 (MeCP2; F157, MBD1; F64) were the most outstanding amino acids between family members, participated in H-bonding or van der Waals interactions between CID3100583, Ebis-Theophylline. However, we detected slight declines in calculated binding energies of MeCP2-/MBD1-ligand complexes, suggesting affinity and inhibition potential of both compounds have remained more specific to MBD2<sub>MBD</sub> *in silico*. These data collectively indicate that docking territories of the compounds were not indiscriminate due to chemical and structural determinants in inhibitor candidates shared between DNA. As another comparison, CID3136570 and CID3100583/Ebis-Theophylline could target the similar subfield within MBD2<sub>MBD</sub> however KCC-07-binding residues were more comparable with other compounds such as CID46959745, CID126001284, CID125615373 (see Supplementary Fig. S5). These overlapping regions between previously patented inhibitors and CID3100583/Ebis-Theophylline compatibly appreciate the value and expected potential of our outcome in a prospective real assay.

Moreover, we have sought out presumptive druggable pockets on the surface or interior of MBD2<sub>MBD</sub> in order to conclude whether ligand-docking sites represented any significance in this sense (see Supplementary Information). *In silico* prediction yielded in a convergence between putative druggable pockets and ligand-anchored pockets (see Supplementary Fig. S6), proposing these ligands have also opted for ideal physicochemical districts on the protein as possible.

Ohki et al. defined a small portion between positions Arg22-Arg30 in human MBD1 (conservatively corresponds to Arg166-Lys174 in human MBD2) as a flexible loop (L1), demonstrating major conformational rearrangements upon DNA binding. Another loop (L2) between residues 44-48 (Lys186-190 in MBD2) was inserted into major groove although the rest of DNA-bound MBD had a stationary architecture in MBD1 and MeCP2 according to unbound state, clearly indicating an induced fit model *in vitro* (Ohki et al., 2001). Similar loops with conserved residues and a congener DNA-bound orientation were well-structured for chicken MBD2<sub>MBD</sub> by Scarsdale and colleagues, as well (Scarsdale et al., 2011). Since we utilized an IFD to further insight to possible structural shifts in especially active binding site in MBD2<sub>MBD</sub>. Consistently, we depicted moderate structural changes in side-chains of Arg166, Ser171, Lys174, Arg188, Ser189, Lys190, all compatible with L1 and L2 regions, and also in their adjacent residues such as Asp176, Val177, Tyr178 upon CID3100583 and Ebis-Theophylline docking (Fig. 3a and 3b, also see Supplementary Movies S1-S2). Hereby, we putatively deduce that CID3100583 and Ebis-Theophylline are not only in contact with crucial amino acids for mCpG recognition, but also may trigger conformational changes in the active site in order to fit better likewise mimicking DNA attachment in fact. Briefly, these data satisfactorily highlight the fact that ligands achieved in our study, especially CID3100583 and 8,8-Ethylenebistheophylline, could actually interfere methylated DNA-binding function of MBD2 by selectively invading of indispensable residues (see Supplementary Fig. S7).

Next, we simulated each ligand–protein complex to test its structural dynamics and stability in a time course. As plotted in Fig. 4a and 5a, CID3100583, CID343482, and Ebis-Theophylline established more stable dimensionality effects or conformational distributions over time, corresponding docking affinities (Table 1). Herein, both ligands (CID3100583 and Ebis-Theophylline) along with protein backbone, showed amenable fluctuations and equilibration timing concurrently with MBD2<sub>MBD</sub> receptor (Fig. 4b and 5b). Consistently, amino acids participated in particularly H-bonding with both CID3100583 and Ebis-Theophylline, such as Val177, Arg188–Lys190, Ser204, Phe206–Arg209, remained non-fluctuating as revealed by RMSF and B-Factor graphics (Fig. 4c and 5c). Actually, regions between 160–166, 175–780, 186–188, coincide with strand, and residues 190–198 corresponds to a helix (K. Liu et al., 2018) (also see Supplementary Fig. S3). As we expected, diminished B-factor along these districts were corroborated with the consent that increased fluctuations and thus higher B-factor is predicted for irregular loop regions (Craveur et al., 2015). More particularly, secondary structures including Val177, Arg188, Lys190 displayed more regularity than part of 160–166, suggesting durability contributed by ligand binding. Surprisingly, even though amino acids between Ser204 and Arg209 have been involved within an irregular loop region (see Supplementary Fig. S3), they possessed lower B-factor values (Fig. 4c and 5c). The magnitude of that reduction may presumably be due to immobilization of the loop region by extensive interactions with either CID3100583 or Ebis-Theophylline over there. As a matter of fact that interactions with CID3100583 were potently concentrated on residues between Ser204 and Arg209, hence ensuring a further decline even rather than strands or helices and corresponding zone in MBD2<sub>MBD</sub>–Ebis-Theophylline complex (Fig. 4c and 5c). Of note, RoG has moderately diminished at around frames 250 and 400 for CID3100583 and Ebis-Theophylline (Fig. 4d and 5d), respectively, postulating a compaction in both protein–ligand systems (Best et al., 2013; Brylinski & Skolnick, 2008). Within this context, spectroscopic analysis revealed that upon methylated CpG binding, dynamic mobility and conformation of the complex have been shifted under the favor of DNA bending by MBD2<sub>MBD</sub> (Pan et al., 2017). Similarly, Liu et al. reported compaction in methylated DNA–bound MeCP2<sub>MBD</sub> detected by atomic force microscopy in parallel with a decrease in RoG correlation (M. Liu et al., 2020). Furtherly, it was experimentally figured out DNA-contacting regions including Arg46 (Arg188 in human), Ser47 (Ser189 in human) Arg67 (Arg209 in human) in chicken MBD2<sub>MBD</sub> was well-structured (Scarsdale et al., 2011), which is consistent with the non-fluctuated residues detected in our MD simulation (Fig. 4c and 5c). This may point to the hypothesis that reduction in RoG and a parallel increment in Q(x) could be resulted from further compaction and folding due to a configuring action of docked CID3100583 or Ebis-Theophylline by mimicking an ordinary MBD–DNA binding. In correlation, acquired energetics data also corroborated our docking results and conclusions so far. Both CID3100583 and Ebis-Theophylline have been ranked amongst the top ligands with the highest binding free energy (see Supplementary Table S2). Energy decomposition for CID3100583 showed Asp207, Arg209, Phe208, and Phe206 appeared in the top five in terms of contribution to overall  $\Delta G_{\text{Bind}}$  (Fig. 6a), supporting rigidity in the loop region discussed above. Besides, per residue energy decomposition (Fig. 6) and interactions in docking analysis (Fig. 2) established a fine correspondence for both ligands. For instance, Phe208, one of the top van der Waals energy–decomposing residues for CID3100583 was also constituted in van der Waals interaction as a

result of docking analysis while Val177 incorporated with Ebis-Theophylline by electrostatic forces via both approaches (Fig. 6b).

