Selection and Binding Validation of Aptamers against Nucleocapsid Protein of SARS-COV-2 Using Capillary Electrophoresis

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Chemistry

Department of Chemistry and Biomolecular Sciences

University of Ottawa

© Yuxuan Gu, Ottawa, Canada 2023

Contents

Acknowledgement

While writing my master's thesis, I have received a great deal of invaluable assistance from numerous individuals whose comments and advice have contributed significantly to its completion. Foremost among these individuals is my supervisor, Professor Dr. Maxim Berezovski, who not only provided me with the opportunity to pursue my graduate studies in Canada but also offered me guidance on various aspects of academic research, such as literature comprehension, instrument operation, presentation skills, and scientific thinking. His teachings have aided me in my graduate studies and proved beneficial to me throughout my life. I would also like to express my sincere gratitude to my lab colleagues, both past and present, who have provided me with support and guidance. Specifically, I would like to thank Dr. Zoran Minic for his teachings on mass spectrometry, Emil for his valuable suggestions and instruction on independent work and self-learning, Kalem, my undergraduate student, for his extensive help and motivation during our experiments, Gurcharan for her assistance with experiments and her valuable experience, Suttinee for her help during difficult times, Abdullah for his instruction on the use of R language and computer issues, Rochelle for her patience and warm heart during CE instruction, and all my colleagues for their support and companionship throughout my research, courses, seminars, and activities. Finally, I am grateful for the unwavering support of my family, both in my studies and life. This graduate study has been an essential and unforgettable part of my life journey.

Abstract

The Coronavirus disease 2019 (COVID-19) pandemic has highlighted the critical need for accurate and sensitive diagnostic tools for detecting the SARS-CoV-2 virus. The nucleocapsid (N) protein is essential for virus replication and plays vital roles in virus assembly, packaging, and RNA transcription. This protein is a crucial component of the viral particle and is less prone to mutations than the other essential proteins in SARS-COV-2. All of these make the N protein a reliable target for virus detection. Aptamers, single-stranded oligonucleotides that can specifically bind to target molecules, have been proposed as a promising alternative to antibodies for detecting and treating viral infections. This study aimed to select DNA aptamers against the N protein of SARS-CoV-2 using capillary electrophoresis (CE) and validate the binding specificity of the aptamers.

After selecting seven clones, a preliminary binding validation was performed, and the two best binding clones were identified as ECK4 and ECK6. The structures of the aptamers were then modified by removing the primer regions from the original sequence, and the binding capacity of the truncated aptamers was confirmed. Dissociation constant (K_D) values were calculated to provide further supportive information for the quality of the two clones. Additionally, Biolayer interferometry (BLI) was used to calculate Apparent K_D as an alternative technique and provided consistent results with CE.

Our results demonstrate the successful selection of aptamers for the N protein of SARS-CoV-2 using CE-SELEX. Confirming the aptamers' binding capacity to N protein paves the way for developing aptamer-based diagnostics for COVID-19.

List of figures:

List of tables:

List of Abbreviations

COVID-19	Coronavirus Disease-2019
PBS	Phosphate Buffer Saline
SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2
MERS	Middle East Respiratory Syndrome
SELEX	Systematic Evolution of Ligands by Exponential Enrichment
CE	Capillary Electrophoresis
EOF	Electroosmotic Flow
N protein	Nucleocapsid protein
PCR	Polymerase Chain Reaction
KCE	Kinetic Capillary Electrophoresis
NECEEM	Nonequilibrium capillary electrophoresis of equilibrium mixtures
BLI	Biolayer Interferometry
SDS	Sodium Dodecyl-sulfate Polyacrylamide
K_D	Dissociation Constant
IPTG	Isopropylthio-β-galactoside
FAM	Fluorescein Amidite
LFA	Lateral Flow Assay

Chapter 1. Introduction

1.1.SARS-CoV-2

1.1.1 Background

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a novel coronavirus that causes COVID-19.[1] The virus was first identified in Wuhan, China, in December 2019 and has since spread globally, causing a pandemic that has significantly impacted public health and the global economy.[2] SARS-CoV-2 is a member of the family Coronaviridae, which also includes the viruses responsible for the severe acute respiratory syndrome (SARS) and Middle East respiratory syndrome (MERS).[2] Like these other coronaviruses, SARS-CoV-2 is an enveloped, positive-sense, single-stranded RNA virus with a characteristic crown-like appearance under electron microscopy.[3]

Since the beginning of the pandemic, the world has seen multiple waves of COVID-19 infections. In 2020, the first wave of the pandemic hit countries worldwide, and many nations implemented strict measures such as lockdowns to curb the spread of the virus. While some countries could control the spread of the virus, others experienced devastating consequences, with hospitals overwhelmed and large numbers of deaths. In 2021, the situation changed with the emergence of several virus variants.[4, 5] These variants, known as variants of concern (VOCs), have mutations that allow the virus to spread more easily or potentially cause more severe disease. The most notable VOCs include the Alpha, Beta, Gamma, and Delta variants.[6] The Delta variant has been responsible for a significant surge in COVID-19 cases in countries worldwide, including the United States and the United Kingdom.[7]

Efforts to control and mitigate the pandemic have included widespread vaccination campaigns, public health measures such as social distancing and mask-wearing, and the development of COVID-19 treatments.[8, 9] Multiple vaccines have been developed and authorized for emergency use in countries around the world,[10] and many countries have been able to vaccinate a significant proportion of their populations. However, the pandemic continues to be an effective global threat, and much must be done to control and mitigate its impact.[11] Ongoing efforts include the development of new and improved COVID-19 treatments, continued research into the virus and its variants, and ongoing public health measures to prevent the spread of the virus.[12, 13]

1.1.2. Nucleocapsid protein

The N protein (N protein) is one of the four structural proteins of SARS-CoV-2.[14-16] The N protein is essential for virus replication and involves several vital functions, including virus assembly[17], packaging, and viral RNA transcription.[18] The N is a highly conserved protein among coronaviruses and is immunogenic in animals and humans.[19]

The SARS-CoV-2 N protein is approximately 419 amino acids long and has a molecular weight of about 51 kDa.[15] The protein is highly basic, has an isoelectric point of about 9.8,

and has a positively charged C-terminal domain. The N protein is highly abundant within infected cells, making up about 90 % of the total viral protein content.[21] During the process of virus assembly, the N protein binds to the viral RNA and forms ribonucleoprotein complexes (RNP) [21, 22], which then interact with the membrane-bound spike (S) protein and envelope (E) protein to form viral particles (**Figure 1.1**).[23] The N protein has also been shown to regulate viral RNA synthesis and transcription, likely by interacting with the viral polymerase. Because of these special biological functions, it has been studied extensively as a potential target for vaccine development and therapeutic interventions.[20]

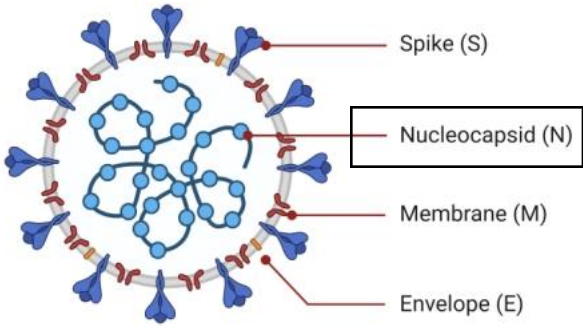


Figure 1.1: Schematic diagram of the SARS-COV-2 and related proteins.

Recent studies have also suggested that the N protein may modulate the host's immune response to the virus.[24-26] The N protein has been shown to bind to multiple host proteins, including the human nucleolar protein nucleophosmin (NPM1) and the host transcription factor Brd4. These interactions may play a role in evading the host immune response by altering the host gene expression profile or regulating the host immune response.[27] The N protein is highly immunogenic and is one of the main targets of the immune response to the virus (Potential antiviral effects of pantetheine against SARS-CoV-2).[28, 29] Antibodies against the N protein have been detected in both convalescent and vaccinated individuals [30, 31], and several studies have demonstrated the potential of the N protein as a target for COVID-19 diagnostics and serological assays.[32-36]

In conclusion, the SARS-CoV-2 N protein is a highly conserved, abundant, and immunogenic protein that plays a crucial role in virus replication and assembly. Its interaction with host proteins and potential role in modulating the host immune response makes it an attractive target for vaccine development and therapeutic interventions. Ongoing research into the N protein and its function in SARS-CoV-2 pathogenesis is crucial for understanding the virus and developing effective treatments and vaccines.

1.1.3. Recent diagnostic methods

Precise and readily available testing for SARS-CoV-2 is crucial for promptly recognizing infections, guiding isolation measures, halting transmission, and enabling timely treatment administration to diminish disease advancement. Several standard test methods will be mentioned in the following paragraphs.

RT-PCR. Reverse transcription-polymerase chain reaction (RT-PCR) is the gold standard for COVID-19 diagnosis. This method involves the detection of viral RNA using reverse transcription to convert RNA to complementary DNA (cDNA), followed by PCR amplification of the viral nucleic acid. RT-PCR has high sensitivity and specificity and can detect viral RNA as early as 1-2 days after symptom onset.[37] However, this method requires specialized laboratory equipment and trained personnel, and results may take several hours to days, which limits its use in mass screening.[38]

Antigen Test. Antigen tests are rapid diagnostic tests that detect viral proteins using immunochromatographic assays.[39, 40] These tests are easy to perform and provide results within 15-30 minutes.[40] Antigen tests have lower sensitivity than RT-PCR but high specificity, especially in individuals with a high viral load. Antigen tests are suitable for mass screening, especially in settings where PCR testing is not readily available.

Antibody Test. Antibody tests detect the presence of antibodies against SARS-CoV-2 in the blood. Antibodies are produced in response to viral infection, and their presence indicates a previous condition.[41] Antibody tests can detect past situations but are unsuitable for diagnosing acute diseases.[41, 42] Antibody tests are easy to perform and provide results within 10-15 minutes. However, the sensitivity and specificity of antibody tests may vary, and the accuracy of the results depends on the time of infection, the severity of the disease, and the type of antibody detected.[42]

Lateral Flow Assay. The lateral flow assay (LFA) is a low-cost, fast, and easy-to-use diagnostic technique that employs gold nanoparticles bound to antibodies for identifying viral antigens.[42] The test is easy to perform and provides results within 15-30 minutes. LFA can be used for mass screening in resource-limited settings where specialized laboratory equipment and trained personnel are unavailable.[43] However, the sensitivity and specificity of LFA may vary, and false negative results may occur, especially in individuals with a low viral load.[44-46]

Based on the current COVID-19 test methods development, developing better methods to detect viruses is crucial for several reasons. First, rapid and accurate identification of infected individuals allows for quick isolation, treatment, and contact tracing, limiting the virus's spread.[47] Second, detecting viruses accurately can improve disease surveillance and facilitate the development of effective vaccines and treatments. Third, identifying viruses quickly and accurately can help prevent pandemics by enabling the implementation of effective measures to control the spread of the virus. Considering all the disadvantages of current methods, people increasingly realize that aptamers could be candidates for better ways of development.[48, 49]

1.2.Aptamers and SELEX

1.2.1. Aptamers and their Applications for SARS-CoV-2

Aptamers have various applications in various fields, such as diagnostics, therapeutics, and biosensing. Due to their unique properties, aptamers have found multiple applications in biomedical research, including detecting and treating viral infections such as SARS-CoV-2.

In the context of SARS-CoV-2, aptamers have been developed as diagnostic tools for detecting the virus. For instance, some researchers have developed aptamers that can bind specifically to the spike protein of SARS-CoV-2.[50] These aptamers can be used in diagnostic tests, such as lateral flow assays or ELISA, to detect the presence of the virus in patient samples.[51] Aptamers have also been explored as potential therapeutics for SARS-CoV-2. For example, some researchers have developed aptamers that can bind to the spike protein and prevent it from binding to human cells, thereby blocking the virus's entry into host cells.[52, 53] Aptamers have neutralized SARS-CoV-2 in vitro. Researchers have identified aptamers that can bind to the virus and prevent it from infecting human cells. This approach could be developed into a therapeutic strategy for COVID-19.

Overall, the unique properties of aptamers, including their high specificity and affinity, make them promising tools for detecting and treating SARS-CoV-2 and other viral infections.

1.2.2. SELEX

Systematic Evolution of Ligands by Exponential Enrichment (SELEX) is a powerful method for identifying aptamers that can specifically bind to a target molecule.[54] SELEX involves iterative rounds of selection and amplification of a pool of random nucleic acid sequences exposed to the target molecule of interest. During each round, the sequences that bind to the target are selectively amplified, while those that do not bind are discarded. This process is repeated several times until a pool of highly specific and high-affinity aptamers is obtained (**Figure 1.2**).

SELEX can be performed using a variety of formats, including solution-phase SELEX, solid-phase SELEX, and cell-based SELEX. In solution-phase SELEX, the target molecule is free in solution. In contrast, in solid-phase SELEX, the target molecule is immobilized on a solid support, such as magnetic beads or a microarray. In cell-based SELEX, the target molecule is presented on the surface of living cells.

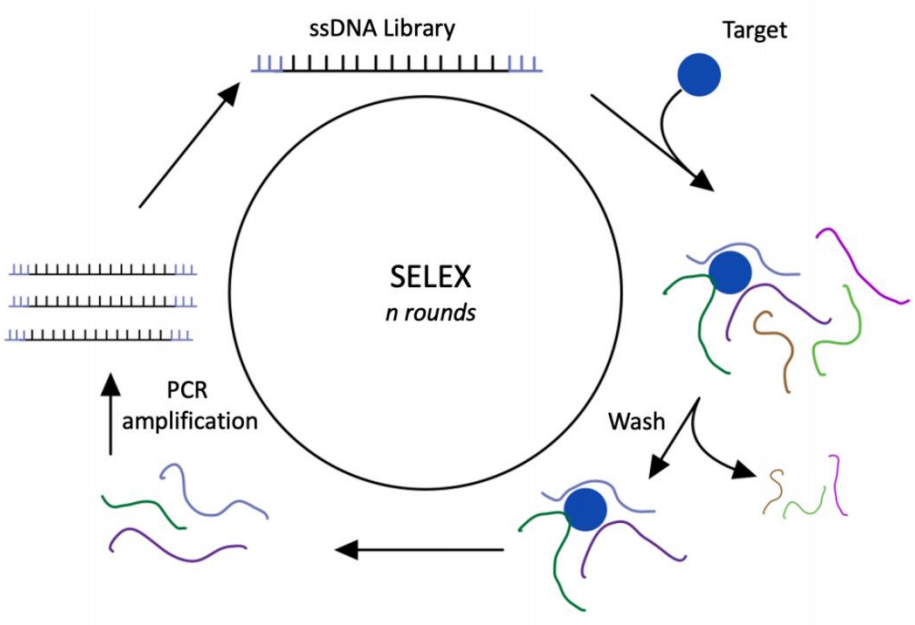


Figure 1.2: Scheme of typical SELEX process.

1.3. Capillary Electrophoresis

1.3.1. Background

Capillary electrophoresis (CE) is a powerful analytical technique that separates and analyzes a wide range of molecules. CE involves the migration of charged species under the influence of an electric field through a narrow capillary filled with an electrolyte solution(**Figure 1.3**).[55] The technique is based on the principle that charged species will migrate at different rates based on their size, shape, and charge, enabling the separation and identification of complex mixtures. CE offers many advantages over traditional electrophoresis techniques, such as higher separation efficiency, faster analysis times, and the ability to analyze small sample volumes. Additionally, CE is highly versatile and can explore a wide range of analytes, including proteins, nucleic acids, carbohydrates, and small molecules.[56]

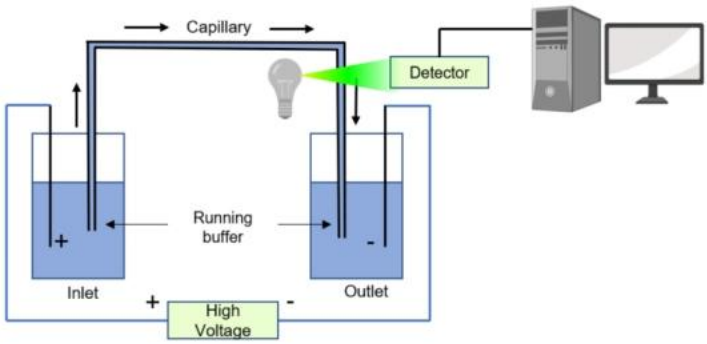


Figure 1.3: Scheme of a typical CE setup. The experimental setup comprises two vials designed for the inlet and outlet. During the separation process using CE, a voltage is applied

to induce movement of the electrolyte buffer. Additionally, a transparent window has been incorporated to facilitate laser detection. The laser employed is typically either laser-induced fluorescence or ultra-violet laser.

1.3.2. Theory

1.3.2.1. Electroosmotic flow

Electroosmotic flow (EOF) is a phenomenon that occurs in CE, a separation technique widely used in analytical chemistry and biochemistry. In CE, a small-diameter capillary tube filled with a buffer solution separates charged molecules based on size and charge(**Figure 1.4**).[57] When an electric field is applied to the buffer solution in the capillary tube, charged molecules in the sample move toward the oppositely charged electrode at different rates based on their size and charge. However, in addition to the movement of the charged molecules, a flow of buffer solution is also generated along the length of the capillary tube due to electroosmosis.[58]

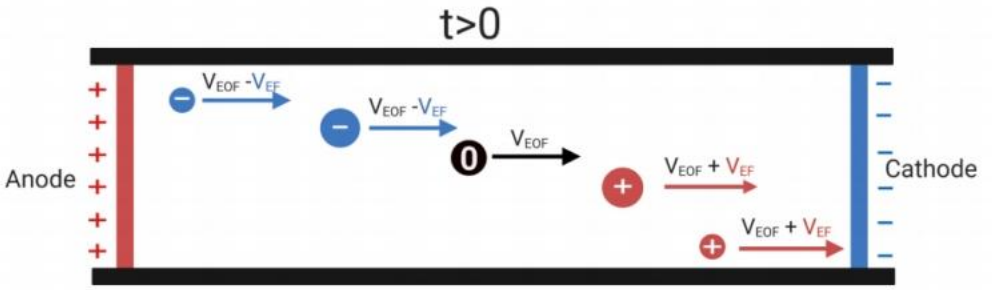


Figure 1.4: Scheme of the movement of analytes in CE based on their size and charge. After the establishment of the electric field, diverse species will experience different movement rates.

Electroosmosis is a phenomenon in which a charged surface, such as the inner surface of the capillary tube, attracts ions in the buffer solution and creates a buffer flow towards the oppositely charged electrode(**Figure 1.4**).[59] The direction and magnitude of the EOF depend on several factors, including the charge and size of the molecules in the buffer solution and the properties of the capillary tube. EOF can have both positive and negative effects on CE separations. On the one hand, EOF can enhance the separation of charged molecules by providing a constant and predictable flow of buffer solution along the capillary tube, which helps to maintain a stable separation environment. On the other hand, EOF can also cause unwanted band broadening and peak distortion if not adequately controlled.

Overall, EOF is a fundamental aspect of CE and can positively and negatively affect the separation of charged molecules. By understanding the factors that affect EOF and employing appropriate strategies to control it, CE can be a highly effective separation technique with numerous applications in analytical chemistry and biochemistry.

