In Silico Design and Validation of DNA-Based Aptamers Targeting Tumors Using Bacterial-Mediated Biotherapies

MADHAV GOPI[1,2] **AND GAURAV SHARMA**[2]

[1]*Lambert High School, 805 Nichols Rd, GA, Suwanee*
[2]*Eigen Sciences, NC, Apex*

Compiled December 27, 2024

Bacteria-mediated biotherapy, also known as bacterial therapy or bacteriotherapy, is a treatment that utilizes live bacteria or bacterial products to treat various medical conditions. Despite their extensive exploration for treating many cancer types, the limited therapeutic efficiency of bacterial-mediated biotherapies, mainly because they do not accumulate tumor-specifically after delivery, has hindered their widespread application. However, this work brings hope by exploring the potential of aptamers in cancer-targeted therapy. Aptamers, nucleotides with the ability to bind targets similarly to antibodies, were the focus of our study. We used computational modeling to understand the interaction mechanism between aptamers and the nucleolin protein, which is overexpressed on the surface of cancer cells. We hypothesize that the aptamers used in this research will show specific binding interactions with the nucleolin protein. The aptamer PDB files were obtained using Vfold2D and Vfold3D programs, and these structures were then docked onto the nucleolin protein. Our studies revealed that the ULF aptamer formed strong interactions with the nucleolin binding site, further validating our results. This will enhance our understanding of the binding mechanism of aptamer to the nucleolin protein. The predicted aptamers can manipulate bacterial behaviors by changing the bacterial surface and can be an effective tumor imaging tool or therapeutic agent against the disease.

1. INTRODUCTION

Bacteria Biotherapy is where bacteria and their products are used to treat various diseases, cancer being one of them (1, 2). Utilizing the natural properties of certain bacteria, bacteria biotherapy aims to track and destroy cancer cells, regulate the immune system, and supply tumors with therapeutic agents (1). Certain bacteria can pick and choose what tumors they colonize due to the unique microenvironment of cancer tissues, which often have hypoxic and

necrotic regions (1). Clostridium and Bifidobacterium are good examples of species (3). By activating immune cells like dendritic cells and T-cells, bacteria can increase the body's ability to detect and attack cancer cells, boosting the immune system (4). Some bacteria can also produce toxins that harm cancer cells, causing apoptosis (5). Bacteria can also be genetically modified to give different genes to cancer cells, which can either kill the cells or enhance treatment (6). Bacteria can also be used for detection if the tumor is in an earlier stage (4).

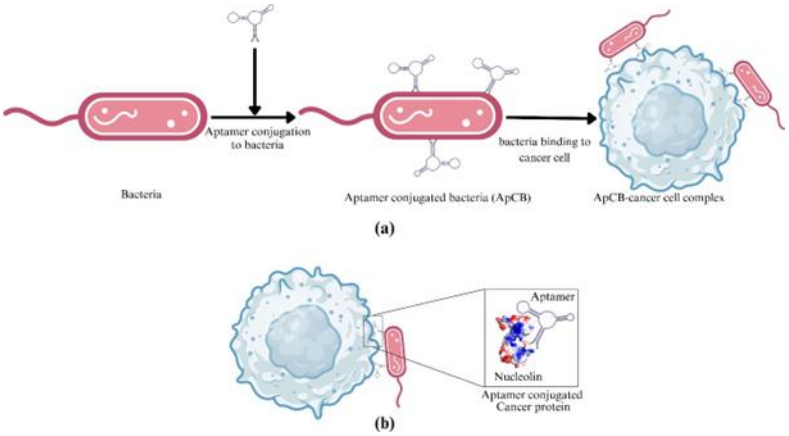

Fig. 1. Mechanism of aptamer interaction with bacteria. (a) Aptamer-assisted bacterial tumor localization is needed for improved biotherapy and (b) aptamer conjugation to the bacterial protein (PD1) on the surface.

