



Article

Drug Discovery Targeting the Disorder-To-Order Transition Regions through the Conformational Diversity Mimicking and Statistical Analysis

Insung Na ¹, Sungwoo Choi ¹, Seung Han Son ¹, Vladimir N. Uversky ^{2,3,*} and Chul Geun Kim ^{1,4,*}

¹ Department of Life Science and Research Institute for Natural Sciences, Hanyang University, Seoul 04763, Korea; nainsung@hanyang.ac.kr (I.N.); cswya@naver.com (S.C.); imsangok12@naver.com (S.H.S.)

² Department of Molecular Medicine, Morsani College of Medicine, University of South Florida, Tampa, FL 33612, USA

³ Institute for Biological Instrumentation of the Russian Academy of Sciences, 142290 Pushchino, Russia

⁴ CGK Biopharma Co. Ltd., 222 Wangshipri-ro, Sungdong-gu, Seoul 04763, Korea

* Correspondence: vuffersky@usf.edu (V.N.U.); cgkim@hanyang.ac.kr (C.G.K.)

Received: 9 June 2020; Accepted: 21 July 2020; Published: 24 July 2020



Abstract: Intrinsically disordered proteins exist as highly dynamic conformational ensembles of diverse forms. However, the majority of virtual screening only focuses on proteins with defined structures. This means that computer-aided drug discovery is restricted. As a breakthrough, understanding the structural characteristics of intrinsically disordered proteins and its application can open the gate for unrestricted drug discovery. First, we segmented the target disorder-to-order transition region into a series of overlapping 20-amino-acid-long peptides. Folding prediction generated diverse conformations of these peptides. Next, we applied molecular docking, new evaluation score function, and statistical analysis. This approach successfully distinguished known compounds and their corresponding binding regions. Especially, Myc proto-oncogene protein (MYC) inhibitor 10058F4 was well distinguished from others of the chemical compound library. We also studied differences between the two Methyl-CpG-binding domain protein 2 (MBD2) inhibitors (ABA (2-amino-N-[[[(3S)-2,3-dihydro-1,4-benzodioxin-3-yl]methyl]-acetamide) and APC ((R)-(3-(2-Amino-acetylamino)-pyrrolidine-1-carboxylic acid tert-butyl ester))). Both compounds bind MBD2 through electrostatic interaction behind its p66 α -binding site. ABA is also able to bind p66 α through electrostatic interaction behind its MBD2-binding site while APC-p66 α binding was nonspecific. Therefore, structural heterogeneity mimicking of the disorder-to-order transition region at the peptide level and utilization of the new docking score function represent a useful approach that can efficiently discriminate compounds for expanded virtual screening toward intrinsically disordered proteins.

Keywords: disorder-to-order transition; computer-aided drug discovery; protein–protein interaction inhibitor; conformational diversity mimicking; peptide docking

1. Introduction

Currently, computer-aided drug discovery represents a usual initiation of the drug development pathway. Among the computer-aided drug discovery approaches, computational hit screening is known as “virtual screening”, which uses ligand- or structure-based strategies and which is complementary to the traditional “high-throughput screening” methods utilized in wet-lab settings. Molecular docking is one of the representative techniques for virtual screening, and machine learning

and deep learning techniques have become powerful tools nowadays [1]. Virtual screening depends on the knowledge of the structure of target proteins determined by X-ray crystallography, solution or solid-state NMR (Nuclear Magnetic Resonance), cryo-EM (Electron Microscopy), and sometimes homology modeling. This means that virtual screening is limited to the set of proteins with known structures, although disease does not always occur due to some problems pertaining to the ordered proteins. Furthermore, blocking such proteins is not always a perfect solution [2,3]. From this viewpoint, the major field of current computer-aided drug discovery process, virtual screening, is restricted to searches within a very limited space at the bottom of the funnel.

Intrinsically disordered proteins (IDPs) do not have defined structures; instead, they exist as highly dynamic conformational ensembles that rapidly change in response to subtle changes in the environment [4,5]. Not only can entire proteins be disordered but also many proteins represent hybrid entities containing ordered domains and intrinsically disordered protein regions (IDPRs) [6]. Some IDPRs function as binding sites for their corresponding partners through disorder-to-order transition, i.e., they utilize the binding-induced folding mechanism [7]. Such binding-induced disorder-to-order transition is characterized by structural transformation of a mostly unfolded unbound state into a specifically folded bound form. Such regions are commonly involved in molecular recognition, and therefore, they are known as molecular recognition features (MoRFs) [8,9]. Although such foldable intrinsically disordered recognition regions represent an attractive target for the discovery of protein–protein interaction inhibitors, for a long time, there were no drugs targeting disorder-based binding sites and there were reliable approaches for the utilization of these specific characteristics (binding-induced folding) in the drug discovery. However, the potential of this shift from ordered to disordered binding sites for the selection of promising drug targets was recently demonstrated by the traditional high-throughput screening-based discovery of the Myc proto-oncogene protein (MYC) inhibitors, including 10058F4 [10,11]. Furthermore, we recently elaborated an approach for the rational discovery of the inhibitors targeting such disorder-to-order transition regions and used this approach to find inhibitors targeting Methyl-CpG-binding domain protein 2 (MBD2) [12].

Earlier, inhibitors targeting disorder-to-order transition region of MYC involved in interaction with Protein Max (MAX) to form a tweezer-like coiled-coil dimer interacting with DNA were discovered using the yeast two-hybrid experiments [10]. Some compounds identified in this study were shown to possess antitumor efficacy *in vitro* and *in vivo* [10]. Later, two members of the library of these compounds, 10058F4 and 10074G5, targeting the MYC binding region were studied by a set of biophysical approaches [11]. The experiment employed mutagenesis and truncations of the MYC disorder-to-order transition region of MYC involved in interaction with MAX [13] and determined that 10058F4 binds to the region 402–412 of MYC whereas residues 363–381 were described as the 10074G5-binding region [11]. A recent comprehensive NMR analysis revealed the presence of subregions of a residual secondary structure within the MYC disorder-to-order transition region and structurally characterized regions affected by the 10058F4 binding [14]. This analysis revealed that the MYC 402–412 region serves as the 10058F4 target site and is embedded within the MYC disorder-to-order transition region responsible for MAX binding [14]. Curiously, according to our previous study, in the MYC per-residue intrinsic disorder profile, this 402–412 segment corresponds to a region with a high positive slope [12].

We also reported the discovery of two inhibitors targeting the disorder-to-order transition regions of MBD2 involved in an interaction with p66 α [12]. We showed that the MBD2 segment that undergoes binding-induced disorder-to-order transition and forms coiled-coil dimer with its partner, p66 α [15] (i.e., acts similarly to the MYC interacting with MAX [13]), is located within the region with a positive slope in the per-residue disorder profile of MBD2 [12]. By molecular docking utilized for virtual screening toward the coiled-coil disorder-to-order transition region of MBD2, we obtained two positive hits, ABA (2-amino-N-[[[(3S)-2,3-dihydro-1,4-benzodioxin-3-yl]methyl]acetamide) and APC (R)-(3-(2-Amino-acetyl-amino)-pyrrolidine-1-carboxylic acid tert-butyl ester). These compounds were shown to possess noticeable anti-metastasis activity due to the inhibition of the MBD2–p66 α complex

formation [12]. However, we did not analyze whether these two compounds can distinguish MBD2 and p66 α (which has a disordered region that undergoes disorder-to-order transition while forming the coiled-coil heterodimer with MBD2).

Here, we developed a new method for escaping the bottom of the funnel that limits the structure-defined protein virtual screening preferentially to ordered proteins. We overcame this limit through segmentation of the target disorder-to-order transition region to a set of overlapping peptides, analysis of the structural preferences of these peptides, and elaboration of a new function for evaluation of the molecular docking score. We used this approach to study the binding mechanism of the MYC inhibitor (10058F4) toward its binding site in MYC. Besides, we also proved that the MBD2 inhibitors (ABA and APC) have different binding preferences toward MBD2 and p66 α . We think that our novel method represents an important breakthrough that will expand the computer-aided drug discovery field, making possible the identification of inhibitors for proteins with unknown structures.

