Supporting Information

Mutation-Resistant Deoxyribozyme OR gate for Highly Selective Detection of Viral Nucleic Acids

Ola Kamar,[a] Sin-Cih Sun,[c] Chieh-Hua Lin,[d] Wan-Yu Chung,[e] Min-Shi Lee,[e] Yu-Chieh Liao,*[d] Dmitry M. Kolpashchikov*[ab] and Min-Chieh Chuang*[c]

[a]Chemistry Department, University of Central Florida, 4000 Central Florida Boulevard, Orlando, FL 32816-2366, USA
[b]Burnett School of Biomedical Sciences, University of Central Florida, Orlando, 32816, Florida, USA
[c]Department of Chemistry, Tunghai University, Taichung, 40704, Taiwan
[d]Institute of Population Health Sciences, National Health Research Institutes, Zhunan, 35053, Miaoli, Taiwan
[e]National Institute of Infectious Diseases and Vaccinology, National Health Research Institutes, Zhunan, 35053, Miaoli, Taiwan
[f]ITMO University, Laboratory of Solution Chemistry of Advanced Materials and Technologies, Lomonosova St. 9, 191002, St. Petersburg, Russian Federation.

*Corresponding authors:
Y. -C. Liao, E-mail: jade@nhri.org.tw; Tel.: (+886) 37 246-166 ext. 36178; Fax: (+886) 37 586-467
D. M. Kolpashchikov, E-mail: Dmitry.Kolpashchikov@ucf.edu; Tel.: (+1) 407-823-6752; Fax: (+1) 407-823-2252
M. -C. Chuang, E-mail: mcchuang@thu.edu.tw; Tel: (+886) 4 2359-0121 ext. 32218; Fax: (+886) 4 2359-0426

Table of Contents

1. Materials and Instrumentation

DNAse/RNAse-free water was purchased from Fisher Scientific, Inc. (Pittsburgh, PA) and used for all buffers and for the stock solutions of oligonucleotides. Fluorogenic substrates (**F-sub**) were synthesised and HPLC purified by TriLink BioTechnologies, Inc. (San Diego, CA). All other oligonucleotides were obtained from Integrated DNA Technologies, Inc. (Coralville, IA). The oligonucleotides were dissolved in water and stored at -20 ^{o}C until needed. Stock concentrations of oligonucleotides were calculated by measuring the absorption of the solutions at 260 nm by using a Perkin–Elmer Lambda 35 UV/Vis spectrometer (San Jose, CA). Extinction coefficients of oligonucleotides were calculated by using OligoAnalyzer 3.1 software (Integrated DNA Technologies, Inc.). Fluorescent spectra were recorded on a Perkin–Elmer LS-55 luminescence spectrometer equipped with a Hamamatsu xenon lamp. Experiments were performed at excitation wavelength of 485 nm. Emission of FAM was monitored at 517 nm. Excitation and emission slits were both 10 nm. For the detection of genuine sequences, fluorescence emission was measured using an Applied Biosystems StepOne real-time PCR.

2. Detailed Experimental Procedures

Fluorescence of sample mixtures were measured in comparison to two blanks, first blank is the 200 nM **F-sub** in the buffer (50 mM HEPES, pH 7.4, 50 mM $MgCl_2$, 20 mM KCl, 120 mM NaCl, 0.03% Triton X-100%, and 1% DMSO) only and the second blank is the 200 nM **F-sub** and deoxyribozyme (DZa and DZb) of the three groups as 7 nM DZa-I, 15 nM DZb-I, 12 nM DZa-II, 30 nM DZb-II, 2 nM DZa-III, and 30 nM DZb-III in the same buffer. Fluorescence of 10 nM of each EV-71 synthetic analytes and 10 nM of each CV-A16 synthetic analytes were measured after 1 hour and 3 hours of incubation at 28 ^{o}C. Limit of detection were determined for EV-III-5 and of EV-I-1 analytes after 1 and 3 hours of incubation at 28 ^{o}C. In the measurement of genuine sequences, the synthetic sequence analytes were replaced with the Lambda exonuclease-treated PCR products. The PCR products (10 μL) were incubated with 6 unit of Lambda Exonuclease (New England Biolabs) at 37 °C for 30 min, followed by enzyme inactivation at 75 °C for 10 min. After the reaction with lambda exonuclease, the ssPCR products (20 μL) were mixed with sodium acetate (final concentration of 0.3 M) and 75 μL ethanol, followed by vortex mix thoroughly. The mixture was stored at −20 °C overnight and subsequently was centrifuged at 14,000 × g for 20 min at 4 °C. The supernatant was taken away; the pellet was slowly rinsed with 200 μL of 95% ethanol. Then the sample was dried under vacuum and was re-suspended in 192.5 μL buffer. The above solution was mixed with the six adaptor strands and **F-sub** for the fluorescence measurements in a total volume of 220 μL. The control sample was prepared by performing same PCR in the absence of viral cDNA.

