

Electronic Supplementary Information (ESI)

for

“Selection of Aptamers Targeting the Sialic Acid Receptor of Hemagglutinin by Epitope-Specific SELEX”

Yeh-Hsing Lao,^{a,b} Hui-Yu Chiang,^a Deng-Kai Yang,^a Konan Peck^b and Lin-Chi Chen^{*a}

^aDepartment of Bio-Industrial Mechatronics Engineering, National Taiwan University, Taipei 10617, Taiwan

^bInstitute of Biomedical Sciences, Academia Sinica, Taipei 11529, Taiwan

*Corresponding author: E-mail: chenlinchi@ntu.edu.tw / Fax: +886-2-2362-7620

S1. ES-SELEX procedure

General Description about the ES-SELEX

In practice, ES-SELEX combined two *in vitro* selection routes to “evolve” HA-specific and SAR-specific aptamers sequentially (Scheme 1, in the text). The “protein-targeting” route was taken in the earlier rounds. The ssDNA library or pool interacted with the Ni-NTA column-bound HA, and then the ssDNAs bound to the column were eluted, amplified and re-prepared as a new ligand pool for the next selection round. It was the same as ordinary SELEX and led to a pool of HA specific ssDNA ligands, constituting an enriched library for subsequent “epitope-targeting” selection. The uniqueness of the SAR epitope-targeting selection was that the ssDNA ligands which specifically bind to the SAR motif on HA were displaced and collected by excess SA ligand (fetuin) induced elution. After rounds, SAR-specific as well as hemagglutination inhibitory aptamers were produced.

Single-stranded DNA library, aptamer and oligonucleotide synthesis

The ssDNA library, the anti-HA aptamers and other oligonucleotides used in this work were synthesized by Purigo Biotech (Taipei, Taiwan). All DNA sequences were confirmed and purified by 7M-urea denatured PAGE. The library was composed of a 10^{15} diversity of ssDNA molecules with a 62-mer sequence containing a centered 30-mer randomized region plus two flanked primer binding domains on both sides. The 62-mer ssDNA library sequence was expressed as follows.

5'-TCCC TACG GCGC TAAC-(N₃₀)-GCCA CCGT GCTA CAAC-3'

Protein-targeting selection

Protein-targeting selection was performed in the first five selection rounds. One microgram of recombinant hexa-His tagged HA1 protein (strain A/New Caledonia/20/99; expressed in the HEK 293 cell lines; Abcam, United Kingdom) was incubated with an ssDNA pool in the selection buffer (50 mM Tris-HCl, pH 7.5 with 4 mM KCl, 140 mM NaCl, 1 mM MgCl₂ and 2.5 mM CaCl₂) for 30 minutes at RT. Subsequently, the protein/ssDNA complex was captured by a His-Select® spin column (Sigma-Aldrich, St. Louis, MO, USA). The protein/ssDNA mixture (100 μL) was loaded into the equilibrated column. By centrifuge at 2000 rpm for 2 minutes, unbound ssDNA molecules

were washed away from the column. The following wash step was performed several times to remove unbound or low affinity ssDNA ligands: the column was loaded with 600 μL selection buffer and then spun at 2900 rpm for 2 minutes.

The elution steps were carried out twice: 50 μL of elution buffer (the selection buffer plus 500 mM imidazole) was added to elute the protein/aptamer complex bound to the column. Afterward, the complex was eluted for two minutes by spinning at 2900 rpm. To avoid the residue imidazole affecting the following-up PCR, the eluted product was buffer-exchanged by a gel-filtration column (CentriSep; Princeton Separations Inc., Adelphia, NJ, USA).

Epitope-targeting selection

Epitope-targeting selection was performed from the sixth to the ninth selection rounds. Basically, protein-targeting and epitope-targeting selection rounds shared the same binding and washing steps. The unique point of the epitope-targeting selection was that the ssDNA ligands, which specifically bound to the sialic acid receptor (SAR) epitope on HA1, were eluted and recovered by fetuin displacement. Fetuin (20.6 μM in selection buffer) was utilized to elute the SAR-specific ssDNA ligands by a spin column with centrifugation at 1600 rpm for 2 minutes. After elution, the product was collected and concentrated by a 10kD molecular cut-off (MWCO) column (Amicon YM-10; Millipore, Billerica, MA, USA). In the ninth round of selection, the SAR-binding ssDNA ligands were further fractionated by a fetuin gradient displacement (33, 165, 824 nM, 4.12 and 20.6 μM), and five fractions of the ssDNA ligand pools with increasing affinity to the SAR epitope were obtained accordingly.

