DEVELOPMENT OF NEW AND MODIFICATION OF EXISTING METHODS FOR THE PRIMARY SCREENING OF APTAMERS AND ANTIBODIES BY MEANS OF CAPILLARY ELECTROPHORESIS

Stanislav Beloborodov

A DISSERTATION SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN CHEMISTRY

YORK UNIVERSITY

TORONTO, ONTARIO

January 2021

ABSTRACT

Non-covalent biomolecular interactions play an important role in metabolism and support life in living organisms. Studying these interactions leads to a better understanding of the nature of the corresponding metabolic processes, their mechanisms and functions. This kind of knowledge helps to distinguish normal from abnormal metabolism and serves as a starting point in disease diagnostic and drug development. Any biomolecular interaction starts with the binding of molecule L (ligand) to molecule T (target) resulting in the formation of the target-ligand complex (TL).

Studying biomolecular interactions often involves testing a variety of ligand candidates for their ability to bind targets (primary screening) and conducting a quantitative assessment of such binding in terms of its kinetic and thermodynamic parameters (secondary screening). Arguably, the best analytical technique for the primary and secondary screening is Capillary Electrophoresis (CE). CE provides homogenous separation of the species, generates minor amount of waste and can serve both analytical and preparative purposes.

Sometimes, CE-based screening approaches fail due to insufficient separation of target–ligand complexes (TL) from non-binding ligands or due to the inability to collect the complexes. The first problem affects all types of ligands, but the second one is especially acute for the oligonucleotides (such as aptamers), since oligonucleotide binders are often required to be collected for the further amplification and sequencing. To address this problem, I developed a mathematical model that allows precise prediction of elution times of protein–aptamer complexes, which guarantees their successful collection. The developed model is especially useful for the cases when the complexes are present in the undetectably small amount. The

developed predictor significantly advances the existing screening approaches, such as NECEEM (Non-equilibrium Capillary Electrophoresis of Equilibrium Mixtures) and Ideal Filter CE (IFCE)

As a continuation of the protein–aptamer binding project, I developed an aptamer-based purification technique that allowed rapid isolation of proteins in their active state from the cell lysate. It is important to have such a technique in the arsenal of analytical tools, since any ligand screening approach benefits from using fully functional proteins.

The problem of inefficient separation of TL complexes from T and L is especially acute for the field of CE-based immunoassays. Particularly, for the assays that study virus-antibody interactions. The existing assays provide a minor (almost negligible) separation of virus-antibody complexes from the source molecules and allow screening mixtures of antibodies and spherical viruses only. In order to expand the field of existing CE-bases immunoassays, I developed an assay for screening bivalent antibodies for their ability to bind another common type of viruses, which are rod-like viruses. The developed assay was able to provide a baseline separation of virus-antibody complexes from free viruses and antibodies. In addition to that, my work resulted in the discovery of the previously unknown phenomenon of abnormally high electrophoretic mobility of the complexes between rod-like virions and bivalent antibodies. I explained this phenomenon and built a simple mathematical model for the prediction of the number of cross-linked virions in the complexes. The developed model can be used as a tool for the evaluation of the antibodies' immunoprecipitation activity with respect to rod-like virions.

DEDICATION

To my parents and grandparents

for giving me the best they could

and

To the Universe,

the mother of everything

AKNOWLEGMENTS

I've been doing my PhD for almost six years and I met a lot of bright and supportive people that helped me during my studies. The first place among them belongs to my supervisor, Professor Sergey N. Krylov. I never met such a smart and talented person like him. It would be impossible for me to finish any of my projects without his wise advices and recommendations. He helped me to develop myself. He turned me into a real scientist and taught me how to write good scientific papers. Thank you for the time that you spent and for the efforts that you applied to raise me as a scientist. Thank you for everything Prof. Krylov! I've learned a lot from you and gained a lot of valuable experience. It will always be an honor for me to use, carry and share (with care and not giving out any secrets) all the knowledge and skills that I got from you during my PhD studies.

Another person that I would like to thank is Dr. Lana Krylova. Sometimes I had disagreements with her, but most of the time she generated very valuable ideas that often helped me to better understand the problem and find right solution. She also made me work harder and I'm grateful to her for that.

I would like to express special thanks to my committee members: Prof. Derek Wilson and Prof. Philip Johnson. They guided me in my research and have always been supportive. Prof. Philip Johnson gave me an opportunity to use his lab equipment for my research needs and Prof. Derek Wilson organized CREATE MS-ESE seminars where I gained a lot of knowledge. Being CREATE member, I also had a unique chance to gain a year of full-time work experience by doing industrial internship at Sciex.

During my stay at Sciex, I met a lot of smart and bright people: Dr. Yves LeBlanc, Dr. Bradley B Schneider, Dr. Igor Chernushevich and others. They introduced me to the world of Mass Spectrometry and helped me with experimental work.

Special thanks go to my lab mates: Dr. Mirzo Kanoatov, Dr. Jiayin Bao, Dr. Agnesa Shala-Lawrence, Dr. Victor Galievsky, Roman Yufa, Dr. Sven Kochmann, Nikita Ivanov and An Le Thi Hoai for their great support and priceless help. Guys, I would not be able to finish my PhD studies without you! Thank you for everything and thank you for the fun time that we had together during New Year's Eves, conferences and in-office birthday meetings.

A special thanks go to my girlfriend Cynthia Suarez for making me work on this thesis and motivating me to not stop working on it. When the dark thoughts came, she always helped me to get over them.

In addition, I would like to thank my collaborators from Moscow: Dr. Vasily G. Panferov, Dr. Irina V. Safenkova and Prof. Boris B. Dzantiev for their valuable contribution to the project on virus-antibody interactions.

I also want to thank my parents, grandparents, cousins, aunts, uncles and other relatives for their warm support from over the ocean. For their wise and life-related advices, for their kindness and for their true love.

TABLE OF CONTENTS

APPENDICES

LIST OF TABLES

LIST OF FIGURES

LIST OF ABRREVIATIONS AND SYMBOLS

Ab	antibody
A_{cyl}	surface area of the particle in the approximation of its cylindrical shape
AFM	atomic force microscopy
BCA	bicinchoninic acid assay
BODIPY	boron-dipyrromethene
BSA	bovine serum albumin
CE	capillary electrophoresis
CGE	capillary gel electrophoresis
CIEF	Capillary Isoelectric Focusing
dd	deionized water
DL	double layer
DNA	deoxyribonucleic acid
DNAzymes	deoxyribozymes
dsDNA	double-stranded DNA
E	applied electrical field
ε	dielectric constant of the media
ε_0	vacuum permittivity
EC_{50}	half-maximum effective concentration
ECEEM	equilibrium capillary electrophoresis of equilibrium mixtures
ELISA	enzyme linked immunosorbent assay
EOF	electroosmotic flow
f	frictional coefficient
FDA	food and drug administration
Fc region	fragment crystallisable region
GFP	green fluorescent protein
I	ionic strength
IFCE	introduced ideal-filter CE
ITP	capillary isotachophoresis
K_d	dissociation constant
k_{on}	rate constant of complex association
k_{off}	rate constant of complex dissociation

L	length of cylinder
LB	lysogeny broth
LIF	laser induced fluorescence
LOD	limit of detection
LOOCV	leave-one-out cross validation
L_{tot}	total length of the capillary
L_{tr}	travel distance from the middle of the sample plug to the detection point
mRNA	messenger ribonucleic acid
MEKC	micellar electrokinetic capillary chromatography
MW	molecular weight
$MW_{P–ssDNA}$	the molecular weight of the complex
μ	electrophoretic mobility
μ_{EOF}	electroosmotic mobility
$\mu_{P–ssDNA}$	mobility of the protein–aptamer complex
μ_{ssDNA}	mobility of ssDNA
μ_{vir}	electrophoretic mobility of an individual virion
η	dynamic viscosity of the solution
n	number of cross-linked particles
N	number of nucleotides in ssDNA
NECEEM	non-equilibrium capillary electrophoresis of equilibrium mixtures
nt	nucleotide
OD	optical density
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
pH	potential of hydrogen
pI	isoelectric point
ppKCE	plug-plug kinetic capillary electrophoresis
PVX	potato Virus X
q	charge
q_{agg}	charge of aggregate
q_{vir}	charge of a single virion

ρ	density
R_{hyd}	hydrodynamic radius
RISC	RNA-induced silencing complex
RNA	ribonucleic acid
R_p	hydrodynamic radius
rpm	revolutions per minute
RSD	residual standard deviation
SDS-MW	protein molecular weight analysis in the SDS-gel electrophoresis mode
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SELEX	systematic Evolution of Ligands by EXponential enrichment
siRNAs	small interfering RNAs
SPR	surface Plasmon Resonance
ssDNA	single-stranded DNA
ssRNA	single stranded RNA
t	migration times
TB	terrific broth
t_{el}	elution times
UV	ultraviolet
v	velocity of the molecule
V	applied voltage
VEGF	vascular endothelial growth factor
V_{EOF}	velocity of electroosmotic flow
V_p	volume of protein
γ	constant coefficient of proportionality
ζ	zeta potential

CHAPTER 1. INTRODUCTION

1.1. THE OVERVIEW

Non-covalent biomolecular interactions play an important role in many metabolic processes. In the majority of cases, these processes involve proteins. For example, interactions of proteins with nucleic acids makes transcription of DNA and translation of RNA possible; interactions of proteins with other proteins play a significant role in immune response and signal transduction; interactions of proteins with small molecules and ions facilitate neural signal transmission, cellular respiration, electron transfer, etc. These examples are only a small part of the processes that are involved in the whole system of metabolism, where proteins are important regulatory molecules. Very often, metabolic disorders are associated with protein malfunctioning, which could be a result of their overexpression, the lack of expression, mutation, inhibition and other abnormal processes. Many treatment strategies require the use of drugs designed to interact with proteins that are directly or indirectly involved in the malfunctioning metabolic pathways. According to FDA, about 70% of all existing drugs target proteins **[1]**. Because of that, proteins are the most common targets in the drug screening approaches.

Discovery of new drugs starts with testing variety of drug candidates (ligands) for their ability to bind target molecules and form target–ligand complexes. This process is known as ligand screening and can be conducted by a variety of analytical techniques that separate target–ligand complexes from non-binders (primary screening) and further measure kinetic and thermodynamic parameters of the target–ligand binding (secondary screening). Analytical techniques that are used for these purposes can be divided in two main categories: heterogeneous techniques, such as Surface Plasmon Resonance (SPR) **[2]** and affinity chromatography **[3]** and

homogenous techniques, such as Thermophoresis **[4]** and Capillary Electrophoresis (CE) **[5,6]**. Homogenous techniques operate with molecules in solution, while heterogeneous techniques require that the target molecules be attached to the surface of the solid media (beads, plates, biosensor). The immobilization of the target molecules to the surface has two major limitations: first, it requires chemical modification of the targets with the linker, and, second, the attached target molecules are sterically hindered and not fully accessible for binding the ligands molecules in solution. Hence, homogenous techniques are preferred over heterogeneous.

Capillary electrophoresis (CE) is arguably the best technique for homogeneous separation. There are several reasons for that. First and most important, CE allows conducting experiments under the conditions that mimic physiological and does not require the presence of organic solvents or volatile compounds in the running buffer. Second, CE requires very small amount of the sample for the analysis and generates minor amount of waste. Third, CE has great separation efficiency. Fourth, CE is based on the physical separation of the species in the capillary and can be used for both, analytical and preparative purposes. Thus, CE not only provides separation of the target–ligand complexes from non-binders, but also allows direct collection of the complexes, which is very important for the further steps of the drug development process.

Successful collection of the target–ligand complexes in CE requires satisfying two conditions: first, the separation efficiency of the complexes from non-binders must be high enough to provide collection of pure complexes with the minimum amount of contamination by non-binders; and second, the elution time of the complexes must be known to provide their precise collection. Estimation of the elution time for the complexes requires the knowledge on their electrophoretic mobilities. The values for the electrophoretic mobilities can be extracted from the experimental electropherograms that represent the dependence between the detected

signal and the migration time. However, it's not always possible to detect signal from the complexes in CE. This problem is often associated with a small amount of binders in the mixture, which results in the formation of an undetectable amount of complexes. In such a case, the only way to find the electrophoretic mobility of the complexes is to predict it.

As it was mentioned above, the most common type of targets in the ligand screening approaches are proteins. To the best of my knowledge, the models for prediction electrophoretic mobilities of protein–ligand complexes were developed only for one category of ligands, which are DNA molecules. The developed models work well for dsDNA **[7,8]** and long free-draining ssDNA **[9]**; however, they are not applicable to ssDNA aptamers, which are considered as promising biological therapeutics. To address this point, I developed a mathematical model that allows accurate mobility prediction for protein–aptamer complexes in the assumption of their spherical shape. The developed model complements CE-based aptamer selection approaches, such as NECEEM **[5]** and IFCE **[10]** and makes them more advanced. The detailed description of the model is given in Chapter 2.

Aptamers are often compared with antibodies since they are both biopolymers that are capable of binding the target molecules that they were developed for. Moreover, aptamers were designed as synthetic analogs of antibodies with a number of advantages, such as possibility to be synthesized *in vitro*, stability in a wide range of temperatures, smaller molecular weight and others. Despite of these advantages, aptamer-based drugs are not as widely used as therapeutic antibodies due to some limitations, such as inability of *in vivo* delivery of the aptamers to the cells, cross-reactivity of the aptamers and others **[11]**. There is a chance that these problems will be solved in the future; however, as for now, there is only one aptamer-based drug approved by the FDA **[12]**, while for the antibodies, this number is equal to 79 **[13]**. The methods for

screening antibodies by means of CE are less developed than those for aptamers. In particular, so-called CE-based immunoassays, which were designed for screening antibodies for their ability to bind viruses, suffer from the inability to provide an efficient separation of virus-antibody complexes from the molecules they are built of **[14]**. In addition, these assays were developed for spherical viruses only **[14**,**15]**. Considering the importance of the antibodies for the pharmaceutical industry, it was important to develop CE-based approach for an effective separation of virus-antibody complexes from the molecules they are built of, and make this approach applicable to other most common type of viruses, which are rod-like viruses. I succeed in the development of such an approach; furthermore, I discovered a previously unknown phenomenon of abnormally high electrophoretic mobility of the complexes between rod-like virions and bivalent antibodies. I explained this phenomenon and proposed a simple empirical formula for predicting the number of cross-linked virus particles in the complexes. The formula can be used for the evaluation of the antibodies' immunoprecipitation activity in respect to the binding of rod-like virions. The explanation for the observed phenomenon and description of the model are provided in Chapter 3.

Any ligand screening approach assumes investigation of the target–ligand interactions *in vitro*, which to some extent can be different from those interactions *in vivo*. Since one of the main goals of drug discovery is to obtain a drug that is fully functional *in vivo*, it is important to mimic *in vivo* conditions in every *in vitro* step of the drug development process. One of the possible ways to narrow the gap between *in vitro* and *in vivo* interactions is to use proteins with the preserved natural activity. The ability of the protein to demonstrate its natural properties (such as catalytic activity, transport activity and so on) under the *in vitro* experimental conditions indicates that *in vitro* state of the protein is identical to that *in vivo*. There are many published

procedures for the expression and purification of proteins; however, only some of them aims at testing the activity of the proteins in the end **[16**,**17]**. Moreover, none of the published procedures aimed to compare the activity of the purified protein to that purified by other procedures. In order to overcome this limitation, I developed an aptamer-based purification technique for the rapid isolation of proteins from the cell lysate. The last step of the technique involved testing the activity of the purified protein and comparing it to that for the same protein purified by affinity column chromatography. Using two proteins, MutS and AlkB, as examples I showed that the purified proteins preserved their natural activity. Moreover, they were more active than the same proteins purified with affinity column chromatography. Therefore, the developed approach can be safely used for the purification of the proteins that are further considered to be targets in the ligand screening process. The detailed description of the developed purification technique is given in Chapter 4.

Many of the existing ligand screening approaches suffer from low reproducibility and robustness. In the case of CE-based approaches, low reproducibility is often associated with the changes in the electroosmotic flow (EOF) from run to run. Therefore, I developed a number of solutions that allow one to achieve excellent control over the EOF. The solutions include reconsideration of the capillary conditioning procedure and elimination of the siphoning effect. Another important detail that I paid attention to was the formation of pseudo-complexes between the analyzed species and NaCl-related ions. This phenomenon was not described in the literature. Thus, I proposed a hypothesis for the formation of such complexes and suggested a simple way for their elimination. The detailed descriptions of the new capillary conditioning procedure and the hypothesis on the formation of salt-related pseudo-complexes are given in Chapter 5.

1.2. NUCLEIC ACIDS AND PROTEINS AS BIOLOGICS

1.2.1. Nucleic acids

1.2.1.1. Aptamers

Aptamers are short single-stranded DNA or RNA molecules with a length of about 20-100 nucleotides, which are capable of forming hairpin-like structures due to the formation of intramolecular hydrogen bonds between the ccomplimentary nucleotides within one molecule. One aptamer may have several different hairpin structures, which are thermodynamically stable at different temperatures. Some of these structures are capable of binding proteins, making the latter inactive. For this reason, aptamers are considered as promising drug candidates with inhibitory mechanism of action **[18]**. Aptamers against specific protein are usually selected from random DNA or RNA oligonucleotide libraries by the process, which is known as SELEX (Systematic Evolution of Ligands by EXponential enrichment) **[19**,**20]**. The detailed mechanism of SELEX will be described in the further paragraphs.

Each oligunocleotide in the DNA or RNA library is made of three sequential regions: initial constant region, middle random region and terminal constant region. Initial and terminal constant regions are the same for all oligunocleotides in the library and serve as priming regions for PCR amplification of the selected aptamers. The random region is unique for each molecule and responsible for the structural diversity of the aptamers. Depending on the length of the oligonucleotides and the complexity of the random region, the number of unique hairpin structures within one library is about 10^{13}-10^{16} **[21]**. Thus, random DNA or RNA libraries can be potential sources of aptamers for any protein. As for now, there is only one FDA-approved aptamer-based drug on the market (Macugen), that is used for treatment of the macular degeneration disorder **[12]**.

1.2.1.2. Antisense oligonucleotides

Antisense therapy aims to supress the synthesis of the proteins involved in the abnormal or disease-related metabolic processes **[22]**. Suppression of the protein's synthesis is achieved by using small single stranded oligonucleotides (antisense agents), which are complimentary to some regions of messenger RNA (mRNA). The binding of the antisense agents to mRNA hinders its interactions with ribosome and prevents protein's synthesis. As for now, there are six commercially available antisense drugs that are approved by FDA **[23]**. The most known among them are Nusinersen **[24]**, which is used in the treatment of muscular atrophy and Fomivirsen **[25]**, which is used as an antiviral agent against cytomegalovirus.

1.2.1.3. Deoxyribozymes and ribozymes

Deoxyribozymes (DNA enzymes, DNAzymes) and Ribozymes are single stranded DNA or RNA molecules that possess catalytic properties. Thus, DNAzymes and Ribozymes are cable of binding target molecules (like aptamers) and turning them into the chemically different products. Selection of Oligonucleotide-based enzymes against specific targets is conducted *in vitro* by SELEX. The first clinically tested ribozyme was Angiozyme **[26]**. It was developed as an anti-cancer drug via inhibiting tumor angiogenesis. The mechanism of its action involves binding and cleaving mRNA which is responsible for the synthesis of the receptors sensitive to Vascular Endothelial Growth Factor (VEGF). Up to date, there are no FDA-approved DNAzymes or Ribozymes on the market; however, many of them are at different stages of clinical trials **[27]**.

1.2.1.4. Gene silencing RNAs

This group of biologics includes MicroRNAs and Small interfering RNAs (siRNAs). MicroRNAs are small single stranded non-coding RNA sequences (20-25 nucleotides long) that exist naturally and play important role in the gene regulation. The molecules of microRNAs are

partially compliment to some regions of mRNA and can bind to them. Such a binding results in the downregulation of the gene expression by one of the following mechanisms: suppression of the protein's synthesis caused by the inhibition of the mRNA's translation; cleavage of the mRNA molecule; detachment of adenosine monophosphate (A) units from poly(A) tail of mRNA. Overexpression of microRNAs in the human body serves as an indicator for the abnormal metabolic processes. For this reason, microRNAs found wide application as biological markers **[28, 29]**, however there are no FDA-approved MicroRNA drugs on the market.

Small interfering RNAs are naturally occurring double stranded RNAs with the length of about 20-25 nucleotides **[30]**. The mechanism of their action is similar to that for microRNAs and involves interfering with complement fragments of mRNA, resulting in the cleavage of mRNA. Cleaved mRNA cannot be further translated into the corresponding protein. Prior to the interaction with mRNA, siRNA goes through the stage of dissociation via RNA-induced silencing complex (RISC). Single stranded siRNA that stays in the complex after the dissociation from its counterpart gains ability to bind and cleave mRNA. As for now, there is only one FDA-approved siRNA-based drug (Patisiran) **[31]**, which is used in the treatment of polyneuropathy.

1.2.2. Proteins and peptides

1.2.2.1. Antibodies

Antibodies (also known as immunoglobulins) is the most abundant class of the drugs among the biologics. As for 2019, the overall number of FDA-approved antibodies was equal to 79 **[13]**. In the current list of 22 new drugs approved by FDA in 2020 (as on June 01), four drugs belong to antibodies.

Antibodies are naturally occurring proteins with molecular weight of ~150 kDa that are produced by white blood cells and fulfill immunological function. The main role of the

antibodies is to bind pathogen agents (such as viruses, photogenic bacteria, cells and others) and cause their neutralization or immunoprecipitation. Production of polyclonal antibodies is a process, which starts with injection of the antigen into the animal body. With the time, the immune system of the animal produces antibodies with different affinity to the antigen **[32]**. Production of identical (monoclonal) antibodies requires isolation of B-cells (responsible for the production of antibodies) from the animal, followed by their fusion with myeloma cancer cells **[33]**. This process results in the formation of the immortalized antibody-producing hybrids with increased survivability and reproductivity. The isolated polyclonal and monoclonal antibodies are then subjected to purification and further laboratory tests aiming to confirm their ability to interact with their antigens.

1.2.2.2. Coagulation factors

Coagulation factors is the group of 12 naturally occurring proteins (factors I-XIII, except IV) that play crucial roles in the blood coagulation processes **[34]**. Recombinant forms of factors VII, IX, X, XIII along with their fusion products with other biomolecules have been approved by FDA as prominent drugs for treatment of hematologic diseases. Some of these factors are known under the following tradenames: NovoSeven, BeneFIX, Rebinyn and others. Factors VII, IX and X belong to the class of serine proteases and factor XIII is the proenzyme that can be activated by thrombin. The activated form of factor XIII (also known as XIIIa) performs crosslinking of fibrin.

1.2.2.3. Fusion proteins

Fusion proteins are chimeric biomolecules composed of two types of proteins (or their parts) conjugated together. The main idea for such fusion is to obtain a chimeric molecule that would combine beneficial properties of both parts. Most existing FDA-approved chimeric

therapeutics are functional proteins conjugated with the fragment crystallizable region (Fc region) of immunoglobulin G **[35]**. The presence of Fc region in the chimeric molecule makes it capable of binding neonatal Fc receptor in the endothelium. Such binding protects the chimeric molecule from the fast degradation and significantly increases its half-life in the blood, giving another part of the chimeric molecule enough time to act as a therapeutic. For this reason, Fc regions are often conjugated with coagulation factors **[36]** and receptor inhibitors **[37]**. Fc-conjugated proteins found wide application in the treatment of cancer, hematological and genetic disorders. Fc-fusion is the most popular but not the only way of extension protein's lifetime in the blood. Thus, the same goal can be achieved by conjugation of the therapeutic proteins with albumin **[38]** or polyethylene glycol (PEG) **[39]**.

1.2.2.4. Enzymes

Enzymes is a large group of proteins that are known for their ability to catalyze biochemical reactions. In most cases, enzymes are targets for other therapeutics; however, recombinant forms of some enzymes and their Fc-conjugates are used in the treatment of metabolic disorders caused by small molecules or proteins, which are targets for the corresponding enzymes and can be neutralized by them. For example, recombinant forms of such enzymes as glucarpidase (Voraxaze) **[40]**, asparaginase erwinia (Erwinaze) **[41]** and ocriplasmin (Jetrea) **[42]** are used in the treatment of cancer, leukemia and eye vitreomacular adhesion, respectively.

1.2.2.5. Hormones and growth factors

Hormones and growth factors are signaling molecules (can be small molecules, proteins or peptides) that are produced by one cells in order to regulate behavior of other cells. The binding of these molecules to the specific receptors located on the target cells leads to a physiological

response. In case of hormones, such a response is usually a metabolic change that is needed for the adaption of the organism to the new environment, new situation or to the new stage in the lifecycle. Thus, hormones may regulate behavior, the state of the immune system, libido and so on. The range of the processes that are regulated by the growth factors is much narrower than that for the hormones and includes growth, differentiation, and proliferation of the cells. With respect to some molecules (such as steroids, for example), the terms hormones and growth factors can be used interchangeably. Synthetic and recombinant forms of some hormones and growth factors are widely used as drugs in the thyroid hormone therapy **[43]**, hormone replacement therapy **[44]**, growth hormone therapy **[45],** and other fields of medicine.

1.3. CAPILLARY ELECTROPHORESIS: A TOOL FOR SCREENING BIOLOGICS

1.3.1. Introduction to capillary electrophoresis

Capillary Electrophoresis (CE) is a homogenous analytical technique that provides efficient separation of the species based on their ability to move with different velocities in the applied electric field **[46]**. In CE, the velocity of the molecule (v) and the applied electric field (E) are related by the following equation:

$$v = \mu E \tag{1-1}$$

Here, μ is an electrophoretic mobility of the molecule. This parameter is not only a coefficient of proportionality between v and E, but also a physical characteristic of the molecule that corresponds to the ratio between the charge of the molecule (q) to its frictional coefficient (f):

$$\mu = q / f \tag{1-2}$$

In turn, the frictional coefficient of the molecule is directly proportional to its hydrodynamic radius (R_{hyd}):

$$f = 6\pi\eta R_{hyd} \tag{1-3}$$

Here, η corresponds to the dynamic viscosity of the solution.

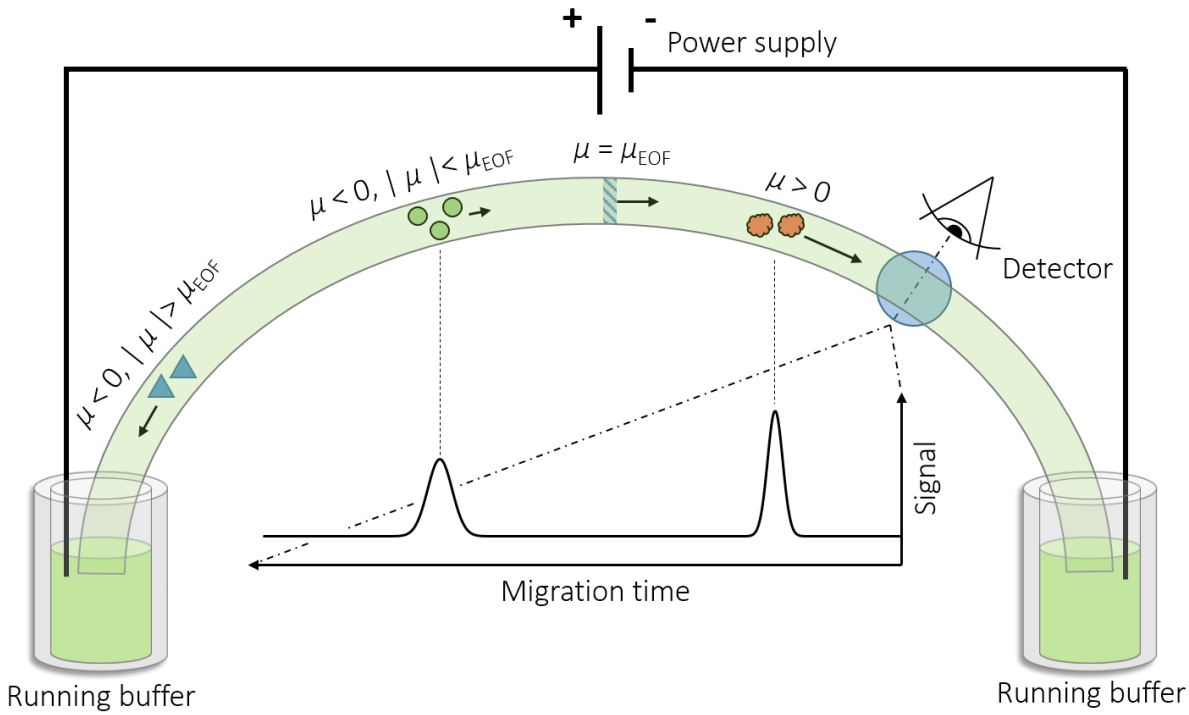

Figure 1.1. Schematic illustration of the process of Capillary Electrophoresis. The left and the right ends of the thin bare silica capillary are placed in the vials with the platinum electrodes inserted in them. Both, the capillary and the vials are filled with the running buffer (background electrolyte). Application of the voltage to the electrodes creates an external electric field and results in the formation of EOF that makes injected species move from the anode to the cathode. All the species that were initially injected as a single plug are separated in the electric field and move with different velocities. The velocity of each specie is proportional to its electrophoretic mobility (μ). When species are passing the point of detection, their optical signal is registered and then plotted versus the passing time, giving the resulting electropherogram.

Any CE experiment starts with an injection of a short plug of the analyzing sample(s) into the narrow capillary that prefilled with the running buffer (**Figure 1.1**). The left and the right ends of the capillary (which are inlet and outlet, respectively) are placed in the separate vials filled the running buffer (further, simply buffer). In a same time, the buffer solution in each vial has a contact with a platinum electrode that connected to the high voltage power supply (1-30 kV). The application of the high voltage to the electrodes creates an external electric field and makes the buffer solution move from the anode (+) to the cathode (-), resulting in the formation of the Electroosmotic Flow (EOF).

The existence of EOF is possible due to the presence of the electric Double Layer (DL) at the interface between the buffer solution and the partially deprotonated silanol groups located on the glass capillary wall **[47]**. A simplified mechanism for the formation of DL and EOF can be described as follows. Deprotonated silanol groups carry negative charge and attract positively charged counterions (cations) from the solution. Attracted cations arrange themselves into the complex structure made of two layers that follow one another: adsorbed layer of counterions and diffuse layer of counterions. The ions in the adsorbed layer are tightly bound to the surface and cannot move, while the ions in the diffuse layer are mobile. These two layers fully compensate the negative electric potential of the wall. With increasing distance from the capillary wall to the bulk of the solution, the electric potential of the capillary wall first drops linearly in the absorbed layer of counterions, and then it drops exponentially in the diffuse layer counterions, until it reaches zero. Beyond the certain point in the diffuse layer, where the electric potential of the system corresponds to so-called zeta-potential (ζ), the ions become responsive to the external electric field and start moving towards the cathode as soon as the external voltage is applied. The

ion motion is further transferred to the upper layers of the buffer and leads to the movement of the entire volume of the buffer, forming EOF.

The main characteristic of Electroosmotic Flow is electroosmotic mobility (μ_{EOF}). Similar to the electrophoretic mobility, this parameter is a coefficient of proportionality between the velocity of electroosmotic flow (v_{EOF}) and applied electric field (E):

$$v_{EOF} = \mu_{EOF} E \tag{1-4}$$

On the other hand, μ_{EOF} is a parameter that characterises electrokinetic properties of the capillary surface and depends on the value of ζ-potential:

$$\mu_{EOF} = \frac{\varepsilon \varepsilon_0 \zeta}{\eta} \tag{1-5}$$

In this equation, ε and ε_0 correspond to the dielectric constant of the media and the vacuum permittivity, respectively.

EOF delivers molecules to the point of detection. If the molecules are neutral ($\mu = 0$), they move along with EOF (i). If the molecules are positively charged ($\mu > 0$), they move faster than EOF (ii). If the molecules are negatively charged ($\mu < 0$), they can either move in the same directions as EOF (iii), or in the opposite direction (iv). For cases (iii), the absolute value of μ of the molecules is smaller than μ_{EOF}, and for case (iv), the absolute value of μ of the molecules is bigger than μ_{EOF} (**Figure 1.1**).

