

Blocker-SELEX: a structure-guided strategy for developing inhibitory aptamers disrupting undruggable transcription factor interactions

Received: 11 January 2024

Accepted: 31 July 2024

Published online: 08 August 2024



Tongqing Li^{1,2,6}, Xueying Liu^{1,6}, Haifeng Qian¹, Sheyu Zhang^{1,3}, Yu Hou^{1,2}, Yuchao Zhang¹, Guoyan Luo¹, Xun Zhu^{1,2}, Yanxin Tao^{1,4,5}, Mengyang Fan¹, Hong Wang², Chulin Sha¹, Ailan Lin¹, Jingjing Qin^{1,2}, Kedan Gu¹, Weichang Chen¹, Ting Fu¹, Yajun Wang¹, Yong Wei^{1,7}✉, Qin Wu^{1,5,7}✉ & Weihong Tan^{1,5,7}✉

Despite the well-established significance of transcription factors (TFs) in pathogenesis, their utilization as pharmacological targets has been limited by the inherent challenges in modulating their protein interactions. The lack of defined small-molecule binding pockets and the nuclear localization of TFs do not favor the use of traditional tools. Aptamers possess large molecular weights, expansive blocking surfaces and efficient cellular internalization, making them compelling tools for modulating TF interactions. Here, we report a structure-guided design strategy called Blocker-SELEX to develop inhibitory aptamers (iAptamers) that selectively block TF interactions. Our approach leads to the discovery of iAptamers that cooperatively disrupt SCAF4/SCAF8-RNAP2 interactions, dysregulating RNAP2-dependent gene expression, which impairs cell proliferation. This approach is further applied to develop iAptamers blocking WDR5-MYC interactions. Overall, our study highlights the potential of iAptamers in disrupting pathogenic TF interactions, implicating their potential utility in studying the biological functions of TF interactions and in nucleic acids drug discovery.

Transcription factors (TFs) represent a pivotal protein class that orchestrates gene transcription, exerting profound influence over cellular identity and fate^{1–3}. TF activity is finely tuned through intricate, dynamic interactions with DNA regulatory elements and TF-binding proteins^{4,5}. Dysregulation of TF activity has been implicated in a plethora of human diseases, encompassing cancer, infectious diseases, and neurodegenerative disorders⁶. Consequently, targeted modulation of TF interactions has emerged as a highly promising therapeutic strategy^{7–9}.

A wide range of molecules, including small molecules, peptide mimics, and antibodies, have been employed as modulators of protein-protein interactions⁹. Small molecules, a traditional approach, are utilized to block interactions that involve proteins with well-defined ligand-binding sites¹⁰; accordingly, small molecules have been successfully used in blocking well-folded TF interactions. A successful example is the development of small molecules capable of disrupting p53-MDM2 (mouse double minute 2) protein interactions, a target in the development of antitumor drugs, such as Idasanutlin and

¹Hangzhou Institute of Medicine, Chinese Academy of Sciences, Hangzhou, China. ²School of Pharmacy, Zhejiang University of Technology, Hangzhou, China.

³School of Life Sciences, Tianjin University, Tianjin, China. ⁴Shanghai Institute of Material Medica, Chinese Academy of Sciences, Shanghai, China. ⁵Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou, China. ⁶These authors contributed equally: Tongqing Li, Xueying Liu.

⁷These authors jointly supervised this work: Yong Wei, Qin Wu, Weihong Tan. ✉e-mail: weiyong@ibmc.ac.cn; wuqin@ibmc.ac.cn; tan@hnu.edu.cn

AMG232, which have been designed and subjected to clinical trials (NCT02545283, NCT02110355, et al.)^{11,12}. However, most TF interactions are intrinsically disordered or lacking in well-folded small-molecule binding pockets, posing challenges for small molecule-based interventions. Peptide mimics, derived from peptide fragments involved in protein interfaces, also serve as modulators of protein interactions^{13,14}. For example, OmoMYC and STRs can disrupt MYC-Ebox interactions, previously considered impossible^{15–17}. However, the clinical translation of peptide mimics has been hampered by intrinsic limitations, including intracellular localization, targeting tissue specificity, and pharmacological potency¹⁸. Recent advancements have witnessed remarkable progress in developing antibodies as modulators of protein-protein interactions^{19,20}. Antibodies blocking PD-1 and PD-L1 interaction, such as pembrolizumab and Opdivo® (nivolumab), have exhibited remarkable efficacy in cancer treatment^{21,22}. Nonetheless, antibodies face limitations in targeting intracellular entities owing to their inherent inability to traverse cellular membranes^{9,23}. Consequently, interest is growing in identifying and developing alternative classes of modulators capable of targeting intracellular TF interactions.

Aptamers are short oligonucleotides demonstrating remarkable specificity and high affinity for diverse targets, encompassing proteins, sugars, phospholipids, and cell metabolites. Their versatility has attracted significant attention based on their potential applications in therapy, diagnostics, bioimaging, and drug delivery^{24,25}. Notably, the U.S. Food and Drug Administration (FDA) has granted authorization for three aptamer drugs, namely Pegaptanib, Apc001PE and Zimura (avacincaptad pegol), all serving as potent modulators of protein-protein interactions. Pegaptanib, approved in 2005 for the treatment of age-related macular degeneration, exerts its therapeutic effect by selectively binding to the surface of vascular endothelial growth factor (VEGF), thereby abrogating its interaction with the cellular receptor VEGFR²⁶. Apc001PE, authorized as an orphan drug in 2019 for osteogenesis imperfecta, functions by disrupting the interaction between sclerostin (SOST) and low-density lipoprotein receptor-related protein 5/6 (LRP5/6)²⁷. Avacincaptad pegol, approved in 2023 for geographic atrophy, inhibits the interaction between Complement C5 and its receptor C5R²⁸. These achievements underscore the therapeutic potential of aptamers that specifically target protein interactions. However, most reported aptamers primarily focus on interactions involving cell membrane proteins. Several natural prokaryotic and eukaryotic examples of RNA aptamers that modulate the activity of TFs can be cited. However, only a few unnatural aptamers have been selected to inhibit TF-DNA interactions, including NF- κ B, TBP, HSF1, and RUNX1²⁹, owing to the lack of selection method for aptamers targeting TF interactions.

To solve this dilemma, we herein present a structure-guided rational design strategy, termed Blocker-SELEX, for the development of inhibitory aptamers (iAptamers) that effectively disrupt TF protein-protein interactions. Through a structure-guided screening of a single-stranded DNA (ssDNA) library against the interfaces derived from TF complex structures, we identify and validate a lead sequence with demonstrated target binding and competitive properties. We further optimize the lead sequence by incorporating supporting nucleotides to optimize affinity and competition efficacy. Using this strategy, we successfully develop iAptamers that inhibit the interaction between SCAF4/SCAF8 and RNAP2 proteins. Treatment of tumor cells with selected iAptamers altered the profile of cell transcripts, dysregulates alternative RNA splicing, ultimately reduces cell proliferation, and increased apoptosis. The versatility of our approach is further validated by successfully developing iAptamers targeting the intrinsically disordered oncogenic MYC protein that interacts with WDR5. Altogether, our study has demonstrated a structure-guided design strategy for the development of iAptamer tools that specifically block intracellular TF interactions, highlighting their promising implications in

investigating the biological functions of TF interactions. Given the rapid advancements in nucleic acid delivery techniques, these iAptamers hold promise for nucleic acid drug therapeutics.

Results

Development of Blocker-SELEX pipeline

To design aptamers able to block protein-protein interactions, we proposed a strategy consisting of four main procedures: ① structure-guided virtual screening, ② lead sequence validation, ③ iterations of sequence optimization, and ④ iAptamer validation (Fig. 1a). Structure-based virtual screening is widely used for lead compound discovery in the early stage of small-molecule drug development. This strategy utilizes the three-dimensional structure of target proteins to dock a collection of chemical molecules into the desired binding site, and a subset of compounds with favorable binding scores is selected for further functional evaluation³⁰. While recent attempts have integrated virtual screening with aptamer SELEX, validation can still be challenged by limitations in accurately predicting 3D structures of aptamers in the library pool^{31,32}.

To address this issue, we propose a screening methodology grounded in a library of short DNA sequences with reduced structural complexity. Structural studies have revealed that aptamer recognition of target molecules often relies on a limited number of key bases, such as the aptamers for Thrombin, GRK2, and Hfq^{33–35}. Thus, a library of 1024 (4⁵) single-stranded DNA (ssDNA) sequences with 5 nucleotides was generated for initial virtual screening. The chemical structures of these sequences were created using JChem software and optimized using the LigPrep tool to achieve optimal chemical structures. The binding potential of each DNA sequence to the target protein was calculated using GlideScore, an empirical scoring function that approximates ligand binding free energy. Binding affinity and blocking efficacy of the top-scoring sequences with the highest binding potentials were determined by both surface plasmon resonance (SPR) and competitive fluorescence polarization assay. DNA sequences exhibiting both target binding and concurrent inhibition of interaction were identified as lead sequences for subsequent optimization.

Traditional aptamers basically consist of two regions: core regions and scaffold regions^{36–38}. Scaffold regions serve as a framework for the core regions, thereby heightening their specificity and affinity towards targeted molecules. To strengthen the affinity and inhibitory efficacy of the identified lead sequence, we iteratively added nucleic acid bases to the lead sequence. In each iteration, three bases were appended to the terminal ends of the lead sequence, resulting in the generation of four types of libraries, each containing 64 (4³) sequences. The overall efficacy of these libraries was assessed to identify the leading library, which exhibited the most significant competitive abilities. The 64 sequences within this leading library were then subjected to detailed assays. The rationale for incorporating three bases in each iteration is rooted in practical: First, the resulting 64 optimized sequences can be handled on a single 96-well plate to determine their affinities to the targeted TFs and their increased competitive capabilities in blocking TF interactions. Second, this strategic approach aligns with cost-effective concerns in nucleic acid synthesis. Upon achieving the desired levels of binding affinity and competitive potency, the fine-tuned aptamers are officially designated as iAptamers, marking the culmination of this optimization process.

Identification of a lead sequence targeting SCAF4-RNAP2 interaction using Blocker-SELEX

To investigate the efficacy of the Blocker-SELEX pipeline in generating iAptamers blocking TF protein-protein interactions, we selected the interaction between SCAF4 and RNAP2 as our target. SCAF4, a TF interacting with the C-terminal domain (CTD) of RNAP2, plays a crucial role in regulating its transcriptional activity. Dysregulation of SCAF4-RNAP2 interaction has been implicated in erroneous poly-A

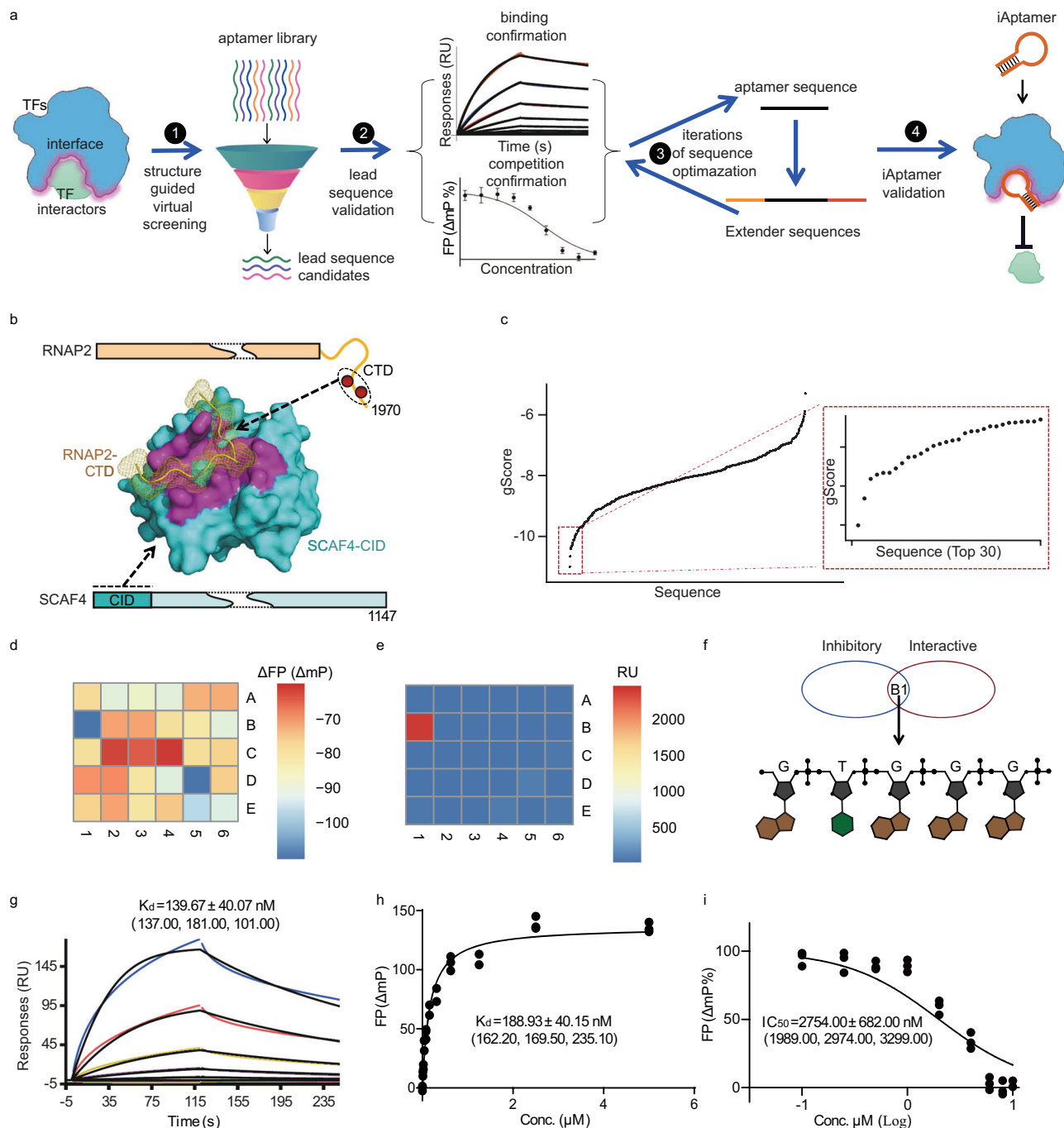


Fig. 1 | Identification of the lead sequence binding to SCAF4 and blocking its interaction with RNAP2. **a** Schematic representation of structure-based design strategy for aptamer inhibitors. Icons from Canva. **b** Surface representation of SCAF4-RNAP2 complex structure re-illustrated from PDB 6XKB. SCAF4 protein is depicted in cyan and RNAP2 in yellow, while the interface is colored in pink. **c** Docking scores of 5-nt sequences to SCAF4 from the Glide pipeline. The top 30 sequences are enlarged in the inserted panel. Source data are provided as a Source Data file. **d** Inhibitory efficacy of top 30 sequences to SCAF4 and RNAP2 interaction are determined by competitive fluorescent polarization assay. Source data are provided as a Source Data file. **e** Direct interactions of top 30 sequences to SCAF4 are determined by SPR spectroscopy. RU, response units. Source data are provided as a Source Data file. **f** Selection of DNA sequences with both binding affinity and inhibitory efficacy. Chemical structure of the identified lead sequence is inserted. **g** Binding of SCAF4_LS to immobilized SCAF4 by standard kinetics SPR. Average affinity value is 139.67 ± 40.07 nM. Affinities of each replicate are indicated

in the bracket (biological replicates $n = 3$, mean $\pm$ SD). Source data are provided as a Source Data file. **h** Fluorescence polarization measurements where 50 nM FAM-labeled SCAF4_LS was titrated with increasing amounts of SCAF4 proteins. The affinity is determined as 188.93 ± 40.15 nM using the FP (Δ mP) signals from three biological replicates, each containing three technical replicates. Affinities of each biological replicate are indicated in the bracket (biological replicates $n = 3$, mean $\pm$ SD). The non-linear fit model and the one-site specific binding to fluorescent population ratio are used for these calculations at the respective aptamer concentrations. Source data are provided as a Source Data file. **i** Dose-response curves of SCAF4_LS showing the inhibition of RNAP2 interaction with SCAF4. The IC_{50} is determined from three independent biological replicates, each containing three technical replicates. IC_{50} values are calculated using GraphPad Prism software (version 7) with the Log(inhibitor) vs. normalized response model. IC_{50} values of all experiments are indicated in the bracket (biological replicates $n = 3$, mean $\pm$ SD). Source data are provided as a Source Data file.

modification, aberrant transcriptional termination, and cell proliferation arrest³⁹. A previous study⁴⁰ of SCAF4-RNAP2 interaction revealed that it is an undruggable pocket, which underscores the challenge in devising traditional pharmacological strategies (Fig. 1b; Supplementary Fig. 1a). Consequently, the absence of specifically designed chemical tools has hindered exploration of the precise biological function associated with this interaction. Therefore, we aimed to design iAptamers targeting the interface between SCAF4 and RNAP2, thereby disrupting their interaction and impeding their functional partnership.