Counter selection

A counter selection against fetuin and human serum albumin (HSA) was carried out after the seventh selection round. Eluted ssDNA ligands from the seventh round were incubated with fetuin and HSA (100 μg each). After 30-minute incubation, a 50 kD molecular MWCO column (Amicon Ultra-0.5; Millipore) was used to separate the fetuin/HSA-bound and unbound ssDNA molecules. The flow-through of 50kD-MWCO column (unbound ssDNA) was collected for the subsequent DNA amplification and single-stranded DNA isolation.

DNA amplification and ssDNA isolation

The amount of eluted ssDNA ligands (as well as the ssDNA bound to the selection target) was determined by qPCR (7900HT; Applied Biosystems, Grand Island, NY, USA). The qPCR reaction mixture (25 μL) was composed of 1 \times SYBR Green Master Mix® (Applied Biosystems), 200 nM primer pairs and the diluted ssDNA elution samples. The qPCR reaction was initiated at 95 $^{\circ}\text{C}$ for 10 minutes to activate the polymerase and then done with a thermal cycle profile of 95 $^{\circ}\text{C}$ for 15 seconds and 60 $^{\circ}\text{C}$ for 1 minute for 35 cycles. The amount of DNA was obtained by a calibration curve obtained from the qPCR process.

In addition to qPCR quantitation, the eluted ssDNA was amplified by conventional PCR. In the PCR process, a PCR mixture contained 200 nM primer pairs (reverse primer contained a biotin-tag), 50 mM Tris – HCl, pH 9.1, 16 mM (NH₄)₂SO₄, 10 mM betaine, 1 % DMSO, 200 μM each dNTPs (Protech Inc, Taipei, Taiwan), 3.5 mM MgCl₂, 150 μg/ml BSA, 2 U of KlenTaq DNA polymerase (Protech Inc) and 5 μL of eluted ssDNA product. The PCR was done by a three-temperature cycling at 95 °C for 30 seconds, 60 °C for 30 seconds and then 72 °C for 20 seconds with a desired cycle number, which was determined by qPCR.

Streptavidin-coated magnetic beads (M-PVA SAV1; PerkinElmer Chemagen Technologie GmbH, Baesweiler, Germany) were washed by 10× SSC buffer for 10 minutes for isolation of the amplified ssDNA products from the PCR amplicons. The PCR amplicons were incubated with the beads in 10× SSC buffer for one hour with gentle shaking. Subsequently, the beads bound with the dsDNA amplicons were washed by 10× SSC buffer twice. The double-stranded amplicons were dissociated into single-stranded DNAs by 100 mM NaOH. The forward strand was isolated with the aid of the bead, and purified with a CentriSep column.

DNA sequencing

The ssDNA ligands in the late selection rounds (which were considered aptamer candidates) obtained from the fifth fraction in the ninth round of selection were ligated to pGEM-T vectors by T4 DNA ligase (Promega, Madison, WI, USA) After DNA ligation, the vectors were transformed to the competent cells (*E. coli* strain DH5α, Real Biotech Corp., New Taipei City, Taiwan). Ninety-six colonies were collected and confirmed by PCR using T7 and SP6 primers. Primers and dNTPs in the PCR products were subsequently digested by Exonuclease I and shrimp alkaline phosphatase at 37 °C for 80 minutes. These clones were then sequenced by Sequencing Core Facility of Institute of Biomedical Sciences, Academia Sinica.

