TOWARD ANTI-CANCER DRUGS:

SELECTION OF ANTI-VEGF AND ANTI-HER2 APTAMER TO INHIBIT CANCER

BIOMARKERS


A THESIS SUBMITTED TO

THE FACULTY OF ENGINEERING AND NATURAL SCIENCES

OF

SABANCI UNIVERSITY


BY


IRENA ROCI


IN PARTIA FULFILMENT OF THE REQUIREMENTS FOR

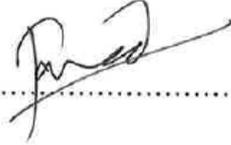THE DEGREE OF MASTER OF SCIENCES


IN


BIOLOGICAL SCIENCES AND BIOENGINEERING


AUGUST 2012

# TOWARD ANTI-CANCER DRUGS:
## SELECTION OF ANTI-VEGF AND ANTI-HER2 APTAMER TO INHIBIT CANCER BIOMARKERS

APPROVED BY

Asst. Prof.  Javed H. Niazi
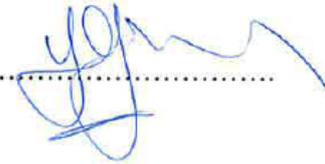(Thesis Supervisor)

Asst. Prof.  Alpay Taralp

Asst. Prof.  Deniz Sezer

Assoc. Prof.  Uğur Sezerman

Prof. Dr.  Yaşar Gürbüz

DATE OF APPROVAL     ..06.08.2012....................

# TOWARD ANTI-CANCER DRUGS:
## SELECTION OF ANTI-VEGF AND ANTI-HER2 APTAMER TO INHIBIT CANCER BIOMARKERS

IRENA ROCI

Biological Sciences and Bioengineering, MS Thesis, 2012

Thesis Supervisor: Dr. Javed H. Niazi

Keywords: *cancer, HER2, VEGF, aptamers, SELEX*

## ABSTRACT

There are ~71 cancer types and most of them exhibit heterogenous phenotypic characteristics. For this reason, studying common physiological alterations in cancer cells can be a successful approach for cancer treatment. Sustaining proliferative signaling and inducing angiogenesis are two of the many acquired characteristics during cancer progression. Both HER2 and VEGF are overexpressed in each of the above conditions, respectively. Therefore, the aim of this study is to develop synthetic ssDNA molecules (aptamers) that can bind to HER2 and VEGF, and inhibit their function, respectively. Even though there are already aptamers for these two targets, we wanted to select our aptamers for which there exist the possibility to bind with higher affinity. These aptamers were developed by using SELEX (Systematic Evolution of Ligands by EXponential Enrichment) technology in which HER2/VEGF immobilized magnetic beads were employed for the selection of specific aptamers. The enriched ssDNA pool was cloned, sequenced and characterized.

The chosen anti-VEGF contained a G-quartet, while anti-HER2 contained a two stem-loops in their structure. We speculate that these features may play a role in specific binding to the target protein. Equilibrium binding assays with anti-VEGF aptamer

showed a dissociation constant ($K_d$) of 315 nM and anti-HER2 showed a $K_d$ of 309 nM. Despite the similar binding, anti-HER2 was less selective toward its target protein. In this study we incorporated the counter selection step toward human serum to exclude the serum proteins binders. The anti-VEGF can potentially be used to detect cancer in blood by using biosensing technologies.

Anti-Kanser İlaçlarına Doğru:

Kanser Biyoişaretleyicilerini Engelleyen Anti-VEGF ve Anti-HER2

Aptamerlerinin Seçimi

İRENA ROÇİ

Biyoloji Bilimleri ve Biyomühendislik, MS Thesis, 2012

Tez danışmanı: Javed H. Niazi

Anahtar Kelimeler: *kanser, HER2, VEGF, aptamer, SELEX*

## ÖZET

Yaklaşık 71 tane kanser tipi vardır ve bu kanser tiplerinin çoğunluğu karma bir yapı gösterirler. Bu nedenle, her kanser tipi için ortak olan fizyolojik değişikler üzerinde çalışmak kanser tedavileri için başarılı bir yaklaşım olabilecektir. Çoğalma sinyalinin geliştirilmesi ve damar gelişiminin fazlalaşması kanser gelişimine sebep olan birçok mekanizma arasında yer alıyorlar. HER2 and VEGF proteinlerinin ekspresyonları yukarıdaki mekanizmalarda, sırasıyla, artmaktadır. Bu nedenle bu çalışmanın amacı yukarıda bahsedilen proteinlere bağlanarak onların fonksiyonlarını engelleyecek olan sentetik tek zincirli DNA molekülleri (aptamer) üretmektir. Bu biyoişaretleyiciler için literatürde aptamerler bulunmakta olsa da, bu calışmada amacımız farklı potansiyeli olan yeni aptamerler seçmektir. Bu spesifik aptamerler, magnetik boncuklar üzerinde hareketsiz bulunan HER2/VEGF proteinler kullanılarak SELEX (Üstel Zenginleştirme aracılığıyla Ligandların Dizgeli Gelişimi) teknolojisi ile geliştirilmiştir. Bunun için zenginleştirilmiş tek zincirli DNA havuzu oluşturulmuş, klonlanmış ve DNA dizi analizi yapılmıştır.

Anti-VEGF aptamerinin ilginç bir özelliği tesbit edilmiştir ki bu da G-dörtlü yapıya sahip olması ve bu yapı sayesinde VEGF proteinine spesifik olarak bağlanma özelliğinin bulunmasıdır. Buna karşılık HER2 aptameri iki gövdeli bir yapıya sahiptir ki bu özelliğide HER2 proteinine bağlanmasında rol oynayabilir. Anti-VEGF aptamer ile ilgili yapılan bağlanma dengesi analizleri $K_d$ (çözünme sabiti) değerini 315 nM göstermiştir. Bu değer VEGF proteinine karşı sıkı bir bağlanma eğilimini gösterir. Bunun yanında anti-HER2 aptameri Kd değerini yaklaşık 309 nM göstermiştir ki bu da zayıf bir bağlanma eğilimine işaret eder. Bu çalışmada insan serumu ile karşı seçim aşamaları kullanılmış, böylece aptamerlerin seçiciliğinin artırılması amaçlanmıştır. VEGF proteinine bağlanan aptamerler çeşitli biyotanıma teknolojileri kullanılarak kanserin kanda teşhisinde kullanılabilirler.

# ACKNOWLEDGEMENT

Two years of project work were finalized with this thesis and I am happy to extent my sincerest thanks to those people who supported me during this period of time and made my life easier, who deserve special acknowledgement and to whom I would like to dedicate this thesis.

My sincere gratitude is to my advisor Dr. Javed H. Niazi, who supported, guided and helped me during masters and finalizing this work. Special thanks to TUBITAK and Sabanci University who supported me financially. I also want to thank Dr. Hayriye Unal who supported me with her invaluable advices and guidance, and to Dr. Meral Yuce for her help. My thanks are to Prof. Dr. Yasar Gurbuz for his support throughout my studies, and to his group for helping me in difficult situations and providing a friendly office environment.

My special thanks are to the jury members: Dr. Alpay Taralp, Dr. Deniz Sezer and Dr. Ugur Sezerman for their patience, precious advices and help for preparing my thesis, and their contributions for reaching a more scientific understanding on the matter.

Finally my gratitude is to my family who supported me with their love and confidence in my ability to succeed; to my friend Husameddin Karahan for his patience and guidance for important decisions; to my friends Bihter Avsar and Reza Pakdaman for their support, especially during thesis preparation.

# TABLE OF CONTENTS

# TABLES LIST

# FIGURES LIST

# ABREVIATIONS

A       adenosine

aa.     aminoacid

A600    absorption at $\lambda = 600$ nm

AMP     Ampicillin

APS     Ammonium persulfate

ATP     adenosine 5'–triphosphate

Bis     N, N'–methylene bisacrylamide

bp      base pair(s)

BSA     bovine serum albumin

C       cytosine

°C      temperature in degrees Celsius

CE      capillary electrophoresis

Da      Dalton

DNA     deoxyribonucleic acid

dNTP    deoxyribonucleotide triphosphate

ds      double–stranded

E. coli Escherichia coli

EDC     1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide,

EDTA    Ethylenediaminetetraacetic Acid

EMSA         electrophoretic mobility shift assay

EtBr    3, 8–diamino–5–Ethyl–6–phenyl phenanthridinium Bromide

FITC    fluorescein isothiocyanate

g       gram

G       guanosine

h       hour(s)

$K_d$     dissociation constant

l       liter

LB      Luria Bertani medium

m       meter

M       mol/l, molar

MES     2-(*N*-morpholino)ethanesulfonic acid

min     minute(s)

mM      micromolar

NHS     *N*-Hydroxysuccinimide

nM      nanomolar

NMR     nuclear magnetic resonance

nt      Nucleotide(s)

OD      optical density

PAGE    polyacrylamide gel electrophoresis

PBS     phosphate buffered saline

PCR     polymerase chain reaction

RNA     ribonucleic acid

rpm     rotations per minute

RT      room temperature

sec     second(s)

SDS     Sodium dodecyl sulfate

SELEX   Serial Evolution of Ligands by Exponential enrichment

SPR     surface plasmon resonance

ss      single–stranded

T       thymine

t       time

TBE     3.3', 5.5'–tetramethylbenzidine

TEMED       N,N,N',N'–Tetramethylethylenediamine

Tris    Tris–(Hydroxymethyl)–Aminomethane

U       units of enzymatic activity

UV      ultraviolet

V       volt (s)

v/v     volume per volume

VEGF    vesicular endothelial growth factor

w/v     weight per volume

μl      micro liter

μM      micro molar

# 1. INTRODUCTION

## 1.1. Mechanisms of Cancer

Cancer is an abnormal condition involving upregulated and uncontrollable cell reproduction. There exist more than 71 different cancer types (1) and most of them are heterogeneous in all the phenotypic characteristics like cellular morphology, gene expression (including the expression of cell surface markers and growth factor and hormonal receptors), metabolism, motility, and angiogenic, proliferative, immunogenic, and metastatic potential (2). On the other hand as described by Hanahan and Weinberg (3) there are similar mechanisms in most of cancers which together cause tumor growth and development. For this reason knowing all the details about cancer might not help. Instead preventing the inducing mechanisms which are common in most cancers is a potential solution.

The knowledge that we have from a vast number of studies suggests some new characteristics of the tumor cells which are in the form of physiological alterations acquired during tumor development and common for most if not all tumor types. For this reason they were called 'acquired capabilities'. These important changes include: self-sufficiency in growth signals, insensitivity to growth-inhibitory (antigrowth signals), evasion of programmed cell death, limitless replicative potential, sustained angiogenesis, and tissue invasion and metastasis. Over years it was understood that the acquired capabilities are made possible by two 'enabling characteristics' like genomic instability and tumor-promoting inflammation. There were also found two characteristics which facilitate the development and progression of most cancers like deregulating cellular energetic and avoiding immune destruction, which are called 'emerging characteristics'.

1
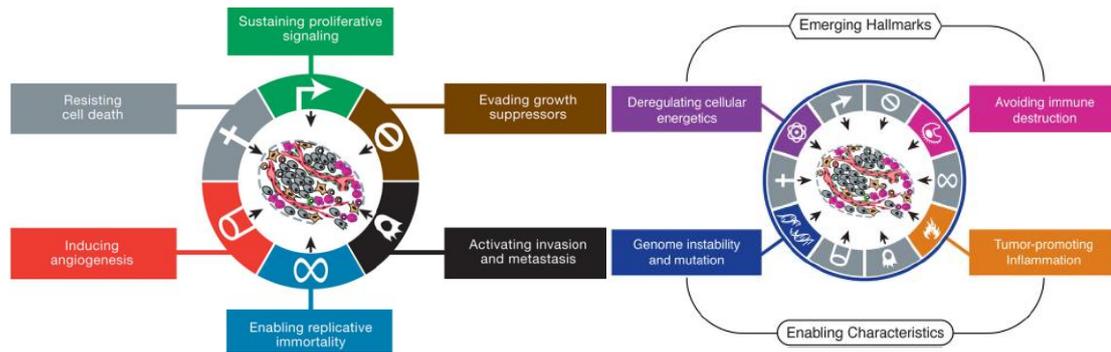
Figure 1. Emerging hallmarks and enabling characteristics of cancer. (3), (4)

### 1.1.1. Self-sufficiency in Growth Signals

Normal cells need growth signals to perform the transition between quiescent and proliferative states. Diffusible growth factors, extracellular matrix components and cell-to-cell adhesion/interaction molecules transmit these signals by binding to transmembrane receptors. Being less dependent on the exogeneous signals than the normal cells, tumor cells achieve autonomy in signaling by altering extracellular growth signals, transcellular transducers or intracellular circuits that receive signals. For example, the cell surface receptors that transduce the signal of growth factors and mainly carry tyrosine kinase activity in their cytoplasmic domain in many cases are deregulated during tumor pathogenesis. In this way the cancer cells become hypersensitive even to low levels of growth factors which normally would not trigger proliferation. A typical example in this case is HER2/neu tyrosine kinase receptor.

### 1.1.2. Insensitivity to Antigrowth Signals

In normal cells everything is balanced. In order to maintain homeostasis growth factors and inhibitory factors are also balanced. The former act by increasing proliferation while the later counterbalance by causing cell senescence. These signals too are received through transmembrane receptors to be transduced to intracellular circuits. In order for the tumor cells to develop they should be resistant to growth inhibitory signals. For example, retinoblastoma protein (pRb) which is a representative of the antiproliferative agents inhibits proliferation by altering the expression of the genes responsible for G1-S phase transition. Cancer cells are made insensitive to anti

2

proliferative agents by disrupting pRb pathway and allowing the expression of S phase transition genes making it uncontrollable.

### 1.1.3. Evasion of Programmed Cell Death

While in normal conditions both cell proliferation and cell death are well programmed, in tumor cells proliferation is higher and cell death is lower than normal. Apoptotic machinery is composed of sensor and effector components. Sensors like IGFs monitor the extracellular and intracellular conditions for any abnormalities like DNA damage or signaling imbalance. On the other hand mitochondria and proteases (caspases) are the effector components which respond to signals favoring apoptosis. In such a case cytochrome C would be released from mitochondria and would in turn activate caspases which destroy the subcellular structures.

Cancer cells acquire cell death evasion mostly by mutation in p53 tumor suppressor gene which indirectly causes the release of cytochrome C from mitochondria. A mutated p53 causes defect in cell death. Similarly defect in PI3-Akt/PKB pathway which sends antiapoptotic signaling is a way to resist apoptosis in cancer cells.

### 1.1.4. Limitless Replicative Potential

Another acquired characteristic of tumor cells is immortality. When tumor suppressors like pRb and p53 are disabled the cells continue to replicate for many other cycles after senescence by entering a stage called crisis which is a massive cell death. 1 in $10^7$ cells that enter crisis emerge into a new variant which multiply without limit becoming immortal. Telomeres found at the ends of chromosomes are composed of several thousands hexanucleotide repeats. Due to the inability of the DNA polymerase to completely replicate the 3' ends of the chromosomes during S phase of each replication around 50-100 bp of telomeric DNA is lost in each cycle. In normal cells the chromosome ends would end up unprotected and participate in end-to-end chromosome fusions after a number of cycles as a result of telomere shortening. But this is not the case for cancer cells. Tumor cell population overcomes this by two mechanisms: 1.

overexpression of the telomerase enzyme which adds hexanucleotide repeats to the end of the chromosomes or 2. Through interchromosomal exchange of sequence information (recombination).

