Electronic Supplementary Material (ESI) for ChemComm.

Supporting information

A DNA-assisted immunoassay for enzyme activity via a DNA-linked, activity based probe

Dongwook Kim, Rachael R. Jetson, Casey J. Krusemark*

Department of Medicinal Chemistry and Molecular Pharmacology, Purdue Center for Cancer Research, Purdue University, West Lafayette, IN 47907, U.S.A

*Email: cjk@purdue.edu

Experimental procedures

Materials and Instrumentation

Azido-fluorophosphonate (FP) was purchased from Thermo Scientific. All modified oligonucleotides and the set of 96 49-mer oligos were purchased from Integrated DNA Technologies (IDT). All 55-mers oligonucleotides were obtained from Bioneer (Seoul, Korea). Bovine Factor Xa, human thrombin, and antibodies for Xa and thrombin were purchased from Haematologic Technologies. Human recombinant DPP4 was purchased from MyBioSource. The DPP-4 antibody (Polyclonal goat IgG, AF1180) was purchased from R&D Systems. Human serum (male AB, sterile-filtered) and citrated human plasma were purchased from Sigma Aldrich.

Oligo conjugation reactions were monitored and purified on a Varian Pro Star HPLC using Agilent Microsorb-MV 300-5 C18 250 x 4.6 mm reverse phase columns using a linear gradient of 0.25% hexafluoroisopropanol (HFIP) adjusted to pH 7.0 with triethylamine with 5 μM EDTA in water (eluent A) and 0.25% hexafluoroisopropanol (HFIP) adjusted to pH 7.0 with triethylamine with 5 μM EDTA in 90% methanol (eluent

B). Routine PCR was conducted on a BIO RAD C1000 Touch Thermo Cycler. qPCR was conducted using an Applied Biosystems ViiA 7 Real Time PCR system using Absolute Blue qPCR SYBR Green Mix (ThermoFisher). Fluorescence gel imaging was performed on GE Heathcare Typhoon Trio+. DNA prior to sequencing was quantified on Agilent Bioanalyzer and then sequenced on an Illumina MiSeq in the Purdue Genomics Core Facility.

Preparation of alkyne oligonucleotide conjugates

A common primer (CP) 20-mer sequence with a 5' amine modification (/5AmMC12/ATGGTATCAAGCTTGCCACA, IDT) was acylated with hexynoic acid, as previously described[1]. Briefly, 40 nmol oligo was acylated while immobilized on DEAE Sepharose (GE Healthcare), eluted with 1.5 M NaCl, and HPLC purified. Collected fractions were reduced by SpeedVac and resuspended in 50 μl H_2O. The oligo was precipitated by addition of an equal volume of 10 M ammonium acetate and 6 volumes of cold ethanol. After overnight incubation at -20 °C, the sample was centrifuged at 16,000 rpm at 4 °C. The pellet was washed with 95% ethanol, dissolved in H_2O, and quantified by UV absorbance (19 nmol, 48%). Mass spectrometry analysis was performed with 1μM solutions in 50% methanol with 1% triethylamine. ESI: $(M-11H)^{11-}$ 586.0 (calcd. 586.1), $(M-10H)^{10-}$ 644.8 (calcd. 644.8), $(M-9H)^{9-}$ 716.5 (calcd. 716.6), $(M-8H)^{8-}$ 806.3 (calcd. 806.3), $(M-7H)^{7-}$ 921.5 (calcd. 921.6), $(M-6H)^{6-}$ 1075.2 (calcd. 1075.4).

For fluorescence detection (Figure 1B, Figure S2) in SDS-PAGE, a 20-mer sequence (/56-FAM/TGTGGCAAGCTTGATACCAT/3AmMO/, IDT) (50 nmol) with a 3' amine

modification and a 5' fluorescein amide (FAM) modification was acylated and purified as with the CP sequence (39 nmol, 78%). ESI: $(M-11H)^{11-}$ 633.2 (calcd. 633.1), $(M-10H)^{10-}$ 696.6 (calcd. 696.5), $(M-9H)^{9-}$ 774.1 (calcd. 774.0), $(M-8H)^{8-}$ 870.8 (calcd. 870.8), $(M-7H)^{7-}$ 995.4 (calcd. 995.4).

Preparation of DNA-linked fluorophosphonate

Click reactions of Azido-FP (30 μM) and alkyne-modified oligos (30 μM) additionally contained 5 mM aminoguanidine, 50 mM phosphate buffer pH 7, 0.4 mM $CuSO_4$, and 2 mM tris[(1-hydroxy-propyl-1*H*-1,2,3-triazol-4-yl)methyl]amine (THPTA) (Sigma-Aldrich), and 5 mM sodium ascorbate with 15% DMSO. The reaction was performed at room temperature (RT) for 30 minutes. Conjugates were then used directly in protein labeling reactions. Oligo-FP conjugation was monitored by HPLC (Fig. S1), and product peaks were collected and characterized by direct electrospray mass spectrometry.

