Leveraging Structural and Computational Biology for Molecular Glue Discovery

Congbao Kang* & Weijun Xu*

Experimental Drug Development Centre, 10 Biopolis Road, #05-01, Chromos, Agency for Science, Technology and Research, Singapore, 138670

Correspondence: CK (kang_congbao@eddc.sg) and WX (xu_weijun@eddc.sg)

Abstract:

The discovery of molecular glues has made significant strides, unlocking new avenues for targeted protein degradation as a therapeutic strategy, thereby expanding the scope of drug discovery into territories previously considered undruggable. Pioneering molecules like thalidomide and its derivatives have paved the way for the development of small molecules that can induce specific protein degradation by hijacking the cellular ubiquitin-proteasome system. Recent advancements have focused on expanding the range of E3 ligases and target proteins that can be modulated by molecular glues. Structural elucidation of E3 ligase in complex with molecular glue and the target of interest, combined with computational modeling, facilitates the understanding of the underlying mechanisms of how molecular glues induce targeted degradation. By leveraging these tools, the next generation of molecular glues are expected to offer unprecedented opportunities for combating a wide range of diseases, including cancer, autoimmune disorders, and neurodegenerative conditions.

Over the past decade, drug discovery has undergone a paradigm shift with the advent of a new class of therapeutic entities - "molecular glue", which are reshaping approaches to targeting "undruggable" proteins. Unlike traditional small molecules that typically inhibit the function of the target proteins by blocking orthosteric or allosteric sites, molecular glues facilitate interactions between E3 ligases and their substrate proteins, stabilizing the complexes transiently to induce target degradation, significantly expanding the druggable proteome [1]. E3 ligases are a diverse family of over 600 enzymes that play a central role in the process of protein degradation within cells. The large diversity of E3 ligases allows for highly specific recognition of target proteins, as each ligase is adapted to bind cognate substrates, often in response to specific cellular conditions.

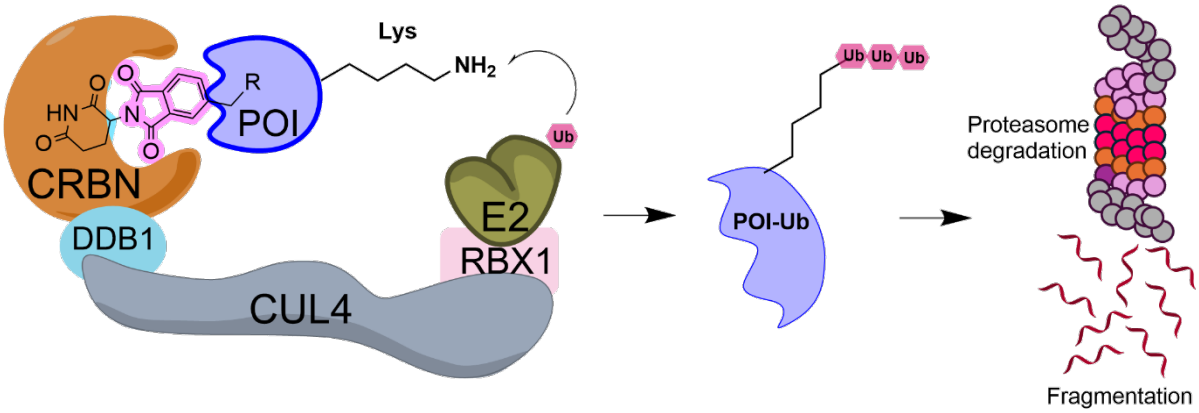

Figure 1. Cartoon representation of targeted degradation by molecular glues using CRBN as the model for E3 ligase. POI, protein of interest; Ub, ubiquitin; CRBN, Cereblon; DDB1, DNA damage-binding protein 1; RBX1, RING-box protein 1.

Today, nearly a dozen molecular glues are in clinical studies (Table 1). The licensing deal alongside molecular glue development has been significantly expanding. Novo Nordisk inked with Neomorph for a licensing agreement of up to $1.46 billion across multiple molecular glue targets. Pfizer taps Triana to discover molecular glues for cancer with a deal worth $1.5B [2]. Novartis acquired VAV-1 glue from Monte Rosa with an upfront payment of $150 million and up to 2.1 billion milestone for the further development. Biogen and Neomorph established partnership worth up to $1.45 billion to co-develop molecular glue degraders (MGDs) for the treatment of Alzheimer's disorders, rare neurological and immunological diseases.

