Computational Modeling of First-In-Class G9a/GLP Protein by Using PROTAC Degrader

NISHITA KULKARNI[1,2] AND GAURAV SHARMA[2]

[1]Lake Norman Charter, NC, Huntersville
[2]Eigen Sciences, NC, Apex

Published Feb, 2025

G9a leads to unusual chromatin structure and gene suppression, disrupting proper cellular functions and contributing to the development of detrimental diseases like cancer—more specifically, colorectal, lung, and head and neck cancer. PROTAC is a technology that attaches to the target protein (G9a in this work) while also attaching to the E3 ubiquitin ligase; the PROTAC role in this is to transmit the ubiquitin molecules to the G9a protein, aiding in the degradation of the target protein. We hypothesize that chemical modulation in the PROTAC structure can help design more potent PROTACs binding G9a protein. P2Rank helped design the prediction of ligand binding sites with the proteins. P2Rank integrates machine learning algorithms and systematic information to help identify the potential binding areas on the protein's surface. For docking, HADDOCK is used, a flexible software aiding in the modeling of biomolecular complexes. HADDOCK prioritizes experimental data supporting the docking process to embody different information to refine complex structures. Current research will help develop PROTACs that can degrade the G9a protein formed in various cancer cells.

1. INTRODUCTION

Lysine methyltransferases G9a and GLP build on methyl groups to a specific spot on histone proteins, specifically histone H3 at lysine 9 (H3K9) (1). Histones contribute to packing DNA into dense structures, and their modification impacts gene expression. When G9a and GLP are hyperactive, they can add excessive methyl groups to H3K9 (2). This excessive methylation affects the form of chromatin (combined DNA and histones) and can cause a suppression of genes, which usually manage cell growth and splitting (2). This disrupts regular cellular functions and contributes to cancer development as it allows unrestrained cell multiplication and livelihood (2). Furthermore, the unusual activity of G9a and GLP creates an environment maintaining the suppression of tumor suppressor genes, which is essential in cancer prevention (3). This suppression leads to the absence of cellular checks and balances, thereby facilitating the growth of cancerous cells (4). Additionally, the alterations in gene expression caused by irregular H3K9 methylation can impact various cellular pathways and functions, including those related to DNA repair, cell cycle regulation, and apoptosis (programmed cell death), all of which contribute to the malignancy of cancer cells (5). G9a is also called euchromatic histone-lysine N-methyltransferase 2 (EHMT2) protein, and its dysregulation may lead to leukemia, colorectal cancer, lung cancer, and head and neck cancer (2).

PROTAC (Proteolysis Targeting Chimera) is a visionary technology that uses the cell's natural protein degradation system to target/turn off specific proteins (6,7). This method involves using small bifunctional molecules that attach to a target protein on one end and an E3 ubiquitin ligase on the other (8). The nearness created by the PROTAC facilitates the transfer of ubiquitin molecules to the target protein, labeling it for degradation by the proteasome (7). In contrast to traditional inhibitors that block protein function, PROTACs lead to the destruction of the protein, providing more effective therapeutic benefits (6). This technology holds a guarantee for treating diseases like cancer by targeting proteins considered "undruggable" (7).

Molecular Docking is a computational practice that predicts the preferred model of a molecule (usually a tiny ligand) to another molecule (usually a protein or enzyme), which forms a stable complex (9). Usually, this method is essential for discovering drugs due to it helping understand the binding relationship and uniqueness of the drug contenders (9). Due to the simulation of the interactions at an atomic level, docking studies are identified as binding sites and aid in predicting the power of the interaction, which is validated through experimental practices (10). The information derived from molecular dockings is crucial as it plays a role in creating and improving new therapeutic medicines (9).