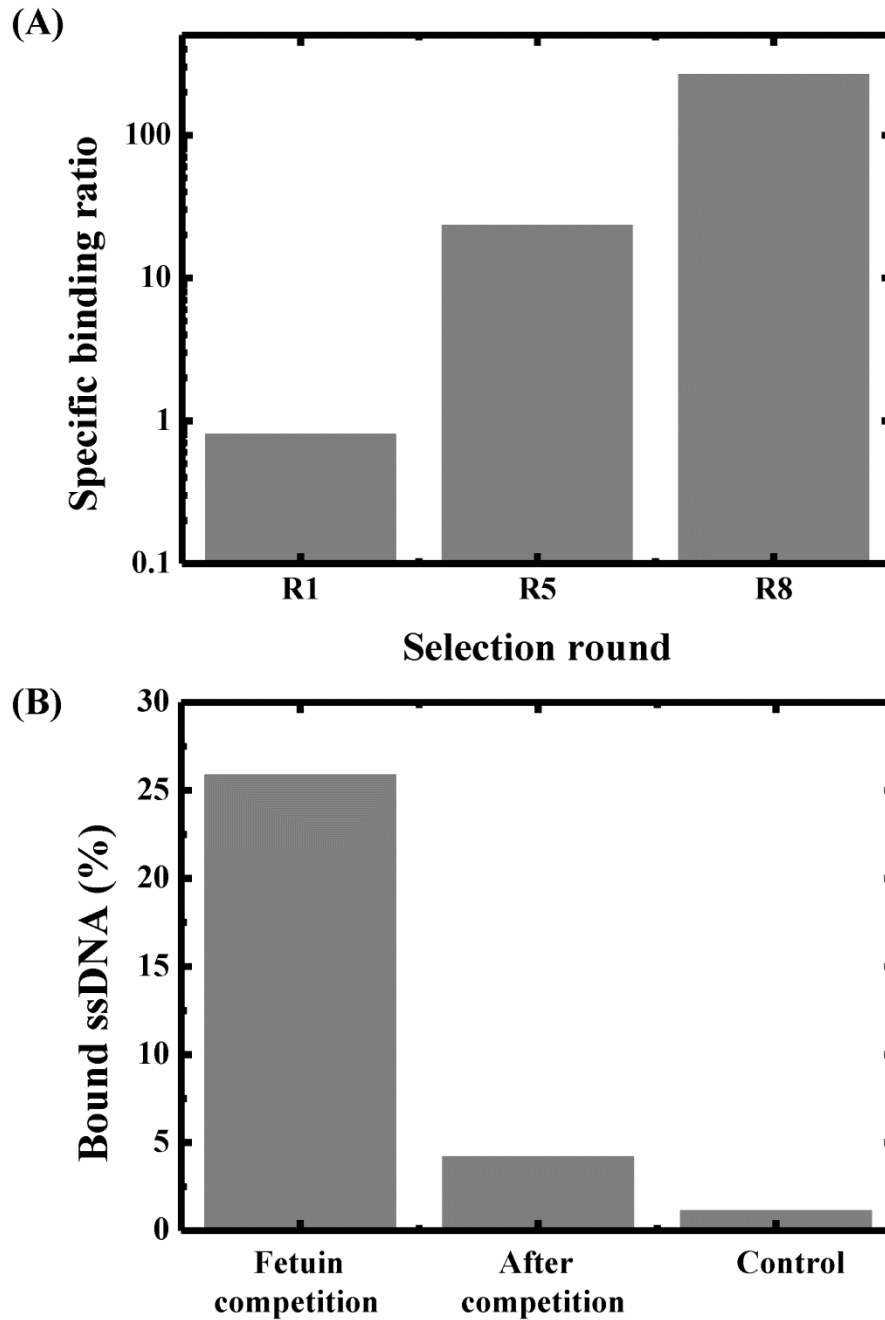


Figure S1. (A) Specific binding ratios of the first, fifth and eighth selection rounds in ES-SELEX. The amount of ssDNA binding to HA was compared to the no-target control to determine the specific binding ratio (= amount of ssDNA bound to HA / amount of background ssDNA binding). (B) The comparison of the bound ssDNA percentages between fetuin competition and imidazole elution (after fetuin competition). (Note: the amount of ssDNA was determined by qPCR.)

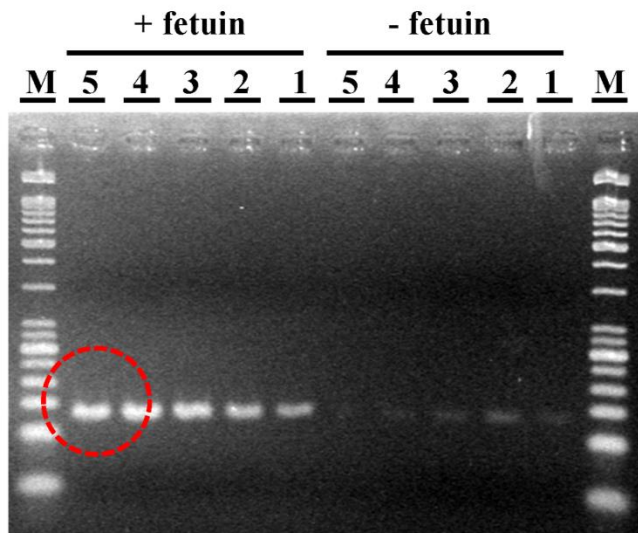


Figure S2. The gel electrophoresis analysis of the fractionated ssDNA ligand pools obtained in the ninth selection round, in which fetuin gradient elution was performed. The PCR product of each fraction was visualized by 3 % agarose gel with ethidium bromide stain (M = 25 bp DNAladder).

S2. Determination of the mean dissociation constant (K_D) for the winning ssDNA ligand pool

The HA1 protein samples (15 pmol) were incubated with serially diluted ssDNA ligand pools from the fifth fetuin-eluted fraction of round 9 for 30 minutes. Then the ssDNA/HA1 protein mixtures were loaded into 30 kD-MWCO columns (Amicon Ultra-0.5; Millipore). Unbound ssDNAs were removed by five spinning-reconstituting cycles. Bound ssDNAs in the retentate were collected and quantitated by qPCR. Fluorescence signals at the 20th qPCR cycle of all samples were plotted against the ssDNA concentration. Finally, the mean K_D of the ssDNA pool was determined by one site total binding model fitting ($Y = (B_{\max} \cdot X) / (K_D + X)$).

