Evolution of abiotic cubane chemistries in a nucleic acid aptamer allows selective recognition of a malaria biomarker

Yee-Wai Cheung, Pascal Röthlisberger, Ariel E Mechaly, Patrick Weber, Fabienne Levi-Acobas, Young Lo, Alvin W C Wong, Andrew B Kinghorn, Ahmed Haouz, G Paul Savage, et al.

Main Manuscript for Professor Chad A. Mirkin

Evolution of abiotic cubane chemistries in a nucleic acid aptamer allows selective recognition of a malaria biomarker

Yee-Wai Cheung[1,5], Pascal Röthlisberger[2,5], Ariel E. Mechaly[3], Patrick Weber[3], Fabienne Levi-Acobas[2], Young Lo[1], Alvin W.C. Wong[1], Andrew B. Kinghorn[1], Ahmed Haouz[3], G. Paul Savage[4], Marcel Hollenstein[2,*] and Julian A. Tanner[1,*]

[1]School of Biomedical Sciences, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong, China.

[2]Institut Pasteur, Laboratory for Bioorganic Chemistry of Nucleic Acids, Department of Structural Biology and Chemistry, CNRS UMR-3523, 28, rue du Docteur Roux, 75015 Paris, France.

[3]Institut Pasteur, Crystallography Platform-C2RT, Department of Structural Biology and Chemistry, CNRS UMR-3528, 25-28 rue du Docteur Roux, 75015, Paris, France.

[4] CSIRO Manufacturing, Bayview Avenue, Victoria, 3168, Australia.

[5]These authors contributed equally: Yee-Wai Cheung and Pascal Röthlisberger

*These authors contributed equally: Marcel Hollenstein and Julian A. Tanner.

Email: marcel.hollenstein@pasteur.fr; jatanner@hku.hk

ORCIDs of authors: Pascal Röthlisberger: 0000-0001-7947-2489; Ahmed Haouz: 0000-0003-1196-1635; Ariel Mechaly: 0000-0002-5305-7495; G. Paul Savage: 0000-0001-7805-8630; Marcel Hollenstein: 0000-0003-0263-9206; Julian Tanner: 0000-0002-5459-1526

Classification

Major: Physical Sciences; minor: Chemistry.

Keywords

Modified aptamer; SELEX; X-ray crystallography; cubane; malaria diagnosis

Author Contributions

J.A.T., M.H., Y.W.C. and P.R. were responsible for the overall study design and data interpretation. Y.W.C. performed all cubamer selections, characterized the cubamers and oversaw all application experiments. A.W.C.W., Y.L., and A.B.K. were involved in aspects of cubamer characterization and in assay design. F.L.-A. expressed and purified the proteins and performed the biochemical characterization of the triphosphate. P.R. performed all chemical syntheses and was involved in aspects of experimental design of cubamer structural analysis. P.R. was also significantly involved in all data interpretation. G.P.S. was involved in steps of synthesis relating to cubanes and advised on study design. P.R., A.E.M., P.W., A.H. carried out the crystallogenesis and crystallographic studies. J.A.T., M.H., Y.W.C. and P.R. drafted the paper with critical feedback from all co-authors. All authors have read, edited and approved the final version of the manuscript.

This PDF file includes:

Main Text

Figures 1 to 6

Abstract

Nucleic acid aptamers selected through systematic evolution of ligands by exponential enrichment (SELEX) fold into exquisite globular structures in complex with protein targets with diverse translational applications. Varying the chemistry of nucleotides allows evolution of non-natural nucleic acids, but the extent to which exotic chemistries can be integrated into a SELEX selection to evolve non-natural macromolecular binding interfaces is unclear. Here, we report the identification of a cubane-modified aptamer (cubamer) against the malaria biomarker *Plasmodium vivax* lactate dehydrogenase (PvLDH). The crystal structure of the complex reveals an unprecedented binding mechanism involving a multi-cubane cluster within a hydrophobic pocket. The binding interaction is further stabilized through hydrogen bonding via cubyl hydrogens, previously unobserved in macromolecular binding interfaces. This binding mechanism allows discriminatory recognition of *Plasmodium vivax* over *Plasmodium falciparum* lactate dehydrogenase thereby distinguishing these highly conserved malaria biomarkers for diagnostic applications. Together, our data demonstrate that SELEX can be used to evolve exotic nucleic acids bearing chemical functional groups which enable novel binding mechanisms which have never been observed in biology. Extending to other exotic chemistries will open a myriad of possibilities for functional nucleic acids.

Significance Statement

We wish to report the identification of a cubane-modified aptamer (cubamer) against the malaria biomarker *Plasmodium vivax* lactate dehydrogenase (PvLDH). The cubamer contains the benzene isostere cubane which is entirely alien to biology. We have resolved the crystal structure of the cubamer-protein complex which reveals a novel binding mechanism involving the formation of a cubane-pocket and an usual C-H···O hydrogen bond. Importantly, the cubamer is capable of distinguishing the PvLDH from the PfLDH despite a very high sequence homology which is not possible for unmodified aptamers. Lastly, we have used the cubamer for the detection of PvLDH in a mimetic clinical situation. This approach blending medicinal chemistry and Darwinian evolution can easily be extended to other non-natural, exotic functional groups.

Main Text
Introduction

Directed evolution of nucleic acids through the iterative process of SELEX(1, 2) (selective evolution of ligands by exponential enrichment) provides routes to aptamers with exquisite structures that cannot be otherwise imagined or designed(3). Even using the simple natural nucleotides of DNA and RNA one can select and evolve aptamers with nanomolar and even picomolar binding affinities and surprising discriminatory specificity(4). Such aptamers are finding applications in therapeutics(5), diagnostics(6), biosensing(7), and nanotechnology(8) and can display numerous advantages relative to antibodies(9). The chemistries of canonical DNA and RNA can be limiting when one compares to antibodies(10). Post-selection chemical modification can aid *in vivo* stability and residence time(11, 12), and can facilitate binding to problematic targets such as glycoproteins or proteins with low isoelectric points(13-15). However, these post-SELEX approaches often come at the expense of a loss of binding affinity and specificity. In contrast, pre-selection chemical modification – where modified nucleotides are directly included in the SELEX process – truly opens up directed evolution to extending the repertoire of aptamer-mediated molecular recognition(16-19). Many chemical modifications have been directed at the level of the sugar-phosphate backbone in order to improve the nuclease resistance of aptamers(18, 20-23). Other approaches include aptamers with artificially expanded genetic information systems (AEGIS) going beyond the simple four building blocks of natural nucleic acids(24), even most recently to the eight building block Hachimoji system(25). Nucleobases have also been modified but have often been restricted to mimics of amino acid side chains(13, 26, 27) or small hydrophobic aromatic moieties(28, 29). Perhaps the most impressive demonstration of translational cross-proteome aptamer specificity are the SOMAmers (Slow Off-rate Modified Aptamers)(30), which enabled quantification of over 3000 plasma proteins to create a genomic atlas of the human plasma proteome(31). Such SOMAmers include a simple modification where instead of dT they bear dU functionalized at the 5-position of the nucleobase with protein-like benzyl/2-napthyl/3-indolyl-carboxamide that confer some

hydrophobic and/or base stacking character to the nucleic acid(10). New, engineered polymerases are not required to cope with unnatural chemistry introduced at position 5 of pyrimidines(32). Base modified nucleotides have also been proposed as elements in an RNA world to promote the catalytic activity of RNA based catalysts(33). Alternatively, click chemistry with alkyne-modified nucleotides can be used to select for nucleobase-modified aptamers(34). Through such approaches, nucleobase-modified aptamers with non-natural chemistries become accessible. One can then take advantage of the exponential enrichment and mutation enabled by the polymerase chain reaction to allow Darwinian molecular evolution, yet inclusion of more exotic unnatural functional groups allows evolutionary experiments well beyond the confined chemistries familiar of biology.

