

PROTAC-induced Protein Structural Dynamics in Targeted Protein Degradation

Reviewed Preprint

v2 • January 31, 2025

Revised by authors

Reviewed Preprint

v1 • September 6, 2024

Kingsley Y Wu, Ta I Hung, Chia-en A Chang 

Department of Chemistry, University of California, Riverside, Riverside, United States • Department of Bioengineering, University of California, Riverside, Riverside, United States

 https://en.wikipedia.org/wiki/Open_access
 Copyright information

eLife Assessment

This study provides **important** computational insights into the dynamics of PROTAC-induced degradation complexes, offering a **convincing** demonstration that differences in degradation efficacy can be linked to linker properties. The analyses address reproducibility considerations comprehensively, reinforcing the study's conclusions. Overall, these findings are significant for advancing cancer treatments and will be of broad interest to both biochemists and biophysicists.

<https://doi.org/10.7554/eLife.101127.2.sa4>

Abstract

PROteolysis TARgeting Chimeras (PROTACs) are small molecules that induce target protein degradation via the ubiquitin-proteasome system. PROTACs recruit the target protein and E3 ligase; a critical first step is forming a ternary complex. However, while the formation a ternary complex is crucial, it may not always guarantee successful protein degradation. The dynamics of the PROTAC-induced degradation complex play a key role in ubiquitination and subsequent degradation. In this study, we computationally modelled protein complex structures and dynamics associated with a series of PROTACs featuring different linkers to investigate why these PROTACs, all of which formed ternary complexes with Cereblon (CRBN) E3 ligase and the target protein bromodomain-containing protein 4 (BRD4^{BD1}), exhibited varying degrees of degradation potency. We constructed the degradation machinery complexes with Cullin-Ring Ligase 4A (CRL4A) E3 ligase scaffolds. Through atomistic molecular dynamics simulations, we illustrated how PROTAC-dependent protein dynamics facilitating the arrangement of surface lysine residues of BRD4^{BD1} into the catalytic pocket of E2/ubiquitin cascade for ubiquitination. Despite featuring identical warheads in this PROTAC series, the linkers were found to affect the residue-interaction networks, and thus governing the essential motions of the entire degradation machine for ubiquitination. These findings offer a structural dynamic perspective on ligand-induced protein degradation, providing insights to guide future PROTAC design endeavors.

Introduction

With the rapid progressive efforts from modern drug discovery and development, numerous promising paradigms for disease treatment are emerging from the wealth of medicinal chemistry and biology. A comprehensive understanding of the molecular mechanisms underlying the function of targeted biological systems expedites the drug development process. Among these innovative approaches, PROteolysis TARgeting Chimeras (PROTACs), also known as hetero-bifunctional degraders, stand out as a revolutionary paradigm for selectively degrading a diverse array of disease-associated targets.^{1,2,3} The formation of a ternary complex involving a PROTAC binding to an E3 ligase and its neo-substrate, commonly referred to as the protein of interest (POI), triggers the cell's native protein degradation machinery (e.g., ubiquitination), which marks the target protein for proteasomal degradation. Previous studies have highlighted the importance of forming a high binding-affinity, stable and long-lived E3-PROTAC-POI complex for achieving potent degradation.⁴ However, it is noteworthy that the formation of a stable ternary complex does not always correlate with the degradation potency of the POI.^{4,5,6}

The final degradation of a target protein, involving the three-step ubiquitination process (activation-E1, conjugation-E2, and ligation-E3) is facilitated by the binding a ternary complex induced by a PROTAC. Recent studies used E3 ligases such as MDM2, IAP, RNF4, and β TRCP to degrade various biological targets.^{7,8,9,10} While stable ternary complexes are often associated with high POI degradation efficiency, discrepancy between stable ternary complexes and degradation have also been observed. For instance, certain von Hippel Lindau (VHL)-based PROTACs exhibit high binding affinity to POIs but fail to induce degradation despite forming stable ternary complexes. Conversely, proteins like p38 α , which exhibit low binding affinity with VHL-based PROTACs, undergo rapid degradation within a short timeframe.^{4,11,12} Therefore, it is evident that the formation of a stable ternary complex does not guarantee high degradation efficiency. Recent studies have suggested that the accessibility of lysine (Lys) residues of the POI after the formation of stable ternary complexes is critical for degradation and requires further research.^{4,13,14}

The formation of a stable E3-PROTAC-POI complex represents the initial crucial step for POI degradation. Several modeling-based approaches have been developed to optimize and rationalize the PROTAC-mediated ternary complex. For example, several studies have employed protein-protein docking, linker conformational searching,^{15,16,17,18} and machine learning approaches^{19,20} to generate ensembles of E3-PROTAC-POI complexes *in silico*. Recent advancement in docking algorithms aimed to provide higher accuracy and energetically favorable ternary complexes.^{21,22} Notably, recent studies have suggested that, in addition to forming a ternary complex, the orientation of E3-PROTAC-POI allowing accessibility of the surface Lys residues of the POI are important for ubiquitination. Because the subsequent ubiquitination process occurs in a ligase complex comprising multiple proteins, recent research efforts have begun to investigate the assembly of the E3-PROTAC-POI complex with the entire ligase complex and key enzymes E2/ubiquitin (E2/Ub) for ubiquitination. For example, two studies applied docking and structure-based modeling approaches to model CDKs-PROTACs and BCL-xL-PROTACs in a Cereblon (CRBN)-based Culling-Ring Ligase 4A (CRL4A) complex and a VHL-based CRL complex, respectively.^{14,23} These studies integrated experimentally determined protein structures to construct a degradation complex using protein structure alignment approaches, and the distance measurements between Lys residues and Ub to show potential ubiquitination. Moreover, enhanced sampling molecular dynamics (MD) techniques have also been employed to investigate the spatial organization of POI within the context of a ligase complex for ubiquitination.²⁴ Nevertheless, challenges remain in accurately modelling the structural dynamics of the degradation machinery complex, and gaining a deeper understanding of the steps involved in POI ubiquitination would aid PROTAC design.

Several existing studies have examined a set of PROTACs targeting the same E3 ligase and POI for structure–activity relationships^{25–27}. Many of these studies have maintained consistent binders for the E3 and POI while varying the linker regions. For example, studies of several PROTACs (i.e., small-molecule degrader of bromodomain and extra-terminal domain [dBET] family) designed for degrading bromodomain-containing protein 4 (BRD4^{BD1}) utilized the same BRD inhibitor (i.e., thieno-triazolo-1,4-diazepine [JO1]) and E3 ligase inhibitor, pomalidomide (or lenalidomide and thalidomide) while varying the linkers^{6,28–30}. Although the crystal structures of these CRBN-dBETs-BRD4^{BD1} ternary complexes have highly similar CRBN-BRD4^{BD1} contacts, significant differences in degradation capabilities (DC_{50/5h}) have been observed^{6,28,31–32}. These findings underscored the need to further understanding of how the linkers and the formation of stable ternary complex may influence degradation capability. Although the linker region may result in minor differences in ligand solubility and/or membrane permeability, the proximity and accessibility of surface Lys residue(s) of the POI to the E2/Ub cascade in the ubiquitination site can be due to different linkers^{11,33}. Thus, investigating the structures of the degradation complexes with PROTAC-induced ubiquitination while considering protein structural dynamics provide valuable insights into PROTAC degradability.

In this study, we employed protein–protein docking, structural alignment, atomistic MD simulations, and post-analysis to model a series of CRBN-dBET-BRD4^{BD1} ternary complexes and the entire degradation machinery complex consisting of BRD4^{BD1} and a CRL4A E3 ligase scaffold (**Figure 1**). These degraders, with different linker properties, were all capable of forming stable ternary complexes, but exhibited different degradation capabilities^{6,28,31–32}. The best degrader, dBET70, had a DC_{50/5h} of about 5 nM, followed by dBET23 (DC_{50/5h} ~50 nM) (**Figure 1D** and Figure S1). Although utilizing the exact same warheads, other degraders such as dBET1 and dBET57 had a DC_{50/5h} of about 500 nM. Our studies identified structural features of the dBETs that contribute to large-scale protein motions and explained the connection between the cellular activities and degradability reported for the dBETs with atomistic details. Because finding energetically stable ternary conformations is a critical first step for protein degradation^{6,17,18}, in addition to protein–protein docking, we also performed MD simulations to thoroughly cover the conformational space and energy calculations to ensure that our modeled ternary complexes were thermodynamically stable.

Assembling these stable CRBN-dBETx-BRD4^{BD1} complexes into multiple modeled CRL4A E3 ligase-based scaffold conformations, we first observed that no surface Lys residue(s) of BRD4^{BD1} was ready for the next ubiquitination step in the modeled degradation machinery complexes. Nevertheless, our unbiased MD simulations illustrated protein structural dynamics of the entire complex and local sidechain arrangements to bring Lys residue(s) to the catalytic pocket of E2/Ub for reactions. Post-analysis revealed the essential motions of the degradation complex and interactions crucial for ubiquitination. Our results relate the structural motion to potential ubiquitination and explain how the linker property affecting the degradation potency. Our results show the importance of the dynamic features in protein structure and how the linker region of a PROTAC may contribute to protein motions to achieve PROTAC-mediated POI degradation.

Results

To understand why PROTACs, although forming similar stable E3-PROTAC-POI ternary complexes, induce different degradation efficacy of POIs, we used an integrative approach that combines docking, structural alignment and atomistic MD simulations (Scheme S1). Due to the limited availability of experimental determined CRBN-dBETx-BRD4^{BD1} conformations, we initially employed protein–protein docking to generate numerous CRBN-dBETx-BRD4^{BD1} ternary complexes for four degraders with different degradation efficacies (Table S1 and Figure S1). Subsequently, we constructed degradation machinery complexes by assembling various ternary complexes into the CRL4A E3 ligase scaffold. MD simulations and subsequent post-analysis were

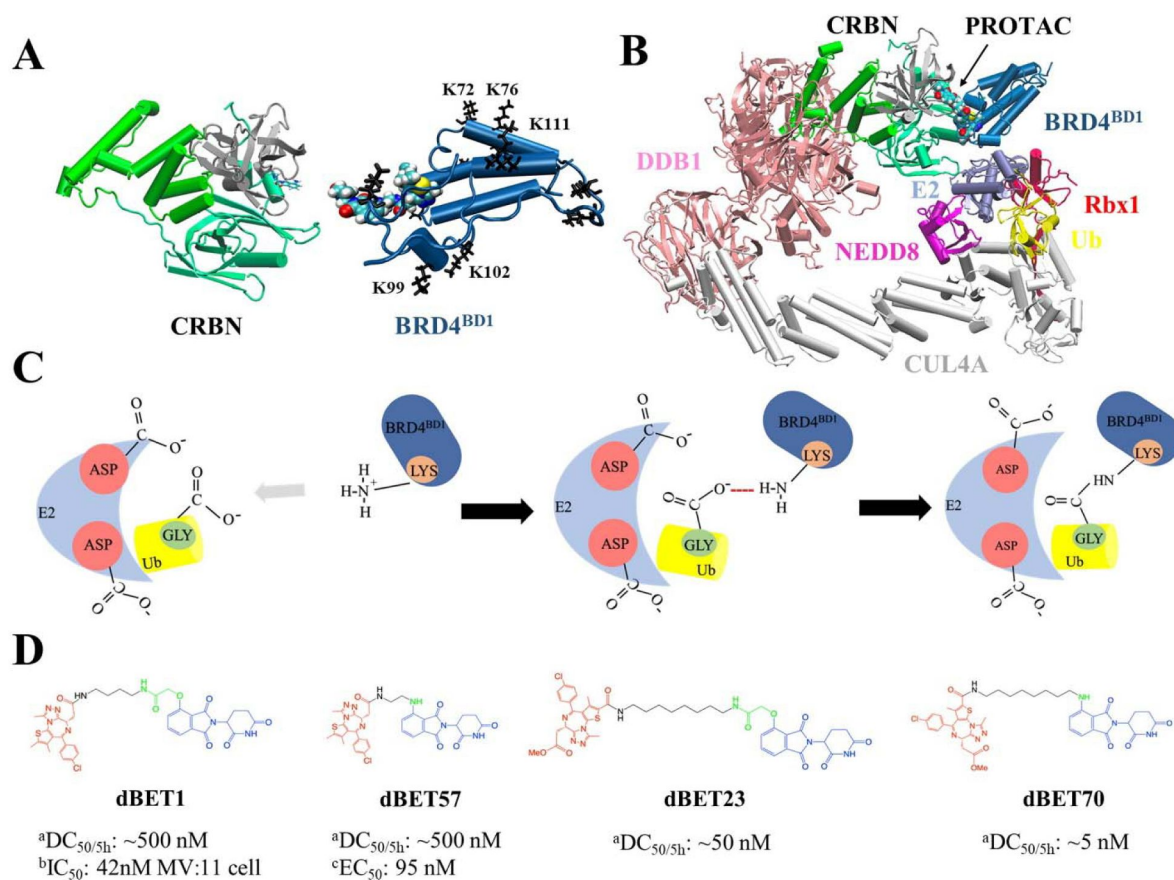


Figure 1

Structures of PROteolysis Targeting Chimera (PROTAC)-mediated degradation machinery complex.