2. Results and Discussion

2.1. Computational Reevaluation of the 10058F4 Binding to the MYC Disorder-To-Order Transition Region

Earlier, we discovered the anti-metastasis activities of two compounds, ABA and APC (see the Section 3 Materials and Methods) by targeting MBD2 in molecular docking environment [12]. In the field of rational drug design, it is a commonly accepted practice to apply molecular docking toward target protein structures determined by X-ray crystallography, NMR, and cryo-EM. However, application of this powerful approach to intrinsically disordered proteins (IDPs) or intrinsically disordered protein regions (IDPRs) is challenging. In our previous study, we used a modified rational drug design approach for the discovery of drug leads based on molecular docking and molecular dynamics (MD) simulations of the IDPRs of target proteins capable of the disorder-to-order transitions (DOTs) [12]. The process started with the analysis of the intrinsic disorder predisposition of a drug target protein of interest followed by the prediction of the presence of potential disorder-based binding regions that can undergo DOTs. In parallel, a Protein Data Bank (PDB) search was conducted to find potential drug-target sites (DOT-based protein–protein interaction regions), and the corresponding structures were retrieved and used for molecular docking. This was combined with the evaluation of the off-target probabilities for the selection of lead compounds from the molecular-docked hit compounds. Finally, MD simulations [12] were used to evaluate the mode and efficiency of the prospected candidate compound binding. Although this approach targeted IDPRs (more precisely IDPRs capable of DOTs), by virtue of utilizing known structure of a target region in its bound state as a starting point for drug discovery, it used logistics of a conventional rational drug design.

In this study, we further developed this approach and moved it closer to the real-life settings by mimicking disorder-to-order transitions. Instead of considering a bound form of a long IDPR that underwent binding-induced folding and gained a specific structure, we focused on the intrinsic disorder nature of this region. To this end, we split each long IDPR of interest to a set of overlapping short peptides, used folding prediction algorithm to generate diverse conformational forms of these peptides, and then subjected these diverse conformations of model peptides to the molecular docking.

In more detail, the target regions used in this study were determined based on the structure for the MYC–MAX (PDB ID 1NKP) [13] complex by the X-ray crystallography. We considered the target DOT of MYC (residues 350–439) and segmented it into the sets of the overlapping 20-amino-acid-long sequences (Table 1). Next, each of the peptides was subjected to structure prediction using the PEPFOLD ver 3.5 web server (<https://mobyli.rpbs.univ-paris-diderot.fr/cgi-bin/portal.py#forms::PEP-FOLD3>) to predict its potential structures [16]. This stage generated a multitude of probable structures, which are exemplified by a set of conformations predicted for the 15 peptides overlapping derived from the MYC 350–439 region. Figure 1 clearly shows that different peptides possess very different folding potentials.

Table 1. Protein target regions and segmented peptides.

Protein	Target Region	Peptides	Sequence (N'–C')
MYC	350–439	MYC_350–369	TEENVKRRRTHNVLERQRRNE
		MYC_355–374	KRRTHNVLERQRRNELKRFS
		MYC_360–379	NVLERQRRNELKRSFFALRD
		MYC_365–384	QRRNELKRSFFALRDQIPEL
		MYC_370–389	LKRSFFALRDQIPELENNEK
		MYC_375–394	FALRDQIPELENNEKAPKVV
		MYC_380–399	QIPELENNEKAPKVVILKKA
		MYC_385–404	ENNEKAPKVVILKKATAYIL
		MYC_390–409	APKVVILKKATAYILSVQAE
		MYC_395–414	ILKKATAYILSVQAEEQKLI
		MYC_400–419	TAYILSVQAEEQKLISEEDL
		MYC_405–424	SVQAEEQKLISEEDLLRKR
		MYC_410–429	EQLISEEDLLRKRREQLKH
		MYC_415–434	SEEDLLRKRREQLKHKLEQL
MYC_420–439	LRKRREQLKHKLEQLRNCSA		
MBD2	359–393	MBD2_359–378	CKAFIVTDEDIRKQEERVQQ
		MBD2_364–383	VTDEDIRKQEERVQQVRKKL
		MBD2_369–388	IRKQEERVQQVRKKLEEALM
		MBD2_374–393	ERVQQVRKKLEEALMADILS
p66 α	136–175	p66 α _136–155	SSPEERERMILKQKLEELRLE
		p66 α _141–160	RERMILKQKLEELRLEEAKLV
		p66 α _146–165	KQKLEELRLEEAKLVLLKKL
		p66 α _151–170	ELRLEEAKLVLLKCLRQSQI
		p66 α _156–175	EAKLVLLKCLRQSQIQKEAT

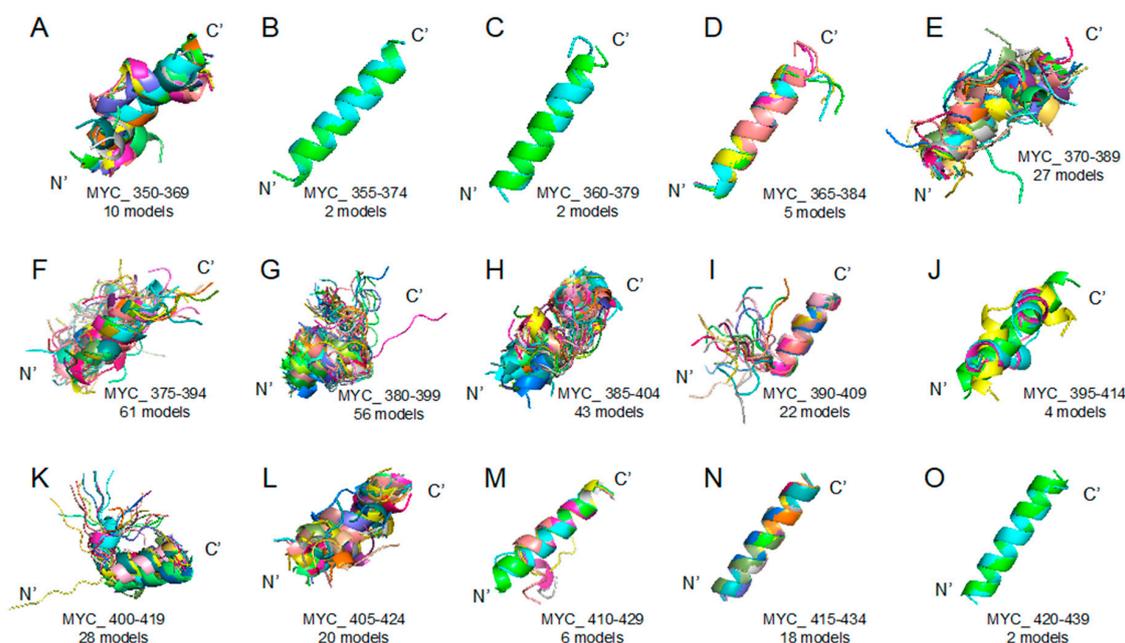


Figure 1. Myc proto-oncogene protein (MYC) 350–439 region segmented peptides: PEPFOLD ver 3.5 prediction result (see the Section 3 Materials and Methods) of peptide structures based on corresponding sequences (please look at the Table 1). (A) MYC_{350–369} (10 models), (B) MYC_{355–374} (2 models), (C) MYC_{360–379} (2 models), (D) MYC_{365–384} (5 models), (E) MYC_{370–389} (27 models), (F) MYC_{375–394} (61 models), (G) MYC_{380–399} (56 models), (H) MYC_{385–404} (43 models), (I) MYC_{390–409} (22 models), (J) MYC_{395–414} (4 models), (K) MYC_{400–419} (28 models), (L) MYC_{405–424} (20 models), (M) MYC_{410–429} (6 models), (N) MYC_{415–434} (18 models), (O) MYC_{420–439} (2 models).

Importantly, we designed a new evaluation score function focusing on the disorder-to-order transition region as described in the Section 3 Materials and Methods. The procedure of using such a segmented peptide for docking is summarized in Figure 2A,B.

