Pacific Northwest
NATIONAL LABORATORY

PNNL-37766

Nanopore Activity Assays for Detection of Biomarker Protease Activity

Design and Testing of Substrates for Both Nanopore Sequencing and PCR-Based Detection Methods

May 2025

David S Wunschel
Jennifer Mobberley
Kai-For Mo
Matt Turner
Natalie Winans
Kristin Victry

U.S. DEPARTMENT OF ENERGY

Prepared for the U.S. Department of Energy
under Contract DE-AC05-76RL01830

DISCLAIMER

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from
the Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062
www.osti.gov
ph: (865) 576-8401
fox: (865) 576-5728
email: reports@osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
or (703) 605-6000
email: info@ntis.gov
Online ordering: http://www.ntis.gov

PNNL-37766

Nanopore Activity Assays for Detection of Biomarker Protease Activity

Design and Testing of Substrates for Both Nanopore Sequencing and PCR-Based Detection Methods

May 2025

David S Wunschel
Jennifer Mobberley
Kai-For Mo
Matt Turner
Natalie Winans
Kristin Victry

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99354

Executive Summary

The work performed in this project has demonstrated the ability to construct proteolytic enzyme substrates that are PCR and sequencing-readable reporter molecules. Specifically, the goal was to detect those reporter molecules via PCR and Oxford Nanopore Technologies MinION sequencing methods following exposure to the biomarker protease thrombin. The assay development focused on binding the constructed peptide-oligonucleotide chimera to immobilized streptavidin. The action of thrombin on the peptide portion of the molecule released the oligonucleotide for detection. Detection of protease activity was demonstrated in a concentration-dependent manner using MALDI-MS, RT-PCR and DNA sequencing. Additional steps to remove background release of reporter molecules during the assay was used to improve the difference in detected oligonucleotide reporter following protease activity.

Additional steps in assay development will be to (1) test the assay in an appropriate matrix, (2) investigate detection using additional DNA sequencing platforms and (3) demonstrate multiplexed detection of multiple protease markers in a single reaction.

Acknowledgments

This research was supported by the Strategic Investment I3T program under the Laboratory Directed Research and Development (LDRD) Program at Pacific Northwest National Laboratory (PNNL). PNNL is a multi-program national laboratory operated for the U.S. Department of Energy (DOE) by Battelle Memorial Institute under Contract No. DE-AC05-76RL01830.

Acronyms and Abbreviations

CHCA = α-cyano-4-hydroxycinnamic acid

DIPEA = N,N-diisopropylethylamine

HATU = hexafluorophosphate

MALDI TOF-MS = Matrix assisted laser desorption /ionization time-of-flight mass spectrometry

NMP = N-methyl-2-pyrrolidone

ONT = Oxford Nanopore Technologies

PCR = Polymerase Chain Reaction

PD = Pull down (for removal of uncleaved substrate)

qPCR = quantitative PCR

RT-PCR = real time PCR

TFA = trifluoroacetic acid

TIPS = triisopropylsilane

HPLC = High performance liquid chromatography

Contents

Figures

Tables

1.0 Introduction

Enzymes drive many of the most important biological processes in living systems and proteases that cleave other proteins and peptides perform essential functions. Furthermore, proteases are recognized as clinically relevant markers for clotting disorders, presence of toxins and certain cancers. More importantly, some proteases are specific for the amino acid sequences they recognize. This aspect has been exploited to develop assays for protease activity using designed peptides that mimic natural protease substrates of specific proteases. The detection of peptide substrate cleavage has been demonstrated using a variety of fluorescence-based assays[1,2] and even mass spectrometry for toxins with proteolytic activity.[3]

Previous project work has demonstrated detection of N-glycosidase enzymatic cleavage of synthetic DNA designed with the enzyme recognition sequence motifs as a substrate. The substrate molecule could be read on a nanopore-based DNA sequencing platform to measure the number of cleaved and intact substrates. A portion of that project work also illustrated that a designed and constructed DNA-peptide chimera can be used as a sequencing template for nanopore-based sequencing platforms. This concept builds on a previous report from the Dekker lab using DNA-peptide conjugates in attempt to obtain amino acid sequences of peptides.[4,5] This previous work illustrates the potential for the DNA-peptide chimera to be used as a substrate for sequencing data collection and detection of the peptide portion of the DNA-peptide construct.

The logical extension of this work is measuring protease activity by using a peptide substrate for a protease attached to an oligonucleotide reporter. Cleavage of the substrate can be detected within DNA sequence read data. This approach will allow for multiple DNA-peptide chimeras recognized by distinct protease biomarkers to be analyzed in a single sample reaction and instrument run. This can be achieved by designing sequence “barcodes” into each DNA molecule that correspond to specific peptide substrates to be recognized by specific proteolytic biomarker targets.

The biomarker detection assay studied in this work is for the protease thrombin. Thrombin levels are normally tested for detecting clotting disorders important in hemophilia and thrombosis.[6] Both Factor Xa and γ-thrombin are essential markers in monitoring the clotting cascade, including those implicated in excessive clotting in COVID-19 disease progression.[7] Peptide substrates have previously been described for thrombin, which cleaves plasminogen to plasmin (as well as activation of Factors V, VIII, XI and XIII) and activated Protein C to cleave (i.e. inactivate) Factor Xa.[8,9] These optimal peptide substrates will serve as the detection module when linked to a synthesized strand of DNA for detection.

1.1 Project Aims

Aim 1: Demonstrate synthesis and cleavage of peptide substrates by clinically relevant clotting factor thrombin as biomarker protease using a synthesized peptide containing the recognition sequence as well as both azide and biotin modifications for further steps.

Aim 2. Construction and testing of the DNA-Peptide substrate chimera. The synthesized peptide substrate was synthesized with an alkyne modification on the C-terminus and a biotin functional group on the N-terminus. This was coupled to a DNA oligomer ordered from a commercial vendor with an azide chemical modification on the 5' terminus to leaving the 3' end free for ligation with the nanopore adapter ligation chemistry. Once "clicked" together as a chimeric molecule and purified, the construct was bound to a magnetic bead with bound streptavidin to immobilize the construct on the bead via the biotin-streptavidin interaction.

Aim 3. Measure peptide cleavage following biomarker protease exposure. The assembled DNA peptide chimera was exposed to thrombin previously shown to cleave the peptide substrate alone. Detection of released DNA molecule as a signature of activity of the biomarker protease. Amplification of the released oligonucleotide allows for detection by PCR (e.g. RT-PCR) or detection of PCR amplicons by ONT sequencing. Tabulating the number of detected oligonucleotide reads with and without cleavage in a protease concentration-dependent manner. General schematic is provided in Figure 1.

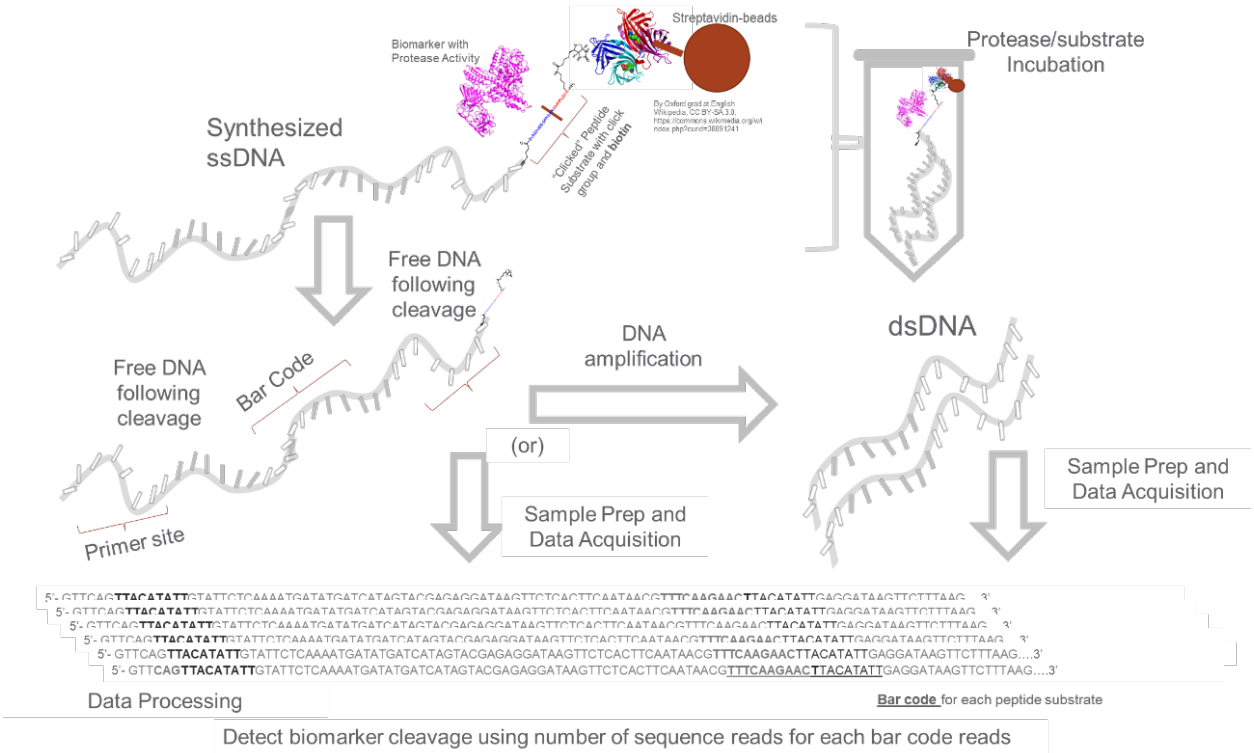


Figure 1. Diagram of assay for protease activity releasing designed DNA with bar code with sequencing alone or PCR amplification and sequencing.

