In Silico Design and Evaluation of DNA-Based Aptamers for Triple-Negative Breast Cancer (TNBC) Diagnosis and Targeted Therapy

SAHASRA Yenumula, GAURAV SHARMA

Published October, 2025

Triple Negative Breast Cancer (TNBC) is a devastating, highly aggressive form of breast cancer with a high recurrence rate originating from carcinoma cells. It poses a significant threat to human health. A distinctive characteristic of TNBC cells is that these cancer cells lack expression of the HER2, progesterone, and estrogen receptors commonly present in other subtypes of breast cancer. Therefore, TNBC does not respond to conventional breast cancer therapy and requires novel treatment approaches. One promising approach focuses on targeting the overexpression of Mucin1 receptors, which represents a valuable target for cancer diagnosis and treatment, given its role in promoting tumor cell formation. Short single-stranded DNA or RNA molecules, known as aptamers, have shown great potential in selectively binding to specific targets with high affinity and specificity. This study investigates the feasibility of designing aptamers that bind to Mucin1 receptors on TNBC cells, paving the way for targeted therapies. In this research I have used various computational methods such as AlphaFold 3, UNAfold, and FARFAR2. The AlphaFold 3 and FARFAR2 was used to model the 3D structure of Mucin1 and aptamer structures, respectively. followed by docking simulations using the HDOCK software. The results showed that the aptamer HIF1 has the strongest binding energy and the highest number of salt bridges with Mucin1, while KVC4 forms the most hydrogen bonds. The current research will help in developing targeted therapies for breast cancer by designing aptamers that can specifically bind to tumor cells, potentially enhancing early detection and treatment strategies.

1. INTRODUCTION

Breast cancer, the most prevalent cancer among women globally, accounts for 2.3 million new cases and 670,000 deaths annually. A pervasive global health concern, breast cancer affects women of all ages post-puberty, with incidence rates escalating in later life (1),(2). Among the various subtypes, Triple Negative Breast Cancer (TNBC) is particularly concerning due to its aggressive nature and lack of targeted therapies, leading to poorer outcomes compared to other breast cancer types. This subtype is characterized by the absence of estrogen, progesterone, and HER2 receptors, which limits treatment options and contributes to its more harmful effects. TNBC is more frequently diagnosed in younger women and is associated with higher recurrence rates. Breast cancer, which is diagnosed when a tumor is formed in the breast region, typically has varying symptoms, including lumps, skin changes, nipple discharge, nipple and breast transformations, and pain. (3) (4) While early-stage breast cancer has relatively high 5-year survival rates (up to 99%), TNBC has significantly lower survival, particularly due to high recurrence and metastasis rates. Standard treatments for breast cancer, such as surgery, chemotherapy, and radiation therapy, (5), are often found to be less effective for TNBC, due to its unique biological characteristics, underscoring the urgent need for more specialized therapeutic approaches. Moreover, the rapid progression of TNBC often leads to late-stage diagnosis, further complicating treatment and negatively impacting patient outcomes. Continued research into the molecular mechanisms of TNBC is essential for developing innovative therapies that can improve survival rates and quality of life for affected individuals.

Mucin 1 is a large, heavily glycosylated protein found on the surface of many human cells. It forms a protective layer on the cell surface.(6) Mucin 1 is found in various tissues, including the gastrointestinal tract, lungs, and breast.(6) Overexpression of Mucin 1 has been linked to the progression of triple-negative breast cancer, as it can help cancer cells evade the immune system and spread to other parts of the body.(6) Triple-negative breast cancer (TNBC) is a subtype of breast cancer that lacks three key receptors: estrogen, progesterone, and the HER2 protein.(7) This makes it more challenging to treat, as it does not respond to hormonal therapies or targeted treatments commonly used for other breast cancer types, leading to poorer outcomes and a higher likelihood of recurrence, as shown in Figure 1.(8)

