Electronic Supplementary Material (ESI) for ChemComm.
This journal is © The Royal Society of Chemistry 2017

Supporting materials

Universal Split Spinach Aptamer (USSA) for Nucleic Acid Analysis and DNA Computation

Nanami Kikuchi[1] and Dmitry M. Kolpashchikov[1,2]

[1]Chemistry Department, University of Central Florida, Orlando, 32816, Florida, USA

[2]National Center for Forensic Science and Burnett School of Biomedical Sciences, University of Central Florida, Orlando, 32816, Florida, USA

Table of content

1. Materials and instruments. All oligonucleotides were purchased from Integrated DNA Technologies, Inc. (Coralville, IA). DNAse/RNAse free water was purchased from Fisher Scientific and used for all assays including buffers, and for dissolution of oligonucleotides. Concentrations of oligonucleotide were determined based on UV light absorption at 260 nm. DFHBI was purchased from Lucerna, Inc. (New York, NY), KCl and $MgCl_2$ were purchased from Fisher Scientific. Trizma Hydrochloride (Tris-HCl), pH 7.40 was purchased from Sigma Aldrich. Two 2× Spinach buffer were prepared: 2 × Spinach Buffer 1 (20 mM Tris-Cl, pH 7.4, 200 mM KCl, 10 mM $MgCl_2$); 2 × Spinach Buffer 2 (20 mM Tris-HCl, pH 7.4, 200 mM KCl, 100 mM $MgCl_2$). All fluorescent spectra were taken using Fluorescence Spectrometer LS55 (PerkinElmer). Otherwise noted, excitation wavelength was set to 450 nm and emission was taken at 500 nm.

2. Table S1: oligonucleotides used in this study

Name	Sequence	Purification
SSA_m	5'-gtg ttg tgt /TEG/ UGG UGA AGG ACG GGU CCA GU	SD
SSA_f	5'-ACU GUU GAG UAG AGU GUG AGC UCC Gca gtg gcc cat acc catg c	SD
Adp_m	aca aga tcg /TEG/ cca ctg aca caa cac	SD
Adp_f	atg ggt atg ggg atct acg ggt t	SD
A_m_DNA	aac ccg tag atc cga **t**ct tgt g	SD
A_{mm}_DNA	aac ccg tag atc cga **a**ct tgt g	SD
A_m_RNA	AAC CCG UAG AUC CGA **U**CU UGU G	SD
A_{mm}_RNA	AAC CCG UAG AUC CGA **A**CU UGU G	SD
Adp_m_NOR1	at ggg tat ggg c gtc ggc gct tgt	SD
Adp_f_NOR1	agcgccgactgttggc atacgtaagtgatcc gcca aca cca ctg aca caa cac	SD
I1	gcg gat cac tta cgt atg cca ac	SD
I2	ccc aca agc gcc gac	SD
***Mtb*. F11**	ggg ttgaccc aca agc gcc gac tgt **t** ggc g ctg	SD
***Mtb*. KZN**	ggg ttgaccc aca agc gcc gac tgt **c** ggc g ctg	SD
f3	gcg cca aca aca caa cac	SD
f4	gcg cca aca a ctg aca caa cac	SD
f5	gcg cca aca ca ctg aca caa cac	SD
f6	gcg cca aca cca ctg aca caa cac	SD
f6peg	gcg cca aca/TEG/ cca ctg aca caa cac	SD
m3	tat ggg cca ctg gtc ggc gct tgt	SD
m4	t ggg tat ggg cc gtc ggc gct tgt	SD
m5	at ggg tat ggg c gtc ggc gct tgt	SD
m6	tat ggg tat ggg gtc ggc gct tgt	SD
m6peg	tat ggg tat ggg /TEG/ gtc ggc gct tgt	SD
Amel Y	c ccagtttaa gctctgatgg tt ggcctcaa gcctgtgttg ctccagcacc ctcctgcctg accattcgga t tgactcttt cctcct aaat **a** tggc tgtaa gtttattcat tcatgaacca ctgctcagga ag gttccatg aaagggcaaa aa	SD
Amel Y_mut (A→ G)	c ccagtttaa gctctgatgg tt ggcctcaa gcctgtgttg ctccagcacc ctcctgcctg accattcgga t tgactcttt cctcct aaat **g** tggc tgtaa gtttattcat tcatgaacca	SD

	ctgctcagga ag gttccatg aaagggcaaa aa	
Adp_f_amelY	at ggg tat ggg agg agg aaa gag tca atc cga atg gtc agg cag gag	SD
Adp_m_amelY	gcca t attt/TEG/cca ctg aca caa cac	SD

TEG-triethylene glycol linkers; SD, standard desalting; DNA sequences are in low case; SNS sites are bold underlined; analyte binding arms are underlined.

