Probing three-dimensional cyclooctatetraene as nucleobase modification in aptamer selection

Greta Charlotte Dahm,[a] Usman Akhtar,[a,b] Alix Bouvier-Müller,[a] Laura Lim,[a] Fabienne Levi-Acobas,[a] Pierre Nicolas Bizat,[a] Germain Niogret,[a] Julian Tanner,[c] Frédéric Ducongé,[d] and Marcel Hollenstein[a]*

[a] Institut Pasteur, Université Paris Cité, CNRS UMR3523, Department of Structural Biology and Chemistry, Laboratory for Bioorganic Chemistry of Nucleic Acids, 28, rue du Docteur Roux, 75724 Paris Cedex 15, France

[b] Department of Pharmacy, Forman Christian College (A Chartered University), Lahore, Pakistan

[c] School of Biomedical Sciences, LKS Faculty of Medicine, The University of Hong Kong, 21 Sassoon Road, Pokfulam, Hong Kong SAR (P. R. China)

[d] CEA, DRF, Institut of biology JACOB, Molecular Imaging Research Center (MIRCen), Université Paris Saclay, CNRS UMR9199, Fontenay aux roses 92335, France

*To whom correspondence should be addressed. E-mail: marcel.hollenstein@pasteur.fr

Abstract

Decoration of aptamers with chemical modifications at the level of nucleobases grants access to alternative binding modes which often result in improved binding properties. Most functional groups involved in such endeavours mimic the side chains of amino acids or are based on sp^2-dominated moieties. While this approach has met undeniable success, trends in modern drug discovery seem to favor sp^3-rich compounds over aromatic derivatives. Here, we report the use of a nucleotide modified with the three dimensional, highly flexible cyclooctatetraene (COT). This nucleotide was engaged in a SELEX experiment against the biomarker PvLDH. Tightly binding aptamers, coined COTmers, were identified which displayed dissociation constants in the low nM range, representing a significant improvement compared to previously identified cubamers. COTmers clearly underscore the usefulness of COT as a bioisostere replacement of aromatic moieties not only in small compounds but also in functional nucleic acids.

Introduction

Aptamers are single-stranded nucleic acid sequences capable of binding to a broad variety of targets with high specificity and affinity[1-4]. These functional nucleic acids are identified by SELEX (systematic evolution of ligands by exponential enrichment) and related methods of related combinatorial methods of *in vitro* selection[5, 6]. Given their intrinsic properties, aptamers have made a significative impact in numerous fields including nanotechnology, sensing, and therapeutics[7-10]. The tremendous (and increasing) interest is reflected by the FDA-approval of two aptamers for the treatment of age-related macular degeneration (Pegaptanib in 2004 and avacincaptad pegol in 2023)[11]. Despite these favorable assets, aptamers consisting of natural DNA or RNA have only access to a very limited number of functional groups (mainly exocyclic amines), especially when compared to protein antibodies to mediate binding to targets. Hence, aptamers have to resort to this limited array of functional groups combined with $\pi-\pi$ stacking, hydrophobic effects (mainly via the nucleobases), hydrogen bonding or van der Waals interactions[12]. This chemical restriction often leads to failures in SELEX or in the identification of aptamers with poor binding affinity and/or specificity. This is particularly the case when more demanding targets (mainly of hydrophobic and anionic nature[13]) are considered such as proteins with low (i.e. <7) pI values[14-16], highly glycosylated proteins[17-21], intrinsically disordered proteins or with little conformational definition[22-25], or small, hydrophobic molecules[26]. Chemical modification of aptamers, either during the selection protocol (mod-SELEX)[15, 27-31] or after identification of binders (post-SELEX)[32-34], can remediate at least some of these shortcomings. Indeed, the addition of chemical modifications can improve binding affinities (K_D values down to nM and even pM[35]), increase circulation half-lives and nuclease resistance, and convey reactivity that is not accessible to unfunctionalized DNA and RNA. Binding affinity is often increased by adding small, hydrophobic moieties which can mimic hydrophobic contacts found in many protein-protein interactions[22, 36-41]. Illustrative examples are SOMAmers (Slow Off-rate Modified Aptamers)[42] which are aptamers equipped with one[43-45] or multiple[46] nucleobase-modified nucleotides that display impressive binding affinities via significant reduction of k_{off} rates[47]. Nonetheless, the chemical space available to SELEX is limited to a very narrow subset of functional groups, mainly inspired by side chains of amino acids. Alternatively, modifications consisting of two-dimensional sp^2-hydrobized entities are appended to improve stacking interactions and hydrophobic contacts. Lovering et al.[48] emitted the hypothesis that including sp^3 scaffolds could, amongst other benefits, increase receptor/ligand complementarity.[49] In a first step towards an escape of flatland chemistry in the aptamer World, we identified aptamers equipped with cubane modified side chains (Fig. 1), the so-called cubamers.[32, 50] These modified aptamers further validated Eaton's hypothesis that cubane **2** was a true bioisostere of benzene **1**[51] and permitted aptamers to distinguish between two closely related protein targets (i.e. the lactate dehydrogenases from *Plasmodium vivax* (PvLDH) and *Plasmodium falciparum* (PfLDH)). Nonetheless, cubane displays steric bulk and a three-dimensional architecture but evidently lacks π character which might be responsible for the moderate binding affinity of cubamers compared to SOMAmers (K_D value of ~400 nM vs low nM range). Cyclooctatetraene (COT, **3**) is a valence isomer of cubane that displays steric bulk, π character, and a shape shifting equilibrium that transits through a planar, antiaromatic structure of D_{4h} geometry (Fig. 1).[52-57] The non-aromatic COT has also been suggested to be capable of interacting with biomolecules in a "skeleton key" type of mechanism[54], yet few examples of such an interaction have been reported.[58] Herein, we have employed a nucleoside triphosphate equipped with a COT moiety in a SELEX experiment to identify aptamers against the malaria biomarker PvLDH. We have identified three individual sequences, coined COTmers, decorated with COT that bind to the target with very high affinity (K_D values in the

low nM range). This represents a substantial gain in affinity compared to corresponding unmodified and cubane-containing aptamers against the same target. The COTmers also highlight the capacity of COT motifs to engage in interactions with biomolecules and further underscore the usefulness of three-dimensional scaffolds in aptamer selection.

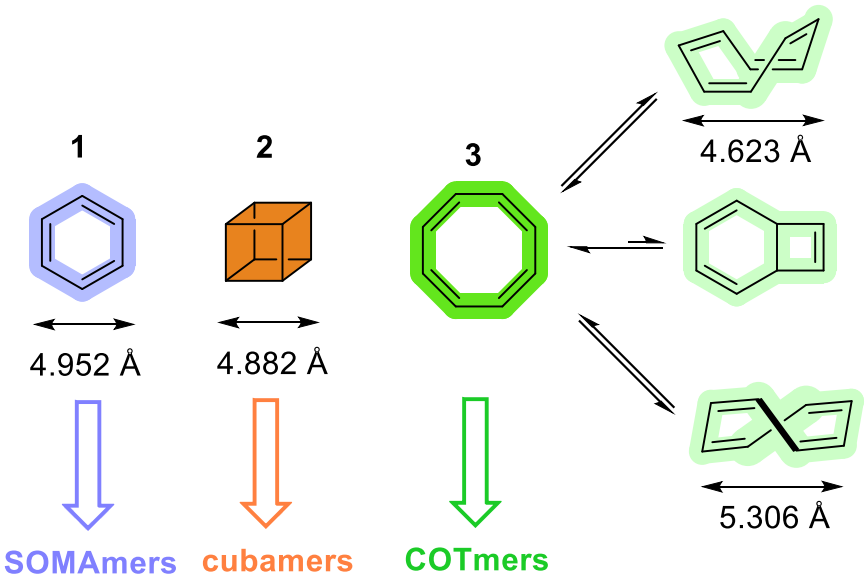


Fig. 1: Chemical structures of benzene (1), cubane (2), and cyclooctatetraene (3) and its valence isomers. Bond distances are given in Å and taken from references [53-55]. The names of aptamers containing these bioisosteres is indicated below each structure.