Recently, a first-in-class PROTAC-based oncogenic G9a/GLP protein degrader was developed (11). We hypothesize that the PROTACs designed bind to the proteins at a specific site and help

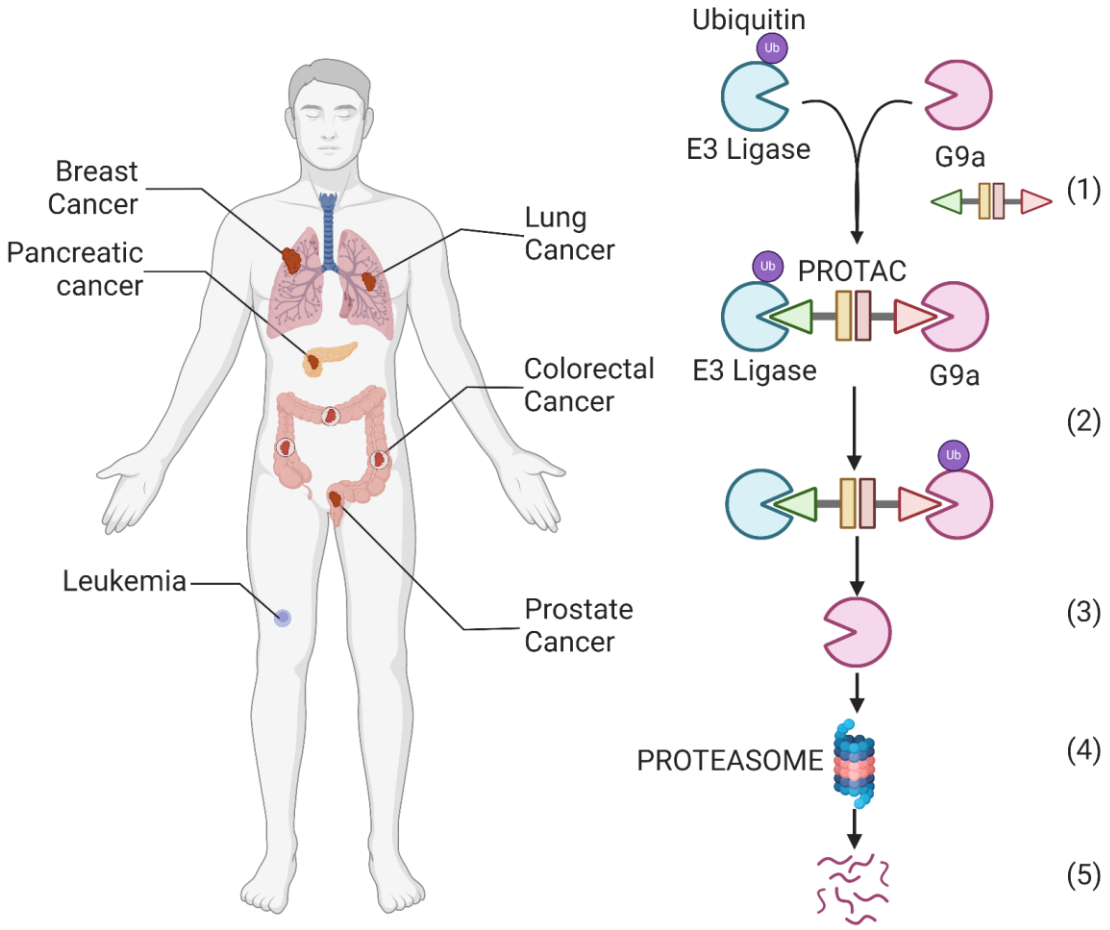

Fig. 1. Schematic of PROTAC functioning. The steps involved in PROTAC functioning are binding the PROTAC molecule to tG9a, integrating E3 ubiquitin ligase, and proceeding with the degradation of the G9a protein. The point of interest would be the specific binding sites of the PROTAC molecule. The initial stage involves binding G9a and E3 ligase in the presence of PROTAC. Subsequently, the ubiquitin ligase, once bound, is transported to the G9a molecule. Finally, G9a is released, leading to its degradation as it enters the proteasome. The figure was generated using BioRender.com

in their ubiquitination. Based on our computational simulations, we have found that the all the PROTAC bind strongly to both proteins. The current research will help in designing novel PROTACs against various cancers.

