

# Identification, Characterization and Application of a G-Quadruplex Structured DNA Aptamer against Cancer Biomarker Protein Anterior Gradient Homolog 2

Jie Wu<sup>1</sup>✉, Chi Wang<sup>1</sup>✉, Xilan Li<sup>1</sup>, Yanling Song<sup>1</sup>, Wei Wang<sup>1</sup>, Cong Li<sup>1</sup>, Jia Hu<sup>1</sup>, Zhi Zhu<sup>1\*</sup>, Jiuxing Li<sup>1</sup>, Weiyun Zhang<sup>1</sup>, Zhongxian Lu<sup>2\*</sup>, Chaoyong James Yang<sup>1\*</sup>

**1** State Key Laboratory of Physical Chemistry of Solid Surface, Key Laboratory for Chemical Biology of Fujian Province, Key Laboratory of Analytical Sciences, and Department of Chemical Biology, College of Chemistry and Chemical Engineering, Xiamen University, Xiamen, People's Republic of China, **2** Xiamen City Key Laboratory of Metabolism Disease, School of Pharmaceutical Sciences, Xiamen University, Xiamen, China

## Abstract

**Background:** Anterior gradient homolog 2 (AGR2) is a functional protein with critical roles in a diverse range of biological systems, including vertebrate tissue development, inflammatory tissue injury responses, and cancer progression. Clinical studies have shown that the AGR2 protein is overexpressed in a wide range of human cancers, including carcinomas of the esophagus, pancreas, breast, prostate, and lung, making the protein as a potential cancer biomarker. However, the general biochemical functions of AGR2 in human cells remain undefined, and the signaling mechanisms that drive AGR2 to inhibit p53 are still not clearly illustrated. Therefore, it is of great interest to develop molecular probes specifically recognizing AGR2 for its detection and for the elucidation of AGR2-associated molecular mechanism.

**Methodology/Principal Findings:** Through a bead-based and flow cytometry monitored SELEX technology, we have identified a group of DNA aptamers that can specifically bind to AGR2 with  $K_d$  values in the nanomolar range after 14 rounds of selections. Aptamer C14B was chosen to further study, due to its high binding affinity and specificity. The optimized and shortened C14B1 has special G-rich characteristics, and the G-rich region of this binding motif was further characterized to reveal an intramolecular parallel G-quadruplex by CD spectroscopy and UV spectroscopy. Our experiments confirmed that the stability of the G-quadruplex structure was strongly dependent on the nature of the monovalent ions and the formation of G-quadruplex structure was also important for the binding capacity of C14B1 to the target. Furthermore, we have designed a kind of allosteric molecule beacon (aMB) probe for selective and sensitive detection of AGR2.

**Conclusion/Significance:** In this work, we have developed new aptamer probes for specific recognition of the AGR2. Structural study have identified that the binding motif of aptamer is an intramolecular parallel G-quadruplex structure and its structure and binding affinity are strongly dependent on the nature of the monovalent ion. Furthermore, with our design of AGR2-aMB, AGR2 could be sensitively and selectively detected. This aptamer probe has great potential to serve as a useful tool for early diagnosis and prognosis of cancer and for fundamental research to elucidate the biochemical functions of AGR2.

**Citation:** Wu J, Wang C, Li X, Song Y, Wang W, et al. (2012) Identification, Characterization and Application of a G-Quadruplex Structured DNA Aptamer against Cancer Biomarker Protein Anterior Gradient Homolog 2. PLoS ONE 7(9): e46393. doi:10.1371/journal.pone.0046393

**Editor:** Pierre Busson, Institute of Cancerology Gustave Roussy, France

**Received:** June 2, 2012; **Accepted:** August 29, 2012; **Published:** September 28, 2012

**Copyright:** © 2012 Wu et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by the National Basic Research Program of China (2010CB732402), National Instrumentation Program (2011YQ03012412), the Natural Science Foundation of Fujian Province for Distinguished Young Scholars (2010 J06004), and National Found for Fostering Talents of Basic Science (J1030415). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: cyyang@xmu.edu.cn (CY); zhuzhi@xmu.edu.cn (ZZ); zhongxian@xmu.edu.cn (ZL)

✉ These authors contributed equally to this work.

## Introduction

Anterior gradient homolog 2 (AGR2) was identified initially as a secretory factor expressed in the anterior region of the dorsal ectoderm in *Xenopus laevis* embryos, where it was postulated to mediate the specification of dorsoanterior ectodermal fate, particularly in the formation of the cement gland [1]. Clinical studies have further shown that the AGR2 protein is overexpressed in a wide range of human cancers, including carcinomas of the esophagus, pancreas, breast, prostate, and lung [2–6]. More

biological studies in these cancer cell lines have indicated a significant role for AGR2 in tumor-associated pathways, including tumor growth, cellular transformation, cell migration, limb regeneration, and metastasis [5,7–9]. However, the general biochemical functions of AGR2 in human cells remain undefined, and the signaling mechanisms that drive AGR2 to inhibit p53 are still not clearly illustrated [10]. Therefore, the development of molecular ligands specifically recognizing AGR2 is of great significance to early diagnosis and prognosis of cancer and to

fundamental research for the elucidation of the biochemical functions of AGR2.

Various ligands have been developed for specific molecular recognition, such as small molecules, antibodies, and peptides [11–13]. More recently, another type of molecular ligand, named aptamer, has drawn significant attention. Aptamers, single-stranded modified or unmodified oligonucleotides (RNA or DNA), are generated through *in vitro* selection process or SELEX (Systematic Evolution of Ligands by EXponential enrichment) with high binding affinity and specificity towards defined targets [14,15]. The selected aptamers can recognize a wide variety of targets, including small molecules, proteins, cells and tissues relying on their diverse tertiary structures. Compared to antibodies, aptamers have low molecular weight, fast tissue penetration rate, high stability and low immunogenesis [16]. They can be chemically synthesized with low cost and modified easily with various reporters [17]. Furthermore, they can be ligated and/or amplified by enzymes *in vitro* [18]. These advantages make aptamers promising ligands for medical and pharmaceutical research, such as drug development, disease diagnosis, and targeted therapy [19].

The possibilities provided by aptamers are enormous, and some aptamers have already shown many important applications in bioanalysis and biomedicine [20–23]. Particularly, several aptamers have been generated against cancer-related proteins, such as PDGF, VEGF, HER3, NF $\kappa$ B, tenascin-C, or PMSA [24–26]. Many aptameric sensors, probes and assays have been developed to allow sensitive and selective detection of these cancer biomarker proteins [27]. For instance, Yang *et al* has reported a light-switching excimer aptamer probes for sensitive quantitative detection of PDGF in cell media [28]. Kwon *et al* have developed a functionalized polypyrrole nanotube with aptamer to build a VEGF biosensor [29]. Aptamers have also been applied for molecular imaging to *in vivo* characterize the complex pathogenic activities that accompany tumor growth for disease early diagnosis and pathogenesis measurement [30–33]. Since the targets for aptamers could be intracellular, extracellular or cell-surface biomolecules, various therapeutic methods have been developed using the aptamers as targeting reagents [34–37], which greatly broaden the range of targeted therapy. In addition, some therapeutically useful aptamers have been found to inhibit protein–protein interactions, such as receptor–ligand interactions, and thereby function as antagonists [38].