Results and Discussion

Synthesis and biochemical characterization of the modified nucleotide

By analogy with the cubamer selection process, we designed a deoxyuridine analog equipped with a COT moiety at position C5 of the nucleobase. As a first step in the preparation of the modified nucleotide, we converted 4-methoxycarbonylcubane-1-carboxylic acid **4**[50] into the corresponding valence isomer **5** by application of a rhodium(I) catalyzed reaction (Fig. 2)[53]. COT **5** was then added onto the nucleobase of commercially available amino-11-dUTP by application of standard amide bond reaction conditions (see Supplementary Figs. 5-9 for characterization of compounds).

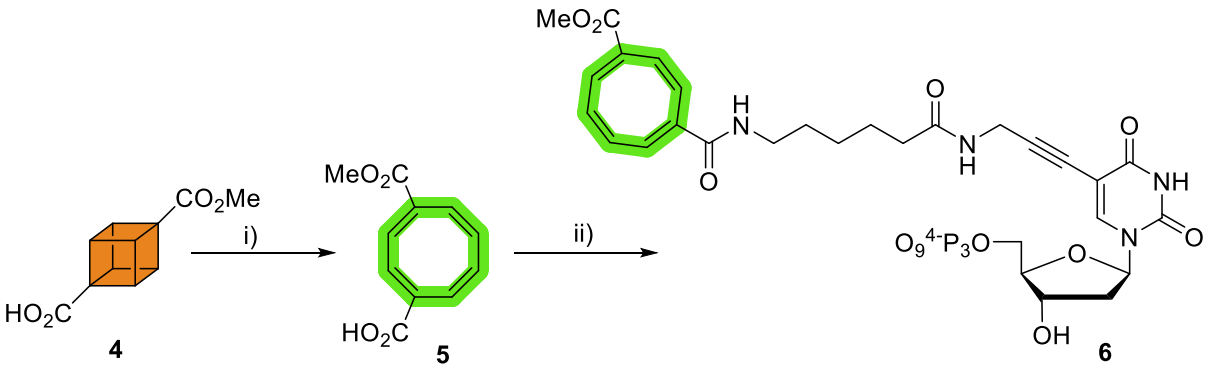


Fig. 2: Synthesis of dU^{COT}TP 6. Reagents and conditions: i) $[Rh(nbd)Cl]_2$, Toluene, 60°C, 20h, 41%; ii) a) HBTU, DIPEA, DMF, RT, 20 min; b) amino-11-dUTP, H_2O, RT, 16h, 8%.

With nucleotide **6** at hand, we evaluated its compatibility with enzymatic DNA synthesis under primer extension (PEX) reaction conditions and PCR. To do so, we carried out PEX reactions on a system consisting of an 18-nt long, 5'-FAM-labelled primer **P1** along with a 71-nt long template **T1** (see Supplementary Table 1 for sequence composition).[59] We included a series of family A (Klenow fragment of *E. coli* DNA polymerase I (Kf (*exo*$^-$)), Taq, Hemo KlenTaq, *Bst*), family B (Phusion, Vent (*exo*$^-$), deep Vent (*exo*$^-$), Therminator, Q5, phi29), and Y family (*Sulfolobus* DNA polymerase IV (Dpo4)) DNA polymerases in different PEX reactions with all dNTPs except for dTTP which was substituted with **dU^{COT}TP 6**. Gel electrophoretic analysis (PAGE 20%) of the reaction products clearly revealed that except for phi29 all polymerases readily accepted the modified nucleotide as a substrate and produced full length products (Fig. 3A). As expected, the modified sequences displayed a lower gel mobility due to the presence of the modifications[36, 59, 60]. Next, we investigated whether nucleotide **6** could also act as a substrate for polymerases under PCR conditions which often reduces the length of the labor-intensive mod-SELEX protocol. To do so, we performed PCR with the 79-mer template **T2**[61] and primers **P2** and **P3** by using five different DNA polymerases to see whether amplicons could be produced when nucleotide **dU^{COT}TP 6** substituted dTTP in the reaction mixture (Fig. 3B). This analysis revealed that nucleotide **6** was also well tolerated as a substrate by a number of polymerases under PCR conditions. Only the reaction catalyzed by Phusion produced a non-identified by-product with a faster electrophoretic mobility.

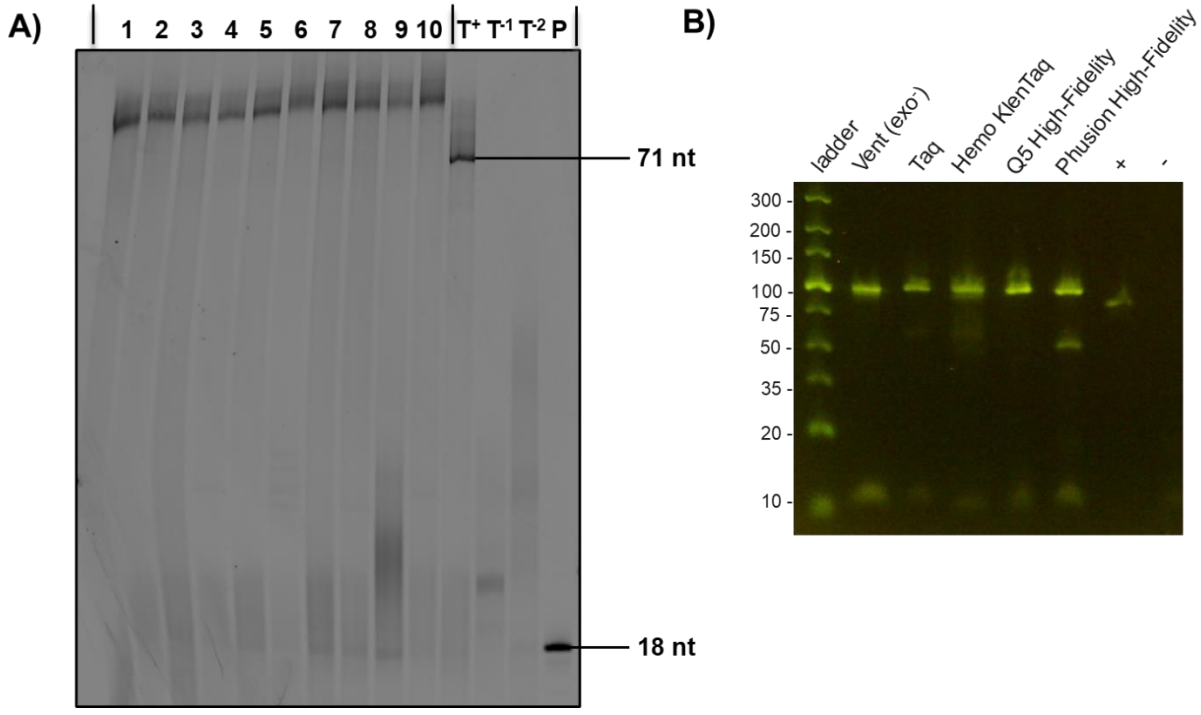


Fig. 3: Biochemical characterization of dU^{COT}TP 6. A) Gel analysis (PAGE 20%) of primer extension reactions. The following types and quantities of polymerases were used: lane 1: Phusion (2 U), lane 2: HemoKlem Taq (8 U), lane 3: Q5 (2 U), lane 4: Bst (8 U), lane 5: Taq (5 U), lane 6: Therminator (2 U), lane 7: Vent (*exo*$^-$) (2 U), lane 8: Dpo4 (2 U), lane 9: Deep Vent (2 U), lane 10: Kf (*exo*$^-$) (5 U). Negative controls: Reaction mixtures containing only dATP and dGTP (T^{-1}) or dATP, dCTP, and dGTP (T^{-2}) and Taq polymerase. Positive control (T^+): with all natural nucleotides and Taq polymerase. All reactions were incubated at adequate reaction temperatures for 1 h in the presence of 200 µM of **dU^{COT}TP 6**. **P** represents unreacted, 5′-FAM-labeled primer. B) Agarose gel (4%) analysis of PCR products obtained with template **T2**,

primers **P2** and **P3**, and a mixture of dATP, dCTP, dGTP, and **dU^{COT}TP 6** (all 200 µM). Control reactions +: PCR with all four natural nucleotides and –: PCR without any polymerase and with all four natural dNTPs.