(A) Cereblon (CRBN) E3 ligase, consists of an N-terminus domain (mint), helical binding domain (green), and C-terminus domain (gray). Notably, CRBN E3 ligase is also simply termed CRBN in the main text. Bromodomain-containing protein 4 (BRD4^{BD1}; targeted protein) consists of several important Lys residues (K72/76/99/102/111 black sticks) for successful degradation. **(B)** Degradation machinery complex is constructed with DDB1, CUL4A, NEDD8, Rbx1, E2 enzyme, ubiquitin (Ub), and CRBN-PROTAC-BRD4^{BD1}. Note, the complex without PROTAC recruited BRD4^{BD1} is termed CRL4A E3 ligase. **(C)** An illustration of the theoretical model of ubiquitination reaction. The catalytic site with multiple Asp residues creates a negatively charged environment that attracts Lys residues to enter the catalytic site. Lys and Asp residues then react with the C-terminus Gly of Ub for future degradation. **(D)** Chemical structure of each dBET PROTAC. ^aDegradation profile DC_{50/5h} for four PROTACs was obtained from EGFP/mCherry reporter assay published in reference #5. ^bData published in reference #31. ^cData published in reference #5.

employed to quantify protein dynamics and to identify crucial hinge regions governing the structure–dynamics–function relationship within the degradation complexes. With quantitative data, we revealed the importance of the structural dynamics of dBETx-induced motions, which arrange positions of the surface lysine residues of BRD4^{BD1} and the entire degradation machinery.

Conformational ensembles of CRBN-

dBETx-BRD4^{BD1} ternary complexes

An effective design of PROTACs relies on a comprehensive understanding how a degrader, such as dBET, yields different conformations of CRBN-dBETx-BRD4^{BD1} ternary complexes, ultimately contributing to degradation efficiency. To generate the conformation ensemble, we first constructed the ternary complexes through protein–protein docking using the Molecular Operating Environment (MOE) program, and removed complexes with atomic clashes (refer to SI for docking data). Figure S1 illustrates the docked CRBN-dBETx-BRD4^{BD1} conformations, with all conformations, except CRBN-dBET57-BRD4^{BD1}, presenting over 160 conformations and displaying various inter-molecular orientations between CRBN and BRD4^{BD1}. Notably, PROTAC dBET57, characterized by a shortest linker, exhibited a constrained protein–protein orientation, resulting in only 24 distinct ternary conformations (Figure S1C and Table S1). To validate the modelled conformations, we superimposed the docking results of CRBN-dBET23-BRD4^{BD1} and CRBN-dBET70-BRD4^{BD1} ternary complexes onto available crystal structures, revealing highly similar conformations. The computed smallest C_α-root mean square deviation (RMSD) values were 1.99 Å, and 2.60 Å, respectively (Figure S2). These ternary ensembles, as depicted in Figure S1, were subsequently utilized to construct degradation machinery complexes, as detailed in the following subsection (See Flowchart at Figure S3).

Furthermore, we conducted multiple MD simulations in explicit solvent for five CRBN-dBETx-BRD4^{BD1} complexes, encompassing four initial conformations derived from our docking model and one from an existing crystal structure (Table 1 [↗](#), total of 15 runs for ternary complexes). Only dBET23 crystal structure is available with the PROTAC and both proteins, while the experimentally determined ternary complexes of dBET1, dBET57 and dBET70 are not available. During 400-ns simulations, the two warheads of a PROTAC bound tightly to CRBN and BRD4^{BD1}, while the linker displayed high flexibility by adopting different conformations and facilitating the interaction between the two proteins across different protein–protein contact surfaces (Figure 2 [↗](#)). The residues contact map throughout the 400ns MD simulation also showed different pattern of protein-protein interactions, indicating that the linkers were able to adopt different conformations (Figure S4). Interestingly, despite the molecular docking results suggested limited fluctuation in CRBN-dBET57-BRD4^{BD1}, our MD simulations revealed considerable variation in ternary conformations. The observed large-scale motions in all CRBN-dBETx-BRD4^{BD1} complexes, as predicted by both docking and MD simulations, underscored the importance of CRBN and BRD4^{BD1} orientations influenced by the presence of dBETs as key factors of contributing to degradation efficiency.

Modelled degradation machinery complexes

To facilitate ubiquitination, once a stable CRBN-dBETx-BRD4^{BD1} complex is formed, led by CRBN, the ternary complex is assembled with an E3 ligase scaffold for ubiquitination. For this purpose, we utilized a widely employed scaffold, CRBN/DDB1/CUL4A/NEDD8/Rbx1/E2/Ub, to assemble our ternary complex, with CRBN binding to the adaptor protein DDB1 to bring the target protein into proximity with E2/Ub (Figure 1B [↗](#)). Given the dynamic nature of the scaffold complex, we selected 12 distinct DDB1 crystal structures to construct multiple scaffold conformations¹⁸ [↗](#). Notably, among these structures, only one of the 12 PDB files had CRBN bound to DDB1 (PDB ID 4TZ4). Using this crystal structure as a template, we further built 12 CUL4A E3 ligase scaffolds: DDB1/CUL4A/NEDD8/Rbx1/E2/Ub/CRBN. Subsequently, three of the resulting 12 complexes were found to contain clashes and were excluded from further analysis (Figure S5). The remaining nine

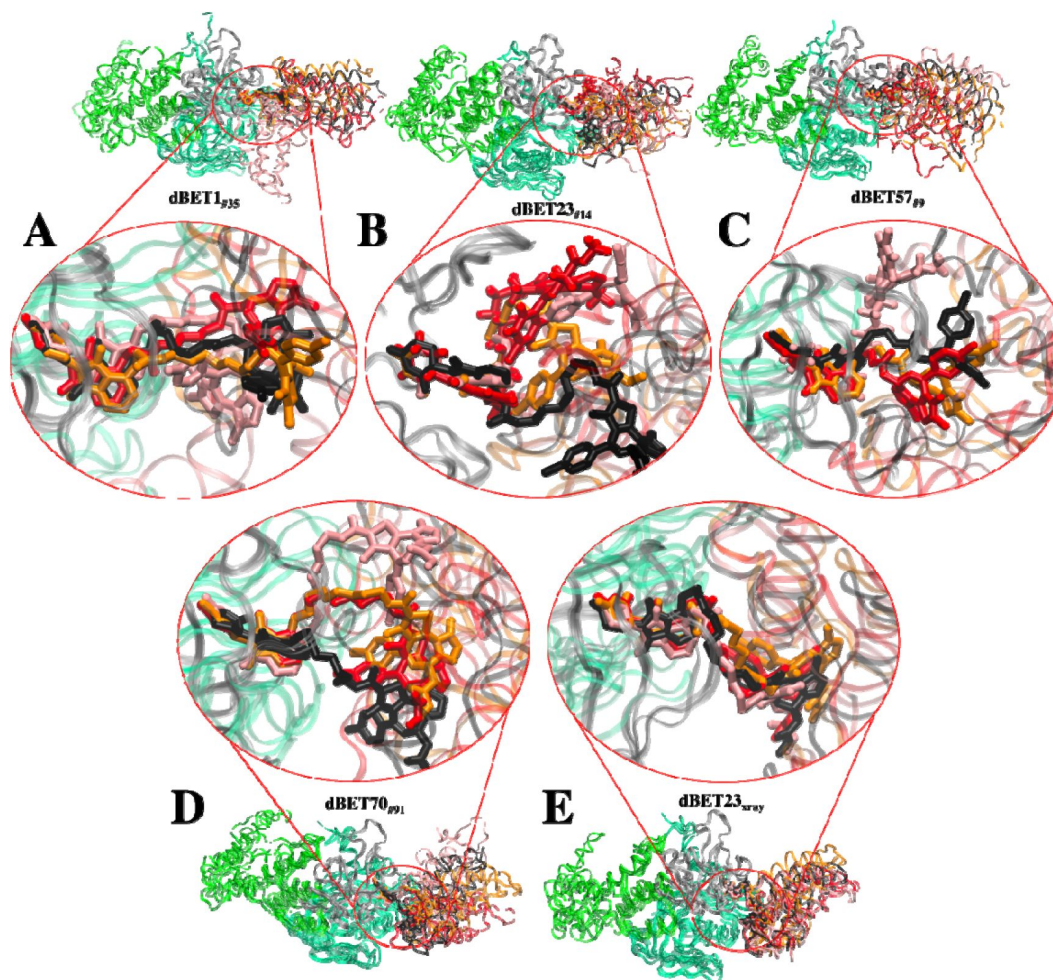


Figure 2

Flexibility of dBET PROTACs.

The last molecular dynamics (MD) frame from three MD runs was aligned with initial frame of the C-terminus domain (gray) of CRBN. Two warheads bound tightly to the respective proteins, whereas the linker is highly flexible, adopting various conformations. MD run 1 (red), MD run 2 (pink), MD run 3 (orange), and initial frame (black). **(A)** CRBN-dBET1_#35-BRD4^{BD1}. **(B)** CRBN-dBET23_#14-BRD4^{BD1}. **(C)** CRBN-dBET57_#9-BRD4^{BD1}. **(D)** CRBN-dBET70_#91-BRD4^{BD1}. **(E)** CRBN-dBET23_{xray}-BRD4^{BD}.

MD index	Run	MD index	Run	MD index	Run
CRBN-dBET1_ _{#35} -BRD4 ^{BD1}	3	A1_dBET1_ _{#35}	1	B1_dBET1_ _{md3}	2
CRBN-dBET23_ _{#14} -BRD4 ^{BD1}	3	A1_dBET23_ _{#14}	1	B1_dBET23_ _{#14}	2
CRBN-dBET57_ _{#9} -BRD4 ^{BD1}	3	A1_dBET57_ _{#9}	1	B1_dBET57_ _{#9} _MD1,	2
				B1_dBET57_ _{#9} _MD2,	1
CRBN-dBET70_ _{#91} -BRD4 ^{BD1}	3	A1_dBET70_ _{#91}	1	B1_dBET70_ _{#91}	2
CRBN-dBET23_ _{xray} -BRD4 ^{BD1}	3	A1_dBET23_ _{xray}	1	B1_dBET23_ _{xray}	1

Table 1.

