Computational Simulation of Aptamer–Protein Interactions Using AptaDock: A Case Study on CD47 Inhibition in Glioblastoma

YASH DEVANG BHUVA[1,2] AND GAURAV SHARMA[2]
[1]Fremont High School, Sunnyvale, CA
[2]Eigen Sciences, Apex, NC

Published September 2025

Glioblastoma multiforme (GBM) is an aggressive brain cancer that originates in the glial cells of the brain. It is the most common malignant brain cancer in adults. CD47 is a transmembrane protein that is overexpressed in several cancers, including glioblastoma multiforme (GBM), where it plays a key role in tumor immune evasion by delivering a 'don't eat me' signal to macrophages via the SIRPα receptor. In GBM, elevated CD47 expression has been correlated with poor prognosis and resistance to immune clearance, making it a promising therapeutic target. Aptamers—short, single-stranded DNA or RNA molecules—have gained increasing recognition for their high specificity and affinity for target molecules. Aptamer-protein affinity studies have revealed promising interactions with biotechnology and molecular medicine applications. In this research work, we have developed a pipeline to select an appropriate aptamer for treating GBM. We hypothesize that aptamer s63 interacts strongly with the predicted binding site of the CD47 receptor and could be used to inhibit the multiplication of GBM cells. We have utilized computational methods, such as AlphaFold 3, to indicate the 3D structure of the CD47 receptor. We have also employed molecular docking simulations using the HDOCK software to investigate the interactions between CD47 receptors and aptamers. The molecular docking simulation shows that the CD47-aptamer complex s66 formed the most hydrogen bonds and salt bridges. On this basis, aptamer s66 was selected as the most appropriate aptamer candidate. The current research will help design novel treatment strategies for GBM and aid in mitigating the global burden of the disease.

1. INTRODUCTION

Glioblastoma multiforme (GBM) is a malignant, rapidly progressing brain cancer that arises from glial cells. [1] Due to its infiltration of the surrounding brain and fierce recurrence, GBM is highly lethal, with a median survival time of only 12 to 18 months after diagnosis. [2] Current treatments include surgery, radiation treatment, and chemotherapy, which are considered palliative and seldom curative. [2] Total surgical removal is impossible; the presence of the blood-brain barrier hinders the effectiveness of systemic chemotherapies as well. Tumors often develop chemotherapy drug resistance, further decreasing the odds of long-term treatment success. These factors underscore the need for innovative strategies against GBM.

Aptamers are synthetic, single-stranded nucleic acids engineered to bind specifically and with high affinity to target molecules. [3] Generated through an optimization selection process known as SELEX, aptamers can fold into diverse shapes and interact with targets through various chemical interactions, allowing them to create binding pockets that fit tightly within a broad array of molecular surfaces. [3] The high specificity, versatility, and ease of modification of aptamers have rendered them to be frequently applied in diagnostics, therapeutics, biosensing, and targeted drug delivery. Computational design further enhances aptamer development by simulating folding sequences into 2D and 3D structures, thereby enabling the prediction and optimization of specific aptamer-target interactions before laboratory testing. Molecular docking methods are employed to further optimize aptamer-target interactions and enhance the predictive capabilities of aptamer designs. [4]

Docking is a computational method used in molecular modeling that provides graphical representations of how two or more molecular structures, often a ligand and a protein, interact by predicting their binding conformations and affinities. [5] Docking is primarily used in the identification of potential drug candidates by screening libraries of compounds to find those that bind effectively to a target protein. [6] The docking software HDOCK was specifically employed in AptaDock's pipeline because it utilizes a hybrid approach that combines template-based and free docking methods, thereby enhancing prediction accuracy and versatility for various biomolecular interactions. Molecular docking simulations play a crucial role in advancing the discovery and development of drugs. Molecular docking has enhanced

treatment for diseases such as cancer and improved effectiveness. Molecular docking has played a pivotal role in accelerating drug discovery pipelines by identifying lead compounds and optimizing their interactions with therapeutic targets. For instance, docking simulations contributed to the development of EGFR inhibitors like Gefitinib and Erlotinib in non-small cell lung cancer. Similarly, docking studies have been used to optimize immune checkpoint inhibitors by predicting interactions with PD-1/PD-L1 complexes. [5] For example, it can identify molecules that inhibit the metabolic pathways of cancer cells, thereby preventing metastasis and further damage. [5] It can also be used to treat COVID-19, inflammatory diseases, and other conditions, making it a significant medicinal discovery.