At the point of detection, the signal coming from the separated molecules is registered and measured. In the majority of cases, the signal has an optical nature and corresponds to the ability of the molecules to absorb light (UV detection, photodiode array detection) or re-emit the light (Laser Induced Fluorescence detection). The measured signal is then plotted versus the migration

times of the species, resulting in a two-dimensional spectrum, which is commonly known as electropherogram.

1.3.2. Using CE for screening DNA/RNA libraries

In comparison with other biologics, nucleic acids are the most suitable molecules for screening by CE. There are two main reasons for that. First, nucleic acids are polyelectrolytes that carry large negative charge and hence have high negative values of electrophoretic mobilities (about -30 $mm^2kV^{-1}s^{-1}$), which are much more negative than that of the majority of other biomolecules **[48]**. Binding nucleic acids to target biomolecules (which are mainly proteins/peptides) results in the formation of complexes, whose electrophoretic mobilities are less negative than that for the nucleic acids and more negative than the mobilities of the targets. In the majority of cases, these differences are sufficient for baseline separation of the nucleic acid–target complexes from non-binding nucleic acids by means of CE. The second reason corresponds to the limitations of CE as a technique. Since it's mainly an analytical technique that only permits operating with small sample volume (5-50 µl), the physical collection of the target–binder complexes followed by the extraction of the binders in a sufficient amount for further investigation becomes challenging. This limitation becomes negligible if the collected binders are amplifiable and their amount can be increased on a post-experimental step. Since nucleic acids are the only amplifiable biomolecules, it makes them perfect candidates for CE-based screening methods.

In the majority of cases, the nucleic acid drug candidates are represented by randomly synthesized libraries of ssDNA or RNA aptamers. The screening of the aptamers for their ability to bind a specific protein starts with mixing the source ssDNA/RNA library with this protein. Within time, the mixture reaches equilibrium. At this point, a small plug of the mixture is

injected into the capillary followed by the electrophoretic separation of the components of the mixture. Usually, such a separation is conducted in non-equilibrium conditions (NECEEM), which lead to the formation of the following zones that come out one after another: pure protein (i), protein-containing dissociation region of the complex (ii), the complex (iii), aptamer-containing dissociation region of the complex (iv) and the zone of non-binding aptamers (v). If the source library of aptamers is fluorescently labeled, the zones (iii), (iv) and (v) can be detected using LIF detection. If the library is unlabeled, the components of the mixture are detected with assistance of UV, however this kind of detection is less sensitive than LIF. Thus, LIF is the most popular method of detection in NECEEM experiments.

The resulting electropherogram of the NECEEM experiment is then used for the establishment of the elution time window of the protein–aptamer complex. Moreover, the presence of the dissociation region (iv) on the electropherogram provides an opportunity for the quantitative assessment of the binding affinity of the source library to the target protein in terms of its half maximum effective concentration (EC_{50}). When the elution time window for the protein–aptamer complex is established, the NECEEM experiment is then conducted again, purposing the physical collection of the protein–aptamer complex and the aptamers from the dissociation region. Collected binders are then purified and amplified by PCR. The overall process of the selection of the aptamers followed by their purification and amplification is called SELEX. The first round of SELEX results in an enriched pool of aptamers, which are then subjected to more rounds of SELEX. Thus, SELEX is repeated multiple times until the pool with desirable EC_{50} is obtained. The aptamers from the final pool are sequenced and then individually tested by NECEEM for their ability to bind the target protein. Such testing assumes evaluation of the binding constant (K_d) and the kinetic parameters (k_{on}, k_{off}).

The efficiency of NECEEM separation of the protein–aptamer complexes from non-binders can be maximized by making protein–aptamer complexes and non-binding aptamers move in opposite directions in CE. This idea was implemented in a newly introduced method called IFCE.

1.3.3. Using CE for screening proteins and peptides

Using CE for screening the libraries of peptides or proteins for their ability to bind target molecules is very uncommon. To the best of my knowledge, there is only one published article that provides an example of screening a peptide library against a small molecular target (vancomycin) by means of CE **[49]**. In the published article, the separation of the vancomycin–peptide complex from non-binding peptides was conducted in the equilibrium conditions, which assumed the presence of the target molecules not only in the sample plug, but also in the running buffer. This approach is known as Equilibrium Capillary Electrophoresis of Equilibrium Mixture (ECEEM) and it is often used to increase the lifetime of the target-ligand complex by stopping it from the dissociation.

The lack of scientific interest in screening peptide and protein libraries by means of CE is reasonable and can be explained as follows. First, peptides and proteins are not amplifiable molecules. Thus, using proteins/peptides libraries as an initial source of binders is limited by the inability to amplify selected binders immediately after the first round of selection by SELEX. The only way of proceeding with the next rounds of SELEX would be identification of the binders from the first round of the selection followed by their *in vitro* or *in vivo* synthesis, which is very expensive and time-consuming. Second, it's much more rational to develop antibodies to the targets or to select high affinity binders from the random libraries of amplifiable oligonucleotides instead of selecting protein-based ligands from the randomly synthesized

mixtures of proteins or peptides. Thus, CE is mainly used as a tool for the quantitative and qualitative assessment of target–protein interactions, rather than for selecting protein ligands from a variety of candidates. In the majority of cases, the target is also a protein and target–protein interactions refer to protein–protein interactions. Assessment of these kind of interactions by means of CE is more challenging than the assessment of the previously described protein–DNA interactions. The main reason for that is very similar electrophoretic mobility values for all existing proteins **[48]**. Thus, efficient separation of the protein–protein complex from its individual components often requires a combination of electrophoretic separation with simultaneous chromatographic separation, which is achieved by adding extra media into the running buffer. For example, adding surfactant molecules into the running buffer turns CE into Micellar Electrokinetic Capillary Chromatography (MEKC) **[50]**, and adding gel matrices turns it into Capillary Gel Electrophoresis (CGE) **[51]**. Both methods are used widely in CE-based immunoassays **[52]**.

CHAPTER 2. SHEPRICAL-SHAPE ASSUMPTION FOR PROTEIN–APTAMER COMPLEXES FACILITATES PREDICTION OF THEIR ELECTROPHORETIC MOBILITY

My contributions to the article were: (i) performing all presented experiments, (ii) developing the mathematical model, (iii) preparing all figures, (iv) interpreting the results and (v) writing the manuscript.

2.1. INTRODUCTION

DNA and RNA aptamers are capable of binding their protein targets with high selectivity and affinity **[19]**; therefore, they are used to develop diagnostic probes, drug-delivery vehicles, and drugs **[53**–**55]**. DNA aptamers are selected from random-sequence ssDNA libraries by incubating the library with the target and then partitioning protein-bound DNA from unbound DNA. More than 70% of attempts to select aptamers fail **[56]**; inefficient partitioning is among the major culprits for these failures **[57]**.

The partitioning of protein-bound DNA from unbound DNA is typically performed with the protein immobilized on the surface of magnetic beads **[58]**. The efficiency of surface-based partitioning is typically below 10^2 **[59]**. This number roughly corresponds to the filtering capacity of the partitioning technique **[10]**. Thus, if the original molar ratio of non-binders to binders is 100:1, the surface-based techniques are capable to reduce this ratio to 1:1 at best. Such a low efficiency of partitioning makes it virtually impossible to select aptamers in a single round.

Capillary electrophoresis (CE) facilitates homogeneous separation, and, thus, provides a powerful alternative for surface-based partitioning. In classical CE-based partitioning, e.g. by nonequilibrium capillary electrophoresis of equilibrium mixtures (NECEEM), protein–DNA complexes move in the same direction as unbound DNA and the efficiency of partitioning is approximately 10^5 **[59]**. In recently introduced Ideal-Filter CE (IFCE), the complexes move towards the capillary outlet while unbound DNA moves backwards to the inlet **[60]**. The efficiency of IFCE-based partitioning reaches 10^9, which is presumed to be sufficient for reliable selection of aptamers in a single step of partitioning.

Both, NECEEM and IFCE require collection of fraction(s) of protein–ssDNA complexes, separated from unbound ssDNA by CE, followed by isolation and amplification of the collected pools of ssDNA. Precise fraction collection the knowledge of complex position in the capillary. If the complex can be optically detected, then collecting fractions containing the complex is simple **[59]**. However, if the complex is optically undetectable (due to its small amount) or if its optical signal interferes with signals from other species in the mixture, then the task becomes much more complicated. In principle, the position of the complex in the capillary can be predicted theoretically if the value of its electrophoretic mobility is known. The mobility of protein–aptamer complexes changes significantly with varying sizes of the protein and DNA, and it should be determined *de novo* for every new combination of the protein and the DNA library. Experimentally measuring the mobility of protein–aptamer complex requires having the aptamers selected, which creates the vicious circle: selecting aptamers requires knowing the mobility while measuring the mobility requires having the aptamers selected. A logical approach to breaking the circle is to find the mobility of protein–aptamer complex theoretically by using a reliable mobility model. A suitable model is, however, not available.

Two major existing mobility models for protein–DNA complexes assume that DNA is either a rigid rod that has minor flexibility or a long free-draining polymer chain, that undergoes no structural changes while moving through the solvent **[7**–**9]**. Further, they assume single-point attachment of the end of DNA molecule to the protein **(Figure 2.1a, b)**. These assumptions are inapplicable to protein–aptamer complexes, in which ssDNA is linked to the protein via multiple bonds and retains a specific aptamer conformation (**Figure 2.1c**). Hence, the existing mobility models for protein–DNA complexes are inapplicable to protein–aptamer complexes.

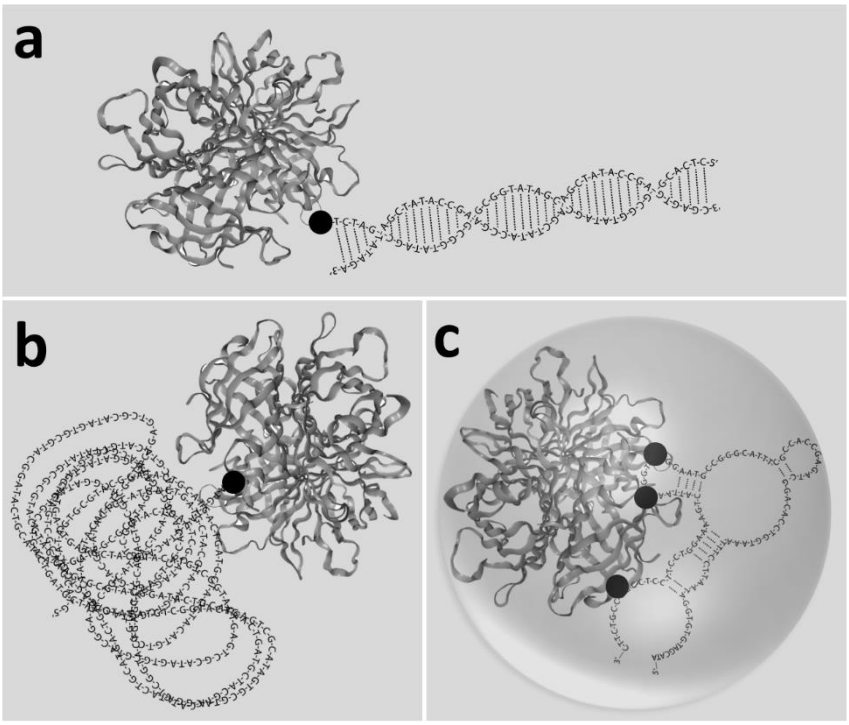

Figure 2.1. Schematic illustration of protein−DNA complexes with respect to three mobility models: **rigid-rod DNA model (a), free DNA model (b), and aptamer model (c). Black dots indicate the attachment points of DNA to the protein.** The image of the protein (streptavidin) was generated using SWISS-MODEL software **[61]**. See text for details.

Here, we present a mobility model that is designed to satisfy the close multi-point contact between a protein and an aptamer (**Figure 2.1c**). Our model assumes a spherical shape for such

tight protein–aptamer complexes. By using basic rules governing migration of molecules in gel-free electrophoresis we derived the dependence of the mobility of a spherical protein–ssDNA complex ($\mu_{\mathrm{P–ssDNA}}$) on the mobility of ssDNA (μ_{ssDNA}), the number of nucleotides in ssDNA (N), the molecular weight of the complex ($\mathrm{MW_{P–ssDNA}}$), and two empirical parameters (A and B):

$$\mu_{\mathrm{P–ssDNA}} = A + B\left(\mu_{\mathrm{ssDNA}}\ N^{0.68}\ \mathrm{MW_{P–ssDNA}}^{-1/3}\right) \qquad (2\text{-}1)$$

where 0.68, the power of N, was previously determined by Slater and co-authors **[62]**. The values of A and B were found by fitting **eq (2-1)** into the experimental dependence of $\mu_{\mathrm{P–ssDNA}}$ on $\{\mu_{\mathrm{ssDNA}}\ \mathrm{N}^{0.68}\ \mathrm{MW_{P–ssDNA}}^{-1/3}\}$ obtained for seven different protein–ssDNA complexes.

The empirical model was successfully cross-validated by excluding one protein–DNA complex from the set of seven and finding pairs of A and B for seven subsets of six protein–DNA complexes (leave-one-out cross validation, LOOCV). The resulting seven pairs of A and B were close to each other, and residual standard deviations of A and B were only 2.2 and 1.7%, respectively. Such a precise determination of these constants allowed accurate mobility prediction for the excluded protein–DNA complexes with only a 3.0% maximum deviation from the experimentally determined mobility values. The model was tested on a set of three protein–aptamer complexes which were not a part of the training/ cross-validation set: human α-thrombin with its 29-nt aptamer, ABH2 with its 80-nt aptamer, and AlkB with its 80-nt aptamer (the complex with AlkB was analyzed at a pH value different from that used for the initial set of experiments with this protein). In all cases, the difference between the predicted mobility and the experimental one was below 5%, which confirmed model validity. The developed model will be indispensable in rational design of CE-based selection of DNA aptamers for protein targets.

Moreover, it strongly suggests the validity of the spherical-shape assumption for the protein-aptamer complexes when considering complex mobility.

2.2. MATERIALS AND METHODS

2.2.1. Chemicals and materials

All chemicals were purchased from Sigma-Aldrich (Oakville, ON), unless otherwise stated. Fused silica capillary with inner and outer diameters of 75 and 360 μm, respectively, was purchased from Polymicro (Phoenix, AZ). Chromeo P503 pyrylium dye was purchased from Active Motif (Burlington, ON). MutS and AlkB proteins were expressed and purified by us according to the published procedures **[63]**. SSB was purchased from Promega (Madison, WI). Streptavidin was purchased from New England Biolabs (Whitby, ON). ABH2 protein was kindly provided by Professor Christopher J. Schofield (University of Oxford, England). All DNA sequences were custom-synthesized by IDT (Oralville, IA). The sequences of ssDNA used in experiments with SSB were as follows: 5′-/Alex488N/TAA TAT TAT TGC AAA TAA ATT TAA A-3′ (25-nt); 5′-/Alex488N/AAA TTA AAG GGA ACG TAT ATA CAA CGC AAA GAA GCT GGA AAA TTG GCG AGA GAA TCT TCT TTC TGT CTA TCG AAG-3′ (75-nt); 5′-/56-FAM/CTT CTG CCC GCC TCC TTC CTG GTA AAG TCA TTA ATA GGT GTG GGG TGC CGG GCA TTT CGG AGA CGA GAT AGG CGG ACA CT-3′ (80-nt); 5′-/56-FAM/CTC CTC TGA CTG TAA CCA CGT GCC TAG CGT TTC ATT GTC CCT TCT TAT TAG GTG ATA ATA GCA TAG GTA GTC CAG AAG CCG ACA ACT G-3′ (88-nt). The sequence of MutS aptamer was exactly the same as 80-nt ssDNA sequence used for experiments with SSB. The sequence of 80-nt AlkB aptamer was 5′-/56-FAM/CTC CTC TGA CTG TAA CCA CGT GCC TAG CGT TTC ATT GTC CCT TCT TAT TAG GTG ATA ATA GCA TAG GTA GTC CAG AAG CC-3′. The sequence of the 80-nt ABH2 aptamer was 5′-/56-FAM/CTC CTC TGA

CTG TAA CCA CGT GAC CGG TTC AGA TGG ATG GGT CCT TCG CAT AGA CCT ACA GCA TAG GTA GTC CAG AAG CC-3′. The aptamer pool for experiments with streptavidin was generated from ssDNA library that contained a random region (N) of 40 nucleobases, using IFCE **[60]**. The sequence of the library was as follows: 5′-/Alex488N/CTA CGG TAA ATC GGC AGT CAN NNN NNN NNN NNN NNN NNN NNN NNN NNN NNN NNN NNN NNN ATC TGA AGC ATA GTC CAG GC-3′. The pool was used instead of an individual aptamer as no individual aptamers had been identified and characterized from this pool before our experiments. IFCE-based partitioning leads to very pure aptamer pools. Therefore, the peak of the protein–ssDNA complex was narrow for the pool, and complex mobility could be determined accurately. Hence, there was no adverse effect of using the aptamer pool instead of an individual aptamer. Tris-HCl (50 mM final concentration, pH 8.0) was used as both a running buffer and a dilution buffer for all CE experiments with SSB. In a case of CE experiments with MutS, AlkB, streptavidin and ABH2, Tris-acetate (50 mM final concentration, pH 8.2) was used as both a running buffer and a dilution buffer. The reason for the buffer change from Tris-HCl to Tris-acetate for protein–aptamer complexes was a noticeably better reproducibility of these experiments conducted in Tris-acetate. Borax (25 mM, pH 9.2) was used as both, a running buffer and a dilution buffer for the additional experiment with the AlkB–aptamer complex. Bodipy (Difluoro{2-[1-(3,5-dimethyl-2H-pyrrol-2-ylidene-N)ethyl]-3,5-dimethyl-1H-pyrrolato-N}boron) and fluorescein were used as internal standards for accurate mobility calculations. All solutions were prepared in deionized water filtered through a 0.22 μm Milipore filter membrane (Nepean, ON).

2.2.2. Capillary electrophoresis

All CE experiments were done with a P/ACE MDQ instrument from Sciex (Brea, CA) using LIF detection with excitation at 488 nm. Fluorescence of FAM- and Alexa-labeled species was detected at 520 nm; fluorescence of Chromeo-labeled proteins was detected at 630 nm. An uncoated fused silica capillary of an 82.2-cm total length and a 72.0-cm distance from the inlet to the detection point was used for all CE experiments with SSB and SSB-ssDNA complexes. For the rest of the proteins and their complexes with aptamers, an uncoated fused silica capillary of a 49.5-cm total length and a 39.5-cm distance from the inlet to the detection point was used. A shorter capillary was chosen for protein–aptamer complexes to ensure that they were intact by the time they reached the detector. New capillaries were preconditioned by washing them with 10 capillary volumes of methanol. In the beginning of each run, the capillary was rinsed with 8 capillary volumes of each of 100 mM HCl, 100 mM NaOH, deionized water, and the running buffer. The samples were injected into the capillary by pressure of 0.5 psi (3.45 kPa) applied for 10 s. Capillary coolant temperature was kept at 25 °C. In the case of an 82.2-cm-long capillary, an electric field of 304 V/cm with a positive electrode at the inlet was applied to carry out the electrophoresis. In the case of 49.5-cm-long capillary, the electric field strength was 505 V/cm. The electric field strength was 404 V/cm in the additional experiments with the AlkB-aptamer complex conducted at pH 9.2.

2.2.3. Sample preparation

MutS, AlkB, SSB and streptavidin (**Table A1**) were labeled with the Chromeo P503 dye according to the dye manufacture's procedure, and were further analyzed with CE separately from ssDNA. The final concentrations of the proteins were 0.6, 1.9, 3.8, and 2 μM for MutS, AlkB, streptavidin, and SSB (per tetramer), respectively. For CE analysis of protein–ssDNA

complexes, an unlabeled protein was mixed with ssDNA and incubated for 1 h at room temperature. A marker of the electroosmotic flow (bodipy) was added to all samples prior to their injection into the capillary. Fluorescein was added as an additional internal standard in experiments with SSB. In our experiments with AlkB, ABH2, and streptavidin, the peak of fluorescein partially overlapped with the peaks of protein–aptamer complexes due to their similar electrophoretic mobilities (−26.64 $mm^2kV^{-1}s^{-1}$ for fluorescein and −25.63, −24.58, and −25.96 $mm^2kV^{-1}s^{-1}$ for AlkB at pH 9.2 and 8.2 and ABH2 at pH 8.2, respectively), which made us exclude fluorescein from these experiments and from all experiments that were conducted in a shorter capillary, for consistency. For CE experiments with Chromeo-labeled MutS, streptavidin, and AlkB, the concentration of bodipy was 1 μM. The concentration of bodipy in the mixtures of unlabeled MutS, AlkB, and strepatavidin with their aptamers was 100 nM. Bodipy concentration was 500 nM and fluorescein concentration was 50 nM for all CE experiments with SSB protein and SSB–ssDNA mixtures. The concentration of any ssDNA (including aptamers) in mixture with a protein was 100 nM, except for the streptavidin–aptamer pool mixture, where the concentration of the aptamer pool was 20 nM. Concentrations of the unlabeled protein in the mixtures with ssDNA (including aptamers) were 280, 250, 317, and 100 nM for MutS, AlkB, streptavidin, and SSB (per tetramer), respectively. In the experiments with ABH2 and AlkB at pH 9.2, concentrations of the protein in the mixture with the corresponding aptamer were 270 nM for ABH2 and 200 nM for AlkB. The concentration of 80-nt ssDNA in the experiment with AlkB at pH 9.2 was 200 nM. Migration times of proteins, ssDNA, and their complexes along with migration time of bodipy were used to find μ_{ssDNA} and $\mu_{P–ssDNA}$. All experiments were performed in triplicates.

2.3. RESULTS AND DISCUSSION

2.3.1. Development of the model

In the following consideration, subscripts P, ssDNA, and P–ssDNA assign parameters to protein, ssDNA, and protein–ssDNA complex, respectively. No distinction between aptamers and other ssDNA is made. The charge of the protein–ssDNA complex is a superposition of charges of the protein and ssDNA:

$$q_{\text{P–ssDNA}} = aq_{\text{P}} + bq_{\text{ssDNA}} \tag{2-2}$$

where all charges take into consideration the ionic atmosphere around the molecule, and coefficients *a* and *b* adjust contributions of the individual charges to the resulting charge in consideration of rearrangements of the ionic atmosphere for each component as a result of complex formation.

A ratio between the charge in **eq (2-2)** and a corresponding frictional coefficient of the species represents its electrophoretic mobility (μ):

$$\mu = q / f \tag{2-3}$$

Expressing charges in **eq (2-2)** through electrophoretic mobilities and frictional coefficients of the components leads to the following equation for the electrophoretic mobility for the complex:

$$\mu_{\text{P–ssDNA}} = \frac{a\mu_{\text{P}} f_{\text{P}} + b\mu_{\text{ssDNA}} f_{\text{ssDNA}}}{f_{\text{P–ssDNA}}} \tag{2-4}$$

In **eq (2-4)**, electrophoretic mobilities of the protein and ssDNA can be easily determined experimentally, while frictional coefficients of the protein, ssDNA, and the complex cannot be easily measured. However, the frictional coefficients can be expressed through known size

characteristics, such as molecular weights of the species and the numbers of nucleotides in ssDNA.

In order to express the frictional coefficients for the protein and the complex through their molecular weights, we use the assumption of spherical shapes for both the protein and the protein–ssDNA complex (**Figure 1.1c**). The volume of the protein can be expressed as a volume of the sphere that encloses the protein **[64**,**65]**:

$$V_{\mathrm{P}} = \frac{4}{3}\pi R_{\mathrm{P}}{}^{3} \tag{2-5}$$

Here, R_P is the protein's hydrodynamic radius which also defines its frictional coefficient according to Stokes' law:

$$f_{\mathrm{P}} = 6\pi\eta R_{\mathrm{P}} \tag{2-6}$$

where η is the dynamic viscosity of solution.

On the other hand, the volume of a globular biopolymer can be expressed as a ratio between its molecular weight (MW) and its density (ρ). Accordingly, we can write for the volume of the protein molecule **[65]**:

$$V_{\mathrm{P}} = \frac{\mathrm{MW_P}}{\rho_{\mathrm{P}}} \tag{2-7}$$

The combination of **eqs (2-5) – (2-7)** links the frictional coefficient of the protein with its molecular weight and density:

$$f_{\mathrm{P}} = 6\pi\eta\left(\frac{3}{4\pi\rho_{\mathrm{P}}}\right)^{\frac{1}{3}}\mathrm{MW_P}^{\frac{1}{3}} \tag{2-8}$$

In the first approximation, the density of the protein and the viscosity of the solution are constant. **Eq (2-8)** can then be rewritten as:

$$f_{\mathrm{P}} = \alpha \mathrm{MW}_{\mathrm{P}}^{\frac{1}{3}},\ \alpha = 6\pi\eta\left(\frac{3}{4\pi\rho_{\mathrm{P}}}\right)^{\frac{1}{3}} \tag{2-9}$$

where α is a constant. Since we assume the protein–ssDNA complex to have a spherical shape, **eqs (2-5) – (2-9)** can also be applied to the complex, and the final equation for f of the complex will be similar to **eq (2-9)** but with a different constant β, which includes the density of the complex:

$$f_{\mathrm{P–ssDNA}} = \beta \mathrm{MW}_{\mathrm{P–ssDNA}}^{\frac{1}{3}},\ \beta = 6\pi\eta\left(\frac{3}{4\pi\rho_{\mathrm{P–ssDNA}}}\right)^{\frac{1}{3}} \tag{2-10}$$

The last unknown parameter — the frictional coefficient of the ssDNA — can be expressed through the number of nucleotides (N) in ssDNA **[62]:**

$$f_{\mathrm{ssDNA}} = \gamma N^{0.68} \tag{2-11}$$

where γ is a constant coefficient of proportionality.

At this point, all frictional coefficients in **eq (2-4)** can be replaced with the corresponding right-hand sides of **eqs (2-9) – (2-11)**, and the equation for $\mu_{\mathrm{P–ssDNA}}$ will be:

$$\mu_{\mathrm{P–ssDNA}} = \frac{\alpha a \mu_{\mathrm{p}} \mathrm{MW}_{\mathrm{P}}^{\frac{1}{3}} + \gamma b \mu_{\mathrm{ssDNA}} N^{0.68}}{\beta \mathrm{MW}_{\mathrm{P–ssDNA}}^{\frac{1}{3}}} \tag{2-12}$$

For simplicity, $\alpha a/\beta$ and $\gamma b/\beta$ can be replaced with empirical constants C and B, respectively, and ($\mathrm{MW_{P–ssDNA}}$) can be represented as a sum of $\mathrm{MW_P}$ and $\mathrm{MW_{ssDNA}}$:

$$\mu_{\text{P–ssDNA}} = \frac{C\mu_{\text{P}}\text{MW}_{\text{P}}^{\frac{1}{3}}}{(\text{MW}_{\text{P}} + \text{MW}_{\text{ssDNA}})^{\frac{1}{3}}} + \frac{B\mu_{\text{ssDNA}}N^{0.68}}{(\text{MW}_{\text{P}} + \text{MW}_{\text{ssDNA}})^{\frac{1}{3}}} \tag{2-13}$$

The empirical constants *C* and *B* can be obtained from **eq (2-13)** using experimental data for electrophoretic mobilities of the species, along with molecular weights of the protein and ssDNA and the number of nucleotides in ssDNA. In such a case, the known parameters in **eq (2-13)** can be grouped as *X* and *Y* and the resulting $\mu_{\text{P–ssDNA}}$ will be a linear combination of *X* and *Y*:

$$\begin{aligned} \mu_{\text{P–ssDNA}} &= BX + CY \\ Y &= \frac{\mu_{\text{P}}\text{MW}_{\text{P}}^{\frac{1}{3}}}{(\text{MW}_{\text{P}} + \text{MW}_{\text{ssDNA}})^{\frac{1}{3}}} \\ X &= \frac{\mu_{\text{ssDNA}}N^{0.68}}{(\text{MW}_{\text{P}} + \text{MW}_{\text{ssDNA}})^{\frac{1}{3}}} \end{aligned} \tag{2-14}$$

In a case of *n* different protein–ssDNA complexes, **eq (2-14)** turns into a group of equations, which can be written in a matrix form:

$$\text{M}\cdot\Phi = \Psi,\ \text{M} = \begin{pmatrix} X_1 & Y_1 \\ X_2 & Y_2 \\ \cdots & \cdots \\ X_n & Y_n \end{pmatrix},\ \Phi = \begin{pmatrix} B \\ C \end{pmatrix} \text{ and } \Psi = \begin{pmatrix} \mu_{\text{P–ssDNA 1}} \\ \mu_{\text{P–ssDNA 2}} \\ \cdots\cdots\cdots \\ \mu_{\text{P–ssDNA }n} \end{pmatrix} \tag{2-15}$$

Φ is a desired vector of solutions, which can be found as:

$$\Phi = [(\text{M}^{\text{T}}\text{M})^{-1}\text{M}^{\text{T}}]\cdot\Psi \tag{2-16}$$

Here, the upper indexes “T” and “−1” corresponds to the transpose and inverse of the corresponding matrices, respectively.

2.3.2. Finding μ_{p}, μ_{ssDNA}, and $\mu_{p\text{-}ssDNA}$ for a set of proteins and ssDNAs

For our initial study, we have chosen seven protein–ssDNA complexes. Among them we had three protein–aptamer complexes for three proteins — MutS, streptavidin and AlkB — and their 80-nucleotide-long aptamers. The remaining four complexes are SSB–ssDNA complexes with ssDNA of different lengths: 88, 80, 75 and 25 nucleotides. ssDNA wraps the SSB molecule and creates a tight complex similar to protein–aptamer complexes (**Figure 2.1c**). Using SSB allowed us to vary the length of ssDNA and ensure that the model is suitable for a range of aptamer lengths. SSB is known for its ability to co-operatively bind ssDNA molecules, resulting in the formation of the complexes with different stoichiometry. These complexes can form when SSB is present in a high molar access in respect to ssDNA **[66]**. In our experiments, we used SSB and ssDNA in equal molar ratios in order to allow for the formation of only SSB-ssDNA complexes with one-to-one stoichiometry. The main three reasons for choosing MutS, AlkB, streptavidin, and SSB for the training set in model development were as follows. First, the range of molecular weights of the chosen proteins (24–95 kDa with a median value of 60 kDa) covers above 80% of proteins found in the eukaryotic proteome (49 ± 48 kDa) **[48**,**67]**. Second, the range of isoelectric point values (pI) of the selected four proteins (starting from 5.61 for MutS and ending up with 8.36 for streptavidin) also covers most of pI values of proteins in the eukaryotic proteome **[48]**. Third, the chosen proteins are suitable for CE analysis in uncoated capillaries since they do not adsorb significantly onto bare fused silica.

We carried out CE experiments to find $\mu_{P,}$ μ_{ssDNA}, and $\mu_{P–ssDNA}$. In one set of experiments, we found μ_{ssDNA} and $\mu_{P–ssDNA}$. In these experiments, a mixture of an unlabeled protein with a corresponding fluorescently-labeled ssDNA was subjected to CE with LIF detection (excitation at 488 nm and emission at 520 nm). Each electropherogram contained two peaks that

corresponded to fluorescently-labeled molecules: protein–ssDNA complex and ssDNA (**Figure 2.2**). In another set of experiments, we labeled proteins with a small electrically-neutral fluorescent dye, ChromeoP503, that did not affect their mobility significantly **[68**,**69]**. The four ChromeoP503-labeled proteins were then run in CE one by one and separately from DNA using LIF detection (excitation at 488 nm, emission at 630 nm) (**Figure 2.2**).