List of molecular dynamics (MD) simulations for ternary complex CRBN-dBETx-BRD4^{BD1} and the degradation machinery complex.

A1 and B2 indicate that the complex was assembled using scaffold cluster A1 and B1, respectively. The subscript number after each dBET degrader indicates the conformation index number from the protein-protein docking results. B1_dBET1__{md3} indicates that a ternary conformation obtained from a CRBN-dBET1__{#35}-BRD4^{BD1} MD run (**Figure 2** [color orange](#)) was used to build the initial conformation.

CRL4A E3 ligase scaffolds all formed a ring-like overall shape. We further analyzed the conformations and clustered them using a distance between the E2 and CRBN interface resulting two distinct groups: the ring-forming cluster A, characterized by a gap distance ranging from ~1.0 Å to 10.0 Å (Figure S6A, clusters A1-5), and the ring-open cluster B, featuring a much larger gap distance ranging from ~14 Å to 35 Å (Figure S6B, clusters B1-4). Subsequently, we allocated 561 modeled CRBN-dBETx-BRD4^{BD1} conformations, along with a crystal structure (CRBN-dBET23-BRD4^{BD1}), into the nine CRL4A E3 ligase scaffolds, labeled A1 to A5 and B1 to B4, resulting in the construction a total of 5058 degradation machinery complexes.

According to the mechanistic studies^{23,34}, both a Lys residue of the POI and a Gly residue in the C-terminal G-G motif in Ub must be positioned within an Asp-rich catalytic site for the ubiquitination reaction to occur, with Ub conjugated to the POI by forming a covalent bond between Lys residue and Gly residue (Figure 1C). Using the distance between one Lys residue of BRD4^{BD1} and the C-terminal Gly residue of Ub as a criterion, we examined each Lys residue of BRD4^{BD1} across the 5058 degradation machinery complexes. Any complex in which the distance between any Lys of BRD4^{BD1} and the C-terminal Gly of Ub was found to be < 16 Å was retained, resulting in the identification of 1226 degradation machinery complexes with diverse conformations (Table 2). For instance, within cluster B1 scaffold, dBET23 yielded 104 conformational degradation complexes (Figure S7). It is noteworthy that in each degradation machinery complex, at least one Lys residue was found within the cutoff criterion. However, due to the wide ring-open conformations of clusters B3 and B4, none of the modelled degradation complexes utilizing these two clusters satisfied our criterion. Given that a distance of 16 Å may appear to be relatively large for catalysis¹⁴, we postulated that the protein dynamics could potentially bring the relevant Lys and Gly residues into close proximity to facilitate the ubiquitination reaction.

Among the docked ensembles of each CRBN-dBETx-BRD4^{BD1}, all top-ranked ternary complexes from protein-protein docking were successfully integrated with the scaffold clusters A and B without encountering clashes (Figure S8). For example, dBET1_{#35} and dBET70_{#91} exhibited the most favorable protein-protein docking energies (Table S2). Intuitively, one might prioritize the ring-forming cluster A due to its smaller gap between the E2 and CRBN interface, potentially facilitating the proximity of Lys and Gly for ubiquitination. However, the conformational ensemble of each CRBN-dBETx-BRD4^{BD1} provided numerous possible orientations for assembly within the degradation complex, regardless of whether the CRL4A E3 ligase scaffold adopts a ring-forming or ring-open conformation.

Moreover, when constructing a degradation complex, all the highest-ranked ternary complexes from docking scores displayed similar conformations for dBETs, implying that the preferred PROTAC conformations may be predicted. These popular ternary complexes exhibited similar protein-protein contacts and linker conformations (Figure S9). Our MD simulations further demonstrated that these initial structures used for modeling degradation complexes effectively directed BRD4^{BD1} toward ubiquitination (see next subsection).

Protein structural dynamics in degradation machinery complex

Among the 1226 distinct degradation machinery complexes constructed, none showed the structure of a Lys residue of BRD4^{BD1} was positioned for ubiquitination. Therefore, for each PROTAC, we selected a degradation machinery complex conformation from clusters A and B with a top docking score CRBN-dBETx-BRD4^{BD1} ternary conformation for MD simulations (Table 1). To further evaluate the docking scores, we also performed protein-protein interaction energy calculations using MD runs initiated by these ternary conformations with top docking scores. The energy calculations confirmed the thermodynamic stability of top-ranked ternary complexes obtained through protein-protein docking, highlighting the robust binding of their associated dBETs (Figure S10). Moreover, they remained structurally stable throughout the MD runs, as evidenced by RMSD plots (see RMSD plots in Figure S11).

		Cluster A					Cluster B			
PROTAC	Conformations from protein–protein docking	Number of conformations for construction of degradation machinery complex								
		A1	A2	A3	A4	A5	B1	B2	B3	B4
dBET1	186	81	51	17	22	38	68	106	0	0
dBET23	168	42	52	87	0	23	104	86	0	0
dBET57	24	19	12	0	5	7	17	17	0	0
dBET70	183	37	73	73	0	10	96	83	0	0

Table 2.

Number of ternary complexes used in the molecular docking and construction of degradation machinery complex.

After confirming that each CRBN-dBETx-BRD4^{BD1} used for constructing the degradation complexes represented popular ternary conformations, we conducted classical MD simulations for these degradation complexes to observe structural dynamics and local interactions that lead to successful ubiquitination. Notably, to ensure that the compatibility of each CRBN-dBETx-BRD4^{BD1} with both clusters A and B of CRL4A E3 ligase scaffolds, we also selected a few ternary complexes from our MD runs (**Table 1**). As depicted in **Figure 1C**, a successful ubiquitination reaction requires protein conformational arrangements that bring a surface Lys of BRD4^{BD1} close to Gly of Ub, where at least one Asp of E2 positioned in proximity to the same Gly to facilitate charge transfer. Therefore, we employed two criteria: (1) ensuring that the distance between Lys (N atom) and Gly (C atom) was less than 10 Å; and (2) confirming that distance between Asp (O atom) of E2 and the same Gly (O atom) was less than 6 Å to determine the likelihood of ubiquitination.

Although none of the initial modelled machinery complexes exhibited a Lys residue of BRD4^{BD1} being ready for ubiquitination, our MD sampling facilitated the movement of different Lys residues of BRD4^{BD1} from distance greater than 16 Å of the catalytic pocket on E2 to establish close contacts with Gly of Ub for each dBET. Specifically, we observed significant large-scale motions within degradation complex B1_dBET1_#35 (cluster B, docked conformation #35), which recruited three Lys residues K72, K76 and K111 at distances greater than 16 Å, positioning them in the catalytic pocket on E2. Concurrently, a conserved charged residue Asp of E2 approached the C-terminus G75 of Ub, potentially facilitating isopeptide bond formation (**Figures 3A** and **B**). Similarly, degradation complex B1_dBET23_#14 (cluster B, docked conformation #14) and B1_dBET70_#91 (cluster B, docked conformation #91) also brought K99 of BRD4^{BD1} into proximity of the catalytic cavity of E2, where one Asp of E2 approached the the C-terminus G75 of Ub as well (**Figure 3C, D, G and H**).

In contrast, degradation complex B1_dBET57_#9MD1 (cluster B, conformation from MD run 1 initiated with docked conformation #9) positioned K102 of BRD4^{BD1} in close proximity to the catalytic pocket on E2 (**Figures 3E** and **3F**). Notably, BRD4^{BD1} of dBET1, dBET23 and dBET70 exhibited a closer proximity to Ub (~6 Å) as compared to dBET57 (~10 Å) (**Figure 3A, C, E and G**). Despite degradation complex B1_dBET1_#35 positioned three Lys residues of BRD4^{BD1} to form stable interactions with the Gly of Ub, the surrounding Asp of E2 was unstable which hindered a successful ubiquitination process (**Figure 3A** and S11). Further analysis revealed that dBET23 (~85%) and dBET70 (~78%) exhibited a higher probability of fulfilling both criteria for potential chemical reactions, whereas dBET1 (~40%) and dBET57 (~22%) had a lower likelihood (Figure S12D). This analysis revealed how dBET23 and dBET70 achieved high degradation efficiency, while dBET1 and dBET57 induced protein motions for bond formation, without achieving all required arrangements simultaneously, resulting in lower degradation efficiency. Our findings are consistent with experimental measurements indicating that all dBETs could degrade BRD4^{BD1} but with varying degradation potency. By considering the Lys–Gly and Asp–Gly distances, our simulations underscored that dBET23 and dBET70 (with DC_{50/5h} values of ~5 nM and ~50 nM, respectively) are more efficient degraders compared to dBET1 and dBET57 (with DC_{50/5h} values of ~500 nM).

To gain further insights into the overall motion of the degradation complexes, we employed principal component analysis (PCA) to extract essential motions within the complex; Notably, the first two PC modes, PC1 and PC2, accounted over 60% of the overall motions in most MD runs (Table S4). While the specific nature of these motions varied depending on the system, the dBETs induced an inter-protein orientation between CRBN and BRD4^{BD1}, with two hinge regions: loops in Rbx1 and DDB1 proteins contributing prominently to the essential motions of the complex (Figure S13). These two hinge regions exhibited twisted and opposite direction motions, facilitating the rearrangement and bringing BRD4^{BD1} closer to the catalytic cavity of E2, which is a critical step for successful ubiquitination. Indeed, the essential motions revealed by the first two PC modes clearly revealed the plasticity between the protein–protein interfaces in the degradation machinery. Notably, the most significant motion in PC1 shifted E2/Ub/Rbx1/NEDD8 closer to BRD4^{BD1} (Figure

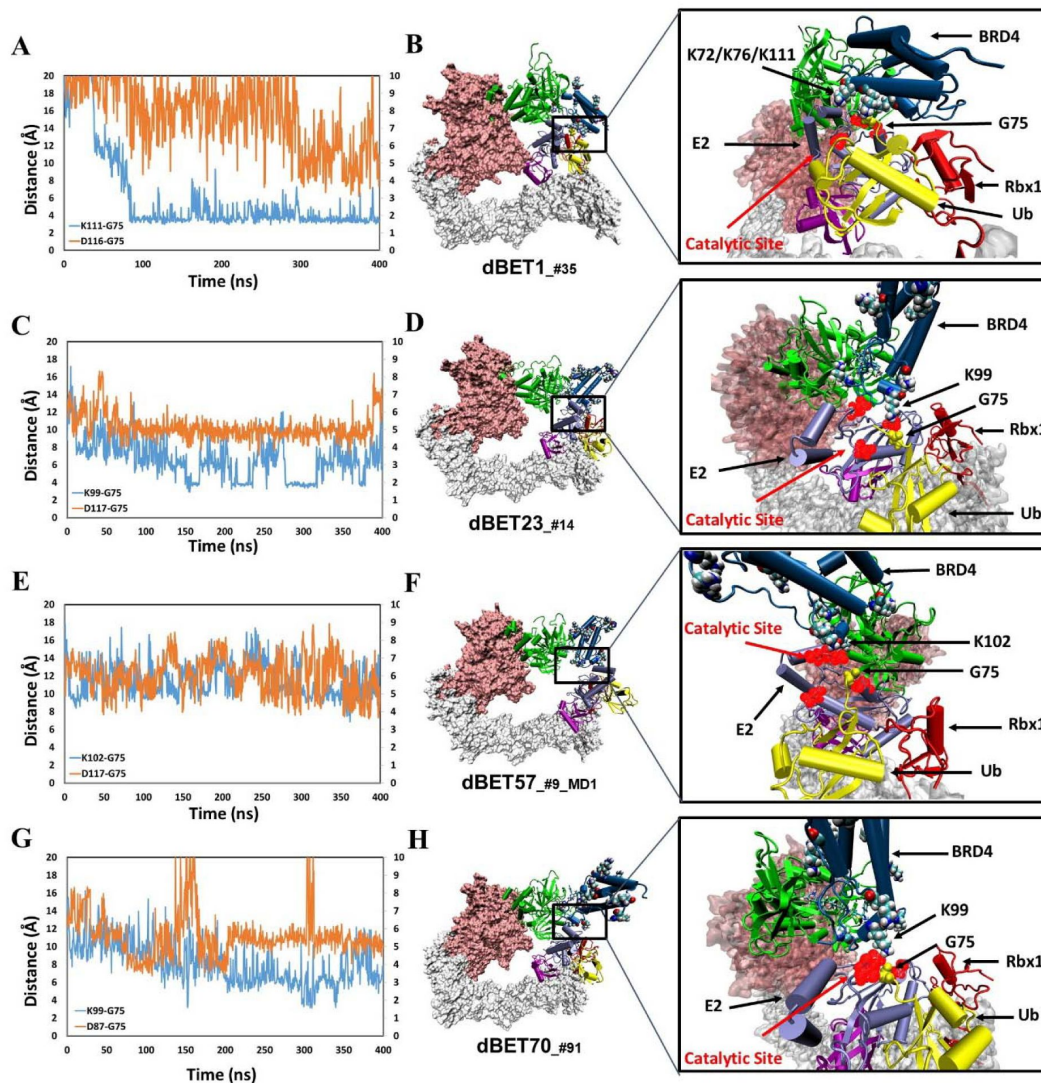


Figure 3.