2. RESULTS

Figure 2 is the structure of the E3 Ligase and G9a protein. It portrays the druggable site, shown in red in the figure's center, with the proteins' electrostatic surface potential portrayed towards the bottom of the figure. We used the software P2rank web server to predict the druggable site on the protein surface. The proteins shown in Figure 2 are E3 ligase and G9a protein. The electrostatic surface potential (ESP) of the protein is displayed in Figure 2. ESP is the division of electric charges located on the molecule's surface; this division portrays information on the spatial arrangement of both positive and negative charges.

In the next step, we performed molecular docking simulation using the HDOCK software. The docking was performed to find the PROTAC that can bind to both G9a and E3 ligase proteins. We ran molecular docking simulations to analyze PROTAC's binding interactions on the protein's surface. This running of the simulation is essential because it aids in the finding of compounds that fix on the earlier analyzed binding site. Regarding this standard, the PROTACs are properly associated with the protein's binding site. Henceforward, the number of interactions and binding energy in the middle of the protein and the PROTAC were utilized to evaluate the molecules. The 2D interaction between the protein and PROTAC is shown in Figure 3. In addition, the 3d structure is displayed in Figure 4. In this image, the purple protein is G9a, while the green protein is E3 Ligase. The center of this protein is the PROTAC; the PROTAC aids in the binding of both the G9a and E3 ligase proteins since the ligand ubitinquizes both proteins, leading to the degradation of the G9a protein, in this case, the G9a. Finally, we computed the binding energy between the protein and PROTAC complex. Ligand III showed the strongest binding energy of -9.82 kcal/mol, while Ligand I and II showed that the binding energy was -9.20 and -9.10 kcal/mol, respectively.

3. DISCUSSION

Due to their specific method of action and capability to target proteins, which were formerly unreachable, PROTACs are a

