Unconventional split aptamers cleaved at functionally essential sites preserve biorecognition capability

Wang, Ruoyu; Zhang, Qiansen; Zhang, Yi; Shi, Hanchang; Nguyen, Kim Truc; Zhou, Xiaohong

2019

Wang, R., Zhang, Q., Zhang, Y., Shi, H., Nguyen, K. T. & Zhou, X. (2019). Unconventional split aptamers cleaved at functionally essential sites preserve biorecognition capability. *Analytical Chemistry*, 91(24), 15811-15817.
<https://dx.doi.org/10.1021/acs.analchem.9b04115>

<https://hdl.handle.net/10356/150833>

<https://doi.org/10.1021/acs.analchem.9b04115>

This document is the Accepted Manuscript version of a Published Work that appeared in final form in *Analytical Chemistry*, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see <https://doi.org/10.1021/acs.analchem.9b04115>

Downloaded on 26 Apr 2023 19:35:18 SGT

Unconventional Split Aptamers Cleaved at Functionally Essential Sites Preserve Biorecognition Capability

Ruoyu Wang,^{1 §} Qiansen Zhang,^{2 §} Yi Zhang,³ Hanchang Shi,¹ Nguyen Kim Truc⁴ and Xiaohong
Zhou^{1 †}

¹Research Centre of Environmental and Health Sensing Technology, School of Environment,
Tsinghua University, Beijing 100084, China.

²Shanghai Key Laboratory of Regulatory Biology, Institute of Biomedical Sciences, School of Life
Sciences, East China Normal University, Shanghai 200241, China.

³School of Mechanical and Aerospace Engineering, Nanyang Technological University, 50
Nanyang Avenue, Singapore 639798

⁴School of Electrical and Electronic Engineering, Nanyang Technological University, 50 Nanyang
Avenue, Singapore 639798

[§]These authors contributed equally to this work

[†]**Corresponding Author:** Email: xhzhou@mail.tsinghua.edu.cn; Tel: +86-10-62796953.

ABSTRACT

Split aptamers (SPAs) are a pair of oligonucleotide fragments generated by cleaving a long parent aptamer. SPAs have many compelling advantages over the parent aptamer such as sandwich target binding, optimized concise structure, and low cost. However, only a limited number of SPAs have been developed so far, because the traditional theory restricts the splitting to the functionally dispensable site that many parent aptamers do not possess. In this work, we challenge the traditional mechanism and hypothesize that SPAs can also be generated by splitting the parent aptamer at the functionally essential site while still preserving the biorecognition capability. To prove our hypothesis, we discovered three SPAs with **B**roken initial small-molecule binding **P**ockets (^{BP}SPAs) and validate their binding capability both in wet lab and *in silico*. An allosteric binding mechanism of ^{BP}SPAs, in which a new binding pocket is formed upon the target binding, is revealed by all-atom microsecond-scale molecular dynamics simulations. Our findings will greatly promote discovery of new SPAs and their applications in molecular sensing in many fields.

INTRODUCTION

Detecting molecules of interest by target-specific affinity probes allows us to gain a better understanding of the roles these molecules play in physiological and pathological processes.^{1, 2, 3} In the past few years, a new class of target-specific affinity probe known as the aptamer has been developed for biosensing, disease diagnostics, environment monitoring and food safety monitoring.^{4, 5} An aptamer, which is an oligonucleotide identified via an *in vitro* selection process, has a high binding affinity to a specific target. Sensitive aptasensing relies on the precise conformational change of the aptamer upon its binding to the target. However, the dynamic conformational variations of the free aptamer may interfere with the true signal resulted from the target-induced conformational transformation, especially in the case of small-molecule targets.^{6, 7} To address this issue, split aptamers (SPAs) are engineered by cleaving the parent aptamer into two oligonucleotide fragments that would assemble together upon the recognition of the common target.⁸ Compared to the long parent aptamer, the short SPAs can be chemically synthesized with a reduced cost and a lower error rate. SPAs also exhibit a higher denaturation/renaturation efficiency and a lower chance of nonspecific conformational transformation caused by non-target molecules.⁶ In addition, the splitting results in additional oligonucleotide termini which provide additional functional groups for multiple conjugations. SPAs have demonstrated excellent capability as the sensing element in single-molecule force spectroscopy,⁹ and optical¹⁰ and electrochemical sensing platforms¹¹. Several SPAs have entered clinical trials for diagnosing cancer, inflammation and macular degeneration.¹²

In spite of these compelling advantages, only fewer than 20 SPAs (**Table S1**) have been reported so far,⁶ because many aptamers are deemed "unsplittable" due to the lack of functionally dispensable site, such as the aptamers against tobramycin, neomycin, histidine, sulforhodamine,

ampicillin and many others in the aptamer database.^{13, 14, 15, 16} It is generally acknowledged that the integrity of conserved motifs and the proximity of essential nucleotides to the binding target in the parent aptamer is reserved during the split (referred to as "reservation rule" for simplicity), which suggest SPAs can only be created by splitting at the functionally dispensable site of the parent aptamer.^{17, 18} The conventional SPAs with **Reserved** initial small-molecule binding **Pockets** (^{RP}SPA) are created in accordance with the "reservation rule" by splitting the parent aptamer at the functionally dispensable loops that are distant from the target binding pocket (**Fig. S1**, ^{RP}SPA). However, we speculate that this empirical "reservation rule", although taken for granted, is unjustified. We hypothesize that the parent aptamer can be cleaved into SPAs at the functionally essential site while preserving its biorecognition capability.