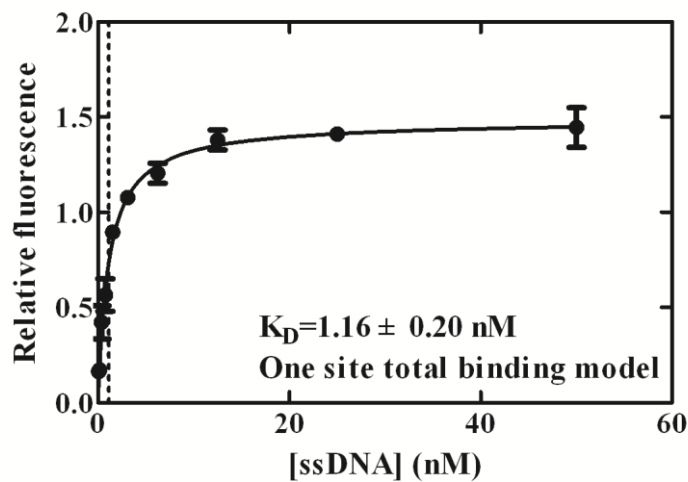
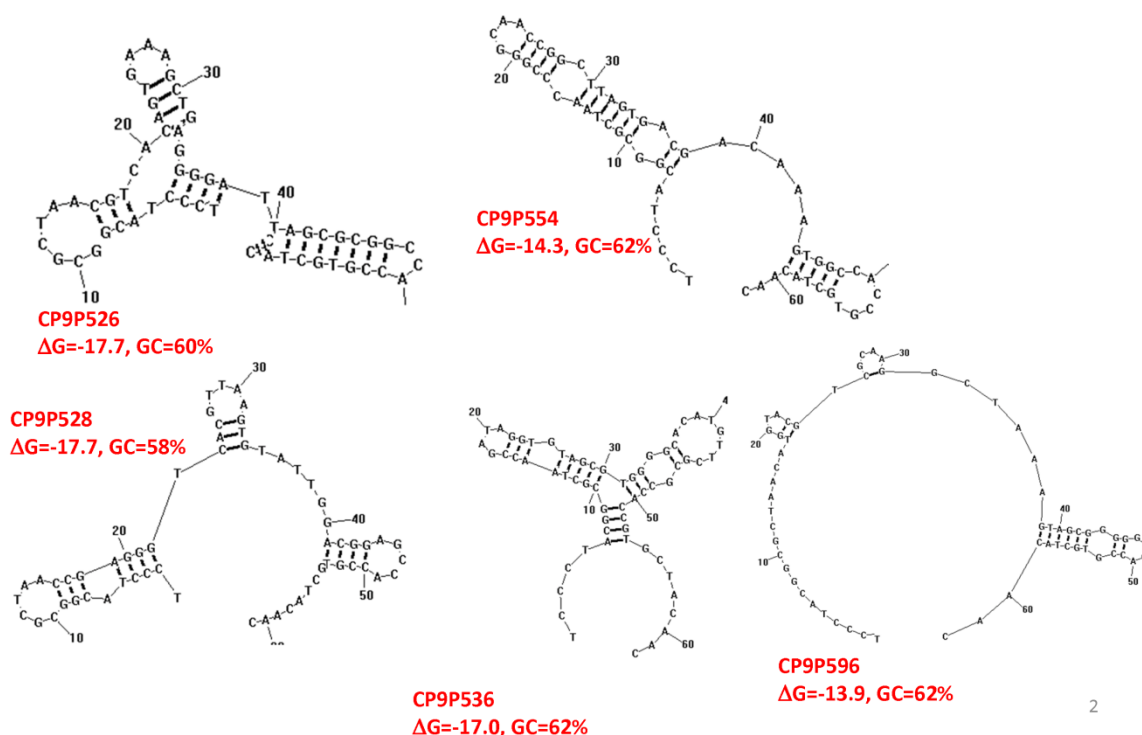


Figure S3. The HA total binding assays for a serial dilution of the winning ssDNA ligand pool from the 5th fetuin-eluted fraction in the ninth selection round. Data points were represented in terms of mean \pm standard deviation ($n = 3$).

S3. Representative secondary structures of the aptamer candidate sequences

The sequencing results of arbitrarily chosen 96 clones of the winning ssDNA ligand pool (Fraction 5 of the 9th round) were analyzed by a multiple sequence alignment software (Vector NTI; version 9.1; Invitrogen). It was found that the sequences could be classified into five groups. The secondary structures and Gibbs Free Energy (ΔG) of the aptamer candidate sequences were then predicted by RNA Structure software (version 5.02)^{S1}. The following representative aptamers, one from each group, were chosen based on the ΔG prediction results.



