Proceedings of the 3rd International Conference on Modern Medicine and Global Health
DOI: 10.54254/2753-8818/82/2025.21546

AI-based Structural Study of CDKN2A Inhibiting MDM2-p53 via AlphaFold2

Wenhao Zhang[1,a,*]

[1]*School of Pharmacy, Dalian Medical University, Dalian Medical University, No. 9, West Section of Lushun South Road, Lushunkou District, Dalian, Liaoning, China*
a. 412634767@qq.com
**corresponding author*

Abstract: The potential role of CDKN2A in substituting MDM2 in its binding with p53 is investigated in this study, emphasizing the significance of wild-type p53 in cancer suppression and its potential contribution to reducing cancer incidence. The protein sequences of CDKN2A, MDM2, and p53 were obtained from the UniProt database and input into AlphaFold2 to predict their three-dimensional structures. Subsequently, potential binding sites within these structures were analyzed using PLIP software. The results provide new insights into the role of CDKN2A in regulating the stability of p53, suggesting that CDKN2A may substitute for MDM2 in its interaction with p53. This research advances the field of structural biology and offers new tools and perspectives for drug discovery and biomedical research.

Keywords: AI (AlphaFold)-based, protein structure prediction, CDKN2A, MDM2, target study

1. Introduction

Cancer remains a global challenge, with its pathogenesis not yet fully elucidated. Researchers have employed various therapeutic strategies, such as chemotherapy, targeted therapy, and immunotherapy, for patient treatment. Among these, targeted therapy aims to target specific oncogenes for treatment. Studies have shown that TP53 is a critical tumor suppressor, playing a key role by inducing cell cycle arrest or apoptosis [1]. The activation of TP53 is detrimental to the initiation and progression of tumors [2]. The synthesis and stability of TP53 are regulated by various modifications, including acetylation and phosphorylation [3]. TP53 degradation primarily occurs through two pathways: MDM2-mediated degradation and autophagy. MDM2, an E3 ubiquitin ligase, is responsible for the degradation of TP53 in wild-type cells [4]. Research indicates that the alternative product of CDKN2A, CDKN2A itself, directly binds to MDM2, blocking MDM2-induced TP53 degradation and enhancing TP53-dependent transactivation and apoptosis to suppress the oncogenic effects of MDM2 [5].

Current research primarily employs experimental methods to explore action sites and protein binding sites. Artificial intelligence (AI) has already established a comprehensive presence in the medical field, creating various cancer screening methods and databases. The use of AI can accelerate the research process, predict potential targets in advance, and provide new ideas for further investigation. As a result, it significantly improves cancer screening and shortens treatment

DOI: 10.54254/2753-8818/82/2025.21546

times. The DeepMind team has developed the first computational method that can predict protein structures with near-experimental accuracy in most cases—AlphaFold2 [6]. This study utilized AlphaFold2 to predict the binding sites of p14ARF with MDM2, thereby preventing the ubiquitination of TP53 and achieving an anti-cancer effect.

- An African-specific variant of TP53 reveals PADI4 as a regulator of p53-mediated tumor suppression
- Targeting p53 pathways: mechanisms, structures, and advances in therapy
- Understanding the complexity of p53 in a new era of tumor suppression
- Drugging p53 in cancer: one protein, many targets
- The alternative product from the human CDKN2A locus, $p14^{ARF}$, participates in a regulatory feedback loop with p53 and MDM2
- Highly accurate protein structure prediction with AlphaFold

2. Method

2.1. AlphaFold2

AlphaFold2, developed by the DeepMind team, is recognized as the most accurate protein 3D structure prediction model to date, revolutionizing the field of biology and demonstrating the potential of AI to expedite scientific advancement. It has been further refined from its predecessor, with a focus on training directly on protein atomic coordinates, which has enhanced the efficiency of elucidating protein structures. AlphaFold2 has achieved precision at the atomic level, matching the measurement accuracy of experimental structural biology.

The architecture of AlphaFold2 is composed of four main components: the input model, the Evoformer, the structural module, and the output module, complemented by iterative loops.

The input model consists of sequence and structural databases. A specific amino acid sequence is initially input, and its homologs are identified within sequence databases such as UniRef 90, BFD, and Mgnify clusters, followed by the performance of multiple sequence alignment (MSA). Feasible 3D structures of homologs are searched for in protein structure databases like PDB and PDB70 clusters [7-9], leading to the construction of a pairwise distance matrix between amino acids. Sequence features and amino acid features are then extracted, and the MSA representation is provided to the Evoformer. The input model serves as the foundation for AlphaFold2's prediction process, directly influencing the accuracy and reliability of subsequent structural predictions.