CD47 is a transmembrane protein commonly referred to as a "don't eat me" signal due to its crucial role in immune evasion. [7] Numerous cells, including healthy ones, express it on their surface, where it interacts with macrophages' Signal Regulatory Protein Alpha (SIRPα) receptor. [7] Normal tissues are shielded from immune attack by this contact, which provides an inhibitory signal that stops the macrophage from engulfing and killing the cell. [8] Many cancer cells, including those of GBM, however, exploit this mechanism by overexpressing CD47, which enables them to evade immune system recognition and removal. Consequently, CD47 has emerged as a promising target for cancer immunotherapy. To develop a new method of treating cancer, researchers aim to restore macrophages' ability to identify and phagocytose cancer cells by inhibiting the CD47-SIRPα interaction. [8] To live and spread, GBM uses a variety of immune evasion techniques. The expression of "don't eat me" signals on the surface of tumor cells is one such technique. [7] By preventing immune cells, such as natural killer cells and macrophages, from phagocytosing the tumor, these proteins enable the tumor to evade immune surveillance. The primary "don't eat me" proteins associated with GBM are listed in table 1, along with an explanation of their mechanisms of action and potential therapeutic applications.

Summary of key 'don't eat me' proteins involved in GBM immune evasion.

In this study, we conducted computational simulations to identify the most suitable aptamer that can interact with the CD47 protein. In this study, we aimed to computationally identify a DNA aptamer that binds with high affinity to the CD47 immune checkpoint receptor—overexpressed in glioblastoma multiforme (GBM)—using molecular docking simulations and interaction analysis. The CD47 protein is expressed on the surface of GBM cells. Therefore, inhibiting these proteins or receptors can also help inhibit the cancer protein in GBM. By targeting CD47—a key immune checkpoint protein overexpressed in GBM—aptamer-based inhibition may prevent immune evasion and facilitate the destruction of tumor cells by macrophages. The current research will pave the way for developing new medicines that can be used in the treatment of GBM.

2. METHOD

We have developed an integrated pipeline for computational simulations of protein-aptamer docking. The pipeline consists of the following steps: (a) Preparation of proteins: Target protein structures were retrieved from the Protein Data Bank (PDB). To ensure compatibility and optimal docking conditions, the proteins were prepared and modified using in-house scripts. These included the addition of missing atoms, correction of structural inconsistencies, and optimization of the conformation to make protein docking feasible. Our scripts also prepared the protein in an appropriate format for subsequent docking simulations.
b) Molecular docking simulation: Molecular docking simulations were performed using the HDOCK software, a powerful tool for predicting the binding modes of protein-protein and protein-nucleic acid complexes. [5] The prepared target proteins underwent aptamer docking to explore the potential binding sites and interaction modes. Post-docking and custom scripts were developed for interaction analysis of the protein-aptamer interactions. The schematic of the computational pipeline is shown in Figure 1.

To predict the structure of the CD47 protein, we began by retrieving the amino acid sequence from the UniProt database. [16] It is a database containing protein amino acid sequences, structures, and properties. The next step involved uploading it to the AlphaFold 3 web server to get the predicted structure for the CD47 protein. [17] AlphaFold 3 is the newest AI tool designed to indicate the 3D shape of proteins based on their genetic sequences. It accurately predicts 3D structures of proteins from their amino acid sequences, aiding biological research and drug discovery. For aptamer 3D structure prediction, we began by modeling the aptamer from 1D to 2D. First, we folded tbinding locationhe DNA by uploading the aptamer sequence we had found earlier. We then downloaded the .ct file for further use. Before proceeding, we converted the .ct file into dot-bracket notation. The next step involved converting the 2D model to a 3D structure. We converted the DNA sequence to RNA and input the corresponding dot-bracket notation, setting the number of clusters to one, and submitted it with the VFOLD tool for 3D aptamer structure prediction. [18] To perform molecular docking, we first accessed the HDOCK server and submitted the protein and ligand data for docking analysis. After completion, we were able to view the results, as shown below. [19] Following the docking process, we proceeded with interaction analysis using the PLIP analysis web server. [20] This tool allowed us to investigate and visualize the molecular interactions formed between the docked molecules.

3. RESULTS

This paper develops a computational pipeline called AptaDock, utilizing HDOCK and analysis, to identify aptamer s66 as the most suitable candidate for inhibiting the CD47 receptor in GBM by forming the most stable interactions with the receptor.