## 1.1.5. Sustained Angiogenesis

Cells would not be able to survive without the supply of nutrients and oxygen which are enabled by the capillary blood vessels. For this reason the cells should reside within 100μm of blood vessels. When cancer cells evacuate a remote place and start growing there (metastasis), they are in need of blood vessels which supply them with the source of life. In order to survive cancer cells have to keep control on the positive and negative angiogenesis signals in the opposite way that normal cells do. To acquire angiogenesis the tumor cells activate an angiogenic switch (5) by favoring the angiogenesis inducers and countervailing inhibitors (Figure 2) (6), (7). The imbalance of these factors causes abnormalities in the vessel pattern. VEGF and FGF which are angiogenesis initiating soluble signals, bind to tyrosine kinase receptors on the surface of the endothelial cells. In many tumor tissues the expression of those two higher compared to normal cells.
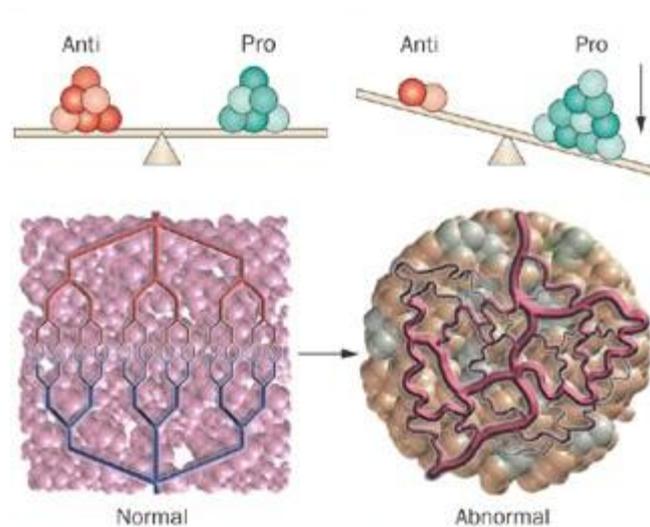


Figure 2. The control of the anti- and pro- angiogenic factors in a) normal cells and b) tumor cells. (6) (7)

### 1.1.6. Tissue Invasion and Metastasis.

The primary tumor cells migrate to new sites with more available space (metastasis). There are several elements which help tumor cell populations to acquire metastasis. Proteins like CAM (mediate cell- cell adhesion) and integrins (keep cells in contact with extracellular matrix substrates) are generally altered in order to facilitate the movement of cancer cells to a remote location. A common example is E-cadherin which is not functioning in most of the tumor cells.

Proteases are another factor which mediate metastasis by making the bed for cancer cells. The genes coding for proteases are mostly upregulated and the protease inhibitory genes are downregulated when cancer cells acquire metastasis.

### 1.1.7. Genomic Instability

Genomic stability in tumor cells generates random mutations including chromosomal rearrangements. Among these are the rare genetic changes that can orchestrate hallmark capabilities.

### 1.1.8. Tumor-promoting Inflammation

Tumor-promoting inflammation is driven by cells of the immune system, some of which serve to promote tumor progression through various means.

### 1.1.9. Deregulating Cellular Energetic

Deregulating cellular energetic involves major reprogramming of cellular energy metabolism in order to support continuous cell growth and proliferation. In this way tumor cells replace the metabolic program that operates in most normal tissues and fuels the physiological operations of the associated cells.

### 1.1.10. Avoiding Immune Destruction

Avoiding immune destruction involves active evasion of the cancer cells from attack and elimination by immune cells. This capability highlights the roles of the immune system that both antagonizes and enhances tumor development and progression. Both of these capabilities may well prove to facilitate the development and progression of many forms of human cancer and therefore can be considered to be emerging hallmarks of cancer.

## 1.2. Therapeutics Targeting

The description of hallmarks principle gives information on possible mechanism-based targeted therapeutics. The success of linking the hallmarks principle with therapeutics stands on the assumption that if a molecule is highly important for tumor development and progression, its inhibition should cause impairment of the same process. In figure 3 are given the suggested targeting (4) of the key molecules in the key pathways for each hallmark. Inhibiting only one molecule might not be enough for preventing cancer, but combination of targeting in parallel might have a higher impact in preventing tumor development.
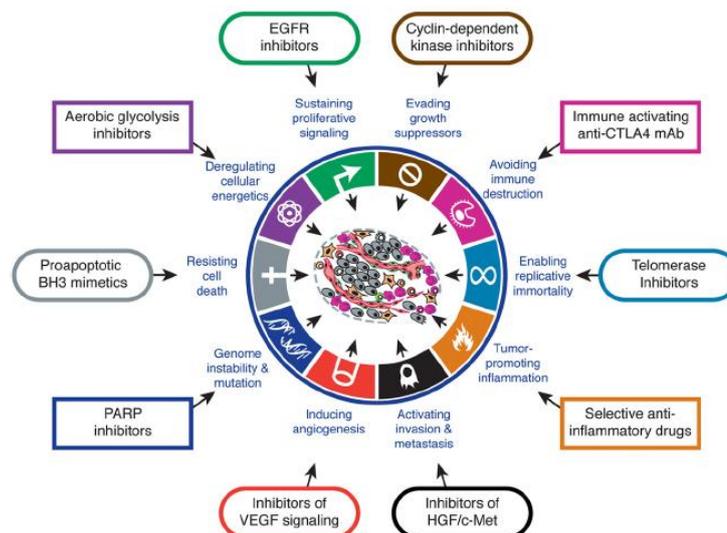


Figure 3. The ten hallmarks of cancer and the possible therapeutics targeting (4).

### 1.2.1. Inhibiting Proliferative Signaling

### 1.2.1.1. HER2

Human epidermal growth factor receptor (HER2) also known as neu / c-erbB2 is a member of HER receptor tyrosine kinase family together with the HER1, HER3 and HER4 (8), (9), (10) and (11). HER2 is a 185 kD transmembrane glycoprotein composed of 1255 aminoacids, encoded by HER2 gene mapped to chromosome 17q21 (12). All the members of the HER family form heterdimers and homodimers with each other to activate many downstream signaling pathways like PI3K/Akt and the Ras/Raf/MEK/MAPK (8), (9) and (11). As a result they regulate important events in the cell which are cell growth, survival, differentiation and migration (10), (11).
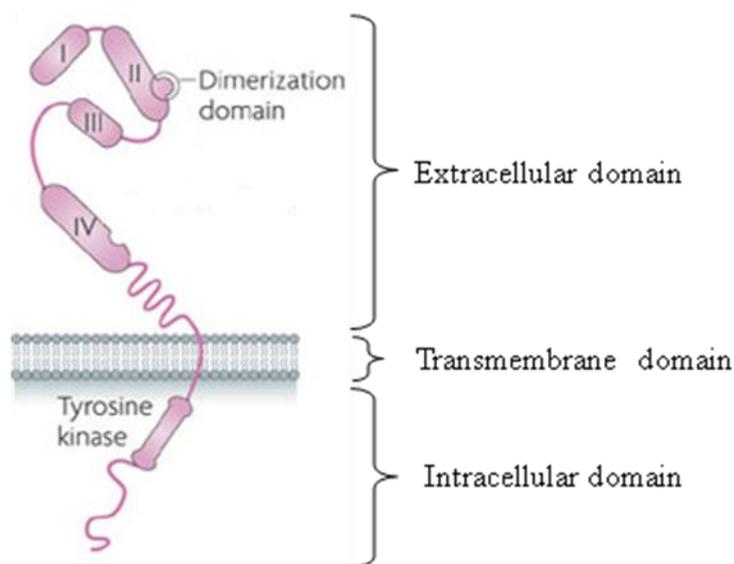


Figure 4. Representative structure of a EGFR family receptor (13)

The receptors of the HER family are transmembrane tyrosine kinase receptors composed of a cytoplasmic tyrosine kinase region, transmembrane region and extracellular domain (10). Except HER2 other members bind to ligands through the extracellular domain and become activated. HER2 which does not bind any ligand is in the activated form. Upon ligand binding the homodimer HER members dimerize with another member, become active and phosphorilate the kinase domain. Dimers show higher stability than monomeric forms of the receptors (12). Phosphorylation in the C terminus fires the respective signaling pathways. The orphan HER2 is the preferred dimer for all other members and the dimers formed by HER2 show higher signaling

than others. The reason for that is the low ligand dissociation from the receptor heterodimer, causing an increase in the signal (14). However EGFR is the preferred partner for HER2 (10).

### 1.2.1.2. The role of HER2 in tumor development

Activating many important downstream signaling cascades any alteration in the HER2 dimers cause defect in cell growth, proliferation, migration, adhesion and survival (8). HER2 is always in the active conformation and can interact with the activated receptors (15). HER2 also has a low dissociation constant to other receptors and has a low rate of endocytosis. For this reason the HER2 containing heterodimers in general show a high potency and prolonged signaling (Figure 5) (12). The dimer with the highest signal and most mitogenic is HER2-HER3 dimer (11). This heterodimer is the activator of PI3K survival and/or MAPK proliferation signaling pathway which leads to growth, proliferation, decreased apoptosis, cellular migration, and angiogenesis (15), (16), (17). Inhibition of these two pathways may cause decrease in proliferation and increase in apoptosis. On the other hand dimers that do not contain HER2 produce normal signaling and do not cause tumor growth. So, a few HER2 containing dimers are formed in normal cells and their number increases in cancer cells where they enhance the signaling.
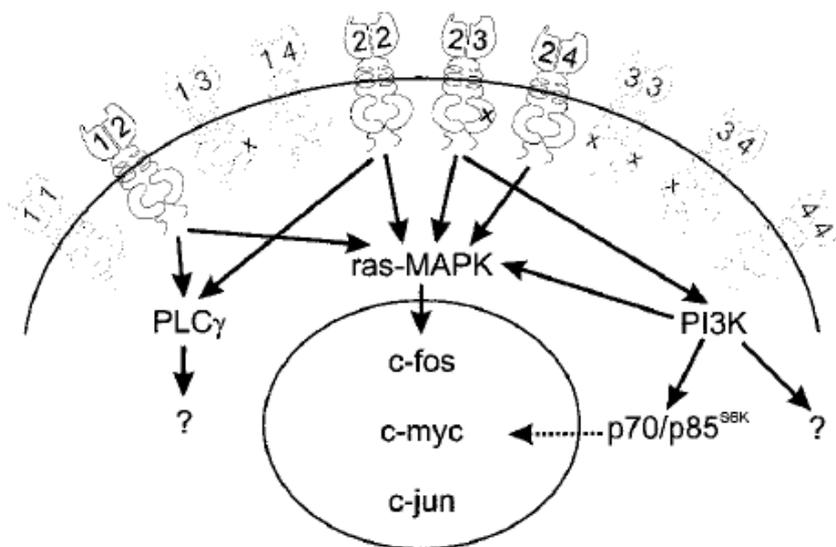


Figure 5. HER2 signal transduction pathway (15).

HER2 overexpression which is mostly caused by gene amplification is seen in 25% of breast cancers and in many other cancers like ovarian, lung, gastric and oral cancers (18), (19) (20). Breast cancers with increased HER2 expression are more aggressive and show higher mortality rates (8) (11) (14).

### 1.2.1.3. HER2 inhibition

Overexpression of HER2 causes altered cell growth, survival, differentiation and migration in tumor cells. Targeting the HER2 receptors and inhibiting their function can impair tumor development. There are several ways of HER2 targeting: extracellular domain, intracellular domain and downstream molecules (9). The extracellular domain is in active conformation and can dimerize with any other ligand bound receptor. Targeting the extracellular domain of the HER2 would prevent the phosphorylation of the kinase domain and then the activation of the signaling pathway. For HER2 inhibition through targeting the intracellular domain and the downstream molecules small molecules that penetrate the cell membrane are used.

### 1.2.1.4. The relationship between HER2 and VEGF

Interestingly, HER2 overexpression is associated metastasis, tumor progression and with overexpression of VEGF which is responsible for tumor angiogenesis (21) and (22). A study in 2004 showed a significant correlation between overexpression of HER2 and 2 VEGF isoforms ($P<0.001$) in a sample pool of 611 patients with primary breast cancer. There is a hypothesis that the aggressive phenotype of HER2-overexpressing tumors may be due in part to VEGF.

### 1.2.2. Inhibiting Angiogenesis

### 1.2.2.1. VEGF

VEGF regulates vascularization in normal cells. In tumor cells altered expression of VEGF helps in tumor progression by increasing angiogenesis which is an important factor (22). VEGF ligand is the main mediator of angiogenesis even though other factors (Figure 6) may become activated as the tumor progresses. On the other hand VEGF is genetically stable and continually expressed. For this reason an efficient strategy for the inhibition of tumor development is direct VEGF ligand inhibition. The inhibition of the other ligands can be employed to enforce the inhibition effect (23), (24).
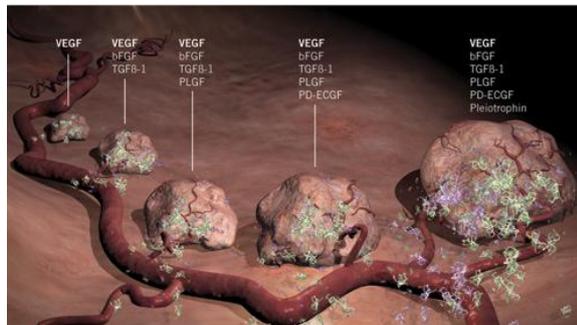


Figure 6. Expression of VEGF and secondary molecules during tumor progression (25), (26).

VEGF protein family is composed of VEGFA, VEGFB, VEGFC, VEGFD and PDGF structurally related proteins. VEGFA is a homodimer, each dimer 23 kDa. As a result of alternative splicing there are 4 VEGF isoforms: VEGF121 VEGF165 VEGF189 VEGF206. VEGF165 is the most abundant one which is secreted in the vascular system by most tumors. It binds to VEGFR1, VEGFR2 and Neuropilin tyrosine kinase receptors. Even though VEGF shows a higher affinity for VEGFR1, VEGFR2 is used for the larger part of signaling which induce vascular endothelial cell permeability, proliferation, migration and survival. Ras-Raf-MEK-ERK and MAPK signaling is responsible for DNA replication and cell replication, PI3-Kinase and Akt/PKB pathway is responsible for cell survival, and PI3Kinase/Akt signaling is responsible for cell migration (5) (24), (26), (27).
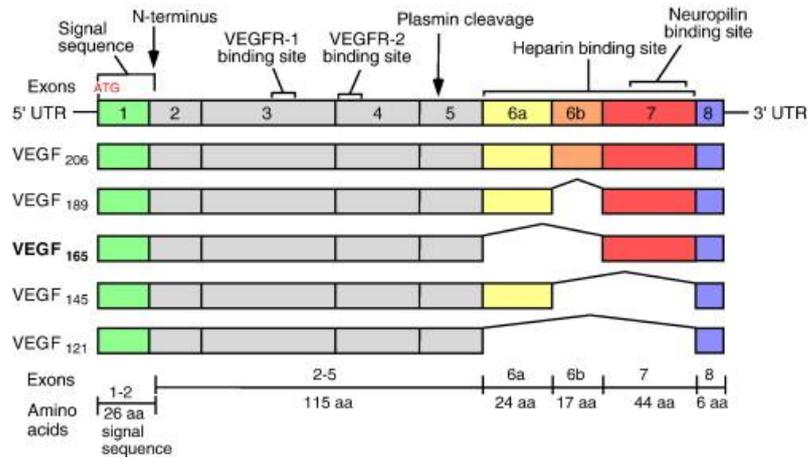
Figure 7. mRNA splice variants of human VEGFA (29)
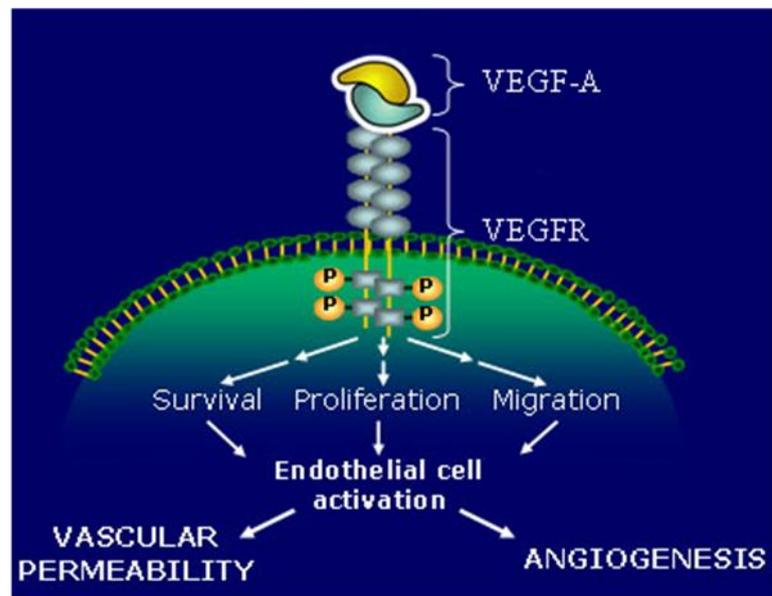


Figure 8. VEGFA signaling pathway (30).