Here, we present an effort to stretch the idea of non-natural chemistry within an evolutionary nucleic acid aptamer selection by applying the in vitro evolutionary technique of SELEX to macromolecular chemistries entirely alien to biology. We perform an aptamer selection incorporating unusual cubane-modified nucleotides, presenting structural data, mechanistic determination, and diagnostic application of cubane-modified nucleic acid aptamers, hereafter called cubamers. The extraordinary platonic solid cubanes have intrigued chemists for decades(35-37), and are now important functional groups in a wide range of pharmaceuticals and agrochemicals(36, 38). Cubanes are benzene bioisosteres with unique properties: i) unlike benzene, cubane is biologically stable and inherently non-toxic; ii) enhanced water solubility compared to benzene due to the disruption of π stacking; iii) formation of unusual C-H···O hydrogen bonds caused by the strong s-character in the C-H bonds itself induced by the p-character of the C-C bonds(39, 40). Synthetic and medicinal chemistry approaches for cubane function discovery can be challenging, so we considered the possibility of developing a Darwinian approach to the evolution of macromolecules functionalized by cubane bioisosteres.

A goal of this study was to evolve cubamers selective for binding to the malaria diagnostic biomarker *Plasmodium vivax* lactate dehydrogenase (PvLDH). We had previously selected canonical DNA aptamers binding to *Plasmodium falciparum* lactate dehydrogenase (PfLDH)(4), then applied these in various diagnostic approaches(41-44). *P. falciparum* is the parasite which causes the most severe disease with the highest mortality whilst *P. vivax* is more widely distributed and causes disease which is complex and recurring. As management and treatment of patients infected with the two different species differs, an ideal point-of-care diagnostic should be able to discriminate the two infections. However, canonical DNA aptamers with the ability to discriminate PfLDH vs PvLDH, had eluded us. We asked whether cubamers might provide a solution.

Results

Synthesis of cubane-modified building blocks for selection and synthesis of cubane-modified aptamers

We developed an approach to select for cubane-modified aptamers through initially synthesizing the cubane-modified deoxyuridine triphosphate dU^{C}TP (SI Appendix, Scheme 2, compound **11**). To incorporate the cubane on the nucleobase we started from the known cubane derivative **1** (SI Appendix, Scheme 1) to prepare the azide (compound **5**) through previously published procedures(45). The 5-position alkyne-modified deoxyuridine derivative (compound **6**) which had been prepared according to previously published procedures(46), was then reacted with the azide **5** through copper(I)-catalyzed azide–alkyne cycloaddition (CuAAC) to form nucleoside compound **10** which was then converted to the corresponding triphosphate (compound **11**). The cubane-modified deoxyuridine phosphoramidite (compound **8**) was synthesized through similar click chemistry approaches to yield compound **9**.

Investigating compatibility of dU^{C}TP with primer extension and polymerase chain reaction

We next asked whether dU^{C}TP **11** (Fig.1b) would be compatible with enzymatic DNA synthesis under primer extension (PEX) reaction and PCR conditions. We used the 15 nucleotide long, 5’ 6-FAM (fluorescein) labelled primer **P1** annealed to template **T1** (see SI Appendix for sequence composition) then investigated the ability of various polymerases to perform a PEX reaction with a mixture of either four natural triphosphates or three natural triphosphates together with dU^{C}TP replacing dTTP (SI Appendix, Fig. S20). This analysis revealed that most polymerases accepted dU^{C}TP as a substrate and the primer was extended by 16 nucleotides (accounting for the incorporation of 4 modified nucleotides). When a longer template was used (the 79-mer template **T2**) which would require extension of primer

P2 by 60 nucleotides a clear higher molecular weight product is formed in the reaction with dU^{C}TP when compared to that with natural triphosphates (SI Appendix, Fig. S21 and Fig. S22). Several polymerases including Therminator, Vent (*exo*$^{-}$), Deep Vent, Taq, *Bst*, Klenow, Q5 DNA polymerase, and phi29 were all capable of forming the expected cubane-modified product. However, Vent (*exo*$^{-}$) appeared to most consistently produce highest yield product (SI Appendix, Fig. S22) so was carried forward to investigation of efficacy in PCR.
A polymerase chain reaction was performed using two simple 19/20 nucleotide DNA primers (**P3** and **P4**) on the same 79 nucleotide template **T2**. Vent (*exo*$^{-}$) was capable of forming the expected amplicon in high yields. Here also, the bands corresponding to cubane-modified DNA presented a lower electrophoretic mobility compared to the amplicon obtained with the natural dNTPs (SI Appendix, Fig. S23). Taq was not able to produce any product where Hemo KlenTaq, Q5 DNA polymerase and Phusion did produce the expected product albeit with lower yields compared to Vent (*exo*$^{-}$). We therefore concluded that Vent (*exo*$^{-}$) was the most appropriate DNA polymerase to use with the dU^{C}TP for subsequent SELEX procedures.

SELEX for cubamers with high affinity and specificity for PvLDH

PvLDH and PfLDH proteins were expressed and purified as we described previously(4, 43). For both proteins, we do not cleave the histidine tag during the purification procedure to facilitate immobilization onto nickel magnetic beads for aptamer selection. A single-stranded DNA library containing a 35-mer random region was used as previously described(4). This was amplified by PCR and transcribed to the cubane-modified pool using Vent (*exo*$^{-}$). The cubane-modified pool was incubated with target PvLDH protein conjugated onto Ni-NTA beads and unbound species removed. Eluted cubamers were reamplified from a forward primer and reverse primer with a biotinylated 5' end. Cubamers were purified through streptavidin magnetic beads followed by alkaline elution to result in the cubamer pool for next round of selection. Four classes of counterselection steps were incorporated: with Ni-NTA magnetic agarose beads after rounds 3 and 9, with PfLDH-immobilized Ni-NTA magnetic agarose beads after rounds 4 and 10, with hLDHB-immobilized Ni-NTA magnetic agarose beads after rounds 6 and 12, and with hLDHA1-immobilized Ni-NTA magnetic agarose beads after rounds 5 and 11. One can observe enrichment of PCR products up to round 15 with the expected depletion of PCR product after round 8 due to counterselection (SI Appendix, Fig. S26a). Amplification of selected pools by either dTTP or by dU^{C}TP showed the enrichment and expected molecular weight size difference to the cubane modification (SI Appendix, Fig. S26b). An enzyme-linked oligonucleotide assay was performed as we previously described(43) to measure binding of the pool to target PvLDH (SI Appendix, Fig. S26c). One can see clear binding of the dU^{C}TP modified pool to PvLDH particularly after round 10 which is not present for the natural nucleotide pool or for other controls. This data indicates that SELEX was successful for evolving a pool of modified oligonucleotides binding to PvLDH, and that the binding is dependent on the cubane modification.
Next Generation Sequencing was performed after round 15 (SI Appendix, Table S2). Two sequences dominated the pool: 1501s with five dU^{C} (23.1% of pool sequences) and 1506s with six dU^{C} (22.4% of pool sequences), then to a lesser extent 1516s (6.0%) and 1526s (4.9%) (SI Appendix, Table S4). These cubane-modified sequences were then synthesized by automated solid phase DNA synthesis using the cubane-modified phosphoramidite **8** (SI Appendix, Fig. S13) and molecular weights were confirmed by mass spectrometry (SI Appendix, Table S1 and Fig. S14-19). Interestingly, all attempts at synthesizing these oligonucleotides by click reaction with azide **5** and EdU-modified sequences failed due to the presence of a distribution of products regardless of the conditions.
We decided to focus our subsequent studies on the most abundant cubamer in our selection, 1501s. We determined the affinity for cubamers binding to PvLDH through surface plasmon resonance (SPR) analysis. SPR was performed through immobilization of PvLDH onto Ni-NTA with the cubamer in the mobile phase. K_D was determined to be 670 ± 9 nM by SPR (Fig. 2a). The cubamer was also observed to be specific for PvLDH relative to PfLDH, PfLDH binding was over 30 times weaker at 26 µM (Fig. 2b)