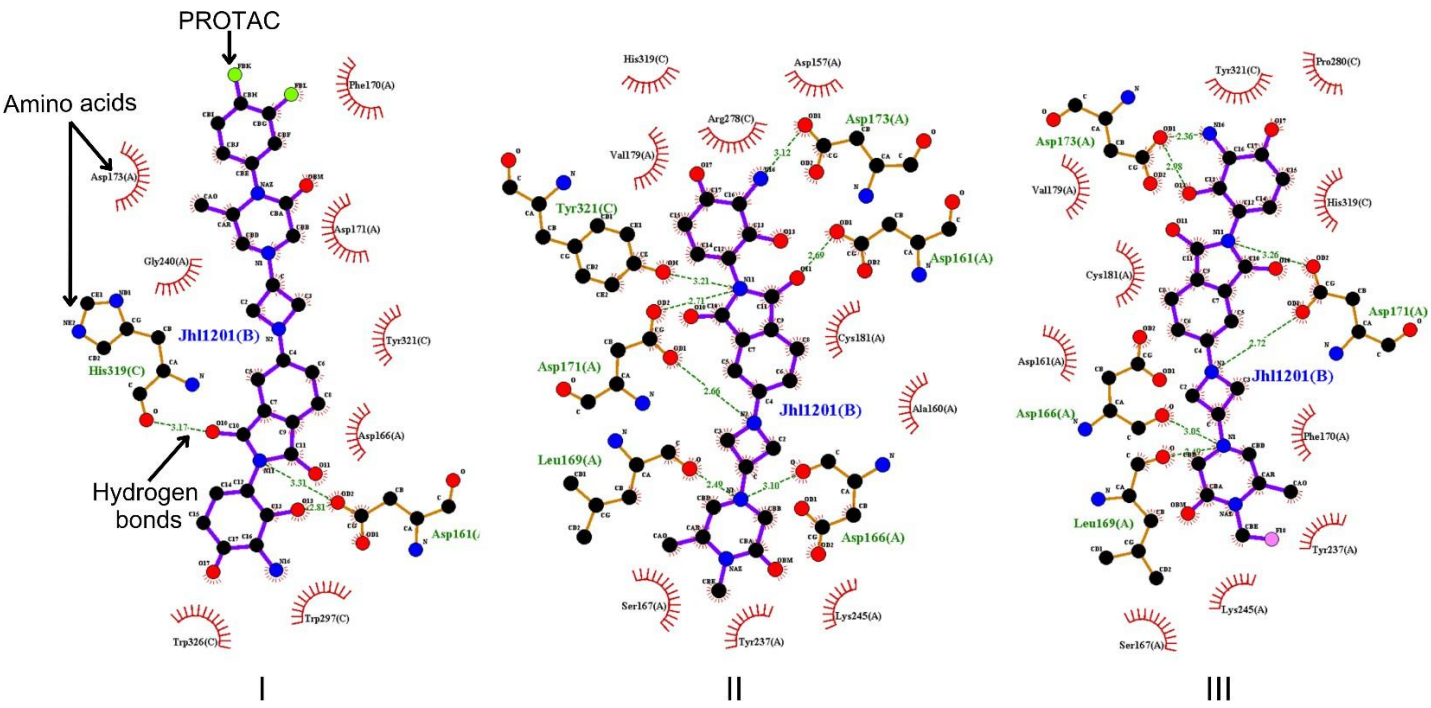

Fig. 2. 2D interactions formed between the protein and PROTAC. The image shows protein amino acids forming hydrogen bonds in green dash lines and red eyebrows showing the hydrophobic interactions.

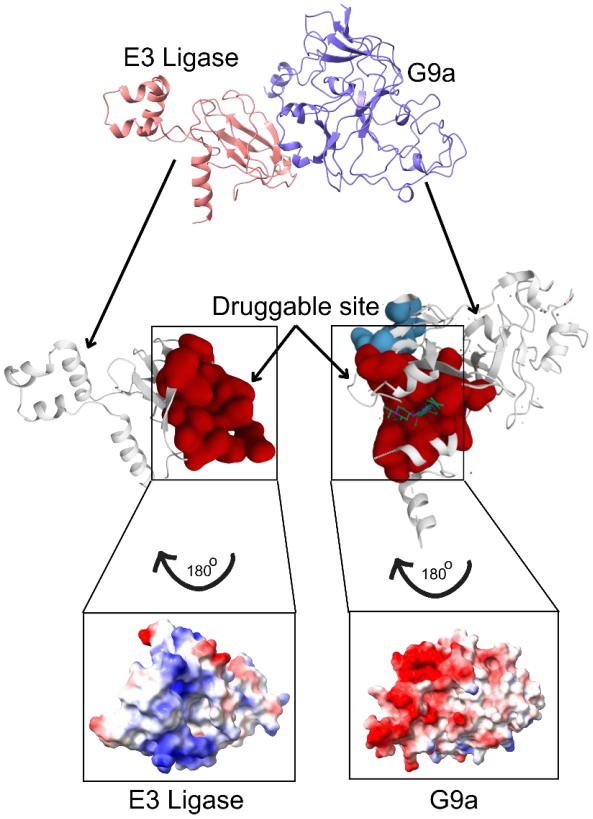

Fig. 3. Structure of E3 Ligase and G9a protein. The druggable site is shown in the red center, and the proteins' electrostatic surface potential is shown at the bottom.