CP 20-mer with 5' FP (/FP/ATGGTATCAAGCTTGCCACA) ESI: $(M-11H)^{11-}$ 618.1 (calcd. 618.0), $(M-10H)^{10-}$ 679.9 (calcd. 679.9), $(M-9H)^{9-}$ 755.4 (calcd. 755.5), $(M-8H)^{8-}$ 849.9 (calcd. 850.1), $(M-7H)^{7-}$ 971.6 (calcd. 971.7), $(M-6H)^{6-}$ 1133.7 (calcd. 1133.8), $(M-5H)^{5-}$ 1360.7 (calcd. 1360.8), $(M-4H)^{4-}$ 1701.1 (calcd. 1701.2).

20-mer with 5' FAM and 3' FP (FAM/TGTGGCAAGCTTGATACCAT/FP/) ESI: $(M-11H)^{11-}$ 609.2 (calcd. 609.4), $(M-12H)^{12-}$ 664.9 (calcd. 664.9), $(M-11H)^{11-}$ 731.5 (calcd. 731.5), $(M-10H)^{10-}$ 812.7 (calcd. 812.9), $(M-9H)^{9-}$ 914.5 (calcd. 914.6), $(M-8H)^{8-}$ 1045.5 (calcd. 1045.4).

Labeling of recombinant serine hydrolases with DNA encoded probe for SDS-PAGE analysis

Each protein (Factor Xa (Figure 1B) or thrombin (Figure S2)) was incubated at approximately 10 μM concentration in PBS with 0.5 mg/mL of BSA and the FP-oligo conjugate at 15 μM (total oligo concentration). The mixture was allowed to react at RT for 3 hours. A fraction of the labeling reaction containing approximately 30 pmol of the target protein was incubated with 5 μg antibody for 45 minutes in 5 μl. This mixture was then mixed with 30 μl protein G slurry (Thermo Scientific Pierce) and incubated for 30 minutes. The supernatant was removed, and the beads were washed 4 times with TBST using 5 μl initially and then 15 μl volumes thereafter. Protein was eluted from the beads with SDS loading buffer. Protein samples were analyzed by SDS-PAGE either on an Any kD Mini-PROTEAN TGX Protein Gel (Bio-rad) (Factor Xa) or a standard 10% gel (thrombin). The gels were scanned using a Typhoon Trio plus variable mode imager (GE Healthcare) set to the manufacture's recommended setting to detect fluorescein fluorescence. After scanning, the gels were stained with Coomassie brilliant blue R-250 (Sigma).

Labeling and activity detection by qPCR

For plus/minus enzyme (Figure 2A) and crosstalk (Figure 2B) experiments, proteins were labeled at 5 μM concentration with FP-oligo (CP 20-mer) at 10 μM (taken directly from click conjugation reaction) for 30 minutes at RT. For enzyme and inhibitor titration experiments (Figure 3), 10 μl samples were prepared with enzymes and inhibitors as

indicated and FP-oligo at 10 μM and allowed to react for 30 minutes at RT. Subsequently, 3 distinct 55-mer DNAs (Table S1) were added to each sample at 1.5 equivalents total to the FP-oligo (5 μM final concentration in PBS). An equimolar portion (to the total of the 55-mers) of a 20-mer complementary to the sample encoding region, as well as an equimolar portion of a 15-mer (GCTCTCTACTGCATA) complementary to the common intervening sequence were also added. The mixture was allowed to hybridize at RT for 30 minutes. All samples were then pooled into a solution containing an excess (10x equivalents to the FP-oligo) of an unmodified CP 20-mer oligo to limit potential crosstalk from oligo melting and reannealing. A small portion of this mixture was reserved for an initial abundance determination by qPCR. Protein G magnetic beads (20 μg protein G, about 2 equivalents over target protein) were pre-blocked with 20 μl PBST-BT (binding buffer) (0.5 mg/mL of BSA and 0.5 mg/mL yeast tRNA). An antibody (2 equivalents over target protein), and the protein-DNA mixture was added to the protein G magnetic beads and incubated at RT for 30 minutes. After removing unbound fraction using a magnetic stand, the sample was washed with 20 μl volumes of PBST-BT at RT for 5 minutes each for a total of three times, and then PBST-BT with 250 mM NaCl (high salt washing buffer) at RT for 5 minutes each for a total of three times. To elute bound DNA, 20 μl pure water was added, and the sample was heated to 95 °C for 3 minutes. Eluted material was analyzed by qPCR after appropriate dilution.