In this study, using the bead-based and flow cytometry monitored SELEX technology, we aimed to obtain specific aptamers to AGR2 and study their structure and potential function. Bead-based SELEX allowed the use of simple, yet effective, flow cytometry analysis to monitor the progress of the selection, avoiding the tedious, time consuming and radioactive EMSA process [39–43]. After 14 rounds of selection, we have identified a group of DNA aptamers that specifically bound to AGR2 with high affinities. Structural studies on one of the aptamer sequences, C14B, revealed an intramolecular parallel G-quadruplex, and its structure and binding affinity to AGR2 depend on K<sup>+</sup> ion intensively. Furthermore, we designed an allosteric molecule beacon AGR2-aMB based on the identified aptamer, which enables simple, sensitive and selective detection of AGR2. The aptamer sequences and AGR2-aMB reported in this study are potentially useful tools for early diagnosis and prognosis of cancer and for fundamental research to elucidate the biochemical functions of AGR2.

## Results and Discussion

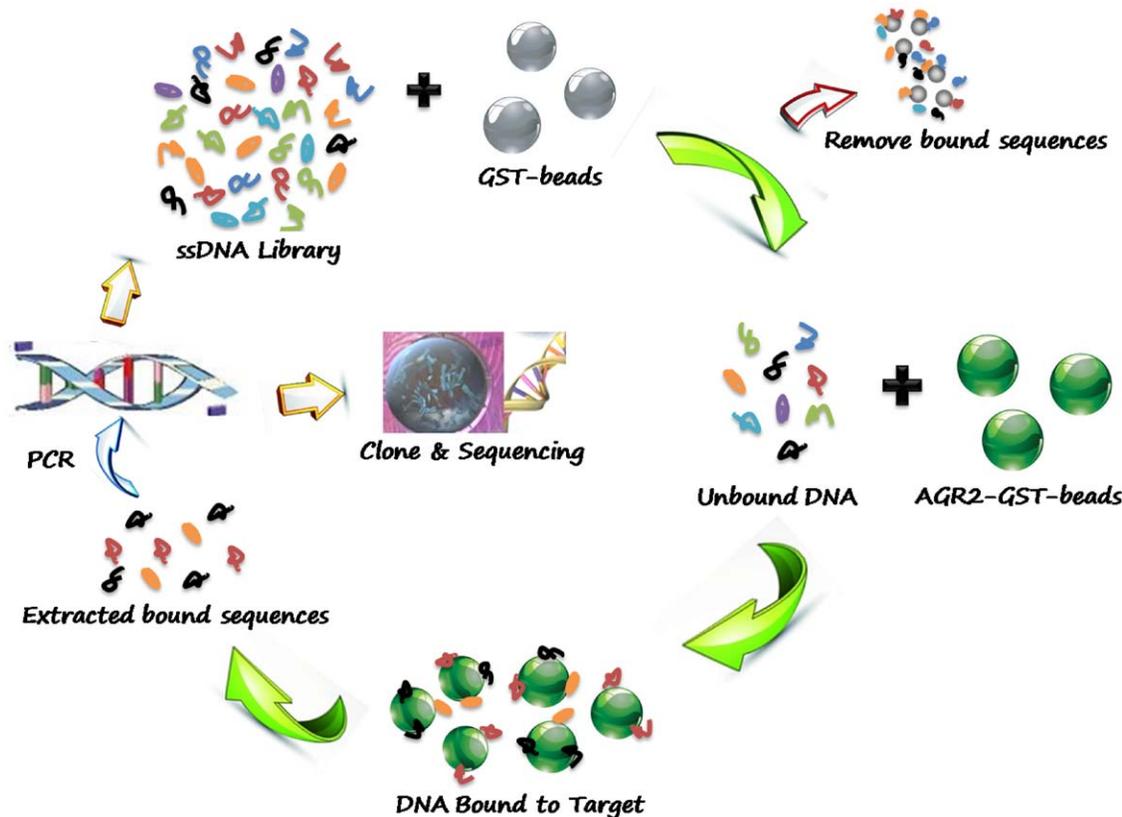
### Selection of DNA aptamers to recognize AGR2

To identify aptamers against AGR2, recombinant AGR2 was fused with glutathione-S-transferase (GST) to facilitate the attachment of the protein to solid supports (Sepharose GSH-beads). The resulting AGR2-GST-beads were used as the positive target in SELEX while the GST-beads as negative control to remove non-specific surface binding sequences. The process of *in vitro* sepharose-bead-based SELEX is schematically illustrated in Figure 1.

An 87-nucleotide (87-nt) single-stranded DNA (ssDNA) library with 45 random bases flanked by two primer sequences (22-nt and 20-nt) was subjected to the SELEX procedure. The library was first allowed to interact with excess negative control beads, and only the DNA sequences that did not bound to the GST-beads were collected. The collected sequences were then incubated with AGR2-GST-beads. After rigorous washing, those sequences that either did not bind, or only weakly bound to the target were discarded. Only the sequences that bound strongly enough were retained on beads, and the bead-ssDNA complexes were collected and amplified by PCR for the next round of selection. After multiple rounds of selection, the subtraction process efficiently reduced the DNA sequences that bound to the GST beads, while those AGR2-specific aptamer candidates were gradually enriched. The progress of the selection process was monitored by flow cytometry. The stronger binding of DNA library to AGR2, the more FAM labeled sequences bound to the beads, thus the higher fluorescence intensity the beads would emit. With the increasing number of selection cycles, steady increases in fluorescence intensity on the target beads were observed (Figure 2a). The binding affinity of the enriched library after 14 rounds of selection was determined to be in the nanomolar range ( $K_d = 64.1 \pm 5.4$  nM), while there was no observable binding of the library to control beads (Figure 2b). These results suggested that the DNA aptamers specifically recognizing AGR2 were enriched during the selection process.

### Characterization of selected aptamer sequences

After 14 rounds of selection, the enriched DNA pool was cloned and sequenced. The sequencing data for clones were analyzed by using sequencing analysis software Clustal W 6.0 [44]. The sequences were grouped based on the homology similarity of the DNA sequences from individual clone. Among the sequences from 62 clones, there were two subfamilies each containing multiple sequences with common features. One subfamily is guanosine-rich sequences (22 clones), and the other is thymine-rich sequences (40 clones) (Figure S1). Four sequences were chosen and synthesized for further characterization (Table S1): three sequences from G-rich subfamily, C14A (appeared 2 times), C14B (appeared 3 times) and C14C (appeared 5 times), and one sequence C14D (appeared 2 times) from T-rich subfamily. Only one T-rich sequence was chosen because of T-rich sequences tend to form non-rigid tertiary structures and the chance of being aptamer was thought to be very low. As shown in Figure 2c, C14B can bind AGR2 with the highest affinity ( $K_d = 13.1 \pm 7.2$  nM). Titration of C14B to GST-beads revealed no observable binding, establishing that the binding target of C14B was indeed AGR2. Other three sequences have similar binding constants to AGR2 but a little weaker than C14B, in which the  $K_d$  of C14A is  $20.9 \pm 5.2$  nM, C14C is  $44.6 \pm 7.0$  nM and C14D is  $48.4 \pm 15.6$  nM. (Figure S2).



