

MOLECULAR BIOLOGY

Rational discovery of antimetastatic agents targeting the intrinsically disordered region of MBD2

Min Young Kim^{1*}, Insung Na^{2*†}, Ji Sook Kim^{1,3}, Seung Han Son¹, Sungwoo Choi¹, Seol Eui Lee¹, Ji-Hun Kim⁴, Kiseok Jang³, Gil Alterovitz⁵, Yu Chen², Arjan van der Vaart⁶, Hyung-Sik Won^{7‡}, Vladimir N. Uversky^{2,8‡}, Chul Geun Kim^{1‡}

Although intrinsically disordered protein regions (IDPRs) are commonly engaged in promiscuous protein-protein interactions (PPIs), using them as drug targets is challenging due to their extreme structural flexibility. We report a rational discovery of inhibitors targeting an IDPR of MBD2 that undergoes disorder-to-order transition upon PPI and is critical for the regulation of the Mi-2/NuRD chromatin remodeling complex (CRC). Computational biology was essential for identifying target site, searching for promising leads, and assessing their binding feasibility and off-target probability. Molecular action of selected leads inhibiting the targeted PPI of MBD2 was validated in vitro and in cell, followed by confirming their inhibitory effects on the epithelial-mesenchymal transition of various cancer cells. Identified lead compounds appeared to potently inhibit cancer metastasis in a murine xenograft tumor model. These results constitute a pioneering example of rationally discovered IDPR-targeting agents and suggest Mi-2/NuRD CRC and/or MBD2 as a promising target for treating cancer metastasis.

INTRODUCTION

Although at least 650,000 protein-protein interactions (PPIs) might occur in humans, only one PPI inhibitor has been approved for clinical use to treat cancers (1), suggesting that the field of PPI inhibitors remains largely unexplored. A variety of proteins and their PPIs have emerged as prospective drug targets to treat tumors because of the extreme heterogeneity and plasticity of cancer (2, 3). Ligands with the potential of binding to a specific site of a target protein with known structure can be efficiently identified by virtual screening. However, the structural plasticity of target proteins usually works against yielding an effective drug candidate. For example, selected compound treatment of cells/organisms often fails to elicit the anticipated effects due to in vivo structural alterations of the target protein caused by various posttranslational modifications (PTMs) and/or unanticipated interactions of the compound and/or its target protein with other molecules (4, 5). Furthermore, many critical proteins regulating various biological processes do not have unique structures as a whole or in some functionally important regions (6, 7). Structures of these intrinsically disordered proteins (IDPs) or IDP regions (IDPRs) are extremely dynamic, depending on the environment, and might change during function (4, 8). Many signaling IDPs/IDPRs undergo characteristic disorder-to-order transitions (DOTs) upon interactions with specific binding partners and/or through PTMs (9, 10). Targeting the IDPs/IDPRs capable of

DOT is generally considered an attractive but challenging task for developing anti-PPI inhibitors. In this regard, a recently identified small-molecule compound, 10058-F4, serves as a pioneering success of anti-PPI inhibitor that binds to an IDPR of c-Myc undergoing a DOT upon binding to its partner Max (11, 12). 10058-F4 was discovered by a random screening using a yeast two-hybrid system (11), followed by experimental identification of its specific binding site (residues 402 to 412 of c-Myc) as an IDPR. Drug leads like 10058-F4 targeting IDPs/IDPRs cannot be found by conventional computational methods that rely on fixed conformations, such as crystallographic structures of target proteins. No computer-aided drug discovery platform is currently available for the systematic exploration of IDPRs as potential drug-target sites (3).

To fill this gap, we developed a novel platform for the discovery of drug leads based on molecular docking and molecular dynamics (MD) simulations of the DOT-associated IDPRs of target proteins. Figure 1A describes this protocol. First, intrinsic disorder pre-dispositions of drug-target proteins are analyzed, and potential disorder-based binding regions that can undergo DOTs are evaluated. A search of the protein structure database [Protein Data Bank (PDB)] is also performed to identify known PPIs and DOTs. Once the potential drug-target sites (DOT-based PPI regions) are determined, the corresponding structures retrieved from the PDB are used for molecular docking with druggable compounds from the ZINC compound library (13). Together with the docking scores, off-target probabilities assessed by the similarity ensemble approach (SEA) (14–16) analysis are also considered for selection of lead compounds from the molecular-docked hit compounds. Last, prospected candidate compounds are suggested via MD simulations that evaluate the mode and efficiency of the compound binding.

The feasibility of the proposed approach was validated in this study by targeting an IDPR of MBD2 that undergoes a DOT upon association with its binding partner p66 α for the integration of the Mi-2/NuRD chromatin remodeling complex (CRC). The integrated Mi-2/NuRD CRC includes one CHD (either CHD3 or CHD4), one HDAC (HDAC1 or HDAC2), two DOC1, three MTA (MTA1, MTA2, and MTA3), six RbAp46/48, two p66 (p66 α or p66 β), and one MBD (MBD2 or MBD3)

¹Department of Life Science and Research Institute for Natural Sciences, Hanyang University, Seoul 04763, Korea. ²Department of Molecular Medicine, Morsani College of Medicine, University of South Florida, Tampa, FL 33612, USA. ³Department of Pathology, Hanyang University College of Medicine, Seoul 04763, Korea. ⁴College of Pharmacy, Chungbuk National University, Cheongju, Chungbuk 28160, Korea. ⁵Boston Children's Hospital/Harvard Medical School, Boston, MA 02115, USA. ⁶Department of Chemistry, University of South Florida, Tampa, FL 33620, USA. ⁷Department of Biotechnology, Konkuk University, Chungju, Chungbuk 27478, Korea. ⁸Institute for Biological Instrumentation of the Russian Academy of Sciences, Pushchino, Moscow Region 142290, Russia.

*These authors contributed equally to this work.

†Present address: Boston Children's Hospital/Harvard Medical School, Boston, MA 02115, USA.

‡Corresponding author. Email: cgkim@hanyang.ac.kr (C.G.K.); wonhs@kku.ac.kr (H.-S.W.); vuffersky@health.usf.edu (V.N.U.)

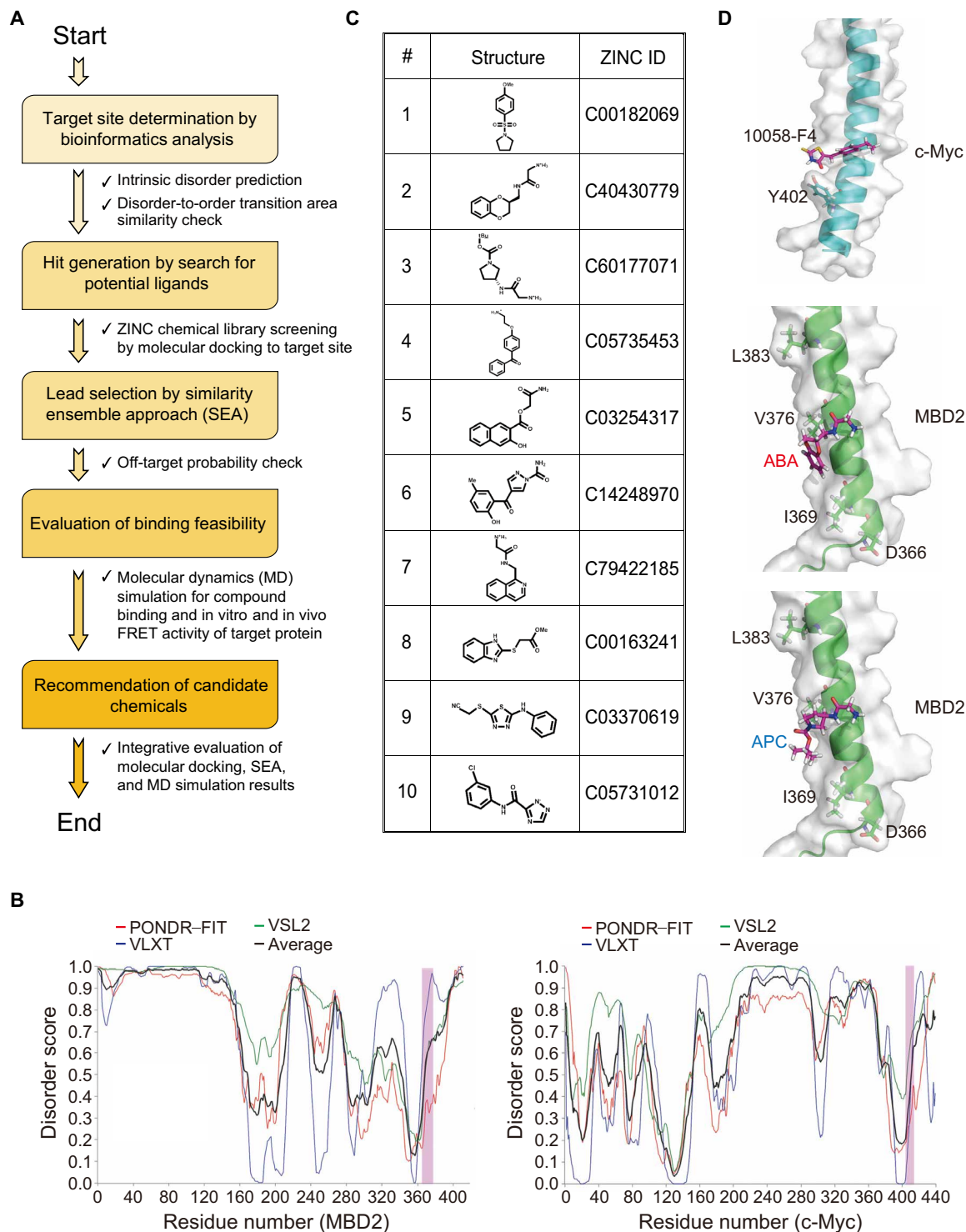


Fig. 1. In silico discovery of the MBD2 IDPR-targeting ligands. (A) Flow chart describing the computational process of ligand discovery. (B) Evaluation of the intrinsic disorder propensity of MBD2 (left) and c-Myc (right); disorder scores 1 and 0 mean fully disordered and fully ordered residues, respectively. Pink bars show positions of the determined DOT sites embedded in residues 360 to 393 for MBD2 and 395 to 430 for c-Myc. (C) Chemical structures of the top 10 compounds showing the most favorable binding to the MBD2 target site in the molecular docking screening of ZINC chemical library. (D) Representative structures of protein-ligand complexes obtained from the molecular docking results (original data file 1 for PDB coordinates): 10058-F4:c-Myc_{C402} (top; control experiment), ABA:MBD2₃₆₉ (middle), and APC:MBD2₃₆₉ (bottom).

molecules (17, 18), where the molecular interaction of MBD2 with p66 α critically mediates the proper assembly of CRC (17, 19). This CRC performs an important epigenetic function in normal development and differentiation by suppressing gene expression by binding directly to the DNA methylation sites and to the DNA methyltransferases (20, 21).