X-Ray crystal structure of cubamers in complex with PvLDH

The crystal structure of PvLDH in complex with the 1501s cubamer was solved at 2.5Å resolution (PDB entry 6TXR; Fig. 1 and SI Appendix, Table S3). As expected, the overall structure of PvLDH is similar to that previously reported (PDB entry 3ZH2) indicating that binding of cubamers does not affect the 3D

structure of this protein. Difference Fourier electron density maps indicate the presence of two cubamers which bind to the target protein at each end of the tetrameric PvLDH (Fig. 1a), in some ways reminiscent of the canonical DNA aptamers we had previously selected binding to PfLDH(4). The 5' end of the cubamer showed two base-paired loops in close interaction with the surface of PvLDH, whilst the 3' end of the cubamer extended away from the protein surface and formed a duplex with an aptamer bound to a second protein molecule in the crystal packing (SI Appendix, Fig. S34). We mapped all the interactions between PvLDH and cubamer in Figure 1c. All of the dU^C units were involved in some hydrogen-bonding within the cubamer. Two of these dU^C (T3-A21 and T17-A8) base-pairing interactions are conventional Watson-Crick base pairs to adenosine (SI Appendix, Fig. S27). Others are single hydrogen bond interactions either between themselves (T5 and T20) or with other noncanonical bases (T13-C11). Four of the cubanes were observed to cluster together through interactions both amongst themselves but then also through interactions with hydrophobic amino acid side chains in the PvLDH. A fifth cubane was somewhat separate from the other four and involved in interactions on another surface. The buried surface between the cubamer and PvLDH (Fig 1d) was 1130 $Å^2$ as calculated using the PISA server (https://www.ebi.ac.uk/pdbe/pisa/).

Mechanism of discrimination of binding recognition by cubamer for PvLDH vs PfLDH

The crystal structure of the complex revealed the binding site of the cubamer 1501s which is in close proximity to an alpha helix where there are some amino acid differences between PvLDH and PfLDH (Fig. 3a). This alpha helix is remarkable with regards to its non-conserved nature, as overall PvLDH and PfLDH are highly conserved with 90% amino acid identity (47). Electrophoretic mobility shift assays demonstrate that our previous selected aptamers 2008s (selected as a canonical DNA aptamer binding to PfLDH) bound more strongly to PfLDH than PvLDH (Fig. 3b) (4). A previously selected pan-specific DNA aptamer 1202s bound to both PfLDH and PvLDH with similar affinity (Fig. 3b). The 1501s cubamer showed preference for PvLDH, with some weak binding to PfLDH (Fig.3b), consistent with the SPR data (Fig. 2b). When the dU^C nucleotides in 1501s were mutated to simple canonical dT units then binding was abolished indicating the necessity of the cubane modification for the specific binding (Fig. 3b).

It was noted that Leucine 232 and Alanine 233 of PvLDH played important roles in the binding directly interacting through hydrophobicity with two of the cubane moieties (Fig. 1c) and being critical part of network of the cubanes in the binding pocket (Fig. 4a). We expressed PvLDH with leucine 232 mutated to glycine, and also with alanine 233 mutated to glycine, then both together. The L232G mutation decreased binding, the A233G mutation decreased binding slightly, and the compound with the double mutation resulted in significant reduction in binding as shown by EMSA (Fig. 4b) and by SPR (Fig. 4c). These experiments provided evidence that cubane interactions were dependent on hydrophobicity of amino acid side chains.

A unique hydrogen bond within the cubane-protein complex structure

Unlike phenyl or other simple arene moieties, cubane can also engage in the formation of stabilizing $C{-}H_{cubane}\cdots O$ bonds due to the rather high acidity of the cubyl H atoms (40). Such non-classical hydrogen bonds also seem to occur between a dU^C unit of the aptamer and a carbonyl unit of amino acids of PvLDH. Indeed, one C-H(2) atom lies in close proximity (2.4-2.6 Å) of the oxygen atom of the carbonyl group of Leu232 and forms a near linear (hydrogen bond angle θ = 147.3-158.7°) angle (Fig. 5). Both values fall within the ranges that have been reported for the formation of *syn* - *anti* catemers of small cubane derivatives (d of 2.2-3.0 Å and θ between 120 and 180°)(39, 40, 48, 49) and that of weak, non-classical hydrogen bonds (cutoff d value of 2.8 Å and θ range comprised between 150° and 180°)(50).

Using the cubamer in diagnostic assays to mimic clinical situation

We have previously developed both enzyme-linked oligonucleotide (ELONA) assays(43) and an aptamer-tethered enzyme capture assay(41, 42, 44) to use our aptamers which bind to LDH as diagnostic approaches for malaria. In the ELONA assay proteins are adhered to a surface, then a biotinylated aptamer/cubamer is bound to the target proteins, then streptavidin/HRP is added to reveal the binding and at the same time the presence or absence of the target protein (Fig. 6a). When using

this assay it could be observed that the test was highly specific for PvLDH over PfLDH (Fig. 6b). Furthermore, when dU^{C}TP was changed to dTTP during enzymatic synthesis of the aptamer, the assay was no longer effective, again proving the requirement for the cubane modifications here in the translational application (SI Appendix, Fig. S29a). The limit of detection for the ELONA assay in human serum buffer was determined as 1.05 ± 0.20 ng/µl (SI Appendix, Fig S29a). The ELONA assay is not clinically relevant in the present format as the attachment is of His-tagged proteins to nickel-coated plates but is relevant for our comparison to previous data(43), and could be implemented through direct adherence of the native protein as per direct ELISA methodologies. Furthermore, the mean levels of PvLDH in *P. vivax* isolates has been reported as 3.9 ± 6.1 ng/µl(51), this ELONA assay would occasionally be outside clinical range as PvLDH in some low parasitemia conditions(51).
Therefore, we switched to the aptamer-tethered enzyme capture assay. In the APTEC assay, the cubamer/aptamer is biotinylated and adhered to a streptavidin surface, then the clinical sample containing the target is applied, then washed, then color developed through the intrinsic catalytic activity of the lactate dehydrogenase(41) (Fig. 6c). Similar observations were made for the APTEC assay as those made for the ELONA assay confirming that the assay was specific for PvLDH (Fig. 6d) and required the cubane moieties (SI Appendix, Fig S29b). The limit of detection of the APTEC assay in human serum buffer was determined as 4.33 ± 1.66 pg/µl (SI Appendix, Fig S29b), which is much below the reported values of clinical range for PvLDH(51). Whilst further experiments in clinical samples would be important to ascertain diagnostic sensitivity particularly at low parasitemia, we have evidence in these mimetic samples that the APTEC cubamer assay can potentially be used for clinical application.