The Evoformer treats protein structure prediction as a graph reasoning problem in three-dimensional space, facilitating the exchange of information between residue pair representations and MSA representations. It employs self-attention mechanisms to process MSA data, enabling the model to consider information from all other residues when predicting the structure of a particular residue. The self-attention mechanism allows for the exchange of information between residue pairs, capturing long-range dependencies within the protein sequence. A gating mechanism within the Evoformer selectively focuses on certain sequences or residue pairs during MSA processing to enhance prediction accuracy. The Evoformer calculates the outer product mean between residue pairs, assisting the model in understanding the relative spatial relationships between residues. It also utilizes a triangular self-attention mechanism, based on the principle that any two sides of a triangle can influence the third side, to further refine spatial relationships between residues. Through iterative updates of residue representations, the Evoformer gradually optimizes the prediction of protein structures, integrating information from different sequences and using evolutionarily conserved patterns to guide structure prediction. The Evoformer transforms information from MSA and pairwise distance matrices into residue representations, which are then

DOI: 10.54254/2753-8818/82/2025.21546

further processed by the structural module. The output of the Evoformer is a refined residue representation, which is passed to the structural module for the prediction of the protein's 3D structure. The design of the Evoformer enables AlphaFold2 to effectively harness sequence information and evolutionary signals for the prediction of protein 3D structures.

The structural module employs a Transformer neural network to convert the abstract representation of the protein structure into specific three-dimensional spatial coordinates. Within this module, each residue is treated as a separate entity, and the output module uses neural networks to predict the rotation angles and spatial translation positions required to place each residue. These predictions are based on the internal abstract representations and sequence features obtained in previous steps. By applying the predicted rotations and translations to the residues, AlphaFold2 generates the 3D coordinates of the protein, representing its structure in space. The Predicted Local Distance Difference Test (pLDDT) is used by AlphaFold2 as a confidence metric to assess the reliability of the predicted structure. A higher pLDDT value indicates that the predicted local structure is closer to the experimentally measured result. AlphaFold2 refines the prediction results through multiple iterative loops, adjusting parameters based on the current predictions to optimize subsequent results. Ultimately, AlphaFold2 outputs the 3D structure of the protein, which can be used for further biological research, drug design, or other applications. The output 3D structure typically requires professional software for visualization and analysis, aiding researchers in better understanding the protein's functional and structural characteristics.

The iterative loop involves three cycles of data refinement to enhance the reliability of the results.

Despite its achievements, AlphaFold2 has certain limitations. Its predictions of protein structures are only about two-thirds as precise as laboratory measurements. It is currently limited to predicting monomeric structures, and further improvements are needed for predicting protein complexes or multimeric structures. The current interpretability of deep learning models is relatively low, which restricts the application of the model in certain situations. AlphaFold2 mainly predicts static structures and is still unable to accurately predict the real-time dynamic changes and functional states of proteins. Its structure prediction is based on MSA data, which requires a large number of evolutionarily related sequences, potentially leading to slower prediction speeds and other issues.

2.2. UniProt

UniProt is a comprehensive protein knowledge database maintained by the UniProt Consortium, which includes the European Molecular Biology Laboratory's European Bioinformatics Institute (EMBL-EBI), the Swiss Institute of Bioinformatics (SIB), and the National Center for Biotechnology Information (NCBI). UniProt contains the following main sub-databases:

UniProtKB (UniProt Knowledgebase): Provides detailed protein function annotations and sequence information.

UniParc: A comprehensive protein sequence database used to track protein sequences across different databases.

UniRef: Reduces sequence redundancy through clustering to facilitate rapid protein sequence searches.

Proteomes: Contains information on proteomes from various organisms.

PIR-PSD (Protein Information Resource-Protein Sequence Database): A historically significant protein sequence database.

Through UniProt, detailed protein sequences for MDM2, P53, and CDKN2A were obtained and input into AlphaFold2 for three-dimensional structure prediction. The official UniProt website is https://www.uniprot.org.