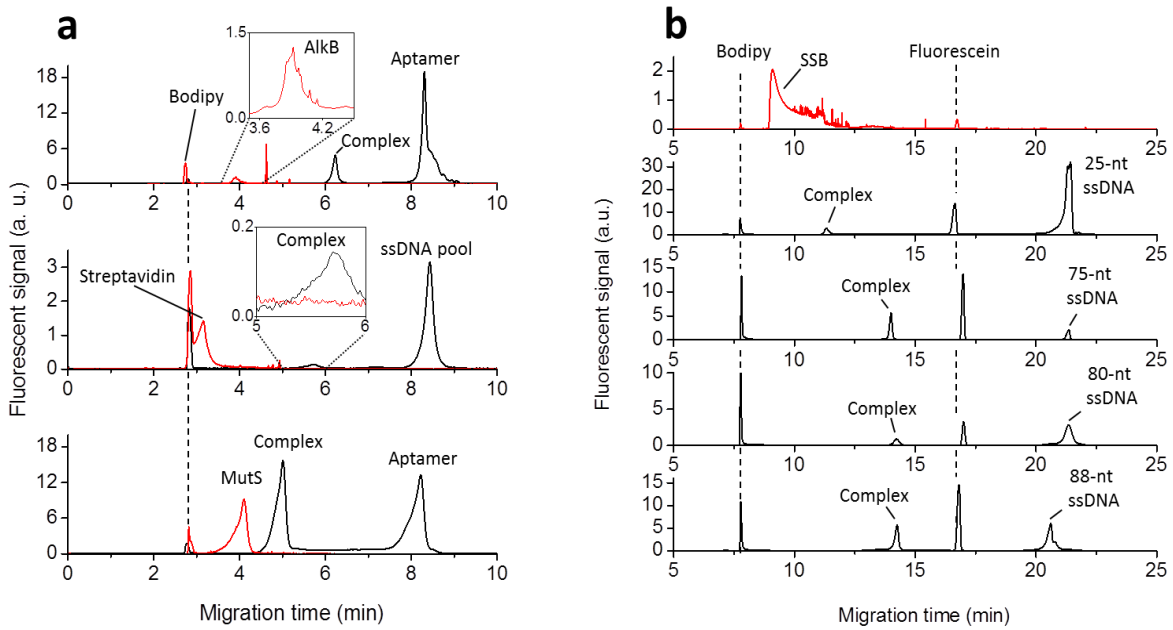

Figure 2.2. Electropherograms utilized for the determination of electrophoretic mobilities of four proteins (red) and seven protein−aptamer complexes and ssDNAs (black). Panel a shows electropherograms obtained for AlkB, MutS, and streptavidin as well as their aptamers and protein− aptamer complexes. Panel b shows electropherograms for SSB, four ssDNAs of different lengths, and SSB−ssDNA complexes. Bodipy was used as an electroosmotic flow marker and internal standard in all experiments; fluorescein was used as an additional internal standard in experiments with SSB. See Materials and Methods for other experimental conditions. Electropherograms were aligned with each other using positions of the internal standards as references. Electrophoretic mobilities were determined using experimental migrations times, capillary lengths, and electric field strengths as described in the text.

We did not mix the labeled proteins with ssDNA to avoid any possible effects of the ChromeoP503 label on complex formation. The obtained electropherograms for the labeled proteins and the electropherograms for the mixtures of unlabeled proteins and ssDNA were overlaid with each other and aligned according to the positions of internal standards. In all CE experiments that were conducted in an 80.2-cm-long capillary, we used peaks of two internal standards (bodipy and fluorescein) for such an alignment. In our CE experiments with a 49.5-cm-long capillary, only the peak of bodipy was used for such an alignment (see Materials and Methods). Being an electrically-neutral molecule, bodipy also served as an EOF marker, and its migration time (t_{bodipy}) was used to calculate the mobility of EOF (μ_{EOF}):

$$\mu_{EOF} = \frac{L}{Et_{bodipy}} \tag{2-17}$$

Here, L is a capillary length from the inlet to the detection point, and E is the electric field strength.

The obtained values for μ_{EOF} along with the migration times (t) of the proteins, ssDNA, and the complexes were then used to find electrophoretic mobilities of the species (**Table A2**) using the following equation:

$$\mu = \frac{L}{Et} - \mu_{EOF} \tag{2-18}$$

The experimentally-found mobility values were used in the determination of empirical constants as explained in the next section.

2.3.3. Finding empirical constants

Empirical constants B and C for a training/ cross-validation set of seven protein–ssDNA complexes (n = 7) were calculated by solving **eq (2-15)** using the experimentally obtained values

of μ_P, μ_{ssDNA}, $\mu_{P-ssDNA}$, along with known $MW_{P-ssDNA}$, MW_{ssDNA}, and N as input parameters (**Table 2.1**, **sub-table a**). Similar calculations were conducted for seven groups of six protein–ssDNA complexes (n = 6). These groups were obtained by excluding one protein–ssDNA complex from the full set of seven. Pairs of B and C for the seven subsets of six protein–ssDNA complexes were determined in order to perform LOOCV of the empirical model (**Table 2.1**, **sub-table b**). It was found that constant B did not vary much across the seven subsets; its residual standard deviation (RSD) was only 4.6%. However, constant C varied significantly resulting in RSD of 55%; such a large RSD indicated that parameter C could not be considered as a constant. The most noticeable deviation of parameter C from its average value was observed for the complex of SSB protein with 25-nt ssDNA. For this complex, parameter C is equal to 0.001, which corresponds to very low values of parameter a. Thus, the rearrangement of the ionic atmosphere of SSB upon binding to short 25-nt ssDNA was negligible in comparison to the rearrangements that occurred in all other cases of complexes formation between the proteins and longer ssDNA molecules (75, 80 and 88-nt long). Our comparison of variables X and Y showed that variable X was greater than Y by more than an order of magnitude (**Table A3**). The large variation of C along with relative smallness of Y suggested that BX dominated over CY in defining the value of $\mu_{P-ssDNA}$ in **eq (2-14)**. Accordingly, we replaced CY in **eq (2-14)** with a constant A and obtained the following simplified equation:

$$\mu_{P\text{-}ssDNA} = A + BX \tag{2-19}$$

Linear fitting of mobility data for seven protein–ssDNA complexes with **eq (2-19)** resulted in B = 0.0929 $kDa^{-1/3}$ and A = −9.95 $mm^2kV^{-1}s^{-1}$ with a correlation coefficient (R^2) of 0.989 (**Figure 2.3**).

Table 2.1. **Parameters of the mobility-prediction model.**

Sub-table a: Calculated constants *C* and *B* for protein-ssDNA complexes, presented in **Figure 2.2**. The constants were calculated with **eq (2-14)** for seven protein–ssDNA complexes, as well as for the cases when one protein–DNA complex was excluded from the complete set of seven. Sub-table b: constants *A* and *B* derived from **eq (2-19)** by the linear regression analysis. Constants *A* and *B* were calculated for seven protein–ssDNA complexes, as well as for the cases when one protein–DNA complex was excluded from the complete set of seven. The row "$\Delta\mu$, %" shows deviations of the calculated mobility from the experimental values for the excluded protein–ssDNA complexes.

Constants	**All complexes**	Cross-validation for $\mu_{P–ssDNA} = BX + CY$ **(sub-table a)**							
		Excluded complexes							
		SSB–25-nt ssDNA	**SSB–75-nt ssDNA**	**SSB–80-nt ssDNA**	**SSB–88-nt ssDNA**	**AlkB–Apt**	**Strepta-vidin–Apt**	**MutS–Apt**	**RSD %**
C	0.236	0.001	0.245	0.245	0.230	0.284	0.332	0.470	55
B, $kDa^{1/3}$	0.151	0.164	0.152	0.152	0.154	0.154	0.145	0.142	4.6
		Cross-validation for $\mu_{P–ssDNA} = A + BX$ **(sub-table b)**							
A, $mm^2kV^{-1}s^{-1}$	–9.95	–9.44	–9.93	–9.97	–9.93	–9.79	–10.00	–10.11	2.2
B, $kDa^{1/3}$	0.0929	0.0966	0.0929	0.0923	0.0932	0.0944	0.0922	0.0924	1.7
R^2	0.988	0.923	0.989	0.991	0.988	0.987	0.989	0.996	
$\Delta\mu$, %		1.9%	0.9%	1.8%	0.5%	1.0%	1.3%	3.0%	

The LOOCV approach that we previously used with **eq (2-14)** was also applied to **eq (2-19)**, resulting in a very small RSD values of 2.2% and 1.7% for constants *A* and *B*, respectively (**Table 2.1**, sub-table b). Such a precise determination of *A* and *B* allowed accurate mobility prediction for the excluded protein–ssDNA complexes with only a 3.0% maximum deviation from the experimentally determined mobility (**Table 2.1**, sub-table b). Considering the obtained values of A and B for seven protein–ssDNA complexes, **eq (2-17)** can be rewritten as follows:

$$\mu_{P–ssDNA} = -9.95 + 0.0929\mu_{ssDNA}\ N^{0.68}\ MW_{P–ssDNA}{}^{-1/3} \qquad (2\text{-}20)$$

Eq (2-20) allows the prediction of $\mu_{\text{P-ssDNA}}$, using only the molecular weight of the complex, the number of nucleotides in ssDNA, and μ_{ssDNA} as input parameters. This equation resembles equations previously obtained for predicting electrophoretic mobilities of proteins, based on the assumption of their spherical shapes **[64**,**70]**. The main difference between **eq (2-20)** and the equations for protein mobilities is in the component that corresponds to the charge of the molecule. In **eq (2-20)**, the charge of the P–ssDNA complex appears to be mainly defined by the charge of ssDNA which depends on N, whereas in the mobility models for proteins, the charge of the protein is considered to be equal to the net charge of its amino acid sequence **[70]**.

As a side note, we would like to mention that **eq (2-20)** can be further simplified by assuming μ_{ssDNA} to be constant for buffers with low ionic strengths. The validity of such an assumption is based on a weak dependence of μ_{ssDNA} on the ionic strength (I) of the running buffer for I ranging between 0 and 0.05 M is **[7]**:

$$\mu_{\text{ssDNA}} = -31.6 + 53.2I - 29.1I^2 \ (\text{mm}^2\text{kV}^{-1}\text{s}^{-1}) \qquad (2\text{-}21)$$

According to **eq (2-21)**, a 2-fold change in I changes μ_{ssDNA} by only a factor of ≈1.04. In our experiments (I = 0.027 M for Tris-HCl and I = 0.021 M for Tris-Acetate), the average experimentally determined value of μ_{ssDNA} was −31.4 $\text{mm}^2\text{kV}^{-1}\text{s}^{-1}$ with RSD of 3.2%. The values of μ_{ssDNA} calculated with **eq (2-21)** for I = 0.021 M and 0.027 M are −30.5 and −30.2 $\text{mm}^2\text{kV}^{-1}\text{s}^{-1}$, respectively. Using either of these values (−31.4 or −30.5 and −30.2 $\text{mm}^2\text{kV}^{-1}\text{s}^{-1}$) in **eq (2-20)** instead of experimentally determined μ_{ssDNA} does not cause significant increase in deviations of calculated $\mu_{\text{P-ssDNA}}$ from the experimentally determined ones (**Table A4**).

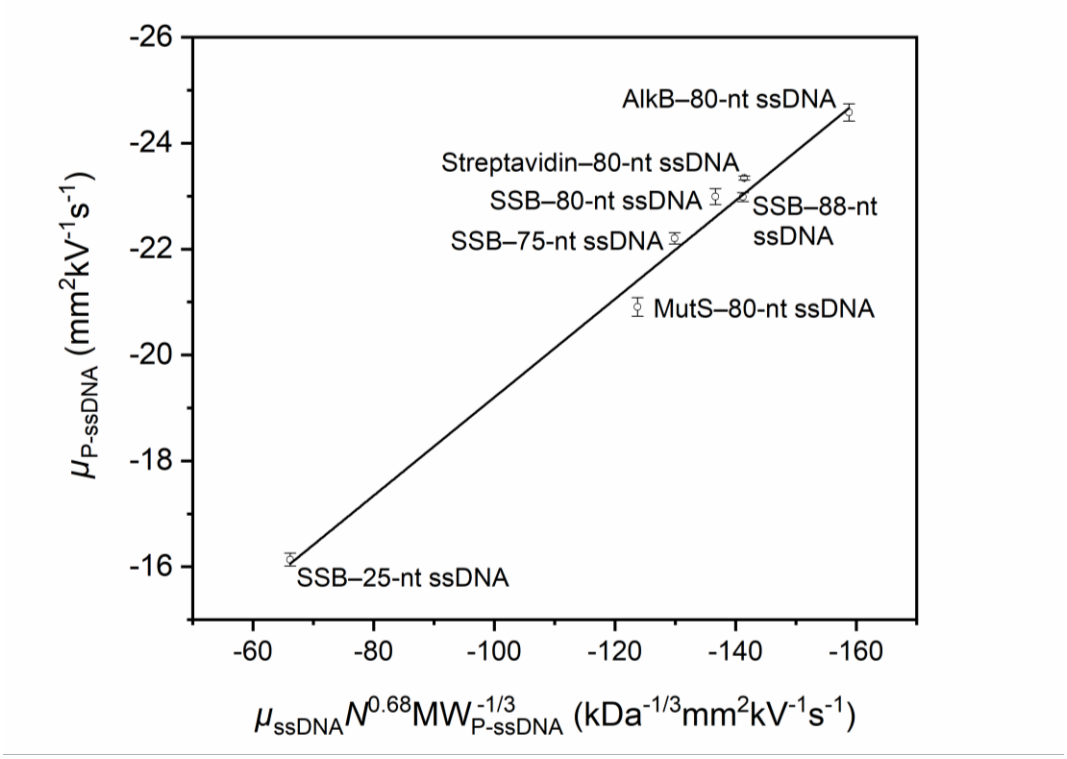

Figure 2.3. Line of the best fit for the electrophoretic mobility of protein−ssDNA complex as a function of X: $\mu_{P-ssDNA} = A + BX$, where X = $\{\mu_{P-ssDNA}\ N^{0.68}\ MW^{-1/3}{}_{P-ssDNA}\}$. Calculated values for *A* and *B* were −9.95 $mm^2\ kV^{-1}\ s^{-1}$ and 0.0929 $kDa^{-1/3}$, respectively. The correlation coefficient was R^2 = 0.988. Error bars correspond to the standard deviations of the experimentally determined electrophoretic mobilities.

In order to test our model and further improve it, we used a set of three additional protein–aptamer complexes (**Table A5**): a complex of human α-thrombin with its 29-nt aptamer at pH 7.5, a complex of ABH2 with its 80-nt aptamer at pH 8.2, and a complex of AlkB with its 80-nt aptamer at pH 9.2. The mobility value for the α-thrombin–aptamer complex was calculated from the literature data on migration times of the complex and an EOF marker **[71]**. The mobility values for the other two complexes were obtained by us experimentally (**Figure A.1**). AlkB with the 80-nt aptamer was analyzed at a higher pH value than the pH value used for the training/ cross-validation set in order to change the calculated charge of the protein (from -1 at pH 8.2 to -

5.3 at pH 9.2). The three predicted mobilities were different from the experimental ones by only 1.8%, 4.6% and 0.1%, for human α-thrombin, ABH2, and AlkB, respectively (**Table A6**). We further added the three new data points to our original set of seven points and recalculated empirical constants A and B (**Figure A.2**). The new constants were A = −9.47 $mm^2kV^{-1}s^{-1}$ and B = 0.0973 $kDa^{-1/3}$ and differed from the original ones by only 4.8% and 4.7%, respectively. The correlation coefficient was R^2 = 0.987. Cumulatively, these results confirm the validity of our mobility predictor.

2.4. CONCLUSIONS

To conclude, we developed an empirical mathematical model for accurate prediction of the electrophoretic mobility of protein–aptamer complexes in the approximation of their spherical shapes. The relationship between the mobility of protein–aptamer complex and the three known parameters (mobility of ssDNA, length of ssDNA, and molecular weight of the complex) was derived using fundamental principles of molecular physics. In this respect, our model differs from predictive models based solely on the empirical correlation between parameters **[72]**. Empirical models require very large training sets to establish the correlation; in our case the correlation is deterministic, and a much smaller training set should suffice finding the slope and intercept of the straight line. Our model was built upon mobility data for seven protein–ssDNA complexes, including four complexes of SSB with ssDNA of different lengths wrapped around SSB and three protein–aptamer complexes for MutS, streptavidin, and AlkB with their 80-nucleotide-long aptamers. The model suggests that the mobility of the protein–ssDNA complex depends linearly on a multiplication product of three terms that can be easily found: the electrophoretic mobility of ssDNA identical in length to the aptamer, the number of nucleotides in the aptamer to the power of 0.68, and a sum molecular weight of the protein and the aptamer

to the power of negative 1/3. The linear dependence includes two empirical parameters (line's intercept *A* and line's slope *B*) which were determined by fitting the model into a data set obtained for the seven protein–ssDNA complexes. The best-fit linear function had a correlation coefficient of 0.988. The developed model will serve as a reliable predictor of protein–aptamer complex mobility. We expect that this model will be applicable to a wide range of pH values because the only pH-dependent parameters in the model are the unknown mobility of the complex and the experimentally-found mobility of DNA. When pH changes, the mobility of DNA changes, and this change leads to a change in the predicted mobility of the complex as was confirmed with the AlkB–aptamer complex studied at pH 8.2 and 9.2 (**Table A4**, **Table A6**). We do not exclude the possibility that the extremely big charge of the protein (similar to the one of DNA) may result in a noticeable difference between the experimental and predicted mobilities of the complex. However, the impact of protein charge on model performance can be minimized by reducing the protein charge via adjusting pH of the running buffer to the value of the protein's pI. The a priori knowledge of complex mobility will allow the accurate blind setting of the time window for complex collection in aptamer partitioning by NECEEM **[59]**. It will also facilitate rational selection of the number of fractions to be blindly collected in the IFCE-based partitioning **[60]**. Overall, this mobility predictor will be an indispensable tool in rational design of CE-based aptamer selection. We expect that our model will be applicable to proteins with molecular weights outside the investigated range; however, it is likely that the accuracy of prediction will be lower for such proteins.

CHAPTER 3. UNEXPECTED ELECTROPHORETIC BEHAVIOUR OF COMPLEXES BETWEEN ROD-LIKE VIRONS AND BIVALENT ANTIBODIES

The presented material was published previously and reprinted with permission from "Beloborodov, S. S.; Panferov, V. G.; Safenkova, I. V.; Krylova, S. M.; Dzantiev, B. B.; Krylov, S. N. Unexpected Electrophoretic Behavior of Complexes between Rod-like Virions and Bivalent Antibodies. *Analytical Chemistry* 2016, 88 (23), 11908–11912." Copyright 2016 American Chemical Society.

My contributions to the article were: (i) performing all presented CE experiments (ii) preparing all figures, (iii) developing empirical formula for predicting the number of cross-linked virions (iv) interpreting results and (v) writing a major part of the manuscript. AFM imaging of the complexes was done by the Dzantiev lab.

3.1. INTRODUCTION

Non-covalent reversible binding of biomolecules plays an important role in cellular regulation, cell-cell recognition, bacterial and viral infections, etc **[73**–**76]**. Reversible binding reactions are the major component of molecular diagnostic tools such as immunoassays and hybridization assays **[77**,**78]**. In addition, most modern drugs are designed to reversibly bind their therapeutic targets **[79**–**81]**. There are multiple approaches to study non-covalent complexes; among them are biosensor-based (e.g. surface plasmon resonance and biolayer interferometry) **[82**,**83]**, calorimetry based (e.g. isothermal titration calorimetry) **[84]**, and separation-based (e.g. capillary electrophoresis (CE)) **[85]**. Every approach is characterized by a unique set of advantages and limitations and, to some extent, different approaches complement each other. In particular, CE requires no immobilization and provides very high sensitivity for

low sample consumption but limited to species with different electrophoretic mobilities. In a typical CE migration pattern, a molecular complex has an electrophoretic mobility intermediate to those of the components it is built of. The reason for this is the approximate additivity of charges and frictional coefficients. The good examples are interactions between: proteins and proteins **[49]**, proteins and DNA **[86**–**88]**, antibodies and spherical virus particles (virions) **[14**,**15]**. To the best of our knowledge, deviations from this rule have not been previously reported. Here we report that the rule of intermediate electrophoretic mobility is broken for complexes of rod-like viruses with bivalent antibodies.

Virions have a variety of shapes with the most common being spherical. Complexes of spherical virions with bivalent antibodies were studied with CE and were found to have mobilities intermediate to those of free virion and free antibody **[14**,**15]**. Some highly pathogenic viruses, including deadly Ebola **[89]** and Marburg **[90]** viruses, are spread by virions of rod-like shapes. This work was originally motivated by the fact that while being a highly-informative tool, CE has not been used to study complexes of rod-like virions with antibodies. We performed such a study using virions of Potato Virus X (PVX) and its monoclonal antibodies (Ab) both native bivalent (full-size) and monovalent derivative (Fab fragments). PVX is a plant virus that belongs to the genus *Potexvirus*, family *Alphaflexiviridae*, containing around 40 species. Potexvirus virions are flexible cylinders of various lengths (470-580 nm) and of an identical diameter of 13 nm **[91]**. The average molecular weight of Potexvirus virions is approximately 35,000 kDa. The RNA:protein weight ratio of virions is 6:94. Each virion contains approximately 1,300 identical coat protein subunits (8.9 subunits per turn of primary helix).

We found that the absolute values of mobilities for complexes of PVX virions with bivalent antibodies were greater than those of unbound virions and antibodies. However, the

same virions with the monovalent derivative of the same antibodies did not result in complexes with greater absolute values of mobilities. We hypothesized that the deviation from the rule of complex mobilities was due to the formation of aggregates of virions cross-linked with bivalent antibodies; the monovalent antibodies cannot do such cross-linking. The theory suggests that if the aggregates are formed largely via side-by-side cross-linking, their frictional coefficient may be not proportional to the number of virions they are built of. On the contrary, the charge of the aggregate is proportional to the number of cross-linked virions. In this case, the electrophoretic mobility, which is defined as a ratio between the charge and friction coefficient, should increase in its absolute value with the growing number of cross-linked virions. We confirmed the formation of the aggregates with prominent side-to-side cross-linking of virions by Atomic Force Microscopy (AFM) and roughly estimated the number of virions in aggregates at different molar ratios of Ab to PVX virions from the electrophoretic data.

The observed phenomenon may support an important analytical application: assessment of antibody's ability to bind virions (antigen). A greater deviation from the expected electrophoretic mobility corresponds to more virions cross-linked by the antibodies, which, in turn, corresponds to higher antigen-binding efficiency of the antibodies. The detailed analysis of electropherograms (taking into account the number of peaks as well as their heights and relative positions) can also provide semi-quantitative information about the size distribution of the complexes and the approximate number of virions in a complex of a specific size. These parameters are important in vaccine development, diseases diagnostics, and immunology, in general. This is especially relevant due to recent progress in the development of vaccines and viral nanoparticles based on plant viruses, in particular PVX **[92**–**94]**.

3.2. MATERIALS AND METHODS

3.2.1. Chemicals and materials

Fused silica capillary (75 μm inner diameter, 360 μm outer diameter) was purchased from Molex (Phoenix, AZ). All chemicals were obtained from Sigma-Aldrich (Toronto, ON, Canada) unless otherwise stated. Chromeo P503 pyrylium dye was purchased from Active Motif (Burlington, ON, Canada). PVX was isolated from artificially contaminated plant leaves and full-size bivalent monoclonal antibodies were obtained and purified as described else-where **[95]**. Purified monovalent antigen-binding fragments (Fab) were obtained from Ab using the Pierce Fab preparation kit from ThermoFisher Scientific (Waltham, MA) according to the manufacturer's instructions. The activity of Fab was confirmed by ELISA (see Appendix B). PVX (7.4 mg/mL) was stored in Phosphate Buffered Saline (PBS) with addition of glycerol (1:1) at -20˚C. Full-size antibodies (1.93 mg/mL) and Fab (0.6 mg/mL) were stored in PBS with addition of 0.02% NaN_3 at +4˚C. All solutions were prepared using double distilled deionized water and filtered through a 0.22 μm filter (Millipore, Nepean, ON, Canada). 50 mM Tris-HCl, pH 7.5 was used both as a run buffer and as a dilution buffer.

3.2.2. Capillary electrophoresis

All CE experiments were done with a P/ACE MDQ instrument from Beckman Coulter (Fullerton, CA, U.S.) using laser-induced fluorescence (LIF) detection with excitation at 488 nm and emission at 605 nm (for Chromeo labeled virions and Ab) and light-absorption detection at 260 and 280 nm for unlabeled PVX and Ab, respectively. Uncoated fused silica capillaries of 80-cm total length and 70-cm distance from the inlet to the detection window were used for all CE experiments. New capillaries were preconditioned with 10 capillary volumes of methanol. Each run was followed by washing the capillary with 7 capillary volumes of each of 100 mM HCl, 100

mM NaOH, deionized water and run buffer. The samples were injected by pressure at 0.5 psi (3.4 kPa) for 5 s. Capillary coolant temperature was kept at +25 °C. An electric field of 310 V/cm with a positive electrode at the injection end was applied to carry out the electrophoresis.

3.2.3. Chromeo P503 labeling

54 μL of the stock solution of antibodies were mixed with 6 μL of Chromeo P503 working solution. The obtained mixture was incubated at +4 °C for 24 h to complete the derivatization reaction, which was accompanied by a color change from blue to red. The final antibody concentration was 11.5 μM.

3.2.4. Atomic force microscopy

All AFM experiments were performed with a SmartSPM-1000 atomic force microscope (AIST-NT, Novato, CA) in a tapping mode on air using fpN01HAR cantilevers (with a tip curvature radius of 10 nm and a resonant frequency of 118-190 kHz). Dilutions of PVX and Ab were done in PBS. 10-15 μL aliquots of PVX (0.14 nM) as well as mixtures of Ab and PVX in ratios 10:1 and 50:1 were incubated for 20 min and immobilized on a freshly cleaved mica and kept for 10 min to facilitate physical adsorption of the particles on the mica surface. The excess of the liquid was carefully removed with a filter paper and the mica was dried with an air flow. The obtained results were analyzed with Gwyddion software **[96]**.

3.3. RESULTS AND DISCUSSION

Separation of particles (molecules) in electrophoresis is possible due to differences in their electrophoretic mobilities μ. In turn, μ depends on charge q and translational friction coefficient f of the particle:

$$\mu = q / f \qquad (3\text{-}1)$$

When two particles (A and B) with different mobilities bind each other and form a complex (C), the latter typically has the mobility intermediate to those of A and B. This simple rule originates from approximate additivities of charges and hydrodynamic sizes (the latter are directly linked with friction coefficients):

$$q_C \approx q_A + q_B, \quad f_C \approx f_A + f_B \tag{3-2}$$

The mobilities of A, B, and C are then:

$$\mu_A = \frac{q_A}{f_A}, \quad \mu_B = \frac{q_B}{f_B}, \quad \mu_C \approx \frac{q_A + q_B}{f_A + f_B} \tag{3-3}$$

If we assume that

$$\mu_A < \mu_B \tag{3-4}$$

we can obtain from **eq (3-2)** and **eq (3-3)** (see section 1 Appendix B for derivations) the following inequality:

$$\mu_A < \mu_C < \mu_B \tag{3-5}$$

Hence, in the case of additivity of charges and friction coefficients, it is expected, that the mobility of complex is intermediate to those of its components.

Our experiments with complexes of PVX virions and bivalent Ab showed a deviation from this rule. In CE, the virions and antibodies migrated as individual zones with electrophoretic mobilities close to each other (**Figure 3.1A,B**). Multiple complexes were formed by mixing bivalent antibodies with PVX virions; the complexes migrated slower than unbound virions and antibodies (**Figure 3.1C**), which is equivalent to complexes having greater absolute values of electrophoretic mobilities. Unbound antibodies and virions as well as their complexes migrate

slower than electroosmotic flow (marked by a peak with negative absorbance at ~9.5 min) and, thus, have negative electrophoretic mobilities. This allows us to operate with absolute values of mobilities through the rest of this paper.

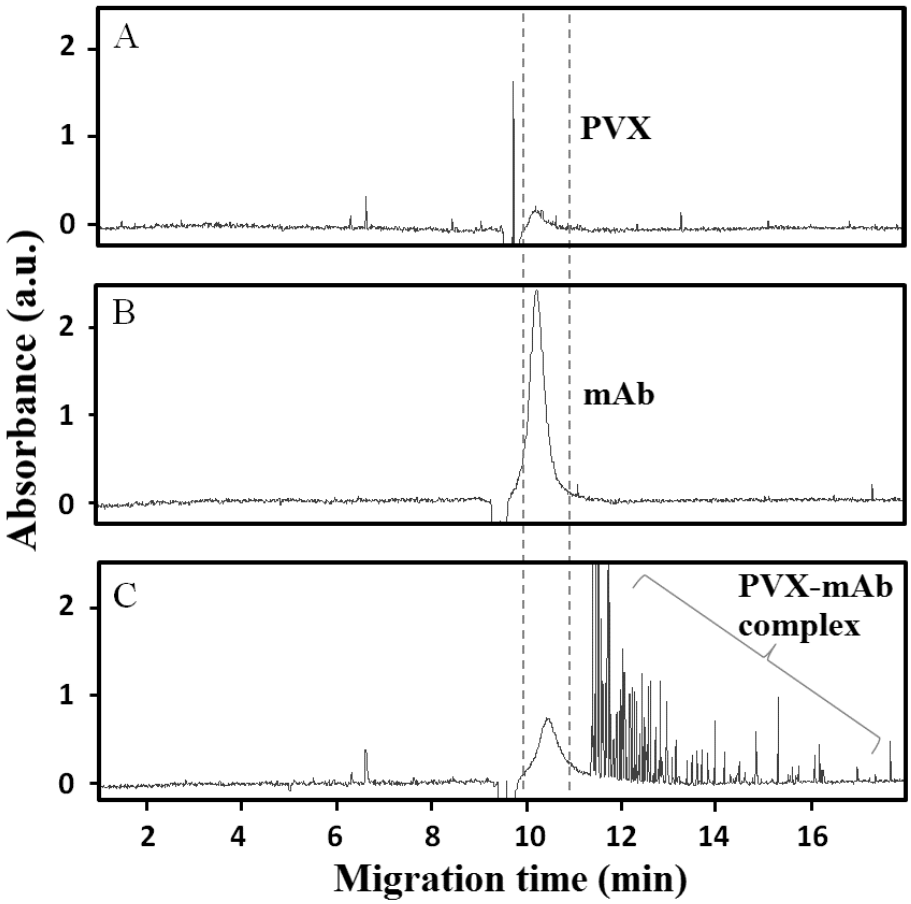

Figure 3.1. Migration of PVX virions, bivalent Ab, and their complexes in CE. The injected samples were: 20 nM PVX (A), 6.4 μM Ab (B), and a mixture of 20 nM PVX and 6.4 μM Ab incubated for 1 h at room temperature (C). Light absorption detection at 260 nm was used for PVX (A) and at 280 nm for Ab and the PVX-Ab mixture (B, C). The dotted lines flank the region of migration of unbound PVX and Ab.

We confirmed that the observed set of multiple peaks was not a result of virion disruption, but corresponded to complexes between the virions and bivalent antibodies, by performing similar CE experiments with Chromeo P503 labeled antibodies and LIF rather than light-absorbance

detection (see section 2 in Appendix B). In this case, only free antibodies and antibody-containing complexes were detected.

When PVX virions were mixed with the monovalent Fab fragments, the complexes could not be separated from unbound virions and antibodies and, thus, had mobility similar to those of unbound virions and antibodies (see section 3 in Appendix B).

To explain the anomalous mobility of the complexes between the virions and bivalent antibodies, we hypothesized that bivalent antibodies cross-link virions into aggregates with disproportionally small friction coefficients. Such aggregates should have significant side-by-side cross-linking of virions with cylinder-like shape of the aggregates (**Figure 3.2**). Individual virions and their aggregates can be considered as cylinders with different diameters.