3. RNA extraction, RT-PCR, and PCR

RNA extraction. Viral RNA was extracted from the clinical specimens (throat swab culture) using a NucleoSpin® RNA Virus (Macherey-Nagel, Germany). Briefly the cell culture supernatant (150 μL) was incubated with RAV1 buffer (600 μL) at 4 °C for 10 min, followed by a mixing with 100%

ethanol (600 μL). An aliquot (650 μL) of the mixed solution was transferred to a NucleoSpin® RNA Virus Columns and was centrifuged at 8,000 rpm for 1 min and the filtrate was removed. Subsequently the procedure containing washing (using 600 μL ethanol-added wash buffer) and centrifugation (13,000 rpm for 3 min) was iterated twice. The RNA was eventually eluted with pre-heated (70 °C) RNAse-free water (30 μL) under 13,000 rpm for 2 min.

Reverse transcription-PCR (RT-PCR). The IQ2 MMLV RT-script & cDNA synthesis accessory kit (Bio Genesis, Taiwan) was used for the reverse transcription-PCR. A primer (5'-CCYTGYTCCATNRCNTCNTC) was designed by using OligoAnalyzer 3.1 and used to produce cDNA given from the templates of EV-E59 and EV-B5/141 viral RNAs. The RNA (400-500 ng) was incubated with the primer (10 μM) at 65 °C for 2 min, then the mixture was immediately put onto ice bath for 1 min. RT-PCR was performed by mixing the RNA/primer mixture (10 μL), dNTP (1 μL, 10 mM), DTT (2 μL, 0.1 M), RNase inhibitor (20 U/μL), and HiScript I Reverse Transcriptase (5 U/μL) in 1X First-Strand Buffer with a total volume of 20 μL, subsequently heated at 42 °C for 90 min and 85 °C for 5 min. The cDNA obtained was stored at −20 °C until a use.

Polymerase Chain Reaction (PCR). The forward (5'-GGGCCATCTACGTGGG) and reverse (5'-/Phos/GGTAATACTCGCTAGCTTCCAC, Phos refers to a phosphorylation terminal end) for amplification of EV-E59 and EV-B5/141 were designed by the OligoAnalyzer 3.1. The PCR product was generated from a 10 μL PCR mixture composed of dNTP (200 μM), FP (0.5 μM), RP (0.5 μM), cDNA (1 μL), and Phusion Hot Start II High-Fidelity DNA Polymerase (0.02 U/μL) in 1X Phusion GC buffer. Reactions were performed with 60 cycles of PCR in MiniOpticon™ Thermal Cycler (Bio-Rad). The amplification cycles included 1 cycle of 2 min at 98 °C, followed by 60 cycles containing 98 °C for 10 s, 61.5 °C for 30 s, and 72 °C for 30 s, and the last extension at 72 °C for 10 min, before returning to 4 °C for 10 min.

