Identification of new breast cancer-related biomarkers via integrated bioinformatics analysis and experimental validation

Inês Lameira[1], Débora Ferreira[2], and Sérgio Sousa[3]

[1] Informatics Department, University of Minho, 4710-057 Braga, Portugal
[2] Centre of Biological Engineering, University of Minho, 4710-057 Braga, Portugal
[3] LAQV/REQUIMTE, BioSIM – Faculty of Medicine, University of Porto, 4200-319 Porto, Portugal

Abstract. Triple-negative breast cancer (TNBC) is a highly aggressive subtype of breast cancer associated with limited therapeutic options and, consequently, poor clinical outcomes. Aptamers have emerged as promising tools in cancer research and targeted therapies due to their high specificity and affinity. This project aimed to structurally validate the Apt2 aptamer, previously identified through cell-SELEX for its binding affinity to the surface of MDA-MB-231 cells, and to characterize its protein targets, previously determined by proteomic studies in this TNBC cell line. Using a comprehensive computational pipeline, including structure prediction, molecular dynamics simulation, protein modeling, molecular docking and interaction analysis, the Apt2 structure was successfully predicted and refined, enabling subsequent docking studies with the modeled protein targets generated using AlphaFold. Some limitations emerged due to low-confidence predictions in certain regions and a limited number of interactions for some targets, suggesting the need for further optimization and additional computational approaches. Nevertheless, this study highlights the importance of integrating in silico methods with experimental data to elucidate aptamer-protein interactions and guide the rational design of Apt2 for potential applications in TNBC research.

Keywords: Aptamer · Molecular docking · Molecular Dynamics · Protein prediction · Triple-negative breast cancer

1 Introduction

1.1 Triple negative breast cancer (TNBC)

Breast cancer (BC) is the most diagnosed cancer and is considered the fifth leading cause of cancer-related deaths worldwide. It is estimated that 2.3 million cases and 685,000 deaths occurred in 2020 [42]. The projections for future trends are even more alarming, with estimates suggesting that BC cases could reach 4.4 million by 2070 [37]. The main risk factors associated with BC include older

age, obesity, exposure to tobacco, physical inactivity, high-fat dietary, early age menarche, late age at first full-term pregnancy, shorter breastfeeding periods, use of hormonal menopausal therapy or oral contraceptives, high breast density and BC family history [6].

BC is a highly heterogenous neoplasm with four subtypes classified based on the immunohistochemical expression of hormone receptors: estrogen receptor positive (ER+), progesterone receptor positive (PR+), human epidermal growth factor receptor positive (HER2+), and triple negative [35]. TNBC was first referred in the mid-2000s to identify the subset of BCs lacking ER, PR, and HER2 [14, 5]. TNBC accounts for approximately 20 % of all BCs and it is most common among women under 40 years old of age [24]. It is characterized by its aggressive nature and early relapse. In addition, TBNC is poorly differentiated, highly proliferative, and a heterogenous neoplasm [30].

To date, chemotherapy remains the standard of care for both early and advanced TBNC, usually involved combination regimens of taxanes, anthracyclines, cyclophosphamide, cisplatin and fluorouracil [32, 45]. Despite the high clinical response rates to neoadjuvant chemotherapy, TNBC patients show an elevated risk of recurrence and visceral metastasis, leading to a significantly higher mortality than any other BC subtypes [9]. The poor outcomes regarding TNBC treatments, demand the research for new therapy strategies, more specific and targeted to improve the current bad prognosis, and ultimately enhance patients' survival chance.

1.2 Aptamers as an emerging strategy in tumor-targeted therapies

The development of high-performance recognition tools for cancer is urgently needed to achieve early diagnosis and precise therapy [47]. In line with this, over the past decades, a varied array of molecular recognition tools has been broadly explored, namely antibodies, peptides, and nucleic acids [48, 36, 20]. An emerging strategy in tumor-targeted therapies is based on the use of specific ligands, such as aptamers, which are capable to bind a variety of targets including proteins, small molecules, viruses, bacteria and live cells with high affinity, specificity and selectivity [52, 40, 53].

The term "aptamer" was first reported by Ellington et al. [12], in 1990. Nucleic acid aptamers are single-stranded DNA/RNA oligonucleotides that bind to target cargos by folding into specific secondary/tertiary conformations [51]. Aptamers are derived from random oligonucleotide libraries through an in vitro iterative method so-called Systematic Evolution of Ligands by EXponential Enrichment (SELEX) [43, 12]. Importantly, the development of cell-SELEX technology has allowed the use of intact living cells as targets to select cell type-specific aptamers [52, 34]. A remarkable feature of cell-SELEX is that it allows the identification of specific aptamers for cells of interest without prior knowledge of the target molecules on cell surface or the available targetable biomarkers [31]. The aptamers developed by cell-SELEX have been widely used in cancer research, including for in vivo tumor imaging [49], as vehicles for the targeted delivery of miRNA/siRNA or anticarcinogens [54], and as anti-cancer agents [55].