To prove our hypothesis, we engineer three SPAs (**Table S2**) that are cleaved at the functionally essential sites with **Broken** initial small-molecule binding **Pockets** (^{BP}SPA, **Fig. 1A**). We first engineer a three-way junction (3WJ) aptamer against tetracycline (TC) into SPAs. The binding capability of the ^{BP}SPA against TC (^{BP}SPA_{TC}), the ^{RP}SPA against TC (^{RP}SPA_{TC}) and the wild-type TC aptamer (wt-TCA) is evaluated using isothermal titration calorimetry (ITC). The binding-induced assembly of ^{BP}SPA_{TC} and the formation of new binding pocket are then validated by Förster resonance energy transfer (FRET) and molecular dynamics (MD) simulation. We also engineer two previously "unsplittable" hairpin aptamers against tobramycin (TOB) and neomycin (NEO) into ^{BP}SPAs and successfully preserve their biorecognition capability. Next, an enzyme-linked apta-sorbent assay (ELASA) against TOB that uses ^{BP}SPA as the affinity probe is developed and validated. Our results provide deep insights into the dynamic adaptive target binding process of engineered aptamers and shed light on the MD-aided design of functional nucleic acids such as SPAs. This work lays the foundation for the engineering mechanism of ^{BP}SPAs that defy the

conventional wisdom and greatly extends the applicability of functional nucleic acids in sensing and monitoring in the near future.

RESULTS AND DISCUSSIONS

Binding Capability of ^{BP}SPA, ^{RP}SPA and wt-Aptamer

Three-way junction (3WJ) represents an important class of nucleic acid structure in aptamers with promising therapeutic and biosensing applications.¹⁹ Investigating the dynamic target binding and post-SELEX engineering of 3WJ-structured aptamers would have important implications for many functional nucleic acid-related systems. Besides, 3WJ is considered a “privileged architecture” for SPA engineering for its availability of functionally dispensable loops.¹⁸ Here, a 3WJ-structured TC aptamer (TCA) with well-characterized target binding properties²⁰ is used as a model system for investigation. As shown in **Fig. S2A**, the secondary structure of wild type TCA (wt-TCA) comprises three helices (P1, P2 and P3), two loops (L2 and L3) and two jointing regions (J1/2 and J2/3). Several nucleotides located in J1/2 and L3 are essential for the high-affinity target binding (labeled with red boxes). The TCA/TC cocrystal structure has been determined using a circularly permuted TCA (cp-TCA, **Fig. S2B**).²¹ In this structure, TC binds to a site at the interface between L3 and the irregular helix formed by J1/2 and J2/3 (**Fig. S2C**). Evidently, L3 plays a crucial role in the TC binding, thus its integrity is strictly maintained according to the “reservation rule”. Therefore, the traditional ^{RP}SPA_{TC} is obtained by cleaving the parent aptamer, *i.e.* wt-TCA, at the functionally dispensable P1 and P2 regions (**Fig. S2D**). But according to our hypothesis, the ^{BP}SPA_{TC} obtained by splitting at L3 (**Fig. 1B**) could also preserve the biorecognition capability of the parent aptamer. This unconventional splitting violates the “reservation rule” and impairs the initial binding pocket, but may shed light on a new way of engineering SPAs.

We compared the binding affinity of our new ^{BP}SPA_{TC} with both the parent wt-TCA and the old-fashioned ^{RP}SPA_{TC} by measuring the dissociation constant (K_d) with ITC. By applying a one-site binding model, the measured K_d of ^{BP}SPA_{TC} was 101.5 ± 0.7 nM ($\Delta G = -9.54 \pm 0.001$ kcal/mol) (**Fig. 2A**). In the absence of either one of the SPA fragment ^{BP}SPA1_{TC} or ^{BP}SPA2_{TC}, no binding to TC was observed under the same titration condition (**Fig. S3**), which provided a strong evidence for the highly synergistic binding of the two ^{BP}SPA_{TC} fragments to TC. In comparison, the K_d of wt-TCA against TC was 130.5 ± 10.6 nM ($\Delta G = -9.34 \pm 0.05$ kcal/mol) (**Fig. S4**). The ^{BP}SPA_{TC} exhibited a slightly higher affinity against TC than wt-TCA, which could be attributed to its preorganized architecture in the absence of TC (> 70% double-stranded ^{BP}SPA_{TC} in equilibrium) and certain degree of conformational flexibility ($\Delta G = -18.81$ kcal/mol) as predicted by NUPACK. The K_d of ^{RP}SPA_{TC} was 45.0 ± 0.3 nM ($\Delta G = -10.03 \pm 0.02$ kcal/mol). Similar to the ^{BP}SPA_{TC}, the binding-induced heat effect was not observed in ITC in the absence of either one of the SPA fragment ^{RP}SPA1_{TC} or ^{RP}SPA2_{TC} (**Fig. S5**). The tight binding between TC and ^{RP}SPA_{TC} was impressive because most SPAs were found to have affinities at least one order of magnitude lower than that of their parent aptamers due to the trimolecular binding.²² The large percentage of well-formed preorganized structures as well as a higher flexibility of the split fragments might help to explain the superior affinity of this ^{RP}SPA_{TC}.