We can theoretically estimate how the mobility of cylindrical aggregates depends on the number (n) of cross-linked particles. The antibodies play a role of "staples" connecting the virions and are, thus, sufficient in a relatively small molar excess to the virions. In addition, the antibodies are much smaller than the virions. Accordingly, both the charge and mobility of the aggregates will be defined by the corresponding parameters of the virions. The charge of the aggregate is a sum of the charges of individual virions it is built of. For similar virions, the charge of the aggregate (q_{agg}) is proportion-al to the number of the cross-liked virions and a charge of a single virion (q_{vir}):

$$q_{agg} \approx nq_{vir} \tag{3-6}$$

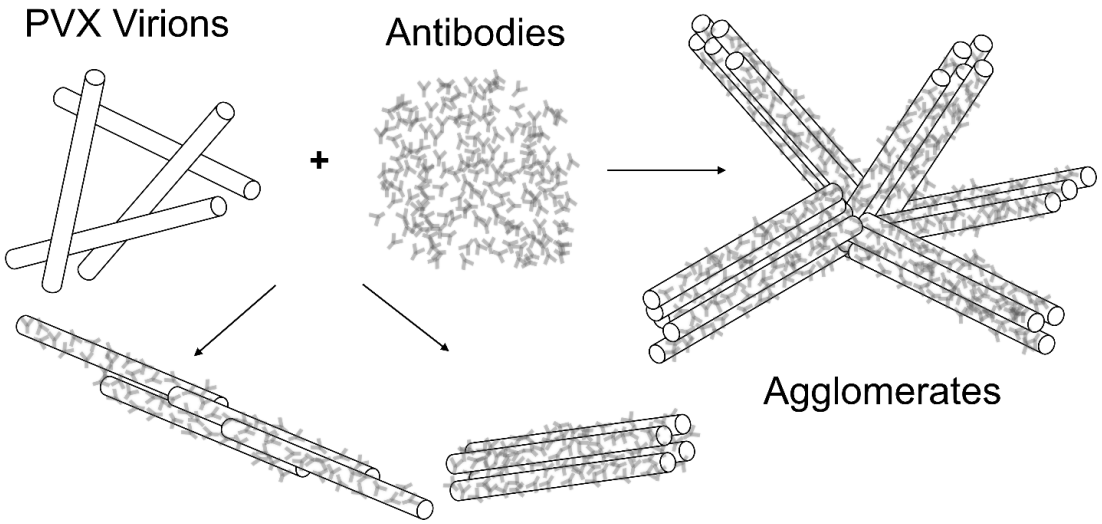

Figure 3.2. Schematic illustration of formation of near-cylindrical aggregates of rod-like virions cross-linked side-by-side via bivalent antibodies.

On contrary, the friction coefficient is a non-additive function in this case, even in the first approximation. The friction coefficient is proportional to the surface area of the particle (A_{cyl}), which, for a cylindrical aggregate (**Figure 3.2**), is proportional to $n^{1/2}$ as shown in the following:

$$f_{agg} \sim A_{cyl} \approx \frac{1}{2}\pi d_{agg}{}^2 + \pi L d_{agg} \approx \frac{1}{2}\pi n d_{vir}{}^2 + \pi L d_{vir} n^{1/2} \approx \pi L d_{vir} n^{1/2} \qquad (3\text{-}7)$$

Here L is the length of the cylinder and d_{agg} and d_{vir} are diameters of its base for the aggregate and the individual virion, respectively. In the last step of derivations, we neglected the area of the cylinder base in the assumption that the cylinder length (around 500 nm) is much greater than its diameter (13 nm): $L >> d_{vir}$. Also, d_{agg} is proportional to $n^{1/2}\, d_{vir}$.

By using eq **(3-1), eq (3-6)** and **eq (3-7),** we can obtain a relation between the electrophoretic mobility and the number of virions in the cylindrical aggregate:

$$\mu_{agg} \sim \frac{nq}{n^{1/2}\pi L d_{vir}} = n^{1/2}\frac{q}{\pi L d_{vir}} \quad (3\text{-}8)$$

Thus, the absolute value of electrophoretic mobility should grow with the number of cross-linked virions proportionally to a square root of *n*. Since q/($\pi L d_{vir}$) is proportional to the mobility of an individual virion (μ_{vir}), we can write:

$$\mu_{agg} \approx n^{1/2}\mu_{vir} \quad (3\text{-}9)$$

Accordingly, complex mobility in this case can be used to estimate roughly the number of cross-linked virions:

$$n \approx \left(\mu_{agg} / \mu_{vir}\right)^2 \quad (3\text{-}10)$$

To estimate *n* for aggregates of PVX virions, we determined the values of μ for free virions and for all of the aggregates at different ratios of full-size antibody:PVX virion (see section 4 Appendix B) and plotted the normalized peak intensity of each fraction as a function of the number of virions (**Figure 3.3**). Expectedly, the number of cross-linked virions in aggregates grew with the antibody concentration. The heterogeneity of sizes also in-creased with increasing concentration of antibodies. We then used AFM imaging to confirm the formation of aggregates with prominent side-to-side cross-linking of virions by bivalent antibodies. In the absence of antibodies, we observed individual virions of the expected dimensions (**Figure 3.4A**). In the presence of molar excess of antibodies, we found aggregates with prominent side-to-side cross-linking. Even a ten time molar ratio between the antibodies and virions was sufficient for aggregate formation (**Figure 3.4B**). Further increase in this ratio led to the formation of larger branching clusters of aggregates still with a significant content of side-to-side binding (**Figure 3.4C**).

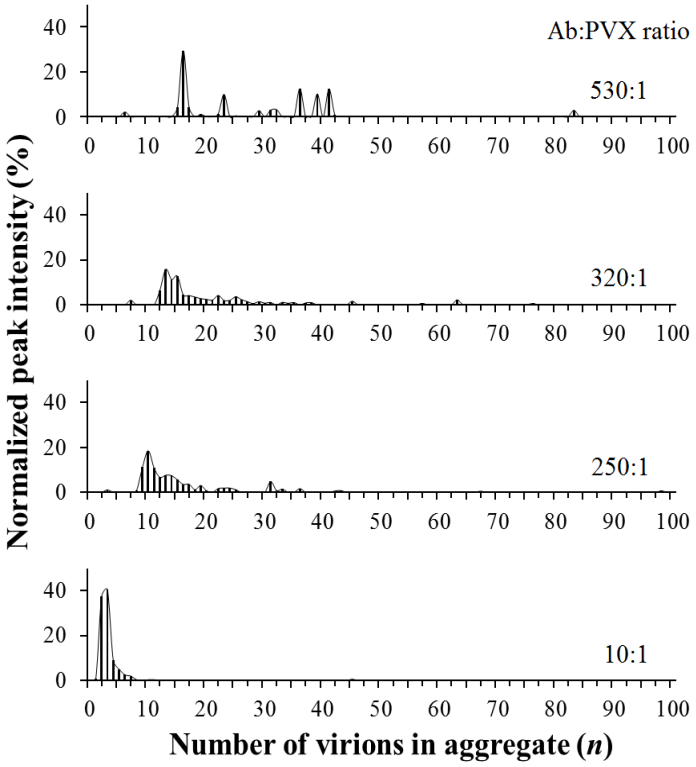

Figure 3.3 Influence of the molar ratio Ab:PVX on the size distribution of aggregates of PVX virions cross-linked with bivalent antibodies. Bars represent fraction distribution profiles of aggregates at different Ab:PVX ratios. The solid line on top shows the same data in a curve form.

When virions were mixed with monovalent Fab fragments no aggregate formation was observed by AFM, thus, confirming the cross-linking nature of aggregate formation in the case of bivalent antibodies.

AFM results also showed that the aggregates of PVX virions crosslinked by bivalent antibodies grew not only in their diameter but also in their length (**Figure 3.4B**). The elongation was due to the expected partial overlap of cross-linked virions. Importantly, such elongation can further increase the absolute value of electrophoretic mobility if the aggregate retains near-linear shape.

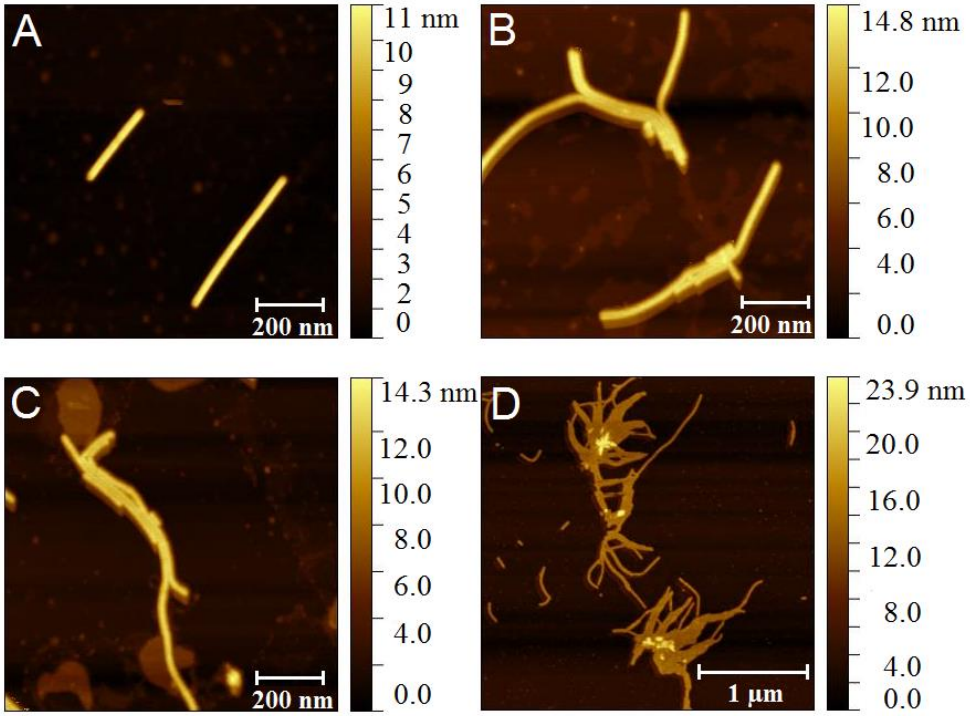

Figure 3.4. AFM images of individual PVX virions (A) and their full-size antibody-cross-linked aggregates of different sizes (B−D). Image A was obtained in the absence of antibodies. The molar ratios of antibodies to virions used to obtain images B and C were 10, and for image D it was 50. The color scale is to indicate the height profile of the structures in the image.

The elongation will increase the dipole moment of the particle and lead to more pronounced orientation of the particle along the electric field and thus along the direction of its translational movement. Such orientation will, in turn, lead to decreasing effective translational friction coefficient and, accordingly, increasing electrophoretic mobility. This hypothetical explanation is strongly supported by the previously observed increase of absolute electrophoretic mobility with increasing length of single-walled carbon nanotubes **[97]**. The in-crease of absolute value of electrophoretic mobility with greater extent of particle orientation along the direction of translational movement was also reported for rod-like viruses **[98]**. Thus, even for elongated aggregates we can still expect absolute values of electrophoretic mobilities greater than that of

individual virions. The in-creased mobility will be characteristic also for clusters of aggregates shown in **Figure 3.4C**.

3.4. CONCLUSIONS

In this work, we report on high absolute electrophoretic mobility of complexes of rod-like virions with bivalent antibodies. The behavior is explained by the formation of cross-linked aggregates of virions in which the charge is proportional to the number of cross-linked virions but the friction coefficient is roughly proportional to the square root of this number. Accordingly, electrophoretic mobility is roughly proportional to the square root of the number of cross-linked virions. The formation of aggregates is clearly seen with AFM and is restricted to complexes of virions with bivalent antibodies; the complexes of virions with monovalent antibodies do not form the aggregates. The larger absolute value of electrophoretic mobility of the aggregates makes it possible to separate them from the unbound virions and antibodies by CE. This, in turn, can, potentially facilitate kinetic and thermodynamic studies of complex formation between the rod-like viruses and antibodies.

CHAPTER 4. APTAMER FACILITATED PURIFICATION OF FUNCTIONAL PROTEINS

The presented material was published previously and reprinted with permission from "Beloborodov, S. S.; Bao, J.; Krylova, S. M.; Shala-Lawrence, A.; Johnson, P. E.; Krylov, S. N. Aptamer Facilitated Purification of Functional Proteins. *Journal of Chromatography B* 2018, 1073, 201–206." Copyright 2018 Elsevier B.V.

My contributions to the article were: (i) performing all presented experiments, (ii) preparing all figures and (iii) writing the manuscript. The major part of the experiments with MutS protein was conducted in collaboration with Dr. J. Bao.

4.1. INTRODUCTION

Purification of recombinant proteins is a key procedure in biochemical research **[99]** and development and production of biologics **[100]**. The gene, which encodes the protein, is cloned and over-expressed in cultured bacterial cells; the cells are lysed to obtain the cell lysate enriched with the target protein **[101]**. The target protein is then separated from the cell lysate based on its unique physicochemical properties such as charge **[102]**, size **[103]**, or affinity to a capture probe (such as Nickel-NTA or Glutathione) **[104**,**105]**. Affinity purification facilitates the highest degree of purity, as it relies on very selective interaction of the target protein with the capture probe. Affinity purification typically requires fusion of the target protein with an affinity tag (such as His or GST tag). An affinity tag may alter protein conformation and, thus, change its physicochemical properties **[106]**, and its activity as a result. Although a tag can be enzymatically removed, the enzyme may cleave fragments of the target protein, rendering it inactive and causing additional contamination of the protein with tag-removal agents **[107]**. Therefore, purification of some proteins requires tag-less affinity techniques.

A commonly used tag-less purification method involves antibodies as capture probes that can specifically bind the target protein itself. However, the utility of this method is restricted by the limitations of antibodies. The development of antibodies is a lengthy and labor-intensive process that cannot be conducted *in vitro*. In addition, antibodies are very sensitive to temperature changes and undergo denaturation over the time. Hence, tag-less purification method would benefit from replacing antibodies with capture probes, which can be easily synthesized and are more stable.

Antibodies can be replaced with oligonucleotide aptamers, which are single-stranded DNA or RNA (ssDNA and ssRNA) that can selectively bind their target protein with high affinity and specificity **[108]**. Aptamers are usually selected from random-sequence DNA (or RNA) libraries *via* the SELEX (Systematic Evolution of Ligands by EXponential enrichment) process, which involves alternating steps of partitioning of protein-DNA complexes from unbound DNA and amplification of the aptamer-enriched DNA library by the polymerase chain reaction **[59]**. Aptamers are more stable than antibodies and can be synthesized *in vitro*. Aptamers can also be synthetically modified with a linker (such as an amine group, thiol group or biotin) **[109]**, which allows their reliable non-covalent or covalent conjugation with the solid-phase substrate. Replacing antibodies with aptamers solves many problems in respect to the usability of the affinity probe; however, it does not solve the problems that may happen to proteins on the pre-purification steps. In particular, proteins often loose their activity as a result of the damage, caused by cell lysis or sonication. Thus, both, aptamers and antibodies are not always the limited factors influencing the stability and purity of the purified proteins. The idea of aptamer facilitated protein purification was introduced in 1999 **[110]**, and since then, a few proof-of-principle studies have been conducted **[111**–**114]**. To the best of our knowledge, protein's

activity was not confirmed in any of the published studies. Also, some studies used aptamers that specifically bind His-tag **[115,116]** (thus, still requiring an affinity tag on a target protein) or involved the use of denaturing agents at the elution step **[117]** (thus potentially leading to protein inactivation). Therefore, our goal was to develop and validate a protocol for aptamer facilitated purification of tag-less proteins with preserved activity.

The developed purification procedure is applied to a crude bacterial cell lysate prepared by sonication and follows a general scheme shown in **Figure 4.1**. The biotinylated aptamer is immobilized onto magnetic beads via streptavidin-biotin interaction. After immobilization, the beads are incubated with the cell lysate containing the target protein. During the incubation, the target protein binds to the immobilized aptamer.

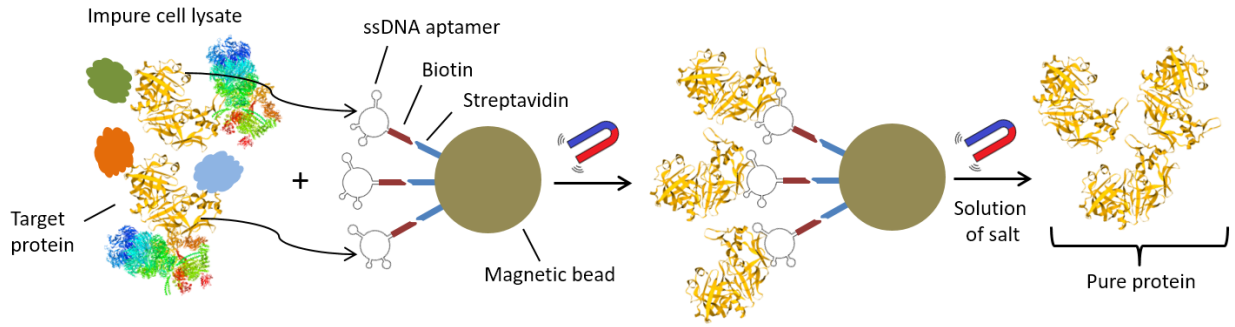

Figure 4.1. Schematic illustration of aptamer-facilitated protein purification. Multiple copies of protein-specific ssDNA aptamer are attached to magnetic beads through the streptavidin-biotin bridge. The aptamers serve as capture probes for the target protein in a crude cell lysate. The beads with captured target protein are separated from the lysate by retaining the beads with a magnet and washing out the lysate. Finally, the target protein is eluted with phosphate-buffered saline (PBS) supplemented with an additional strong electrolyte (NaCl or $MgCl_2$) to increase the ionic strength.

The unbound components of the lysate are then removed and the bound protein is eluted by a buffer containing a high concentration of a strong electrolyte (0.5 M NaCl or 2.0 M $MgCl_2$) with physiological pH at +4 °C (the high ionic strength induces reversible structural alteration of DNA aptamer without denaturing the target protein). The procedure takes 4 h to complete.

The developed procedure was validated with the AlkB **[118]** and MutS **[119]** proteins. Both proteins are involved in DNA repair: AlkB (23.9 kDa) catalyzes dealkylation of ssDNA, and MutS (95 kDa) recognizes and binds to mismatches in double stranded DNA (dsDNA). Protein activity after purification was confirmed by two capillary electrophoresis based assays: an enzymatic assay for AlkB and a DNA-binding assay for MutS. These assays showed that both proteins retained their specific activity. Thus, our protocol can be used for isolation of the tagless proteins in their active state and can also be combined with a previously-developed approach of aptamer selection for non-purified protein in a crude cell lysate **[113]**. This approach involves multiple rounds of the following alternating steps: incubation of the DNA library with the target containing cell lysate, collection of the bound DNA, incubation of the collected DNA with target-free cell lysate and collection of the unbound DNA, followed by the PCR amplification and purification. The selected polyclonal aptamers can then be used for purification of the target protein.

4.2. MATERIALS AND METHODS

4.2.1. Chemicals and materials

E. coli strain MG1655 was kindly provided by Coli Genetic Stock Center (CT, USA). The pET-24a(+) vector and NovaBlue competent cells were purchased from EMD Millipore (PA, USA). *E. coli* BL21-Gold (DE3) competent cells were purchased from Agilent Technologies (CA, USA). A biotin labeled anti-AlkB aptamer **[88]** (biotin-AlkB-aptamer): 5’-/5BioTEG/ CTC

CTC TGA CTG TAA CCA CGT GCC TAG CGT TTC ATT GTC CCT TCT TAT TAG GTG ATA ATA GCA TAG GTA GTC CAG AAG CC-3' and AlkB primers: forward 5'-GGTGGTCATATGTTGGATCTGTTTGCC-3' and reverse 5'-GGTGGTGGATCCTTATTCTTTTTTACCTGC-3'were custom-synthesized by Integrated DNA Technologies (IA, USA). The AlkB methylated substrate TTCmTTTTTTTTTTT3'-FAM (TTCm) was custom-synthesized by ATDBio (Southampton, UK). The restriction enzymes NdeI and BamHI, blue prestained protein standard, and T4 DNA ligase were purchased from New England Biolabs (MA, USA).

The pETMutS plasmid #13245 was obtained from Addgene (MA, USA). The *E. coli* Rossetta-gamiTM 2 (DE3) competent cells were purchased from EMD Millipore (PA, USA). A biotin labeled anti-MutS aptamer **[120]** (biotin-MutS-aptamer): 5'- /5BioTEG/CTT CTG CCC GCC TCC TTC CTG GTA AAG TCA TTA ATA GGT GTG GGG TGC CGG GCA TTT CGG AGA CGA GAT AGG CGG ACA CT -3' and fluorescently-labeled dsDNA with a G-T mismatch (dsDNA-GT) with forward sequence of 5'-CTT CTG CCC GCC TCC TTC CTT CCA ACC TTC ATC GGC CAC CC-3' (the mismatch is indicated with an underline) were custom synthesized by Integrated DNA Technologies (IA, USA). Streptavidin-coupled magnetic beads (MagnaBind Streptavidin Beads) were purchased from Thermo Scientific (IL, USA). Mono Q anion exchange column (1 mL bed bolume) was purchased from Amersham (Amersham Pharmacia Biotech, currently GE Healthcare Life Sciences, IL, USA).

The isopropyl-β-D-thiogalactoside (IPTG) was purchased from Gold Biotechnology (MO, USA). A protease inhibitor cocktail (cOmplete^{TM}) was purchased from Roche (CA, USA). A bicinchoninic acid (BCA) assay kit was obtained from Pierce Biotechnology (IL, USA). A fused silica capillary (75 μm inner diameter, 360 μm outer diameter) was purchased from Molex (AZ,

USA). All other chemicals were purchased from Sigma-Aldrich (ON, Canada). All solutions were prepared with deionized water filtered through a 0.22 μm Millipore filter (ON, Canada).

4.2.2. Immobilization of biotinylated aptamer on streptavidin-beads

To immobilize the aptamers on magnetic beads we took 1 mL of streptavidin-coated magnetic beads (which can bind 2 μg or 8.19 nanomoles of biotin) and washed them 3 times with deionized water. The washed beads were mixed with 0.8 mL deionized water and 105 μL of 100 μM aptamer-biotin solution. After a 2-h incubation at room temperature, the supernatant was removed and the beads were washed with deionized water three times and then resuspended in 0.5 mL of deionized water. The aptamer occupied 92% of the streptavidin molecules on the beads, which was determined by absorbance measurements of the aptamer-containing solution at 260 nm before and after the binding. In other words, 1 mL of the original beads suspension bound 7.53 nanomoles of biotin and aptamer, accordingly. A suspension of aptamer-free beads was prepared in a way similar to preparation of the suspension of aptamer-containing beads, but the aptamer solution was replaced with water at the mixing step and the final suspension was made in 0.5 mL of the resuspension buffer instead of deionized water.

4.2.3. Protein expression and preparation of cell lysate

AlkB was expressed as the following. The AlkB gene from *E. coli* strain MG1655 was amplified by PCR. The product of PCR amplification was then digested with NdeI and BamHI enzymes and cloned into the pET-24a(+) vector. The pET-24a(+)-AlkB vector was then transfected into NovaBlue cells, amplified, reextracted and purified by using the GenElute Plasmid Miniprep Kit. The purified plasmid was then transfected into *E. coli* BL21-Gold (DE3) competent cells. The culture was first grown in 25 mL Lysogeny Broth (LB) media with 50 μg/mL kanamycin at 37 °C overnight (at 250 rpm shaking). Then 1 mL of the culture was

transferred to 100 mL of Terrific Broth (TB) media (with 50 µg/mL kanamycin) and was incubated with shaking at 250 rpm for 6 hours at 37 °C until the Optical Density value of 0.8 was achieved at 600 nm (OD^{600}). Then, IPTG was added to 100 mL of the cell culture in LB media to achieve 1 mM final concentration, and the mixture was shaken overnight under room temperature at 250 rpm. Next, the cells were pelleted by centrifugation at ~ 5000 ×g (Allegra 21R centrifuge with S4180 rotor (Beckman Coulter, ON, Canada)) for 30 min at 4°C. The supernatant was removed, and the cells were resuspended in 10 mL of the resuspension buffer (100 mM 2-(N-morpholino)ethanesulfonic acid (MES), 1 mM Dithiothreitol (DTT) and 19 µL of protease P8340 for 10 mL solution, pH 5.8), and sonicated on ice with 15-s on/off intervals for 7 min at 20% of maximum power (Fisher Scientific, Sonic Dismembrator, Model 500). Next, the mixture was centrifuged for 30 min at ~20000 ×g (Eppendorf 5417R centrifuge with F45-30-11 rotor (Fisher scientific, PA, USA)) at 4 °C. The supernatant was collected.

MutS was expressed in a similar way to that of AlkB expression **[113,121]**. The competent *E. coli* Rossetta-gami™ 2 (DE3) cells were transformed with the plasmid pETMutS. The transformed cells were grown in 1 L of LB medium in the presence of 50 µg/mL ampicillin at 37 °C with 250 rpm shaking to reach an OD^{600} value of 0.8. Protein expression was induced by adding IPTG to 1 mM in the cell culture. The culture was shaken overnight at room temperature at 250 rpm. The cells were pelleted by centrifugation at ~ 5000 ×g (Allegra 21R centrifuge with S4180 rotor (Beckman Coulter, ON, Canada)) for 1 h at 4 °C. The cell pellet was resuspended in 10 mL Tris buffer (20 mM tris-HCl pH 7.4) that contained a cOmplete protease inhibitor cocktail tablet and sonicated on ice with 15-s on and 55-s off intervals for 4 min at 60% of maximum power of the sonicator. The cell debris was pelleted by centrifugation at 5000 ×g for 1 h at 4 °C. The supernatant was incubated at 70 °C for 45 min followed by a new round of centrifugation for

1 h at 4 °C. For pre-affinity purification, the protein sample was applied onto a Mono Q anion exchange column, which was equilibrated with Tris buffer, and eluted with a linear gradient of 0.1 to 1.0 M NaCl at a flow rate of 1 mL/min. The MutS protein was eluted at approximately 0.4 M NaCl. Fractions containing MutS protein were pooled and dialyzed extensively against Tris buffer by using an Amicon Ultra-15 Centrifugal Filter Unit with Ultracel-10 membrane (Millipore, Ireland).

4.2.4. Protein purification by aptamer-beads

In the case of AlkB purification, 4 mL of protein-containing supernatant was mixed with 0.5 mL of bead suspension (on ice) and incubated at continuous shaking for 2 h at 250 rpm. The beads were removed and sequentially incubated for 10 min with 1 mL of each of the following buffers: resuspension buffer, 0.5 M NaCl solution (in PBS) and 1 M NaCl solution in PBS. The eluates were analyzed with CE by using ProteomeLab SDS-MW Analysis Kit (BeckmanCoulter) according to the manufacturer's instructions. The resuspension buffer eluate contained traces of the cell lysate. The eluate of 0.5 M NaCl in PBS contained the largest amount of pure AlkB and then was subjected to a buffer exchange with AlkB protein storage buffer (50 mM Tris-HCl, 500 mM NaCl, 1 mM DTT, 0.02% NaN_3) by using Amicon Ultra-0.5 Centrifugal Filter Unit with Ultracel-10 membrane (Millipore, Ireland). The eluate of 1 M NaCl in PBS contained only traces of AlkB and was not used. AlkB-containing cell supernatant was also used for the negative control experiments. In these experiments, 8 μL of the supernatant was mixed with 200 μL of the aptamer-free beads suspension and incubated on ice with continuous shaking at 250 rpm for 2 h. The same was done for 100, 50, 25 and 0 μL of the aptamer-free beads suspension with addition of an appropriate volume of the resuspension buffer to keep the final volume of the solution equal to 208 μL. The beads were then removed and the supernatant was analyzed with CE by

using ProteomeLab SDS-MW Analysis Kit (Beckman-Coulter) according to the manufacturer's instructions.

For purification of MutS protein, 0.5 mL of MutS-protein-containing solution was incubated with 0.5 mL of aptamer-bead solution on ice for 1 h under gentle agitation. After the incubation, the beads were collected and the clear solution was removed. For MutS protein elution, the beads were sequentially incubated (10 min each) and eluted with 500 μL of 0.25, 0.5, 0.75, 1, 2 and 3 M $MgCl_2$ solution at room temperature. The fraction which eluted with 2 M $MgCl_2$ contained the largest amount of pure MutS and was further used. The eluted protein solutions were dialyzed extensively against Tris buffer (20 mM tris-HCl pH 7.4) by using Amicon Ultra-15 Centrifugal Filter Unit with Ultracel-10 membrane (Millipore, Ireland). The concentrations of both AlkB and MutS were determined by using the BCA assay with bovine serum albumin (BSA) as a standard.

4.2.5. Capillary electrophoresis

CE experiments were carried out with a P/ACE MDQ instrument from SCIEX (Fullerton, CA, USA). Laser-induced fluorescence (LIF) detection with excitation at 488 nm and emission at 520 nm was used for AlkB enzymatic activity assay and dsDNA-MutS binding assay. Light absorption detection at 214 nm was used for protein molecular weight analysis in the SDS-gel electrophoresis mode (SDS-MW). New capillaries were preconditioned by washing them with 10 capillary volumes of methanol. For CE experiments with LIF detection each run was followed by washing the capillary with 7 capillary volumes of each of 100 mM HCl, 100 mM NaOH, deionized water (dd), and running buffer. Capillary coolant temperature was kept at +25 °C. An electric field of 310 V/cm with a positive electrode at the injection end was used to run electrophoresis. Uncoated fused silica capillaries with a total length of 50 cm and a 40-cm

distance from the inlet to the detection window and with inner diameter of 75 μm were used for detection of the output products of the AlkB enzymatic assay. Uncoated fused silica capillaries of a 30-cm total length and a 20-cm distance from the inlet to the detection window (50 μm inner diameter) were used for the MutS-DNA mismatch binding assay. In a case of SDS-MW analysis, the total length of capillaries was 30 cm (20 cm from the inlet to the detection window) and the inner diameter was 50 μm.

4.2.6. SDS-PAGE analysis

SDS-PAGE assay was conducted in the freshly made 12% polyacrylamide gel. Twenty microliters of MutS-containing solutions (cell lysate and MutS-containing eluates) were mixed with 80 μL of 4× SDS-PAGE loading buffer (62.5 mM Tris, 10% Glycerol, 0.05% Bromphenol Blue, 2% SDS and 5 % 2-mercaptoethanol) and 10 μL of each mixture was injected into the gel-plate vials. SDS-PAGE electrophoresis was conducted at 200 V for 50 min. Coommassie Blue dye was used for visualization of the protein bands.

4.2.7. AlkB enzymatic activity assay

The demethylation assay was carried out by mixing AlkB enzyme with the fluorescently labeled homopolymeric sequence TTCm at 20 and 200 nM final concentrations, respectively, in 50 mM Tris-HCl pH 7.5 buffer in the presence of 4 mM ascorbic acid, 0.16 mM α-Ketoglutaric acid (2OG), 0.05 mM BSA, 2100 units of catalase and 0.08 mM Fe (II) sulphate. The aliquots of the mixture (15 μL each) were added to 5 μL of 20 mM EDTA (iron chelator) at different time intervals (0, 0.5, 1, and 2 min) and kept at 70 °C on a hot plate to denature the protein and stop the reaction. Each sample was then injected into the capillary by pressure pulse of 0.5 psi (3.45 kPa) for 6 s and analyzed with CE at an electric field of 300 V/cm. The methylated

DNA(unreacted substrate) and demethylated DNA (product) were separated in a Borax buffer (20 mM Borax, 60 mM SDS, pH 8.2).

4.2.8. MutS-DNA mismatch binding assay based on plug-plug kinetic capillary electrophoresis (ppKCE)

A plug of 1 μM fluorescently-labeled dsDNA-GT solution was injected into the capillary by a 10 s pressure pulse of 0.5 psi (3.45 kPa). Next, a plug of the separation buffer was injected at 0.5 psi (3.45 kPa) for 30 s. Finally, the plug of 1.5 μM purified MutS protein solution was injected at 0.5 psi (3.45 kPa) for 10 s. The ends of the capillary were inserted into the inlet and outlet reservoirs, filled with the running buffer (50 mM Tris-HCl and 1 mM $MgCl_2$, pH 8.0) and electrophoretic separation was conducted at an electric field of 300 V/cm. The control experiments were done at the same temperatures by using a solution of dsDNA-GT void of MutS protein.