The discovery of molecular glues using traditional methods often relied on a combination of high-throughput screening (HTS), biological validation and medicinal chemistry optimization [3]. Early approaches typically involved screening libraries of small molecules for their ability to induce degradation of targeted protein via proteasome-dependent pathway in cell-based assays, followed by mechanistic validation to confirm the E3 ligases and protein substrate being degraded [4]. One of the most seminal successes was the discovery of thalidomide in the 1950s, which was originally known for its teratogenic effects but found to act as a molecular glue of E3 ligase CRBN 60 years later [5]. While early discovery for molecular glues stemmed from serendipity, the studies laid the groundwork and potential for rational approaches, including structure-based drug design and computational modeling, which are anticipated to accelerate the discovery of new molecular glues with broader therapeutic applications.

Cryo-EM and X-ray crystallography are powerful tools for resolving the structures of ternary complexes formed by a protein of interest, a molecular glue, and an E3 ligase. The structural knowledge allows medicinal chemists to identify potential "hot spots" on target proteins and E3 ligases where molecular glues might effectively mediate and enhance protein-protein interactions, which are essential for selective protein degradation. Recent advancements in these fields have shed light on how molecular glues engage both E3 ligase and target protein for degradation at high resolution atomic level [6]. For example, the molecular signature (**Figure 2**) of the CRBN E3 ligase, specifically its recognition of the G-motif in target proteins, plays a crucial role by which molecular glues facilitate target degradation. The G-motif, typically characterized by a stretch of amino acid residues that adopt a beta-hairpin structural architecture, is recognized by the CRBN ligand-binding cleft, enabling the recruitment of the target protein to the glue-bound "neosurface" of the E3 ligase. This binding event bridges the CRBN ligase to the target protein, facilitating its ubiquitination by the E3

ligase and subsequent degradation via the proteasome. Importantly, the presence of the G-motif is crucial for conferring the specificity of the molecular glue, as the binding of the glue to CRBN effectively reprograms the ligase's substrate specificity [7], creating the "neosubstrate" known as "degron". Such molecular interaction highlights the importance of understanding the structural details of the E3-substrate interaction for the design of next-generation molecular glues.

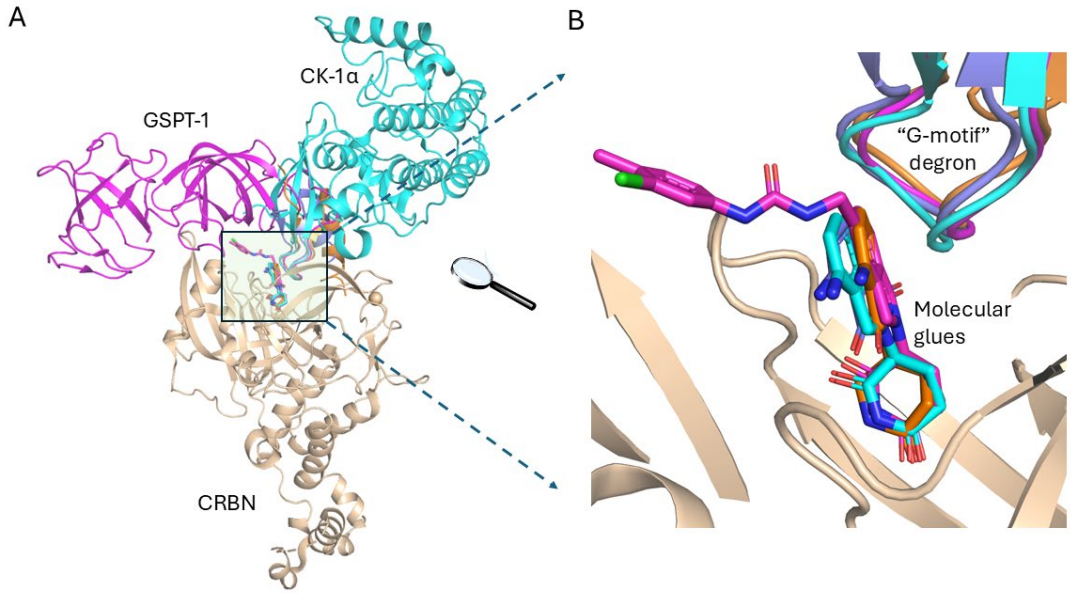

Figure 2. Superimposed ternary structures of CRBN-glue-substrate complex. A). The following structures from protein data bank are used. 5fqd: CRBN-lenalidomide-CK1a (cyan) complex; 5hxb: CRBN-CC-885-GSPT1 (purple); 6h0g: CRBN-pomalidomide-ZNF692 (orange); 6uml: CRBN-pomalidomide-SALL4 (slate); CRBN is colored in wheat, carbon atoms from molecular glues are colored as the same scheme for the substrate proteins B). Zoom-in view at the protein-protein interaction interface bridged by molecular glues sticking to β-hairpin G-motif degron that is recognized as a molecular signature by CRBN molecular glues.