The ITC measurements also provided insights into the binding mechanism of the aptamer/TC complex by distinguishing the enthalpic (ΔH) and entropic (ΔS) contributions of the binding process. As listed in **Table S3**, all of the three aptamer/TC binding events featured favorable enthalpy changes ($\Delta H < 0$) and unfavorable entropy changes ($\Delta S < 0$). Therefore, the dominant driving forces of these binding events were attributed to the formation of new inter- and intra-molecular bonds, stacking effects and Van der Waals forces.

To better understand the interaction between the ^{BP}SPA_{TC} and TC, the distance between the two oligonucleotide termini at the splitting site (U18 in ^{BP}SPA1_{TC} and A1 in ^{BP}SPA2_{TC}) was investigated by FRET. As shown in **Fig. 2B**, ^{BP}SPA1_{TC} and ^{BP}SPA2_{TC} were labeled with a fluorophore (FAM, F) at U18 and a quencher (BHQ1, Q) at A1, respectively (^{BP}SPA_{TC}-F and ^{BP}SPA_{TC}-Q in **Table S2**). To rule out the false positive signals resulted from non-specific FRET, the fluorescent (FL) intensity of FAM was measured in the presence of TC at 2.5-fold molar excess. As shown in **Fig. S6**, the FL intensity of FAM was almost the same as that measured in the absence of TC, indicating no non-specific FRET. For each FRET assay, ^{BP}SPA_{TC}-F and ^{BP}SPA_{TC}-Q were heated and annealed prior to the addition of TC. Gradual decrease in FL intensity was observed during the annealing process (**Fig. S7**), which was consistent with the previous assumption of the formation of pre-organized structure by free ^{BP}SPA_{TC}. The annealed ^{BP}SPA_{TC}-F and ^{BP}SPA_{TC}-Q (^{BP}SPA_{TC}-F/Q) was then allowed to bind to TC. In the presence of as little as 0.25 μM of TC, an evident decrease in FL intensity was observed, indicating the occurrence of the binding-induced FRET (**Fig. 2C**). In contrast, the annealed ^{BP}SPA_{TC}-F/Q showed a steady FL intensity against time in the absence of TC. The monotonous quantitative relationship between the FL intensity and TC concentration (**Fig. 2D**, **Fig. S8**) further reinforced the assumption that the two termini of the SPAs were brought into close proximity during the adaptive binding of TC to ^{BP}SPA_{TC}. In other words, TC accommodated itself to a new binding pocket in ^{BP}SPA_{TC}. The LOD toward TC is calculated to be 34 nM with a linear range from 96 nM to 3.5 μM. This ^{BP}SPA beacon exhibits a sensitivity comparable to that of a multi-step competitive enzyme-linked aptamer assay (ELAA) that uses the parent wt-TCA (LOD = 21 nM) as the affinity probe,²³ but takes much less time (30 min) than ELAA (over 18 h).

Following the reservation rule, traditional SPAs are obtained by splitting at the functionally dispensable loop of the parent aptamer. To the best of our knowledge, almost all SPAs reported up to date (**Table S1**) obey this empirical "reservation rule". As a result, a large number of existing aptamers are deemed "unsplittable" because they do not carry any functionally dispensable loops. Here we present the evidence that unconventional ^{BP}SPAs that are engineered by splitting at the functionally essential loop preserve their biorecognition capability. Remarkably, ^{BP}SPA_{TC} has exhibited a slightly higher affinity ($K_d = 101.5 \pm 0.7$ nM) than its parent aptamer ($K_d = 130.5 \pm 10.6$ nM), which may be attributed to the formation of ^{BP}SPA_{TC} pre-organized structure and the higher flexibility of the split fragments. Our study proves the feasibility of engineering SPAs by splitting at the functionally essential loops, which suggests that the empirical "reservation rule" is no longer valid in the development of reliable SPAs. Thus, our study has significantly expanded the boundary of SPA engineering, and a large number of new ^{BP}SPAs can be expected in the near future.

MD Simulations of SPA Binding Mechanism

In order to examine the binding mechanisms of the SPA_{TC}, all-atom MD simulations were conducted to provide a dynamic view of the induced-fit process of the TC binding to SPA_{TC} at the molecular level. MD simulations of ^{BP}SPA_{TC} with the TC molecule (lasting 1 μ s) or without the TC molecule (lasting 4 μ s) were carried out with a focus on the dynamic binding of TC to ^{BP}SPA_{TC}. For each simulation system, two independent MD simulations were performed with different initial velocities. The weak root-mean-square deviation (RMSD) is an accepted index to measure the conformational changes of the MD simulation systems.

Figs. 3A and B display the time evolution of RMSDs for ^{BP}SPA_{TC} in the simulations without and with TC, respectively. Globally, a plateau is reached after a certain duration (about 0.5 μ s in