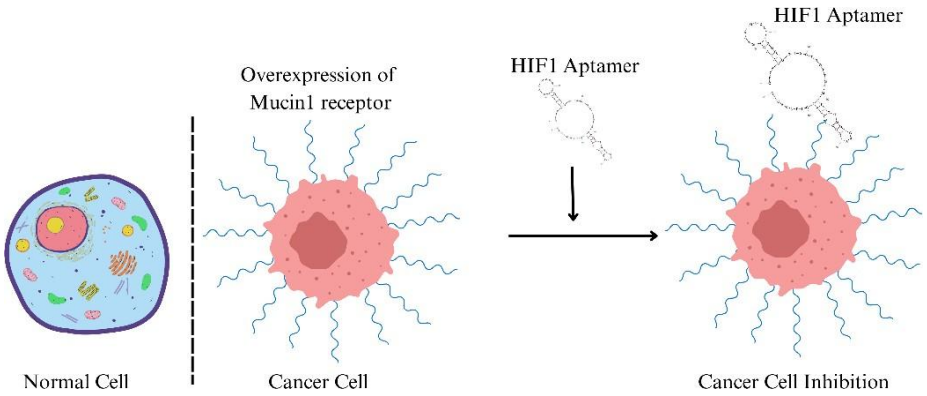

Fig. 1. In cancer cell the Mucin1 receptor gets overexpressed . Anticancer drug conjugated aptamer binds to these receptors and results in the destruction of these cancer cells.

Aptamers are short, single-stranded DNA or RNA molecules that can bind to specific targets with high affinity and specificity.(9) Through a process called SELEX, aptamers have a promising future in medicine and research. (9) Due to the removal of complex biological processes in the pathogen termination, aptamers have a promising future in medicine and research.(9) For instance, aptamer-based biosensors can detect and quantify biomolecules like proteins and small molecules, while aptamers aid in identifying diseases or pathogens in diagnostics.(10) Moreover, aptamers can target and deliver drugs to specific cells or tissues, offering massive prospective benefits in drug delivery and therapy.

Molecular docking simulation predicts how a ligand will interact with a receptor, such as a protein, by simulating how the ligand will fit onto the protein's binding site.(11) Using this computational biology tool, I can discover the ligand's best possible orientation and binding, allowing for maximal effects.(11) This technique is crucial in drug discovery, as it helps identify potential drug candidates by evaluating their binding affinities and specific interactions with target proteins.(11) Additionally, molecular docking can provide insights into the mechanisms of action of variouss compounds, aiding in the design of more effective drugs.(11) By utilizing scoring functions and algorithms, researchers can rank different ligand conformations, streamlining the selection process for experimental validation. Molecular docking is a powerful approach to enhance our understanding of molecular interactions in biological systems.

Recently, Eskandani et al. designed aptamers that can specifically interact with MUC1 receptors on the surface of breast cancer cells. (12) I hypothesize that the aptamers used in this research will bind to the specific site of the mucin receptor. The computational simulations revealed that the design binds to the active site of the receptors and could be used in the detection and targeted therapy of the cancer cells. The current research will lay the groundwork for further research in the development of targeted therapy for breast cancer.

2. RESULTS

Aptamer secondary structure modeling:

Surface properties of mucin1: The binding site predictions were made by two different software: ScanNet and GrASP. The binding site is a specific region on a protein target that the aptamer binds to. Figures 2a and 2b indicate the binding site, visualized by two different software, ScanNet and GrASP. Electrostatic Surface potential is a 3D visualization of a molecule's charge distribution that shows regions of varying charges. To get the ESP, I used the software ChimeraX. Figure 3c in the diagram below depicts the ESP of the aptamer, indicating the charge distribution in various areas.

