

# In Silico Analysis of Inhibitors Against the RNA-binding motif 10 (RBM10) Protein in Non-Small Cell Lung Cancer

VIVAAN RANADIVE<sup>1,2</sup> AND GAURAV SHARMA<sup>2</sup>

<sup>1</sup>Green Level High School, NC, Cary

<sup>2</sup>Eigen Sciences, NC, Apex

Published October, 2025

With other 2 million cases around the globe, lung cancer is perhaps the second most frequency cancer. RBM10 protein is a tumor suppressor protein and helps in controlling cell growth and preventing cells from becoming cancerous. Mutations or dysregulation of these genes can contribute to the development and advancement of lung cancer, such as non-small cell lung cancer (NSCLC). Dysregulation of this protein results in the aberrant RNA processing, leading to tumor growth and metastasis, hence rendering it an exceptional target for drug development. However, due to a highly flexible nature of RBM10, targeting the entire protein is very impractical. Therefore, in the current work we have targeted only the RRM1 segment of the complete RBM10 protein. We hypothesize that inhibitors will bind strongly to the RNA binding region RRM1 domain of the RBM10 protein complex, thereby inhibiting the RBM10 functioning by restricting the RNA binding to the RBM10 protein. We have used in silico simulations to predict ligands that bind strongly to the RRM1 protein. Specifically, molecular docking simulations were used to scan 3154 compounds and best ligands were selected for further analysis. The results are in agreement in our hypothesis since the ligand specifically binds to the active site predicted by the graph neural network. Since it binds to the active site the interactions of the RNA to the RRM1 and hence RBM10 will be inhibited.

## 1. INTRODUCTION

Lung cancer is caused by changes in the DNA of the lung cells worldwide (1). These damaged cells get the freedom to multiply uncontrollably and diminishes lung functioning. In normal cell functioning RNA-binding motif 10 (RBM10) protein binds to c-Myc protein and inhibits the oncogenic activity of c-Myc (2). However, mutation in RBM10 protein fails to do so; instead, it promotes lung cancer (3). Due to high flexibility of the RBM10 protein the 3D model of Rbm10 is not known (4). Finding the 3D structure of the protein is crucial in understanding its mechanism and hence designing therapeutic strategies against the disease. c-Myc, a proto-oncogene, is a protein playing an important role in regulating cell proliferation, growth, and differentiation (4). c-Myc is a protein that plays an instrumental role in cell division, cell growth, and cell specialization (5). However, the protein can facilitate the development and progression of cancerous tumors when an abnormal amount of it becomes present (6). Hence, targeting the c-Myc proto-oncogene holds promise for suppressing the development of various cancers, namely lung cancer.

The c-Myc protein's complex structure and involvement in regulatory mechanisms, pathways, and cellular functions makes it challenging to isolate and then inhibit its contribution to cancer progression (7). Currently there are existing cancer treatment strategies that have seen some success (8). In chemotherapy and radiation therapy, cancer cells are killed or their growth is stopped (9). Immunotherapy involves identifying and analyzing a patient's genetic makeup and using that information to target mutations or abnormalities (10). Precision medicine encompasses the identification and analysis of a patient's genetic makeup and employing that information to target abnormalities (11). Targeted therapy, which is the focus of this paper, involves utilizing drugs to block abnormalities within cancer cells.

Computational biochemistry is a field that involves the use of computers to simulate biomolecular processes. Molecular docking is a highly conventional method that has been utilized to discover interactions between ligands and receptors (proteins) for the first step of computer aided drug design (12-15). This technique involves acquiring a 3D structure of a protein and docking a ligand on that protein (12). The software utilizes data on the shape, size, and biochemical structure of the compounds

to predict the best binding site (12). The AutoDock software was predominantly used in this work, however there are other available docking software (16). Understanding the molecular mechanism of the proteins involved in the process can help in designing novel therapeutics against the disease. We hypothesize that inhibitors will bind strongly to the active site of the protein, thereby inhibiting the protein's functioning by restricting the RNA binding to the RBM10 protein. In the current work we have used computational molecular docking tools to predict the interaction between the RBM10 protein and c-Myc protein. Based on our studies we have found that the two proteins form strong interaction with each other. In addition, we have also performed AlphaFold 3 and have predicted the interactions between different components of RBM10 containing RRM1, RRM2, Zinc-finger complex. The current study will not only help in understanding the RNA binding mechanism to these proteins but also help in designing effective therapeutic strategies against the disease.