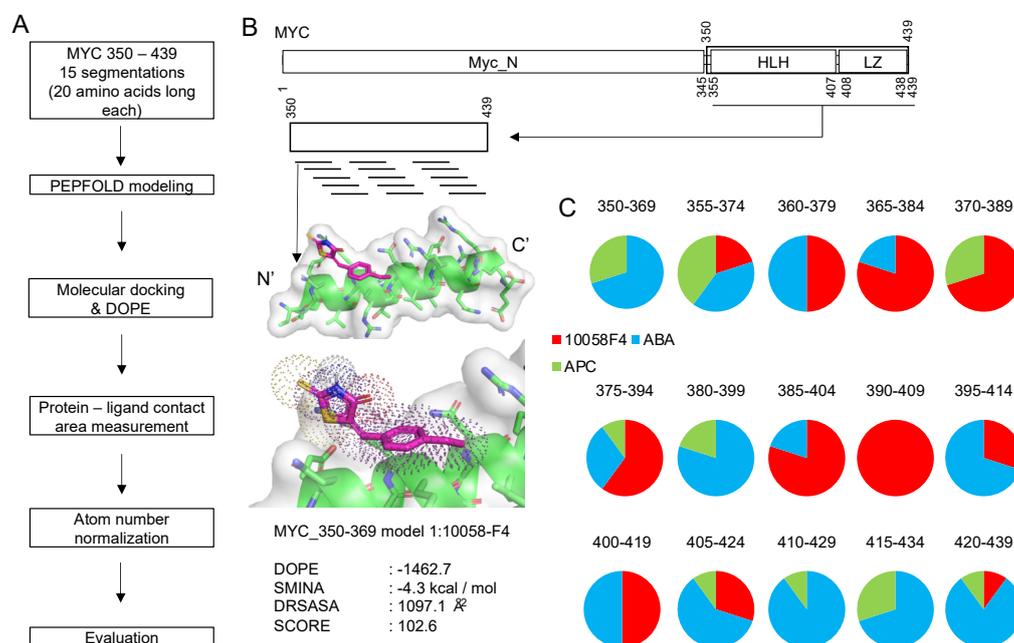


Figure 2. Peptide molecular docking procedure and its application to 10058F4 and other disorder-to-order transition region inhibitors (ABA and APC targeting MYC 350–439 region 15 peptide groups (Figure 1). (A) The procedure of target region segmentation and molecular docking toward generated peptides. (B) A workflow diagram following the procedure described in plot A. (C) Application of the method in the analysis of 10058F4, ABA, and APC interactions with 15 peptide groups derived from the MYC 350–439 region and pie charts showing distribution of the 10 best docking results for each peptide group.

To validate the applicability of this new approach for accurate discovery of small molecules inhibiting DOT-based protein–protein interactions, we aimed to distinguish a MYC disorder-to-order transition coiled-coil region inhibitor (10058F4) from other inhibitors (ABA and APC). Recent an NMR study described secondary structure regions in MYC and annotated structural changes induced by 10058F4 in this protein [14]. We expected that our approach would show results that would be similar or comparable with the outputs of this recent NMR study, especially for the higher rank docking results. After molecular docking of 10058F4 to 15 peptide groups, we discovered that the best 10 molecular docking outputs (see the Section 3 Materials and Methods) from each of the 15 groups showed two distinguishable interaction parts separated by the 380–399 region (Figure 2C). These observations follow the results of a recent NMR study [14].

According to the previous research on MYC inhibitors, the MYC 363–381 region is involved in 10074G5 binding and the 402–412 region plays a role in 10058F4 binding [11]. Our peptide docking study revealed that the 10 best positions from the MYC 390–409 region interact only with 10058F4. Because our peptide segmentation slides 5 amino acids, surrounding peptide docking results also showed higher 10058F4 occupancy in the 10 best docking results (Figure 2C). This region overlaps with the already known 10058F4 binding site (residues 402–412). Especially Y402 was discovered as the major binding residue from the previous circular dichroism and NMR experiment, which revealed Y402 activity in the MYC 402–412 under 10058F4 treatment [11], and two molecular simulation studies confirmed the role of Y402 in binding [17,18]. These previous studies support our discovery of the regions of major occupancy of 10058F4 in MYC 385–424 region peptides.

In summary, molecular docking traditionally relied on already determined target protein structures. As a result, it was hard to apply the technique toward the disorder-to-order transition region in noncomplex forms. Another problem is that intact protein corresponding binding site in MYC is so long to study disorder-to-order transition. To overcome the problems and to follow disorder-to-order transition definition, it was necessary to predict possible conformations in shorter length peptides derived from the binding site. Thus, we utilized web algorithm to predict diverse forms of binding site originated peptides. Molecular docking was applied toward peptides in diverse conformation. Finally, we proved well-known MYC inhibitor (10058F4) bindings toward corresponding disorder-to-order transition region.

2.2. MYC and MDM2 Inhibitors Docking Evaluation Shows Each Compound Binds to Peptide Corresponding Regions, and 10058F4 Interacts with MYC Y402

In the previous study, we focused on the similarity of the intrinsic disorder score pattern (positive slope) within the disorder-to-order transition coiled-coil regions of MYC and MBD2 [12]. Interestingly, the higher occupancy rank of 10058F4 was discovered for the MYC 360–394 region (Figure 2C), which is known as the 10074G5 binding site [11]. The higher occupancy of 10058F4 among the MYC-derived peptides could be due to their high protein-ligand activity, as a recent NMR discovered [14]. To compare with its original binding partner, we tested the interaction of 10074G5 with these peptides. Furthermore, we further tested the relevance using another well-known protein intrinsic disorder advantage for drug discovery, RG7112, an MDM2 inhibitor [19,20]. We determined the RG7112 binding site according to the structure of the humanized *Xenopus* MDM2, which showed that the corresponding human MDM2 binding site is located at the 50–99 region [21].

We applied the same molecular docking environment to 10074G5 and RG7112 (see the Section 3 Materials and Methods) with corresponding targets (10074G5–MYC and RG7112–MDM2). From each peptide docking result, we selected the 10 best positions (samples) and obtained T scores for them. We considered each compound docking result as a population for statistic calculation.

Importantly, our segmented peptide docking and evaluation of newly derived T scores showed distinguishable patterns between 10058F4 in MYC 390–409 and 10074G5 in MYC 365–384 (Figure 3A), which is in the agreement with the results of previous CD and NMR studies [14]. Besides, RG7112 binding site at MDM2 50–69 showed the highest T score among segments derived from the 50–99 region (Figure 3B). The MDM2 50–69 region contains a α -helix between residues 50 and 64, whereas residues 65–69 form a loop. This region directly contacts RG7112 [21]. The majority of the remaining regions in the 70–90 segments go away from the RG7112. However, a short loop (93–96) comes back to the binding site and contacts RG7112 [21]. Therefore, the already known structural biology study results of the active compound to the intrinsically disordered region are well coordinated with our observations.

This evidence supports our method validity for IDPR binding site inhibition, not limited to direct IDPR inhibition. Another merit of our method is that it does not depend on binding site structure. Although we did not apply our method for fuzzy complex inhibition, we propose that our method is applicable to block such fuzzy complex. NUPR1-MSL1 forms a fuzzy complex, and a chemical compound inhibitor (ZZW-115) was discovered for blocking their complex formation [22,23]. We suggest that our next goal is to validate this compound binding toward the corresponding target regions in NUPR1 [23].

To study binding properties of compounds, we analyzed peptide contacts from molecular docking and compared those contact sites with sites involved in interaction with their target protein structures. To do this from molecular docking result, we evaluated the number of contacts between the peptides and compounds. For double evaluation, we used two contact distance cutoffs (4 Å and 3.6 Å) for contact decision. After mapping those contact residues, we compared the residue positions with the corresponding structures of MYC (PDB ID: 1NKP chain A) in its complex with MAX (Figure 3C). Newly evaluated contact scores from the analyzed docking cases showed significantly higher values than the results of other molecular docking experiments (Figure 3D).