1.3.2.2. KCE and NECEEM

Kinetic capillary electrophoresis (KCE) is a technique that utilizes capillary electrophoresis to evaluate the affinity of interacting species. This method involves the interaction of two or more species, resulting in the formation of distinct peaks. By analyzing these peaks, which usually correspond to each species and interaction, the association and dissociation rate constants and the measure of affinity (a dissociation constant, K_D) can be determined.[60]

KCE is a broad term encompassing various techniques using a dynamic capillary wall coating to improve separation efficiency. Nonequilibrium Capillary Electrophoresis of Equilibrium Mixtures (NECEEM), considered a sub-type of KCE, is a powerful technique used in analytical chemistry to determine the binding constants (**Figure 1.5**)[61]. Unlike conventional CE methods, which are typically utilized for separating a diverse range of analytes, NECEEM is used to investigate the equilibrated mixtures of a ligand and a target to determine the binding affinity and stoichiometry of the complex.[61-63] This research primarily employs NECEEM as the CE mode for all following experiments.

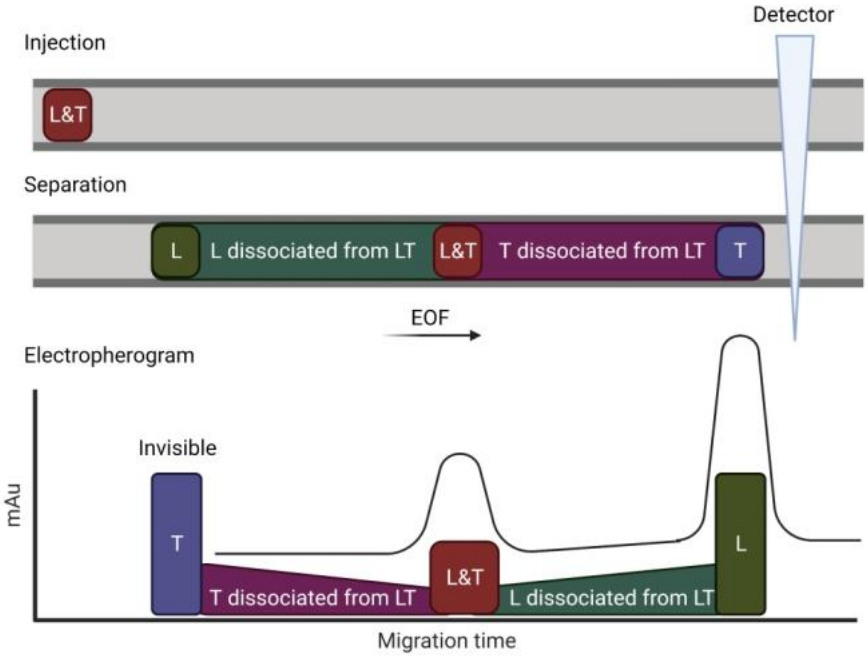


Figure 1.5: A theoretical scheme of NECEEM. After injection, the sample will undergo separation based on its charge/size ratio, resulting in two visible peaks on the electropherogram. Theoretically, the sample can be separated into five components: protein (T), aptamer-protein complex (L&T), free aptamer (L), left smear indicating the protein dissociated from complex, and right smear indicating aptamer dissociated from the complex. In Laser-Induced mode CE, only the smear between the complex (L&T) and the accessible aptamer-only peak (L) will likely be observed.

1.3.3. CE-SELEX

CE-SELEX (Capillary Electrophoresis-Systematic Evolution of Ligands by Exponential Enrichment) is a powerful technique to select aptamers that bind to a specific target molecule with high affinity and specificity.

The CE-SELEX (Capillary Electrophoresis-SELEX) and conventional SELEX use the same basic principles to isolate bound aptamers from the pool. However, CE-SELEX incorporates capillary electrophoresis, a separation technique that separates charged molecules based on their size and charge, to increase the efficiency of the selection process. (**Figure 1.6**). In CE-SELEX, a library of random oligonucleotides with a specific fluorescent label is incubated with the target molecule, and the resulting complexes are separated using CE[64].

Most importantly, CE-SELEX has several advantages over traditional SELEX methods: **Increased speed**: CE-SELEX can be completed in a shorter time than conventional SELEX methods because it allows for the selection of aptamers in a single round. **High resolution**: CE separates nucleic acid molecules based on size and charge, providing high resolution and accuracy in selecting aptamers[65, 66]. **Automated**: CE-SELEX can be fully automated, allowing for the selection and amplification of aptamers without manual intervention[67, 68]. **Less sample requirement**: CE-SELEX requires smaller amounts of starting material than traditional SELEX methods, making it suitable for targets that are rare or difficult to isolate. **Reduced interference**: CE to separate the bound and unbound oligonucleotides can eliminate interference from other molecules in the sample, leading to higher specificity in the selection of aptamers[64, 69, 70].

Various research groups have employed the CE-SELEX technique to select aptamers for specific targets. Huang and his team focused on the recombinant human LRP-6, identifying four candidate aptamers exhibiting a stem-loop structure[73]. Andrade's group used CE-SELEX to isolate an ZIKV60 aptamer specific to the Zika virus non-structural protein 1 (NS1)[74]. Similarly, Qu and her colleagues applied the CE-SELEX method to obtain an ssDNA aptamer directed against thyroglobulin[75].

Overall, CE-SELEX is a fast, efficient, and automated method for selecting aptamers that can save time and labour, increase specificity, and reduce interference in the selection process.

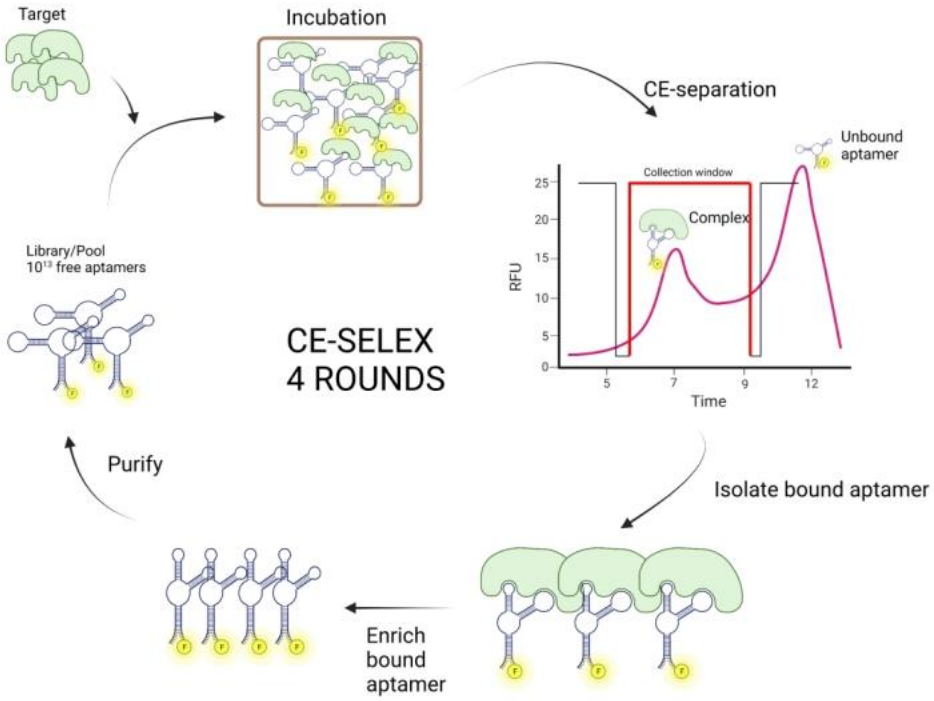


Figure 1.6. Scheme of selection of aptamers by CE-SELEX. NECEEM mode is employed for isolating the binding aptamer. In this approach, the complex containing the binding aptamer will be identified based on its distinctive peak in the electropherogram and subsequently collected for amplification in the next round.

1.4. Biolayer Interferometry

Biolayer interferometry (BLI) is a technique that can be used to analyze biomolecular interactions in real time without the need for labelling molecules. BLI is based on the interference of light waves reflected from two surfaces, one coated with immobilized biomolecules. BLI is a valuable analytical tool that can determine the binding kinetics of different biomolecules, including proteins, RNA, DNA, and small molecules.

In BLI, a biosensor tip is coated with a layer of biomolecules, and then the information is exposed to a sample containing the biomolecule of interest. The layer's thickness and refractive index change when the target biomolecule binds to the immobilized biomolecules on the sensor surface. This change is detected by measuring the interferogram of light waves reflected from the biosensor surface, and the binding kinetics of the interaction can be calculated (**Figure 1.7**).

In this study, we will use BLI to determine the binding affinity of aptamers to proteins and calculate the Apparent K_D value. And the result will be a powerful support to values obtained from CE.

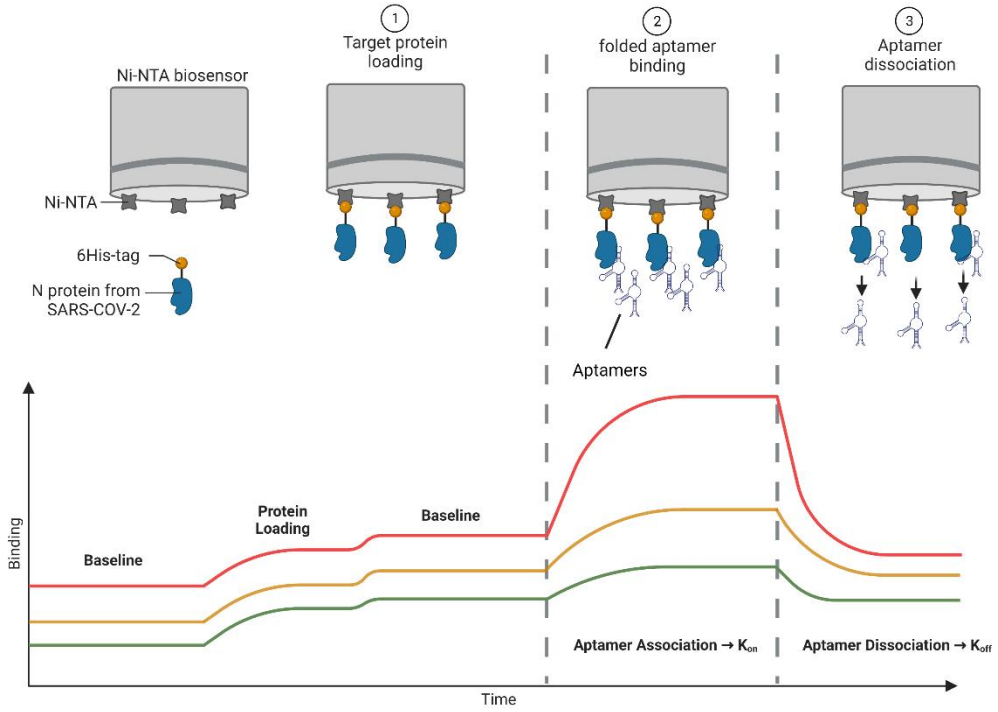


Figure 1.7: Scheme of an aptamer-based BLI setup. Different colors represent different aptamer clones. The protein is immobilized on the sensor and immersed in an aptamer buffer-filled vial. Once binding has stabilized, the sensor is submerged in a sterile buffer vial to facilitate dissociation.

1.5. Thesis Objectives, Hypothesis and Methodology

The primary objective of this study is to identify novel aptamers targeting the N protein of SARS-CoV-2, ensuring these sequences are distinct from those reported in previous research. To characterize the binding specificity of the selected aptamers, the dissociation constant was measured, providing a quantitative assessment of the aptamer's binding affinity to the protein. For a holistic evaluation, aptamers from existing literature were employed as positive controls, tested under the same experimental conditions. Furthermore, we employed the Bio-Layer Interferometry (BLI) method — a heterogeneous technique reminiscent of real-world applications such as antigen tests — to validate the binding efficacy of the aptamers. This approach not only evaluates the aptamers' viability for diagnostic tool development but also offers insights into their interactions under varied conditions.

The central hypothesis of this study revolves around two core assertions. First, aptamers selected via Capillary Electrophoresis (CE) are anticipated to display sequences differing from those documented in previous research. Moreover, it's expected that the dissociation constant (K_D) of clones derived from the CE-SELEX process, as assessed through CE, will be notably lower than that of the negative control, which hasn't been subjected to the SELEX procedure. The second assertion posits that when evaluated using Bio-Layer Interferometry (BLI), the K_D values for these selected clones should remain lower than the control, aligning

with the CE-derived findings.

Regarding the methodology, the process begins with the use of CE to differentiate between the aptamer-protein complex and free aptamers. Following this separation, Next-Generation Sequencing (NGS) will be employed to determine their sequences. With the sequences identified, initial binding affinities of the aptamers will be gauged using Bio-Layer Interferometry (BLI). Subsequently, the K_D will be ascertained using CE, adhering to the conditions set during the CE-SELEX stage. To fortify and expand on these findings, a re-evaluation of the K_D values will be undertaken using BLI, thus validating the outcomes obtained from CE and elucidating the diagnostic potential of the aptamer.

1.6. Chapters Overview

The specific objectives which the studies aimed to address are described in Chapters 2 - 4 of the thesis. Beginning with Chapter 2, the quality of purchased N protein was checked using direct infusion mass spectrometry and the selection of ssDNA aptamers targeting SARS-CoV-2 N protein using CE-SELEX was explained, then the binding affinity was screened for the three pools went through the SELEX. In Chapter 3, the aptamer sequences were identified using a software tool named FASTAptamer. There were six DNA sequences passed through the criteria used for choosing the sequences targeting the N protein, BLI was conducted to screen their binding affinity. Two clones ECK4 and ECK6 were selected for downstream experiment. Then a pre-truncation experiment was performed to ensure their binding ability without primer region. After that, Chapter 4 describes 2 methods for measuring the K_D of truncated ECK4 and truncated ECK6, BLI result was used to compare with documented clones and provide insights to their potential utility.

Chapter 2. Selection of DNA aptamers to N protein using CE-SELEX.

2.1. Objective

In the second chapter of this thesis, we describe the SELEX process and protein purity validation. Our objective is to successfully perform the CE-SELEX and enhance the binding with number of the SELEX rounds increasing. To accomplish this, we present a comprehensive CE-SELEX process, which includes optimizing the CE method, performing the PCR used in the screening process, and the binding of the enriched DNA pools were screened under the BLI experiment.

2.2. Materials and Method

Phosphate Buffer Saline (PBS) without Ca^{2+} - Mg^{2+} (Cat. No. 10010031) and PBS with Ca^{2+} - Mg^{2+} (Cat. No. 14040141) were purchased from Gibco. The 0.75 um capillary was purchased from Beckman-Coulter (Cat. No. 1398972). Phire Hot Start II DNA Polymerase Kit (Cat. No. 2398793), lambda exonuclease reaction kit (Cat. No. EN0581), and Ultra Low Range DNA Ladder (Cat. No. 10597012) were purchased from ThermoFisher Scientific. The DNA clean-up kit (Cat. No. 12596-4) was purchased from Qiagen. The DNA (N40) Library (5'-CTCCTCTGACTGTAACCACG-(N40)-GCATAGGTAGTCCAGAAGCC-3'), 5' Cyanine5

(Cy5) dye labeled forward primer (5'-/5CY5/CTCCTCTGACTGTAACCACG-3'), and 5' phosphorylated reverse primer (5'-/5Phos/ GGCTTCTGGACTACCTATGC-3') were purchased from Integrated DNA Technologies (Newark, NJ, USA), SARS-CoV-2 nucleocapsid protein was purchased from ACRO biosystems (Newark, DE).

2.2.1. Protein purity validation

Analysis was performed with a quadrupole time of flight spectrometer (Waters Synapt G2 HDMS, Massachusetts, USA) for direct infusion ESI-MS after diluting the eluant in Milli-Q water. MS detection settings were adjusted to a capillary voltage of 2.25 kV, sampling cone and source offset at 140, a source temperature of 150 ℃, a desolvation temperature of 600 ℃, a cone gas flow at 20 L/h, desolvation gas flow set at 600 L/h, and collision energy set at 6 eV with data collected over 1 minute. Mass spectra were manually processed using MaxEnt1 deconvolution with a resolution of 20,000 over a range of 1000-4000 *m/z* at a rate of 10 Hz.

2.2.2. CE-SELEX method

Library preparation: Prepare a 50 μL solution containing a 25 nM N40 DNA library or aptamer pool and 180 nM protein. Snap-cool (95 ℃ for 5 min and 4 ℃ for 10 min) the mixture and transferred it into a sample vial.

EOF preparation: An appropriate EOF must be created to achieve a stable NECEEM partitioning of aptamer and aptamer-protein complex. The CE method was set up based on Table S2.1. In the CE method, the initial 20 steps involve creating the EOF, including acid, base, and buffer fill. Increasing the duration of the acid wash ensures that all silanol groups on the capillary wall are protonated, which is an essential requirement for stable partitioning.

Aptamer-collection window: To determine the migration time of any species to the end of the capillary, multiply the migration time to the detector by the conversion factor f, which is calculated as $L_{total}/L_{detector}$, where L_{total} and $L_{detector}$ represent the length from the capillary inlet to its outlet and detector, respectively. In this case, f equals 1.2, calculated as 60 cm/50 cm.

Inject and separate: Capillary electrophoresis selection was performed using a 60 cm-long (with 50 cm to the detector), 75 μm-inner diameter, 360 μm-outer diameter uncoated fused silica capillary (Polymicro Technologies, Phoenix, AZ) on a P/ACE MDQ capillary electrophoresis system (Beckman Coulter Inc., Fullerton, CA). The mixture was separated under 25 kV voltage (normal polarity) in 10 mM ammonium acetate buffer at 25 °C and the separation was monitored under LIF detection (at 488 nm for Cy5-labeled DNA targets). The N protein-aptamer complex migrated off the capillary before the unbound ssDNA and was collected into a vial containing 48 μL separation buffer at the capillary outlet. After collecting the bound sequences, the voltage was turned off and the unbound sequences were washed off the capillary into a waste container vial. The method was set up based on **Table S2.2.**

2.2.4. Aptamer pool preparation

2.2.4.1. Symmetric PCR

The PCR master mix was made following the manufacturer's guidelines, and the other necessary components listed in **Table 2.3** were added. All PCR reactions involved in this experiment were done in 25 cycles. The samples were then subjected to electrophoresis on a 1 % agarose gel for 30 minutes at 120 V after staining with 1 X Gel Red (Invitrogen, Cat No. 41003). The resulting gel was imaged under the RGB channel to evaluate the optimal number of cycles. The ideal number of cycles was determined by selecting the lane with the most intense and well-defined band at the appropriate migration rate relative to the control while minimizing the presence of by-product bands. The symmetric PCR was conducted in a thermocycler machine. The template was denatured at 95 ℃ for 2 minutes, followed by primer annealing at 55 ℃ for 15 s and DNA elongation at 72 ℃ for 5 s.

Table 2.3. PCR master mix recipe.

Final Concentration		Volume (μL) per 10 reactions	Catalogue Number
ddH2O	---	266	---
5x Phire Buffer[3]	1X	100	F-527L
dNTPs	0.4 μM	0.4	R1121
DMSO	7 %	35	---
Forward primer[1]	0.4 μM	2	Integrated DNA Technologies
Reverse primer[2]	0.4 μg	2	Integrated DNA Technologies
Phire II Hot Start DNA polymerase	1 U/25 μL	5	F122L
DNA template	2.5 pg-25 ng	---	---

[1]Forward primer: 5'-/5CY5/ CTC CTC TGA CTG TAA CCA CG-3'

[2]Reverse primer: 5'-/5Phos/ GGC TTC TGG ACT ACC TAT GC-3'

[3]Buffer components: The buffer component is not revealed in the manual.

After PCR amplification, all samples were combined, and 10x exonuclease reaction buffer and lambda exonuclease were added to the combined samples according to Table 2.4. The samples were re-divided into PCR tubes, and the digestion protocol was performed in MasterCycler (Eppendorf) at 37 ℃ for 400 minutes, at 95 ℃ for 25 minutes and at 4 ℃ for 10 minutes.