Molecular docking simulations: In the next step, I performed molecular docking simulations to understand the aptamer-muc1 binding interactions. This step will help us select the better binding aptamer. Molecular docking is a computational biology technique that predicts how molecules fit together and their binding affinity. For molecular docking, I used the HDOCK2.0 software. The aptamer-mucin1 binding structures obtained from molecular docking simulations are shown in Figure 4.

Molecular docking analysis: Binding energy is the energy associated with the interaction between the aptamer and its target, indicating the strength and stability of their binding. I used the tool, PDA-Pred, short for Protein-DNA complex Binding Affinity Prediction to find the binding energies for the various aptamers. The lowest binding energy (most negative) indicates the strongest, best aptamer, which was revealed to be HIF1. Protein-ligand interaction Profiler analysis identifies and characterizes interactions between proteins and ligands, including

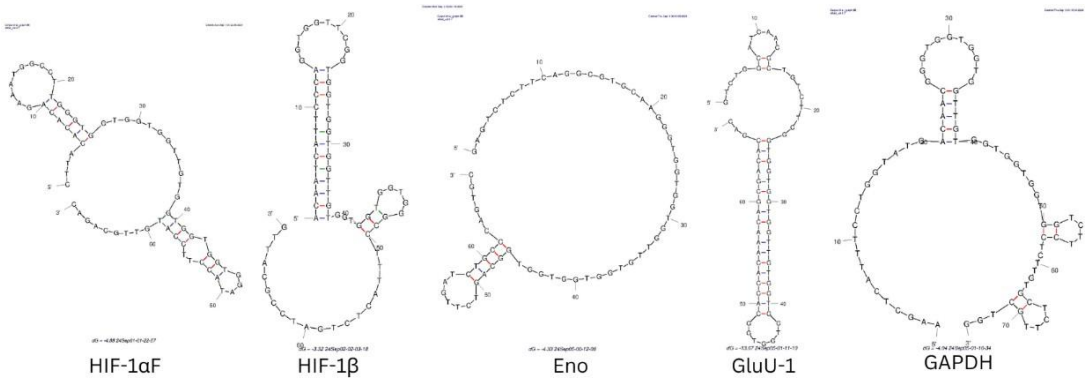

Fig. 2. Secondary Structures (2D) of aptamer used in this research. These 2D structures show the loops, stems, and tail. These structures are crucial in understanding their functional properties for diverse applications.