Discussion

We have used *in vitro* selection to evolve aptamers that are functionalized with the benzene isostere cubane. The so-called cubamers were selected against PvLDH and selectively recognized this protein target from PfLDH despite high sequence homology and bound with an appreciable affinity (K_D of 670 nM). The binding affinity of the cubamer is ~10-fold lower than a previously reported, unmodified aptamer(4) but this increase in K_D is largely compensated by the capacity of discriminating PvLDH from PfLDH. The high sequence homology between both LDHs might have reduced the number of epitopes available for binding and thus limited the binding capacity. Importantly, we resolved the crystal structure of the binary aptamer-protein complex which revealed the formation of a cubane cluster that interacts with hydrophobic residues of PvLDH. These findings mirror the functional interaction interface observed in crystal structures of SOMAmers in complex with their targets where hydrophobic interactions between the small aromatic moieties of the aptamer and of amino acid side chains were largely responsible for the binding mechanism (52, 53). As for SOMAmers and HFNAPs (sequence-defined highly functionalized nucleic acid polymers (17, 54, 55)), the specific location and sequence context of cubane moieties is crucial for binding to the protein target since it enables the formation of critical hydrophobic interactions accompanied by a C−H$_{cubane}$···O hydrogen bond.
In addition, cubamer 1501s was applied for the detection of the malaria biomarker PvLDH with excellent limit of detection in the APTEC assay, underscoring the usefulness of cubane-modified aptamers in practical applications. Therefore, by merging medicinal chemistry approaches with traditional SELEX, we could isolate a cubamer that outperforms functional nucleic acids selected with canonical nucleotides that are unable to distinguish between both LDHs.
Molecular evolution approaches provide routes to molecules which can have translational application but which could otherwise never be imagined or designed. SELEX has provided an experimental approach towards evolving nucleic acid ligands by the blind watchmaker of evolution entirely without need for a living organism through the replicative process of the polymerase chain reaction. There have been many extensions to SELEX in recent years through subtle differences of chemistry, but here we have demonstrated how entirely unnatural artificial functional group chemistries never observed in life can be subject to evolution under SELEX selection pressures.
This work demonstrates that the inclusion of nucleotides bearing the benzene isostere cubane on their nucleobases in a SELEX experiment has profound implications on the resulting functional nucleic acids. Besides opening the possibility of forming new binding modes with proteins via the formation of cubane-based pockets combined with C−H$_{cubane}$···O hydrogen bonding interactions, the modifications massively

improve the specificity of the cubamers for binding a PvLDH target enabling detection of this malaria biomarker in a mimetic clinical situation.
One can imagine steps beyond this work for completely new chemistries which could significantly extend such selection experiments where the selection pressure is not just binding, but could be selection pressure for catalysis or other function. This is particularly the case when one considers expanded genetic codes of 6-letter DNA alphabets(56-58), or even of 8-letter DNA(25, 59). In addition, the increase in diversity through unusual chemistries coupled to the replicative evolutionary power enabled by the polymerase chain reaction would provide a route to entirely new horizons for nucleic acid chemistry and give rise to hitherto unknown binding mechanisms. The introduction of exotic functional groups during SELEX may confer additional functional properties not accessible to canonical nucleic acids. This may for example be important for portable biosensors to detect biomarkers for pathogens in low-resource settings. Evolution of macromolecules incorporating chemistries entirely unknown in biology will have important repercussions across a plethora of applications.

Materials and Methods

Chemical syntheses

Synthetic procedures of the cubane-modified building blocks are given in the Supplementary Note available online.

*Pv*LDH and *Pf*LDH expression and purification

Protein expression and purification was carried out at the platform for production and purification of recombinant proteins of Institut Pasteur. PvLDH gene sequence was inserted into the pET28a (Novagen) vector and then recombinant PvLDH was transformed into *E. coli* strain BL21 (DE3) pLysS. Bacteria were incubated at 37 degrees in Luria-Bertani (LB) broth until OD_{600} reaches 0.5. After cooling down to room temperature, protein expression was induced by adding 0.25mM IPTG (isopropyl β-D-1-thiogalactopyranoside). Cells were grown at 25°C for 4 hours with agitation (220 rpm). Cells were then harvested by centrifuging the culture at 4000 rpm for 45 minutes. The supernatant was discarded and the cell pellet was resuspended by using 1× PBS. The cell suspension was transferred to 50mL tubes and centrifuged at 4000 rpm for 30 minutes. After discarding the supernatant, the cell pellet was resuspended in lysis buffer (100 mM Tris-HCl, 10 mM imidazole, 0.3 M NaCl, pH 8) in a ratio 1:50 lysis buffer to the original culture volume. After sonication, cell debris was removed by centrifugation at 19000 rpm rotor for1 h at 4°C. Supernatants were loaded onto a 5 ml Protino Ni-NTA column (Macherey Nagel) on an AKTA pure system (GE Healthcare) and eluted with 50 mM Tris-HCl, 0.3 M NaCl, 500 mM imidazole, pH 8. The eluted fractions were then purified on a Hiload 16/60 (120 ml) superdex 75 column using 50 mM Tris-HCl, 0.3 M NaCl, pH 8. After dialysis, the purified protein was stored in 25 mM Tris-HCl, 0.1 M NaCl, 40% glycerol, pH 8. 120 mg of purified protein were obtained at a concentration of 20 mg/mL. PfLDH was expressed and purified in a similar manner.

Cloning, expression and purification of *Pv*LDH mutants

The mutants of *PvLDH* were amplified by using overlap extension PCR. For PvLDH-M1 that consists a mutation at L232 become G232, two fragments that containing the mutations corresponded to PvLDH-M1 were firstly amplified by *PvLDH-S*, ATTATTGCTAGCATGACGCCGAAACCCAAAATTGTGCTC, and *PvLDH-L232G-AS*, TGGGGCAACATAAGGAGAGGCACCGAGGTTCACAATCTCCAA, to become *M1-a*; and PvLDH-L232G-S, TTGGAGATTGTGAACCTC**GGT**GCCTCTCCTTATGTTGCCCCA, and *PvLDH-AS*, ATTATTCTCGAGTTAAATGAGCGCCTTCATCCTTTTAGTCTC, to become M1-b. The ORF corresponded to PvLDH-M1 was further amplified by *PvLDH-S*, ATTATTGCTAGCATGACGCCGAAACCCAAAATTGTGCTC, and *PvLDH-AS*, ATTATTCTCGAGTTAAATGAGCGCCTTCATCCTTTTAGTCTC, using *M1-a* and *M1-b* as template. The PvLDH-M1 ORF then was ligated into the *Nhe*I/*Xho*I digested *pET28a-PvLDH* plasmid that constructed in our previous study to become *pET28a-PvLDH-M1*. For PvLDH-M2 that consists a mutation at A233 become G233, two fragments that containing the mutations corresponded to PvLDH-M1 were firstly amplified by *PvLDH-S*, ATTATTGCTAGCATGACGCCGAAACCCAAAATTGTGCTC, and *PvLDH-L232G-AS*, TGGGGCAACATAAGGAGAACCACCGAGGTTCACAATCTCCAA, to become