Table 2.4. Exonuclease digestion for single-stranded DNA recipe.

Final Concentration	Volume (μL)per 50 μL reaction		Catalogue Number
Exonuclease	1 X	2	EN0561

reaction buffer			
Lambda Exonuclease 10U/μL	10 U	1	EN0561
ddH2O	---	47	---

In the previous steps, the single-stranded DNA product was loaded onto a 1 % agarose gel with the library and primer controls and stained with Gel Red to a final concentration of 1X. The samples were then run on the gel in 1X TBE buffer at 120 V for 35 minutes, with only one large well prepared for the DNA product. The product was then cut into small cubes using a scalpel, transferred to a falcon tube, and incubated at room temperature in ddH2O at 200 rpm overnight. The liquid was carefully removed from the gel cubes. The samples were then dehydrated at 60 ℃ until dry and resuspended in 50 μL of ddH2O.

2.2.4.2. Concentration determination

To determine the volume of DNA sample needed to achieve a final concentration of 150 nM, mix 20 μL of the DNA sample with 180 μL of water, and transfer the mixture into a 0.5 mL tube. Then, use nanodrop to measure the ssDNA concentration and convert the measurement to molarity.

2.2.5. Binding validation of selected aptamer pools

The binding affinity of aptamer pools against SARS-CoV-2 N protein (ACRO biosystems, Newark, DE) was performed using Ni-NTA biosensors and a BLI instrument named Octet N1 (Sartorius, Bohemia, NY). Firstly, the DNA-enriched pools were prepared in the Assay Buffer (PBS with 0.5mM $MgCl_2$ and 1mM $CaCl_2$, pH 7.4, Gibco) at a concentration of 500 nM and heated at 95 °C for 5 min, then 4 °C for 10 min cooled. The binding assay was performed in 0.5 mL light-blocking tubes at 25 °C with a total volume of 250 μL. In the next steps, 20 μg/mL of his-tagged N protein was loaded on Ni-NTA biosensor tips. During the protein binding assay, a baseline was established in the Binding Buffer (the Assay Buffer with 0.05 % Tween20) followed by monitoring the association of the DNA-enriched pools with the N protein on the biosensors. The association step includes four samples: N40 library, Round 1, Round 2, and Round 3, N40 Library was selected as negative control since it hadn't undergone any SELEX process. The assay is performed at room temperature, with a 30 s baseline, 120 s loading, 100 seconds association, and 150 s dissociation.

Table 2.4. Workflow of BLI.

Step	Time	Purpose
Baseline	30 s	Create a baseline for the loading step
Loading	120 s	For His-tag protein binding to the Ni-NTA sensor
Baseline	30 s	Stabilize the binding of protein to the sensor
Association	100 s	Monitor the binding of aptamers and protein
Dissociation	150 s	Monitor the binding strength of aptamers and proteins

2.3. Results and discussion

2.3.1. Identify the Purity of the Commercial N protein

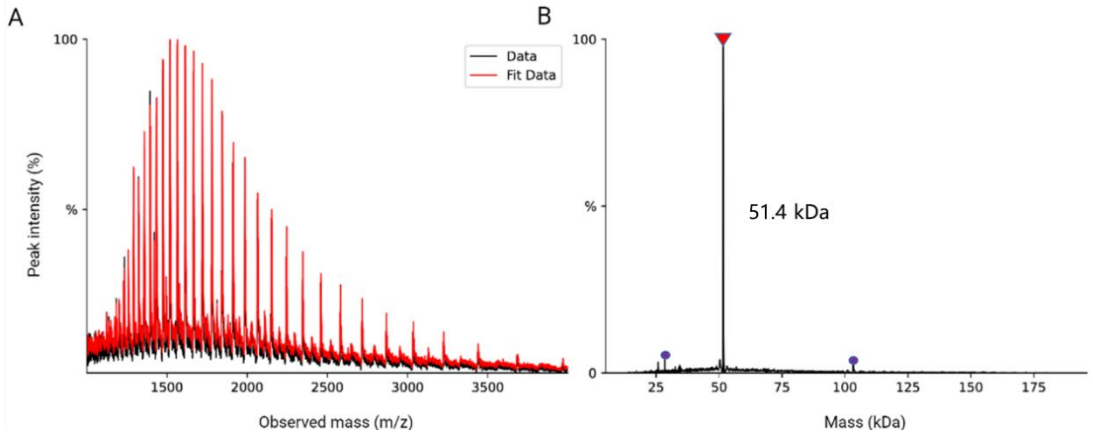


Figure 2.1. Mass spectra and deconvoluted MS of purchased N protein. (A) Convoluted Mass spectrum. (B) Deconvoluted Mass spectrum. The red arrow indicates the sample of interest which was normalized to 100 % of intensity, while the purple sphere denotes other observed signals and their intensities are around.

SARS-CoV-2 Nucleocapsid protein (NUN-C51H9) purchased from ACRO biosystems company is expressed from *E. coli (Escherichia coli)* cells. It contains amino acids Met 1 - Ala 419. This protein carries a poly-histidine tag at the N-terminus and the protein has a calculated MW of 51.4 kDa. Here is the reported protein sequence in FASTA format (https://www.ncbi.nlm.nih.gov/protein/QHO62115.1).

>sp|P0DTC9|NCAP_SARS2 Nucleoprotein OS=severe acute respiratory syndrome coronavirus 2 OX=2697049 GN=N PE=1 SV=1.
MSDNGPQNQRNAPRITFGGPSDSTGSNQNGERSGARSKQRRPQGLPNNTASWFTALT
QHGKEDLKFPRGQGVPINTNSSPDDQIGYYRRATRRIRGGDGKMKDLSPRWYFYYL
GTGPEAGLPYGANKDGIIWVATEGALNTPKDHIGTRNPANNAAIVLQLPQGTTLPKG
FYAEGSRGGSQASSRSSSRSRNSSRNSTPGSSRGTSPARMAGNGGDAALALLLLDRL
NQLESKMSGKGQQQQGQTVTKKSAAEASKKPRQKRTATKAYNVTQAFGRRGPEQT
QGNFGDQELIRQGTDYKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTYTGAIKL
DDKDPNFKDQVILLNKHIDAYKTFPPTEPKKDKKKKADETQALPQRQKKQQTVTLLP
AADLDDFSKQLQQSMSSADSTQA

The MS results indicated a high purity for the N protein, with other impurity signals being negligible in comparison. The deconvoluted protein size was determined to be 51.4 kDa, matching the calculated mass of the N protein as provided by both the company and NCBI data.

2.3.2. CE-SELEX

In CE, two necessary signals are usually analyzed: the current and fluorescence intensity. The electropherogram depicted in **Figure 2.2** shows two significant peaks. Although the complex size is larger than the aptamer, it has a more positive charge. Proteins have amino acid residues on their surface that can be positively or negatively charged, depending on the pH of

the surrounding environment. Consequently, the complex can move more rapidly through the capillary due to its higher charge-to-size ratio.

The current indicates that the collection window was created to collect the complex with the target aptamers. The voltage was terminated midway because the outlet vial needed to be replaced with the collection vial. In **Figure 2.2**, the blue lines represent the current, while the electropherograms show the samples' RFU and time of CE separation. The separation was halted twice: first, to change the separation vial to the collection vial to make sure the complex (binding aptamer) is collected, and second, to change the collection vial to the free aptamer collection vial because it is required that only the complex will be composed eventually.

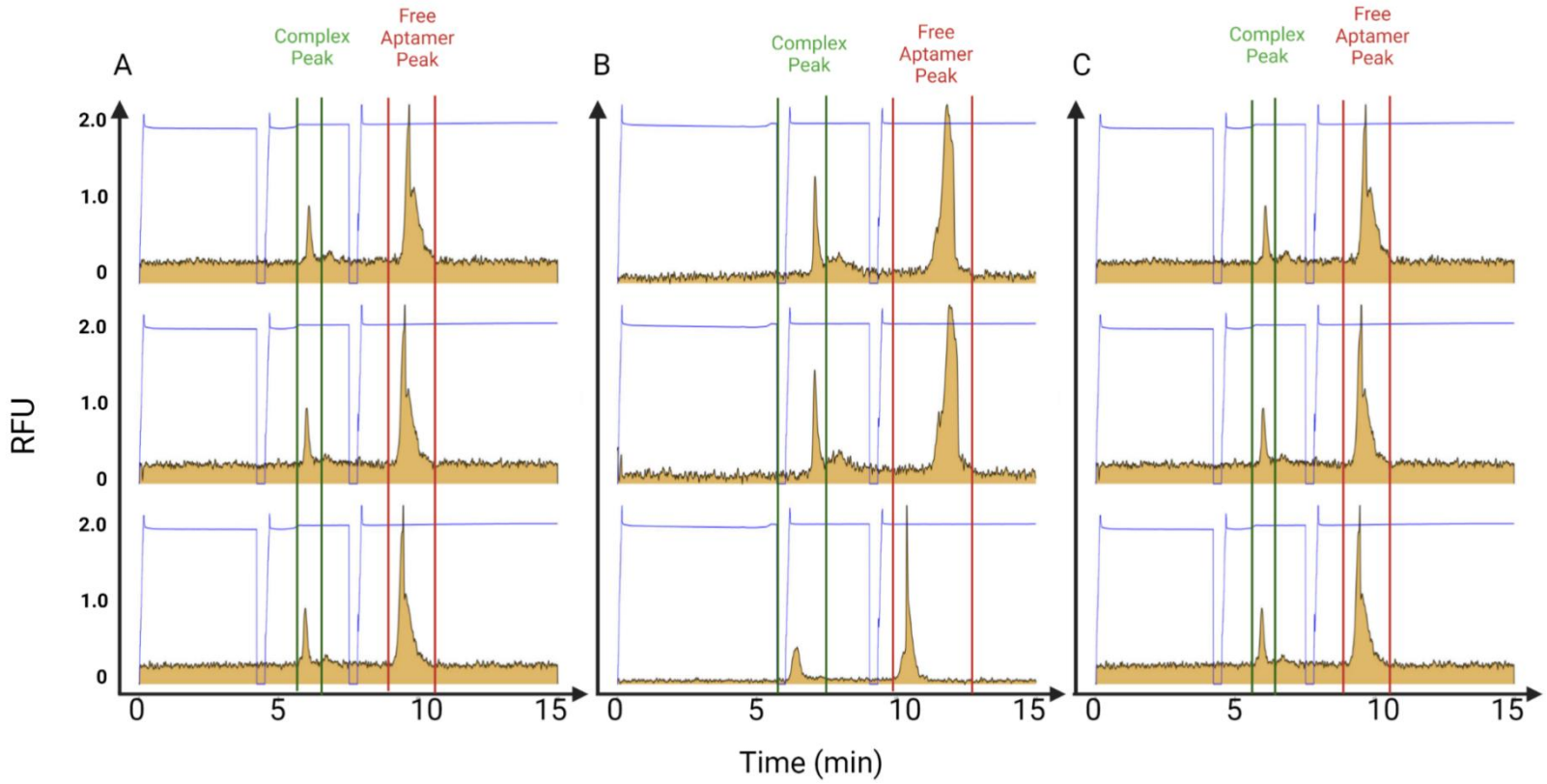


Figure 2.2. Representative electropherograms for three rounds of CE-SELEX. When synthesizing the N40 DNA library and primers for this experiment, a FAM (Fluorescein amidites) label was attached to the DNA, and in CE, with laser-induced fluorescence (LIF) detection (488 nm excitation, 520 nm emission) was employed. An uncoated fused-silica capillary was used with the following dimensions: 60 cm × 75 µm i.d. × 350 µm o.d. The length from the injection end to the detection window was 50 cm(A) 25 nM of aptamer and 180 nM protein. (B) 25 nM of aptamer and 160 nM protein. (C) 25 nM of aptamer and 100 nM protein.

2.3.3. Symmetric PCR

To prepare for the next round of aptamer selection, the collected aptamer needs to be dissociated from the DNA-protein complex and amplified using PCR. This amplification process will produce more potential binders for the subsequent selection rounds.

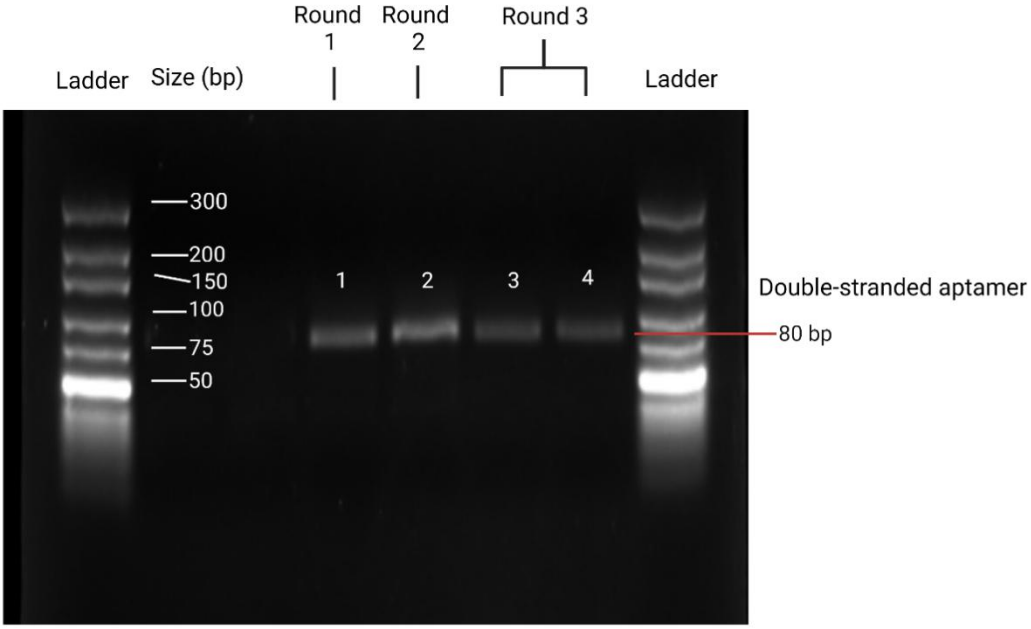


Figure 2.3. Agarose gel separation of double-stranded DNA products after PCR. The PCR product bands in lanes 1&2, 3, and 4 correspond to rounds 1, 2, and 3, respectively. The most left and most right lanes are the Ultra low DNA ladder (Thermo Fisher Scientific). The PCR samples were run on 2 % agarose gel at 110 V for 60 min. The band on the gel was observed with lengths of 80 bp and it is the exact size of a double stranded aptamer.

PCR produces double-stranded DNA, but aptamer binding requires single-stranded DNA for proper folding. Lambda exonuclease is an enzyme derived from the bacteriophage lambda. It is well known for its ability to digest the 5'-phosphorylated strand of double-stranded DNA (dsDNA), leaving the complementary strand intact. in this experiment, the reverse primer includes a phosphate group on the coding strand during PCR. With the help of Mg^{2+}, lambda exonuclease efficiently digests the complementary reverse strand in a 5' to 3' direction. This process allows for the recovery of the initial coding strand. The single-stranded DNA is 80 nt; however, the size indicated by the ladder is about double-stranded DNA. Theoretically, the size should be divided by half to demonstrate the extent of single-stranded DNA, as shown in **Figure 2.4**. The digested and purified DNA will be used for the next round or NGS (Next Generation Sequencing).

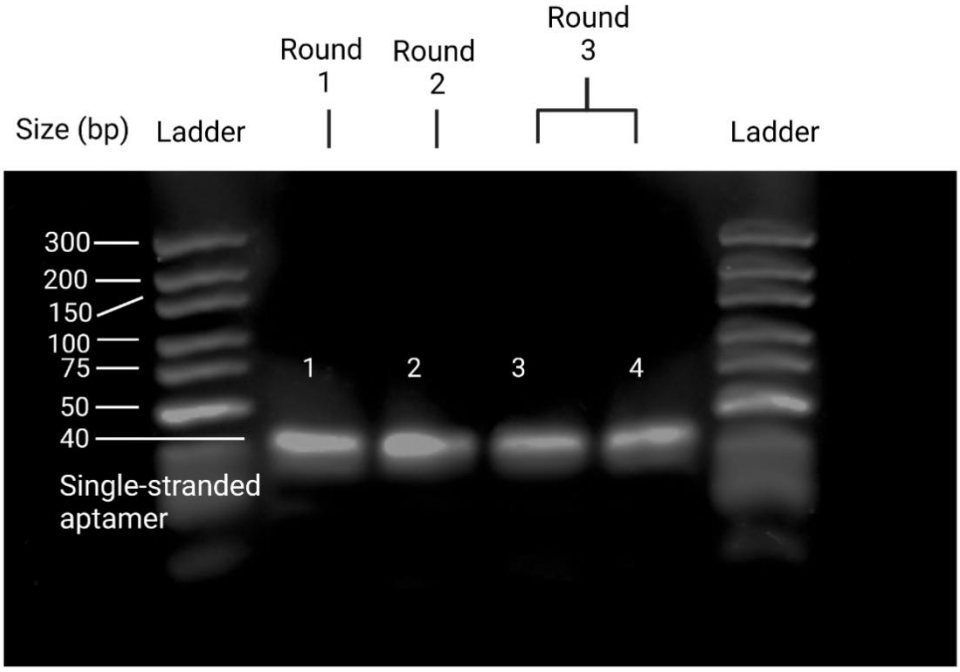


Figure 2.4. Agarose gel electrophoresis of single-stranded DNA product after being digested by Lambda nuclease. The Lambda Nuclease products are visualized in an agarose gel. The Ultra low DNA ladder (Thermo Fisher Scientific) is the most left and right lanes, and bands in lanes 1&2, 3, and 4 correspond to rounds 1, 2, and 3, respectively. The digested samples were run on 2 % agarose gel at 110 V for 60 min. The band on the gel was observed with lengths of 80 nt and it is the exact size of a single stranded aptamer.

2.3.4. Selected aptamer pool binding validation

After three rounds of selection, the enriched pools were tested for their preliminary binding affinity using BLI. The first and second pools showed the same binding or even worse than the N40 library (**Figure 2.5**). However, the third pool binds much better than in previous rounds. As a result, all enriched pools were sent to next-generation sequencing (NGS) analysis to obtain the DNA sequence.