To prepare the docking receptor, the structure of SCAF4 was extracted from the SCAF4-RNAP2 complex (PDB code: 6XKB, chain A). Specifically, we selected the protein-protein interaction (PPI) interface of SCAF4 that recognizes the RNAP2-CTD domain as the docking target (Fig. 1b). 4⁵ ssDNA molecules from the library were docked against the prepared receptor of SCAF4 to evaluate the binding potentials using GlideScore (gScore) (Fig. 1c). The gScore data from virtual screening showed promising results, with gScores <−8 for 283 sequences, instilling confidence in the outcomes. Drawing upon insights from the experience of virtual screening and subsequent validation of the top 30–500 small molecule inhibitors depending on the library size^{41,42}, we chose to validate the top 30 sequences, about 3% of our library. Consequently, the top 30 ssDNA sequences with the highest Glide scores were selected for synthesis and experimental validation to assess their inhibitory efficacies against SCAF4-RNAP2 interaction, and their binding to SCAF4, using SPR. To investigate the potential inhibitory efficacy of these sequences on SCAF4-RNAP2 interaction, we established a competitive screening assay based on fluorescence polarization. Initially, the interaction between SCAF4 and RNAP2-CTD was captured with a phosphorylated CTD peptide (pS2pS5) of RNAP2 labeled with an NH2 terminal FITC group for polarization measurements. The affinity between phosphorylated CTD and SCAF4 was determined as 365.60 nM (Supplementary Fig. 1b). For the competition screening, FP signals (> 60 ΔmP) were generated by combining 2 μM SCAF4-CID with 40 nM FITC-labeled CTD peptide solutions. Reductions in FP signals were observed with the addition of the top-scoring sequences. Notably, sequences 5'-GTGGG-3' and 5'-CTGGG-3' exhibited significant signal reduction, suggesting their competitive roles in blocking SCAF4-RNAP2 interactions (Fig. 1d). Concurrently, we investigated the interactions of these top-scoring sequences with SCAF4. As summarized in Fig. 1e, ssDNA 5'-GTGGG-3' demonstrated significant binding to SCAF4 with a response unit exceeding 2000 (RU), while the remaining sequences displayed only weak, or negligible, responses (Fig. 1e). This observation shows that not all sequences effectively bind to SCAF4, underscoring the target specificity of 5-nt ssDNA. Based on these findings, ssDNA 5'-GTGGG-3' with both binding affinity and inhibitory efficacy was selected for further determination (Fig. 1f). The binding affinity of ssDNA 5'-GTGGG-3' to SCAF4 was further determined to be 139.67 ± 40.07 nM by SPR (Fig. 1g; Supplementary Fig. 1c). A fluorescence polarization assay verified its affinity to be 188.93 ± 40.15 nM (Fig. 1h; Supplementary Fig. 1d). The inhibitory capacity (IC₅₀) of ssDNA 5'-GTGGG-3' was calculated as 2754.00 ± 682.00 nM, confirming its competitive role in disrupting SCAF4-RNAP2 interaction (Fig. 1i; Supplementary Fig. 1e). Due to its potent inhibitory efficacy and robust binding affinity, ssDNA 5'-GTGGG-3' (hereinafter referred to as the SCAF4 lead sequence (SCAF4_LS)), was selected as the primary iAptamer in this study.

Specific interaction of SCAF4_LS binding to SCAF4

To gain insight into the interaction between SCAF4_LS and SCAF4, a molecular dynamics (MD) simulation was carried on the initial complex model from docking. The production phase simulations were performed for a total time of 500 ns at a constant temperature of 298 K. The binding position of SCAF4_LS on SCAF4 shifted but still stayed tightly associated, achieving a stable binding pose after 300 ns. The binding poses of SCAF4_LS in the last 100 ns (400–500 ns) were

summarized in Fig. 2a. The RMSF values of each nucleotide from all SCAF4_LS conformations were plotted against the simulation time from 400 ns to 500 ns for stability analysis (Fig. 2b). The trajectory clearly shows that the 3' terminus exhibited greater flexibility compared to its 5' terminus, while G1, T2, and G3 bases at the 5' terminus exhibited stable binding with SCAF4 protein, as evidenced by the RMSF (Fig. 2b). This observation was further corroborated by plotting the RMSD values of individual nucleotides against the simulation time between 400 to 500 ns (Fig. 2c). The simulation results suggest that SCAF4_LS binds to SCAF4 in a stable manner and that the first three nucleotides exhibit heightened interactions with the protein, while the last two nucleotides slide over the protein surface (Fig. 2a).

SCAF4_LS exhibits a marked predilection for engaging with a positively charged groove on the PPI interface of SCAF4, as demonstrated in the molecular simulation assay (Fig. 2d). Superposition of the simulated SCAF4_LS aptamer-SCAF4 model and RNAP2-SCAF4 complex structure (6XKB) reveals an overlap between the RNAP2-CTD and SCAF4_LS aptamer. Our analysis identified 17 amino acids of SCAF4 involved in interactions with SCAF4_LS, while 19 amino acids of SCAF4 were involved in interactions with RNAP2, with 13 amino acids shared between the two interfaces (Supplementary Fig. 1f). Projection of these residues onto the 3D structure of SCAF4 highlights a substantial overlap (1012.27 Å², solvent accessible surface area (SASA)) in the contact areas between the two interfaces. This overlap encompasses approximately 75.47% of the SCAF4_LS interface (1341.31 Å², SASA) and 65.74% of the RNAP2 interface (1539.90 Å², SASA), indicating the competitive role of SCAF4_LS in blocking the interaction between SCAF4 and RNAP2 (Supplementary Fig. 1f). To determine the binding specificity of SCAF4_LS to the PPI interface, critical residues on the SCAF4 interface were substituted, and their binding to SCAF4_LS was characterized using a fluorescence polarization assay. Three arginine residues (R23, R71, and R112), known to be involved in SCAF4 recognition of RNAP2⁴⁰, were substituted with alanines, resulting in weaker interaction with SCAF4_LS and reducing the binding affinity by over two-fold (Fig. 2e; Supplementary Fig. 1g–i).

In parallel, to investigate the base specificity of the SCAF4_LS-SCAF4 interaction, a series of nitrogenous base-group deletions within the SCAF4_LS sequence was synthesized, while preserving the pentose sugar and phosphate group (Fig. 2f). These deletions were incorporated into a seven-based backbone sequence (5'-GGTGGGG-3') containing two additional nucleotides, required in DNA synthesis. The affinity of each base deletion of SCAF4_LS to SCAF4 was then assessed by SPR (Fig. 2g; Supplementary Fig. 2a–f). Significantly, the elimination of guanine groups from nucleotides G1, G3, and G5 drastically abrogated interaction between SCAF4_LS and SCAF4, underscoring the critical roles of these three guanine base groups in recognizing SCAF4. Additionally, removing the guanine group from G4 led to reduced affinity. However, SCAF4_LS lacking the thymine group at T2 still maintained its affinity to SCAF4. These findings collectively underscore that the importance of nitrogenous bases in the recognition between SCAF4 and SCAF4_LS. Sequences lacking specific bases exhibited no binding affinity to SCAF4, further confirming the sequence-specific nature of this interaction. These results highlight the capacity of different combinations of base groups within 5-nt ssDNA to facilitate the recognition of target proteins and validate the rationale behind utilizing 5-nt ssDNA as the initial screening sequence.

Optimization of SCAF4_LS

Having identified the initial SCAF4_LS aptamer, further optimization was performed by introducing three nucleotides randomly to the termini of the lead sequence, resulting in four libraries: nGTGGGnn, nnGTGGGn, nnnGTGGG and GTGGGnnn, each containing 64 (4³) sequences (Fig. 3a). We assessed the overall efficacy of these libraries in blocking SCAF4-RNAP2 interactions using a competitive fluorescence polarization assay (Fig. 3b). Among them, the 'nGTGGGnn'

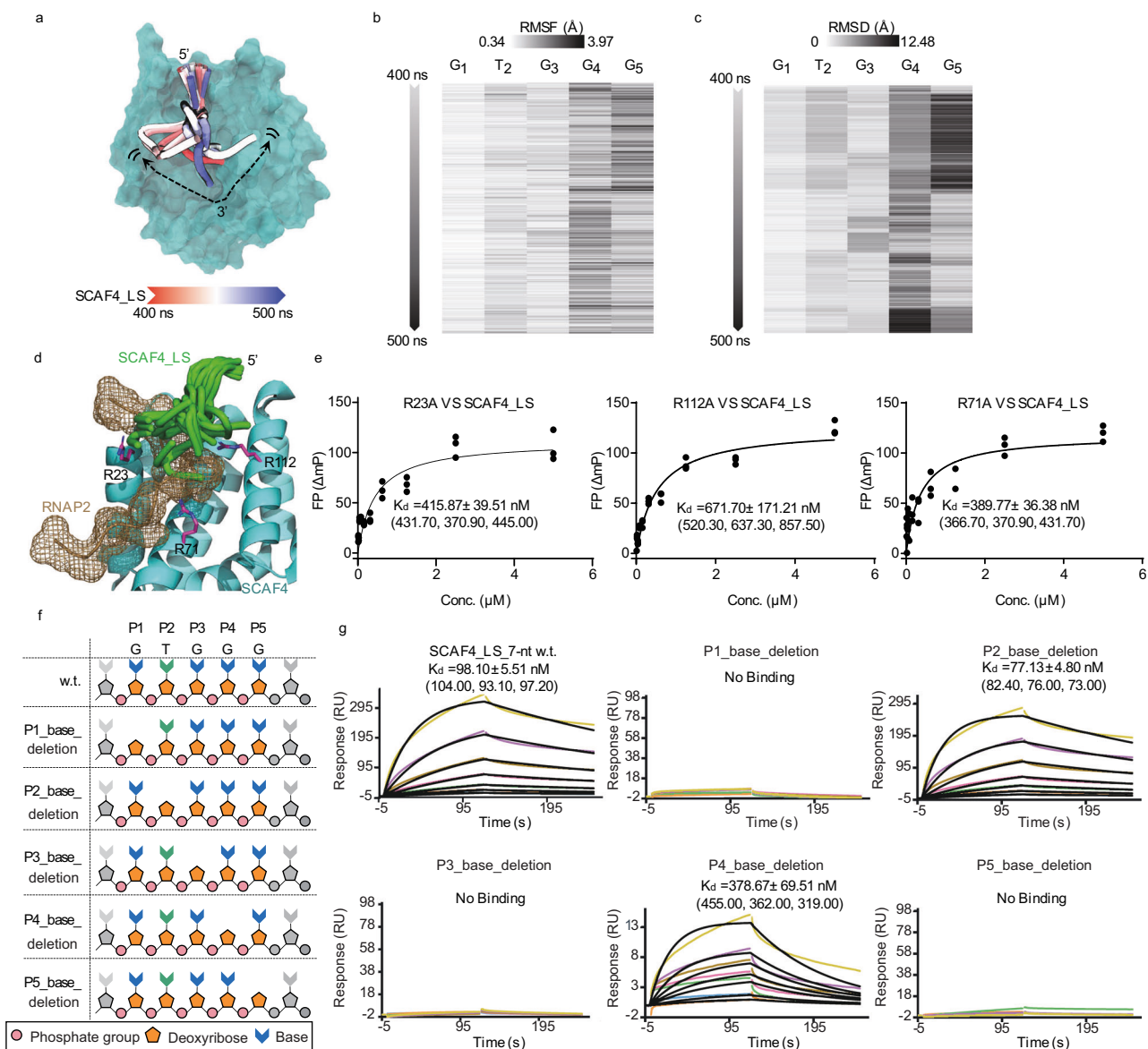


Fig. 2 | SCAF4_LS binding at the interface between SCAF4 and RNAP2.

a Molecular simulation of SCAF4_LS binding to SCAF4 protein. The colors of SCAF4_LS are color-coded based on the time interval from 400 ns to 500 ns. **b** Trajectory plot showing the RSMF of five nucleotides over the 400–500 ns time scale. Source data are provided as a Source Data file. **c** Trajectory plot showing the RMSD of five nucleotides over the 400–500 ns time scale. Source data are provided as a Source Data file. **d** Overlay of SCAF4_LS sequence and RNAP2 on the surface of SCAF4 protein. Conformations of SCAF4_LS are represented in green cartoon, while RNAP2 peptide is depicted in sand-colored mesh. **e** Fluorescence polarization measurements where 50 nM FAM-labeled SCAF4_LS was titrated with increasing