Labeling of serum and plasma protein with FP-oligo probe (Figure S4B)

DPP-4 antibody (6 ug) was added into 10 μL of pre-blocked protein G magnetic beads (1x PBST), and incubated at RT for 30 min. Human plasma (2 μg at 10 mg/mL) was added into the antibody-protein G beads at RT for 30 minutes. After washing twice with

20 µl volumes of binding buffer, the beads were suspended in 20 µl volume containing the CP 20-mer FP probe (7.5 µM) and three different 55-mers per each sample (8 µM total) and incubated at RT for 30 minutes. After washing with 20 µl volumes of binding (3x) and washing buffer (2x), the remaining DNA was eluted by 20 µl pure water after heating at 95 °C for 3 minutes.

Activity detection by DNA sequencing

For the control biotinylation experiment (Figure 4A), 48 of the 49-mers (20 nM each) (Table S3) were mixed in PBS with 0.2 equivalents of a CP 5' biotinylated 20-mer ((/biotin/ATGGTATCAAGCTTGCCACA, IDT), 1.8 equivalents of an unmodified CP 20-mer, and 1 equivalent of a 20-mer (AGATCGGAAGAGCGGTTCAG) complementary to the 5' portion of the 49-mers. Similarly, the remaining 48 49-mers were mixed with 2 equivalents of the unmodified CP 20-mer and 1 equivalent to the additional 20-mer. The samples were individually heated to 95 °C for 3 minutes and then allowed to cool slowly to RT over 30 minutes. Both samples were then pooled into a solution containing an excess (100 µM) of the unmodified CP 20-mer, and a small portion of this mixture was removed for determination of initial abundance by DNA sequencing. This mixture was then incubated with 10 µg streptavidin magnetic beads (NanoLink Streptavidin Magnetic Beads, 0.8 µm diameter) for 30 minutes at RT. The beads were decanted and washed with 20 µl volumes 3x with PBST-BT, 3x with PBST-BT with 250 mM NaCl, and finally 3x with PBS. Bound DNA was eluted with in pure water by heating to 95 °C for 3 minutes. Both the pre-selection and post-selection (elution) pools were PCR amplified with 10 cycles using adaptor primers TS-Index18-F (pre) or TS-Index19-F (post) as

forward primers and TS-Uni-9mer-R (Table S4) as the reverse primers. PCR products were submitted for sequencing on an Illumina MiSeq at the Purdue Genomics Facility. Reads were matched to barcodes are previously described[1] and considered positive only for exact 9 base matches. Barcode read numbers are shown in Table S5.

For the thrombin inhibition assay (Figure 4B), thrombin (2 μM) was labeled with 4 μM CP 20-mer FP probe in the presence or absence of 1μM argatroban for 30 minutes at RT. Both samples were diluted 4-fold and made to 20 nM each of 48 of the 49mers (roughly 1 equivalent total to the CP 20-mer) and 1 μM 5' end 49-mer complement oligo in PBS. The samples were incubated at RT for 30 minutes. The samples were then pooled and a portion reserved for pre-selection abundance determination by DNA sequencing. The pool was then made to 2 μM of thrombin antibody and purified with 10 μg of protein G magnetic beads as with the qPCR assays. Starting material and elution DNAs were PCR amplified with 10 cycles of PCR using TS-Index18-F (pre) or TS-Index19-F (post) as forward primers and TS-Uni-9mer-R (Table S4) as the reverse primers. PCR products were then sequenced and aligned to barcodes as with the biotinylation experiment. Barcode read numbers are shown in Table S6.