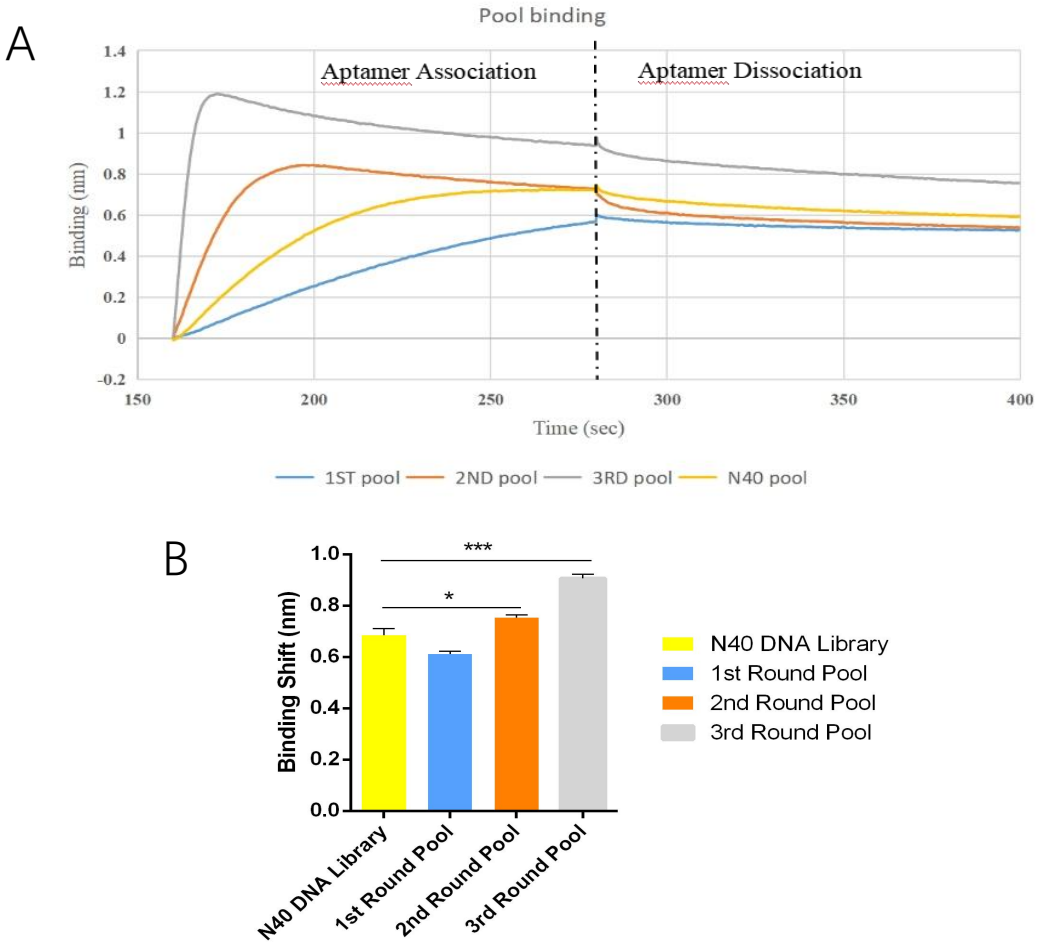


Figure 2.5. BLI results for pool binding screening. In the evaluation of aptamer pools from three selection rounds conducted via CE-SELEX, binding screening was executed using BLI, employing the N40 DNA library as a reference control. (A) The BLI Sensorgram indicated a progressive increase in binding affinity correlating with the number of selection rounds. (B) The binding shift of three DNA-enriched pools. Quantitative Analysis: The depicted error bars represents the mean values ± standard deviations (SDs) based on three independent trials. Statistical significance was established with *$p < 0.05$ and ***$p < 0.001$, as determined by the Student's t-test.

2.4. Discussion

This chapter focuses on using CE to perform CE-SELEX to generate DNA-enriched pools with enhanced binding affinity with N protein. Firstly, the protein's purity was confirmed, and its identity was verified through mass spectrometry, which revealed that the measured size aligns with the documented size of the protein, however, when conducting SELEX with the *E. coli*-expressed N protein of SARS-CoV-2, potential differences in post-translational modifications (PTMs), protein folding, and structural variances can impact outcomes. Absent PTMs in *E. coli* might alter the binding affinity and specificity of selected ligands, potentially limiting their relevance to the native viral protein. Differences in protein folding within *E. coli* can expose or conceal regions, leading to oligonucleotide selections that don't accurately target the native state. Structural disparities between the expressed and native proteins further risk identifying ligands that may not efficiently bind in a viral context. Therefore, using *E. coli*-expressed proteins necessitates careful validation to ensure the identified ligands' applicability to the native viral protein.

Secondly, CE is chosen due to its automation, reproducibility, and simple sample preparation and operation. The separation of complex and free aptamers is evident in the clear separation of two distinct peaks in the resulting graph produced by CE. Aptamers with binding ability are isolated from the complex peak and amplified to create a pool for subsequent rounds of screening. After implementing the CE-SELEX method, three rounds of aptamer pools were identified. To pinpoint a specific aptamer clone, it is essential to conduct a preliminary screening for binding affinity. BLI was employed for this affinity screening, with the binding shift distance serving as a visual metric to illustrate enhanced binding. This binding shift distance mirrors the variations in the thickness of the ligand layer. Notably, the thickness indicates the volume of ligand molecules adhering to the target layer. The results indicated an increase in binding affinity compared to the binding of the N40 DNA library. Moreover, binding efficacy seemed to rise with each successive round. These outcomes suggest that CE-SELEX played a role in enriching aptamers with a strong binding capacity to the target.

Traditional SELEX often requires multiple iterative rounds of binding, separation, and amplification, which can be time-consuming and introduce biases, such as beads-SELEX and filter-SELEX. In contrast, CE-SELEX offers higher resolution and faster separation of bound and unbound aptamer sequences, often reducing the required selection rounds. Here, we performed 3 rounds of selection within 4 days; typically, beads-SELEX needs to perform more than 6 rounds to ensure the specificity of the selected aptamer. While CE-SELEX presents several advantages for aptamer selection, it does have its challenges in the experimental context. Firstly, the choice of buffer for incubation must be compatible with capillary electrophoresis. Yet, this very requirement might mean that the buffer isn't appropriate to those typically used in real body environments or diagnostic tools, which means aptamers are selected from a completely different environment; for instance, 10 mM ammonium acetate with PBS, this topic will be discussed in next chapter. Furthermore, CE-SELEX requires specialized equipment and expertise, which is not required for more straightforward techniques. Operating the machinery involved in CE-SELEX can be more challenging than using simpler tools like pipettors or magnetic beads.

However, there still needs to be more evidence to validate the hypothesis that three rounds of CE-SELEX is efficient to select aptamer clones with strong binding capacity. Therefore, in Chapter 3, we will send the DNA-enriched pools for Next Generation Sequencing and validate the binding of selected clones.

Chapter 3. Next-generation sequencing, Binding evaluation, and structure optimization of anti-N aptamer clones

3.1. Objective

This chapter focused on utilizing Next-Generation Sequencing (NGS) technology to identify aptamers that are the most abundant in the three DNA enriched pools, and a preliminary binding check will be conducted to validate the binding. Based on the binding shift of the BLI result, the top two aptamers who exhibit highest binding shift will be selected to do a pre-truncation experiment to ensure their binding after the structure truncation. The primer region of the aptamer will be hybridized by a complementary primer sequence to inhibit them forming secondary structure and involving in the binding process with protein. The binding affinity will be validated using BLI for primer-hybridized aptamers and normal aptamers.

3.2. Method

3.2.1. Next-generation sequencing

Sample preparation

The sample prepared for NGS needs to have an adaptor sequence on both sides, and it requires at least 500 ng double-stranded DNA total, normalized to a concentration of 20 ng/μL, stored in a microcentrifuge tube. For three pools, a PCR was performed with the same protocol as the '**symmetric PCR**' section in Chapter 2 but with different primers, and then it was purified using a purification kit. The samples were stained with 1X Gel Red (Invitrogen, Cat. No. 41003) and run on 1 % agarose gel at 120 V for 30 minutes. The gel was imaged in the ethidium bromide channel. The specific band was cut from the gel, and extract DNA by using a GeneJET Gel Extraction Kit (Thermo Fisher, USA). The final samples were stored in a microcentrifuge tube, sealed with parafilm, and delivered to the sequencing company. The NGS analysis was carried out by AZENTA company.

Forward adaptor-primer:
ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTCTCCTCTGACTGTAACCACG
Reverse adaptor-primer:
GACTGGAGTTCAGACGTGTGCTCTTCCGATCTTTTGGCTTCTGGACTACCTATGC.

Data analysis

Next-generation sequencing and data analysis After three selection rounds, each ssDNA pool was reamplified by symmetric PCR and purified before NGS. The fastq raw data were analyzed following the guideline from FASTAptamer as shown in Table 3.1. The data representing each aptamer pool were categorized into its cluster, followed by using the FASTAptamer-enrich tool to retrieve the most enriched sequences. From the result obtained from AZENTA companies, the mean quality scores of 3 pools are all higher than 36 and less than 40.
.

Table 3.1. Commander codes for analyzing NGS data.

Steps	Code in strawberry Perl	Purpose
1	C:\Users\Lab>cd Documents\FASTAptamer	Load a modular collection of scripts that can extract the information from the dataset.
2	>Perl fastaptamer_count -i YOUR RESULT.fastq -o NR1_R1.fasta	The Fastq file needs to be converted to a Fasta file. To count the reads for each pool.
3	>perl fastaptamer_cluster -i NR2_R1.fasta -o NR2_cluster.fasta - d 5 -f 4	Cluster families of the sequence.
4	>perl fastaptamer_enrich -x S1_Count_fasta/S1-R1_cluster.fasta -y S1_Count_fasta/S1-R2_cluster.fasta -o S1_Count.fasta/S1_vs_S2_enrich.tsv	The code can be used to obtained the enriched sequences.

3.2.2. Biolayer interferometry

BLI sensors were hydrated for 10 minutes before use. All aptamer clones were diluted to 150 nM with a total volume of 250 μL in PBS; the protein was diluted to 20 g/mL with a total volume of 250 μL in PBS. The protocol for the experiment is consistent with the one mentioned in the '**Binding validation of selected aptamer pools'** section from Chapter 2. For structurally modified aptamer, the same protocol was used for the binding validation. All

experiments are carried out in room temperature. The actual measurable quantity in BLI is a shift in the interference pattern due to the accumulation or loss of material (molecular binding) on the sensor tip, so in this experiment the binding shift will be reflected to the accumulation of aptamer clones that is observed. The shift in wavelength, usually measured in nanometers, is proportional to the change in thickness or density of the molecules on the sensor.

3.2.3. Aptamer structure modification

A 150 nM concentration of ECK4 was mixed with both 150 nM forward and 150 nM reverse primer sequences. The mixture was then heated to 95 °C for 5 minutes and subsequently cooled to 4 °C for 10 minutes to ensure the primers hybridized to the aptamer sequences. All samples were prepared in Assay Buffer, which consisted of PBS supplemented with 0.5 mM $MgCl_2$ and 1 mM $CaCl_2$, adjusted to pH 7.4. The same experimental protocol was repeated with ECK6. Subsequent BLI binding screening experiments were conducted following the methodology described in section 2.2.5 of Chapter 2. For negative controls, 80 nt non-aptamer DNA, non-primer-hybridized ECK4, and non-primer-hybridized ECK6 were utilized. See **Figure 3.1** for the scheme of the set-up for this pre-truncation experiment.

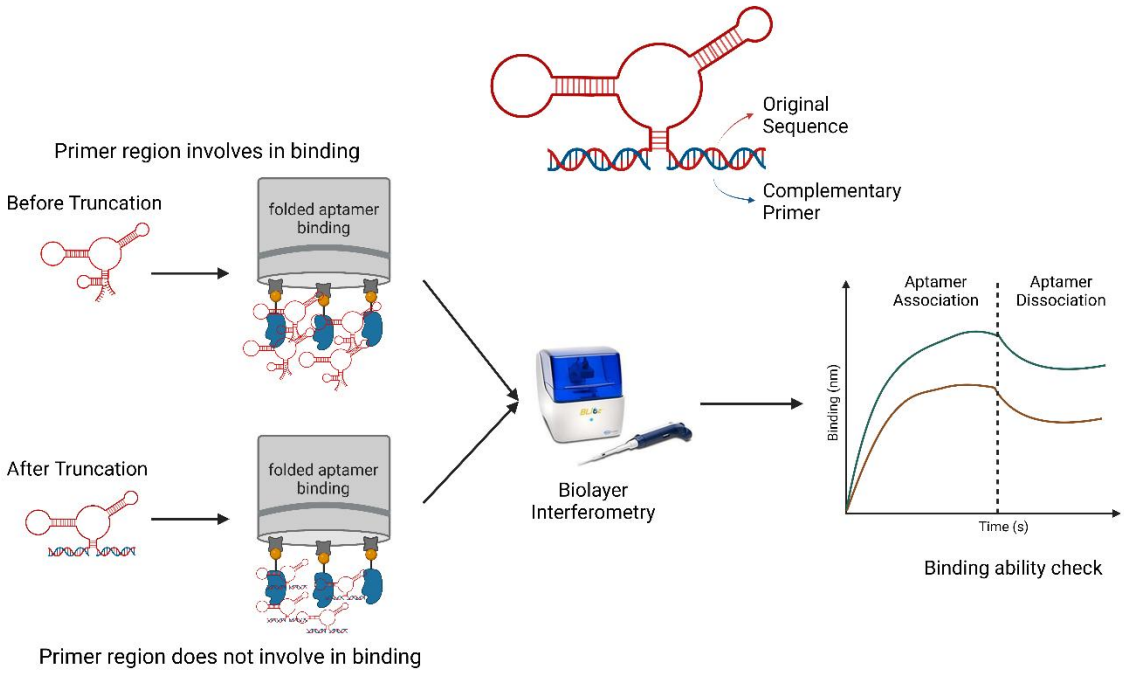


Figure 3.1. Scheme of the set-up for aptamer structure modification experiment. The aptamer before and after pre-truncation experiment were used in the experiment to check the binding ability. In the scheme, different curve colour represents different aptamer clone used in the experiment.

3.3. Results and Discussion

3.3.1. Next-generation sequencing

DNA purity

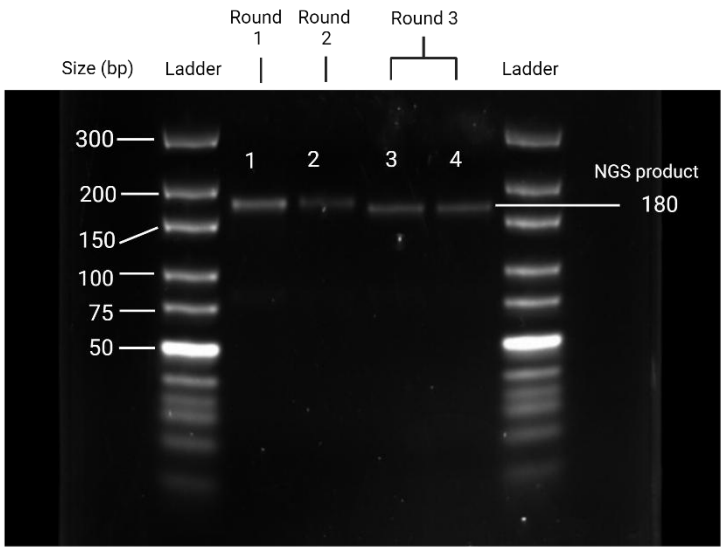


Figure 3.2. Agarose gel for NGS sample. 1 μL of the sample was diluted in 9 μL of water and loaded on the gel; lanes 1&2, 3 and 4 represented rounds 1, 2, and 3, respectively.

In **Figure 3.2**, four distinct bands were observed, all appearing to be the same size. However, it was discovered that the size of bands was larger than expected. Specifically, the band appeared to be approximately 180 nt, whereas it was expected to be 150 nt based on the molecular weight ladder. Despite this discrepancy, the sample was still submitted for Next Generation Sequencing (NGS), and the results revealed that the anomaly was an artifact and not indicative of a genuine mistake. A detailed explanation of this phenomenon will be provided in the discussion section of the thesis.

Sequencing result

Table 3.2. Full-length sequences (80 nt) of the selected aptamer against N protein.

Abundance rank	Clone sequence	Name
1	CTCCTCTGACTGTAACCACGTCATCGCTTTG CATATTCTATGTCCTGTCCATACTCGACAGC ATAGGTAGTCCAGAAGCC	Apt-N ECK1
2	CTCCTCTGACTGTAACCACGTTCACCGCTGC CCATCGGCATGGTTACCCGCGGTCCTTTTGC ATAGGTAGTCCAGAAGCC	Apt-N ECK2
3	CTCCTCTGACTGTAACCACGCCTCCTATTGC GCCCCAACGCATAGGCAATGGCCTACAGGG CATAGGTAGTCCAGAAGCC	Apt-N ECK3
4	CTCCTCTGACTGTAACCACGTTCGAACGAG GTTGCTCTACCGGCCAGGCCCTAGGTGCGT GCATAGGTAGTCCAGAAGCC	Apt-N ECK4
5	CTCCTCTGACTGTAACCACGGTTATACAGT GTTGTGACCGAGGCACATCCGCTATCACAG GCATAGGTAGTCCAGAAGCC	Apt-N ECK5
6	CTCCTCTGACTGTAACCACGAGTCAGGCGC TCATTCAAAATGAGGATCGTCATCATCCAA GCATAGGTAGTCCAGAAGCC	Apt-N ECK6

The NGS results for all six pools were processed using the FASTAptamer. Sequences were

counted and clustered into families based on the neighboring distance and were further subjected to the enrichment tool to calculate the fold-enrichment of individual sequences across three pools. The retrieved sequences were chosen based on three criteria, including (i) the total length of 80-82 nucleotides containing 40-42 nt random region flanked with forward and reverse primers, (ii) the identical sequences enriched in more than one pool, and (iii) the sequence composition without primer-dimer formation in the random region. The count-enrichment tool resulted in six unique sequences, and they are synthesized from IDT company.[71]

3.3.2. Bindings of aptamer clones

After obtaining the synthesized aptamer clones, a BLI was performed to screen the binding against the N protein. In our study, we reference an 80 nt non-aptamer DNA sequence. This sequence is distinct from the conventional aptamer construction (20 nt each for the forward and reverse primer regions, with a central 40 nt randomly generated region). Crucially, this 80 nt sequence has not been subjected to the SELEX evolutionary process, differentiating it further from standard aptamers. As expected, all clones showed different binding capacities, and 80 nt non-aptamer DNA didn't show significant binding like other binders.

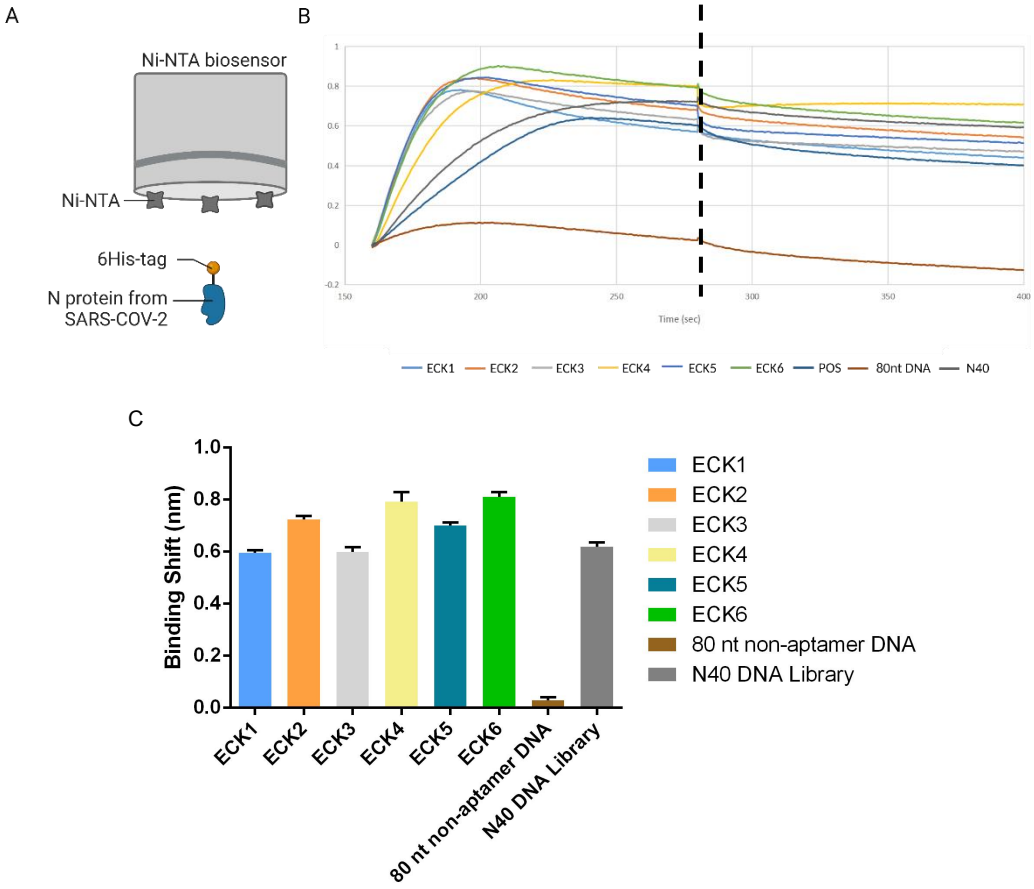


Figure 3.3. BLI results for 6 synthesized clones. The protein concentration was set up to 20 mg/mL, and the clones' concentration was set to 150 nM. All samples are snap-cooled before incubating with each other. (A) Scheme of aptamer-protein experiment set-up. (B) Different colors represent different samples in this experiment. (C) The binding shift of 6 clones and 80 nt non-aptamer DNA, ECK4 and ECK6 showed highest binding shift and all clones have higher binding shift than non-aptamer DNA.