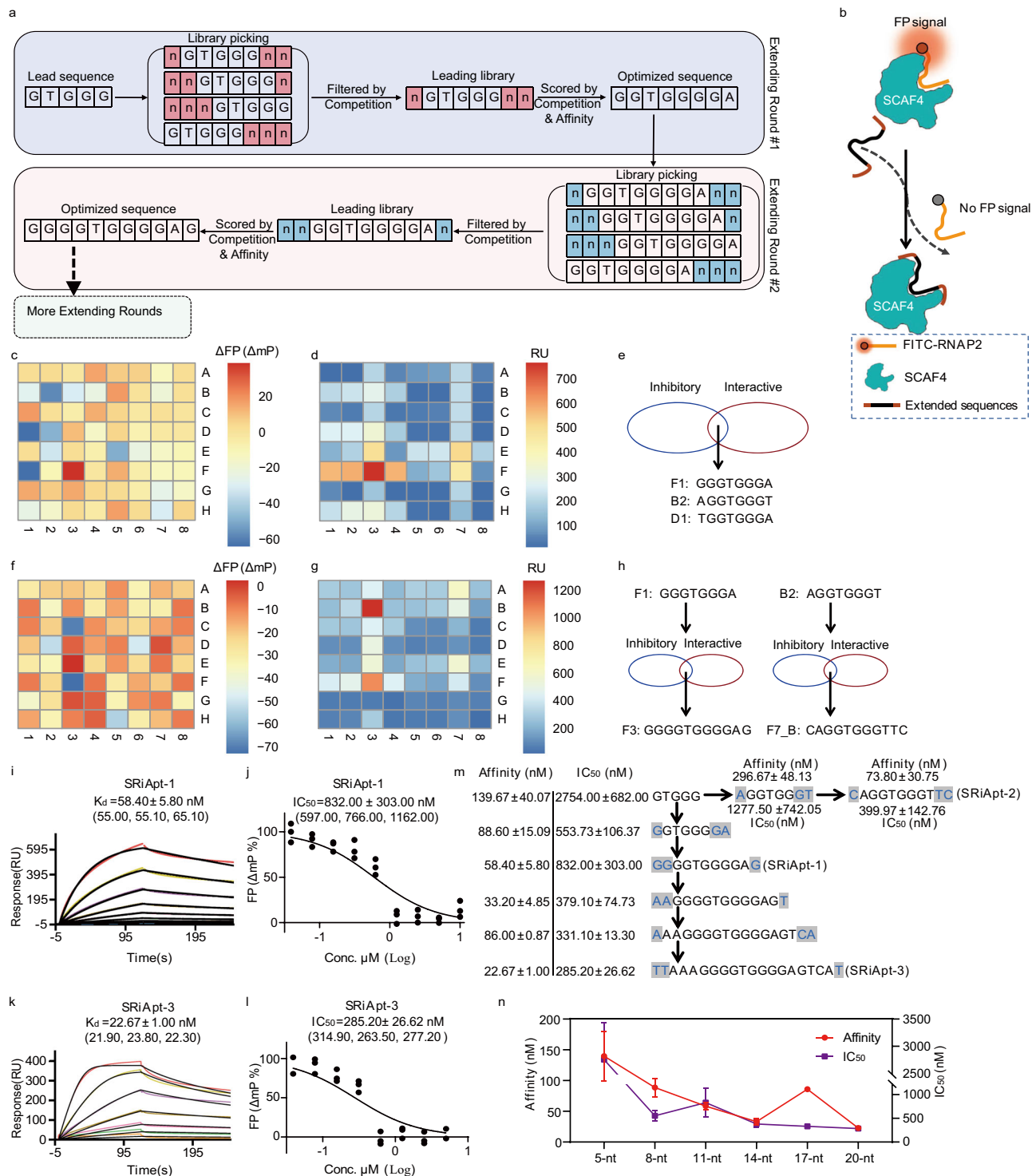
amounts of SCAF4 mutants. All experiments are performed with three biological replicates, each containing three technical replicates. Affinities of each biological replicate are indicated in the bracket (biological replicates $n = 3$, mean $\pm$ SD). The non-linear fit model and the one-site specific binding to fluorescent population ratio are used for these calculations. Source data are provided as a Source Data file. **f** Schematic diagram of SCAF4_LS with deletion of base groups at each position. **g** SPR results showing binding of wild type and base group-deleted SCAF4_LS to SCAF4. Average affinities are generated from three biological replicates, with affinities of each experiment indicated in the bracket (biological replicates $n = 3$, mean $\pm$ SD). Source data are provided as a Source Data file.

library demonstrated the highest overall competition ability and was selected for further investigation, with its 64 sequences subjected to subsequent assays (Supplementary Fig. 3a).

These expanded sequences were then evaluated for their inhibitory efficacy using the competitive fluorescence polarization assay (Fig. 3b). Three sequences (B2_8-nt, D1_8-nt, and F1_8-nt) displayed significant signal reductions, indicating their pronounced competitiveness in blocking SCAF4-RNAP2 interactions (Fig. 3c). The IC_{50} s of B2_8-nt, D1_8-nt, and F1_8-nt on SCAF4-RNAP2 interactions were 1040.47 ± 418.56 nM, 1265.67 ± 864.02 nM and 553.73 ± 106.37 nM, respectively, by competitive FP assays (Supplementary Fig. 3b). Concurrently, the interactions of these 64 expanded sequences with

SCAF4 were investigated using SPR. Multiple sequences, including B2_8-nt, D1_8-nt, and F1_8-nt, showed significant binding to SCAF4 with response units (RU) over 200 (Fig. 3d). The binding affinities of F1_8-nt and B2_8-nt to SCAF4 were further quantified as 88.60 ± 15.09 nM and 296.67 ± 48.13 nM, respectively, by SPR (Supplementary Fig. 3c, d). Based on these findings, two sequences, F1_8-nt and B2_8-nt, were selected for further optimization (Fig. 3e).

In the subsequent round of aptamer refinement from F1_8-nt, two remarkably inhibitory aptamer sequences (C3_11-nt and F3_11-nt) were identified through a competitive FP assay, mirroring the prior optimization process (Fig. 3f). Concurrently, F3_11-nt exhibited the most substantial binding to SCAF4 in the SPR screening assay (Fig. 3g).



Consequently, the F3_11-nt sequence was selected for further evaluation (Fig. 3h). The binding affinity of F3_11-nt to SCAF4 was determined to be 58.40 ± 5.80 nM (Fig. 3i; Supplementary Fig. 3e). The inhibitory efficacy of SCAF4_F3_11-nt aptamer on SCAF4-RNAP2 interactions was verified through competitive FP assay, showing an IC_{50} value of 832.00 ± 303.00 nM, showcasing its enhanced inhibitory potential compared to the parent SCAF4_LS aptamer (Fig. 3j; Supplementary Fig. 3f). Despite the guanine-rich nature of F3_11-nt, G-quadruplex structures were not observed within this sequence across various K^+ concentrations (Supplementary Fig. 3g). The binding specificity of SCAF4_F3_11-nt to SCAF4 was confirmed by examining its interactions with three unrelated proteins (streptavidin (SA), γ SPA, and Pepsin),

revealing no observable interactions, thus affirming its binding specificity (Supplementary Fig. 3h). The stability of SCAF4_F3_11-nt was evaluated using an in vitro serum stability assay, with a half-life of 15.04 h against 20% serum (Supplementary Fig. 3i).

Simultaneously, extended sequences from B2_8-nt underwent screening via competitive FP assay. Many sequences demonstrated substantial reductions in FP signal, leading to the selection of three sequences (F7_11-nt_B, H7_11-nt_B, and G8_11-nt_B) for further analysis. The inhibitory potency of each sequence towards SCAF4-RNAP2 interaction was determined to be 399.97 ± 142.76 nM, 437.50 ± 129.49 nM, and 459.57 ± 188.98 nM, respectively, through competitive FP assays (Supplementary Fig. 4a–c). Moreover, the

Fig. 3 | Optimization of SCAF4_LS to improve its affinity and inhibitory capacity. **a** Workflow illustrating the sequence optimization procedures through the addition of supporting nucleotides to SCAF4_LS. **b** Schematic representation of the competition assay used for the selection of extended sequences. **c** Evaluation of the inhibitory efficacy of the expanded 64 sequences, utilizing a competitive FP assay. **d** Direct interactions of the expanded 64 sequences with SCAF4 assessed by SPR spectroscopy. **e** Selection of DNA sequences exhibiting both binding affinity and inhibitory efficacy. **f** Assessments of the inhibitory efficacy of the second round of expanded 64 sequences on the SCAF4 and RNAP2 interaction. **g** Direct interactions of the second round of expanded 64 sequences to SCAF4. **h** Selection of DNA sequences demonstrating both binding affinity and inhibitory efficacy. **i** Standard kinetics SPR assay demonstrating the binding of SCAF4_F3_11-nt (SRiApt-1) to immobilized SCAF4. Average affinity value is 58.40 ± 5.80 nM (biological replicates $n = 3$, mean $\pm$ SD). Source data are provided as a Source Data file. **j** Dose-response curves depicting the inhibition of RNAP2 interaction with SCAF4 by SCAF4_F3_11-nt

(SRiApt-1). The IC_{50} is determined from three biological replicates, each containing three technical replicates (biological replicates $n = 3$, mean $\pm$ SD). Source data are provided as a Source Data file. **k** Standard kinetics SPR assay demonstrating the binding of SRiApt-3 to immobilized SCAF4. Average affinity value is 22.67 ± 1.00 nM (biological replicates $n = 3$, mean $\pm$ SD). Source data are provided as a Source Data file. **l** Dose-response curves depicting the inhibition of RNAP2 interaction with SCAF4 by SRiApt-3. The IC_{50} is determined from three biological replicates, each containing three technical replicates (biological replicates $n = 3$, mean $\pm$ SD). Source data are provided as a Source Data file. **m** Summary of aptamers optimized from the lead sequence. Both affinities and IC_{50} s of selected aptamers are labeled on the left, and the sequence information was indicated on the right. Residues iterated from the previous sequence are labeled in gray. **n** Iterations of affinities and IC_{50} s for aptamers, with affinity shown in red and IC_{50} displayed in blue (biological replicates $n = 3$, mean $\pm$ SD). Source data are provided as a Source Data file.

binding affinity between each sequence, F7_11-nt_B, H7_11-nt_B, and G8_11-nt_B, and SCAF4 was quantified via SPR as 73.80 ± 30.75 nM, 62.03 ± 14.32 nM, and 57.03 ± 1.65 nM, respectively (Supplementary Fig. 4d–f). Owing to the substantial affinity of each sequence to SCAF4 and potent inhibitory efficacy on SCAF4-RNAP2 interactions, SCAF4_F3_11-nt and SCAF4_F7_11-nt_B were designated as SCAF4-RNAP2 inhibitory aptamer 1 (SRiApt-1) and SRiApt-2, respectively (Fig. 3h, m).

At last, SRiApt-1 underwent further optimization to maximize its binding and inhibitory capabilities. Through additional three rounds of iterative refinement, we identified a 20-base DNA sequence with an affinity of 22.67 ± 1.00 nM to SCAF4 and an IC_{50} of 285.20 ± 26.62 nM for blocking SCAF4-RNAP2 interactions (Fig. 3k, l; Supplementary Figs. 5a, b, 6a). Consequently, this aptamer was designated as SCAF4-RNAP2 inhibitory aptamer 3 (SRiApt-3). The binding specificity of SRiApt-3 was investigated by assessing its interactions with three unrelated proteins (streptavidin (SA), γ SPA, and Pepsin), revealing no observable interactions and thereby affirming its binding specificity (Supplementary Fig. 6b). Furthermore, the stability of SRiApt-3 was assessed using an in vitro serum stability assay, revealing a half-life of 13.01 h without ACTA and 20.12 h with ACTA packaging (Supplementary Fig. 6c).

In summary, employing the Blocker-SELEX pipeline and a series of iterative optimizations on the lead sequence, we achieved a continuous enhancement in both affinity and inhibitory capacity of the aptamers (Fig. 3m). While there were occasional fluctuations in affinity or inhibitory trends, the overall effectiveness of the iterative optimization approach to improve the performance of aptamers was validated (Fig. 3n).

SCAF4_LS sequence was recaptured in developing aptamers for SCAF8

SCAF8, a crucial paralog of SCAF4, shares a comparable CID domain with an overall RMSD of 0.266 Å and a PPI interface for RNAP2 recognition^{40,43} (Supplementary Fig. 7a). Concomitant knockout of human SCAF4 and SCAF8 altered poly-A selection, leading to truncated mRNAs and proteins, causing cell death³⁹. We aimed to select inhibitory aptamers to block SCAF8-RNAP2 interaction using the Blocker-SELEX pipeline. Interface details were obtained from the SCAF8-RNAP2 complex structure (PDB code: 3D9M). The binding potential of 5-nt ssDNA sequences was scored by Glide docking. As anticipated, the structural similarity between SCAF4-CID and SCAF8-CID caused the SCAF4_LS sequence to be a top-scoring sequence (Fig. 4a).

To evaluate the inhibitory efficacy of these sequences, SCAF8-RNAP2 interaction was recapitulated using a fluorescence polarization assay, yielding an affinity of 622.40 nM (Supplementary Fig. 7b). The top 20 sequences were screened for their inhibitory efficacy by adding 10 μ M of each into the FP solutions. Among the top 20 sequences, D2_5-nt (SCAF4_LS) and A5_5-nt exhibited the most potent competition with SCAF8-RNAP2 interaction, as determined by

competitive fluorescence polarization assay (Fig. 4b). The binding of each of the 20 sequences to SCAF8 was assessed using SPR, revealing D2_5-nt (SCAF4_LS) bound to SCAF8 with RU over 1000, while the remaining sequences displayed no significant responses (Fig. 4c). Based on these results, D2_5-nt (SCAF4_LS) was selected for further investigation (Fig. 4d). The affinity between D2_5-nt (SCAF4_LS) and SCAF8 was quantified as 76.40 ± 7.33 nM, using SPR (Fig. 4e; Supplementary Fig. 7c). Moreover, the inhibitory efficacy of D2_5-nt (SCAF4_LS) on SCAF4-RNAP2 interaction was determined with an IC_{50} value of 1.96 ± 0.24 μ M (Fig. 4f; Supplementary Fig. 7d). Collectively, D2_5-nt (SCAF4_LS), with significant binding affinity and inhibitory efficacy, was selected as the lead sequence for targeting SCAF8-RNAP2 interaction.

To assess the contribution of each nucleotide base of D2_5-nt (SCAF4_LS) to SCAF8 binding, the affinity of each base group-deleted variant of D2_5-nt (SCAF4_LS) was investigated (Figs. 2f, 4g–k; Supplementary Fig. 8). Similar to the interaction between SCAF4_LS and SCAF4, removing guanine groups at positions P1, P3, P4, and P5 resulted in substantial reductions in affinity, indicating the critical role of these four guanine groups in D2_5-nt (SCAF4_LS) recognition. We further examined the binding affinities of optimized SRiApt-1/2/3 to SCAF8, determined as 56.27 ± 7.42 nM, 90.50 ± 25.80 nM and 69.90 ± 13.96 nM, respectively (Supplementary Fig. 7e–g). The IC_{50} s of SRiApt-1/2/3 in blocking SCAF8-RNAP2 interaction were determined as 1.50 ± 0.69 μ M, 2.69 ± 1.43 μ M and 291.1 ± 19.68 nM, respectively (Supplementary Fig. 7h–j). These findings support that SRiApt-1/2/3 aptamers simultaneously block SCAF4 or SCAF8-RNAP2-CTD interactions.