Structural dynamics and local interaction of each degradation machinery.

(A) K111 and D116 of dBET1_#35 (Distance of K72/76 is reported in Figure S20). The left y-axis in blue presents the distance between Lys (N atom) of BRD4^{BD1} and Gly (C atom) of Ub, and the right y-axis in red presents the distance between Asp (O atom) of E2 and Gly (O atom) of Ub. The same left and right y-axis representations are used in C, E and G. (B) K72/76/111 of BRD4^{BD1} reaches the negatively charged catalytic site of E2. (C) K99 and D117 of dBET23_#14. (D) K99 of BRD4^{BD1} reaches the negatively charged catalytic site of E2. (E) K102 and D117 of dBET57_#9_MD1. (F) K102 of BRD4^{BD1} reaches the negatively charged catalytic site of E2. (G) K99 and D87 of dBET70_#91. (H) K99 of BRD4^{BD1} reaches the negatively charged catalytic site of E2. (refer to Figure S21 for data of second seed of MD simulation)

S14). Furthermore, the pairwise force matrix between individual residues within the degradation complex offered insights into the non-covalent interactions governing the dynamic behavior of the complex. This interaction network, starting from the dBETs' linker and extending to the hinge loop of DDB1 and Rbx1 (Figure S15), indicated a clear correlation between linker rotation and the movement of the entire degradation complex. Our study highlights the importance of structural dynamics of the degradation complex during the ubiquitination processes, which is influenced by a dBET ligand.

To quantify the correlation between linker rotation and the movement of the degradation complex, we selected four atoms shown in **Figure 4A** to construct a pseudo dihedral angle and used histograms to depict the population of these pseudo angles. Deviations between the histograms from different time periods indicated the motion of the degradation complex during the MD simulations. Across the first and last 100 ns of an MD run, the degradation complexes large motions with the presence of each dBET (**Figure 4B**). Given that the only distinction among the four dBETs was the linker, we observed that the complex motions correlated to the C-C-N3-C dihedral angles in the linkers of dBET1, dBET23, dBET70, and dBET57 (**Figure 5A**). Notably, all four dBETs exhibited substantial shifts of BRD4^{BD1} of approximately 10 and 15 degrees, which underscores their flexibility in promoting extensive motion of the BRD4^{BD1} protein for Lys–Gly interaction. The various motions from different degradation complexes also resulted in different sets of Lys residues potentially forming contacts with E2 and Ub. For instance, the degradation complexes dBET23_{#14} and dBET70_{#91} brought K99 of BRD4^{BD1} closer to G75 of Ub (**Figure 5B**), while dBET57_{#9MD1} placed K102 in proximity to G75 of Ub. Despite allowing for different Lys–Gly interactions, dBET57 exhibited weaker interactions. The pairwise force analysis revealed a strong attraction between K99 and G75 of dBET23 and dBET70, and a weaker attraction between K102 and G75 of dBET57_{#9MD1} (Figure S16).

To further examine the correlation between PROTAC rotation and the Lys–Gly interaction, we performed a time-dependent correlation analysis. This analysis showed that PROTAC rotation translates motion over time, leading to the Lys–Gly interaction, with a correlation peak around 60–85 ns, marking the time of the interaction (**Figure 6** and Figure S17). In addition, the pseudo dihedral angles also showed a high correlation (0.85 in the case of dBET1) with Lys–Gly distance. This indicated that degradation complex undergoes structural rearrangement and drives the Lys–Gly interaction.

Given the presence of numerous local energy minima conformations in the degradation complex, MD simulations could not sample all conformations comprehensively. Furthermore, as evidenced by the protein–protein docking results (Figure S1C), the short linker of dBET57 yielded a less flexible ternary complex, suggesting challenges in sampling various conformations using MD runs. Our dihedral entropies analysis showed that dBET57 has ~0.3 kcal/mol lower entropies than the other dBETs with three different linkers, suggesting that dBET57 is less flexible than other PROTACs (Figure S18). To address this, we selected another distinctly different ternary complex of dBET57 (Figure S19B) to construct another degradation complex, B1_dBET57_{#9_MD2}, for MD simulation. Unlike B1_dBET57_{#9_MD1}, B1_dBET57_{#9_MD2} exhibited significantly more rigidity, with only minimal rotation (< 5 degrees) in the pseudo dihedral angle. This rigidity limits the structural dynamics of the whole complex to bring BRD4^{BD1} closer to Ub for Lys–Gly interactions (Figure S19). This observation underscores the crucial role of the linker in guiding the motion of the degradation complex, thereby influencing its capacity for ubiquitination. In summary, the linker induces conformations of ternary complexes but also guides the motion of the degradation complex, which allows an effective Lys–Gly interaction for ubiquitination. Our studies highlight the critical role of linker flexibility in facilitating protein structural dynamics necessary for ubiquitination.

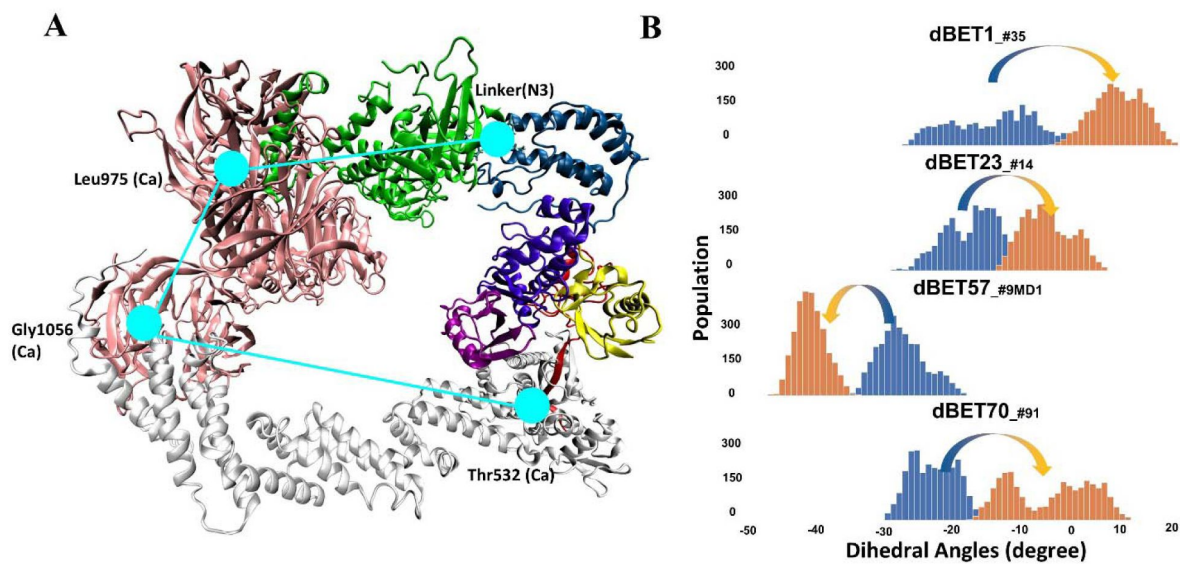


Figure 4.

Quantifying the motion of degradation complex.

(A) Defining the pseudo dihedral angles, Linker(N3)-Leu975(Ca)-Gly1056(Ca)-Thr532(Ca) to capture the essential motion of the degradation complex. **(B)** Dihedral angle histogram shows the population distribution of the first 100 ns (blue) and last 100 ns (orange) of the MD simulation. The distribution shows a ~10-15° dihedral angle shift.

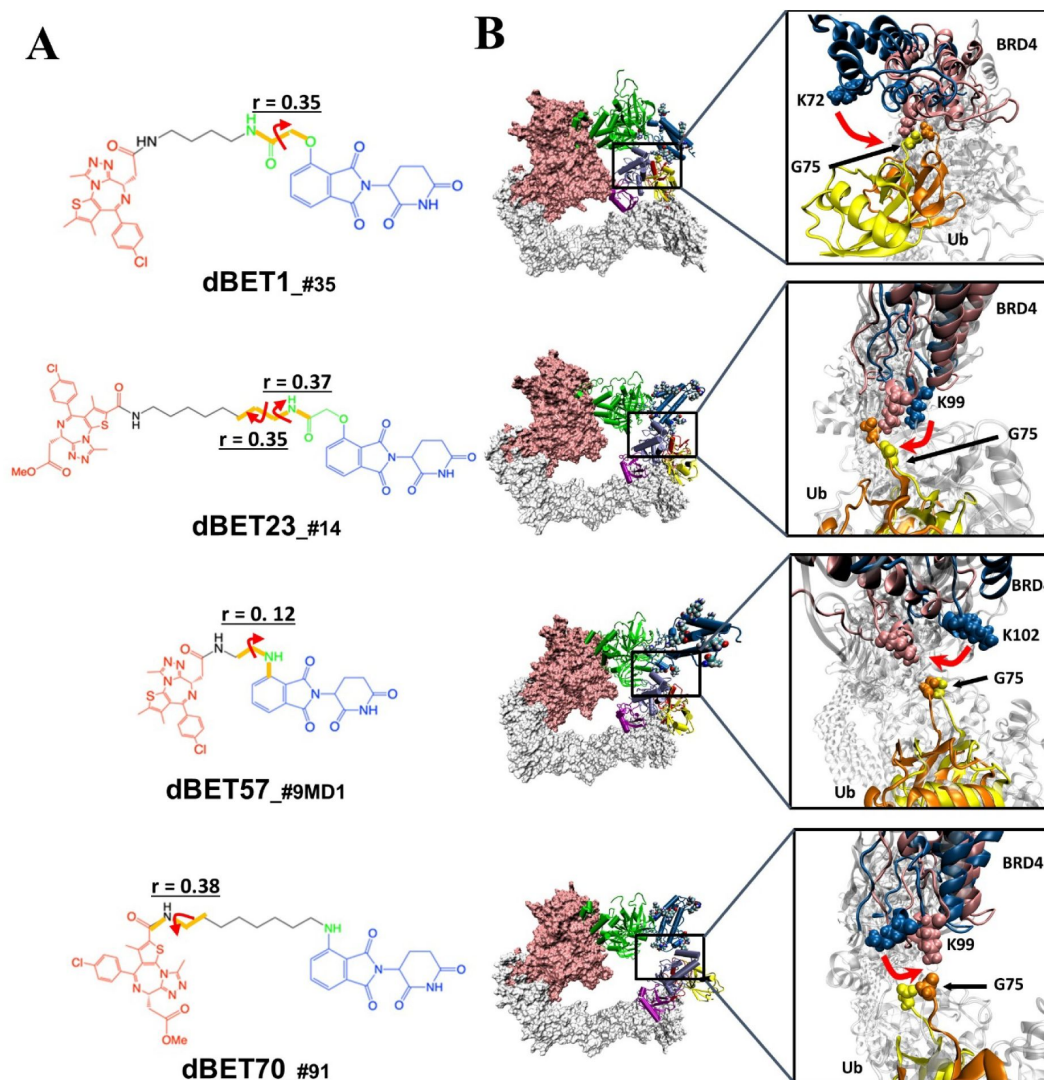


Figure 5.