Consequently, the presented computational study has established prosperous and crosschecking evidence for certain ligands in a comprehensive manner. CID3100583 and 8,8-Ethylenebistheophylline have successfully been invented as novel and most likely MBD2 inhibitors in the light of effective binding affinity, and negligible off-target probability predicted by ADMET analysis (Xiong et al., 2021) (see Supplementary Table S3). Though *in silico* approach was utilized within the scope of this paper, we have provided a consolidated basis for the next experimental studies testing their actual druggability. In conclusion, we rationally propose that especially CID3100583 and 8,8-Ethylenebistheophylline are highly versatile with the capability to be notably used as anti-cancer drugs and agents in iPSC reprogramming cocktails by impeding MBD2 function in related pathways.

## Methods

### Data Sources and Acquisition of 3D Protein Model and Ligand Structure Files

3D model file of DNA-binding domain of human MBD2 (MBD2<sub>MBD</sub>) was obtained from RCSB PDB Protein Databank (<https://www.rcsb.org/>). The experimentally resolved structure of Chromatin-binding domains in the NuRD deacetylase complex (PDB: 7A08) (Millard et al., 2020) was downloaded. Residues between 149–212 (Chain C) corresponding to MBD2<sub>MBD</sub> receptor was extracted from the whole assembly by removal of other chains in AutoDock tools.

Total 1293 tranches in ZINC15 database (Sterling & Irwin, 2015) were filtered by molecule charge (between 0 and +2), molecular weight (MW, between 200 and 400 Da), and logP value (between -1 and +4.5) and thereby over 15,000 ligands could be extracted in .sdf format. Afterwards, ligand structure files were converted into .pdbqt file format via Open Babel toolbox (O'Boyle et al., 2011) for the further docking analysis.

### High-throughput Virtual Screening

Structure-based virtual screening was accomplished by Vina wizard in PyRx program (Dallakyan & Olson, 2015) to discover possible binders for the related domain of MBD2 after ligand libraries including 15,893 compounds and MBD2<sub>MBD</sub> (7A08; C149–C212) was pre-processed. Grid box was adjusted with centering coordinates 121.02, 178.69, 185.85 and sizes with 38.57 Å x 35.46 Å x 31.60 Å for X-, Y-, Z- dimensions, respectively, in order to involve the entire structure. Before starting screening, all compounds in the library were energy minimized. Results were sorted in descending according to affinity and the top 10 compounds with the best scores were evaluated for *in silico* docking approach.

### *In Silico* Molecular Docking

Molecular docking analysis was performed by AutoDock 4.2 (version 1.5.6), a widely used automated program for the prediction of molecular interactions between protein and ligand structure, based on flexible-ligand/rigid-receptor model (Morris et al., 2009). Prior to docking run, all protein models (PDB: 7A08, Chain C for MBD2<sub>MBD</sub>; 6OGJ for MeCP2<sub>MBD</sub>; 6D1T for MBD1<sub>MBD</sub>) and each ligand were structured and prepared using the mentioned software. Accordingly, removal of any other existed ligands/ions/heteroatoms and water molecules to obtain a crude protein structure, addition of polar hydrogens and missing atoms, calculation of Kollman charges, assignment of search parameter (Genetic Algorithm) and output format (Lamarckian GA) were all implemented as previously reported (Çalışkaner, 2021). As distinct from, we have played 150 runs (conformations) with a population size of 300 as rigid docking for each ligand-receptor complex. Besides, grid box was adjusted in the same coordinates and sizes as used in virtual screening with 0.375 Å spacing for blind docking since any possible binding residues in Methyl-DNA Binding Domain (MBD) of human MBD2 (MBD2<sub>MBD</sub>) were not missed out. The first ranked conformation with the best binding energy and inhibition constant (Pantsar & Poso, 2018) was proceeded for further examination.

Induced-Fit-Docking (IFD) was performed by using AutoDockFR, a new docking engine based on the AutoDock4 scoring function (Ravindranath et al., 2015). Protein and each ligand were prepared according to the program instructions and stored with .pdbqt extension. All ligand interacting amino acids from prior AutoDock docking (rigid receptor; flexible ligand) were selected as active site residues. After affinity maps were generated for each atom type, IFD (flexible-ligand; flexible-receptor) was run by scripts presented by the developer.

### **Mapping Molecular Interactions and Distances**

Intermolecular interactions such as hydrogen bonds (H-bonds), hydrophobic interactions, salt bridges, halogen bonds, etc. in the best docked ligand-protein pose with calculated distances, were extracted and visually rendered using PLIP-Protein Ligand Interaction Profiler tool (Salentin et al., 2015) and PyMOL Molecular Graphics System, Version 2.4.1 Schrödinger, LLC. Additionally, 2D interaction diagrams were illustrated by BIOVIA Discovery Studio Visualizer, v17.2.0.16349, Dassault Systèmes.

### **Molecular Dynamics (MD) Simulation**

Dynamic behaviors of the docked protein-ligand complexes were computed through VMD ver1.9.4a51 software (Humphrey et al., 1996), a parallel molecular dynamics code NAMD (Phillips et al., 2020), and CHARMM22 force field (MacKerell et al., 1998). The parameters and protein topology files were generated automatic PSF builder inside VMD while ligand structures were utilized using CHARMM-GUI Ligand Reader and Modeler tool (S. Kim et al., 2017). The systems were solvated with an octahedron box of TIP3P water environment (Price & Brooks, 2004) extended 10 Å in each direction and Na<sup>+</sup>/Cl<sup>-</sup> counterions were added to neutralize the system. Upon energy minimization in 3000 steps, the whole system was heated to 300K, in 30 ps, at 1 atm pressure under periodic boundary conditions to avoid edge effects. Finally, a production run of MD simulation was carried out for a 20-ns duration. Flexibility and dynamics

of each trajectory were evaluated by the query of Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), Radius of Gyration (RoG), B-factor (Debye–Waller factor), fraction of native contacts (Qx), and estimated MM-PB(GB)SA binding free energy ( $\Delta E_{PB}/\Delta E_{GB}$ ). Resulted values were plotted and rendered by GraphPad Prism 6 graphic and statistics software.