Ultimately, our study highlights the reproducibility of the Blocker-SELEX approach in generating iAptamers that specifically disrupt SCAF4/SCAF8 and RNAP2-CTD interactions.

Binding of SRiApt-1/3 aptamers to SCAF4/SCAF8 at the cellular level

To investigate the interactions between the selected inhibitory aptamers and SCAF4/SCAF8 within cells, an immunoprecipitation (IP) assay was performed using SRiApt-1/3. The IP assays demonstrated that the SRiApt-1/3 aptamer-loaded beads captured SCAF4 or SCAF8, while the control sequences showed no significant enrichment, indicating direct interaction within cells (Fig. 5a). As SCAF4/8-RNAP2 interactions predominantly occur within the nucleus, it was crucial to evaluate the nuclear accumulation of the selected inhibitory aptamers. Hence, fluorescence localization assays examined the subcellular distributions of SRiApt-1/3. FAM-labeled SRiApt-1/3 aptamers were introduced into living cells through ACTA-LNPs, and their distribution was monitored over time. Within 6 h of transfection, the FAM-labeled SRiApt-1/3 aptamers exhibited significant accumulation in the nucleus, as indicated by strong green fluorescence signals. This observation confirms the rapid nuclear localization of the SRiApt-1/3 aptamers (Fig. 5b; Supplementary Fig. 9a, b). Furthermore, to investigate whether the

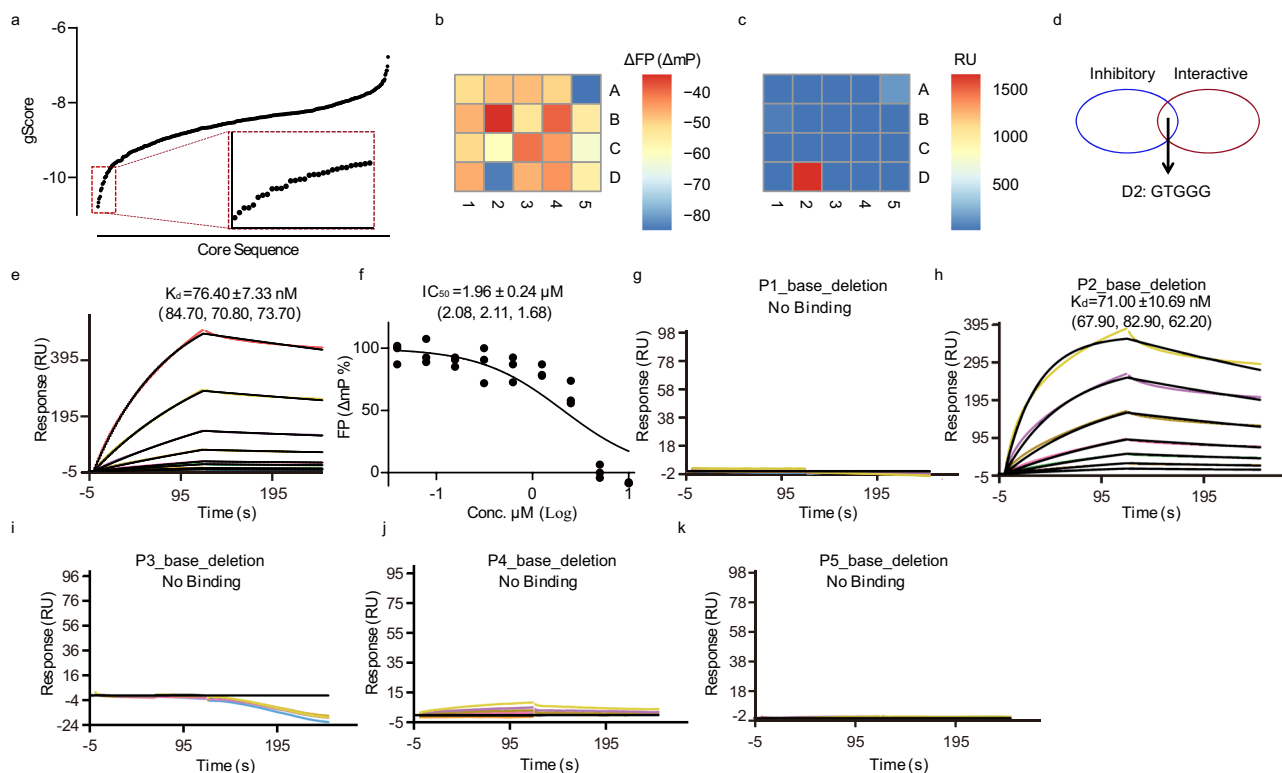


Fig. 4 | Identification of the lead sequence binding to SCAF8 and blocking its interaction with RNAP2. **a** Docking scores of 5-nt sequences to SCAF8 generated by the Glide pipeline. The top 20 sequences are enlarged in the inserted panel. Source data are provided as a Source Data file. **b** Assessments of the inhibitory efficacy of the top 20 sequences on the interaction between SCAF8 and RNAP2 via competitive fluorescent polarization assay. **c** Evaluation of the direct interactions between the top 20 sequences and SCAF8 using SPR spectroscopy. RU, response units. **d** Selection of DNA sequences exhibiting both binding affinity and inhibitory efficacy. **e** Standard kinetics SPR assay demonstrating the binding of SCAF4_LS (D2_5-nt) to immobilized SCAF. This experiment is performed with three biological replicates. Average affinity value is 76.40 ± 7.33 nM. Affinities of each replicate are

indicated in the bracket (biological replicates $n = 3$, mean $\pm$ SD). Source data are provided as a Source Data file. **f** Dose-response curves of SCAF4_LS (D2_5-nt) depicting the inhibition of RNAP2 interaction with SCAF8. The IC_{50} is determined from three biological replicates, each containing three technical replicates. The averaged IC_{50} value is calculated from all experiments indicated in brackets (biological replicates $n = 3$, mean $\pm$ SD). Source data are provided as a Source Data file. **g–k** SPR results illustrating the binding of base group-deleted SCAF4_LS (D2_5-nt) sequences to SCAF8. Average affinities are generated from three independent experiments, with affinities of each experiment indicated in the bracket. Source data are provided as a Source Data file.

SRiApt-1/3 aptamers could disrupt SCAF4/8-RNAP2 interaction in a cellular context, a CoIP assay was performed, using SCAF4 as an example. HA-tagged SCAF4-CID was exogenously overexpressed, and endogenous RNAP2 protein was co-precipitated using an anti-HA antibody, but the co-precipitation was significantly reduced upon SRiApt-1/3 treatment. These results conclusively validate the inhibitory potential of SRiApt-1/3 in blocking SCAF4-RNAP2 interaction (Fig. 5c).

SRiApt-1 aptamer altered gene transcription and RNA splicing

Previous studies have demonstrated that simultaneous depletion of SCAF4 and SCAF8 disrupts the gene transcriptome and promotes alternative splicing events³⁹. To validate the effect of those aptamers on gene transcription and mRNA splicing by inhibiting RNAP2-SCAF4/SCAF8 interactions, we employed the SRiApt-1 aptamer. Initially, we enhanced the stability of SRiApt-1 by introducing two 2'-O-Me modifications, resulting in SRiApt-1-m (5'-G(2'-O-Me)GGGTGGGGA(2'-O-Me)G-3'), with an increased half-life from 15.04 to 18.65 h (Supplementary Fig. 3i). Subsequently mRNA sequencing was conducted on HCT116 cells treated with SRiApt-1-m, or the reverse sequence for 36 h. RNA-seq analysis revealed significant transcriptome alterations in HCT116 cells upon SRiApt-1-m treatment (Fig. 5d). Specifically, 97 genes were differentially expressed with 72 genes upregulated and 25 genes downregulated ($p < 0.1$, Fig. 5d). Gene set enrichment analysis (GSEA) was performed, showing upregulation of the spliceosome pathway upon SRiApt-1-m treatment (Fig. 5e). Further examination of

alternative splicing events based on the RNA-seq data revealed a substantial increase, totaling 5821 events, with over 50% classified as skipped exons (SE) after SRiApt-1-m treatment (Fig. 5f, g). Additionally, numerous non-coding RNAs (ncRNAs) were identified in the RNA-seq data (Fig. 5h). These findings align with a recent study that demonstrated significant alternative splicing events resulting from the double knockout of SCAF4 and SCAF8³⁹.

Since one of the crucial cellular processes influenced by alternative splicing is programmed cell death, or apoptosis, we investigated the RNA-seq data and observed activation of the apoptosis signaling pathway following treatment with SRiApt-1-m (Fig. 6a). To validate this observation, a cell apoptosis assay was performed using HCT116 cells. Cells were exposed to SRiApt-1-m, or the reverse sequence, using vehicles at a final concentration of $1 \mu\text{M}$. Subsequently, apoptotic cells were labeled with Annexin V, a marker of apoptosis. A noticeable increase in apoptotic cells was observed compared to the control, confirming SRiApt-1-m's ability to promote apoptosis in HCT116 cells (Fig. 6b, Supplementary Fig. 9c). Furthermore, the CCK8 assay further confirmed the inhibitory efficacy of SRiApt-1/3 or SRiApt-1-m on HCT116 cell proliferation (Fig. 6c, Supplementary Fig. 9d). Collectively, these findings substantiate the potential therapeutic utility of SRiApt-1-m inhibitor in impeding tumor growth. Consequently, these observations support the on-target effect of SRiApt-1 and establish its significance as a valuable chemical tool for investigating RNAP2-SCAF4/SCAF8 interactions.

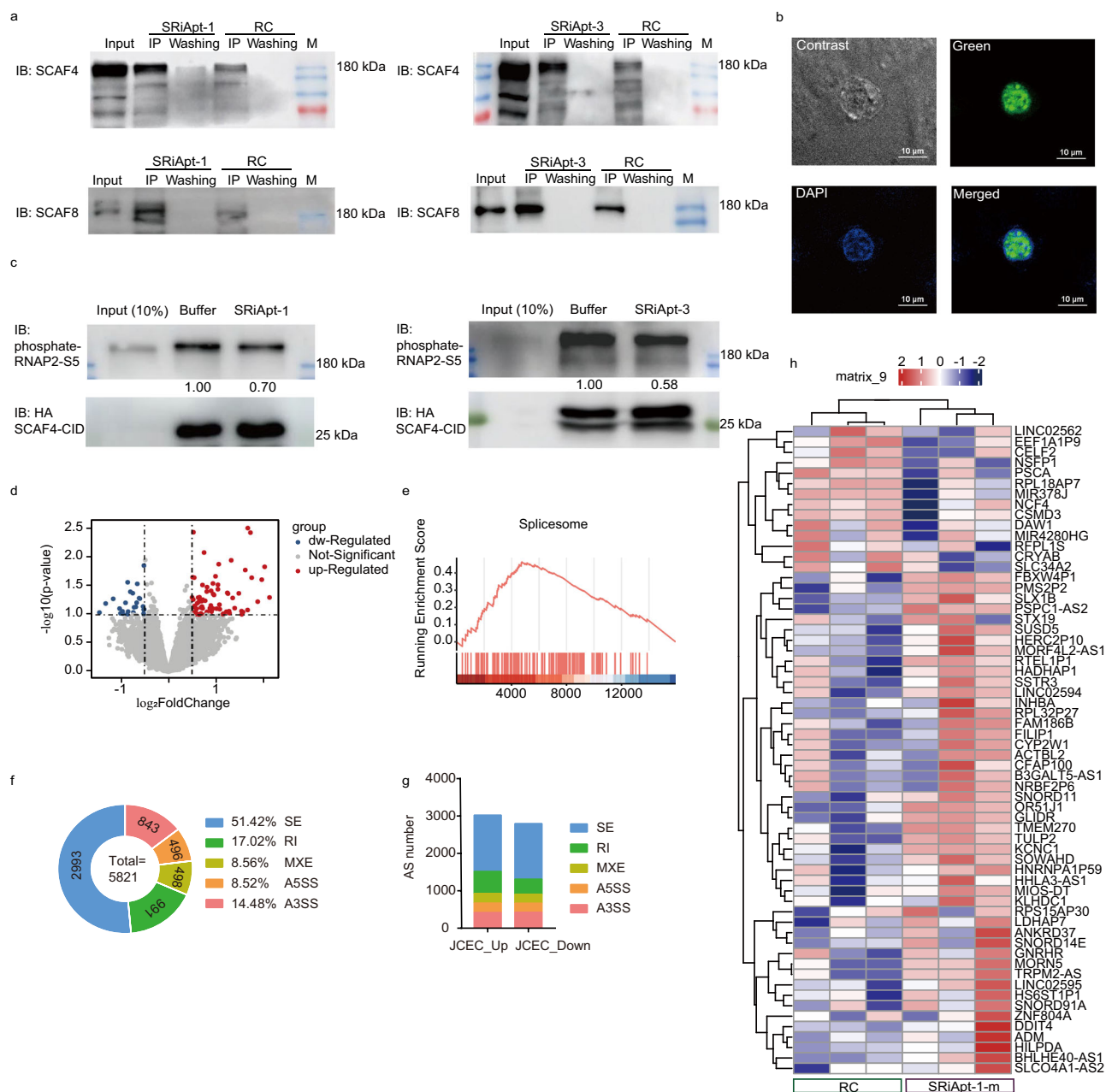


Fig. 5 | Biological characterizations of the SRiApt-1/3 aptamers. **a** Co-IP assays using Biotin-SRiApt-1/3 aptamers or Biotin-RC sequences (two biological replicate). Source data are provided as a Source Data file. **b** Intracellular localization of FAM-labeled SRiApt-1 aptamer following transfection with LNPs. Aptamers are highlighted in green, while DAPI staining is shown in blue. Scale bar: 10 μ m. (biological replicates $n = 2$) **c** Co-IP assays revealing the inhibitory role of SRiApt-1/3 aptamers in blocking the interaction between SCAF4 and RNAP2 (one biological replicate). Source data are provided as a Source Data file. **d** A volcano plot representing the differential expression of genes between the treated group with SRiApt-1 aptamers to that in the control group treated with the RC sequence. Genes that exhibited significant changes are selected using $\log_2$ FoldChange criteria (>0.5 or <-0.5) and p value threshold (<0.1 , Wald test). Each dot on the plot represents an individual gene, and their color indicates the corresponding regulation as specified in the

legend. **e** Gene Set Enrichment Analysis (GSEA) for KEGG enrichment (NES = 1.665, $p < 0.05$) of the Spliceosome pathway after treatment with SRiApt-1 aptamers. Data analyzed by one-tailed Fisher's exact test. **f** Pie plot displaying the number and ratio of ASEs belonging to each of the main alternative splicing categories. **g** Bar plot showing the number and ratio of up-/down-regulated ASEs belonging to each of the main alternative splicing categories. The color represents various AS categories. The horizontal axis represents upregulation or downregulation of AS events, and the vertical axis represents the number of alternative splicing events for each type. JCEC represents AS event detection using both Junction Counts and reads on target. **h** Heatmap representing significant upregulated and downregulated mRNA or non-coding RNA with SRiApt-1 aptamers treatment (absolute value of $\log_2$ fold change >1), compared to the RC sequence treating.