Preparation of modified library and *in vitro* selection

The biochemical analysis with nucleotide **dU^{COT}TP 6** suggests that libraries for SELEX can be prepared either by PEX reactions or PCR since various polymerases accepted the modified nucleotide as a substrate under these conditions (Fig. 3). Hence, we set out to prepare a modified library suitable for SELEX by PCR. We used a library consisting of a 30 nt long randomized region flanked by two primer binding regions for PCR amplification (see Supplementary Table 1 for sequence composition). Based on the results displayed in Figure 3B, we set out PCR experiments to evaluate the most suitable combination of reaction conditions and polymerase for amplification of the library in the presence of **dU^{COT}TP 6** (Supplementary Fig. 1). PCR reactions under these conditions revealed the formation of faster running bands corresponding to side-products when Q5 was employed but not with Taq polymerase (Supplementary Fig. 1A). After optimization, we obtained good PCR amplification of the library and observed the expected shift in gel mobility compared to a naïve library obtained without modified nucleotides (Supplementary Fig. 1B). Next, we proceeded to identify COT-modified aptamers against PvLDH by application of a modified version of a previously reported protocol[50]. Briefly, after producing a modified dsDNA library, we removed the 5'-phosphorylated template by digestion with λ-exonuclease. The positive selection step included incubation of the resulting ssDNA naïve library with target protein that was not conjugated to Ni-NTA coated magnetic beads. We expected this to favor binding to the target protein[62]. After two hours of incubation, the library-protein complex was immobilized on Ni-NTA coated magnetic particles. Eluted sequences were then reamplified using an on-beads PCR protocol[62]. The stringency of the selection protocol was controlled by including a counter-selection step against empty beads for each round of SELEX and by gradually decreasing the quantity of PvLDH (expressed and purified as described previously[32, 50]). After 10 rounds of SELEX, we evaluated the binding capacity of selected, enriched pools against the protein target with an enzyme-linked oligonucleotide assay (ELONA). This analysis (Fig. 4A) clearly revealed a strong increase in binding of the libraries to PvLDH as the SELEX proceeds, thus suggesting an enrichment of the library with modified species capable of interacting with the target.

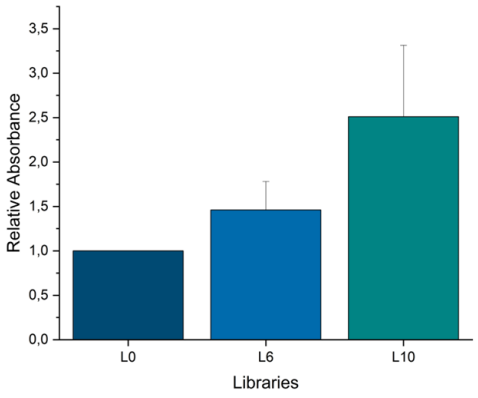
NGS analysis of populations of the SELEX

After ten rounds of selection, the naïve library (L0) and libraries from the 3rd, 6th and 10th rounds were sequenced by NGS to investigate if any enrichment could be observed and confirm the results obtained by ELONA. Around 100,000 reads were sequenced and analyzed per library and data were analyzed as previously described[63]. This analysis first confirmed that **dU^{COT}TP 6** is an excellent substrate for polymerases since in the naïve modified library, the fraction of dT (corresponding to **6**) ranges between 20 and 30% of all bases across the 30 positions of the randomized region (see Supplementary Fig. 4). Furthermore, this analysis also shows that the frequency of dT does not drop much during the progress of the SELEX experiment and without inducing any shrinking of the length of the randomized region since over 90% still contain 30±1 nucleotides in the population of round 10 (see Supplementary Fig. 4 and Supplementary Table 2). Among the 290,547 unique sequences identified, 811 sequences had a frequency that was superior to 0,01% in at least one round. These sequences were retrieved

and grouped into 650 clusters according to a Levenshtein distance of 6, meaning that within one family sequences have a maximum of six mutations compared to the lead sequence. Importantly, a significant enrichment can be observed for several clusters, 82 clusters have a frequency of at least 0,1% in the library of the 10th round, while their frequency is around 100 times lower in the library of the 6th round (see Supplementary Table 3).. For example, cluster 0 accounts for only 0.02% in the library of the 6th round (Fig. 4B) but for over 5% of the sequences in the 10th round library. Moreover, the alignment of the lead sequences of the top 200 clusters shows that different "motifs" are shared by several different clusters (Supplementary Table 4).

We decided to choose 10 clusters and to select their lead sequences to test their capacity to bind the PvLDH protein (Fig. 4C). These clusters were chosen according to two parameters: their enrichment inside the library and the presence of the different motifs. For each motif, we decided to select the best enriched cluster for ELONA binding assays. The majority of these enriched clusters contain between four and seven **dU^{COT}** again suggesting that dTs are not depleted during *in vitro* evolution.

A)



B)

Cluster	Naive library	Round 3	Round 6	Round 10
0	0,0022	0,0000	0,0201	5,2695
1	0,0033	0,0000	0,0222	3,2835
2	0,0022	0,0000	0,0127	2,2667
3	0,0011	0,0000	0,0074	1,7862
4	0,0011	0,0000	0,0180	1,9706
5	0,0000	0,0000	0,0180	1,5342
6	0,0000	0,0000	0,0095	1,1536
7	0,0022	0,0000	0,0074	0,9583
8	0,0000	0,0000	0,0095	0,9862
9	0,0022	0,0000	0,0042	0,8108
10	0,0011	0,0000	0,0159	0,7928
11	0,0011	0,0011	0,0137	0,7352
12	0,0000	0,0000	0,0095	0,7028
13	0,0000	0,0000	0,0095	0,6587
14	0,0000	0,0000	0,0116	0,4949
15	0,0000	0,0000	0,0021	0,4760
16	0,0000	0,0000	0,0095	0,4553
17	0,0000	0,0000	0,0106	0,4229
18	0,0000	0,0000	0,0085	0,4211
19	0,0000	0,0000	0,0095	0,4004
20	0,0000	0,0000	0,0063	0,3590
21	0,0000	0,0000	0,0042	0,3257
22	0,0000	0,0000	0,0063	0,3095
23	0,0011	0,0000	0,0021	0,3068
24	0,0011	0,0000	0,0021	0,2960
25	0,0011	0,0011	0,0053	0,2897
26	0,0000	0,0000	0,0021	0,2664
27	0,0000	0,0000	0,0032	0,2888
28	0,0000	0,0000	0,0085	0,2439
29	0,0000	0,0000	0,0085	0,2430
30	0,0000	0,0000	0,0042	0,2349
31	0,0011	0,0000	0,0032	0,2331
32	0,0000	0,0000	0,0032	0,2268
33	0,0000	0,0011	0,0032	0,4202
34	0,0000	0,0000	0,0011	0,3176
35	0,0000	0,0000	0,0042	0,2124
36	0,0000	0,0000	0,0032	0,2124
37	0,0000	0,0011	0,0021	0,2079
38	0,0000	0,0000	0,0032	0,2052
39	0,0000	0,0000	0,0021	0,1998
40	0,0011	0,0000	0,0053	0,1989
41	0,0000	0,0000	0,0032	0,1980
42	0,0000	0,0000	0,0053	0,1971
43	0,0000	0,0000	0,0011	0,1944
44	0,0000	0,0011	0,0011	0,1800
45	0,0000	0,0011	0,0032	0,2178
46	0,0000	0,0000	0,0042	0,1755
47	0,0011	0,0011	0,0053	0,3590
48	0,0000	0,0000	0,0011	0,1728
49	0,0000	0,0000	0,0021	0,2097
50	0,0000	0,0000	0,0021	0,1620

C)

Random sequence part	Name	Motif
GCA**XX**ACGCAGCGC**X**GCGCAGCGAG**X**ACCG	N0	–
CAG**X**GCAGACGA**X**GC**X**GGG**X**CGCCAGGGA**X**	N1	TGGGTCGCCA
AA**X**CGGAA**XX**CC**X**GAGCCG**X**GCC**X**CCGAG**X**	N2	TGAGCCG
CAAGCGAAG**XX**G**X**GCGCCCGCG**X**CG**X**GG**X**C	N5	–
GACGGAAGAGAACGCCCACAG**X**G**X**C**X**CG**X**C	N6	GCCCACAGTGTC
AGGA**XXX**CC**X**GGACC**X**GCGCGCAGG**X**CAC**X**	N7	TTCCTGGACC
GCCCCG**X**CCG**X**GCAGACA**X**GCACG**XX**CCCC	N10	GCCCCGT
AGG**X**CC**X**C**X**CGCCCCCAG**X**GAC**XX**GCG**X**G	N15	GTCCTCTC
GCCCG**X**AC**X**CG**X**CGG**X**GGACGGCACC**X**C**X**C	N42	GTCGGTGG
AA**X**GCG**X**GCCCCGC**X**GAC**X**CGA**X**GC**X**ACG**X**	N46	GATGCTAC