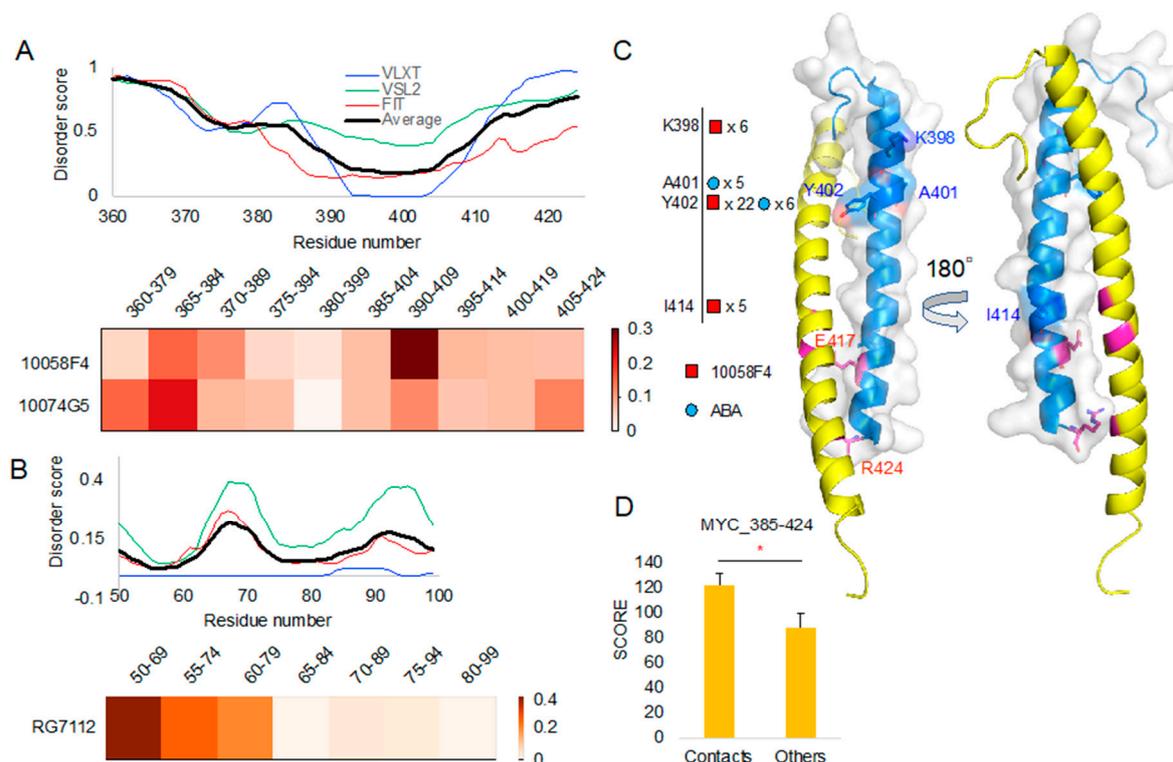


Figure 3. A protein intrinsic disorder score plot of peptide docking region and the corresponding region docking T score heatmap (normalized T scores): (A) MYC protein intrinsic disorder score and corresponding region T score heatmap of 10058F4 and 10074G5. (B) MDM2 protein intrinsic disorder score and corresponding region T score heatmap of RG7112. (C) MYC contact residues shown on MYC–Protein max (MAX) complex structure (PDB ID 1NKP) [13]. (D) *T*-test on the new docking evaluation score (*y*-axis) between contact mapping cases and others. * *p*-value < 0.001.

Our structural comparison study of 10058F4 and MYC confirmed that Y402 serves as an important contact point (Figure 3C, Supplementary Figure S1). According to the previous studies on the 10058F4 and MYC, Y402 was considered as the main contact residue of 10058F4 [11,17,18]. In corresponding analyses, two studies focused on the 402–412 region and used it in the molecular dynamics simulation [17,18]. Because we segmented the 385–424 region, K398 also appeared as the contact residue near Y402. Furthermore, Y402 is embedded within the T400–S405 region between two polar amino acid residues. Therefore, our study suggests that the major MYC–10058F4 binding mechanism is centered at Y402 and the polarity of its environment.

Although they were less frequent than 10058F4–MYC interactions, we also found contacts between ABA and the MYC A401–Y402 region (Figure 3C, Supplementary Figure S1). In our previous study, we applied the Similarity Ensemble Approach (SEA) database to search for the off-targets [24]. Our novel method showed distinguishable contacts for 10058F4, 10074G5, and RG7112 inside their known binding regions as well as the direct/indirect contacts in the disorder-to-order transition region (Figure 3A,B). Although SEA was focusing on structural data, our new method considers diverse forms of disorder-to-order transition regions. MYC A401–Y402 contacts of ABA may provide an evidence of MYC inhibition, which constitutes a new discovery.

2.3. Chemical Compound Library Docking to MYC 390–409 Shows Better 10058F4 Scores

Generally, molecular docking aims to find hits from the chemical compound libraries, and this is the reason why the procedure is called “virtual screening”. Instead of virtual screening, we studied the statistical significance of 10058F4 docking through comparison with the results of docking of the members of a valid chemical compound library (Diversity set V of National Cancer Institute). Besides,

because of utilization of our new evaluation score function characteristics (Section 3 Materials and Methods), we hypothesized that there is a positive correlation between the docking-involved atom number and a new evaluation score. From the compounds (10058F4, ABA, and APC) and MYC peptide dockings, we discovered that the 10 best results of the MYC 390–409 docking contain only 10058F4 (Figure 2C). Therefore, we applied molecular docking focused on the MYC 390–409 peptide as described in the Section 3. Materials and Methods.

To check the correlation between the docking-involved atom numbers and the new evaluation scores, we generated linear regression equation from each population (sample 1, sample 5, and sample 10) (Figure 4A,B). The 22 models of MYC 390–409 peptide docking with 10058F4 (220 results per sample) or without 10058F4 did not show meaningful differences in linear regression. T scores of the 10 best 10058F4 molecular docking results from 22 models, considering each sample as population, were like each other among different populations (Figure 4C). The other remaining 10058F4 docking results (210 results) did not cause significant differences in the T score distribution (Figure 4C).

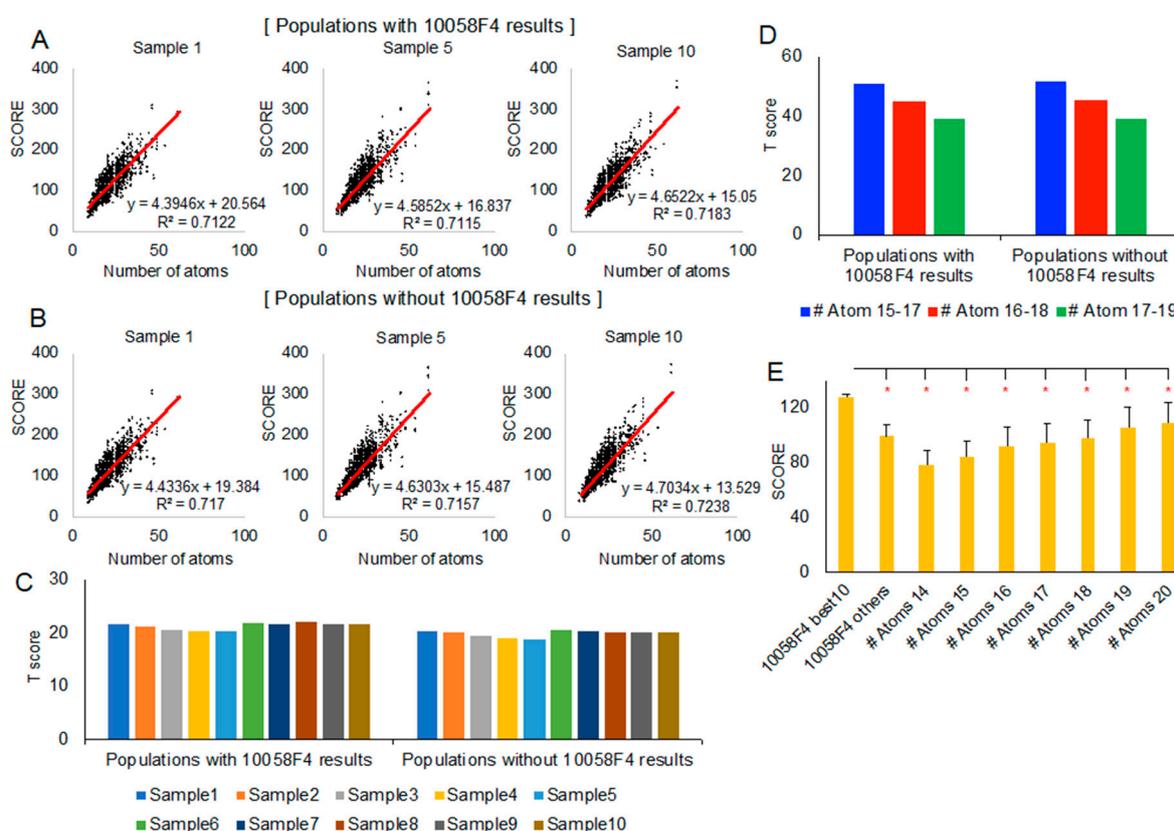


Figure 4. Statistical analysis of the new docking evaluation score focusing on 10058F4–MYC 390–409 region peptides: (A,B) Linear regression between docking-involved atom number (x -axis) and the new docking evaluation score (y -axis) with 10058F4 docking results in the population (A) and without 10058F4 docking results in the population (B). (C) The 10 best 10058F4 docking result T scores in 10 random samples (considering each sample as a population) with/without 10058F4 docking results. (D) The 10 best 10058F4 docking result T scores in three different docking-involved atom number groups (15–17, 16–18, and 17–19) with/without 10058 docking results from all 10 random samples. (E) T -test between the 10 best 10058F4 docking results and each of the docking-involved atom number groups (14–20) from all 10 random samples. * p -value < 0.001.

Because of a new docking evaluation score function with docking-involved atom number normalization, we also grouped all samples (sample 1–10) according to each compound docking-involved atom number. Since the 10058F4 docking-involved atom number was 17, we selected three groups, 15–17, 16–18, and 17–19, for the analysis. As shown in Figure 4D,

the population of the 10058F4 docking results did not cause differences in trend. However, the grouped docking-involved atom number increase caused a decrease in the T score of the 10 best 10058F4 docking results. This means that the significance of the score of the 10 best 10058F4 docking results is reduced along the docking-involved atom number increase.