CRC also contributes to the development of human diseases, including cancer (22, 23); for example, the epigenetic regulation by Mi-2/NuRD CRC includes multiple tumor suppressor genes (23, 24), and several CRC components, including MBD2, were also observed to be oncogenic and/or closely correlated with the aggressiveness of several cancers (23, 25, 26). In particular, the function of Mi-2/NuRD CRC is known to be associated with the cellular process of epithelial-mesenchymal transition (EMT; the conversion of adhesive epithelial cells into migratory, invasive mesenchymal cells) that drives wound healing and cell migration and invasion (27, 28). In cancer, EMT necessarily mediates the metastasis of cancers and thus also enables carcinoma cells to acquire cancer stem cell (CSC) properties, malignancy-associated traits, and drug resistance (29–31). Given that the metastasis is responsible for more than 90% of contemporary cancer deaths and yet no marketed antimetastatic drug is currently available (32), developing these drugs to target the cancer spreading is an essential and urgent task for oncological therapy. In this context, functional inhibition of CRC or modulation of its individual components might serve as a novel strategy for effective anticancer therapy to prevent the progression of cancer to metastatic stage. In particular, it has been observed that down-regulation of MBD2 and/or p66 α , which triggered derepression of epithelial regulators via epigenetic reprogramming of the Mi-2/NuRD CRC into the MBD2-free or disengaged CRC, resulted in promoted epithelial differentiation and loss of tumor-initiating ability. Therefore, targeting MBD2 specifically at its IDPR would be a promising approach to the development of antimetastatic agents by inhibiting its DOT-based PPI with p66 α that is essential for the integration of CRC and thus for its critical function in EMT. In addition, no noticeable adverse effects displayed by MBD2 inhibitors can be expected from the fact that down-regulation of MBD2 expression is essential for normal cell differentiation (33), and yet, MBD2 knockout (MBD2^{-/-}) mice exhibit normal survival and reproduction (34).

Hence, in this study, the MBD2 IDPR and its DOT-based interaction with p66 α for the CRC integration were selected as a highly promising target system to evaluate the efficiency of our platform for rational drug discovery. Using this novel approach, we identified two small-molecule compounds capable of inhibiting the PPI of MBD2 and thereby efficiently suppressing the cancer metastatic potentials. In vivo efficacy of both leads in inhibiting cancer metastasis was also evident in a murine xenograft tumor model. Therefore, our novel method renders IDPRs available for rational discovery of anticancer drugs targeting DOT-based PPIs. In particular, the identified compounds provide a basis for the development of previously unidentified inhibitors capable of controlling metastasis of various carcinomas.

RESULTS

In silico analysis for determination of target site and search for potential ligands

As our study was inspired by the discovery of 10058-F4, which binds to the c-Myc IDPR to inhibit its DOT for interaction with Max (11, 12), we compared the PPI site of MBD2 with that of c-Myc using our

computational platform. Sequence analysis (see fig. S1 for sequence and structure information) revealed that disorder profiles of the PPI site of MBD2 (residues 360 to 393 for p66 α interaction) (17, 35) closely resembled that of c-Myc (residues 400 to 434 for Max interaction) (36, 37) (Fig. 1B), characterized by a positive slope in its disorder profile. As both MBD2 and c-Myc are folded in complexes with their cognate partners (p66 α and Max, respectively) (17, 35, 37), this analysis suggests that the PPI sites of MBD2 and c-Myc could undergo a similar type of DOT upon complex formation.

Subsequently, a nuclear magnetic resonance (NMR) ensemble structure of MBD2_{360–393} in its complex with p66 α _{138–178} (PDB ID: 2L2L) was retrieved, and the lowest-energy conformation of the ensemble was extracted for molecular docking analysis using the four residues (D₃₆₆, I₃₆₉, V₃₇₆, and L₃₈₃) of MBD2_{360–393} engaged in the coiled-coil interaction, with p66 α (35) as the initial target site in the molecular docking. From the molecular docking-based virtual screening of 2×10^6 chemical compounds in the ZINC library, 10 promising compounds (compounds #1 to #10 in Fig. 1C) capable of interaction with MBD2 at the designated target site were selected. As a control, the Y₄₀₂-targeted molecular docking of 10058-F4 to c-Myc_{395–430} (Fig. 1D; note that the key residue for the c-Myc interaction with 10058-F4 is Y₄₀₂) (35) was compared with the MBD2_{360–393} docking of the 10 selected hit compounds (table S1). The MBD2₃₆₉-targeted docking of two compounds {compounds #2 and #3 in Fig. 1D named herein ABA [2-amino-N-(2,3-dihydro-benzo[1,4]dioxin-2-ylmethyl)-acetamide] and APC [3-(2-amino-acetylaminopyrrolidine-1-carboxylic acid tert-butyl ester), respectively} was found as the most favorable. In ABA:MBD2₃₆₉ and APC:MBD2₃₆₉ dockings, these compounds formed three intermolecular hydrogen bonds and had relatively low DOCK scores (–35.2 and –33.3 kcal mol⁻¹, respectively) of the DOCK binding. These binding features could be compared favorably with those of the 10058-F4:c-Myc₄₀₂ docking, which showed the DOCK score of –6.77 kcal mol⁻¹ and just one intermolecular hydrogen bond (table S1).

Selection of lead compounds by in silico assessment of off-target probability and activity test in cells

Concerning the potential side effects of the selected hit compounds, their off-target probabilities were assessed by the SEA analysis (14, 16), which has served as an eminent bioinformatics resource aiding in target identification for drug development by profiling multiple protein targets of chemical compounds as probes (15). For this analysis, the c-Myc inhibitor 10058-F4 and two anticancer drugs, imatinib (Gleevec) and sorafenib (Nexavar), were also compared as controls, and 2060 human proteins in the database were searched as potential targets. Given that a significant binding is feasible when both the Max T_c value more than 0.5 and E value lower than 10^{-10} are relevant, no suggestible off-target was predicted for 7 of the 10 hit compounds including both ABA and APC, whereas four proteins were found as the probable 10058-F4 targets (Fig. 2A and table S2). Two of the other compounds also showed a small number of putative off-target proteins (six and two proteins for compounds #4 and #10, respectively), whereas 35 and 26 targets were suggested for imatinib and sorafenib, respectively (fig. S2A and table S2). Therefore, we screened nine compounds with low off-target probability for cellular activity dysregulating MBD2. In particular, the cell migration assay was used for this preliminary test of the compounds on the basis of the previous observation that knockdown of MBD2 in cancer cell lines resulted in decreased migration of the cells. The

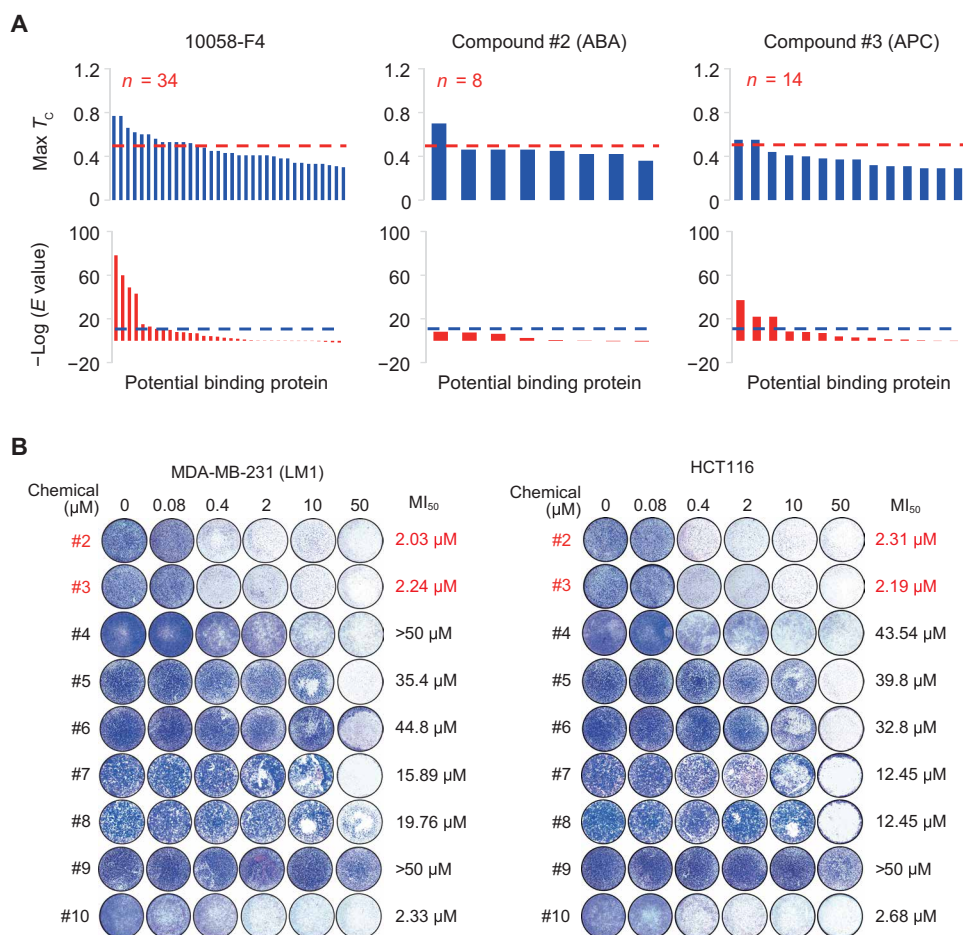


Fig. 2. Lead selection from hit compounds. (A) Computational analysis for off-target probabilities of the 10058-F4 (control experiment) and two selected lead compounds (ABA and APC). Max T_c and E value of the predicted binding are plotted for the n (number of potential targets predicted) off-target candidates yielded from SEA using 2060 human proteins in the database. See fig. S2 for the other hit compounds. (B) Cell migration inhibition by the hit compounds. The LM1 and HCT116 cancer cells were fixed and stained after 48 hours of Transwell migration in the presence of indicated concentrations of individual compounds, followed by counting the number of migrated cells ($n = 2$) to yield MI_{50} value.