### 1.2.2.2. Role of VEGF in tumor development

Tumor cells behave more independent than normal cells and stimulate most of the vital molecules by themselves or stimulate the environment to help them survive. In order to stimulate new vessel formation they secrete angiogenic molecules, and activate endothelial cells which produce VEGF and induce angiogenesis. (22) VEGF overexpression help tumor cells for enhancing survival of the existing cells, inducing vascular abnormalities, stimulating vessel formation and impairing immune response (31).

11

So VEGF is a survival factor for tumor cells. VEGF helps to establish, grow and survive tumor vessels. Without the supply of the nutrients cancer cells would 'die' through programmed cell death. Upon vascular maturation VEGF is not crucial for tumor survival any more. However it continues to be important for regulation of tumor angiogenesis throughout life cycle of the tumor cells. (32)

Supporting vascular abnormalities like vascular permeability is another way VEGF enhancing tumor development. As the activity of VEGF increases in tumor development, the permeability increases and causes leakage of plasma protein and formation of extravascular fibrin gel which forms a suitable environment of endothelial cell growth. So in tumors, the vasculature is excessively permeable and leaky, causing uneven delivery of nutrients, and oxygen. (31)

### 1.2.2.3. The inhibition of VEGF

In many studies of cultured cancer cells it was observed that the VEGF inhibition maintains anti-angiogenic effects, forcing the tumor cells tumor cells to grow and spread less. On the other hand preclinical studies show that withdrawal of VEGF suppression makes tumor cells behave as in the untreated cells and restart the abnormal vascularization (Figure 9). (33)
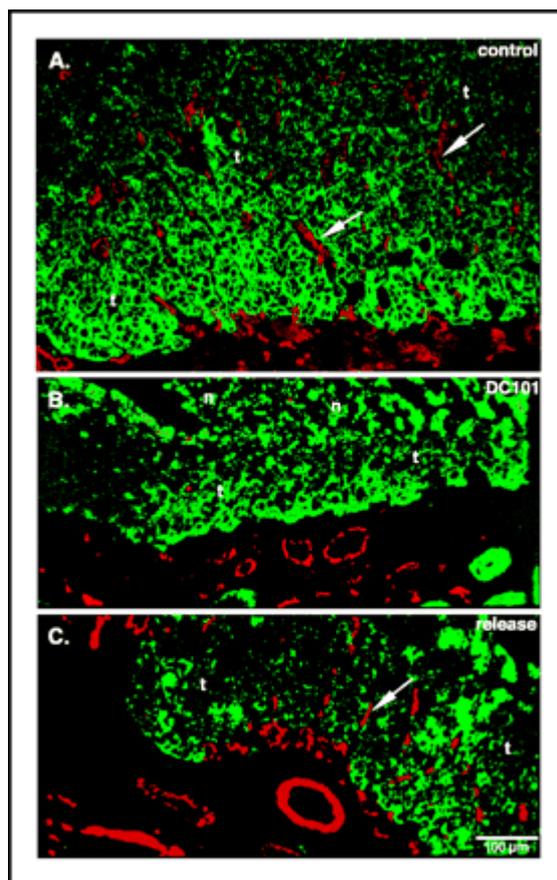
Figure 9. Vascularization a) untreated, b) upon treatment of c) discontinuation of treatment of tumor cells (33).

## 1.3.  In Vitro Selection

### 1.3.1.  Aptamers

Aptamers are high sensitivity, stability and specificity biological recognition elements which also can be used in medical diagnostics. In addition they are advantageous over antibodies because they are produces more simply by chemical synthesis (in a test tube), are more flexible in terms of storage conditions, are little or non immunogenic and can be modified to give desired properties. Aptamers are short, single stranded oligomers of 70-90 bases long which adopt complex, sequence dependent secondary and tertiary structures, which enables specific interactions with their targets.

Unstructured in solution, aptamers undergo structure change and fold upon target binding (34), (35). The sequential evolution of aptamers in each cycle leads to selection of aptamers which are trained to bind specifically to a certain target. This is reflected in the tertiary structure of the aptamers which fold around the ligand through specific recognition. By means of stacking, hydrophobic interactions, hydrogen bonds and/or electrostatic interactions they form unique structures around their specific targets (35). Different from the natural nucleic acids which serve several functions in the cell, in the in vitro selection the evolutionary pressure act to select only the aptamers which show high specificity and affinity. While it is the target that adapts to the structure of the natural nucleic acids, the selected aptamers during the in vitro process are the ones that bind to their targets by adaptive recognition. (35)

### 1.3.1.1. Structural features of aptamers

It is observed that the specific aptamer-target interactions mediated by the unpaired nucleotide sites and other stable secondary structures support the functional motif formation of the aptamer. Stems, internal loops, hairpins, purine-rich bulges, pseudoknots, and quadraplexes are the structural motifs encountered in most aptamers' structure (36). The most common motifs seen in aptamers are hairpins. Pseudoknots are formed as a result of complementary interactions of the hairpin loop with the sequence around the loop. Four-stranded structure (quadruplexes) are composed of layers of 4 guanine (G-quartet) nucleotides which form hydrogen bonds with each other. The least quadruplex contains 2 G-quartets. These structure confer higher stability than other structures and are more typical for DNA aptamers. The stability is increased in the presence of ions like potassium in between the quartet layers. In many studied cases there are the loops in the quadruplex that play role in target recognition. (37)

Aptamers are selected in vitro from combinatorial oligomer pool. The possible targets studied up to now are proteins, peptides, nucleic acids, polysaccharides, small organic molecules (aa., nucleotides and other metabolites), virus particles, whole cells and tissues. Because of the larger surface and consequently higher potential of hydrogen bond formation, aptamers show higher affinities for protein targets (36). Aptamers can be used for studying the protein-nucleic acids interaction, to inhibit target proteins, to

detect different target molecules, etc. Even though mostly it is expected that aptamers have inhibitory effect on their targets, in some cases it was seen that aptamers enhance target activity. (37)

Aptamers confer high affinity and specificity for their targets. In general aptamers selected for protein targets show higher affinity than aptamers selected for smaller molecules. This is most probably because of the larger area of interaction between aptamer and the protein, when compared to the smaller molecules. The same logic follows for higher specificity toward larger molecules. The higher the interacting surface between the aptamer and the ligand, the easier it is to distinguish the small alterations. At the same time it depends on the site of the protein where aptamer is binding. If the aptamer binds on the site that is unique from the other closely related proteins it shows high specificity. If the other way around happens, the aptamer will not be selective for the target. Nevertheless, many aptamers are able to distinguish their target from another very similar protein. For example protein kinase C is different from its isoform only by 4% and its aptamer can distinguish those two by showing a higher affinity for kinase C (37). However, both affinity and specificity can be improved by modification of the structure of the aptamer.

The selection of high affinity and specificity aptamers from vast combinatorial oligomer pool is done by means of a technique called Serial Evolution of Ligands by Exponential erichment (SELEX), through selective and competitive binding.

### 1.3.2. SELEX

SELEX is the selection and exponential enrichment of target-specific candidates through iterative cycles $(5 - 15)$ of: (i) incubation of the large combinatorial library of aptamers with the target molecule, (ii) separation of bound from unbound aptamers; (iii) elution of bound aptamers; (iv) PCR amplification of the binding aptamers for further selection round; (v) cloning of the potential candidates. Apart from these, counter and negative selection steps are included to select the candidates that have both strong binding ability as well as selectivity. This is called toggle-SELEX and is achieved by toggling the target with other molecules.
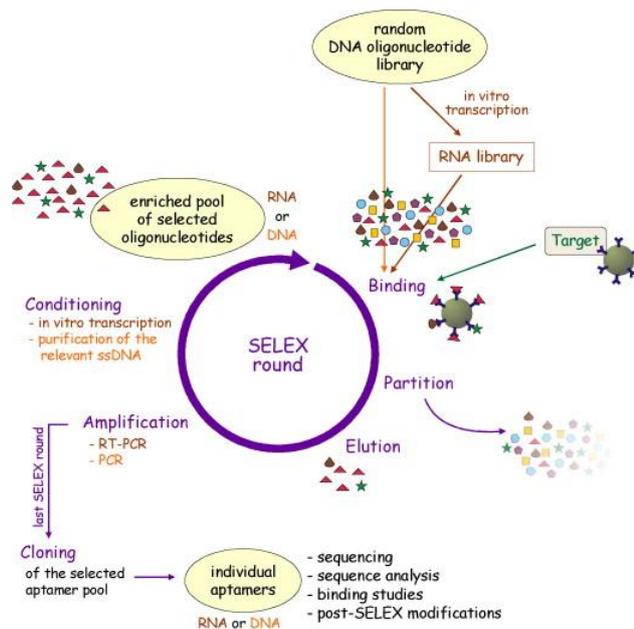
Figure 10.  A representative picture of the SELEX process. (38)

The starting pool is composed of oligomers containing a random region ranging from 20-200 nucleotides and two constant sequences 18-30 nucleotides long flanking at both ends (39). The constant sequences play role in the molecular biology processes like amplification of the selected candidates and the random region makes the combinatorial library.

Separation of the unbound oligomers from bound-target complex is an important issue during aptamer selection. There exist several ways (37) for the separation step: i. protein-bound aptamer complex is separated through nitrocellulose filters and the unbound pass through the filter; ii. protein is immobilized and the unbound are washed away; iii. gel electrophoresis; iv. capillary electrophoresis; and v. centrifugation (in the case of cell/virus selex.

After amplification of the bound aptamers, the dsDNA is separated into ssDNA by several methods (37) like: i. asymmetric PCR; ii. size separation using electrophoresis (5` labeling of one of the primers with biotin, fluorescent dyes, etc); iii. cleavage of the phosphorylated strand of dsDNA with the phage λ endonuclease (phosphate group is introduced into 5` end of one of the primers); and iv. column separation by using streptavidin and biotin labeled dsDNA (one of the primers is labeled with biotin in its 5` ends).

Though significant enrichment is observed after the first rounds of selection. The selection pool is enriched gradually and generally 5-15 cycles are applied. After the last cycle the enriched pool is cloned.

### 1.3.3. Post-SELEX Analysis of Selected Aptamers

After the SELEX cycles individual clones are sequenced and analyzed. They are classified according to the sequence homology and each class is tested for target affinity. The best candidate is considered for further analysis.

Secondary structure of the chosen aptamer candidates is predicted and they are later checked by sequence truncation. It was observed that in many cases minimizing the aptamer dimensions results in higher specificity and in lower cost of production of the aptamer (37). This is especially important in diagnostics. Later the binding characteristics of the aptamer to the protein (target) are investigated.

Further analyses include structure analysis of the aptamer-target complex. Site directed mutagenesis can be implemented to reveal the specific nucleotides that play role in the interaction.

### 1.3.4. The Characteristics of Binding

### 1.3.4.1. Binding specificity

Fitting of the aptamer to the protein binding site and their specific interaction through discriminatory intermolecular interactions depends on the nature of the nucleic acids and aminoacids from which aptamers and proteins (respectively) are made from (34). As a result of different combination of the 20 aa. there form different combinations which are high in number and produce a more defined substrate-binding site. They favor hydrogen bonds and acid base interactions with the ligand. On the other hand the nucleotides which are less in number (4 nucleotides) form a less diverse set of interactions and structures. The planar nature of the nucleotides prefers stacking

interactions with the ligand. Considering the characteristics of both proteins and aptamers the interactions favored by both of them act in a protein-aptamer complex.

Nuclear magnetic resonance spectroscopy is used to get information about the structure dynamics of the aptamer molecules and their behavior in the presence and/or absence of the target protein in solution. It was observed that aptamers form partially structured and interconverting bulge and loop conformations when they are in solution. On the other hand the same sequences form defined secondary structures when they bind to a high affinity and specific ligand. Shape complementarity, hydrogen bonds, electrostatic interactions and stacking interactions between aromatic compounds (eg. small molecule targets) (36) are the interactions that maintain the target-aptamer complex and count mostly for specific recognition. As a result of the adaptive structure formation upon target binding, most aptamers create a specific binding site in their tertiary structure. In some cases base-pair mismatches and triples are formed. However the structure adaption is not limited with the aptamer. Like in the HIV-1 Rev peptide-RRE RNA and BIV Tat peptide – TAR RNA complexes, sometimes proteins are the ones that undergo adaptive binding upon complex formation. (40)

In both cases of molecular adaptation the RNA aptamers form a deep groove from the adaptive formation of non-Watson-Crick purine-purine pairs and U-A-U base tripes where HIV RRE penetrates. The specific binding is due to the hydrogen bonds formation between the guanines and the guanidinium groups at the groove edges with the arginine residues of the peptide, which are the most specific contacts of the protein-DNA complexes (41). Other motifs that contribute to the specific binding include the non-Watson-Crick purine-purine base pair interaction with asparagines, and pyrimidine base stacking with tryptophan residues. Additionally specific protein-RNA interactions include non-Watson-Crick base pairs and triples which provide unique hydrogen bonds and distort the RNA deep groove. There are also the nonspecific intermolecular interactions between the arginine guanidium groups of the peptide and the phosphate groups of the RNAs that stabilize the complex. (34)

The deep groove formation is a characteristic of RNA aptamers which can accommodate secondary structure elements of proteins such as α helices and β sheets. This characteristic architecture is linked to bulges, non-Watson-Crick base pair and triple alignments. On the other hand proteins that bind to DNA aptamers do not room

18

protein secondary structure elements, but two or more of these are required for binding. (34)

Likewise, proteins like MS2 coat proteins which do not undergo structural change upon aptamer binding  form specific interactions that are mediated by 3 unpaired bases. The intermolecular hydrogen bonds which stabilize the unpaired nucleotides that take part in the interaction are formed by 2 unpaired adenines, and an unpaired cytosine which stacks on a tyrosine side chain. (34)

In some cases it was observed that the constant region participate in the conformations that take part in binding. From a sequence analysis (39) of 2000 aptamers binding to 141 different targets it was found that the constant regions commonly play a minor role in binding to the target. To test the involvement of the constant region in the secondary structure, the folding of the random region in the presence and absence of the constant region are compared and the results were as expected.