Nevertheless, the 10 best 10058F4 docking results which had many contacts with Y402 (Figure 3C, Supplementary Figure S1) as well as the majority of binding with MYC 390–409 (Figures 2C and 3A) showed greater new evaluation scores than each of the docking-involved atom number 14–20 group average scores (Figure 4E). In the study describing the 10058F4 discovery, the compound was found using the yeast two-hybrid assay, applying 10,000 drug-like small molecules [10]. The following study revealed that the specific 10058F4 binding region in MYC is the 402–412 region [11]. We showed that the 10 best docking results of 10058F4 toward MYC 390–409 have higher scores than other compounds with similar docking-involved atom numbers. Although the binding mechanism between 10058F4 and MYC 402–412 is unclear, we confirmed the previous discovery of specific binding of 10058F4 to Y402 [11,17,18], applying the original definition of the disorder-to-order transition.

2.4. ABA and APC Both Have Electrostatic Interaction with MBD2, but Only ABA Shows Specific Electrostatic Interaction with p66 α

We studied ABA and APC interactions with MBD2 and p66 α , applying the same procedure used in the 10058F4 analysis. Although we discovered these two compounds from molecular docking to MBD2 and both compounds were shown to efficiently block the MBD2–p66 α complex formation [12], it was unclear if there is a binding difference between those proteins. To see their difference using our novel method, the MBD2–p66 α (PDB ID 2L2L) [15] complex structure was referred to determine the ABA and APC binding region as a disorder-to-order transition region. Then, we segmented MBD2 (residues 359–393) and p66 α (residues 136–175) into the sets of the overlapping 20-amino-acid-long sequences (Table 1) and obtained diverse structures of each peptide region. From diverse forms of peptides (Figure 5), we applied molecular docking and analyzed the result as described in the Materials and Methods section.

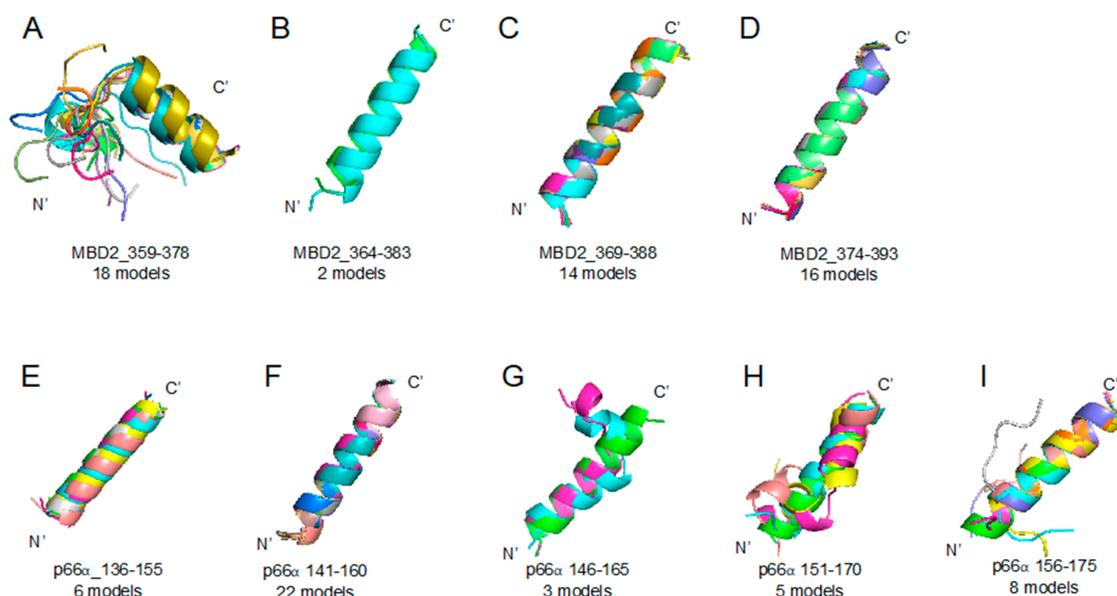


Figure 5. Methyl-CpG-binding domain protein 2 (MBD2) 359–393 region (A–D) and p66 α 136–175 region (E–I) segmented peptides: PEPFOLD ver 3.5 prediction result of peptide structures based on corresponding sequences (please look at the Table 1). (A) MBD2_359–378 (18 models), (B) MBD2_364–383 (2 models), (C) MBD2_369–388 (14 models), (D) MBD2_374–393 (16 models), (E) p66 α _136–155 (6 models), (F) p66 α _141–160 (22 models), (G) p66 α _146–165 (3 models), (H) p66 α _151–170 (5 models), (I) p66 α _156–175 (8 models).

ABA appears to have high T score for regions 359–378/369–388 in MBD2 and for the 136–155 region in p66 α (Figure 6A,B). Protein intrinsic disorder score patterns did not show meaningful correlation with the T scores as it was in the 10058F4 result (Figure 3A,B). ABA contact mapping revealed residues from the MBD2 backside. These mainly include charged and polar residues, such as R (Arg, +1), K (Lys, +1), E (Glu, -1), and Q (Gln) (Figure 6C, Supplementary Figure S2). These results suggested that electrostatic interactions constitute a main binding mechanism that defines the MBD2–ABA interaction. The p66 α structural comparison with the contact mapping also revealed the presence of charged and polar residues, R(Arg, +1), K(Lys, +1), E(Glu, -1), and Q (Gln) from behind the MBD2 binding site (Figure 6D, Supplementary Figure S3). There was no overlap with MBD2–p66 α binding site residues [15]. Although we discovered ABA and APC from targeting p66-binding residues of MBD2 [12], the majority of contacts is located behind the binding site. It suggests that ABA interacts with MBD2 behind the p66 α binding site and that this interaction prevents MBD2 from p66 α binding, likely by inducing the formation of a conformation incapable of MBD2 binding. In p66 α as well, ABA binds p66 α behind its MBD2-binding site. Therefore, it may also cause the structural changes of p66 α , which do not allow MBD2 binding.

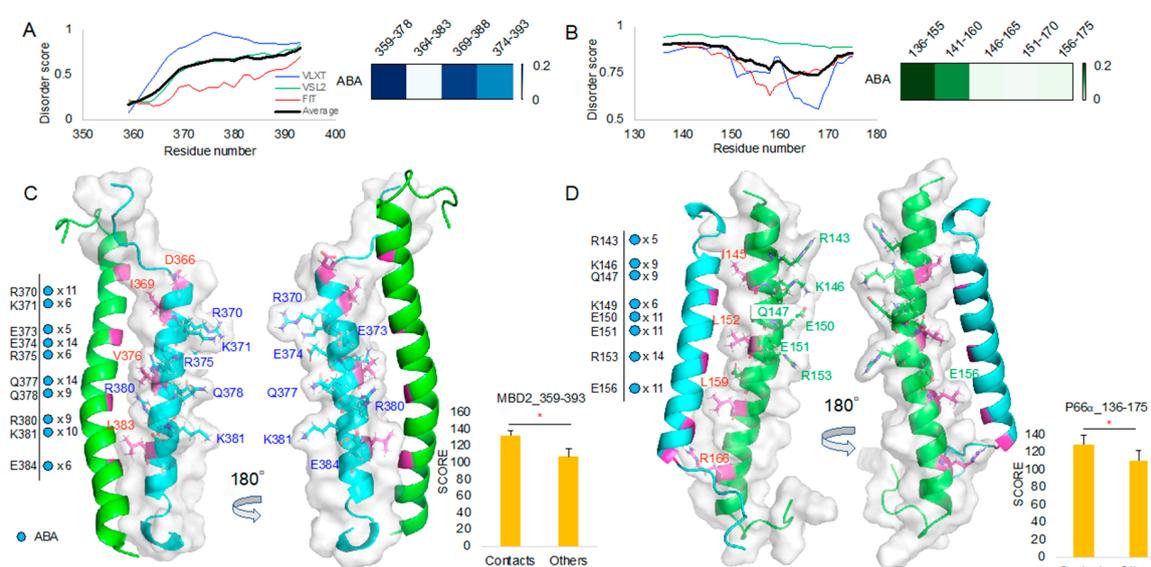


Figure 6. ABA molecular docking toward MBD2 and p66 α : (A,B) Protein intrinsic disorder score plot and the 10 best target region segment peptides ABA docking result T score heatmap of MBD2 (A) and p66 α (B). (C) MBD2 (cyan) contact residues shown on the MBD2–p66 α complex (PDB ID 2L2L) [15] and the corresponding new docking evaluation score (y-axis) T-test between contact cases and others. (D) p66 α (green) contact residues shown on the MBD2–p66 α complex (PDB ID 2L2L) [15] and the corresponding new docking evaluation score (y-axis) T-test between contact cases and others. * *p*-value < 0.001.

