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

Min Young Kim1*, Insung Na2*, Ji Sook Kim1,3, Seung Han Son1, Sungwoo Choi1, Seol Eui Lee1, Ji-Hun Kim4, Kiseok Jang5, Gil Alterovitz6, Yu Chen2, Arjan van der Vaart6, Hyung-Sik Won7*, Vladimir N. Uversky2,8*, Chul Geun Kim1

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 p66α 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 (p66α or p66β), and one MBD (MBD2 or MBD3)
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-Myc402 (top; control experiment), ABA:MBD2369 (middle), and APC:MBD2369 (bottom).
CRC is known to be associated with the cellular process of epithelial-antimetastatic agents by inhibiting its DOT-based PPI with p66α. Tangled CRC, resulted in promoted epithelial differentiation and loss of MBD2 and p66α might serve as a novel strategy for effective anticancer therapy to prevent the progression of CRC or modulation of its individual components 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 p66α, which triggered derepression of epithelial regulators via epigenetic reprogramming of the Mi-2/NuRD CRC into the MBD2-free or disentangled 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 antimitastatic agents by inhibiting its DOT-based PPI with p66α 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 and reproduction.

Hence, in this study, the MBD2 IDPR and its DOT-based interaction with p66α 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 cell lines resulted in decreased migration of the cells. The potential off-target proteins (six and two proteins for compounds #4 and #3, respectively) were found as the probable 10058-F4 targets (Fig. 2A and table S2). Concerning the potential side effects of the selected hit compounds, off-target probability and activity test in cells 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 most favorable. In ABA:MBD2 and APC:MBD2 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-Myc dockering that 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 Tc 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 potential ligands.

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 p66α 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 (p66α 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 and c-Myc was observed (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 and c-Myc engaged in the coiled-coil interaction, with p66α (35) as the initial target site in the molecular docking. From the molecular docking–virtual screening of 2 × 10⁶ 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 MBD2 IDPR and its DOT-based docking of 10058-F4 to c-Myc (36, 37) (Fig. 1D; note that the key residue for the c-Myc interaction with 10058-F4 is Y402) (35) was compared with the MBD2 IDPR docking of the 10 selected hit compounds (table S1). The MBD2 IDPR-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-acetylamino)-pyrrolidine-1-carboxylic acid tert-butyl ester], respectively} was found as the most favorable. In ABA:MBD2 and APC:MBD2 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-Myc dockering that showed the DOCK score of −6.77 kcal mol⁻¹ and just one intermolecular hydrogen bond (table S1).

Kim et al., Sci. Adv. 2019; 5 : eaav9810  20 November 2019
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:MBD2\(_{360–393}\), APC:MBD2\(_{360–393}\), and 10058-F4:c-Myc\(_{402}\) 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-Myc\(_{402}\) but showed noticeable fluctuations for ABA:MBD2\(_{360}\) and APC:MBD2\(_{360}\), particularly during the first half of the simulation period, suggesting that the binding of ABA or APC to MBD2\(_{360–393}\) might be less persistent than the 10058-F4–c-Myc\(_{395–430}\) interaction. However, heatmaps representing intermolecular contacts for individual residues (Fig. 3B) indicated frequent contacts of the ABA/APC–MBD2\(_{360–393}\) interaction comparable to that of the 10058-F4–c-Myc\(_{395–430}\) interaction. In particular, the highest contact density value at the most contacted residue (D368 contact) in the ABA:MBD2\(_{369}\) trajectory was higher than that (L404 contact) in the 10058-F4:c-Myc\(_{402}\) trajectory, suggesting stronger binding. Next, MD simulations for the ligand:MBD2\(_{360–393}\) complex were extended to include D\(_{366}\)-, V\(_{376}\)-, and L\(_{383}\)-targeted docking (Fig. 3C). Consistent with the ABA:MBD2\(_{369}\) trajectory, D\(_{366}\) was the most contacted residue in the heatmaps for heavy atom contacts of the ABA:MBD2\(_{366}\) trajectory, although no preferential contact was found in the other ABA:MBD2\(_{360–393}\) trajectories and in the APC:MBD2\(_{360–393}\) MD simulation sets. Collectively, the MD simulation indicated that the actual binding of ABA and APC to MBD2\(_{360–393}\) would be as promising as the 10058-F4 binding to c-Myc\(_{395–430}\), 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. 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 \(\tau_c\) 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 \(\text{MI}_{50}\) value.
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 (D369) 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 p66α-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 p66α138–178-bound MBD2360–393. This reflects the structural instability of MBD2360–393 in the absence of bound molecules (or, conversely, DOT upon complex formation). Notably, the backbone fluctuation was also different between compound- and p66α138–178-bound MBD2360–393, especially at the p66α138–178-contacting D366 and I369 residues, reflecting the compound-specific local conformational perturbation in MBD2360–393. The presence of this compound-specific perturbation was also obvious from torsion angle distributions of the p66α138–178-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-p66α (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 p66α138–178-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 MBD2360–393 in different states (apo-, compound-, and p66α138–178-bound) suggested the compound-specific induction of local conformational perturbation of MBD2, especially at its p66α-interacting site, which would most likely interfere with the MBD2-p66α interaction. Therefore, we next examined whether these leads can actually inhibit the PPI of MBD2, with p66α 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 p66α 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 p66α1–206 (33) for in vitro FRET. Unfortunately, the full-length p66α 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-p66α 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-p66α1–206 complex (Fig. 4A and fig. S4A). The FRET analysis in 293T cells by transient cotransfection of eYFP-MBD2 and mCherry-p66α expression constructs also showed the noticeable FRET reduction, which was dependent on the concentrations of the compounds used for the treatment
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 μg kg⁻¹) and APC (20 μg 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 antimitastatic 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 antimitastatic 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 MB22 (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
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 p66α, 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 p66α 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.
increase in normal tissues (Fig. 6, E and F).