The binding shift measurements obtained from BLI, ECK4, and ECK6 display the highest values,

approximately 0.80 nm and 0.81 nm. These results indicate a more substantial binding of ECK4 and ECK6 aptamer molecules to the protein, with all aptamers exhibiting higher binding shifts compared to an 80 nt non-aptamer DNA.

3.3.3. Bindings of aptamer clones after structure modification

After performing pre-validation of the binding of all aptamer clones, ECK4 and ECK6 were selected for truncation experiments. To ensure that the binding was not affected by the absence of primer regions, additional BLI assays were performed, and it was confirmed that the binding remained strong. The results, depicted in Figure 3.4, show that all primer region-modified aptamer clones exhibited superior binding performance compared to the full-length aptamer.

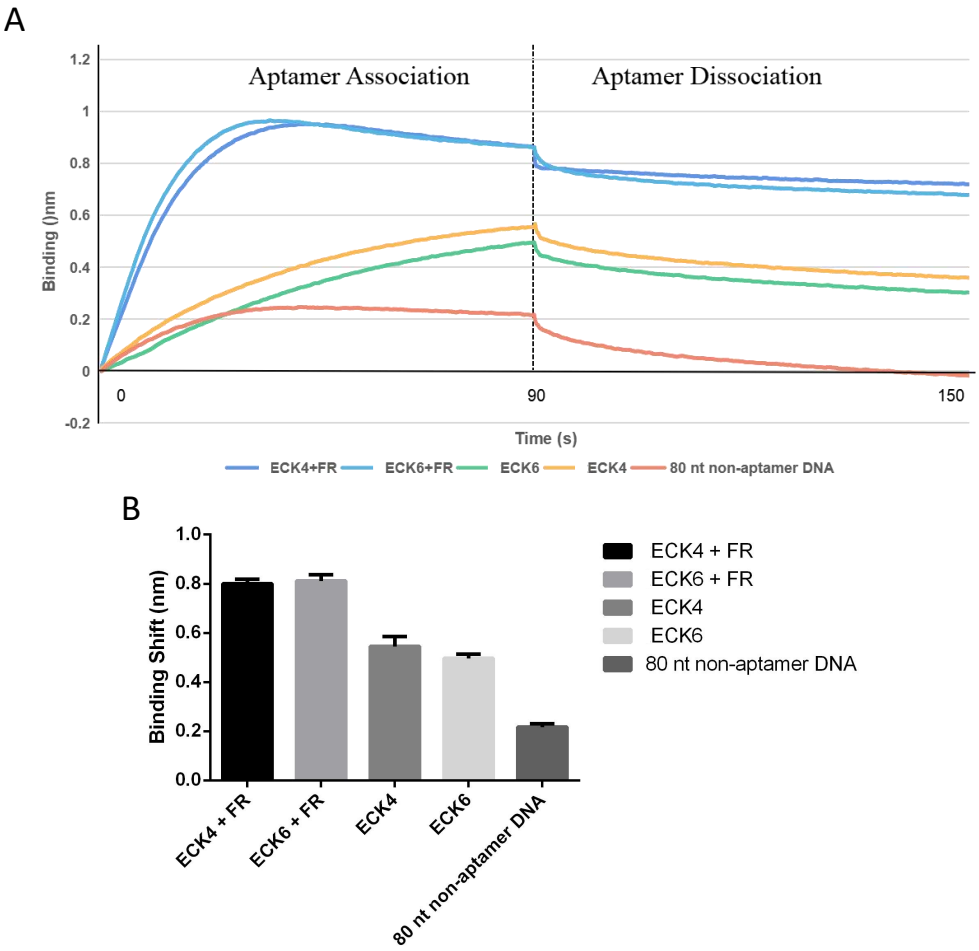


Figure 3.4. BLI results for double-primer-hybridized aptamer. BLI sensorgram (A) and Binding shift (B) for primer-hybridized ECK4 and ECK6, ECK4 and ECK6 with original sequences and 80 nt non-aptamer DNA. Different colors represent different clones used in this experiment. FR represents both forward primer and reverse primer. Single-primer aptamer BLI results will be present in the appendix.

The binding shift measured by BLI for primer-hybridized tECK4 is around 0.8 nm and for primer-hybridized tECK6 is 0.81. they are both higher than the binding shift of tECK4 and tECK6 which is 0.58 nm and 0.54 nm, respectively. And all the clones have a higher binding shift of 80 nt non-aptamer DNA which is 0.21 nm. After confirming the binding of the primer region modified aptamer, 40 nt ECK4, and ECK6 were purchased without the primer region from IDT. The newly synthesized aptamers were labelled with biotin and FAM. The truncated ECK4 and ECK6 were renamed to tECK4 and tECK6.

3.3.4. Discussion

In this chapter, we validated the purity of the commercial protein and conducted aptamer sequencing and identified the top candidates through analysis of the BLI results. For the DNA purity experiment, we observed that the PCR product showed a size greater than anticipated. Several factors could account for this deviation: Foremost, the ladder used might have been expired, casting doubt on its size representation accuracy; this is bolstered by the consistent positioning of all three PCR products at an unexpectedly large size. Additionally, non-specific amplification could have led to the inadvertent amplification of unintended sequences. Furthermore, there's the possibility of an error in primer design, wherein the primers, projected to be 55 nucleotides, may have been inaccurately designed for 70 nucleotides. The selection criteria for the aptamer sequences were primarily designed to negate the influence of PCR artifacts. Consequently, sequences from the final selection round with high abundance were de-emphasized if they were incorrectly structured. Such flawed sequences contained repetitive dimeric or trimeric sequences from the forward and reverse primer binding regions, positioned either preceding or following the randomized section, deviating from the intended structure of the initial N40 DNA library. Therefore, it is still not clear about the exact reasons caused the wrong size of the PCR product. To figure this out, a titration PCR is necessary for determining the optimal amplification round and minimize the effects from the artifact. Other than that, primers will be sent for NGS to obtain its sequence to confirm the sequence is correct.

All clones showed higher binding shift in the BLI binding experiment, which indicated a successful enrichment of the clones. ECK4 and ECK6 exhibit higher binding shift than other clones so they were selected for the downstream experiments.

For aptamer structure modification, we found that all clones showed better binding with the complementary primer. However, we can not assume that the primer-hybridized aptamers bind better than original ones, because the size of the molecule increased, which will lead to an increase of ligand layer which reflected as a bigger change on the light wavelength shift. Nonetheless, this is still strong evidence that the aptamers will exhibit binding even without primer regions. To further confirm the binding of those truncated aptamer, we need to measure the K_D value for them, which indicates the overall binding strength of those aptamers. The experiment of K_D measurement will be performed in chapter 4.

Chapter 4. Apparent K_D determination for truncated anti-N aptamers

4.1. Objective

Through the previous chapter's content, we have determined that structurally modified aptamers can bind to N protein. But to better evaluate aptamers, we need to measure the K_D of tECK4 and tECK6. The Apparent K_D value, the dissociation constant, represents the target molecule concentration required to occupy half of the available binding sites on the aptamer. Understanding the Apparent K_D value of an aptamer is crucial for researchers as it offers valuable insights into the strength of the interaction between the aptamer and its target molecule.

4.2. Method A: Biolayer interferometry

In this section, BLI was the primary experiment method to use for Apparent K_D determination; BLI is a label-free technique that measures the binding affinity by immobilizing the aptamer on the surface of a biosensor and then measuring the changes in the interference pattern of light caused by the binding of the target molecule to the aptamer. These changes in the interference pattern are directly proportional to the binding affinity between

the aptamer and the target molecule and can be used to calculate the Apparent K_D value.

4.2.1. BLI experiment method

The binding affinity of tECK4 and tECK6 aptamers and SARS-CoV-2 N protein was performed using Octet N1 (Sartorius, Bohemia, NY). To determine the dissociation constant (K_D), the biotinylated aptamers were immobilized on streptavidin biosensor tips. The serial dilutions of the N protein (7.69, 15.38, 30.77, 61.54 and 123.1 nM) were prepared in the optimized assay buffer (PBS, with 0.5mM $MgCl_2$ and 1mM $CaCl_2$, pH 7.4, added 0.2 % BSA and 0.05 % Tween 20). The binding affinity was performed by dipping a streptavidin biosensor (Sartorius, Bohemia, NY), which was initially immobilized with the biotinylated aptamer for 60 s into the N protein spiked PBST buffer. The assay was carried out at room temperature with 180 s for association and 120 s for dissociation. The assay buffer was used to wash out the unbound biomolecules from the biosensor tip during the baseline step for 60 s. The Apparent K_D of each aptamer was calculated by Octet N1 software (version 1.3.0.5) using a 1:1 global binding model. The equation used for measuring K_D is $\mathbf{K_D = k_{off} / k_{on}}$, k_{on} (Association rate constant) represents the rate at which two molecules bind and k_{off} (Dissociation rate constant) represents the rate at which the bound molecules come apart. Those kinetic values were automatically generated by the software.

4.2.2. Results

To determine the binding affinity of the aptamer to N protein by BLI, the biotinylated aptamer was used as a ligand on the streptavidin biosensor to validate its interaction with the N protein. As shown in **Figure 4.1 (A), (B), and (C)**, the binding at different concentrations of the N protein readily fits a single exponential (1:1) model. The Apparent K_D values of tECK4 and tECK6 aptamers were calculated as 8 ± 1 and 5.42 ± 0.04 nM, shown in **Table 4.1**.

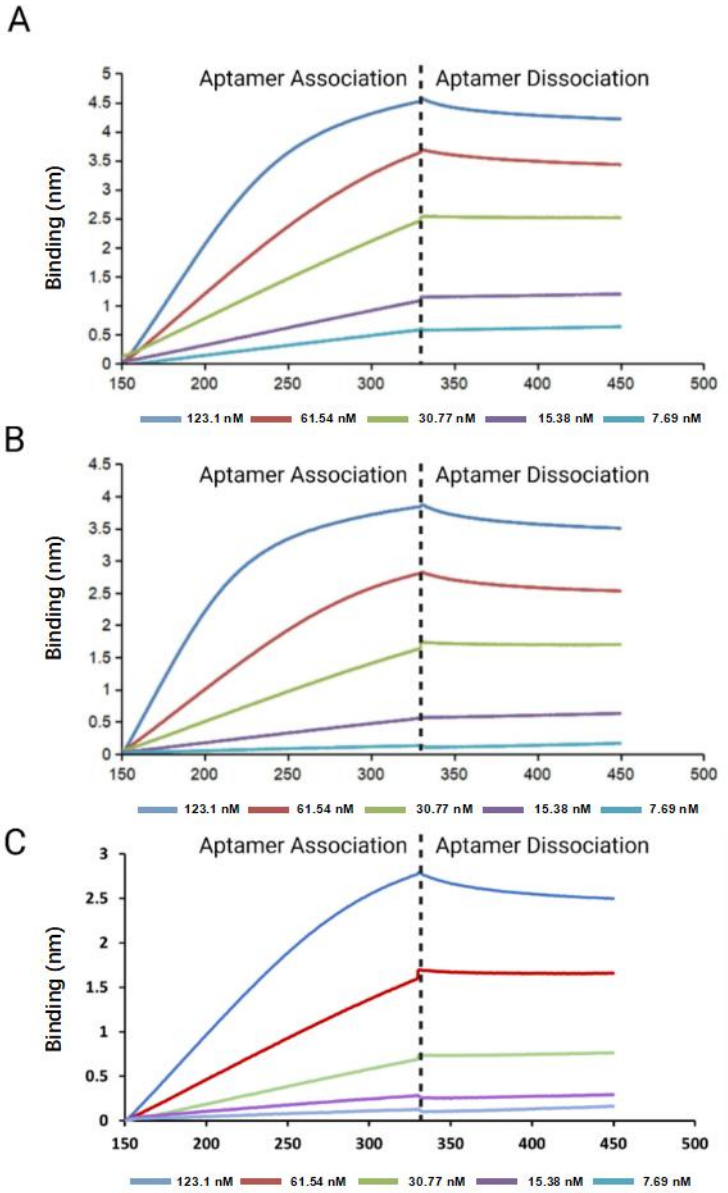


Figure 4.1. The binding affinity analysis of tECK4 and tECK6 aptamers to SARS-CoV-2 N protein. The BLI analysis of the tECK4 (A), tECK6 (B), and 40 nt non-aptamer DNA (C) aptamers binding to N protein. The concentrations of the SARS-CoV-2 N protein were 7.69, 15.38, 30.77, 61.54 and 123.1 nM, respectively, from the bottom to the top.

Table 4.1 Binding affinity of tECK4, tECK6, and 40 nt non-aptamer DNA to SARS-CoV-2 N protein. The experiment was done with 3 technical replicates, and quantification data are shown as means ± SDs.

Aptamer	Apparent K_D (nM)	R^2	k_{on} (1/Ms)	k_{off} (1/s)
tECK4	8 ± 1	0.997	6.42×10^4	4.77×10^{-4}
tECK6	5.42 ± 0.04	0.996	1.03×10^5	5.59×10^{-4}
40 nt non-aptamer	74 ± 2	0.987	3.11×10^3	2.30×10^{-3}

4.3. Method B: Capillary Electrophoresis

4.3.1. CE experiment method

A solution containing a fixed concentration of 15 nM FAM-labelled aptamer, and various concentrations of N protein was prepared using an incubation buffer of 5 mM ammonium acetate. The protein concentration was systematically varied in the range of 0-300 nM, with increments of 25 nM up to 100 nM and 50 nM after that. The samples were placed in vials coated with 2 % BSA for 30 minutes. The aptamer solution was snap-cooled before adding the desired amount of N protein. The mixture was incubated at room temperature for 10 minutes, and the binding was measured using CE.

4.3.2. Apparent K_D calculation

The peak areas of the free aptamer peak were calculated using Origin (MicroCal Inc, Northampton, Massachusetts), and the graph was created that indicated the percentage of aptamer bound against protein concentration by GraphPad (GraphPad Software Inc, California). The calculation of Apparent K_D is based on the equation in **Figure 4.2**.

$$K_D = \frac{[T][L]}{[TL]}$$

$$a = \frac{[LT]}{[L]+[LT]}$$

$$\rightarrow K_D = \left(\frac{1}{a/100\%} - 1\right) \cdot [T]$$

a: Percentage of Binding

[L]: Concentration of Free Aptamer (Ligand)

[T]: Concentration of Protein (Target)

[LT]: Concentration of Aptamer-Protein complex

Figure 4.2. Equations used to determine K_D in CE-K_D measurement.

4.3.3 Results and Discussion

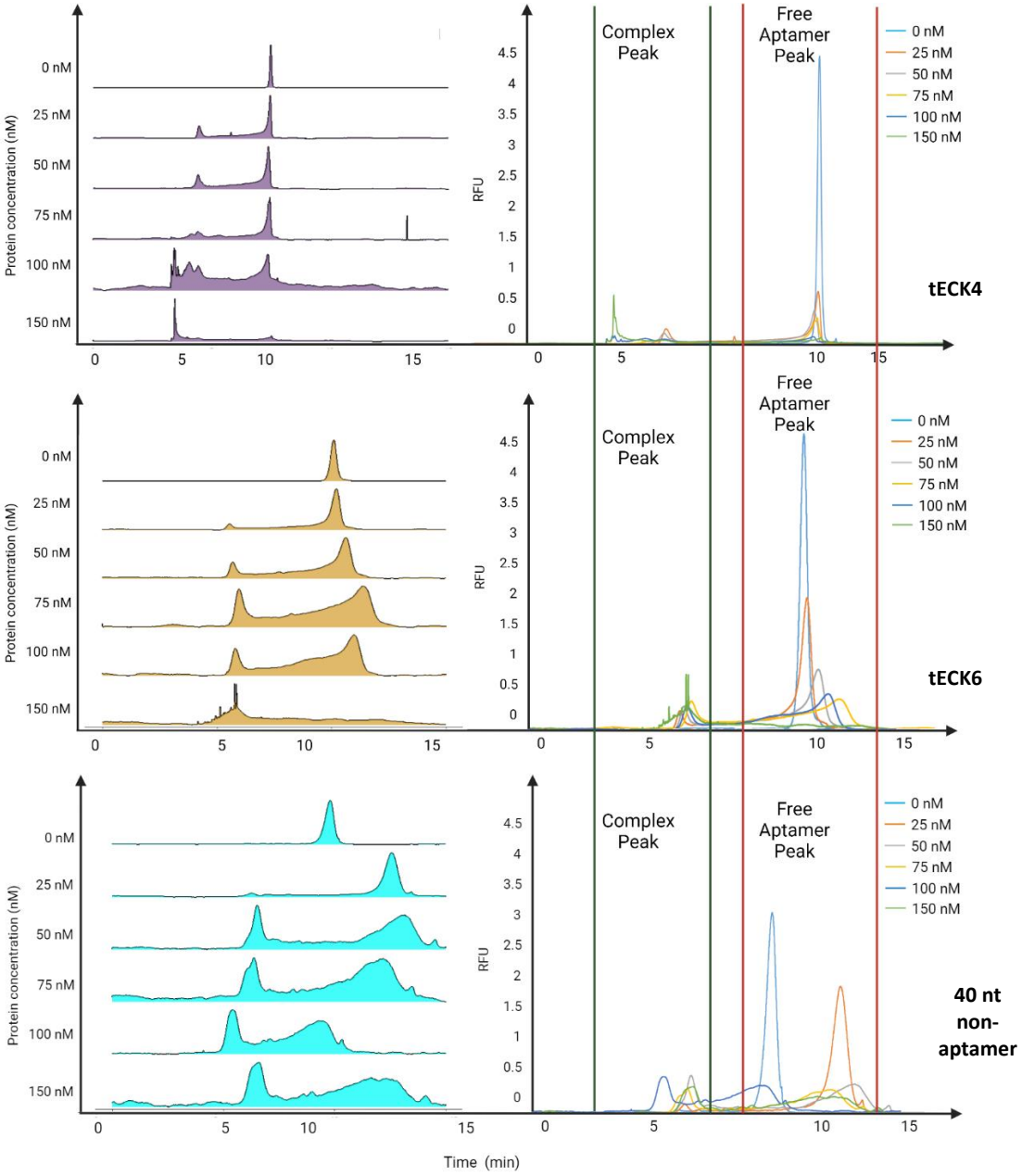


Figure 4.3. Representative CE electropherograms for Apparent K_D determination. Each graph represents a specific protein concentration and is presented on the left side. The experimental setup involved incubating 15nM tECK4, 40 nt non-aptamer, and tECK6 aptamers with varying concentrations of N protein (0 nM, 25 nM, 50 nM, 75 nM, 100 nM, and 150 nM). The electropherograms were stacked and normalized to the accessible aptamer peak from 0 nM N protein to highlight the differences in the peak scale of free aptamers.