M2-a; and *PvLDH-A233G-S*, TTGGAGATTGTGAACCTCCTT**GGT**TCTCCTTATGTTGCCCCA, and *PvLDH-AS*, ATTATTCTCGAGTTAAATGAGCGCCTTCATCCTTTTAGTCTC, to become *M2-b*. The ORF of PvLDH-M2 was further amplified by *PvLDH-S*, ATTATTGCTAGCATGACGCCGAAACCCAAAATTGTGCTC, and *PvLDH-AS*, ATTATTCTCGAGTTAAATGAGCGCCTTCATCCTTTTAGTCTC, using *M2-a* and *M2-b* as template. The PvLDH-M2 ORF then was ligated into the *Nhe*I/*Xho*I digested *pET28a-PvLDH* plasmid that constructed in our previous study to become *pET28a-PvLDH-M2*. For PvLDH-M3 that consists mutations at L233 and A233 that became G233 and G233 respectively, two fragments that containing the mutations corresponded to PvLDH-M3 were firstly amplified by *PvLDH-S*, ATTATTGCTAGCATGACGCCGAAACCCAAAATTGTGCTC, and *PvLDH-L232A233G-AS*, TGGGGCAACATAAGGAGAACCACCGAGGTTCACAATCTCCAA, to become *M3-a*; and *PvLDH-L232A233G-S*, TTGGAGATTGTGAACCTC**GGTGGT**TCTCCTTATGTTGCCCCA, and *PvLDH-AS*, ATTATTCTCGAGTTAAATGAGCGCCTTCATCCTTTTAGTCTC, to become *M3-b*. The ORF of PvLDH-M3 was further amplified by *PvLDH-S*, ATTATTGCTAGCATGACGCCGAAACCCAAAATTGTGCTC, and *PvLDH-AS*, ATTATTCTCGAGTTAAATGAGCGCCTTCATCCTTTTAGTCTC, using *M3-a* and *M3-b* as template. The PvLDH-M3 ORF then was ligated into the *Nhe*I/*Xho*I digested *pET28a-PvLDH* plasmid that constructed in our previous study to become *pET28a-PvLDH-M3*. The plasmids of the mutants were transformed into *E. coli* BL21 (DE3) pLysS for isopropyl ・-D-1- thiogalactopyranoside (IPTG)-induced expression. PvLDH mutants were expressed and purified according to our previous study (1, 2). Bacterial culture was incubated at 37 °C until OD_{600} = 0.5, then 0.25 mM IPTG was added followed by 4-h expression at 25 °C. Pellets were collected, sonicated, and expressed proteins were purified by using HisTrap chromatography (GE Healthcare).

Cubamer selection

A single-stranded DNA (ssDNA) library containing a 35-mer random region flanked by two 18-mer priming regions (5'-CGTACGGTCGACGCTAGC-[N35]-CACGTGGAGCTCGGATCC-3') was used as starting material for *in vitro* selection. The ssDNA library was amplified by forward primer (5'-CGTACGGTCGACGCTAGC-3') and reverse primer with biotinylated 5'-end (5'-biotin-GGATCCGAGCTCCACGTG-3') using Vent (*exo*-) DNA Polymerase (NEB), $dU^{c}TP$, dATP, dGTP and dCTP. PCR conditions were as follows: denaturation at 95 °C for 15 s, annealing at 50 °C for 30 s and elongation at 72 °C for 15 s for 10 cycles. The amplified library was purified by streptavidin magnetic beads followed by a wash of 100 mM to elute the non-biotinylated complementary strand as the cubane modified ssDNA library. Two namomoles of cubane modified ssDNA library was incubated with 1 nmol of PvLDH that was conjugated with Ni-NTA magnetic beads. The unbound species were removed, and the ssDNA-protein-magnetic bead complexes were suspended in 10 µL of water for PCR amplification of PfLDH-bound species by using forward primer (5'-CGTACGGTCGACGCTAGC-3') and reverse primer with biotinylated 5'-end (5'-biotin-GGATCCGAGCTCCACGTG-3') using Vent *(exo-)* DNA polymerase. PCR conditions were as follows: denaturation at 95 °C for 15 s, annealing at 50 °C for 30 s and elongation at 72 °C for 15 s for 10 cycles. Enriched DNA aptamer pool was purified by streptavidin magnetic beads followed by a wash of 100 mM NaOH to remove the non-biotinylated complementary strand. The resultant ssDNA pool was used for the next round of selection. Counterselections by using Ni-NTA magnetic agarose beads (after rounds 3 and 9), PfLDH-immobilized Ni-NTA magnetic agarose beads (after rounds 4 and 10), hLDHB-immobilized Ni-NTA magnetic agarose beads (after rounds 6 and 12) and hLDHA1-immobilized Ni-NTA magnetic agarose beads (after rounds 5 and 11) were incorporated in between the SELEX cycles.

Cubamer biophysical characterization by surface plasmon resonance (SPR)

SPR measurement was collected using a Biacore X100 instrument (GE Healthcare). PvLDH was captured on the surface of NTA sensor chip (GE Healthcare). A running buffer containing 25 mM Tris-HCl, pH7.5, 100 mM NaCl, 20 mM imidazole, 0.005% (v/v) Tween 20 and 0.5 mg/mL BSA was used for ligand capturing. The surface of NTA chip was primed with running buffer and PvLDH cubamer, 1501s, was injected in at a flow rate of 30 µL min^{-1} and at 25 °C. All data were referenced for surface

without captured cubamer and blank injections of buffer. Sensorgrams were analyzed with Biacore X100 Plus Package evaluation software (GE Healthcare).

Crystallization and data collection and structure refinement

Crystallization screening trials of the protein-aptamer complex (14mg/ml) were carried out by the sitting drop vapor-diffusion method with a Mosquito automated nanoliter dispensing system (TTP Labtech, Melbourn, UK) at 291K. Sitting drops of 400 nL were set up in Greiner plates for 672 commercially available screening solutions using a 1:1 mixture of protein sample and reservoir well solution (150 µL). The plates were stored in a RockImager (Formulatrix, Bedford, USA) automated imaging system to monitor crystal growth. The crystals appeared after 3 months in crystallization condition 52 of Index HT kit (Hampton Research) containing 45% (v/v) 2-Methyl-2,4-pentanediol, 0.2 M Ammonium acetate and 0.1 M Hepes pH 7.5. The fished crystals were flash-cooled in liquid nitrogen directly, as the condition of crystallization served also as cryoprotectant. Diffraction data were collected at beamline PROXIMA-1 (SOLEIL synchrotron, St. Auban, France) and processed with XDS(60) and Aimless(61).
The structure of the PvLDH-cubamer complex was solved by the molecular replacement technique using a PvLDH tetramer (PDB entry 3ZH2) as search model with Phaser(62). A complete model of the complex was obtained through interactive cycles of manual model building with Coot(63) and reciprocal space refinement with Buster(64) and phenix.refine(65). The complex's DNA moiety was gradually traced into difference electron density maps during this process. X-ray data collection and model refinement statistics are summarized in Table S2. Figures showing the crystallographic model were generated and rendered with Pymol (Schrodinger, LLC) and/or Chimera(66). The atomic coordinates and structure factors for the PvLDH-cubamer complex have been deposited in the RCSB Protein Data Bank under the accession code 6TXR.