Administration of the suggested compounds raised no significant tox-
icity in actual usage. In subsequent studies, mRNA-Seq
 results in cells (Fig. 5J) were consistent with the SEA result (Fig. 2A)
that predicted no significant off-target probability, and in vivo ad-
verse effects in actual usage. In subsequent studies, mRNA-Seq

study using the SEA, which was practical to assess off-target proba-

duced in detail.

Our computational platform also contains an additional in silico
study using the SEA, which was practical to assess off-target proba-

between ABA- or APC-treated cells and MBD2 or p66

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 antime-
tastatic drug development (43, 44). Concerning this study, we have recently
identified the MBD2-p66α molecular system in Mi-2/NuRD CRC
as a promising target for EMT modulation by observing the induc-
tion of MET (conversed process of EMT) by knockdown of MBD2
and/or p66α in cancer cells. Together with this parallel effort, the
present discovery of novel antimeastatic agents targeting a com-
ponent of Mi-2/NuRD CRC validates that this epigenetic machin-
ery 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 in-
ducing epithelial differentiation of the more aggressive CSCs. Last,

(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 p66α interaction, although the exact mode of bind-
ing of our compounds to the MBD2 IDPR, which can ultimately
underlie their PPI inhibition mechanism, remains to be character-
ized in detail.

Our computational platform also contains an additional in silico
study using the SEA, which was practical to assess off-target proba-
bility of the suggested compounds that is potentially associated with
adverse effects in actual usage. In subsequent studies, mRNA-Seq
results in cells (Fig. 5I) were consistent with the SEA result (Fig. 2A)
that predicted no significant off-target probability, and in vivo ad-
inistration of the suggested compounds raised no significant tox-
icity in normal tissues (Fig. 6, E and F).
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 p66α 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 antinmetastatic 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 ± 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 \( P (P \text{ value}) \) 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 MI₃₀ 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