Distinctive from traditional small molecule drug discovery targeting protein-ligand binary complex, molecular glues work by remodeling conformational changes at the protein-protein interaction interfaces, stabilizing interaction between two proteins by enhancing the intrinsically transient interactions [8]. Thus, the rational discovery and design of molecular glues face specific challenges. First, it is imperative to identify protein-protein interactions that are transient or natively present between the target of interest and an E3 ligase. These interfaces are often diverse and difficult to address with small molecules, further complicating the design process. Second, the rational selection of compound library for screening novel molecular glue

scaffolds is limited by deficiency in prior knowledge towards the chemical space that is likely enriched for "glue-like" scaffolds. Third, the effective discovery and molecular glues at early discovery stage requires the use of robust assays for screening. The identification of a "*bona fide*" glue molecule is further implicated by the complexity of the mechanisms of protein degradation, which often involve proteome-wide interactions and the proteasomal machinery. Nanoluciferase fusion proteins and the NanoBiT technology represent two popular screening platforms to monitor protein abundance in high-throughput screening campaign. However, concerns have arisen regarding potential artifacts due to the presence of lysine residues on the tagging molecules (HiBiT), warranting the development of improved methods for screening [9] . Fourth, the complex molecular basis makes it difficult to identify molecular glues for a novel combination of a target protein and an E3 ligase. The mechanism of action from novel molecular glues requires intensive biochemical, biophysical and structural characterization. Fifth, the association (K_{on}) and dissociation (K_{off}) kinetics of molecular glues are crucial for the efficiency of protein degradation. Achieving the right balance between stability and reversibility is essential for proper complex formation and disassembly [10]. Lastly, most cell-based assays monitor target protein levels using western blot, which indicates protein level as a net balance of protein synthesis and degradation, careful assessment on whether the protein abundance is reduced due to the synthesis or degradation is needed to confirm on-target effect of the molecular glues. Together, these challenges emphasize the need for a detailed structural understanding of protein-protein interactions, robust computational modeling tools, and advanced screening methods for characterizing the formation of ternary complexes, including dynamics and kinetics analysis. Furthermore, the availability and curation of structural information for new classes of E3 will be key to the understanding of conserved motif that constitutes the "degron" for the rest of E3s.

Today, only a limited fraction of known E3 ligase ligands have been reported, including thalidomide, lenalidomide, and pomalidomide, which primarily bind to the CRBN E3 ligase. Additional examples include ligands that bind to KEAP1, DCAF, VHL, MDM2, IAP, and TRIM family E3 ligases. However, the majority of the E3 ligases remained un-drugged. In a recent blog [11] , Ian Churcher raised a question on how many E3 ligases may be repurposed for targeted degradation like CRBN, which is a "housekeeping" E3 ligase that recognizes an imide moiety. According to Monte Rosa's prediction [12], there are about 1400 proteins that feature the G-motif, positioning these proteins putative substrates for CRBN mediated degradation. However, whether the rest of E3 ligases may behave like CRBN remains enigmatic. Despite