Aptamers are small, single-stranded oligonucleotides or peptides that bind with attention to detail and coordinate with a target molecule (7). The selection of aptamers is a meticulous process using the SELEX method (8). Aptamers are versatile in their ability to bind to various molecules, differing by structure, function, and form (7). They can be preferable to antibodies for their excellent composition, endurance, and lack of denaturing (7). Their small size allows them to get inside tissue with greater efficiency. Aptamers are also produced through chemical processes, making them easy to make, multiply, and change (7).

Regarding bioinformatics, docking uses a computer to show the binding between a small binding molecule and a large protein to get a read on the binding conditions at the active site. First, it finds the protein's binding pocket, which it then uses to compute the algorithms of different molecule orientations. The computer then reads the various configurations and analyzes the probability of each, ranking them in order of most likely to least likely. Nucleolin protein, found in the nucleolus, is involved in many DNA and RNA functions (9). It is also between the shuttle between the cytoplasm and the nucleus (9). The protein is sent to the cell surface in many cancers as part of cell signaling. It has been suggested as a viable biomarker in detecting and targeting cancer.

We hypothesize that the aptamers used in this research will show specific binding interactions with the nucleolin protein, Figure 1. This innovative approach, explored in the current work, involves first predicting the tertiary

structures of DNA-based aptamers and then performing molecular docking simulations. The nucleolin-aptamer interactions are explained based on the surface electrostatics and shape complementary. These results were further validated by experimental work performed in the previous study. Finally, we have also performed aptamer mutations, out of which aptamer APT-M1 and APT-M2 has stronger interactions from the wt aptamers. The current results will help develop novel bacterial-based cancer-targeted therapy.

2. METHOD

The Protein Data Bank (PDB ID = 1FJE) was the source of the receptor's three-dimensional structures. The aptamer's main nucleotide sequence underpinned structural predictions. Vfold2D first determined this sequence's secondary structure (10). After obtaining the nucleotide sequence, it was converted to a secondary structure, and the images are in Figure 3. After predicting the secondary structure, the tertiary structure was modeled using Vfold3D (11). ChimeraX visualized the 3D structures that were obtained (12). These 3D structures were docked on nucleolin protein by the HDOCK software (13). The complete singular structure of DNA-bound nucleolin protein was obtained from AlphaFold 3, a deep-learning model that predicts the 3D models of the protein and DNA (14).

3. RESULTS

In this study, we successfully demonstrated the enhanced tumor localization of bacteria using aptamer-assisted targeting. We designed in silico aptamers using computational tools and then docked these structures onto bacterial proteins. Two sets of aptamers from the previous research have been taken in this research: AS1411 and ULS11a. Later, mutant forms of these aptamers AS1411_M1, AS1411_M2, and AS1411_M3 for AS1411 and ULS11a_M1 and ULS11a_M2 for ULS11a were developed. The DNA-based aptamers and their respective mutants () dot and bracket notations are shown in Table 1. Understanding the surface chemistry of the protein has dramatically increased the understanding of protein aptamer binding functions. Firstly, GrASP utilized binding site prediction using GNN analysis. Figure 2a shows the binding site. The binding site is displayed in yellow and green. However, in this figure, the aptamer cannot access it because the site is too small. Figures 2b and c show the receptor's ESP estimated with ChimeraX. It provides a look into the nature of protein interactions with other molecules. Positives are blue, negatives red, and neutrals white.

To understand the nucleolin-aptamer binding interactions, we had to perform molecular docking simulations, and to do that after 3d structure should be obtained. Therefore, we have performed these two steps to get the aptamer 3D structure. First, Vfold2D predicted aptamers' secondary structures, as seen in Figure 3 and Table 1. The 3D shape of the aptamer depends on its base pair interactions (hydrogen bonds) in its secondary structure. Nucleic acids, particularly aptamers, are often represented by dot-bracket notation. This notation indicates paired and unpaired bases using dots, brackets, and other symbols. Second, the 2D images also provide dot-bracket notations of the specific aptamers, which helps determine their 3D structure. Many researchers use these 2D images to predict the similarity/difference between aptamers based on the