Aim 4. Measure peptide cleavage events in ONT sequencing data

The detection of protease released DNA reporters through PCR amplicons ultimately will be performed using DNA sequencing. The ONT MinION sequencing platform will be used for detection of amplicons, and specifically “barcode” sections of the DNA reporter to perform read counts. Detection of DNA barcode in amplicon reads with the barcode following protease cleavage and relative counts in samples without were evaluated.

2.0 Methods

2.1 Selection and Synthesis of peptide substrates

There were 4 peptides selected for testing that contain the known cleavage site for thrombin. Two were native peptide sequence of fibrinogen alpha and beta chain and two were the modified peptides where an amino acid in the native sequence was replaced with a proline (Figure S1). It is expected that replacing the amino acid next to the cleavage site with proline will increase the reactivity of the substrate against thrombin.[10]

Synthesis of the target peptides was performed on an automatic solid-phase peptide synthesizer (Prelude® X) using standard Fmoc-based protocol with 1-[bis(dimethylamino)-methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate (HATU) and N,N-diisopropylethylamine (DIPEA) as the activator system. Each peptide was prepared using Rink Amide AM resin LL (Novabiochem, 70 mg, 0.021 mmol, loading: 0.30 mmol/g). The amino acid for each coupling step was 10 equivalents to the amount of resin. The synthesis was programmed to coupling time for 10 minutes with heating at 70 °C. The amounts of activator HATU and DIPEA were set to 10 and 20 equivalents, respectively, relative to the amount of the resin.

After the synthesis was completed by the peptide synthesizer, the resin was collected, and the fully assembled peptide was cleaved from the resin while simultaneously deprotecting the side chains using a cocktail solution (2 mL) of trifluoroacetic acid (TFA, 1.75 mL), water (0.1 mL), thioanisole (0.1 mL) and triisopropylsilane (TIPS, 0.05 mL) for 2 hours at room temperature. The resin was filtered off and the filtrate was added dropwise to cold (0 °C) diethyl ether (40 mL) to precipitate out the peptide. The crude peptide was collected by centrifugation (20 minutes at 3000 rpm) and the supernatant was decanted. The crude product was dissolved in 50% acetonitrile in water with 0.1% TFA and was lyophilized to give a white powered. The solid was reconstituted in 50% acetonitrile in water (1 mL) with 0.1% TFA for HPLC purification. The crude peptide solution (0.05 mL) was injected into a HPLC system (Agilent 1260 Infinity II) equipped with an autosampler, a fraction collector, and a UV detector and was purified on a semi-preparative C18 column (Eclipse XDB-C18, 9.4 × 250 mm, 5 µm particle size) using a linear gradient from 100% water (HPLC grade with 0.1% TFA) to 100% acetonitrile (HPLC grade with 0.1% TFA) over 20 min with a constant flow rate at 3 mL/min. The eluent was examined at a wavelength of 215 nm. The fraction that gave the right molecular weight of the target peptide was collected and the solution was lyophilized to give the pure peptide as an amorphous white solid. Mass spectra of peptides were recorded on a Bruker MALDI-TOF mass spectrometery using positive ion mode. The MALDI matrix was α-cyano-4-hydroxycinnamic acid (CHCA) with concentration at 5 mg/mL in 50% acetonitrile in water with 0.1% TFA.

2.2 Thrombin activity assay verification for the peptide substrate

Thrombin was chosen as the biomarker protease to design the assay around and the 37 kDa recombinant form of human thrombin was obtained from SignalChem Biotech as a 1 mg/mL solution (27 µM). Working solutions were aliquoted and reaction concentrations from 25 to 250 nM. Briefly, the thrombin assay reaction involved a 1.5 mL microcentrifuge tube with 18 µL

reaction buffer[1] (1 or 2), 1 µL peptide substrate (1-4, 1 and 5 nmole) and 1 µL thrombin enzyme diluted to the desired final concentration. The resulting reaction mixture was incubated at 37 °C with gentle mixing (400 rpm) for 4 hours. After incubation, 2 µL of the reaction mixture was mixed with 8 µL of MALDI matrix (CHCA) and this sample was analyzed by MALDI-TOF MS in positive ion mode.

2.3 Synthesis of the DNA-peptide substrate chimera through click chemistry

2.3.1 Synthesis of the peptide for the "click" reaction

The first amino acid, lysine functionalized with a biotin moiety, was coupled manually to the Rink Amide AM resin LL (Novabiochem, 70 mg, 0.021 mmol, loading: 0.30 mmol/g). The resin was first swollen in N-methyl-2-pyrrolidone (NMP, 2 mL) for 20 minutes. The solvent was removed by filtration and the resin was treated with 20% piperidine in dimethylformamide (DMF, 2 mL) for 5 minutes twice to remove the Fmoc protecting group. The resin was then washed thoroughly with NMP (4 x 2 mL) and was ready for the coupling with the lysine-biotin. To a solution of Fmoc-Lys(biotin)-OH (125 mg, 0.21 mmol) in NMP (2 mL) was added HATU (79.8 mg, 0.21 mmol) and DIPEA (0.073 mL, 0.42 mmol). After premixing for 5 minutes, the pre-activated lysine solution was added to the pretreated Rink Amide AM resin and the resulting mixture was agitated with a stream of nitrogen gas for 6 hours. The solution was filtered off and the resin was washed successively with NMP (4 x 2 mL) and methanol (4 x 2 mL). The resin was dried under vacuum and submitted to the automatic peptide synthesizer to finish the remaining amino acid sequences and the last alkyne functional group.

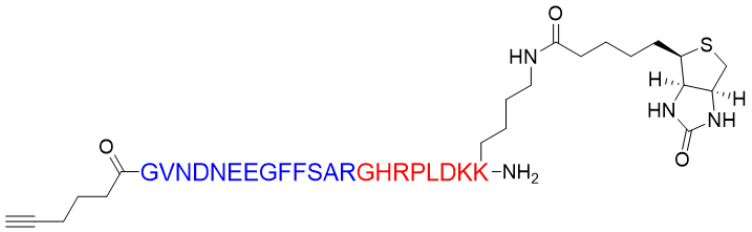


Figure 2. The structure of peptide 4 with a biotin at the N-terminus (for streptavidin magnetic bead capture) and an alkyne group at C-terminus (for "click" reaction with oligo SN-1 azide).

A solution of oligo SN1-azide (Integrated DNA Technologies, 20 µM, 300 µL, 6 nmol) and a solution of biotin-peptide-alkyne (1 mM, 60 µL, 60 nmol) was mixed in a 1.5 mL microcentrifuge tube and the resulting solution was lyophilized to give an amorphous white solid. The resulting solid mixture was reconstituted into tert-butanol/water (100uL, 1/1, v/v) and was added copper(II) sulfate solution (30 mM, 2 µL, 60 nmol). The resulting mixture was flushed with nitrogen gas and was added sodium ascorbate (48 mM, 5 µL, 240 nmol). The final reaction mixture was agitated at 700 rpm at 37°C for 24 hours. 1 µL of the reaction mixture was mixed with 4 µL of 3-

[1] Reaction buffer 1: 50 mM Tris, 10 mM CaCl2, 0.15 M NaCl, pH 7.6, Reaction buffer 2: 20 mM HEPES, 2 mM CaCl2, 0.14M NaCl, pH 7.6

hydroxypicolinic acid (3-HPA) matrix[2] and this sample was analyzed by MALDI-TOF MS in positive ion mode to confirm the completion of the reaction. The unreacted peptide and reaction solvent (tert-butanol) were removed by using MWCO filter (10,000) and the chimeric oligo-peptide conjugate solution was collected and lyophilized to give an amorphous white solid. The product was reconstituted into water (0.2 mL) and the final concentration was determined by using nanodrop.