Binding site prediction: To determine the binding site, the protein's binding site was identified using ScanNet. [21] ScanNet is a deep learning-based method that predicts the probability of amino acids binding to other amino acids or biomolecules. The ScanNet-predicted sites are typically used for comparison with the binding sites obtained from molecular docking simulations. It enables molecular docking to determine the optimal placement of ligands. Figure 2 shows the binding site of the protein.

Molecular docking simulation: Molecular docking simulations have been performed using the HDOCK software, a powerful tool for predicting the binding modes of protein-protein and protein-nucleic acid complexes. [5]

Molecular docking analysis: Subsequently, the web tool Protein-Ligand Interaction Profiler (PLIP) was employed to analyze the protein-aptamer complexes (Salentin, Schreiber, Haupt, Adasme, & Schroeder, 2015). The hydrogen bonds with the CD47 protein indicated the strength of specific aptamers and helped identify the optimal sequences for intervention. The analysis shows that s66 forms the most stable and intense interaction

Table 1. Summary of key 'don't eat me' proteins involved in GBM immune evasion.

Protein	Receptor	Mechanism	Explanation
CD47 [9]	SIRPα on macrophages	Inhibits phagocytosis by delivering a "don't eat me" signal	Most well-studied; highly expressed in GBM stem-like cells.
PD-L1 [10]	PD-1 on T-cells	Suppresses T-cell activation (indirect "don't attack me" signal)	Immune checkpoint; commonly targeted in immunotherapy.
CD24 [11]	Siglec-10 on macrophages	Inhibits macrophage-mediated phagocytosis	Recently identified in GBM and other solid tumors.
MHC Class I [12]	LILRB1 and others	Inhibits natural killer (NK) cell activity	Normal function, but overexpressed to avoid immune detection.
B7-H3 (CD276) [13]	Unknown	Inhibits T-cell and NK cell activation	Overexpressed in gliomas; involved in immune evasion.
Calreticulin (CALR) [14]	LRP1 on macrophages	Usually acts as an "eat me" signal, but suppressed in GBM	Low CALR expression helps maintain immune evasion.
CD200 [15]	CD200R on myeloid cells	Suppresses macrophage and dendritic cell activity	Found in glioma stem cells and promotes an immunosuppressive microenvironment.

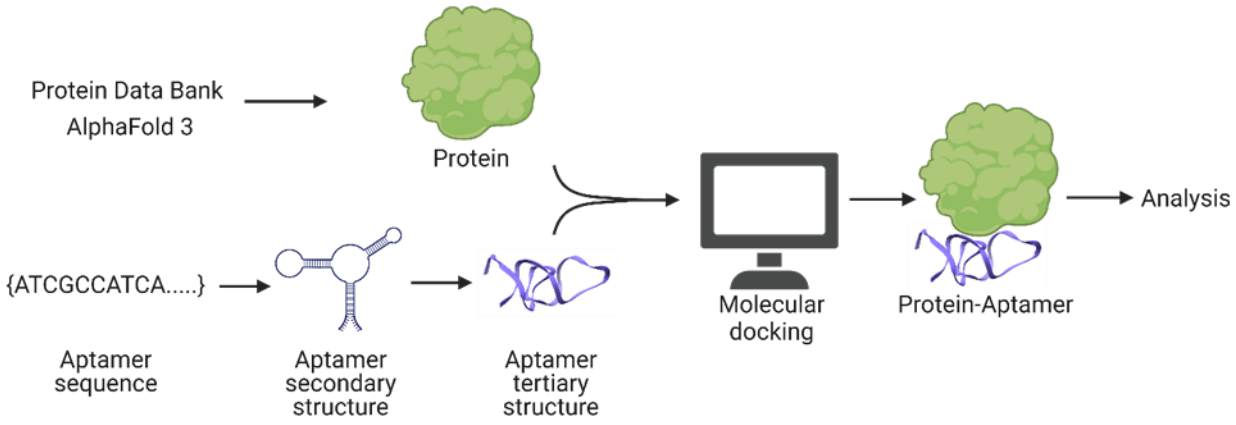

Fig. 1. Schematic of AptaDock. This figure illustrates a computational pipeline for simulating the interaction between a pro- tein and an RNA or DNA aptamer through molecular docking. The workflow begins with the acquisition of the protein's three-dimensional structure, either obtained directly from the Protein Data Bank or predicted using AlphaFold 3. In parallel, the aptamer modeling process starts with its nucleotide sequence, which is used to generate the secondary structure. This secondary structure is then folded into a three-dimensional tertiary structure using aptamer modeling tools. Once both the protein and aptamer structures are prepared, they are subjected to molecular docking simulations to predict their binding interactions. The resulting protein–aptamer complex is then analyzed to evaluate binding affinity, interaction sites, and potential biological relevance.

with the CD47 protein. Based on the PLIP analysis, aptamer s66 was selected as the most appropriate aptamer with highest number of interactions as displayed in Figure 3.