To compare ABA and APC binding, we further analyzed the 10 best docking results of the MBD2 359–393 region and p66 α 136–175 region in comparison with 10058F4. It turned out that ABA occupies the majority of the peptides derived from the MBD2 364–393 region (Figure 7A) and the p66 α 136–175 region (Figure 7B). APC was not ranked among the 10 best docking results for any of the MBD2 359–393 peptides (Figure 7A), but it ranked in the 10 best docking results for the majority of the p66 α 136–175 peptides except for the p66 α 146–165 peptide (Figure 7B).

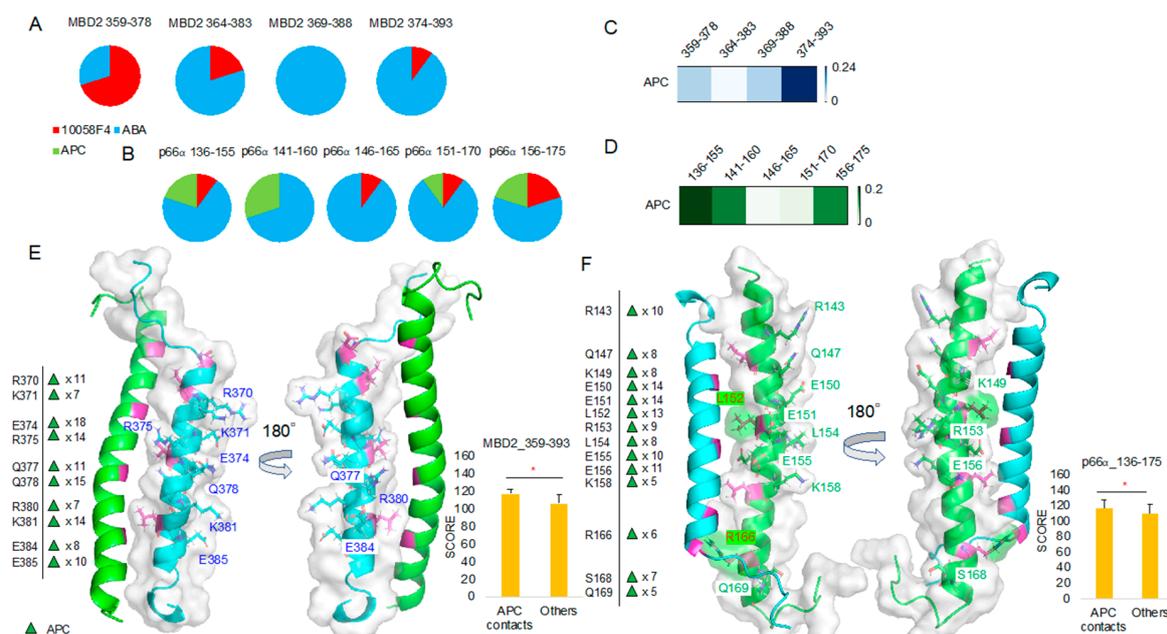


Figure 7. Molecular docking toward MBD2 and p66 α , and APC focused docking result analysis: (A,B) 10058F4, ABA, and APC molecular docking toward MBD2 (A) and p66 α (B) and the corresponding 10 best result distribution pie charts. (C,D) APC-focused docking result T score heatmap from MBD2 docking (C) and p66 α docking (D). (E) MBD2 (cyan) contact residues shown on the MBD2–p66 α complex (PDB ID 2L2L) [15] and the corresponding new docking evaluation score ((y-axis)) *t*-test between contact cases and others. (F) p66 α (green) contact residues shown on the MBD2–p66 α complex (PDB ID 2L2L) [15] and the corresponding new docking evaluation score ((y-axis)) *t*-test between contact cases and others. * *p*-value < 0.001.

Because APC docking scores were lower than those of ABA as shown in Figure 7A,B, we collected only APC docking results and further analyzed them. The heatmap derived using the APC binding T score matrix proved APC preference for the MBD2 374–393 (Figure 7C). However, there was no APC among the 10 best results of the MBD2 374–393 peptide docking. In the case of p66 α binding, T score was relatively higher in p66 α 136–155 and 141–160 peptides, and APC occupied 25% of the 10 best positions in the p66 α 136–155 and 141–160 regions (Figure 7B,D).

MBD2–APC contact mapping (Figure 7E, Supplementary Figure S4) revealed results like the outputs of MBD2–ABA contact mapping (Figure 6C). Contact residues appeared from the MBD2 backside (opposite from p66 α -facing side) and were mostly charged and polar ones, R (Arg, +1), K (Lys, +1), E (Glu, –1), and Q (Gln). However, p66 α –APC contact mapping (Figure 7F, Supplementary Figure S5) was different from that of p66 α –ABA (Figure 6D), and in addition to charged and polar residues R (Arg, +1), K (Lys, +1), E (Glu, –1), and Q (Gln), hydrophobic residue L (Leu) and polar residue S (Ser) were found in the analysis. Despite the fact that no specific interaction mechanisms were identified and no inside–backside distinctions were found, L152 and R166 contact was discovered. Furthermore, the 10 best molecular docking results of 10058F4 and ABA with MBD2 (4 groups of peptides) and p66 α (5 groups of peptides) without APC showed greater score than the APC docking results discovered based on the significant contacts of this compound with MBD2 and p66 α (Supplementary Figure S6).

3. Materials and Methods

3.1. Segmented Peptide Preparation, Molecular Docking, and Evaluation Score Update

For mimicking characteristics of intrinsically disordered region, we obtained diverse structures of target disorder-to-order transition region after its segmentation into a set of

overlapping 20-amino-acid-long peptides. This length of segmentation was used because a 20-amino-acid-long peptide is optimal for the unfolding and folding studies [25]. We segmented protein sequences into subsets (Table 1) and then uploaded them on PEPFOLD v3.5 web server (<https://mobylipe.rpbs.univ-paris-diderot.fr/cgi-bin/portal.py#forms::PEP-FOLD3>) to predict peptide structures [16]. From the corresponding results (Figures 1 and 5), we obtained representative structures of clusters and applied molecular docking with custom function (smina) of Autodock vina (The Scripps Research Institute, Molecular Biology, La Jolla, CA 92037, USA), targeting the central residue (every peptide tenth residue) C α coordinates [26,27]. Chemical compound structures in 3-dimensional SDF formats were obtained from the PubChem database (<https://pubchem.ncbi.nlm.nih.gov/>). Chemical compounds used in this study are 10058F4 (5-[4-Ethylphenyl)methylene]-2-thioxo-4-thiazolidinone, ZINC12406714, PubChem CID: 1271002), 10074G5 (4-nitro-N-(2-phenylphenyl)-2,1,3-benzoxadiazol-7-amine, ZINC3879010, PubChem CID: 2836600), RG7112 ([[(4S,5R)-2-(4-tert-butyl-2-ethoxyphenyl)-4,5-bis(4-chlorophenyl)-4,5-dimethylimidazol-1-yl]-[4-(3-methylsulfonyl propyl)piperazin-1-yl]methanone, ZINC96270381, PubChem CID: 57406853), ABA (2-Amino-N-[[[(3S)-2,3-dihydro-1,4-benzodioxin-3-yl]methyl]acetamide, ZINC40430779, PubChem CID: 93602182), and APC ((R)-3-(2-Amino-acetyl-amino)-pyrrolidine-1-carboxylic acid tert-butyl ester, ZINC60177071, PubChem CID: 66563909). Recent reverse docking study revealed that Autodock vina is biased (shows negative correlation) due to the contact surface area [28]. Therefore, we measured the protein–compound contact surface area [29] and normalized for the number of the molecular docking-involved atoms. We also evaluated the peptide model quality applying DOPE (discrete optimized protein energy) [30] because we obtained structures from the modeling web server. Finally, we designed a new custom score for intrinsically disordered region direct targeting drug discovery evaluation, combining all the aforementioned scores.

$$SCORE = \sum_{i < j} f_{t_i t_j}(r_{ij}) \times -k_B T \ln g^{(n)}(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n) / \left(\sum_{i=0}^n CSA_i \div n \right)$$

where $\sum_{i < j} f_{t_i t_j}(r_{ij})$ is the smina (Autodock vina) score function [28], $-k_B T \ln g^{(n)}(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n)$ is the DOPE score function [30], and $\sum_{i=0}^n CSA_i$ is protein–compound contact surface area calculation algorithm (dr_sasa) score [29]. Normalization of the dr_sasa score with the molecular docking-involved atom number (n) was empirical adjustment to calculate atomic contribution.