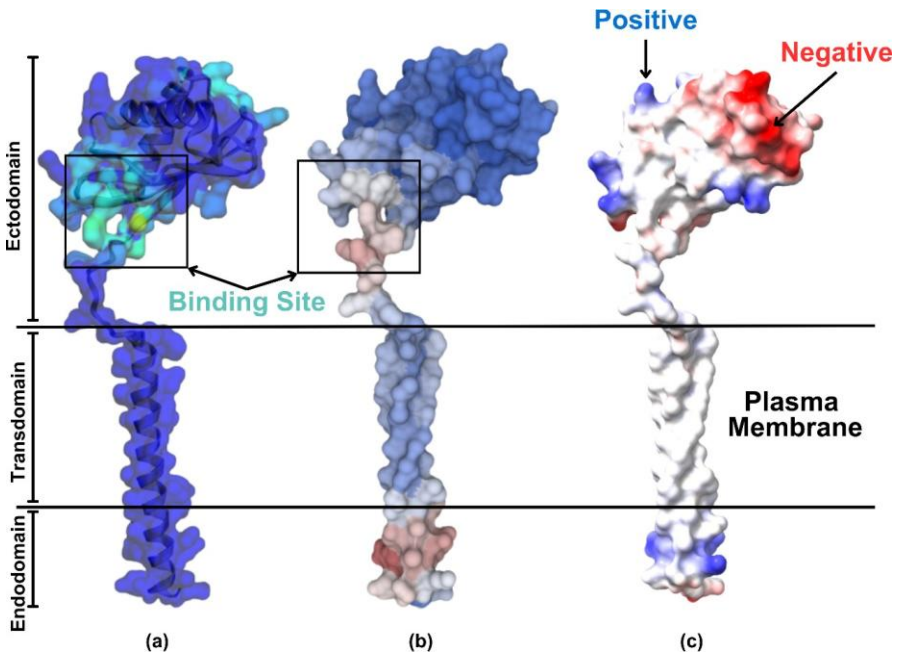

Fig. 3. Surface Properties of Mucin1. (a) it depicts the binding sites of the aptamer, visualized using GrASP and pinpointed in the boxes; (b) similar to (a), (b) also depicts the binding sites of the aptamer within the box, and was visualized using ScanNet; and (c) it shows depicts the ESP of the mucin1, visualized using ChimeraX.

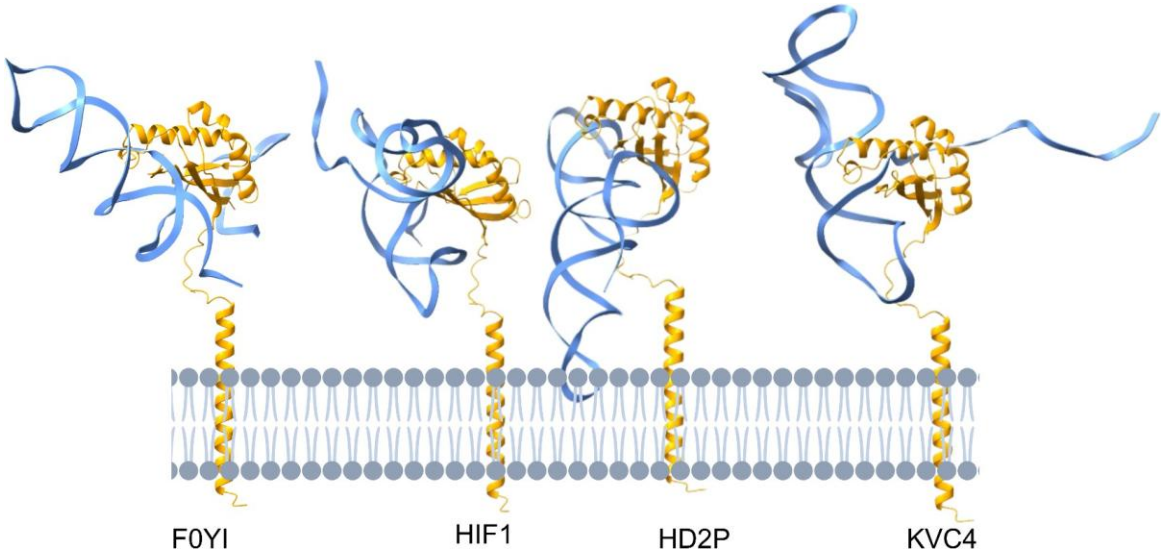

Fig. 4. Molecular Docking Structure of Mucin and Aptamers : All the aptamers bind to the druggable site of the protein, as shown in Figures 2a and b.

hydrogen bonds, hydrophobic interactions, and salt bridges, to understand binding mechanisms and optimize aptamer design for higher affinity and specificity. HIF1 had the most salt bridges, while KVC4 had the most hydrogen bonds.

3. DISCUSSION

Typical breast cancer treatment methods primarily include surgery, radiation therapy, and chemotherapy.(13). However, these treatments often encounter challenges due to their limited effectiveness and the risk of harming healthy tissue, leading to significant side effects. Surgical resection aims to remove as much of the tumor as possible, but it often leaves residual cancerous cells, leading to the recurrence of the cancer.(13) Radiation therapy can effectively target tumor cells in Triple Negative Breast Cancer; however, it may also cause significant damage to surrounding healthy tissue, leading to adverse side effects and complications.(13) Chemotherapy, while useful, often has limited efficacy against Triple Negative Breast Cancer due to the tumor's aggressive nature and the lack of targeted receptors, which complicates effective drug delivery. These conventional treatments often struggle against Triple Negative Breast Cancer due to the tumor's unique biology and the absence of hormone receptors that typically guide treatment strategies. The best treatment for TNBC would be more comprehensive, newer approaches such as immunotherapy and targeted therapies. These methods could potentially improve outcomes by specifically addressing the unique molecular features of TNBC and enhancing the immune response while minimizing damage to healthy tissues.