## 2. RESULTS

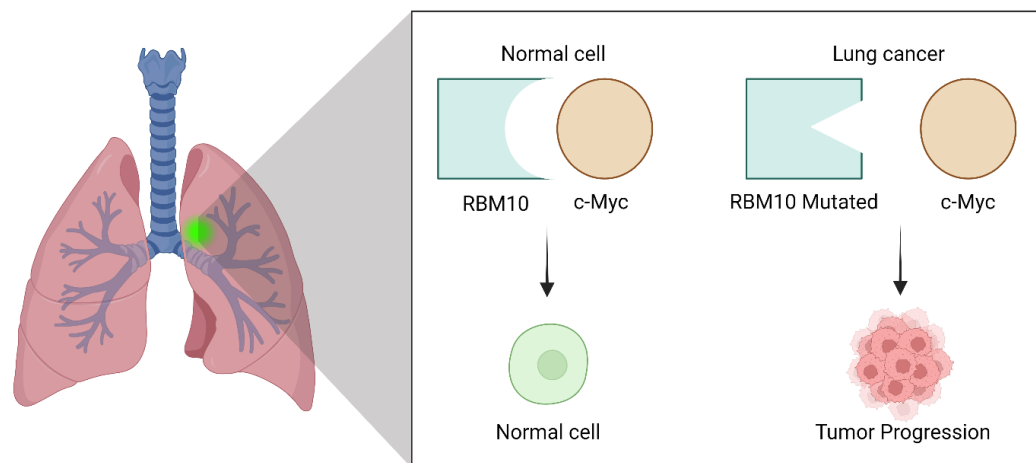
RNA-binding protein (RBM10) is involved in development and is commonly altered in the relation of human disease (17). This protein distinguishes different types of RNA motifs and regulated its processing (17). Due to the complex nature of RBM10 it was very difficult to crystallize the protein; hence, the 3D structure of this protein is not available. To obtain a 3D structure of these complex protein AI technology AlphaFold was used. In Figure 3a, the 3D structure of RBM10 protein bound with RNA is shown. Based on our knowledge this is the first time the RBM10 protein binds with RNA has been predicted. This structure shows that RBM10 protein is a complex and highly flexible protein that has many disordered regions. This property is common in many RNA-binding proteins and is essential for its role in RNA recognition and binding; therefore, RBM10 protein must adopt multiple conformations. In Figure 3a the protein is in three different colors which determines its surface electrostatics (surface charge) called as electrostatic surface potential (ESP). The area in red is negatively charged, blue is positive, while white is neutral. Based on the ESP the negative charged RNA binds to the positive region of the protein. Apart from flexible regions the protein also contains stable, rigid, and crystallizable regions with well-defined conformations. These regions are RRM1, RRM2, and zinc finger domains. Among these RNA recognition motifs (RRM1 and RRM2) are well known for their RNA recognition and binding. Our AlphaFold 3 predicted structure also showed that the RNA binds to the positive region (blue) of the RRM1 protein, further validating our results (Figure 3b). Mutation in the RRM1 gene is associated with lung cancer. This is because the mutation impacts the cell's ability to replicate DNA accurately and repair DNA damage resulting in the uncontrolled cell growth and cancer. Inhibiting RRM1 protein can be a therapeutic strategy to target these cancer cells. Therefore, RRM1 protein was used in this study. We also predicted the druggable binding site of the RRM1 protein using graph neural network (GNN) and the yellow region is the druggable binding site. To validate the AlphaFold structure of RBM10 protein we docked the three components to the RBM10 protein and computed the root mean square distribution (RMSD) shown in Table 1. The AlphaFold 3D model of the complete RBM10 protein is shown in Figure 2. The structure was obtained from the sequence of three components RRM1, RRM2, and Zinc Finger Domain. Based on the 3D structure the complex is highly flexible.