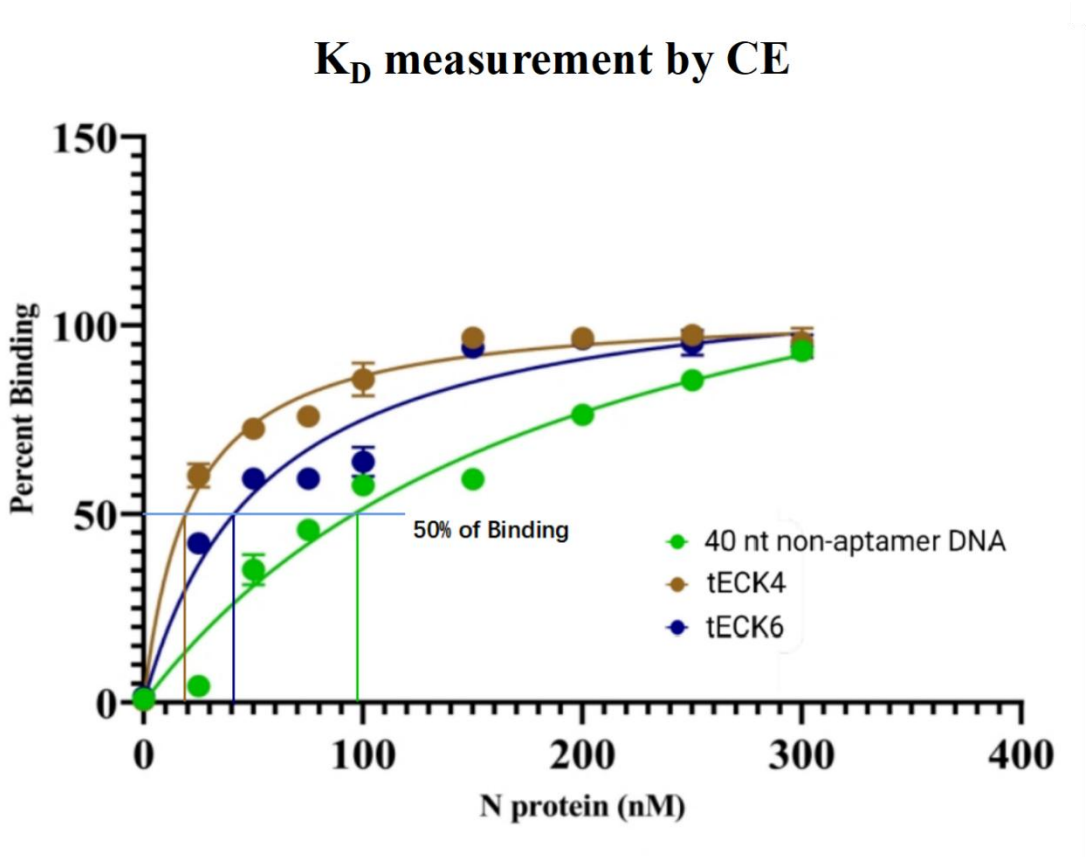


Figure 4.4. Apparent K_D determination by CE. Binding curve of tECK4, tECK6 aptamers and 40 nt non-aptamer DNA (a negative control) with 5 concentrations of SARS-CoV-2 N protein.

The Apparent K_D value, representing the affinity between an aptamer and its target protein (N protein), was determined by creating a binding curve, plotting the fraction of bound aptamer against the N protein concentration. As more aptamer-N protein complexes formed, the peak corresponding to the free aptamer decreased, allowing quantification of complex formation by measuring the change in free aptamer peak area. This approach facilitated the calculation of the Apparent K_D value, which was taken at the point where 50 % of the aptamer was bound. In Figure 4.3, the Apparent K_D values for tECK4 and tECK6 were found to be 19 ± 4 nM and 41 ± 7 nM, respectively, indicating their higher affinity for N protein compared to a 40 nt non-aptamer DNA with an Apparent K_D value of 99 ± 6 nM **(Table 4.2)**. Notably, tECK4 exhibited stronger binding to the protein than tECK6, consistent with the results obtained from BLI experiments. It's worth mentioning that maximum binding was observed with as little as 150 nM N protein for tECK4 and tECK6, while the 40 nt non-aptamer DNA required up to 300 nM N protein to achieve a similar level of binding, further emphasizing the enhanced binding affinity of the aptamers.

Table 4.2. Apparent K_D of tECK4, tECK6, and 40 nt non-aptamer DNA to SARS-CoV-2 N protein, measured using CE-LIF and calculated from the binding curves by GraphPad. The experiment was done with 3 technical replicates, and quantification data are shown as means ± SDs.

Clones	Apparent K_D, nM	R^2
tECK4	19 ± 4	0.9839
tECK6	41 ± 7	0.9521
40 nt non-aptamer	99 ± 6	0.9057

4.4. Discussion

4.4.1. CE-SELEX Efficiency

In this study, we utilized capillary electrophoresis (CE) for the SELEX process of aptamers targeting the N protein. Typically, it takes only 1–4 rounds of screening in CE-SELEX to obtain aptamers with high specificity and high affinity from a random oligonucleotide library with a capacity of 10^{13}–10^{15} molecules[76-79]. The process was limited to three rounds, and a marked increase in the binding affinity of enriched aptamer pools to the N protein was observed. Two of the aptamers demonstrated similar or even lower K_D compared to validated clones. These findings suggest that three rounds might be sufficient for pool evolution and aptamer selection specific to the target. Nevertheless, it remains uncertain if additional rounds might have enhanced the binding further.

For comparison, clone tNSP3 (selected through 6 rounds of asymmetric PCR-SELEX) and A48 (selected through 6 rounds beads-SELEX) were used as positive controls[80, 81]. Both controls, as documented in the literature, target the N protein of SARS-CoV-2, using different SELEX methodologies. These methodologies aim to enhance the traditional SELEX process efficiency and specificity and the buffer used in the SELEX process are all PBS + 0.05 % Tween 20, which is different with ammonium acetate.

Table 4.3. K_Ds of CE-SELEX aptamers (tECK4, t ECK6), Strip-SELEX (tNSP3), and Beads-SELEX (A48) binding to SARS-CoV-2 N protein determined using BLI.

Aptamer	Apparent K_D (nM)	R^2
tNSP3（Lab+literature）	3 ± 1	0.90
A48（Literature）	7 ± 1	0.97
tECK4 (Lab）	8 ± 1	0.98
tECK6 (Lab)	5.42 ± 0.04	0.95
40 nt non-aptamer DNA	74 ± 2	0.91

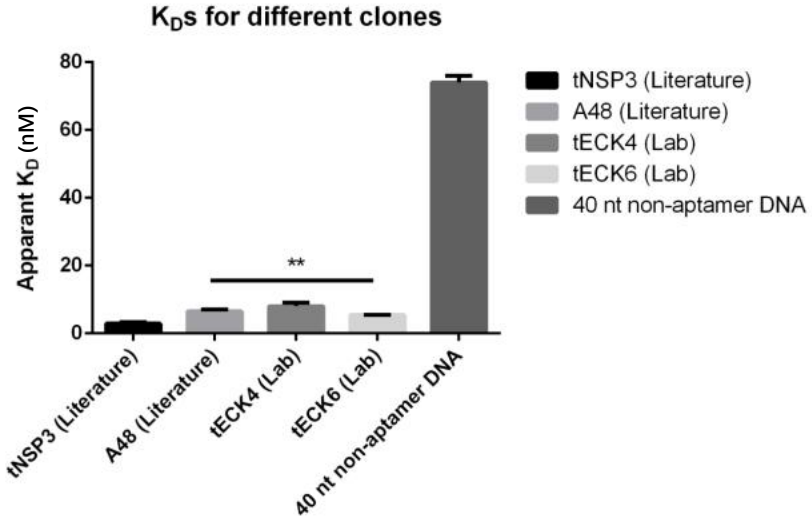


Figure 4.5. Apparent K_Ds for clones. Two aptamers from literature targeting N protein were selected to compare with the clones selected by CE-SELEX. After three independent experiments, quantification data are shown as means ± SDs，**p< 0.01. Student's t test.

A BLI-K_D determination experiment was conducted using the same conditions as the method A depicted in Chapter 4 for all aptamers, ensuring the K_D values are comparable. Results indicate that tECK6 possesses a lower K_D than A48, and tECK4 has no significant difference with A48. Those results suggest that our selected aptamers have similar binding efficiency with documented aptamers. Although tNSP3 and A48 were not derived from CE-SELEX, the data implies that aptamers from CE-SELEX exhibit promising potential. Comparing the

selection rounds between different SELEX methods can be challenging, particularly when the methodologies differ significantly from CE-SELEX. Even if both approaches utilized six rounds of selection, the efficacy of the selection process is heavily influenced by the stringency of the selection conditions, encompassing factors such as target concentration, washing procedures, and elution conditions. These variables can have a profound impact on the rate of aptamer enrichment, thus potentially leading to variations in the number of rounds required. However, it's important to note that despite these differences, the aptamers generated through these distinct methods have been validated for their strong binding affinity to the N protein, as evidenced by both existing literature data and our own experimental results using BLI. Six rounds of SELEX were carried out to identify both A48 and tNSP3. Nevertheless, comparing the number of rounds between these SELEX methods is challenging due to their execution under different mechanisms.

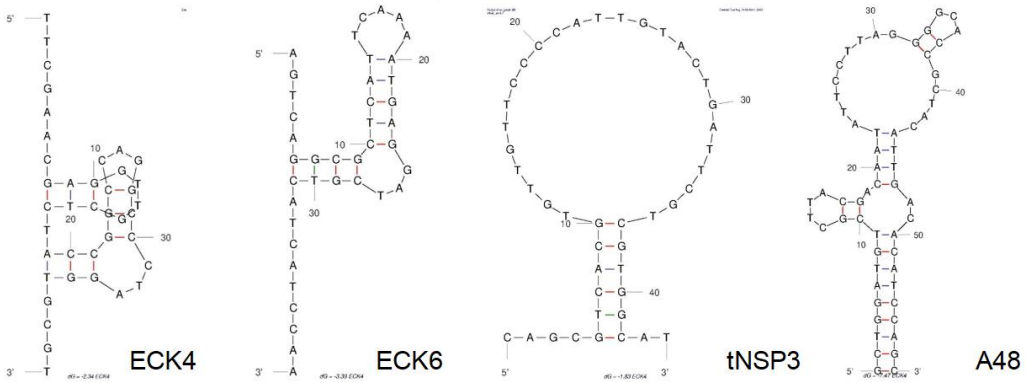


Figure 4.5. Predicted secondary structures for all aptamers.

Along with the K_D measurement experiment, the aptamer structures were predicted by MFOLD web server (http://www.unafold.org/) for a preliminary analysis[82]. From the prediction, we observe no similarity. Based on the overall result, we assume that the CE-SELEX successfully isolate aptamer which showed high binding affinity and with unique structures.

4.4.2. Conditions in K_D measurement and CE-SELEX

4.4.2.1. Buffers in K_D measurement and CE-SELEX

In this research, tECK4 and tECK6 are selected under a buffer condition of (10 mM ammonium acetate with pH 9.2), however the K_D was measured under the buffer of PBS. PBS is commonly used because it maintains a stable pH and mimics the ionic strength of the human body, which is essential for maintaining the natural conformation of many proteins[83]. On the other hand, 10 mM ammonium acetate stands out as a proper buffer choice for CE-SELEX because its low ionic strength can enhance the differential migration of aptamer species, whether they are bound to the target or not, ensuring distinct separations during the process. Crucially, the volatility of ammonium acetate is pivotal when transitioning to mass spectrometry (MS) analyse[84, 85]. The difference in ionic strength and composition between the two buffers can influence aptamer-protein binding as well. High ionic strength solutions like PBS can potentially shield or mitigate electrostatic interactions, which might be crucial for certain aptamer-protein interactions. In contrast, a low ionic strength environment provided by 10 mM ammonium acetate can enhance these interactions. Therefore, the choice of buffer can significantly impact the observed binding affinities and should be carefully considered based on the specific properties of the aptamer and protein being studied.

4.4.2.2. Immobilization methods in K_D measurements and CE-SELEX

In K_D measurement experiment, the difference in immobilization methods between BLI and CE can affect the measured binding affinity of aptamer-protein interactions. BLI involves surface immobilization of the capture molecule, which can impact its orientation and stability, potentially influencing the binding affinity[86]. CE, on the other hand, does not involve immobilization and eliminates this factor. Instead, the aptamer and protein are allowed to interact in a homogeneous solution, and their interactions are detected based on their electrophoretic mobility and separation within the capillary. The absence of immobilization in CE eliminates the potential influence of the mass transfer of the ligand to a surface, steric hindrance on biomolecular interactions. As a result, a discrepancy in apparent K_D value is expected in the BLI and CE experiments[87]. Despite these differences, the trend in both experiments is consistent, with the selected aptamers demonstrating better binding ability than 40 nt non-aptamer DNA. Notably, the concentration of protein that was used for CE and BLI experiments are different, so the comparison may not be accurate in this regard.

In experiments, the aptamer demonstrated interaction with the N protein in a 10 mM ammonium acetate environment. Theoretically, under the conditions stipulated by the original SELEX method, particularly with regards to binding buffer composition and immobilization methodology, this should suggest a superior binding affinity, manifesting as a lower K_D value. However, when comparing the K_D values derived from both BLI and CE methods, the BLI-derived values were not only lower but were also in line with results from established positive controls. This implies that the binding affinity between the selected aptamer and the protein was not significantly impacted by factors such as buffer composition and immobilization technique. Such robustness in binding suggests that the aptamer could be versatile across diverse application environments, even those underpinned by different principles. A practical implication of this could be the simplification of the preparation process in serological antigen tests, eliminating the need to modify the molecular environment of blood samples.

In summary, the comparable K_D with documented aptamer clones indicated the potential of diagnostic tool development of tECK4 and tECK6, and with these two aptamers selected out and the binding is validated that the K_Ds are higher than the negative control, the efficiency of 3 rounds CE-SELEX has been proved.

Limitation and Future Direction

Limitation:

In our conducted research, aptamers were selected within a 10 mM ammonium acetate buffer. This selection process has raised concerns about potential disparities when applying these aptamers in actual diagnostic settings. Specifically, the use of a 10 mM ammonium acetate buffer without the inclusion of detergents, stabilizers, or conventional physiological buffers like PBS presents challenges for real-world COVID-19 detection[88-90]. Additionally, it's worth noting that the protein utilized in this research was procured from a commercial source and originated from *E. coli* expression. This introduces a significant difference between the target used in our research and the actual target in the real world. For a detailed discussion on this aspect, please refer to Chapter 2.

The N protein, which exhibits a natural affinity for RNA, is not typically known for binding to single-stranded DNA (ssDNA) with high specificity. While there may be some degree of non-specific interaction between the N protein and ssDNA due to their shared electrostatic properties, the primary function of the N protein is to interact with viral RNA, particularly the viral genomic RNA. Its structural and functional features are tailored for RNA binding, including conserved domains like the N-terminal domain (NTD) and the C-terminal domain (CTD), which are optimized for RNA recognition. This possible interaction between the N protein and ssDNA introduces a background signal that can complicate the isolation of specific aptamers. During the SELEX process, where a diverse ssDNA library is subjected to

multiple rounds of selection and amplification, non-specific binding of the N protein to the ssDNA library can lead to the enrichment of sequences that bind promiscuously to ssDNA rather than specifically to the protein of interest. This phenomenon can result in a higher likelihood of obtaining false-positive sequences that do not possess genuine binding affinity to the N protein. For the development of real aptamer-based diagnostic tool, the natural binding ability can lead to the recognition of unrelated DNA sequences, potentially resulting in incorrect diagnoses.

Furthermore, the unique secondary structures of these aptamers suggest the potential for binding to different epitopes, opening the door for the development of an enhanced N protein detection method. However, it is essential to note that the confirmation of whether the epitopes bound by different aptamers are indeed distinct remains pending. It's also crucial to consider the possibility that these epitopes may include the His-tag that accompanies the protein during the *E. coli* protein expression process which is not a part of the protein that expressed in a real virus.

Discussion and Future Direction:

Addressing these limitations and advancing the research opens up several promising avenues for enhancing the adaptability and effectiveness of these aptamers in practical diagnostic applications. One significant consideration involves incorporating buffers that resemble PBS, enriched with detergents and stabilizers, into the CE-SELEX process. This approach holds the potential to provide a dual benefit by improving the performance of these aptamers. The obstacle for accomplishing this is that the electrophoresis may not be able to carried out properly in the buffer of PBS or any other buffers with a high salt concentration. And the introduction of detergent can also be problematic for the electrophoresis.

Moreover, in our effort to validate the aptamer's binding capabilities and its suitability for real-world diagnostic tools, we turned to BLI as part of our research. BLI, much like Lateral Flow Assays (LFA), relies on a solid phase for immobilizing specific molecules. The buffer used for determining K_D values in BLI comprises PBS to ensure isotonicity, supplemented with 0.5 mM $MgCl_2$, 1 mM $CaCl_2$, 0.2 % BSA as a stabilizing agent, and 0.05 % Tween 20 to simulate the presence of detergents for lysis purposes. While this composition may not perfectly replicate real diagnostic tools, it offers a more practical buffer composition compared to what was employed in capillary electrophoresis (CE).

The comparable K_D values obtained through BLI, in conjunction with those reported in the literature, provide further validation for the feasibility of using these aptamers in future diagnostic tools aimed at enhancing virus detection. This suggests that the aptamers exhibit adaptability across a range of application environments.

The robust binding performance underscores the aptamers' potential adaptability across diverse application contexts, even those governed by different principles. A practical implication of this adaptability could be the streamlining of the preparation process for serological antigen tests, potentially eliminating the need for extensive modifications to the molecular environment of blood samples. However, it remains crucial to validate this binding capability in more challenging conditions, such as human saliva and blood.

To mitigate the potential for natural binding interference, a comprehensive analysis of binding epitopes is essential. Mass spectrometry analysis will be employed, involving the immobilization of the N protein with the aptamer, subsequent protein digestion, and elution of the retained epitope on the aptamer. The acquired sequence will then undergo mass spectrometry analysis. Additionally, Molecular Docking Simulations will be conducted to gain deeper insights into the binding interactions between the aptamer and the N protein. These simulations will be performed for both aptamers and a 40 nt non-aptamer DNA

sequence to ensure distinctive binding patterns.