These aptamers were also tested for thermodynamic stability whichh is another characteristic of the aptamers which describes their endurability to mutations and environment. Even though it was shown that the constant regions play a minor role in the structure, in some cases the random regions also contribute minimally to the thermostability of the structure. Constant regions participate more in the secondary structures of the aptamers with short random regions, which means that both random and constant regions are important for secondary structure formation of these aptamer. Anti-isoleucine aptamer case which was selected from a library of varying lengths oligomers showed that the ones with shorter random region (in some cases even 50 nucleotides long) showed a higher participation of the constant region in the secondary structure. On the other hand constant regions do not have the same effect on the function of the aptamers. Random regions are more important in the functional structures of the aptamers. (39)

"The tyranny of short motifs" was a hypothesis that describes the fact that short, less information-rich and functional motifs emerge from selection experiments. If this was always the case constant regions would play role on functional structures by favoring the shorter aptamers. Actually there is such an example where anti-arginine

aptamer with a random region of 25 bases where the constant regions play a role both in structure and function of the aptamer. The prediction of "the tyranny of short motifs" is challenged by the higher number of cases in which preference random regions are preferred for the structure and function of the aptamer. (39)

There are several reasons for why the aptamers do not prefer the inclusion of the constant regions in the functional secondary structure of the high affinity and specificity aptamers. The first is more a speculation that the information-rich random regions are more functional than the information-poor constant regions. This is based on the observation that the functional conformations of the random region dominate over the ones in the constant regions. Secondly, located at the ends of the aptamers the constant regions have a lower probability to be included in the structural motifs. Thirdly, the role of the constant regions in the selection process is another factor. Constant regions are implemented for amplification of the nucleic acids by molecular biology methods. The functional structures where the random regions interact with the constant regions are under evolutionary pressure which might suppress any advantage due to high specific binding or stable secondary structure. (39)

### 1.3.4.3. Binding affinity

The equilibrium reactions for target-aptamer binding are shown by dissociation ($K_d$) or association ($K_a$) constant equations. $K_d$ (1.2) gives information on affinity of the aptamer-target, not on the kinetics of the reaction. A 1:1 equilibrium reaction of binding is represented with the equation 1.1.

$$A + T \rightleftharpoons AT \qquad (1.1)$$

Where A is the aptamer, T is the target and AT is the complex.

$$K_d = \frac{[A][T]}{[AT]} = \frac{1}{Ka} \qquad (1.2)$$

In order to measure the $K_d$ value in an equilibrium reaction of complex formation, one of the components is kept constant while the concentration of the other is increased gradually. The graph in Figure 11 represents the binding plot of a ligand (A) to a target

(B). The concentration of the target keeps increasing and at the point where half of the ligand is free and half of the ligand is bound to target, $K_d$ equals target concentration.
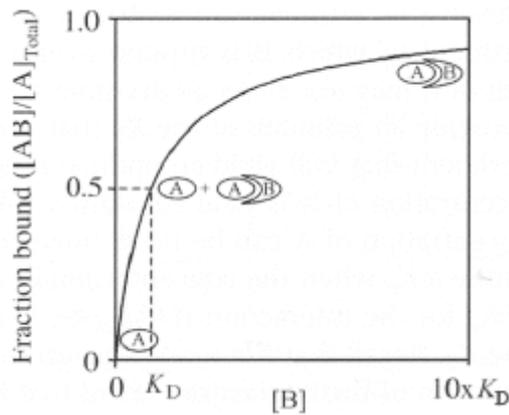


Figure 11. The binding plot of ligand (A) – target (B) (42)

A real example is illustrated in Figure 12 where is shown the 1:1 binding stoichiometry of RNA with the protein. Here the fixed concentrations of the protein are titrated with the increasing concentrations of the protein. The $K_d$ ($10^{-8}$ M) value of the complex is equal to the concentration of the protein when half of the RNA is bound to the protein. (43)
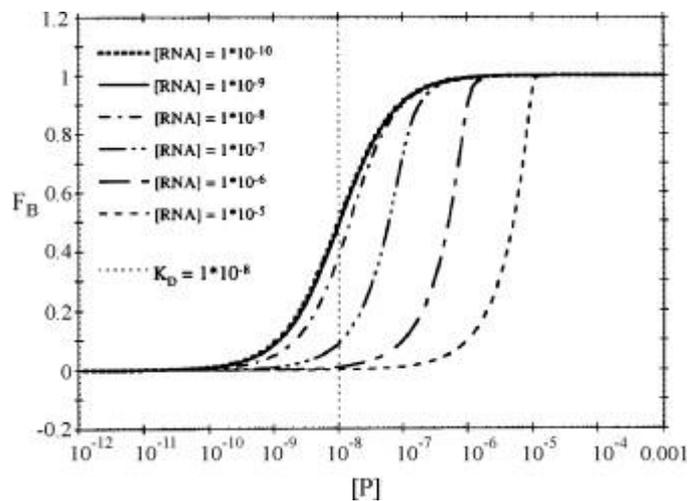


Figure 12. Simulated binding plots of an RNA–protein interaction. (43)

There are several techniques that are used for measuring the binding reactions quantitatively (44) (45).

1) Assays that separate complexes from a solution

   a) Filter-binding (Use of nitrocellulose filtering)

   b) Cell association assays (For binding of small molecule ligands)

   c) Gel-filtration chromatography (Separation based on size and shape)

   d) Electrophoretic mobility shift assays (EMSAs, makes use of gel shift in the native gels)

2) Assays that detect complexes in solution

   a) Fuorescence (FRET, induced fluorescence, fluorescence anisotropy, fluorescence quenching)

   b) Protection assays

3) Assays in which a biomolecule is immobilized

   a) Affinity resins (use of Sepharose, agarose beads for immobilization; separated by centrifugation)

   b) Surface plasmon resonance (commercially available machine that detects the binding of the molecules due to mass change near the surface of the sensor)

1.3.4.3.1. Fluorescence intensity

By measuring the fluorescence signal when the aptamer and protein are at equilibrium the $K_d$ can be measured under various buffer conditions independent of the separation techniques, and the association and dissociation rates. There are also other advantages like being a faster method than separation based ones, availability of many fluorescent probes with different lifetimes and emission wavelengths and well characterized labeling methods which make the fluorescence intensity based measuring methods a better choice to measure $K_d$ at the equilibrium reaction.

## 2. OBJECTIVE

The aim of this thesis was to select anti-HER2 and anti-VEGF aptamers for inhibition of Her2 and VEGF cancer biomarkers respectively, by in vitro selection. The motivation for this work was selection of aptamers with higher specificity and selectivity when compared to the existing ones.

Several challenges were faced during the process. Work with short aptamers (around 55 bases), gel extraction, separation of ssDNA and lack of resources were some of the difficulties encountered during the process. On the other hand use of magnetic beads for separation of the bound to unbound ssDNA and implementation of negative selection were two strategies that made the work easier to select more successful candidates.

At the end of this work relatively high affinity and specificity aptamers were selected. However this work has to be continued with structure modification to obtain more successful candidates. It is a common procedure in aptamer selection to challenge the obtained aptamer with post-SELEX modification in order to emphasize the desired characteristics and to decrease the cost of production.

The aim for the future (perspective) is to use the selected aptamers for biosensing technologies to detect the presence of cancer. By combining the two aptamers it is possible to inhibit the overexpressed Her2 and VEGF which support tumor cells.

# 3. METHODS

## 3.1. SELEX

Her2 and VEGF were used as the target proteins for selection of aptamers using SELEX technique. In each cycle, the proteins were introduced to a random pool of ssDNA strands and the ones that bind were selected and amplified to be used in the next cycle. The selected oligonucleotide strands formed the pool for the next cycle. The target proteins were immobilized on magnetic beads in order to fascilitate separation of the protein from the unbound ssDNA molecules.

### 3.1.1. Immobilization of Proteins on Surface Activated Magnetic Beads

$10^6$ magnetic beads containing carboxylic acid groups were used in each cycle. The commercially available beads were subjected to pre-activation of the carboxylic acid groups and target coating before they were used (46). Beads were washed twice using 1X binding buffer (100 mM NaCl, 20 mM Tris-HCl pH 7.6, 2 mM $MgCl_2$, 5 mM KCl, 1 mM $CaCl_2$, 0.02% Tween 20) (47) before and after each treatment.

### 3.1.1.1. Magnetic beads activation

The beads from stock were washed with 25 mM MES (2-(*N*-morpholino)ethanesulfonic acid), pH 5. Equal volumes of 50 mg/ml EDC (1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide) and NHS (*N*-Hydroxysuccinimide) respectively in 25 mM MES was added to the tube containing the magnetic beads that were used for coating. The solution was incubated in the room temperature in tilted rotation and then washed twice using MES. A magnetic stand was used for beads separation.

### 3.1.1.2. Covalent coupling

After the carboxylic acid groups on the surface of magnetic beads were activated, the coating with the target protein was performed. The required amount of target protein in 25 mM MES was added to the separated beads. The solution was incubated again for 30 min at room temperature with tilted rotation. After incubation the beads were separated from the supernatant and washed twice using 25 mM MES. The unreacted carboxylic acid groups were quenched with 50 mM ethanol in PBS (Phosphate Buffer Saline) pH 8.0 for 60 minutes. After separation, the beads were washed with PBS and were ready to be used.

A schematic representation showing the details of the protein immobilization process (magnetic beads coating and covalent coupling of protein) is as shown in Figure 13. The process occurs in two steps step using EDC/NHS crosslinker.
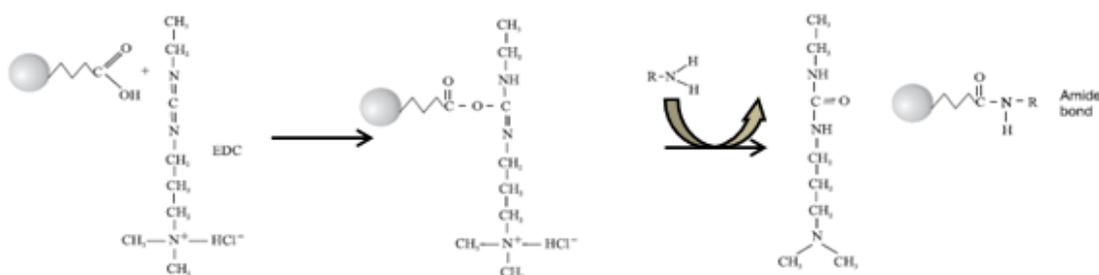


Figure 13. Covalent coupling of proteins on magnetic beads. (46)

### 3.1.2. DNA Pretreatment

A random library of ssDNAs (25 μl) was used for each SELEX process. This pool had a concentration of ~60 pmol/μl. Since the secondary structure of the oligonucleotide strands is important for binding to the target, heat treatment named as pretreatment was important to separate dimers or non-specific folders, and allow ssDNA to form stable secondary structures upon target binding. This step was performed using a PCR machine with a following program: $95^{o}$C for 10 min, $4^{o}$C for 15 min and $25^{o}$C for 7 min. The pool used was composed of ssDNA strands with sequence GGGCCGTTCGAACACGAGCATG(N)$_{40}$GGACAGTACTCAGGTCATCCTAGG. Letter N stands for the random nucleotides and for this reason the 40 nucleotide region is called random

sequence. The sequences of the forward and reverse primers are as following: ZGGGCCGTTCGAACACGAGCATG and GGACAGTACTCAGGTCATCCTAGG, respectively. The letter Z in front of the forward primer stands for fluorescein, which is bound to the 5' end of the forward primer. The primer binding sites which flank the random region are called constant regions (sequences).

### 3.1.3. Negative Selection

Negative selection step was employed both in anti-VEGF and anti-HER2 aptamers selection process. The importance of this process stands in the selection of more specific aptamers which selectively bind to the target. In other words, the ssDNA pool is first trained with non-target proteins. So the ssDNA which had a higher affinity to the non-target protein were excluded from the target incubation, while the supernatant was used for selection with the target protein. In this case, ethanolamine, BSA and 1:10 diluted whole human serum were used for negative selection. Before starting the first cycle, the ssDNA pool was incubated with ethanolamine coated beads and the supernatant was used to continue with selection with the target protein (VEGF/HER2 biomarker proteins). Negative selection with BSA coated beads was performed in the same way before the second cycle. In the sixth cycle, serum was added to the incubation reaction so that it was diluted 10 times in the final volume. In this step, the ssDNA which show higher affinity to serum proteins were separated after the incubation. So, only ssDNA that bound target protein stronger were enriched, amplified and used for the next round of selection.

### 3.1.4. Selection

In each selection step, approximately $10^6$ VEGF and HER2 coated beads were used, respectively. Assuming that all the protein was bound to the beads during coating procedure, about 21 pmol of VEGF and 12 pmol of HER2 protein were introduced to the ssDNA pool in each SELEX cycle, respectively. Each time the pool was allowed to incubate with the target coated beads and the ssDNAs that bound to the target protein were separated by using magnet while the unbound ones were discarded. In order for the most specific candidates to bind to the target protein, several stringent conditions

were introduced. For example, the number of beads used was 0.25 x $10^6$ (4 times less) in the fourth and fifth cycle for both targets, and $10^5$ beads were used in the eighth cycle for HER2 SELEX. Then, in the seventh cycle, the incubation time was decreased to 10 min from the initial incubation time of 1 h in the first cycle.

### 3.1.5. Elution

The unbound ssDNA remained in the supernatant while the target ssDNAs was bound to the surface of the protein. The bound ones were recovered by simply drawing the magnetic beads bound with aptamer-protein complex using a magnet. In the elution process, 100 µL of elution buffer (40 mM Tris-HCl pH 8, 10 mM EDTA, 3.5 M urea, 0.02% Tween 20) (47) was added to the beads coated with target protein and with the candidate aptamers bound. They were incubated at 80 $^o$C for 10 min mixed by vortexing every 2 min for 30 min. This step was repeated twice. The supernatants in which the candidate aptamers were found were purified and recovered by ethanol precipitation. Elution step was performed in the first 3 cycles for HER2 and in the first 2 cycles for VEGF. The main aim of the elution step is to use the ssDNA in the PCR for amplification. Since some ssDNA was lost during the elution step, in the later cycles the target protein coated beads bound ssDNAs were directly used for PCR amplification.

### 3.1.6. Ethanol Precipitation

All ssDNA samples after elution were subjected to purification by ethanol precipitation. This step was also performed to concentrate the ssDNA obtained from the elution step. After ethanol precipitation, the initial volume of the eluted ssDNA was concentrated (from ~250 to ~30 µl). First 1/10$^{th}$ volume of sodium acetate was added to the ssDNA tube and then 3 volumes of 100% cold ethanol, and 1 volume isopropanol were added. The samples were incubated at -80$^o$C overnight for cold precipitation. The tubes were centrifuged for 20 min at 13000 rpm, the supernatant was decanted gently and the pellet was washed with 1 volume 70% cold ethanol followed by centrifugation at 13000 rpm for 10 min. This washing step which aims to dessolve the salts was repeated twice. The supernatant was decanted and centrifuged again for 1 min to

sediment the pellet at the bottom of the tube. The pellet was air dried and resuspended in 1X EB buffer (commercially available).

### 3.1.7. PCR Amplification of ssDNA Traces

PCR reaction was used to amplify the selected ssDNA strands. The number of cycles varied initially from 35 down to 21 cycles in the last cycle. The PCR program contained following steps: preheating at 95 $^{o}$C for 5 min, denaturation at 95 $^{o}$C for 20 sec, extension at 72 $^{o}$C for 20 sec and finally, 10 min extension at 72 $^{o}$C. Annealing temperature varied from 45 in the first cycle to 55 $^{o}$C in the last cycle of Her2 SELEX. In the case of VEGF, the annealing temperature was increased from 44 $^{o}$C in the first cycle to 47 $^{o}$C in the last cycle. For a 50 µL reaction, 1 µM of forward fluorescent and reverse primers each, along with a commercially available Taq premix (Qiagen) were added. Around 30 ng of ssDNA template was added to each reaction tube for amplification.

### 3.1.8. Agarose

Agarose electrophoresis was used to run the PCR product, and visualize and discriminate the band of interest from any other contamination. 2% agarose was used to run double stranded DNAs shorter than 100 bp. The expected band was predicted to be right below the last band of a 100 bp ladder. The unstained gel was initially inspected for occurrence of a green fluorescent band for documentation, which was followed by ethidium bromide staining to discriminate any non-specific DNA bands. The staining was performed after electrophoresis in order not to interfere with running of the sample. The bands of interest appear green in color before and orange after staining. Both the gel and the running buffer were composed of 1X TBE (Tris/Borate/EDTA) buffer and the gel was run at 100 V (50 mA).