no-TC simulations, and about 2.7 μ s in with-TC simulations) in each trajectory, indicating that the simulation systems have reached the equilibrium. In the representative trajectory, the RMSD of TC shows a large variation ranging from ~ 30 Å to a final stable value of ~ 2 Å, indicating a huge difference between the final stable TC binding conformation and the initial TC binding conformation. Snapshots of $^{BP}SPA_{TC}$ structures at different time points also show the allosteric binding behavior of TC. The TC molecule is initially placed in the original binding pocket revealed by the crystal structure of the parent aptamer. Subsequently, the binding-induced conformational reorganization of the $^{BP}SPA_{TC}$ makes the TC molecule hop from the original pocket to a new and more stable binding pocket (**Fig. 3B**). As shown in **Fig. 3C**, the dynamic change in distance between the TC molecule and nucleotide residues implies the allosteric binding behavior of the TC molecule to $^{BP}SPA_{TC}$. The distances between the TC molecule and the nucleotides residues (U11 and U18 from $^{BP}SPA1_{TC}$, A1 and C2 from $^{BP}SPA2_{TC}$) exhibit pronounced fluctuations. Over the period between 0 to 2.7 μ s, $^{BP}SPA_{TC}$ undergoes allosteric reconfiguration until the aptamer structure stabilizes and brings the TC molecule close to the nucleotide residues, leading to the formation of a new binding pocket (**Fig. 3C, left**). This new pocket is stabilized by hydrogen bonds between the TC molecule and the nucleotide residues in $^{BP}SPA_{TC}$ (**Fig. 3C, right**). As a result, both the binding pocket and the essential nucleotide residues that interact with the TC molecules are different in the $^{BP}SPA_{TC}$ and $^{RP}SPA_{TC}$ derived from the same parent aptamer.

It is noteworthy that $^{BP}SPA_{TC}$ experience a weaker enthalpy-driven force ($\Delta H = -24.3$ kcal/mol) than that of $^{RP}SPA_{TC}$ ($\Delta H = -32.3$ kcal/mol) (**Table S3**), which may explain its longer calculation time (μ s level, **Fig. 3B**) costed in MD simulations. Furthermore, the binding-induced coupling of the two $^{BP}SPA_{TC}$ fragments can also be explained by the MD simulation results. In the presence of TC, the average distance between U18 and A1 decreases gradually to less than 10 Å

191 in the stable ^{BP}SPA_{TC}/TC complex (**Figs. S9A** and **B**), indicating that FAM is brought into close
192 proximity of the quencher which suppresses the FL signals due to FRET. In contrast, in the absence
193 of TC, the distance between U18 and A1 is dynamic, and the average distance is greater than 18
194 Å (**Figs. S9A** and **C**). In such a case, the FL quenching effect is not prominent as the distance
195 between the fluorophore and the quencher is too large.

196 The MD simulation of the ^{RP}SPA_{TC}/TC system is also conducted. As shown in **Fig. S10A**, the
197 RMSD values of ^{RP}SPA_{TC} and TC in the 200-ns MD simulation indicate that there is no large
198 conformational variation upon the target binding. The representative snapshots from the simulation
199 is presented in **Fig. S10B**. In both the initial and final snapshots, we notice that the TC molecule
200 interacts with the same nucleotide residues of ^{RP}SPA_{TC}, and the distance between them remains
201 constant. All the results above indicate that the TC molecule binds stably to ^{RP}SPA_{TC} in the
202 conventional pocket, which is in consistent with the crystal structure.²¹The *in silico* investigations
203 above complement the experimental studies and elucidate the mechanisms of the dynamic adaptive
204 target binding of engineered aptamers. MD is a useful tool for predicting and interpreting
205 biological processes.²⁴ It has been recently applied to the studies of aptamers on their
206 conformational dynamic properties²⁵ as well as ligand recognition and binding²⁶ at atomic level.
207 For the first time, the formation of a new binding pocket in the SPAs has been discovered using
208 MD simulations, providing the critical evidence to the contrary of the empirical "reservation rule"
209 that is strictly abided by traditional SPA design. Notably, the MD simulation results agree well
210 with the ITC and FRET experiments. Overall, our work highlights the important role of MD
211 simulations in SPAs engineering, and implies a broader applicability of MD simulations to
212 monitoring the structures and the dynamics of other functional nucleic acids, and to predicting the
213 ligand binding potency at the molecular level.

^{BP}SPA from Previously "Unsplittable" Aptamers

Besides 3WJ, hairpin is another dominant secondary structure in aptamers.²⁷ Many aptamers with hairpin structures are deemed "unsplittable" according to the traditional "reservation rule" due to the lack of functionally dispensable loops. For example, the two hairpin aptamers against tobramycin (TOB) and neomycin (NEO) shown in in **Figs. 4A** and **B** contain only one loop which is functionally involved in the target binding. Therefore, they cannot be split into ^{BP}SAPs from the traditional point of view because the "reservation rule" cannot be satisfied. Nevertheless, according to our hypothesis, these aptamers can be engineered into ^{BP}SPAs while preserving their biorecognition capability by splitting at the functionally essential loop (splitting site indicated by the red arrows in **Figs. 4A** and **B**). The K_d values of the two parent aptamers are at the low micromolar level as verified by ITC (**Fig. S11**). The engineered ^{BP}SPA_{TOB} and ^{BP}SPA_{NEO} show a reasonably good binding affinity to their respective target with a K_d value of $52 \pm 10 \mu\text{M}$ and $414 \pm 9 \mu\text{M}$, respectively (**Fig. 4C**). Although this increase in binding affinity has not been observed in TOB and NEO aptamers which has a low percentage of pre-organized structure, enhanced small-molecule binding may be achieved by elongating the stem region with a well-defined pre-organized structure.

The two ^{BP}SPAs were validated by MD simulations. Two simulation systems (^{BP}SPA_{TOB}/TOB and ^{BP}SPA_{NEO}/NEO) were built based on the reported crystal structure (PDB ID: 2TOB and 1NEM). Three independent MD simulations were performed for each system with different initial velocities. By comparing the initial and the final configuration, we found that both TOB and NEO bound steadily to the binding pocket of the ^{BP}SPAs (**Fig. 4D** and **E**). We also monitored the contact distances between the small-molecule ligands and the essential nucleotide residue. The contact

distances stayed at a constant level below 4 Å in the MD simulations, suggesting that TOB and NEO bound stably to their ^{BP}SPAs.