The docking results are shown in Figure 3 and 4. All the ligands binds to the druggable site of the protein. this is the same site where the RNA binds; hence, the ligands proposed in this study have a potential to inhibit the mutated TDP43 protein functioning and could be used in mitigating the cell proliferation and cancer progression. The PLIP website was utilized to compute the RRM1-ligand binding. Based on the PLIP web server, Ligand I formed four hydrophobic interactions with Ile3, Phe46, Phe48, and Asp81 at 3.74, 3.93, 3.84, and 3.85 Å, consequently. Additionally, it also formed 2 hydrogen bonds with Gln42 at 2.19 and 2.63 Å. Ligand II formed hydrophobic interactions with Ile3 twice, Phe46 once, and Asp81 once at 3.67, 3.69, 3.56, and 3.72 Å, correspondingly. It also formed 2 hydrogen bonds with Ser40 at 2.33 and 2.52 Å. Ligand III formed hydrophobic interactions with Ile3 twice, Phe46 once, and Phe 48 once, at 3.79, 3.73, 3.70, and 3.89 Å, consequently. It also formed 2 hydrogen bonds with Arg33 at a 2.79 Å and Gln at 2.73 Å. Ligand IV formed four hydrophobic interactions with Ile3, Phe46, Phe48, and Asp81 at 3.79, 3.64, 3.57, and 3.74 Å, accordingly. Furthermore, it formed two hydrogen bonds with Arg33 at 2.71 and 2.74 Å. Ligand V formed hydrogen bonds with Phe46, Pro84, and Ile86 Å, at 3.96, 3.72, and 3.81 Å. It also formed hydrogen bonds with Arg33 twice and Asp 81 at 2.90, 2.75, and 3.08 Å, correspondingly. Finally, Ligand VI formed hydrophobic interactions with Pro84 and Ile86 at 3.73 and 3.76 Å, consequently and it formed hydrogen bonds with Arg33 twice and Asp81 once at respective 2.87, 2.73, and 3.03 Å.

Finally, the pharmaceutical properties of the selected ligands are shown in Table 3. This shows that all the ligands have high gastrointestinal absorption (GI absorption), high blood brain barrier passing, and high drug likeness. Passing all these properties make them highly likely to be a drug.

## 3. DISCUSSION

AlphaFold 3 was used to obtain the 3D model of the RBM10 protein as shown in Figure 2. RBM10 protein is a big, flexible, and complex protein. Since it contains loose strands, it is highly difficult to crystallize and obtain a 3D model. However, the three components of the RBM10 proteins; (1) RRM1; (2) RRM2; and (3) ZnF are rigid and the 3D structure of these proteins are available (18). The electrostatic surface potential of the RBM10 protein is shown in Figure 2. This shows that the RRM1 protein region has a more positive (blue) region as compared to the rest of the protein. Therefore, it is highly likely that this could be the binding region (close to RRM1 protein) of the negatively charged RNA. The molecular docking simulations were performed only to the RRM1 protein. The molecular docking simulations depicts that the ligand binds only to the drug binding site of the RRM1 which was even more validated by using Graph Neural Network analysis called as GrASP. The compounds that were docked exhibited a major overlap with RNA interaction regions and replicated various connections between RNA base pairs and RRM1 domain. As illustrated by Clery et al, targeting the mutated RRM1 protein can lead to cancer cell apoptosis (19). Finding chemical compounds that can bind strongly to the RRM1 protein can potentially inhibit the RBM10-RNA interactions which in turn can inhibit cell proliferation. To achieve this goal, molecular docking has been performed to find the inhibitor interactions with the RRM1 protein. Based on our results the ligands bind to the RNA binding site of the RRM1 protein. The RNA binding site is well determined by previous studies. We further confirmed that the RNA binding site can also be a druggable binding site by using



**Fig. 1. Disease progression in lung cancer.** In normal cells RBM10 binds to the c-Myc and performs the normal cell process; however, in lung cancer the RBM10 gets mutated and fails to bind to the c-Myc which in turn results in the tumor progression.