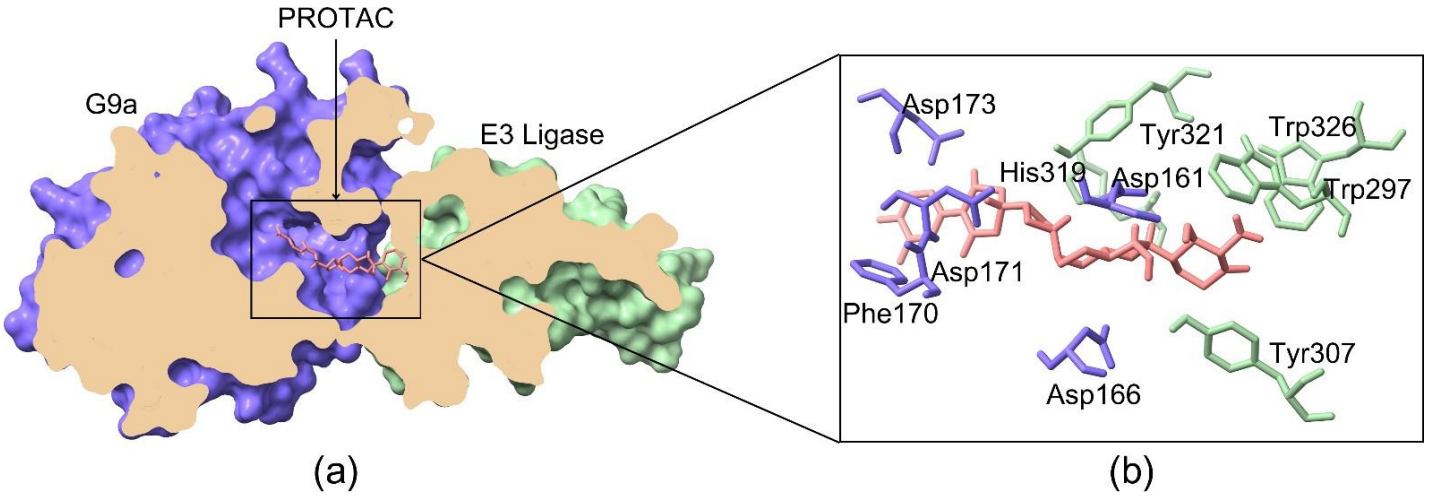

Fig. 4. G9a-PROTAC I-E3 Ligase interactions:(a) PROTAC I interacting with both G9a and E3 ligase; (b) amino acids (in red and green) interacting with PROTAC I (pink); and (c) 2D image of PROTAC IX (in purple) and the hydrophobic residues are in red eyebrows.

type of restorative medicine that has become very attractive due to their interest in the field of oncology research. PROTACs are amphiphilic molecules that bind to specific target proteins, attracting E3 ubiquitin ligases, leading to the protein being ubiquitinated and then impaired by proteases. The field of oncology research has thoroughly delved into the innovation and results of PROTACs. These medications might be able to recover from the drawbacks of traditional small-molecule inhibitors, which repeatedly find it challenging to precisely target specific proteins, particularly those involved in signaling pathways vital to the development of cancer. PROTACs present a novel approach that could result in long-term and more effective cancer treatments by using the cell's protein degradation machinery to target and mainly remove distinct proteins. Targeting receptor tyrosine kinases and non-receptor kinases, which are essential in various cancer types, has been an area of awareness in PROTAC research for cancer. According to a research analysis, it has been proposed that different proteins like Src kinase and the insulin-like growth factor 1 receptor (1GF-1R) that are related to types of cancer could be targeted for degradation from the method PROTAC.

Furthermore, scientists explored the potential of using PROTACs in cancer research to assess other molecular imbalances and genetic and epigenetic changes. Researchers are looking at novel strategies to construct and optimize these compounds to improve their strength, selectivity, and effectiveness in cancer treatment while the PROTAC Research field is developing. By focusing on the carcinogenic G9a/GLP proteins, the work reveals the exciting possibility of PROTAC technology in cancer therapy; the use of molecular docking simulations, certain PROTACs were found to attach to the proteins efficiently, leading to the cell's natural processes impairing them. This method provides a powerful and innovative alternative to traditional inhibitors, possibly leading to longer-lasting cancer treatments. These outcomes inspire further research into PROTAC-based therapies, which have the prospect of improving cancer targeting and treatment significantly.