Fig. 4: Characterization of enrichment of the *in vitro* selection experiment by ELONA and HTS. A) Binding studies of three libraries by ELONA assays. Shown is the average and standard deviation of three ELONA replicates. B) Evolution of Top50 clusters during the SELEX. This figure presents for the Top50 clusters, the frequency of the whole cluster inside each library (in %). For instance, all the sequences belonging to the cluster 0 are representing 0,0201% of the library of round 6, and 5,2695 % of the library of round 10. Clusters colored in light pink are clusters selected for the binding tests. C) Sequences chosen for binding tests based on the NGS analysis. The exact motif present in their sequence is also given. For the sake of clarity, the primer binding regions were omitted. Red, bold-face **X** indicate the position of the modified nucleotide.

Binding affinity determination using flow cytometry

Based on the NGS analysis, we performed a first screening assay of the different aptamer candidates by ELONA (see Supplementary Fig. 3). This analysis revealed that some sequences such as N1, N2, N6, N42, and N46 displayed little (if any) propensity at binding to PvLDH. On the other hand, other sequences such as N0, N10, or N15 might interact strongly with target protein. However, despite many attempts, the error on the outcome of the ELONA assay remained important. Hence, we turned to flow cytometry to evaluate the binding capacity of the aptamer candidates. We hypothesized that flow cytometry would require only low amounts of each modified sequence which is compatible with enzymatic production and abrogates the need for chemical synthesis. Hence, we based our analysis on a recently developed flow cytometry approach for unmodified aptamers [64]. To do so, we first produced suitable 5'-FAM-labbelled, modified ssDNA sequences corresponding to the motifs displayed in Figure 4C using PCR with primers **P1** and **P2** and 5'-phosphorylated templates (See Supplementary Fig. 2 and Supplementary Table 1). The resulting PCR products where then converted to COT-modified ssDNA by λ-exonuclease digestion of the unmodified templates. The resulting sequences were then first incubated with PvLDH in binding buffer for 1 h. The resulting complexes were then subjected to flow cytometry analysis (Fig. 5). This analysis confirmed that most of the sequences identified by ELONA were capable of binding to target protein.

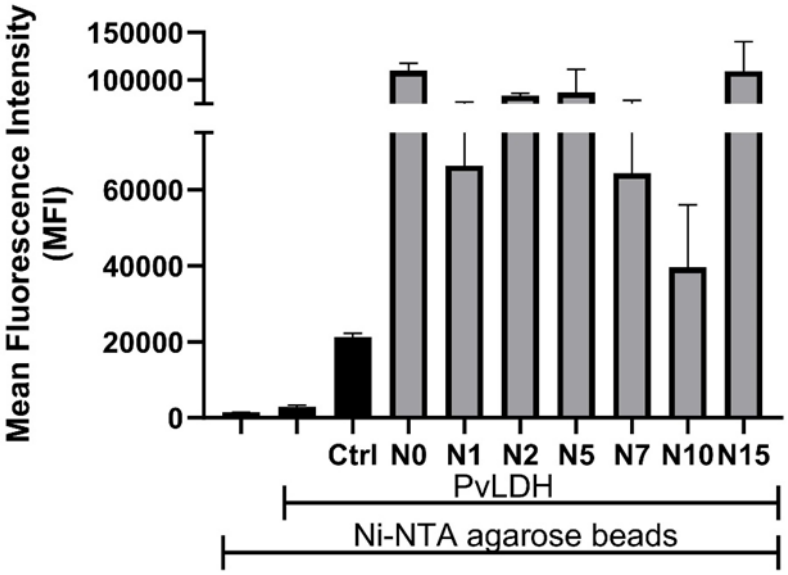


Fig. 5: Flow cytometry binding analysis of aptamer candidates with PvLDH. 10 nM 5'-FAM-labelled sequences along with a control sequence (Ctrl) were incubated with His-tagged PvLDH (10 µg) in binding buffer at room temperature for 1 hour. After washing and preparation of Ni-NTA agarose beads, the beads were incubated with the aptamer-PvLDH solution for 30 minutes and analyzed by flow cytometry using the Attune NxT Flow Cytometer.

Based on this analysis, we further characterized the binding affinity of the most promising candidates, namely N0, N5, and N15. To do so, we prepared modified sequences as well as their unmodified counterparts using PCR with either a mixture of natural and modified dNTPs or only unmodified nucleotides, respectively. We then subjected the resulting sequences to a flow cytometry analysis using a range of concentration (0, 2.5, 5, 10, 20, 60, and 100 nM).

Aptamer N0 displayed the highest affinity for PvLDH with a K_D value of 3.7 ± 0.8 nM while in the range of concentrations that were evaluated the sequence devoid of COT-modified nucleotides did not show a typical binding curve suggesting poor and/or non-specific binding to PvLDH (Fig. 6A). A similar pattern was observed with aptamer N15 since a K_D value of 5.6 ± 1.7 nM was determined and a non-typical binding curve was observed with the unmodified sequence with a K_D value estimate of 208 nM (Fig. 6C). Aptamer N5 displayed a slightly reduced binding affinity compared to N0 and N15 (K_D value of 14.4 ± 7.9 nM) but the unmodified sequence seems to be poorly binding to target (Fig. 6B). Overall, all three aptamers (or COTmers), N0, N5, and N15 displayed significantly improved (> 150-fold) binding affinities compared to a previously identified cubamer. All aptamers strongly depend on the presence of the modified nucleotides to interact with the target protein.

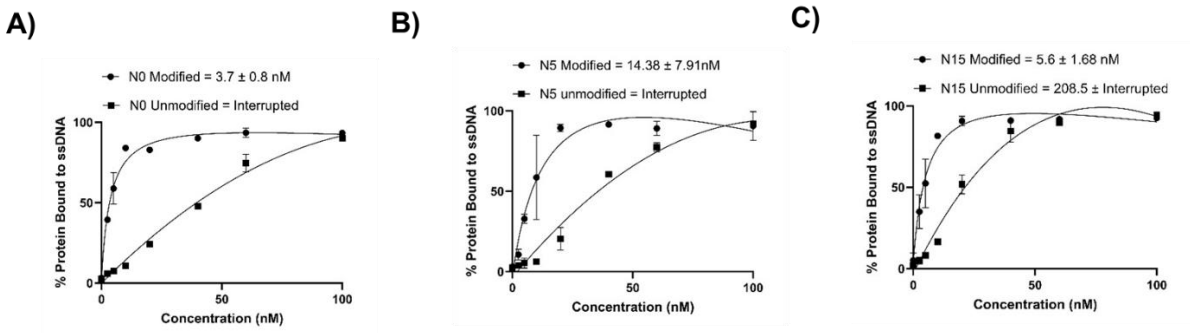


Fig. 6: Dissociation constant (K_D) determination of aptamers. Various concentrations of 5'-FAM-labeled modified and unmodified aptamers corresponding to sequences A) N0, B) N5, and C) N15 were used to evaluate binding interactions. The binding curves for modified aptamers were analyzed using GraphPad Prism with nonlinear regression (curve fit) based on a one-site total binding model. The calculated K_D values and their 95% confidence intervals were used for standard deviation calculations. K_D values for the unmodified aptamers were found to be undefined, preventing the calculation of a mean value during the analysis.

We next sought to determine whether COTmers N0, N5, and N15 retained the specificity for PvLDH displayed by the cubamer. To do so, we performed a similar flow cytometry binding analysis with the modified sequences N0, N5, and N15 but with PfLDH instead of PvLDH (Fig. 7). This analysis revealed that the sequence N15 displayed a similar propensity at interacting with both proteins. A similar result was observed with N5, while N0 binds significantly less to PfLDH than PvLDH but is still capable of interacting with both proteins.