3.2. MYC, MDM2, MBD2, and p66 α T Score Heatmap

The new function score was employed for molecular docking evaluation. Each of the 10 best scores of the compound–peptide pair was collected, and we considered each compound docking score from all peptide dockings as a population. Then, we obtained T scores of each compound–peptide pair for the 10 best scores (the 10 best results for each compound) according to the equation below.

$$T = \frac{\bar{x} - \mu}{S / \sqrt{n}} \quad (1)$$

where \bar{x} is each compound's 10 best docking results mean score, μ is the population mean score, S is the standard deviation of each compound's 10 best docking results, and \sqrt{n} is the square root of the number of best 10 docking results ($\sqrt{10}$). After normalization of each compound T score divided by the corresponding compound of all peptide docking T scores sum, we generated heatmaps using Python 3.7 (Python Software Foundation, Wilmington, DE 19801, USA) and matplotlib 3.2 (Plotly, Montreal, QC H2T 2A3, Canada).

3.3. MYC, MBD2, and p66 α Docking Contact Residue Mapping

Each peptide's 10 best docking results from MYC 385–424 (5 peptide groups), MBD2 359–393 (4 peptide groups), and p66 α 136–175 (5 peptide groups) were collected. Next, we analyzed the contacts between the peptide residue atoms and the compound docking-involved atoms. We applied two distance cutoffs (4 Å and 3.6 Å) for double evaluation (Supplementary Figures S1–S5). If the atom pair exists, we determined it the contact of the corresponding residue. After locating all residues from highest rank (1) to lowest rank (10), bottom-up order, we counted all contacts at a distance ≤ 4 Å from segmented peptides and summed up the counts. If the summed-up count of a residue was ≥ 5 , the corresponding residue was chosen. The chosen residue contact count sum were placed along the protein target region and compared with the X-ray crystallography determined structure of MYC–MAX (PDB ID 1NKP) [13] and with the NMR structure of MBD2–p66 α (PDB ID 2L2L) [15].

3.4. MYC 390–409 Peptide Molecular Docking with 10058F4 and Chemical Compound Library (NCI Diversity Set V)

We utilized National Cancer Institute chemical compound library Diversity set V (https://dtp.cancer.gov/organization/dscb/obtaining/available_plates.htm) for molecular docking toward MYC 390–409 peptide 22 models (Figure 11) to compare with 10058F4. It generated about 350,000 docking results, and we obtained 10 groups of each 3000 docking results (sample size) from random sampling and, then, considered each group as a population for the statistic (T score) calculation based on the new docking evaluation scores.

To see the new score docking-involved atom number correlation, we employed linear regression toward three random samples (sample 1, sample 5, and sample 10) using Microsoft Excel 16.37. For T score calculation, we collected the 10 best 10058F4 docking scores from each random sample (1–10), and considered each random sample with/without 10085F4 docking results as the population. Docking-involved atom number groupings into 15–17, 16–18, and 17–19 were pursued, and we collected all docking results from 10 random samples. The 10058F4 docking-involved atom number is 17; thus, we selected the group range in this manner. In this case, each group with/without 10085F4 docking results was considered as a population. Then, we obtained the T scores of the 10 best 10058F4 docking results according to the equation below.

$$T = \frac{\bar{x} - \mu}{S / \sqrt{n}} \quad (2)$$

where \bar{x} is the 10 best 10058F4 docking results mean score, μ is the population mean score, S is the standard deviation of the 10 best 10058F4 docking results, and \sqrt{n} is the square root of the number of 10 best 10058F4 docking results ($\sqrt{10}$).

Finally, we compared the 10 best 10058F4 docking results with docking-involved atom number groups (14, 15, 16, 17 without 10058F4 results, 18, 19, and 20) using *t*-test. For the analysis, again, we collected docking results according to the docking-involved atom number from all of random samples (sample 1–10).

4. Conclusions

The target binding region segmentation generated sets of the overlapping 20-amino-acid-long peptides, and diverse structural forms of these peptides were predicted. We used this step of structural diversity mimicking to study the disordered binding sites that undergo binding-induced folding (i.e., the disorder-to-order transition regions). Molecular docking to these peptides and the new evaluation score successfully distinguished already known compounds and their specific binding sites. Furthermore, we discovered that 10058F4 binds specifically to MYC Y402 and polar residues near this key residue. These observations suggest that the MYC–10058F4 interaction is mainly driven by polarity. Statistically, the new docking score is correlated with the number of contact atoms due to the normalization inside the new docking evaluation score function. Nevertheless,

10058F4 is distinguished rather well from compounds with similar number of atoms randomly chosen from a chemical compound library, NCI Diversity set V. We applied the same procedure for the analysis of other inhibitors targeting disorder-to-order transition regions, ABA and APC. ABA showed contacts with charged residues of MBD2 and p66 α located behind each protein partner-binding site. APC also showed contacts with charged residues behind the p66 α -binding site of MBD2, but no specific binding mechanism was discovered from the analysis of the p66 α docking. In summary, we mimicked structural heterogeneity of the disorder-to-order transition regions based on their native characteristics and evaluated compound binding status with a new score function. We suggest this method as a novel protein intrinsic disorder utilizing a drug discovery platform and hope that it will expand the field of the computer-aided drug discovery.

Supplementary Materials: Supplementary materials can be found at <http://www.mdpi.com/1422-0067/21/15/5248/s1>. Figure S1. MYC binding region peptide docking result contact mapping, Figure S2. MBD2 binding region peptide docking result contact mapping, Figure S3. p66 α binding region peptide docking result contact mapping, Figure S4. APC-focused MBD2 binding region peptide docking result contact mapping, Figure S5. APC-focused p66 α binding region peptide docking result contact mapping, Figure S6. APC-focused docking score comparison between APC contact residue mapping result and the other 10 best peptide docking results (MBD2, 4 peptides; p66 α , 5 peptides) without APC (wo APC).

Author Contributions: Conceptualization, I.N. and C.G.K.; methodology, I.N., S.C., and S.H.S.; validation, I.N. and S.C.; formal analysis, I.N.; investigation, I.N.; resources, C.G.K.; data curation, I.N.; writing—original draft preparation, I.N.; writing—review and editing, I.N., V.N.U., and C.G.K.; visualization, I.N.; supervision, V.N.U. and C.G.K.; project administration, C.G.K.; funding acquisition, C.G.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Basic Science Research Program (NRF-2020R1A2C2009112 to C.G.K.) and the Bio & Medical Technology Development program (NRF-2017M3A9C8027975 and NRF-2020M3A9G4021041 to C.G.K.), National Research Foundation (NRF), Republic of Korea.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

10058F4	10058-F4: 5-[(4-Ethylphenyl)methylene]-2-thioxo-4-thiazolidinone, ZINC12406714, PubChem CID: 1271002
10074G5	10074-G5: 4-nitro-N-(2-phenylphenyl)-2,1,3-benzoxadiazol-7-amine, ZINC3879010, PubChem CID: 2836600
ABA	2-Amino-N-[[[(3S)-2,3-dihydro-1,4-benzodioxin-3-yl]methyl]acetamide, ZINC40430779, PubChem CID: 93602182
APC	(R)-3-(2-Amino-acetyl-amino)-pyrrolidine-1-carboxylic acid tert-butyl ester, ZINC60177071, PubChem CID: 66563909
CD	Circular dichroism
CSA	Contact surface area
DOPE	Discrete optimized protein energy
DOT	Disorder-to-order transition
EM	Electron microscopy
IDP	Intrinsically disordered protein
IDPR	Intrinsically disordered protein region
MAX	Protein Max
MBD2	Methyl-CpG-binding domain protein 2
MDM2	E3 ubiquitin-protein ligase Mdm2
MoRF	Molecular recognition feature
MYC	Myc proto-oncogene protein
NCI	National Cancer Institute
NMR	Nuclear magnetic resonance
p66 α	Transcriptional repressor p66-alpha

RG7112 RG-7112:
 [(4S,5R)-2-(4-tert-butyl-2-ethoxyphenyl)-4,5-bis(4-chlorophenyl)-4,5-dimethylimidazol-1-yl]-
 -[4-(3-methylsulfonyl propyl)piperazin-1-yl]methanone, ZINC96270381, PubChem CID:
 57406853