above mentioned challenges, computational tools have significantly transformed the way of ligand discovery from large chemical space (billions of compounds), leveraging on available experimental structural information or high quality predictive molecular models. Artificial intelligence, or deep learning-based tools like “MASIF Site”, “MASIF Search”, AlphaFold Multimer, PPIMiner [13], and protein-protein docking are revolutionizing the field of molecular glue discovery by enhancing the ability to predict and understand the hotspots at the interface of protein-protein interactions (PPIs) with an unprecedented scale and accuracy [14]. MASIF site and MASIF [15] search utilize machine learning to predict the binding sites of small molecules on protein surfaces, enabling researchers to identify potential regions where molecular glues can be designed to engage E3 ligases and their targets. AlphaFold Multimer [14], with its ability to predict protein structures and complexes, offers valuable insights into the 3D architecture of binary or ternary protein complexes, including the identification of interfaces between E3 ligases and their substrates. PPI miner, leveraging on deep learning models, predicts the formation of protein-protein complex model. Physics-based protein-protein docking and molecular dynamic simulations sample conformational ensembles between an E3 ligase and its target protein, guiding the design and optimization of molecular glue. Free energy perturbation (FEP) simulations on ternary complexes comprising the target protein, E3 ligase, and the molecular glue should in principle play a critical role in advancing molecular glue discovery during the hit-to-lead phase. While FEP has been widely used for predicting binding affinity between small molecule-protein binary complexes, its application for ternary complexes is less reported. This method enables the optimization of molecular glue candidates by identifying key residues that contribute to binding specificity and affinity, ultimately guiding structure-activity relationship (SAR), an area that is not well established for hit-to-lead stage during molecular glue discovery campaign. Collectively, these tools allow the mining of hypothetical match-pairing between E3 ligases and their substrates structurally, leading to further analysis for a putative site “sandwiched” in between E3 ligases and their cognate substrates, enabling multiple computational approaches (e.g. virtual screening or generative AI) to be employed with the aim to discover novel molecular glues, and refine their design for experimental validation. It is worth noting that biophysical techniques such as NMR spectroscopy can pinpoint binding pockets likely occupied by small molecules, streamlining assay development by providing “pre-validated” structures of molecular glue-E3 ligase complexes. Furthermore, combining fragment-based and structure-based drug design allows researchers to identify and refine molecular glues with greater specificity and efficiency, ultimately creating a new or expanding compound libraries. Despite challenges faced with ternary structure determination,

NMR spectroscopy represents a relatively high throughput approach that provides unique insights into protein dynamics, conformational shifts, protein-protein interaction and ternary complex formation.

Looking forward, the rapid technological evolution in both structural and computational biology combined with the strengths of biophysical methods such as NMR spectroscopy and ITC offers unprecedented opportunities for the rational design of molecular glues. The integration of these approaches should be implemented across different stages, including target identification, E3 ligase and substrate selection, druggability evaluation of the E3 ligase, assay development, hit discovery, and subsequent optimization. The synergy of these techniques paves the avenue towards successful discovery and development of next-generation molecular glues that would serve as therapeutic modalities to treat diseases associated with aberrant protein functions.

Table 1. **Selected examples (non-exhaustive) of molecular glue degraders in clinical studies. The information is collected from the company's website.**

Target	Target function	Molecule	Company	Indication	Clinical status
GSPT1	GTPase, translation termination factor	MRT-2359 CC-90009	Monte Rosa	MYC-driven cancer	Phase 1/2; discontinued
NA	Transcription factor	HbF-activating CELMoD BMS-986470	Bristol-Myers Squibb (BMS)	Sickle cell disease	Phase 1/2
WIZ	Transcription factor	NA (glue)	Novartis	Sickle cell disease	Preclinical
IKZF2	Transcription factor	Helios CELMoD	BMS	Cancer	Phase 1
IKZF2	Transcription factor	DKY709	Novartis	Cancer	Discontinued
RBM39	Splicing factor	NA	Seed Therapeutics	Cancer	Phase 1 in 2025
NEK7	Kinase	MRT-8102 (glue);	Monte Rosa	Inflammation	Preclinical
HuR	RBP	NA	Degron	Cancer	Preclinical
Vav1	Guanine nucleotide exchange factor	MRT-6160	Monte Rosa	Autoimmunity	Phase 1