4. Sequence selection

Genome of enterovirus consists of approximately 7500 nucleotides that include an open reading frame (ORF) and untranslated regions (URTs) franked by 5' and 3' terminal. The ORF was divided into three domains, P1, P2, and P3, which encodes polyproteins containing VP1-4, 2A-2C and 3A-3D, respectively [1]. To select the targeted fragments specific to EV71, all the enterovirus sequences deposited in GenBank were downloaded. Genotype classification of these enterovirus sequences were examined by comparing them with the classification results obtained from EVIDENCE [2]. Incongruent genotype assignments between GenBank and EVIDENCE were excluded. The obtained sequences were further separated into EV71 and non-EV71 sequences to produce 8455 EV71 sequences. Subsequently the nucleotide regions corresponding to the 11 polyproteins in each sequence were defined. It was found that 15-base was the longest nucleotide length which was conserved in EV71 sequences. As a result, more than 387,000 oligonucleotide fragments (each is 15-nt) were obtained from the 8455 EV71 sequences. However, the enterovirus sequences are highly mutated; neither the 15-nt fragment was able to

represent all EV71 isolates. We therefore performed three-fragment combination (3-combination) to represent EV71, that is basically the OR gate concept of the study. Since there produced great number of the combinations, we further reduced the data size by excluding the fragments which either represent less than 30% of whole EV71 isolates or are present in non-EV71 sequences, which produced the no. of fragments and 3-combinations revealed in Table S1. Coverage rate of the 3-fragment combination in each protein region was also estimated. Notably, although the highest coverage rate (93.3%) was obtained for VP4, one of its targeted fragments, GTTTGCAAATCCTGT, was present in the lately submitted CV-A16 isolate (KU163608.1). Similarly, the fragments selected for VP2 and VP3 also appeared in CV-A16 and CVA10 isolates, respectively. To ensure a high selectivity of the sensors, we selected the fragments obtained from 2A (CTTTAGAGTGGTCAA, TTGGGAAGACAGCTC, and ACCCAGCCTCATATA) to be the EV71-specific fragments which achieved a 90.5%- coverage rate. It should be furthermore noted that the 15-nt targeted fragments were insufficiently long to bind with the two adaptor strands (DZa and DZb) stably, which hampered the formation of BiDZ catalytic core. We therefore made 3-nt extensions at both 5’ and 3’ terminals of the selected targeted fragments to produce 21-nt targets. Because the 1st and the 19th nucleotide (counted by 5’ $\rightarrow$ 3’) of the 21-nt targets were generally variable, different targets containing two nucleotide substitutions were designed in accordance with the varied nucleotides present in the isolates. The corresponding GenBank accession numbers are depicted in Table S2. In addition, degenerate probes of DZa and DZb were also designed for high hybridization efficiency.

For the selectivity study, eight Coxsackievirus sequences containing one to three nucleotide variations from EV71 isolates were selected. The corresponding GenBank accession numbers are indicated in Table S2.

Table S1. Result summary of sequence selection

Protein	No. of sequences	No. of fragments	No. of 3-combinations	Target sequence 1	Target sequence 2	Target sequence 3	Maximum coverage rate (%)
VP4	1478	121	287980	GTTTGCAAATCCTGT	ATCTTCACTGAAATG	TTTCACTGAAATGGC	93.37
VP2	1105	432	13343760	ATGGTGAGTGGCCTT	ACACCCGTACGTGCT	AACGCGCCCAGATGT	92.04
VP3	1036	443	14391741	ACAGCCTATATAATA	ATCTGGGATTTTGGG	AGTGATATCCTGCAG	88.51
VP1	6682	627	40885625	CTTTTGTTGCGTGCA	ACAGCTGAGACCACT	ACACACAGGTGAGCA	92.01
2A	1208	105	187460	TTGGGAAGACAGCTC	CTTTAGAGTGGTCAA	ACCCAGCCTCATATA	90.56
2B	677	188	1089836	GATAGGTTGTCATGG	ATGATATGGTTACCC	AACAGGCTTCACTGA	77.70
2C	657	374	8649124	ACCAGGCACTGGGAA	AGAGTGGGTTTCCAA	ATGTCTCATCATTAG	83.11
3A	641	117	260130	AGCCCCAGACGCTAT	AAGTTCAGGCCAATT	CTATTAGCGATCTCC	80.03
3B	8	41	10660	AACAGCCACGGTCCA	AAGAAGCCCGTGTTA	AGCATATTCCGGCGC	N.D.
3C	1371	180	955860	ACATCAGTCGGTAAG	ATTCACATTGGTGGC	GACATCTGTTGGGAA	74.47
3D	1163	395	10193765	TAAATGATTCAGTGT	ATCCCTTCCCAATTG	ATGCCAATGAGGGAG	60.53