result implicated most of the hit compounds in actual suppression of the migration of breast adenocarcinoma MDA-MB-231 (LM1) and colorectal carcinoma HCT116 cells (Fig. 2B and fig. S2B). In particular, ABA (compound #2) and APC (compound #3), which accomplished the most favorable target binding in the aforementioned molecular docking experiments, also showed the least MI_{50} (concentration for half-inhibition of cell migration) values. Therefore, these two molecules were selected as lead compounds for subsequent evaluation in detail.

In silico analysis of target binding for selected lead compounds

To assess target-binding feasibility and mode of binding of the two selected leads, we conducted MD simulation using the structures resulting from the ABA:MBD₂₃₆₉, APC:MBD₂₃₆₉, and 10058-F4:c-Myc₄₀₂ docking (Fig. 1D) as starting points. In 50-ns MD trajectories, the number of the compound-protein contacts (Fig. 3A) and the compound-protein interaction energies (fig. S3A) over time were steady for 10058-F4:c-Myc₄₀₂ but showed noticeable fluctuations for ABA:MBD₂₃₆₉ and APC:MBD₂₃₆₉, particularly during the first half of the simulation period, suggesting that the binding of ABA or APC

to MBD₂₃₆₀₋₃₉₃ might be less persistent than the 10058-F4:c-Myc₃₉₅₋₄₃₀ interaction. However, heatmaps representing intermolecular contacts for individual residues (Fig. 3B) indicated frequent contacts of the ABA/APC-MBD₂₃₆₀₋₃₉₃ interaction comparable to that of the 10058-F4:c-Myc₃₉₅₋₄₃₀ interaction. In particular, the highest contact density value at the most contacted residue (D368 contact) in the ABA:MBD₂₃₆₉ trajectory was higher than that (L404 contact) in the 10058-F4:c-Myc₄₀₂ trajectory, suggesting stronger binding. Next, MD simulations for the ligand:MBD₂₃₆₀₋₃₉₃ complex were extended to include D₃₆₆-, V₃₇₆-, and L₃₈₃-targeted docking (Fig. 3C). Consistent with the ABA:MBD₂₃₆₉ trajectory, D₃₆₈ was the most contacted residue in the heatmaps for heavy atom contacts of the ABA:MBD₂₃₇₆ trajectory, although no preferential contact was found in the other ABA:MBD₂₃₆₀₋₃₉₃ trajectories and in the APC:MBD₂₃₆₀₋₃₉₃ MD simulation sets. Collectively, the MD simulation indicated that the actual binding of ABA and APC to MBD₂₃₆₀₋₃₉₃ would be as promising as the 10058-F4 binding to c-Myc₃₉₅₋₄₃₀, although detailed interaction modes can be different between the two compounds. Therefore, it was subsequently examined whether the targeted binding of the compounds to MBD2 would influence specific PPI of the protein.

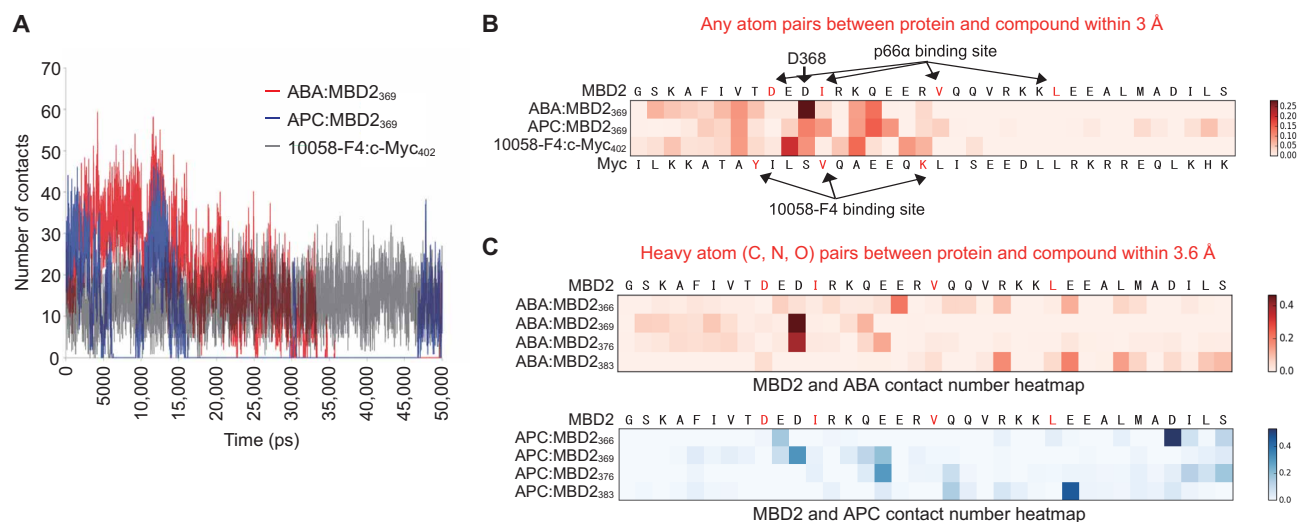


Fig. 3. In silico analysis of the lead compound binding to target site. (A) Time-course alterations of the number of intermolecular contacts within 3 Å cutoff in MD simulations. (B) Heatmap describing the number of simulated compound-protein contacts during 50-ns trajectory for individual residues. Each value of a number of contacts was normalized by dividing it by the total number of contacts in each simulation. The already-known critical residues for PPI are shown in darker red. (C) Heatmap of the intermolecular heavy atom contacts between the lead compounds and target proteins during 50-ns trajectory. Number of contacts was normalized by the total number of contacts in each simulation. MBD2 N-terminal two residues, G and S, were from the NMR structure (PDB ID: 2L2L). MBD2 sequence starts from K360, after G, and S.

Inhibition of the DOT-mediated PPI of MBD2 by the lead compounds

It has been suggested that 10058-F4 evokes a local conformational change (36) or conformational equilibrium shift (38, 39) of the c-Myc IDPR at its binding sites, and this small but significant alteration is critically involved in the functional inhibition of the DOT-mediated PPI of c-Myc with Max. Detailed inspection of the MD simulation results suggested that the MBD2 IDPR could also undergo a local conformational perturbation upon the binding of ABA and APC. For instance, in the ABA:MBD2₃₆₉ and APC:MBD2₃₆₉ trajectories, both Φ and Ψ backbone torsion angles of the most contacted residue (D₃₆₈) in the compound-contacting states were significantly (*t* test, *P* < 0.05) different from those in the noncontacting states (fig. S3B). The compound-bound conformation also appeared to be different between ABA and APC, as the D₃₆₈ Ψ angles in the compound-contacting states were significantly different in between ABA:MBD2₃₆₉ and APC:MBD2₃₆₉ trajectories, although Φ angle differences were not significant (*t* test, *P* = 0.574). Therefore, to further analyze the possible conformational perturbation, we compared the compound-bound ABA:MBD2₃₆₉ and APC:MBD2₃₆₉ trajectories with the apo-MBD2 and p66 α -MBD2 trajectories (fig. S3C). The backbone root mean square fluctuation values of individual residues (fig. S3D) showed that apo-MBD2 underwent stronger backbone fluctuations than compound- or p66 α _{138–178}-bound MBD2_{360–393}. This reflects the structural instability of MBD2_{360–393} in the absence of bound molecules (or, conversely, DOT upon complex formation). Notably, the backbone fluctuation was also different between compound- and p66 α _{138–178}-bound MBD2_{360–393}, especially at the p66 α _{138–178}-contacting D₃₆₆ and I₃₆₉ residues, reflecting the compound-specific local conformational perturbation in MBD2_{360–393}. The presence of this compound-specific perturbation was also obvious from torsion angle distributions of the p66 α _{138–178}-interacting D₃₆₆, I₃₆₉, V₃₇₆, and L₃₈₃ residues (fig. S3E), as the backbone Φ/Ψ torsion angles in both ABA:MBD2₃₆₉ and APC:MBD2₃₆₉ trajectories were

different from those in apo-MBD2 and MBD2-p66 α (tables S3 and S4). In addition, comparison between ABA:MBD2₃₆₉ and APC:MBD2₃₆₉ MD trajectories revealed that the two compounds likely evoked different local conformational changes at the p66 α _{138–178}-interacting residues of MBD2. In particular, significant difference in Ψ of I₃₆₉ and Φ/Ψ of V₃₇₆ and L₃₈₃ (table S4), which is distinguished from the similarity in Φ/Ψ of D₃₆₆ and Φ of I₃₆₉, suggested that I₃₆₉ served as a turning point for the observed torsion angle differences more evident in its C-terminal region from I₃₆₉. Collectively, comparative MD simulations of MBD2_{360–393} in different states (apo-, compound-, and p66 α _{138–178}-bound) suggested the compound-specific induction of local conformational perturbation of MBD2, especially at its p66 α -interacting site, which would most likely interfere with the MBD2-p66 α interaction. Therefore, we next examined whether these leads can actually inhibit the PPI of MBD2, with p66 α both in vitro and in cell, by fluorescence resonance energy transfer (FRET) and co-immunoprecipitation (co-IP) assay.