### 3.1.9. Agarose Extraction

The specific bands on the agarose gel were excised and transferred to an eppendorf tube. They were weighed and 3 times the volume of the gel pieces QX1 buffer (Qiagen) was added (as per the manufacturer's instructions). The mixture was added with 10 µl of QXII buffer and the solution was heated at 50 $^{o}$C until the gel was

dissolved. The tubes were centrifuged for 30 seconds at 14000 rpm and the supernatant was discarded. The above process was repeated by adding another 500 µl of QX1 buffer and centrifuged for 30 sec. The pellet was washed twice using PE buffer (Qiagen). After the last centrifugation, the pellet was dried until it becomes lime white and then resuspended in EB buffer (Qiagen). The concentration of the ss/dsDNA was measured using Nanodrop spectrophotometer at $A_{260}$.

### 3.1.10. Denaturing PAGE

The double stranded DNA cannot be used for selection because it does not attain a proper secondary structure for binding to the protein. For this reason, Denaturing Polyacrylamide Gel Electrophoresis was performed to separate the dsDNA into single stranded DNA. Two layer PAGE gel was used to run the dsDNA samples. The top layer was composed of native gel (with no denaturating agent) while the bottom layer is composed of denaturing gel where the forward fluorescent strand is separated from the other strand, as a result of 7 M urea and 20% formamide present. PAGE was performed using 1X TBE as the running buffer. To provide a better separation, the samples are run in low power (50 V) for around 2 h. After the tracking dye reached at the bottom of the glass plate, the gel was separated from the glass plates and visualized under transilluminator. The green fluorescent bands representing the ssDNA of interest was excised for next step.

### 3.1.11. Extraction of ssDNAs from PAGE

The gel was frozen at -80 $^{o}$C for 30 min and the gel was crushed and excised the fluorescent bands using a pestle in an eppendorf tube. Diffusion buffer (0.5 M ammonium acetate, 10 mM magnesium acetate, 1 mM EDTA, pH 8.0, 0.1% SDS) was added in twice the mass of the crushed gel and the tubes were incubated at 50 $^{o}$C for 1 h in tilting rotation. The samples were centrifuged and the supernatant was transferred to a new tube. The above process was repeated for the remaining gel pieces until all the ssDNA traces were recovered from the gel. Then three times the volume of supernatant QX1 buffer and 10 µl of QXII buffer was added. The mixture was then incubated for 10 min at room temperature and centrifuged 10,000 rpm for 30 sec. The pellet was resuspended for a second time in QX1 buffer and washed twice with 500 µl PE buffer (Qiagen). The pellet was lime dried and resuspended in EB buffer. After centrifugation

29

for 30 sec, the ssDNA concentration was measured using a Nanodrop spectrophotometer at $A_{260}$. Thus, obtained ssDNA was used as a next selection pool for the further SELEX process.

## 3.2. Cloning of ssDNA Pool from SELEX

TOPO TA Cloning kit was used for all the cloning experiments (36). Prior to cloning steps, the ssDNA pool obtained after the completion of SELEX process was amplified by using unlabeled primers. The dsDNA pool was then cloned and subjected to sequencing.

### 3.2.1. Screening of the positive clones

The pCR4-TOPO plasmid vector was used for cloning. This vector contains single 3' thymidine (T) overhangs and topoisomerase (function) is covalently bound to the vector. Considering this characteristic, dsDNA strands with A overhangs are enough for template insertion. This property can be easily incorporated into the DNA strands by Taq polymerase which adds one deoxyadenosine (A) at the 3' end of any PCR product by using its terminal transferase activity (Figure 14).



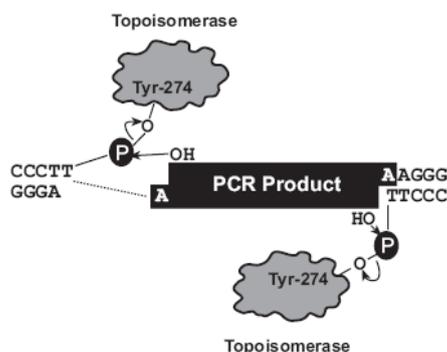Figure 14.  The template insertion reaction mediated by topoisomerase (48)

For this step, dsDNA product from a PCR reaction was directly used. The dsDNA from the last cycle was reamplified with unlabelled forward and reverse primers until the fluorescently labeled forward primer is diluted. The reason for that is the possible inhibition that the fluorescene might cause for direct template insertion into the vector (Figure 15).
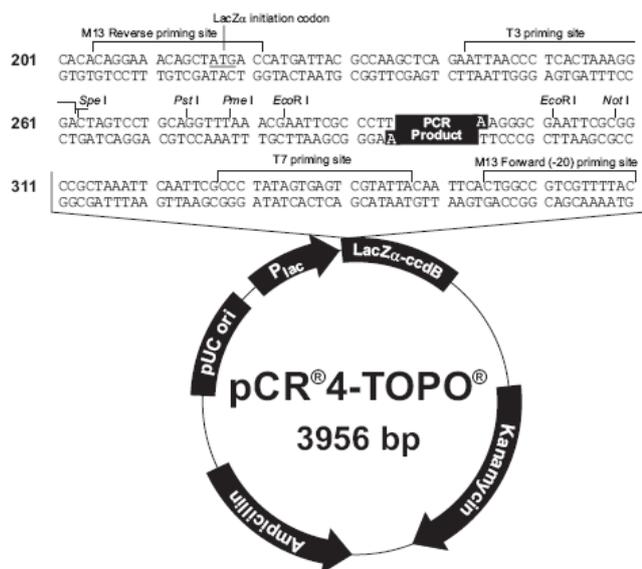
Figure 15. The TOPO vector map that was used for cloning of the candidate aptamers.
(48)

### 3.2.2. TOPO Cloning Reaction for Transformation of Competent Cells

For the transformation to be successful, it was important to have freshly prepared DNA. The A overhangs in the PCR products can be lost by time and this can inhibit transformation to the competent cells. The cloning reaction contained fresh PCR product (from anti-HER2 and anti-VEGF, respectively) 3 μl, salt 1 μl and TOPO vector 2 μl which made a total reaction volume of 6 μl.

The cloning reaction mixture was incubated at the room temperature for 15 min and transformed into 45 μL of *E. coli* DH5α competent cells (as recommended by the manufacturer). About 200 μl of LB was added to each reaction tube and then incubated for 1 h at 37 °C in tilted rotation. The transformation reaction mixture was spread on LB-agar plates containing 50 μg/ml ampicillin and incubated for 10 h at 37 °C.

### 3.2.3. Screening for Positive Clones

About 45 individual colonies were obtained in total from anti-HER2 and 63 colonies from anti-VEGF cloning. Before plasmid isolation, each of the colonies was cultured in LB amp+ media for overnight. After culturing for 12 h about 20 ml from each cell culture with $OD_{600}$=~0.1 was plated in a separate LB-amp+ plate. Again about

3 colonies from each plate were cultured in individual tubes containing 4 mL LB broth and incubated for another 12-16 hours. These cultures were used for plasmid extraction.

### 3.2.3.1. Plasmid extraction

The plasmids were isolated using a commercially available plasmid purification kit (MiniPrep kit, Qiagen). The cells from the LB broth were first harvested by centrifugation for 15 min at 4 $^{o}$C. The other steps were followed as recommended by the manufacturer. The bacterial pellet was resuspended in P1 buffer. After adding an equal volume of P2 buffer the sample was mixed by inverting and incubated at RT for 5 min. 350 µL of prechilled N3 buffer was added, mixed thoroughly by inverting, and incubated at RT for another 5 min. The samples were centrifuged for 10 min at max speed and the supernatant was pipetted to spin columns. After centrifugation 0.5 mL of PB buffer was added to the column and centrifuged again. The column was washed with PE buffer, and the DNA was eluted with EB buffer and quantified by Nanodrop spectrophotometer at A$_{260}$.

### 3.2.3.2. Restriction digestion and confirmation of positive clones carrying aptamer inserts

The plasmids were analyzed for inserts by digesting with restriction enzyme *Eco*RI. As it is shown in the plasmid map (Figure 15), the *Eco*RI sites are present in the plasmid that was used for the reactions. Digestion enzyme together with the digestion buffer was added to the plasmid samples and incubated at 37 $^{o}$C for 1 h. From the agarose gel analysis, it was observed that the inserts are ~50 base pairs in size. Since the number of clones that could be sequenced was limited, the bands which showed longer fragments were considered for sequencing.

### 3.3. Sequencing

The extracted plasmids from the selected colonies were sent for sequencing to a commercial sequencing company in US.

## 3.4. Sequence Analysis

### 3.4.1. Secondary Structure

For predicting the secondary structure of the potential aptamer candidates, free energy minimization algorithm (Mfold server) was used. The structures that had lowest negative $\Delta G$ were selected as stable secondary structures of the ssDNA sequences.

### 3.4.2. Phylogenic Relationship of the Aptamer Sequences

Mafft server was used to align the cloned sequences and to build the phylogenic trees which were visualized in JALVIEW.

## 3.5. Binding Assay

Binding assays were performed to determine the dissociation constant which describes the interaction between protein and ligand by calculating the ratio of bound to unbound ligands. Fluorospectrometer and protein coated beads were used to perform these assays. In both anti-VEGF and anti-HER2 binding assays the fluorescently labeled ssDNA volume (50µl) and concentration (373 nM) were kept constant while the protein concentration was increased. The decrease in the fluorescent signal measured by fluorespectrometer indicated the binding of the ssDNA to the protein (VEGF/HER2). In order not to disturb the equilibrium of binding, the beads were drawn to the walls of the reaction tube by magnetic stand and the samples were obtained from the supernatant. The binding assays were performed in triplicates using three independent samples for each concentration for error calculation. To verify the specificity of the aptamers to VEGF/HER2, BSA protein was used and the measurements were taken in identical conditions.

In order to obtain the equilibrium binding curve, the obtained fluorescence signal was normalized by formula (3.1) and the results were fitted using the One Site Binding fit. Origin Pro 8.1 program was used for both calculations and curve fitting.

$$\% \ change \ in \ fluorescence \ = \ Fo - \frac{F}{Fo} \times 100 \qquad (3.1)$$

# 4. RESULTS

## 4.1. SELEX

The exponential enrichment process was performed for the selection of anti-HER2 and anti-VEGF aptamers separately. In this process, the magnetic beads were immobilized with target proteins. Their SEM images appeared rough in surface and round in shape (Figure 16). SEM images were actually taken to compare the bare beads with target coated and ssDNA treated beads respectively. Since there was not observed any difference between those samples, SEM images were used only to show the structure of the magnetic beads used during the selection process.
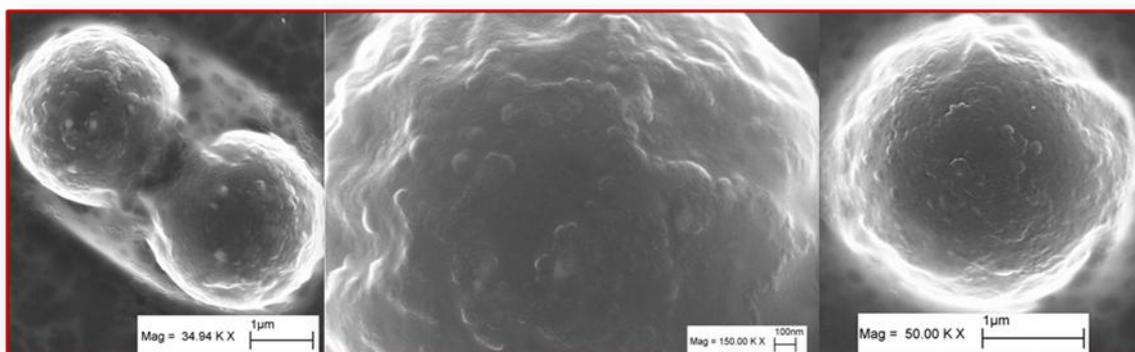


Figure 16. SEM images of the magnetic beads immobilized with target biomarker proteins.

The inherent magnetic properties of the beads made it easy for the separation of the protein bound ssDNAs from unbound molecules by simply drawing the beads with a magnet. The steps involved in each cycle of aptamer selection were as following: i. selection of target protein bound ssDNA; ii. Separation of the bound ssDNA from unbound; iii. amplification of the bound which produced dsDNA and iv. separation of the dsDNA into ssDNA for the next cycle. After the enrichment process was complete,

35

the bound ssDNA pool was cloned and sequenced. The obtained sequences were analyzed by free energy minimization algorithm using Mfold program and one candidate was chosen among the pool of anti-HER2 and anti-VEGF aptamers, respectively. The results obtained after each step of the aptamer selection process is described below.

## 4.2. Elution of Enriched ssDNA Bound to Target Protein

The aim of the bound ssDNA elution was to separate ssDNA from target protein and use them in the amplification step. Another aim of the elution process was to obtain statistical data on the enrichment of the specific aptamer pool. An increase in the ratio of eluted ssDNA from the initial pool indicates that the ssDNA pool obtained is composed of more specific candidates. However it was observed that some of the ssDNA can be lost during the elution. Therefore, the elution step was performed only during the first 2-3 cycles for anti-VEGF and anti-HER2 aptamer selection. During the later rounds, the beads containing protein-aptamer complex were used directly as ssDNA template for the PCR reaction.

As it is shown in Figure 17, the percentage of the bound ssDNA was too low after the first cycle but it increased in the consequent cycles in both anti-HER2 and anti-VEGF aptamer selection processes. This is an evidence for the enrichment of the ssDNA sequences that bind specifically to the target protein.
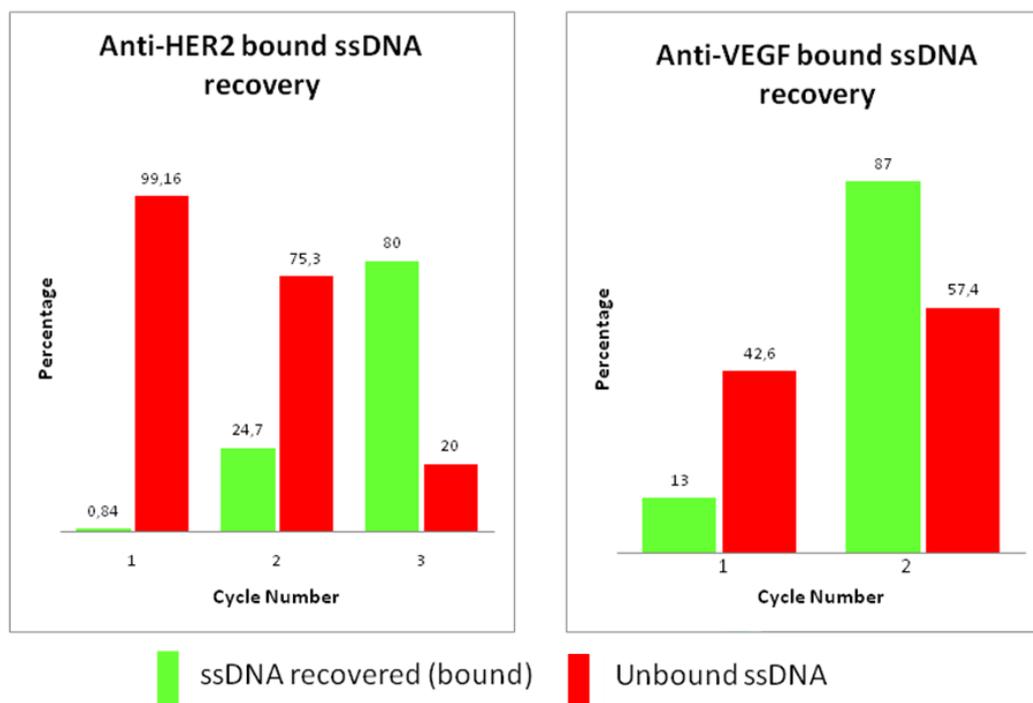
Figure 17. The graphical representation of the ssDNA recovery.