Electrophoretic mobility shift assay (EMSA)

EMSAs were carried out according to our previous study (4, 43). Protein was 2-fold serial diluted in binding buffer (25 mM Tris·HCl at pH 7.5 containing 0.1 M NaCl and 20 mM imidazole). The nucleotides, including 1501s, 2008s, 1202s or corresponding cubamer variants, was mixed with the serial diluted proteins to a final concentration of 25 mM, followed by incubating at room temperature. The reactions were loaded on 12% native polyacrylamide gels and visualized by SYBR gold nucleic acid gel staining (Invitrogen).

Cubamer-Integrated enzyme linked oligonucleotide assay (ELONA)

A biotinylated complementary strand (3'–CTCCCGTGCCGTTTT–5') was annealed to 1501s cubamer by incubating equal molar volumes in buffer (10 mM Tris, pH 7.5, 50 mM NaCl, 1 mM EDTA) heating at 95 °C for 5 min, then allowed to cool at room temperature for 60 min before use. Annealing effectiveness was assessed by 12% TAE gel. Protein samples were prepared by diluted in phosphate buffered saline (PBS) with 0.05% Tween 20 to concentrations of 10, 5, 2.5, 1.25, 0.63, 0.31, 0.16, 0 µg/mL which is 10000, 5000, 2500, 1250, 630, 310, 160, 0 pg/µL, respectively. 100 µL of each protein were incubated in Ni-NTA HisSorb Plate (QIAGEN) wells for 1 hour at room temperature. Wells were then washed 3 times with 200 µL PBST (0.05% Tween 20), and 100 µL of 50 nM biotinylated cubamer was added and incubated for 1 hour at room temperature. PBST (0.05% Tween 20) was used as the cubamer binding buffer, while a 1:1 ratio of human AB plasma (Sigma-Aldrich) to PBS (0.05% Triton X-100) was used for the human serum binding buffer. Subsequently, wells were washed 3 times with PBST, then 100 µL of 1:10000 Streptavidin-HRP (Abcam) diluted in PBST were added and incubated for 30 min at room temperature. Following 3 times wash of 200 µL with PBST, 100 µL of Pierce 1-Step Ultra TMB ELISA Substrate (Thermo Fisher Scientific) was added and incubated for 15 min at room temperature. The reaction was stopped by adding 2M H_2SO_4 and absorbance was read at 450nm.

Cubamer-integrated aptamer-tethered enzyme capture assay (APTEC)

100 µL of biotinylated cubamer was added to pre-washed (PBST 0.1% Tween 20) Pierce streptavidin coated plates (Thermo Fisher Scientific) for 2 hrs. Following a 3 times PBST wash, 100 µL of sample was added to each cubamer-decorated well for 1 hour, then washed 5 times with PBST. For color development, 100 µL of pre-prepared L-lactate/NTB solution: 12 mL L-lactate buffer (0.2 M sodium L-

lactate, 100 mM Tris HCl, 0.2% Triton X-100, pH 9.1), 158 µL NAD^+ solution (50 mg/mL in water), 48 µL NBT solution (25 mg/mL in water) and 25 µL PES solution (5 mg/mL in water) was added and left to incubate for 45 min with mild shaking in the dark. Acetic acid (5%) was added to stop the reaction and the absorbance was read at 570nm. For recombinant pLDH samples, 2x PBS was used as the binding buffer. For spiked human serum samples, 50 µL of human AB plasma (Sigma-Aldrich) was mixed with 50 µL PBS (0.5% Triton X-100) prior to adding protein.

Data Availability

All data used in the study are included in the manuscript and in the SI Appendix. The crystal structure of the cubamer-target complex has been deposited in the PDB (ID code 6TXR).

Acknowledgments

We thank the Shirley Boyde Trust for their generous support through a Shirley Boyde Foundation Grant to J.A.T. and acknowledge Institut Pasteur funding to M.H., A.H., P.R., and F.L.-A.. J.A.T. also acknowledges the University of Hong Kong Outstanding Young Researcher Award 2015-16 and the University of Hong Kong Outstanding Research Student Supervisor Award 2016-2017. J.A.T. also acknowledges HKU Seed Fund for Basic Research grants 201711159185 and 201611159269. We thank the staff of the platform for production and purification of recombinant proteins of Institut Pasteur for technical assistance during the expression and purification of the proteins steps. The authors are grateful to the staff of the crystallography platform at Institut Pasteur for carrying out robot-driven

crystallization screening. We acknowledge synchrotron SOLEIL (Saint-Aubin, France) for granting access to their facility and the staff of Proxima1 for helpful assistance during the data collection.

Author Competing Interest

Julian A Tanner is a scientific advisor for Zio Health and Julian A Tanner and Andrew B. Kinghorn are founders of Jushan Bio.

References

1. Ellington AD & Szostak JW (1990) In vitro selection of RNA molecules that bind specific ligands. *Nature* 346(6287):818-822.
2. Tuerk C & Gold L (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* 249(4968):505-510.
3. Gelinas AD, Davies DR, & Janjic N (2016) Embracing proteins: structural themes in aptamer-protein complexes. *Current opinion in structural biology* 36:122-132.
4. Cheung YW*, et al.* (2013) Structural basis for discriminatory recognition of Plasmodium lactate dehydrogenase by a DNA aptamer. *Proceedings of the National Academy of Sciences of the United States of America* 110(40):15967-15972.
5. Kaur H, Bruno JG, Kumar A, & Sharma TK (2018) Aptamers in the Therapeutics and Diagnostics Pipelines. *Theranostics* 8(15):4016-4032.
6. Liang S*, et al.* (2019) Measuring luteinising hormone pulsatility with a robotic aptamer-enabled electrochemical reader. *Nat Commun* 10(1):852.
7. Nguyen VT, Kwon YS, & Gu MB (2017) Aptamer-based environmental biosensors for small molecule contaminants. *Curr Opin Biotechnol* 45:15-23.
8. Sakai Y*, et al.* (2018) DNA Aptamers for the Functionalisation of DNA Origami Nanostructures. *Genes (Basel)* 9(12).
9. Gold L (2015) SELEX: How It Happened and Where It will Go. *Journal of molecular evolution* 81(5-6):140-143.
10. Rohloff JC*, et al.* (2014) Nucleic Acid Ligands With Protein-like Side Chains: Modified Aptamers and Their Use as Diagnostic and Therapeutic Agents. *Molecular therapy. Nucleic acids* 3:e201.
11. Dunn MR, Jimenez RM, & Chaput JC (2017) Analysis of aptamer discovery and technology. *Nat. Rev. Chem.* 1:0076.
12. Röthlisberger P & Hollenstein M (2018) Aptamer Chemistry. *Adv. Drug Deliv. Rev.* 134:3-21.
13. Gawande BN*, et al.* (2017) Selection of DNA aptamers with two modified bases. *Proc. Natl. Acad. Sci. U. S. A.* 114(11):2898-2903.
14. Kimoto M, Yamashige R, Matsunaga K, Yokoyama S, & Hirao I (2013) Generation of high-affinity DNA aptamers using an expanded genetic alphabet. *Nat. Biotechnol.* 31(5):453-457.
15. Pinheiro VB*, et al.* (2012) Synthetic Genetic Polymers Capable of Heredity and Evolution. *Science* 336(6079):341-344.
16. Taylor AI, Houlihan G, & Holliger P (2019) Beyond DNA and RNA: The Expanding Toolbox of Synthetic Genetics. *Cold Spring Harb Perspect Biol* 11(6).
17. Lichtor PA, Chen Z, Elowe NH, Chen JC, & Liu DR (2019) Side chain determinants of biopolymer function during selection and replication. *Nat. Chem. Biol.* 15(4):419-426.
18. Arangundy-Franklin S*, et al.* (2019) A synthetic genetic polymer with an uncharged backbone chemistry based on alkyl phosphonate nucleic acids. *Nat. Chem.* 11(6):533-542.
19. Yu HY, Zhang S, & Chaput JC (2012) Darwinian evolution of an alternative genetic system provides support for TNA as an RNA progenitor. *Nat. Chem.* 4(3):183-187.
20. Liu ZX, Chen TJ, & Romesberg FE (2017) Evolved polymerases facilitate selection of fully 2 '-OMe-modified aptamers. *Chem. Sci.* 8(12):8179-8182.
21. Eremeeva E*, et al.* (2019) Highly stable hexitol based XNA aptamers targeting the vascular endothelial growth factor. *Nucleic Acids Res.* 47(10):4927-4939.
22. Mei H*, et al.* (2018) Synthesis and Evolution of a Threose Nucleic Acid Aptamer Bearing 7-Deaza-7-Substituted Guanosine Residues. *J. Am. Chem. Soc.* 140(17):5706-5713.
23. Rose KM*, et al.* (2019) Selection of 2 '-Deoxy-2 '-Fluoroarabino Nucleic Acid (FANA) Aptamers That Bind HIV-1 Integrase with Picomolar Affinity. *ACS Chem. Biol.* 14(10):2166-2175.