This analysis portrayed how PRTOAC technology can target the oncogenic proteins G9a and GLP. These proteins encourage extra methylation of histone H3 at lysine 9 and contribute to cancer development. The analysis discovered specific PROTAC molecules binding these proteins effectively, causing their destruction to occur through the use of molecular docking simulations. This specialized method portrays a feasible substitute for traditional cancer therapy, with the ability to overcome flaws of conventional inhibitors, offering a concentrated means of inhibiting tumor maturation. The results can give hope for better cancer therapy since it opens a way to create novel PROTAC-based therapeutics.

4. METHOD

After receiving the E3 Ligase-PROTAC-PIM1 protein complex from the Protein Data Bank, the following steps were taken: The PROAC molecules were obtained in Mol2 format, then converted into fourteen different PROTAC mutations in SMILES format. The SMILES were then changed into PDB format for analysis. By examining the complexes through the software, the relationships between proteins and the PROTAC molecules inspired new designs for new pharmacological agents by indicating the preferred binding conformations and evaluating the resilience of the resulting complexes. The PLIP software

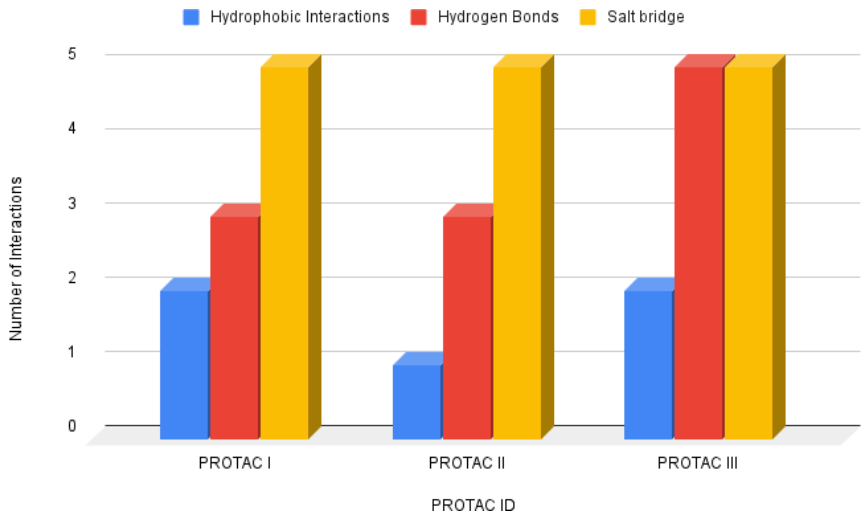

Fig. 5. Number of interactions formed between protein and PROTAC.

was utilized to analyze the binding interactions of the output protein-PROTAC complexes (12). Additionally, a machine learning program called P2rank was used to calculate the druggable site on the two proteins based on the local properties of the protein's surface (13). P2Rank employs a machine learning algorithm to predict ligand-binding sites on protein structures. It is intended to be fast and accurate, aiding drug discovery and other bioinformatic applications. Finally, all the complexes were evaluated using the ChimeraX software (14). To obtain a PROTAC 3D structure, we have used the following steps: Smiles is the abbreviation for "Simplified Molecular Input Line Entry System"; this is a specific line notation used to describe a chemical species structure. We have utilized the following linkhttps: //www.novoprolabs.com/tools/smiles2pdb to transform the SMILES notation to 3D PDB files to perform molecular docking simulations. Download the file and modify the PDB file names based on their serial numbers for a more methodical format. Confirm the security of PDB files for the docking process. Employ the HADDOCK software to dock PROTAC PDB files with protein. This step allows for predicting the best-fitted orientation and arrangement of molecules (15). This step is crucial as it aids in the understanding of the potential interaction of a PROTAC molecule and the target protein. Identify the precise active residues for the protein (Molecule 1) and Protac (Molecule 2), which will be used for molecular docking simulations. To find the protein PROTAC binding interaction, we performed molecular docking simulations using the HADDOCK software. The G9a-PROTAC-E3 Ligase image is shown in Figure 3. The PROTAC binds to the G9a and E3 Ligase protein because it forms strong bonds, as shown in Table 1.