2.4 Thrombin assay of chimeric oligo-peptide conjugate

Working solutions were aliquoted and reaction concentrations from 10 to 250 nM. The thrombin assay reaction involved a 1.5 mL microcentrifuge tube with 18 µL reaction buffer (20 mM HEPES, 2 mM CaCl2, 0.14M NaCl, pH 7.6), 1 µL chimeric oligo-peptide substrate (20 pmol) and 1 µL thrombin enzyme diluted to the desired final concentration. The resulting reaction mixture was incubated at 37°C with gentle mixing (400 rpm) for 4 hours. After incubation, 2 µL of the reaction mixture was mixed with 8 µL of 3-HPA matrix and this sample was analyzed by MALDI-TOF in positive ion mode.

2.5 Magnetic bead capture of chimeric oligo-peptide conjugate

Magnetic beads with streptavidin functional group were used to capture the chimeric oligo-peptide conjugates by binding to the biotin group present on the C-terminus of the target peptide. Dynabeads™ MyOne™ Streptavidin T1 and C1 (10 mg/mL) were evaluated for this purpose. Both beads were evaluated for performance, and ultimately the T1 beads were selected for ease of use. Two wash buffers were also evaluated, PBS, pH 7.4, and 10mM Tris-HCl, 1mM EDTA, 2M NaCl, pH 7.5. The ability of the wash buffers to remove oligo sequences that were unsuccessfully clicked onto the biotin containing peptide were evaluated. To evaluate the effectiveness of the wash buffers to remove unreacted oligo, 20 pmol of oligo-peptide conjugate or unreacted oligo were added to the beads and allowed to react at room temperature for 1 hour with gentle agitation. 10 µL of beads were used, and washed 3X with 200 µL of the respective wash solution prior to addition of 20 µL of 1 µM chimeric oligo-peptide conjugate or unreacted oligo solution. Following the incubation, the beads were placed on a magnetic separator and the supernatant was saved. The beads were then washed 3X with 200 µL of the respective wash solution, and each wash was saved. To evaluate the effectiveness of the wash solutions at disrupting nonspecific interactions between unreacted oligo and the beads, endpoint PCR and gel analysis was used to determine if oligo was still present in the initial supernatant and three washes (see Section 9.8.1 for endpoint PCR and gel details).

2.6 Thrombin assay of oligo-peptide conjugate on beads

Thrombin exposures were conducted with the oligo-peptide conjugate attached to beads. Prior to the exposure assay, 10 µL of beads were added, and washed 3X with 200 µL with 10mM Tris-HCl, 1mM EDTA, 2M NaCl, pH 7.5 washing buffer. 20 µL of 1 µM chimeric oligo-peptide conjugate

[2] Matrix solution: 250 mM 3-hydroxypicolinic acid, 40 mM ammonium citrate, and 5 mM ammonium tartrate in 50% acetonitrile in water.

was added to the beads and allowed to react for 1 hour at room temperature with gentle agitation. Following the incubation, the beads were placed on a magnetic separator and the supernatant was removed. The supernatant was measured by nanodrop to determine the extent of retention for the construct by the beads. The beads were then washed 3X with 200 µL of washing buffer. After the oligo-peptide conjugate was cleaned, the washing buffer was exchanged for reaction buffer for thrombin assay, with 3X rinses with 200 µL of reaction buffer 2. Reactions were conducted in 20 µL of reaction buffer with thrombin concentrations of 250, 100, and 25 nM, and a no-thrombin control. The reaction was allowed to proceed at 37°C on a Thermomixer at 400rpm agitation. After incubation, the supernatant was collected using a magnetic separator and placed in a new tube. After the exposure, a second bead capture was used to remove unreacted oligo-peptide conjugates that disassociated from the beads without digestion. For this, 5 µL of beads were used and wash 3X with reaction buffer. 10 µL of the reaction product was added to the beads and was incubated at room temperature for 1 hour. The supernatant was then retained, and both reaction products that had underwent secondary pulldown and those that did not were subjected to endpoint PCR and gel analysis, qPCR, and nanopore sequencing.

2.7 Nucleic acid detection with Quantitative Polymerase Chain Reaction (qPCR)

All samples were diluted according to the highest measured concentration (2.7 ng/µL) to 20 fg/ µL in PCR-grade Tris-EDTA (TE) pH 8.0, then 5 µL of that dilution was used in the qPCR reaction. If no measured concentration was available, all samples were diluted 1:10,000. In addition, the 20 µL total volume qPCR reaction contained 10 µL of the 2X commercial master mix, 1 µL of the 20X assay and 4 µL of molecular-grade water. The commercial master mix was Applied Biosystems TaqMan Fast Universal PCR Master Mix No AmpErase UNG. The assay was designed using the IDT Realtime PCR tool for the oligo1 nucleotide sequence. The components of the resulting IDT PrimeTime Std qPCR primer/probe mix were: forward primer (5'-GAGGTGAGCGACTATAGTTTCGAG-3'), reverse primer (5'-CCGTGCCATTCTGTGCGATA-3'), and probe (5'-/56-FAM/TCGTCTGGCAAATAGCCGATCCAG/3BHQ_1/-3'). An Applied Biosystems QuantStudio 7-Flex instrument was used with the cycling parameters of one cycle for 20 seconds at 95°C, followed by 40 cycles of 1 second at 95°C and 20 seconds at 60°C. The amplification data were analyzed with auto-baseline and auto-threshold instrument software settings.

2.8 Nucleic acid detection with the Oxford Nanopore MinION Sequencer

Following thrombin exposure, the Oxford Nanopore MinION sequencer was used to detect DNA reporter sequences both directly from the exposure assay and indirectly following endpoint PCR amplification of the DNA reporter.

2.8.1 Preparation of Oligo1-peptide for direct sequencing

Following thrombin exposure, the sample was cleaned and concentrated to a final volume of 15 µL using the Oligo Clean & Concentrator per the manufacturer's instructions (Zymo). For direct sequencing of the unamplified samples a concentration of 1000 fmol per sample was used with

the SQK-LSK114 with the Ligation Sequencing Kit (Oxford Nanopore Technologies). Briefly, the 1000 fmol of amplicon is added to a thin-walled PCR tube containing 1.75 uL of Ultra II end-prep reaction buffer and 1.5 uL of Ultra II end-prep reaction mix and brought to a final volume of 30 uL with molecular grade water, then incubated at 20°C for 5 min and 65°C for 5 min using thermocycler. Following incubation, the reaction was transferred to an Eppendorf tube and AMPure XP beads are added to achieve a 1.8:1 beads: reaction product ratio to optimize recovery of smaller fragments. Beads were washed twice with 80% ethanol and eluted with 30 uL of nuclease free water. For the adapter ligation and clean-up step, 12.5 µL Ligation buffer, 5 µL of Quick T4 ligase (New England Biolabs), and 2.5 µL of Ligation Adapter are added to the 30 µL end-prep library, and incubated at room temperature for 10 min. The ligation product is then cleaned using 45 µL of Ampure XP beads (0.9:1 beads:ligation reaction) and Short Fragment Buffer was used for all washes. The ligated substrate is eluted using 7 uL of Elution Buffer.

2.8.2 End-point polymerase chain reaction (PCR)

End-point PCR was used to amplify reporter sequences prior to analysis by nanopore sequencing and gel electrophoresis. A VWR XT96 thermocycler was used with the cycling parameters of one cycle for 60 seconds at 94°C, followed by 30 cycles of 30 seconds at 94°C, 30 seconds at 57°C, and 30 seconds at 72°C. Platinum Taq polymerase, PCR buffer, dNTP, and magnesium chloride were all from Invitrogen. Electrophoresis was conducted on a E-Gel Power Snap (Invitrogen) using E-Gel EX 2% or 4% agarose (Invitrogen), with E-Gel Ultra Low Range DNA Ladder from 10-300 bp (Invitrogen) for reference.

2.8.3 Preparation of amplicon for MinION sequencing

For endpoint PCR amplification, the amplicon was cleaned and concentrated using Oligo Clean & Concentrator per the manufacturer's instructions (Zymo). The concentration of the amplicon was determined using Nanodrop One Microvolume UV-Vis Spectrometer (Fisher Scientific). The amplicon (100 fmol) was prepared for sequencing as described for sequencing amplicons by the SQK-LSK114 Ligation Sequencing Kit (Oxford Nanopore Technologies). The procedure was done as described in section 2.8.1 except for the starting with 100 fmol of amplicon.

2.8.4 Sequencing Data Analysis

All data analyses were performed in R version 4.3.1. Sequencing data, in the form of FASTQ files, were parsed using the readFastq function from the package *microseq*. Each sequenced read was aligned to oligo-1, our synthetic nucleotide substrate, using the pairwiseAlignment function from the Bioconductor package *Biostrings*. This function performs a local alignment using the Smith-Waterman algorithm. Two metrics, aligned sequence length and read coverage, were derived from the alignment data and used to detect released oligonucleotide. The distributions of the aligned sequence lengths were compared among samples exposed to thrombin at different concentrations and control groups. Read coverage, defined as the proportion of reads that aligned to the query sequence at a given base location, was calculated using the coverage function from the *Biostrings* package. Based on analysis of the coverage data, we identified a region with consistent high coverage across datasets, corresponding to bases 52 through 60 in the oligonucleotide substrate. Counts of aligned reads containing this "barcode" region were used as a detection metric, allowing comparison of samples across datasets.