^{BP}SPA-Based Enzyme-Linked Apta-Sorbent Assay (ELASA)

Aptamers have compelling advantages over antibodies as the affinity probes for small molecules with limited number of epitopes. We applied the ^{BP}SPA developed by us to a sandwich-type aptasensing assay for the detection of the small molecule target TOB. In the proposed ELASA (**Fig. 5A**), ^{BP}SPA1_{TOB} with an amine-modified 12-carbon flexible linker (Amine-^{BP}SPA1_{TOB}) was covalently conjugated to a maleic anhydride-activated 96-well plate. In the presence of TOB, the biotinylated Bio-^{BP}SPA2_{TOB} and the surface-immobilized ^{BP}SPA1_{TOB} coupled into ^{BP}SPA_{TOB} and formed a sandwich assembly with TOB. The biotinylated Bio-^{BP}SPA2_{TOB} captures the peroxidase-streptavidin (HRP-SA) conjugate via the strong biotin-SA interaction, which in turn catalyzed the development of the chromogenic substrate tetramethylbenzidine (TMB). In contrast, in the absence of TOB, the binding between Bio-^{BP}SPA2_{TOB} and the surface-immobilized ^{BP}SPA1_{TOB} is too weak to stand the washing process, resulting in no color development. The reaction was stopped by adding a strong acid, and the acidified TMB products appeared yellow with a peak absorbance at 450 nm (A₄₅₀). A standard curve was obtained by plotting the A₄₅₀ against the TOB concentration. A naked-eye limit of detection (LOD) toward TOB down to 1 μM was achieved using the ^{BP}SPA-based ELASA. It is noteworthy that the A₄₅₀ signal increases by 114% in the presence of 1 μM of TOB, whereas an electrochemical sensor that employs the full-length parent TOB aptamer shows a merely 1.3% signal change upon the addition of 1 μM of TOB.²⁸ The sensitivity of this TOB ELASA is quite impressive considering that the *K_d* of ^{BP}SPA_{TOB} is only 52 ± 10 μM and the fact that this method is based purely on the affinitive interaction between the ^{BP}SPA and TOB, whereas

earlier reported aptamer-based enzyme-linked assays require other types of affinity probe²⁹ or covalent SPA ligation.³⁰ In comparison, the ^{BP}SPA_{TOB}-based ELASA is simple in both design and operation, and can be extended to other types of aptamer-based assays for biosensing and clinical diagnostics.

Currently, there is an urgent need for facile screening and quantification of small molecules in many fields. However, it is extremely challenging to develop a sandwich-type small-molecule sensor using traditional antibodies as the affinity probe, because rarely could a small molecule be recognized by two antibodies concurrently due to the lack of epitopes. Engineered from *in vitro*-selected functional nucleic acids, new ^{BP}SPAs developed in this work serve as a new class of bio-affinity probe that holds a great promise in the sandwich-type small-molecule assay. Our results have proved that ^{BP}SPA is a competent affinity probe for sandwich-type ELASA-based small molecule sensing, which may spur further applications in other sensing modalities for facile small-molecule analysis.

CONCLUSIONS

Although SPA is an attractive affinity probe in many sensing applications, the SAP engineering technology is still in its infancy. In the traditional view of aptamer engineering, the chemical environment of nucleotides/motifs that are functionally essential should be changed as little as possible in order to guarantee the successful reconstruction of the initial target binding pocket (i.e., the "reservation rule").³¹ We have challenged this conventional theory on SPA engineering that dictates strict abidance to the "reservation rule" which asserts that SPAs can only be generated by splitting at the functionally dispensable site. In this work, we have hypothesized and proved that SPAs can be engineered by splitting at the functionally essential sites while preserving the biorecognition capability of the aptamers. The ^{BP}SPAs designed based on our new

theory have been validated both in wet lab and *in silico*. For the first time, the mechanism of small-molecule binding to ^{BP}SPAs ~~has been discovered using~~ is observed based on the all-atom MD simulations, which suggest that ^{BP}SPA would undergo a conformational change and form a new binding pocket in the presence of the small-molecule target. We have demonstrated the ^{BP}SPA “aptamer beacon” for the detection of small molecules by conjugating the two SPA fragments with a fluorescent report and a quencher. Moreover, a facile sandwich-type small-molecule sensing method has been established using ^{BP}SPA as the affinity probe. These findings expand our knowledge on aptamer engineering and suggest a broader role of MD simulation in the prediction of nucleic acid functions than previously recognized.

METHODS

Materials and Assay Protocols

Detailed information on chemical reagents, instruments and MD simulation conditions are provided in the *SI* text.

Important aspects of ITC, FRET and ELASA are briefly described in this section. Detailed protocols of these experiments are available in the *SI* text.

Statistical Information

In this work, error bars in all figures represent the standard deviations from three individual experiments.