3. Detailed Experimental Procedure

General Fluorescent assay for miDNA and miRNA. For DNA analyte, the reaction mixtures contained **DFHBI** (1 μM) **SSA_m** (3.6 μM) and **SSA_f** (2 μM) strands, **Adp_m** (0.8 μM) and **Adp_f** (4.0 μM) and the indicated concentration of analyte in 60 μL of 1 × Spinach 1 or Spinach 2 buffers. Control samples contained only **DFHBI** (1 μM) or **DFHBI** (1 μM) and **SSA_m** (3.6 μM) and **SSA_f** (2 μM) and **Adp_m** (0.8 μM) and **Adp_f** (4.0 μM) strands. For RNA analyte, **DFHBI** (0.5 μM) **SSA_m** (3.6 μM) and **SSA_f** (2.0 μM) and the indicated concentration of analyte in 60 μL of Spinach 2 buffers. Control samples contained only **DFHBI** (0.5 μM) or **DFHBI** (0.5 μM) **SSA_m** (3.6 μM) and **SSA_f** (2.0 μM) and **Adp_m** (0.8 μM) and **Adp_f** (4.0 μM) strands. All samples were incubated at room (22.5^{o}C) temperature. Fluorescent spectra were recorded at 430-600 nm, with excitation at 450 nm after indicated incubation times. Data of three independent experiments were processed using Microsoft Excel.

General Fluorescent assay for Mtb.F11 and Mtb.KZN. **DFHBI** (2 μM) **SSA_m** (3.6 μM) and **SSA_f** (2 μM) strands, **Adp_m** (0.8 μM) and **Adp_f** (4.0 μM) and the indicated concentration of analyte in 60 μL of 1 × Spinach 1 or Spinach 2 buffers. Control samples contained only **DFHBI** (2 μM) or **DFHBI** (2 μM) and **SSA_m** (3.6 μM) and **SSA_f** (2 μM) strands. All samples were incubated at room (22.5^{o}C) temperature. Fluorescent spectra were recorded at 430-600 nm, with excitation at 450 nm after indicated incubation times. Data of three independent experiments were processed using Microsoft Excel.

Time dependence of USSA fluorescent response. **DFHBI**, **SSA_m** and **SSA_f** strands, **Adp_m** and **Adp_f** and 400 nM for DNA analyte, 100 nM for RNA analyte in 60 μL of 1 × Spinach 2 buffers. Fluorescence measurements were taken after 0, 1, 2, 5, 10, 25 and 40 min. Control samples were ran parallel with the samples. Data of three independent experiments were processed using Microsoft Excel.

***Determining Limit of Detection*.** **DFHBI**, **SSA_m** and **SSA_f** strands, **Adp_m** and **Adp_f** matched DNA (1.4, 5.5, 14, 28, 138, 275 nM or 5.5 μM) or RNA (0.5, 2, 5, 10, 50, 100, 250 nM, 1 μM) analytes in 60 μL of 1 × Spinach 2 buffers. matched or mismatched analyte was added for a total volume of 60 μL reaction mixture. Fluorescent spectra were measured after 30 min. Control samples were ran parallel with the samples. Data of three independent experiments were processed using Microsoft Excel.

Selectivity Assay. For miDNA and miRNA, **DFHBI**, **SSA_m** and **SSA_f** strands, **Adp_m** and **Adp_f**, and 100 nM matched or mismatched analytes in 60 μL of 1 × Spinach 2 buffers. For

Mtb.F11 and *Mtb*.KZN, **DFHBI**, **SSA_m** and **SSA_f** strands, **Adp_m** and **Adp_f**, and the indicated concentration of analyte in 60 µL of 1 × Spinach 2 buffers.. Control samples were ran parallel with the samples. Fluorescent spectra were measured after 30 min. Data of three independent experiments were processed using Microsoft Excel.

NOT and NOR logic gates. **DFHBI** (0.5 µM), **SSA_m** and **SSA_f** strands (2.0 µM), **Adp_f_NOR1** (1.0 µM) and **Adp_m_NOR1** (0.3 µM), were added to 30 µL of 2 × Spinach 2 buffer. Total volume was adjusted to 50 µL by H_2O and 5 µL each of Input **I1** and **I2** (1.0 µM) were added for a total volume of 60 µL reaction mixture. Control samples were ran parallel with the samples. Fluorescent spectra were measured after 30 min. Data of three independent experiments were processed using Microsoft Excel.