As the coiled-coil interaction between MBD2 and p66 α occurs in an antiparallel fashion (17), MBD2 was fused with a FRET acceptor protein dTomato at its N terminus, whereas the FRET donor enhanced yellow fluorescent protein (eYFP) was C-terminally fused to p66 α ^{1–206} (33) for in vitro FRET. Unfortunately, the full-length p66 α was not available for the in vitro FRET studies due to the inclusion body formation in the *Escherichia coli* system for the recombinant production. The in vitro FRET result evidenced that both ABA and APC efficiently interfere with the MBD2-p66 α interaction by provoking significant reduction of FRET, which, at 1 to 1.5 equimolar concentrations of the compounds, reached half of the value recorded for the MBD2-p66 α ^{1–206} complex (Fig. 4A and fig. S4A). The FRET analysis in 293T cells by transient cotransfection of eYFP-MBD2 and mCherry-p66 α expression constructs also showed the noticeable FRET reduction, which was dependent on the concentrations of the compounds used for the treatment

(Fig. 4B and fig. S4B). Furthermore, the half maximal inhibitory concentration (IC_{50}) values determined in this in-cell FRET experiments (1.93 and 1.75 μM for ABA and APC, respectively; see Fig. 4B) were in good agreement with the MI_{50} values determined in the migration assay (2.03 and 2.24 μM for ABA and APC, respectively; Fig. 2B). Last, the results of the co-IP assay to capture the endogenous MBD2-p66 α complex corroborated the fact that ABA and APC inhibit the MBD2-p66 α association with the sub-micromolar IC_{50} (Fig. 4C). Therefore, as the interruption of the MBD2-p66 α interaction is anticipated to result in the prevention of the proper assembly of Mi-2/NuRD CRC, we subjected the compounds to an in-depth evaluation of biological activities targeting the function of Mi-2/NuRD CRC in cellular EMT and thereby in cancer metastasis.

Specific action of the two identified leads suppressing metastatic potentials of cancer cells

The cellular EMT process that drives cell migration and invasion is critical not only for wound healing but also for cancer metastasis, including promotion of CSC and drug-resistant properties of cancer cells (29–31). As we have previously observed that the MBD2 and/or p66 α down-regulation in cancer cell lines resulted in the depressed EMT and conversely promoted epithelial differentiation, we reasoned that disrupted PPI between MBD2 and p66 α by the ABA and APC compounds could result in suppression of metastatic potentials of cancer cells by regulating the Mi-2/NuRD CRC-mediated EMT. In agreement with these hypotheses, in mesenchymal type of cancer cells (triple-negative and basal-type breast cancers and aggressive colon cancers) treated with ABA or APC, the increased levels of epithelial markers (CDH1 and CTNNB1) were appreciable, whereas the mesenchymal marker (VIM, SNAIL, SLUG, and CDH2) expressions were suppressed. On the other hand, such an alteration indicative of mesenchymal-epithelial transition (MET) was not apparent in the epithelial cancer cells (luminal breast cancers and less aggressive colon cancer) (Fig. 5, A and B, and fig. S5A). Subsequent analyses confirmed that the compounds suppressed wound healing and migration/invasion abilities of the cancer cells (Fig. 5, C and D, and fig. S5B). In addition, flow cytometric measurements of the cell surface markers CD44 and CD24 indicated that the LM1 cells of the stem-like phenotype (CD44^{hi}/CD24^{lo}) were switched over to the nonstem phenotype (CD44^{lo}/CD24^{lo}) by the compound treatments (Fig. 5E), although the compounds did not induce significant alterations in the proliferation rates and cell cycle progression of the cells tested (Fig. 5, F and G, and fig. S5, C and D). Furthermore, the compound-treated cancer cells showed reduced capability of mammosphere formation (Fig. 5H and fig. S5E), thereby resulting in enhanced susceptibility of the cells to chemotherapeutic drugs including doxorubicin and cisplatin (Fig. 5I and fig. S5F). Last, mRNA sequencing (mRNA-Seq) results showed that global gene expression profiles of the ABA- or APC-treated cells were highly comparable to those of MBD2- or p66 α -knockdown cells but markedly discriminated from the profiles of nontreated wild-type cells (Fig. 5J), supporting no significant off-target effects as initially predicted by SEA (Fig. 2A). Together, these observations established antimetastatic activity of the lead compounds, ABA and APC, by demonstrating that the compounds acted so specifically on the MBD2-p66 α PPI system that the EMT process was efficiently modulated to induce transition of CSC-like cells from a mesenchymal-like state to a bona fide epithelial state.

Verification of antimetastatic efficacy of the two selected lead compounds

Antimetastatic efficacy of the two selected lead compounds in vivo was analyzed using xenograft mice transplanted with the LM1 cells, which were chosen for its potent ability to readily metastasize to lung in mice (40). Here, ABA (10 $\mu\text{g kg}^{-1}$) and APC (20 $\mu\text{g kg}^{-1}$) compounds were administered by intravenous injection six times every 3 days from day 10 after the subcutaneous injection of the green fluorescent protein (GFP)-tagged LM1 cells, followed by sacrifice of the mice (after 4 days of the last administration) for subsequent analysis of tumors (Fig. 6, A and B). Notably, although growth inhibition of original tumor was not significant (Fig. 6, A, C, and D), both ABA and APC compounds exhibited a potent inhibition of the cancer metastasis to lung (represented by the number of nodules developed in lung; Fig. 6C), with no significant effects on body weight of the xenograft mice (Fig. 6B). It was also confirmed by immunohistochemistry that the injected LM1 cells were responsible for the origination of tumor and the metastasized tumor nodules in lung (Fig. 6D). In contrast, histological properties of major organs (Fig. 6E) and complete blood count (CBC) result (Fig. 6F) of the compound-administered mice remained normal. Thus, both ABA and APC appear to be promising antimetastatic agents that are unlikely to cause adverse effects in normal tissues.

DISCUSSION

IDPs/IDPRs are important not only for normal cellular processes but also for the development of various human diseases. In particular, proteins validated as potential drug targets have been increasingly identified to contain IDPRs crucial for PPI mediation. However, the dynamic structure of IDPs/IDPRs limits their use in rational structure-based drug discovery. There are some successful examples of finding of compounds that can bind to and regulate the IDPR-containing proteins (e.g., the c-Myc IDPR-targeting compound 10058-F4). However, most of the current approaches to discover compounds targeting functional IDPR are based on random screening. Meanwhile, because many IDPRs undergo characteristic DOTs upon specific PPIs (9, 10), related structural information can be retrieved from their complexed structures. This, together with the in-depth insights into the compound binding modes (38) and the rapidly accumulating knowledge of the IDPR structural properties (6, 7), suggests the possibility for utilization of the structure-based rational approach as a feasible route for efficient discovery of drug leads targeting specific IDPRs engaged in DOT-based PPIs.

The present novel approach to an antimetastatic agent development provides a prime example of a collaborative work of in silico, in vitro, in cell, and in vivo analyses to discover the drug candidates targeting a pharmacologically important IDPR. In particular, we propose here a three-step computational platform for finding these drug leads. First, IDPRs with DOT potential are selected as potential drug-target sites. We speculate that these regions can be identified based on the characteristic features of their intrinsic disorder predisposition profiles similar to those observed in the known DOT-based PPI regions of MBD2 (residues 360 to 393) and c-Myc (residues 395 to 430) (Fig. 1B). Second, for virtual screening, ordered conformation is taken from the structure of selected IDPR complexed with binding partner. Third, MD simulation is conducted for the selected drug leads targeting IDPRs. Because the structure of target IDPR is dynamic (6, 7) and because the presumably entropy-driven compound