2

Figure S4. The secondary structures of the five representative aptamers determined by a multiple sequence alignment and a minimum free energy analysis.

S4. Hemagglutination-inhibition (HAI) assay

The hemagglutinin-inhibition (HAI) assay was slightly modified from the WHO standard protocol^{S2}. Fresh human erythrocytes (type O) were briefly washed by the selection buffer or 0.1 M PBS (100 mM phosphate, 150 mM NaCl, pH 7.2) and then spun down at 2000 rpm for 10 minutes. By three washes, erythrocytes pellets were diluted to a final concentration of 0.75%.

Full-length HA protein samples (with 4 HAU activity) were incubated with the aptamers of this work or commercial antibody (Anhi-H1 antibody 1.B.408; Abcam) with a desired concentration in the selection buffer or 0.1 M PBS for 30 minutes. Afterward, the HA/aptamer (or antibody) mixtures (in 50 μ L) were transferred to a U-type microplate and incubated with 50 μ L of 0.75% erythrocytes for another 1 hour. Hemagglutinin inhibition by aptamer was visually observed (refer to Figure 2 and Figure 3 in the text). By contrast, the commercial antibody did not display significant HAI function, as shown below.

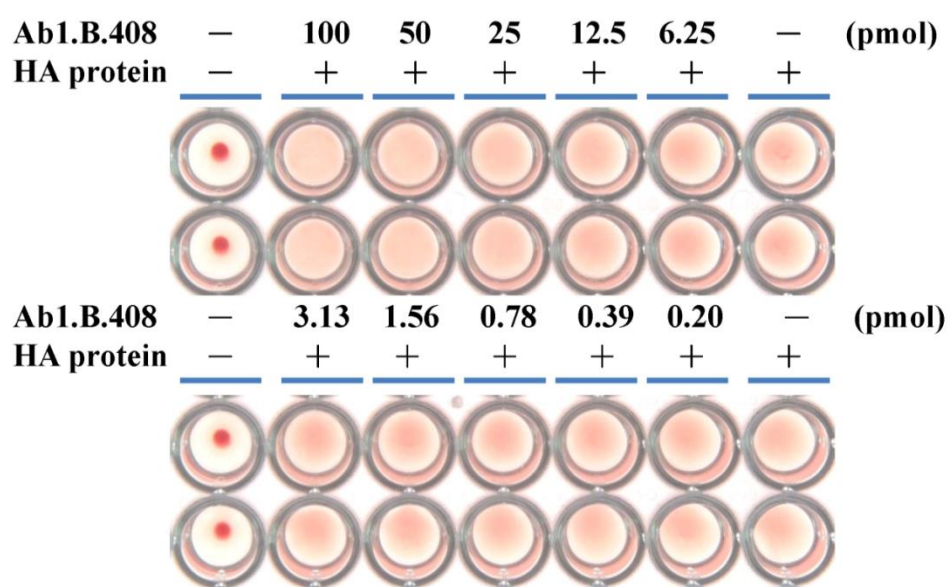


Figure S5. The HAI result of antibody 1.B.408.

Table S1. Comparison of the HAI efficacies of anti-HA aptamers reported in the literature and those of aptamer CP9P536 and commercial antibody determined in this work.

	Target HA	Effective Dose	Reference
Aptamer CP9P536	H1	6.25 pmol	This work
Aptamer A22	H3	12.5 pmol	S3
Aptamer HAS15-5	H5	34.1 pmol (1 μ g)	S4
Antibody 1.B.408	H1	>100 pmol	This work

S5. Microarray analysis for identifying the responsible binding domain of aptamer CP9P536

The binding domain of anti-hemagglutinin aptamer CP9P536 was investigated by a high-density microarray analysis. A “base-by-base shift” design^{S5} was utilized to probe the sequence information of the full-length aptamer. By this approach, the full-length aptamer sequence was divided into a variety of truncated forms (subsequences) with the length of 25 to 40 mer. Every truncated aptamer probe was combined with a 20 oligo-dT spacer at the 3'end. All aptamer probes were *in situ* synthesized in duplicates by an Agilent 2×105 K custom array.

The protocol of the aptamer microarray assay followed that of our previous work^{S6}. Three major steps of pre-incubation process were performed. First, the surface of the aptamer microarray was blocked with bovine serum albumin (BSA) solution (0.1 mg/ml BSA, 0.1 % SDS in $5\times$ SSC buffer) at 42°C for one hour. Second, the microarray was washed with diluted Corning Pronto! solution for three times at room temperature. Third, the microarray was immersed in the selection buffer for 30 minutes with gentle shaking.