Blocker-SELEX pipeline is feasible against WDR5-MYC interaction

To assess the versatility of Blocker-SELEX in selecting aptamers targeting TF interactions, we aimed to develop inhibitory aptamers blocking the interactions of the intrinsically disordered oncogenic MYC protein. Among the cofactors known to regulate MYC, WD40-

repeat protein 5 (WDR5) is a key determinant for recruiting MYC to chromatin, a process essential for MYC's oncogenic properties⁴⁴. Disrupting this interaction impedes MYC binding at approximately 80% of its chromosomal locations, deactivating its oncogenic properties⁴⁴. Utilizing Blocker-SELEX, we prepared the WDR5-MYC binding interface from the crystal structure of WDR5-MYC complex (PDB code: 4Y7R)

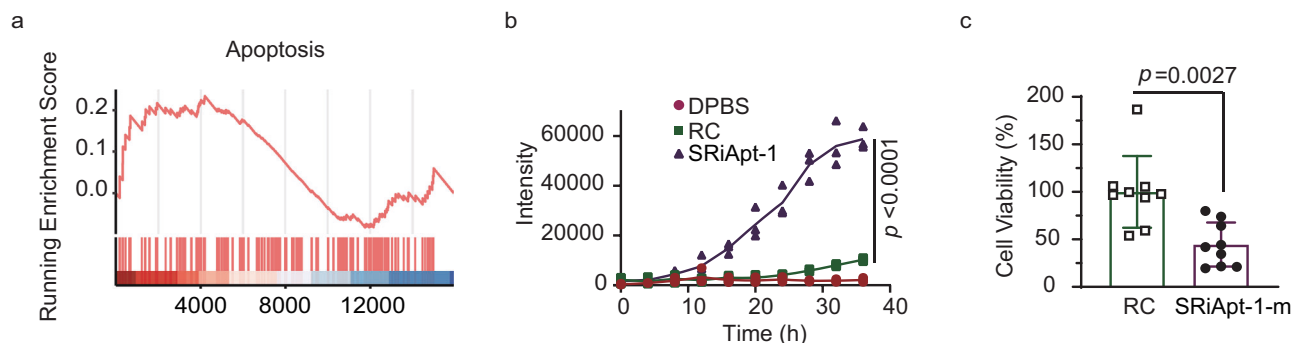


Fig. 6 | Function of the SRiApt-1 aptamer on cell phenotypes. **a** GSEA for KEGG gene sets associated with the apoptosis pathway (NES = 0.83, $p < 0.05$) after treated with SRiApt-1 aptamers. Data were analyzed by one-tailed Fisher's exact test. **b** Representative apoptotic curves of HCT116 cells treated with SRiApt-1 or RC control for 3 days under Annexin V staining. ($n = 3$, mean $\pm$ SD. Two-sided Student's t -test, 95% confidence interval = 41223–55705) Source data are provided as a Source

Data file. **c** Cell viability assessed using the CCK8 assay when cells were treated with SRiApt-1-m or RC aptamers at a working concentration of $1 \mu\text{M}$. The result is determined from three biological replicates, each containing three technical replicates (biological replicates $n = 3$, mean $\pm$ SD, Two-sided Student's t -test). The 95% confidence interval is -78.25 to -19.65 . Source data are provided as a Source Data file.

(Supplementary Fig. 10a). Then we calculated the binding potentials of ssDNA sequences to the interface of WDR5 (Supplementary Fig. 10b). The top-ranked sequences were assessed for their competition to WDR5-MYC interactions via a competitive FP assay. In this assay, we utilized an FITC-labeled MYC peptide and purified WDR5 proteins to establish the initial FP signal (Supplementary Fig. 10c).

In our experimental validation, we initially assessed the top 30 sequences, mirroring the approach for SCAF4-RNAP2 interaction. However, no sequence binding to WDR5 or disrupting the WDR5-MYC interaction was identified. Consequently, we expanded our validation to include additional 10 sequences, totaling the top 40 sequences for further analysis. We found that the 32nd top-scoring sequence, 5'-GGACC-3', exhibited a modest yet significant FP signal reduction, thus being selected as the lead sequence (Fig. 7a). The inhibitory efficacy of the lead sequence was quantified, with an IC_{50} value of $8.50 \pm 4.25 \mu\text{M}$, indicating roughly two-thirds reduction in the FP signal at $10 \mu\text{M}$ (Fig. 7b; Supplementary Fig. 10d). The sequences underwent further optimization by incorporating supporting bases to the lead sequence per our protocol (Supplementary Fig. 10e–g). Through five rounds of iteration using the competitive FP assay, a stem-loop folded sequence, 5'-GGGGGCTGGACCCCTCAACT-3', was identified (Fig. 7c). Secondary structure analysis indicated a stem-loop folding with ΔG of -4.30 kcal/mol . This sequence, designated as WDR5-MYC-iAptamer-1 (WMiApt-1), exhibited a substantially lower IC_{50} of $431.30 \pm 133.61 \text{ nM}$ in disrupting WDR5-MYC interaction (Fig. 7d; Supplementary Fig. 10f). Then, we tried to improve the inhibitory capacity by stabilizing the stem region with four variants with improved ΔG values, yet all exhibited weaker inhibitory capacity, suggesting the stem couldn't accommodate change (Fig. 7e; Supplementary Fig. 10h). The affinity between WMiApt-1 and WDR5 was determined using the FP assay, yielding a value of $797.37 \pm 95.16 \text{ nM}$ (Fig. 7f; Supplementary Fig. 10i). This binding was also validated through SPR with a value of $138.20 \pm 86.73 \text{ nM}$ (Fig. 7g; Supplementary Fig. 10j).

Furthermore, a comprehensive model of WDR5-WMiApt-1 complex was generated by HADDOCK^{45,46}, and optimized through MD simulations, supporting the competitive role of WMiApt-1 in the WDR5-MYC binding (Fig. 7h). To validate this simulation model, three mutations (N225A, V268E and L240K) previously identified critical to the WDR5-MYC interaction were introduced to WDR5⁴⁴ (Fig. 7i). The pivotal roles of these three residues in recognizing MYC were reaffirmed through FP assays; all three mutations attenuated the affinity between WDR5 and MYC (Fig. 7j; Supplementary Fig. 11a–c). Additionally, two residues (Y228, L249) in this region were mutated to alanine, and their impact on recognizing MYC was confirmed; Y228A weakened the interaction between WDR5 and MYC (Fig. 7j;

Supplementary Fig. 11d, e). Compared to the wild-type WDR5, which exhibited an affinity of $210.00 \pm 20.00 \text{ nM}$ to WMiApt-1 (Supplementary Fig. 10c), four of the five residues (N225, V268, Y228 and L249), showing reduced affinities to WMiApt-1, emerged as pivotal for WMiApt-1 binding (Fig. 7j; Supplementary Fig. 11f–j). This validation supports the simulation model, confirming the binding of WMiApt-1 on the WDR5-MYC interface. Although L240K mutation improved the affinity, it is reasonable to posit that despite an overlap between WMiApt-1 and MYC on the WDR5 interface, distinctions in the recognition patterns to WDR5 persist. Subsequently, we explored the binding specificity of WMiApt-1 to WDR5 by investigating its interactions with three unrelated proteins, SA, γSPA and Pepsin (Supplementary Fig. 11k). This analysis revealed no specific interactions with these unrelated proteins, underscoring WMiApt-1's binding specificity to WDR5. These findings collectively highlight the potential of WMiApt-1 aptamer in disrupting WDR5-MYC interaction at the cellular level.

To experimentally validate the inhibitory efficacy of WMiApt-1 on WDR5-MYC interaction within cells, CoIP studies were conducted. By exogenous over-expression of MYC protein in HEK293T cells, the WDR5-MYC complex was captured by anti-MYC antibody but disrupted upon WMiApt-1 treatment (Fig. 8a; Supplementary Fig. 11l). The stability of WMiApt-1 was assessed, revealing a half-life of 2.07 h without ACTA and 12.97 h with ACTA (Supplementary Fig. 11m). This experimental confirmation solidified the inhibitory capabilities of WMiApt-1 in blocking WDR5-MYC interactions in a cellular context.

Previous studies showed that disrupting WDR5-MYC interaction can arrest MYC-driven tumorigenesis⁴⁴. To validate the function and specificity of WMiApt-1 at the cellular level, we selected cell lines with varying levels of MYC and WDR5 expression, including HCT116, MDA-MB-468, and SKNB2, to evaluate their sensitivity to WMiApt-1. Our results revealed a significant correlation between the cytotoxicity of these aptamers and the expression levels of MYC and WDR5. Specifically, HCT116 cells, with relatively higher MYC and WDR5 levels, displayed heightened sensitivity to WMiApt-1 (Fig. 8b, c). These findings strongly support the selectivity of WMiApt-1 in targeting the WDR5-MYC interaction at the cellular level. Furthermore, to investigate the relationship between IC_{50} and binding affinity, we selected a 14-nt intermediate sequence from the third optimization round for comparison. Our analysis revealed that the 20-nt aptamer, WMiApt-1, exhibited higher binding affinity and lower IC_{50} (Supplementary Fig. 11n). Additionally, WMiApt-1 demonstrated the most potent cytotoxic effect on HCT116 cells (Supplementary Fig. 11o). These findings suggest a consistent relationship between IC_{50} s and binding affinities,

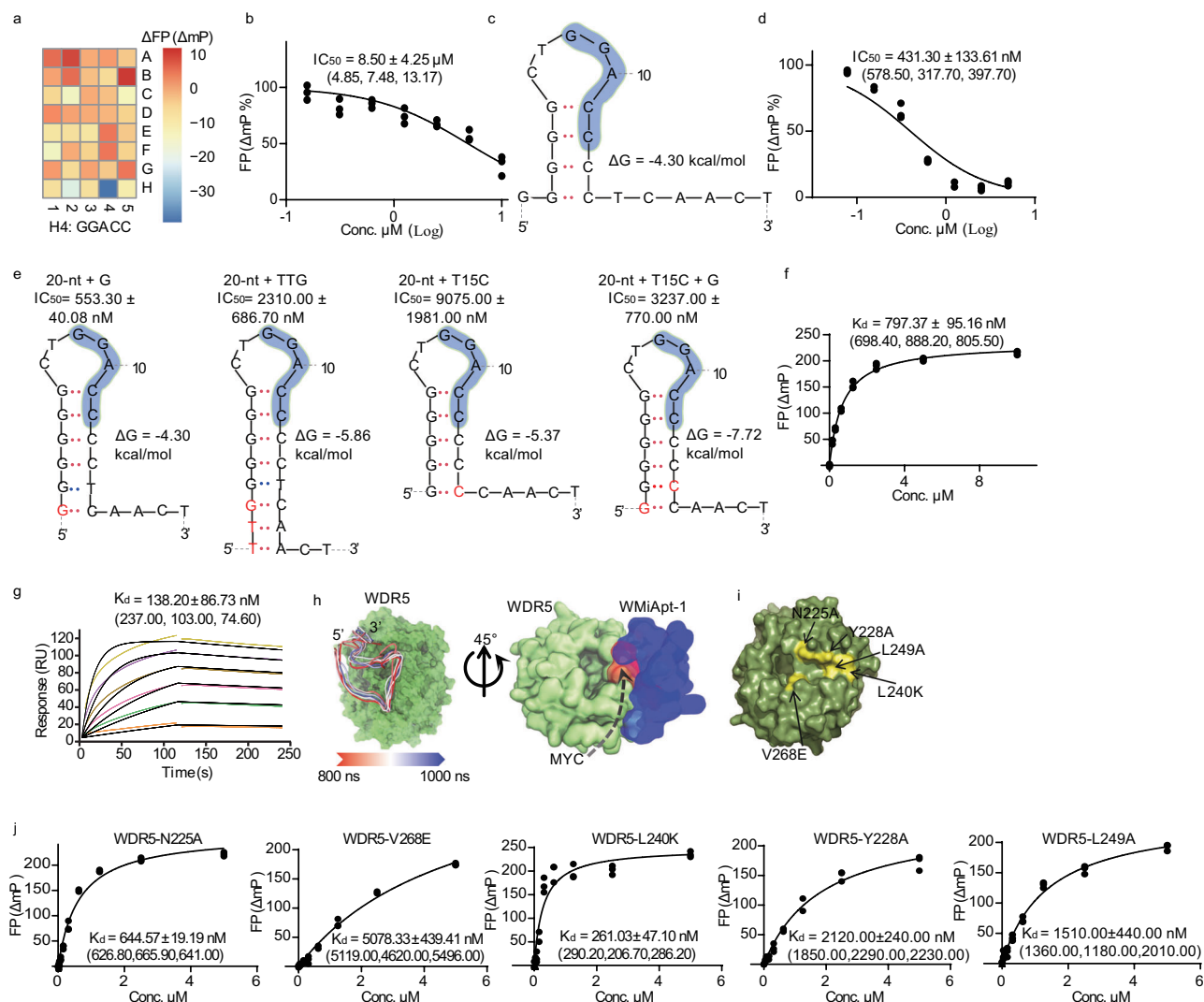


Fig. 7 | Design of inhibitory aptamers targeting the oncogenic WDR5-MYC interaction. **a** Assessments of the inhibitory efficacy of the top 40 sequences on WDR5-MYC interaction. Source data are provided as a Source Data file. **b** Dose-response curves depicting the inhibition of WDR5-MYC interaction by the lead sequence. The IC_{50} is determined from three biological replicates, each containing three technical replicates (biological replicates $n = 3$, mean $\pm$ SD). Source data are provided as a Source Data file. **c** The secondary structure of WMiApt-1 generated by the DNA folding form with a ΔG of -4.30 kcal/mol and melting temperature of $62.8^\circ C$. The lead sequence is highlighted in blue. **d** Inhibition of WMiApt-1 to MYC-WDR5 interaction. The IC_{50} is determined from three biological replicates, each containing three technical replicates (biological replicates $n = 3$, mean $\pm$ SD). Source data are provided as a Source Data file. **e** Summary of IC_{50} s of WMiApt-1 variants in blocking WDR5-MYC interactions. The mutated bases are labeled in red. IC_{50} values are generated from three biological replicates, each containing three technical replicates. Source data are provided as a Source Data file. **f** Fluorescence

polarization measurements of the interaction between WDR5 and WMiApt-1. The affinity is determined as 797.37 ± 95.16 nM with three biological replicates, each containing three technical replicates (biological replicates $n = 3$, mean $\pm$ SD). Source data are provided as a Source Data file. **g** Standard kinetics SPR assay demonstrating the binding of WMiApt-1 to immobilized WDR5. This experiment is performed with three biological replicates. The average affinity value was 138.20 ± 86.73 nM. Affinities of each replicate in the bracket (biological replicates $n = 3$, mean $\pm$ SD). Source data are provided as a Source Data file. **h** A model of WDR5-WMiApt-1 complex with WMiApt-1 color-coded based on the time interval from 800 ns to 1000 ns. Its overlap with MYC is shown with WDR5 in limon, WMiApt-1 in blue, and MYC in red. **i** The locations of five key residues in WDR5-MYC interface in yellow. The interaction between WMiApt-1 and WDR5 mutants is determined with three biological replicates, each containing three technical replicates. Affinities of each replicate in the bracket. Source data are provided as a Source Data file.

further supporting the efficacy of WMiApt-1 in disrupting the WDR5-MYC interaction at the cellular level.