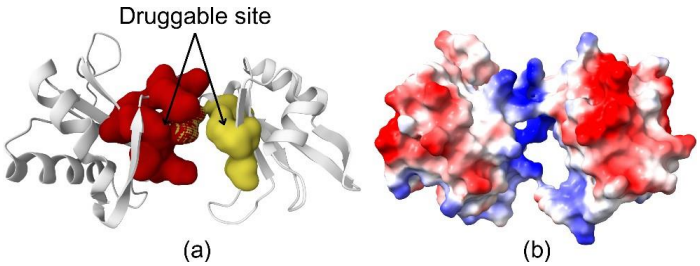

Fig. 2. Binding site of the protein.(a) The binding site on the nucleolin is shown in yellow-red color; (b) Nucleolin is both negative (red) and positive (blue) in nature, according to the protein's ESP. The central region is positive and should be the RNA binding site.

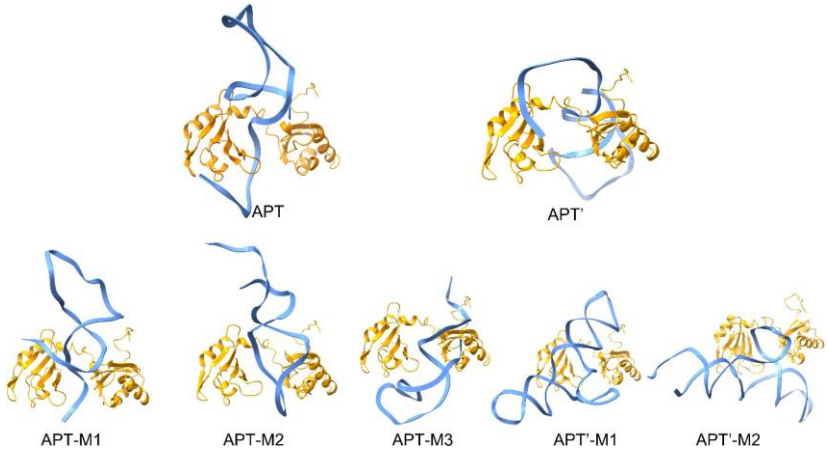

Fig. 3. Binding of aptamer toNucleolin . The aptamers were docked on the nucleolin to get the interactions, and the negative aptamers bind to the positively charged region of the protein.

hypothesis that similar 2D aptamers should have similar binding patterns. In this situation, all aptamer secondary structures have distinct binding patterns. The obtained 3D aptamers were docked on nucleolin protein in the next step. Figure 3 shows HADDOCK web server-predicted molecular docking structures. This study used HADDOCK, which has been thoroughly validated for protein-nucleic acid complexes. First, 10,000 rigid body dockings were performed with rigid proteins and aptamers. Four hundred semi-flexible refinement dockings added protein and aptamer flexibility in the second stage. Last, water was added to the docking procedure. The docking simulations showed that all

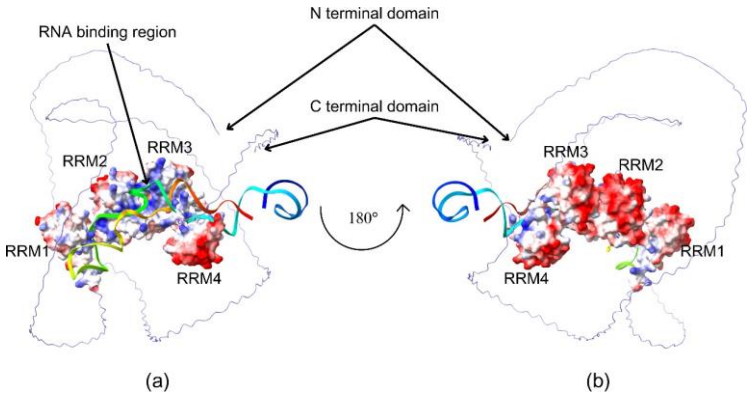

Fig. 4. Nucleolin -bound RNA. (a) Front view and (b) Back view. The negatively (red) charged RNA binds to the four domains' positive (blue) region.