24. Biondi E & Benner SA (2018) Artificially Expanded Genetic Information Systems for New Aptamer Technologies. *Biomedicines* 6(2).
25. Hoshika S*, et al.* (2019) Hachimoji DNA and RNA: A genetic system with eight building blocks. *Science* 363(6429):884-887.
26. Renders M, Miller E, Lam CH, & Perrin DM (2017) Whole cell-SELEX of aptamers with a tyrosine-like side chain against live bacteria. *Org. Biomol. Chem.* 15(9):1980-1989.
27. Battersby TR*, et al.* (1999) Quantitative analysis of receptors for adenosine nucleotides obtained via in vitro selection from a library incorporating a cationic nucleotide analog. *J. Am. Chem. Soc.* 121(42):9781-9789.
28. Gupta S*, et al.* (2014) Chemically Modified DNA Aptamers Bind Interleukin-6 with High Affinity and Inhibit Signaling by Blocking Its Interaction with Interleukin-6 Receptor. *J. Biol. Chem.* 289(12):8706-8719.
29. Ren XM, Gelinas AD, von Carlowitz I, Janjic N, & Pyle AM (2017) Structural basis for IL-1 alpha recognition by a modified DNA aptamer that specifically inhibits IL-1 alpha signaling. *Nat. Commun.* 8:13.
30. Gold L, Walker JJ, Wilcox SK, & Williams S (2012) Advances in human proteomics at high scale with the SOMAscan proteomics platform. *N Biotechnol* 29(5):543-549.
31. Sun BB*, et al.* (2018) Genomic atlas of the human plasma proteome. *Nature* 558(7708):73-79.
32. Vaught JD*, et al.* (2010) Expanding the chemistry of DNA for in vitro selection. *J Am Chem Soc* 132(12):4141-4151.
33. Wolk SK*, et al.* (2020) Modified nucleotides may have enhanced early RNA catalysis. *Proceedings of the National Academy of Sciences of the United States of America* 117(15):8236-8242.
34. Pfeiffer F*, et al.* (2018) Identification and characterization of nucleobase-modified aptamers by click-SELEX. *Nat Protoc* 13(5):1153-1180.
35. Biegasiewicz KF, Griffiths JR, Savage GP, Tsanaktsidis J, & Priefer R (2015) Cubane: 50 years later. *Chem Rev* 115(14):6719-6745.
36. Reekie TA, Williams CM, Rendina LM, & Kassiou M (2019) Cubanes in Medicinal Chemistry. *J. Med. Chem.* 62(3):1078-1095.
37. Auberson YP*, et al.* (2017) Improving Nonspecific Binding and Solubility: Bicycloalkyl Groups and Cubanes as para-Phenyl Bioisosteres. *ChemMedChem* 12(8):590-598.
38. Houston SD*, et al.* (2019) The cubane paradigm in bioactive molecule discovery: further scope, limitations and the cyclooctatetraene complement. *Org Biomol Chem* 17(28):6790-6798.
39. Kuduva SS, Craig DC, Nangia A, & Desiraju GR (1999) 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(9):1936-1944.
40. Flanagan KJ, Bernhard SSR, Plunkett S, & Senge MO (2019) Not Your Usual Bioisostere: Solid State Study of 3D Interactions in Cubanes. *Chem.-Eur. J.* 25(28):6941-6954.
41. Dirkzwager RM, Kinghorn AB, Richards JS, & Tanner JA (2015) APTEC: aptamer-tethered enzyme capture as a novel rapid diagnostic test for malaria. *Chem Commun (Camb)* 51(22):4697-4700.
42. Dirkzwager RM, Liang S, & Tanner JA (2016) Development of aptamer-based point-of-care diagnostic device for malaria using 3D printing rapid prototyping. *ACS Sensors* 1(4):420-426.
43. Cheung YW*, et al.* (2018) Aptamer-mediated Plasmodium-specific diagnosis of malaria. *Biochimie* 145:131-136.
44. Fraser LA*, et al.* (2018) A portable microfluidic Aptamer-Tethered Enzyme Capture (APTEC) biosensor for malaria diagnosis. *Biosens Bioelectron* 100:591-596.
45. Chalmers BA*, et al.* (2016) Validating Eaton's Hypothesis: Cubane as a Benzene Bioisostere. *Angewandte Chemie* 55(11):3580-3585.
46. Röthlisberger P, Levi-Acobas F, & Hollenstein M (2017) New synthetic route to ethynyl-dUTP: A means to avoid formation of acetyl and chloro vinyl base-modified triphosphates that could poison SELEX experiments. *Bioorg Med Chem Lett* 27(4):897-900.
47. Turgut-Balik D*, et al.* (2004) Cloning, sequence and expression of the lactate dehydrogenase gene from the human malaria parasite, Plasmodium vivax. *Biotechnol. Lett.* 26(13):1051-1055.
48. Kuduva SS, Blaser D, Boese R, & Desiraju GR (2001) Crystal Engineering of Primary Cubanecarboxamides. Repetitive Formation of an Unexpected N−H···O Hydrogen-Bonded Network. *J. Org. Chem.* 66(5):1621-1626.
49. Das D, Jetti RKR, Boese R, & Desiraju GR (2003) Stereoelectronic effects of substituent groups in the solid state. Crystal chemistry of some cubanecarboxylic and phenylpropiolic acids. *Cryst. Growth Des.* 3(5):675-681.