### 4.2.1. Agarose Electrophoresis

Agarose electrophoresis was performed after amplification of the bound ssDNA which was then converted to dsDNA. The agarose electrophoresis allowed size based separation of the dsDNA of interest from side products and extraction of dsDNA for further experiments. In this work, the band of interest was predicted to be around 60-70 base pairs.

Figure 18 shows that amplified dsDNA was well below the 100 bp standard band, and above the primers that appeared green that are ~25 bases long. The gels shown below were stained with EtBr. The dsDNA bands appeared green before staining and turned red/orange after EtBr staining (depending on incubation time), which enabled discriminating the target fluorescent bands from the other fragments. The bands of interest were carefully excised, extracted and purified from the agarose gel.
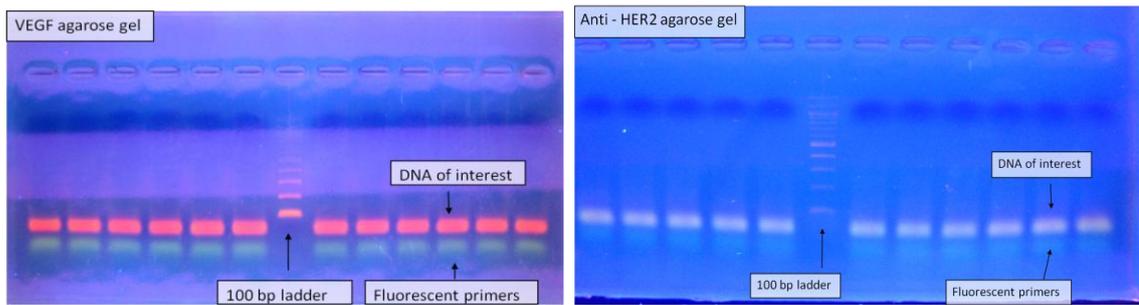
Figure 18. Representative images of dsDNA separation in agarose gel. a) Anti-VEGF dsDNA b) anti-HER2.

## 4.3. Denaturing PAGE

In order to continue with the next selection round, the dsDNA was separated into ssDNA. Denaturing PAGE was used to perform separation. The discrimination of the target ssDNA was guided by the fluorescein molecule attached to it. The highly denaturing conditions created by presence of 7M urea and 20% formamide made the dsDNA strands to separate from each other. The presence of the fluorescein in the forward strand made it easy to discriminate from the unlabeled ssDNA (Figure 19). The presence of the fluorescein makes the molecular weight of the forward strand 560 g/mol higher than the reverse strand. The green bands that were observed in the unstained gel picture are the forward ssDNA strands of interest which were excised and extracted for later use.



Figure 19.  PAGE gels of dsDNA samples. The green bands are ssDNA of interest for A) anti-VEGF, B. anti-HER2.

38

## 4.4. Cloning of the Final Selected ssDNA Pool and Screening of Positive Clones

The cloning process was performed before sequencing in order to obtain numerous copies of the same sequence. The picked colonies were cultured in LB and plated in selection media (amp⁺ LB). The positive clones were screened b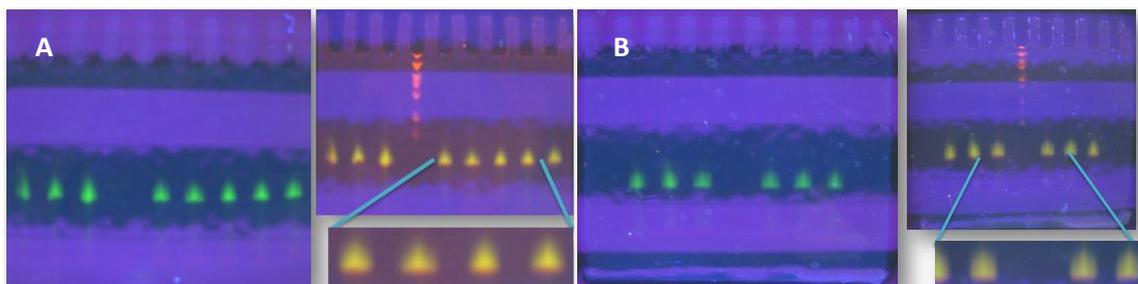y restriction enzyme digestion followed by agarose electrophoresis gel. Expected size of the fragment after digestion of the plasmid construct was confirmed. Also it was important to make sure that the candidates sequenced were at least ~60 base pairs in size. As a result of cloning reaction, 62 positive clones as anti-VEGF and 45 for anti-HER2 were obtained. The plasmids from these clones were digested separately and run in agarose gel as shown in Figure 20. The remaining plasmid DNA backbone without the insert was ~4000 bp as expected, and the inserts of interest were less than 100 bp. It appears that some of the clones did not contain the insert (for eg., numbers 2, 8, 28, 31, 50, 51, 56,57 and 61 for anti-VEGF clones and number 3, 13, 14, 16, 17 38 and 40 anti-HER2 clones), some bands were truncated to18, 42 and 45 bp for anti-HER2 clones), and some showed multiple bands (eg., numbers 26, 43, 44, 52 anti-VEGF clones and numbers 29, 35, 42 anti-HER2 clones). This data made it easier for the selection of the clones that were sequenced. Anti-VEGF clones with numbers 1, 3, 7, 14, 18, 21, 25, 27, 33, 35, 41, 54, 57, 58 and 61, and anti-HER2 clones with numbers 7, 11, 13, 18, 22, 23, 24, 26, 30, 32, 34, 38, 43 and 44 were selected for cloning (see lanes in Figure 20).
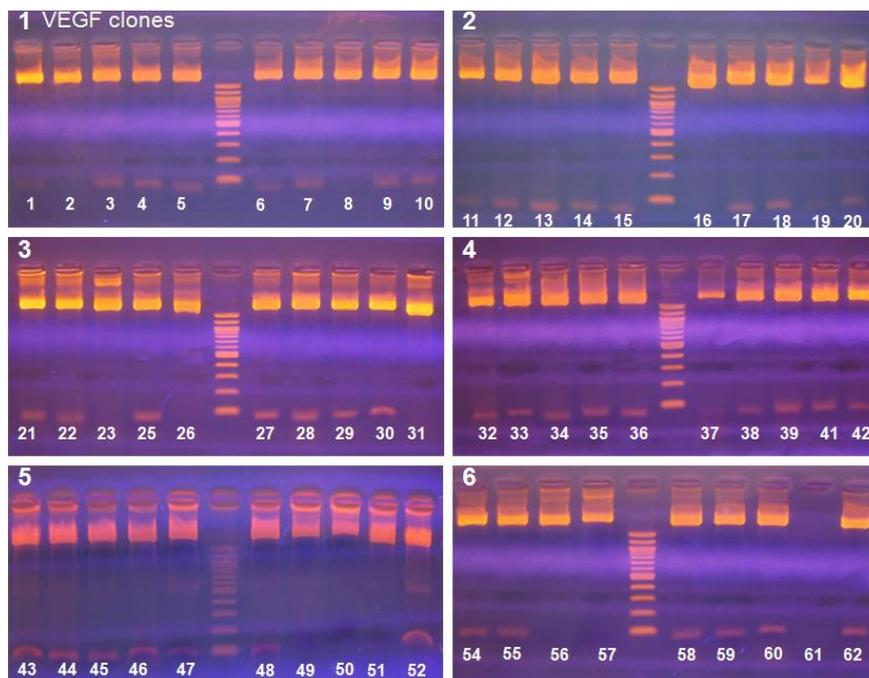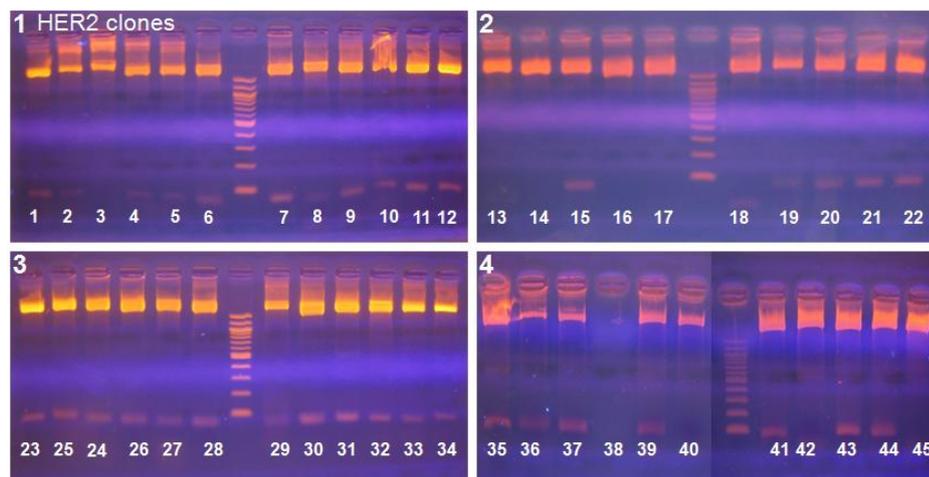
Figure 20. Screening of positive clones in agarose gel for A) anti-VEGF and B) anti-HER2.

## 3.4. Sequencing Results

The selection pool was composed of ssDNA which contained 40 bases long random sequences flanked by constant sequences. The forward primer can bind to any complementary sequence in the random region and as a result, the sequences obtained have various lengths. The length of anti-VEGF aptamer sequences vary from 50~58 bases, whereas the length of anti-HER2 clones vary from 49~55 bases long (Table 1 and Table 2).

| No | Sequences of anti-VEGF aptamers | $\Delta G^o$ |
|---|---|---|
| 1 | GGGCCGTTCGAACACGAGCATG**GGGG**CCTAGGATGACCTGAGTACTGTCC | -5.04 |
| 2 | GGGCCGTTCGAACACGAGCATG**GCGGAGCTG**CCTAGGATGACCTGAGTACTGTCC | -7.03 |
| 3 | GGGCCGTTCGAACACGAGCATG**GGCCGTCAG**CCTAGGATGACCTGAGTACTGTCC | -6.50 |
| 4 | GGGCCGTTCGAACACGAGCATG**GTGTCGAG**CCACCTAGGATGACCTGAGTACTGTCC | -6.50 |
| 5 | GGGCCGTTCGAACACGAGCATG**CGGG**CCTAGGATGACCGAGTACTGTCC | -6.38 |
| 6 | GGGCCGTTCGAACACGAGCATG**CGGCACGAC**CCTAGGATGACCTGAGTACTGTCC | -7.24 |
| 7 | GGGCCGTTCGAACACGAGCATG**GCAGTGTGC**CCTAGGATGACCTGAGTACTGTCC | -7.22 |
| 8 | GGGCCGTTCGAACACGAGCATG**CGGCACGAC**CCTAGGATGACCTGAGTACTGTCC | -7.24 |
| 9 | GGGCCGTTCGAACACGAGCATG**GTGGGTGGTGGC**CCTAGGATGACCTGAGTACTGTCC | -6.08 |
| 10 | GGGCCGTTCGAACACGAGCATG**CAGCGTA**CCTAGGATGACCTGAGTACTGTCC | -4.95 |
| 11 | GGGCCGTTCGAACACGAGCATG**GGGGTTGCA**CCTAGGATGACCTGAGTACNTCC | -5.01 |
| 12 | GGGCCGTTCGAACACGAGCATG**CGGG**CCTAGGATGACCGAGTACTGTCC | -6.38 |
| 13 | GGGCCGTTCGAACACGAGCATG**GCAA**CCTAGGATGACCTGAGTACTGTCC | -5.15 |
| 14 | GGGCCGTTCGAACACGAGCATG**CATCAC**CCTAGGATGACCTGAGTACTGTCC | -4.26 |

Table 1. The sequences of anti-VEGF aptamer clones

| No. | Sequences of anti-HER2 aptamers | $\Delta G^o$ |
|---|---|---|

40

| | | |
|---|---|---|
| 1 | GGGCCGTTCGAACACGAGCATG**GGCGGG**CCTAGGATGACCTGAGTCTGTCC | -7.23 |
| 2 | GGGCCGTTCGAACACGAGCATG**GGG**CCTAGGATGACCTGAGTACTGTCC | -5.24 |
| 3 | GGGCCGTTCGAACACGAGCATG**GCGGGT**CCTAGGATGACCTGAGTACTGTCC | -5.47 |
| 4 | GGGCCGTTCGAACACGAGCATG**GGG**CCTAGGATGACCTGAGTACTGTCC | -5.24 |
| 5 | GGGCCGTTCGAACACGAGCATG**GTGC**CCTAGGATGACCTGAGTACTCC | -4.56 |
| 6 | GGGCCGTCGAACACGAGCATG**GTGCGTGGA**CCTAGGATGACCTGAGTACTGTCC | -5.80 |
| 7 | GGGCCGTTCGAACACGAGCATG**GGG**CCTAGGATGACCTGAGTACTGTCC | -5.24 |
| 8 | GGGCCGTTCGAACACGAGCATG**ATA**CCTAGGATGACCTGAGTACTGTCC | -4.02 |
| 9 | GGGCCGTTCGAACACGAGCATG**ATA**CCTAGGATGACCTGAGTACTGTCC | -4.02 |
| 10 | GGGCCGTCGAACACGAGCATG**GTGCGTGGA**CCTAGGATGACCTGAGTACTGTCC | -5.80 |
| 11 | GGGCCGTTCGAACACGAGCATG**GTG**CCTAGGATGACCTGAGTACTGTCC | -4.56 |
| 12 | GGGCCGTTCGAACACGAGCATG**GGTGTGACA**CCTAGGATGACCTGAGTACTGTCC | -6.14 |
| 13 | GGGCCGTTCGAACACGAGCATG**GGCGGG**CCTAGGATGACCTGAGTCTGTCC | -7.23 |

Table 2.  The sequences of anti-HER2 clones

The multiple sequence alignment was performed in MAFFT server in order to have an idea on the sequence similarity. According to Figure 21, both anti-VEGF and anti-HER2 aptamer sequences are significantly diverse both in terms of window length (the sequence between the primers) and nucleotide composition. From this data, the sequence number 9 from anti-VEGF aptamers and the sequence number 10 from the anti-HER2 aptamer were selected for binding assays. Free energy values (-$\Delta G$) were also calculated. Since it was not economically reasonable to do binding assays for all of them, only one of the aptamers from each pool was chosen for further analysis.