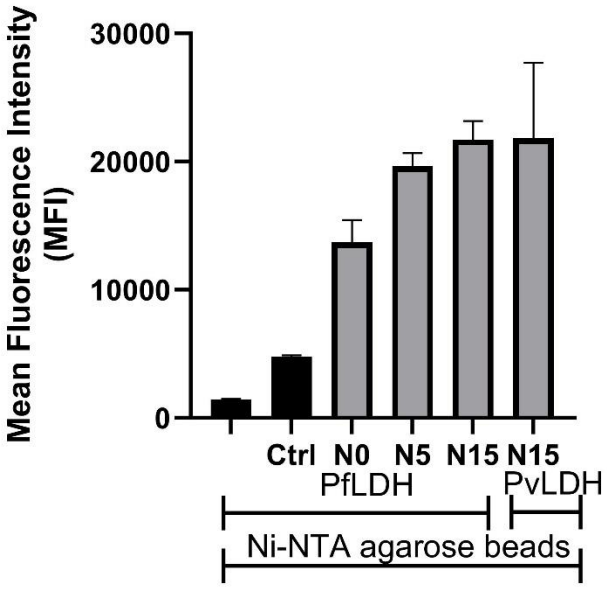


Fig. 7: Flow cytometry binding analysis of aptamers with PfLDH. 5' FAM-labeled aptamers N0, N5, and N15 (10 nM) along with a control sequence (Ctrl) were incubated with His-tagged PfLDH (10 µg) in binding buffer at room temperature for 1 hour. After washing and preparation of Ni-NTA agarose beads, the beads were incubated with the aptamer-PfLDH solution for 30 minutes and analyzed by flow cytometry using the Attune NxT Flow Cytometer. PvLDH was used as a control for comparison of binding interaction of aptamers with PfLDH.

Discussion

We have used *in vitro* selection to raise aptamers modified with cyclooctatetraene moieties which adopts a three-dimensional, saddle-like architecture. The resulting COTmers were raised against the malaria biomarker PvLDH to provide direct comparison with previously identified cubamers. After ten rounds of SELEX, three COTmers (N0, N5, and N15) were identified that do not share any sequence homology with the cubamer but bear a similar amount (N0 and N5) or slightly more (N15) modified nucleotides. The COTmers displayed a superior binding affinity for the same target than the cubamer (100 to 150-fold improvement) and unmodified DNA aptamers (2-10 fold improvement) [65]. This surge in binding avidity might be a direct consequence from chemical and structural differences between cubane and COT. Even though both structural motifs are non-planar and display similar lipophilicities [53, 54], cubane only contains sp^3-hybridized carbons and COT has a strong π-character. In addition, cubane adopts a rather rigid cage structure while COT can navigate between various structural conformations (Fig. 1). This combination of steric bulk, π-stacking capacity, and conformational flexibility make COT an even better phenyl ring bioisostere than cubane. In addition to these effects, the differential nature of the linker connecting the nucleobase to COT/cubane might have an incidence on the binding affinity of the resulting aptamers. While no specific studies have been dedicated to this topic, it is believed that longer and less rigid linker arms reduce the efficiency of functional nucleic acids [37, 66-68]. This would imply that the structural and chemical differences between COT and cubane might be strong enough to counterbalance the negative impact of the longer and more flexible linker arm present in **dU^{COT}TP 6**. Taken together, in addition to improving the bioactivity of small, pharmaceutical and agrochemical compounds[53, 54, 58], the COT also provides aptamers with higher binding affinities. This is

evidenced by the substantial improvement of dissociation constants of COTmers compared to cubamers and unmodified DNA aptamers.

On the other hand, COTmers seem to have lost the capacity of cubamers to discriminate PvLDH from PfLDH (90% amino acid identity). This lack of specificity might arise from the absence of a negative counterselection step including incubation with PfLDH which we have used for the identification of the cubamer. Additionally, due to ring strain cubyl hydrogen atoms are ~10^5 more acidic than those present on phenyl rings and comparable to that of NH_3[69, 70]. This acidity enables cubane derivatives to engage in non-classical C-H_{cubane}···O hydrogen bonding interactions[70]. Such a hydrogen bonding interaction was observed in the crystal structure of the cubamer binding to PvLDH. Importantly, this C-H_{cubane}···O interaction between a cubane and the carbonyl of Leu232 was believed to participate in the discrimination capacity of the cubamer. Such a hydrogen bonding capacity is obliterated by swapping cubane with COT modifications and might be partially responsible for the lack of specificity.

Future work will encompass structural elucidation of the PvLDH-COTmer complex. This will permit to shed light into the binding mechanism of the COTmers. Importantly, such work will also determine the conformational preference of the COT moieties in this context. Indeed, it is currently unclear which geometry the COT substituents adopt upon binding to the target. It is likely that the COT adopt the nonplanar geometry (D_{2d} symmetry) of the ground state but other conformations (e.g. planar antiaromatic D_{4h} or the bicyclic valence isomer) could be imposed upon binding to the protein [71, 72]. Hence, in addition to prompting binding to targets, we foresee that endowing aptamers with other exotic functional groups combined with structural resolution could provide insights into long-standing questions on the conformation adopted by organic compounds such as COT, annulenes [73], or boroles [57, 74].

Conclusions

Darwinian evolution combined with modified nucleotides represents an alluring strategy to improve the properties of functional nucleic acids. The potency of this approach is showcased by SOMAmers which are capable of binding to targets with K_D values in the low nM/high pM range [22, 42, 44, 46] and DNAzymes capable of cleaving amide bonds[75] or hydrolyzing RNA in the absence of M^{2+}-cofactors [76, 77]. Nonetheless, most nucleotides that have been engaged in SELEX experiments are endowed with amino acid-like residues or flat, aromatic moieties. On the other hand, modern trends in drug discovery and medicinal chemistry tend to favor sp^3-rich compounds[48, 78]. Following this precept, we report herein the synthesis of a nucleotide modified with a cyclooctatetraene (COT) moiety and its application to *in vitro* selection. The COT substitution allowed to identify highly potent aptamers coined COTmers that bind to the malaria biomarker PvLDH with low nM binding affinity. This represents a significant gain in binding affinity compared to a previously reported cubamer and unmodified DNA aptamers. Taken together, these results indicate the beneficial effect of nucleotides modified with substituents capable of adopting three dimensional conformations in aptamer selection. We also further demonstrate the usefulness of the COT motif in bioactive molecule discovery. Future work will encompass structural elucidation of the COTmer-PvLDH complex and SELEX with two nucleotides equipped with three dimensional substituents.

Methods

Chemical syntheses

Detailed protocols for the synthesis of all nucleoside and nucleotide analogs can be found in the Supporting Information of this article.

Protocol for primer extension (PEX) reactions

5′-FAM-labelled primer **P1** (10 pmol) was hybridized with the corresponding template **T1** (15 pmol) in DNase/RNase-free ultrapure water. This was achieved by elevating the temperature to 95°C and then allowing it to gradually cool down to room temperature over an hour. Subsequently, DNA polymerase (0.5 to 1 µL), suitable reaction buffer, and the required dNTP(s) were added to yield a 10 µL reaction mixture. This mixture underwent incubation at the polymerase-specific optimal temperature for given times. The reactions were quenched by adding 10 µL of a solution containing formamide (70%), EDTA (50 mM), bromophenol (0.1%), and xylene cyanol (0.1%). The resulting reaction mixtures were analyzed by gel electrophoresis in a denaturing 20% polyacrylamide gel, complemented with 1× TBE buffer (pH 8) and urea (7 M). Visualization of PAGE gels was performed by fluorescence imaging using a Typhoon Trio phosphorimager from Cytiva.

Protocol for PCR

The PCR mixtures were obtained by adding primers **P2**/**P3** (6 µM each), template **T2** (0.1 µM), modified and natural dNTPs (200 µM), polymerase (0.4 µL), and polymerase buffer in a total volume of 20 µL. PCR cycles were dependent on the nature of the DNA polymerase: denaturation at 95°C for 30 s, annealing at 55°C for 30 s, and elongation at 72°C for 60 s (Vent (*exo*$^-$) DNA Polymerase, Hemo KlenTaq, and Taq DNA Polymerase) or denaturation at 98°C for 10 s, annealing at 61°C for 30 s, and elongation at 72°C for 60 s (Phusion High-Fidelity DNA Polymerase and Q5 High-Fidelity DNA Polymerase). After PCR amplification with 25 cycles, the reaction products were analyzed by 4% agarose gels, supplemented with 1× E-GEL sample loading buffer (loading: 1 to 5 pmol).