**Figure 1. Schematics of systematic evolution of DNA aptamers against AGR2.** In sepharose-beads-based SELEX process for protein AGR2, the GST-beads were incubated with the ssDNA library for counter-selection to remove nonspecific sequences. The unbound DNAs were then incubated with AGR2-GST-beads for target-selection. After harshly washing, the AGR2-specific DNA sequences were subsequently amplified by PCR for the next round of selection, or for cloning and sequencing to identify individual aptamers after flow cytometry analysis. doi:10.1371/journal.pone.0046393.g001

### Sequence optimization of aptamer C14B

Since C14B is the best aptamer we obtained from the four sequences tested, it was chosen for future optimization and characterization. The length of C14B is 87mer, which is disadvantageous for future applications because it is inconvenient and expensive to synthesize such a long sequence. To identify the binding region of the aptamer, the marginal sequences of C14B were gradually truncated. Two truncated sequences C14B0 and C14B1 from C14B were shown in Table 1. Subsequent binding affinity experiments revealed that both C14B0 ( $8.5 \pm 3.6$  nM) and C14B1 ( $19.1 \pm 5.1$  nM) have similar  $K_d$  to AGR2 as that of original C14B ( $13.1 \pm 7.2$  nM). The majority of eliminated sequences were primer sequences, implying that the binding region of the aptamer was the middle of random sequences and the primer sequences do not or contribute little to the binding affinity of aptamer. C14B1 has five portions of poly-G and we designed five truncated sequences by removing one of poly-G portion each (Table S2). Removal of any poly-G portion would destroy its binding affinity to certain extent (Figure S3), suggesting the whole G-motif is required for aptamer binding. Taking together, these results indicate that C14B1, which was only 33 mer, was the essential binding region to AGR2. Thus, C14B1 was applied for further characterization and probe design.

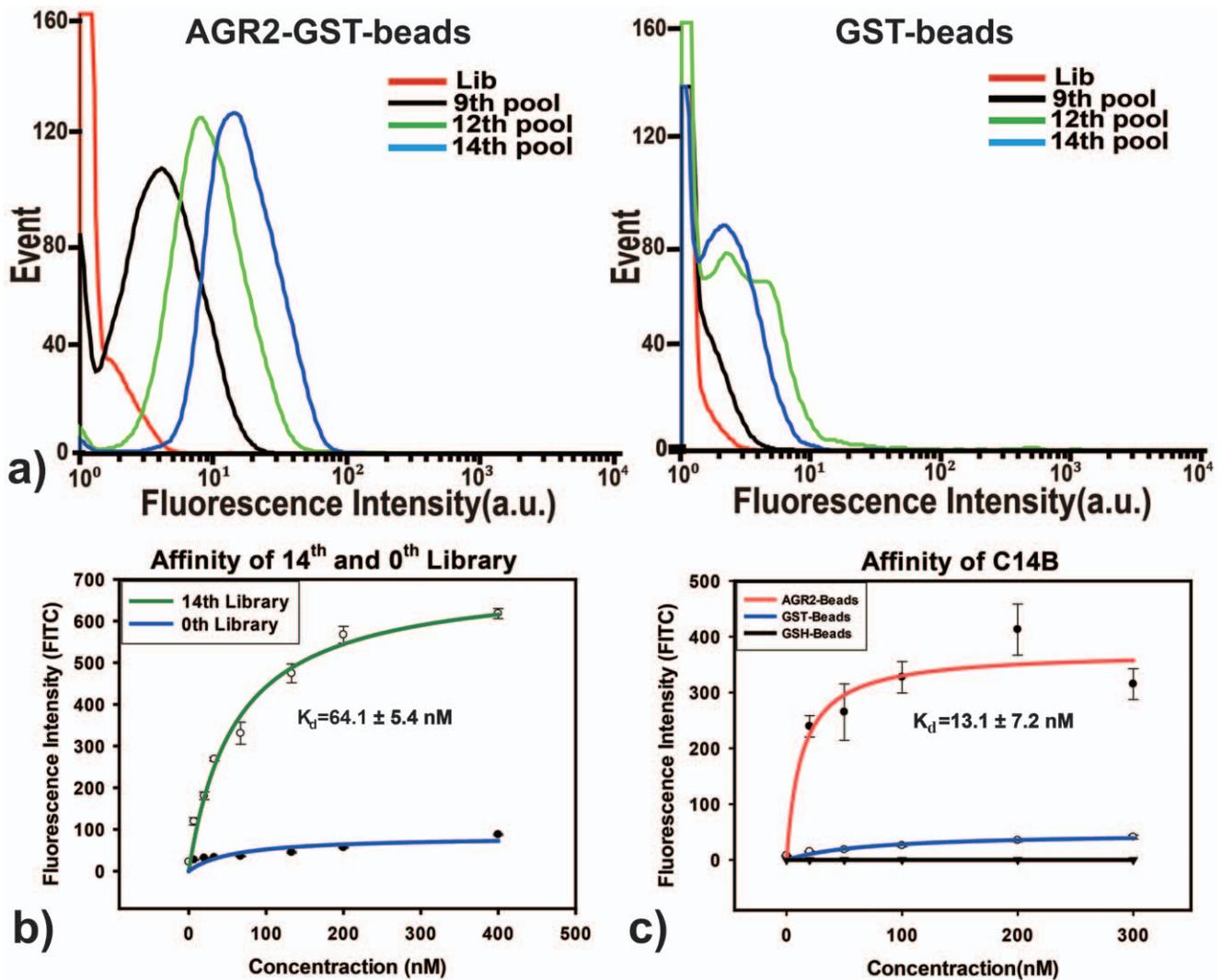
### Selectivity of Aptamer C14B1

In order to demonstrate the specific interaction between AGR2 and C14B1, three control proteins BSA, thrombin and trypsin

were coupled with the NHS-sepharose-beads and then were incubated with C14B1. As shown in Figure 3, C14B1 could bind to target AGR2 strongly, while there were no or little binding affinity towards thrombin, BSA and trypsin, demonstrating the high selectivity of the aptamer.