To test the anti-tumor effect of WMiApt-1 in vivo, we premixed 4T1 cells with either WMiApt-1 or RC encapsulated in ACTA were bilaterally implanted into the abdominal regions of the same mice (Fig. 8d, Supplementary Fig. 11p). Notably, by the fifth day post-implantation, 4T1 cells treated with RC exhibited substantial tumor growth, whereas those treated with WMiApt-1 displayed attenuated tumorigenesis (Fig. 8e). This experiment illustrates the potent anti-tumor efficacy of WMiApt-1 in vivo, although further investigations are imperative for its clinical translation and application.

Collectively, the structure-based design strategy, Blocker-SELEX, has facilitated the development of inhibitory aptamers disrupting interactions involving transcription factors, thereby opening avenues for exploring inhibitory aptamer-based therapeutics.

Discussion

Aptamers, distinguished by their several distinctive merits, have emerged as powerful chemical tools with broad applicability. However, the development of aptamers tailored to specific application scenarios remains a formidable challenge. Traditional aptamer-selection strategies, such as SELEX and Cell-SELEX⁴⁷, are labor-intensive and

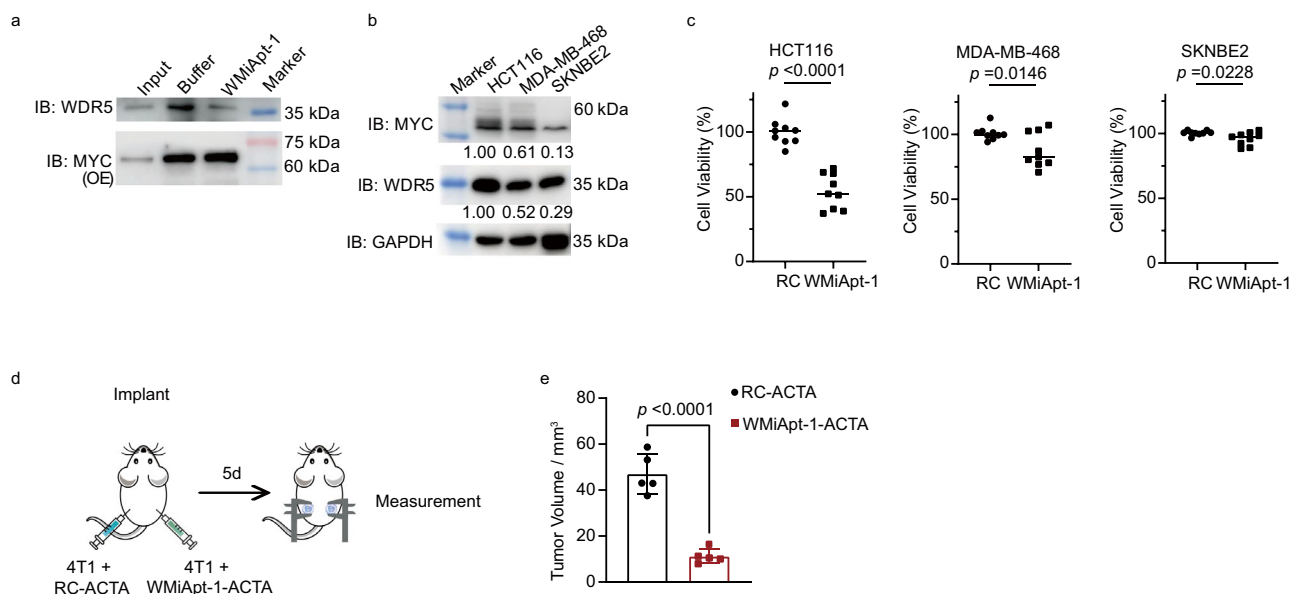


Fig. 8 | Biological characterizations of the WMiApt-1 aptamer. **a** CoIP assays of WMiApt-1 in blocking the interaction between WDR5 and MYC (biological replicates $n = 2$). Source data are provided as a Source Data file. **b** Protein levels of MYC and WDR5 of three human-derived cell lines. The result is determined from three biological replicates, each containing three technical replicates (biological replicates $n = 3$, mean $\pm$ SD, Two-sided Student's t -test). The 95% confidence interval are -56.79 to -36.85 ,

-23.20 to -2.945 and -8.267 to -0.7118 . Source data are provided as a Source Data file. Source data are provided as a Source Data file. **d** Schematic diagram illustrating the anti-tumorigenesis experiments in-vivo. Icons from Canva. **e** Tumor volumes of 4T1 cells treated with WMiApt-1 or control on the fifth day post-implantation (biological replicates $n = 5$, mean $\pm$ SD, Two-sided Student's t -test). The 95% confidence interval is -8.267 to -0.7118 . Source data are provided as a Source Data file.

time-consuming, requiring multiple cycles of affinity selection to obtain high-affinity aptamers. Recent advancements have enhanced SELEX methodologies, including Pro-SELEX, which leverages microfluidic technology to enable the quantitative isolation of aptamers with programmable binding affinities⁴⁸. Additionally, a functional group-guided approach has been proposed for small molecules⁴⁹. Nonetheless, a significant research gap exists in developing methods explicitly designed for aptamers to bind to discrete protein regions and modulate protein-protein interactions.

In this study, we introduce Blocker-SELEX, a structure-guided aptamer design pipeline. This approach efficiently generates competitive iAptamers that selectively bind to desired target interfaces, blocking the associated interactions without relying on computationally intensive processes or accurate 3D structure predictions. Blocker-SELEX assesses the binding potential of individual sequences to the target interface, enabling the scoring of competitive capabilities for each aptamer candidate in terms of target interactions. Moreover, Blocker-SELEX serves as a valuable starting point for constructing potent aptamers within specific binding regions. This reverse process of traditional aptamer SELEX effectively addresses the challenge of identifying essential and nonessential regions of aptamers post-SELEX, facilitating the development of aptamers with enhanced functionality and specificity⁵⁰.

By employing this strategy, we have effectively engineered inhibitory aptamers to disrupt the interactions between RNAP2 and SCAF4 or SCAF8, which were previously considered “undruggable” using small-molecule inhibitors due to the lack of “conventional drug-binding” pockets at the interface of SCAF4/SCAF8. Notably, aptamers exhibit target recognition capabilities that do not necessitate the presence of a “conventional drug-binding” pocket^{51–54}. This makes aptamers ideal chemical tools for intervening in “undruggable” protein-protein/DNA interactions. By utilizing this approach, we validated the feasibility of structure-guided aptamer design and screening. Given the advancements in artificial intelligence within the structural domain and the increasing computational power, it is now foreseeable

we will be capable of de novo designing longer aptamers based on their structural attributes.

An alternative method for modulating “undruggable” transcription factors is Aptamer-based PROTACs^{55,56}. By utilizing aptamers to recognize target proteins without the need for classical drug-binding pockets, this approach overcomes the limitation of lacking available small molecules to degrade “undruggable” transcription factors⁵⁷. Over the past few years, several aptamer-based Targeted Protein Degradation (TPD) strategies have been developed, successfully achieving targeted degradation of proteins such as MYC, TP53 and Nucleolin et al., expanding the spectrum of PROTAC targets⁵⁸. However, for many proteins with multifaceted functional domains involved in regulating various cellular processes, especially in fundamental metabolic pathways, simply modulating their protein levels may yield unintended consequences. Notably, proteins such as P53, GPCRs, and NF- κ B exemplify this complexity, as targeting them for degradation could disrupt vital cellular processes⁵⁹. PPIs represent specific mechanisms through which proteins execute diverse biological functions⁶⁰. Aberrant PPIs are associated with various diseases, spanning cancer, infectious diseases, and neurodegenerative diseases⁶¹. Consequently, targeting PPIs represents a tailored approach specific to distinct protein signaling pathways. In contrast, TPD strategies indiscriminately disrupt all biological functions of the target protein. Therefore, leveraging inhibitory aptamers for targeted modulation of PPIs offers a more precise functional intervention. Ultimately, the choice between aptamer-based PROTACs and direct interference with PPIs necessitates careful evaluation, taking into account the specific biological context and desired therapeutic outcomes.

The instability of aptamers to nuclease degradation presents a significant challenge for their clinical application²⁵. Various chemical modifications and conjugations have been developed to improve the pharmacokinetic properties of aptamer-based therapeutics⁶². Nevertheless, enhancing the stability of inhibitory aptamers remains a critical research direction to sustain their inhibitory with enhanced therapeutic applications.

Overall, Blocker-SELEX offers a significant advantage in designing functional aptamers targeting specific regions of a protein with distinct functions, blocking protein-protein interactions, transcription factor-DNA/RNA interactions, and even extending to the inhibition of enzyme-substrate interactions or acting as allosteric modulators. This capability enables the precise regulation of signal pathways through inhibitory aptamers, thereby presenting a promising avenue for advancements in drug discovery. By harnessing the full potential of the Blocker-SELEX pipeline, our objective is to revolutionize aptamer design and facilitate the creation of tailor-made aptamers for precise applications, thereby expanding their versatility across diverse fields.

Methods

Ethical statement

All experiments were conducted following the guidelines of the Hangzhou Institute of Medicine, Chinese Academy of Sciences, Hangzhou, China. All animal experiments undergo a strict approval process adhere to the use and care guidelines of experimental animals of Zhejiang Province. The License number is SYXK(ZHE)2023-0011.

Virtual screening

A library containing 1024 ssDNA sequences with a length of 5 nucleotides was prepared by JChem software (<http://www.chemaxon.com>) and optimized using the LigPrep module from Schrodinger suite software (release 2019-02)⁶³. Atomic protonation was adjusted to pH 7.0 with the Epik software, and geometric optimization was performed using the OPLS_2005 Force Field software. Protein structures used for docking were retrieved from the Protein Data Bank (SCAF4: 6XKB; SCAF8: 3D9M chain B; WDR5: 4Y7R chain A) (<https://www.rcsb.org/>). Protein preparation was performed with the Protein Preparation Wizard from the Schrödinger platform using default settings to remove all crystallographic water molecules, add hydrogen atoms, assign partial charges, and minimize structures. Targeted docking was carried out on SCAF4, SCAF8 or WDR5 proteins. Grid boxes were set to 25 Å × 25 Å × 25 Å for ligand docking centered at each bound ligand by using the receptor grid generation pipeline. Molecular docking was performed with the ligand docking pipeline with ligand sampling set to flexible in a standard precision scale.

Protein expression and purification

The Human SCAF4-CID domain fragment comprising residues 1-139 and SCAF8-CID domain fragment comprising residues 1-139 were subcloned into a modified pET28-MHL vector, generating N-terminal His-tagged fusion proteins. The hSCAF4-CID mutants were obtained via Quick site-directed mutagenesis utilizing the hSCAF4-CID (residues 1-139) expression construct as the template. Primer sequence (5' → 3'): SCAF4-R23A (F: ATCTCTGCAGCCAAGATGATTCTCATC, R: CTTGGCTGCAGAGATGGGAGGTTTCAT); SCAF4-R71A (F: AGACTGTGCCACAATTGAGTCAATTAC, R: ATTGTGGCAGCTCTCGTCATCAGTTT); SCAF4-R112A (F: ATAGTTGCTGTGCTGAACCTTTGGCAA, R: CAGCACAGCAACTATTTACTCTTATC). The recombinant protein was expressed in BL21 (DE3) cells and induced with 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at 16 °C overnight. Cells were collected via centrifugation, resuspended in lysis buffer containing DPBS buffer, 0.05 mM EDTA, and 5 mM imidazole, and sonicated. The supernatant was obtained after centrifugation at 16,000 × g for 1 h and further purified via the Ni-NTA Beads Gravity Column (Changzhou Smart-Lifesciences Biotechnology Co., China). His-tag was cleaved using TEV Protease, and the protein was subsequently purified using a Superdex75 gel-filtration column (GE Healthcare, Chicago, IL, USA). Finally, purified proteins were concentrated to 2 mg/mL in DPBS buffer supplemented with 5% glycerol. Human WDR5 proteins was tagged with 6xHis-SUMO and expressed in BL21 (DE3) cells upon reaching an optical density of OD = 0.8 via induction with 0.5 mM IPTG at 16 °C for 18 h. Recombinant WDR5 protein was purified to homogeneity using Ni-NTA. The His-SUMO-tag was removed by Thrombin

protease cleavage during dialysis and subsequent subtractive second nickel-column. The WDR5 mutants were obtained via Quick site-directed mutagenesis utilizing the 6xHis-SUMO-tagged WDR5 expression construct as the template. Primer sequence (5' → 3'): WDR5-N225A (F: AGCCCGGCCGCAAATACATCCTGGCA, R: TTTGCCGCGCGGGCTG AATTTCACAAA); WDR5-Y228A (F: GGCAAAGCCATCCTGGCAGTACTC TG, R: CAGGATGGCTTTGCCGTTTCGGGCTGAA); WDR5-L240A (F: CTGA AAGCGTGGGATTACTCTAAGGGT, R: ATCCCACGCTTTCAGAGTGTTA TCCAG); WDR5-L249A (F: AAATGCGCGAAAACCTATACCGCCAT, R: GGTTTTCGCGCATTTACCCTTAGAGTA; WDR5-V268A (F: TTTTCTGCTA CCGGTGGTAAATGGATC, (R: ACCGGTAGCAGAAAAGTTAGCGAAGAT).

Surface plasmon resonance experiments

Lead sequence screening: Identification of lead sequences was carried out using surface plasmon resonance (SPR) analysis, employing a BIAcore 8 K instrument with CM5 chips (GE Healthcare) at ambient temperature (25 °C). Aptamer candidates were diluted to a concentration of 10 μM and then passed over the chip surface to measure response units. Aptamers exhibiting high response levels were deemed as lead sequences with one biological replicates.

Affinity determination: Measurement of binding affinity was carried out using SPR analysis with a BIAcore 8 K instrument and CM5 chips (GE Healthcare) at ambient temperature (25 °C). Aptamers were diluted serially to a range of concentrations, and the analytes were passed over the chip surface to measure response units. Binding kinetics was analyzed using the Biacore Insight Evaluation Software (v5.0.18.22102) with a 1:1 Langmuir binding model. For interactions between aptamers and SCAF4/8 proteins, the SPR buffer consisted of DPBS supplemented with 0.002% (v/v) Tween-20. In the case of interactions between aptamers and the WDR5 protein, the SPR buffer comprised MES buffer (30 mM MES buffer, pH 6.5, 25 mM NaCl, 0.002% (v/v) Tween-20). The affinity is determined using the responses (RU) with three biological replicates.