4.3. RESULTS AND DISCUSSION

4.3.1. AlkB expression and purification by aptamer beads

The AlkB expression procedure was adapted from a previously developed protocol **[121]**. The main deviation from the protocol was the use of TB media instead of LB media. This replacement allowed us to increase the amount of the expressed protein. The crude AlkB-containing cell lysate was analyzed with SDS-MW in CE (**Figure 4.2**, bottom trace). The AlkB peak was identified according to its molecular weight of 23.9 kDa, using molecular weight markers (**Figure 4.2**, upper trace). Based on the area analysis it was established that AlkB is the most abundant protein (30% of the total protein yield). By using BCA, it was found that the total protein concentration in the lysate was equal to 20 mg/mL (± 10%) and the concentration of AlkB protein was approximately 6 mg/mL. Next, we applied the developed purification protocol

to the crude cell lysate. By using 1 mL of aptamer-saturated bead suspension (containing 7.53 nanomoles of aptamer at 92% occupancy) we have obtained 0.21 mg of pure AlkB protein in a form of solution with the total volume of 70 μL. The purified eluate was analyzed by SDS-MW in CE (**Figure 4.2**, middle trace). Based on the area analysis of peaks in this trace, it was established that the purity of AlkB protein was more than 85%. In theory, we could obtain only 0.18 mg of the protein with this amount of the aptamer, but due to non-specific binding of AlkB to the silanol groups located on the surface of the beads, we were able to extract more protein.

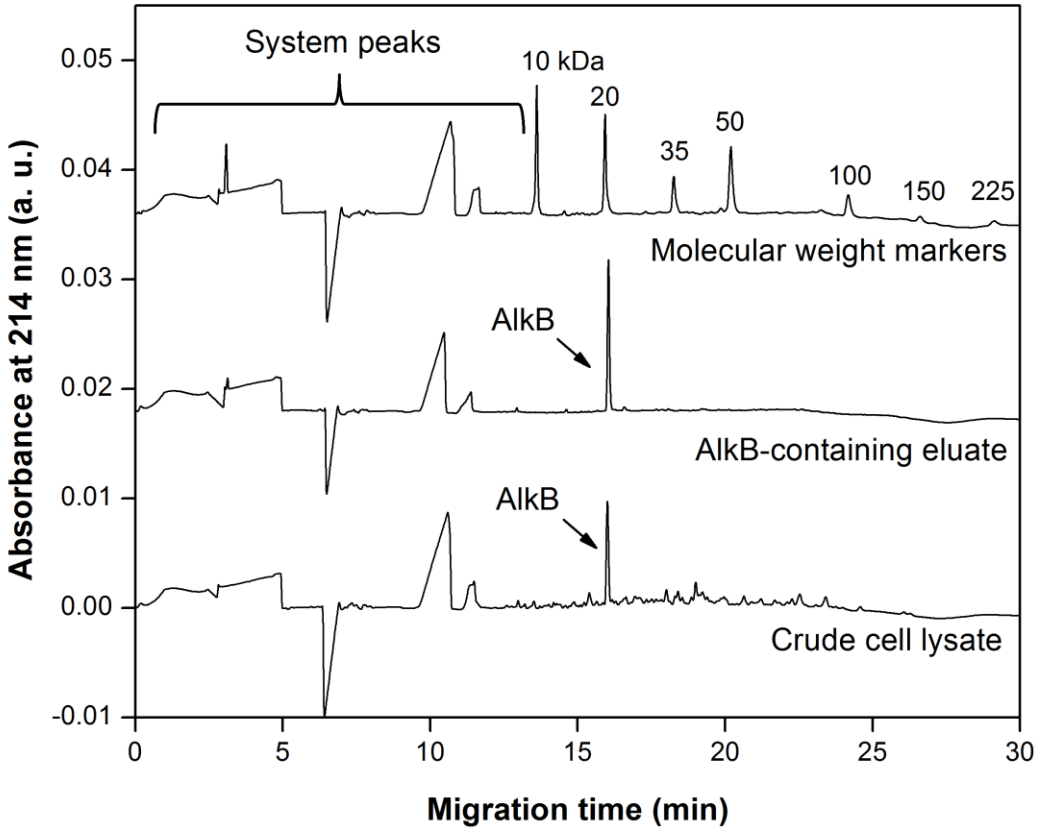

Figure 4.2. CE based SDS-MW analysis of crude lysate and purified AlkB protein, confirming expression and purity of the product. The traces correspond to the following samples (from bottom to top): crude cell lysate, AlkB-containing eluate, and a mixture of molecular weight markers. The traces are vertically offset for clarity of presentation.

The non-specific binding was confirmed in a series of negative control experiments with aptamer-free beads, which showed that these beads are capable of binding proteins in the cell lysate, but do not have specificity to any particular protein. According to SDS-MW electropherograms (not shown), the areas corresponding to all the proteins were gradually decreasing with increasing the number of beads in the system. It was found that approximately 200 μL of the aptamer-free beads suspension was sufficient to bind all the proteins in 8 μL of the cell lysate.

4.3.2. AlkB activity assay

To confirm that the aptamer-purified AlkB protein retained its activity, we conducted a comparative demethylation assay. As a reference for comparison, we used another AlkB protein, which was purified by a previously published approach **[121]**, based on cation exchange column chromatography. The methylated DNA substrate was incubated with AlkB protein in the presence of Fe (II) and 2OG as cofactors. The concentration of the DNA in a mixture was 10 times higher than the concentration of the protein. The demethylation reaction was terminated at different time intervals from reaction initiation. Then, CE was used to separate the methylated DNA substrate from the demethylated DNA product. As follows from the electropherograms (**Figure 4.3**) the concentration of the methylated DNA is decreasing over the time, while the concentration of demethylated product becomes greater, which indicates that both aptamer-purified and cation-exchange chromatography-purified AlkB proteins were active. However, the AlkB that was purified by the aptamers led to an almost complete demethylation of the substrate in 1 min (**Figure 4.3B**), while the reference protein demethylated less than a half of the substrate during the same time (**Figure 4.3A**).

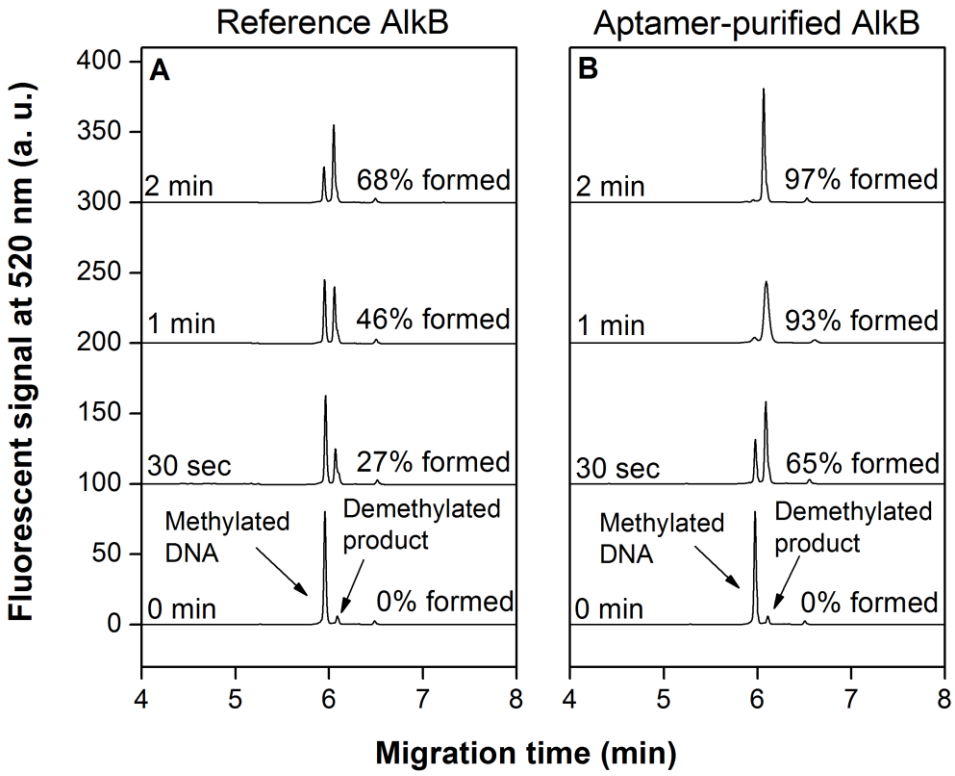

Figure 4.3. Activity comparison of the reference AlkB (A) and aptamer-purified AlkB (B). The traces in each panel correspond to different reaction times ranging from 0 to 2 min. The percentage corresponds to relative amount of demethylated product formed by the time the reaction is stopped. The traces are offset vertically for clarity of presentation. The additive concentration of the methylated and de-methylated DNA corresponds to 85 nM on all of the electropherograms.

4.3.3. MutS expression and purification by aptamer beads

The MutS expression was induced with 1 mM IPTG, similar to induction of AlkB expression. The cell lysate before and after induction was analyzed by slab-gel SDS-PAGE (**Figure 4.4**, lanes b, c). A large amount of undesirable non-target proteins was removed by heating the cell lysate at 70 °C for 45 min, consequently denaturing those proteins (**Figure 4.4**, lane d); MutS was not affected as it a temperature-stable protein. The soluble portion of heat treated lysate was loaded onto an anion exchange column. The anion exchange pre-purification

was performed in order to minimize the nonspecific interactions between MutS and endogenous bacterial DNA (**Figure 4.4**, lane e).

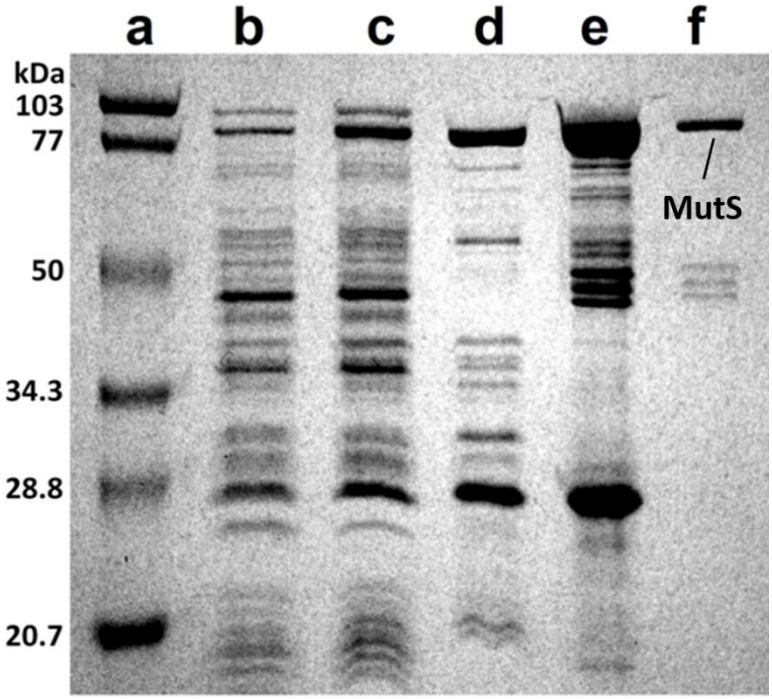

Figure 4.4. Expression and aptamer facilitated purification of MutS protein. The lanes correspond to the following (from left to right): molecular weight markers (a), un-induced cell culture (b), IPTG-induced cell culture (c), temperature-treated supernatant (d), an eluate after anion exchange column purification (e) and MutS eluted by $MgCl_2$(f).

The final step of protein isolation was done by incubating the MutS-containing eluate with aptamer-beads on ice for 1 h. The target protein was retained by aptamers on a surface of the beads, while the solution was removed. We eluted the bound protein with a high-ionic-strength solution of salt. We first tried the monovalent cation Na^+ , however, the interaction between MutS and aptamer could not be disrupted even by 3M NaCl. We then tested the divalent cation Mg^{2+} and found that 2 M $MgCl_2$ could elute 82%-pure MutS (based on the color densities of the bands in the SDS-PAGE slab). A total amount of obtained MutS was 0.24 mg (according to the BCA assay).

4.3.4. MutS-DNA mismatch binding assay

The DNA-binding activity of purified MutS was assessed by a ppKCE-based binding assay. Short plugs of dsDNA-GT and MutS were sequentially injected one after another by low-pressure pulses. An electric field was then applied, causing movement of MutS with a faster velocity that dsDNA-GT. When passing through dsDNA-GT, MutS could form a complex with it. As separation continued, the zones of the complex and unbound dsDNA-GT and MutS protein were separated from each other (**Figure 4.5**).

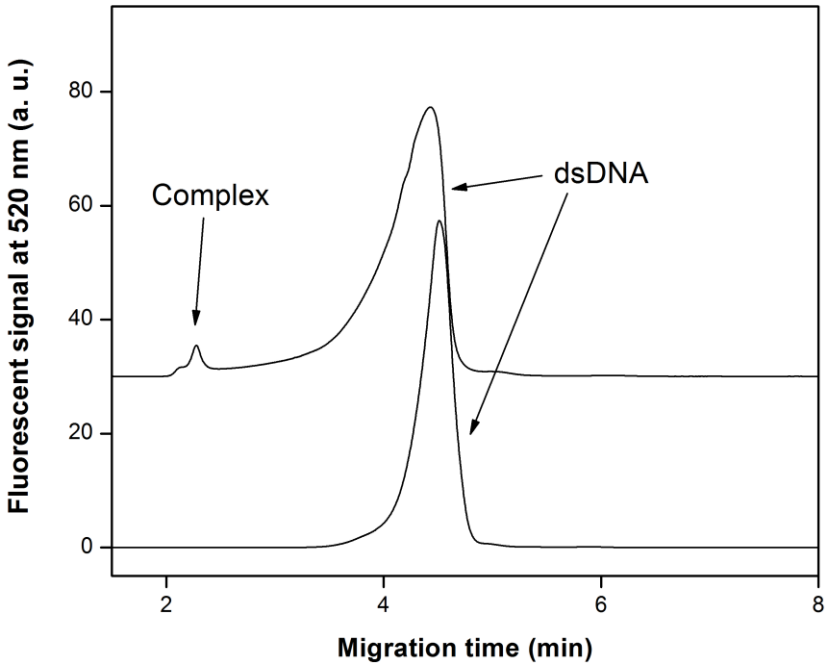

Figure 4.5. CE-based activity assay of MutS protein, confirming the formation of the MutS-dsDNA-GT complex. The upper trace represent the electropherogram of dsDNA in the presence of MutS protein. The lower traces represents dsDNA-GT only. The traces are offset vertically for clarity of presentation (leave just one temperature).

Based on the area analysis, we found a KD value of the complex of 326 ± 26 nM at 25°C. The obtained value is within the range of previously reported KD values, lying between 35 ± 6 nM (G in the mismatch is surrounded by G fragments only) **[122]** and 420 ± 30 nM (G in the

mismatch is surrounded by AT fragments) **[122]**. Therefore, we conclude that aptamer-purified MutS protein retained its binding activity.

4.4. CONCLUSIONS

To conclude, we have successfully developed and validated an aptamer facilitated protein purification protocol. Using AlkB protein and MutS protein as two models, we demonstrated that our approach could be used to achieve fast and reliable purification of active proteins. To confirm the AlkB activity, we have conducted the DNA demethylation assay and to confirm the functional activity of MutS protein, we have conducted a CE-based dsDNA-GT binding assay. Both assays showed that proteins retained their respective activities. Thus, the developed protocol can be used for purification of tag-less proteins in their active state. It should also be noted that despite the approximate equality of the dissociation constants, K_d, for the utilized AlkB and MutS aptamers (20 and 15 nM, respectively), the optimum elution procedures were different for AlkB and MutS proteins. It indicates that each case of the aptamer-facilitated protein purification is unique and may require its own specific conditions, especially at the elution step.

CHAPTER 5. GENERAL TIPS FOR THE DEVELOPMENT OF ANY CE-BASED SEPARATION APPROACH

In this chapter, I summarized all the tips and tricks gained from optimizing a new aptamer selection approach called Ideal Filter CE (IFCE). My colleague An Le introduced this approach and implemented it to select aptamers for MutS protein in one round **[60]**. Later, we found that the developed approach requires very precise control of Electroosmotic Flow (EOF) in respect to the selection of aptamers for low molecular weight proteins. Even small deviations in EOF could lead to an unsuccessful collection of the complexes due to the strong dependence between their elution time and EOF. In this part of my research, I investigated the impact of different capillary conditioning techniques on the EOF and introduced my own technique that resulted in a very stable and reproducible EOF. I also noticed that the difference in the height of the buffer solution at the inlet and outlet of the capillary led to a siphoning effect. The impact of this effect on the apparent EOF was so strong that it led to the loss of protein–aptamer complex in some of our experiments. Another phenomenon that I noticed and explained was the effect of the presence of salt (NaCl) in the sample plug on the formation of the complexes between the analyzed species and the salt-related ions. Such complexes were detectable in CE and could be mistakenly considered as a protein-aptamer complexes. I investigated the behavior of these salt-related pseudo-complexes, described the mechanism of their formation, and proposed a simple method for their elimination.

5.1. INTRODUCTION

Most drug screening approaches start by testing the binding affinity of potential drug candidates to target molecules, followed by the isolation of the drug candidates with highest affinity. Often, the experimental implementation of this procedure entails screening a drug

library against target molecules followed by the separation of binders from non-binders (step #1). Further, binders are collected, identified, and amplified (step #2). Steps #1 and #2 are sequentially repeated until the binders with sufficient affinity to the target molecules are isolated. The number of repetitions depends on the ratio of binders to non-binders in the original library and on the separation efficiency of the method. The first parameter has a statistical nature and it is difficult to control, however the second parameter depends on the physical principles of the chosen method and can be adjusted. In the case of CE-based methods, the separation efficiency depends not only on the differences in the electrophoretic mobilities (μ) of the species **[123**,**124]**, but also on experimental conditions such as the length of the capillary, its inner diameter, the temperature of the separation buffer, the length of the sample plug, the way of the sample injection, the value of the applied electric field and others. Achieving high separation efficiency by manipulating these conditions usually makes the method more complex and imposes several limitations. For instance, increasing the resolution by increasing the field strength **[125]** leads to a higher production of heat in CE, which is also known as the Joule heating effect **[126**,**127]**. The excess of heat increases the temperature along the capillary **[128]** and results in a more rapid dissociation of the target complexes **[129]**. Using longer capillaries with smaller inner diameter may solve the problem of overheating **[130]**, but on the other hand, a smaller diameter decreases the limit of detection (LOD) and worsens the sensitivity of the method **[131]**; in addition, longer capillaries increase the overall time of the experiment **[132]**. Hence, any improvement in one parameter worsens another. The only solution to this "double-edged sword" problem is to compromise and wisely sacrifice less important parameters for the benefit of more important ones. In our particular case, we sacrificed the robustness and applicability of our newly introduced Ideal Filter CE (IFCE) method for maximizing the separation efficiency between the

protein–ssDNA complexes and unbound ssDNA by bringing the system to conditions at which the target complexes move in one direction and all non-binding ssDNA molecules move in the opposite direction **[60]**. Such conditions were achieved by adjusting the ionic strength of the running buffer and by this setting the value of Electroosmotic Flow (EOF) to be somewhere between the absolute values for electrophoretic mobilities of the protein–ssDNA complex and unbound ssDNA. Moreover, it should be noted that the closer the mobility of the complex to the value of EOF, the higher the elution time of the complex (**Figure 5.1**). In other words, the complex elution time asymptotically tends to infinity when mobility of the complex is equal to EOF. From here, it becomes clear that the acceptable range of EOF for the elution of protein–ssDNA complex at a reasonable time is case specific and strongly depends on the mobility of the complex, which in turn depends on the molecular weight of the protein **[133]**. As we already mentioned in chapter 2, the higher the molecular weight of the protein, the smaller the absolute value for the electrophoretic mobility of its complex. Therefore, the acceptable EOF range for the elution of the complexes of heavy proteins is wider than that for lighter proteins. For example, if the ionic strength is fixed at 146 mM, as it was done in the original IFCE experiment, then the resulting EOF will correspond to 23.5 mm^2/kVs^{-1} **[133]**. This value of EOF is sufficient for the elution of the complexes of proteins such as Muts (95 kDa) and AlkB (24 kDa) with 80-nt long ssDNA at a reasonable time (**Figure 5.1**). For this experiment, the elution times (t_{el}) at different EOF were calculated with the following equation:

$$t_{el} = L_{tr} / \left((\mu_{EOF} + \mu_{P-ssDNA}) \cdot \frac{V}{L_{tot}} \right) \qquad (5\text{-}1)$$

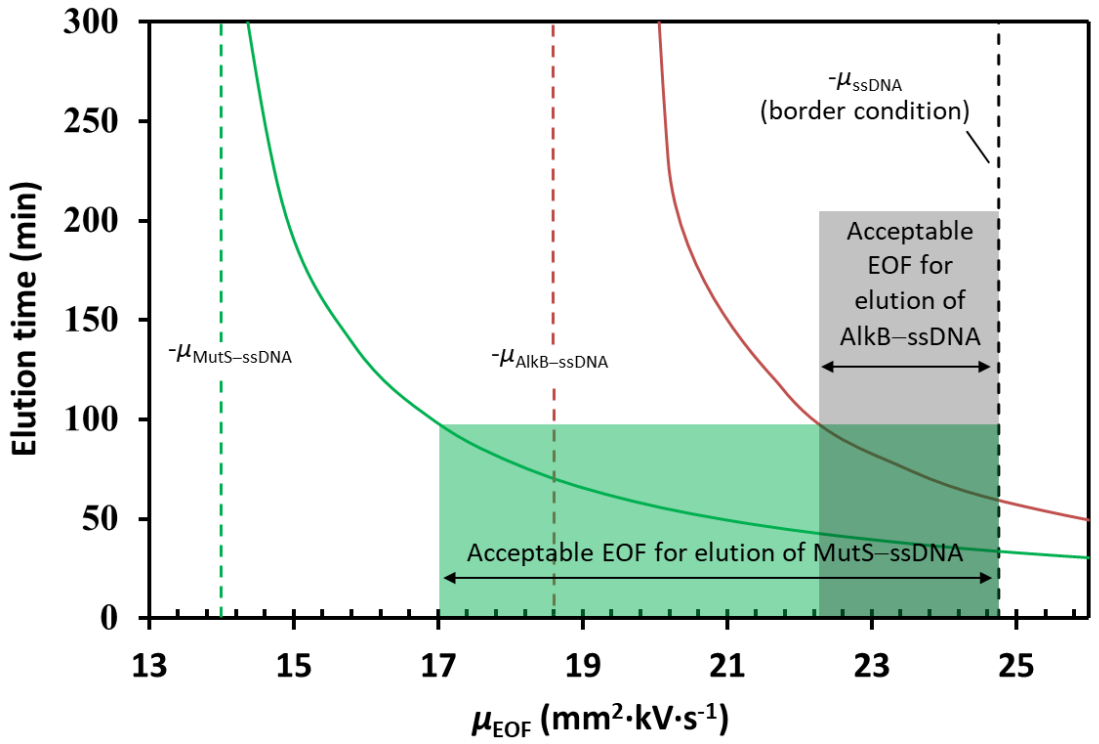

Figure 5.1. Elution time for the protein–aptamer complex as a function of the mobility of EOF (μ_{EOF}). The black-dotted line corresponds to the absolute value of the electrophoretic mobility of ssDNA. Green and red-dotted lines correspond to the absolute values of the experimental electrophoretic mobilities for Muts–ssDNA and AlkB–ssDNA complexes, respectively (electropherograms are not shown); green and red-curved lines correspond to the calculated elution time for these complexes as a function of μEOF. Green and grey rectangular areas correspond to the acceptable EOF range for the elution of the Muts–ssDNA and AlkB–ssDNA complexes at reasonable time. The left border of both areas corresponds to the EOF at which the elution happens at 100 minutes and the right border corresponds to the condition when μEOF = |μssDNA|, indicating minimum elution times for Muts–ssDNA and AlkB–ssDNA, that are 34 and 60 min, respectively.

Here, L_{tr} corresponds to the travel distance from the middle of the sample plug to the detection point and L_{tot} corresponds to the total length of the capillary; V corresponds to the applied voltage. Mobilities of EOF and the protein–aptamer complex are assigned as μ_{EOF} and

$\mu_{P-ssDNA}$, respectively. Mobilities of Muts–ssDNA and AlkB–ssDNA were measured experimentally (electropherograms are not shown).

For this example, the value of EOF is considered as acceptable since it's smaller than the absolute mobility of ssDNA (which is equal to 24.5 mm^2/kVs^{-1} and corresponds to the maximum allowed value of EOF) and higher than the value of EOF at which the complexes would be eluted at 100 minutes. An elution time of 100 minutes was chosen manually as the maximum allowable time of the experiment and can be adjusted. However, the higher the elution time, the more likely the loss of binders due to the dissociation of the complex in the capillary or due to the absorption of the collected binders on the vial's walls. The elution time of 100 minutes corresponds to EOF values of 17.0 mm^2/kVs^{-1} and 22.3 mm^2/kVs^{-1} for MutS–80-nt ssDNA and AlkB–80-nt ssDNA complexes, respectively. Thus, for the chosen experimental conditions, the acceptable EOF ranges for the complexes of MutS and AlkB are 17.0-24.5 and 22.3-24.5 mm^2/kVs^{-1}, respectively (**Figure 5.1**). Percentage-wise, the allowable deviations of the EOF from its experimental value are [-28%, +4%] for the complex of MutS and only [-5%, +4%] for the complex of AlkB. From here, it becomes clear, that even a small drift in EOF may ruin the whole IFCE experiment, especially in the case of aptamer selection for proteins with small molecular weights. The only solution to this problem is to minimize the experimental error in EOF and increase reproducibility by adequately conditioning the capillary.

In this chapter, I provided a comparison of several different capillary conditioning techniques and selected the best one with the most stable and reproducible EOF. I also noticed that the capillary siphoning effect could significantly change the apparent value of EOF especially in IFCE conditions. Solving the siphoning problem by equalizing the height of the buffers at the inlet and outlet allowed us to achieve a better control of EOF. Achieving

reproducible and controllable EOF was the first part of the IFCE optimization procedure. In the second part, I explained the possible reasons for the formation of pseudo-complexes between the analytes and salt-related ions (such as NaCl). This phenomenon was observed when the concentration of salt in the sample plug was much higher than that of the rest of the buffer in the capillary. Salt-related pseudo-complexes were well-detected in CE and could be mistakenly considered as a protein-aptamer complexes. In this chapter, I proposed the mechanism of formation for such complexes and suggested a simple strategy for their elimination.

5.2. MATERIALS AND METHODS

5.2.1. Chemicals and materials

All chemicals were purchased from Sigma-Aldrich (Oakville, ON), unless otherwise stated. Fused silica capillaries with inner diameter of 75 μm and outer diameters of 360 μm were purchased from Polymicro (Phoenix, AZ). Chromeo P503 pyrylium dye was purchased with Active Motif (Burlington, ON). AlkB and MutS proteins were expressed and purified in our lab according to the published procedure **[63]**. Recombinant *A. victoria* GFP protein was purchased from abcam (Cambridge, MA). All DNA sequences were purchased from IDT (Oralville, IA). The sequences of MutS and AlkB ssDNA aptamers were 5′-/56-FAM/CTT CTG CCC GCC TCC TTC CTG GTA AAG TCA TTA ATA GGT GTG GGG TGC CGG GCA TTT CGG AGA CGA GAT AGG CGG ACA CT-3′ and 5′-/56-FAM/CTC CTC TGA CTG TAA CCA CGT GCC TAG CGT TTC ATT GTC CCT TCT TAT TAG GTG ATA ATA GCA TAG GTA GTC CAG AAG CC-3′. Tris-HCl (50 mM final concentration, pH 7.0) with 100 mM NaCl was used as a dilution buffer and as a running buffer in CE experiments with MutS and AlkB proteins, in the experiments with capillary conditioning and in the experiments related to the siphoning effect. Tris-Acetate (50 mM final concentration, pH 8.2) and Tris-HCl (50 mM final

concentration, pH 7.0) with varying concentrations of NaCl were used as a running buffer and as a dilution buffer in all experiments with salt-related (NaCl-related) pseudo-complexes. Bodipy (Difluoro{2-[1-(3,5-dimethyl-2H-pyrrol-2-ylidene-N)ethyl]-3,5-dimethyl-1H-pyrrolato-N}boron) was used as internal standard for accurate mobility calculations. All solutions were prepared in deionized water filtered through a 0.22 μm Milipore filter membrane (Nepean, ON).

5.2.2. Capillary electrophoresis

All CE experiments were conducted using P/ACE MDQ instrument from Sciex (Brea, CA) equipped with LIF detection system. The excitation wavelength had been set at 488 nm. Fluorescent signal from the species was detected at 520 nm and 630 nm. The wavelength of 630 nm was only used in the experiments with Chromeo-labeled proteins. In all other experiments the fluorescent detection was conducted at 520 nm. Uncoated fused silica capillaries with a total length of 50.0 cm were used for all CE experiments. For all presented experiments, with the exception of the siphoning effect experiments, the sample plug with a length of 1 cm was injected into the capillary by applying a short pressure pulse of 0.5 psi (3.45 kPa) for 10 seconds. Another pressure pulse of 0.3 psi (2.07 kPa) for 90 seconds was applied to propagate the sample plug through the uncooled region (5.9 cm), so the travel distance from the middle of the plug to the detection point was equal to 33.9 cm; and the travel distance from the middle of the plug to the capillary outlet was equal to 44.1 cm. For the experiments with capillary siphoning effect, the travel distance from the middle of the plug to the detection point was 5.7 cm. For the experiments where we investigated the impact of the salt concentration on the formation of complexes between the salt-related ion and other molecules (such as proteins, bodipy, fluorescein), the applied electric field was 500 V/cm; in all other cases, CE was conducted at the applied electric field of 200 V/cm. Capillary coolant temperature was kept at 15°C for all CE

experiments, unless otherwise stated. For all presented experiments, capillary inlet was located near the positive electrode and capillary outlet was located near the negative electrode. In these conditions, the EOF was directed from the inlet to the outlet. The level of the running buffer in the inlet's vial was equal to the level of the running buffer in the outlet's vial, unless otherwise stated.

5.3. CAPILLARY CONDITIONING PROCEDURES

5.3.1. Protocol 1. Standard conditioning procedure

Brand new bare silica capillary with inner diameter of 75 μm and total length of 50 cm was rinsed with methanol for 10 minutes at 20 psi (137.9 kPa, 50 capillary volumes) and then with each of the following solutions: 0.1M HCl, 0.1M NaOH and deionized water for 3 minutes at 20 psi each (137.9 kPa, 15 capillary volumes per solution). Next, capillary was rinsed with the running buffer (50 mM Tris-HCl + 100 mM NaCl, pH 7.0) for 10 minutes at 10 psi (69.0 kPa, 25 capillary volumes). Upon completion of the rinsing, electric field of 200 V/cm was applied to capillary for 10 min to finish the equilibration. At every step, the temperature of the capillary was kept at 15 °C. The same procedure was also performed with the running buffer without salt (as a separate experiment).

5.3.2. Protocol 2. Overnight conditioning procedure

Brand new bare silica capillary with inner diameter of 75 μm and total length of 50 cm was conditioned as described in protocol 1. Then, capillary prefilled with the running buffer was left overnight (approximately for 10 hours) at 25 °C.

5.3.3. Protocol 3. Modified conditioning procedure

Step 1. Brand new bare silica capillary with inner diameter of 75 μm and total length of 50 cm was sequentially rinsed with each of the following solutions: methanol for 10 minutes at 20

psi (137.9 kPa, 50 capillary volumes), 0.1M HCl for 3 minutes at 20 psi (137.9 kPa,15 capillary volumes), 0.1M NaOH for 6 minutes at 20 psi (137.9 kPa, 30 capillary volumes), deionized water for 3 minutes at 20 psi (137.9 kPa, 15 capillary volumes) and running buffer (50 mM Tris-HCl + 100 mM NaCl, pH 7.0) for 40 minutes at 40 psi (275.9 kPa, 410 capillary volumes). At step 1, the temperature of the capillary was kept at 15 °C.
Step 2. After completion of step 1, capillary was rinsed with each of the following solutions: 0.1M HCl, 0.1M NaOH, deionized water and running buffer for 3 minutes at 20 psi each (137.9 kPa, 15 capillary volumes per solution). Upon completion of step 2, electric field of 200 V/cm was applied to capillary for 30 min to finish the equilibration. At step 2, the temperature of the capillary was kept at 25 °C.

5.3.4. Rinsing between runs

Upon completion of each conditioning procedure, the resulting EOF was estimated by measuring the migration time of the neutral marker. For the following runs within one conditioning procedure, capillary was rinsed with the running buffer for 3 minutes at 20 psi (137.9 kPa, 15 capillary volumes).