### Aptamer C14B1 is an Intramolecular Parallel G-quadruplex

Further examination revealed that C14B1 has multiple stretches of guanines (CGGGTGGGAGTTGTGGGGGGGGTGG-GAGGGT). It is well-known that guanine-rich sequences can fold into four-stranded secondary structures called quadruplexes. According to G-Quadruplex prediction formula  $d(G_{3+}N_{1-7}G_{3+}N_{1-7}G_{3+}N_{1-7}G_{3+})$  [45], C14B1 had great probability to form a G-Quadruplex structure. In the G-quadruplex, the planar square arrangement of four guanines (tetrad) is stabilized by Hoogsteen hydrogen bonding. Depending on the direction of the strands or parts of a strand that form the tetrads, structures may be described as parallel or antiparallel. Moreover, these quadruplexes can form through the intermolecular association of four or two DNA molecules, or by the intramolecular folding of a single strand containing four blocks of guanines [46] CD experiments were performed to investigate the secondary structure of aptamer C14B1. The CD spectrum of C14B1 showed a negative peak near 240 nm and a positive peak near 260 nm, a typical spectrum of a parallel G-quadruplex [47]. In order to investigate whether C14B1 forms an intermolecular 4 stranded or intramolecular



**Figure 2. Enrichment of DNA sequences that bind to AGR2-GST-beads.** a) Flow cytometry assay to monitor the binding of a selected pool with AGR2 (target protein) and GST (control protein). The red curve represents the background binding of unselected library. For the target protein AGR2, there was an increase in binding capacity of the pool as the selection progressed, while there was no observable change for the control protein GST. The final concentration of the selected pool in binding buffer was 100 nM. b) Determination of the dissociation constant of the enriched library and unselected library to AGR2. c) Fluorescence measurements to determine the dissociation constant of selected aptamer C14B. Two negative controls, GST-beads and GSH-beads, were carried out. The data were the average of triplicate experiment results. doi:10.1371/journal.pone.0046393.g002

single stranded G-quadruplex structure, the dependence of melting temperature on the concentration of C14B1 was studied [48]. The melting temperature of C14B1 was found to be 59°C, which was independent of oligonucleotide concentration (Figure S4), indicating that the aptamer forms an intramolecular G-quadruplex.

#### Binding Affinity and Structure Stability of C14B1 is $K^+$ Dependent

Many DNA aptamers have been found to form G-quadruplex structures [49–51]. It is well established that the existence of a monovalent cation (especially potassium) in the center of these tetrads can significantly stabilize G-quadruplexes [52–54]. Thus, we investigated how a cation would affect the structural stability of C14B1 and its binding activity towards AGR2. As shown in Figure 4a, the stability of the C14B1 G-quadruplex structure is strongly dependent on the presence of the monovalent ion. With

60 mM of KCl, strong CD peaks were observed, suggesting the formation of stable G-quadruplex structure. Replacing KCl with the same concentration of LiCl, NaCl, and  $MgCl_2$  led to a dramatic intensity decrease in CD.

We further studied the effect of  $K^+$  concentration on the stability of G-quadruplex. In Figure 4b, addition of 0.1 mM  $K^+$  in phosphate buffer dramatically increased the CD intensity at 240 and 260 nm. CD absorption intensity enhanced with the increase of  $K^+$  concentration and reached a plateau when  $K^+$  concentration was higher than 20 mM. The binding affinity of C14B1 at different concentrations of  $K^+$  was also investigated. As the flow cytometry results shown in Figure 4c, very weak binding of C14B1 towards AGR2-beads was observed when there was no  $K^+$  in the buffer, and the binding affinity kept increasing with the addition of  $K^+$ . The binding constants of C14B1 were measured and compared in the presence or absence of  $K^+$  (Figure 4d). The  $K_d$  value in the presence of  $K^+$  was determined to be  $6.2 \pm 1.9$  nM.

**Table 1.** The sequences of C14B, C14B0 and C14B1, and their  $K_d$  to AGR2.

Aptamer	Sequences	K <sub>d</sub> s
<b>C14B</b>	TCTCGGACGCGTGTGGT <b>CGGGTGGGAGTTGTGGGGGGGGTGGGAGGGTT</b> CTTTGTTTATCTTTCTCGCT GCCTGGCCCTAGAGTG	$K_d = 13.1 \pm 7.2$ nM
<b>C14B0</b>	<b>CGGGTGGGAGTTGTGGGGGGGGTGGGAGGGTT</b> CTTTGTTTATCTTTCTCGCTGCCTGGCCCTAGAGTG	$K_d = 8.5 \pm 3.6$ nM
<b>C14B1</b>	<b>CGGGTGGGAGTTGTGGGGGGGGTGGGAGGGTT</b>	$K_d = 19.1 \pm 5.1$ nM

doi:10.1371/journal.pone.0046393.t001

The results demonstrated that formation of G-quadruplex is important for the binding capability of aptamer C14B1 to its target, which is highly  $K^+$  dependent.

### Allosteric molecular beacons for detection AGR2

Clinical studies have already shown that the AGR2 protein is overexpressed in a wide range of human cancers. The sensitive and selective detection of AGR2 is thus of great importance to early cancer diagnostics. Herein, by applying the selected and optimized aptamer C14B1, we designed and developed an allosteric molecular beacon against AGR2, named AGR2-aMB, which converts the molecule recognition property of aptamer to fluorescence flow cytometry signal for AGR2 sensing [55]. Figure 5 illustrates the working principle of AGR2-aMB. An AGR2-aMB is a ssDNA consisting of an streptavidin (SA) aptamer sequence [56], a C14B1 sequence, a short sequence complimentary to a small part of the SA aptamer sequence, and a fluorophore. A stable hairpin structure is formed by intramolecular hybridization between the SA aptamer sequence and the complementary sequence, temporarily disabling the probe's ability to bind with SA beads. Consequently, when incubated with SA beads, no probe can bind to SA beads, and the beads display very low fluorescence. In the presence of AGR2, however, the C14B1 sequence in the loop of aMB binds to the target sequence, which in turn disrupts the hairpin structure to free the SA binding sequence, thereby activating the probe's binding affinity to SA beads. Thus, the AGR2-bound probe will bind to SA beads, which will fluoresce strongly because the probe is FAM labeled. Target molecules can

be detected and quantified by reading the fluorescence intensity of SA beads through flow cytometer.