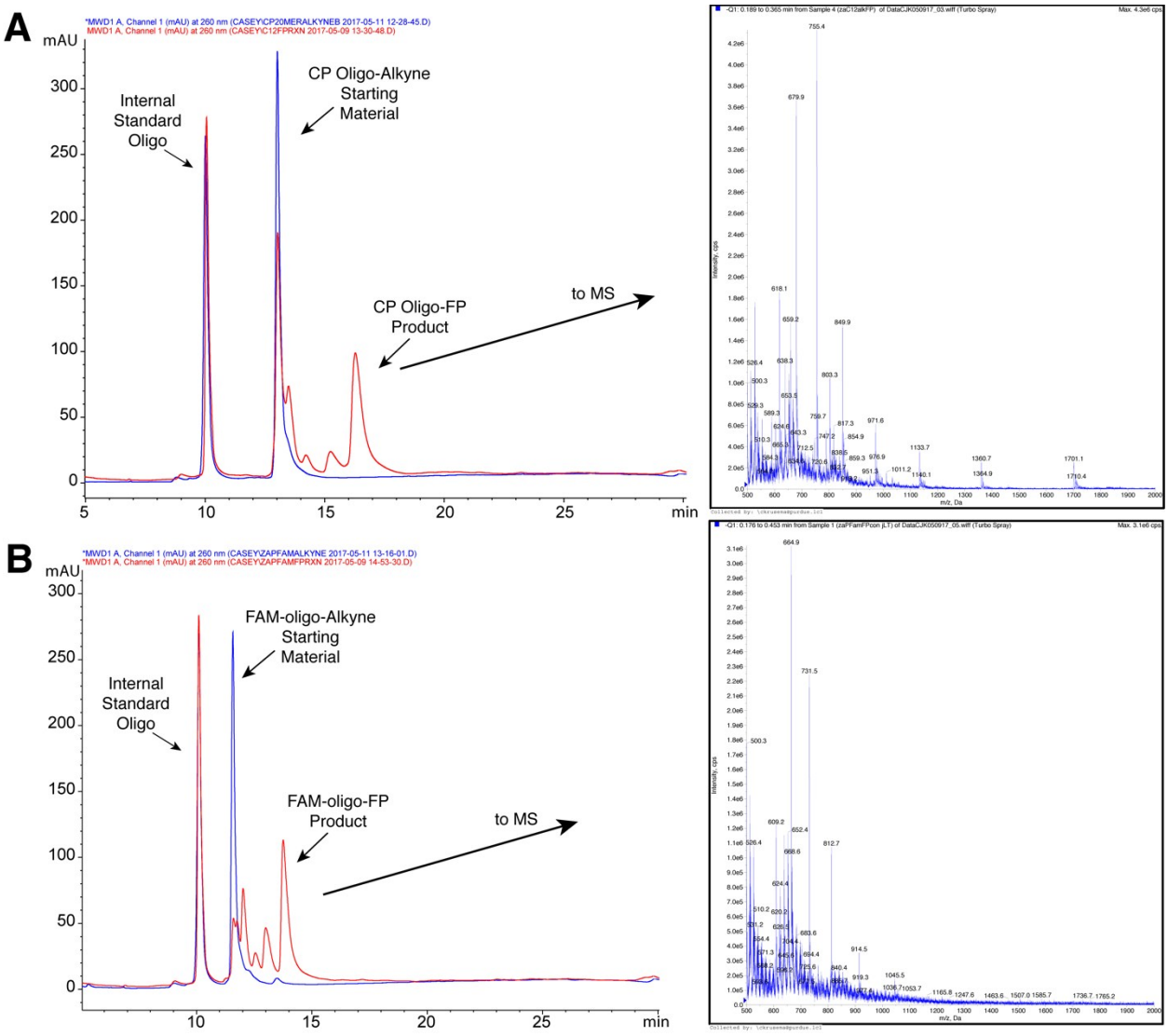

Figure S1. HPLC reaction monitoring of fluorophosphonate-N_3 conjugation to alkyne oligonucleotides and mass spectra of product peaks. Crude reaction chromatograms (in red) are overlaid with starting material chromatograms (in blue). **A**) Click reaction with CP oligonucleotide with 5' alkyne modification. **B**) Click reaction with a 5' FAM, 3' alkyne 20-mer oligonucleotide.