aptamers bind to a distinct receptor area. To get the interactions of these aptamers, we have also calculated the nucleolin-protein dining interactions. Based on the PLIP web server, AS1411 formed ten hydrogen bonds with Asn49 (3.75), Gln91 (2.58 and 2.72), Asp92 (2.64), Thr98 (3.61 and 3.68), Asn102 (3.41), His107 (3.05), and Ser109 (3.04 and 3.05) and a salt bridge with Arg96 at a distance of 4.75 Å. AS1411 Mutation 1 formed five hydrogen bonds with Met31 (2.37), Met31 (2.68), Asn33 (3.13), Ser57 (2.69), and Asn58 (3.41). It also forms a hydrophobic interaction with Trp32 (3.48). In addition, it also formed a salt bridge with Glu61 (4.70), Arg104 (5.30), and Arg104 (5.5). AS1411 Mutation 2 forms four hydrogen bonds with Asn49 (3.17), Gln99 (2.6), His107 (3.18), and Ser109 (3.25). It formed two salt bridges with Arg96 (4.41) and Arg96 (4.56). Finally, to understand the strength of the interaction, we have also computed the nucleolin-aptamer binding energy, Table 1. This shows that the mutant aptamers (APT-M3 and APT'-M2) generated by us have more interactions than the native aptamers (APT and APT').

Finally, we have computed the complete 3D RNA-bound nucleolin structure using the AlphaFold 3 software, as shown in Figure 4. Due to nucleolin's complex and flexible nature, obtaining the entire 3D structure of nucleolin using experimental techniques is very difficult. Only the rigid regions of the protein have been elucidated and were used in this study. Based on this structure, in the absence of RNA, the four domains of the protein are arranged in a square shape; however, in the presence of RNA, the protein fluctuates, and to accommodate RNA, the four domains rearrange themselves in a linear shape. In addition, the negatively charged RNA is specifically attached to the positive charged region of all four domains. The positive area of the protein is blue, while the negative is red. The 3D structure of the protein will provide new insights into the RNA binding mechanism of this protein

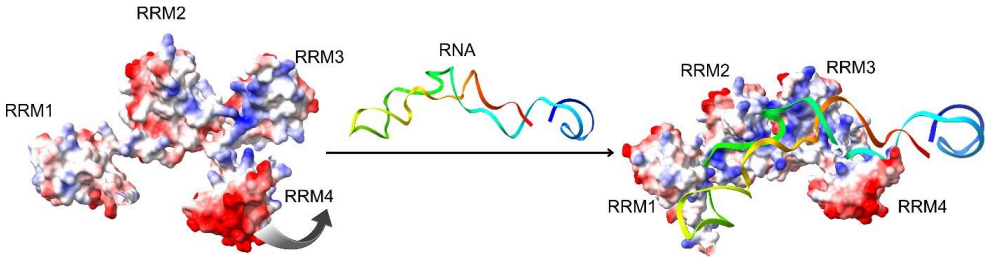

Fig. 5. Flexibility of Nucleolin protein. In the absence of RNA, the four domains are square, and in the presence of RNA, significant fluctuations were observed to accommodate the protein.