Table S2. Targeted fragments (and their GenBank accession numbers) selected in this study

Name	Sequence 5' → 3'	GenBank number (exemplified)
EV-I-1	TAA CTT TAG AGT GGT CAA TCG	KF982854.1
EV-I-2	CAA CTT TAG AGT GGT CAA TCG	KU159442.1
EV-I-3	TAA CTT TAG AGT GGT CAA CCG	KJ686293.1
EV-I-4	CAA CTT TAG AGT GGT CAA CCG	KY582572.1
Cox-I-1	TAA CTA TAG AGT GGT CAA TCG	JQ746659.1 (CV-A16)
Cox-I-2	TAA CTT TAG AGT GGT AAA TCG	KX064305.1 (CV-A6)
EV-II-1	AGT TTG GGA AGA CAG CTC ACG	N/A
EV-II-2	AGT TTG GGA AGA CAG CTC CCG	AB550332.1
EV-II-3	AGT TTG GGA AGA CAG CTC TCG	HQ825317.1
EV-II-4	CGT TTG GGA AGA CAG CTC ACG	N/A
EV-II-5	CGT TTG GGA AGA CAG CTC CCG	KT354870.1
EV-II-6	CGT TTG GGA AGA CAG CTC TCG	KU159436.1
EV-II-7	TGT TTG GGA AGA CAG CTC ACG	KP308443.1
EV-II-8	TGT TTG GGA AGA CAG CTC CCG	KX372324.1
EV-II-9	TGT TTG GGA AGA CAG CTC TCG	KY425527.1
Cox-II-1	TGT TTG GGA AGA CAG CAC CCG	KX156360.1 (CV-A2)
Cox-II-2	CGT GTG GGA AGA CAG CTC TAG	JQ746674.1 (CV-A16)
Cox-II-3	CGT GTG GGA GGA CAG CTC TAG	KY674980.1 (CV-A16)
EV-III-1	AAA ACC CAG CCT CAT ATA TGT	AB747373.1
EV-III-2	CAA ACC CAG CCT CAT ATA TGT	KP274877.1
EV-III-3	TAA ACC CAG CCT CAT ATA TGT	KX372319.1
EV-III-4	AAA ACC CAG CCT CAT ATA CGT	KR045300.1
EV-III-5	CAA ACC CAG CCT CAT ATA CGT	KP691658.1
EV-III-6	TAA ACC CAG CCT CAT ATA CGT	KX372331.1
Cox-III-1	CAA ACC CAG CCT GAT ATT TGT	KF055242.1 (CV-A16)
Cox-III-2	TAA ACC CAG CCT CGT CTT CGT	KX768158.1 (CV-A10)
Cox-III-3	CAA GCC CAG CCT GAT ATA TGT	LT617104.1 (CV-A16)

Table S3. Sequences of deoxyribozyme and the DNA analyte mimic sequences used in this study