MD Simulations

The simulation systems of SPA_{TC} were built based on the X-ray structure of cp-TCA (PDB ID: 3EGZ),²¹ where TC binds to the pseudoknot-like fold interface as a magnesium ion chelator. Based

on the X-ray coordinate of cp-TCA, three simulation systems were built: (i) ^{BP}SPA_{TC} free-state system. (ii) ^{BP}SPA_{TC}/TC system. and (iii) ^{RP}SPA_{TC}/TC system. All MD simulations were performed using the GROMACS 5.1 package³² with the isobaric-isothermal (NPT) ensemble and periodic boundary condition. The CHARMM36 force field³³ was applied, and the force-field parameters of small-molecules TC, TOB, and NEO in the simulations were generated based on the CHARMM General Force Field (CGenFF).³⁴ Energy minimizations were first performed to relieve unfavorable contacts, followed by equilibration steps of 11 ns in total to equilibrate the solvent with restraints (isotropic force constant $\kappa = 1 \times 10^3 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$) on TC and the phosphorus atoms in SPA_{TC}.

For each equilibrated system, at least two replicates of all-atom MD simulations were generated with different initial velocities. The temperature of each system was maintained at 300 K using the v-rescale method³⁵ with a coupling time of 0.1 ps. The pressure was kept at 1 bar using the Parrinello-Rahman³⁶ with $\tau_p = 2.0$ ps and a compressibility of $4.5 \times 10^{-5} \text{ bar}^{-1}$. LINCS³⁷ constraints were applied to the hydrogen-involving covalent bonds in water molecules and other molecules, respectively, and the time step was set to 2 fs. Electrostatic interactions were calculated with the Particle-Mesh Ewald (PME) algorithm³⁸ with a real-space cutoff of 1.2 nm. Van der Waals interactions used a 1.0-nm cutoff. Analysis of the trajectories was performed using Gromacs analysis tools. PyMOL (Version 1.3, Schrödinger, LLC) was used to visualize the structural models and generate figures.

Isothermal Titration Calorimetry (ITC)

Each titration procedure consisted of 20 successive 2- μL injections of the ligand every 150 s (the first injection volume was 0.5 μL) to the aptamer solutions. The temperature was kept at 25 °C.

TC-Induced FRET.

To investigate the effect of TC molecules on the FL intensity of FAM, FAM and TC were gently mixed and incubated for different durations before the FL measurements. Unless otherwise mentioned, all experimental results were based on measurements from at least three separate experiments. ^{BP}SPA_{TC}-F and ^{BP}SPA_{TC}-Q were annealed in advance to ensure the formation of pre-organized ^{BP}SPA_{TC} structure.

ELASA.

The amine-modified Amine-^{BP}SPA1_{TOB} strands were immobilized to the maleic anhydride-activated plate. For ELASA-based detection, Bio-^{BP}SPA2_{TOB} and TOB solutions of different concentrations were mixed and incubated in the TOB binding buffer in each well for 1 h. Next, each well was successively incubated with SA-HRP and TMB solutions for 1 h each to ensure SA-biotin binding and TMB colorimetric reaction, respectively. In the end, H₂SO₄ was added as stop solution. The absorbance was measured at 450 nm using a microplate reader.

Acknowledgments

The authors acknowledge Prof. Ai-Qun Liu from Nanyang Technological University and Prof. Xinhui Lou from Capital Normal University for valuable discussions. This research was funded by National Nature Science Foundation of China (21677082), Shanghai Sailing Program (18YF1406600). We acknowledge the Tsinghua University Branch of China National Center for Protein Sciences for providing the facility support. We also thank the Computer Center of East China Normal University for computational resources.

Author Contributions

R.W., Q.Z. and X.Z. designed research; R.W. and Q.Z. performed research; R.W., Q.Z., X.Z. and H.S. analyzed data; Q.Z., Y.Z and X.Z discussed the results; R.W., Q.Z., N.K.T, Y.Z and X.Z. wrote the paper. All authors have given approval to the final version of the manuscript, and declare no competing financial interest.

Additional information

Supplementary Information accompanies this paper is available.

Competing interests: The authors declare no competing interests.

References

1. Li, H., Dauphin-Ducharme, P., Ortega, G., Plaxco, K.W. Calibration-free electrochemical biosensors supporting accurate molecular measurements directly in undiluted whole blood. *J. Am. Chem. Soc.* **139**, 11207-11213 (2017).
2. Petrone, J., DeFrancesco, L. Small molecules get the message. *Nat. Biotechnol.* **36**, 787-790 (2018).
3. Warner, K.D., Hajdin, C.E., Weeks, K.M. Principles for targeting RNA with drug-like small molecules. *Nat. Rev. Drug Disco.* **17**, 547-558 (2018).
4. Meng, H.M., Liu, H., Kuai, H., Peng, R., Mo, L., Zhang, X.B. Aptamer-integrated DNA nanostructures for biosensing, bioimaging and cancer therapy. *Chem. Soc. Rev.* **45**, 2583-2602 (2016).
5. Gray, B.P., *et al.* Tunable cytotoxic aptamer-drug conjugates for the treatment of prostate cancer. *Proc. Natl. Acad. Sci. USA.* **115**, 4761-4766 (2018).
6. Chen, A., Yan, M., Yang, S. Split aptamers and their applications in sandwich aptasensors. *TrAC-Trend. Anal. Chem.* **80**, 581-593 (2016).
7. Wang, Z., *et al.* Introducing structure-switching functionality into small-molecule-binding aptamers via nuclease-directed truncation. *Nucleic Acids Res.* **46**, e81 (2018).