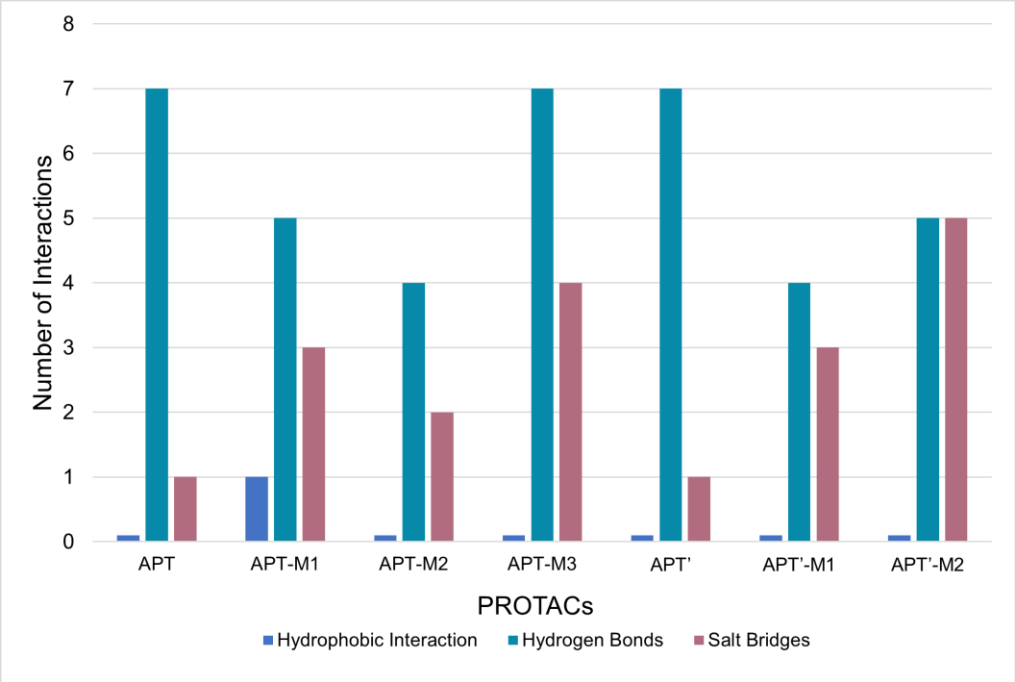

Fig. 6. Nucleolin-aptamer interactions. The graph shows that mutant aptamers (APT-M3 and APT'M2) form more interactions than the wt aptamers (APT and APT').

4. DISCUSSION

Bacterial biotherapy has gained significant attention in the field of cancer and challenges the traditional methods of diagnostic and therapeutic interventions. Tumor tropism can selectively target them and accumulate within the cancer cells (15). They can target tumors by modulating the immune system or de- livering therapeutic payload. Bioconjugation is a technique in which bacteria are used as a vehicle to transport drugs to a specific region of the body (16). Bacteria have unique properties that make them well-suited for drug transport, such as traversing various physiological barriers and selectively targeting spe- cific diseased cells (tumor cells). Nucleolin (NCL) is a multifunctional protein primarily located in the nucleolus (9, 17).

However, it can also be present on the cell surface in various cancer diseases. NCL on the plasma membrane acts as a coreceptor for attaching multiple single-stranded RNA, bacteria, and toxins (18). The protein is primarily present inside the cell and helps in ribosomal production, DNA repair, DNA transcription, and RNA splicing (19). In cancer cells, these proteins are also present on the cell's surface and can be used as cell surface markers for cancer detection and treatment.

Although NCL is a nuclear protein; however, it contains cancer cells that move to the surface of cancer cells and can be used as a biomarker of these cells (9). Using the alpha fold technique, we have elucidated the 3D structure of nucleolin bound with RNA (20). Since, NCL shares a complex and flexible nature getting the complete 3D structure is very difficult, and only a tiny segment of the protein has yet been elucidated. Computational-based 3D structure (elucidated by AlphaFold 2) has been reported in a review article by Tonello et al. in 2022 (18). We have not identified any RNA structures bound to the nucleolin protein to date. In this research work, we have elucidated, for the first time, the complex structure of RNA-bound nucleolin protein, Figure 4 and 5. Significant structural changes were observed in the free nucleolin protein vs RNA-bound nucleolin protein. In the free state, the domains are close to each other, forming a square; however, in the presence of the RNA, the four domains arrange themself in a linear form to accommodate the RNA. The protein's electrostatic surface potential (ESP) plays a significant role in RNA binding because only the protein's positive portion interacts with the negatively charged RNA. In nucleolin, all four domains bind to RNA, similar to other RNA-binding proteins that also engage multiple domains in RNA interaction.