4. Figure S1. Design of USSA probe

Figure S1. Design of USSA with miRNA analyte. Dotted black line is the triethylene glycol linkers. Ribunucleotides are shown in upper case; deoxyribonucleotide are in low case. SNP location on miRNA underlined.

We initially designed the adapter strands as indicated in Figure S2A but showed no selectivity towards fully matched analyte. Therefore strand **Adp_m** in Figure S1 and the main paper crossed over and was made complementary to **SSA_f** by 4 nucleotides which reduced the fluorescence signal of single mismatched analyte (Figure S2B). We also inserted a triethyleneglycol (TEG) linker connecting the SSA binding arm and the analyte binding arm which allowed great increase in the selectivity of the USSA (Figure S3C). Magnesium concentration in Spinach buffer was increased from 5 mM to 50 mM which allowed greater stability for hybridization which inducing higher signal for fully matched analyte (Figure S4). For any TEG containing strand combination, lower concentration of adapter-f was used which allowed lowering single mismatch analyte signal (Figure S4). We concluded that USSA in which strand_f is complementary to SSA_f by 6 nucleotides produced the greatest analyte dependent fluorescent increase (Figure S4).

5. *Figure S2: Suboptimal designs of USSA analyzing Mtb. F11/KZN analyte*

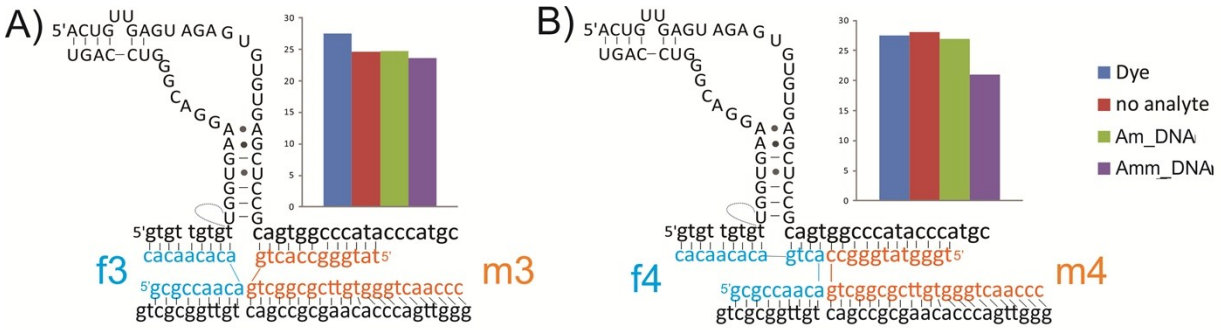

Figure S2. Fluorescence complexes of USSA with the DNA analyte and responses of the two different combinations of stands m and f. Data for each combination is shown on the right side of each panel. Samples contained 3.6 µM **SSA_f**; 2.0 µM **SSA_m**; 4.0 µM **Adp_f**; 4.0µM **Adp_m**; 6.0 µM analyte; 20 µM DFHBI; in Spinach Buffer 2). Emission (λ_{ex} = 450 nm) was registered 500 nM after 30 min at 22.5˚C. A, B) The bar graph shows the fluorescence intensity of samples containing DFHBI dye, no analyte (SSA+Adp), Am_DNA (SSA+Adp+matched analyte), and Amm_DNA (SSA+Adp+mismatched analyte).