Fluorescence polarization

Experimental procedures were carried out using a Perkin Elmer EnVision® 2104 Multilabel Reader (BioTek, USA, v1.14) equipped with 485 nm excitation and 535 nm emission filters for the FITC. FP measurements were performed using a Bioland 96-well plate (product #PB06-96S). Each well was loaded with 100 μL assay solution containing FITC-labeled peptides or aptamers (FITC-labeled, 50 μL) in the concentration of 50 nM and proteins (50 μL). Protein samples were serially diluted to various concentrations. After a 30-min incubation period at room temperature, FP measurements were taken. Both parallel and perpendicular fluorescence intensity ($F_{||}$ and $F_{\perp}$) relative to linearly polarized excitation light were determined to calculate the FP signal. The affinity experiments were conducted in three biological replicates, each containing three technical replicates. The average affinity value was determined using the GraphPad Prism 7 program (GraphPad Software, Inc., USA) through curve fitting. The non-linear fit model and the one-site specific binding to fluorescent population ratio are used for these calculations at the respective aptamer concentrations. The FP assay buffer was constituted by 30 mM MES buffer pH 6.5, 25 mM NaCl, 2 mM β-ME, 1 mM CHAPS, and 0.002 mg/mL BSA.

Simulation of LS-SCAF4 complex

The docked complex of SCAF4 and SCAF4_{LS} was used as the initial starting point for the MD (Molecular Dynamics) simulation study for assessing their stability as a complex in terms of Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF) and other parameters. MD Simulation was performed by Gromacs 2021.7 software⁶⁴. CHARMM36m force-field was used for protein and DNA in complex⁶⁵. The initial binding conformation was obtained via template-guided molecular docking using S2, S5-quadruphosphorylated CTD peptide as the template (PDBID:6XKB). Solvent

water, as charmm-tip3p water model, and ions (Na^+ , K^+ and HPO_4^{2-}) were added via Gromacs to emulate PBS buffer. After energy minimization, temperature of the system was heated from 0 K to 298 K via a 500 ps annealing simulation (NPT ensemble) with position restraints on the backbone of protein and DNA. Then, final equilibrium NPT ensemble simulation of 1 ns without any restraint at 298 K was done before production phase simulation. A 500 ns NPT simulation was conducted after equilibrium, and the last 100 ns of trajectory were extracted for post-simulation analysis. RMSD of DNA in the trajectory was calculated by superimposing the protein and comparing the result to the final conformation. The RMSF was calculated in the same way but compared to the average conformation. Temperature and pressure of the system were controlled by Bussi-Parinello (Stochastic) Velocity Rescaling⁶⁶ and Stochastic Cell Rescaling⁶⁴ during all MD simulation processes.

Simulation of WMiApt-1-WDR5 complex

Given that WMiApt-1 is a 20-nt aptamer obtained after multiple rounds of optimization and exhibits a certain degree of structural complexity, the HADDOCK server⁴⁵ was applied to generate the complex structure of WDR5 and WMiApt-1. This involved loading the cleaned WDR5 structure (PDBID: 4R7Y) and WMiApt-1 onto the web server and conducting docking with standard parameters through a hybrid algorithm of template-based and template-free docking. The docked complex structure was used as the initial starting point for the MD (Molecular Dynamics) simulation study for assessing their stability in terms of Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF) and other parameters. MD Simulation was performed using the same protocol as LS-SCAF4 complex simulation. Due to the large complex WDR5-WMiApt, we have performed 1 μs of simulation and the last 200 ns of trajectory was extracted.

Competition assays of inhibitory aptamers

Competition assays of inhibitory aptamers for protein-protein interaction were performed using Fluorescence Polarization (FP) assays. In brief, FP signals resulting from the interaction were detected by combining 40 nM of FITC-labeled peptides with SCAF4/8 or WDR5 proteins in the FP competition buffer (30 mM MES pH 6.0, 25 mM NaCl, 2 mM β -ME, 1 mM CHAPS, and 0.002 mg/mL BSA). Each well of a 96-well plate was loaded with 100 μL of the assay solution. Aptamers were serially diluted to varying concentrations and added to each well for the competition assay. Experiments were performed in three biological replicates, each containing three technical replicates, and the average inhibition constant values were determined by performing curve fitting and data analysis using GraphPad Prism 7 (GraphPad Software, Inc., USA) with the Log(inhibitor) vs. normalized response model.

Inhibitory aptamer screening by competition assay

Experiments were conducted using a Perkin Elmer EnVision® 2104 Multilabel Reader (BioTek, USA) with 485 nm excitation and 535 nm emission filters for the FITC. The competition assay was performed using a Bioland 96-well plate (product #PBO6-96S). To generate the FP signal, 50 nM FITC-labeled peptides or aptamers were mixed with 1 μM target proteins and plated into wells of a 96-well plate at a volume of 100 μL per well. The FP signal was measured for each well, and optimized aptamers were introduced into each well for the competition assay. The FP signal was measured again, and reduction of the FP signal was calculated. Aptamers that exhibited the most significant FP signal reduction were regarded as highly competitive with one biological replicates. The assay buffer was constituted by 30 mM MES buffer pH 6.0, 25 mM NaCl, 2 mM β -ME, 1 mM CHAPS, and 0.002 mg/mL BSA.

Immunoprecipitation

Immunoprecipitation (IP) was performed following the manufacturer's instructions using the BeyoMag™ Streptavidin Magnetic Beads IP Kit.

Biotinylated aptamers, or reverse complementary sequences (30 μg), were mixed with 20 μL of BeyoMag™ Streptavidin Magnetic Beads in a tube containing 20 μL of PBST and incubated at room temperature for 2 h. Biotinylated aptamers or the reverse complementary sequences-loaded beads were then obtained by removing the supernatant on a magnetic separator. To prepare whole cell extracts, cell pellets were lysed using cell lysis buffer containing 50 mM Tris-HCl pH 7.5, 500 mM NaCl, 2 mM EDTA, 0.5% (v/v) Triton X-100, PhosSTOP (Roche, 04906837001) and Protease Inhibitor Cocktail (Roche, 05056489001). Samples were centrifuged, and the supernatants were mixed with the biotinylated immunomagnetic beads to form an immunomagnetic beads-aptamer-antigen complex. Beads were washed three times with PBS and subjected to Western blotting analysis.

Cells transfected with indicated plasmids (SCAF4-CID-3xHA-pCDNA3.1 or MYC-pCDNA3.1) were lysed in cell lysis buffer. The cell lysate was treated with PBS buffer (control) or inhibitory aptamers in 10 μM and incubated with anti-HA antibody (Cell Signaling Technology, 3724 S, Rabbit, 1:50 dilution) or anti-MYC antibody (Sangon, D199941-0100, Mouse, 1:50 dilution) overnight at 4 °C. Normally, 600 μg of total proteins were used for Co-IP in a total volume of 700 μL . Protein A/G conjugated beads were added to precipitate the antibodies. After removing the supernatant and washing with 700 μL of lysis buffer five times, the immunoprecipitates were processed for Western blotting.

Western blotting

Proteins captured on beads were solubilized by adding SDS-loading buffer, subjecting to electrophoresis, and transferring onto a Polyvinylidene Fluoride (PVDF) membrane (Millipore, ISEQ00010). Membranes were subsequently incubated with the appropriate primary antibodies: anti-SCAF4 (Sangon, D153368-0100, Rabbit, 1:1000 dilution), anti-SCAF8 (Abclonal, A19467, Rabbit, 1:1000 dilution), anti-Phospho-RNAP2-S5 (Abclonal, AP0828, Rabbit, 1:1000 dilution), anti-WDR5 (Abclonal, A3259, Rabbit, 1:1000 dilution), anti-GAPDH (HUABIO, ET1601-4, Rabbit, 1:20000 dilution) and anti-MYC antibody (Abclonal, A1309, Rabbit, 1:1000 dilution), followed by incubation with horseradish peroxidase (HRP)-conjugated secondary antibodies (WanleiBio, China, 1:3000 dilution). Proteins of interest were visualized using an enhanced chemiluminescence detection kit (Biosharp, China).

Aptamer-LNPs assembly

The preparation of Aptamer-LNPs was carried out using the ethanol dilution method described elsewhere⁶⁷, using DOTAP, D-Lin-MC3-DMA, DOPE, cholesterol and DMG-PEG2k as the reagents with molar ratio of 50: 25: 5: 19.3: 0.8. Final mass ratio of lipid material to Aptamer was 40:1, and the volume ratio of ethanol to aqueous was 1:3. The assembled LNPs were dialyzed in 1×PBS using a 3500 Da dialysis bag for 2 h, followed by centrifugation at 4000 × g at 4 °C. The concentrated Aptamer-LNPs were then stored in the dark at 4 °C. In the same way, NC-LNPs were prepared by replacing Aptamer with a negative control sequence. ACTA, derived from SM-102, was carried out as described⁶⁸, using SM-102, DSPC, beta-sitosterol, DMG-PEG2000 as the reagents with molar ratio of 50: 10: 39: 1.

Cell proliferation assay

Homo sapiens colon cancer cell line HCT116 was cultured in RPMI-1640 (YEASEN, 41402ES76) with 10% FBS (Sunrise, SR100180.03) and 1% penicillin/streptomycin (Biosharp) and incubated in a 5% CO_2 incubator at 37 °C. HCT116 cell line was generously gifted from Zhejiang Cancer Hospital. For inhibitory aptamer treatment, cells were seeded into a 96-well plate at a density of 15 K cells per well and exposed to 1 μM SRIapt aptamer or NC packed in LNPs for 5 days. Cell viability was determined by Cell Counting Kit-8 assay (Biosharp, BS350B). *Homo sapiens* breast cancer cell line MDA-MB-468 was cultured in DMEM

(YEASEN, 41401ES76) with 10% FBS (Sunrise, SR100180.03) and 1% penicillin/streptomycin (Biosharp) and incubated in a 5% CO₂ incubator at 37 °C. *Homo sapiens* neuroblastoma cell line SKNB2 was cultured in DMEM/F-12 (gibco, C11330500BT) with 10% FBS (Sunrise, SR100180.03) and 1% penicillin/streptomycin (Biosharp) and incubated in a 5% CO₂ incubator at 37 °C. For inhibitory aptamer treatment, cells were seeded into a 96-well plate at a density of 30 K cells per well and exposed to 1 µg/mL WMIapt-1 aptamer or RC packed in ACTA for 36 h. *Mus musculus* breast cancer cell line 4T1 was cultured in RPMI-1640 (YEASEN, 41402ES76) with 10% FBS (Sunrise, SR100180.03) and 1% penicillin/streptomycin (Biosharp) and incubated in a 5% CO₂ incubator at 37 °C. For inhibitory aptamer treatment, cells were seeded into a 96-well plate at a density of 3 K cells per well and exposed to 1 µg/mL WMIapt-1 aptamer or RC packed in ACTA for 3 days. Cell viability was determined by Cell Counting Kit-8 assay (Biosharp, BS350B). Experiments were performed in three biological replicates, each containing three technical replicates (biological replicates $n = 3$, mean $\pm$ SD). Data analysis using GraphPad Prism 7 (GraphPad Software, Inc., USA) with the Two-sided Student's *t*-test. HCT116 (CCL-224) cell line was generously gifted from Zhejiang Cancer Hospital. MDA-MB-468 (HTB-132), HEK293T (ACS-4500), SKNB2 (CRL-2271), 4T1 (CRL-2539) cell lines were purchased from ATCC.

Animal experiments

All animal experiments undergo a strict approval process with the use and care guidelines of experimental animals of Zhejiang Province. The License number is SYXK(ZHE)2023-0011. The maximum tumor volume allowed by the institution is 1500 mm³, and we declare that we do not exceed the maximal tumor size. 2×10^5 viable 4T1 breast cancer cells were premixed with 40 µl of RC-ACTA-PBS or WMIapt-1-ACTA-PBS, resulting a final concentration of 0.5 µg/ml for ssDNA. Those cell mixtures were subcutaneously implanted on the upper left or right side of the same female Balb/c nude mice. Five females Nude mice, SPF level, 4–6 weeks years old, 20 ± 2 g. Mice were housed under social conditions (5 mice per cage) on a standard 12-h dark/12-h light cycle, ambient temperature 21 ± 1 degrees centigrade, and humidity $50\% \pm 10\%$. All mice were housed in a pathogen-free animal facility with standard food and water. Five female Nude mice represent five biological replicates, were included in this experimental setup. On the fifth day post-implantation, tumor volumes were assessed using slide calipers and calculated using the formula: tumor volume (mm³) = length $\times$ (width)² / 2. Data analysis using GraphPad Prism 7 (GraphPad Software, Inc., USA) with the Two-sided Student's *t*-test.

Apoptosis

HCT116 cells (15 K cells/well) were seeded in wells of a 96-well plate and treated with 1 µM SCAF4_F3_11-nt aptamer or NC packed in LNPs for 3 days. Annexin V reagent (Vazyme) was added and analyzed with the Incucyte SX5 live cell imaging device (Sartorius, Germany, v2022REV2). Experiments were performed in three biological replicates (biological replicates $n = 3$, mean $\pm$ SD). Data analysis using GraphPad Prism 7 (GraphPad Software, Inc., USA) with the Two-sided Student's *t*-test.

Confocal imaging

HCT116 cells were seeded and treated with SRiApt-1 or SRiApt-3 aptamer-SM102 in a final concentration of 1 µM for 6 h. After treatment, cells were washed with PBS and then fixed in 4% multigrade formaldehyde (Biosharp) for 1 h at room temperature. Cells were then washed with PBS and blocked in 5% bovine serum albumin (Biosharp) for 1 h at room temperature. Cells were washed with PBS, and DAPI (Beyotime) was added and incubated for 1 min. Images were detected on a single-photon confocal microscope (NIKON, A1 HD25, v5.12) and analyzed by ImageJ (v1.8.0). Experiments were performed in two biological replicates.

RNA-seq analysis

HCT116 cells were subjected to treatment with SRiApt-1 aptamers or NC at a final concentration of 1 µM for 36 h, followed by collection for Bulk RNA-seq. Three batches of samples were prepared for RNA-seq. Library construction and sequencing were conducted by GENEWIZ (GENEWIZ Biotechnology Co., Ltd., Suzhou, China) on an Illumina HiSeq instrument with a 2×150 paired-end (PE) configuration in accordance with the manufacturer's instructions. The reference genome sequences were indexed using Hisat2 (v2.0.1), and clean data were aligned to the reference genome using the same software. The known gff annotation files were converted into transcripts in fasta format and indexed appropriately. With the reference gene file, HTSeq (v0.6.1) estimated gene and isoform expression levels from the pair-end clean data. The DESeq2 package (<https://github.com/mikelove/DESeq2>) was used to perform gene expression analysis and rank the results based on stat value. The Hallmark and KEGG gene sets were obtained from MsigDB (<https://www.gsea-msigdb.org/gsea/msigdb/>). The clusterProfiler R package (v4.2.0) was used to perform GSEA (gene set enrichment analysis), and the enrichplot R package was employed to generate pathway enrichment plots.