References

(1) Liu, Y.; Yang, J.; Wang, T.; Luo, M.; Chen, Y.; Chen, C.; Ronai, Z. e.; Zhou, Y.; Ruppin, E.; Han, L. Expanding PROTACtable genome universe of E3 ligases. *Nature Communications* **2023**, *14* (1), 6509. DOI: 10.1038/s41467-023-42233-2.
(2) https://www.fiercebiotech.com/biotech/triana-sticks-landing-pfizer-molecular-glue-discovery-deal-potentially-worth-15b.
(3) Sasso, J. M.; Tenchov, R.; Wang, D.; Johnson, L. S.; Wang, X.; Zhou, Q. A. Molecular Glues: The Adhesive Connecting Targeted Protein Degradation to the Clinic. *Biochemistry* **2023**, *62* (3), 601-623. DOI: 10.1021/acs.biochem.2c00245.
(4) Ishida, T.; Ciulli, A. E3 Ligase Ligands for PROTACs: How They Were Found and How to Discover New Ones. *SLAS Discovery* **2021**, *26* (4), 484-502. DOI: 10.1177/2472555220965528 (acccessed 2025/01/01).
(5) Ito, T.; Ando, H.; Suzuki, T.; Ogura, T.; Hotta, K.; Imamura, Y.; Yamaguchi, Y.; Handa, H. Identification of a Primary Target of Thalidomide Teratogenicity. *Science* **2010**, *327* (5971), 1345-1350. DOI: doi:10.1126/science.1177319.
(6) Haubrich, K.; Spiteri, V. A.; Farnaby, W.; Sobott, F.; Ciulli, A. Breaking free from the crystal lattice: Structural biology in solution to study protein degraders. *Current Opinion in Structural Biology* **2023**, *79*, 102534. DOI: https://doi.org/10.1016/j.sbi.2023.102534.
(7) Oleinikovas, V.; Gainza, P.; Ryckmans, T.; Fasching, B.; Thomä, N. H. From Thalidomide to Rational Molecular Glue Design for Targeted Protein Degradation. *Annual Review of Pharmacology and Toxicology* **2024**, *64* (Volume 64, 2024), 291-312. DOI: https://doi.org/10.1146/annurev-pharmtox-022123-104147.
(8) Miñarro-Lleonar, M.; Bertran-Mostazo, A.; Duro, J.; Barril, X.; Juárez-Jiménez, J. Lenalidomide Stabilizes Protein–Protein Complexes by Turning Labile Intermolecular H-Bonds into Robust Interactions. *Journal of Medicinal Chemistry* **2023**, *66* (9), 6037-6046. DOI: 10.1021/acs.jmedchem.2c01692. Konstantinidou, M.; Arkin, M. R. Molecular glues for protein-protein interactions: Progressing toward a new dream. *Cell Chemical Biology* **2024**, *31* (6), 1064-1088. DOI: 10.1016/j.chembiol.2024.04.002 (acccessed 2025/01/07).
(9) Lin, H.; Riching, K.; Lai, M. P.; Lu, D.; Cheng, R.; Qi, X.; Wang, J. Lysineless HiBiT and NanoLuc Tagging Systems as Alternative Tools for Monitoring Targeted Protein Degradation. *ACS Medicinal Chemistry Letters* **2024**, *15* (8), 1367-1375. DOI: 10.1021/acsmedchemlett.4c00271.
(10) Kiely-Collins, H.; Winter, G. E.; Bernardes, G. J. L. The role of reversible and irreversible covalent chemistry in targeted protein degradation. *Cell Chemical Biology* **2021**, *28* (7), 952-968. DOI: 10.1016/j.chembiol.2021.03.005 (acccessed 2025/01/01).
(11) https://www.janusdrugdiscovery.com/241023-ligases/.
(12) Petzold, G.; Gainza, P.; Annunziato, S.; Lamberto, I.; Trenh, P.; McAllister, L.; Demarco, B.; Schwander, L.; Bunker, R. D.; Zlotosch, M.; et al. Mining the CRBN Target Space Redefines Rules for Molecular Glue-induced Neosubstrate Recognition. *bioRxiv* **2024**, 2024.2010.2007.616933. DOI: 10.1101/2024.10.07.616933.
(13) Wang, L.; Li, F.-l.; Ma, X.-y.; Cang, Y.; Bai, F. PPI-Miner: A Structure and Sequence Motif Co-Driven Protein–Protein Interaction Mining and Modeling Computational Method. *Journal of Chemical Information and Modeling* **2022**, *62* (23), 6160-6171. DOI: 10.1021/acs.jcim.2c01033.
(14) Abramson, J.; Adler, J.; Dunger, J.; Evans, R.; Green, T.; Pritzel, A.; Ronneberger, O.; Willmore, L.; Ballard, A. J.; Bambrick, J.; et al. Accurate structure prediction of

biomolecular interactions with AlphaFold 3. *Nature* **2024**, *630* (8016), 493-500. DOI: 10.1038/s41586-024-07487-w.
(15) Gainza, P.; Sverrisson, F.; Monti, F.; Rodolà, E.; Boscaini, D.; Bronstein, M. M.; Correia, B. E. Deciphering interaction fingerprints from protein molecular surfaces using geometric deep learning. *Nature Methods* **2020**, *17* (2), 184-192. DOI: 10.1038/s41592-019-0666-6.