5.3.5. Sample preparation

The concentration of bodipy (neutral marker) was equal to 500 nM in all experiments. In the experiments with Muts–ssDNA and AlkB–ssDNA complexes, the concentration of both proteins was equal to 500 nM, and the concentration of the corresponding 80-nt long aptamer (ssDNA) was equal to 100 nM. For the experiments where we investigated the impact of the salt concentration on the formation of complexes between the salt-related ions and other molecules, the concentrations of GFP, chromeo-labeled MutS, fluorescein and rhodamine-6G in the sample plug were 200 nM, 130 nM, 200 nM and 400 nM, respectively. The concentration of NaCl in the

sample plug varied from 0 to 6 M. MutS protein was labeled with the Chromeo P503 dye according to the dye manufacture's procedure.

5.4. RESULTS AND DISCUSSION

5.4.1. The impact of the capillary conditioning on EOF

In this section, we investigated the impact of several different capillary conditioning procedures on EOF. The goal was to find the best procedure to achieve stable and reproducible EOF in a short period of time. The first procedure tested was the standard conditioning procedure. We have been using this procedure in our lab for the past 10 years and considered it the standard to follow, however never evaluated its impact on EOF. To test this procedure, we used two running buffers with different concentrations of NaCl in them: 50 mM Tris-HCl + 100 mM NaCl, pH 7.0 and 50 mM Tris-HCl + 0 mM NaCl, pH 7.0. These two buffers were chosen because we usually control EOF in CE by varying the salt concentration in a range between 0 and 100 mM. Upon completion of the capillary conditioning with the standard conditioning procedure, we performed six sequential CE runs (20 minutes each), using bodipy as a neutral marker to measure the EOF (**Figure 5.2**). After each run, the capillary was rinsed with the corresponding running buffer. As follows from **Figure 5.2**, the stabilization of the electroosmotic flow only happened at run 6 for both buffers, and the value of EOF at run 6 corresponds to EOF at the equilibrium between the running buffer and capillary walls. This means that the standard conditioning procedure is not sufficient for the establishment of the full equilibrium between the buffer and capillary walls. The presence of NaCl in the buffer solution had significant impact on the variation of EOF from run to run. Thus, in the case of the running buffer with 100 mM NaCl in it, the EOF ranged from 15.56 to 21.16 mm^2/kVs^{-1} for runs 1-6; and for the running buffer without NaCl, the EOF ranged from 27.55 to 30.65 mm^2/kVs^{-1} for the same number of runs.

Such a large difference in the EOF ranges can be explained by the well-known ability of sodium ions to intercalate glass and change its surface properties.

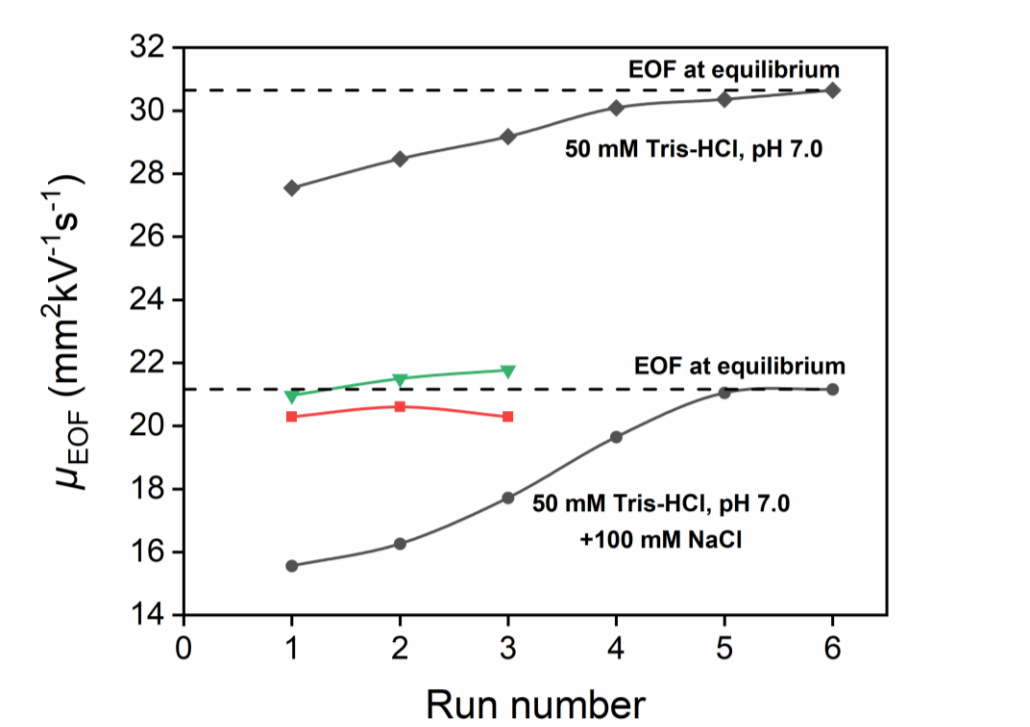

Figure 5.2. **The impact of capillary conditioning procedures on EOF.** The lower (●) and upper curves (♦) correspond to the changes in EOF from run to run for the capillary that was conditioned with 50 mM Tris-HCl + 100 mM NaCl at pH 7.0 and 50 mM Tris-HCl at pH 7.0, respectively, using standard conditioning procedure. Dash lines correspond to the values of EOF for the fully equilibrated capillaries. The green and red curves represent the changes in EOF from run to run for the capillaries that were conditioned with 50 mM Tris-HCl + 100 mM NaCl at pH 7.0, using modified conditioning procedure (▼) and overnight conditioning procedure (■), respectively.

For our further experiments, we assumed that the capillary conditioning procedure that allowed to reach stable EOF in the presence of 100 mM NaCl would also work for the buffers with smaller concentrations of salt. In order to fully equilibrate the capillary and confirm that EOF at plateau corresponds to EOF in equilibrium, we filled capillary with 50 mM Tris-HCl +

100 mM NaCl at pH 7.0 and left it overnight (about 10 hours). The next day, we performed three sequential CE runs (**Figure 5.2**, red trace**)** and found that average EOF value was equal to 20.30 ± 0.18 mm^2/kVs^{-1}, which deviated from the EOF at equilibrium by only 3.62%. Thus, overnight conditioning procedure is a simple way to fully equilibrate the capillary with the running buffer and reach a stable EOF; however, this procedure is very lengthy, which limits its applicability. To overcome this limitation, we developed a modified capillary conditioning procedure. The idea of this procedure was to reach the solution-glass equilibrium by increasing the number of rinses with the running buffer (see materials and methods for details). Increasing this number to 410 allowed us to fully equilibrate the capillary in just two hours. Three subsequent CE runs (**Figure 5.2**, green trace) resulted in the average EOF value of 21.41 ± 0.41 mm^2/kVs^{-1}, which was different from the EOF at equilibrium only by 1.20%.

Very often, CE experiments require the use of several buffers. Thus, we decided to test if it's possible to re-equilibrate the used capillary with another buffer. To accomplish that, we first took a brand new capillary and equilibrated it with 50 mM Tris-HCl + 75 mM NaCl at pH 7.0, using modified conditioning procedure. Upon completion of the conditioning, the resulting EOF was equal to 23.33 mm^2/kVs^{-1}. Next, we equilibrated another brand new capillary with a running buffer that contained higher concentration of salt: 50 mM Tris-HCl + 100 mM NaCl at pH 7.0, and finally, implemented modified conditioning procedure on the same capillary, using original buffer with 75 mM NaCl in it. The resulted EOF for the reconditioned capillary was equal to 21.43 mm^2/kVs^{-1}, which was different from the original EOF for the buffer with 75 mM NaCl by 8.11%. Therefore, it is impossible to re-condition a used capillary with another buffer using the modified conditioning procedure. Thus, to keep EOF stable and experiments reproducible, it is better to use a new capillary every time when a buffer change is needed.

5.4.2. Capillary siphoning effect

The nature of the capillary siphoning effect is well-known and arises from the difference of the buffer's heights at the outlet and inlet of the capillary. This difference creates a weak pressure-driven hydrodynamic flow that coexists along with EOF and can move to the same or opposite direction as EOF. In both cases, it makes the measured value of EOF different from the real value. The siphoning effect can be easily eliminated by equalizing the heights of the buffers at the inlet and the outlet; however, it's not always possible especially if the experiment involves collection of the buffer's fractions in a very small volume of the running buffer at the capillary outlet. It is preferable to use a small volume of the running buffer since it allows to minimize the dilution factor of the species in the collected fractions. For instance, in the case of our IFCE experiments, the volumes of the running buffer at the outlet and inlet are usually equal to 20 and 500 μL, respectively (**Figure 5.3a**). This difference in volumes corresponds to the 1.5 cm difference in the buffers' heights, which results in the apparent EOF value of 22.71 mm^2/kVs^{-1} for 50 mM Tris-HCl + 100 mM NaCl, pH 7.0. To assess the contribution of the siphoning effect to the apparent EOF, we equalized the heights of the buffer at the inlet and outlet by removing the disposable 0.6 mL CE buffer vial at the inlet and filling the vacant space of non-disposable supporting CE vial with the running buffer (**Figure 5.3b**). This improvement resulted in the real value of EOF being equal to 20.11 mm^2/kVs^{-1}, which is smaller than the apparent value by 11.45%. Such a huge difference in EOF may result in a significant shift of the migration times of the species, whose electrophoretic mobilities are close to EOF. For instance, the AlkB–aptamer complex with electrophoretic mobility of -18.59 mm^2/kVs^{-1} was detected 11.5 minutes in the first experiment with apparent EOF (**Figure 5.3c**, upper trace); however the same complex did not show up even after 30 minutes, when the siphoning effect was eliminated and the experiment was conducted at real EOF (**Figure 5.3c**, lower trace).

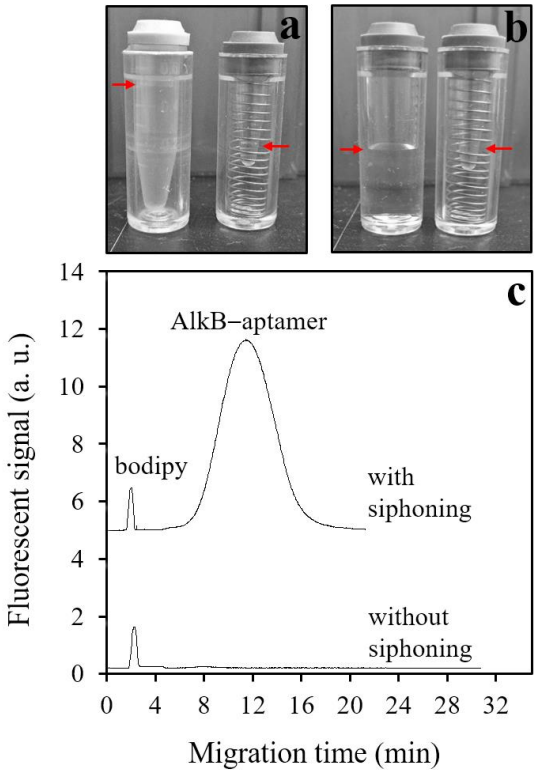

Figure 5.3. Detection of the AlkB–aptamer complex in the presence and absence of the capillary siphoning effect. Panel **a** shows the experimental setup that results in the capillary siphoning effect. Panel **b** shows the experimental setup that eliminates the effect of the capillary siphoning. For both panels, the left and right vials correspond to the inlet and outlet, respectively. Red arrows indicate the levels of the buffer in the vials: 500 µl and 20 µl for panel **a**; and 660 µl and 20 µl for panel **b**. The upper trace on panel **c** corresponds to CE experiment with the setup illustrated on panel **a**, and the lower trace corresponds to the CE experiment with setup illustrated on panel **b**. In both experiments, the analyzing mixture contained a neutral EOF marker (bodipy) and the AlkB–aptamer complex (see Materials and Methods for details). For the presented experiments, the apparent and real EOF values were equal to 22.71 and 20.11 mm^2/kVs^{-1}, respectively. The higher value corresponds to the apparent EOF in the presence of the capillary siphoning and the smaller value corresponds to the real EOF in the absence of the siphoning effect. The AlkB–aptamer complex was only detected at higher EOF (in the presence of capillary siphoning); at lower EOF, the migration time of the complex was beyond the time of the experiment.

As it can be seen from **Figure 5.3**, such a minor detail as the capillary siphoning effect may lead to irreproducible experiments. In addition, the presence of the capillary siphoning effect corrupts the measured EOF. Therefore, we strongly recommend conducting any CE experiment in the absence of the capillary siphoning effect, especially for experimental conditions in which the mobilities of the analyzing species are close to the value of EOF

5.4.3. Formation of salt-related pseudo-complexes

As it was mentioned previously, formation of the complexes between the target molecules and the binders (potential drug candidates) is the first and most important step in any drug screening approach. The success of this step depends on parameters such as affinity of the binders to the target molecules, abundance of the binders in the original pool and the absence of non-specific binding that may result in the formation of complexes that are different from the target-binder complexes. Non-specific binding usually happens when the binders are conjugated with the tags, which may interact with the target molecules and form complexes, or when the mixture of the targets with binders contains third-party molecules/ions (such as buffer's components or contaminants) that have affinity to the target molecules. Using untagged binders and components with no affinity to the targets minimizes the risk of the formation of non-specific complexes, but does not exclude the risk completely. In this section, we describe a special case of the formation of complexes between the analyzing species and components of the buffer that had no affinity to the targets. Such a case has already been described in the literature for DNA-Na^+ complexes **[134]**, but has never been reported for the proteins and small molecules. We could not find any reasonable explanation for the mechanism of the formation of such complexes in the literature and proposed our own hypothesis. Understanding the nature of these complexes will help avoiding them in the future.

The complexes between the analyzing species (GFP, Chromeo-labeled MutS, fluorescein) and the counterions were formed in the presence of high concentrations of sodium chloride in the sample plug (**Figure 5.4**). At the same time, the concentration of sodium chloride in the running buffer was equal to zero. If the sample buffer and the running buffer were identical and contained the same concentration of sodium chloride, the formation of the complexes was not observed. As can be seen from **Figure 5.4**, the complexes of the analyzing species with counterions have the same migration time as a neutral marker, which indicates the absence of charge on them. Noticeably, the neutral marker by itself did not form such complexes and was migrating as a single peak at different concentration of sodium chloride. Therefore, one of the necessary conditions for the formation of the complexes is the presence of non-zero charge on the analyzing species in the reaction mixture. All species in **Figure 5.4** had negative charges (except for the neutral marker) and, most likely, formed complexes with Na^+ ions. When a positively charged molecule such as Rhodamine-6G was used (data not shown), complex formation was also observed; however, in this case the role of the counterion belonged to Cl^-. The same hypothesis of the tight condensation of counterions on the surface of the charged species was also proposed for the similar DNA-Na^+ complexes **[134]**. The hypothesis seems to be correct from the point of view of the electrostatic interactions; however, it does not explain the stability of such complexes. It does not answer the question of why these complexes exist in the non-dissociative form. Especially it is very questionable for the fluorescein sodium salt (**Figure 5.4c**), since this compound is a strong electrolyte and should exist in the aqueous solutions in a fully dissociated form. These discrepancies made us revise the proposed hypothesis. Thus, we formulated another hypothesis that represents the sample plug as a big induced dipole. To explain this hypothesis, we can refer to the example of the polarization of the

dielectric in an external electric field. Application of an external field leads to the redistribution of charges inside of the dialectic, which in turn results in the formation of an internal electric field that compensates for the external one.

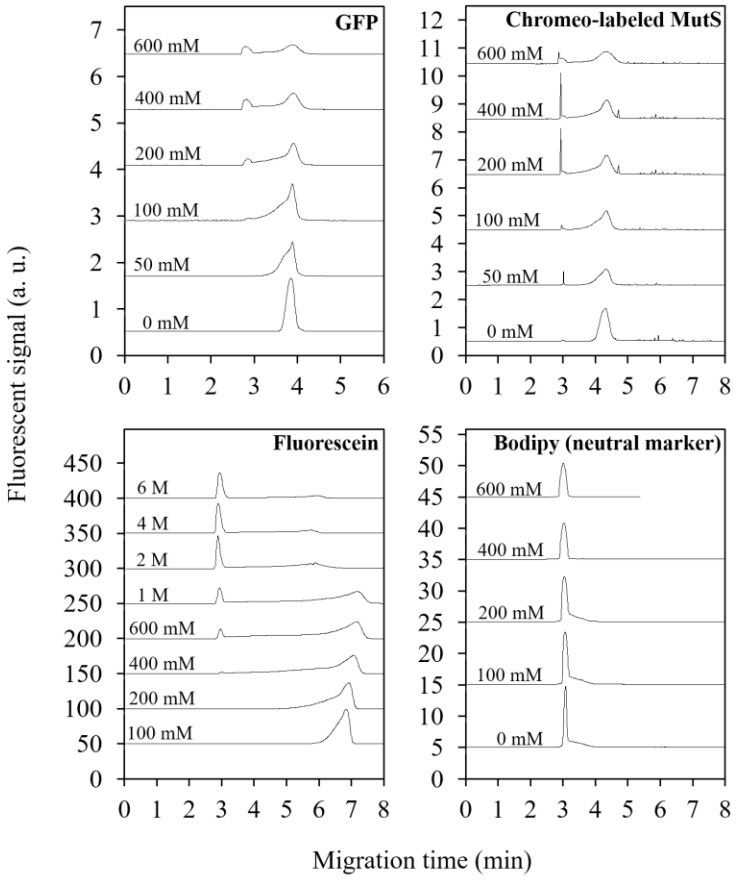

Figure 5.4. **Detection of salt-related pseudo-complexes in CE. All panels correspond to the CE analysis of the species (GFP, Chromeo-labeled MutS, fluorescein, bodipy) in the presence of sodium chloride in the sample plug**. The concentration of sodium chloride for each particular experiment is written on the left. The running buffer and the dilution buffer was 50 mM Tris-Acetate, pH 8.2 for all presented experiment. Sodium chloride was only presented in the sample plug, but not in the running buffer. NaCl-related complexes have neutral charge and migrate along with EOF. The original form of all presented species (with exception for bodipy) is negatively charged. See materials and methods for more details. Y-axis offset was applied for the clarity of presentation.

The degree of the compensation is determined by the nature of the dielectric and can be numerically expressed by the dielectric constant. Depending on the nature of the particles that the dielectric is made of, the mechanism of charge redistribution can be different. Thus, if the dielectric is made of dipoles, they change their position in space, aligning themselves along the external field (orientational polarization). If the dielectric is made of neutral particles, they first gain the dipole moment as because of the displacement of the electron density relative to the nuclei (electron polarization) and only then align themselves along the external field. If the dielectric has an ionic crystal lattice, the application of the electric field results in the displacements of the nodes of the crystal lattice and/or in the displacement of the ions in the structural cavities (ionic polarization). By the analogy of the redistribution of charges in the dielectrics, similar processes may occur inside of the sample plug. We think that the application of the electric field may polarize the sample plug by making clouds of Na^+ and Cl^- ions move to the borders of the plug. Upon reaching a certain distance, clouds stop moving relative to each other. At this point, the external electric field is completely compensated by the internal electric field created between the clouds inside of the plug. The formed macro dipole moves along with EOF, holding all the molecules that were caught inside of it, including the analyzing species. Thus, the enclosed molecules are shielded from the external electric field and move along with EOF. Depending on the original concentration of the sodium chloride in the plug, the size of the dipole (and the dipole moment) may vary. We assume that the size of the dipole is directly proportional to the concentration of NaCl. Thus, at small NaCl concentrations, the dipole is so small, that no visible signal from the encapsulated species can be found on the electropherograms (**Figure 5.4**). However, at very high concentrations of salt (**Figure 5.4c**), the analyzing species are almost fully encapsulated and no longer exist in a free form. We assume

that the analyte molecules that are located in the clouds (but not between) may not be completely shielded from the external field: the closer they are to the outer border of the cloud, the more they "sense" the external field. Thus, these molecules can migrate beyond the macro dipole, but the values of their electrophoretic mobilities are somewhere between the value of EOF (which corresponds to the fully encapsulated molecule) and the mobility of the free form of the original molecule. This suggestion may explain the existence of the dissociation region between the free and encapsulated forms of the molecules on the electropherograms (**Figure 5.4**). We also think that the dissociation region may contain molecules that could diffuse through the ion clouds to the bulk of the solution. We believe that not only NaCl, but also other highly dissociative salts may encapsulate the analytes and drug them to the point of detection. These stable dipole-like structures with analytes encapsulated in them were called salt-related pseudo-complexes. In this work, the word "salt" refers to NaCl only.

The observed phenomenon has great fundamental importance, as it has not been previously described. From a practical point of view, formation of salt-related pseudo-complexes is an undesirable event that must be avoided in any CE-based separation technique because the formed complexes may be misinterpreted as binders-target complexes. Avoiding high concentration of salts (and other strong electrolytes) in the sample plug and using only one buffer system across the capillary can prevent formation of salt-related pseudo-complexes. Ideally, the concentration of highly-dissociative salts in the sample plug should be kept at the lowest possible value.

5.5. CONCLUSIONS

In this chapter, I described several important details that must be considered in the development of any CE-based separation approach. First, I tested several different capillary conditioning techniques, assessed their impact on EOF and proposed my own technique that

allowed to achieve stable and reproducible EOF. Second, I paid attention to the capillary siphoning effect and showed how impactful is on the apparent value of EOF. For our experiments, elimination of the siphoning effect resulted in the change of EOF by 11.45%. Therefore, I proposed a simple vials' setup that allows conducting both, preparative and analytical CE experiments in the absence of siphoning effect. Finally, I observed formation of the complexes between NaCl-related ions and the analyzing species. Formation of these complexes is undesirable because they may be mistakenly interpreted as other species in CE and the signal from them on an electropherogram may overlap with signals from other species. To the best of my knowledge, the formation of such complexes has never been described in the literature. I suggested a mechanism for their formation, based on the consideration of the sample plug with NaCl as an induced dipole. The hypothesis about their nature has an important fundamental value and can help avoiding them in future research.

LIMITATIONS

The main limitation of the developed mathematical model for predicting electrophoretic mobilities of protein-aptamer complexes is the assumption of their spherical shape. This assumption worked perfectly for predicting electrophoretic mobilities of 10 protein-aptamer complexes (see Appendix A for details); however we do not exclude the possibility that the developed model can be less accurate in respect to the prediction of the electrophoretic mobilities of the complexes, whose shape cannot be approximated as a sphere. A good example of such molecules could be complexes of long cylindrical proteins and short aptamers. In addition to that, the model was developed for the complexes with simplest stoichiometry, which considers binding of one aptamer molecule to one protein molecule. For the complexes with stoichiometry different from one-to-one, the predicted mobilities can be different from the real ones. Another limitation of the developed model is the fact that it's empirical. Empirical models are known to be less reliable than analytical models; however, their reliability can be improved by increasing the number of experimental data points that the model is built on.

In our work on the explanation of abnormal electrophoretic behaviour of the complexes between bivalent antibodies and rod-like virions, we proposed a simple mathematical model for predicting the number of cross-linked virus particles in the complex with antibodies using electrophoretic mobilities of the complex and the virions as input parameters. The main limitation of the model arises from the assumption of negligible contribution of the antibodies to the molecular weight and the charge of the complex. Such an assumption was made on the experimental confirmation of the ability of antibodies (Ab) to cross-link virions even at the very small Ab:virion ratio, which was equal to 10:1. At this ratio, the binding of 10 Ab molecules to one virion would increase its molecular weight by 4%. As for the charge contribution, we

assumed it to be negligible, since antibodies had electrophoretic mobilities somewhere close to the mobility of EOF, which indicated their nearly zero charge. However, we do not exclude the possibility that for some other cases of the complex formation between Ab and rod-like virions, the impact of antibodies on the electrophoretic mobility of the complex may be significant. Thus, If the mass contribution of antibodies to the molecular weight of the complex is higher than 10% and/or the charge of antibodies is different from zero and can't be neglected, the deviations from the proposed model can be observed. Also, we cannot exclude the chance that virus-antibody complexes may interact with capillary walls. Such an interaction may alter their migration times and calculated electrophoretic mobilities. In order to evaluate the impact of these interactions on the electrophoretic mobilities of the complexes, an additional experiment with coated capillary must be conducted. The coating should prevent the adsorption of the complexes, virus and antibodies on capillary walls, and thus, will allow to find the unbiased migration times and mobilities of the species. CE experiments for predicting the number of cross-linked virions (illustrated in Figures 3.3) were conducted just once. Because of that, the predicted numbers should be considered as semi-quantitative. In order to get more reliable numbers, the presented experiments should be repeated.

The main limitation of the aptamer assisted purification technique that we introduced in the current work was the inability to produce purified proteins in high amounts. The reason for that was a limited amount of initial reagents, which were streptavidin-coated magnetic beads and biotinylated aptamers. It's possible to increase the output amount of the purified proteins by increasing the amount of the initial reagents, however it's not a rational approach since these reagents are very expensive. Thus, it is better to use conventional chromatographic techniques for the bulk production of the proteins. Another limitation of our procedure is the fact that it was

validated on example of only two proteins. Thus, it would benefit from showing its validity on the example of another protein.

CONCLUDING REMARKS

The main achievements of the presented work can be summarized in the following sentences. We developed an empirical model that allowed for the accurate prediction of electrophoretic mobilities of protein–aptamer complexes (i). We explained an abnormal electrophoretic behaviour of the complexes between rod-like virions and bivalent antibodies, and proposed an empirical formula for calculating the number of virions in the corresponding complexes (ii). We developed a fast protein purification technique where aptamers attached to magnetic beads were used as affinity probes. We also proved that reversible binding of the aptamers to the proteins did not affect their native properties and thus can be used for their fast extraction from the cell lysate (iii). We discovered a previously unknown phenomenon of formation of neutral pseudo-complexes between any charged molecules and counterions in the presence of high concentration of salt in the sample plug (iv). For macromolecules (GFP and MutS), this phenomenon was observed at the salt concentrations of 100 mM or higher; for small molecule (Fluorescein), this phenomenon became noticeable at higher concentrations of salt (400 mM and higher).

All presented achievements have great scientific potential and can complement the existing method for screening biologics for their ability to bind proteins. The developed mathematical model for predicting electrophoretic mobility of protein–aptamer complexes can be used as an important addition to the existing IFCE and NECEEM methods of aptamer selection.

The discovered phenomenon of unusual electrophoretic behavior of the complexes between rod-like virions and bivalent antibodies complements the existing CE-based immunoassays. Moreover, the proposed empirical formula for predicting the number of cross-

linked virions in the complexes turns the observed phenomenon into a quantitative method for the evaluation of the antibodies' immunoprecipitation activity.

The importance of using active proteins in their native state in any drug screening approach led us to the development of a fast and reliable protein purification protocol. In our protocol, aptamers attached to the magnetic beads acted as an affinity probe for extracting proteins from the cell lysate. The proteins purified by the developed protocol were more active than that purified by standard chromatographic techniques.

Conducting experiments with high concentration of NaCl in the sample plug (100 mM and higher) led to the discovery of the analyte's peak splitting phenomenon. The resulting electropherogram contained two peaks: one peak corresponded to the unbiased peak of the analyte and another one corresponded to the neutral form of the analyte, which moved along with EOF. To the best of our knowledge, the observed phenomenon has not been described in the literature. We proposed that the plug of NaCl works as a macrodipole (or polarized dielectric) with an internal electric field that compensates for the external one. Thus, species being encapsulated in the macrodipole do not sense the external field and move along with EOF. These complexes can be mistakenly considered as target–ligand complexes, and must be avoided. Awareness of this phenomenon is very important for any CE-based drug screening approach since it allows distinguishing salt-related pseudo-complexes from the real target-ligand complexes.

FUTURE PLANS

The developed empirical model for predicting electrophoretic mobility of the protein-aptamer complexes significantly advanced the newly introduced IFCE method for the screening and selection of aptamers; however, this method is still not perfect and can not guarantee selection of high affinity aptamers for any protein in a single round. Despite of its name, which is "Ideal Filter CE", this method still allows for some non-binding ssDNA molecules to pass through. These molecules migrate in the same direction with EOF and form the ssDNA background. The presence of the background limits selection of high affinity aptamers and may even result in the failure of the whole selection process, especially if the amount of non-binding DNA in the background is higher than the amount of binding aptamers. We believe that DNA background has the same nature as the observed phenomenon of the formation of salt-related pseudo-complexes, since any undialyzed solution of ssDNA contains high concentration of alkali metal salts *a priori*. Thus, our future research will be focused on the deeper investigation of this phenomenon and on the development of the ways for complete elimination of the DNA background in IFCE conditions. Our most recent step towards the better understanding of the nature of the background was the establishment of the correlation between the length of the sample plug and the amount of leaked ssDNA molecules in IFCE conditions (**Figure C1**). We found that longer sample plugs resulted in a higher level of ssDNA background at the same travel distance as that for the shorter sample plugs. Thus, one of the possible ways of minimizing the level of ssDNA background is decreasing the length of the sample plug.

Turning IFCE into the ideal separation technique that would allow selection of high affinity aptamers to any protein in one round is our main future goal. The aptamers selected by

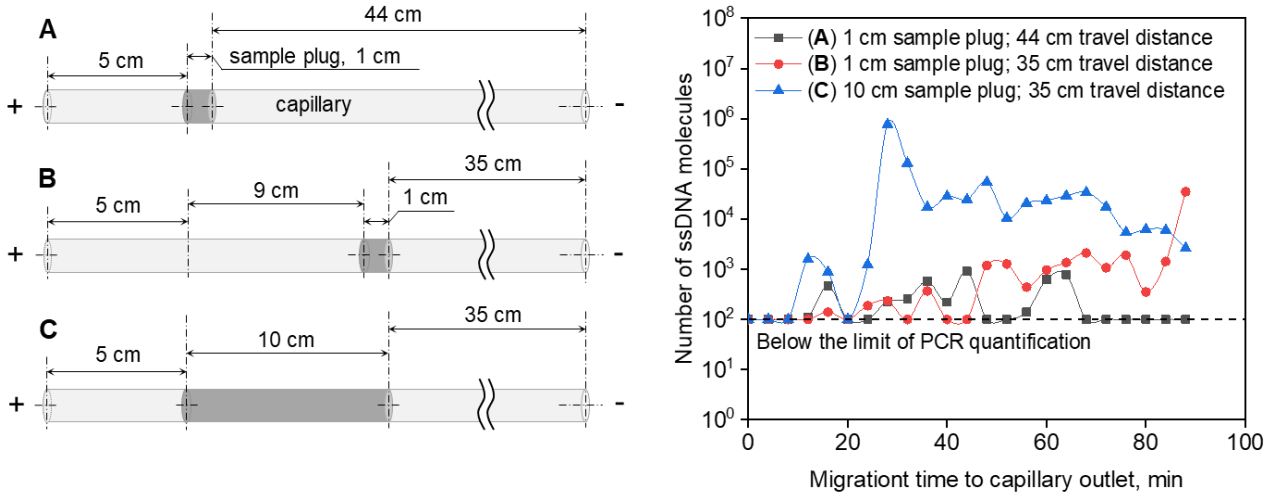

Figure C1. **The length of the sample plug and its position in the capillary affect the level of ssDNA background in IFCE conditions.** The panel on the left corresponds to the experimental CE setups that were tested. The panel on the right demonstrates the level of ssDNA background for the corresponding experimental setups. The number of ssDNA molecules on the right panel was established by PCR amplification of the collected fractions (4 min per fraction). In the experiments (A) and (B) the length of the sample plug was 1 cm. In the experiment (C) the length of the sample plug was 10 cm. The travel distance from the right border of the plug to the capillary outlet was equal to 44 cm in the experiment (A) and to 35 cm in the experiments (B) and (C), respectively. In all presented experiments, the first interval of 5 cm corresponds to the uncooled segment of the capillary. This segment is always avoided by propagating the sample plug deeper unto the capillary. For all presented experiments, both, the running buffer and the sample buffer were 50 mM Tris-HCl with addition of 100 mM NaCl (pH = 7.0, IRB = 146). CE was carried out in the presence of external electric field of 200 V/cm. The concentration of ssDNA in the sample plug was 1 μM.

this method can be further used as affinity probes for the purification of the corresponding proteins in their native state. In addition to that, we want to diversify the random ssDNA libraries that we are going to use as an initial source of aptamers. We believe that chemically modified

libraries and the libraries with altered number of nucleotides will favour the selection of high affinity aptamers.