1.3 Aptamers in TBNC therapies

Aptamers targeting BC cells have been successfully used in research for the development of novel therapeutic strategies, including for TNBC [2, 21, 23, 13]. For example, a study performed by Wan et al. [45] demonstrated that the synthetic aptamer PDGC21T, previously identified and used to recognize poorly differentiated gastric cancer tissue, was capable of binding to the surface of TNBC cells with high affinity. Camorani et al. [8] published another remarkable study in which they identified six aptamers capable of distinguishing TBNC from non-TNBC cells and non-malignant cells. In another work, Luo et al. [25] selected the aptamer 5TR1, which binds to the tumor biomarker MUC1 in MDA-MB-231 cells and conjugated it to doxorubicin to enhance specificity and mitigate known side effects such as cardiotoxicity . Similarly, He et al. [15] developed an aptamer-drug-conjugate with high specificity and cytotoxicity against the MDA-MB-231 cell line, demonstrating in vivo anti-TNBC efficacy with negligible side effects on healthy organs.

Among these promising advances in aptamer research for TNBC therapy, the focus of this project builds upon a previous study initiated by Ferreira et al. [13]. In that study, the authors identified novel aptamers targeting TNBC cells, identified through cell-SELEX with Apt2 outperforming the other identified candidates. Apt2 demonstrated high binding affinity to the surface of MDA-MB-231 cells, a highly metastatic TNBC cell line, highlighting its potential for targeted therapy and diagnostic applications in TNBC [13]. However, fully realize this potential, additional key studies are required. As cell-SELEX was employed, the specific target protein recognized by this aptamer remains unknown. This is precisely where the current project is positioned – leveraging computational approaches to identify and characterize Apt2's interaction with potential protein targets on TNBC cells, thus advancing its development toward clinical applications.

2 Objectives

This work focused on studying how the experimentally developed aptamer, Apt2, binds to TNBC cells. In particular, this project aimed to analyze potential targets identified in the lab and determine the most suitable candidates for binding to Apt2.To achieve this, computational methodologies - including structure prediction AI-models, docking, and molecular dynamics – were leveraged to complement and validate the experimental findings.

The specific aims for this project were:

- Determine 2D and 3D models of Apt2 using bioinformatics tools and molecular modeling techniques to ensure its stability and functionality. Secondary structures were predicted both with and without the primer zones used during its experimental generation;
- Study the 3D structure of target proteins expressed on the surface of TNBC cells to estimate their binding sites. Additionally, some modeling work was

performed to search to better understand the structural features of these target proteins;
– Characterize the molecular interactions between Apt2 and target proteins by analyzing the Apt2–protein complexes generated from molecular docking simulations to identify and classify the types of contacts involved.

3 Methodology

To accomplish the proposed specific aims, this work was divided into the following tasks:

3.1 Generation of Apt2 aptamer secondary structure

The first step in generating the three-dimensional structure for the aptamer involves predicting its secondary structure, following a well-established methodology previously described [18]. Initially, the DNA sequence of the aptamer was used as input into the Mfold web server, which is currently part of the UnaFold web server [56]. The secondary structure was predicted using free energy minimization algorithms. The sequence was treated as linear, with the folding temperature set at 37 ° C. The default ionic conditions included 187 mM Na^+ and 0.5 mM Mg^{2+}, at pH 7.4, with adjustments made for oligomers [13]. Only foldings within 5 % of the minimum free energy were calculated, with a maximum of 50 foldings computed. No constraints were applied regarding the maximum distance between paired bases. The resulting structure with the lowest free energy was utilized in subsequent steps. Additionally, the secondary structure was extracted in dot-bracket notation (Vienna format) for further analysis in the next task.

3.2 Generation of three-dimensional structures for Apt2

The resulting secondary structure was used as input for the 3dRNA/DNA tool to predict aptamer's three-dimensional structure, as described by [50]. This tool uses a template-based or a distance-geometry method to model three-dimensional conformations from the secondary structure. The model outputs ten structures, each scored through an internal scoring. The best scoring structure was considered and downloaded for the further analysis in subsequent steps.

3.3 Aptamer refinement by molecular dynamics

To prepare Apt2 for molecular docking simulations, its conformational flexibility was evaluated using the LEAP module of the AMBER20 software package [10]. Aptamer flexibility was studied under the effect of an explicit solvent by molecular dynamics simulations. In this case, the aptamer was solvated using a TIP3P rectangular box of water molecules [27], ensuring that each atom is at least 12 Angstrom from the edge. To achieve physiological concentration of

NaCl within the simulation, the SPLIT method was used [26], allowing precise calculation of the total number of Na^+ and Cl^- ions. The system was described using the BSC1 force field, developed for atomistic DNA simulations [17]. After solvation, a sequence of minimization steps of the system were undertaken to remove atomic clashing during molecular dynamics: (1) water molecules; (2) hydrogen atoms; (3) side chains of the DNA structure; and (4) the complete system. Following a step of system equilibration, the final stage was the production run using an NPT ensemble. To constrain all covalent bonds involving hydrogen atoms, the SHAKE algorithm was employed [28].

3.4 Target proteins data integration

A previous proteomic study identified six potential target proteins of interest to which Apt2 binds: P09493, P13984, Q14141, Q9BTC0, Q9NSK0, and Q8N8S7 (primary protein accession IDs). Before generating the three-dimensional structure of these targets in the next task, preliminary modeling work was conducted. This step aimed to perform a data retrieval process from databases such as UniProt, NCBI Protein, and PDB, integrating information on important features and key aspects that may influence the modeling process itself.

3.5 Three-dimensional protein structure prediction

To predict the three-dimensional structure for each protein, AlphaFold3, a state-of-the-art deep learning-based method for protein structure prediction developed by DeepMind [1], was used. The amino acid sequences of the selected proteins served as input for the AlphaFold prediction server. By leveraging AlphaFold's advanced neural network architecture and extensive training data, highly accurate 3D structural models of the proteins were generated.