Targeted therapy for TNBC is a highly promising approach that aims to selectively kill cancerous cells while sparing normal cells, minimizing collateral damage. One potential strategy involves targeting the Mucin receptor, which is found to be more prevalent in TNBC cells. This specificity enhances the therapeutic effect while reducing side effects associated with conventional treatments, making it a viable and compelling treatment alternative. However, there are notable limitations to this approach. Currently, there is a lack of experimental validation in laboratory settings, which is crucial for establishing the efficacy and safety of this targeted therapy. Future work should prioritize conducting rigorous lab experiments to confirm these findings and optimize treatment protocols. Additionally, the limited number of available aptamers poses a challenge for this strategy. Expanding the repertoire of aptamer screenings will be essential to identify more effective candidates for targeting carcinoma cells. Addressing these limitations will pave the way for more effective and innovative treatment options for TNBC patients, ultimately improving patient outcomes and quality of life.

4. CONCLUSION: In this research, I computationally modeled and analyzed aptamer binding to the Mucin1 receptor using advanced AI and molecular docking techniques. Through comprehensive analysis, I found that the HIF1 aptamer emerged as the most promising candidate, demonstrating the lowest binding energy of -14.95 kcal/mol and the most salt bridges. The primary application of this research is to revolutionize early detection and treatment of TNBC by designing aptamers that selectively bind to the carcinoma cells. Computationally identified promising aptamer candidates that warrant further experimental validation have been researched in this work. Moving forward, the future direction involves translating these insights into targeted therapeutic agents, incorporating nanoparticle-based