### Calculation of MM-PB(GB)SA Binding Free Energy and Energy Decomposition

The Poisson–Boltzmann or generalized Born and surface area continuum solvation (MM/PBSA and MM/GBSA) approaches are utilized in certain drug design strategies to estimate the free energy binding of small ligands to receptors (Genheden & Ryde, 2015; Yang et al., 2020). First, a total of 200 snapshots from simulated MD trajectory for 20 ns were extracted to explore MM-PB(GB)SA free energy and energy decomposition ( $E_{DEC}$ ) per residue. Then, van der Waals force ( $\Delta E_{VDW}$ ), electrostatic interaction ( $\Delta E_{ELE}$ ), total (polar+nonpolar) solvation free energy ( $\Delta E_{SOL}$ ) contributions, and total estimated binding free energy ( $\Delta G_{Bind}$ ) were employed as post-processing approach, using two different web servers AMMIS (Wu et al., 2018) and farPPI (Wang et al., 2019), by following recommended tool instructions.

## References

Ali, I. A. I., Al-Masoudi, I. A., Aziz, N. M., & Al-Masoudi, N. A. (2008). New acyclic quinoxaline nucleosides. Synthesis and anti-HIV activity. *Nucleosides, Nucleotides and Nucleic Acids*, 27(2), 146–156.

<https://doi.org/10.1080/15257770701795920>

Aloia, L. (2021). Epigenetic Regulation of Cell-Fate Changes That Determine Adult Liver Regeneration After Injury. In *Frontiers in Cell and Developmental Biology* (Vol. 9).

<https://doi.org/10.3389/fcell.2021.643055>

Antunes, D. A., Devaurs, D., & Kavraki, L. E. (2015). Understanding the challenges of protein flexibility in drug design. *Http://Dx.Doi.Org/10.1517/17460441.2015.1094458*, 10(12), 1301–1313.

<https://doi.org/10.1517/17460441.2015.1094458>

Ballestar, E., Yusufzai, T. M., & Wolffe, A. P. (2000). 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(24), 7100–7106. <https://doi.org/10.1021/bi0001271>

Baubec, T., Ivánek, R., Lienert, F., & Schübeler, D. (2013). Methylation-dependent and -independent genomic targeting principles of the mbd protein family. *Cell*, 153(2). <https://doi.org/10.1016/j.cell.2013.03.011>

Bell, C. G., Lowe, R., Adams, P. D., Baccarelli, A. A., Beck, S., Bell, J. T., Christensen, B. C., Gladyshev, V. N., Heijmans, B. T., Horvath, S., Ideker, T., Issa, J. P. J., Kelsey, K. T., Marioni, R. E., Reik, W., Relton, C. L., Schalkwyk, L. C., Teschendorff, A. E., Wagner, W., ... Rakyan, V. K. (2019). DNA methylation aging clocks: Challenges and recommendations. In *Genome Biology* (Vol. 20, Issue 1). <https://doi.org/10.1186/s13059-019-1824-y>

- Best, R. B., Hummer, G., & Eaton, W. A. (2013). Native contacts determine protein folding mechanisms in atomistic simulations. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(44), 17874–17879. <https://doi.org/10.1073/pnas.1311599110>
- Brandt, B., Rashidiani, S., Bán, Á., & Rauch, T. A. (2019). DNA methylation-governed gene expression in autoimmune arthritis. In *International Journal of Molecular Sciences* (Vol. 20, Issue 22). <https://doi.org/10.3390/ijms20225646>
- Brylinski, M., & Skolnick, J. (2008). Q-dock: Low-resolution flexible ligand docking with pocket-specific threading restraints. *Journal of Computational Chemistry*, *29*(10), 1574–1588. <https://doi.org/10.1002/jcc.20917>
- Buchmuller, B. C., Kosel, B., & Summerer, D. (2020). Complete Profiling of Methyl-CpG-Binding Domains for Combinations of Cytosine Modifications at CpG Dinucleotides Reveals Differential Read-out in Normal and Rett-Associated States. *Scientific Reports*, *10*(1), 1–9. <https://doi.org/10.1038/s41598-020-61030-1>
- Çalışkaner, Z. O. (2021). Determination of binding potential of HCV protease inhibitors against SARS-CoV-2 Papain-like protease with computational docking. *Letters in Drug Design & Discovery*, *18*(10), 949–960. <https://doi.org/http://dx.doi.org/10.2174/1570180818666210531092605>
- Cedar, H., & Bergman, Y. (2009). Linking DNA methylation and histone modification: Patterns and paradigms. In *Nature Reviews Genetics* (Vol. 10, Issue 5). <https://doi.org/10.1038/nrg2540>
- Cheishvili, D., Chik, F., Li, C. C., Bhattacharya, B., Suderman, M., Arakelian, A., Hallett, M., Rabbani, S. A., & Szyf, M. (2014). Synergistic effects of combined DNA methyltransferase inhibition and MBD2 depletion on breast cancer cells; MBD2 depletion blocks 5-aza-2'-deoxycytidine-triggered invasiveness. *Carcinogenesis*, *35*(11), 2436. <https://doi.org/10.1093/CARCIN/BGU181>
- Cheng, R. K. Y., Segala, E., Robertson, N., Deflorian, F., Doré, A. S., Errey, J. C., Fiez-Vandal, C., Marshall, F. H., & Cooke, R. M. (2017). Structures of Human A1 and A2A Adenosine Receptors with Xanthines Reveal Determinants of Selectivity. *Structure*, *25*(8), 1275-1285.e4. <https://doi.org/10.1016/j.str.2017.06.012>
- Cramer, J. M., Scarsdale, J. N., Walavalkar, N. M., Buchwald, W. A., Ginder, G. D., & Williams, D. C. (2014). Probing the dynamic distribution of bound states for methylcytosine-binding domains on DNA. *Journal of Biological Chemistry*, *289*(3), 1294–1302. <https://doi.org/10.1074/jbc.M113.512236>
- Craveur, P., Joseph, A. P., Esque, J., Narwani, T. J., Noël, F., Shinada, N., Goguet, M., Leonard, S., Poulain, P., Bertrand, O., Faure, G., Rebehmed, J., Ghozlane, A., Swapna, L. S., Bhaskara, R. M., Barnoud, J., Téletchéa, S., Jallu, V., Cerny, J., ... de Brevern, A. G. (2015). Protein flexibility in the light of structural alphabets. In *Frontiers in Molecular Biosciences* (Vol. 2, Issue MAY). Frontiers Media S.A. <https://doi.org/10.3389/fmolb.2015.00020>