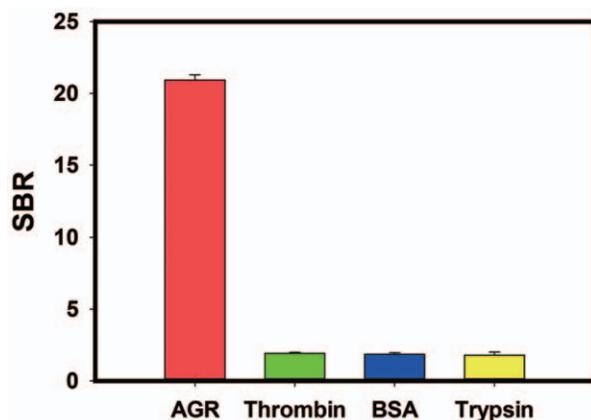
We designed AGR2-aMB, with the following sequence: **CGACGCACCGATCGCAGGTTCTGGGATTTTCGGGTGGGAGTTGTGGGGGGGGTGGGAGGGTT**-FAM, where the AGR2 binding region is underlined and the SA aptamer sequence is in bold. Stable hairpin structure is formed by intramolecular hybridization (Figure S5). In our experiment, when only AGR2-aMB incubated with SA beads, no probe can bind to SA beads, and the beads fluoresced very weakly. In the presence of AGR2, however, the fluorescence intensity of SA beads increase significantly, suggesting the binding of AGR2-bound probe to SA beads. With the increase of AGR2 concentration, steady increases in fluorescence intensity on the target beads were observed. AGR2 at 100 nM concentration can be easily detected (Figure 6a). A series of proteins including BSA, trypsin, thrombin, and IgG were employed as controls (500 nM), and no distinct fluorescent signals were observed (Figure 6b), indicating an excellent specificity of the AGR2-aMB towards target molecule. The results demonstrated the sensitivity and specificity of the allosteric molecular beacon probe, implying its potential for application in real sample analysis, such as protein function study and disease diagnosis.

In conclusion, we have developed new aptamer probes for specific recognition of the AGR2. In the SELEX process, AGR2-GST fused protein was used as the target protein, and linked to sepharose beads through the mild, yet specific, noncovalent GST-Glutathione interaction. Beads-based SELEX allowed the use of simple, yet effective, flow cytometry analysis to monitor the progress of the selection, avoiding the tedious, time consuming and radioactive EMSA process. Through multiple rounds of selection with GST as a control, we have identified aptamers that selectively recognize AGR2 with nanomolar  $K_d$  values. CD measurements and melting-temperature assays demonstrated that the optimal aptamer C14B1 forms an intramolecular parallel G-quadruplex structure and its structure and binding affinity are strongly dependent on the nature of the monovalent ion. Furthermore, with our design of AGR2-aMB, AGR2 could be sensitively and selectively detected. The aptamer sequences and sensors reported here has great potential to serve as a useful tool for early diagnosis and prognosis of cancer and for fundamental research to elucidate the biochemical functions of AGR2.

## Materials and Methods

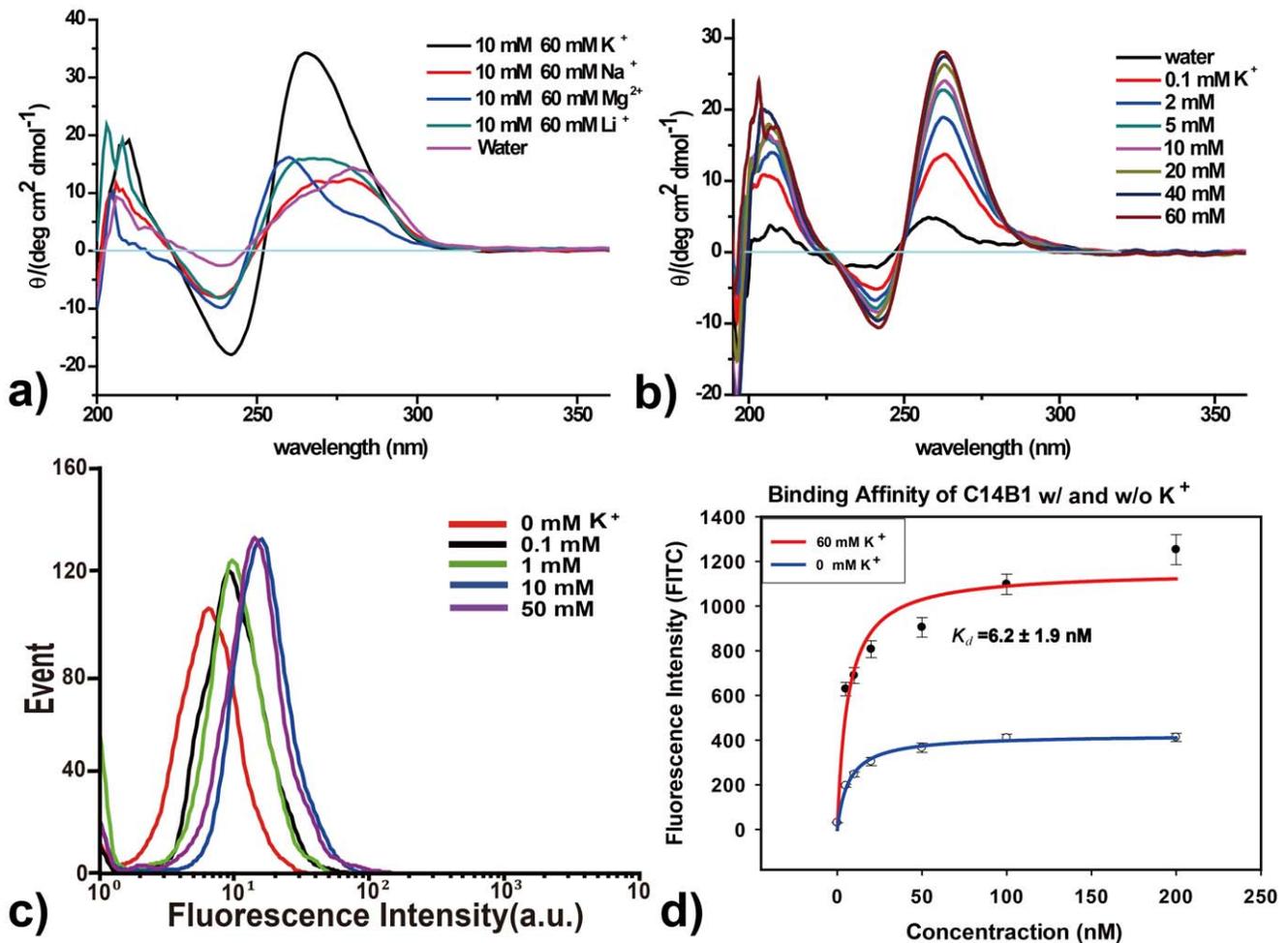
### Initial library design

The HPLC-purified library containing a central randomized sequence of 45 nucleotides flanked by two 20-nt and 22-nt primer hybridization sites. Initial library was: 5'-TCT CGG ACG CGT GTG GTC GG-N45-C TCG CTG CCT GGC CCT AGA GTG-3', forward primer: 5'-FAM-TCT CGG ACG CGT GTG

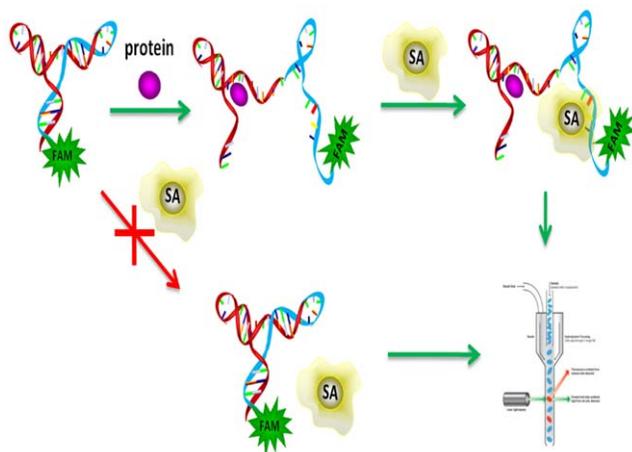


**Figure 3. Specificity of aptamer C14B1 against AGR2 and three control proteins, BSA, Trypsin and Thrombin.** The signal-to-background ratio (SBR) is defined as the fluorescence signal ratio of FAM-labeled aptamer to FAM-labeled random sequence when treating with protein modified beads. The data were the average of triplicate experiment results.