5. REFERENCES

1. Krishnan, S. et al. "Structure and function of histone h3 lysine 9 methyltransferases and demethylases." vol. 12, no. 2, 2011, pp. 254-263.

2. Sampath, S.C. "Studies on the histone methyltransferase g9a." 2006.

3. Casciello, F. et al. "Functional role of g9a histone methyltransferase in cancer." vol. 6, 2015, p. 487.

4. Wing, K. et al. "Regulatory t cells exert checks and balances on self tolerance and autoimmunity." vol. 11, no. 1, 2010, pp. 7-13.

5. Zhang, Y.-J. et al. "The dual functions of the extreme n-terminus of tdp-43 in regulating its biological activity and inclusion formation." vol. 22, no. 15, 2013, pp. 3112-3122.

6. Weng, G. et al. "Protac-db 2.0: An updated database of protacs." vol. 51, no. D1, 2023, pp. D1367-D1372.

7. Wang, C. et al. "New-generation advanced protacs as potential therapeutic agents in cancer therapy." vol. 23, no. 1, 2024, p. 110.

8. Bricelj, A. et al. "E3 ligase ligands in successful protacs: An overview of syntheses and linker attachment points." vol. 9, 2021, p. 707317.

9. Fan, J. et al. "Progress in molecular docking." vol. 7, 2019, pp. 83-89.

10. Ghasemi, J.B. et al. "Molecular docking challenges and limitations." Pharmaceutical sciences: Breakthroughs in research and practice, IGI Global, 2017, pp. 770-794.

11. Velez, J. et al. "Discovery of the first-in-class g9a/glp protac degrader." vol. 67, no. 8, 2024, pp. 6397-6409.

12. Salentin, S. et al. "Plip: Fully automated protein–ligand interaction profiler." vol. 43, no. W1, 2015, pp. W443-W447.

13. Krivák, R. et al. "P2rank: Machine learning based tool for rapid and accurate prediction of ligand binding sites from protein structure." vol. 10, 2018, pp. 1-12.

14. Pettersen, E.F. et al. "Ucsf chimerax: Structure visualization for researchers, educators, and developers." vol. 30, no. 1, 2021, pp. 70-82.

15. De Vries, S.J. et al. "The haddock web server for data-driven biomolecular docking." vol. 5, no. 5, 2010, pp. 883-897.

Table 1. Type of interactions and their bond distances between protein and PROTAC.

	PROTAC I	
	Residue	Distance (Å)
Hydrophobic Interactions	Phe170	3.48
	Pro82	3.29, 3.84
	Asp161	3.09
Hydrogen Bonds	Tyr307	2.98
	His319	2.55
	Asp161	4.59
	Asp161	3.34
Salt bridge	Asp171	2.76
	Asp171	3.92
	Asp173	4.34
	PROTAC II	
Hydrophobic Interactions	Phe170	3.48
	Asp161	3.09
Hydrogen Bonds	Tyr307	2.98
	His319	2.55
	Asp161	4.59
	Asp161	3.34
Salt Bridges	Asp171	2.76
	Asp171	3.92
	Asp173	4.34
	PROTAC III	
Hydrophobic Interactions	Tyr237	3.50
	Tyr321	3.18
	Asp166	2.33
	Ser167	3.69
Hydrogen Bonds	Asp171	2.92
	Asp173	3.58
	Lys245	2.44
	Asp166	5.16
	Asp171	4.97
Salt Bridges	Asp171	3.64
	Asp171	2.73
	As173	3.92