41
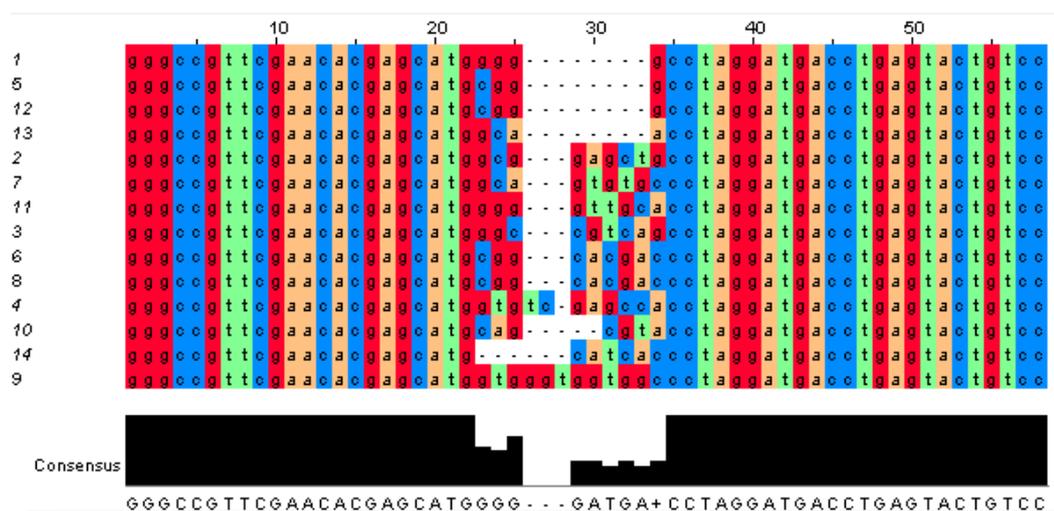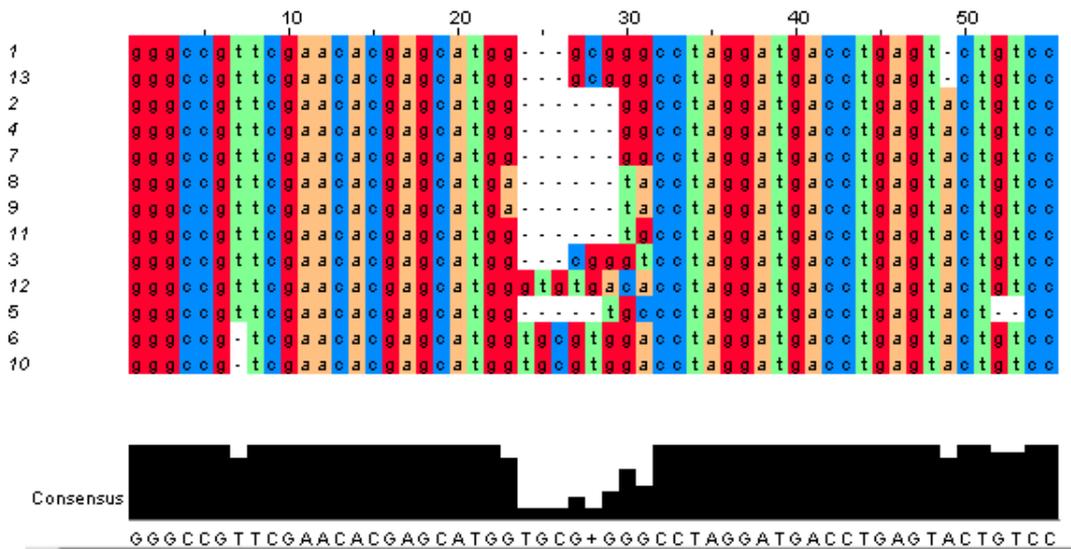
Figure 21. The multiple alignments of: A) anti-VEGF and B) anti-HER2 aptamers.

Phylogenic tree representation of the sequences was performed using PID (percentage identity) in JALVIEW. The aim of building the phylogenic tree was to create an idea about the relation of the aptamer sequences between each other. In this case, only one aptamer candidate was chosen for binding assays from anti-VEGF and anti-HER2 group, respectively. This phylogenic representation would be more helpful if more than one candidate was chosen for binding assay. In that case, choosing the most diverse ones according to percentage identity would give a better idea for further analysis. As it can be seen from the number of branches in each tree, the sequences of anti-VEGF apamers (Figure 22) were more diverse than the anti-HER2 sequences (Figure 23).
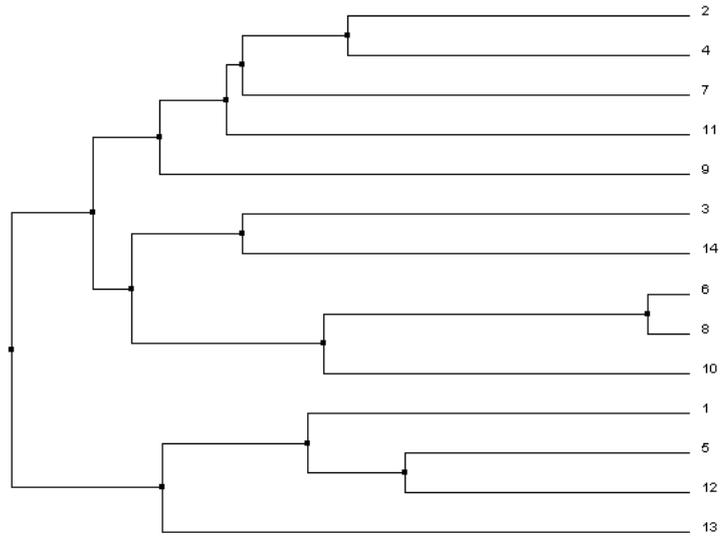
Figure 22.  Phylogenic tree representation of anti-VEGF aptamers.
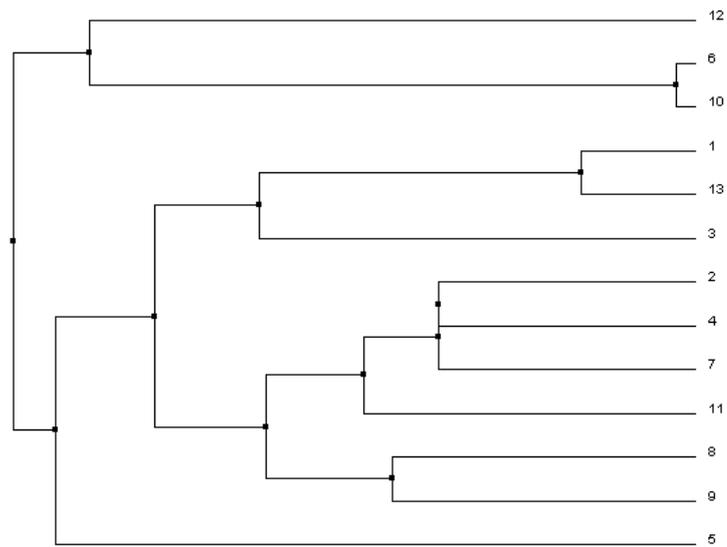


Figure 23. Phylogenic tree representation of anti-HER2 aptamers.

Figure 24 shows the reoccurrence of each random sequence within the set window of <50~60 nucleotides. In both cases, as it was verified above by multiple alignment and phylogeny representation, the sequences were well diverse and the number of reoccurred sequences was constrained. Since the highest number of reoccurrence was 2 among 14 sequences for anti-VEGF aptamers and 3 among 13 sequences for anti-HER2, this information was not very helpful for sequence selection, aptamers. Each

sequence has a similar probability for being the most specific candidate to bind to the target protein.
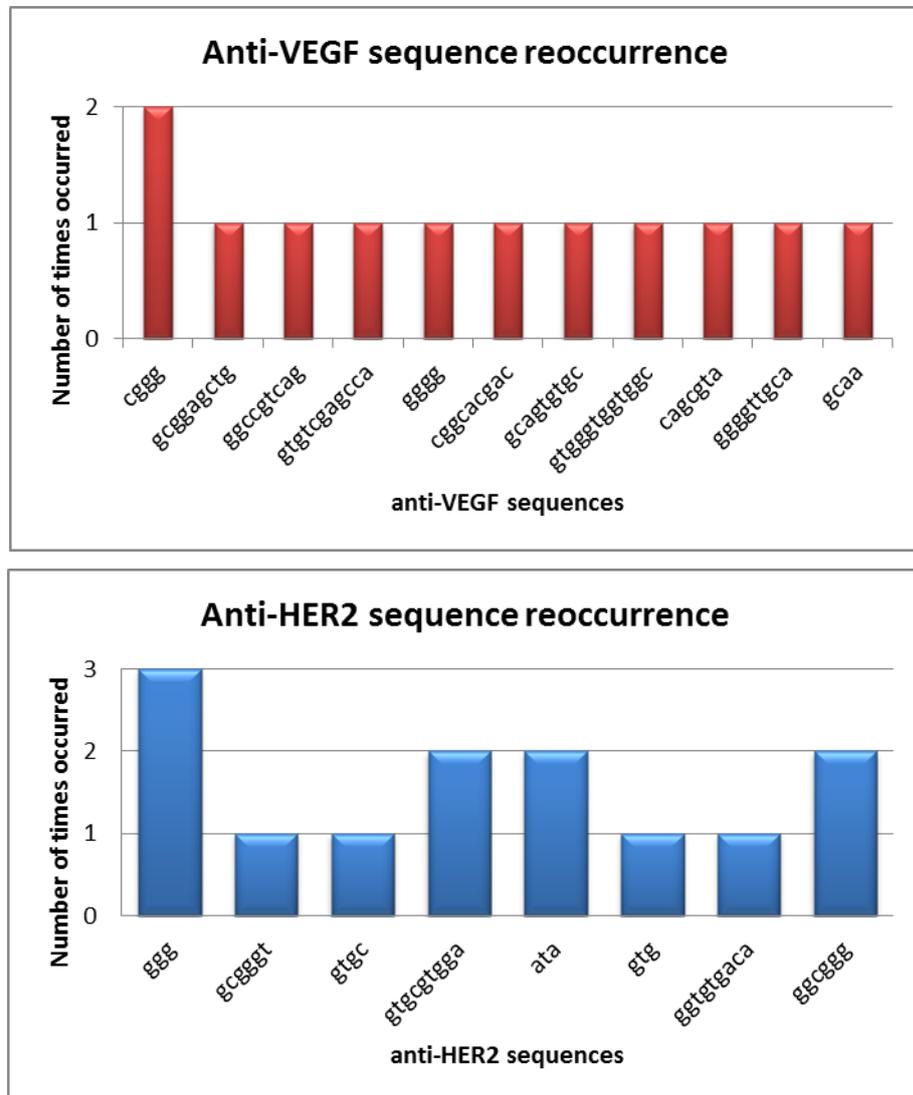


Figure 24. Graphical representation of sequence reoccurrence for A. anti-VEGF and B. anti-HER2.

### 4.5.1. Secondary Structure Analysis

The secondary structures of the obtained sequences were calculated using Mfold server for DNA folding, which considers the ionic content of buffer and the temperature information. Here, the temperature was set as 25 °C, $Na^+$ concentration as 50 mM and $Mg^{++}$ concentration as 1 mM according to the binding buffer content. The aim of building the secondary structure was to have an idea of how the ssDNA folds in its

secondary structure. The aptamer candidates selected in this studied is highlighted by encircling in **Error! Reference source not found.**. As it can be observed from **Error! Reference source not found.**, all the secondary structures contained stem-loops in their structure.
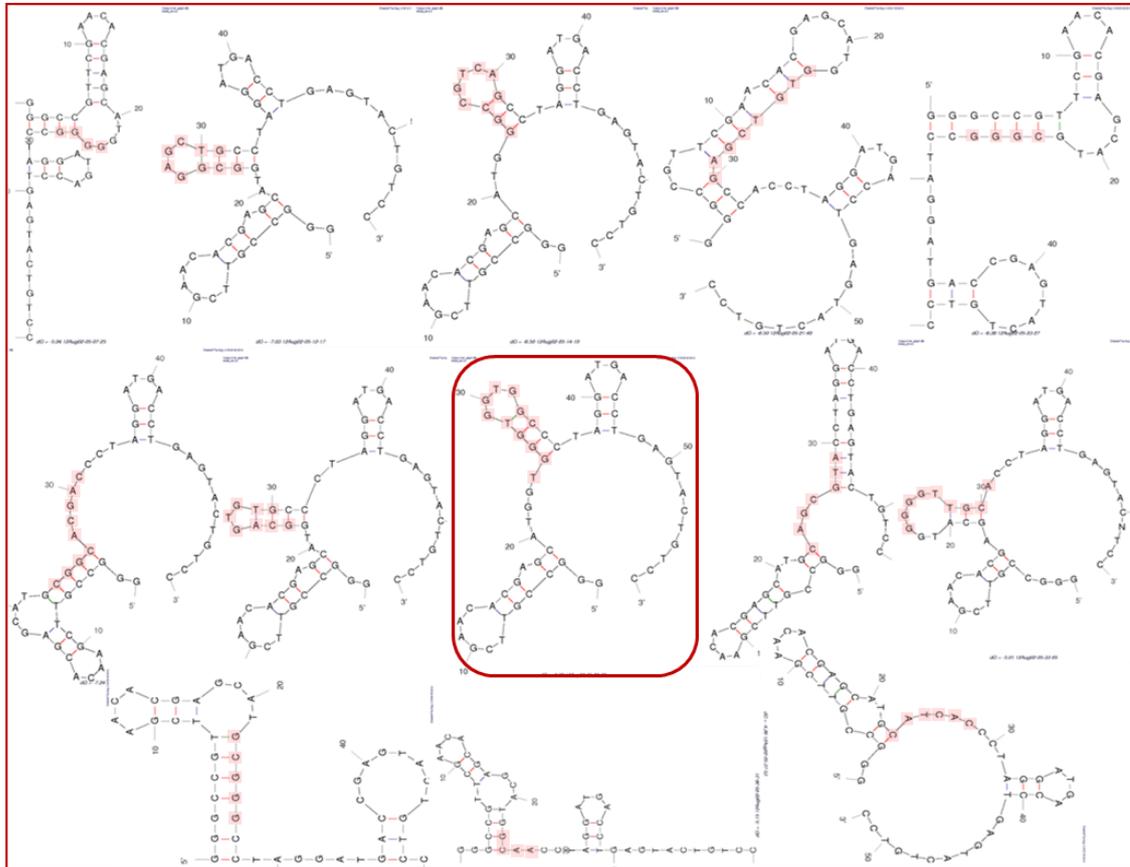


Figure 25. Secondary structures of anti-VEGF sequences. Window sequences are in pink.
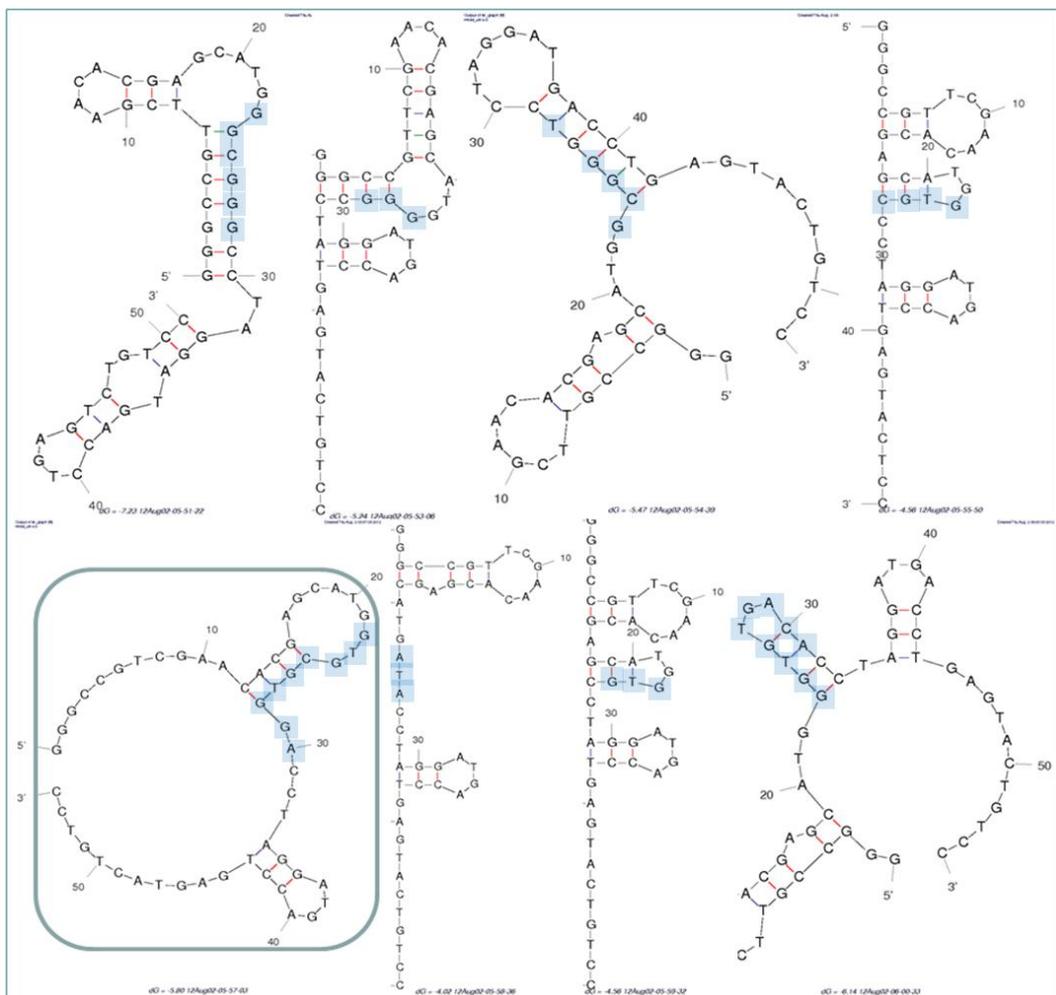
Figure 26. Secondary structures of anti-HER2 sequences. Window sequences are in blue.