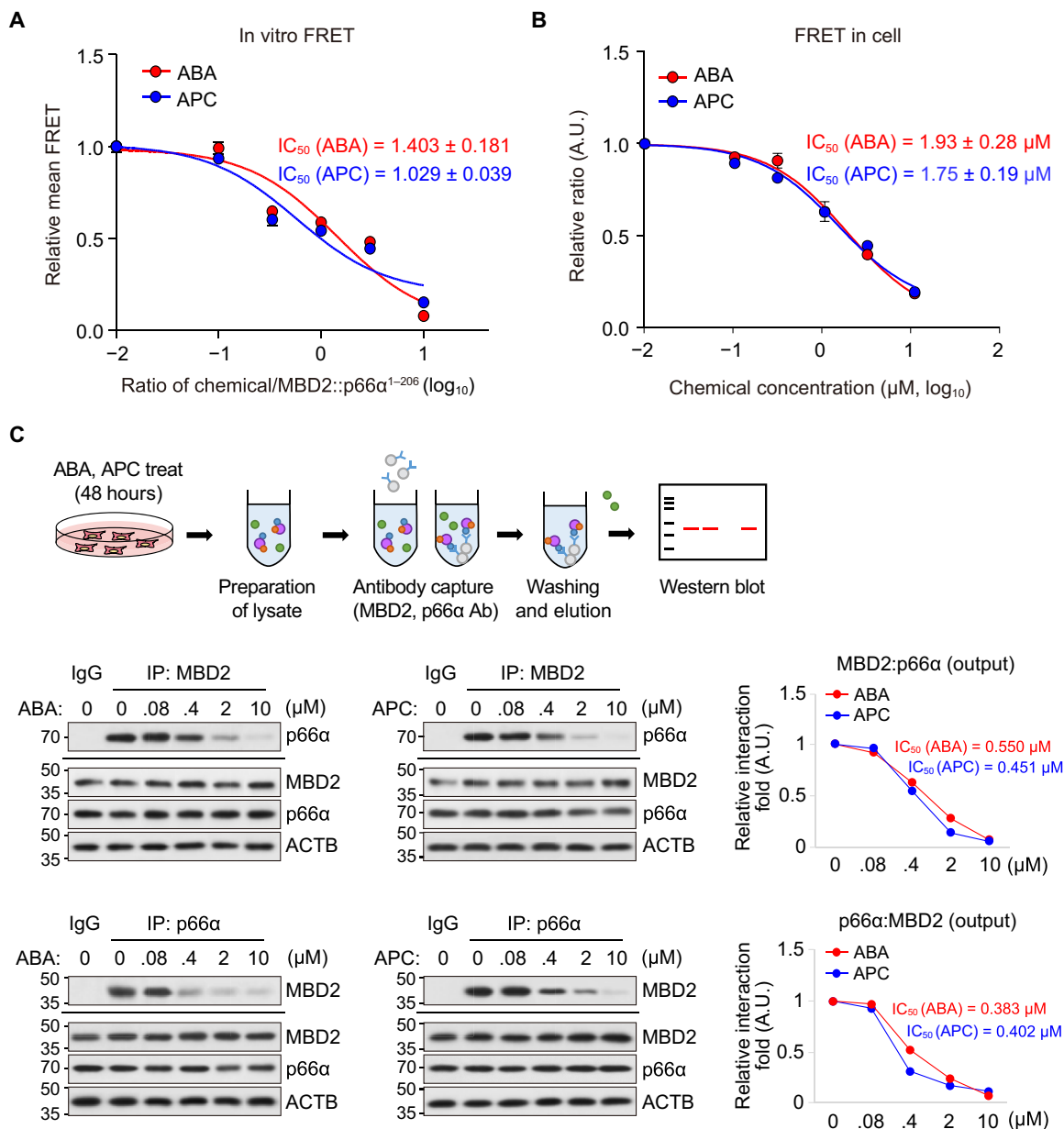


Fig. 4. Two lead compounds efficiently disrupt MBD2 interaction with p66 α in vitro and in cell. (A) Inhibition of in vitro FRET dynamics of MBD2 interaction with p66 α by ABA and APC. Relative mean FRET values for the corresponding ratios of chemical concentration over MBD2::p66 α^{1-206} were plotted. See fig. S4A for the original data. $n = 3$. (B) Inhibition of FRET dynamics of MBD2 interaction with p66 α by ABA and APC in cells. Quantified FRET activities of mock- and compound-treated samples were obtained, and the relative FRET ratios for compounds were calculated by $\text{FRET}_{\text{comp}}/\text{FRET}_{\text{mock}}$ (see Materials and Methods). See also fig. S4B for representative immunofluorescence microscopic photos of cells. $n = 2$. (C) Dose-dependent suppression of the endogenous MBD2-p66 α association by the ABA and APC compounds revealed by in vivo co-IP. Relative fold changes of MBD2 interaction with p66 α (right) were obtained by the quantification of immunoblots (left). Data (means \pm SD) in (A) and (B) were analyzed using Student's t test. Ab, antibody; IgG, immunoglobulin G.

binding also occurs in a dynamic fashion (38), MD simulations of the compound-target complex structures are essential for detailed evaluation of the binding feasibility. In this study, MD simulation indicated the compound binding-specific conformational perturbations of MBD2, particularly at its critical PPI site with p66 α , which could provide a structural basis for the molecular inhibition of the DOT-based PPI of MBD2. In general, specific molecular interactions of IDPs/IDPRs are known to be accomplished in distinctive ways such as DOT, avidity, allovalency, and fuzzy binding; the last

three involves multivalent binding sites, whereas the first represents a simple two-state binding involving a single binding site (41, 42). The present MD simulation result suggests that the ABA and APC binding of the MBD2 IDPR resembled a dynamic, multivalent interaction at low entropic cost, rather than the DOT-based interaction relevant to its p66 α binding. The entropy-driven compound binding and structural multiplicity of the compound-bound IDPR have been identified earlier in the case of 10058-F4 binding to c-Myc₄₀₂₋₄₁₂, which also requires just a few stable atomic interactions

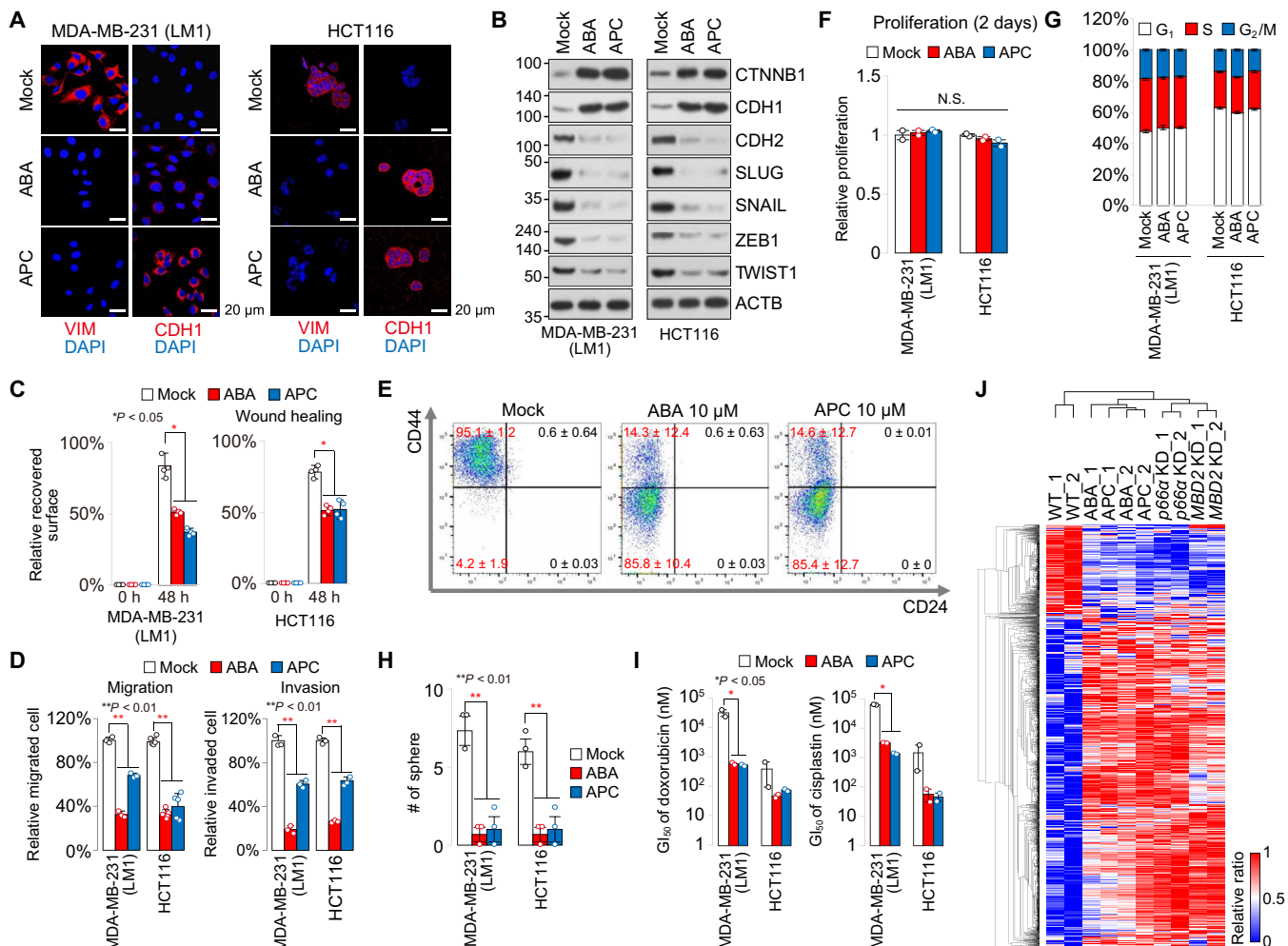


Fig. 5. Induction of a MET on treatment of mesenchymal type of cancer cells with ABA or APC. (A) Representative images showing immunofluorescent signals for VIM or CDH1 (red) and 4',6-diamidino-2-phenylindole (DAPI) (blue) in LM1 (left) and HCT116 (right) cells treated with 10 μ M ABA or APC. Photo credit: S.H.S., Hanyang University. (B) Immunoblots showing the expression levels of EMT markers 48 hours after compound (10 μ M) treatment. ACTB was used as a loading control. A.U., arbitrary units. (C) Effects on wound healing, estimated by the recovered surface areas of scraped cell monolayers, 48 hours after treatment with 10 μ M ABA or APC. $n = 4$. (D) ABA and APC (10 μ M) impact on cell migration (left) and invasion (right) represented by the number of migrated and Matrigel-invaded cells in Transwell plates 48 hours following compound treatments. $n = 3$. (E) Relative proliferation rates quantified by 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay after 2 days. Cells were treated with 10 μ M ABA or APC. $n = 2$. (F) Cell cycle analysis by fluorescence-activated cell sorter (FACS). Cells were treated with 10 μ M ABA or APC. $n = 2$. (G) Number of spheres counted by the naked eye after 5 days. Cells were treated with 10 μ M ABA or APC. $n = 3$. (H) Representative cell population images for the stem-like CD44^{hi} profile of the ABA- or APC-treated LM1 cells analyzed by FACS. Data from one experiment are shown as averages of two technical replicates. (I) Sensitivity to doxorubicin (left) and cisplatin (right) of the 10 μ M ABA- or APC-treated cells quantified by MTT assay. $n = 2$. (J) Heatmap of mRNA-Seq data, which demonstrates similarity in gene expression between ABA- or APC-treated cells and MBD2 or p66 α knockdown LM1 cells. Data (means \pm SD) in (E) to (I) were analyzed using Student's t test. ** $P < 0.01$ and * $P < 0.05$.