In addition, we have also identified similar proteins linked to the CD47 protein, as shown in Figure 4a. Among the proteins listed—CD47, SIRPA, SIRPB1, SIRPG, SCARB1, SCARB2, CD36, THBS1, CD200R1, RHAG, and GYPB—several have been associated with glioblastoma (GBM), a highly aggressive brain tumor. CD47, known as the "don't eat me" signal, is overexpressed in various cancers, including GBM. Its interaction with SIRPA on macrophages inhibits phagocytosis, allowing tumor cells to evade immune surveillance. Blocking the CD47-SIRPA interaction has been shown to increase phagocytosis of GBM cells and reduce tumor burden in preclinical models. SIRPA, as the receptor for CD47, plays a role in the immune evasion mechanism of GBM cells. Therapeutic strategies targeting the CD47-SIRPA axis are being explored to enhance anti-tumor immune responses in GBM. CD36, a fatty acid translocase involved in the uptake of long-chain fatty acids and oxidized lipoproteins, also interacts with THBS1 (thrombospondin-1), an extracellular matrix protein. The CD36-THBS1 interaction has been implicated in GBM pathogenesis, particularly in modulating tumor cell invasion and angiogenesis. THBS1, being a multifunctional protein, influences tumor progression by interacting with various receptors, including CD36 and CD47. In GBM, THBS1's interaction with CD47 has been shown to affect tumor cell invasion and angiogenesis, contributing to the tumor's aggressive nature. The roles of SIRPB1, SIRPG, SCARB1, SCARB2, CD200R1, RHAG, and GYPB in GBM are not well-established based on current literature.

4. DISCUSSION

Proposed mechanism of aptamer mediated inhibition of GBM cells. Figure 4 illustrates how cancer cells evade immune clearance and how aptamer-based therapy can prevent this from occurring. GBM cancer cells overexpress CD47, a surface protein that binds to macrophages' SIRPα receptor and serves as a "don't eat me" signal. This connection enables the tumor to evade immune surveillance by sending an inhibitory signal that

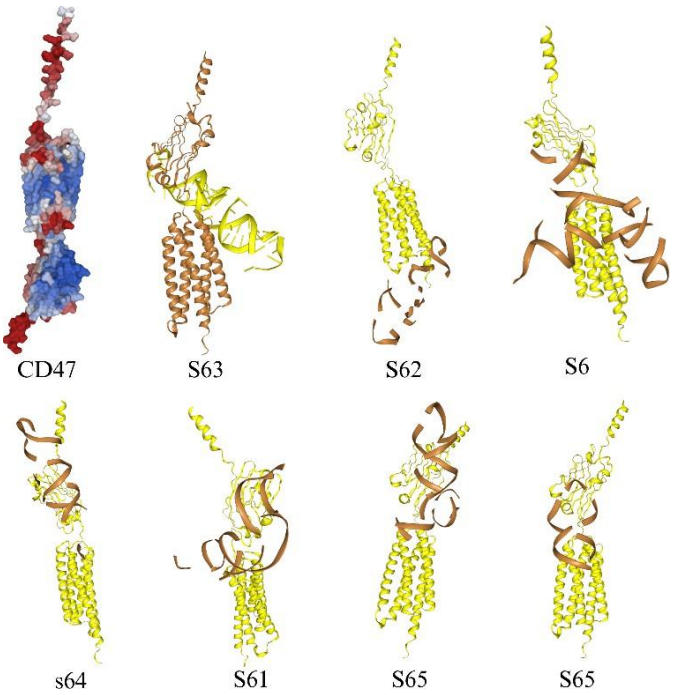

Fig. 2. Molecularly docked structures of CD47 with various aptamers. The leftmost panel shows the surface representation of CD47, followed by ribbon models of CD47–aptamer complexes (S63, S62, S6, S64, S61, and S65). Each structure highlights the interaction interface, illustrating the binding orientations and conformational differences among the docked aptamer candidates.