3.0 Results

3.1 Determination of optimal peptide to use for the thrombin activity detection assay

Of the four peptides and reaction buffer condition that were tested, peptide 4, which was the proline modified form of the fibrinogen beta chain in reaction buffer 2, was determined to be the optimal peptide substrate and reaction conditions based on the ratio of N-terminus product to intact substrate (Table 1). The intact peptide has a m/z 2399 and the cleaved form m/z 1467 (Figure 3).

Table 1. Thrombin assay of four synthetic peptides with four reaction conditions. The ratio of N-terminus product to intact substrate was determined based on the peak height of their corresponding peaks measured from the MALDI-TOF MS. The higher the ratio indicated the higher reactivity of the peptide against thrombin. Reaction buffer 1: 50 mM Tris, 10 mM CaCl2, 0.15 M NaCl, pH 7.6, Reaction buffer 2: 20 mM HEPES, 2 mM CaCl2, 0.14M NaCl, pH 7.6

Buffer	Amount of peptide (nmole) in the assay	Peptide	Ratio of N-terminus product to intact substrate
1	1	1	1.29
		2	1.28
		3	0.72
		4	8.21
	5	1	2.02
		2	0.93
		3	0.83
		4	6.66
2	1	1	3.36
		2	3.62
		3	3.26
		4	6.99
	5	1	3.80
		2	3.56
		3	2.36
		4	12.29

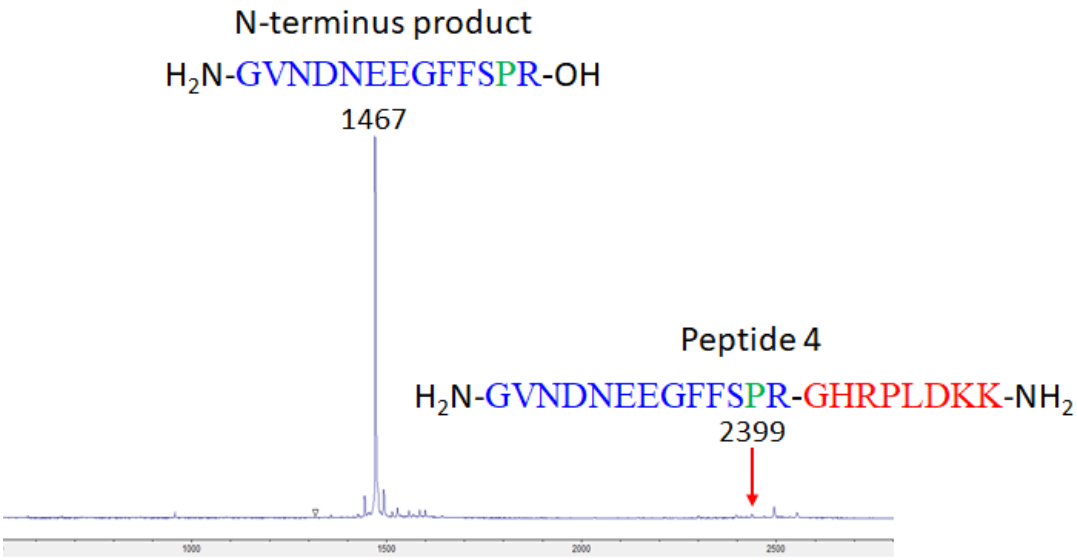
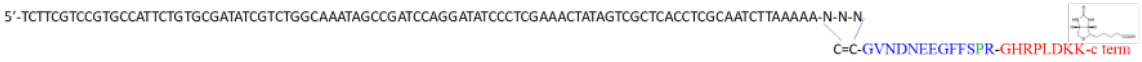


Figure 3. MALDI-TOF mass spectrometry assay of peptide 4 (5nmole) with 50nM thrombin enzyme in reaction buffer 2 (20 mM HEPES, 2 mM CaCl2, 0.14M NaCl, pH 7.6)

3.2 Click chemistry was successful in construction of the chimeric DNA-Peptide substrate for detecting thrombin activity

Click chemistry was used to construct the chimeric DNA-peptide substrate using peptide 4 and a synthetic single stranded DNA oligonucleotide. The oligonucleotide (oligo1) is a 97-nucleotide sequence that has no significant similarity to existing DNA sequences and was designed to minimize secondary structure. We successfully detected oligo1 (without the azide) on the MinION (Figure S2A). The assembled chimeric molecule is depicted below with the biotin molecule attached to the C-terminal end of the molecule (Figure 4A). The MALDI-TOF mass spectrum of the unpurified click reaction is shown in Figure 4B. The absence of a 29680-peak showed that all the oligo1-azide was reacted. After removal of unreacted peptide, the concentration of chimeric conjugate solution was 405 ng/uL (0.2 mL) based on the nanodrop measurement. The reaction yield of the click reaction was 42%. We were able to detect this construct with the MinION and it had a different coverage distribution than Oligo1 alone (Figure S2B).

A



B

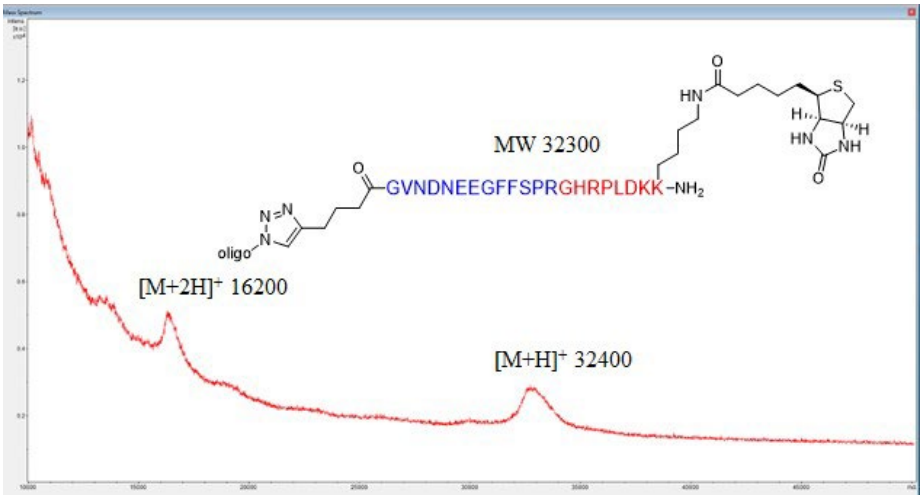


Figure 4. A. Complete oligonucleotide and peptide sequences of the DNA-peptide substrate. B. MALDI-TOF mass spectrum of the crude “click” reaction mixture.

3.3 Magnetic bead capture of chimeric oligo-peptide conjugate minimized unreacted substrates

To clean-up the click reaction mixture to ensure most of the substrate was the DNA-peptide and not the oligo alone we took the approach of using streptavidin beads to capture the biotin on the peptide side to enrich for the target substrate. It was determined that 10mM Tris-HCl, 1mM EDTA, 2M NaCl, pH 7.5 more effectively removes unreacted oligo from the Streptavidin beads than PBS, pH 7.4, and the Tris based washed was used for subsequent experiments (Figure 3). 65-70% of the oligo-peptide conjugate was retained by the beads following the initial pull-down (PD), while 25-30% of the unreacted oligo was retained as observed by nanodrop in the supernatant following pull-down (Data not shown).

The persistence of bands in all washes for both solutions indicate loss of the oligo-peptide conjugate throughout the washing process using PCR amplification (Figure 5). Non-specific binding of the oligo to the streptavidin beans may lead to spurious positive results following thrombin exposure assays, and this was addressed by an additional pulldown using the streptavidin beads to remove uncleaved peptide-oligo1 chimera washed off the beads in the initial digestion step. The additional post-digestion pulldown limited the amount of amplified target in thrombin negative samples such that it was not detected by gel analysis (Figure 6), and this step was incorporated into all subsequent experiments. The absence of bands in thrombin negative controls indicate that secondary pulldown removed intact substrate that washed off the beads during the exposure incubation without the cleavage by thrombin.