Quantification of the correlated motion between the linker and degradation complex.

(A) Dihedral angles correlated with the degradation complex. r is the dihedral correlation coefficient (See Methods for detailed calculation). **(B)** Motion of BRD4^{BD1} and Ub indicates a shift between the initial frames (BRD4^{BD1}, blue; Ub, yellow) and final frames (BRD4^{BD1}, pink; Ub, orange) of MD simulation. K72, K99 and K102 engaged in the interaction with G75 of Ub in each degradation complex, which implies their potential for degrading BRD4^{BD1}. For visualization purposes we present only K72 of dBET1_#35, but K76 and K111 can also engage in interaction with G75 (See [Figure 3](#) [↗](#)).

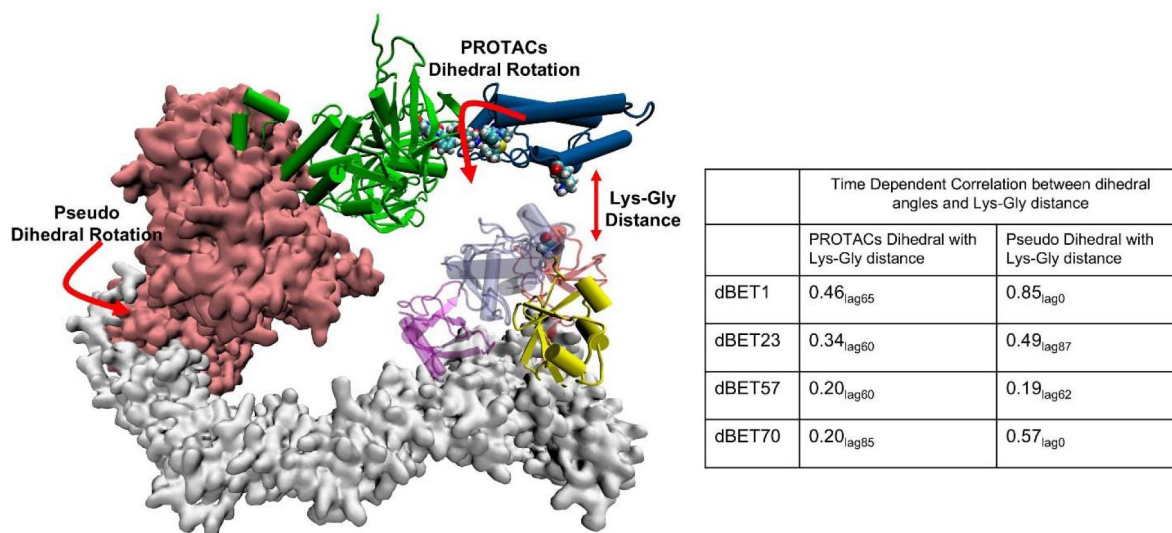


Figure 6.

Time dependent correlation between PROTACs/Pseudo dihedral angles and Lys-Gly interaction.

The peak values of time dependent correlation were reported. For instance, 0.46_{lag65} represents the correlation coefficient at 65ns of the MD simulation.

Discussion

As bifunctional small-molecule ligands, PROTACs play a pivotal role in recruiting a neo-substrate of E3 ligase, namely the target protein POI, to form a ternary E3-PROTAC-POI complex by stabilizing the protein–protein interactions. By forming an E3-PROTAC-POI complex, E3 ligase brings the POI close to an E2 enzyme for protein ubiquitination, with both E2 and E3 proteins belonging to a larger Ub ligase machinery complex. While forming a stable E3-PROTAC-POI complex represents a critical step in ubiquitination, it alone is insufficient for subsequent ubiquitination and degradation by proteasomes. Several studies suggest that positioning the POI's Lys residues close to the E2 for ubiquitination is crucial for successful POI degradation^{11,24,35}. However, due to the lack of experimentally determined structures for POI-bound degradation complexes, directly observing the proximity of Ub to POI's Lys residues remains challenging. In addition, proteins exhibit dynamic behaviors, and the fluctuation in the ternary E3-PROTAC-POI complex and the degradation machine can either promote or impede catalytic activity. To investigate the degradability of PROTACs, we employed a strategy combining protein docking, MD simulations, and structural and energy analysis tools. Our focus was on a series of dBET degraders, differing in their linkers, which all induced stable CRBN-dBETx-BRD4^{BD1} complexes with different degradation efficiencies.

One straightforward approach to promote ubiquitination is to obtain a stable E3-PROTAC-POI ternary complex to allow pre-organized orientation placing a Lys residue(s) of the POI close to the E2/Ub catalytic site in the degradation machine. However, experimentally determined structures for degradation complexes are scarce. Protein–protein docking can reveal low energy ternary complexes and popular conformational ensembles, providing insights into potential orientations for ubiquitination. To observe whether the preferred conformations are pre-organized for ubiquitination. Although protein–protein docking is usually quick (i.e., <50 hr to sample 5000 protein–protein complexes using one CPU in computation), the docked E3-PROTAC-POI conformations may not be the functionally active form to induce successful ubiquitination. Nevertheless, results from protein–protein docking are highly informative and suggest potential flexibility of a ternary complex, as shown in Figure S1. However, MD simulations or other conformational search methods can further generate more stable ternary conformations, and the sampling from protein–protein docking may not be thorough. As illustrated by dBET57, protein–protein docking generated only similar ternary complex conformations (Figure S1C). In contrast, classical MD simulations using all atoms and explicit water molecules sampled additional low energy conformations (Figure 3C).

Using the protein–protein docking results, multiple thermodynamically stable ternary E3-PROTAC-POI conformations were placed in the degradation machine to assess whether surface Lys residues of POI were appropriately oriented for ubiquitination. However, relying solely on a handful of static conformations for predicting POI degradability is not insufficient. For example, our docking results showed that none of the Lys residues of BRD4^{BD1} was ready for ubiquitination. Notably, our modelled CRL4A E3 ligase scaffolds in Figure S6 show two distinct conformations: a ring-forming conformation (cluster A) with a gap distance approximately 1.0 Å to 10.0 Å, and a ring-open conformation (cluster B) with a much larger gap distance, approximately 14 Å to 35 Å. When docking the ternary complex into the ring-forming scaffold (Figure S6A), the orientations of BRD4^{BD1} tended to be sterically confined, limiting their range of motion to bring one Lys residue of BRD4^{BD1} closer to the Gly of Ub. In contrast, using MD to simulate protein dynamics, CRL4A E3 ligase scaffolds with ring-open conformations provided a larger space that may allow an assembled BRD4^{BD1} to properly rearrange its conformations for successful ubiquitination, even though the initial distance between BRD4^{BD1} and Ub is far. Our studies demonstrated that using CRBN-dBETs-BRD4^{BD1} conformations sampled by protein–protein docking and/or MD simulations provided reasonably good initial structures to model the structural dynamics of the degradation

machinery. The essential motions were captured by PCA, which illustrated that all dBETs studied can produce motions to bring BRD4^{BD1} closer to Ub, thus supporting their degradability. Atomistic simulations revealed that specific Lys residue(s) of BRD4^{BD1} (i.e., K99, K102 and K72/76/111 of dBET23 and dBET70, dBET57 and dBET1, respectively) were attracted by polar residues around Ub and E2.

The analysis of how a PROTAC, particularly the linker region, affects the overall dynamics and specific arrangements to bring Lys close to the catalytic site offers further information on linker design. The slight variations in linker length and composition that may significantly alter the degradability of a PROTAC was puzzling. Even though PROTACs can strengthen interactions between E3 and POI, protein degradation is not always associated with forming stable ternary complexes. One hypothesis is that some ternary complexes have a preferred E3/POI orientation in the degradation machinery complex to result in efficient ubiquitination. Here we demonstrated that by using accurately modeled initial conformations, performing a few hundred nanosecond classical MD simulations could illustrate how protein structural dynamics orient surface Lys residues of BRD4^{BD1} and Asp of E2 close to the Gly of Ub, thus offering information for predicting protein degradation. We examined the hinges and used pseudo-dihedral angles of the degradation complexes to describe the large-scale swinging motions near the end of the MD runs (300–400 ns). To understand the role of the linker of different dBETs, we compared conformation ensembles from the first and last 100 ns of MD simulations. The movement of the degradation machinery correlated with rotations of specific dihedrals of the linker region in dBETs (**Figure 5** [↗](#)). In addition, the, dihedral rotation of PROTACs shows correlation with the Lys-Gly interaction (**Figure 6** [↗](#) and Figure S17). This phenomenon shows that dBETs contribute to the structural dynamics of degradation complexes and suggests that linkers with different properties may lead to various large-scale motions.

The dynamics of the degradation machinery, induced by PROTACs, has implications in linker design and degradation prediction. Because the “warhead” (thieno-triazolo-1,4-diazepine [JQ1]) used to bind to an E3 ligase and the “warhead” (pomalidomide) used to bind to a POI are tightly bound ligands to each protein, the two warheads typically have limited motions when they are in the binding site of E3 or POI. Clearly, the length and composition of the linker are critical in degradation efficiency, and sites of conjugation to each warhead also affect the stability of the ternary complexes. Despite no general rules that apply to every system for PROTAC linker design, guidelines used in designing small-molecule drugs binding to their target proteins still hold. First, the linker should have sufficient length to form a ternary complex. Although longer and more flexible linkers typically provide enhanced opportunities for E3-PROTAC-POI complexes to orient suitable conformations for ubiquitination, shorter and/or rigid linkers may provide a preorganized E3-PROTAC-POI for efficient catalysis. However, knowing in advance the best related orientations is challenging, but the shorter linker in our systems, dBET57, had disadvantages in accommodating a Lys residue in the E2/Ub catalytic region for forming stable interactions for catalysis. In contrast, dBET23 and dBET70, with greater flexibility, allowed BRD4^{BD1} to adopt diverse conformations for catalysis, ultimately leading to more efficient degradation. Notably, although dBET1 recruited multiple Lys residues in the E2/Ub catalytic region and K111 was highly stable. The unstable Asp from the surrounding prevented charge transfer, and thus hindered the isopeptide bond formation for degradation. Notably, dBET1 and dBET57 still induce stable CRBN-dBET57-BRD4^{BD1} and CRBN-dBET1-BRD4^{BD1} complexes, which effectively assemble the neo-substrate BRD4^{BD1} with the degradation machine. However, the either Lys-Gly or Asp-Gly attractive force was unstable, which reduced the chances for successful ubiquitination.