3.6 Prediction of Apt2 binding site by docking

To identify representative aptamer structures from the molecular dynamics' simulations suitable for molecular docking with the target proteins, cluster analysis were conducted. The K-means clustering method was used, and clusters were generated based on the root mean square deviation (RMSD) of all non-hydrogen atoms. The representative structure of each cluster were selected for further analysis. Next, the complexes between proteins and Apt2 were predicted using the HADDOCK 2.4 web server [16]. Ambiguous interaction restraints (AIRs) were randomly determined from residues. The ranking of the different complexes for each protein were determined based on the HADDOCK score (HS), which is calculated as a linear weighted sum of energetic terms and buried surface area:

$$HS = 1.0Evdw + 0.2Eelec + 1.0Edesol + 0.1Eair \quad (1)$$

Each term in Equation 1 represents van der Waals (Evdw), Coulomb electrostatics (Eelec), desolvation (Edesol), and restraints energy (Eair), respectively.

3.7 Characterization of docked Apt2–Protein Complexes using RING

The most likely Apt2–protein complexes, as suggested by molecular docking studies, were further analyzed using the RING tool to characterize intermolecular contacts. RING identifies and classifies non-covalent interactions such as hydrogen bonds, van der Waals forces, π–π stacking, salt bridges, and hydrophobic contacts within macromolecular complexes [11]. This allowed the identification of key contact residues at the binding interface.

4 Results

4.1 Secondary structure prediction of Apt2

Apt2-core. The aptamer, Apt2, is constituted by a sequence of 49 nucleotides with the following composition:

```
5‘- ATGTTGTTGCCGGGACGCCTCCTTCACCAAAGTTGGTGTCCCCACCTAC -3’
```

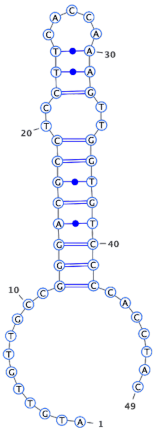
Apt2's secondary structure was predicted using the Mfold web server, which applies thermodynamic models to nucleic acid folding. The resulting structure displays a Gibbs free energy (ΔG) of -5.98 kcal/mol, suggesting a relatively stable conformation under the given conditions. The prediction, illustrated in figure 1a, reveals the presence of a single prominent stem-loop structure. The stem region contributes to the overall thermodynamic stability, while the loop region introduces flexibility that may facilitate molecular interactions [46]. This conformation provides a foundation for modeling the tertiary structure and evaluating potential binding interfaces through molecular docking.

Apt2-primers. To evaluate the impact of primer regions on the folding of Apt2, the secondary structure of the full-length aptamer, including the constant primer binding regions used during the cell-SELEX process described by Ferreira et al. (2021), was also predicted using the Mfold web server. The nucleotide sequence including primer binding regions is following described:

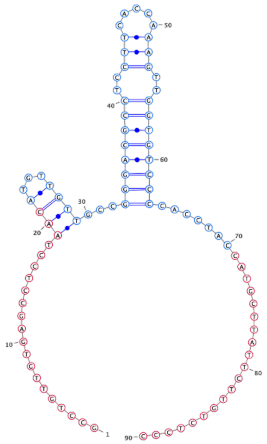
```
5‘- GCCTGTTGTGAGCCTCCTAAC-nt49-CATGCTTATTCTTGTCTCCC -3’
```

The folding simulation was performed under the same previous conditions resulting in a structure with a minimum free energy of -6.87 kcal/mol, indicating a sightly more stable conformation that the core sequence alone (which exhibited a ΔG = -5.98 kcal/mol). The predicted secondary structure is shown in figure 1b.

The inclusion of the primer sequences resulted in an additional stem region, potentially influencing the global conformation of the aptamer. Nevertheless, the core stem-loop motif remained intact and structurally conserved. Therefore,



(a) Core-Apt2 sequence (49 nt); predicted ΔG = –5.98 kcal/mol.



(b) Full-length Apt2 including primer regions (colored in red); predicted ΔG = –6.87 kcal/mol.

Fig. 1: Comparison of Apt2 secondary structure predictions with and without the flanking primer regions. Structures were predicted using the Mfold web server under physiological ionic conditions and visualized with VARNA (v3.93).

three-dimensional structure prediction and docking simulations, were based only on the 49-nucleotide randomized core of Apt2. This approach ensures that the modeling focuses exclusively on the functional region of the aptamer. So, the flanking primer regions were excluded from subsequent steps of the project, since they could interfere with the modeling outcomes.

4.2 Tertiary structure prediction and refinement of Apt2

The three-dimensional structure of Apt2 was predicted using the 3dRNA/DNA server, based on the secondary structure previously obtained for the Apt2-core. The predicted structure having the best score associated was chosen, and the structure was subsequently visualized to assess the spatial arrangement and folding patterns of the aptamer. As shown in figure 2a, the tertiary structure preserves the central stem-loop identified in the secondary structure prediction. The stem region forms a stable helical segment, while the loop adopts a flexible and exposed configuration that extends outward from the helical axis. This three-dimensional model was refined and employed in docking simulations with target proteins found in TNBC.