doi:10.1371/journal.pone.0046393.g003



**Figure 4. Characterization of the aptamer C14B1.** a) CD spectra of the aptamer C14B1 in the presence of different cations. This aptamer has a positive band near 260 nm that can be assigned to a parallel quadruplex structures. b) The structural stability of the aptamer relies heavily on  $[K^+]$ . c) Fluorescent signal migration of C14B1 binding to AGR2-beads at different concentration of  $K^+$  was investigated. d) The binding constants of C14B1 in the buffer w/ and w/o  $K^+$ . The data were the average of triplicate experiment results. doi:10.1371/journal.pone.0046393.g004



**Figure 5. Working principle of allosteric molecular beacon probe for sensitive and selective detection of AGR2 based on fluorescent flow cytometry analysis.** doi:10.1371/journal.pone.0046393.g005

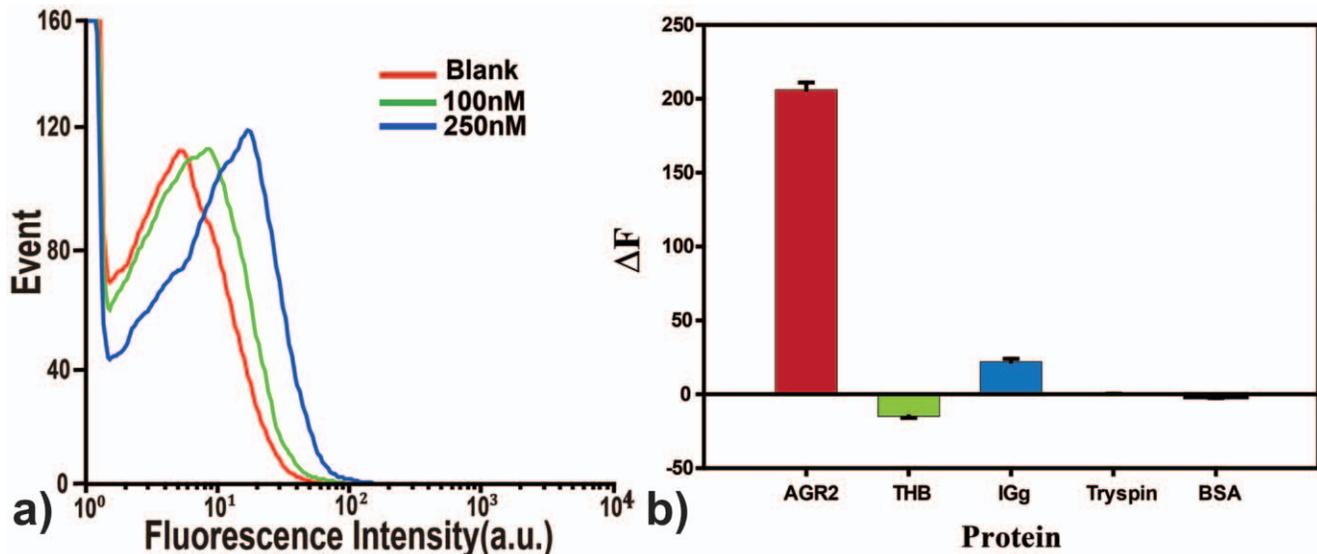
GTC GG-3', reverse primer: 5'-Biotin-CAC TCT AGG GCC AGG CAG CGA G-3'.

#### Preparation of AGR2-GST fused protein

The plasmid pGEX-GST-AGR2 was transformed into the engineering strain BL-21 and the GST-tagged AGR2 protein was expressed. After purification with glutathione sepharose beads (GE healthcare) by affinity chromatography, the GST-AGR2 fused protein was linked to sepharose beads for positive selection. After being washed several times with W1 buffer (25 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.07%  $\beta$ -mercaptoethanol, 1% Triton X-100), the final purified AGR2-GST-beads were stored in sterilized PBS buffer. Flow cytometry analysis with TAMRA-labeled anti-AGR2 monoclonal antibody (Santa Cruz) and SDS-PAGE indicated that the GST-AGR2 fused protein was successfully linked to the sepharose beads.

#### Selection of AGR2 binding aptamers

The procedures of selection were as follows. The ssDNA pool (200 pmol) dissolved in binding buffer (200  $\mu$ L, 1 $\times$ PBS buffer containing 137 mM NaCl, 2.7 mM KCl, 10 mM  $Na_2HPO_4$  and 2.0 mM  $KH_2PO_4$ , pH 7.4) was denatured by heating at 95 $^\circ$ C for



**Figure 6. Evaluation of the performance of AGR2-aMB.** a) Response of AGR2-aMB to different concentrations of AGR2 in Tris-HCl buffer. b) Response of AGR2-aMB to different targets including AGR2 (100 nM), BSA, IgG, Trypsin and Thrombin (500 nM each for control proteins). The data were the average of triplicate experiment results.

doi:10.1371/journal.pone.0046393.g006

5 min, then quickly cooled on ice for 10 min, and subsequently incubated for another 10 min at room temperature before binding. The ssDNA pool was then incubated with negative GST beads ( $1.0 \times 10^5$  beads) for counter selection to remove sequences non-specifically binding to GST beads. After filtering with a homemade filter column, the filtrate was incubated with positive AGR2-GST-beads ( $1.0 \times 10^5$  beads) at 37°C for 45 min. The unbound or nonspecifically bound oligoes were removed by filtration. The sequences bound to the target-coated beads were then amplified by PCR with FAM and biotin-labeled primers (5–15 cycles of 0.5 min at 94°C, 0.5 min at 53°C, and 0.5 min at 72°C, followed by 5 min at 72°C; the Easytaq plus polymerase and dNTPs were obtained from Transgen Beijing). After denaturation in NaOH (0.1 M), the selected sense ssDNA was separated from the biotinylated antisense ssDNA strand on streptavidin-coated sepharose beads (GE healthcare) and used for next round selection. For the first-round selection, the amount of initial ssDNA pool was 5 nmol, dissolved in 500  $\mu$ L binding buffer, and the counter-selection step was eliminated. To acquire aptamers with high affinity and specificity, the selection strength was enhanced gradually by increasing the number of washes (from three to ten times with 200  $\mu$ L 1 $\times$ PBS buffer each) and decreasing the amount of the ssDNA library per round (from 200 to 150 pmol).