- Dallakyan, S., & Olson, A. J. (2015). Small-molecule library screening by docking with PyRx. *Methods in Molecular Biology*, 1263, 243–250. [https://doi.org/10.1007/978-1-4939-2269-7\\_19](https://doi.org/10.1007/978-1-4939-2269-7_19)
- Dean, W., Lucifero, D., & Santos, F. (2005). DNA methylation in mammalian development and disease. In *Birth Defects Research Part C - Embryo Today: Reviews* (Vol. 75, Issue 2). <https://doi.org/10.1002/bdrc.20037>
- Du, Q., Luu, P. L., Stirzaker, C., & Clark, S. J. (2015). Methyl-CpG-binding domain proteins: Readers of the epigenome. In *Epigenomics* (Vol. 7, Issue 6, pp. 1051–1073). Future Medicine Ltd. <https://doi.org/10.2217/epi.15.39>
- Fieger, S. M., & Wong, B. J. (2010). Adenosine receptor inhibition with theophylline attenuates the skin blood flow response to local heating in humans. *Experimental Physiology*, 95(9), 946–954. <https://doi.org/10.1113/expphysiol.2010.053538>
- Fraga, M. F., Ballestar, E., Montoya, G., Taysavang, P., Wade, P. A., & Esteller, M. (2003). The affinity of different MBD proteins for a specific methylated locus depends on their intrinsic binding properties. In *Nucleic Acids Research* (Vol. 31, Issue 6, pp. 1765–1774). Oxford University Press. <https://doi.org/10.1093/nar/gkg249>
- García, B., Garcia-Tojal, J., Ruiz, R., Gil-García, R., Ibeas, S., Donnadieu, B., & Leal, J. M. (2008). Interaction of the DNA bases and their mononucleotides with pyridine-2-carbaldehyde thiosemicarbazonecopper(II) complexes. Structure of the cytosine derivative. *Journal of Inorganic Biochemistry*, 102(10), 1892–1900. <https://doi.org/10.1016/j.jinorgbio.2008.06.013>
- Genheden, S., & Ryde, U. (2015). The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. In *Expert Opinion on Drug Discovery* (Vol. 10, Issue 5, pp. 449–461). Informa Healthcare. <https://doi.org/10.1517/17460441.2015.1032936>
- Ghosh, R. P., Nikitina, T., Horowitz-Scherer, R. A., Gierasch, L. M., Uversky, V. N., Hite, K., Hansen, J. C., & Woodcock, C. L. (2010). Unique physical properties and interactions of the domains of methylated DNA binding protein 2 (MeCP2). *Biochemistry*, 49(20), 4395. <https://doi.org/10.1021/BI9019753>
- Ginder, G. D., & Williams, D. C. (2018). Readers of DNA methylation, the MBD family as potential therapeutic targets. In *Pharmacology and Therapeutics* (Vol. 184, pp. 98–111). Elsevier Inc. <https://doi.org/10.1016/j.pharmthera.2017.11.002>
- Giovinazzo, H., Reichert, Z. R., Bergman, A., Lin, X., Wyhs, N., Esopi, D., Vaghasia, A., Liu, J., Jain, Y., Bhamidipati, A., Steinberg, R., Speed, T., Vaughn, M., Zhang, Y., Brennen, N., Deweese, T., Yegnasubramanian, S., & Nelson, W. G. (2018). Abstract 5881: Novel inhibitors of the epigenetic reader protein MBD2. *Cancer Research*, 78(13 Supplement), 5881–5881. <https://doi.org/10.1158/1538-7445.AM2018-5881>

Gowher, H., & Jeltsch, A. (2019). Erratum: Mammalian DNA methyltransferases: New discoveries and open questions (Biochemical Society Transactions (2018)46 (1191-1202)Doi: 10.1042/BST20170574). In *Biochemical Society Transactions* (Vol. 47, Issue 3). [https://doi.org/10.1042/BST-2017-0574C\\_COR](https://doi.org/10.1042/BST-2017-0574C_COR)

Hendrich, B., & Bird, A. (1998). Identification and Characterization of a Family of Mammalian Methyl-CpG Binding Proteins. *Molecular and Cellular Biology*, 18(11), 6538–6547. <https://doi.org/10.1128/mcb.18.11.6538>

Huang, S.-Y., & Zou, X. (2010). Advances and Challenges in Protein-Ligand Docking. *International Journal of Molecular Sciences*, 11(8), 3016. <https://doi.org/10.3390/IJMS11083016>

Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: Visual molecular dynamics. *Journal of Molecular Graphics*, 14(1), 33–38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5)

Humphreys, B. D., Virginio, C., Surprenant, A., Rice, J., & Dubyak, G. R. (1998). Isoquinolines as antagonists of the P2X7 nucleotide receptor: High selectivity for the human versus rat receptor homologues. *Molecular Pharmacology*, 54(1), 22–32. <https://doi.org/10.1124/mol.54.1.22>

Irvine, R. A., Lin, I. G., & Hsieh, C.-L. (2002). DNA Methylation Has a Local Effect on Transcription and Histone Acetylation. *Molecular and Cellular Biology*, 22(19). <https://doi.org/10.1128/mcb.22.19.6689-6696.2002>

Jin, S. G., Jiang, C. L., Rauch, T., Li, H., & Pfeifer, G. P. (2005). MBD3L2 interacts with MBD3 and components of the NuRD complex and can oppose MBD2-MeCP1-mediated methylation silencing. *Journal of Biological Chemistry*, 280(13), 12700–12709. <https://doi.org/10.1074/jbc.M413492200>

Jurkowska, R. Z., & Jeltsch, A. (2016). Enzymology of mammalian DNA methyltransferases. *Advances in Experimental Medicine and Biology*, 945. [https://doi.org/10.1007/978-3-319-43624-1\\_5](https://doi.org/10.1007/978-3-319-43624-1_5)

Kim, M. Y., Na, I., Kim, J. S., Son, S. H., Choi, S., Lee, S. E., Kim, J.-H., Jang, K., Alterovitz, G., Chen, Y., Vaart, A. van der, Won, H.-S., Uversky, V. N., & Kim, C. G. (2019). Rational discovery of antimetastatic agents targeting the intrinsically disordered region of MBD2. *Science Advances*, 5(11), eaav9810. <https://doi.org/10.1126/SCIADV.AAV9810>

Kim, S., Lee, J., Jo, S., Brooks, C. L., Lee, H. S., & Im, W. (2017). CHARMM-GUI ligand reader and modeler for CHARMM force field generation of small molecules. *Journal of Computational Chemistry*, 38(21), 1879–1886. <https://doi.org/10.1002/jcc.24829>

Kribelbauer, J. F., Lu, X. J., Rohs, R., Mann, R. S., & Bussemaker, H. J. (2020). Toward a Mechanistic Understanding of DNA Methylation Readout by Transcription Factors. In *Journal of Molecular Biology* (Vol. 432, Issue 6). <https://doi.org/10.1016/j.jmb.2019.10.021>

Landells, L. J., Jensen, M. W., Orr, L. M., Spina, D., O'Connor, B. J., & Page, C. P. (2000). The role of adenosine receptors in the action of theophylline on human peripheral blood mononuclear cells from

healthy and asthmatic subjects. *British Journal of Pharmacology*, 129(6), 1140–1144.