50. Desiraju GR (1996) The C−H···O Hydrogen Bond: Structural Implications and Supramolecular Design. *Accounts Chem. Res.* 29(9):441-449.
51. Jang JW, Cho CH, Han ET, An SS, & Lim CS (2013) pLDH level of clinically isolated Plasmodium vivax and detection limit of pLDH based malaria rapid diagnostic test. *Malar J* 12:181.
52. Ren XM, Gelinas AD, von Carlowitz I, Janjic N, & Pyle AM (2017) Structural basis for IL-1 alpha recognition by a modified DNA aptamer that specifically inhibits IL-1 alpha signaling. *Nat. Commun.* 8.
53. Gelinas AD, Davies DR, & Janjic N (2016) Embracing proteins: structural themes in aptamer-protein complexes. *Curr. Opin. Struct. Biol.* 36:122-132.
54. Chen Z, Lichtor PA, Berliner AP, Chen JC, & Liu DR (2018) Evolution of sequence-defined highly functionalized nucleic acid polymers. *Nat. Chem.* 10(4):420-427.
55. Kong DH, Yeung W, & Hili R (2017) In Vitro Selection of Diversely Functionalized Aptamers. *J. Am. Chem. Soc.* 139(40):13977-13980.
56. Kimoto M, Matsunaga K, & Hirao I (2016) DNA Aptamer Generation by Genetic Alphabet Expansion SELEX (ExSELEX) Using an Unnatural Base Pair System. *Methods in molecular biology* 1380:47-60.
57. Malyshev DA*, et al.* (2014) A semi-synthetic organism with an expanded genetic alphabet. *Nature* 509(7500):385-388.
58. Zhang L*, et al.* (2015) Evolution of functional six-nucleotide DNA. *J Am Chem Soc* 137(21):6734-6737.
59. Dien VT, Holcomb M, & Romesberg FE (2019) Eight-Letter DNA. *Biochemistry* 58(22):2581-2583.
60. Kabsch W (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* 66:133-144.
61. Evans PR & Murshudov GN (2013) How good are my data and what is the resolution? *Acta Crystallogr. Sect. D-Biol. Crystallogr.* 69:1204-1214.
62. McCoy AJ*, et al.* (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* 40:658-674.
63. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D-Struct. Biol.* 60:2126-2132.
64. Bricogne G*, et al.* (2011) BLUSTER (Global Phasing Ltd, Cambridge, UK), 2.11.1.
65. Afonine PV*, et al.* (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. Sect. D-Struct. Biol.* 68:352-367.
66. Pettersen EF*, et al.* (2004) UCSF chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* 25(13):1605-1612.

Figures and Tables

Figure 1. Structure of PvLDH **in complex with cubamer a,** Crystal structure of tetrameric PvLDH binds to two cubamers. The four subunits of PvLDH are coloured in yellow, green, cyan and pink. The two 1501s cubamers are coloured in grey and the cubane modification coloured in blue. **b,** Protected cubane-modified deoxyuridine phosphoramidite (left) and cubane-modified deoxyuridine triphosphate (right). **c,** Details of interactions between 1501s and PvLDH. Light blue indicates hydrogen bonding interactions between cubamer and protein. Blue indicates hydrophobic interactions between cubamer and protein. Orange indicates residues within interaction range of 3 Å. Connected nucleotides are within 3.8 Å. Below, the sequence of the 1501s cubamer is shown with **X** representing cubane-modified deoxyuridine. **d,** Buried surface of PvLDH-1501s complex. We pinpointed by different colours the regions of interactions between the monomers A and B of PvLDH and a 1501 cubamer.

Figure 2. Determination of binding affinity and specificity of cubamer towards PvLDH. a, Surface plasmon resonance (Biacore) analysis for the interaction between 1501s and PvLDH. PvLDH was immobilised to a Ni-NTA chip, then the response was observed at multiple concentration of 1501s as shown in the figure. Raw data are shown background with the best fit overlaid in the thick solid line. For 1501s binding to PvLDH, the association (k_{on}), dissociation (k_{off}) and the equilibrium constant (K_D) were 2.2 x 10^4 $M^{-1}s^{-1}$, 1.4 x 10^{-2} s^{-1} and 670 nM, respectively. **b,** Specificity of 1501s binding to PvLDH relative to PfLDH. For 1501s binding to PfLDH, the association (k_{on}), dissociation (k_{off}) and the equilibrium constant (K_D) were 6.8 x 10^3 $M^{-1}s^{-1}$, 0.18 s^{-1} and 26 µM, respectively.

Figure 3. The cubamer discriminates PvLDH **from *Pf*LDH through binding to a variable region. a,** Superimposition of PvLDH and *Pf*LDH shows they are highly conserved with some difference on the variable region. The variable region is coloured in red, whereas the conserved region is in light green. The insert shows the cubanes on the modified nucleotides interact with the variable region. **b,** Comparing the affinity of 1501s to PvLDH and *Pf*LDH with the control of *Pf*LDH specific aptamer, 2008s, and pan-LDH specific aptamer, 1202s. EMSA analysis showed 1501s is highly specific to PvLDH. **c,** EMSA analysis showed cubanes on modified triphosphates involved in the binding between 1501s and PvLDH. With the replaced cubane modified triphosphate by dU^{C}TP, dU1501s, and by dTTP, dT1501s, the affinity of the cubamers are significantly decreased.

Figure 4. The cubamer interacts with PvLDH by hydrophobic interactions. a, Structure of PvLDH-1501s complex shows the cubane modified triphosphates, T1, T5, T20, T13 and T17 formed a hydrophobic cluster. The hydrophobic surface of PvLDH was illustrated with a colour code according to the polarity of the respective amino acids. Hydrophobic amino acids were coloured in red whereas hydrophilic residues are in white. **b,** EMSA reveals the binding of PvLDH and 1501s along with amino acid side chains. The affinities of 1501s to three mutants of PvLDH, $PvLDH_{L232G}$, $PvLDH_{A233G}$ $PvLDH_{L232G/A233G}$, and the wide-type PvLDH were estimated. **c,** SPR comparison of 1501s cubamer binding to PvLDH and the three mutants $PvLDH_{L232G}$, $PvLDH_{A233G}$ and $PvLDH_{L232G/A233G}$. Results from **b** and **c** indicate the affinities among all the mutants are significantly decreased by mutating the hydrophobic amino acids to glycine.

Figure 5. A unique hydrogen bond of the cubamer with PvLDH. Close up views of the interaction between a dU^C unit (T5) of the two cubamers 1501s with PvLDH (Leu232) showing the respective angles and bond lengths of the unique C-H(2) ···O hydrogen bond. The grey mesh corresponds to a Sigma-A weighted mFo-DFc difference map (contoured at 3 σ), in which Leu232 and the cubane moiety of T5 were omitted from the model before map calculation.

Figure 6. Determination of affinity and specificity of the 1501s cubamer against pLDHs in human serum (HS) buffer by Enzyme linked oligonucleotide assay (ELONA) and aptamer-tethered enzyme capture (APTEC) assay. a, Diagram illustrating the ELONA. Target pLDHs are immobilized onto a microtiter plate, then biotinylated cubamers are added to recognize pLDHs. Streptavidin-HRP is added to recognize the biotin-end of the cubamer, and TMB solution is used to allow for colour development. **b,** The response of the ELONA against clinical range of PvLDH and PfLDH in human serum (HS) buffer. The LOD of 1501s-PvLDH in the ELONA assay was determined to be 1.05 ± 0.20 $ng \cdot \mu L^{-1}$ (Fig. S29). **c,** Schematic diagram illustrating the APTEC assay. Biotinylated cubamers are immobilized onto streptavidin-coated wells, then pLDHs are added and captured by the cubamer. L-lactate/NTB solution is then added for colour development. **d,** The response of the APTEC assay against clinical range of PvLDH and PfLDH in HS buffer. The LOD of 1501s-PvLDH in the APTEC assay was determined to be 4.33 ± 1.66 $pg \cdot \mu L^{-1}$ (Fig. S29). Measurements were done in triplicate, reported in standard error (SE).