Protocol for SELEX of modified aptamers

A ssDNA library with a 30 nucleotide long randomized region was used to prepare a naïve library for SELEX using PCR with primers **P2**/**P3** (6 µM each), Taq as polymerase, and in the presence of **dU^{COT}TP 6**. After removing the 5'-phosphorylated template by digestion with λ-exonuclease, we carried out a counter-selection step. To do so, the library (100 pmol) was incubated with Ni-NTA Magnetic Agarose Beads (from Jena Bioscience) for 30 min at room temperature. The unbound fraction was recovered and incubated with free *Pv*LDH for 2 h at room temperature in binding buffer (100 mM NaCl, 5 mM $MgCl_2$, 25 mM Tris-HCl, pH 8.0). The resulting protein-library mixture was added to fresh Ni-NTA Magnetic Agarose Beads and incubated for 30 min at room temperature. The supernatant was discarded, and the beads were washed 3 X 100 µL binding buffer before being suspended in 100 µL. We then carried out on-beads PCR amplification. To do so, the bound sequences were amplified using Taq DNA Polymerase (0.5 U/µL), a mixture of natural dATP, dCTP, and dGTP (each at 75 µM) and **dU^{COT}TP 6** (75 µM), along with Taq Standard buffer, $MgSO_4$ (2 mM), forward primer (**P2**, 5 µM) and 5'phosphorylated reverse primer (**P3**, 5 µM). The PCR cycles consisted of a program that started at 95°C for 30 s, then at 55°C for 30 s, and finally at 72°C for 60 s. The number of cycles was adjusted after each round of selection and varied from 8 to 11 cycles. Following a

purification with MinElute® PCR Purification Kit (QIAGEN), the phosphorylated strand of the dsDNA library was digested with Lambda Exonuclease (NEB) to yield an ssDNA library that can be used in a subsequent round of SELEX. The stringency of the SELEX was increased by decreasing the amount of protein over the selection rounds (for rounds 1 and 2 we used 150 µg of protein, then 100 µg of protein for rounds 3 to 7 and finally 50 µg of protein for rounds 8 to 10).

Protocol for ELONA

For ELONA binding tests, DNA libraries of rounds 0, 6, and 10 (10 nM each) or templates corresponding to individual sequences were amplified using *Taq* DNA Polymerase (0.5 U/µL), natural dATP, dCTP, and dGTP (each at 75 µM) and **dU^{COT}TP 6** (75 µM), 5′-biotinylated forward primer P4 and 5′-phosphorylated reverse primer **P3** (0.5 µM each). The following PCR conditions were used: denaturation at 95°C for 30 s, annealing at 55°C for 30 s, and elongation at 72°C for 60 s for 10 cycles. After purification with the MinElute® PCR Purification Kit (QIAGEN), the phosphorylated strands were digested with λ-exonuclease (NEB). In parallel, His_6-*Pv*LDH (2 µg/mL, 100 µL) was incubated in Ni-NTA HisSorb™ Plates (QIAGEN) wells for 1.5 h at room temperature. Wells were washed four times with PBS (0.05% Tween 20). The modified, 5'-biotinylated ssDNA libraries or individual sequences (50 nM, 50 µL) were then incubated in the protein coated wells for 1 h at room temperature. After washing three times with binding buffer (100 mM NaCl, 5 mM $MgCl_2$, 25 mM Tris-HCl, pH 8.0), streptavidin-HRP (50 µL, Abcam) was added to the wells and allowed to interact for 25 min at room temperature. Wells were then washed twice with the binding buffer. Three minutes after the addition of tetramethylbenzidine (50 µL, SigmaAldrich) the reaction was stopped by adding H_2SO_4 (1 M, 50 µL). Absorption was measured at 450 nm.

Next-generation sequencing (NGS)

For this SELEX, aliquots of the library from the naïve library and rounds 3, 6 and 10 were analyzed by NGS on a iSeq 100 Sequencing System (Illumina) as previously described [63]. Approximately 100,000 sequencing reads were analyzed for each round of SELEX. using several home-made scripts that were used sequentially to analyze the results and generate the corresponding graphs (Excel and GraphPad Prism). Briefly, the primer binding sites were removed to recover only the sequences corresponding to the randomized region. Sequences having a randomized region ranging between 25 and 32 nucleotides in between the primer binding sites were recovered because it is very common for sequences to undergo deletions or insertions of a few nucleotides during SELEX. The frequency of each sequence in each round was then calculated. All sequences with a frequency of at least 0.01% in one round were retrieved and clustered into families based on a Levenstein distance of 6 (i.e. all sequences with less than six substitutions, deletions or insertions are grouped into the same family). The frequency of each cluster was then calculated for each round (Supplementary Table S3). Multiple alignment of theTop200 clusters was performed by MultAlin[79] and the conservation of these motifs was analyzed using MEME (Supplementary Table S4) [80].

Protocol for Flow Cytometry Binding Analysis

To study the binding interaction of aptamers with PvLDH, 10 nM of 5'-FAM-labelled modified aptamer sequences (obtained by PCR) were mixed with 10 µg of His-tagged PvLDH in the binding buffer (25 mM Tris, pH 8.0, 100 mM NaCl, 5 mM $MgSO_4$) and incubated for 1 hour at

room temperature. Meanwhile, Ni-NTA agarose beads were washed three times with 200 µl of binding buffer. After the final centrifugation (500 x g, 5 min), the beads were resuspended in the binding buffer and incubated with the aptamer-PvLDH solution for 30 minutes at room temperature to immobilize the His-tagged PvLDH. The beads were then washed three times with 200 µl of binding buffer, and after the final centrifugation, the supernatant was removed. The beads were resuspended in the buffer for FACS analysis using the Attune NxT Flow Cytometer. The data were analysed using FlowJo software.

For the determination of the dissociation constant (K_D) of the aptamers, various concentrations of 5'-FAM-labelled modified and unmodified aptamers (0, 2.5, 5, 10, 20, 60, and 100 nM) were used, following the same protocol for bead preparation and analysis as described above. GraphPad Prism was employed to analyse the binding curves using nonlinear regression (curve fit) based on a one-site total binding model. The K_D value calculated with the 95% confidence intervals was used to determine the standard deviation (±). The K_D values for the unmodified aptamers were undefined, preventing the calculation of a mean value during the analysis.

Data availability

The authors declare that all data supporting the findings of this study are available within the article and the Supplementary Information.

Author Contributions

G.C.D., L.L., G.N., and P.N.B. synthesized the modified nucleotides and G.C.D. and F.L.A. carried out the biochemical characterization of the nucleotides. G.C.D. carried out the SELEX experiment with an important contribution from F.L.A.. A.B.M. and F.D. carried out the sequencing of the libraries and the bioinformatic analysis. L.L., U.A., and P.N.B. synthesized the modified and natural sequences. U.A. carried out the binding studies by flow cytometry. J.T. and M.H. designed the study, and M.H. analyzed the results and wrote the paper. All authors (G.C.D., U.A., A.B.M., L.L., F.L.A., P.N.B., G.N., J.T., F.D., and M.H.) have read and approved the final version of the manuscript.

Acknowledgements

The authors gratefully acknowledge financial support from Institut Pasteur. P.N.B. and M.H. acknowledge funding from the CNRS/ANR grant PEPR MolecularXiv (ANR-22-PEXM-0002). A.B.M. and M.H. acknowledge funding via a grant from the Emergency COVID-19 Fundraising Campaign of Institut Pasteur (project AptaCoV). U.A. is thankful for funding from Ligue Contre le Cancer. G. N. gratefully acknowledges a fellowship from the doctoral school MTCI from Université Paris Cité.