<https://doi.org/10.1038/sj.bjp.0703177>

Li, L., Li, N., Liu, N., Huo, F., & Zheng, J. (2020). MBD2 Correlates with a Poor Prognosis and Tumor Progression in Renal Cell Carcinoma. *OncoTargets and Therapy*, 13, 10001–10012.

<https://doi.org/10.2147/OTT.S256226>

Liu, K., Xu, C., Lei, M., Yang, A., Loppnau, P., Hughes, T. R., & Min, J. (2018). Structural basis for the ability of MBD domains to bind methyl-CG and TG sites in DNA. *Journal of Biological Chemistry*, 293(19), 7344–7354. <https://doi.org/10.1074/jbc.RA118.001785>

Liu, M., Movahed, S., Dangi, S., Pan, H., Kaur, P., Bilinovich, S. M., Faison, E. M., Leighton, G. O., Wang, H., Williams, D. C., & Riehn, R. (2020). DNA looping by two 5-methylcytosine-binding proteins quantified using nanofluidic devices. *Epigenetics and Chromatin*, 13(1), 1–10. <https://doi.org/10.1186/s13072-020-00339-7>

Locke, W. J., Guanzone, D., Ma, C., Liew, Y. J., Duesing, K. R., Fung, K. Y. C., & Ross, J. P. (2019). DNA Methylation Cancer Biomarkers: Translation to the Clinic. In *Frontiers in Genetics* (Vol. 10).

<https://doi.org/10.3389/fgene.2019.01150>

Lu, Y., Loh, Y. H., Li, H., Cesana, M., Ficarro, S. B., Parikh, J. R., Salomonis, N., Toh, C. X. D., Andreadis, S. T., Luckey, C. J., Collins, J. J., Daley, G. Q., & Marto, J. A. (2014). Alternative splicing of MBD2 supports self-renewal in human pluripotent stem cells. *Cell Stem Cell*, 15(1), 92–101.

<https://doi.org/10.1016/j.stem.2014.04.002>

Luscombe, N. M., Laskowski, R. A., & Thornton, J. M. (2001). Amino acid–base interactions: a three-dimensional analysis of protein–DNA interactions at an atomic level. *Nucleic Acids Research*, 29(13), 2860–2874. <https://doi.org/10.1093/NAR/29.13.2860>

MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., ... Karplus, M. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *Journal of Physical Chemistry B*, 102(18), 3586–3616.

<https://doi.org/10.1021/jp973084f>

Martin, V., Jørgensen, H. F., Chaubert, A. S. B., Berger, J., Barr, H., Shaw, P., Bird, A., & Chaubert, P. (2008). MBD2-Mediated Transcriptional Repression of the p14ARF Tumor Suppressor Gene in Human Colon Cancer Cells. *Pathobiology*, 75(5), 281–287. <https://doi.org/10.1159/000151708>

Menafrá, R., & Stunnenberg, H. G. (2014). MBD2 and MBD3: Elusive functions and mechanisms. In *Frontiers in Genetics* (Vol. 5, Issue DEC). Frontiers Research Foundation.

<https://doi.org/10.3389/fgene.2014.00428>

- Mian, O. Y., Wang, S. Z., Zhu, S. Z., Gnanapragasam, M. N., Graham, L., Bear, H. D., & Ginder, G. D. (2011). Methyl-binding domain protein 2-dependent proliferation and survival of breast cancer cells. *Molecular Cancer Research*, *9*(8), 1152–1162. <https://doi.org/10.1158/1541-7786.MCR-11-0252>
- Millard, C. J., Fairall, L., Ragan, T. J., Savva, C. G., & Schwabe, J. W. R. (2020). The topology of chromatin-binding domains in the NuRD deacetylase complex. *Nucleic Acids Research*, *48*(22), 12972–12982. <https://doi.org/10.1093/nar/gkaa1121>
- Moore, L. D., Le, T., & Fan, G. (2013). DNA methylation and its basic function. In *Neuropsychopharmacology* (Vol. 38, Issue 1). <https://doi.org/10.1038/npp.2012.112>
- Morris, G. M., Ruth, H., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). Software news and updates AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry*, *30*(16), 2785–2791. <https://doi.org/10.1002/jcc.21256>
- Nelson, W. G., Yegnasubramanian, S., Lin, X., Speed, T. J., & Reichert, Z. (2010). *Agents for Reversing Epigenetic Silencing of Genes*.
- Ng, H. H., Zhang, Y., Hendrich, B., Johnson, C. A., Turner, B. M., Erdjument-Bromage, H., Tempst, P., Reinberg, D., & Bird, A. (1999). MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nature Genetics*, *23*(1), 58–61. <https://doi.org/10.1038/12659>
- Nikitina, T., Shi, X., Ghosh, R. P., Horowitz-Scherer, R. A., Hansen, J. C., & Woodcock, C. L. (2007). Multiple Modes of Interaction between the Methylated DNA Binding Protein MeCP2 and Chromatin. *Molecular and Cellular Biology*, *27*(3), 864. <https://doi.org/10.1128/MCB.01593-06>
- O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., & Hutchison, G. R. (2011). Open Babel: An Open chemical toolbox. *Journal of Cheminformatics*, *3*(10), 33. <https://doi.org/10.1186/1758-2946-3-33>
- Ohki, I., Shimotake, N., Fujita, N., Jee, J. G., Ikegami, T., Nakao, M., & Shirakawa, M. (2001). Solution structure of the methyl-CpG binding domain of human MBD1 in complex with methylated DNA. *Cell*, *105*(4), 487–497. [https://doi.org/10.1016/S0092-8674\(01\)00324-5](https://doi.org/10.1016/S0092-8674(01)00324-5)
- Pan, H., Bilinovich, S. M., Kaur, P., Riehn, R., Wang, H., & Williams, D. C. (2017). CpG and methylation-dependent DNA binding and dynamics of the methylcytosine binding domain 2 protein at the single-molecule level. *Nucleic Acids Research*, *45*(15), 9164–9177. <https://doi.org/10.1093/nar/gkx548>
- Pantsar, T., & Poso, A. (2018). Binding affinity via docking: Fact and fiction. In *Molecules* (Vol. 23, Issue 8, p. 1DUMMY). MDPI AG. <https://doi.org/10.3390/molecules23081899>
- Parra, S., Laurent, F., Subra, G., Deleuze-Masquefa, C., Benezech, V., Fabreguettes, J. R., Vidal, J. P., Pocock, T., Elliott, K., Small, R., Escale, R., Michel, A., Chapat, J. P., & Bonnet, P. A. (2001). Imidazo[1,2-