The structural refinement of the Apt2 aptamer was achieved by molecular dynamics (MD), using the predicted three-dimensional structure as the initial model. After the simulation, the final representative structure was extracted based on its prevalence throughout the MD simulation and visualized in PyMOL. The final conformation for Apt2 is shown figure 2b. The overall folding of the

aptamer remained stable throughout the simulation. The stem region preserved its helical structure, indicating strong base-pairing and structural rigidity. The loop region exhibited slight conformational rearrangement, which are expected due to its inherent flexibility. No major unfolding or loss of global structure was observed, suggesting that the aptamer maintains a compact and viable conformation under simulated physiological conditions.

(a) Initial tridimensional model of Apt2 before MD simulation.

(b) Representative conformation of Apt2 after MD simulation.

Fig. 2: Comparison of the Apt2 aptamer structure before and after MD simulation. Structures were visualized in PyMOL. The stem-loop architecture is preserved, with the stem maintaining helical integrity and the loop showing slight conformational adjustments after simulation.

4.3 Structural modeling of candidate target proteins

The structural models of the six target proteins were generated using AlphaFold, guided by curated annotations from protein databases such as UniProt, as well as literature-supported evidence regarding their physiological oligomeric state and functional conformation. The selected oligomeric assemblies correspond to forms reported to be biologically relevant or experimentally supported for each target. The resulting models are illustrated in figure 3.

Overall, the models predicted reproduced key structural features such as stable core domains and disordered or flexible regions, reflecting the biological diversity and dynamic behavior of these selected targets. Some models (notably the T2FB/T2FA heterodimer and DIDO1) exhibited extended low-confidence segments — commonly referred to as "spaghetti" regions — which appear as unstructured loops due to their low predicted pLDDT scores (typically <50). These regions often correspond to intrinsically disordered or highly flexible parts of the protein, lacking a well-defined structure and their presence reduces the structural reliability of those segments [33, 44]. Therefore, any specific conformation predicted in these "spaghetti" loops should be carefully interpreted, as it may not reflect a unique stable structure but rather an artifact of the prediction process [19]. In a recent study by Bergonzo et al. (2025) involving RNA and DNA structure predictions using AlphaFold, flexible regions, noncanonical interactions, and loop conformations showed inconsistencies when compared to

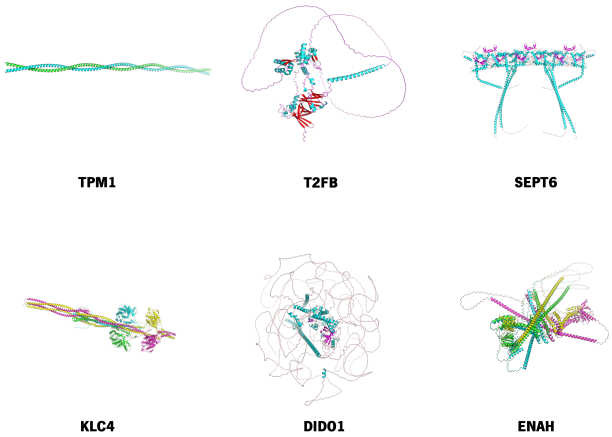


Fig. 3: Structural models generated using AlphaFold based on UniProt sequence and annotation data. From left to right, top to bottom: TPM1 (homodimer), T2FB (heterodimer with T2FA), SEPT6 (hexameric complex with SEPT2 and SEPT7), KLC4 (heterotetramer with two kinesin heavy chains, KIF5B, and two light chains, KLC4), DIDO1 (monomer), and ENAH (homotetramer). The models illustrate a diversity of structural features, including elongated coiled-coils, globular domains and multimeric assemblies.

experimental data [4]. Importantly, although AlphaFold has demonstrated outstanding performance in predicting protein structures, its accuracy significantly decreases when the sequences to be modeled are greatly different from those used in its training data. This limitation arises from its deep learning foundation, which relies on recognizing pattern from evolutionarily related proteins [7].

4.4 Molecular interaction analysis between Apt2 and target proteins

Finally, the recognition interfaces between Apt2 and its experimentally determined targets were investigated. Interaction profiling for each target protein was studied with Ring 3.0 to identify contact residues and interaction types (e.g., hydrogen bonds, π–π stacking, Van der Waals). A summary of direct aptamer–protein contacts is presented in Table 1. Representative docking poses of Apt2 with the six target proteins are shown in figure 4, highlighting the structural diversity of the complexes and the predicted binding sites. The docking results for the targets composed by KLC4, TPM1 and ENAH exhibit a relatively low number of interactions with the aptamer. This observation requires particular attention: while these proteins are large and possess extended or multimeric structures, the aptamer is of medium size [29], and limited conformational reach.

Table 1: Summary of interactions between Apt2 and each predicted protein target. For each complex, residues of the aptamer and protein involved in direct contacts are listed, along with the types of non-covalent interactions identified using RING. Abbreviations: HBOND - hydrogen bond; VDW - Van der Waals interaction; PIPISTACK – π–π stacking interaction.