References

- Lin, X.; Li, X.; Lin, X. A Review on applications of computational methods in drug screening and design. *Molecules* **2020**, *25*, 1375. [[CrossRef](#)] [[PubMed](#)]
- Uversky, V.N.; Oldfield, C.J.; Dunker, A.K. Intrinsically disordered proteins in human diseases: Introducing the D2 concept. *Annu. Rev. Biophys.* **2008**, *37*, 215–246. [[CrossRef](#)] [[PubMed](#)]
- Babu, M.M.; Van Der Lee, R.; De Groot, N.S.; Gsponer, J. Intrinsically disordered proteins: Regulation and disease. *Curr. Opin. Struct. Boil.* **2011**, *21*, 432–440. [[CrossRef](#)]
- Uversky, V. p53 Proteoforms and Intrinsic Disorder: An Illustration of the Protein Structure–Function Continuum Concept. *Int. J. Mol. Sci.* **2016**, *17*, 1874. [[CrossRef](#)] [[PubMed](#)]
- Uversky, V.N. Intrinsically Disordered Proteins and Their “Mysterious” (Meta)Physics. *Front. Phys.* **2019**, *7*, 10. [[CrossRef](#)]
- Romero, P.; Obradovic, Z.; Kissinger, C.R.; E Villafranca, J.; Garner, E.; Guillot, S.; Dunker, A. Thousands of proteins likely to have long disordered regions. *Pac. Symp. Biocomput. Pac. Symp. Biocomput.* **1998**, 437–448.
- Garner, E.; Romero, P.; Dunker, A.K.; Brown, C.; Obradovic, Z. Predicting Binding Regions within Disordered Proteins. *Genome Inform. Ser. Workshop Genome Inform.* **1999**, *10*, 41–50.
- Oldfield, C.J.; Cheng, Y.; Cortese, M.S.; Romero, P.; Uversky, V.N.; Dunker, A.K. Coupled folding and binding with α -helix-forming molecular recognition elements. *Biochemistry* **2005**, *44*, 12454–12470. [[CrossRef](#)]
- Mohan, A.; Oldfield, C.J.; Radivojac, P.; Vacic, V.; Cortese, M.S.; Dunker, A.K.; Uversky, V.N. Analysis of Molecular Recognition Features (MoRFs). *J. Mol. Boil.* **2006**, *362*, 1043–1059. [[CrossRef](#)]
- Yin, X.; Giap, C.; Lazo, J.S.; Prochownik, E.V. Low molecular weight inhibitors of Myc–Max interaction and function. *Oncogene* **2003**, *22*, 6151–6159. [[CrossRef](#)]
- Follis, A.V.; Hammoudeh, D.I.; Wang, H.; Prochownik, E.V.; Metallo, S.J. structural rationale for the coupled binding and unfolding of the c-Myc oncoprotein by small molecules. *Chem. Boil.* **2008**, *15*, 1149–1155. [[CrossRef](#)] [[PubMed](#)]
- Kim, M.Y.; Na, I.; Kim, J.S.; Son, S.H.; Choi, S.; Lee, S.E.; Kim, J.-H.; Jang, K.; Alterovitz, G.; Chen, Y.; et al. Rational discovery of antimetastatic agents targeting the intrinsically disordered region of MBD2. *Sci. Adv.* **2019**, *5*, eaav9810. [[CrossRef](#)] [[PubMed](#)]
- Nair, S.K.; Burley, S.K. X-ray structures of Myc-Max and Mad-Max recognizing DNA. Molecular bases of regulation by proto-oncogenic transcription factors. *Cell* **2003**, *112*, 193–205. [[CrossRef](#)]
- Panova, S.; Cliff, M.; Macek, P.; Blackledge, M.; Jensen, M.R.; Nissink, J.W.M.; Embrey, K.J.; Davies, R.; Waltho, J.P. Mapping hidden residual structure within the Myc bHLH-LZ domain using chemical denaturant titration. *Structure* **2019**, *27*, 1537–1546. [[CrossRef](#)]
- Gnanapragasam, M.N.; Scarsdale, J.N.; Amaya, M.L.; Webb, H.D.; Desai, M.A.; Walavalkar, N.M.; Wang, S.Z.; Zu Zhu, S.; Ginder, G.D.; Williams, J.D.C. p66 -MBD2 coiled-coil interaction and recruitment of Mi-2 are critical for globin gene silencing by the MBD2-NuRD complex. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 7487–7492. [[CrossRef](#)]
- Lamiable, A.; Thevenet, P.; Rey, J.; Vavrusa, M.; Derreumaux, P.; Tufféry, P. PEP-FOLD3: Faster de novo structure prediction for linear peptides in solution and in complex. *Nucleic Acids Res.* **2016**, *44*, W449–W454. [[CrossRef](#)]
- Michel, J.; Cuchillo, R. The impact of small molecule binding on the energy landscape of the intrinsically disordered protein C-Myc. *PLoS ONE* **2012**, *7*, e41070. [[CrossRef](#)]
- Heller, G.T.; Aprile, F.A.; Bonomi, M.; Camilloni, C.; De Simone, A.; Vendruscolo, M. sequence specificity in the entropy-driven binding of a small molecule and a disordered peptide. *J. Mol. Boil.* **2017**, *429*, 2772–2779. [[CrossRef](#)]
- Vassilev, L.T.; Vu, B.; Graves, B.; Carvajal, D.; Podlaski, F.; Filipovic, Z.; Kong, N.; Kammlott, U.; Lukacs, C.; Klein, C.; et al. *In vivo* activation of the p53 pathway by small-molecule antagonists of MDM2. *Science* **2004**, *303*, 844–848. [[CrossRef](#)]

20. Vu, B.; Wovkulich, P.; Pizzolato, G.; Lovey, A.; Ding, Q.; Jiang, N.; Liu, J.-J.; Zhao, C.; Glenn, K.; Wen, Y.; et al. discovery of RG7112: A small-molecule MDM2 inhibitor in clinical development. *ACS Med. Chem. Lett.* **2013**, *4*, 466–469. [[CrossRef](#)]
21. Tovar, C.; Graves, B.; Packman, K.; Filipovic, Z.; Xia, B.H.M.; Tardell, C.; Garrido, R.; Lee, E.; Kolinsky, K.; To, K.-H.; et al. MDM2 small-molecule antagonist RG7112 activates p53 signaling and regresses human tumors in preclinical cancer models. *Cancer Res.* **2013**, *73*, 2587–2597. [[CrossRef](#)] [[PubMed](#)]
22. Neira, J.L.; Bintz, J.; Arruebo, M.; Rizzuti, B.; Bonacci, T.; Vega, S.; Lanas, A.; Velazquez-Campoy, A.; Iovanna, J.; Abián, O. Identification of a drug targeting an intrinsically disordered protein involved in pancreatic adenocarcinoma. *Sci. Rep.* **2017**, *7*, 39732. [[CrossRef](#)]
23. Santofimia-Castaño, P.; Rizzuti, B.; Xia, Y.; Abian, O.; Peng, L.; Velázquez-Campoy, A.; Neira, J.L.; Iovanna, J. Targeting intrinsically disordered proteins involved in cancer. *Cell. Mol. Life Sci.* **2019**, *77*, 1695–1707. [[CrossRef](#)] [[PubMed](#)]
24. Keiser, M.J.; Setola, V.; Irwin, J.J.; Laggner, C.; Abbas, A.I.; Hufeisen, S.J.; Jensen, N.H.; Kuijjer, M.B.; Matos, R.C.; Tran, T.B.; et al. Predicting new molecular targets for known drugs. *Nature* **2009**, *462*, 175–181. [[CrossRef](#)]
25. Sluysmans, D.; Willet, N.; Thévenot, J.; Lecommandoux, S.; Duwez, A.-S. Single-molecule mechanical unfolding experiments reveal a critical length for the formation of α -helices in peptides. *Nanoscale Horizons* **2020**, *5*, 671–678. [[CrossRef](#)] [[PubMed](#)]
26. Trott, O.; Olson, A.J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2009**, *31*, 455–461. [[CrossRef](#)]
27. Koes, D.R.; Baumgartner, M.P.; Camacho, C.J. Lessons learned in empirical scoring with smina from the CSAR 2011 Benchmarking exercise. *J. Chem. Inf. Model.* **2013**, *53*, 1893–1904. [[CrossRef](#)]
28. Luo, Q.; Zhao, L.; Hu, J.; Jin, H.; Liu, Z.; Zhang, L. The scoring bias in reverse docking and the score normalization strategy to improve success rate of target fishing. *PLoS ONE* **2017**, *12*, e0171433. [[CrossRef](#)]
29. Ribeiro, J.; Ríos-Vera, C.; Melo, F.; Schüller, A. Calculation of accurate interatomic contact surface areas for the quantitative analysis of non-bonded molecular interactions. *Bioinformatics* **2019**, *35*, 3499–3501. [[CrossRef](#)]
30. Shen, M.-Y.; Sali, A. Statistical potential for assessment and prediction of protein structures. *Protein Sci.* **2006**, *15*, 2507–2524. [[CrossRef](#)] [[PubMed](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).