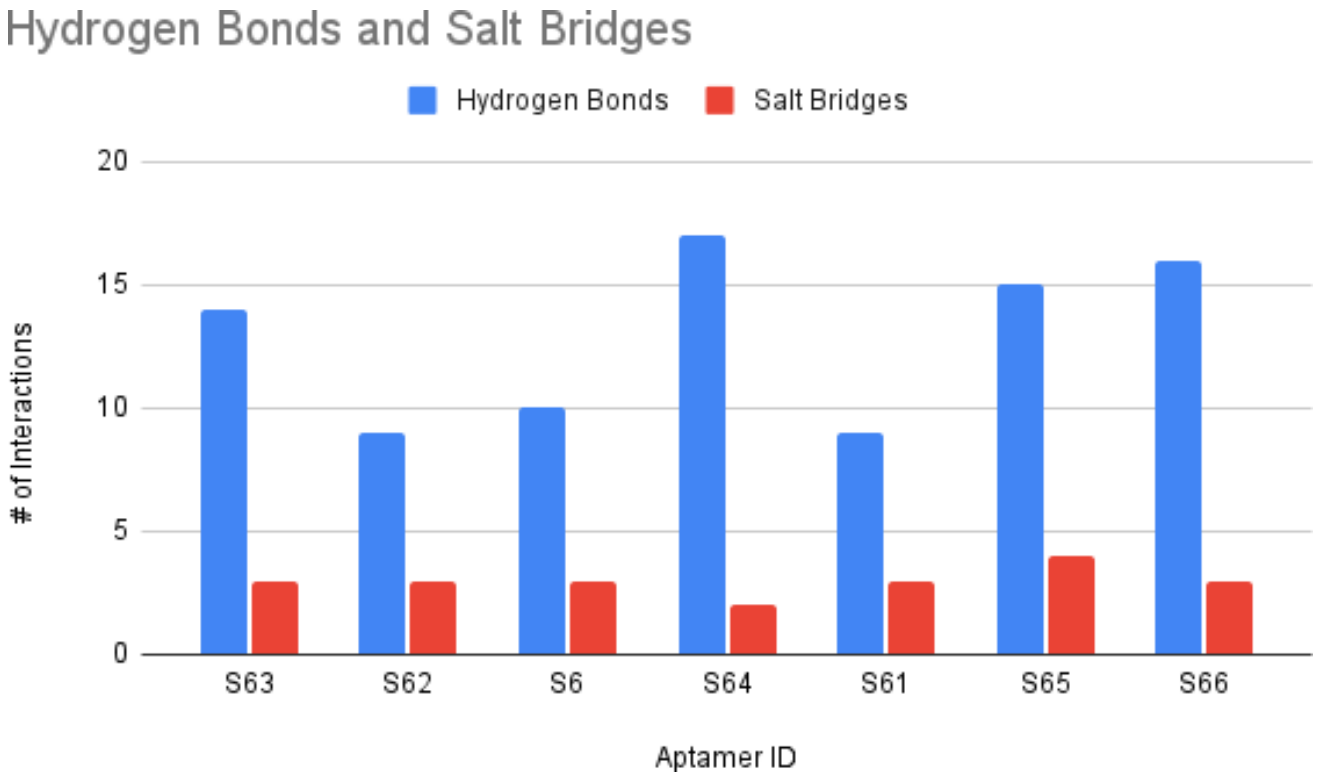

Fig. 3. CD47-aptamer interactions table computed using the PLIP web server. The interaction analysis shows that aptamer s66 formed the most interactions and was selected as the best aptamer.