GrASP a graphical neural network web server.

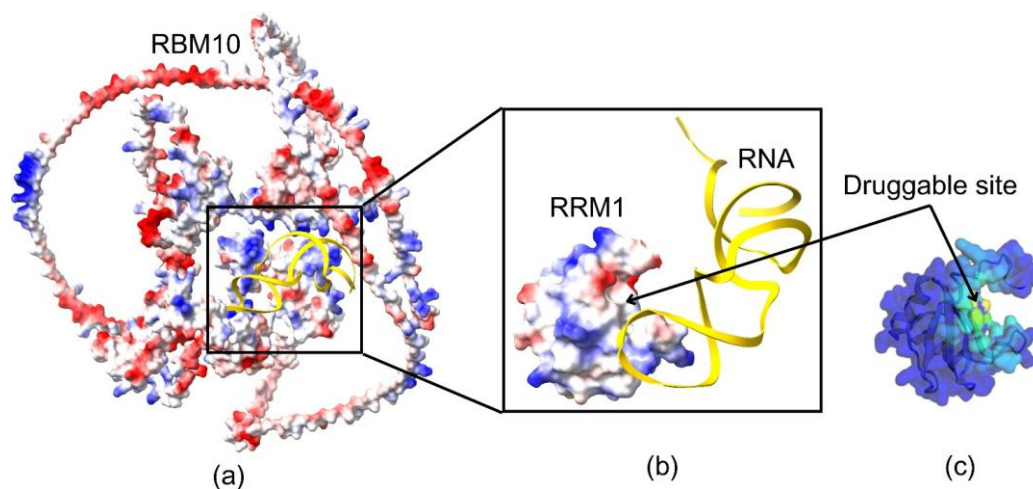
RBM10 protein binds to RNA to form ribonucleoprotein particles and help in cell division. Since the RBM10 protein is highly flexible it is very difficult to get it crystallized hence there are no 3D structure of the protein available (18). Only alpha fold generated 3D structure of the RBM10 protein is available. Due to the recent advent of AlphaFold3 we have generated the predicted 3D structure of RBM10 complexes with RNA. Based on our knowledge, this is the first time the 3D structure RBM10-RNA has been developed. The 3D structure of these complexes is shown in Figure 2. Based on this structure the apo-RBM10 protein is a complex structure with both rigid and loose ends. Since the rigid ends have already been crystallized, we have also compared the predicted and X-ray sequence and have found high similarity with low RMSD Table 1. This validates the AlphaFold predicted structure. The apo-RBM10 protein has both positive (blue) and negative (red) regions. The positive region is where the negatively charged RNA binds as shown in Figure 2. The negatively charged RNA is binding at the center of the protein and we assume for the RNA processing and splicing the RNA passes through the protein. Although the RNA splicing is complex and tightly regulated by different elements. The prediction of RBM10-RNA complex protein will provide shed in the DNA processing and splicing process.