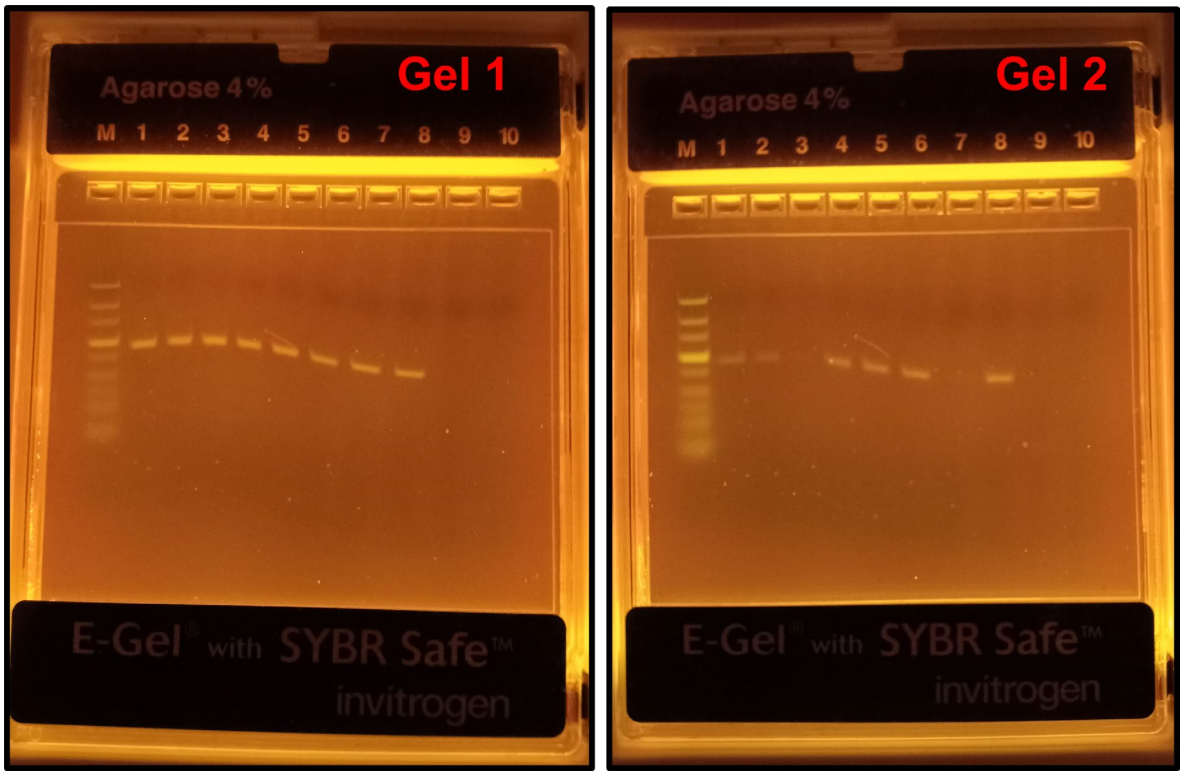


Figure 5. Gel analysis of end-point PCR products run on a 4% agarose E-Gel EX. The amplification product matches the anticipated size of 97 bp. Gel lanes are as follows: **Gel 1**: Lane 1 , peptide-oligo chimera pre-PD*; Lane 2, oligo alone pre-PD; Lane 3, chimera supernatant post-PD; Lane 4, oligo supernatant post-PD; Lane 5; chimer Tris-wash 1; Lane 6, chimera PBS-wash 1; Lane 7; oligo Tris-wash 1; Lane 8, oligo PBS-wash 1; Lane 9, NTC; Lane 10, blank. **Gel 2**: Lane 1, chimera Tris-wash 2; Lane 2, chimera PBS-wash 2; Lane 3, oligo Tris-wash 2; Lane 4, oligo PBS-wash 2; Lane 5; chimera Tris-wash 3; Lane 6, chimer PBS-wash 3; Lane 7; oligo Tris-wash 3; Lane 8, oligo PBS-wash 3; Lane 9, NTC; Lane 10, blank. *PD refers to bead pull-down. No bands present in Lanes 3 and 7 of Gel 2 indicate that 10mM Tris-HCl, 1mM EDTA, 2M NaCl, pH 7.5 more effectively removes unreacted oligo.

Figure 6. Gel analysis of end-point PCR products from additional streptavidin pull-down run on a 2% agarose E-Gel EX. The amplification product matches the anticipated size of 97 bp. Gel lanes are as follows: Lane 1, thrombin+ Rep1; Lane 2, thrombin+ Rep2; Lane 3, thrombin- Rep1; Lane 4, thrombin- Rep2; Lane 5; Blank; Lane 6, thrombin+ PD* Rep 1; Lane 7; thrombin+ PD Rep 2; Lane 8, thrombin- PD Rep 1; Lane 9, thrombin- PD Rep 2; Lane 10, blank. *PD refers to secondary bead pulldown following thrombin exposure. No bands present in Lanes 8 and 9 indicate that secondary pulldown step successfully removed intact substrate that washed off the beads during the exposure assay without thrombin cleavage of the peptide substrate.

3.4 Direct measurement of peptide cleavage of the chimeric substrate with MALDI-TOF MS

Cleavage of the chimeric oligo-peptide substrate by thrombin was confirmed using MALDI-TOF MS (Figure 7). The molecular weight of the intact peptide coupled to the oligonucleotide was detected as a ~32,400 *m/z* peak. Following incubation with low (e.g. 5 nM and 10 nM) concentrations of thrombin, the molecular weight of substrate remained unchanged indicating that no reaction occurred. A shift to a lower *m/z* peak at ~ 31,300 representing the oligonucleotide coupled to a cleaved form of the peptide (Figure 7A) following incubation of higher concentrations of thrombin at 25 nM, 50 nM, 100 nM and 250 nM (Figure 7B). The product peak is observed following exposure to 25 – 250 nM thrombin, while no product was detected at lower thrombin concentrations (5 and 10 nM).

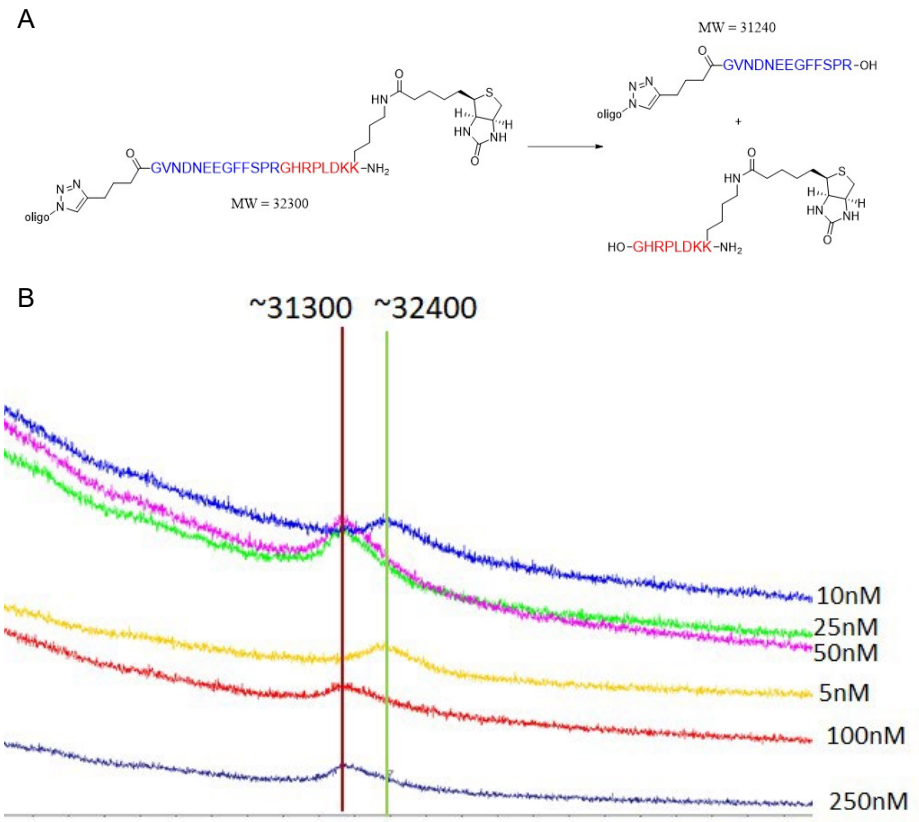


Figure 7. A. Reaction scheme of thrombin on the chimeric oligo-peptide substrate. B. MALDI assay showing substrate (MW 32400) and product (MW 31300) following exposure to 5 – 250 nM concentration of thrombin. No reaction occurred at low thrombin concentrations at 5 and 10 nM.

3.5 Measurement of peptide cleavage of the chimeric substrate with qPCR

The first qPCR experiment examined different substrate amplification between samples with and without thrombin exposure as well as with and without a post-reaction PD step using the immobilized streptavidin beads to remove uncleaved substrate. The goal was to reduce the amount of oligonucleotide with uncleaved peptide and reduce the amount of false positive detection. The increase in cycle count for the thrombin negative control from ~21 to ~ 32 is observed following the post-reaction PD step (Figure 8). Without the clean-up pulldown to remove unbound substrate, it was difficult to see a difference between samples with and without thrombin. Since the target of the qPCR assay is present in both bound and unbound substrate, removing the unbound substrate with the PD step demonstrates more clearly the activity of the thrombin using the qPCR assay.