a]quinoxalines: Synthesis and cyclic nucleotide phosphodiesterase inhibitory activity. *European Journal of Medicinal Chemistry*, 36(3), 255–264. [https://doi.org/10.1016/S0223-5234\(01\)01213-2](https://doi.org/10.1016/S0223-5234(01)01213-2)

Phillips, J. C., Hardy, D. J., Maia, J. D. C., Stone, J. E., Ribeiro, J. V., Bernardi, R. C., Buch, R., Fiorin, G., Hénin, J., Jiang, W., McGreevy, R., Melo, M. C. R., Radak, B. K., Skeel, R. D., Singharoy, A., Wang, Y., Roux, B., Aksimentiev, A., Luthey-Schulten, Z., ... Tajkhorshid, E. (2020). Scalable molecular dynamics on CPU and GPU architectures with NAMD. *Journal of Chemical Physics*, 153(4), 044130. <https://doi.org/10.1063/5.0014475>

Price, D. J., & Brooks, C. L. (2004). A modified TIP3P water potential for simulation with Ewald summation. *Journal of Chemical Physics*, 121(20), 10096–10103. <https://doi.org/10.1063/1.1808117>

Ravindranath, P. A., Forli, S., Goodsell, D. S., Olson, A. J., & Sanner, M. F. (2015). AutoDockFR: Advances in Protein-Ligand Docking with Explicitly Specified Binding Site Flexibility. *PLOS Computational Biology*, 11(12), e1004586. <https://doi.org/10.1371/JOURNAL.PCBI.1004586>

Röthlisberger, P., Levi-Acobas, F., Sarac, I., Marlière, P., Herdewijn, P., & Hollenstein, M. (2017). On the enzymatic incorporation of an imidazole nucleotide into DNA. *Organic and Biomolecular Chemistry*, 15(20), 4449–4455. <https://doi.org/10.1039/c7ob00858a>

Rugowska, A., Starosta, A., & Konieczny, P. (2021). Epigenetic modifications in muscle regeneration and progression of Duchenne muscular dystrophy. In *Clinical Epigenetics* (Vol. 13, Issue 1). <https://doi.org/10.1186/s13148-021-01001-z>

Salentin, S., Schreiber, S., Haupt, V. J., Adasme, M. F., & Schroeder, M. (2015). PLIP: Fully automated protein-ligand interaction profiler. *Nucleic Acids Research*, 43(W1), W443–W447. <https://doi.org/10.1093/nar/gkv315>

Samblas, M., Milagro, F. I., & Martínez, A. (2019). DNA methylation markers in obesity, metabolic syndrome, and weight loss. In *Epigenetics* (Vol. 14, Issue 5). <https://doi.org/10.1080/15592294.2019.1595297>

Scarsdale, J. N., Webb, H. D., Ginder, G. D., & Williams, D. C. (2011). Solution structure and dynamic analysis of chicken MBD2 methyl binding domain bound to a target-methylated DNA sequence. *Nucleic Acids Research*, 39(15), 6741–6752. <https://doi.org/10.1093/NAR/GKR262>

Sterling, T., & Irwin, J. J. (2015). ZINC 15 - Ligand Discovery for Everyone. *Journal of Chemical Information and Modeling*, 55(11), 2324–2337. <https://doi.org/10.1021/acs.jcim.5b00559>

Stirzaker, C., Song, J. Z., Ng, W., Du, Q., Armstrong, N. J., Locke, W. J., Statham, A. L., French, H., Pidsley, R., Valdes-Mora, F., Zotenko, E., & Clark, S. J. (2017). Methyl-CpG-binding protein MBD2 plays a key role in maintenance and spread of DNA methylation at CpG islands and shores in cancer. *Oncogene*, 36(10), 1328–1338. <https://doi.org/10.1038/onc.2016.297>

- Stricker, S. H., & Götz, M. (2018). DNA-methylation: Master or slave of neural fate decisions? In *Frontiers in Neuroscience* (Vol. 12, Issue FEB). <https://doi.org/10.3389/fnins.2018.00005>
- Vangone, A., Schaarschmidt, J., Koukos, P., Geng, C., Citro, N., Trellet, M. E., Xue, L. C., & Bonvin, A. M. J. J. (2019). Large-scale prediction of binding affinity in protein–small ligand complexes: the PRODIGY-LIG web server. *Bioinformatics*, *35*(9), 1585–1587. <https://doi.org/10.1093/BIOINFORMATICS/BTY816>
- Vogt, G. (2017). Evolution of epigenetic mechanisms in animals and their role in speciation. In *Handbook of Epigenetics: The New Molecular and Medical Genetics*. <https://doi.org/10.1016/B978-0-12-805388-1.00027-4>
- Wang, Z., Wang, X., Li, Y., Lei, T., Wang, E., Li, D., Kang, Y., Zhu, F., & Hou, T. (2019). FarPPI: A webserver for accurate prediction of protein-ligand binding structures for small-molecule PPI inhibitors by MM/PB(GB)SA methods. *Bioinformatics*, *35*(10), 1777–1779. <https://doi.org/10.1093/bioinformatics/bty879>
- Watano, T., Matsuoka, I., & Kimura, J. (2002). Characteristics of ATP-induced current through P2X7 receptor in NG108-15 cells: Unique antagonist sensitivity and lack of pore formation. *Japanese Journal of Pharmacology*, *88*(4), 428–435. <https://doi.org/10.1254/jjp.88.428>
- Wood, K. H., Johnson, B. S., Welsh, S. A., Lee, J. Y., Cui, Y., Krizman, E., Brodtkin, E. S., Blendy, J. A., Robinson, M. B., Bartolomei, M. S., & Zhou, Z. (2016). Tagging methyl-CpG-binding domain proteins reveals different spatiotemporal expression and supports distinct functions. *Epigenomics*, *8*(4), 455–473. <https://doi.org/10.2217/epi-2015-0004>
- Wood, K. H., & Zhou, Z. (2016). Emerging molecular and biological functions of MBD2, a reader of DNA methylation. In *Frontiers in Genetics* (Vol. 7, Issue MAY). Frontiers Media S.A. <https://doi.org/10.3389/fgene.2016.00093>
- Wu, F. X., Wang, F., Yang, J. F., Jiang, W., Wang, M. Y., Jia, C. Y., Hao, G. F., & Yang, G. F. (2018). AIMMS suite: A web server dedicated for prediction of drug resistance on protein mutation. *Briefings in Bioinformatics*, *21*(1), 318–328. <https://doi.org/10.1093/bib/bby113>
- Wyhs, N., Walker, D., Giovinazzo, H., Yegnasubramanian, S., & Nelson, W. G. (2014). Time-resolved fluorescence resonance energy transfer assay for discovery of small-molecule inhibitors of methyl-CpG binding domain protein 2. *Journal of Biomolecular Screening*, *19*(7), 1060–1069. <https://doi.org/10.1177/1087057114526433>
- Xiong, G., Wu, Z., Yi, J., Fu, L., Yang, Z., Hsieh, C., Yin, M., Zeng, X., Wu, C., Lu, A., Chen, X., Hou, T., & Cao, D. (2021). ADMETlab 2.0: an integrated online platform for accurate and comprehensive predictions of ADMET properties. *Nucleic Acids Research*, *1*. <https://doi.org/10.1093/nar/gkab255>