In the current work, we have utilized in silico investigation to

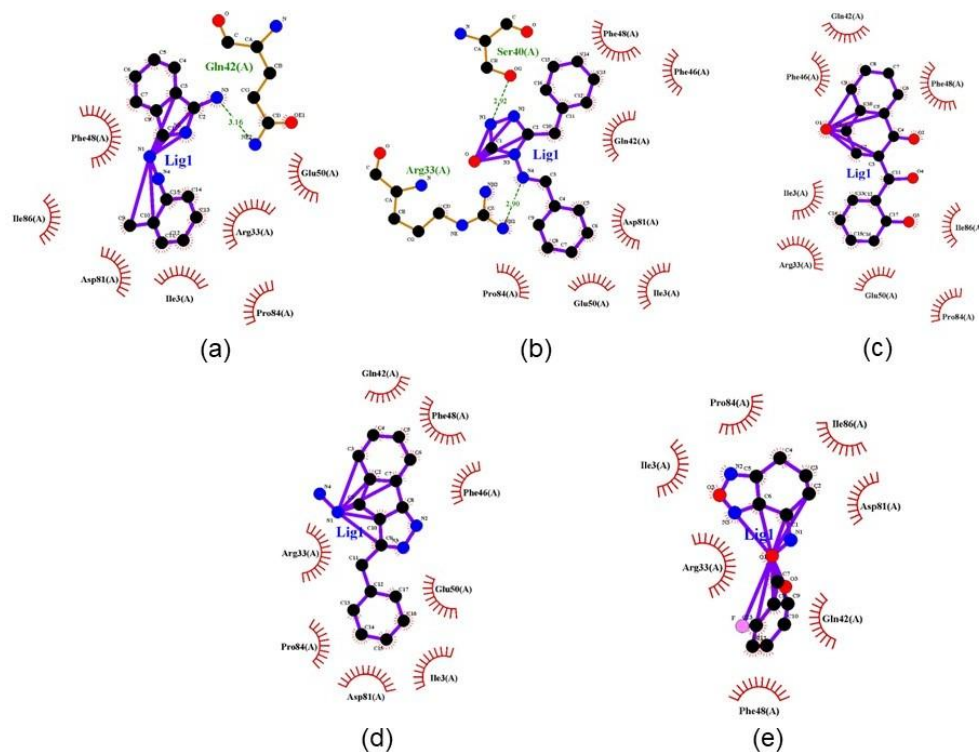
study RBM10 protein and how it can be inhibited. Mutations in RBM10 proteins result in uncontrollable working of c-Myc protein resulting in lung cancer. Therefore, the inhibitors found in this study can be potential inhibitors against the protein. Moreover, we have also performed alpha fold simulations to find the 3D model on RBM10 contains RRM 1, RRM 2, and zinc finger. Based on our model the RBM10 protein is highly flexible and undergoes structural modification during the RNA binding. The current therapeutic strategy will help in designing therapeutic strategies against lung cancer, and neurodegenerative diseases.

#### 4. METHOD

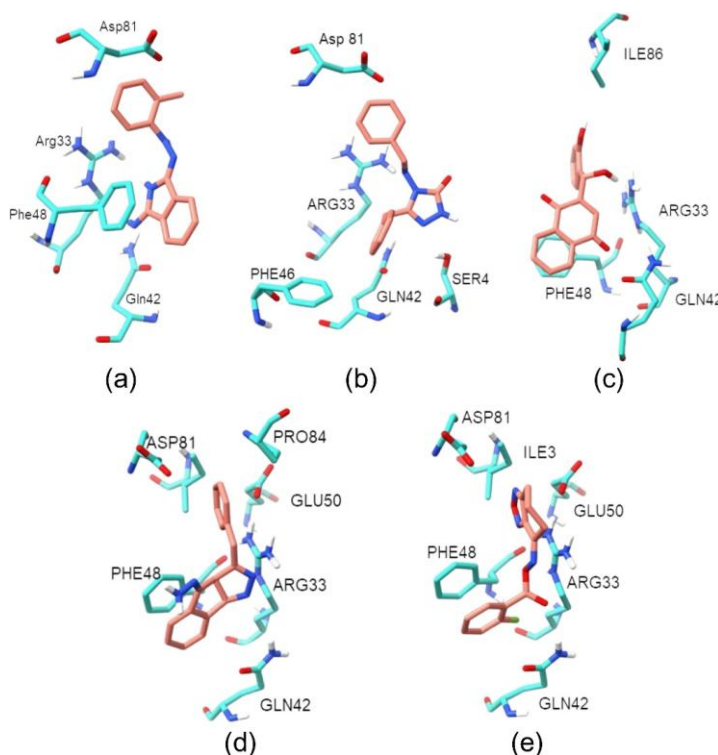
The 3D structure of the RRM1 protein was selected and downloaded via the Protein Data Bank while the 3154 3D ligand compounds were acquired from the Zinc20 database (20). The PDB ID of the protein was 2LXI. The ligands were selected on the basis of their LogP values because the orally active drugs have LogP value of 2. All the ligands were downloaded in .pdbqt file format. Molecular docking simulations were performed using AutoDock Vina 1.5.6 (16) software and 10 binding orientations were obtained for each ligand. Among these 3154 compounds, top five candidates were selected based on the binding energy score provided by autodock vina software. Protein-Ligand Interaction Profiler was utilized to compute the bonds formed