#### Cloning and DNA sequencing

The resulting pool from the 14<sup>th</sup> round was PCR amplified, cloned and sequenced (Shanghai Sangon sequencing facility). The resulting 62 sequences were subjected to multiple sequence alignment analysis by using Clustal W 6.0 software to discover highly conserved motifs in groups of selected DNA sequences. The discovered consensus sequences with high repeats among selected pools were then chemically synthesized for further testing.

#### Flow cytometric analysis

To monitor the enrichment of aptamers after selection, the FAM-labeled ssDNA pool was incubated with  $5 \times 10^4$  AGR2-GST-beads or GST-beads in binding buffer (200  $\mu$ L) at 37°C for

45 min. Beads were washed three times with 200  $\mu$ L binding buffer by means of filtration, and suspended in binding buffer (250  $\mu$ L). The fluorescence intensity of the resulting beads was monitored with a FACSAria cytometer (Becton Dickinson Immuno cytometry systems) by counting 10000 events. The binding affinities of aptamers were determined by incubating AGR2-GST-beads ( $5 \times 10^4$ ) with various concentrations of FAM-labeled aptamers (pre-heat-treated) in binding buffer (200  $\mu$ L) at 37°C for 45 min in the dark. Beads were then washed three times with the binding buffer, then resuspended in binding buffer (250  $\mu$ L) and subjected to flow cytometry analysis. The FAM-labeled unselected ssDNA library was used as negative control for the nonspecific binding evaluation. All binding experiments were repeated two to four times. The mean fluorescence intensity of target protein labeled by aptamers was used to evaluate binding affinity by subtracting the mean fluorescence intensity of nonspecific binding produced by unselected ssDNA library. The dissociation constants ( $K_d$ ) of the fluorescent ligands were obtained by fitting the dependence of fluorescence intensity of specific binding on the concentration of the ligands to the equation (1):  $Y = B_{max}X/(K_d+X)$  using SigmaPlot software.

#### Circular dichroism spectroscopy

CD measurements were carried out on a Jasco J-810 spectropolarimeter equipped with a programmable temperature-control unit (Julabo HP-4). The concentration of DNA samples were 2 mM. Before the CD spectrum measurement, the DNA samples were annealed by heating to 95°C for 5 min, then rapidly cooled on ice for 10 min, and incubated for another 10 min at room temperature. The spectra from 400 to 200 nm were obtained by using 1 nm slit width and 0.1 nm scanning resolution. Each CD spectrum was an average of 8 scans with the buffer background subtracted.

#### UV thermal-denaturation experiment

UV absorbance and melting studies were carried out on an Agilent 8453 spectrophotometer equipped with a programmable temperature-control unit (Agilent 89090A). Melting temperatures ( $T_m$ ) were taken as the temperature of half-dissociation of the

quadruplex and were obtained from the maximum of the first derivative dA/dT plots at 295 nm. The heat-treated DNA solutions at several concentrations were introduced into a quartz cuvette and overlaid with a thin layer of silicone oil to prevent evaporation. The optical path length was 1 cm. Absorbance and temperature were recorded every 2°C.

### Evaluation of AGR2-aMB

AGR2-aMB (40 nM) was annealed by heating to 95°C for 5 minutes in 200 µL Tris-HCl buffer (25 mM Tris-HCl, 120 mM NaCl, 0.5 mM KCl, 2 mM MgCl<sub>2</sub> at pH 7.4), and then rapidly cooled on ice for 10 min, and subsequently waiting for another 10 min at room temperature before use. To the solution of AGR2-aMB, various concentrations of AGR2 were added and the resulting solutions were allowed to incubate at 37°C for 30 min. To the mixture, 1 µL of streptavidin beads (about 5×10<sup>4</sup> beads) were added and the mixture was allowed to set for 45 min in dark. After washing twice with Tris-HCl buffer, SA beads were filtered to eliminate unbound AGR2-aMB and resuspended in Tris-HCl buffer solution 250 µL before flow cytometry analysis. In selectivity tests, BSA (Tagene Biotechnology Xiamen), trypsin (Dingguo Beijing), thrombin (alfa) and IgG (Dingguo Beijing) were used.

### Supporting Information

**Figure S1 The sequences of 62 clones.** One subfamily is guanosine-rich sequences (22 clones), and the other is thymine-rich sequences (40 clones).

(TIF)

**Figure S2 The dissociation constant measurement of C14A, C14B, C14C and C14D against AGR2 GST and GSH.**

(TIF)

### References

- Bradley L, Wainstock D, Sive H (1996) Positive and negative signals modulate formation of the *Xenopus* cement gland. *Development* 122: 2739–2750.
- Fritzsche FR, Dahl E, Pahl S, Burkhardt M, Luo J, et al. (2006) Prognostic relevance of AGR2 expression in breast cancer. *Clinical Cancer Research* 12: 1728–1734.
- Riener MO, Pilarsky C, Gerhardt J, Grutzmann R, Fritzsche FR, et al. (2009) Prognostic significance of AGR2 in pancreatic ductal adenocarcinoma. *Histol Histopathol* 24: 1121–1128.
- Fritzsche FR, Dahl E, Dankof A, Burkhardt M, Pahl S, et al. (2007) Expression of AGR2 in non small cell lung cancer. *Histol Histopathol* 22: 703–708.
- Zhang JS, Gong A, Chevillie JC, Smith DI, Young CYF (2005) AGR2, an androgen-inducible secretory protein overexpressed in prostate cancer. *Genes Chromosomes & Cancer* 43: 249–259.
- Liu D, Rudland PS, Sibson DR, Platt-Higgins A, Barraclough R (2005) Human homologue of cement gland protein, a novel metastasis inducer associated with breast carcinomas. *Cancer Research* 65: 3796–3805.
- Kumar A, Godwin JW, Gates PB, Garza-Garcia AA, Brookes JP (2007) Molecular basis for the nerve dependence of limb regeneration in an adult vertebrate. *Science* 318: 772–777.
- Pohler E, Craig AL, Cotton J, Lawrie L, Dillon JF, et al. (2004) The Barrett's antigen anterior gradient-2 silences the p53 transcriptional response to DNA damage. *Molecular & Cellular Proteomics* 3: 534–547.
- Wang Z, Hao Y, Lowe AW (2008) The adenocarcinoma-associated antigen, AGR2, promotes tumor growth, cell migration, and cellular transformation. *Cancer Research* 68: 492–497.
- Maslon MM, Hrstka R, Vojtesek B, Hupp TR (2010) A Divergent Substrate-Binding Loop within the Pro-oncogenic Protein Anterior Gradient-2 Forms a Docking Site for Reptin. *Journal of Molecular Biology* 404: 418–438.
- Schreiber SL (2011) Organic synthesis toward small-molecule probes and drugs. *Proc Natl Acad Sci U S A* 108: 6699–6702.
- Montagnani M, Marangoni A, Roda A, Azzaroli F, Mazzella G, et al. (2009) Generation of a novel antibody probe to the apical sodium-dependent bile acid transporter that inhibits ileal bile acid absorption. *Mol Pharm* 6: 1012–1018.
- Dover JE, Hwang GM, Mullen EH, Prorok BC, Suh SJ (2009) Recent advances in peptide probe-based biosensors for detection of infectious agents. *J Microbiol Methods* 78: 10–19.