Yang, J. F., Wang, F., Chen, Y. Z., Hao, G. F., & Yang, G. F. (2020). LARMD: Integration of bioinformatic resources to profile ligand-driven protein dynamics with a case on the activation of estrogen receptor. *Briefings in Bioinformatics*, *21*(6), 2206–2218. <https://doi.org/10.1093/bib/bbz141>

Zhong, J., Agha, G., & Baccarelli, A. A. (2016). The Role of DNA Methylation in Cardiovascular Risk and Disease: Methodological Aspects, Study Design, and Data Analysis for Epidemiological Studies. In *Circulation Research* (Vol. 118, Issue 1). <https://doi.org/10.1161/CIRCRESAHA.115.305206>

Zhu, D., Hunter, S. B., Vertino, P. M., & Meir, E. G. Van. (2011). Overexpression of MBD2 in Glioblastoma Maintains Epigenetic Silencing and Inhibits the Anti-Angiogenic Function of the Tumor Suppressor Gene BAI1. *Cancer Research*, *71*(17), 5859. <https://doi.org/10.1158/0008-5472.CAN-11-1157>

Zhu, D., Osuka, S., Zhang, Z., Reichert, Z. R., Yang, L., Kanemura, Y., Jiang, Y., You, S., Zhang, H., Devi, N. S., Bhattacharya, D., Takano, S., Gillespie, G. Y., Macdonald, T., Tan, C., Nishikawa, R., Nelson, W. G., Olson, J. J., & Van Meir, E. G. (2018). BAI1 Suppresses Medulloblastoma Formation by Protecting p53 from Mdm2-Mediated Degradation. *Cancer Cell*, *33*(6), 1004-1016.e5. <https://doi.org/10.1016/J.CCELL.2018.05.006>

Zhu, H., Wang, G., & Qian, J. (2016). Transcription factors as readers and effectors of DNA methylation. *Nature Reviews Genetics*, *17*(9). <https://doi.org/10.1038/nrg.2016.83>

Zou, X., Ma, W., Solov'yov, I. A., Chipot, C., & Schulten, K. (2012). Recognition of methylated DNA through methyl-CpG binding domain proteins. *Nucleic Acids Research*, *40*(6), 2747–2758. <https://doi.org/10.1093/NAR/GKR1057>

## Declarations

### Author Contributions

Z.O.Ç. solely employed throughout the workflow of manuscript including conceptualization, investigation, data acquisition and interpretation, figure formatting, writing original draft, and proofreading manuscript.

### Competing Interests

The author declares no competing interests.

### Additional Information

**Supplementary information** is available.

**Correspondence** and requests should be addressed to Z.O.Ç.

## Tables

**Table 1.** Calculated binding energies, RMSD values and inhibition constants ( $K_i$ ) of the best enzyme-drug docking poses.

| PubChem ID<br>or Annotation   | Binding Energy<br>(kcal/mol) | RMSD<br>(Å°) | Estimated $K_i$ (nM) |
|-------------------------------|------------------------------|--------------|----------------------|
| CID3100583                    | -11.42                       | 275.20       | 4.22                 |
| CID343482                     | -10.14                       | 277.80       | 37.09                |
| CID136748749                  | -10.01                       | 274.46       | 45.59                |
| CID46959745                   | -9.88                        | 281.55       | 57.48                |
| CID126001284                  | -9.67                        | 277.00       | 80.96                |
| CID125615373                  | -9.41                        | 274.33       | 127.17               |
| CID28689118                   | -9.26                        | 274.75       | 162.21               |
| CID25826758                   | -9.20                        | 279.74       | 180.51               |
| CID323153                     | -8.34                        | 276.53       | 766.71               |
| CID24241                      | -8.29                        | 277.420      | 833.56               |
| 8,8'-Ethylenebis-Theophylline | -10.32                       | 278.82       | 27.30                |
| Regadenoson                   | -9.83                        | 285.73       | 62.67                |
| Sanguinarine                  | -8.76                        | 277.98       | 308.44               |
| GHMCytosine                   | -8.18                        | 277.95       | 1001.00              |
| Golgicide A                   | -7.81                        | 278.05       | 1880.00              |