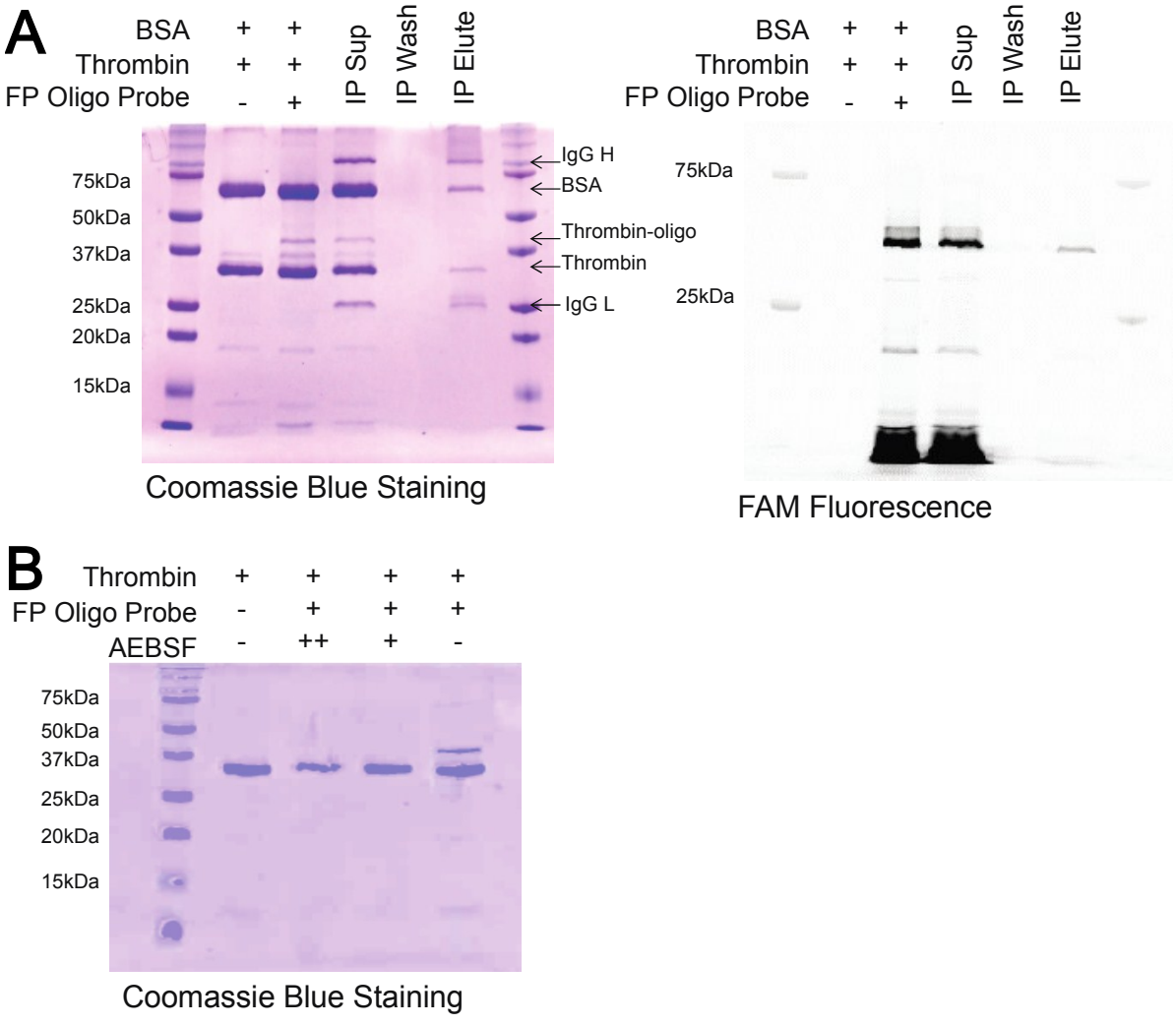

Figure S2. Thrombin labeling with FP oligo conjugate. **A**) Selective labeling of thrombin and subsequent IP. **B**) Inhibition of thrombin labeling with FP probe by AEBSF. ++ denotes 100 μM and + denotes 20 μM.

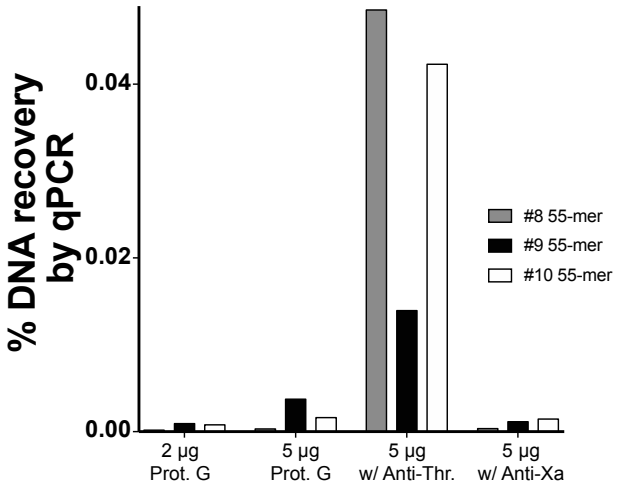

Figure S3. Nonspecific DNA binding to protein G beads or thrombin or Xa antibodies. Three 55-mers (C8, C9, and C10) were used for nonspecific binding test to antibodies. Each 55mer was incubated with protein G magnetic beads only (2 µg of protein G: 2G, or 5 µg of protein G: 5G). Antibodies to thrombin or Factor Xa (3 µg each) were incubated with protein G magnetic beads (5 µg of protein G (5GThr or 5GXa)), and then incubated with 1 µM of each 55-mer. Beads were washed with 1x PBST-BT for a total of three times, and then washed three times with 1x PBST-BT with 250 mM NaCl (high salt washing buffer). After washing, 55-mers were eluted by heating to 95 °C for 3 minutes.