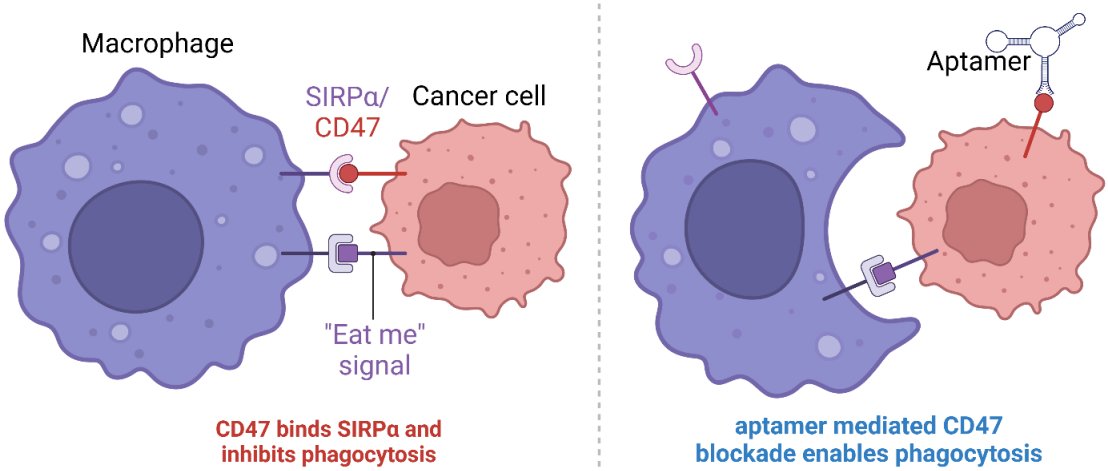

Fig. 4. This figure illustrates how aptamer-mediated blockade of CD47 restores phagocytosis. On the left, the interaction between CD47 on the cancer cell and SIRPα on the macrophage sends a “don’t eat me” signal, preventing engulfment despite the presence of pro-phagocytic (“eat me”) cues. On the right, an aptamer binds to CD47, blocking its interaction with SIRPα. This removes the inhibitory signal, allowing the macrophage to recognize the “eat me” signal and proceed with phagocytosis of the cancer cell. The image was generated using Biorender.com.

prevents macrophages from engulfing the GBM cells. However, this interaction is prevented when a rationally designed aptamer with high affinity for CD47. The disruption of the inhibitory signal caused by CD47’s inability to interact with SIRP*α* allows the macrophage to identify the cancer cell using its "eat me" signals and begin phagocytosis. A possible immunotherapeutic approach to reestablish the immune system’s capacity to recognize and destroy cancer cells is the aptamer-mediated inhibition of CD47.

The interaction landscape and expression profile of CD47 and similar proteins are depicted in Figure 5, highlighting their role in immune evasion, which is particularly relevant in GBM. Using the STRING database, a protein–protein interaction (PPI) network was created in panel (a), with CD47 at its core. Essential proteins that interact with CD47 are highlighted in this network, including its receptor SIRP*α* (SIRPA), SCARB1, SCARB2, CD36, THBS1, CD200R1, SIRPG, SIRPB1, RHAG, and GYPA/GYPB. [22] These linkages, which encompass both functional associations and physical binding, suggest a larger immunoregulatory module that CD47 coordinates. Gene expression data, most likely from single-cell RNA sequencing or spatial transcriptomics, is superimposed on top of a UMAP or t-SNE plot in panel (b). Red markers indicate the expression of CD47 and its interaction partners, while individual black dots represent solitary cells. The coordinated expression and possible co-regulation of this immune checkpoint module within particular tumor or immune cell populations are highlighted by the colocalization of these genes in particular cellular clusters. This data is essential in improving GBM treatment by helping scientists understand how the tumor evades immune system attacks. The network of proteins centered around CD47 reveals how GBM cells send “don’t eat me” signals to immune cells, such as macrophages, enabling them to survive and proliferate. By studying which of these proteins are most active in different parts of the tumor, researchers can identify the most effective targets for new treatments. For example, blocking CD47 or other related proteins, such as PD-L1 or CD24, could help immune cells recognize and destroy the cancer. Since GBM employs several immune evasion strategies, this data also suggests that combining different therapies may be more effective than using a single one. It also helps scientists design more effective treatments, such as engineered immune cells (e.g., CAR-T cells), that can evade being blocked by the tumor’s defense signals.

(a) Protein-protein interaction network showing CD47 and its interacting partners (e.g., SCARB1, SCARB2, SIRPA, THBS1, etc.) and (b) UMAP plot highlighting query genes (e.g., CD47, SCARB1, SCARB2, SIRPA) within a broader gene expression dataset.

GBM is a fast-growing, highly malignant version of a brain tumor. [1] It tends to infiltrate the surrounding normal brain