After pre-incubation process, the microarray was incubated either with Cy3-labeled hemagglutinin (total 2 μ g hemagglutinin with ~3:1 dye-protein ratio) or with Cy3-labeled allantoic fluid (inoculated with influenza A strain A/New Caledonia/0/99; Advanced ImmunoChemical Inc., Long beach, CA, USA) solutionS (containing target proteins and 1 % Roche Blocking reagent in the selection buffer) at room temperature for overnight. Subsequently, the microarray was washed with the selection buffer plus 0.05 % Tween-20 for three times and was then scanned by a GenePix 4000B microarray scanner with a 650V PMT gain setting (Molecular Devices, Sunnyvale, CA, USA). The fluorescence results were analyzed by GenePix PRO 6.0 (Axon Instruments, Union City, CA, USA). The truncated sequences on the microarray showing higher fluorescence intensity than the full-length aptamer were picked out and aligned by Vector NTI.

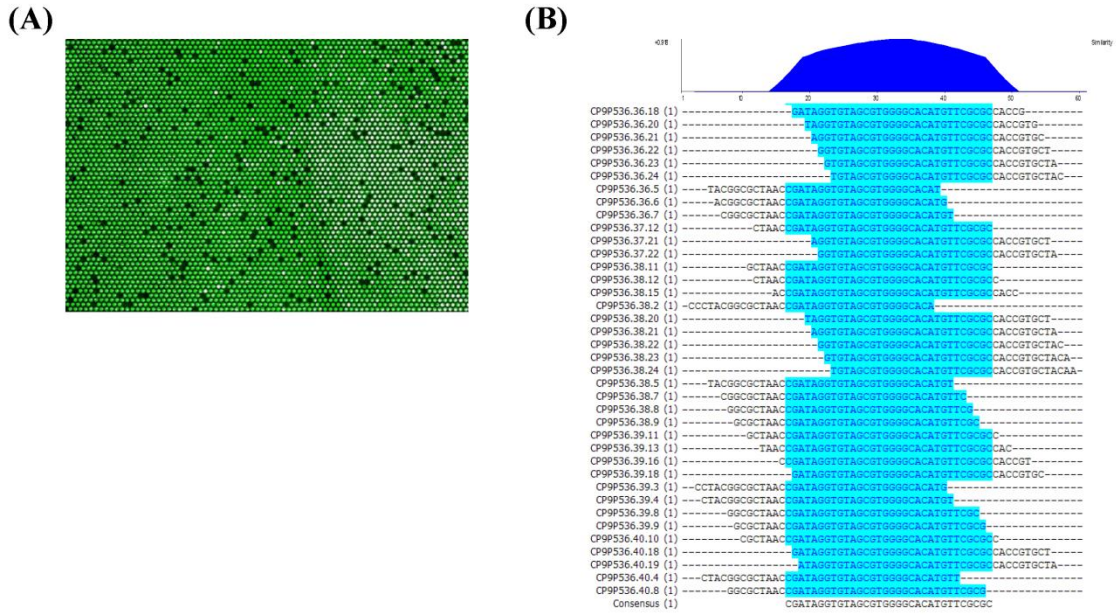


Figure S6. (A) An image glance of the Cy3-HA hybridized aptamer microarray in this study. (B) A partial result of the sequence alignment of CP9P536 truncates.