The observed phenomenon for the unusual electrophoretic behaviour of the complexes between rod-like virions and bivalent antibodies inspired us on the possibility of using heavy drag tags, whose mobilities depend on their shapes. Thus, these tags can serve as CE mobility shifters for the separation of complexes between the target protein and the small molecules. For example, labeling different molecules with single-wall nanotubes of different lengths will make them distinguishable in CE, even if the original molecules had almost identical values of electrophoretic mobilities. Addition of the target protein to the mixture of the labeled molecules will also allow simultaneous detection and identification of the corresponding complexes in the case of their formation.

Finally, we want to combine CE with mass spectrometry and perform the screening of drug candidates (either small molecules, DNA or proteins) for their ability to bind target proteins in the conditions of Capillary Isoelectric Focusing (CIEF). Such an approach will be capable of separating species even with almost identical values of electrophoretic mobilities and will be able to provide detailed information about the composition of the target-ligand complexes.

LIST OF PUBLICATIONS

1. Beloborodov, S.S.; Krylova, S.M.; Krylov, S.N. Spherical-shape assumption for protein–aptamer complexes facilitates prediction of their electrophoretic mobility. *Analytical Chemistry* **2019**, *91*, 12680–12687.

2. Beloborodov, S.S.; Bao, J.; Krylova, S.M.; Shala-Lawrence, A.; Johnson, P.E.; Krylov, S.N. Aptamer facilitated purification of functional proteins. *Journal of Chromatography B* **2018**, *1073*, 201–206.

3. Shala-Lawrence, A.; Decker, M.J.; Krylova, S.M.; Gangopadhyay, S.S.; Beloborodov, S.S.; Hougland, J.L.; Krylov, S.N. Simultaneous analysis of non-lipidated protein and its lipidated counterpart: enabling quantitative investigation of protein lipidation's impact on cellular regulation. *Analytical Chemistry* **2017**, *89*, 13502–13507.

4. Beloborodov, S.S.; Panferov, V.G.; Safenkova, I.V.; Krylova, S.M.; Dzantiev, B.B.; Krylov, S.N. Unexpected electrophoretic behaviour of complexes between rod-like virions and bivalent antibodies. *Analytical Chemistry* **2016**, *88*, 11908–11912.

REFERENCES

1. Santos, R.; Ursu, O.; Gaulton, A.; Bento, A. P.; Donadi, R. S.; Bologa, C. G.; Karlsson, A.; Al-Lazikani, B.; Hersey, A.; Oprea, T. I.; Overington, J. P. A Comprehensive Map of Molecular Drug Targets. *Nature Reviews Drug Discovery* **2017**, *16* (1), 19–34.
2. Misono, T. S.; Kumar, P. K. R. Selection of RNA Aptamers against Human Influenza Virus Hemagglutinin Using Surface Plasmon Resonance. *Analytical Biochemistry* **2005**, *342* (2), 312–317.
3. Ng, W.; Dai, J.-R.; Slon-Usakiewicz, J. J.; Redden, P. R.; Pasternak, A.; Reid, N. Automated Multiple Ligand Screening by Frontal Affinity Chromatography–Mass Spectrometry (FAC-MS). *Journal of Biomolecular Screening* **2007**, *12* (2), 167–174.
4. Rainard, J. M.; Pandarakalam, G. C.; McElroy, S. P. Using Microscale Thermophoresis to Characterize Hits from High-Throughput Screening: A European Lead Factory Perspective. *SLAS DISCOVERY: Advancing the Science of Drug Discovery* **2018**, *23* (3), 225–241.
5. Krylov, S. N. Nonequilibrium Capillary Electrophoresis of Equilibrium Mixtures (NECEEM): A Novel Method for Biomolecular Screening. *Journal of Biomolecular Screening* **2006**, *11* (2), 115–122.
6. Krylov, S. N. Kinetic CE: Foundation for Homogeneous Kinetic Affinity Methods. *ELECTROPHORESIS* **2007**, *28* (1–2), 69–88.
7. Desruisseaux, C.; Long, D.; Drouin, G.; Slater, G. W. Electrophoresis of Composite Molecular Objects. 1. Relation between Friction, Charge, and Ionic Strength in Free Solution. *Macromolecules* **2001**, *34* (1), 44–52.
8. Meagher, R. J.; Won, J.-I.; McCormick, L. C.; Nedelcu, S.; Bertrand, M. M.; Bertram, J. L.; Drouin, G.; Barron, A. E.; Slater, G. W. End-Labeled Free-Solution Electrophoresis of DNA. *ELECTROPHORESIS* **2005**, *26* (2), 331–350.
9. Bao, J.; Krylova, S. M.; Cherney, L. T.; Hale, R. L.; Belyanskaya, S. L.; Chiu, C. H.; Shaginian, A.; Arico-Muendel, C. C.; Krylov, S. N. Predicting Electrophoretic Mobility of Protein–Ligand Complexes for Ligands from DNA-Encoded Libraries of Small Molecules. *Analytical Chemistry* **2016**, *88* (10), 5498–5506.
10. Le, A. T. H.; Krylova, S. M.; Krylov, S. N. Ideal-filter Capillary Electrophoresis: A Highly Efficient Partitioning Method for Selection of Protein Binders from Oligonucleotide Libraries. *ELECTROPHORESIS* **2019**, *40* (18–19), 2553–2564.
11. Lakhin, A. V.; Tarantul, V. Z.; Gening, L. V. Aptamers: Problems, Solutions and Prospects. *Acta Naturae* **2013**, *5* (4), 34–43.
12. Ng, E. W. M.; Shima, D. T.; Calias, P.; Cunningham, E. T.; Guyer, D. R.; Adamis, A. P. Pegaptanib, a Targeted Anti-VEGF Aptamer for Ocular Vascular Disease. *Nature Reviews Drug Discovery* **2006**, *5* (2), 123–132.
13. Lu, R.-M.; Hwang, Y.-C.; Liu, I.-J.; Lee, C.-C.; Tsai, H.-Z.; Li, H.-J.; Wu, H.-C. Development of Therapeutic Antibodies for the Treatment of Diseases. *Journal of Biomedical Science* **2020**, *27* (1), 1–30.
14. Okun, V. M.; Moser, R.; Blaas, D.; Kenndler, E. Complexes between Monoclonal Antibodies and Receptor Fragments with a Common Cold Virus: Determination of

Stoichiometry by Capillary Electrophoresis. *Analytical Chemistry* **2001**, *73* (16), 3900–3906.

15. Okun, V. M.; Ronacher, B.; Blaas, D.; Kenndler, E. Affinity Capillary Electrophoresis for the Assessment of Complex Formation between Viruses and Monoclonal Antibodies. *Analytical Chemistry* **2000**, *72* (19), 4634–4639.
16. Chong, S.; Mersha, F. B.; Comb, D. G.; Scott, M. E.; Landry, D.; Vence, L. M.; Perler, F. B.; Benner, J.; Kucera, R. B.; Hirvonen, C. A.; Pelletier, J. J.; Paulus, H.; Xu, M.-Q. Single-Column Purification of Free Recombinant Proteins Using a Self-Cleavable Affinity Tag Derived from a Protein Splicing Element. *Gene* **1997**, *192* (2), 271–281.
17. Turner, P. R.; Bourne, K.; Garama, D.; Carne, A.; Abraham, W. C.; Tate, W. P. Production, Purification and Functional Validation of Human Secreted Amyloid Precursor Proteins for Use as Neuropharmacological Reagents. *Journal of Neuroscience Methods* **2007**, *164* (1), 68–74.
18. Missailidis, S.; Hardy, A. Aptamers as Inhibitors of Target Proteins. *Expert Opinion on Therapeutic Patents* **2009**, *19* (8), 1073–1082.
19. Ellington, A. D.; Szostak, J. W. In Vitro Selection of RNA Molecules That Bind Specific Ligands. *Nature* **1990**, *346* (6287), 818–822.
20. Tuerk, C.; Gold, L. Systematic Evolution of Ligands by Exponential Enrichment: RNA Ligands to Bacteriophage T4 DNA Polymerase. *Science* **1990**, *249* (4968), 505–510.
21. Imashimizu, M.; Takahashi, M.; Amano, R.; Nakamura, Y. Single-Round Isolation of Diverse RNA Aptamers from a Random Sequence Pool. *Biology Methods and Protocols* **2018**, *3* (1), 1–13.
22. Crooke, S. T. Molecular Mechanisms of Action of Antisense Drugs. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1999**, *1489* (1), 31–43.
23. Sharad, S. Antisense Therapy: An Overview. *IntechOpen* **2019**, 1–11.
24. Hoy, S. M. Nusinersen: First Global Approval. *Drugs* **2017**, *77* (4), 473–479.
25. Geary, R. S.; Henry, S. P.; Grillone, L. R. Fomivirsen: Clinical Pharmacology and Potential Drug Interactions. *Clinical Pharmacokinetics* **2002**, *41* (4), 255–260.
26. Weng, D. E.; Usman, N. Angiozyme: A Novel Angiogenesis Inhibitor. *Current Oncology Reports* **2001**, *3* (2), 141–146.
27. Burnett, J. C.; Rossi, J. J. RNA-Based Therapeutics: Current Progress and Future Prospects. *Chemistry & Biology* **2012**, *19* (1), 60–71.
28. Hu, L.; Anand, M.; Krylova, S. M.; Yang, B. B.; Liu, S. K.; Yousef, G. M.; Krylov, S. N. Direct Quantitative Analysis of Multiple MicroRNAs (DQAMmiR) with Peptide Nucleic Acid Hybridization Probes. *Analytical Chemistry* **2018**, *90* (24), 14610–14615.
29. Nofech-Mozes, R.; Khella, H. W. Z.; Scorilas, A.; Youssef, L.; Krylov, S. N.; Lianidou, E.; Sidiropoulos, K. G.; Gabril, M.; Evans, A.; Yousef, G. M. MicroRNA-194 Is a Marker for Good Prognosis in Clear Cell Renal Cell Carcinoma. *Cancer Medicine* **2016**, *5* (4), 656–664.
30. Phipps, K.; Martinez, A.; Lu, J.; Heinz, B.; Zhao, G. Small Interfering RNA Molecules as Potential Anti-Human Rhinovirus Agents: In Vitro Potency, Specificity, and Mechanism. *Antiviral Research* **2004**, *61* (1), 49–55.
31. Hoy, S. M. Patisiran: First Global Approval. *Drugs* **2018**, *78* (15), 1625–1631.

32. Waltz, E. Polyclonal Antibodies Step out of the Shadows. *Nature Biotechnology* **2006**, *24* (10), 1181–1181.

33. Olsson, L.; Kaplan, H. S. Human-Human Hybridomas Producing Monoclonal Antibodies of Predefined Antigenic Specificity. *Proceedings of the National Academy of Sciences* **1980**, *77* (9), 5429–5431.

34. Periayah, M. H.; Halim, A. S.; Saad, A. Z. M. Mechanism Action of Platelets and Crucial Blood Coagulation Pathways in Hemostasis. *International Journal of Hematology-Oncology and Stem Cell Research* **2017**, *11* (4), 319–327.

35. Strohl, W. R. Fusion Proteins for Half-Life Extension of Biologics as a Strategy to Make Biobetters. *BioDrugs* **2015**, *29* (4), 215–239.

36. Peters, R. T.; Toby, G.; Lu, Q.; Liu, T.; Kulman, J. D.; Low, S. C.; Bitonti, A. J.; Pierce, G. F. Biochemical and Functional Characterization of a Recombinant Monomeric Factor VIII-Fc Fusion Protein: *RFVIIIFc Biochemical Characterization*. *Journal of Thrombosis and Haemostasis* **2013**, *11* (1), 132–141.

37. Joo, K.; Park, S. J.; Choi, Y.; Lee, J. E.; Na, Y. M.; Hong, H. K.; Park, K. H.; Kim, H. M.; Chung, J.-Y.; Woo, S. J. Role of the Fc Region in the Vitreous Half-Life of Anti-VEGF Drugs. *Investigative Opthalmology & Visual Science* **2017**, *58* (10), 4261.

38. Kratz, F. Albumin as a Drug Carrier: Design of Prodrugs, Drug Conjugates and Nanoparticles. *Journal of Controlled Release* **2008**, *132* (3), 171–183.

39. Alconcel, S. N. S.; Baas, A. S.; Maynard, H. D. FDA-Approved Poly(Ethylene Glycol)–Protein Conjugate Drugs. *Polymer Chemistry* **2011**, *2* (7), 1442–1448.

40. Rattu, M. A.; Shah, N.; Lee, J. M.; Pham, A. Q.; Marzella, N. Glucarpidase (Voraxaze), a Carboxypeptidase Enzyme for Methotrexate Toxicity. **2013**, *38* (12), 732–744.

41. Keating, G. M. Asparaginase Erwinia Chrysanthemi (Erwinaze®): A Guide to Its Use in Acute Lymphoblastic Leukemia in the USA. *BioDrugs* **2013**, *27* (4), 413–418.

42. Singh, R. P.; Li, A.; Bedi, R.; Srivastava, S.; Sears, J. E.; Ehlers, J. P.; Schachat, A. P.; Kaiser, P. K. Anatomical and Visual Outcomes Following Ocriplasmin Treatment for Symptomatic Vitreomacular Traction Syndrome. *British Journal of Ophthalmology* **2014**, *98* (3), 356–360.

43. Stott, D. J.; Rodondi, N.; Kearney, P. M.; Ford, I.; Westendorp, R. G. J.; Mooijaart, S. P.; Sattar, N.; Aubert, C. E.; Aujesky, D.; Bauer, D. C.; Baumgartner, C.; Blum, M. R.; Browne, J. P.; Byrne, S.; Collet, T.-H.; Dekkers, O. M.; den Elzen, W. P. J.; Du Puy, R. S.; Ellis, G.; Feller, M.; Floriani, C.; Hendry, K.; Hurley, C.; Jukema, J. W.; Kean, S.; Kelly, M.; Krebs, D.; Langhorne, P.; McCarthy, G.; McCarthy, V.; McConnachie, A.; McDade, M.; Messow, M.; O'Flynn, A.; O'Riordan, D.; Poortvliet, R. K. E.; Quinn, T. J.; Russell, A.; Sinnott, C.; Smit, J. W. A.; Van Dorland, H. A.; Walsh, K. A.; Walsh, E. K.; Watt, T.; Wilson, R.; Gussekloo, J. Thyroid Hormone Therapy for Older Adults with Subclinical Hypothyroidism. *New England Journal of Medicine* **2017**, *376* (26), 2534–2544.

44. Gambacciani, M.; Cagnacci, A.; Lello, S. Hormone Replacement Therapy and Prevention of Chronic Conditions. *Climacteric* **2019**, *22* (3), 303–306.

45. Lee, V. M.; Nelly, M. Growth Hormone Therapy in Adults and Children. *The New England Journal of Medicine* **1999**, *341* (16), 1206–1216.

46. Ewing, A. G.; Wallingford, R. A.; Olefirowicz, T. M. Capillary Electrophoresis. *Analytical Chemistry* **1989**, *61* (4), 292A-303A.

47. Rathore, A. S.; Horváth, Cs. Capillary Electrochromatography: Theories on Electroosmotic Flow in Porous Media. *Journal of Chromatography A* **1997**, *781* (1), 185–195.

48. Kozlowski, L. P. Proteome-PI: Proteome Isoelectric Point Database. *Nucleic Acids Research* **2017**, *45* (D1), D1112–D1116.

49. Chu, Y.-H.; Lees, W. J.; Stassinopoulos, A.; Walsh, C. T. Using Affinity Capillary Electrophoresis To Determine Binding Stoichiometries of Protein-Ligand Interactions. *Biochemistry* **1994**, *33* (35), 10616–10621.

50. Hancu, G.; Simon, B.; Rusu, A.; Mircia, E.; Gyéresi, Á. Principles of Micellar Electrokinetic Capillary Chromatography Applied in Pharmaceutical Analysis. *Advanced Pharmaceutical Bulletin* **2013**, *3* (1), 1–8.

51. Danish, A.; Lee, S.-Y.; Müller, C. E. Quantification of Green Fluorescent Protein-(GFP-) Tagged Membrane Proteins by Capillary Gel Electrophoresis. *Analyst* **2017**, *142* (19), 3648–3655.

52. Yeung, W. S. B.; Luo, G. A.; Wang, Q. G.; Ou, J. P. Capillary Electrophoresis-Based Immunoassay. *Journal of Chromatography B* **2003**, *797* (1), 217–228.

53. Ciancio, D.; Vargas, M. R.; Thiel, W. H.; Bruno, M. A.; Giangrande, P. H.; Mestre, M. B. Aptamers as Diagnostic Tools in Cancer. *Pharmaceuticals* **2018**, *11* (3), 86.

54. Kruspe, S.; Mittelberger, F.; Szameit, K.; Hahn, U. Aptamers as Drug Delivery Vehicles. *ChemMedChem* **2014**, *9* (9), 1998–2011.

55. Proske, D.; Blank, M.; Buhmann, R.; Resch, A. Aptamers—Basic Research, Drug Development, and Clinical Applications. *Applied Microbiology and Biotechnology* **2005**, *69* (4), 367–374.

56. Gold, L.; Ayers, D.; Bertino, J.; Bock, C.; Bock, A.; Brody, E.; Carter, J.; Cunningham, V.; Dalby, A.; Eaton, B.; Fitzwater, T.; Flather, D.; Forbes, A.; Foreman, T.; Fowler, C.; Gawande, B.; Goss, M.; Gunn, M.; Gupta, S.; Halladay, D.; Heil, J.; Heilig, J.; Hicke, B.; Husar, G.; Janjic, N.; Jarvis, T.; Jennings, S.; Katilius, E.; Keeney, T.; Kim, N.; Kaske, T.; Koch, T.; Kraemer, S.; Kroiss, L.; Le, N.; Levine, D.; Lindsey, W.; Lollo, B.; Mayfield, W.; Mehan, M.; Mehler, R.; Nelson, M.; Nelson, S.; Nieuwlandt, D.; Nikrad, M.; Ochsner, U.; Ostroff, R.; Otis, M.; Parker, T.; Pietrasiewicz, S.; Resnicow, D.; Rohloff, J.; Sanders, G.; Sattin, S.; Schneider, D.; Singer, B.; Stanton, M.; Sterkel, A.; Stewart, A.; Stratford, S.; Vaught, J.; Vrkljan, M.; Walker, J.; Watrobka, M.; Waugh, S.; Weiss, A.; Wilcox, S.; Wolfson, A.; Wolk, S.; Zhang, C.; Zichi, D. Aptamer-Based Multiplexed Proteomic Technology for Biomarker Discovery. *Nature Precedings* **2010**, 1–1.

57. Vant-Hull, B.; Payano-Baez, A.; Davis, R. H.; Gold, L. The Mathematics of SELEX against Complex Targets11Edited by F. Cohen. *Journal of Molecular Biology* **1998**, *278* (3), 579–597.

58. Gopinath, S. C. B. Methods Developed for SELEX. *Analytical and Bioanalytical Chemistry* **2006**, *387* (1), 171–182.

59. Berezovski, M.; Drabovich, A.; Krylova, S. M.; Musheev, M.; Okhonin, V.; Petrov, A.; Krylov, S. N. Nonequilibrium Capillary Electrophoresis of Equilibrium Mixtures: A Universal Tool for Development of Aptamers. *Journal of the American Chemical Society* **2005**, *127* (9), 3165–3171.

60. Le, A. T. H.; Krylova, S. M.; Kanoatov, M.; Desai, S.; Krylov, S. N. Ideal-Filter Capillary Electrophoresis (IFCE) Facilitates the One-Step Selection of Aptamers. *Angewandte Chemie International Edition* **2019**, *58* (9), 2739–2743.
61. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T. SWISS-MODEL: Homology Modelling of Protein Structures and Complexes. *Nucleic Acids Research* **2018**, *46* (W1), W296–W303.
62. Nkodo, A. E.; Garnier, J. M.; Tinland, B.; Ren, H.; Desruisseaux, C.; McCormick, L. C.; Drouin, G.; Slater, G. W. Diffusion Coefficient of DNA Molecules during Free Solution Electrophoresis. *ELECTROPHORESIS* **2001**, *22* (12), 2424–2432.
63. Beloborodov, S. S.; Bao, J.; Krylova, S. M.; Shala-Lawrence, A.; Johnson, P. E.; Krylov, S. N. Aptamer Facilitated Purification of Functional Proteins. *Journal of Chromatography B* **2018**, *1073*, 201–206.
64. Erickson, H. P. Size and Shape of Protein Molecules at the Nanometer Level Determined by Sedimentation, Gel Filtration, and Electron Microscopy. *Biological Procedures Online* **2009**, *11* (1), 32–51.
65. Rickard, E. C.; Strohl, M. M.; Nielsen, R. G. Correlation of Electrophoretic Mobilities from Capillary Electrophoresis with Physicochemical Properties of Proteins and Peptides. *Analytical Biochemistry* **1991**, *197* (1), 197–207.
66. Wan, Q.-H.; Le, X. C. Studies of Protein−DNA Interactions by Capillary Electrophoresis/Laser-Induced Fluorescence Polarization. *Analytical Chemistry* **2000**, *72* (22), 5583–5589.
67. Brocchieri, L.; Karlin, S. Protein Length in Eukaryotic and Prokaryotic Proteomes. *Nucleic Acids Research* **2005**, *33* (10), 3390–3400.
68. Wojcik, R.; Swearingen, K. E.; Dickerson, J. A.; Turner, E. H.; Ramsay, L. M.; Dovichi, N. J. Reaction of Fluorogenic Reagents with Proteins: I. Mass Spectrometric Characterization of the Reaction with 3-(2-Furoyl)Quinoline-2-Carboxaldehyde, Chromeo P465, and Chromeo P503. *Journal of Chromatography A* **2008**, *1194* (2), 243–248.
69. Swearingen, K. E.; Dickerson, J. A.; Turner, E. H.; Ramsay, L. M.; Wojcik, R.; Dovichi, N. J. Reaction of Fluorogenic Reagents with Proteins: II: Capillary Electrophoresis and Laser-Induced Fluorescence Properties of Proteins Labeled with Chromeo P465. *Journal of Chromatography A* **2008**, *1194* (2), 249–252.
70. Basak, S. K.; Ladisch, M. R. Correlation of Electrophoretic Mobilities of Proteins and Peptides with Their Physicochemical Properties. *Analytical Biochemistry* **1995**, *226* (1), 51–58.
71. Bai, Y.; Zhao, Q. Investigation of the Effects of Metal Ions in Sample Buffer on Capillary Electrophoresis Coupled with Laser-Induced Fluorescence Analysis of Thrombin Using a Dye-Labeled 29-Mer DNA Aptamer. *Analytical Methods* **2017**, *9* (38), 5684–5690.
72. Rosenblatt, R.; Halámková, L.; Doty, K. C.; de Oliveira, E. A. C.; Lednev, I. K. Raman Spectroscopy for Forensic Bloodstain Identification: Method Validation vs. Environmental Interferences. *Forensic Chemistry* **2019**, *16*, 100175.
73. Liu, S.; Thomas, S. M.; Woodside, D. G.; Rose, D. M.; Kiosses, W. B.; Pfaff, M.; Ginsberg, M. H. Binding of Paxillin to α 4 Integrins Modifies Integrin-Dependent Biological Responses. *Nature* **1999**, *402* (6762), 676–681.

74. Butcher, E. C. Leukocyte-Endothelial Cell Recognition: Three (or More) Steps to Specificity and Diversity. *Cell* **1991**, *67* (6), 1033–1036.

75. Aderem, A.; Ulevitch, R. J. Toll-like Receptors in the Induction of the Innate Immune Response. *Nature* **2000**, *406* (6797), 782–787.

76. Sauter, N. K.; Hanson, J. E.; Glick, G. D.; Brown, J. H.; Crowther, R. L.; Park, S. J.; Skehel, J. J.; Wiley, D. C. Binding of Influenza Virus Hemagglutinin to Analogs of Its Cell-Surface Receptor, Sialic Acid: Analysis by Proton Nuclear Magnetic Resonance Spectroscopy and x-Ray Crystallography. *Biochemistry* **1992**, *31* (40), 9609–9621.

77. Liu, S.; Zhang, X.; Wu, Y.; Tu, Y.; He, L. Prostate-Specific Antigen Detection by Using a Reusable Amperometric Immunosensor Based on Reversible Binding and Leasing of HRP-Anti-PSA from Phenylboronic Acid Modified Electrode. *Clinica Chimica Acta; International Journal of Clinical Chemistry* **2008**, *395* (1–2), 51–56.

78. Wegman, D. W.; Krylov, S. N. Direct Quantitative Analysis of Multiple MiRNAs (DQAMmiR). *Angewandte Chemie (International Ed. in English)* **2011**, *50* (44), 10335–10339.

79. Shieh, Y.-A.; Yang, S.-J.; Wei, M.-F.; Shieh, M.-J. Aptamer-Based Tumor-Targeted Drug Delivery for Photodynamic Therapy. *ACS Nano* **2010**, *4* (3), 1433–1442.

80. Flego, M.; Ascione, A.; Cianfriglia, M.; Vella, S. Clinical Development of Monoclonal Antibody-Based Drugs in HIV and HCV Diseases. *BMC Medicine* **2013**, *11* (1), 4.

81. Saylor, C.; Dadachova, E.; Casadevall, A. Monoclonal Antibody-Based Therapies for Microbial Diseases. *Vaccine* **2009**, *27*, G38–G46.

82. Urusov, A. E.; Kostenko, S. N.; Sveshnikov, P. G.; Zherdev, A. V.; Dzantiev, B. B. Ochratoxin A Immunoassay with Surface Plasmon Resonance Registration: Lowering Limit of Detection by the Use of Colloidal Gold Immunoconjugates. *Sensors and Actuators B: Chemical* **2011**, *156* (1), 343–349.

83. Hong, M.; Lee, P. S.; Hoffman, R. M. B.; Zhu, X.; Krause, J. C.; Laursen, N. S.; Yoon, S.; Song, L.; Tussey, L.; Crowe, J. E.; Ward, A. B.; Wilson, I. A. Antibody Recognition of the Pandemic H1N1 Influenza Virus Hemagglutinin Receptor Binding Site. *Journal of Virology* **2013**, *87* (22), 12471–12480.

84. Weber, P. C.; Salemme, F. R. Applications of Calorimetric Methods to Drug Discovery and the Study of Protein Interactions. *Current Opinion in Structural Biology* **2003**, *13* (1), 115–121.

85. Kanoatov, M.; Galievsky, V. A.; Krylova, S. M.; Cherney, L. T.; Jankowski, H. K.; Krylov, S. N. Using Nonequilibrium Capillary Electrophoresis of Equilibrium Mixtures (NECEEM) for Simultaneous Determination of Concentration and Equilibrium Constant. *Analytical Chemistry* **2015**, *87* (5), 3099–3106.

86. Bao, J.; Krylova, S. M.; Cherney, L. T.; Hale, R. L.; Belyanskaya, S. L.; Chiu, C. H.; Arico-Muendel, C. C.; Krylov, S. N. Prediction of Protein–DNA Complex Mobility in Gel-Free Capillary Electrophoresis. *Analytical Chemistry* **2015**, *87* (4), 2474–2479.

87. de Jong, S.; Krylov, S. N. Protein Labeling Enhances Aptamer Selection by Methods of Kinetic Capillary Electrophoresis. *Analytical Chemistry* **2011**, *83* (16), 6330–6335.

88. Krylova, S. M.; Karkhanina, A. A.; Musheev, M. U.; Bagg, E. A. L.; Schofield, C. J.; Krylov, S. N. DNA Aptamers for as Analytical Tools for the Quantitative Analysis of DNA-Dealkylating Enzymes. *Analytical Biochemistry* **2011**, *414* (2), 261–265.

89. Kiley, M. P.; Regnery, R. L.; Johnson, K. M. Ebola Virus: Identification of Virion Structural Proteins. *Journal of General Virology*, **1980**, *49* (2), 333–341.

90. Geisbert, T. W.; Jahrling, P. B. Differentiation of Filoviruses by Electron Microscopy. *Virus Research* **1995**, *39* (2), 129–150.

91. Atabekov, J.; Dobrov, E.; Karpova, O.; Rodionova, N. Potato Virus X: Structure, Disassembly and Reconstitution. *Molecular Plant Pathology* **2007**, *8* (5), 667–675.

92. Wen, A. M.; Steinmetz, N. F. Design of Virus-Based Nanomaterials for Medicine, Biotechnology, and Energy. *Chemical Society Reviews* **2016**, *45* (15), 4074–4126.

93. Uhde-Holzem, K.; McBurney, M.; Tiu, B. D. B.; Advincula, R. C.; Fischer, R.; Commandeur, U.; Steinmetz, N. F. Production of Immunoabsorbent Nanoparticles by Displaying Single-Domain Protein A on Potato Virus X. *Macromolecular Bioscience* **2016**, *16* (2), 231–241.

94. Lico, C.; Benvenuto, E.; Baschieri, S. The Two-Faced Potato Virus X: From Plant Pathogen to Smart Nanoparticle. *Frontiers in Plant Science* **2015**, *6*.

95. Safenkova, I. V.; Zherdev, A. V.; Dzantiev, B. B. Correlation between the Composition of Multivalent Antibody Conjugates with Colloidal Gold Nanoparticles and Their Affinity. *Journal of Immunological Methods* **2010**, *357* (1), 17–25.

96. Nečas, D.; Klapetek, P. Gwyddion: An Open-Source Software for SPM Data Analysis. *Open Physics* **2012**, *10* (1), 181–188.

97. Doorn, S. K.; Fields, R. E.; Hu, H.; Hamon, M. A.; Haddon, R. C.; Selegue, J. P.; Majidi, V. High Resolution Capillary Electrophoresis of Carbon Nanotubes. *Journal of the American Chemical Society* **2002**, *124* (12), 3169–3174.

98. Grossman, P. D.; Soane, D. S. Orientation Effects on the Electrophoretic Mobility of Rod-Shaped Molecules in Free Solution. *Analytical Chemistry* **1990**, *62* (15), 1592–1596.

99. Karikari, T. K.; Turner, A.; Stass, R.; Lee, L. C. Y.; Wilson, B.; Nagel, D. A.; Hill, E. J.; Moffat, K. G. Expression and Purification of Tau Protein and Its Frontotemporal Dementia Variants Using a Cleavable Histidine Tag. *Protein Expression and Purification* **2017**, *130*, 44–54.

100. Jozala, A. F.; Geraldes, D. C.; Tundisi, L. L.; Feitosa, V. de A.; Breyer, C. A.; Cardoso, S. L.; Mazzola, P. G.; Oliveira-Nascimento, L. de; Rangel-Yagui, C. de O.; Magalhães, P. de O.; Oliveira, M. A. de; Pessoa Jr, A.; Jozala, A. F.; Geraldes, D. C.; Tundisi, L. L.; Feitosa, V. de A.; Breyer, C. A.; Cardoso, S. L.; Mazzola, P. G.; Oliveira-Nascimento, L. de; Rangel-Yagui, C. de O.; Magalhães, P. de O.; Oliveira, M. A. de; Pessoa Jr, A. Biopharmaceuticals from Microorganisms: From Production to Purification. *Brazilian Journal of Microbiology* **2016**, *47*, 51–63.

101. Hunt, I. From Gene to Protein: A Review of New and Enabling Technologies for Multi-Parallel Protein Expression. *Protein Expression and Purification* **2005**, *40* (1), 1–22.

102. Knudsen, H. L.; Fahrner, R. L.; Xu, Y.; Norling, L. A.; Blank, G. S. Membrane Ion-Exchange Chromatography for Process-Scale Antibody Purification. *Journal of Chromatography A* **2001**, *907* (1), 145–154.