References

1. Ji, C. et al. Aptamer–Protein Interactions: From Regulation to Biomolecular Detection. *Chem. Rev.* **123**, 12471-12506 (2023).
2. Dunn, M.R., Jimenez, R.M. & Chaput, J.C. Analysis of aptamer discovery and technology. *Nat. Rev. Chem.* **1**, 0076 (2017).
3. Dong, Y.H. et al. Aptamer-based assembly systems for SARS-CoV-2 detection and therapeutics. *Chem. Soc. Rev.* **53**, 6830-6859 (2024).
4. Zhang, Y. & Li, Y. Clinical Translation of Aptamers for COVID-19. *J. Med. Chem.* **66**, 16568-16578 (2023).
5. Tuerk, C. & Gold, L. Systematic Evolution of Ligands by Exponential Enrichment: RNA Ligands to Bacteriophage T4 DNA Polymerase. *Science* **249**, 505-510 (1990).
6. Ellington, A.D. & Szostak, J.W. In vitro selection of RNA molecules that bind specific ligands. *Nature* **346**, 818-822 (1990).
7. Bouvier-Müller, A. et al. Aptamer binding footprints discriminate α-synuclein fibrillar polymorphs from different synucleinopathies. *Nucleic Acids Res.* **52**, 8072-8085 (2024).
8. Lu, W., Lou, S., Yang, B., Guo, Z. & Tian, Z. Light-activated oxidative capacity of isoquinoline alkaloids for universal, homogeneous, reliable, colorimetric assays with DNA aptamers. *Talanta* **279**, 126667 (2024).
9. Wen, X. et al. Development of an aptamer capable of multidrug resistance reversal for tumor combination chemotherapy. *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2321116121 (2024).
10. Wang, B. et al. Functional Selection of Tau Oligomerization-Inhibiting Aptamers. *Angew. Chem. Int. Ed.* **63**, e202402007 (2024).
11. Gao, J. et al. Unlocking the Potential of Chemically Modified Nucleic Acid Therapeutics. *Adv. Therap.* **7**, 2400231 (2024).
12. Hermann, T. & Patel, D.J. Adaptive Recognition by Nucleic Acid Aptamers. *Science* **287**, 820-825 (2000).
13. Manimala, J.C., Wiskur, S.L., Ellington, A.D. & Anslyn, E.V. Tuning the Specificity of a Synthetic Receptor Using a Selected Nucleic Acid Receptor. *J. Am. Chem. Soc.* **126**, 16515-16519 (2004).
14. Blanco, C., Bayas, M., Yan, F. & Chen, I.A. Analysis of Evolutionarily Independent Protein-RNA Complexes Yields a Criterion to Evaluate the Relevance of Prebiotic Scenarios. *Curr. Biol.* **28**, 526-537.e525 (2018).
15. Renders, M., Miller, E., Lam, C.H. & Perrin, D.M. Whole cell-SELEX of aptamers with a tyrosine-like side chain against live bacteria. *Org. Biomol. Chem.* **15**, 1980-1989 (2017).
16. Ahmad, K.M. et al. Probing the limits of aptamer affinity with a microfluidic SELEX platform. *PloS one* **6**, e27051 (2011).
17. Li, M. et al. Selecting Aptamers for a Glycoprotein through the Incorporation of the Boronic Acid Moiety. *J. Am. Chem. Soc.* **130**, 12636-12638 (2008).
18. Díaz-Martínez, I., Miranda-Castro, R., de-los-Santos-Álvarez, N. & Lobo-Castañón, M.J. Lectin-Mimicking Aptamer as a Generic Glycan Receptor for Sensitive Detection of Glycoproteins Associated with Cancer. *Anal. Chem.* **96**, 2759-2763 (2024).
19. Li, W., Ma, Y., Guo, Z., Xing, R. & Liu, Z. Efficient Screening of Glycan-Specific Aptamers Using a Glycosylated Peptide as a Scaffold. *Anal. Chem.* **93**, 956-963 (2021).
20. Díaz-Fernández, A. et al. Aptamers targeting protein-specific glycosylation in tumor biomarkers: general selection, characterization and structural modeling. *Chem. Sci.* **11**, 9402-9413 (2020).

21. Yoshikawa, A.M. et al. Discovery of indole-modified aptamers for highly specific recognition of protein glycoforms. *Nat. Commun.* **12**, 7106 (2021).
22. Rohloff, J.C. et al. Nucleic Acid Ligands With Protein-like Side Chains: Modified Aptamers and Their Use as Diagnostic and Therapeutic Agents. *Mol. Ther. Nucleic Acids* **3**, e201 (2014).
23. Murakami, K., Izuo, N. & Bitan, G. Aptamers targeting amyloidogenic proteins and their emerging role in neurodegenerative diseases. *J. Biol. Chem.* **298** (2022).
24. Li, T. et al. Blocker-SELEX: a structure-guided strategy for developing inhibitory aptamers disrupting undruggable transcription factor interactions. *Nat. Commun.* **15**, 6751 (2024).
25. Wu, D. et al. Flow-Cell-Based Technology for Massively Parallel Characterization of Base-Modified DNA Aptamers. *Anal. Chem.* **95**, 2645-2652 (2023).
26. Rosenthal, M., Pfeiffer, F. & Mayer, G. A Receptor-Guided Design Strategy for Ligand Identification. *Angew. Chem. Int. Ed.* **58**, 10752-10755 (2019).
27. Latham, J.A., Johnson, R. & Toole, J.J. The application of a modified nucleotide in aptamer selection: novel thrombin aptamers containing -(1 -pentynyl)-2′-deoxyuridine. *Nucleic Acids Res.* **22**, 2817-2822 (1994).
28. Tarasow, T.M., Tarasow, S.L. & Eaton, B.E. RNA-catalysed carbon–carbon bond formation. *Nature* **389**, 54-57 (1997).
29. Vaish, N.K., Larralde, R., Fraley, A.W., Szostak, J.W. & McLaughlin, L.W. A Novel, Modification-Dependent ATP-Binding Aptamer Selected from an RNA Library Incorporating a Cationic Functionality. *Biochemistry* **42**, 8842-8851 (2003).
30. Gordon, C.K.L. et al. Click-Particle Display for Base-Modified Aptamer Discovery. *ACS Chem. Biol.* **14**, 2652-2662 (2019).
31. Minagawa, H. et al. A high affinity modified DNA aptamer containing base-appended bases for human β-defensin. *Anal. Biochem.* **594**, 113627 (2020).
32. Niogret, G. et al. Interrogating Aptamer Chemical Space Through Modified Nucleotide Substitution Facilitated by Enzymatic DNA Synthesis. *ChemBioChem* **25**, e202300539 (2024).
33. Majumdar, B., Sarma, D., Yu, Y., Lozoya-Colinas, A. & Chaput, J.C. Increasing the functional density of threose nucleic acid. *RSC Chem. Biol.* **5**, 41-48 (2024).
34. Eaton, B.E. et al. Post-SELEX combinatorial optimization of aptamers. *Bioorg. Med. Chem.* **5**, 1087-1096 (1997).
35. Kimoto, M. et al. Strict Interactions of Fifth Letters, Hydrophobic Unnatural Bases, in XenoAptamers with Target Proteins. *J. Am. Chem. Soc.* **145**, 20432-20441 (2023).
36. Mulholland, C. et al. The selection of a hydrophobic 7-phenylbutyl-7-deazaadenine-modified DNA aptamer with high binding affinity for the Heat Shock Protein 70. *Commun. Chem.* **6**, 65 (2023).
37. Kohn, E.M. et al. Terminal Alkyne-Modified DNA Aptamers with Enhanced Protein Binding Affinities. *ACS Chem. Biol.* **18**, 1976-1984 (2023).
38. Kong, D., Yeung, W. & Hili, R. In Vitro Selection of Diversely Functionalized Aptamers. *J. Am. Chem. Soc.* **139**, 13977-13980 (2017).
39. Lozoya-Colinas, A., Yu, Y. & Chaput, J.C. Functionally Enhanced XNA Aptamers Discovered by Parallelized Library Screening. *J. Am. Chem. Soc.* **145**, 25789-25796 (2023).
40. Paul, A.R., Falsaperna, M., Lavender, H., Garrett, M.D. & Serpell, C.J. Selection of optimised ligands by fluorescence-activated bead sorting. *Chem. Sci.* **14**, 9517-9525 (2023).
41. Lichtor, P.A., Chen, Z., Elowe, N.H., Chen, J.C. & Liu, D.R. Side chain determinants of biopolymer function during selection and replication. *Nat. Chem. Biol.* **15**, 419-426 (2019).
42. Gold, L. et al. Aptamer-Based Multiplexed Proteomic Technology for Biomarker Discovery. *PloS one* **5**, e15004 (2010).