Veverka and his research team investigated the mechanisms by which the N protein links the viral genome to the viral membrane. They found that the N-terminal RNA binding domain (N-NTD) is essential for capturing the RNA genome, while an intrinsically disordered region anchors the ribonucleoprotein complex to the viral membrane. In their analysis of the NTD (***PDB ID: 6YI3***), they identified a positively charged cleft serving as a potential RNA binding site (**Figure S4**). Through computer simulations guided by NMR data, they constructed atomic models of N-NTD in complex with both single-stranded and double-stranded RNA, leading to the identification of two specific binding complexes (***PDB IDs: 7ACS and 7ACT***). Molecular docking revealed that RNA duplexes and ssRNA bind similarly to the positively charged canyon in N-NTD, both 7mer and 10mer ssRNA affected the same N-NTD residues. The significantly perturbed residues (L56, G60, K61, K65, F66, A90, R93, I94, R95, K102, D103, L104, T165, T166, G175 and R177) formed a U-shaped binding epitope on the N-NTD surface circumventing the base of the positively charged finger (**Figure S5**). These findings will serve as the basis for comparisons with the tECK4 and tECK6 models, as well as for analyzing epitope sequences obtained from mass spectrometry in relation to NTD binding sites. Furthermore, the molecular docking model will be compared to the one generated in this research to assess whether the nucleotides occupy the same binding domain and if the same amino acid residues are involved. We anticipate discovering additional binding sites beyond just the NTD domain, and we acknowledge that the binding mechanism may not solely depend on electrostatic forces but may also involve more conformational changes[91].
(The amino acid sequence of N-terminal RNA binding domain of N protein:
GAMGLPNNTASWFTALTQHGKEDLKFPRGQGVPINTNSSPDDQIGYYRRATRRIRGG
DGKMKDLSPRWYFYYLGTGPEAGLPYGANKDGIIWVATEGALNTPKDHIGTRNPAN
NAAIVLQLPQGTTLPKGFYAEGSRGGS).

Utilizing these aptamers in tandem can provide deeper insights into the structural and functional aspects of the N protein, enabling comprehensive protein characterizations and in-depth studies. The presence of a unique aptamer suggests that the iterative rounds of selection and amplification in SELEX have successfully yielded molecules with distinctive binding properties while maintaining a similar binding affinity.

Conclusion

This thesis primarily focuses on the development of single-stranded DNA aptamers designed to capture the SARS-CoV-2 Nucleocapsid protein, with the potential application of creating a label-free biosensor for COVID-19 diagnosis.

Chapter 2: In this chapter, we successfully selected novel DNA aptamers targeting the SARS-CoV-2 Nucleocapsid protein. This selection process utilized CE-SELEX, and we employed FASTAptamer to identify DNA sequences with the highest enrichment. Through NGS data analysis, we identified six aptamer candidates that specifically target the N protein. These candidates met the criteria for aptamer selection and underwent pre-validation through BLI experiments. Among these candidates, ECK4 and ECK6 exhibited the most significant binding shifts, indicating their robust binding capabilities compared to others.

Chapter 3: In this chapter, we focused on aptamer truncation. Prior to truncation, we conducted a pre-truncation experiment to confirm the binding of truncated aptamers. Our rationale was to utilize complementary primer regions to cover the original aptamer's primer regions, preventing the formation of secondary structures. BLI experiments confirmed the binding of these primer-hybridized aptamers.

Chapter 4: This chapter delves into the determination of K_D values using both CE and BLI

K_D measurement experiments. Since CE was employed in the aptamer selection process, its results were expected to be more reflective. BLI, on the other hand, closely resembles real-world virus detection methods like lateral flow assays. Interestingly, the CE results indicated higher K_D values across all tested aptamers compared to BLI. This difference may be attributed to variations in the experimental conditions for aptamer binding and the methods used for immobilization.

While there are differences in K_D values obtained via CE and BLI measurements, it is important to highlight that aptamers selected through CE-SELEX exhibit superior binding performance compared to 40-nt non-aptamer DNA sequences that have not undergone SELEX procedures. This indicates the effectiveness of the CE-SELEX process with only 3 rounds of enrichment. Additionally, it's worth noting that all aptamers selected through CE-SELEX demonstrate K_D values that closely match those of documented aptamers designed for detection purposes and validated with actual virus particles. This strong resemblance suggests that the aptamers developed in this study possess a binding capacity comparable to established counterparts, warranting further exploration of their potential applications in practical diagnostic tools.

For the future work, one significant consideration involves incorporating buffers resembling PBS, enriched with detergents and stabilizers, into the CE-SELEX process, potentially improving aptamer performance. However, challenges related to electrophoresis in high-salt buffers and detergent introduction need to be addressed. Moreover, further validation in challenging conditions such as human saliva and blood is crucial. To mitigate potential binding interference, a comprehensive analysis of binding epitopes is essential, employing mass spectrometry analysis and Molecular Docking Simulations, with comparisons to non-aptamer DNA sequences. Exploring additional binding sites beyond the N-terminal RNA binding domain (NTD) and understanding potential conformational changes in the binding mechanism are also vital. Finally, utilizing multiple aptamers in tandem can provide deeper insights into the structural and functional aspects of the N protein, enabling comprehensive protein characterizations while maintaining similar binding affinities.

Reference

[1] Hasoksuz M, Kilic S, Sarac F. Coronaviruses and SARS-COV-2[J]. Turk J Med Sci, 2020,50(SI-1):549-556.

[2] Salzberger B, Buder F, Lampl B, et al. Epidemiology of SARS-CoV-2[J]. Infection, 2021,49(2):233-239.

[3] Ravi V, Saxena S, Panda P S. Basic virology of SARS-CoV 2[J]. Indian J Med Microbiol, 2022,40(2):182-186.

[4] SARS-CoV-2 and its infected world[J]. J Mol Cell Biol, 2020,12(12):913.

[5] Actis G C, Pellicano R, Ribaldone D G. SARS-CoV-2: Impacts and Echoes[J]. Rev Recent Clin Trials, 2022,17(2):69-72.

[6] Chu H, Yuen K Y. Pathogenicity of SARS-CoV-2 Omicron[J]. Clin Transl Med, 2022,12(5):e880.

[7] Wiggen T D, Bohn B, Ulrich A K, et al. SARS-CoV-2 seroprevalence among healthcare workers[J]. PLoS One, 2022,17(4):e266410.

[8] Creech C B, Walker S C, Samuels R J. SARS-CoV-2 Vaccines[J]. JAMA, 2021,325(13):1318-1320.

[9] Krammer F. SARS-CoV-2 vaccines in development[J]. Nature, 2020,586(7830):516-527.

[10] Dong Y, Dai T, Wei Y, et al. A systematic review of SARS-CoV-2 vaccine candidates[J]. Signal Transduct Target Ther, 2020,5(1):237.

[11] Dong Y, Dai T, Wang B, et al. The way of SARS-CoV-2 vaccine development: success and challenges[J]. Signal Transduct Target Ther, 2021,6(1):387.

[12] Joe C, Jiang J, Linke T, et al. Manufacturing a chimpanzee adenovirus-vectored SARS-CoV-2 vaccine to meet global needs[J]. Biotechnol Bioeng, 2022,119(1):48-58.

[13] Pinana J L, Lopez-Corral L, Martino R, et al. SARS-CoV-2 vaccine response and breakthrough infection rate in patients with hematological disorders[J]. J Hematol Oncol, 2022,15(1):54.

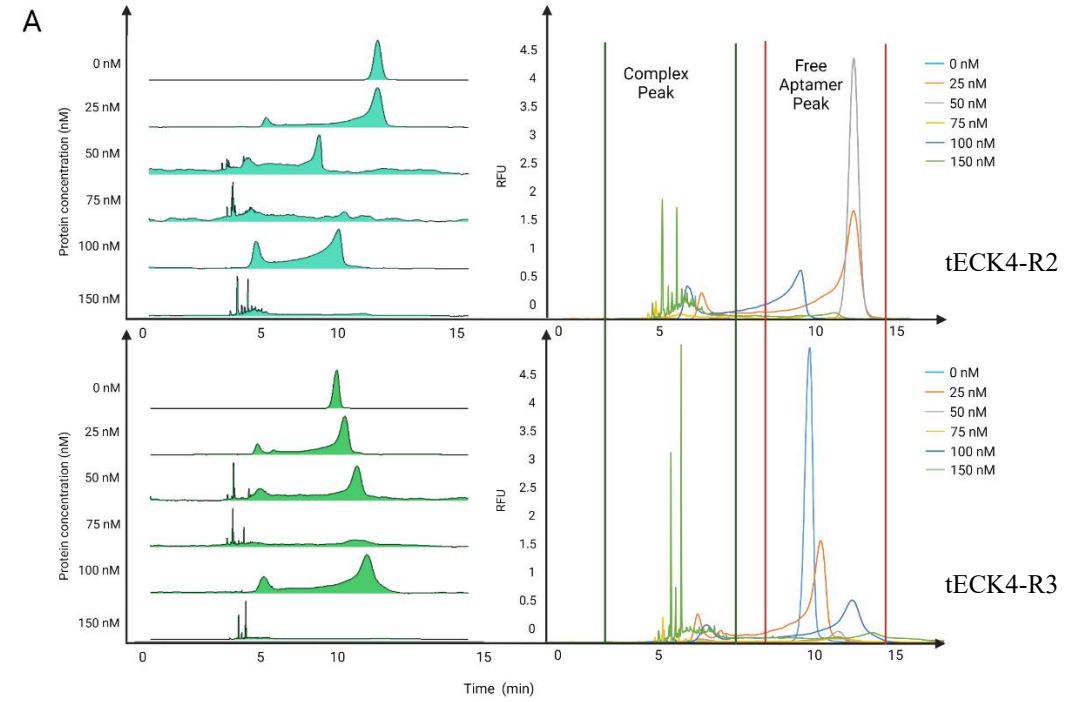
[14] Parashar N C, Poddar J, Chakrabarti S, et al. Repurposing of SARS-CoV nucleocapsid protein specific nuclease resistant RNA aptamer for therapeutics against SARS-CoV-2[J]. Infect Genet

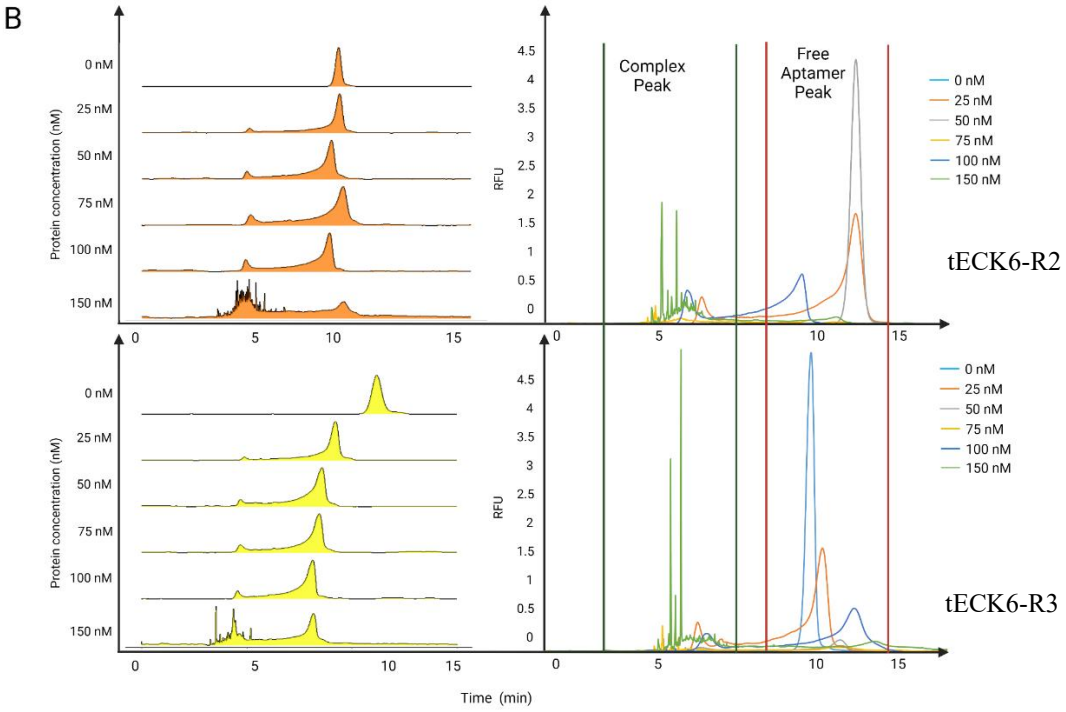
Evol, 2020,85:104497.
[15] Zeng W, Liu G, Ma H, et al. Biochemical characterization of SARS-CoV-2 nucleocapsid protein[J]. Biochem Biophys Res Commun, 2020,527(3):618-623.
[16] Bai Z, Cao Y, Liu W, et al. The SARS-CoV-2 Nucleocapsid Protein and Its Role in Viral Structure, Biological Functions, and a Potential Target for Drug or Vaccine Mitigation[J]. Viruses, 2021,13(6).
[17] Yao H, Song Y, Chen Y, et al. Molecular Architecture of the SARS-CoV-2 Virus[J]. Cell, 2020,183(3):730-738.
[18] Park G J, Osinski A, Hernandez G, et al. The mechanism of RNA capping by SARS-CoV-2[J]. Nature, 2022,609(7928):793-800.
[19] Sette A, Crotty S. Adaptive immunity to SARS-CoV-2 and COVID-19[J]. Cell, 2021,184(4):861-880.
[20] Castro D X, Ols S, Lore K, et al. Immunity to SARS-CoV-2 induced by infection or vaccination[J]. J Intern Med, 2022,291(1):32-50.
[21] Peng Y, Du N, Lei Y, et al. Structures of the SARS-CoV-2 nucleocapsid and their perspectives for drug design[J]. EMBO J, 2020,39(20):e105938.
[22] Cai T, Yu Z, Wang Z, et al. Arginine methylation of SARS-Cov-2 nucleocapsid protein regulates RNA binding, its ability to suppress stress granule formation, and viral replication[J]. J Biol Chem, 2021,297(1):100821.
[23] Wu W, Cheng Y, Zhou H, et al. The SARS-CoV-2 nucleocapsid protein: its role in the viral life cycle, structure and functions, and use as a potential target in the development of vaccines and diagnostics[J]. Virol J, 2023,20(1):6.
[24] Lee E, Sandgren K, Duette G, et al. Identification of SARS-CoV-2 Nucleocapsid and Spike T-Cell Epitopes for Assessing T-Cell Immunity[J]. J Virol, 2021,95(6).
[25] Zhao Y, Sui L, Wu P, et al. A dual-role of SARS-CoV-2 nucleocapsid protein in regulating innate immune response[J]. Signal Transduct Target Ther, 2021,6(1):331.
[26] Kang S, Yang M, Hong Z, et al. Crystal structure of SARS-CoV-2 nucleocapsid protein RNA binding domain reveals potential unique drug targeting sites[J]. Acta Pharm Sin B, 2020,10(7):1228-1238.
[27] Lu Y, Michel H A, Wang P H, et al. Manipulation of innate immune signalling pathways by SARS-CoV-2 non-structural proteins[J]. Front Microbiol, 2022,13:1027015.
[28] Thorne L G, Bouhaddou M, Reuschl A K, et al. Evolution of enhanced innate immune evasion by SARS-CoV-2[J]. Nature, 2022,602(7897):487-495.
[29] Rashid F, Xie Z, Suleman M, et al. Roles and functions of SARS-CoV-2 proteins in host immune evasion[J]. Front Immunol, 2022,13:940756.
[30] Silva E, Bomfim C G, Barbosa A P, et al. Immunization with SARS-CoV-2 Nucleocapsid protein triggers a pulmonary immune response in rats[J]. PLoS One, 2022,17(5):e268434.
[31] Cerna K, Duricova D, Hindos M, et al. Cellular and Humoral Immune Responses to SARS-CoV-2 Vaccination in Inflammatory Bowel Disease Patients[J]. J Crohns Colitis, 2022,16(9):1347-1353.
[32] Guo L, Ren L, Yang S, et al. Profiling Early Humoral Response to Diagnose Novel Coronavirus Disease (COVID-19)[J]. Clin Infect Dis, 2020,71(15):778-785.
[33] Bourassa L, Perchetti G A, Phung Q, et al. A SARS-CoV-2 Nucleocapsid Variant that Affects Antigen Test Performance[J]. J Clin Virol, 2021,141:104900.
[34] Peng Y, Du N, Lei Y, et al. Structures of the SARS-CoV-2 nucleocapsid and their perspectives for drug design[J]. EMBO J, 2020,39(20):e105938.
[35] Wang S, Dai T, Qin Z, et al. Targeting liquid-liquid phase separation of SARS-CoV-2 nucleocapsid protein promotes innate antiviral immunity by elevating MAVS activity[J]. Nat Cell Biol, 2021,23(7):718-732.
[36] Li X, Xiong M, Deng Q, et al. The utility of SARS-CoV-2 nucleocapsid protein in laboratory diagnosis[J]. J Clin Lab Anal, 2022,36(7):e24534.
[37] Wu J, Liu X, Zhou D, et al. Identification of RT-PCR-Negative Asymptomatic COVID-19 Patients via Serological Testing[J]. Front Public Health, 2020,8:267.
[38] Sule W F, Oluwayelu D O. Real-time RT-PCR for COVID-19 diagnosis: challenges and prospects[J]. Pan Afr Med J, 2020,35(Suppl 2):121.
[39] Yamayoshi S, Sakai-Tagawa Y, Koga M, et al. Comparison of Rapid Antigen Tests for COVID-19[J]. Viruses, 2020,12(12).
[40] Khalid M F, Selvam K, Jeffry A, et al. Performance of Rapid Antigen Tests for COVID-19 Diagnosis: A Systematic Review and Meta-Analysis[J]. Diagnostics (Basel), 2022,12(1).
[41] Lv Y, He J, Liu R, et al. China NMPA perspective on clinical evaluation of SARS-CoV-2 antibody test reagents in the process of emergency approval[J]. Bioanalysis, 2021,13(2):69-76.