Specific information is as follows:

DOI: 10.54254/2753-8818/82/2025.21546

MDM2 has the UniProt ID Q00987, and its sequence is:

MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLFYLG
QYIMTKRLYDEKQQHIVYCSNDLLGDLFGVPSFSVKEHRKIYTMIYRNLVVVNQQESSDSG
TSVSENRCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELSGERQR
KRHKSDSISLSFDESLALCVIREICCERSSSSESTGTPSNPDLDAGVSEHSGDWLDQDSVSDQ
FSVEFEVESLDSEDYSLSEEGQELSDEDDEVYQVTVYQAGESDTDSFEEDPEISLADYWKC
TSCNEMNPPLPSHCNRCWALRENWLPEDKGKDKGEISEKAKLENSTQAEEGFDVPDCKKT
IVNDSRESCVEENDDKITQASQSQESEDYSQPSTSSSIIYSSQEDVKEFEREETQDKEESVESS
LPLNAIEPCVICQGRPKNGCIVHGKTGHLMACFTCAKKLKKRNKPCPVCRQPIQMIVLTYFP

P53 has the UniProt ID P04637, and its sequence is:

MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLMLSPDDIEQWFTEDPGP
DEAPRMPEAAPPVAPAPAAPTPAAPAPAPSWPLSSSVPSQKTYQGSYGFRLGFLHSGTAKS
VTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGTRVRAMAIYKQSQHMTEVVRRCPHHE
RCSDSDGLAPPQHLIRVEGNLRVEYLDDRNTFRHSVVVPYEPPEVGSDCTTIHYNYMCNSS
CMGGMNRRPILTIITLEDSSGNLLGRNSFEVRVCACPGRDRRTEEENLRKKGEPHHELPPGS
TKRALPNNTSSSPQPKKKPLDGEYFTLQIRGRERFEMFRELNEALELKDAQAGKEPGGSRA
HSSHLKSKKGQSTSRHKKLMFKTEGPDSD

CDKN2A has the UniProt ID Q8N726, and its sequence is:

MVRRFLVTLRIRRACGPPRVRVFVVHIPRLTGEWAAPGAPAAVALVLMLLRSQRLGQQPL
PRRPGHDDGQRPSGGAAAAPRRGAQLRRPRHSHPTRARRCPGGLPGHAGGAAPGRGAAG
RARCLGPSARGPG

AlphaFold2 is an artificial intelligence program developed by DeepMind, dedicated to the precise prediction of protein three-dimensional structures. Through deep learning and complex computational models, AlphaFold2 is capable of predicting protein conformations at atomic resolution. It finds significant applications in protein structure prediction, biological research, drug design, disease research, protein engineering, and bioinformatics. The three-dimensional structures of MDM2, TP53, and CDKN2A were predicted using AlphaFold2, and potential binding sites within these structures were explored.

2.3. PLIP (Protein-Ligand Interaction Profiler)

PLIP is an advanced tool designed for the analysis of protein-ligand complexes, automatically extracting and characterizing various interactions between proteins and ligands from their three-dimensional structures. Developed by researchers at the University of Gothenburg, PLIP is maintained as an open-source tool and is continuously updated by an active community[10].

The tool can identify and visualize a wide range of interaction types, including hydrogen bonds, water bridges, salt bridges, halogen bonds, hydrophobic interactions, π-stacking, π-cation interactions, and metal complexes. Through its intuitive graphical interface, PLIP provides detailed insights into the structural and interactional aspects of protein-ligand complexes, thereby facilitating a deeper understanding of molecular recognition and interaction mechanisms. This tool is considered invaluable for advancing research in fields such as drug design, structural biology, and molecular modeling.

3. Results

Predicted Structures and Binding Sites of MDM2, P53, and CDKN2A

The amino acid sequences of MDM2, P53, and CDKN2A proteins were retrieved from UniProt. These sequences were then input into the AlphaFold2 program, which employs deep learning and complex computational models to predict their structures. The predicted structures revealed the

DOI: 10.54254/2753-8818/82/2025.21546

binding sites between MDM2 and P53, as well as between CDKN2A and MDM2. The interaction between MDM2 and P53 is characterized by the involvement of α-helices, β-strands, and disordered regions. Similarly, the interaction between CDKN2A and MDM2 is also characterized by the presence of α-helices, β-strands, and disordered regions in Figures 1 and 2.