(38, 39). In this regard, increased fuzziness of the MBD2 IDPR by the compound binding may conversely lead to decreased propensity for DOT for its p66 α interaction, although the exact mode of binding of our compounds to the MBD2 IDPR, which can ultimately underlie their PPI inhibition mechanism, remains to be characterized in detail.

Our computational platform also contains an additional *in silico* study using the SEA, which was practical to assess off-target probability of the suggested compounds that is potentially associated with adverse effects in actual usage. In subsequent studies, mRNA-Seq results in cells (Fig. 5) were consistent with the SEA result (Fig. 2A) that predicted no significant off-target probability, and *in vivo* administration of the suggested compounds raised no significant toxicity in normal tissues (Fig. 6, E and F).

It is generally appreciated that identifying and understanding molecular regulation and signaling network involved in the EMT process are essential to provide a molecular basis for antimetastatic drug development (43, 44). Concerning this study, we have recently identified the MBD2-p66 α molecular system in Mi-2/NuRD CRC as a promising target for EMT modulation by observing the induction of MET (conversed process of EMT) by knockdown of MBD2 and/or p66 α in cancer cells. Together with this parallel effort, the present discovery of novel antimetastatic agents targeting a component of Mi-2/NuRD CRC validates that this epigenetic machinery can serve as an emerging target system for efficient antimetastatic drug developments. Both ABA and APC disrupting the specific PPI of MBD2 were able to suppress cellular EMT processes, thereby inducing epithelial differentiation of the more aggressive CSCs. Last,

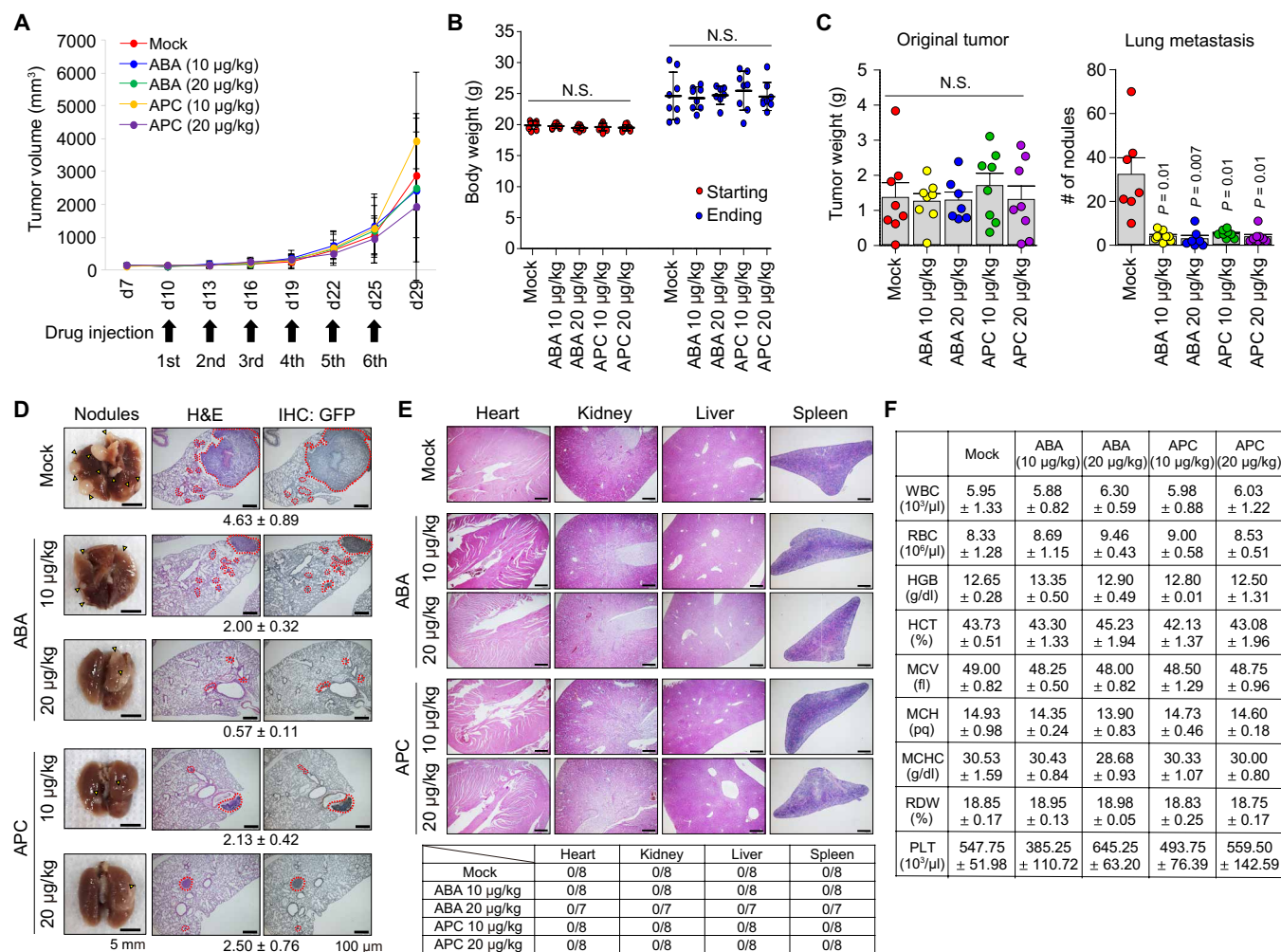


Fig. 6. In vivo antimetastatic efficacy of ABA and APC. (A) Estimated volume (means \pm SEM; *P* value for significance test by ANOVA) of original tumor developed during the experimental period with and without the drug administration. *n* = 8 for each group. (B) Body weights of mice monitored at the starting and ending point of experiment. (C) Effects of the compound administration on the xenograft tumor and its metastasis. Estimated tumor weights are presented for the original tumors, whereas the number of nodules developed by lung metastasis is plotted. (D) Representative photographs for lung nodules acquired 29 days after injection of the LM1 cells. Images of metastasized lung tissue sections illustrated by hematoxylin and eosin (H&E) staining and GFP immunohistochemistry (IHC). Yellow arrowhead represents the tumor nodule, and red dotted area indicates the tumor region. Numbers below the H&E-stained tissue sections indicate the average number of tumor nodules in all mice of the same group. Photo credit: M.Y.K. and S.C., Hanyang University. (E) Representative images of H&E-stained tissue sections for the major organs derived from the xenograft NOD-Prkdc^{scid} IL2rg^{-/-} (NPG) mice after completion of the metastasis inhibition tests with the ABA and APC administration (top). Histological scoring (tumor-bearing mice/total mice) for the H&E-stained major organs of the xenograft mice (bottom). Scale bars, 500 μ m. Photo credit: M.Y.K. and S.C., Hanyang University. (F) CBC analysis of the ABA- and APC-treated xenograft mice. WBC, white blood cell count; RBC, red blood cell count; HGB, hemoglobin; HCT, hematocrit; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; RDW, red cell distribution width; PLT, platelet count; N.S., not significant. Data (means \pm SD) in (B) to (D) and (G) were analyzed using Student's *t* test.

our compounds potently inhibited the cancer metastasis in vivo. Furthermore, considering that they raised no noticeable adverse effects on blood and normal tissues, the present results provide a basis for a novel safe control of cancer metastasis. Hence, found in this study, low-molecular weight (<250 g mol⁻¹) compounds constitute a pioneering example of antimetastatic agents acting on a specific Mi-2/NuRD CRC component. In addition, the present observation that the compound treatments rendered the cancer cells more sensitive to anticancer drugs (Fig. 5I) provides important implications in combination therapy for cancer.

In conclusion, this study successfully used a rational approach to search for the novel antimetastatic agents acting via inhibition of the DOT-based PPI in an IDPR. As IDPs/IDPRs play crucial roles

in diverse cellular processes (6, 7), we believe that this platform can be applied for the discovery of innovative drug leads targeting DOT-based PPI regions in proteins associated with various cancers and other diseases.

MATERIALS AND METHODS

Study design

This study was designed to develop a novel platform for the discovery of drug leads based on molecular docking and MD simulations of the DOT-associated IDPRs of target proteins and, as a proof of concept, to identify candidate drugs, suppressing metastatic potentials of cancer cells in vitro and in vivo, by targeting an IDPR of MBD2

that undergoes a DOT upon association with its binding partner p66 α for the integration of the Mi-2/NuRD CRC. These objectives were addressed by (i) analyzing intrinsic disorder predispositions of drug-target proteins and evaluating potential disorder-based binding regions (45), (ii) doing molecular docking with druggable compounds from the ZINC compound library to the potential drug-target sites, (iii) selecting two lead compounds based on the docking scores and off-target probabilities and experimental validation of target binding, (iv) evaluating the mode and efficiency of the compound binding via MD simulations, (v) assessing the identified leads for biological effects suppressing metastatic potentials of cancer cells, and (vi) verifying antimetastatic efficacy in a murine xenograft tumor model.

In animal studies, mice were randomly assigned to treatment and control groups. Numbers of tested mice were specified in each figure. Outliers were removed only if mice died at an early stage of the treatment according to the Hanyang University Institutional Animal Care and Use Committee (IACUC) dimension guideline. The primary end points were tumor size and cancer metastasis to lung. Mice were euthanized when moribund or at the end of the prespecified treatment period. All procedures were performed in accordance with institutional protocols approved by the IACUC of the Hanyang University. Pathology analysis was performed in a blinded fashion.

Statistical analysis

Data were presented as means \pm SE. The sample size for each experiment, n , was included in Results and the associated figure legend. Everywhere in the text, the difference between two subsets of data was considered statistically significant if the one-tailed Student's t test gave a significance level P (P value) less than 0.05. Multiple comparisons, more than two means, were performed using a univariate analysis of variance (ANOVA), where a Scheffe posttest was performed in some cases or Kruskal-Wallis test. GraphPad Prism was used to generate MI₅₀ curves for cell lines treated with ABA and APC in vitro. In addition, IC₅₀ curves for FRET assay were also generated by GraphPad Prism. Statistical analyses were performed using IBM SPSS statistics 23.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <http://advances.sciencemag.org/cgi/content/full/5/11/eaav9810/DC1>

Supplementary Materials and Methods

Fig. S1. Structural information on MBD2 and c-Myc.