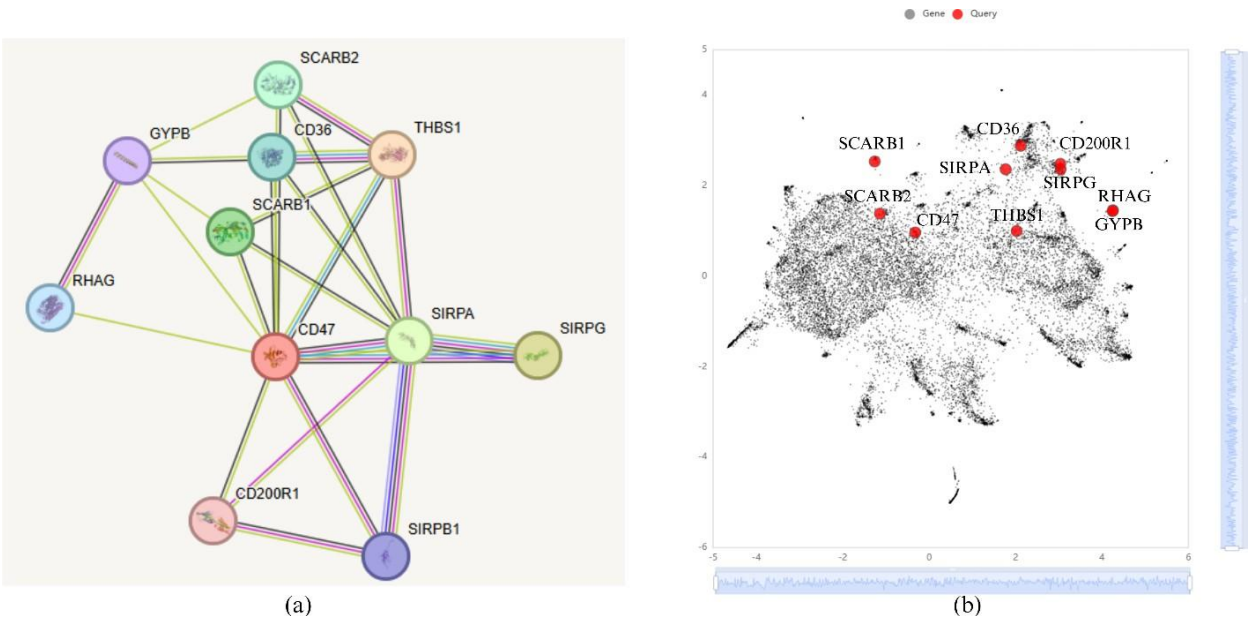

Fig. 5. (a) Protein-protein interaction network showing CD47 and its interacting partners (e.g., SCARB1, SCARB2, SIRPA, THBS1, etc.) and (b) UMAP plot highlighting query genes (e.g., CD47, SCARB1, SCARB2, SIRPA) within a broader gene expression dataset.

tissue and is conventionally challenging to treat. This type of tumor is typically resistant to therapies, such as surgery, radiation, and chemotherapy, due to its complex genetic and cellular heterogeneity, which forms barriers that protect cancer cells from drug molecules. Moreover, GBM is an invasive tumor; hence, complete surgical resection is almost impossible, and residual cancer cells frequently give rise to recurrence. The inefficiency of existing therapeutic modalities, coupled with the dismal prognosis of GBM—15 months median survival rate post-diagnosis—emphasizes the pressing requirement for new targeted therapeutic strategies.

Limitations: The Computational docking of aptamers on proteins has several limitations. One primary challenge is the inherent flexibility of aptamers, which can adopt multiple conformations, making it difficult to accurately predict their binding modes. [5] Additionally, nucleic acids have specific interactions with the model, further complicating the docking process. Another limitation is the potential for non-specific binding due to unintended interactions, which could render an inaccurate reference aptamer potentially detrimental.

In future studies, we will also design, compute, and perform chemical programs that can be used to convert aptamers from their primary structure to secondary structure, and then to tertiary structure. This process can be utilized in the development of aptamer 3D structures and for molecular docking simulations. Connecting the AlphaFold server to the docking process will be a speedy approach within this pipeline. Additionally, utilizing diffusion-based deep learning methods, such as DiffDock, can also accelerate the docking process and further enhance computational speed.

To enhance the predictive accuracy and efficiency of our aptamer-protein interaction pipeline, we propose integrating advanced structure prediction and flexible docking tools. Specifically, AlphaFold 3 will be further automated within the pipeline to streamline the generation of high-confidence 3D structures for protein targets, eliminating manual intervention and improving scalability across multiple receptors. Additionally, we plan to incorporate DiffDock, a diffusion-based deep learning model for molecular docking, to address one of the core limitations of current docking methods: ligand flexibility. In our proposed workflow, DiffDock will replace or complement HDOCK by performing flexible docking of aptamer 3D structures to CD47 and related immune checkpoint receptors. DiffDock's generative diffusion architecture enables it to better sample the conformational space and predict multiple plausible binding poses. Following docking, PLIP or similar interaction profilers will still be used to assess hydrogen bonds, salt bridges, and π-stacking interactions. Moreover, we will explore DiffDock's internal scoring functions for ranking aptamer candidates and compare them with traditional scoring methods. This integration will allow more accurate modeling of flexible aptamer conformations and binding poses, thereby improving the robustness of aptamer screening for GBM and other cancers.