Revealing structural dynamics of the entire degradation complex that bring surface Lys residue(s) of BRD4^{BD1} for target ubiquitination explains the degradability of these dBETs. Notably, the classical MD trajectories with atomistic details offer information about overall protein motions and specific Lys residues critical for the degradability of BRD4^{BD1} for further investigation. Theoretically speaking, excessively long MD simulations should be able to sample a wide range of

protein conformations. However, long MD simulations can still stick in local minima and are computationally expensive. We demonstrated the use of the most stable ternary CRBN-dBETx-BRD4^{BD1} conformations from protein-protein docking to facilitate sampling the correct orientation of the surface Lys residue(s) of BRD4^{BD1} for ubiquitination. Of note, CRBN-dBET57-BRD4^{BD1} formed by the more rigid linker of dBET57 could be more easily trapped in a local energy minimum, and we need to select conformations from both protein-docking (dBET57_{#9}) and MD-sampling results (CRBN-dBET57_{#9}-BRD4^{BD1}) to construct the degradation machinery complex for structural dynamics studies. Different E3 ligases may also result in significantly different related orientation when bound to the same POI. Although static conformations from experimentally determined structures may suggest whether a dBET orients CRBN-dBETx-BRD4^{BD1} conformations suitable for degradation, real motions cannot be predicted directly from static conformations. Examining the dynamic nature of the degradation complex provides a more complete picture of the conformations of both ternary CRBN-dBETx-BRD4^{BD1} and machinery complexes that determine Lys accessibility for POI ubiquitination, an important factor for consideration when optimizing a PROTAC to improve its potency.

Acknowledgements

We thank Drs. Eric Fischer and Radosław Nowak for helpful discussions and for providing experimental data for protein binding and degradation efficiency, and Dr. Kevin Kou for helpful discussions regarding the mechanisms of isopeptide bond formation. This publication was made possible by Grant No. GM109045 and GM151651 from the National Institute of General Medical Sciences (NIGMS) of the NIH.

Additional information

Synopsis

Our study showcased mechanistic insights with atomistic details into PROTAC-induced protein structural dynamics within degradation machinery complex with multi-level modeling approaches.

Additional files

Supporting Material 

References

- (1) Dale B., Cheng M., Park K.-S., Kaniskan H. Ü., Xiong Y., Jin J (2021) **Advancing Targeted Protein Degradation for Cancer Therapy** *Nat Rev Cancer* **21**:638–654
- (2) Békés M., Langley D. R., Crews C. M (2022) **PROTAC Targeted Protein Degraders: The Past Is Prologue** *Nat Rev Drug Discov* **21**:181–200
- (3) Ramachandran S., Ciulli A (2021) **Building Ubiquitination Machineryes: E3 Ligase Multi-Subunit Assembly and Substrate Targeting by PROTACs and Molecular Glues** *Curr Opin Struct Biol* **67**:110–119 <https://doi.org/10.1016/j.sbi.2020.10.009>
- (4) Bondeson D. P., Smith B. E., Burslem G. M., Buhimschi A. D., Hines J., Jaime-Figueroa S., Wang J., Hamman B. D., Ishchenko A., Crews C. M (2018) **Lessons in PROTAC Design from Selective Degradation with a Promiscuous Warhead** *Cell Chem Biol* **25**:78–87
- (5) Huang H.-T., Dobrovolsky D., Paulk J., Yang G., Weisberg E. L., Doctor Z. M., Buckley D. L., Cho J.-H., Ko E., Jang J (2018) **A Chemoproteomic Approach to Query the Degradable Kinome Using a Multi-Kinase Degradator** *Cell Chem Biol* **25**:88–99
- (6) Nowak R. P. *et al.* (2018) **Plasticity in Binding Confers Selectivity in Ligand-Induced Protein Degradation** *Nat Chem Biol* **14**:706–714 <https://doi.org/10.1038/s41589-018-0055-y>
- (7) Zhao Q., Lan T., Su S., Rao Y (2019) **Induction of Apoptosis in MDA-MB-231 Breast Cancer Cells by a PARP1-Targeting PROTAC Small Molecule** *Chemical Communications* **55**:369–372 <https://doi.org/10.1039/C8CC07813K>
- (8) Tinworth C. P. *et al.* (2019) **PROTAC-Mediated Degradation of Bruton's Tyrosine Kinase Is Inhibited by Covalent Binding** *ACS Chem Biol* **14**:342–347 <https://doi.org/10.1021/acscchembio.8b01094>
- (9) Ward C. C. *et al.* (2019) **Covalent Ligand Screening Uncovers a RNF4 E3 Ligase Recruiter for Targeted Protein Degradation Applications** *ACS Chem Biol* **14**:2430–2440 <https://doi.org/10.1021/acscchembio.8b01083>
- (10) Ottis P., Toure M., Cromm P. M., Ko E., Gustafson J. L., Crews C. M (2017) **Assessing Different E3 Ligases for Small Molecule Induced Protein Ubiquitination and Degradation** *ACS Chem Biol* **12**:2570–2578 <https://doi.org/10.1021/acscchembio.7b00485>
- (11) Smith B. E., Wang S. L., Jaime-Figueroa S., Harbin A., Wang J., Hamman B. D., Crews C. M (2019) **Differential PROTAC Substrate Specificity Dictated by Orientation of Recruited E3 Ligase** *Nat Commun* **10**:1–13
- (12) Lai A. C., Toure M., Hellerschmied D., Salami J., Jaime-Figueroa S., Ko E., Hines J., Crews C. M (2016) **Modular PROTAC Design for the Degradation of Oncogenic BCR-ABL** *Angewandte Chemie International Edition* **55**:807–810 <https://doi.org/10.1002/anie.201507634>
- (13) Han B (2020) **A Suite of Mathematical Solutions to Describe Ternary Complex Formation and Their Application to Targeted Protein Degradation by Heterobifunctional Ligands** *Journal of Biological Chemistry* **295**:15280–15291 <https://doi.org/10.1074/jbc.RA120.014715>

- (14) Bai N., Riching K. M., Makaju A., Wu H., Acker T. M., Ou S.-C., Zhang Y., Shen X., Bulloch D., Rui H (2022) **Modeling the CRL4A Ligase Complex to Predict Target Protein Ubiquitination Induced by Cereblon-Recruiting PROTACs** *Journal of Biological Chemistry* **101**:653
- (15) Drummond M. L., Williams C. I (2019) **In Silico Modeling of PROTAC-Mediated Ternary Complexes: Validation and Application** *J Chem Inf Model* **59**:1634–1644
- (16) Drummond M. L., Henry A., Li H., Williams C. I (2020) **Improved Accuracy for Modeling PROTAC-Mediated Ternary Complex Formation and Targeted Protein Degradation via New in Silico Methodologies** *J Chem Inf Model* **60**:5234–5254
- (17) Zaidman D., Prilusky J., London N (2020) **PROsettaC: Rosetta Based Modeling of PROTAC Mediated Ternary Complexes** *J Chem Inf Model* **60**:4894–4903
- (18) Bai N., Miller S. A., Andrianov G. V., Yates M., Kirubakaran P., Karanickolas J (2021) **Rationalizing PROTAC-Mediated Ternary Complex Formation Using Rosetta** *J Chem Inf Model* **61**:1368–1382 <https://doi.org/10.1021/acs.jcim.0c01451>
- (19) Rao A., Tunjic T. M., Brunsteiner M., Müller M., Fooladi H., Gasbarri C., Weber N (2023) **Bayesian Optimization for Ternary Complex Prediction (BOTCP)** *Artificial Intelligence in the Life Sciences* **3** <https://doi.org/10.1016/j.ailsci.2023.100072>
- (20) Zheng S., Tan Y., Wang Z., Li C., Zhang Z., Sang X., Chen H., Yang Y (2022) **Accelerated Rational PROTAC Design via Deep Learning and Molecular Simulations** *Nat Mach Intell* **4**:739–748
- (21) Ignatov M. *et al.* (2023) **High Accuracy Prediction of PROTAC Complex Structures** *J Am Chem Soc* **145**:7123–7135 <https://doi.org/10.1021/jacs.2c09387>
- (22) Pereira G. P., Jiménez-García B., Pellarin R., Launay G., Wu S., Martin J., Souza P. C. T (2023) **Rational Prediction of PROTAC-Compatible Protein–Protein Interfaces by Molecular Docking** *J Chem Inf Model* <https://doi.org/10.1021/acs.jcim.3c01154>
- (23) Lv D., Pal P., Liu X., Jia Y., Thummuri D., Zhang P., Hu W., Pei J., Zhang Q., Zhou S (2021) **Development of a BCL-XL and BCL-2 Dual Degradator with Improved Anti-Leukemic Activity** *Nat Commun* **12**:1–14
- (24) Dixon T. *et al.* (2022) **Predicting the Structural Basis of Targeted Protein Degradation by Integrating Molecular Dynamics Simulations with Structural Mass Spectrometry** *Nat Commun* **13** <https://doi.org/10.1038/s41467-022-33575-4>
- (25) Liu Z. *et al.* (2022) **An Overview of PROTACs: A Promising Drug Discovery Paradigm** *Molecular Biomedicine* **3** <https://doi.org/10.1186/s43556-022-00112-0>
- (26) Bricelj A., Steinebach C., Kuchta R., Gütschow M., Sosič I (2021) **E3 Ligase Ligands in Successful PROTACs: An Overview of Syntheses and Linker Attachment Points** *Front Chem* **9** <https://doi.org/10.3389/fchem.2021.707317>
- (27) Lee J., Lee Y., Jung Y. M., Park J. H., Yoo H. S., Park J (2022) **Discovery of E3 Ligase Ligands for Target Protein Degradation** *Molecules* **27** <https://doi.org/10.3390/molecules27196515>
- (28) Zhou B., Hu J., Xu F., Chen Z., Bai L., Fernandez-Salas E., Lin M., Liu L., Yang C.-Y., Zhao Y (2018) **Discovery of a Small-Molecule Degradator of Bromodomain and Extra-Terminal (BET) Proteins with Picomolar Cellular Potencies and Capable of Achieving Tumor Regression** *J Med Chem* **61**:462–481

- (29) Lu J., Qian Y., Altieri M., Dong H., Wang J., Raina K., Hines J., Winkler J. D., Crew A. P., Coleman K (2015) **Hijacking the E3 Ubiquitin Ligase Cereblon to Efficiently Target BRD4** *Chem Biol* **22**:755–763
- (30) Filippakopoulos P., Qi J., Picaud S., Shen Y., Smith W. B., Fedorov O., Morse E. M., Keates T., Hickman T. T., Felletar I (2010) **Selective Inhibition of BET Bromodomains** *Nature* **468**:1067–1073
- (31) Winter G. E., Buckley D. L., Paulk J., Roberts J. M., Souza A., Dhe-Paganon S., Bradner J. E (2015) **Phthalimide Conjugation as a Strategy for in Vivo Target Protein Degradation** *Science* **348**:1376–1381
- (32) Qin C., Hu Y., Zhou B., Fernandez-Salas E., Yang C.-Y., Liu L., McEachern D., Przybranowski S., Wang M., Stuckey J (2018) **Discovery of QCA570 as an Exceptionally Potent and Efficacious Proteolysis Targeting Chimera (PROTAC) Degradator of the Bromodomain and Extra-Terminal (BET) Proteins Capable of Inducing Complete and Durable Tumor Regression** *J Med Chem* **61**:6685–6704
- (33) Zhang W., Burman S. S. R., Chen J., Donovan K. A., Cao Y., Shu C., Zhang B., Zeng Z., Gu S., Zhang Y (2022) **Machine Learning Modeling of Protein-Intrinsic Features Predicts Tractability of Targeted Protein Degradation** *Genomics Proteomics Bioinformatics*
- (34) Valimberti I., Tiberti M., Lambrughli M., Sarcevic B., Papaleo E (2015) **E2 Superfamily of Ubiquitin-Conjugating Enzymes: Constitutively Active or Activated through Phosphorylation in the Catalytic Cleft** *Sci Rep* **5**:1–13
- (35) Petzold G., Fischer E. S., Thomä N. H (2016) **Structural Basis of Lenalidomide-Induced CK1α Degradation by the CRL4CRBN Ubiquitin Ligase** *Nature* **532**:127–130

Author information

Kingsley Y Wu[#]

Department of Chemistry, University of California, Riverside, Riverside, United States
ORCID iD: [0000-0002-7989-1584](https://orcid.org/0000-0002-7989-1584)

[#]These authors contributed equally to this work

Ta I Hung[#]

Department of Chemistry, University of California, Riverside, Riverside, United States,
Department of Bioengineering, University of California, Riverside, Riverside, United States

[#]These authors contributed equally to this work

Chia-en A Chang

Department of Chemistry, University of California, Riverside, Riverside, United States
ORCID iD: [0000-0002-6504-8529](https://orcid.org/0000-0002-6504-8529)

For correspondence: chiaenc@ucr.edu

Editors

Reviewing Editor

Donald Hamelberg

Georgia State University, Atlanta, United States of America

Senior Editor

Qiang Cui

Boston University, Boston, United States of America

Reviewer #1 (Public review):

This study by Wu et al. provides valuable computational insights into PROTAC-related protein complexes, focusing on linker roles, protein-protein interaction stability, and lysine residue accessibility. The findings are significant for PROTAC development in cancer treatment, particularly breast and prostate cancers.