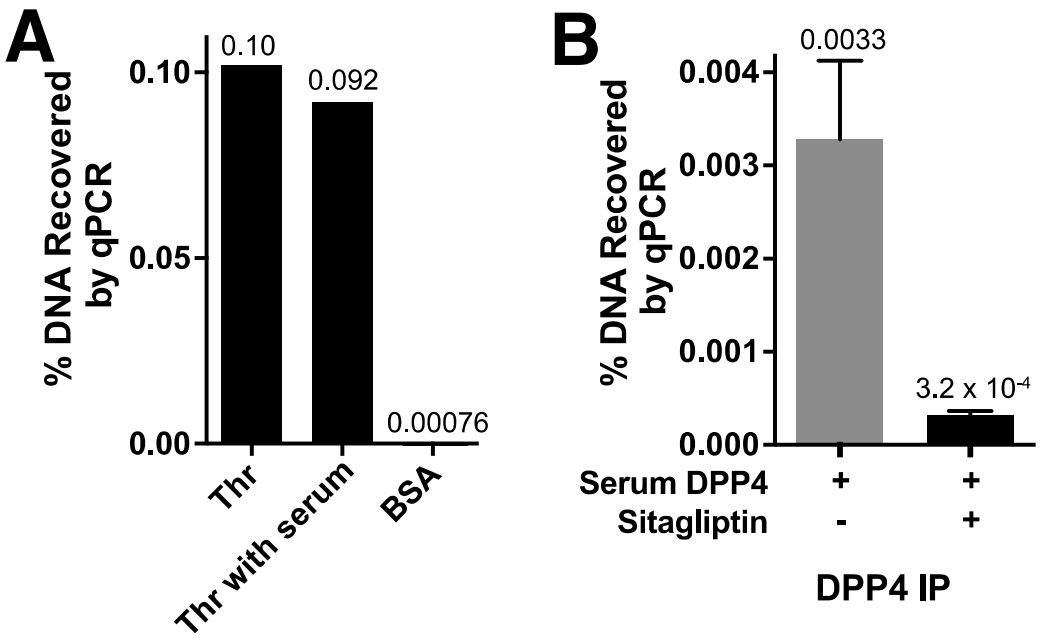

Figure S4. **A**) Detection of rThrombin activity using the FP oligo probe in buffer (1x PBS) or human serum (30 mg/ml protein). **B**) Detection of DDP-4 activity from human serum in the absence or presence of inhibitor sitagliptin. Error bars represent one standard deviation in the percent recovery for four unique DNA amplicons per sample.

Supplemental Table 1. 55-mer sequences. The sequence in blue indicates a 15-base intervening sequence. The sequence indicates the 20-base region complementary to the CP 20-mer FP probe. 20-mers of the same sequence as the encoding region were used as qPCR primers. 20-mers of the complementary sequence to the encoding region were used to duplex the 5' portion of the 55-mers prior to selection.

General 55-mer Structure	XXXXXXXXXXXXXXXXXXXXTATGCAGTAGAGAGCTGTGGCAAGCTTGATACCAT
Oligo Name	**Sample Encoding 20-base Sequences**
C8	CTGATGCGACTCCTAGACTT
C9	ATCAGTAGTCCAGTGCAGTG
C10	GTTCCCATTCCATGTGATCG
C11	TACCTACAGTTCCTGAAGCG
C12	GGGAAGAAATTAACCGTCCG
C20	AGAGACGAGGCATTCCATAG
C21	CTCCGTTGAATCCCATTGAG
C22	CACCTATCAGGAGACGTGAG
C23	AATAGCAGATACGGCCTGAG
C24	ATGTACCAGCCAGATACCAG
C32	CCCATTGTTTCAAGCATTGC
C33	ACCATAGCTGAATACCAGGC
C34	GCATACATCAGCACTAAGGC
C35	ACATAGGTAACAGGAGTCGC
C36	CCTCAGGAAGGAGAAGTCAC
C44	TAGATACCACCAGAGGGTCC
C45	TGTCCACACAGGTAGAACAC
C46	GGGAGACACATTTCACCAAC
C47	GCCCTCCGACAGAATAAGTA
C48	GGCCAGGAACTAACGATCTA
C58	AATTTCTAGTCACCGGGCTT
C59	AGACCGGGTTCATGGTTATT
C60	TAGACCTTCCTAAGTTGCCG
C68	CCGTTGGAATAAGCAGGAAC

Supplemental Table 2. qPCR data for individual 55-mer constructs from Figure 3.