Alternative splicing analysis

The input for the analysis consisted of RNA-seq generated bam files aligned using the HISAT2 aligner. The computational tool rMats v4.1.0 was employed to identify differential alternative splicing events from the RNA-seq data. To ensure high confidence, all alternative splicing events were filtered based on an FDR of ≤ 0.05 and an inclusion level difference of > 0.02 . The number of events per comparison was then calculated.

Statistics & reproducibility

Affinity data measured by SPR were subjected to three biological replicates. The affinity and IC₅₀ data obtained by FP were also subjected to three biological replicates, each containing three technical replicates. Cell proliferation assays were subjected to three biological replicates, each containing three technical replicates. RNA-seq was performed in three biological replicates. In the animal experiment, five female nude mice, representing five biological replicates, were included in this experimental setup. No data were excluded from the analyses. The experiments were randomized, and the investigators were blinded to group allocation during experiments and outcome assessment.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All data needed to evaluate the conclusions in the paper are present in the manuscript and/or the Supplementary Materials. The RNA sequencing data generated in this study have been deposited in the Gene Expression Omnibus (GEO) database under accession code [GSE254651](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE254651). Source data are provided as a Source Data with this paper. Protein structures used in this work include SCAF4: PDB ID: [6XKB](https://www.rcsb.org/structure/6XKB), SCAF8: PDB ID: [3D9M](https://www.rcsb.org/structure/3D9M), WDR5: PDB ID: [4Y7R](https://www.rcsb.org/structure/4Y7R). Source data are provided with this paper.

References

- Lambert, S. A. et al. The human transcription factors. *Cell* **172**, 650–665 (2018).
- Ng, A. H. M. et al. A comprehensive library of human transcription factors for cell fate engineering. *Nat. Biotechnol.* **39**, 510–519 (2021).

3. Spitz, F. & Furlong, E. E. M. Transcription factors: from enhancer binding to developmental control. *Nat. Rev. Genet.* **13**, 613–626 (2012).
4. Göös, H. et al. Human transcription factor protein interaction networks. *Nat. Commun.* **13**, 766 (2022).
5. Wang, Z. et al. Interplay between cofactors and transcription factors in hematopoiesis and hematological malignancies. *Signal Transduct. Target Ther.* **6**, 24 (2021).
6. Lee, T. I. & Young, R. A. Transcriptional regulation and its misregulation in disease. *Cell* **152**, 1237–1251 (2013).
7. Zhuang, J.-J., Liu, Q., Wu, D.-L. & Tie, L. Current strategies and progress for targeting the ‘undruggable’ transcription factors. *Acta Pharm. Sin.* **43**, 2474–2481 (2022).
8. Henley, M. J. & Koehler, A. N. Advances in targeting ‘undruggable’ transcription factors with small molecules. *Nat. Rev. Drug Discov.* **20**, 669–688 (2021).
9. Lu, H. et al. Recent advances in the development of protein-protein interactions modulators: mechanisms and clinical trials. *Signal Transduct. Target Ther.* **5**, 213 (2020).
10. Scott, D. E., Bayly, A. R., Abell, C. & Skidmore, J. Small molecules, big targets: drug discovery faces the protein-protein interaction challenge. *Nat. Rev. Drug Discov.* **15**, 533–550 (2016).
11. Zhao, Y., Aguilar, A., Bernard, D. & Wang, S. Small-molecule inhibitors of the MDM2-p53 protein-protein interaction (MDM2 Inhibitors) in clinical trials for cancer treatment. *J. Med. Chem.* **58**, 1038–1052 (2015).
12. Vousden, K. H. & Lu, X. Live or let die: the cell’s response to p53. *Nat. Rev. Cancer* **2**, 594–604 (2002).
13. Groner, B., Weber, A. & Mack, L. Increasing the range of drug targets: interacting peptides provide leads for the development of oncoprotein inhibitors. *Bioengineered* **3**, 320–325 (2012).
14. Leabeater, B. & Kingston, J. Cefuroxime and metronidazole in gynaecological surgery. *Br. J. Obstet. Gynaecol.* **88**, 543–547 (1981).
15. Soucek, L. et al. Omomyc, a potential Myc dominant negative, enhances Myc-induced apoptosis. *Cancer Res.* **62**, 3507–3510 (2002).
16. Speltz, T. E. et al. Targeting MYC with modular synthetic transcriptional repressors derived from bHLH DNA-binding domains. *Nat. Biotechnol.* **41**, 541–551 (2023).
17. Villanueva, M. T. Long path to MYC inhibition approaches clinical trials. *Nat. Rev. Cancer* **19**, 252 (2019).
18. Sorolla, A. et al. Precision medicine by designer interference peptides: applications in oncology and molecular therapeutics. *Oncogene* **39**, 1167–1184 (2020).
19. Socinski, M. A. et al. Atezolizumab for first-line treatment of metastatic nonsquamous NSCLC. *N. Engl. J. Med.* **378**, 2288–2301 (2018).
20. Li, D., Sempowski, G. D., Saunders, K. O., Acharya, P. & Haynes, B. F. SARS-CoV-2 neutralizing antibodies for COVID-19 prevention and treatment. *Annu. Rev. Med.* **73**, 1–16 (2022).
21. Goodman, A., Patel, S. P. & Kurzrock, R. PD-1-PD-L1 immune-checkpoint blockade in B-cell lymphomas. *Nat. Rev. Clin. Oncol.* **14**, 203–220 (2017).
22. Polesso, F. et al. PD-1-specific ‘Blocking’ antibodies that deplete PD-1(+) T cells present an inconvenient variable in preclinical immunotherapy experiments. *Eur. J. Immunol.* **51**, 1473–1481 (2021).
23. Weiner, L. M., Surana, R. & Wang, S. Monoclonal antibodies: versatile platforms for cancer immunotherapy. *Nat. Rev. Immunol.* **10**, 317–327 (2010).
24. Breaker, R. R. Natural and engineered nucleic acids as tools to explore biology. *Nature* **432**, 838–845 (2004).
25. Zhou, J. & Rossi, J. Aptamers as targeted therapeutics: current potential and challenges. *Nat. Rev. Drug Discov.* **16**, 181–202 (2017).
26. Ng, E. W. M. et al. Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. *Nat. Rev. Drug Discov.* **5**, 123–132 (2006).
27. Wang, L. et al. Therapeutic aptamer targeting sclerostin loop3 for promoting bone formation without increasing cardiovascular risk in osteogenesis imperfecta mice. *Theranostics* **12**, 5645–5674 (2022).
28. Kanwar, J. R. et al. Nucleic acid-based aptamers: applications, development and clinical trials. *Curr. Med. Chem.* **22**, 2539–2557 (2015).
29. Mondragón, E. & Maher, L. J. 3rd Anti-transcription factor RNA aptamers as potential therapeutics. *Nucleic Acid Ther.* **26**, 29–43 (2016).
30. Cavasotto, C. N. Homology models in docking and high-throughput docking. *Curr. Top. Med. Chem.* **11**, 1528–1534 (2011).
31. Wondergem, J. A. J., Schiessel, H. & Tompitak, M. Performing SELEX experiments in silico. *J. Chem. Phys.* **147**, 174101 (2017).
32. Hamada, M. In silico approaches to RNA aptamer design. *Biochimie* **145**, 8–14 (2018).
33. Troisi, R. et al. A terminal functionalization strategy reveals unusual binding abilities of anti-thrombin anticoagulant aptamers. *Mol. Ther. Nucleic Acids* **30**, 585–594 (2022).
34. Tesmer, V. M., Lennarz, S., Mayer, G. & Tesmer, J. J. G. Molecular mechanism for inhibition of G protein-coupled receptor kinase 2 by a selective RNA aptamer. *Structure* **20**, 1300–1309 (2012).
35. Someya, T. et al. Crystal structure of Hfq from *Bacillus subtilis* in complex with SELEX-derived RNA aptamer: insight into RNA-binding properties of bacterial Hfq. *Nucleic Acids Res.* **40**, 1856–1867 (2012).
36. Qi, S. et al. Strategies to manipulate the performance of aptamers in SELEX, post-SELEX and microenvironment. *Biotechnol. Adv.* **55**, 107902 (2022).
37. Canoura, J. et al. Accelerating post-SELEX aptamer engineering using exonuclease digestion. *J. Am. Chem. Soc.* **143**, 805–816 (2021).
38. Xu, G. et al. Structure-guided post-SELEX optimization of an ochratoxin A aptamer. *Nucleic Acids Res.* **47**, 5963–5972 (2019).
39. Gregersen, L. H. et al. SCAF4 and SCAF8, mRNA anti-terminator proteins. *Cell* **177**, 1797–1813.e18 (2019).
40. Zhou, M. et al. Structural basis for the recognition of the S2, S5-phosphorylated RNA polymerase II CTD by the mRNA anti-terminator protein hSCAF4. *FEBS Lett.* **596**, 249–259 (2022).
41. Zhu, T. et al. Hit identification and optimization in virtual screening: practical recommendations based on a critical literature analysis. *J. Med. Chem.* **56**, 6560–6572 (2013).
42. Singh, N., Chaput, L. & Villoutreix, B. O. Virtual screening web servers: designing chemical probes and drug candidates in the cyberspace. *Brief. Bioinforma.* **22**, 1790–1818 (2021).
43. Becker, R., Loll, B. & Meinhardt, A. Snapshots of the RNA processing factor SCAF8 bound to different phosphorylated forms of the carboxyl-terminal domain of RNA polymerase II. *J. Biol. Chem.* **283**, 22659–22669 (2008).
44. Thomas, L. R. et al. Interaction with WDR5 promotes target gene recognition and tumorigenesis by MYC. *Mol. Cell* <https://doi.org/10.1016/j.molcel.2015.02.028> (2015).
45. Dominguez, C., Boelens, R. & Bonvin, A. M. J. J. HADDOCK: A protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* <https://doi.org/10.1021/ja026939x> (2003).
46. Navien, T. N., Thevendran, R., Hamdani, H. Y., Tang, T.-H. & Citartan, M. In silico molecular docking in DNA aptamer development. *Biochimie* **180**, 54–67 (2021).
47. Sefah, K., Shangguan, D., Xiong, X., O’Donoghue, M. B. & Tan, W. Development of DNA aptamers using Cell-SELEX. *Nat. Protoc.* **5**, 1169–1185 (2010).
48. Chang, D. et al. A high-dimensional microfluidic approach for selection of aptamers with programmable binding affinities. *Nat. Chem.* **15**, 773–780 (2023).

49. Yang, K. et al. A functional group-guided approach to aptamers for small molecules. *Science* **380**, 942–948 (2023).
50. Gao, S., Zheng, X., Jiao, B. & Wang, L. Post-SELEX optimization of aptamers. *Anal. Bioanal. Chem.* **408**, 4567–4573 (2016).
51. Kim, J. et al. Functional selectivity of insulin receptor revealed by aptamer-trapped receptor structures. *Nat. Commun.* **13**, 6500 (2022).
52. Cheng, E. L. et al. Discovery of a transferrin receptor 1-binding aptamer and its application in cancer cell depletion for adoptive T-cell therapy manufacturing. *J. Am. Chem. Soc.* **144**, 13851–13864 (2022).
53. Grau, F. C., Jaeger, J., Groher, F., Suess, B. & Muller, Y. A. The complex formed between a synthetic RNA aptamer and the transcription repressor TetR is a structural and functional twin of the operator DNA-TetR regulator complex. *Nucleic Acids Res.* **48**, 3366–3378 (2020).
54. Troisi, R., Balasco, N., Santamaria, A., Vitagliano, L. & Sica, F. Structural and functional analysis of the simultaneous binding of two duplex/quadruplex aptamers to human α -thrombin. *Int. J. Biol. Macromol.* **181**, 858–867 (2021).
55. Liu, Y. et al. Aptamer-based targeted protein degradation. *ACS Nano* **17**, 6150–6164 (2023).
56. Wang, C., Zhang, Y., Yu, W., Xu, J. & Xing, D. PROTAC-biomacromolecule conjugates for precise protein degradation in cancer therapy: a review. *Int. J. Biol. Macromol.* **261**, 129864 (2024).
57. Li, Y., Song, J., Zhou, P., Zhou, J. & Xie, S. Targeting undruggable transcription factors with PROTACs: advances and perspectives. *J. Med. Chem.* **65**, 10183–10194 (2022).
58. Bhole, R. P. et al. Unlocking the potential of PROTACs: a comprehensive review of protein degradation strategies in disease therapy. *Bioorg. Chem.* **139**, 106720 (2023).
59. Zanzoni, A., Ribeiro, D. M. & Brun, C. Understanding protein multifunctionality: from short linear motifs to cellular functions. *Cell Mol. Life Sci.* **76**, 4407–4412 (2019).
60. Stelzl, U. et al. A human protein-protein interaction network: a resource for annotating the proteome. *Cell* **122**, 957–968 (2005).
61. Rosell, M. & Fernández-Recio, J. Hot-spot analysis for drug discovery targeting protein-protein interactions. *Expert Opin. Drug Discov.* **13**, 327–338 (2018).
62. Ni, S. et al. Recent progress in aptamer discoveries and modifications for therapeutic applications. *ACS Appl Mater. Interfaces* **13**, 9500–9519 (2021).
63. Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R. & Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput Aided Mol. Des.* **27**, 221–234 (2013).
64. Abraham, M. J. et al. Gromacs: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
65. Huang, J. et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14**, 71–73 (2017).
66. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 14101 (2007).
67. Cheng, Q. et al. Selective organ targeting (SORT) nanoparticles for tissue-specific mRNA delivery and CRISPR-Cas gene editing. *Nat. Nanotechnol.* **15**, 313–320 (2020).
68. Hou, X., Zaks, T., Langer, R. & Dong, Y. Lipid nanoparticles for mRNA delivery. *Nat. Rev. Mater.* **6**, 1078–1094 (2021).

Acknowledgements

We thank the model of organized joint research of HIM. This work was supported by grants from the National Key R&D Program of China (2022YFC3400400 to Y.W.), National Natural Science Foundation of China (No. 32201010 to Y.W., No. 82103287 to Q.W., YQ2022HW01 to Q.W.), Zhejiang Provincial Natural Science Foundation of China (LR22B050001 to Q.W., 2023SDYXS0003 to Q.W., YXD24B0701 to Y.W., 2023SDYXS0002 to Y.W.), and Zhejiang Province Postdoctoral Research Project Merit-based Funding Project (ZJ2022048 to X.L.).

Author contributions

W.T. and Q.W. conceived the project. T.L. and X.L. performed the bulk of experiments. Y.W., Y.T., and M.F. performed virtual screening. Y.Z. and C.S. performed Molecular Simulation and Modeling. A.L., T.L., and K.G. performed the assembly of Aptamer-LNPs. J.Q., W.C., Y.H., G.L., X.Z. helped. X.L., S.Z. performed the bioinformatics analysis. H.Q., Y.W., and H.W. performed chemical synthesis. Y.W., X.L., and T.F. wrote the manuscript. W.T. and Q.W. supervised the work. All authors discussed the results and contributed to the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-51197-w>.

Correspondence and requests for materials should be addressed to Yong Wei, Qin Wu or Weihong Tan.

Peer review information *Nature Communications* thanks Ge Zhang and the other, anonymous, reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024