43. Ren, X., Gelinas, A.D., von Carlowitz, I., Janjic, N. & Pyle, A.M. Structural basis for IL-1α recognition by a modified DNA aptamer that specifically inhibits IL-1α signaling. *Nat. Commun.* **8**, 810 (2017).
44. Ren, X. et al. Evolving A RIG-I Antagonist: A Modified DNA Aptamer Mimics Viral RNA. *J. Mol. Biol.* **433**, 167227 (2021).
45. Gelinas, A.D. et al. Broadly neutralizing aptamers to SARS-CoV-2: A diverse panel of modified DNA antiviral agents. *Mol. Ther. Nucleic Acids* **31**, 370-382 (2023).
46. Gawande, B.N. et al. Selection of DNA aptamers with two modified bases. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 2898-2903 (2017).
47. Gelinas, A.D., Davies, D.R. & Janjic, N. Embracing proteins: structural themes in aptamer–protein complexes. *Curr. Opin. Struct. Biol.* **36**, 122-132 (2016).
48. Lovering, F., Bikker, J. & Humblet, C. Escape from Flatland: Increasing Saturation as an Approach to Improving Clinical Success. *J. Med. Chem.* **52**, 6752-6756 (2009).
49. Tsien, J., Hu, C., Merchant, R.R. & Qin, T. Three-dimensional saturated C(sp3)-rich bioisosteres for benzene. *Nat. Rev. Chem.* **8**, 605-627 (2024).
50. Cheung, Y.W. et al. Evolution of abiotic cubane chemistries in a nucleic acid aptamer allows selective recognition of a malaria biomarker. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 16790-16798 (2020).
51. Chalmers, B.A. et al. Validating Eaton's Hypothesis: Cubane as a Benzene Bioisostere. *Angew. Chem. Int. Ed.* **55**, 3580-3585 (2016).
52. Houston, S.D. et al. Cyclooctatetraenes through Valence Isomerization of Cubanes: Scope and Limitations. *Chem. Eur. J.* **25**, 2735-2739 (2019).
53. Houston, S.D. et al. The cubane paradigm in bioactive molecule discovery: further scope, limitations and the cyclooctatetraene complement. *Org. Biomol. Chem.* **17**, 6790-6798 (2019).
54. Xing, H. et al. Cyclooctatetraene: A Bioactive Cubane Paradigm Complement. *Chem. Eur. J.* **25**, 2729-2734 (2019).
55. Liu, Y. et al. Cubane and Cyclooctatetraene Pirfenidones – Synthesis and Biological Evaluation. *Asian J. Org. Chem.* **12**, e202300238 (2023).
56. Li, L. et al. Stabilizing a different cyclooctatetraene stereoisomer. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 9803-9808 (2017).
57. Lavendomme, R. & Yamashina, M. Antiaromaticity in molecular assemblies and materials. *Chem. Sci.* **15**, 18677-18697 (2024).
58. Xing, H. et al. In search of herbistasis: COT-metsulfuron methyl displays rare herbistatic properties. *Chem. Sci.* **16**, 649-658 (2025).
59. Niogret, G. et al. A toolbox for enzymatic modification of nucleic acids with photosensitizers for photodynamic therapy. *RSC Chem. Biol.* **5**, 841-852 (2024).
60. Spampinato, A. et al. ABNOH-Linked Nucleotides and DNA for Bioconjugation and Cross-linking with Tryptophan-Containing Peptides and Proteins. *Chem. Eur. J.* **30**, e202402151 (2024).
61. Niogret, G. et al. Facilitated Synthetic Access to Boronic Acid-Modified Nucleoside Triphosphates and Compatibility with Enzymatic DNA Synthesis. *Synlett* **35**, 677-683 (2024).
62. Pantier, R., Chhatbar, K., Alston, G., Lee, H.Y. & Bird, A. High-throughput sequencing SELEX for the determination of DNA-binding protein specificities in vitro. *STAR Protoc.* **3**, 101490 (2022).
63. Nguyen Quang, N., Bouvier, C., Henriques, A., Lelandais, B. & Ducongé, F. Time-lapse imaging of molecular evolution by high-throughput sequencing. *Nucleic Acids Res.* **46**, 7480-7494 (2018).
64. Civit, L. et al. A Multi-Faceted Binding Assessment of Aptamers Targeting the SARS-CoV-2 Spike Protein. *Int. J. Mol. Sci.* **25**, 4642 (2024).
65. Choi, S.-J. & Ban, C. Crystal structure of a DNA aptamer bound to PvLDH elucidates novel single-stranded DNA structural elements for folding and recognition. *Sci. Rep.* **6**, 34998 (2016).

66. Vaught, J.D. et al. Expanding the Chemistry of DNA for in Vitro Selection. *J. Am. Chem. Soc.* **132**, 4141-4151 (2010).
67. Hipolito, C.J., Hollenstein, M., Lam, C.H. & Perrin, D.M. Protein-inspired modified DNAzymes: dramatic effects of shortening side-chain length of 8-imidazolyl modified deoxyadenosines in selecting RNaseA mimicking DNAzymes. *Org. Biomol. Chem.* **9**, 2266-2273 (2011).
68. Hollenstein, M. Nucleic acid enzymes based on functionalized nucleosides. *Curr. Opin. Chem. Biol.* **52**, 93-101 (2019).
69. Hare, M., Emrick, T., Eaton, P.E. & Kass, S.R. Cubyl Anion Formation and an Experimental Determination of the Acidity and C−H Bond Dissociation Energy of Cubane. *J. Am. Chem. Soc.* **119**, 237-238 (1997).
70. Kuduva, S.S., Craig, D.C., Nangia, A. & Desiraju, G.R. Cubanecarboxylic Acids. Crystal Engineering Considerations and the Role of C−H···O Hydrogen Bonds in Determining O−H···O Networks. *J. Am. Chem. Soc.* **121**, 1936-1944 (1999).
71. Wu, J.I., Fernández, I., Mo, Y. & Schleyer, P.v.R. Why Cyclooctatetraene Is Highly Stabilized: The Importance of “Two-Way” (Double) Hyperconjugation. *J. Chem. Theory Comput.* **8**, 1280-1287 (2012).
72. Deslongchamps, G. & Deslongchamps, P. Bent Bonds (τ) and the Antiperiplanar Hypothesis—The Chemistry of Cyclooctatetraene and Other C8H8 Isomers. *J. Org. Chem.* **83**, 5751-5755 (2018).
73. Deslongchamps, G. & Deslongchamps, P. Bent Bond/Antiperiplanar Hypothesis and the Chemical Reactivity of Annulenes. *J. Org. Chem.* **85**, 8645-8655 (2020).
74. Braunschweig, H. & Kupfer, T. Recent developments in the chemistry of antiaromatic boroles. *Chem. Commun.* **47**, 10903-10914 (2011).
75. Zhou, C. et al. DNA-Catalyzed Amide Hydrolysis. *J. Am. Chem. Soc.* **138**, 2106-2109 (2016).
76. Perrin, D.M., Garestier, T. & Hélène, C. Bridging the gap between proteins and nucleic acids:: A metal-independent RNAseA mimic with two protein-like functionalities. *J. Am. Chem. Soc.* **123**, 1556-1563 (2001).
77. Hollenstein, M., Hipolito, C.J., Lam, C.H. & Perrin, D.M. A self-cleaving DNA enzyme modified with amines, guanidines and imidazoles operates independently of divalent metal cations (M^{2+}). *Nucleic Acids Res.* **37**, 1638-1649 (2009).
78. Lyon, W.L., Wang, J.Z., Alcázar, J. & MacMillan, D.W.C. Aminoalkylation of Alkenes Enabled by Triple Radical Sorting. *J. Am. Chem. Soc.* **147**, 2296-2302 (2025).
79. Corpet, F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **16**, 10881-10890 (1988).
80. Bailey, T.L. & Elkan, C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **2**, 28-36 (1994).