Group	Name	Sequence 5'→ 3'
	F-Sub	AAG GTT^{FAM} TCC TCg uCCC TGG GCA-BHQ1
I	**DZa-I**	CGR TTG ACC AC*A CAA CGA* GAGGAAAC
	DZb-I	CCA GGG A *GGCTAGCT* TCT AAAG TTR
	EV-I-1	**T**AA CTT TAG AGT GGT CAA **T**CG
	EV-I-2	**C**AA CTT TAG AGT GGT CAA **T**CG
	EV-I-3	**T**AA CTT TAG AGT GGT CAA **C**CG
	EV-I-4	**C**AA CTT TAG AGT GGT CAA **C**CG
	Cox-I-1	TAA CT**A** TAG AGT GGT CAA TCG
	Cox-I-2	TAA CTT TAG AGT GGT **A**AA TCG
II	**DZa-II**	CGD GAGCTGTC *ACAACGA* GAGGAAAC
	DZb-II	TGCCCA GTG A *GGCTAGCT* TTC CCA AAC D
	EV-II-1	**A**GT TTG GGA AGA CAG CTC **A**CG
	EV-II-2	**A**GT TTG GGA AGA CAG CTC **C**CG
	EV-II-3	**A**GT TTG GGA AGA CAG CTC **T**CG
	EV-II-4	**C**GT TTG GGA AGA CAG CTC **A**CG
	EV-II-5	**C**GT TTG GGA AGA CAG CTC **C**CG
	EV-II-6	**C**GT TTG GGA AGA CAG CTC **T**CG
	EV-II-7	**T**GT TTG GGA AGA CAG CTC **A**CG
	EV-II-8	**T**GT TTG GGA AGA CAG CTC **C**CG
	EV-II-9	**T**GT TTG GGA AGA CAG CTC **T**CG
	Cox-II-1	TGT TTG GGA AGA CAG C**A**C CCG
	Cox-II-2	CGT **G**TG GGA AGA CAG CTC T**A**G
	Cox-II-3	CGT **G**TG GGA **G**GA CAG CTC T**A**G
III	**DZa-III**	ACR TAT ATG AG *ACAACGA* GAGGAAAC
	DZb-III	CCA GGG A *GGC TAG CTG* CTG GGT TTD
	EV-III-1	**A**AA ACC CAG CCT CAT ATA **T**GT
	EV-III-2	**C**AA ACC CAG CCT CAT ATA **T**GT
	EV-III-3	**T**AA ACC CAG CCT CAT ATA **T**GT
	EV-III-4	**A**AA ACC CAG CCT CAT ATA **C**GT
	EV-III-5	**C**AA ACC CAG CCT CAT ATA **C**GT
	EV-III-6	**T**AA ACC CAG CCT CAT ATA **C**GT
	Cox-III-1	CAA ACC CAG CCT **G**AT AT**T** TGT
	Cox-III-2	TAA ACC CAG CCT C**G**T **C**T**T** CGT
	Cox-III-3	CAA **G**CC CAG CCT **G**AT ATA TGT

BHQ1, Black Hole Quencher 1; FAM, fluorescein label; **F-Sub** binding arms are underlined; the SNS sites are bold underlined; nucleotides part of the DZ catalytic core are in italic; ribonucleotides are in low case; analyte binding arm of DZa and DZb are in blue color; R is A or G; D is A, G or T

Schematics illustrating the interactions of the most and the least fluorescing analytes with DZ sensors

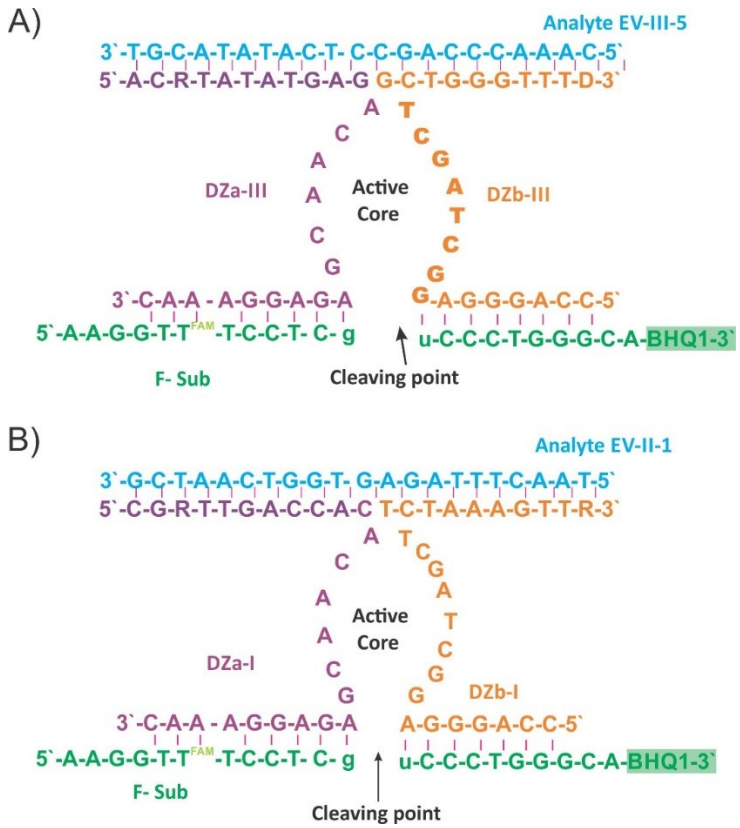
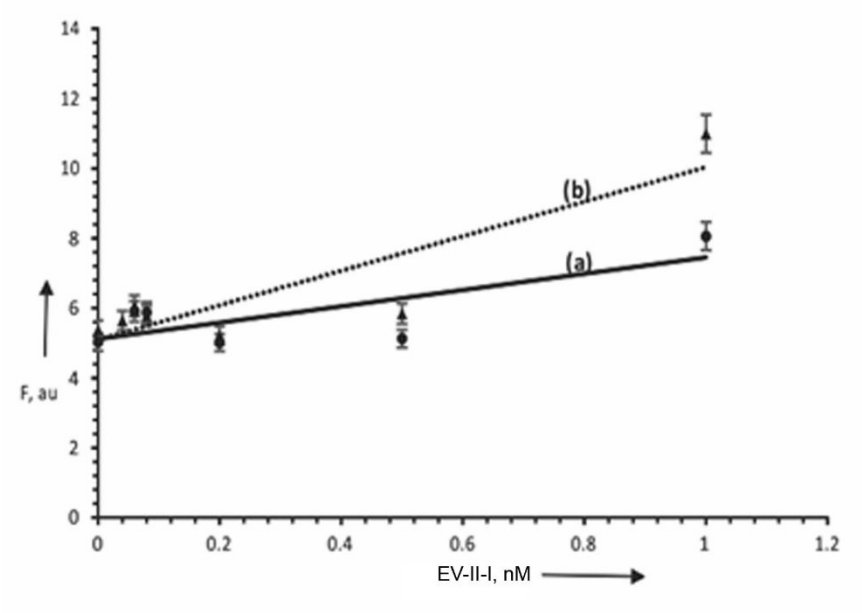