Strengths:

- (1) Comprehensive computational analysis of PROTAC-related protein complexes.
- (2) Focus on critical aspects: linker role, protein-protein interaction stability, and lysine accessibility.

Weaknesses:

- (1) Limited examination of lysine accessibility despite its stated importance.
- (2) Use of RMSD as the primary metric for conformational assessment, which may overlook important local structural changes.

The authors' claims about the role of PROTAC linkers and protein-protein interaction stability are generally supported by their computational data. However, the conclusions regarding lysine accessibility could be strengthened with more in-depth analysis. The use of the term "protein functional dynamics" is not fully justified by the presented work, which focuses primarily on structural dynamics rather than functional aspects.

Comments on revisions:

The authors have addressed the questions raised substantially.

<https://doi.org/10.7554/eLife.101127.2.sa3>

Reviewer #2 (Public review):

Summary:

The manuscript reports the computational study of the dynamics of PROTAC-induced degradation complexes. The research investigates how different linkers within PROTACs affect the formation and stability of ternary complexes between the target protein BRD4BD1 and Cereblon E3 ligase, and the degradation machinery. Using computational modeling, docking, and molecular dynamics simulations, the study demonstrates that although all PROTACs form ternary complexes, the linkers significantly influence the dynamics and efficacy of protein degradation. The findings highlight that the flexibility and positioning of Lys residues are crucial for successful ubiquitination. The results also discussed the correlated motions between the PROTAC linker and the complex.

Strengths:

The field of PROTAC discovery and design, characterized by its limited research, distinguishes itself from traditional binary ligand-protein interactions by forming a ternary complex involving two proteins. The current understanding of how the structure of PROTAC influences its degradation efficacy remains insufficient. This study investigated the atomic-level dynamics of the degradation complex, offering potentially valuable insights for future research into PROTAC degradability.

Comments on revisions:

All my questions have been addressed.

<https://doi.org/10.7554/eLife.101127.2.sa2>

Reviewer #3 (Public review):

The authors offer an interesting computational study on the dynamics of PROTAC-driven protein degradation. They employed a combination of protein-protein docking, structural alignment, atomistic MD simulations, and post-analysis to model a series of CRBN-dBET-BRD4 ternary complexes, as well as the entire degradation machinery complex. These degraders, with different linker properties, were all capable of forming stable ternary complexes but had been shown experimentally to exhibit different degradation capabilities. While in the initial models of the degradation machinery complex, no surface Lys residue(s) of BRD4 were exposed sufficiently for the crucial ubiquitination step, MD simulations illustrated protein functional dynamics of the entire complex and local side-chain arrangements to bring Lys residue(s) to the catalytic pocket of E2/Ub for reactions. Using these simulations, the authors were able to present a hypothesis as to how linker property affects degradation potency. They were able to roughly correlate the distance of Lys residues to the catalytic pocket of E2/Ub with observed DC50/5h values. This is an interesting and timely study that presents interesting tools that could be used to guide future PROTAC design or optimization.

<https://doi.org/10.7554/eLife.101127.2.sa1>

Author response:

The following is the authors' response to the original reviews.

Reviewer #1 (Public review):

This study by Wu et al. provides valuable computational insights into PROTAC-related protein complexes, focusing on linker roles, protein-protein interaction stability, and lysine residue accessibility. The findings are significant for PROTAC development in cancer treatment, particularly breast and prostate cancers.

The authors' claims about the role of PROTAC linkers and protein-protein interaction stability are generally supported by their computational data. However, the conclusions regarding lysine accessibility could be strengthened with more in-depth analysis. The use of the term "protein functional dynamics" is not fully justified by the presented work, which focuses primarily on structural dynamics rather than functional aspects.

Strengths:

(1) Comprehensive computational analysis of PROTAC-related protein complexes.

(2) Focus on critical aspects: linker role, protein-protein interaction stability, and lysine accessibility.

Weaknesses:

(1) Limited examination of lysine accessibility despite its stated importance.

(2) Use of RMSD as the primary metric for conformational assessment, which may overlook important local structural changes.

Reviewer #1 (Recommendations for the authors):

(1) The authors' claims about the role of PROTAC linkers and protein-protein interaction stability are generally supported by their computational data. However, the conclusions regarding lysine accessibility could be strengthened with more in-depth analysis. Expand the analysis of lysine accessibility, potentially correlating it with other structural features such as linker length.

We thank the reviewers for the suggestions! We performed time dependent correlation analysis to correlate the dihedral angles of the PROTACs and the Lys-Gly distance (Figures 6 and S17). We included detailed explanation on page 16:

“To further examine the correlation between PROTAC rotation and the Lys-Gly interaction, we performed a time-dependent correlation analysis. This analysis showed that PROTAC rotation translates motion over time, leading to the Lys-Gly interaction, with a correlation peak around 60-85 ns, marking the time of the interaction (Figure 6 and Figure S17). In addition, the pseudo dihedral angles also showed a high correlation (0.85 in the case of dBET1) with Lys-Gly distance. This indicated that degradation complex undergoes structural rearrangement and drives the Lys-Gly interaction.”

(2) The use of the term "protein functional dynamics" is not fully justified by the presented work, which focuses primarily on structural dynamics rather than functional aspects. Consider changing "protein functional dynamics" to "protein dynamics" to more accurately reflect the scope of the study.

Thanks to the reviewer for the suggestion to use the more accurate terminology! We agreed with the reviewer that if we keep “protein functional dynamics” in the title, we should focus on how the “overall protein dynamic” links to the “function” – The function is directly related to PROTAC-induced structural dynamics which is commonly seen in “protein-structural-function” relationship, but it is not our main focus. Therefore, we changed the title to replace “functional” by “structural”.

(3) Incorporate more local and specific characterization methods in addition to RMSD for a more comprehensive conformational assessment.

We thank the reviewer for the suggestion. We performed time dependent correlation analysis to understand how the rotation of PROTACs can translate to the Lys-Gly interaction. In addition, we performed dihedral entropies analysis for each dihedral angle in the linker of the PROTACs to better examine the flexibility of each PROTAC.

We included detailed explanation at page 18: “Our dihedral entropies analysis showed that dBET57 has ~0.3 kcal/mol lower entropies than the other three linkers, suggesting dBET57 is less flexible than other PROTACs (Figure S18).”

Reviewer #2 (Public review):

Summary:

The manuscript reports the computational study of the dynamics of PROTAC-induced degradation complexes. The research investigates how different linkers within PROTACs affect the formation and stability of ternary complexes between the target protein BRD4BD1 and Cereblon E3 ligase, and the degradation machinery. Using computational modeling, docking, and molecular dynamics simulations, the study demonstrates that although all PROTACs form ternary complexes, the linkers significantly influence the dynamics and efficacy of protein degradation. The findings highlight that the flexibility and positioning of Lys residues are crucial for successful ubiquitination. The results also discussed the correlated motions between the PROTAC linker and the complex.

Strengths:

The field of PROTAC discovery and design, characterized by its limited research, distinguishes itself from traditional binary ligand-protein interactions by forming a ternary complex involving two proteins. The current understanding of how the structure of PROTAC influences its degradation efficacy remains insufficient. This study investigated the atomic-level dynamics of the degradation complex, offering potentially valuable insights for future research into PROTAC degradability.

Reviewer #2 (Recommendations for the authors):

(1) Regarding the modeling of the ternary complex, the BRD4 structure (3MXF) is from humans, whereas the CRBN structure in 4CI3 is derived from Gallus gallus. Is there a specific reason for not using structures from the same species, especially considering that human CRBN structures are available in the Protein Data Bank (e.g., 8OIZ, 4TZ4)?

We appreciate the reviewer's insightful comment regarding the choice of crystal structures of BRD4 and CRBN structures from two species. Our initial selection of 4CI3 for CRBN structure was based on its high resolution and publication in Nature journal. Furthermore, the *Gallus gallus* CRBN structure shares high degree of sequence and structural similarity with *Homo sapiens* CRBN, especially in the ligand binding region. At the time of our study, we were aware of 4TZ4 as *Homo sapiens* CRBN, however, we did not use this structure since no publication or detailed experimental was associated with it. Additionally, PDB 8OIZ, was not publicly available yet for other researchers to use at the time.

(2) Based on the crystal structure (PDB ID: 6BNB) discussed in Reference 6, the ternary complex of dBET57 exhibits a conformation distinct from other PROTACs, with CRBN adopting an "open" conformation. Using the same CRBN structure for dBET57 as for other PROTACs might result in inaccurate docking outcomes.

Thank you for the reviewer's comment! As noted by the authors in Reference 6, the observed open conformation of CRBN in the dBET57 ternary complex may result from the high salt crystallization conditions, which could drive structural rearrangement, and crystal contacts that may induce this conformation. The authors also mentioned that this open conformation could, in part, reflect CRBN's intrinsic plasticity. However, they acknowledged that further studies are needed to determine whether this conformational flexibility is a characteristic feature of CRBN that enables it to accommodate a variety of substrates. Despite these observations, we believe that the compatibility of the observed BRD4^{BD1} binding conformation with both open and closed CRBN states suggests that these conformational changes are all possible. Therefore, we believe using the same initial CRBN structure for

dBET57 as for other PROTACs can still reasonably reveal the dynamic nature of the ternary complex and would not significantly affect the accuracy of our docking outcomes either.

(3) Figure 2 displays only a single frame from the simulations, which might not provide a comprehensive representation. Could a contact frequency heatmap of PROTAC with the proteins be included to offer a more detailed view?

We thank the reviewer for the suggestion! We performed the contact map analysis to observe the average distance between PROTACs and BRD4^{BD1} over 400ns of MD simulation (new Figure S4 added).

We included detailed explanation at page 8 and 9: “The residues contact map throughout the 400ns MD simulation also showed different pattern of protein-protein interactions, indicating that the linkers were able to adopt different conformations (Figure S4).”

(4) The conclusions in Figure 3 and S11 are based on a single 400 ns trajectory. The reproducibility of these results is therefore uncertain.

We thank the reviewer for the suggestion! We added one more random seed MD simulation for each PROTAC to ensure the reproducibility of the results. The Result is shown in Figure S21 and the details for each MD run are updated in Table 1.

(5) Figure 4 indicates significant differences between the first and last 100 ns of the simulations. Does this suggest that the simulations have not converged? If so, how can the statistical analysis presented in this paper be considered reliable?