103. Goyon, A.; Beck, A.; Colas, O.; Sandra, K.; Guillarme, D.; Fekete, S. Evaluation of Size Exclusion Chromatography Columns Packed with Sub-3μm Particles for the Analysis of Biopharmaceutical Proteins. *Journal of Chromatography A* **2017**, *1498*, 80–89.
104. Karakus, C.; Uslu, M.; Yazici, D.; Salih, B. A. Evaluation of Immobilized Metal Affinity Chromatography Kits for the Purification of Histidine-Tagged Recombinant CagA Protein. *Journal of Chromatography B* **2016**, *1021*, 182–187.
105. Scheich, C.; Sievert, V.; Büssow, K. An Automated Method for High-Throughput Protein Purification Applied to a Comparison of His-Tag and GST-Tag Affinity Chromatography. *BMC Biotechnology* **2003**, *3* (1), 12.
106. Wu, J.; Filutowicz, M. Hexahistidine (His6)-Tag Dependent Protein Dimerization: A Cautionary Tale. *Acta Biochimica Polonica* **1999**, *46* (3), 591–599.
107. Arnau, J.; Lauritzen, C.; Petersen, G. E.; Pedersen, J. Current Strategies for the Use of Affinity Tags and Tag Removal for the Purification of Recombinant Proteins. *Protein Expression and Purification* **2006**, *48* (1), 1–13.
108. Berezovski, M. V.; Musheev, M. U.; Drabovich, A. P.; Jitkova, J. V.; Krylov, S. N. Non-SELEX: Selection of Aptamers without Intermediate Amplification of Candidate Oligonucleotides. *Nature Protocols* **2006**, *1* (3), 1359–1369.
109. Balamurugan, S.; Obubuafo, A.; Soper, S. A.; Spivak, D. A. Surface Immobilization Methods for Aptamer Diagnostic Applications. *Analytical and Bioanalytical Chemistry* **2008**, *390* (4), 1009–1021.
110. Romig, T. S.; Bell, C.; Drolet, D. W. Aptamer Affinity Chromatography:: Combinatorial Chemistry Applied to Protein Purification. *Journal of Chromatography B: Biomedical Sciences and Applications* **1999**, *731* (2), 275–284.
111. Forier, C.; Boschetti, E.; Ouhammouch, M.; Cibiel, A.; Ducongé, F.; Nogré, M.; Tellier, M.; Bataille, D.; Bihoreau, N.; Santambien, P.; Chtourou, S.; Perret, G. DNA Aptamer Affinity Ligands for Highly Selective Purification of Human Plasma-Related Proteins from Multiple Sources. *Journal of Chromatography A* **2017**, *1489*, 39–50.
112. Qiao, L.; Lv, B.; Feng, X.; Li, C. A New Application of Aptamer: One-Step Purification and Immobilization of Enzyme from Cell Lysates for Biocatalysis. *Journal of Biotechnology* **2015**, *203*, 68–76.
113. Javaherian, S.; Musheev, M. U.; Kanoatov, M.; Berezovski, M. V.; Krylov, S. N. Selection of Aptamers for a Protein Target in Cell Lysate and Their Application to Protein Purification. *Nucleic Acids Research* **2009**, *37* (8), e62–e62.
114. Ahirwar, R.; Nahar, P. Development of an Aptamer-Affinity Chromatography for Efficient Single Step Purification of Concanavalin A from Canavalia Ensiformis. *Journal of Chromatography B* **2015**, *997*, 105–109.
115. Bartnicki, F.; Kowalska, E.; Pels, K.; Strzalka, W. Imidazole-Free Purification of His3-Tagged Recombinant Proteins Using SsDNA Aptamer-Based Affinity Chromatography. *Journal of Chromatography A* **2015**, *1418*, 130–139.
116. Gädke, J.; Kleinfeldt, L.; Schubert, C.; Rohde, M.; Biedendieck, R.; Garnweitner, G.; Krull, R. In Situ Affinity Purification of His-Tagged Protein A from Bacillus Megaterium Cultivation Using Recyclable Superparamagnetic Iron Oxide Nanoparticles. *Journal of Biotechnology* **2017**, *242*, 55–63.

117. Song, A.; Zhao, C.; Zhang, Y.; Li, J.; Blackburn, M. R.; Grenz, A.; Eltzschig, H.; Kellems, R. E.; Xia, Y. Functions and Regulation of Erythrocyte Equlibrative Nucleoside Transporter 1 (ENT1) in Acute Hypoxia Mediated Tissue Injury. *Blood* **2014**, *124* (21), 2666–2666.
118. Karkhanina, A. A.; Mecinović, J.; Musheev, M. U.; Krylova, S. M.; Petrov, A. P.; Hewitson, K. S.; Flashman, E.; Schofield, C. J.; Krylov, S. N. Direct Analysis of Enzyme-Catalyzed DNA Demethylation. *Analytical Chemistry* **2009**, *81* (14), 5871–5875.
119. Drabovich, A. P.; Krylov, S. N. Identification of Base Pairs in Single-Nucleotide Polymorphisms by MutS Protein-Mediated Capillary Electrophoresis. *Analytical Chemistry* **2006**, *78* (6), 2035–2038.
120. Drabovich, A.; Berezovski, M.; Krylov, S. N. Selection of Smart Aptamers by Equilibrium Capillary Electrophoresis of Equilibrium Mixtures (ECEEM). *Journal of the American Chemical Society* **2005**, *127* (32), 11224–11225.
121. Welford, R. W. D.; Schlemminger, I.; McNeill, L. A.; Hewitson, K. S.; Schofield, C. J. The Selectivity and Inhibition of AlkB. *Journal of Biological Chemistry* **2003**, *278* (12), 10157–10161.
122. Groothuizen, F. S.; Fish, A.; Petoukhov, M. V.; Reumer, A.; Manelyte, L.; Winterwerp, H. H. K.; Marinus, M. G.; Lebbink, J. H. G.; Svergun, D. I.; Friedhoff, P.; Sixma, T. K. Using Stable MutS Dimers and Tetramers to Quantitatively Analyze DNA Mismatch Recognition and Sliding Clamp Formation. *Nucleic Acids Research* **2013**, *41* (17), 8166–8181.
123. Nowak, P. M.; Woźniakiewicz, M.; Gładysz, M.; Janus, M.; Kościelniak, P. Improving Repeatability of Capillary Electrophoresis—a Critical Comparison of Ten Different Capillary Inner Surfaces and Three Criteria of Peak Identification. *Analytical and Bioanalytical Chemistry* **2017**, *409* (18), 4383–4393.
124. Sursyakova, V. V.; Burmakina, G. V.; Rubaylo, A. I. Strategy for Non-Target Ionic Analysis by Capillary Electrophoresis with Ultraviolet Detection. *Analytical and Bioanalytical Chemistry* **2017**, *409* (4), 1067–1077.
125. Demorest, D.; Dubrow, R. Factors Influencing the Resolution and Quantitation of Oligonucleotides Separated by Capillary Electrophoresis on a Gel-Filled Capillary. *Journal of Chromatography A* **1991**, *559* (1), 43–56.
126. Petersen, N. J.; Nikolajsen, R. P. H.; Mogensen, K. B.; Kutter, J. P. Effect of Joule Heating on Efficiency and Performance for Microchip-Based and Capillary-Based Electrophoretic Separation Systems: A Closer Look. *Electrophoresis* **2004**, *25* (2), 253–269.
127. Rathore, A. S. Joule Heating and Determination of Temperature in Capillary Electrophoresis and Capillary Electrochromatography Columns. *Journal of Chromatography A* **2004**, *1037* (1), 431–443.
128. Nowak, P. M.; Śpiewak, K.; Woźniakiewicz, M.; Kościelniak, P. Minimizing the Impact of Joule Heating as a Prerequisite for the Reliable Analysis of Metal-protein Complexes by Capillary Electrophoresis. *Journal of Chromatography A* **2017**, *1495*, 83–87.
129. Musheev, M. U.; Filiptsev, Y.; Krylov, S. N. Noncooled Capillary Inlet: A Source of Systematic Errors in Capillary-Electrophoresis-Based Affinity Analyses. *Analytical Chemistry* **2010**, *82* (20), 8637–8641.

130. Rasmussen, H. T.; McNair, H. M. Influence of Buffer Concentration, Capillary Internal Diameter and Forced Convection on Resolution in Capillary Zone Electrophoresis. *Journal of Chromatography A* **1990**, *516* (1), 223–231.
131. Galievsky, V. A.; Stasheuski, A. S.; Krylov, S. N. "Getting the Best Sensitivity from on-Capillary Fluorescence Detection in Capillary Electrophoresis" – A Tutorial. *Analytica Chimica Acta* **2016**, *935*, 58–81.
132. Zhu, Y.; Li, Z.; Wang, P.; Shen, L.; Zhang, D.; Yamaguchi, Y. Factors Affecting the Separation Performance of Proteins in Capillary Electrophoresis. *Journal of Chromatography. B, Analytical Technologies in the Biomedical and Life Sciences* **2018**, *1083*, 63–67.
133. Beloborodov, S. S.; Krylova, S. M.; Krylov, S. N. Spherical-Shape Assumption for Protein–Aptamer Complexes Facilitates Prediction of Their Electrophoretic Mobility. *Analytical Chemistry* **2019**, *91* (20), 12680–12687.
134. Musheev, M. U.; Kanoatov, M.; Krylov, S. N. Non-Uniform Velocity of Homogeneous DNA in a Uniform Electric Field: Consequence of Electric-Field-Induced Slow Dissociation of Highly Stable DNA–Counterion Complexes. *Journal of the American Chemical Society* **2013**, *135* (21), 8041–8046.

APPENDICES

APPENDIX A

Table A1. Properties of the proteins.

The table below shows the charges of proteins at the corresponding pH (pH 8.2 for MutS, streptavidin and AlkB; pH 8.0 for SSB), the number of amino (N_{aa}) acids and their Isoelectric points (pI). Isoelectric points and charges of the proteins were calculated with ProteinCalculator v3.4 software (http://protcalc.sourceforge.net/) using amino acid sequence of each protein as an imputing parameter. The presented isoelectric points and charges correspond to monomeric units for all proteins.

Protein	Isoelectric point (pI)	Charge at pH 8.2 (MutS, streptavidin, AlkB) and at pH 8.0 (SSB)	Number of Amino acids (N_{aa})
MutS	5.61	-19.6	811
Streptavidin	8.36	+0.2	183
AlkB	7.87	-1	216
SSB	5.59	-2.6	178

Table A2. Electrophoretic mobilities and molecular weights of the species.

Molecular weights of MutS and AlkB were found in the literature **[63]**; molecular weights of SSB, streptavidin, and all ssDNA molecules were provided by the vendor (see Materials and Methods section in the main text). The presented molecular weights of the proteins correspond to the monomeric molecule (AlkB), dimeric molecule (MutS), tetrameric molecule (SSB), and homotetrameric molecule (streptavidin).

System	Protein		ssDNA			Complex
	MW_P, kDa	$-\mu_P$, $mm^2kV^{-1}s^{-1}$	MW_{ssDNA}, kDa	N, nt	$-\mu_{ssDNA}$, $mm^2kV^{-1}s^{-1}$	$-\mu_{P\text{-}ssDNA}$, $mm^2kV^{-1}s^{-1}$
MutS–ssDNA	95.0	14.53 ± 0.06	25.3	80	31.03 ± 0.20	20.91 ± 0.17
Streptavidin–ssDNA	52.8	4.23 ± 0.10	25.4	80	30.72 ± 0.10	23.34 ± 0.04
AlkB–ssDNA	23.9	13.64 ± 0.07	25.1	80	29.62 ± 0.18	24.59 ± 0.16
SSB–ssDNA	75.5	7.65 ± 0.14	27.5	88	31.52 ± 0.05	22.98 ± 0.09
			25.3	80	32.31 ± 0.30	22.99 ± 0.15
			24.0	75	31.94 ± 0.07	22.20 ± 0.11
			8.4	25	32.44 ± 0.10	16.14 ± 0.11

Table A3. Calculated variables *X* and *Y* for protein–ssDNA complexes.

The table below demonstrates variables *X* and *Y* that were calculated for seven protein–ssDNA complexes using **(2-14)** (main text). The column *X*/*Y* shows the ratio between *X* and *Y* for each particular case and the column %*Y* of *X* shows $Y/X \cdot 100$.

Protein–ssDNA	**–*X***	**–*Y***	***X*/*Y***	**%*Y* of *X***
MutS–apt	123.72	13.43	9.21	10.85
Streptavidin–apt	141.42	3.71	38.13	2.62
AlkB–apt	159.39	10.74	14.84	6.74
SSB–88-nt	141.22	6.89	20.48	4.88
SSB–80-nt	136.67	6.95	19.68	5.08
SSB–75-nt	129.88	6.98	18.62	5.37
SSB–25-nt	66.15	7.38	8.96	11.16

Table A4. Comparison of electrophoretic mobilities calculated for different μ_{ssDNA}.

The table below demonstrates mobilities calculated for each protein–ssDNA complex using **eq (2-19)** (main text) and different values of μ_{ssDNA}: the values obtained directly from the electropherograms for each particular case, the average value of −31.37 ± 1.00 obtained from all electropherograms, and the values of −30.50 and −30.18, calculated by the polynomial **eq (2-20)** (main text) for I = 0.021 M (experiments with MutS, streptavidin, and AlkB) and 0.027 M (experiments with SSB), respectively.

Protein–ssDNA	**Experimental values**	**Using μ_{ssDNA} from electropherograms**		**Using average μ_{ssDNA} of −31.37 ± 1.00**		**Using polynomial μ_{ssDNA} of −30.50, −30.18**	
	$-\mu_{P-ssDNA}$, $mm^2kV^{-1}s^{-1}$	$-\mu_{P-ssDNA}$, $mm^2kV^{-1}s^{-1}$	$\Delta\mu_{P-ssDNA}$, %	$-\mu_{P-ssDNA}$, $mm^2kV^{-1}s^{-1}$	$\Delta\mu_{P-ssDNA}$, %	$-\mu_{P-ssDNA}$, $mm^2kV^{-1}s^{-1}$	$\Delta\mu_{P-ssDNA}$, %
MutS–80-nt	20.91 ± 0.17	21.44	2.6	21.57	3.2	21.25	1.6
Streptavidin–80-nt	23.34 ± 0.04	23.09	1.1	23.37	0.1	23.00	1.5
AlkB–80-nt	24.59 ± 0.16	24.76	0.7	25.63	4.2	24.60	2.4
SSB–88-nt	22.98 ± 0.09	23.07	0.4	23.01	0.1	22.51	2.1
SSB–80-nt	22.99 ± 0.15	22.65	1.5	22.28	3.1	21.81	5.1
SSB–75-nt	22.20 ± 0.11	22.02	0.8	21.80	1.8	21.35	3.8
SSB–25-nt	16.14 ± 0.11	16.10	0.3	15.89	1.5	15.67	2.9

Table A5. Properties of proteins used for testing the predictive model.

The table below shows physical properties of the proteins and their aptamers that have been used to test the developed model. Molecular weights, isoelectric points, and charges of the proteins were calculated with ProteinCalculator v3.4 software (http://protcalc.sourceforge.net/) using an amino acid sequence of each protein as an input parameter. Molecular weights of aptamers for ABH2 and AlkB were provided by the supplier and the molecular weight of the human α-thrombin aptamer was calculated based on its number of nucleotides.

Protein	**Isoelectric Point (pI)**	**Charge**	**Running buffer pH**	**Number of Amino acids (N_{aa})**	**MW_p, kDa**	**MW_{ssDNA}, kDa**	**N_{ssDNA}, nt**
Human α-thrombin	8.08	+2.4	7.5	295	36.7	8.9	29
ABH2	9.65	+9.2	8.2	261	29.3	25.1	80
AlkB	7.87	−5.3	9.2	216	23.9	25.1	80

Table A6. Experimental and predicted electrophoretic mobilities for complexes of human α-thrombin, ABH2, and AlkB.

The table below demonstrates experimental and predicted electrophoretic mobilities for complexes of human α-thrombin, ABH2, and AlkB. Electrophoretic mobilities for these complexes we predicted using **(2-20)** (main text). Experimental values for the mobilities of the complexes of ABH2 and AlkB were obtained by us. The value for the electrophoretic mobility of the complex of human α-thrombin was calculated with **(2-17)** and **(**2-18) (main text), using literature values **[71]** for the migration times of the aptamer (2.5 min), complex (2.2 min) and EOF marker (TMR, 1.7 min), along with the values of *L* (34 cm) and *E* (500 V/cm).

Complex	**Experimental** $-\mu_{\text{P–ssDNA}}$**,** $\text{mm}^2\text{kV}^{-1}\text{s}^{-1}$	**Predicted** $-\mu_{\text{P–ssDNA}}$**,** $\text{mm}^2\text{kV}^{-1}\text{s}^{-1}$	$\Delta\mu_{\text{P–ssDNA}}$**, %**
Human α-thrombin–29-nt aptamer, pH 7.5	15.15	15.43	1.8
ABH2–80-nt aptamer, pH 8.2	25.96 ± 0.21	24.76	4.6
AlkB–80-nt aptamer, pH 9.2	25.63 ± 0.08	25.59	0.1

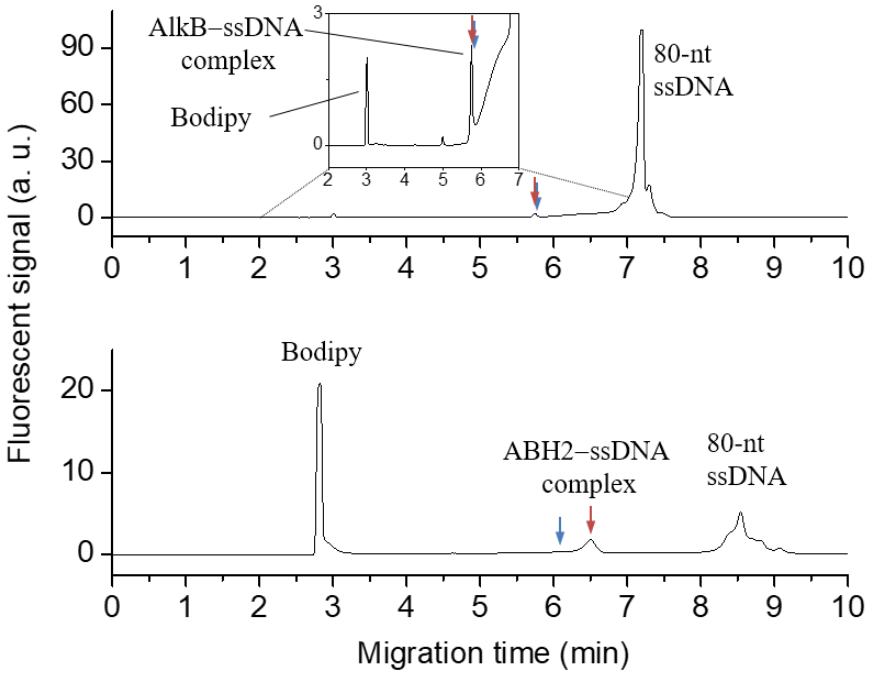

Figure A.1. Electropherograms used for the determination of electrophoretic mobilities for complexes of AlkB with its 80-nt aptamer at pH 9.2 and ABH2 with its 80-nt aptamer. The upper trace shows the peaks of AlkB–80-nt aptamer complex and free aptamer. The bottom trace shows the peaks of ABH2–80-nt aptamer complex and free aptamer. The peak of bodipy (internal standard) is present in both electropherograms. Red arrows indicate the actual migration times of the complexes, while blue arrows indicate their predicted migration times. The predicted migration times were calculated using the predicted values for electrophoretic mobilities determined with **(2-20)** (main text). Experimental conditions can be found in Materials and Methods (main text).

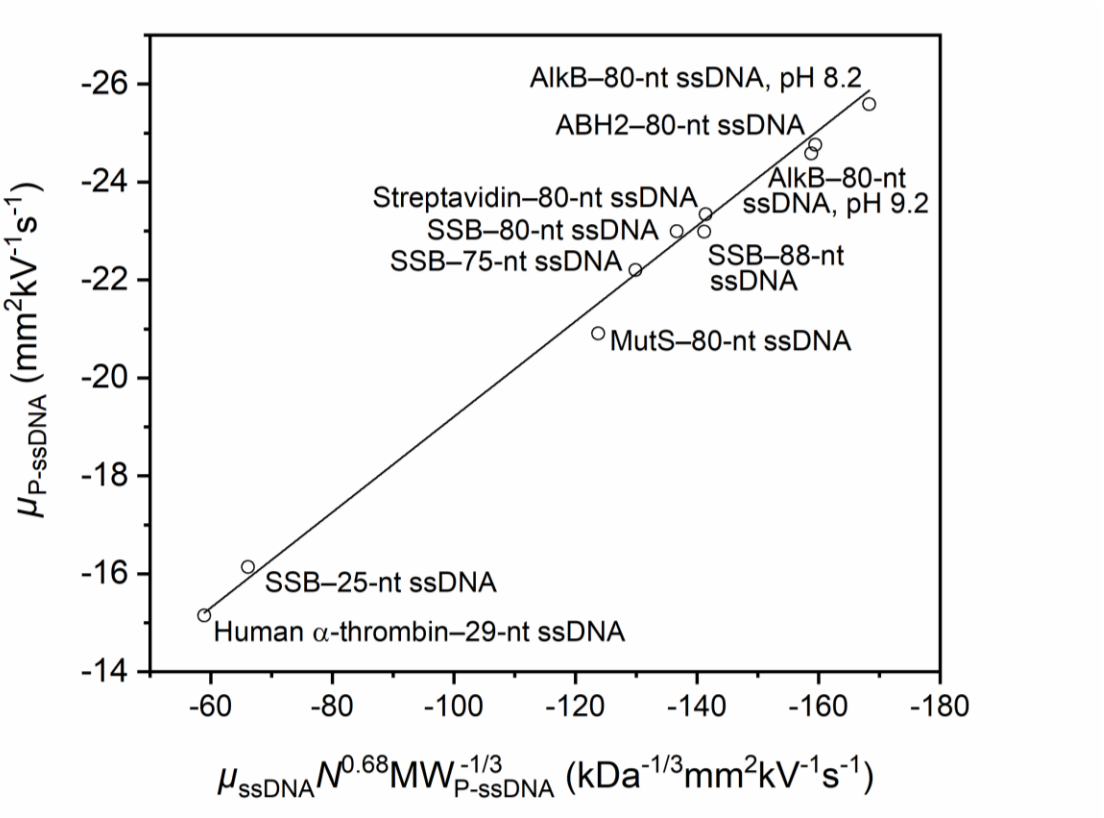

Figure A.2. Line of the best fit for electrophoretic mobilities of ten protein–ssDNA complexes as a function of X: $\mu_{P-ssDNA} = A + BX$, where $X = \{\mu_{P-ssDNA} N^{0.68} MW^{-1/3}_{P-ssDNA}\}$. The line of the best fit includes 10 data points for the training/ cross-validation set of protein–ssDNA complexes and 3 data points for complexes of aptamers with human α-thrombin, ABH2, and AlkB. The calculated values for A and B were -9.47 $mm^2kV^{-1}s^{-1}$ and 0.0973 $kDa^{-1/3}$, respectively. The correlation coefficient was $R^2 = 0.987$. The new constants A and B for ten points were different from those for seven points only by 4.8% and 4.7%, respectively.

APPENDIX B

1. Derivation of inequality for electrophoretic mobility of the complex

To obtain inequality **(3-5)** in main text, we rewrite inequality $\mu_A < \mu_B$ in a following manner:

$$\frac{q_A}{f_A} < \frac{q_B}{f_B} \quad \text{(S1)}$$

$$q_A f_B < q_B f_A \quad \text{(S2)}$$

By adding $q_A f_A$ to both sides in (S2) we get:

$$q_A f_B + q_A f_A < q_B f_A + q_A f_A \quad \text{(S3)}$$

$$q_A (f_B + f_A) < f_A (q_B + q_A) \quad \text{(S4)}$$

$$\frac{q_A}{f_A} < \frac{(q_B + q_A)}{(f_B + f_A)} \quad \text{(S5)}$$

By adding $q_B f_B$ to both sides in (S2) we get:

$$q_A f_B + q_B f_B < q_B f_A + q_B f_B \quad \text{(S6)}$$

$$f_B (q_A + q_B) < q_B (f_A + f_B) \quad \text{(S7)}$$

$$\frac{(q_B + q_A)}{(f_B + f_A)} < \frac{q_B}{f_B} \quad \text{(S8)}$$

$$\frac{q_A}{f_A} < \frac{(q_B + q_A)}{(f_B + f_A)} < \frac{q_B}{f_B} \quad \text{(S8)}$$

or

$$\mu_A < \mu_C < \mu_B \quad \text{(S9)}$$

2. CE experiments with Chromeo P503-labeled antibodies

To confirm the formation of multiple complexes and their atypical migration we performed a number of CE experiments with Chromeo P503 labeled antibodies. It was found that the labeled antibodies, being analyzed separately, formed two peaks on an electropherogram (**Figure B.1**, lower traces). Upon adding the virions, the higher peak did not change, but the smaller one decreased its height (the decrease was more pronounced at a 100:1 Ab:PVX ratio (**Figure B.1-B**) than at a 1000:1 ratio (**Figure B.1-A**). Accordingly, the higher peak was assigned to the fluorescently labeled products of the Ab deamidation and the smaller one to the fluorescently labeled antibodies by itself.

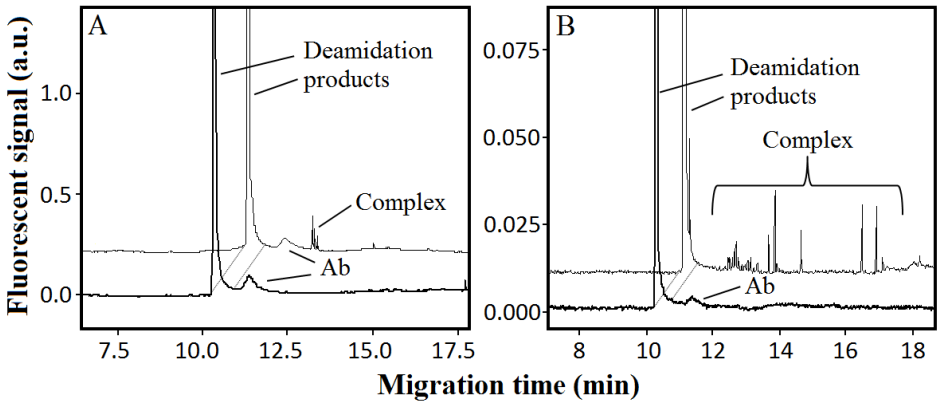

Figure B.1. CE-LIF analysis of mixtures of Chromeo P503 labeled antibodies (Ab) and unlabeled virus at different antibody:virions ratios: 1000:1 (A) and 100:1 (B). The bottom traces correspond to Ab-only controls (at 1 μM and 100 nM concentrations, respectively). For clarity of presentation, a time offset of 1 min and signal offsets of 0.25 (A) and 0.01 (B) units were applied. LIF detection was at 605 nm.

3. Investigation of the mixtures containing virions and monovalent antigen-binding fragments

As a negative control in our CE experiments we used monovalent fragments (Fab) of original bivalent monoclonal antibodies (Ab). Fab is unable to cross-bind viral particles. We found that the complex of virions with Fab migrated as a uniform zone and resulted in a single peak on the electropherogram. The mobility of this complex was equal to that of the virion within the limits of experimental error (**Figure B.2**).

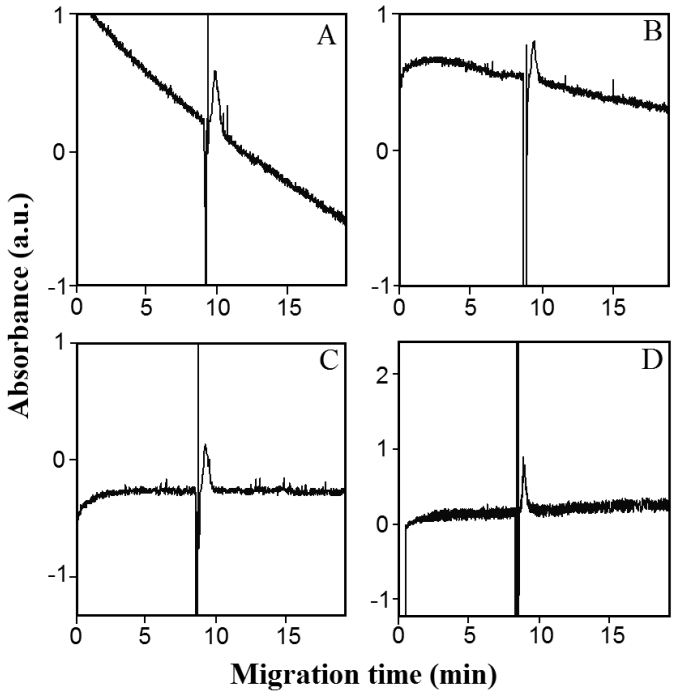

Figure B.2. Detection of 63 nM PVX in the presence of monovalent antibodies (Fab) at 260 nm. (A) represents PVX only; (B) represents Fab:PVX in a ratio of 2:1; (C) represents Fab:PVX in a ratio of 10:1; (D) represents Fab:PVX in a ratio of 135:1. Absorbance at 260 nm.

To confirm the antigen-binding ability of Ab we performed ELISA. Digested Fab was used as a negative control.Microplate wells were filled with 100 μL solution of PVX in PBS (1 μg/ml) and incubated at 37 °C for 2 h to complete the immobilization. The microplate was

washed four times with a mixture of PBS and 0.05% Triton X-100 (PBST). Series of dilutions of Ab (2.00, 1.00, 0.500, 0.250, 0.125, 0.063, 0.031, 0.017 μg/mL) and Fab (5.00, 2.50, 1.25, 0.625, 0.313, 0.156, 0.078, 0.039 μg/mL) were added to immobilized virions and incubated at 37 °C for 1 hour followed by washing procedure that described above. Then portions of secondary antibodies conjugated with peroxidase were added to the wells and incubated at 37°C for 1 h followed by washing. To determine the peroxidase activity, the substrate solution (100 mL of 0.4 mM 3,3′,5,5′-tetramethyl-benzidine solution in 40 mM sodium citrate buffer, pH 4.0, containing 3 mM H_2O_2) was added to each microplate well and incubated for 15 min at room temperature. Then reaction was terminated by adding 1 M H_2SO_4 (50 mL) and optical density measurements were conducted using EnSpire Multimode Plate Reader at 450 nm wavelength (room temperature). Each sample was studied in triplicates. The statistical treatment was done by Origin Pro 9.0 (Origin Lab, Northampton, MA, USA). Results showed that Ab are able to bind virus particles (**Figure B.3**).

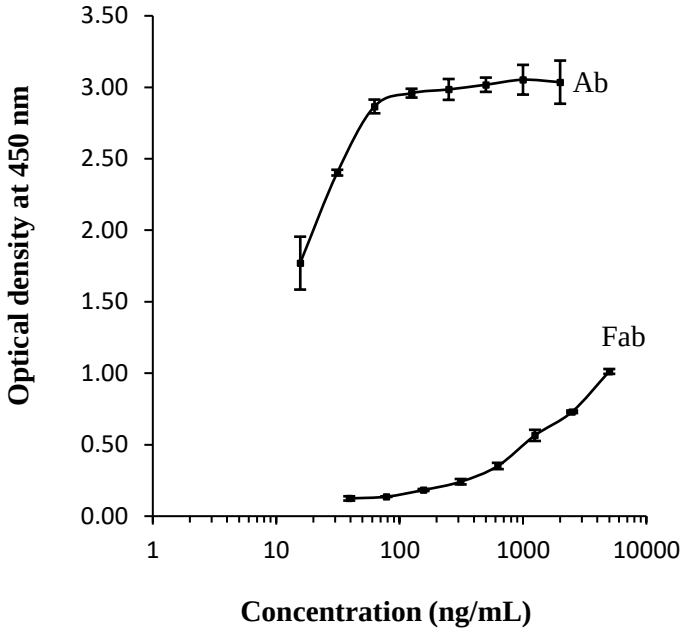

Figure B.3. Optical density of Ab and Fab solutions as a function of their concentrations (in logarithmic scale). Ab clearly demonstrates the ability to bind virions. Fab also demonstrates this ability, but to a lesser extent. Ideally, Fab should not demonstrate any binding activity, since fragmentation (hydrolysis) of a parent antibody (Ab) to Fab leads to the loss of Fc region that contains antigenic determinants responsible for the binding with secondary antibody. Thus, the residual activity of Fab can be explained by the incomplete hydrolysis of parent Ab.