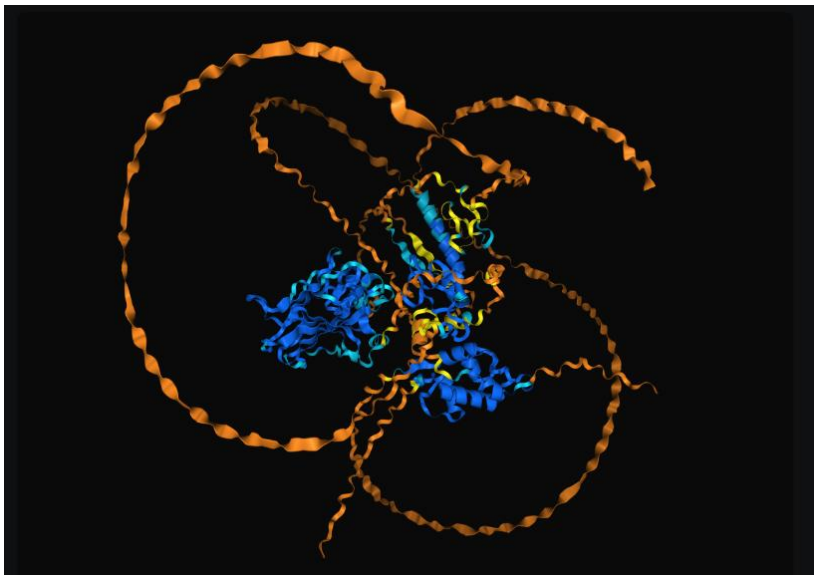


Figure 1: The structure and binding sites of MDM2 and P53, along with their confidence scores. The predicted local Distance Difference Test (plDDT) scores are represented as follows: orange (plDDT < 50), yellow (50 < plDDT < 70), light blue (70 < plDDT < 90), and dark blue (plDDT > 90).

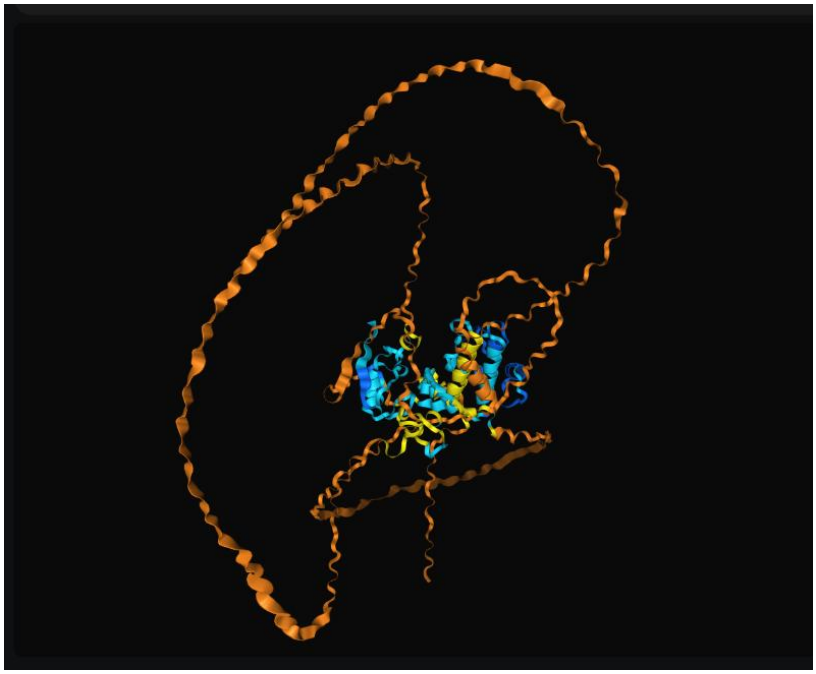
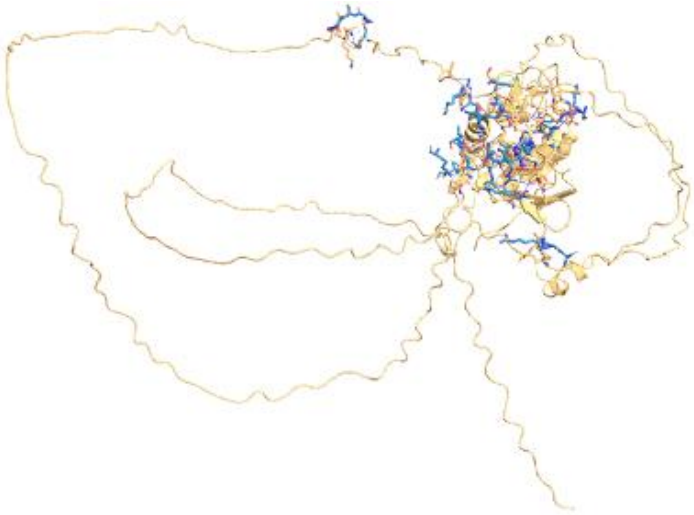


Figure 2: The structure and binding sites of MDM2 and CDKN2A, along with their confidence scores. The predicted local Distance Difference Test (plDDT) scores are represented as follows: orange (plDDT < 50), yellow (50 < plDDT < 70), light blue (70 < plDDT < 90), and dark blue (plDDT > 90).