We thank the reviewers for the question. The simulation was initiated with a 10-15Å gap between BRD4 and Ub to monitor the movement of degradation machinery and Lys-Gly interaction. The significant changes in pseudo dihedral in Figure 4 shows that the large-scale movement of the degradation complex can initiate the Lys-Gly binding. It does not relate to unstable sampling because the system remains very stable when BRD4 comes close to Ub.

(6) In Figure 5, the dihedral angle of dBET57_#9MD1 is marked on a peptide bond. Shouldn't this angle have a high energy barrier for rotation?

We thank the reviewers for catching the error! Indeed, it was an error that the dihedral angles were marked on the peptide bond. We reworked the figure and double checked our dihedral correlation analysis. The updated correlate dihedral angle selection and the correlation coefficient is shown in Figure 5.

(7) Given that crystal structures for dBET 70, 23, and 57 are available, why is there a need to model the complex using protein-protein docking?

We thank the reviewer for the feedback. Only dBET23 has the ternary complex available in a crystal structure, which has the PROTAC and both proteins, while dBET1, dBET57 and dBET70 are not completed as ternary complexes. Although dBET70 has a crystal structure, its PROTAC's conformation is not resolved, and thus we decided to still perform protein-protein docking with dBET70.

We included the explanation at page 8: “Only dBET23 crystal structure is available with the PROTAC and both proteins, while the experimentally determined ternary complexes of dBET1, dBET57 and dBET70 are not available.”

(8) On page 9, it is mentioned that "only one of the 12 PDB files had CRBN bound to DDB1 (PDB ID 4TZ4)." However, there are numerous structures of the DDB1-CRBN complex

available, including those used for docking like 4CI3, as well as 4CI1, 4CI2, 8OIZ, etc.

We thank the reviewers for the comment! We acknowledged the existence of several DDB1-CRBN complex crystal structures, such as PDB IDs 4CI1, 4CI2, 4CI3, and the more recent 8OIZ. For our study, we chose to use 4TZ4 to maintain consistency in complex construction and to align with the methodology established in a previously published JBC paper (<https://doi.org/10.1016/j.jbc.2022.101653>), which successfully utilized the same structure for a similar construct. At the time our study was conducted, the 8OIZ structure had not yet been released. We appreciate your suggestion and will consider incorporating alternative structures in future studies to further investigate our findings.

(9) Table 2 is first referenced on page 8, while Table 1 is mentioned first on page 10. The numbering of these tables should be reversed to reflect their order of appearance in the text.

We thank the reviewer for catching the error! We switched the order of Table 1 and Table 2.

Reviewer #3 (Public review):

The authors offer an interesting computational study on the dynamics of PROTAC-driven protein degradation. They employed a combination of protein-protein docking, structural alignment, atomistic MD simulations, and post-analysis to model a series of CRBN-dBET-BRD4 ternary complexes, as well as the entire degradation machinery complex. These degraders, with different linker properties, were all capable of forming stable ternary complexes but had been shown experimentally to exhibit different degradation capabilities. While in the initial models of the degradation machinery complex, no surface Lys residue(s) of BRD4 were exposed sufficiently for the crucial ubiquitination step, MD simulations illustrated protein functional dynamics of the entire complex and local side-chain arrangements to bring Lys residue(s) to the catalytic pocket of E2/Ub for reactions. Using these simulations, the authors were able to present a hypothesis as to how linker property affects degradation potency. They were able to roughly correlate the distance of Lys residues to the catalytic pocket of E2/Ub with observed DC50/5h values. This is an interesting and timely study that presents interesting tools that could be used to guide future PROTAC design or optimization.

Reviewer #3 (Recommendations for the authors):

(1) My most important comment refers to the MM/PBSA analysis, the results of which are shown in Figure S9: binding affinities of -40 to -50 kcal/mol are unrealistic. This would correspond to a dissociation constant of 10^{-37} M. This analysis needs to be removed or corrected.

We thank the reviewer for the comment! MM/PBSA analysis indeed cannot give realistic binding free energy. It does not include the configurational entropy loss which should be a large positive value. In addition, while the implicit PBSA solvent model computes solvation free energy, the absolute values may not be very accurate. However, because this is a commonly used energy calculation, and some readers may like to see quantitative values to ensure that the systems have stable intermolecular attractions, we kept the analysis in SI. We edited the figure legend, moved the Figure S10 in SI page 19, and added sentences to clearly state that the calculations did not include configuration entropy loss “Note that the energy calculations focus on non-bonded intermolecular interactions and solvation free energy calculations using MM/PBSA, where the configuration entropy loss during protein binding was not explicitly included.”

(2) I think that the analysis of what in the different dBETx makes them cause different degradation potency is underdeveloped. The dihedral angle analysis (Figure 4B) did not explain the observed behavior in my opinion. Please add additional, clearer analysis as to what structural differences in the dBETx make them sample very different conformations.

We thank the reviewer for the suggestions! Based on the suggestion, we further performed dihedral entropy analysis for each dihedral angle in the linker part of the PROTAC to examine the flexibility of each PROTAC. Because each PROTAC has a different linker, we now clearly label them in a new Figure S18 in SI page 27. Low dihedral entropies indicate a more rigid structure and thus less flexibility to make a PROTAC more difficult to rearrange and facilitate the protein structural dynamic necessary for ubiquitination.

We added detailed explanation on page 18: “Our dihedral entropy analysis showed that dBET57 has ~0.3 kcal/mol lower configuration entropies than the other dBETs with three different linkers, suggesting that dBET57 is less flexible than the other PROTACs (Figure S18).”

(3) "The movement of the degradation machinery correlated with rotations of specific dihedrals of the linker region in dBETs (Figure 5).": this is not sufficiently clear from the figure. Definitely not in a quantitative way.

We thank the reviewers for the suggestions! To further understand the correlation between PROTACs dihedral angles and the movement of degradation machinery, we performed time dependent correlation analysis to correlate the dihedral angles of the PROTACs and the Lys-Gly distance (Figures 6 and S17).

We included detailed explanation on page 16:

“To further examine the correlation between PROTAC rotation and the Lys-Gly interaction, we performed a time-dependent correlation analysis. This analysis showed that PROTAC rotation translates motion over time, leading to the Lys-Gly interaction, with a correlation peak around 60-85 ns, marking the time of the interaction (Figure 6 and Figure S17). In addition, the pseudo dihedral angles also showed a high correlation (0.85 in the case of dBET1) with Lys-Gly distance. This indicated that degradation complex undergoes structural rearrangement and drives the Lys-Gly interaction.

(4) Cartoons are needed at multiple stages throughout the paper to enhance the clarity of what the modeled complexes looked like (e.g. which subunits they contained).

We thank the reviewers for the suggestions. We added and remade several Figures with cartoons to better represent the stages. We also used higher resolution and included clearer labels for each protein system.

(5) The difference between CRL4A E3 ligase and CRBN E3 ligase is not clear to the non-expert reader.

Thanks for the reviewer’s comment! To clarify the terms “CRL4A E3 ligase” and “CRBN E3 ligase”, which refer to different levels of description for the protein complexes, we added a couple of sentences in the Figure 1 legend. As a result, the non-expert readers can clearly know the differences.

As illustrated in Figure 1,

- CRL4A E3 ligase refers to the full E3 ligase complex, which includes all protein components such as CRBN, DDB1, CUL4A, and RBX1.
- CRBN E3 ligase, on the other hand, is a more colloquial term typically used to describe just the CRBN protein, often in isolation from the full CRL4A complex.

(6) Figure 1, legend: unclear why it's E3 in A and E2 in B.

We thank the reviewer for the question! E3 ligase in Figure 1A refers to CRBN E3 ligase, where researchers also simply term it CRBN. We have added a sentence to specify that CRBN E3 ligase is also termed CRBN for simplicity. In Figure 1B, E2 was unclear in the sentences. The full name of E2 should be E2 ubiquitin-conjugating enzyme. Because the name is a bit long, researchers also call it E2 enzyme. We have corrected it and used E2 enzyme to make it clearer.

(7) "Although the protein-protein binding affinities were similar, other degraders such as dBET1 and dBET57 had a $DC_{50/5h}$ of about 500 nM". It's unclear what experimental data supports the assertion that the protein-protein binding affinities are similar.

We thank reviewer for the question. Indeed, the statement is unclear.

We corrected the sentence in page 6: "Although utilizing the exact same warheads, other degraders such as dBET1 and dBET57 had a $DC_{50/5h}$ of about 500 nM."

(8) Was the construction of the degradation machinery complex guided by experimental data (maybe cryo-EM or tomography)? If not, what is the accuracy of the starting complex for MD? This may impact the reliability of the obtained results.

Thank you for your insightful comments! Yes, the construction of the degradation machinery complex was guided by available high-resolution crystal structures, which was selected to maintain consistency and align with the methodology established in a previously published JBC paper (<https://doi.org/10.1016/j.jbc.2022.101653>).

We acknowledged that static crystal structures represent only a single snapshot of the system and may not capture the full conformational flexibility of the complex. To address this limitation, we performed MD simulations using multiple starting structures. This approach allowed us to explore a broader conformational landscape and reduced the dependence on any single starting configuration, thereby enhancing the reliability of the results.

We hope this clarifies the robustness of our methodology and the steps taken to ensure accuracy in our simulations.

(9) "With quantitative data, we revealed the mechanism underlying dBETx-induced degradation machinery": I think this may be too strong of an assertion. The authors may have developed a mechanistic hypothesis that can be tested experimentally in the future.

We thank the reviewer for the suggestion. This is indeed a strong assertion and needs to be modified. We edited the sentence in page 7: "With quantitative data, we revealed the importance of the structural dynamics of dBETx-induced motions, which arrange positions of surface lysine residues of BRD4^{BD1} and the entire degradation machinery."

(10) Figure S2: are the RMSDs calculated over all residues? Or just the BRD4 residues? Given that the structures are aligned with respect to CRBN, the reported RMSD numbers might be artificially low since there are many more CRBN residues than there are BRD4

residues. Also, why weren't the crystal structures used for dBET 23 and 70 for the modeling? Wouldn't you want to use the most accurate possible structures? Simulations were run for 23. Why not for 70?

We thank the reviewer for the suggestion. We added a sentence to more clearly explain the RMSD calculations in Figure S2: “The structural superposition is performed based on the backbone of CRBN and RMSD calculation is conducted based on the backbone of BRD4^{BD1}.”

Although dBET70 has crystal structure, its PROTAC structure is not resolved, and thus we decided to still perform protein-protein docking with dBET70. dBET1 and dBET57 do not have a crystal structure for the ternary complexes.

We included the explanation at page 8: “Only dBET23 crystal structure is available with the PROTACs and both proteins, while the experimentally determined ternary complexes of dBET1, PROTACs of dBET57 and dBET70 are not available.”

a. And there are no crystal structures available for 1 and 57? If so, please clearly say that. Otherwise please report the RMSD.

We thank the reviewer for the suggestion. We included the explanation at page 8: “Only dBET23 crystal structure is available with the PROTACs and both proteins, while the experimentally determined ternary complexes of dBET1, PROTACs of dBET57 and dBET70 are not available.”

(11) Table 2 is referenced before Table 1.

We thank the reviewer for catching the error! We switched the order for Table 1 and Table 2.

(12) Figure S3 is not referenced in the main paper.

We thank the reviewer for catching the error! We now referred Figure S3 on page7.

(13) Minor comments on grammar and sentence structure:

a. It should be "binding of a ternary complex"

b. "Our shows the importance": word missing.

c. "...providing insights into potential orientations for ubiquitination. observe whether the preferred conformations are pre-organized for ubiquitination." Word or words missing.

We thank reviewer for catching the errors! We corrected grammatical errors and unclear sentences throughout the entire paper and revised the sentences to make them easily understandable for non-expert readers.

<https://doi.org/10.7554/eLife.101127.2.sa0>