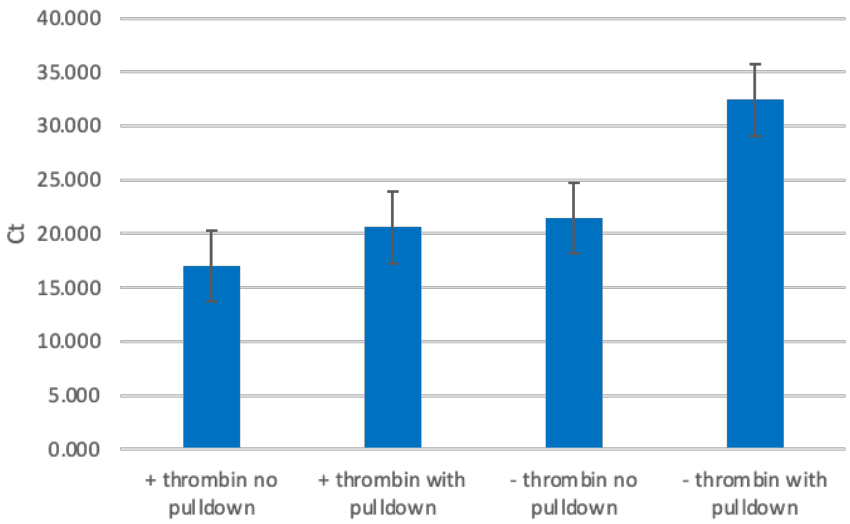


Figure 8. qPCR detection of oligo1-peptide substrate between thrombin exposed samples with no additional pulldown (no pulldown) and an additional pulldown (with pulldown).The average of each bar is an average of qPCR reactions in duplicate. The error bars are Excel-calculated standard error.

The next experiment explored the qPCR detection thrombin activity assay across different concentrations of thrombin with the pull-down method. The concentration-dependent release of detectable oligonucleotide can be seen as decreasing PCR cycle count (Ct) for 0 nM, 25 nM, 100 nM and 250 nM thrombin concentrations in Figure 9. The 0 nM reaction had an average Ct value of 35.5 vs. a Ct value of 24.5 for 250 nM thrombin reaction.

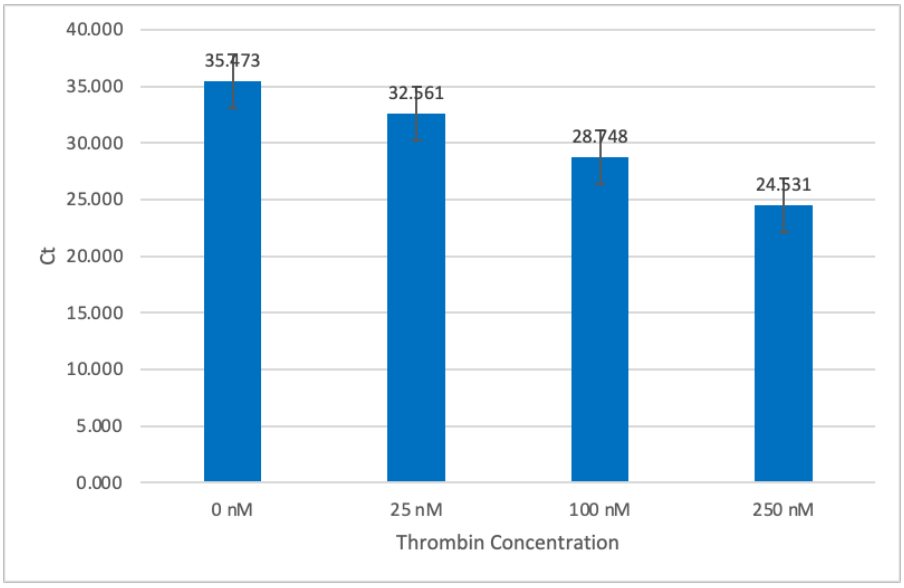


Figure 9. qPCR detection of oligo1-peptide substrate over a range of thrombin exposures. Each bar is an average of the three replicate PCR reactions at each concentration. and represent one standard error.

3.6 Measurement of peptide cleavage of the chimeric substrate with ONT sequencing

The initial experiment focused on determining if thrombin activity (250 nM) could be detected with direct sequencing of the oligo1-peptide or with the amplicon of the oligo1-peptide. The sequencing results indicate that cleaved and uncleaved forms of the oligo1-peptide and the amplicon are detected (Figure 10). While the sequence length distributions for oligo1-peptide show some differences between treatments these samples produced fewer reads than the amplicon sequences (data not shown). There was no difference between amplicons generated from cleaved and uncleaved oligo1-peptide. However, amplicon sequences had increased coverage between 20-30 nt and 50-62 nt that was not observed in the unamplified samples (Figure 10B). Due to time and resource constraints, additional experiments were performed using the amplicon approach. As mentioned previously, the presence of intact oligo-peptide chimera during washes and exposure may lead to an increased false positive detection of oligonucleotide without enzyme present. The removal of intact chimera was addressed by an additional pulldown following exposure and analyzed by nanopore sequencing (Figure 11). Increased sequence coverage between 20-30 nt and 50-62 nt consistently observed for the oligo1-peptide amplification product is absent in thrombin negative controls when subjected to secondary pulldown. This indicates the necessity of the secondary pulldown in reducing false positive detection in the thrombin negative control. Furthermore, there was no change in the coverage profile between thrombin positive sample prior to and after secondary pulldown, demonstrating that cleaved oligo1-peptide is not removed during this step.

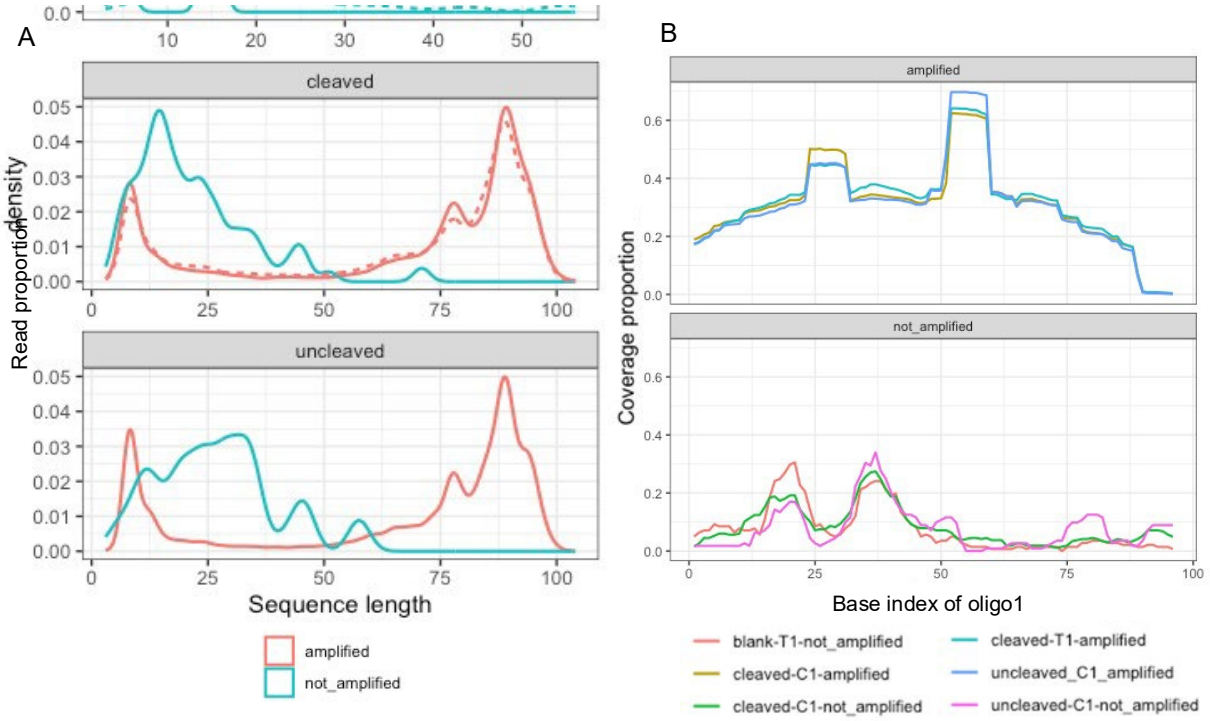


Figure 10. MinION read characteristics of oligo1-peptide substrate exposed to 250 nM of thrombin. No additional streptavidin pulldown was performed. A. Sequence length distributions. Read proportion corresponds to the relative number of reads at each sequence length. B. Proportion of reads that mapped to the oligo1 sequence.