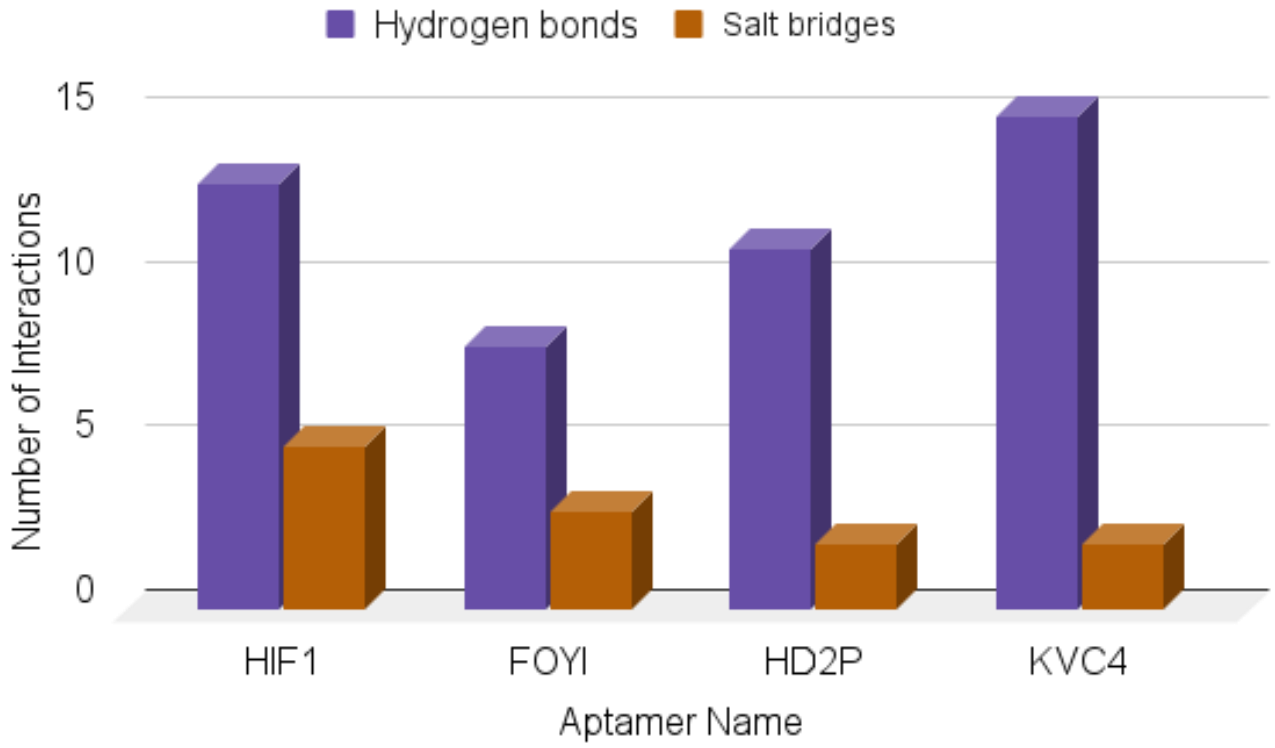

Fig. 5. Visualization and Comparison of Identified Hydrogen Bonds and Salt Bridges in Each Aptamer through PLIP Analysis. The number of interactions formed between the receptors and antibodies formed.

drug delivery systems that specifically target cancer cells with an overexpression of Mucin1 overexpressing cancer cells. In vitro and in vivo testing will evaluate the aptamer's efficacy in delivering therapeutic payloads, while exploring combination therapies to enhance specificity and effectiveness, particularly for triple-negative breast cancer. This research has the potential to significantly impact cancer treatment, offering a promising avenue for overcoming the challenges associated with this devastating disease.

5. METHODS: We first obtained the protein sequence using the UniProt software to predict the Mucin 1 receptor structure and design bindable aptamers.(14) A protein sequence is the protein's order of amino acids, which determines its structure and function. Then, I used AlphaFold 3 to predict the 3D structure of the receptor.(15) AlphaFold3 is a deep learning-based system that can accurately predict protein structures from amino acid sequences.(15) Next, I identified the surface binding site using GrASP, a computational tool to predict protein-protein binding sites based on the 3D structures of the proteins involved.(16) A binding site on the surface of the Mucin 1 receptor is a specific region or pocket where other molecules, such as aptamers or proteins, can bind. The following method was then used to obtain the aptamers' 3D structures: I first converted the 1-dimensional sequence into a 2D structure by uploading the aptamer sequence into a DNA Folding Form software. Using the output ct file, I then obtained a representation of the sequence through dot-bracket notation. Next, using the RNA sequence, I converted the 2D structure into a 3D model using VFOLD3D.(17) Finally, the Mucin1 and aptamers binding interactions were computed by molecular docking simulations using the HDOCK web server.(18) Molecular docking is a computational technique to predict a molecule's binding mode and affinity to a target molecule. To complete molecular docking, I used HDOCK, a software tool developed by a research group at Huazhong University of Science and Technology.