8. Stojanovic, M.N., de Prada, P., Landry, D.W. Fluorescent sensors based on aptamer self-assembly. *J. Am. Chem. Soc.* **122**, 11547-11548 (2000).
9. Nick, T.A., *et al.* Stability of a split streptomycin binding aptamer. *J. Phys. Chem. B* **120**, 6479-6489 (2016).
10. Wang, Q., *et al.* Surface plasmon resonance detection of small molecule using split aptamer fragments. *Sensor. Actuat. B-Chem.* **156**, 893-898 (2011).
11. Zuo, X., Xiao, Y., Plaxco, K.W. High specificity, electrochemical sandwich assays based on single aptamer sequences and suitable for the direct detection of small-molecule targets in blood and other complex matrices. *J. Am. Chem. Soc.* **131**, 6944-6945 (2009).
12. Zhou, J., Rossi, J. Aptamers as targeted therapeutics: current potential and challenges. *Nat. Rev. Drug Disco.* **16**, 181-202 (2017).
13. Jiang, L., Patel, D.J. Solution structure of the tobramycin–RNA aptamer complex. *Nat. Struct. Mol. Biol.* **5**, 769-774 (1998).
14. Jiang, L., Majumdar, A., Hu, W., Jaishree, T., Xu, W., Patel, D.J. Saccharide–RNA recognition in a complex formed between neomycin B and an RNA aptamer. *Structure* **7**, 817-S817 (1999).
15. Song, K-M., Jeong, E., Jeon, W., Cho, M., Ban, C. Aptasensor for ampicillin using gold nanoparticle based dual fluorescence–colorimetric methods. *Anal. Bioanal. Chem.* **402**, 2153-2161 (2012).
16. Lee, J.F., Hesselberth, J.R., Meyers, L.A., Ellington, A.D. Aptamer database. *Nucleic Acids Res.* **32**, D95-D100 (2004).
17. Smith, K.D., Lipchock, S.V., Ames, T.D., Wang, J., Breaker, R.R., Strobel, S.A. Structural basis of ligand binding by a c-di-GMP riboswitch. *Nat. Struct. Mol. Biol.* **16**, 1218-1223 (2009).
18. Kent, A.D., Spiropulos, N.G., Heemstra, J.M. General approach for engineering small-molecule-binding DNA split aptamers. *Anal. Chem.* **85**, 9916-9923 (2013).
19. Shu, D., Shu, Y., Haque, F., Abdelmawla, S., Guo, P. Thermodynamically stable RNA three-way junction for constructing multifunctional nanoparticles for delivery of therapeutics. *Nat. Nanotechnol.* **6**, 658-667 (2011).
20. Müller, M., Weigand, J.E., Weichenrieder, O., Suess, B. Thermodynamic characterization of an engineered tetracycline-binding riboswitch. *Nucleic Acids Res.* **34**, 2607-2617 (2006).

21. Xiao, H., Edwards, T.E., Ferré-D'Amaré, A.R. Structural basis for specific, high-affinity tetracycline binding by an in vitro evolved aptamer and artificial riboswitch. *Chem. Biol.* **15**, 1125-1137 (2008).
22. Zou, R., *et al.* Highly specific triple-fragment aptamer for optical detection of cocaine. *RSC Adv.* **2**, 4636-4638 (2012).
23. Jeong, S., Rhee Paeng, I. Sensitivity and selectivity on aptamer-based assay: the determination of tetracycline residue in bovine milk. *The Scientific World J.* **2012**, Article ID 159456 (2012).
24. Tan, D., Piana, S., Dirks, R.M., Shaw, D.E. RNA force field with accuracy comparable to state-of-the-art protein force fields. *Proc. Natl. Acad. Sci. USA.* E1346-E1355 (2018).
25. Bottaro, S., Bussi, G., Kennedy, S.D., Turner, D.H., Lindorff-Larsen, K. Conformational ensembles of RNA oligonucleotides from integrating NMR and molecular simulations. *Sci. Adv.* **4**, eaar8521 (2018).
26. Sun, L.Z., Zhang, D., Chen, S.J. Theory and modeling of RNA structure and interactions with metal ions and small molecules. *Ann. Rev. Biophys.* **46**, 227-246 (2017).
27. Durand, G., *et al.* A combinatorial approach to the repertoire of RNA kissing motifs; towards multiplex detection by switching hairpin aptamers. *Nucleic Acids Res.* **44**, 4450-4459 (2016).
28. Schoukroun-Barnes, L.R., Wagan, S., White, R.J. Enhancing the analytical performance of electrochemical RNA aptamer-based sensors for sensitive detection of aminoglycoside antibiotics. *Anal. Chem.* **86**, 1131-1137 (2014).
29. Toh, S.Y., Citartan, M., Gopinath, S.C., Tang, T-H. Aptamers as a replacement for antibodies in enzyme-linked immunosorbent assay. *Biosen. Bioelectron.* **64**, 392-403 (2015).
30. Sharma, A.K., Kent, A.D., Heemstra, J.M. Enzyme-linked small-molecule detection using split aptamer ligation. *Anal. Chem.* **84**, 6104-6109 (2012).
31. Long, S.B., Long, M.B., White, R.R., Sullenger, B.A. Crystal structure of an RNA aptamer bound to thrombin. *RNA* **14**, 2504-2512 (2008).
32. Abraham, M.J., *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1**, 19-25 (2015).
33. Klauda, J.B., *et al.* Update of the CHARMM all-atom additive force field for lipids: Validation on six lipid types. *J. Phy. Chem. B* **114**, 7830-7843 (2010).