Figure S3: Suboptimal designs of USSA analyzing Mtb. F11/KZN analyte

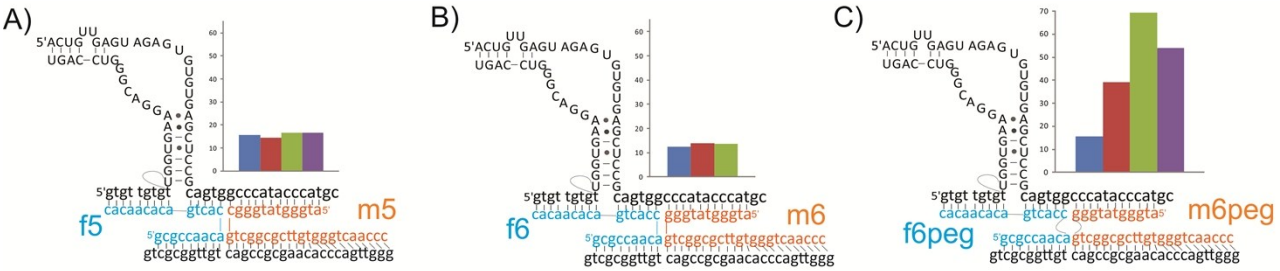

Figure S3. Fluorescence complexes of USSA with the DNA analyte and responses of the three different combinations of stands m and f. Data for each combination is shown on the right side of each panel. Samples contained 3.6 µM **SSA_f**; 2.0 µM **SSA_m**; 4.0 µM **Adp_f**; 4.0µM **Adp_m**; 6.0 µM analyte; 10 µM **DFHBI**; in 2 × Spinach Buffer 1. Emission spectrum (λ_{ex} = 450 nm) were collected after 30 min at 22.5˚C. A, B, C) The bar graph shows the fluorescence intensity of samples containing **DFHBI** dye, no analyte (SSA+Adp), **Am_DNA** (SSA+Adp+matched analyte), and **Amm_DNA** (SSA+Adp+mismatched analyte). C) Dashed lines represent triethylene glycol (TEG) linkers.

Figure S4: Suboptimal designs of USSA analyzing Mtb. F11/KZN analyte

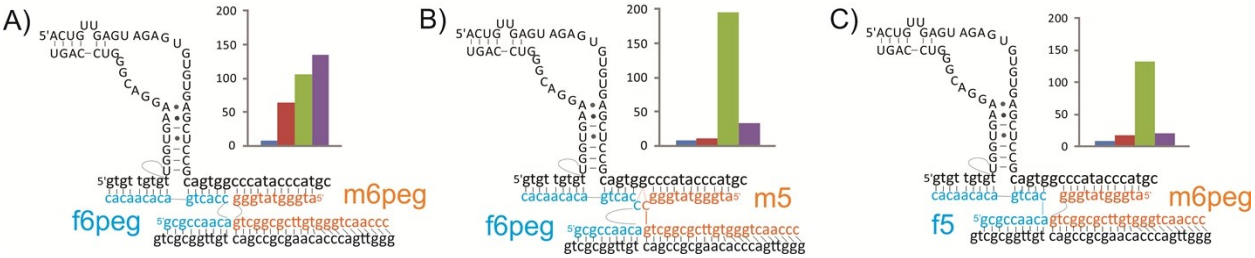

Figure S4. Fluorescence complexes of USSA with the DNA analyte and responses of the two different combinations of stands m and f. Data for each combination is shown on the right side of each panel. Samples contained 3.6 µM **SSA_f**; 2.0 µM **SSA_m**; 0.4 µM **Adp_f**; 4.0µM **Adp_m**; 6.0 µM analyte; 5 µM **DFHBI**; in 2 × Spinach Buffer 2. Emission spectrum (λ_{ex} = 450 nm) were collected after 30 min at 22.5˚C. A, B, C) The bar graph shows the fluorescence intensity of samples containing **DFHBI** dye, no analyte (SSA+Adp), **Am_DNA** (SSA+Adp+matched analyte), and **Amm_DNA** (SSA+Adp+mismatched analyte). Deshed lines represents TEG linkers.