	3a. Thrombin Titration		3b. Factor Xa Titration		3c. Thrombin Inhibition		3d. Factor Xa Inhibition	
55-mer	Units	% rec.	Units	% rec.	log μM	% rec.	log μM	% rec.
C8	0	0.0037	0	0.0019	0	0.31	0	0.079
C9	0	0.0039	0	0.0018	0	0.29	0	0.081
C10	0	0.012	0	0.00068	0	0.39	0	0.091
C11	0.068	0.046	0.019	0.0038	1	0.26	1	0.066
C12	0.068	0.035	0.019	0.0037	1	0.21	1	0.067
C20	0.068	0.067	0.019	0.0026	1	0.31	1	0.068
C21	0.14	0.048	0.038	0.0091	2	0.11	2	0.045
C22	0.14	0.074	0.038	0.011	2	0.12	2	0.046
C23	0.14	0.086	0.038	0.0073	2	0.2	2	0.051
C24	0.27	0.19	0.075	0.014	3	0.054	3	0.032
C32	0.27	0.14	0.075	0.011	3	0.064	3	0.034
C33	0.27	0.12	0.075	0.015	3	0.039	3	0.044
C34	0.54	0.13	0.15	0.037	4	0.018	4	0.034
C35	0.54	0.18	0.15	0.036	4	0.02	4	0.021
C36	0.54	0.19	0.15	0.02	4	0.028	4	0.028
C44	1.08	0.32	0.3	0.049	5	0.028	5	0.034
C45	1.08	0.23	0.3	0.036	5	0.031	5	0.029
C46	1.08	0.3	0.3	0.046	5	0.031	5	0.039
C47	2.16	1.08	0.6	0.13				
C48	2.16	0.89	0.6	0.16				
C58	2.16	1.01	0.6	0.11				

Supplemental Table 3. 49-mer sequences. 9-base barcodes[2] were placed between the CP 20-mer complementary sequence (red) and a portion of an Illumina adaptor sequence (blue).

General 49-mer Structure	CTGAACCGCTCTTCCGATCTXXXXXXXXXTGTGGCAAGCTTGATACCAT				
Oligo Number	**Encoding 9-mer**	**Oligo Number**	**Encoding 9-mer**	**Oligo Number**	**Encoding 9-mer**
1	CAACATATT	33	CAAGCAAGC	65	ACAGTTCCG
2	ATAGCACGG	34	ATGTAATGG	66	AGTCTGTAC
3	AACTCATTG	35	CAATAGGGT	67	CAGTGCCCG
4	CAATACAAG	36	ACACGAGAT	68	AACCGCCTC
5	ATATTGTAA	37	AACATCCAT	69	CACGCCACT
6	CAATTGATC	38	CACAGCAGG	70	ATGTCGCAC
7	CACAAAGGC	39	AACCCTCGG	71	ATTAAGTGC
8	AATTATAAC	40	CAGTTTCAC	72	ACGAACCAG
9	AACGAGAAG	41	AACGGTGGC	73	ATTAGGCAT
10	ATTATTACC	42	CACTGGGCC	74	ACATGTCAC
11	CAAACCGCC	43	ACACACAAA	75	AGCCCAATC
12	ATGCTTGAT	44	AACTTCTCC	76	CACGATCCG
13	AAGGGTCAG	45	ATTACCAAG	77	ACCAAGGAC
14	ATATGCGTT	46	AAGAGGGTT	78	CACTTATGT
15	AGTATCTGT	47	CAGCCGCTG	79	AAGACATGC
16	ACCAAAATG	48	AAGCCTTCT	80	CACCGTTAC
17	CAACGATCT	49	ATCATTCTG	81	ATGAACGTC
18	ATTCATCGT	50	AAGGACTAT	82	AGGCGTTGG
19	AAGGCGATC	51	CACCGGCGT	83	ACATAAGCG
20	CAGGCATTT	52	AGCCATCAC	84	AGAGGGTAT
21	AACTGTATT	53	AGTAAAGCC	85	ACTTTAACT
22	CACAAGTAT	54	CACCTACCC	86	TAGGGATAC
23	ATTTAGATG	55	AGGGCTGGT	87	AGCGACACC
24	CAGCACTTC	56	ATCAGAGCT	88	ACGACCACC
25	ACAATATGG	57	AGCACCGGC	89	ATCGAGTCT
26	CAAGTTTCC	58	AGAGAAGTT	90	AGACTCTCG
27	AAGCTACTT	59	ATCGGTACG	91	AGGTTCGCT
28	ACACCTTAG	60	ACGCTGTTG	92	GAGGTAGCC
29	AAGGCCGCA	61	ATGAGTTAC	93	AGCTATGTG
30	ATACTTAGG	62	ACACGCCGG	94	AGGAGGCGC
31	CAGGGAGGG	63	CAGCCTGGC	95	CACGTCGTC
32	AAGCTCAAC	64	AAGCAGGGG	96	AGCCGCAAG

Supplemental Table 4. Sequencing adaptor and indexing primers for Illumina DNA sequencing.