Fig. S2. SEA and cell migration analysis for the nine selected hit compounds targeting MBD2.

Fig. S3. MD simulations of the selected compound-docked structures of MBD2 and c-Myc.

Fig. S4. FRET dynamics of ABA and APC to the MBD2-p66 α interaction.

Fig. S5. Effects of ABA and APC on the expression of EMT markers and CSC properties in various breast and colon cancer cells.

Table S1. Molecular docking result (H-bond, hydrogen bond; N/A, not available).

Table S2. Selection of compound by in silico assessment of off-target probability by SEA analysis.

Table S3. Backbone torsion angle variations (95% confidence interval) of the four key residues in the four different MD simulations of MBD2.

Table S4. T test and P values on the backbone torsion angle summarized in table S3.

Table S5. Primer sets for vector construction.

Original data file S1. Figure 1D PDB files.

References (46–69)

[View/request a protocol for this paper from Bio-protocol.](#)

REFERENCES AND NOTES

- D. R. Green, A BH3 mimetic for killing cancer cells. *Cell* **165**, 1560 (2016).
- A. L. Hopkins, C. R. Groom, The druggable genome. *Nat. Rev. Drug Discov.* **1**, 727–730 (2002).
- E. E. Scott, C. R. Wolf, M. Otyepka, S. C. Humphreys, J. R. Reed, C. J. Henderson, L. A. McLaughlin, M. Paloncova, V. Navratilova, K. Berka, P. Anzenbacher, U. P. Dahal, C. Barnaba, J. A. Brozik, J. P. Jones, D. F. Estrada, J. S. Laurence, J. W. Park, W. L. Backes, The role of protein-protein and protein-membrane interactions on P450 function. *Drug Metab. Dispos.* **44**, 576–590 (2016).
- V. N. Uversky, A. K. Dunker, Understanding protein non-folding. *Biochim. Biophys. Acta* **1804**, 1231–1264 (2010).
- V. N. Uversky, A decade and a half of protein intrinsic disorder: Biology still waits for physics. *Protein Sci.* **22**, 693–724 (2013).
- A. K. Dunker, E. Garner, S. Guilliot, P. Romero, K. Albrecht, J. Hart, Z. Obradovic, C. Kissinger, J. E. Villafranca, Protein disorder and the evolution of molecular recognition: Theory, predictions and observations. *Pac. Symp. Biocomput.* 473–484 (1998).
- V. N. Uversky, Dancing protein clouds: The strange biology and chaotic physics of intrinsically disordered proteins. *J. Biol. Chem.* **291**, 6681–6688 (2016).
- V. N. Uversky, Unusual biophysics of intrinsically disordered proteins. *Biochim. Biophys. Acta* **1834**, 932–951 (2013).
- H. J. Dyson, P. E. Wright, Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **6**, 197–208 (2005).
- V. N. Uversky, Intrinsic disorder-based protein interactions and their modulators. *Curr. Pharm. Des.* **19**, 4191–4213 (2013).
- X. Yin, C. Giap, J. S. Lazo, E. V. Prochownik, Low molecular weight inhibitors of Myc-Max interaction and function. *Oncogene* **22**, 6151–6159 (2003).
- V. N. Uversky, Intrinsically disordered proteins and novel strategies for drug discovery. *Expert Opin. Drug Discov.* **7**, 475–488 (2012).
- J. J. Irwin, T. Sterling, M. M. Mysinger, E. S. Bolstad, R. G. Coleman, ZINC: A free tool to discover chemistry for biology. *J. Chem. Inf. Model.* **52**, 1757–1768 (2012).
- S. Helander, M. Montecchio, R. Pilstål, Y. Su, J. Kuruvilla, M. Elvén, J. M. E. Ziauddin, M. Anandapadamanaban, S. Cristobal, P. Lundström, R. C. Sears, B. Wallner, M. Sunnerhagen, Pre-anchoring of Pin1 to unphosphorylated c-Myc in a fuzzy complex regulates c-Myc activity. *Structure* **23**, 2267–2279 (2015).
- M. J. Keiser, B. L. Roth, B. N. Armbruster, P. Ernsberger, J. J. Irwin, B. K. Shoichet, Relating protein pharmacology by ligand chemistry. *Nat. Biotechnol.* **25**, 197–206 (2007).
- M. J. Keiser, V. Setola, J. J. Irwin, C. Laggner, A. I. Abbas, S. J. Hufeisen, N. H. Jensen, M. B. Kujler, R. C. Matos, T. B. Tran, R. Whaley, R. A. Glennon, J. Hert, K. L. H. Thomas, D. D. Edwards, B. K. Shoichet, B. L. Roth, Predicting new molecular targets for known drugs. *Nature* **462**, 175–181 (2009).
- M. N. Gnanaprasadam, J. N. Scarsdale, M. L. Amaya, H. D. Webb, M. A. Desai, N. M. Walavalkar, S. Z. Wang, S. Zu Zhu, G. D. Ginder, D. C. Williams Jr., p66 α -MBD2 coiled-coil interaction and recruitment of Mi-2 are critical for globin gene silencing by the MBD2-NuRD complex. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 7487–7492 (2011).
- S. L. Kloet, H. I. Baymaz, M. Makowski, V. Groenewold, P. W. T. C. Jansen, M. Berendsen, H. Niaz, G. J. Kops, M. Vermeulen, Towards elucidating the stability, dynamics and architecture of the nucleosome remodeling and deacetylase complex by using quantitative interaction proteomics. *FEBS J.* **282**, 1774–1785 (2015).
- M. A. Desai, H. D. Webb, L. M. Sinanan, J. N. Scarsdale, N. M. Walavalkar, G. D. Ginder, D. C. Williams Jr., An intrinsically disordered region of methyl-CpG binding domain protein 2 (MBD2) recruits the histone deacetylase core of the NuRD complex. *Nucleic Acids Res.* **43**, 3100–3113 (2015).
- R. L. dos Santos, L. Tosti, A. Radziszewska, I. M. Caballero, K. Kaji, B. Hendrich, J. C. Silva, MBD3/NuRD facilitates induction of pluripotency in a context-dependent manner. *Cell Stem Cell* **15**, 102–110 (2014).
- A. Laugesen, K. Helin, Chromatin repressive complexes in stem cells, development, and cancer. *Cell Stem Cell* **14**, 735–751 (2014).
- P. A. Jones, S. B. Baylin, The epigenomics of cancer. *Cell* **128**, 683–692 (2007).
- V. L. Malisetty, V. Penugurti, P. Panta, S. K. Chitta, B. Manavathi, MTA1 expression in human cancers—Clinical and pharmacological significance. *Biomed. Pharmacother.* **95**, 956–964 (2017).
- K. Günther, M. Rust, J. Leers, T. Boettger, M. Scharfe, M. Jarek, M. Bartkuhn, R. Renkawitz, Differential roles for MBD2 and MBD3 at methylated CpG islands, active promoters and binding to exon sequences. *Nucleic Acids Res.* **41**, 3010–3021 (2013).
- L. Xia, W. Huang, M. Bellani, M. M. Seidman, K. Wu, D. Fan, Y. Nie, Y. Cai, Y. W. Zhang, L. R. Yu, H. Li, C. A. Zahnow, W. Xie, R. W. Chiu Yen, F. V. Rassool, S. B. Baylin, CHD4 has oncogenic functions in initiating and maintaining epigenetic suppression of multiple tumor suppressor genes. *Cancer Cell* **31**, 653–668.e7 (2017).
- C. Stirzaker, J. Z. Song, W. Ng, Q. du, N. J. Armstrong, W. J. Locke, A. L. Statham, H. French, R. Pidsley, F. Valdes-Mora, E. Zotenko, S. J. Clark, 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**, 1328–1338 (2017).
- O. Y. Mian, S. Z. Wang, S. Z. Zhu, M. N. Gnanaprasadam, L. Graham, H. D. Bear, G. D. Ginder, Methyl-binding domain protein 2-dependent proliferation and survival of breast cancer cells. *Mol. Cancer Res.* **9**, 1152–1162 (2011).
- J. Fu, L. Qin, T. He, J. Qin, J. Hong, J. Wong, L. Liao, J. Xu, The TWIST/Mi2/NuRD protein complex and its essential role in cancer metastasis. *Cell Res.* **21**, 275–289 (2011).