This study shows how aptamer-assisted tumor localization can potentially improve the precision of aptamer-assisted bacteriotherapy. By crafting tumor-specific aptamers and by simulations of molecules, aptamer binding to the bacterial protein's process has been simulated. These findings show one of the significant drawbacks to bacterial-based therapies—non-specific distribution—can be overcome by aptamer targeting, thus providing a pathway for better and more precise treatment approaches. Improving aptamer-bacteria interactions and judging the therapeutic results in various tumor models to test this potentially helpful approach further should be done in the future.

Table 1. Aptamer used in this study and nucleolin-aptamer binding energy. The most muscular binding energy computed was of APT'M1 and APT'- M2. Dot and bracket notations are two common ways to represent the secondary structure of nucleic acids, such as RNA and DNA aptamers. Dots represent unpaired nucleotides (.). Matching parentheses represent paired nucleotides: Opening parenthesis (indicates the beginning of a base pair, and Closing parenthesis)indicates the end of a base pair.

Aptamer	**Sequence Dot and bracket**	Binding Energy (Kcal/mol)
APT	GGTGGTGGTGGTTGTGGTGGTGGTGGTTTTTTTTT TTT .....((((.....)))).	-9.27
APT-M1	GGTGGTGGTGGTTGTGGTGGTGGTGGTTTCTTTTT TTT .....((((.....)))).....	-9.26
APT-M2	GGTGGTGGTGGTGGTGGTGGTGGTGGTTTTTTTTT TTT .....((((.....)))).....	-9.32
APT-M3	GGTGGTGGTGGTTGGGGTGGTGGTGGTTTTTTTTT TTT .....((((((.....))))))...	-9.99
APT'	ACAGCATCCCCATGTGAACAATCGCATTGTGATTG TTACGGTTTCCGCCTCATGGACGTGCTGTTT (((((((.(((((((.((((((((((...)))))))))).((((....)))))))))).))))))..	-10.12
APT'-M1	ACAGCATCCCCATTTGAACAATCGCATTGTGATTG TTACGGTTTCCGCCTCATGGACGTGCTGTTT (((((((.(((((((.((((((((((...)))))))))).((((....)))))))))).))))))..	-10.97
APT'-M2	ACAGCATCCCCATGTGGACAATCGCATTGTGATTG TTACGGTTTCCGCCTCATGGACGTGCTGTTT (((((((.(((((((.((((((((((...)))))))))).((((....)))))))))).))))))..	-11.04

Table 2: Interactions between Nucleolin and aptamers. The Nucleolin-aptamer binding was due to the bonds formed between the two complexes.

APT		
Hydrogen Bonds	Asn49	3.75
	Gln91	2.58, 2.72
	Asp92	2.64
	Thr98	3.61, 3.68
	Asn102	3.41
	His107	3.05
	Ser109	3.04, 3.05
Salt Bridge	Arg96	4.75
APT-M1		
Hydrogen Bonds	MET	2.37
	MET	2.68
	ASN	3.13
	SER	2.69
	ASN	3.41
Salt Bridges	GLU	4.7
	ARG	5.3
	ARG	5.5
Hydrophobic interactions	TRP	3.48
APT-M2		
Hydrogen Bonds	ASN	3.17
	GLN	2.60
	HIS	3.18
	SER	3.25
Salt Bridges	ARG	4.41
	ARG	4.56
APT-M3		
Hydrogen Bonds	MET	2.7
	ASN	3.14
	ASN	2.59