Fig. S1. Structures of BiDZ sensors in complex with cognate analytes. (A) DZ sensor EV-III-5 that produces the highest S/B ratio. (B) DZ sensor EV-II-1, which produces the lowest S/B ratio.

5. **Limit of detection**

A)



B)

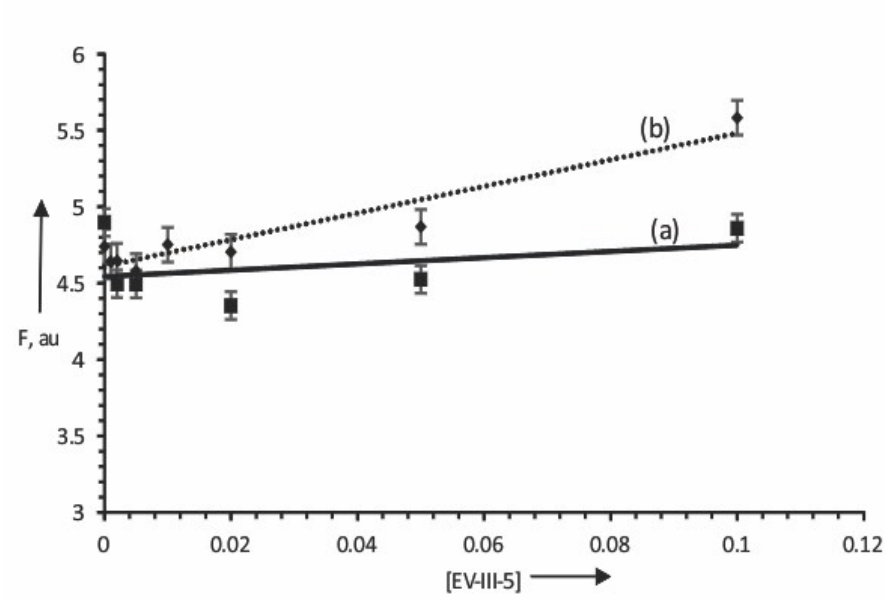


Fig. S2. Limits of detection for the DZ OR gate and two selected analytes, EV-III-5 (A) and EV-II-1 (B). **F_sub** (200 nM), DZa-I (7 nM), DZb-I (15 nM), DZa-II (12 nM), DZb-II (30 nM), DZa-III (2 nM), and DZb-III (30 nM) were incubated in the absence or presence of different concentrations of (i) EV-II-1 analyte (A) of EV-III-5 (B) in 50 mM HEPES, pH 7.4, 50 mM $MgCl_2$, 20 mM KCl, 120 mM NaCl, 0.03% Triton X-100, 1% DMSO at 28^{o}C. Fluorescent intensities were measured at 517 nM (excitation at 485 nm) after 1 hr (a) and 3 hrs of incubation (b). The data of 3 independent experiments with standard deviations are presented.