DOI: 10.54254/2753-8818/82/2025.21546

PLIP Prediction of CDKN2A and MDM2 Interactions



.

Figure 3: The interaction interface of p53 and MDM2. Blue: Protein, Orange: Ligand, Purple: Water, Yellow: Charge center, White: Aromatic ring center, Pink: Metal ion, Dark Blue: Hydrogen bond, Green: π-Stacking, Orange-Yellow: Salt bridge.

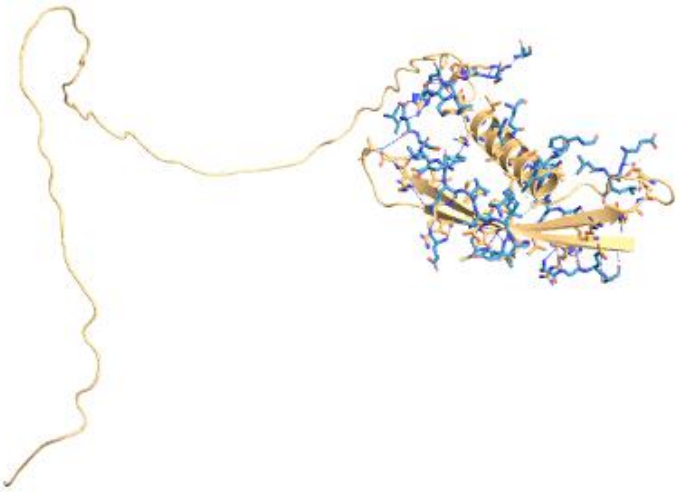


Figure 4: The interaction interface of CDKN2A and MDM2. Blue: Protein, Orange: Ligand, Purple: Water, Yellow: Charge center, White: Aromatic ring center, Pink: Metal ion, Dark Blue: Hydrogen bond, Orange-Yellow: Salt bridge.

PLIP results suggest that the binding site between p53 and MDM2 may involve hydrophobic interactions, hydrogen bonds, and salt bridges as shown in Table 1 and Figure 3. Similarly, the binding site between CDKN2A and MDM2 also involves hydrophobic interactions, hydrogen bonds, and salt bridges as shown in Table 1 and Figure 4. The interference of CDKN2A with the MDM2 binding site may occur at hydrophobic interaction sites at amino acids 19 and 280, and at hydrogen bond sites at amino acids 16 and 283, while no salt bridge interaction is involved.

In summary, the structural predictions for MDM2, P53, and CDKN2A proteins, derived through AlphaFold2, elucidate critical binding interactions that are essential for understanding their biological functions. The analyses indicate that the interactions between MDM2 and P53, as well as between CDKN2A and MDM2, are characterized by key structural elements, including α-helices, β-strands, and disordered regions. Additionally, the PLIP results highlight the significance of

DOI: 10.54254/2753-8818/82/2025.21546

hydrophobic interactions, hydrogen bonds, and salt bridges in these binding sites. Notably, CDKN2A's interference with MDM2 binding may specifically target hydrophobic and hydrogen bond interactions at designated amino acid positions, underscoring the complex regulatory mechanisms involved in these protein interactions.