**Fig. 2. RBM10 protein interaction with RNA.** (a) AI generated RBM10 protein binds to RNA. The Electrostatic Surface Potential (ESP) of RBM10 shows the positive (in blue), negative (in red), and neutral (white) charge; (b) The RNA binds to the positively charged RRM1 protein; (b) The druggable site is show in yellow as predicted by Graph Neural Network.



**Fig. 3. 2D protein-ligand interactions.** Based on this image most of the amino acids are forming hydrophobic bonds with the ligands.



**Fig. 4. 3D Protein-ligand interactions.** The ligands are shown in pink and amino acids are shown in cyan.

between the protein and ligands(21). All the complexes were visualized and analyzed using ChimeraX, and PyMol software (22,23). LigPlot software was used to make the 2D images. Swis- sADME was used to find the pharmaceutical properties of the selected ligands. The druggable site was predicted using GrASP web server, which uses graph neural network to find the best possible drug/ligand binding site (24). The 2D structure of complete 3D structure of DNA bound RBM10 protein was obtained using AlphaFold 3 and based on our knowledge this is the first time the complex structure has been elucidated (25).

#### 4. METHOD

The 3D structure of the RRM1 protein was selected and downloaded via the Protein Data Bank while the 3154 3D ligand compounds were acquired from the Zinc20 database (20). The PDB ID of the protein was 2LXI. The ligands were selected on the basis of their LogP values because the orally active drugs have LogP value of 2. All the ligands were downloaded in .pdbqt file format. Molecular docking simulations were performed using AutoDock Vina 1.5.6 (16) software and 10 binding orientations were obtained for each ligand. Among these 3154 compounds, top five candidates were selected based on the binding energy score provided by autodock vina software. Protein-Ligand Interaction Profiler was utilized to compute the bonds formed

#### 5. REFERENCES

- Gridelli, C. et al. "Non-small-cell lung cancer." vol. 1, no. 1, 2015, pp. 1-16.
- Ncube, S.M. "The c-myc/tbx3/nucleolin/hsc70 signalling axis in breast cancer." 2019.
- Cao, Y. et al. "Rbm10 regulates tumor apoptosis, proliferation, and metastasis." vol. 11, 2021, p. 603932.
- Koshre, G.R. et al. "Star-pap rna binding landscape reveals novel role of star-pap in mrna metabolism that requires rbm10-rna association." vol. 22, no. 18, 2021, p. 9980.
- Albihn, A. Roles of myc and mad in cell cycle and apoptosis. Karolinska Institutet (Sweden), 2006.
- Testa, U. et al. "Colorectal cancer: Genetic abnormalities, tumor progression, tumor heterogeneity, clonal evolution and tumor-initiating cells." vol. 6, no. 2, 2018, p. 31.
- Liao, D.J. et al. "Perspectives on c-myc, cyclin d1, and their interaction in cancer formation, progression, and response to chemotherapy." vol. 13, no. 2, 2007.
- Das, S.K. et al. "Gene therapies for cancer: Strategies, challenges and successes." vol. 230, no. 2, 2015, pp. 259-271.
- Baskar, R. et al. "Cancer and radiation therapy: Current advances and future directions." vol. 9, no. 3, 2012, p. 193.
- Tran, E. et al. "Final common pathway of human cancer immunotherapy: Targeting random somatic mutations." vol. 18, no. 3, 2017, pp. 255-262.
- König, I.R. et al. "What is precision medicine?" vol. 50, no. 4, 2017.
- Fan, J. et al. "Progress in molecular docking." vol. 7, 2019, pp. 83-89.
- Nowak, R.P. et al. "Plasticity in binding confers selectivity in ligand-induced protein degradation." *Nature chemical biology*, vol. 14, no. 7, 2018, pp. 706-714, doi:10.1038/s41589-018-0055-y.
- Batta, I. et al. "Molecular docking simulation to predict inhibitors against zinc transporters." 2024.
- Arif, M. et al. "Molecular docking and molecular dynamics simulation to predict inhibitors against hiv envelope 1 protein." 2024.
- Morris, G.M. et al. "Autodock." 2001.
- Warraich, S.T. et al. "Tdp-43: A DNA and rna binding protein with roles in neurodegenerative diseases." vol. 42, no. 10, 2010, pp. 1606-1609.
- Inoue, A.J.G. "Rbm10: Structure, functions, and