[42] Cobb B L, Lloyd M, Hock KG, et al. Clinical Performance of a Lateral Flow SARS-CoV-2 Total Antibody Assay[J]. J Appl Lab Med, 2022,7(4):827-833.
[43] Grant B D, Anderson C E, Alonzo L F, et al. A SARS-CoV-2 coronavirus nucleocapsid protein antigen-detecting lateral flow assay[J]. PLoS One, 2021,16(11):e258819.
[44] Maglaras P, Lilis I, Paliogianni F, et al. A Molecular Lateral Flow Assay for SARS-CoV-2 Quantitative Detection[J]. Biosensors (Basel), 2022,12(11).
[45] Michel M, Bouam A, Edouard S, et al. Evaluating ELISA, Immunofluorescence, and Lateral Flow Assay for SARS-CoV-2 Serologic Assays[J]. Front Microbiol, 2020,11:597529.
[46] Xie X, Nielsen M C, Muruato A E, et al. A SARS-CoV-2 lateral flow assay was evaluated using the plaque reduction neutralization test[J]. Diagn Microbiol Infect Dis, 2021,99(2):115248.
[47] Recker A, White E M, Yang X, et al. Factors Affecting SARS-CoV-2 Test Discordance in Skilled Nursing Facilities[J]. J Am Med Dir Assoc, 2022,23(8):1279-1282.
[48] Tokuda Y, Shibuya K, Oguro K. Priority of SARS-CoV-2 test, trace, and isolation in Japan[J]. J Gen Fam Med, 2021,22(1):1-2.
[49] Phuong J, Hyland S L, Mooney S J, et al. Sociodemographic and clinical features predictive of SARS-CoV-2 test positivity across healthcare visit types [J]. PLoS One, 2021,16(10):e258339.
[50] Lin Y C, Chen W Y, Hwu E T, et al. In-Silico Selection of Aptamer Targeting SARS-CoV-2 Spike Protein[J]. Int J Mol Sci, 2022,23(10).
[51] Yang L F, Kacherovsky N, Panpradist N, et al. Aptamer Sandwich Lateral Flow Assay (AptaFlow) for Antibody-Free SARS-CoV-2 Detection[J]. Anal Chem, 2022,94(20):7278-7285.
[52] Schmitz A, Weber A, Bayin M, et al. A SARS-CoV-2 Spike Binding DNA Aptamer that Inhibits Pseudovirus Infection by an RBD-Independent Mechanism*[J]. Angew Chem Int Ed Engl, 2021,60(18):10279-10285.
[53] Sun M, Liu S, Wei X, et al. Aptamer Blocking Strategy Inhibits SARS-CoV-2 Virus Infection[J]. Angew Chem Int Ed Engl, 2021,60(18):10266-10272.
[54] Kohlberger M, Gadermaier G. SELEX: Critical factors and optimization strategies for successful aptamer selection[J]. Biotechnol Appl Biochem, 2022,69(5):1771-1792.
[55] Geiger M, Hogerton A L, Bowser M T. Capillary electrophoresis[J]. Anal Chem, 2012,84(2):577-596.
[56] Holland L. Capillary electrophoresis of biomolecules[J]. Anal Bioanal Chem, 2015,407(23):6909-6910.
[57] Yan X P, Yin X B, Jiang D Q, et al. Speciation of mercury by hydrostatically modified electroosmotic flow capillary electrophoresis coupled with volatile species generation atomic fluorescence pectrometry[J]. Anal Chem, 2003,75(7):1726-1732.
[58] Kang J, Wistuba D, Schurig V. Fast enantiomeric separation with vancomycin as chiral additive by co-electroosmotic flow capillary electrophoresis: increased detection sensitivity by the partial filling technique[J]. Electrophoresis, 2003,24(15):2674-2679.
[59] Terao K, Kondo S. AC-Electroosmosis-Assisted Surface Plasmon Resonance Sensing for Enhancing Protein Signals with a Simple Kretschmann Configuration[J]. Sensors (Basel), 2022,22(3).
[60] Petrov A, Okhonin V, Berezovski M, et al. Kinetic capillary electrophoresis (KCE): a conceptual platform for kinetic homogeneous affinity methods[J]. J Am Chem Soc, 2005,127(48):17104-17110.
[61] Krylov S N. Nonequilibrium capillary electrophoresis of equilibrium mixtures (NECEEM): A novel method for biomolecular screening[J]. J Biomol Screen, 2006,11(2):115-122.
[62] Kochmann S, Le ATH, Hili R, et al. Predicting efficiency of NECEEM-based partitioning of protein binders from nonbinders in DNA-encoded libraries[J]. Electrophoresis, 2018,39(23):2991-2996.
[63] Kanoatov M, Galievsky V A, Krylova S M, et al. Using nonequilibrium capillary electrophoresis of equilibrium mixtures (NECEEM) for simultaneous determination of concentration and equilibrium constant[J]. Anal Chem, 2015,87(5):3099-3106.
[64] Mosing R K, Bowser M T. Isolating aptamers using capillary electrophoresis-SELEX (CE-SELEX)[J]. Methods Mol Biol, 2009,535:33-43.
[65] Eaton R M, Shallcross J A, Mael L E, et al. Selecting DNA aptamers for ovarian cancer biomarker HE4 using CE-SELEX and high-throughput sequencing[J]. Anal Bioanal Chem, 2015,407(23):6965-6973.
[66] Sattari R, Palizban A, Khanahmad H. Single-Strand DNA-Like Oligonucleotide Aptamer Against Proprotein Convertase Subtilisin/Kexin 9 Using CE-SELEX: PCSK9 Targeting Selection[J]. Cardiovasc Drugs Ther, 2020,34(4):475-485.
[67] Ruff P, Pai R B, Storici F. Real-Time PCR-Coupled CE-SELEX for DNA Aptamer Selection[J].

ISRN Mol Biol, 2012,2012:939083.
[68] Dong L, Tan Q, Ye W, et al. Screening and Identifying a Novel ssDNA Aptamer against Alpha-fetoprotein Using CE-SELEX[J]. Sci Rep, 2015,5:15552.
[69] Kasahara Y, Irisawa Y, Ozaki H, et al. 2',4'-BNA/LNA aptamers: CE-SELEX using a DNA-based library of full-length 2'-O,4'-C-methylene-bridged/linked bicyclic ribonucleotides[J]. Bioorg Med Chem Lett, 2013,23(5):1288-1292.
[70] Morihiro K, Hasegawa O, Kasahara Y, et al. Azobenzene-modified DNA aptamers evolved by capillary electrophoresis (CE)-SELEX method[J]. Bioorg Med Chem Lett, 2021,31:127607.
[71] Alam, K. K., Chang, J. L., & Burke, D. H. (2015). FASTAptamer: A Bioinformatic Toolkit for High-throughput Sequence Analysis of Combinatorial Selections. Molecular therapy. Nucleic acids, 4(3), e230.
[72] McBride, R., van Zyl, M., & Fielding, B. C. (2014). The coronavirus nucleocapsid is a multifunctional protein. Viruses, 6(8), 2991–3018.
[73] Zhang, X., et al., Screening and Identification of ssDNA Aptamers for Low-Density Lipoprotein (LDL) Receptor-Related Protein 6. Molecules, 2023. 28(9).
[74] Almeida, N., et al., DNA aptamer selection and construction of an aptasensor based on graphene FETs for Zika virus NS1 protein detection. Beilstein J Nanotechnol, 2022. 13: p. 873-881.
[75] Zhu, C., et al., Selection and characterization of an ssDNA aptamer against thyroglobulin. Talanta, 2021. 223(Pt 1): p. 121690.
[76] Mendonsa, S.D. and M.T. Bowser, In vitro evolution of functional DNA using capillary electrophoresis. J Am Chem Soc, 2004. 126(1): p. 20-1.
[77] Zhu, C., et al., Evolution of multi-functional capillary electrophoresis for high-efficiency selection of aptamers. Biotechnol Adv, 2019. 37(8): p. 107432.
[78] Bayat, P., et al., SELEX methods on the road to protein targeting with nucleic acid aptamers. Biochimie, 2018. 154: p. 132-155.
[79] Darmostuk, M., et al., Current approaches in SELEX: An update to aptamer selection technology. Biotechnol Adv, 2015. 33(6 Pt 2): p. 1141-61.
[80] Poolsup, S., et al., Discovery of DNA aptamers targeting SARS-CoV-2 nucleocapsid protein and protein-binding epitopes for label-free COVID-19 diagnostics. Mol Ther Nucleic Acids, 2023. 31: p. 731-743.
[81] Zhang, L., et al., Discovery of sandwich type COVID-19 nucleocapsid protein DNA aptamers. Chem Commun (Camb), 2020. 56(70): p. 10235-10238.
[82] Zuker, M., Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res, 2003. 31(13): p. 3406-15.
[83] Guo, F., J.T. Babauta and H. Beyenal, The effect of additional salinity on performance of a phosphate buffer saline buffered three-electrode bioelectrochemical system inoculated with wastewater. Bioresour Technol, 2021. 320(Pt A): p. 124291.
[84] Rodrigues, K.T., M. Tavares and A. Van Schepdael, CE-MS for the Analysis of Amino Acids. Methods Mol Biol, 2018. 1730: p. 305-313.
[85] Haselberg, R., G.J. de Jong and G.W. Somsen, CE-MS for the analysis of intact proteins 2010-2012. Electrophoresis, 2013. 34(1): p. 99-112.
[86] Schoukroun-Barnes, L.R., E.P. Glaser and R.J. White, Heterogeneous Electrochemical Aptamer-Based Sensor Surfaces for Controlled Sensor Response. Langmuir, 2015. 31(23): p. 6563-9.
[87] Nevidalova, H., L. Michalcova and Z. Glatz, Capillary electrophoresis-based immunoassay and aptamer assay: A review. Electrophoresis, 2020. 41(7-8): p. 414-433.
[88] Yuce, M., E. Filiztekin and K.G. Ozkaya, COVID-19 diagnosis -A review of current methods. Biosens Bioelectron, 2021. 172: p. 112752.
[89] Zhou, Y., et al., Point-of-care COVID-19 diagnostics powered by lateral flow assay. Trends Analyt Chem, 2021. 145: p. 116452.
[90] Zhang, Y., et al., Recent Progress on Rapid Lateral Flow Assay-Based Early Diagnosis of COVID-19. Front Bioeng Biotechnol, 2022. 10: p. 866368.
[91] Dinesh, D.C., et al., Structural basis of RNA recognition by the SARS-CoV-2 nucleocapsid phosphoprotein. PLoS Pathog, 2020. 16(12): p. e1009100.

Appendix


A
Protein concentration (nM)
0 nM
25 nM
50 nM
75 nM
100 nM
150 nM
RFU
Complex Peak
Free Aptamer Peak
tECK4-R2
tECK4-R3
Time (min)


B
Protein concentration (nM)
0 nM
25 nM
50 nM
75 nM
100 nM
150 nM
RFU
Complex Peak
Free Aptamer Peak
tECK6-R2
tECK6-R3
Time (min)

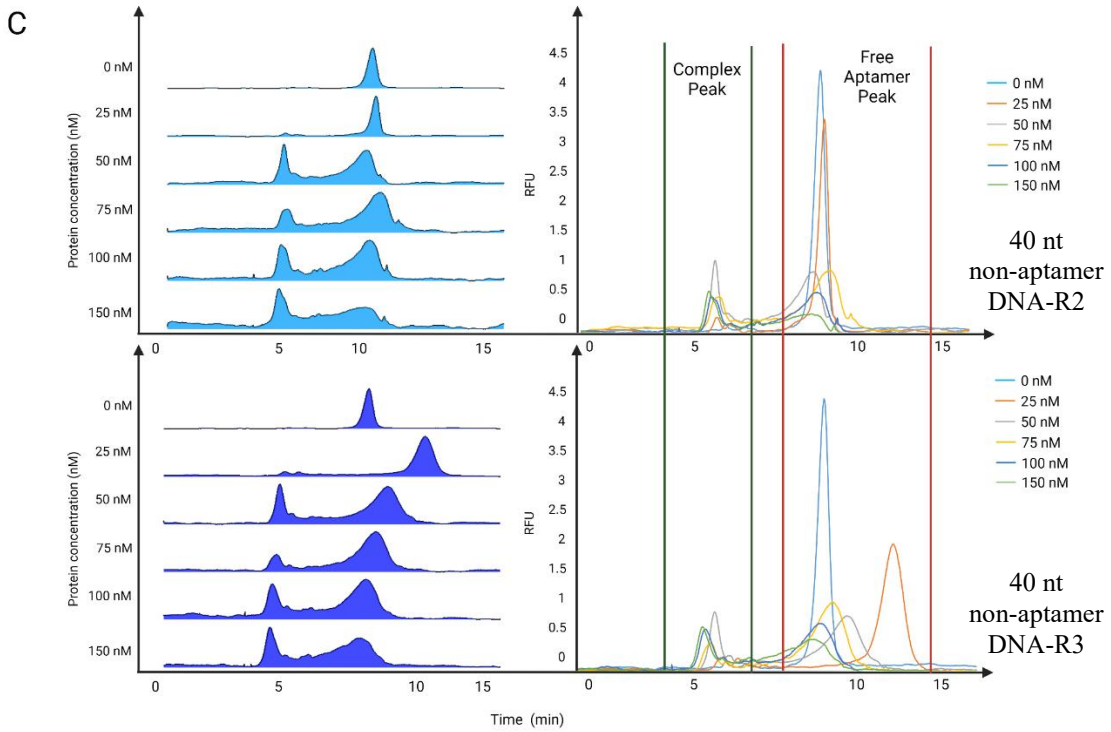


Figure S1. Additional replicates for tECK4 (A), tECK6 (B) and 40 nt non-aptamer DNA (C) in the experiment for Apparent K_D determination.

Table S2.1. The method for CE-SELEX.

	Time [min]	Event	Value	Duration	Inlet vial	Outlet vial	Summary
1		Capillary tip clean		0.00 min	B1:F4	B0:F4	DNAaway dip
2		Capillary tip clean		0.00 min	Bl: F5	B0:F5	water dip
3		Capillary tip clean		0.00 min	Bl: F6	B0:F6	water dip
4		Rinse-Pressure	20.0 psi	1.00 min	B1:A1	B0:A1	SDS wash
5		Capillary tip clean		0.00 min	Bl: F5	B0:A1	water dip
6		Capillary tip clean		0.00 min	B1:F6	B0:A1	Water dip
7		Rinse-Pressure	20.0 psi	1.00 min	B1:A2	B0:A1	DNAaway dip
8		Capillary tip clean		0.00 min	B1:F5	B0:Al	water dip
9		Capillary tip clean		0.00 min	Bl: F6	B0:A1	water dip
10		Rinse-Pressure	20.0 psi	1.00 min	BlA3	B0:A1	Acid wash
11		Capillary tip clean		0.00 min	B1:F5	B0:A1	water dip
12		Capillary tip clean		0.00 min	B1:F6	B0:A1	water dip
13		Rinse-Pressure	20.0 psi	1.00 min	Bl: A4	B0:A1	Water wash
14		Capillary tip clean		0.00 min	Bl: F5	B0:A1	water dip
15		Capillary tip clean		0.00 min	B1: F6	B0:A1	water dip
16		Rinse-Pressure	20.0 psi	2.00 min	B1:A5	B0-A1	Base wash
17		Capillary tip clean		0.00 min	Bl: F5	B0-A1	water dip
18		Capillary tip clean		0.00 min	B1:F6	B0:A1	water dip
19		Rinse-Pressure	20.0 psi	1.00 min	Bl: B6	B0:A1	Buffer wash
20		Capillary tip clean		0.00 min	B1:F5	B0:A1	water dip
21		Capillary tip clean		0.00 min	B1:F6	B0:A1	water dip

Table S2.2. The method of injection, separation, and collection of target aptamer.

	Time [min]	Event	Value	Duration	Inlet vial	Outlet vial	Summary
21		Capillary tip clean		0.00 min	B1:F6	B0:A1	water dip

22		Inject-Pressure	0.5 psi	30 s	B1:C1	B0:F4	Sample injection
23		Capillary tip clean		0.00 min	Bl: F5	B0:F5	water dip
24		Capillary tip clean		0.00 min	Bl: F6	B0:F6	water dip
25	0.00	Separate-Voltage	25.0 kV	3.80 min	B1:B1	B0:B1	Separation
26	0.10	Autozero					
27	3.80	Separate-Voltage	25.0 kV	3.50 min	B1:B1	B0:C1	Complex collection
28	7.30	Separate-Voltage	25.0 kV	2.70 min	B1:B1	B0:D1	Free aptamer collection
29		Capillary tip clean		0.00 min	Bl:F4	B0:A1	DNAaway dip
30		Capillary tip clean		0.00 min	B1:F5	B0:A1	water dip
31		Capillary tip clean		0.00 min	B1:F6	B0:A1	water dip

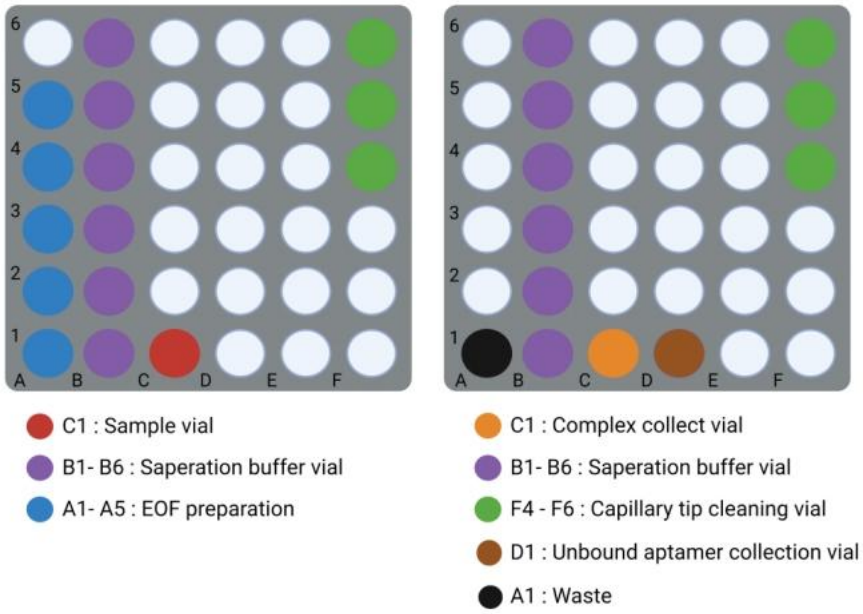


Figure S2. Vials' position for CE experiment.

Table S4.1. the binding percentage calculated based on the peak areas for all CE-K_D measurement experiments

N protein (nM)	40 nt non-aptamer -R1	40 nt non-aptamer -R2	40 nt non-aptamer-R2	tECK4-R1	tECK4-R2	tECK4-R3	tECK6-R1	tECK6-R2	tECK6-R3
0	0.797693916	0.591237053	1.388930969	0.850839996	0.425544645	0.425295351	2.289314239	1.422932158	0.86638208
25	5.658303072	3.503196402	3.97257205	62.4039907	61.64863025	56.64280924	42.71183329	40.11980171	43.91742403
50	36.97564502	38.00843501	30.61066057	73.11365557	71.95693199	72.70680797	57.38823097	58.79578628	61.85210686
75	47.06721318	47.07952268	43.31551543	77.25941271	75.46449681	74.97413576	59.46249885	60.23822638	58.2874323
100	55.3507819	58.48269355	59.00467882	85.87750198	81.24362743	89.95744554	64.09965115	67.54108143	59.81823189
150	58.93334863	61.37828351	57.41768777	95.27338813	97.55118626	97.4113324	93.78040944	93.98696411	94.73974112

200	73.19874	77.2138974	78.4082375	96.91854987	99.20398578	93.91874591	94.98745184	96.942357	97.21857821
250	83.19451312	88.095791	85.12083748	96.10923499	98.08417591	98.12873412	95.02381748	92.23041832	98.8417983
300	93.20138974	95.23814733	91.2983741	97.20837419	98.1293234	91.23174	91.2013974	97.12938742	95.2103842

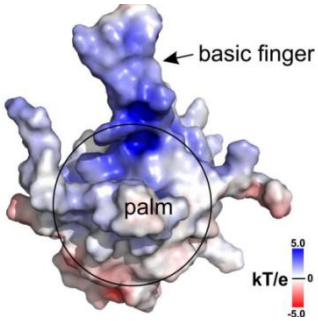


Figure S3. The N-NTD molecular surface electrostatic potentials revealed a basic patch extended between the finger and the palm, with a positively charged surface shown in blue and negatively charged surface in red.

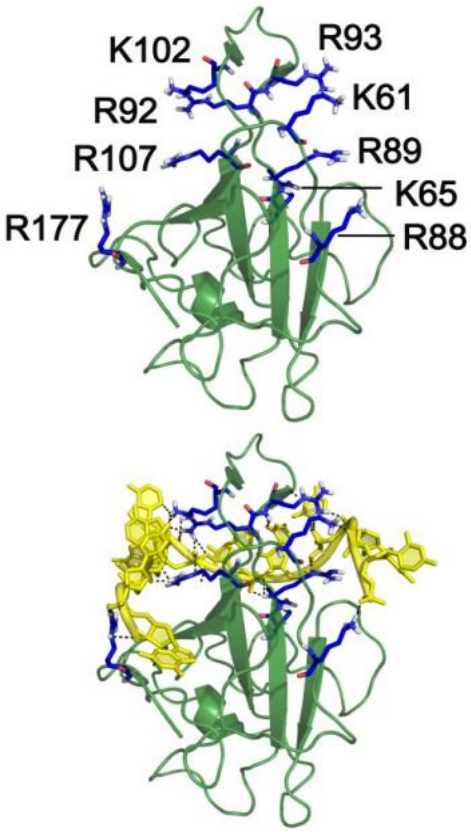


Figure S4. Cartoon representation of NTD of N protein highlighting all the available arginine and lysine residues in the interaction interface. The residues are shown as blue sticks in the upper panel, and the lower panel displays the ssRNA-10mer docked model in same orientation.