Table 1: Potential binding sites of p53 and MDM2 (left); Potential binding sites of MDM2 and CDKN2A (right).

p53 and MDM2			MDM2 and CDKN2A		
Hydrophobic Interactions			Hydrophobic Interactions		
1	**17A**	**GLU**	1	**18B**	**GLN**
2	**17A**	**GLU**	2	**19B**	**ILE**
3	**19A**	**PHE**	3	**19B**	**ILE**
4	**19A**	**PHE**	4	**19B**	**ILE**
5	**19A**	**PHE**	5	**21B**	**ALA**
6	**19A**	**PHE**	6	189B	ILE
7	**23A**	**TRP**	7	189B	ILE
8	**23A**	**TRP**	8	191B	LEU
9	**23A**	**TRP**	9	191B	LEU
10	**25A**	**LEU**	10	193B	PHE
11	**25A**	**LEU**	11	193B	PHE
12	26A	LEU	12	195B	GLU
13	26A	LEU	13	197B	LEU
14	27A	PRO	14	199B	LEU
15	28A	GLU	15	199B	LEU
16	30A	ASN	16	202B	ILE
17	32A	LEU	17	204B	GLU
18	32A	LEU	18	205B	ILE
19	32A	LEU	19	245B	PHE
20	49A	ASP	20	245B	PHE
21	51A	GLU	21	245B	PHE
22	52A	GLN	22	254B	LEU
23	52A	GLN	23	274B	GLU
24	136A	GLN	24	274B	GLU
25	137A	LEU	25	276B	TYR
26	181A	ARG	26	278B	VAL
27	**280A**	**ARG**	27	**280B**	**VAL**
28	**280A**	**ARG**	28	**280B**	**VAL**
Hydrogen Bonds			29	285B	GLU
1	15A	SER	30	323B	TRP
2	16A	GLN	31	323B	TRP
3	**17A**	**GLU**	32	323B	TRP
4	**17A**	**GLU**	33	439B	VAL
5	**17A**	**GLU**	Hydrogen Bonds		
6	**17A**	**GLU**	1	**16B**	**THR**
7	**19A**	**PHE**	2	**16B**	**THR**
8	**21A**	**ASP**	3	**16B**	**THR**
9	**25A**	**LEU**	4	**18B**	**GLN**

DOI: 10.54254/2753-8818/82/2025.21546

Table 1: (continued).

10	30A	ASN	5	**19B**	**ILE**
11	32A	LEU	6	**22B**	**SER**
12	49A	ASP	7	**22B**	**SER**
13	49A	ASP	8	**24B**	**GLN**
14	120A	LYS	9	196B	SER
15	120A	LYS	10	196B	SER
16	136A	GLN	11	201B	VAL
17	178A	HIS	12	201B	VAL
18	181A	ARG	13	203B	ARG
19	181A	ARG	14	203B	ARG
20	181A	ARG	15	205B	ILE
21	239A	ASN	16	206B	CYS
22	241A	SER	17	206B	CYS
23	241A	SER	18	208B	GLU
24	243A	MET	19	208B	GLU
25	243A	MET	20	255B	ASP
26	**276A**	**ALA**	21	**273B**	ASP
27	**277A**	**CYS**	22	**275B**	VAL
28	**280A**	**ARG**	23	**275B**	VAL
29	**280A**	**ARG**	24	**276B**	TYR
30	**280A**	**ARG**	25	**277B**	GLN
31	**280A**	**ARG**	26	277B	GLN
32	**280A**	**ARG**	27	277B	GLN
33	**280A**	**ARG**	28	**279B**	**THR**
34	**280A**	**ARG**	29	**279B**	**THR**
35	**283A**	**ARG**	30	**279B**	**THR**
36	337A	ARG	31	**281B**	**TYR**
π-Stacking			32	**281B**	**TYR**
1	178A	HIS	33	**283B**	**ALA**
Salt Bridges			34	**284B**	**GLY**
1	120A	LYS	35	**284B**	**GLY**
2	139A	LYS	36	286B	SER
3	**273A**	**ARG**	37	286B	SER
4	**280A**	**ARG**	38	442B	GLN
5	333A	ARG	Salt Bridges		
			1	204B	GLU
			2	248B	GLU
			3	274B	GLU
			4	274B	GLU
			5	285B	GLU
			6	285B	GLU

4. Conclusion

The UniProt database was utilized in this study to determine the protein sequences of CDKN2A, MDM2, and p53. Subsequently, the protein sequences of MDM2 in complex with p53, as well as

DOI: 10.54254/2753-8818/82/2025.21546

CDKN2A in complex with MDM2, were input into AlphaFold 2 to obtain their three-dimensional structures [11]. Thereafter, the PLIP software was employed to analyze the resulting three-dimensional structures in order to identify potential binding sites [12]..