**REFERENCES AND NOTES**

1. D. R. Green, A BH3 mimic for killing cancer cells. Cell 165, 1560 (2016).
2. A. L. Hopkins, C. R. Groom, The druggable genome. Nat. Rev. Drug Discov. 1, 727–730 (2002).
3. 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).
4. V. N. Uversky, A. K. Dunker, Understanding protein non-folding. Biochim. Biophys. Acta 1804, 1231–1264 (2010).
5. V. N. Uversky, A decade and a half of protein intrinsic disorder: Biology still waits for physics. Protein Sci. 22, 693–724 (2013).
6. A. K. Dunker, E. Garner, S. Guilliot, P. Romero, K. Albrecht, J. Hart, Z. Obadovics, K. Kissinger, J. E. Villafranca, Protein disorder and the evolution of molecular recognition: Theory, predictions and observations. Pac. Symp. Biocomput. 473–484 (1998).
7. V. N. Uversky, Dancing protein clouds: The strange biology and chaotic physics of intrinsically disordered proteins. J. Biol. Chem. 291, 6681–6688 (2016).
8. V. N. Uversky, Unusual biophysics of intrinsically disordered proteins. Biochem. Biophys. Acta 1834, 932–951 (2013).
9. H. J. Dyson, P. E. Wright, Intrinsically unstructured proteins and their functions. Nat. Rev. Mol. Cell Biol. 6, 197–208 (2005).
10. V. N. Uversky, Intrinsic disorder-based protein interactions and their modulators. Curr. Pharm. Des. 19, 4191–4213 (2013).
11. 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).
12. V. N. Uversky, Intrinsically disordered proteins and novel strategies for drug discovery. Expert Opin. Drug Discov. 7, 475–488 (2012).
13. J. J. Irwin, T. Sterling, M. M. Myssinger, E. S. Bolstad, R. G. Coleman, ZINC: A free tool to discover chemistry for biology. J. Chem. Inf. Model. 52, 1757–1768 (2012).
14. S. Haveland, M. Montecchio, R. Příští, 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).
15. 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).
16. M. J. Keiser, V. Setola, J. J. Irwin, C. Laggner, A. I. Abbas, S. J. Hufniesen, N. H. Jensen, M. B. Kuiper, R. C. Matos, T. B. Tran, R. Whaley, R. A. Glennon, J. Hinthorn, 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).
17. M. N. Graingargam, J. N. Scarsdale, M. L. Amaya, H. D. Webb, M. A. Desai, N. M. Walvalker, S. Z. Wang, S. Z. Zhu, G. D. Ginder, D. C. Williams Jr., p66α-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).
18. S. L. Kloet, H. I. Baymaz, M. Makowski, V. Groenewold, P. W. T. C. Jansen, M. Berendsen, H. Niazi, 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).
19. M. A. Desai, H. D. Webb, L. M. Sinanan, J. N. Scarsdale, N. M. Walvalker, 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).
20. R. L. dos Santos, L. Tosti, A. Radziszewskaya, 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).
21. A. Laugesen, K. Helin, Chromatin repressive complexes in stem cells, development, and cancer. Cell Stem Cell 14, 735–751 (2014).
22. P. A. Jones, S. B. Baylin, The epigenomics of cancer. Cell 128, 683–692 (2007).
23. V. L. Malisetty, V. Penuguri, P. Panta, S. K. Chitta, B. Manavath, MTA1 expression in human cancers—Clinical and pharmacological significance. Biomed. Pharmacother. 95, 956–964 (2017).
24. K. Günter, M. Rust, J. Leers, T. Boettger, M. Scharfe, M. Bartkuhn, R. Renkawitz, The role of protein-protein and protein-membrane interactions on P450 function. Drug Metab. Dispos. 41, 3010–3021 (2013).
25. 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).
26. C. Stirzaker, J. Z. Song, W. Ng, Q. du, N. J. Armstrong, W. J. Locke, A. L. Statham, H. French, R. Pidsey, F. Valdes-Mora, E. Zotenko, S. J. Clark, Methyl-CpG-binding protein MBBD2 plays a key role in maintenance and spread of DNA methylation at CpG islands and shores in cancer. Oncogene 36, 1328–1338 (2017).
27. O. Y. Maan, S. Z. Wang, S. Z. Zhu, M. N. Graingargam, 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).
28. J. Fu, L. Qin, T. He, J. Qin, H. Hong, L. Wang, J. Xu, The TWIST/Mi2/NuRD protein complex and its essential role in cancer metastasis. Cell Res. 21, 275–289 (2011).
58. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison

32. T. N. Seyfried, L. C. Huysentruyt, On the origin of cancer metastasis.

5

Sci. Adv.
etal

29. T. Brabletz, R. Kalluri, M. A. Nieto, R. A. Weinberg, EMT in cancer.

35. N. M. Walavalkar, N. Gordon, D. C. Williams Jr., Unique features of the anti-parallel,

38. G. T. Heller, F. A. Aprile, M. Bonomi, C. Camilloni, A. de Simone, M. Vendruscolo,

39. G. T. Heller, M. Bonomi, M. Vendruscolo, Structural ensemble modulation upon

40. A. J. Minn, G. P. Gupta, P. M. Siegel, P. D. Bos, W. Shu, D. D. Giri, A. Viale, A. B. Olshen,

43. C. A. Klein, Parallel progression of primary tumours and metastases.

45. E. Anzai, K. Hirata, M. Shibazaki, C. Yamada, M. Morii, T. Honda, N. Yamaguchi,

288, 3419–3427 (2013).

46. W. L. DeLano, Use of PYMOL as a communications tool for molecular science.

50. K. Peng, P. Radivojac, S. Vucetic, A. K. Dunker, Z. Obradovic, Length-dependent prediction

53. K. Vanommeslaeghe, A. D. MacKerell Jr., Automation of the CHARMM General Force Field (CGenFF) II: Assignment of bonded parameters and partial atomic

molecular-dynamics implementation.

91

charges.

General Force Field targeting improved sampling

of the additive CHARMM all-atom protein force field.

56. M. Frigo, S. G. Johnson, The design and implementation of FFTW3. Proc. IEEE 93, 216–231 (2005).

59. G. Busi, D. Donadio, M. Parrinello, Canonical sampling: A next molecular dynamics method. J. Appl. Phys. 81, 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.

62. B. Hess, P-LINCS: A parallel linear constraint solver for molecular simulation.

63. T. Darden, York, L. Pedersen, Particle mesh Ewald: An n·log(N) method for Ewald sums in large systems. J. Chem. Phys. 100, 1089–1099 (1993).

64. U. Esmann, 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. 14, 33–38 (1996).

66. J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kelve, K. Schulten, Scalable molecular dynamics with NAMD. 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, PIA51 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. W. Tilder, A. Nan, A. D. MacKerell Jr, W. R. Smythe, S. Fletcher, Relaxation of the rigid backbone of Bcl-xL 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 (2014R1A2A1A01054452 and NRF-2017M3A9B2079775 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. Authority 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.E., and S.E.L. performed cellular, molecular, and animal experiments. K.J. evaluated the tissue grades and analyzed the data. G.A., Y.C., 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 CN2017180072772, 28 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 p66α 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-p66α-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, H. S. 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).