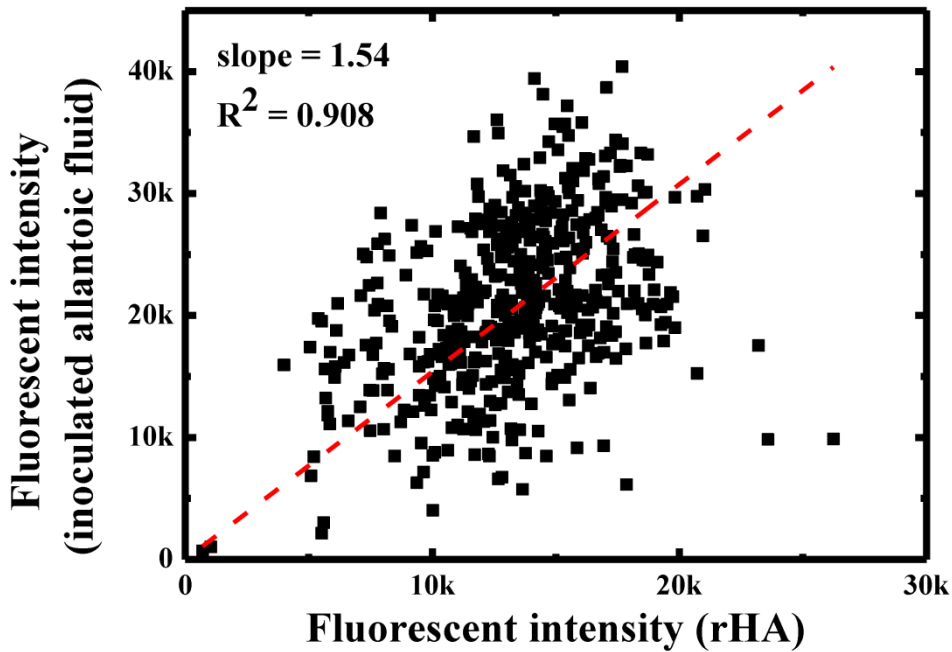


Figure S7. Correlation analysis of two aptamer microarray experimental results that assayed the subsequences of CP9P536 with the recombinant HA protein and with the inoculated allantoic fluid.

S6. *In silico* aptamer-HA docking simulation

An *in silico* aptamer-protein docking pipeline was developed to see if the aptamer CP9P536 could exactly dock on the SAR epitope of an influenza hemagglutinin structure released in PDB. To do this, the three dimensional structure of the aptamer was generated *in silico*. In the beginning, the secondary structure of aptamer CP9P536 was predicted by the software RNAstructure, and the secondary structure defined using dot-bracket notation was then sent to the software RNAcomposer for tertiary structure prediction. Finally, PatchDock was used to carry out the *in silico* molecular docking between H1 protein (PDB ID: 1RUZ) and CP9P536. For prediction of the minimal aptamer binding region, a perl script was utilized to calculate the distances between the atomic pairs and to identify the nucleotide bases containing at least one atom within 5Å with any atom in the amino acid moiety of H1's SAR epitope. As a result, the molecular simulation result supports that aptamer CP9P536 could perfectly dock on the SAR epitope of H1 and predicts a minimal aptamer binding region of CP9P536 for the SAR domain: "GCGTG GGGCA CATGT TC".

RNAstructure: <http://rna.urmc.rochester.edu/RNAstructure.html>

RNAcomposer: <http://rnacomposer.cs.put.poznan.pl/>

PatchDock: <http://bioinfo3d.cs.tau.ac.il/PatchDock/>

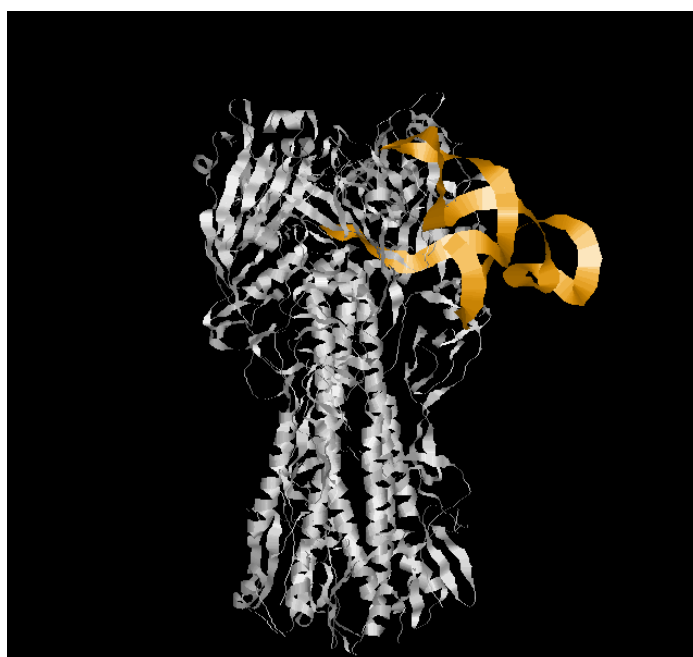


Figure S8. Graphical presentation of the molecular docking between hemagglutinin H1 protein (3D structure from PDB) and aptamer CP9P536 (3D structure from RNAstructure and RNAcomposer) done by PatchDock.

S7. Application of ES-SELEX to direct selection of anti-HA aptamers against H1N1 virus