Artificial intelligence techniques were applied in this study to explore whether CDKN2A can replace the binding of MDM2 with p53, thereby reducing p53 degradation [13, 14]. AlphaFold 2, an artificial intelligence tool capable of predicting protein three-dimensional structures with experimental-level accuracy, was employed, significantly enhancing the efficiency of scientific research. Compared to traditional experimental methods, AlphaFold 2 not only saves time and costs but also provides precise predictive models for many proteins that are challenging to resolve through experimental means. However, certain limitations are associated with AlphaFold's application in dynamic structures and complex biological systems[15]. Overall, as an AI tool, AlphaFold excels in efficiently and accurately predicting protein structures; nonetheless, there is still room for improvement in the prediction of dynamic structures and the practical handling of complex biological systems.

The research findings suggest that p53 may interact with CDKN2A at specific sites through different types of molecular interactions: primarily hydrophobic interactions at amino acid positions 19 and 280 of p53, and predominantly hydrogen bonding at positions 16 and 283. This discovery provides new insights into the potential role of CDKN2A in regulating p53 stability. The developers of AlphaFold were awarded the 2024 Nobel Prize in Chemistry, marking a significant breakthrough for artificial intelligence in the scientific domain. This award recognizes the contributions of the DeepMind team in utilizing deep learning techniques to address the protein folding problem. This achievement not only advances the field of structural biology but also offers powerful tools for drug discovery and biomedical research.

This trend reflects the increasingly important role of AI in scientific research, particularly in data analysis and the resolution of complex problems [16]. It has the potential to transform our understanding and methodologies in scientific inquiry.

References

[1] *Indeglia, Leung, Miller et al. An african-specific variant of tp53 reveals padi4 as a regulator of p53-mediated tumor suppression [J]. Cancer Discov, 2023, 13(7): 1696-1719.*

[2] *Wang, Guo, Wei et al. Targeting p53 pathways: Mechanisms, structures, and advances in therapy [J]. Signal Transduct Target Ther, 2023, 8(1): 92.*

[3] *Liu, Su, Tavana et al. Understanding the complexity of p53 in a new era of tumor suppression [J]. Cancer Cell, 2024, 42(6): 946-967.*

[4] *Hassin, Oren. Drugging p53 in cancer: One protein, many targets [J]. Nat Rev Drug Discov, 2023, 22(2): 127-144.*

[5] *Stott, Bates, James et al. The alternative product from the human cdkn2a locus, p14(arf), participates in a regulatory feedback loop with p53 and mdm2 [J]. Embo j, 1998, 17(17): 5001-5014.*

[6] *Jumper, Evans, Pritzel et al. Highly accurate protein structure prediction with alphafold [J]. Nature, 2021, 596(7873): 583-589.*

[7] *Berman, Westbrook, Feng et al. The protein data bank [J]. Nucleic Acids Research, 2000, 28(1): 235-242.*

[8] *Burley, Bhikadiya, Bi et al. Rcsb protein data bank (rcsb.Org): Delivery of experimentally-determined pdb structures alongside one million computed structure models of proteins from artificial intelligence/machine learning [J]. Nucleic Acids Research, 2022, 51(D1): D488-D508.*

[9] *Berman, Henrick, Nakamura. Announcing the worldwide protein data bank [J]. Nat Struct Biol, 2003, 10(12): 980.*

[10] *Adasme, Linnemann, Bolz et al. Plip 2021: Expanding the scope of the protein-ligand interaction profiler to DNA and rna [J]. Nucleic Acids Res, 2021, 49(W1): W530-w534.*

[11] *Vassilev, Vu, Graves et al. In vivo activation of the p53 pathway by small-molecule antagonists of mdm2 [J]. Science, 2004, 303(5659): 844-848.*

[12] *Salentin, Schreiber, Haupt et al. Plip: Fully automated protein-ligand interaction profiler [J]. Nucleic Acids Res, 2015, 43(W1): W443-447.*

DOI: 10.54254/2753-8818/82/2025.21546

[13] *Kamb, Gruis, Weaver-Feldhaus et al. A cell cycle regulator potentially involved in genesis of many tumor types [J]. Science, 1994, 264(5157): 436-440.*

[14] *Serrano, Lin, McCurrach et al. Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16ink4a [J]. Cell, 1997, 88(5): 593-602.*

[15] *Silver, Schrittwieser, Simonyan et al. Mastering the game of go without human knowledge [J]. Nature, 2017, 550(7676): 354-359.*

[16] *Baek, DiMaio, Anishchenko et al. Accurate prediction of protein structures and interactions using a three-track neural network [J]. Science, 2021, 373(6557): 871-876.*