Oligo name	Sequences
TS-Uni-9mer-R	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
TS-Index18-F	AATGATACGGCGACCACCGAGATCTACACGCGGACGTTCAGATGGTATCAAGCTTGCCACA
TS-Index19-F	AATGATACGGCGACCACCGAGATCTACACTTTCACGTTCAGATGGTATCAAGCTTGCCACA

Supplemental Table 5. DNA sequencing read numbers for Figure 4A.

Hyb'd to 10% biotinylated CP 20-mer			Hyb'd to control CP 20-mer		
Barcode	**Reads Pre**	**Read Post**	**Barcode**	**Reads Pre**	**Read Post**
1	354	27538	49	2573	0
2	707	32209	50	2385	4
3	616	32660	51	4321	8
4	667	46576	52	3952	2
5	294	24703	53	2868	2
6	425	29707	54	4247	7
7	376	21086	55	3499	1
8	277	22523	56	2112	1
9	398	25818	57	4497	4
10	318	20051	58	2536	4
11	11	427	59	5139	2
12	382	17436	60	3348	6
13	685	44641	61	5495	6
14	498	32262	62	3394	5
15	550	35399	63	5374	4
16	873	60544	64	3714	3
17	758	49318	65	4955	1
18	811	41737	66	5449	2
19	515	28942	67	4600	4
20	828	53965	68	2487	1
21	489	35706	69	3098	2
22	402	21188	70	6962	7
23	377	28526	71	4722	2
24	498	27843	72	3512	2
25	471	31930	73	3357	6
26	1178	64332	74	5518	2
27	1440	79210	75	5236	6
28	1047	47631	76	1881	2
29	1244	49570	77	358	0
30	879	63832	78	2946	6
31	1093	73002	79	1899	5
32	1216	64030	80	5663	14
33	246	5776	81	6329	2
34	227	12902	82	6210	5
35	862	66523	83	4399	2
36	670	41104	84	4516	6
37	1558	96837	85	597	1
38	1256	62110	86	4144	3
39	1748	88884	87	3942	6
40	1440	65962	88	2990	4
41	1656	59773	89	4227	7
42	2219	54048	90	1236	2
43	1236	82123	91	4135	5
44	1157	61638	92	5802	3
45	1326	86463	93	4392	4
46	1041	74196	94	2583	4
47	1053	24751	95	2116	3
48	2234	127787	96	4702	6

Supplemental Table 6. DNA sequencing read numbers for Figure 4B.

Control			+Argatroban		
Barcode	**Reads Pre**	**Read Post**	**Barcode**	**Reads Pre**	**Read Post**
1	2467	34150	49	11123	3337
2	5333	42608	50	10053	3033
3	4434	50246	51	17337	8579
4	4825	59692	52	11895	3267
5	1998	26117	53	9287	2026
6	3075	34019	54	17987	5355
7	2677	29816	55	14700	3579
8	1956	26040	56	7903	1914
9	2818	28242	57	15638	4438
10	2204	27117	58	10136	3023
11	142	1401	59	19013	3912
12	2834	22573	60	14931	5785
13	5120	61200	61	22020	6051
14	3632	38535	62	12632	7275
15	3747	40553	63	17421	6821
16	7336	124221	64	15802	7536
17	5140	64730	65	21018	5323
18	5872	49633	66	22203	6256
19	3446	29318	67	15051	4735
20	5233	60906	68	10239	3127
21	3868	40479	69	7081	2430
22	2613	20245	70	26796	6557
23	3395	56069	71	17370	4988
24	3395	37037	72	15465	3915
25	4444	57671	73	13711	4116
26	7235	63226	74	21762	6382
27	9281	82039	75	18736	5667
28	7246	52786	76	6631	1908
29	7735	52611	77	1557	520
30	7492	87577	78	13604	4932
31	8787	148979	79	7295	2132
32	7785	58026	80	22544	6403
33	1428	11927	81	24737	5415
34	1993	19731	82	27229	7260
35	6849	84629	83	17794	6131
36	4801	48314	84	19367	6471
37	10973	111195	85	2584	892
38	9500	76152	86	16083	4875
39	13378	105081	87	12823	3731
40	8993	69635	88	10827	3522
41	10131	66168	89	15586	4213
42	13490	120149	90	5504	1752
43	7189	76103	91	16957	3926
44	7970	81391	92	18589	4566
45	10631	103402	93	18662	4943
46	7294	81427	94	9320	2115
47	6783	45145	95	8717	2624
48	15457	146974	96	16101	4478

References

1 R. R. Jetson and C. J. Krusemark, *Angew. Chem. Int. Ed. Engl.*, 2016, **55**, 9562–9566.

2 M. A. Quail, M. Smith, D. Jackson, S. Leonard, T. Skelly, H. P. Swerdlow, Y. Gu and P. Ellis, *BMC Genomics*, 2014, **15**, 110.