Target	**Target Residue(s)**	**Interaction Type(s)**	**Aptamer Residue(s)**
TPM1	LYS 140	VDW	DC 41
	GLN 144	HBOND	DC 42
	GLU 150	VDW	DC 16
T2FB	PHE 266	VDW	DG 14
	TYR 273	HBOND, PIPISTACK	DA 48, DC 49
	SER 280	VDW	DC 10
	TYR 683	VDW, VDW	DG 13, DG 14
	GLU 684	VDW	DG 14
	LYS 687	HBOND, VDW	DG 14, DG 14
SEPT6	GLN 1800	VDW	DC 21
	GLN 1804	VDW	DC 21
	LYS 1807	VDW, VDW	DT 20, DG 36
	HIS 2117	HBOND, VDW	DC 22, DG 36
KLC4	LEU 175	VDW	DA 1
	LEU 178	VDW	DA 1
	GLY 203	VDW	DC 18
DIDO1	PHE 1683	PIPISTACK, VDW	DG 6, DT 7
	THR 1684	VDW	DG 6
	PRO 1686	HBOND, VDW	DG 3
	CYS 1685	HBOND	DT 4
	HIS 1815	HBOND	DT 34
	ILE 1786	VDW	DT 47
ENAH	PRO 939	VDW	DA 31
	PRO 1504	VDW	DC 49
	PRO 1507	VDW	DC 49

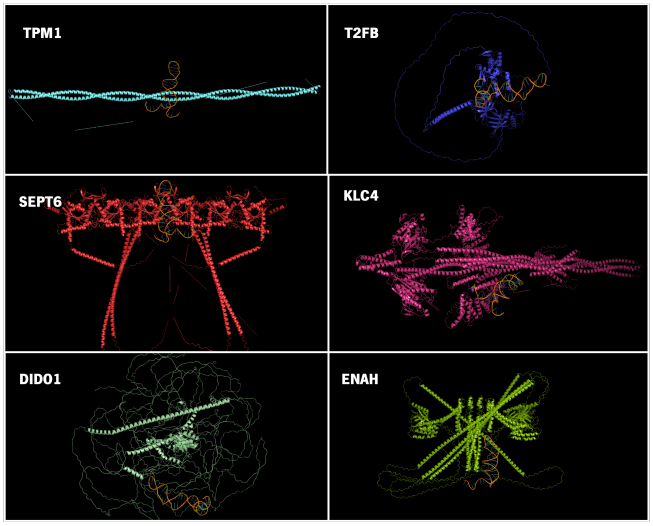


Fig. 4: Representative docking poses of the Apt2 aptamer (colored in orange) with the six predicted protein targets. The complexes target-Apt2 were visualized in PyMOL. The panel includes: (top row) TPM1 and T2FB; (middle row) SEPT6 and KLC4; (bottom row) DIDO1 and ENAH. The conformations illustrate the predicted binding orientations of Apt2 across different protein surfaces.

This results in the binding interface that may be insufficient to establish multiple strong contacts, especially in the absence of a clearly defined binding pocket or surface topology favoring aptamer anchoring [41, 38]. This observation highlights a potential limitation in the targeting of large multidomain proteins with the aptamer, which may require further optimization or engineering of aptamer loops or extensions to improve binding affinity and specificity [3].

Apt2 formed chemically diverse interactions with its predicted targets, although the number and strength of contacts varied across complexes. However, docking onto low-confidence or intrinsically disordered regions, such as in T2FB and DIDO1, must be interpreted carefully, when interactions are predicted in segments lacking ordered secondary structure. Such "spaghetti" regions correspond to biologically flexible segments that are often functionally relevant in regulatory proteins [33, 22] but inherently challenging to model with high accuracy [44, 19]. Among the non-covalent forces detected, hydrogen bonds and π–π stacking stand out for their directionality and higher specificity compared to weaker Van der Waals contacts [39]. Their presence suggests more stable or functionally meaningful interactions, particularly when observed in structured regions of the protein surface.

5 Conclusion and Future Perspectives

The primary goal of this project - structurally validating the interaction between the Apt2 aptamer and experimentally identified targets on the surface of MDA-MB-231 cells - was achieved, supporting and complementing previous proteomic data. However, some limitations emerged, and the conclusions must be interpreted with caution. First, the limited structural reliability of some AlphaFold models, particularly in disordered regions, introduces uncertainty in the predicted docking conformations. Additionally, the modest number of interactions observed in certain complexes may reflect intrinsic structural constraints and suggests that further aptamer optimization could enhance target engagement. These challenges highlight the importance of integrating experimental validation and complementary modeling approaches to overcome the constraints of purely computational predictions. Future work should include biochemical binding assays or energy-based docking refinements to strengthen and deepen the understanding of Apt2's potential as a multi-target tool in TNBC.

References

1. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., Bodenstein, S.W., Evans, D.A., Hung, C.C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., Beattie, C., Bertolli, O., Bridgland, A., Cherepanov, A., Congreve, M., Cowen-Rivers, A.I., Cowie, A., Figurnov, M., Fuchs, F.B., Gladman, H., Jain, R., Khan, Y.A., Low, C.M., Perlin, K., Potapenko, A., Savy, P., Singh, S., Stecula, A., Thillaisundaram, A., Tong, C., Yakneen, S., Zhong, E.D., Zielinski, M., Žídek, A., Bapst, V., Kohli, P., Jaderberg, M., Hassabis, D., Jumper, J.M.: Accurate structure prediction of biomolecular interactions with alphafold 3. Nature 2024 630:8016 **630**, 493–500 (5 2024). https://doi.org/10.1038/s41586-024-07487-w
2. Agnello, L., d'Argenio, A., Nilo, R., Fedele, M., Camorani, S., Cerchia, L.: Aptamer-based strategies to boost immunotherapy in tnbc. Cancers **15**, 2010 (4 2023). https://doi.org/10.3390/cancers15072010
3. Alkhamis, O., Byrd, C., Canoura, J., Bacon, A., Hill, R., Xiao, Y.: Exploring the relationship between aptamer binding thermodynamics, affinity, and specificity. Nucleic Acids Research **53**, gkaf219 (3 2025). https://doi.org/10.1093/nar/gkaf219
4. Bergonzo, C., Grishaev, A.: Critical assessment of rna and dna structure predictions via artificial intelligence: The imitation game. Journal of Chemical Information and Modeling (4 2025). https://doi.org/10.1021/acs.jcim.5c00245
5. Brenton, J.D., Carey, L.A., Ahmed, A., Caldas, C.: Molecular classification and molecular forecasting of breast cancer: ready for clinical application? Journal of clinical oncology : official journal of the American Society of Clinical Oncology **23**, 7350–7360 (2005). https://doi.org/10.1200/JCO.2005.03.3845
6. Britt, K.L., Cuzick, J., Phillips, K.A.: Key steps for effective breast cancer prevention (8 2020). https://doi.org/10.1038/s41568-020-0266-x
7. Callaway, E.: Deepmind's ai predicts structures for a vast trove of proteins. Nature **595**, 635–635 (7 2021). https://doi.org/10.1038/d41586-021-02025-4
8. Camorani, S., Granata, I., Collina, F., Leonetti, F., Cantile, M., Botti, G., Fedele, M., Guarracino, M.R., Cerchia, L.: Novel aptamers selected on living