- 474 34. Vanommeslaeghe, K., *et al.* CHARMM general force field: A force field for drug-like
475 molecules compatible with the CHARMM all-atom additive biological force fields. *J.*
476 *Comput. Chem.* **31**, 671-690 (2010).
477
- 478 35. Bussi, G., Donadio, D., Parrinello, M. Canonical sampling through velocity rescaling. *J.*
479 *Chem. Phys.* **126**, (2007).
480
- 481 36. Parrinello, M., Rahman, A. Polymorphic transitions in single crystals: A new molecular
482 dynamics method. *J. Appl. Phys.* **52**, 7182-7190 (1981).
483
- 484 37. Hess, B., Bekker, H., Berendsen, H.J.C., Fraaije, J.G.E.M. LINCS: A linear constraint
485 solver for molecular simulations. *J. Comput. Chem.* **18**, 1463-1472 (1997).
486
- 487 38. Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H., Pedersen, L.G. A smooth
488 particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577-8593 (1995).
489
- 490 39. Rose, A.S., Hildebrand, P.W. NGL Viewer: a web application for molecular visualization.
491 *Nucleic Acids Res.* **43**, W576-W579 (2015).
492

Figures

Fig. 1. Schematic illustration of engineering SPAs with Broken initial small-molecule binding Pockets (^{BP}SPAs) from the parent aptamer. (A) Unconventional ^{BP}SPAs are created from “unsplittable” atpamers without functionally dispensable loop. (B) Secondary structure of ^{BP}SPA_{TC} by unconventional split. In ^{BP}SPA_{TC}, functionally dispensable loops like L2 and U1A are removed, and the P1 sequence from cp-TCA (AGGUA) were adopted to guarantee a more accurate MD starting point.

Fig. 2. Verification of the target-binding capability of $^{BP}SPA_{TC}$. (A) ITC. Upper panel: power required to maintain the temperature of $^{BP}SPA_{TC}$ solution (25 μM) in the cell recorded over a period of 20 injections of TC (0.25 mM) until saturation; Lower panel: integrated heat of injection against TC/ $^{BP}SPA_{TC}$ molar ratio applying the one-site binding model. (B) Schematic illustration of TC-induced FRET. (C) Normalized fluorescence (FL) of the $^{BP}SPA_{TC}$ -F/Q system against time in the presence and absence of TC. (D) The monotonous quantitative relationship between the FL and TC concentration fitted by the four-parameter logistic model ($R^2 = 0.992$, **Table S4**). Each data point represents the average value (with standard deviation bars) of triplicate experiments.

Fig. 3. Study of ^{BP}SPA-TC structural conformation and binding mechanism using MD simulations. (A) Time evolution of RMSD for ^{BP}SPA_{TC} in the simulations without TC, calculated based on all ^{BP}SPA_{TC} simulations that are aligned to the last snapshot. The first and last snapshots from the trajectories are shown around the RMSD plot. ^{BP}SPA_{1TC} and ^{BP}SPA_{2TC} are illustrated in red and blue, respectively. (B) Time evolution of RMSD for the TC molecule and the ^{BP}SPA_{TC} in the presence of TC. The representative snapshots of ^{BP}SPA_{TC} are shown as in (B), and the yellow spheres indicate the position of the TC molecule as it hops to the new pocket. (C) Characterization of the new TC binding pocket in ^{BP}SPA_{TC}. (left) The distances between TC and nucleotide residues

520 that interact electrostatically with TC during the simulation. (*right*) Highlight of the interaction
521 between the TC molecule and nucleotides in the new binding pocket observed in the last snapshot
522 from the ^{BP}SPA_{TC}-with-TC simulation. For clarity, only the TC molecule and the interacting
523 nucleotides are shown.

Fig. 4. ^{BP}SPAs against tobramycin (TOB) and neomycin (NEO) obtained by unconventional splitting. Secondary structures of aptamers against TOB (A) and NEO (B) and their ligand-bound crystal structures shown in NGL Viewer.³⁹ Two unconventional splitting sites are labeled with red

528 arrows. (C) ITC results of the two ^{BP}SPAs against their targets of TOB (*left*) and NEO (*right*)
529 under the conditions of 60 μM ^{BP}SPA, 3 mM TOB and 3 mM NEO. (D, E) Monitoring of TOB or
530 NEO bindings in ^{BP}SPA_{TOB} (D) or ^{BP}SPA_{NEO} (E) simulations, respectively. (*top*) the distances
531 change between TOB or NEO and the corresponding nucleotide residues during MD simulations.
532 (*bottom*) The first and last snapshots from the MD trajectories. ^{BP}SPA_{TOB} and ^{BP}SPA_{NEO} are shown
533 in red (chain 1) and blue (chain 2). The interacting nucleotide residues are shown as red and blue
534 sticks. Small molecules TOB and NEO are shown as yellow sticks.

Fig. 5. A $^{\text{BP}}\text{SPA}$ -based enzyme-linked apta-sorbent assay (ELASA) for tobramycin (TOB).

(A) Schematic illustration of the sensing principle. (B) Calibration curve of the developed ELASA against TOB fitted by the four-parameter logistic model ($R^2 = 0.996$, **Table S5**). Each data point represents the average A_{450} (with standard deviation bars) of triplicates. Inset: corresponding photographic images. The naked-eye LOD toward TOB is labeled with a red box.