**Table 2.** Number of hydrogen bonds (H-bond) and the list of interacted amino acids with inhibitor.

| <i>Compound</i>               | <i>Inhibition Constant</i> | <i>Number of H-bonds</i> | <i>Amino acids involved in interaction</i>  |
|-------------------------------|----------------------------|--------------------------|---|
| CID3100583                    | 4.22 nM                    | 9                        | <b>Val177, Arg188</b> , Lys190, Ser204, <b>Phe206</b> , Asp207, <b>Phe208, Arg209</b> |
| CID343482                     | 37.04 nM                   | 9                        | <b>Val177</b> , Lys190, <b>Phe206</b> , Asp207, <b>Phe208, Arg209</b>                 |
| CID136748749                  | 45.59 nM                   | 6                        | Lys190, <b>Phe206, Phe208, Arg209</b>   |
| CID46959745                   | 57.48 nM                   | 5                        | <b>Asp151, Cys152, Leu155</b> , Gly158  |
| CID126001284                  | 80.96 nM                   | 2                        | <b>Ser205, Phe206</b>   |
| CID125615373                  | 127.17 nM                  | 1                        | <b>Ala154</b>   |
| CID28689118                   | 162.21 nM                  | 2                        | <b>Asp176, Phe208</b>   |
| CID25826758                   | 180.51 nM                  | 5                        | <b>Val177, Arg188</b> , Lys190, <b>Arg209</b>   |
| CID323153                     | 766.71 nM                  | 2                        | <b>Arg209</b>   |
| CID24241                      | 833.56 nM                  | 4                        | <b>Phe208, Arg209</b>   |
| 8,8'-Ethylenebis-Theophylline | 27.30 nM                   | 5                        | <b>Val177, Phe208, Arg209</b>   |
| Regadenoson                   | 62.67 nM                   | 9                        | Trp159, Ser183, Tyr196, Thr200  |
| Sanguinarine                  | 308.44 nM                  | -                        | -   |
| GHMCytosine                   | 1.01 $\mu$ M               | 9                        | <b>Val177, Arg188</b> , Lys190, <b>Phe206, Phe208, Arg209</b>                         |
| Golgicide A                   | 1.88 $\mu$ M               | 2                        | <b>Phe208, Arg209</b>   |

- *Certain residues are colored in association with druggable pockets.*

**Table 3.** Number of hydrophobic and other interactions and the list of interacted amino acids with inhibitor.

| <i>Compound</i>                            | <i># of Hydrophobic Interactions</i> | <i># of <math>\pi</math>-Stacking/Cation</i> | <i># of Salt Bridges</i> | <i># of Halogen Bonds</i> | <i>Interacting Amino Acids</i>                |
|--|--------------------------------------|--|--------------------------|---------------------------|---|
| CID3100583                                 | 2                                    | -  | -                        | -                         | <b>Phe208, Arg209</b>                         |
| CID343482                                  | -                                    | 2  | 1                        | -                         | Asp207, <b>Phe208</b>                         |
| CID136748749                               | 2                                    | -  | -                        | -                         | <b>Asp176, Arg209</b>                         |
| CID46959745                                | 6                                    | -  | -                        | -                         | <b>Leu155, Pro156, Pro157, Val201, Leu203</b> |
| CID126001284                               | 6                                    | -  | -                        | -                         | <b>Leu155, Pro156, Val201, Leu203</b>         |
| CID125615373                               | 5                                    | -  | -                        | -                         | <b>Ala154, Val201, Leu203, Phe206, Lys212</b> |
| CID28689118                                | 4                                    | 2  | -                        | -                         | Lys190, Asp207, <b>Phe208</b>                 |
| CID25826758                                | 4                                    | -  | -                        | -                         | Lys190, <b>Phe208</b>                         |
| CID323153                                  | 6                                    | -  | -                        | -                         | <b>Val177, Lys190, Asp207, Phe208</b>         |
| CID24241                                   | 6                                    | -  | -                        | -                         | <b>Asp176, Lys190, Phe208</b>                 |
| Theophylline, 8,8'-Ethylenebis-Regadeneson | 3                                    | -  | -                        | -                         | <b>Val177, Phe208, Arg209</b>                 |
|  | -                                    | 1  | -                        | -                         | Trp159  |
| Sanguinarine                               | 4                                    | -  | -                        | -                         | <b>Val177, Phe208</b>                         |
| GHMCytosine                                | 1                                    | -  | -                        | -                         | <b>Arg209</b>                                 |
| Golgicide A                                | 3                                    | 1  | -                        | 1                         | <b>Asp176, Val177, Lys190, Arg209</b>         |

- *Certain residues are colored in association with druggable pockets.*

## Figures

Figure 1

The number of interactions contributed by common amino acids which have distributed within top ten non-annotated and five annotated compounds from virtual screening.

## Figure 2

Illustration of docked MBD2<sub>MBD</sub>-Ligand complexes. The best docking pose with interactions between atoms of MBD2<sub>MBD</sub> (blue) and CID3100583 (yellow) **(a)**, visual binding surface **(b)**, and depiction in 2D diagram **(c)**. Docking attitude of MBD2<sub>MBD</sub>-Ebis-Theophylline complex in spatial **(d)** and 2D representation **(e)**. Hydrogen bonds and hydrophobic interactions are indicated by solid blue and dashed grey lines, respectively, with the distances between the atoms in angstrom.

## Figure 3

Apo (unbound) receptor and holo (ligand-bound) complex was superimposed to show side-chain motions upon ligand docking. Side-chains in CID3100583-bound **(a)** and Ebis-Theophylline-bound **(b)** holo complexes were displayed in magenta sticks while side-chains in apo state were colored in green sticks and ligands were depicted as black lines. Re-docking by IFD resulted in similar residues in CID3100583 and Ebis-Theophylline binding pockets **(c-d)**. See also Supplementary Movies for a dimensional view of superimposed conformations.

## Figure 4

MD simulation output of MBD2<sub>MBD</sub>-CID3100583 system. RMSD values of all top non-annotated ligands along the whole trajectory **(a)**. Overall fluctuation of ligand and protein backbone **(b)**, atomic fluctuations and B-factor per residue **(c)**, and RoG with Q(x) **(d)** are plotted for solely CID3100583. Tiny squares on RMSF and B-factor curves, corresponds to Val177, Arg188, Lys190, residues between Ser204-Arg209, respectively.

## Figure 5

MD simulation output of MBD2<sub>MBD</sub>-Ebis-Theophylline system. RMSD values of certain top annotated ligands along the whole trajectory **(a)**. Overall fluctuation of ligand and protein backbone **(b)**, atomic fluctuations and B-factor per residue **(c)**, and RoG with Q(x) **(d)** are plotted for solely 8,8-

Ethylenebistheophylline. Tiny squares on RMSF and B-factor curves, corresponds to residues between Asp176–Tyr178, Arg188, residues between Asp207–Arg209, respectively.

## Figure 6

Detailed binding free energetics of both MD simulation output of MBD2<sub>MBD</sub>–Ligand systems. Graphics of top ten residues (top panels) with the highest energy decomposition per residue for MBD2<sub>MBD</sub>–CID3100583 **(a)** and MBD2<sub>MBD</sub>–Ebis-Theophylline **(b)** complexes. Bottom panels **(a, b)** infer different energetic contributions of all residues within each system.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [RevisedSupplementaryInformation.pdf](#)