cells for specific recognition of triple-negative breast cancer. iScience **23** (4 2020). https://doi.org/10.1016/j.isci.2020.100979

9. Carey, L.A., Dees, E.C., Sawyer, L., Gatti, L., Moore, D.T., Collichio, F., Ollila, D.W., Sartor, C.I., Graham, M.L., Perou, C.M.: The triple negative paradox: primary tumor chemosensitivity of breast cancer subtypes. Clinical cancer research : an official journal of the American Association for Cancer Research **13**, 2329–2334 (4 2007). https://doi.org/10.1158/1078-0432.CCR-06-1109
10. Case, D.A., Aktulga, H.M., Belfon, K., Cerutti, D.S., Cisneros, G.A., Cruzeiro, V.W.D., Forouzesh, N., Giese, T.J., Götz, A.W., Gohlke, H., Izadi, S., Kasavajhala, K., Kaymak, M.C., King, E., Kurtzman, T., Lee, T.S., Li, P., Liu, J., Luchko, T., Luo, R., Manathunga, M., Machado, M.R., Nguyen, H.M., O'Hearn, K.A., Onufriev, A.V., Pan, F., Pantano, S., Qi, R., Rahnamoun, A., Risheh, A., Schott-Verdugo, S., Shajan, A., Swails, J., Wang, J., Wei, H., Wu, X., Wu, Y., Zhang, S., Zhao, S., Zhu, Q., Cheatham, T.E., Roe, D.R., Roitberg, A., Simmerling, C., York, D.M., Nagan, M.C., Merz, K.M.: Ambertools. Journal of Chemical Information and Modeling **63**, 6183–6191 (10 2023). https://doi.org/10.1021/acs.jcim.3c01153
11. Del Conte, A., Camagni, G.F., Clementel, D., Minervini, G., Monzon, A.M., Ferrari, C., Piovesan, D., Tosatto, S.E.: Ring 4.0: faster residue interaction networks with novel interaction types across over 35,000 different chemical structures. Nucleic Acids Research **52**, W306–W312 (7 2024). https://doi.org/10.1093/nar/gkae337, https://academic.oup.com/nar/article/52/W1/W306/7660079
12. Ellington, A.D., Szostak, J.W.: In vitro selection of rna molecules that bind specific ligands. Nature **346**, 818–822 (1990). https://doi.org/10.1038/346818A0
13. Ferreira, D., Barbosa, J., Sousa, D.A., Silva, C., Melo, L.D., Avci-Adali, M., Wendel, H.P., Rodrigues, L.R.: Selection of aptamers against triple negative breast cancer cells using high throughput sequencing. Scientific Reports **11** (12 2021). https://doi.org/10.1038/s41598-021-87998-y
14. Foulkes, W.D., Smith, I.E., Reis-Filho, J.S.: Triple-negative breast cancer. The New England journal of medicine **363**, 1938–1948 (11 2010). https://doi.org/10.1056/NEJMra1001389
15. He, J., Peng, T., Peng, Y., Ai, L., Deng, Z., Wang, X.Q., Tan, W.: Molecularly engineering triptolide with aptamers for high specificity and cytotoxicity for triple-negative breast cancer. Journal of the American Chemical Society **142**, 2699–2703 (2 2020). https://doi.org/10.1021/jacs.9b10510
16. Honorato, R.V., Trellet, M.E., Jiménez-García, B., Schaarschmidt, J.J., Giulini, M., Reys, V., Koukos, P.I., Rodrigues, J.P., Karaca, E., van Zundert, G.C., Roel-Touris, J., van Noort, C.W., Jandová, Z., Melquiond, A.S., Bonvin, A.M.: The haddock2.4 web server for integrative modeling of biomolecular complexes. Nature protocols **19** (11 2024). https://doi.org/10.1038/S41596-024-01011-0
17. Ivani, I., Dans, P.D., Noy, A., Pérez, A., Faustino, I., Hospital, A., Walther, J., Andrio, P., Goñi, R., Balaceanu, A., Portella, G., Battistini, F., Gelpí, J.L., González, C., Vendruscolo, M., Laughton, C.A., Harris, S.A., Case, D.A., Orozco, M.: Parmbsc1: a refined force field for dna simulations. Nature Methods 2016 13:1 **13**, 55–58 (11 2015). https://doi.org/10.1038/nmeth.3658
18. Jeddi, I., Saiz, L.: Three-dimensional modeling of single stranded dna hairpins for aptamer-based biosensors. Scientific Reports 2017 7:1 **7**, 1–13 (4 2017). https://doi.org/10.1038/s41598-017-01348-5
19. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C.,