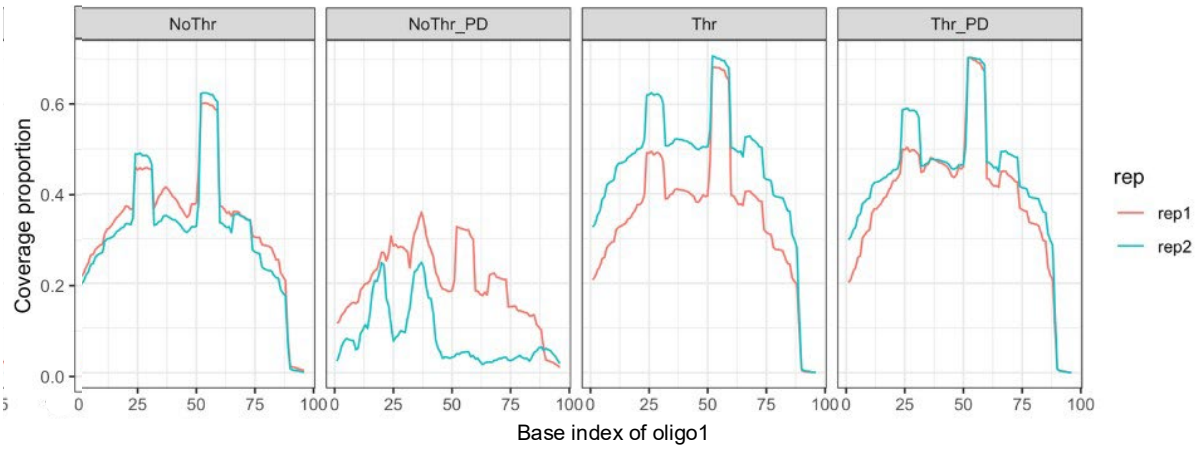


Figure 11. Coverage proportion of reads from MinION sequencing with the additional streptavidin pulldown step (PD) prior to end-point PCR of the oligo1-peptide substrate. Proportion of reads that mapped to the oligo1 sequence.

Following the establishment of the additional streptavidin pull-down step, the next experiment looked at MinION sequencing detection of the amplicon produced from samples exposed to a concentration range of thrombin. This was performed on the same thrombin exposure samples that underwent MALDI-TOF (section 3.4) and qPCR analysis (section 3.5). The sequencing results show that there were differences in the read characteristics between the samples not exposed to thrombin and those exposed to thrombin (Figure 12). The read coverage metrics indicate that the samples not exposed to thrombin had different regions of the oligo1 that had increased coverage compared to the samples that were exposed to thrombin (Figure 12A). The coverage proportion showed differences between treatments where fewer reads in the no thrombin samples mapped to the amplicon while those thrombin exposed samples showed a higher number of reads mapping in a more even matter (Figure 12B). The samples not exposed to thrombin (0 nM) also had fewer reads compared to those exposed to thrombin (Table S2). There was no distinct difference in sequencing characteristics for the samples exposed to different concentrations of thrombin. It should be noted the rep1 from 100 nM and rep1 from 250 nM had very low read counts such that they should be excluded from the analysis. Together these data indicate that with the current protocol, MinION sequencing of the amplicon can discriminate between samples not exposed or exposed to thrombin down to 25 nM but cannot discriminate between different concentrations of thrombin (Figure 12B).

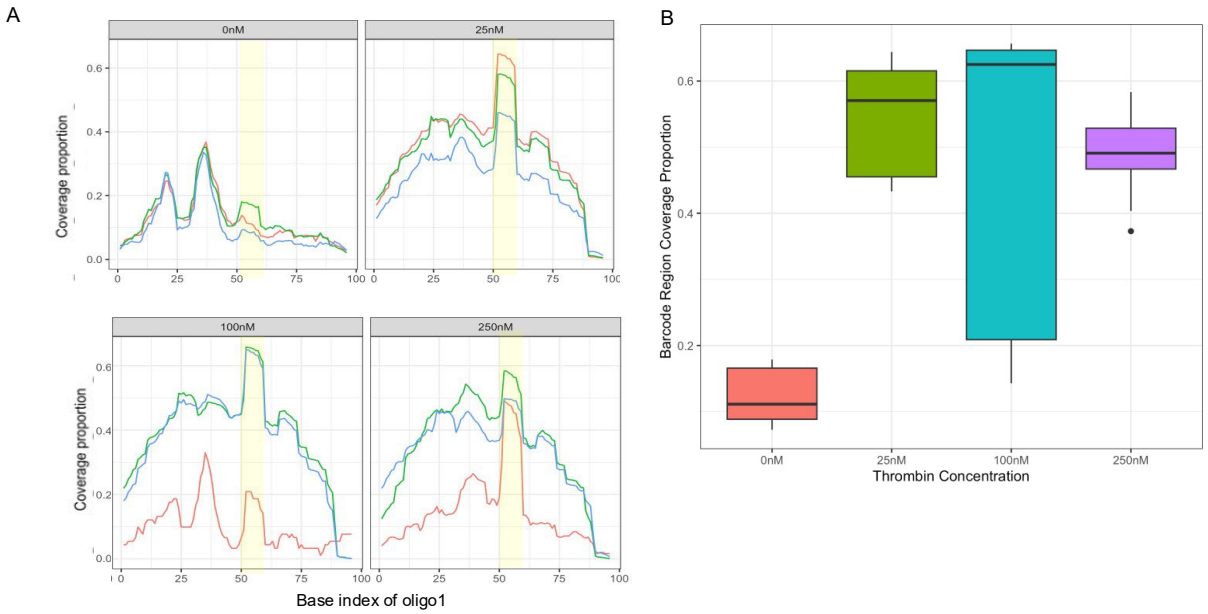


Figure 12. MinION sequencing of the end-point PCR amplicon from oligo1-peptide substrate exposed to a range of thrombin concentrations (0, 25 nM, 100 nM, 250 nM). A. Coverage proportion of the reads that mapped to the sequence of oligo1. The yellow highlight shows the region of interest for identifying a barcode. B. Boxplot showing the distribution of coverage proportion in the barcode region, identified as bases 52 to 60 of the query sequence (highlighted in yellow in Fig. 12A). This region reads with particularly high fidelity across samples and was selected as a basis for a reliable detection metric.

4.0 Discussion and Conclusions

Detection of biomarkers is a common method for diagnosis, prognosis, and prediction of future disease. Towards that end, this work sought to demonstrate the ability to construct proteolytic enzyme substrates linked to oligonucleotides that are PCR and sequencing readable. Initial efforts were to synthesis and cleave a peptide by a clinically relevant protease by identifying a substrate which thrombin has a high reactivity against. These included human fibrinogen alpha and beta chains. Once the activity of thrombin on the substrate was demonstrated and variations in sequence and reaction buffer were investigated, peptide substrate #4 was selected for coupling to the oligonucleotide into a chimeric molecule.

Coupling the peptide substrate with the oligonucleotide was achieved through a click chemistry reaction. The coupled chimeric molecule was purified from unreacted oligonucleotide and peptide using two different strategies. Molecular weight cutoff (MWCO) filtration was used to remove unreacted peptide from the product. Initially, HPLC was used to separate the chimeric oligonucleotide molecule from any remaining unreacted oligonucleotide. This was essential to prevent any false detection of the reporter oligonucleotide portion. Ultimately capture using immobilized streptavidin was used to remove unreacted oligonucleotide from the peptide-oligonucleotide chimera product. Removing the reliance on HPLC purification of the peptide-oligonucleotide chimera product greatly enhanced throughput.

Detection of thrombin cleavage of the peptide-oligonucleotide chimera was achieved on three different platforms: MALDI-TOF MS, qPCR, and ONT nanopore sequencing. Detection of cleavage of the peptide-oligonucleotide chimera while bound to streptavidin beads was demonstrated using both qPCR or PCR amplification followed by ONT nanopore sequencing. Both qPCR and DNA sequencing results illustrate a difference in released oligonucleotide with thrombin activity on the peptide substrate. While qPCR showed correlation between thrombin concentration and the determined Ct value, there was no distinct difference in sequencing characteristics for the samples exposed to varying thrombin concentration. This maybe ameliorated by optimization of sample preparation methods, such as the number of cycles used during endpoint PCR amplification of cleaved products, or by direct sequencing following exposure. The variability between the individual nanopore runs and the high standard deviations are partially due to the relatively high variability amongst the MinION flow cells themselves. Therefore, additional replicates of the thrombin concentration curve assay may result in significant differences in the results observed between different thrombin concentrations.

4.1 Next steps

The next steps in the process of developing this technology is to test the approach in an appropriate complex matrix, such as blood using both types of oligonucleotide detection methods, RT-PCR and DNA sequencing. A complementary step will be to demonstrate that additional sequencing platforms can be used for detection of the oligonucleotide reporter. Finally, demonstrating the multiplexed detection of targets in a sample with unique, designed oligonucleotide reporter markers.

It needs to be noted that a follow-on project to the Field Forward Toxin Diagnostics project is externally funded and anticipated to support continued work on this approach to protease detection.

5.0 References

1. James J. Schmidt, Robert G. Stafford. (2003) Fluorogenic Substrates for the Protease Activities of Botulinum Neurotoxins, Serotypes A, B, and F. **Appl. Env. Micro.** 69 (1): 297–303.DOI: 10.1128/AEM.69.1.297–303.2003

2. J. Oliver Mcintyre, Barbara Fingleton, K. Sam Wells, David W. Piston, Conor C. Lynch, Shiva Gautum, Lynn M. Matrisian. (2004) Development of a novel fluorogenic proteolytic beacon for in vivo detection and imaging of tumour-associated matrix metalloproteinase-7 activity. **Biochem. J**. 377:617–628.