5. CONCLUSION

In our research, various computational simulation techniques, including AlphaFold 3 for structural modeling and HDOCK for molecular docking, were employed to investigate the potential of aptamers to bind and inhibit CD47, a key immune checkpoint receptor overexpressed in glioblastoma multiforme (GBM). Binding site prediction using ScanNet, followed by docking and interaction analysis with PLIP, identified aptamer s66 as exhibiting the strongest predicted binding affinity and the highest number of stabilizing interactions with CD47 among the candi-

dates studied. Based on these results, aptamer s66 is a promising candidate for further experimental validation as a potential therapeutic agent for targeting CD47-mediated immune evasion in GBM. Additionally, the AptaDock pipeline developed in this study demonstrates an effective computational workflow for aptamer screening and can be extended to other cancer-associated targets. Continued refinement of aptamer modeling and integration with experimental studies will be essential to advance these findings toward clinical application.

REFERENCES

1. N. H. Rekers, P. Sminia, and G. J. Peters, "Towards tailored therapy of glioblastoma multiforme," J Chemother **23**, 187–99 (2011).
2. C. Alifieris and D. T. Trafalis, "Glioblastoma multiforme: Pathogenesis and treatment," Pharmacology & Therapeutics **152**, 63–82 (2015).
3. A. D. Keefe, S. Pai, and A. Ellington, "Aptamers as therapeutics," Nature Reviews Drug Discovery **9**, 537–550 (2010).
4. G. T. Rozenblum, V. G. Lopez, A. D. Vitullo, and M. Radrizzani, "Aptamers: current challenges and future prospects," Expert Opin Drug Discov **11**, 127–162 (2016).
5. J. Fan, A. Fu, and L. Zhang, "Progress in molecular docking," Quantitative Biology **7**, 83–89 (2019).
6. O. Korb, "Potential and Limitations of Ensemble Docking," Journal of Chemical Information and Modeling **52**, 1262–1274 (2012).
7. A. Russ, "Blocking "don't eat me" signal of CD47-SIRPα in hematological malignancies, an in-depth review," Blood Reviews **32**, 480–489.
8. H. Wu, J. Liu, Z. Wang, W. Yuan, and L. Chen, "Prospects of antibodies targeting CD47 or CD24 in the treatment of glioblastoma," CNS Neuroscience & Therapeutics **27**, 1105–1117 (2021).
9. A, "Phagocytosis Checkpoints in Glioblastoma: CD47 and Beyond," Current Issues in Molecular Biology **46**, 7795–7811 (2024).
10. Y. Han, D. Liu, and L. Li, "PD-1/PD-L1 pathway: current researches in cancer," Am J Cancer Res **10**, 727–742 (2020).
11. A. A. Barkal, "CD24 signalling through macrophage Siglec-10 is a target for cancer immunotherapy," Nature **572**, 392–396 (2019).
12. A. A. Barkal, "Engagement of MHC class I by the inhibitory receptor LILRB1 suppresses macrophages and is a target of cancer immunotherapy," Nat Immunol **19**, 76–84 (2018).
13. W. T. Zhou and W. L. Jin, "B7-H3/CD276: An Emerging Cancer Immunotherapy," Front Immunol **12**, 701006–701006 (2021).
14. M. P. Chao, "Calreticulin is the dominant pro-phagocytic signal on multiple human cancers and is counterbalanced by CD47," Sci Transl Med **2**, 63–94 (2010).
15. A. Shao and D. M. Owens, "The immunoregulatory protein CD200 as a potentially lucrative yet elusive target for cancer therapy," Oncotarget **14**, 96–103 (2023).
16. U. Consortium, "UniProt: a hub for protein information," Nucleic acids research **43**, 204–212 (2015).
17. J. Abramson (2024).
18. X. Xu, P. Zhao, and S.-J. J. P. O. Chen (2014).
19. Y. Yan, D. Zhang, P. Zhou, B. Li, and S.-Y. J. N. Huang (2017).
20. S. Salentin, S. Schreiber, V. J. Haupt, M. F. Adasme, and M. J. N. Schroeder (2015).
21. J. Tubiana, D. Schneidman-Duhovny, and H. J. Wolfson, "ScanNet: an interpretable geometric deep learning model for structure-based protein binding site prediction," Nature Methods **19**, 730–739.
22. D. Szklarczyk, "The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest," Nucleic Acids Res **51**, 638–646 (2023).