Kohl, S.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., Hassabis, D.: Highly accurate protein structure prediction with alphafold. Nature 2021 596:7873 **596**, 583–589 (7 2021). https://doi.org/10.1038/s41586-021-03819-2

20. Lee, S., Xie, J., Chen, X.: Peptide-based probes for targeted molecular imaging. Biochemistry **49**, 1364–1376 (2 2010). https://doi.org/10.1021/bi901135x
21. Li, X., Zhang, W., Liu, L., Zhu, Z., Ouyang, G., An, Y., Zhao, C., Yang, C.J.: In vitro selection of dna aptamers for metastatic breast cancer cell recognition and tissue imaging. Analytical Chemistry **86**, 6596–6603 (7 2014). https://doi.org/10.1021/ac501205q
22. Liu, J., Perumal, N.B., Oldfield, C.J., Su, E.W., Uversky, V.N., Dunker, A.K.: Intrinsic disorder in transcription factors. Biochemistry **45**, 6873–6888 (6 2006). https://doi.org/10.1021/bi0602718
23. Liu, M., Yang, T., Chen, Z., Wang, Z., He, N.: Differentiating breast cancer molecular subtypes using a dna aptamer selected against mcf-7 cells. Biomaterials Science **6**, 3152–3159 (11 2018). https://doi.org/10.1039/C8BM00787J
24. Loibl, S., Gianni, L.: Her2-positive breast cancer. Lancet (London, England) **389**, 2415–2429 (6 2017). https://doi.org/10.1016/S0140-6736(16)32417-5
25. Luo, S., Wang, S., Luo, N., Chen, F., Hu, C., Zhang, K.: The application of aptamer 5tr1 in triple negative breast cancer target therapy. Journal of cellular biochemistry **119**, 896–908 (1 2018). https://doi.org/10.1002/JCB.26254
26. Machado, M.R., Pantano, S.: Split the charge difference in two! a rule of thumb for adding proper amounts of ions in md simulations. Journal of Chemical Theory and Computation **16**, 1367–1372 (3 2020). https://doi.org/10.1021/acs.jctc.9b00953
27. Mark, P., Nilsson, L.: Structure and dynamics of the tip3p, spc, and spc/e water models at 298 k. Journal of Physical Chemistry A **105**, 9954–9960 (11 2001). https://doi.org/10.1021/jp003020w
28. Miyamoto, S., Kollman, P.A.: Settle: An analytical version of the shake and rattle algorithm for rigid water models. Journal of Computational Chemistry **13**, 952–962 (10 1992). https://doi.org/10.1002/jcc.540130805
29. Ni, S., Zhuo, Z., Pan, Y., Yu, Y., Li, F., Liu, J., Wang, L., Wu, X., Li, D., Wan, Y., Zhang, L., Yang, Z., Zhang, B.T., Lu, A., Zhang, G.: Recent progress in aptamer discoveries and modifications for therapeutic applications. ACS Applied Materials & Interfaces **13**, 9500–9519 (3 2021). https://doi.org/10.1021/acsami.0c05750, https://pubs.acs.org/doi/10.1021/acsami.0c05750
30. Orrantia-Borunda, E., Anchondo-Nuñez, P., Acuña-Aguilar, L.E., Gómez-Valles, F.O., Ramírez-Valdespino, C.A.: Subtypes of breast cancer. Breast Cancer pp. 31–42 (8 2022). https://doi.org/10.36255/exon-publications-breast-cancer-subtypes
31. Pang, X., Cui, C., Wan, S., Jiang, Y., Zhang, L., Xia, L., Li, L., Li, X., Tan, W.: Bioapplications of cell-selex-generated aptamers in cancer diagnostics, therapeutics, theranostics and biomarker discovery: A comprehensive review. Cancers **10**, 47 (2 2018). https://doi.org/10.3390/cancers10020047
32. Park, J.H., Ahn, J.H., Kim, S.B.: How shall we treat early triple-negative breast cancer (tnbc): From the current standard to upcoming immuno-molecular strategies. ESMO Open **3**, e000357 (1 2018). https://doi.org/10.1136/esmoopen-2018-000357
33. Piovesan, D., Monzon, A.M., Tosatto, S.C.E.: Intrinsic protein disorder and conditional folding in alphafolddb. Protein Science **31** (11 2022). https://doi.org/10.1002/pro.4466