3. Suzanne R. Kalb, Michael C. Goodnough, Carl J. Malizio, James L. Pirkle, John R. Barr. (2005) Detection of Botulinum Neurotoxin A in a Spiked Milk Sample with Subtype Identification through Toxin Proteomics. **Anal. Chem.** 77:6140-6146

4. Laura Restrepo-Pérez, Chirlmin Joo Cees Dekker. (2018) Paving the way to single-molecule protein sequencing. **Nature Nanotech** 786 (13):786–796

5. Henry Brinkerhoff, Albert S. W. Kang, Jingqian Liu, Aleksei Aksimentiev, Cees Dekker. (2021) Multiple rereads of single proteins at single–amino acid resolution using nanopores. **Science** 10.1126/science.abl4381.

6. C. M. Danforth, T. Orfeo, Stephen J. Everse2, Kenneth G. Mann, Kathleen E. Brummel-Ziedins. Defining the Boundaries of Normal Thrombin Generation: Investigations into Hemostasis. 2012 7(2): e30385

7. Ling-Zhi Hong, Zhang-Xuan Shou,De-Ming Zheng, Xue Jin. (2021) The most important biomarker associated with coagulation and inflammation among COVID-19 patients. **Mol Cell Biochem**. 476(7): 2877–2885

8. S. Krishnaswamy. (2005) Exosite-driven substrate specificity and function in coagulation. **J. Thromb. Haemost.** 3: 54–67.

9. S. Modrzycka, S. Kołt, S. G. I. Polderdijk, T. E. Adams, S. Potoczek, J. A. Huntington, P. Kasperkiewicz, M. Drag. (2022) Parallel imaging of coagulation pathway proteases activated protein C, thrombin, and factor Xa in human plasma. **Chem. Sci.** 13:6813-6829.

10. J.Y. Chang, S.S. Alkan, N. Hilschman, D.G. Braun. (1985) Thrombin specificity. Selective cleavage of antibody light chains at the joints of variable with joining regions and joining with constant regions. **Eur J Biochem**. 151:225-30.

Appendix A – Supplementary Figures and Tables

Table S2. [M+H]+ of thrombin target peptides and their corresponding products. Substrates 1 and 3 are the native peptide sequence of fibrinogen alpha and beta chain. Substrates 2 and 4 are the modified peptide sequence with one amino acid replaced by proline.

Origin peptide	Role	Peptide sequence	$[M+H]^+$
Fibrinogen alpha chain	Substrate 1	H_2N-GEGDFLAEGGGVR-GPRVVERK-NH_2	2185
	Products	H_2N-GEGDFLAEGGGVR-OH	1264
		H_2N-GPRVVERK-NH_2	940
	Substrate 2	H_2N-GEGDFLAEGGGPR-GPRVVERK-NH_2	2183
	Products	H_2N-GEGDFLAEGGGPR-OH	1262
		H_2N-GPRVVERK-NH_2	940
Fibrinogen beta chain	Substrate 3	H_2N-GVNDNEEGFFSAR-GHRPLDKK-NH_2	2373
	Products	H_2N-GVNDNEEGFFSAR-OH	1441
		H_2N-GHRPLDKK-NH_2	950
	Substrate 4	H_2N-GVNDNEEGFFSPR-GHRPLDKK-NH_2	2399
	Products	H_2N-GVNDNEEGFFSPR-OH	1467
		H_2N-GHRPLDKK-NH_2	950

Table S2. MinION sequencing statistics for the thrombin exposure assay

		Median Sequence Length		Read Count		
Conc	Rep	Fail	Pass	Fail	Pass	Total
0nM	rep1	19	5	14,641	5	14,646
0nM	rep2	16	9	12,813	9	12,822
0nM	rep3	17	6	10,839	6	10,845
25nM	rep1	11	8	72,916	739	73,655
25nM	rep2	11	8	21,244	667	21,911
25nM	rep3	13	9	46,824	1,036	47,860
100nM	rep1	16	5	1,126	5	1,131
100nM	rep2	10	8	99,892	1,916	101,808
100nM	rep3	10	6	25,399	81	25,480
250nM	rep1	13	20	3,223	20	3,243
250nM	rep2	11	6	112,616	20	112,636
250nM	rep3	11	8	122,041	3,070	125,111

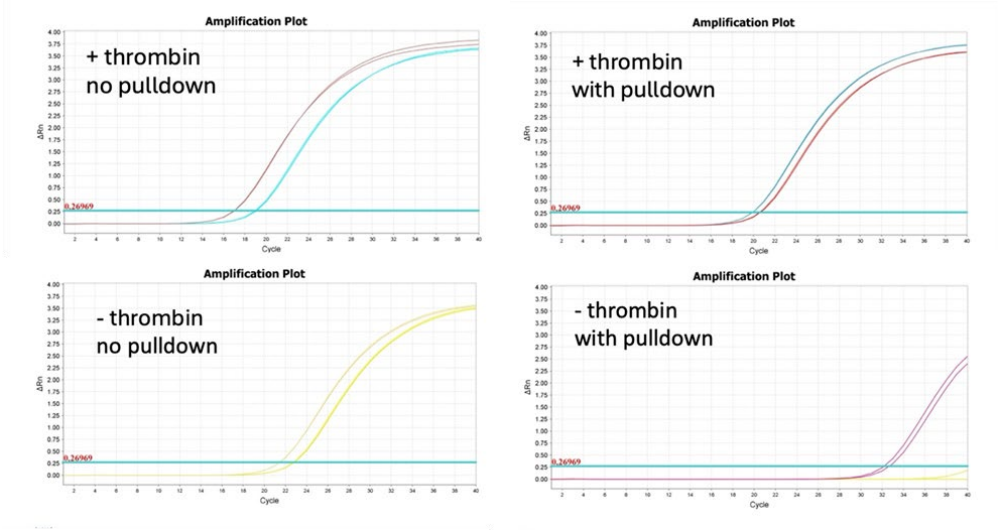


Figure S1. RT-PCR detection curves for the oligonucleotide following cleavage of the peptide substrates by thrombin. Curves following cleavage using thrombin without pulldown (top left) and with streptavidin bead pulldown cleanup (top right). Fluorescent detection curve of reactions without thrombin without pulldown cleanup (lower left) and with streptavidin bead cleanup pulldown (Lower right).

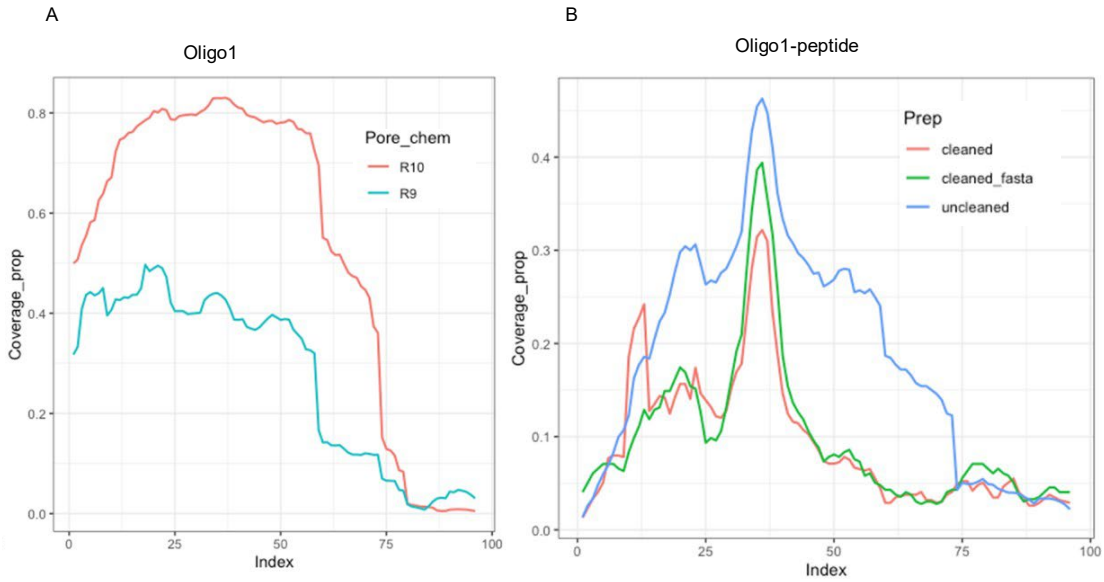


Figure S2. Oxford Nanopore MinION sequencing of substrates. A. Oligo 1 was sequenced on two different pore chemistries R9 and R10. B. Chimeric Oligo1-peptide substrate underwent additional clean steps following click chemistry but prior to sequencing.

Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354

1-888-375-PNNL (7665)

www.pnnl.gov