associated diseases." vol. 783, 2021, p. 145463.

19. Cléry, A. et al. "Structure of srsf1 rrm1 bound to rna reveals an unexpected bimodal mode of interaction and explains its involvement in smn1 exon7 splicing." vol. 12, no. 1, 2021, p. 428.

20. Irwin, J.J. et al. "Zinc20—a free ultralarge-scale chemical database for ligand discovery." vol. 60, no. 12, 2020, pp. 6065-6073.

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

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

23. Yuan, S. et al. "Using pymol as a platform for computational drug design." vol. 7, no. 2, 2017, p. e1298.

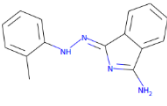
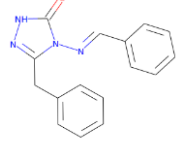
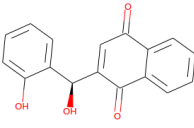
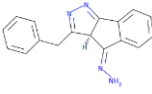
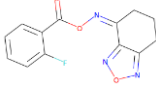
24. Smith, Z. et al. "Graph attention site prediction (grasp): Identifying druggable binding sites using graph neural networks with attention." vol. 64, no. 7, 2024, pp. 2637-2644.

25. Abramson, J. et al. "Accurate structure prediction of biomolecular interactions with alphafold 3." 2024, pp. 1-3.

**Table 2: RMSD between RBM10 predicted by AlphaFold 3 and experimentally derived RBM10 components.** This shows that the AlphaFold 3 predicted structures were highly accurate.

|            |      |       |
|------------|------|-------|
| TDP43-RRM1 | RRM1 | 1.131 |
| TDP43-RRM2 | RRM2 | 1.192 |
| TDP43-ZnF  | ZnF  | 1.108 |

**Table 3: Properties of ligands:** Pharmaceutical characteristics of the compounds are shown in this table.

|                               | 290                                                                                 | 854                                                                                 | 930                                                                                  | 1031                                                                                  | 1756                                                                                  |
|-------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
| Formula,<br>molecular weight  | C <sub>15</sub> H <sub>14</sub> N <sub>4</sub><br>250.30 g/mol                      | C <sub>16</sub> H <sub>14</sub> N <sub>4</sub> O<br>278.31 g/mol                    | C <sub>17</sub> H <sub>12</sub> O <sub>4</sub><br>280.27 g/mol                       | C <sub>17</sub> H <sub>14</sub> N <sub>4</sub><br>274.32 g/mol                        | C <sub>13</sub> H <sub>10</sub> FN <sub>3</sub> O <sub>3</sub><br>275.24 g/mol        |
| GI absorption                 | High                                                                                | High                                                                                | High                                                                                 | High                                                                                  | High                                                                                  |
| BBB permeation                | Yes                                                                                 | Yes                                                                                 | Yes                                                                                  | Yes                                                                                   | Yes                                                                                   |
| Drug likeliness<br>(Lipinski) | Yes                                                                                 | Yes                                                                                 | Yes                                                                                  | Yes                                                                                   | Yes                                                                                   |
| 2D structure                  |  |  |  |  |  |