34. Sefah, K., Shangguan, D., Xiong, X., O'Donoghue, M.B., Tan, W.: Development of dna aptamers using cell-selex. Nature Protocols **5**, 1169–1185 (6 2010). https://doi.org/10.1038/nprot.2010.66
35. Shaath, H., Elango, R., Alajez, N.M.: Molecular classification of breast cancer utilizing long non-coding rna (lncrna) transcriptomes identifies novel diagnostic lncrna panel for triple-negative breast cancer. Cancers **13**, 5350 (11 2021). https://doi.org/10.3390/cancers13215350
36. Singh, H., Tiwari, K., Tiwari, R., Pramanik, S.K., Das, A.: Small molecule as fluorescent probes for monitoring intracellular enzymatic transformations. Chemical Reviews **119**, 11718–11760 (11 2019). https://doi.org/10.1021/acs.chemrev.9b00379
37. Soerjomataram, I., Bray, F.: Planning for tomorrow: global cancer incidence and the role of prevention 2020-2070. Nature reviews. Clinical oncology **18**, 663–672 (10 2021). https://doi.org/10.1038/s41571-021-00514-z
38. Song, K.M., Lee, S., Ban, C.: Aptamers and their biological applications. Sensors **12**, 612–631 (1 2012). https://doi.org/10.3390/s120100612
39. Stefan Loverix, Jan Steyaert, P.G.P.M.: Influence of the - interaction on the hydrogen bonding capacity of stacked dna/rna bases. Nucleic Acids Research **33**, 1779–1789 (4 2005). https://doi.org/10.1093/nar/gki317
40. Stoltenburg, R., Reinemann, C., Strehlitz, B.: Selex-a (r)evolutionary method to generate high-affinity nucleic acid ligands. Biomolecular Engineering **24**, 381–403 (10 2007). https://doi.org/10.1016/j.bioeng.2007.06.001
41. Sun, H., Zu, Y.: A highlight of recent advances in aptamer technology and its application. Molecules **20**, 11959–11980 (6 2015). https://doi.org/10.3390/molecules200711959
42. Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A., Bray, F.: Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians **71**, 209–249 (5 2021). https://doi.org/10.3322/caac.21660
43. Tuerk, C., Gold, L.: Systematic evolution of ligands by exponential enrichment: Rna ligands to bacteriophage t4 dna polymerase. Science **249**, 505–510 (8 1990). https://doi.org/10.1126/science.2200121
44. Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Žídek, A., Bridgland, A., Cowie, A., Meyer, C., Laydon, A., Velankar, S., Kleywegt, G.J., Bateman, A., Evans, R., Pritzel, A., Figurnov, M., Ronneberger, O., Bates, R., Kohl, S.A., Potapenko, A., Ballard, A.J., Romera-Paredes, B., Nikolov, S., Jain, R., Clancy, E., Reiman, D., Petersen, S., Senior, A.W., Kavukcuoglu, K., Birney, E., Kohli, P., Jumper, J., Hassabis, D.: Highly accurate protein structure prediction for the human proteome. Nature 2021 596:7873 **596**, 590–596 (7 2021). https://doi.org/10.1038/s41586-021-03828-1
45. Wan, Q., Zeng, Z., Qi, J., Zhao, Y., Liu, X., Chen, Z., Zhou, H., Zu, Y.: Aptamer targets triple-negative breast cancer through specific binding to surface cd49c. Cancers **14**, 1570 (3 2022). https://doi.org/10.3390/cancers14061570
46. Wang, K., Wang, M., Ma, T., Li, W., Zhang, H.: Review on the selection of aptamers and application in paper-based sensors. Biosensors **13**, 39 (12 2022). https://doi.org/10.3390/bios13010039
47. Wu, L., Zhang, Y., Wang, Z., Zhang, Y., Zou, J., Qiu, L.: Aptamer-based cancer cell analysis and treatment (10 2022). https://doi.org/10.1002/open.202200141
48. Yang, L., Mao, H., Wang, Y.A., Cao, Z., Peng, X., Wang, X., Duan, H., Ni, C., Yuan, Q., Adams, G., Smith, M.Q., Wood, W.C., Gao, X., Nie, S.:

Single chain epidermal growth factor receptor antibody conjugated nanoparticles for in vivo tumor targeting and imaging. Small **5**, 235–243 (1 2009). https://doi.org/10.1002/smll.200800714
49. Yuan, B., Jiang, X., Chen, Y., Guo, Q., Wang, K., Meng, X., Huang, Z., Wen, X.: Metastatic cancer cell and tissue-specific fluorescence imaging using a new dna aptamer developed by cell-selex. Talanta **170**, 56–62 (8 2017). https://doi.org/10.1016/j.talanta.2017.03.094
50. Zhang, Y., Xiong, Y., Xiao, Y.: 3ddna: A computational method of building dna 3d structures. Molecules **27**, 5936 (9 2022). https://doi.org/10.3390/molecules27185936
51. Zhang, Y., Du, Y., Zhuo, Y., Qiu, L.: Functional nucleic acid-based live-cell fluorescence imaging. Frontiers in chemistry **8** (12 2020). https://doi.org/10.3389/fchem.2020.598013
52. Zhong, Y., Zhao, J., Li, J., Liao, X., Chen, F.: Advances of aptamers screened by cell-selex in selection procedure, cancer diagnostics and therapeutics. Analytical Biochemistry **598** (6 2020). https://doi.org/10.1016/j.ab.2020.113620
53. Zhou, J., Rossi, J.: Aptamers as targeted therapeutics: Current potential and challenges. Nature Reviews Drug Discovery **16**, 181–202 (3 2017). https://doi.org/10.1038/nrd.2016.199
54. Zhou, J., Rossi, J.J.: Cell-specific aptamer-mediated targeted drug delivery. http://www.liebertpub.com/oli **21**, 1–10 (2 2011). https://doi.org/10.1089/oli.2010.0264
55. Zueva, E., Rubio, L.I., Ducongé, F., Tavitian, B.: Metastasis-focused cell-based selex generates aptamers inhibiting cell migration and invasion. International Journal of Cancer **128**, 797–804 (2 2011). https://doi.org/10.1002/ijc.25401
56. Zuker, M.: Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Research **31**, 3406–3415 (7 2003). https://doi.org/10.1093/nar/gkg595