A preliminary viral SELEX experiment was carried out to further demonstrate the niches of the proposed ES-SELEX. A real H1N1 viral particle contains several different surface proteins, in which HA and NA are the main two groups. Hence, an ordinary SELEX protocol would be difficult in generating aptamers only targeting to a specific viral surface protein (e.g., HA) when the whole H1N1 viral particles were used as the SELEX targets. In our preliminary viral SELEX experiment, real H1N1 virus (A/Swine/Changhua/415-7/2009) was used the SELEX target, and the experiment was carried out in a biosafety level 2⁺ (P2plus) laboratory in the School of Veterinary Medicine at National Taiwan University (done by Ms. Ching-Chun Huang, supervised by Prof. Chung-Hsi Chou and Prof. Hsiang-Jung Tsai). The ES-SELEX approach was used to direct select anti-HA aptamers in the viral SELEX using SA for competitive elution of the ssDNA ligands bound to the HA of the H1N1 virus. Here an aptamer BR33 from the viral ES-SELEX is served as the example. The HA-specificity of BR33 was proven by an ELAA experiment assayed with a pure hemagglutinin H1 recombinant protein (prepared by Prof. Je-Ruei Liu's Lab in the Institute of Biotechnology, National Taiwan University). Two hundred μ l of H1 protein (5 μ M) and BSA (5 μ M) in carbonate buffer (pH 8.0) was coated in the wells of a 96-well microtiter plate, respectively. Biotinylated aptamers were then added into the wells coated by H1 protein after refolding of aptamer BR33. After washes, the secondary antibody reagent horseradish peroxidase-streptavidin (HRP-SA) was used to recognize biotinylated aptamers bound to H1. Finally, 3,3',5,5'-tetramethylbenzidine (TMB) substrate was added into the wells for reporting the binding event (see Fig. S9). The specificity of BR33 to H1 is observed, while the aptamers from non-ES, ordinary viral SELEX did not show HA specificity (data not shown here).

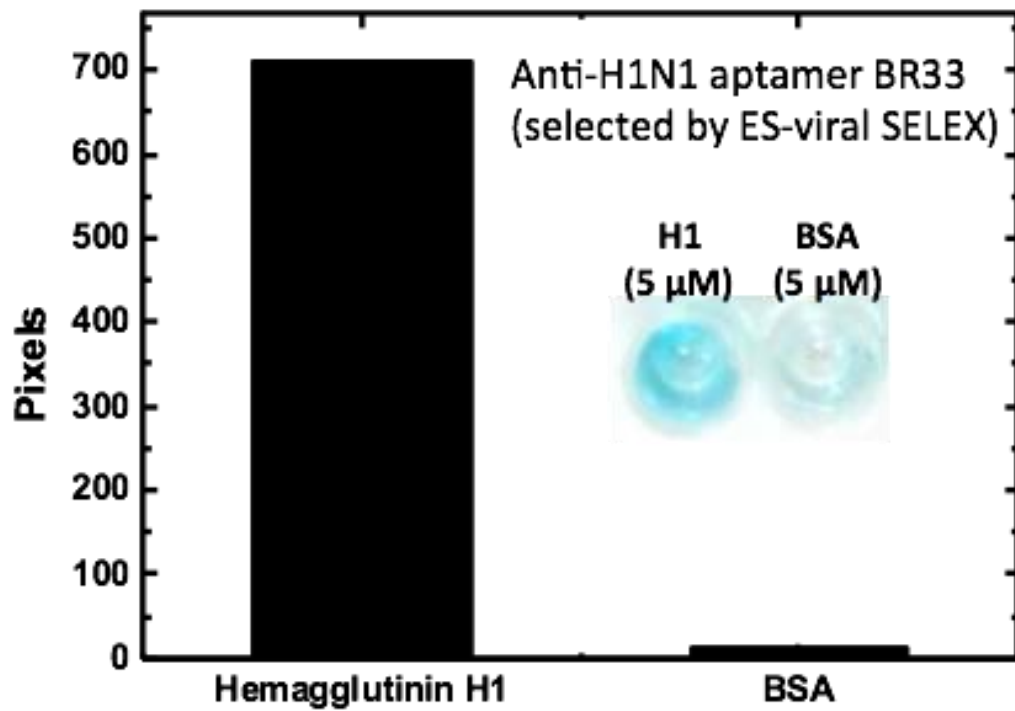


Figure S9. Enzyme-linked aptamer assay was used to examine the binding specificity between the aptamer BR33 from ES viral SELEX and hemagglutinin H1 recombinant protein.

References for electronic supplementary information

- S1. J. S. Reuter and D. H. Mathews, *BMC bioinformatics*, 2010, 11, 129.
- S2. K. G. Nicholson, J. M. Wood and M. Zambon, *Lancet*, 2003, 362, 1733-1745.
- S3. S. H. Jeon, B. Kayhan, T. Ben-Yedidia and R. Arnon, *J Biol Chem*, 2004, 279, 48410-48419.
- S4. S. Choi, C. Lee, K. Lee, S. Y. Choe, I. Mo, R. Seong, S. Hong and S. Jeon, *Mol Cells*, 2011, 32, 527-533.
- S5. Y. C. Chang, W. C. Kao, W. Y. Wang, W. Y. Wang, R. B. Yang and K. Peck, *Faseb J*, 2009, 23, 3078-3088.
- S6. Y. H. Lao, K. Peck and L. C. Chen, *Anal Chem*, 2009, 81, 1747-1754.