**Figure S3 Flow cytometry assay to monitor the binding of C14B1 and its five truncated sequences with a) AGR2 (target protein) and b) GST (control protein).**

(TIF)

**Figure S4 UV thermal-denaturation experiment of C14B1.** Denaturation profiles obtained at 295 nm for the aptamer at three different concentrations (2 µM, 4 µM, 8 µM). The T<sub>m</sub> (59°C) at 295 nm is independent of oligonucleotide concentration, indicating that the aptamer forms an intramolecular G-quartet.

(TIF)

**Figure S5 The secondary structure of AGR-aMB.** Stable hairpin structure is formed by intramolecular hybridization between the SA aptamer sequence and the complementary sequence of C14B<sub>1</sub>.

(TIF)

**Table S1 The sequences of C14A, C14B, C14C and C14D, and their K<sub>d</sub> to AGR2.**

(DOC)

**Table S2 The truncated sequences by removing one of poly-G portion each from C14B1.**

(DOCX)

### Author Contributions

Conceived and designed the experiments: YC ZZ LZ WJ WC. Performed the experiments: WJ WC LX SY WW HJ ZW LC LJ. Analyzed the data: YC ZZ LZ JW WC WJ CW LX SY WW HJ ZW LC LJ. Contributed reagents/materials/analysis tools: YC LZ ZZ. Wrote the paper: YC ZZ LZ WJ WC.

29. Kwon OS, Park SJ, Jang J (2010) A high-performance VEGF aptamer functionalized polypyrrole nanotube biosensor. *Biomaterials* 31: 4740–4747.
30. Massoud TF, Gambhir SS (2003) Molecular imaging in living subjects: seeing fundamental biological processes in a new light. *Genes Dev* 17: 545–580.
31. Backer MV, Levashova Z, Patel V, Jehning BT, Claffey K, et al. (2007) Molecular imaging of VEGF receptors in angiogenic vasculature with single-chain VEGF-based probes. *Nature Medicine* 13: 504–509.
32. Weissleder R, Tung CH, Mahmood U, Bogdanov A, Jr. (1999) In vivo imaging of tumors with protease-activated near-infrared fluorescent probes. *Nat Biotechnol* 17: 375–378.
33. Jiang T, Olson ES, Nguyen QT, Roy M, Jennings PA, et al. (2004) Tumor imaging by means of proteolytic activation of cell-penetrating peptides. *Proc Natl Acad Sci U S A* 101: 17867–17872.
34. Bell SD, Denu JM, Dixon JE, Ellington AD (1998) RNA molecules that bind to and inhibit the active site of a tyrosine phosphatase. *Journal Of Biological Chemistry* 273: 14309–14314.
35. Jenison RD, Jennings SD, Walker DW, Bargatzke RF, Parma D (1998) Oligonucleotide inhibitors of P-selectin-dependent neutrophil-platelet adhesion. *Antisense Nucleic Acid Drug Dev* 8: 265–279.
36. Gal SW, Amontov S, Urvil PT, Vishnuvardhan D, Nishikawa F, et al. (1998) Selection of a RNA aptamer that binds to human activated protein C and inhibits its protease function. *European Journal Of Biochemistry* 252: 553–562.
37. Blind M, Kolanus W, Famulok M (1999) Cytoplasmic RNA modulators of an inside-out signal-transduction cascade. *Proc Natl Acad Sci U S A* 96: 3606–3610.
38. Keefe AD, Pai S, Ellington A (2010) Aptamers as therapeutics. *Nature Reviews Drug Discovery* 9: 537–550.
39. Raddatz MS, Dolf A, Endl E, Knolle P, Famulok M, et al. (2008) Enrichment of cell-targeting and population-specific aptamers by fluorescence-activated cell sorting. *Angew Chem Int Ed Engl* 47: 5190–5193.
40. Tang Z, Shanguan D, Wang K, Shi H, Sefah K, et al. (2007) Selection of aptamers for molecular recognition and characterization of cancer cells. *Analytical Chemistry* 79: 4900–4907.
41. Chen HW, Medley CD, Sefah K, Shanguan D, Tang Z, et al. (2008) Molecular recognition of small-cell lung cancer cells using aptamers. *ChemMedChem* 3: 991–1001.
42. Shanguan D, Meng L, Cao ZC, Xiao Z, Fang X, et al. (2008) Identification of liver cancer-specific aptamers using whole live cells. *Analytical Chemistry* 80: 721–728.
43. Li Y, Jiang Z, Chen H, Ma WJ (2004) A modified quantitative EMSA and its application in the study of RNA–protein interactions. *J Biochem Biophys Methods* 60: 85–96.
44. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24: 1596–1599.
45. Huppert JL, Balasubramanian S (2005) Prevalence of quadruplexes in the human genome. *Nucleic Acids Research* 33: 2908–2916.
46. Rachwal PA, Fox KR (2007) Quadruplex melting. *Methods* 43: 291–301.
47. Paramasivan S, Rujan I, Bolton PH (2007) Circular dichroism of quadruplex DNAs: applications to structure, cation effects and ligand binding. *Methods* 43: 324–331.
48. Mergny JL, Phan AT, Lacroix L (1998) Following G-quartet formation by UV-spectroscopy. *Febs Letters* 435: 74–78.
49. Tasset DM, Kubik MF, Steiner W (1997) Oligonucleotide inhibitors of human thrombin that bind distinct epitopes. *Journal of Molecular Biology* 272: 688–698.
50. Shum KT, Tanner JA (2008) Differential Inhibitory Activities and Stabilisation of DNA Aptamers against the SARS Coronavirus Helicase. *Chembiochem* 9: 3037–3045.
51. Macaya RF, Schultze P, Smith FW, Roc JA, Feigon J (1993) Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution. *Proc Natl Acad Sci U S A* 90: 3745–3749.
52. Simonsson T (2001) G-quadruplex DNA structures—variations on a theme. *Biol Chem* 382: 621–628.
53. Burge S, Parkinson GN, Hazel P, Todd AK, Neidle S (2006) Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Research* 34: 5402–5415.
54. Davis JT (2004) G-quartets 40 years later: From 5'-GMP to molecular biology and supramolecular chemistry. *Angewandte Chemie-International Edition* 43: 668–698.
55. Song Y, Cui L, Wu J, Zhang W, Zhang WY, et al. (2011) Allosteric molecular beacons for sensitive detection of nucleic acids, proteins, and small molecules in complex biological samples. *Chemistry* 17: 9042–9046.
56. Bing T, Yang X, Mei H, Cao Z, Shanguan D (2010) Conservative secondary structure motif of streptavidin-binding aptamers generated by different laboratories. *Bioorg Med Chem* 18: 1798–1805.