29. T. Brabletz, R. Kalluri, M. A. Nieto, R. A. Weinberg, EMT in cancer. *Nat. Rev. Cancer* **18**, 128–134 (2018).
30. M. A. Nieto, R. Y. Huang, R. A. Jackson, J. P. Thiery, EMT: 2016. *Cell* **166**, 21–45 (2016).
31. J. Yang, R. A. Weinberg, Epithelial-mesenchymal transition: At the crossroads of development and tumor metastasis. *Dev. Cell* **14**, 818–829 (2008).
32. T. N. Seyfried, L. C. Huysenruyt, On the origin of cancer metastasis. *Crit. Rev. Oncog.* **18**, 43–73 (2013).
33. M. Y. Kim, J. S. Kim, S. H. Son, C. S. Lim, H. Y. Eum, D. H. Ha, M. A. Park, E. J. Baek, B.-Y. Ryu, H. C. Kang, V. N. Uversky, C. G. Kim, Mbd2-CP2c loop drives adult-type globin gene expression and definitive erythropoiesis. *Nucleic Acids Res.* **46**, 4933–4949 (2018).
34. B. Hendrich, J. Guy, B. Ramsahoye, V. A. Wilson, A. Bird, Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development. *Genes Dev.* **15**, 710–723 (2001).
35. N. M. Walavalkar, N. Gordon, D. C. Williams Jr., Unique features of the anti-parallel, heterodimeric coiled-coil interaction between methyl-cytosine binding domain 2 (MBD2) homologues and GATA zinc finger domain containing 2A (GATAD2A/p66 α). *J. Biol. Chem.* **288**, 3419–3427 (2013).
36. A. V. Follis, D. I. Hammoudeh, H. Wang, E. V. Prochownik, S. J. Metallo, Structural rationale for the coupled binding and unfolding of the c-Myc oncoprotein by small molecules. *Chem. Biol.* **15**, 1149–1155 (2008).
37. S. K. Nair, S. K. Burley, X-ray structures of Myc-Max and Mad-Max recognizing DNA. Molecular bases of regulation by proto-oncogenic transcription factors. *Cell* **112**, 193–205 (2003).
38. G. T. Heller, F. A. Aprile, M. Bonomi, C. Camilloni, A. de Simone, M. Vendruscolo, Sequence specificity in the entropy-driven binding of a small molecule and a disordered peptide. *J. Mol. Biol.* **429**, 2772–2779 (2017).
39. G. T. Heller, M. Bonomi, M. Vendruscolo, Structural ensemble modulation upon small-molecule binding to disordered proteins. *J. Mol. Biol.* **430**, 2288–2292 (2018).
40. A. J. Minn, G. P. Gupta, P. M. Siegel, P. D. Bos, W. Shu, D. D. Giri, A. Viale, A. B. Olshen, W. L. Gerald, J. Massagué, Genes that mediate breast cancer metastasis to lung. *Nature* **436**, 518–524 (2005).
41. J. G. Olsen, K. Teilum, B. B. Kragelund, Behaviour of intrinsically disordered proteins in protein-protein complexes with an emphasis on fuzziness. *Cell. Mol. Life Sci.* **74**, 3175–3183 (2017).
42. T. Flock, R. J. Weatheritt, N. S. Latysheva, M. M. Babu, Controlling entropy to tune the functions of intrinsically disordered regions. *Curr. Opin. Struct. Biol.* **26**, 62–72 (2014).
43. C. A. Klein, Parallel progression of primary tumours and metastases. *Nat. Rev. Cancer* **9**, 302–312 (2009).
44. P. S. Steeg, Tumor metastasis: Mechanistic insights and clinical challenges. *Nat. Med.* **12**, 895–904 (2006).
45. E. Anzai, K. Hirata, M. Shibazaki, C. Yamada, M. Morii, T. Honda, N. Yamaguchi, N. Yamaguchi, FOXA1 induces E-cadherin expression at the protein level via suppression of slug in epithelial breast cancer cells. *Biol. Pharm. Bull.* **40**, 1483–1489 (2017).
46. W. L. DeLano, Use of PYMOL as a communications tool for molecular science. *Abstr. Pap. Am. Chem. Soc.* **228**, U313–U314 (2004).
47. F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J. D. Thompson, D. G. Higgins, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
48. B. Xue, R. L. Dunbrack, R. W. Williams, A. K. Dunker, V. N. Uversky, PONDR-FIT: A meta-predictor of intrinsically disordered amino acids. *Biochim. Biophys. Acta* **1804**, 996–1010 (2010).
49. P. Romero, Z. Obradovic, X. Li, E. C. Garner, C. J. Brown, A. K. Dunker, Sequence complexity of disordered protein. *Proteins* **42**, 38–48 (2001).
50. K. Peng, P. Radivojac, S. Vucetic, A. K. Dunker, Z. Obradovic, Length-dependent prediction of protein intrinsic disorder. *BMC Bioinformatics* **7**, 208 (2006).
51. D. M. Lorber, B. K. Shoichet, Hierarchical docking of databases of multiple ligand conformations. *Curr. Top. Med. Chem.* **5**, 739–749 (2005).
52. Y. Chen, B. K. Shoichet, Molecular docking and ligand specificity in fragment-based inhibitor discovery. *Nat. Chem. Biol.* **5**, 358–364 (2009).
53. K. Vanommeslaeghe, A. D. Mackerell Jr., Automation of the CHARMM General Force Field (CGenFF) I: Bond perception and atom typing. *J. Chem. Inf. Model.* **52**, 3144–3154 (2012).
54. K. Vanommeslaeghe, E. P. Raman, A. D. Mackerell Jr., Automation of the CHARMM General Force Field (CGenFF) II: Assignment of bonded parameters and partial atomic charges. *J. Chem. Inf. Model.* **52**, 3155–3168 (2012).
55. H. J. C. Berendsen, D. Vandespoel, R. Vandrunen, Gromacs: A message-passing parallel molecular-dynamics implementation. *Comput. Phys. Commun.* **91**, 43–56 (1995).
56. M. Frigo, S. G. Johnson, The design and implementation of FFTW3. *Proc. IEEE* **93**, 216–231 (2005).
57. R. B. Best, X. Zhu, J. Shim, P. E. Lopes, J. Mittal, M. Feig, A. D. Mackerell Jr., Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ , ψ and side-chain χ_1 and χ_2 dihedral angles. *J. Chem. Theory Comput.* **8**, 3257–3273 (2012).
58. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
59. G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
60. M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
61. S. Nosé, M. L. Klein, Constant pressure molecular-dynamics for molecular-systems. *Mol. Phys.* **50**, 1055–1076 (1983).
62. B. Hess, P-LINCS: A parallel linear constraint solver for molecular simulation. *J. Chem. Theory Comput.* **4**, 116–122 (2008).
63. T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An N.log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
64. U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577–8593 (1995).
65. W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics. *J. Mol. Graph. Model.* **14**, 33–38 (1996).
66. J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, K. Schulten, Scalable molecular dynamics with NAMM. *J. Comput. Chem.* **26**, 1781–1802 (2005).
67. H. C. Kang, J. H. Chae, J. Jeon, W. Kim, D. H. Ha, J. H. Shin, C. G. Kim, C. G. Kim, PIAS1 regulates CP2c localization and active promoter complex formation in erythroid cell-specific α -globin expression. *Nucleic Acids Res.* **38**, 5456–5471 (2010).
68. K. N. Gibson-Corley, A. K. Olivier, D. K. Meyerholz, Principles for valid histopathologic scoring in research. *Vet. Pathol.* **50**, 1007–1015 (2013).
69. J. L. Yap, X. Cao, K. Vanommeslaeghe, K. Y. Jung, C. Peddaboina, P. T. Wilder, A. Nan, A. D. Mackerell Jr., W. R. Smythe, S. Fletcher, Relaxation of the rigid backbone of an oligoamide-foldamer-based α -helix mimetic: Identification of potent Bcl-x_L inhibitors. *Org. Biomol. Chem.* **10**, 2928–2933 (2012).

Acknowledgments: We thank S.-J. Lee, I. Shin, and J.-M. Choi for sharing research materials. We also thank Y. P. Kim and H. Lee for invaluable encouragement for and discussions of the research. **Funding:** This work was supported by the Basic Science Research Program (2014R1A2A1A11054432 and NRF-2017M3A9C8027975 to C.G.K.), the National Research Foundation (NRF), and the Research Fund of Hanyang University (HY-2018 to M.Y.K.), Republic of Korea. Computer time was provided by USF Research Computing, sponsored, in part, by NSF MRI CHE-1531590 to A.v.d.V. The RNA sequencing data have been deposited with the Gene Expression Omnibus (www.ncbi.nlm.nih.gov/geo) under accession no. GSE130637. **Author contributions:** I.N., V.N.U., and C.G.K. designed and supervised the project. I.N. performed molecular docking and MD simulations. M.Y.K., J.S.K., S.H.S., S.C., and S.E.L. performed cellular, molecular, and animal experiments. K.J. evaluated the tissue grades and analyzed the data. G.A., Y.C., A.v.d.V., and J.-H.K. provided database and/or gave critical advice. M.Y.K., I.N., H.-S.W., V.N.U., and C.G.K. analyzed the data and wrote the manuscript. **Competing interests:** I.N., M.Y.K., V.N.U., and C.G.K. are inventors on patent applications related to the method of discovering new drug candidate targeting disorder-to-order transition region and a platform for discovering new drug candidate, filed by Industry-University Foundation, Hanyang University Cooperation (KR patent nos. KR1020160157162, 24 November 2016 and KR1020170157889, 24 November 2017), Industry-University Foundation, Hanyang University Cooperation and University of South Florida (PCT, U.S., EU, and CN patent nos. PCT/KR2017/013489, 24 November 2017; US16463022, 22 May 2019; EP7516668, 28 May 2019; and CN201780072772.8, 24 May 2019). C.G.K., M.Y.K., V.N.U., I.N., Y.C., and A.v.d.V. are inventors on patents and patent applications related to the composition comprising compound inhibiting interactions of MBD2 and p66 α for antimetastasis and prevention and treatment of cancer disease, filed by Industry-University Cooperation Foundation, Hanyang University (KR patent nos. KR102022154, 9 September 2019 and KR102026516, 23 September 2019), Industry-University Foundation, Hanyang University Cooperation, and University of South Florida (U.S. patent no. US 15/821,984, 28 December 2017). All other authors declare that they have no competing interest. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Python 3 codes were designed to improve analysis of molecular docking and molecular dynamics simulation data. Although the codes are not central to the conclusion, codes are accessible from GitHub repository (<https://github.com/InsungNa/MBD2-p66a-Inhibitor-Discovery>). Raw RNA sequencing data are available through the Gene Expression Omnibus (www.ncbi.nlm.nih.gov/geo) under accession no. GSE130637. Additional data related to this paper may be requested from the authors.

Submitted 7 November 2018

Accepted 20 September 2019

Published 20 November 2019

10.1126/sciadv.aav9810

Citation: M. Y. Kim, I. Na, J. S. Kim, S. H. Son, S. Choi, S. E. Lee, J.-H. Kim, K. Jang, G. Alterovitz, Y. Chen, A. van der Vaart, H.-S. Won, V. N. Uversky, C. G. Kim, Rational discovery of antimetastatic agents targeting the intrinsically disordered region of MBD2. *Sci. Adv.* **5**, eaav9810 (2019).