6. REFERENCES:

1. Waks AG, Winer EP. Breast Cancer Treatment: A Review. JAMA. 2019;321(3):288-300.
2. Eden CM, Johnson J, Syrnioti G, Malik M, Ju T. The Landmark Series: The Breast Cancer Burden of the Asian American Population and the Need for Disaggregated Data. Annals of Surgical Oncology. 2023;30(4):2121-7.
3. Wang L. Early Diagnosis of Breast Cancer. Sensors. 2017;17(7):1572.
4. Bish A, Ramirez A, Burgess C, Hunter M. Understanding why women delay in seeking help for breast cancer symptoms. Journal of Psychosomatic Research. 2005;58(4):321-6.
5. Hortobagyi GN, de la Garza Salazar J, Pritchard K, Amadori D, Haidinger R, Hudis CA, et al. The Global Breast Cancer Burden: Variations in Epidemiology and Survival. Clinical Breast Cancer. 2005;6(5):391-401.
6. Mukherjee P, Madsen CS, Ginardi AR, Tinder TL, Jacobs F, Parker J, et al. Mucin 1-Specific Immunotherapy in a Mouse Model of Spontaneous Breast Cancer. Journal of Immunotherapy. 2003;26(1).
7. Irvin WJ, Carey LA. What is triple-negative breast cancer? European Journal of Cancer. 2008;44(18):2799-805.
8. Kumar P, Aggarwal R. An overview of triple-negative breast cancer. Archives of Gynecology and Obstetrics. 2016;293(2):247-69.
9. Dunn MR, Jimenez RM, Chaput JC. Analysis of aptamer discovery and technology. Nature Reviews Chemistry. 2017;1(10):0076.
10. Seok Kim Y, Ahmad Raston NH, Bock Gu M. Aptamer-based nanobiosensors. Biosensors and Bioelectronics. 2016;76:2-19.

11. Fan J, Fu A, Zhang L. Progress in molecular docking. Quantitative Biology. 2019;7(2):83-9.
12. Eskandani M, Mohabbat A, Karimiyan A, Dadashi H, Adibkia K, Sanaat Z, et al. MUC1 aptamer-conjugated niclosamide-loaded PLGA-PEG nanoparticles attenuate HIF-1 stabilization upon hypoxia in MCF7 breast cancer cells. Journal of Drug Delivery Science and Technology. 2024;92:105278.
13. Alifieris C, Trafalis DT. Glioblastoma multiforme: Pathogenesis and treatment. Pharmacology & Therapeutics. 2015;152:63-82.
14. Consortium TU. UniProt: a hub for protein information. Nucleic Acids Research. 2014;43(D1):D204-D12.
15. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature. 2024;630(8016):493-500.
16. Smith Z, Strobel M, Vani BP, Tiwary P. Graph Attention Site Prediction (GrASP): Identifying Druggable Binding Sites Using Graph Neural Networks with Attention. Journal of Chemical Information and Modeling. 2024;64(7):2637-44.
17. Li J, Zhang S, Zhang D, Chen S-J. Vfold-Pipeline: a web server for RNA 3D structure prediction from sequences. Bioinformatics. 2022;38(16):4042-3.
18. Yan Y, Zhang D, Zhou P, Li B, Huang S-Y. HDOCK: a web server for protein–protein and protein–DNA/RNA docking based on a hybrid strategy. Nucleic Acids Research. 2017;45(W1):W365-W73.

Table 1. Basepair sequences of aptamers used in this research.

Gene	Sequence	Binding Energy (kcal/mol)
HIF1	CTTACACACAGAAATGGCCTTGGGTGGTGGTGGTTGTGGTGGTGGTGGATACCTTCCATGTTGCAGAC ....(((.((..........)).)))............((((.(((....))).)))).....	-14.95
HD2P	ACAATCATTCCCAGGTGGTTCGGTGGTGGTGGTTGTGGTGGTGGTGGGCCTTTACTCTGATCCGCATTG (((((((((..(((..........)))..)))))))))...(((.....))).....	-13.68
FOYI	GTCTGGCATCAACGCTGTCTTCGGTGGTGGTGGTTGTGGTGGTGGTGGCACCACAAACAGCGACACGAC .....((.....))........(((.((.((.(((((((((.......)))))))).)).)).)))...	-14.24
KVC4	AAGCTCATTTCCTGGTATGACAACGGGTGGTGGTGGTTGTGGTGGTGGTGGTCTTCCTCTTGTGCTCTTGCTGG .....(((((..........)))))..........((....))......((....))...	-14.75