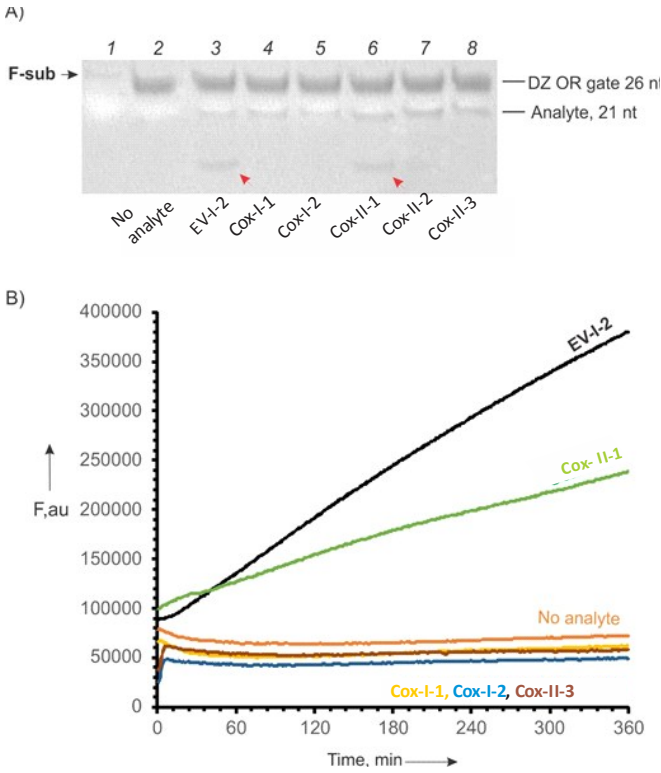


Fig. S3. Analysis of DZ OR gate performance by gel electrophoresis and kinetic of fluorescent signal accumulation. A) Polyacrylamide gel electrophoresis (PAGE) analysis of reaction mixtures. All reaction mixtures contained 200 nM F-sub, 7 nM DZaI, 15 nM DZbI, 12 nM DZaII, 30 nM DZbII, 2 nM DZaIII, 30 nM DZbIII, 10 nM analyte. Samples in lanes 3-8 in addition contained correspondent analyte strands. The samples were analyzed by 20% PAGE containing 7 M urea followed by staining with SYBR Gold (Thermofisher) and imaging using BioRad GelDoc XR+ system. The mobility of stands in the gel are indicated by arrows on the side of the picture. The fluorescent product of F-sub cleavage is indicated by the red arrows. B) The samples were composed as described for panel A followed by fluorescent measurement using QuantStudio™ 6 at a constant temperature of 28°C every 1 min over 6 hrs.

We further analyzed the reaction mixtures containing OR gates using denaturing polyacrylamide gel electrophoresis (Fig. S3 A). The product of F-sub cleavage was observed only in the presence of complementary EV-1-2 and mismatched Cox-II-1 analytes (indicated by red arrows in Fig. S3A), which is consistent with the fluorescent results reported in the main text (Fig.2). No substrate cleavage was observed in the absence of analytes (Lane 2) or in the presence of mismatched (Cox-I-1, Cox-I-2, Cox-II-2 and Cox-II-3, lanes 4, 5, 7 and 8, respectively).

Measurement of kinetics of fluorescent response revealed significant fluorescent increase over time only in the presence of fully complementary EV-1-2 and mismatched Cox-II-1 analytes in accordance with data of Fig 2 and Fig. S3, panel A.

6. References

1. Steven, M., Penaranda, S., Maher, Kaija, and Pallansch, M. Journal of General Virology, 2004, 85, 1597-1607.
2. Lin CH, Wang YB, Chen SH, Hsiung CA, Lin CY. Precise genotyping and recombination detection of Enterovirus. BMC Genomics. 2015;16 Suppl 12:S8.