	ASN	3.27
	LYS	2.58
	GLN	3.35
	LYS	3.35
Salt Bridges	ARG	4.48
	LYS	4.82
	LYS	4.02
	LYS	3.89
APT’		
Hydrogen Bonds	Trp32	2.42
	Thr36	2.69
	Ala40	2.78
	Val43	2.42
	Asn49	2.77, 3.04
	His107	2.46
	Arg147	2.60
Salt Bridge	Asp48	4.45
APT’-M1		
Hydrogen Bonds	PRO	3.39
	ARG	2.8
	GLN	3.24
	LEU	2.25
Salt Bridges	ARG	4.53
	LYS	4.79
	LYS	4.89
APT’-M2		
Hydrogen Bonds	GLY	3.05
	ASN	2.67
	SER	3.13
	THE	2.42

	LEU	2.27
Salt Bridges	ARG	5.17
	ARG	5.17
	ARG	5.47
	ARG	5.16
	ARG	5.47

REFERENCES

1. K. H. Gupta, C. Nowicki, E. F. Giurini, A. L. Marzo, and A. Zloza (2021).
2. Y. Zhang, "The role of bacteria and its derived biomaterials in cancer radiotherapy," Acta Pharmaceutica Sinica B **13**, 4149–4171 (2023).
3. C. Roffe, "Biotherapy for antibiotic-associated and other diarrhoeas," Journal of Infection **32**, 1–10 (1996).
4. Y. Yu, "Bacteria-driven bio-therapy: From fundamental studies to clinical trials," Nano Today **48**, 101731–101731 (2023).
5. A. J. Lax and W. Thomas, "How bacteria could cause cancer: one step at a time," Trends in Microbiology **10**, 293–299 (2002).
6. S. C. Nallar, D. Q. Xu, and D. V. Kalvakolanu, "Bacteria and genetically modified bacteria as cancer therapeutics: Current advances and challenges," Cytokine **89**, 160–172 (2017).
7. M. R. Dunn, R. M. Jimenez, and J. C. Chaput, "Analysis of aptamer discovery and technology," Nature Reviews Chemistry **1**, 76–76 (2017).
8. M. Darmostuk, S. Rimpelova, H. Gbelcova, and T, "Current approaches in SELEX: An update to aptamer selection technology," Biotechnology Advances **33**, 1141–1161 (2015).
9. H. Ginisty, H. Sicard, B. Roger, and P. Bouvet, "Structure and functions of nucleolin," Journal of Cell Science **112**, 761–772 (1999).
10. J. Li, S. Zhang, D. Zhang, and S. J. Chen, "Vfold-Pipeline: a web server for RNA 3D structure prediction from sequences," Bioinformatics **38**, 4042–4043 (2022).
11. X. Xu and S. J. Chen, "Rna scaffolds: Methods and protocols, l. ponchon," (Springer US, 2021), pp. 1–11.
12. E. F. Pettersen (2021).
13. Y. Yan, D. Zhang, P. Zhou, B. Li, and S. Y. Huang, "HDOCK: a web server for protein-protein and protein-DNA/RNA docking based on a hybrid strategy," Nucleic Acids Research **45**, 365–373 (2017).
14. M. Goedert, "Alpha-synuclein and neurodegenerative diseases," Nature Reviews Neuroscience **2**, 492–501 (2001).
15. J. M. Mathis, M. A. Stoff-Khalili, and D. T. Curiel, "Oncolytic adenoviruses - selective retargeting to tumor cells," Oncogene **24**, 7775–7791 (2005).
16. X. Zhao, "A rapid bioassay for single bacterial cell quantitation using bioconjugated nanoparticles," Proceedings of the National Academy of Sciences **101**, 15027–15032 (2004).
17. F. Mongelard and P. Bouvet, "Nucleolin: a multiFACeTed protein," Trends in Cell Biology **17**, 80–86 (2007).
18. F. Tonello, M. L. Massimino, and C. Peggion, "Nucleolin: a cell portal for viruses, bacteria, and toxins," Cellular and Molecular Life Sciences **79**, 271–271 (2022).
19. L. Shkreta and B. Chabot, "The RNA Splicing Response to DNA Damage," Biomolecules **5**, 2935–2977 (2015).
20. J. Abramson, "Accurate structure prediction of biomolecular interactions with AlphaFold 3," Nature **630**, 493–500 (2024).