6) Figure S5. Analysis of long analyte using USSA probe

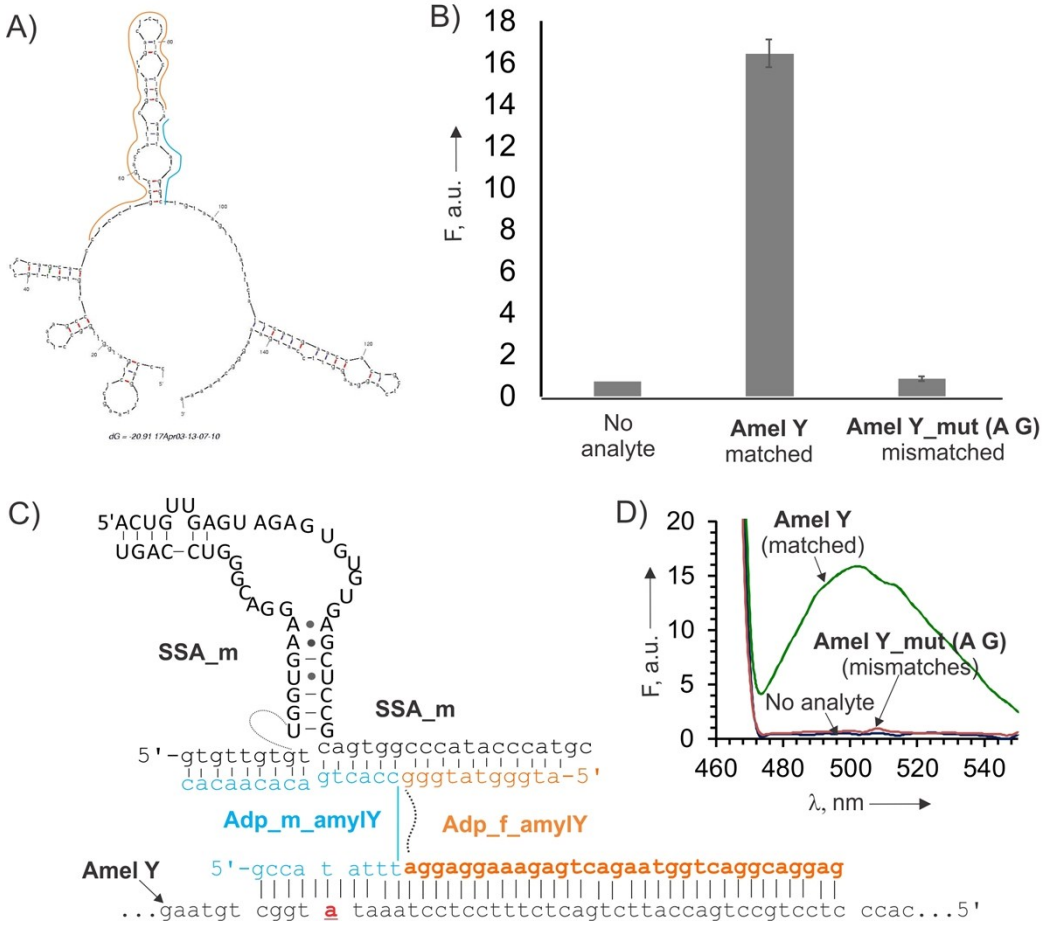

Figure S5. Detection of a long DNA analyte by USSA_AMELY probe. A) Secondary structure of 152 nt AMELY analyte predicted by mfold.[1] The fragment bound by **Adp_f_amelY** and **Adp_m_amelY** are indicated by cyan and orange lines respectively. B) Fluorescent response of the USSA_AMELY probe in the presence of fully matched **AMEL Y** or single base mismatched **Amel Y_mut (A→ G)** analyte. 'No analyte' is the probe response in the absence of analytes (negative control). The assay conditions: **DFHBI** (1 μM), **SSA_m** (2 μM), **SSA_f** (2 μM), **Adp_f_amelY** (2 μM), **Adp_m_amelY** (0.5 μM) in Spinach Buffer 2. C) The predicated secondary structure of USSA_AMELY probe with fully matched **AMELY** analyte. Single nucleotide substitution site is red underlined. For full sequence of Amel Y analyte see Table S1. D) Representative spectrums of USSA_AMELY probe in the presence of matched AmelY (green) or mismatched (**Amel Y_mut (A→ G)**) analytes. Blue line is USSA_AMELY probe is the absence of the analyte (negative control)

In order to prove that SSA is suitable for the detection of long structure-forming analytes, we designed a probe for detection of AMELY (Figure S5A). DNA fragments of AMEL X and AMEL Y genes, which code for the amelogenin – gene involved in the development of the enamel.[2] Different amelogenin sequences are located in X and Y chromosomes, which makes this gene to be a good marker for sex determination, in forensic and anthropological applications.[3] The analyte binding arm of the adaptor strand **Adp_m_amelY** was designed to unwind the secondary structure

of the analyte (Figure 3SA). The analyte binding arm of **Adp_m_amelY** was short to enable highly specific hybridization of signal nucleotide substitution. The signal-to-background ratio was ~20 under optimized concentrations of the USSA components (Figure S5B). Importantly, a single base substituted **Adp_m_amelY** produced fluorescence only slightly above the background (Figure S5B, last bar).

7) References

1) M. Zuker, Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.,* 2003, **31**, 3406-3415.

2) B. A. Álvarez-Sandoval, L. R. Manzanilla, R. Montiel, Sex determination in highly fragmented human DNA by high-resolution melting (HRM) analysis. *PLoS One,* 2014, **9**, e104629.

3) H. Nogami, H. Tsutsumi, T. Komuro, R. Mukoyama Rapid and simple sex determination method from dental pulp by loop-mediated isothermal amplification. *Forensic Sci Int Genet.* 2008, **2**, 349-353.