## 4.6.  Binding Assay

Binding assay was performed to test the candidate aptamer for affinity and selectivity toward the target protein. The fluorescence measurements were made using fluorospectrometer to probe the changes occurred before and after binding of fluorescein labeled aptamers to target protein. Equation 4.1 was applied to normalize the fluorescence data obtained from binding of aptamers to target.

$$\% \ change \ in \ fluorescence \ = \ Fo - \frac{F}{Fo} \times 100 \qquad (4.1)$$

where $f_0$ is the initial fluorescence and $f$ is the fluorescence after binding. The target VEGF and HER2 proteins were used to test the affinity of the anti-VEGF and anti-

HER2 aptamers respectively, while BSA (Bovine Serum Albumin) protein was used to test the selective ability of these aptamers compared to target protein. Figure 27 first shows an increase in the relative fluorescence value with an increase in target (VEGF) concentration up to 500 nM and attains a saturation. This is probably because all of the available aptamer binding sites are occupied and thus no further binding occurs. This data was then used to calculate $K_D$ (dissociation constant) by fitting the data to the one site bind formula using equation 4.2.

$$\frac{Bmax \ [L]}{[L] + Kd} \tag{4.2}$$

where $B_{max}$ is the maximum number of binding sites, [L] is ligand concentration and $K_d$ is the dissociation constant. As a result the $K_d$ was 315 nM Figure 28. On the other hand, anti-VEGF aptamer showed high selectivity by showing no binding to BSA.
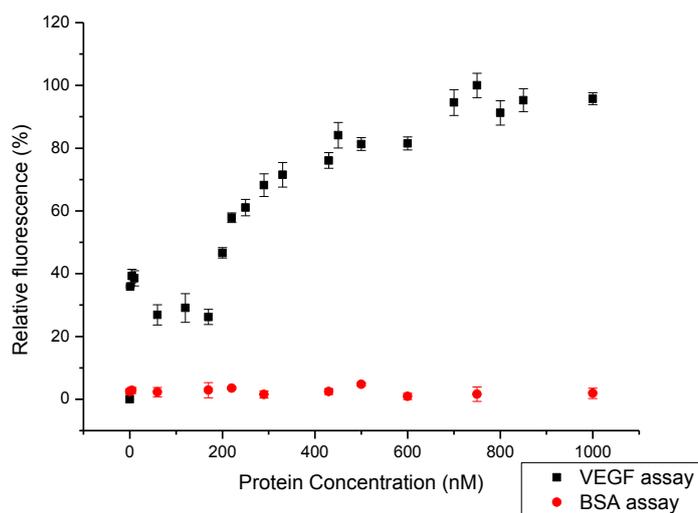


Figure 27. Relative fluorescence (%) versus protein concentration for VEGF (black) and BSA (red) binding assays.

| Equation | y = Bmax * x / (k + x) | | |
|---|---|---|---|
| Adj. R-Squ | 0,71123 | | |
| | | Value | Standard Er |
| B | Bmax | 132,006 | 21,35794 |
| B | k1 | 314,924 | 128,23912 |

Figure 28.  One Site Bind Fit of Relative fluorescence (%) vs protein concentration data.

Anti-HER2 binding data did not show the same success as seen with anti-VEGF aptamers (Figure 29). Although there was a weak binding showing some affinity for HER2 target, BSA binding data showed that anti-HER2 is not selective toward its target. For this reason, the binding data of anti-HER2 was not fitted to one site binding.
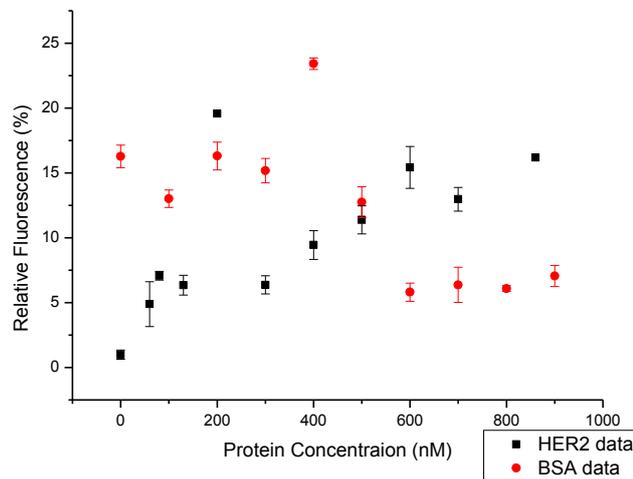


Figure 29.  Relative fluorescence (%) versus protein concentration for HER2 (black) and BSA (red) binding assays.

## 5.  DISCUSSION

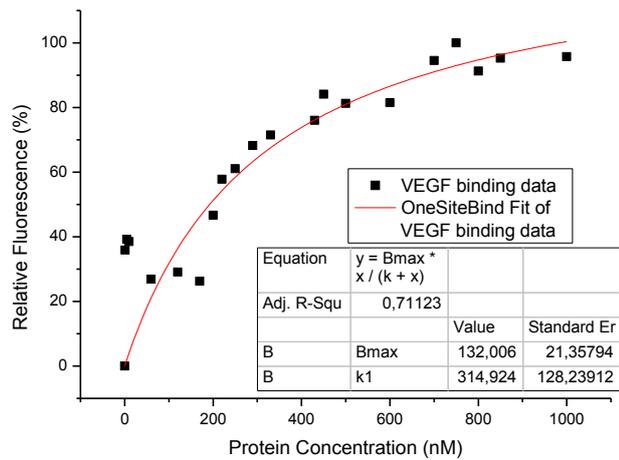The aim of this thesis was to select ssDNA aptamers that bind to two of the most important cancer biomarkers which are VEGF and HER2. The hypothesis was that these aptamers bind to respective target proteins with high affinity and selectivity, and inhibit their function. Overexpressed VEGF and HER2 help tumor progression, mainly by enhancing cell proliferation and vascularization. Inhibition of these two molecules may lead to cancer prevention (21). For this reason the selected aptamers are assumed to be used for cancer treatment. Even though there have been many attempts to block these molecules in cancerous cells by means of antibodies, still aptamers are expected to be more successful. Several characteristics like stability, easiness to be selected and produced, cost and possibility of structure and function modifications, makes aptamers advantageous over antibodies.  There are already some examples like Macugen, an approved drug, where anti-VEGF aptamer was used to target VEGF in macular degeneration.

In this study VEGF(A)165 and the extracellular domain of HER2 protein were chosen for the in vitro selection process. VEGF-A which was used here, is a member of 6 structurally related proteins, and it is the one which plays the most important role in angiogenesis. VEGF-A itself has many isoforms and VEGF165 which is one them, was chosen for aptamer selection. The reason for choosing VEGF165 was that it is the predominant isoform of the two which are secreted into the circulation system in tumor conditions. This feature of VEGF165 is especially important for sensing the presence of tumor by using only blood samples. This works by measuring the signal of VEGF165

49

binding to another specific molecule with the help of biosensors (anti-VEGF might be a good candidate).

The same strategy of protein selection was followed for HER2. There are many strategies for blocking the HER2 function. From the interaction with the available drugs, it has been observed that the most efficient inhibitors are the molecules which block the extracellular region and prevent the phosphorylation of the tyrosine kinase domain precluding the activation of the pathway. With the assumption that the selected aptamer inhibits HER2 function upon binding, the plan was to select an aptamer that binds to HER2 extracellular region. This seems to be the most plausible solution considering that purified protein was used as a target.

Incubation of the target with the ssDNA library and the separation of the unbound ssDNA is often a challenge and an important issue during selection. This was solved by use of magnetic beads. They are commercially available and are provided together with the validated protein immobilization procedures. The magnetic beads are round in structure and rough in surface. This was also verified by SEM analysis of magnetic beads samples (Figure 16). Despite the clear images of the beads' structure this analysis was not sufficient to discriminate between the protein coated and uncoated beads and, ssDNA treated and untreated samples. The reason for that is the limit in magnification which is not high enough to visualize the tiny structures of protein and ssDNA which are a few nm.

The proteins were immobilized on the surface of the beads by the active carboxylic groups which interact with the amine groups of the protein. So, a stable amide bond mediates the link between the bead and the protein. The magnetic property of the beads made it easy for washing the sample after ssDNA incubation without losing the protein but only the unbound ssDNA. The protein - ssDNA pool was performed in binding buffer, which has salt content and pH similar to body fluid. This was also an important choice which is linked to the future perspectives of this work. The secondary structure of the ssDNA changes according to pH, temperature and salt concentration. An aptamer which is selected in such binding conditions would behave in a very similar way when injected to the blood or when a sample of blood is dropped on it. This serves the future perspectives to use the selected aptamers as potential drugs or as biosensing

tools. On the other hand, the detergent (Tween20) was used to provide stringency to ssDNA binding on the surface of the target.

Information on the recovery of ssDNA after incubation with the target gave an idea on binding of specific aptamers. The recovery after the first cycle was too low. At the beginning of SELEX a highly diverse and large ssDNA pool was introduced to the target and it was observed that the ratio of specific to total ssDNA was low. This was expected because only a few from all the library will bind specifically to the target, which is actually a good sign. The specific to total ssDNA ratio was expected to increase in the following cycles because after each selection the bound ssDNA is amplified. With the implementation of the negative selection with other proteins, and the stringency conditions incorporated in each cycle, more specific strands are enriched. The ssDNA recovery record which was according to the expectations, was an important data to show that the enrichment is going in the right way. The ssDNA pool recovery was also expected to increase fast in the first three cycles and more slowly in the later cycles.

After amplification with PCR the dsDNA was separated and extracted by means of electrophoretic methods, which are well known and simple to handle. Simplicity in selection is one of the most important advantages of aptamers over antibodies. Here, the importance stands on producing very specific molecules in a short time rather than using complicated methods. Fluorescein labeled forward primer made this even easier. As a result amplification with fluorescent primer, the forward strand was fluorescently labeled, whereas the reverse strand not. The labeled dsDNA was easy to distinguish and excise. It also played an important role for strand separation in denaturing polyacrylamide electrophoresis procedures (denaturing PAGE). The labeled strand appeared to be larger in size and migrated more slowly than the reverse strand. The fluorescently labeled band facilitated band excision without mixing with the reverse one. On the other hand in highly denaturing conditions such as 7 M urea and 20% formamide it is highly improbable to have dsDNA.

The same process of selection, amplification and separation was repeated for 7 rounds for anti-VEGF and 8 rounds for anti-HER2 selection. Here the aim was to enrich the most specific strands by repeating the cycles, and employing negative selection and stringency conditions in each cycle. For this reason 7/8 cycles were considered

sufficient for selection of candidate aptamers. At the end of SELEX procedure the dsDNA pool was cloned in *E. coli* cells. With the plasmid digestion and selection of positive clones in agarose gel, it was possible to predict the size and separate the false positives. The samples showing the band at around ~60 base pairs were selected for further sequencing.

The sequencing results showed that the pool of candidate aptamers was highly diverse in terms of both content and length of the sequences. The VEGF sequences varied from 50 to 58 bases whereas the HER2 sequences varied from 49 to 55 bases. These results were expected and scientifically possible. The ssDNA library was composed of 86 bases long candidates, which contained 40 bases of random region flanked by constant regions 46 bases long in total. During amplification the reverse primer binds to the constant region flanking to the right of the random region. On the other hand the forward primer has to bind to any complementary sequence within the random region. With this information in mind it was obvious that the selected aptamers would be at most around 60 bases long. The size reduction happens after the amplification step during the first cycle. The obtained sample became the selection pool for the second cycle. In the following cycle there is no size reduction, but the sample obtained in the first amplification is selected and enriched until the end of SELEX.

At this point choosing the best candidates for further analysis is a challenging step. What is more important, due to financial issues only one candidate should be selected for binding assays. This means that the sequence information is the only source to build a reasonable explanation and choose the candidate that might bind with the highest affinity and specificity. The best way to do that was to classify the candidates according to their sequence identity and choose the representative candidate for the sequence that appears the most. This lies on the assumption that a sequence that appears more frequently in the pool is the one that is enriched the most and is the preferred one for binding. Fist the limited number of the sequenced clones was a challenge, but most importantly the diversity of the sequences was the main obstacle. Anyway some attempts were made to make use of any information that could be extracted from these sequences.

First the sequences of anti-VEGF and anti-HER2 where aligned separately. The aim was to have a clear understanding on how these similar these sequences are to each

other. Anti-VEGF sequences were had sequence similarity at the beginning of the random region, while anti-HER2 sequences showed sequence similarities both at the beginning and at the end of the random region. The short random regions and the gaps formed due to the diversity in length were two factors which prevented from arriving to a conclusion. On the other hand it is obvious that sequence conservation in the constant regions is not meaningful in terms of specific binding because they were used for amplification.

Second, the sequence alignment data was used to build the phylogenic trees of the sequences. Phylogenic trees give information on how the sequences relate to each other and group them accordingly. In this way a kind of classification or grouping of the sequences is done. From this representation of the data it was observed that anti-VEGF sequences were more diverse than anti-HER2 sequences. However both pools were separated into two main groups of classification. This would be an important data if two aptamers were chosen for binding assays. In that case one aptamer would be chosen from each group and the affinity binding results would be compared.

Thirdly, an analysis of the sequence reoccurrence was performed. It was shown that only one anti-VEGF sequence appeared twice. On the other hand half of the anti-HER2 sequences appeared twice and one sequence appeared three times. Since the members in one sequence pool appear similarly, it was hard to choose the 'preferred' sequence that binds to each target. In this case it can be predicted that these different sequences bind to the respective target with similar affinity.

Finally the sequences were subjected to secondary structure formation given the ionic concentration similar to the selection conditions. In the literature there are some cases in which it was shown that the secondary structure motif rich sequences confer higher affinity for their target. One assumption is that increase in the secondary structure motifs like loops, results in higher interaction sites with the target protein. However not all the motifs in a secondary structure serve for binding. Some of the motifs / sequence serve for supporting the functional motifs. Another factor is the free energy. Thermodynamic stability which is measured by change in free energy measures the resistance of aptamers to mutations and environment. From analysis of different aptamers it was observed that constant regions commonly do not have high impact in the aptamer structure. There is an exception for the aptamers with short random regions,

which is the case for most sequences here. In aptamers with shorter random region, the constant sequences have higher impact to the thermodynamic stability of the aptamers when compared to others. It has to be mentioned that the secondary structures are taken into consideration with the assumption that aptamer are the ones that undergo adaptive conformation upon binding. It is also true that in some cases it is the target protein that undergo adaptive binding, but this case is not considered here.

Another plan was protein – aptamer candidates docking analysis, which was actually the first one. The docking servers require pdb files of both aptamer and the protein. It was not possible to form the pdb files of the former. To my concern up to date there was not any bioinformatic tool that could construct the pdb file of the ssDNA taking in consideration the secondary structure motifs which form upon target binding. However it would be a crucial information since all the candidates would be analyzed for binding to their respective target.

Finally one aptamer candidate was chosen from anti-VEGF and anti-HER2 sequence pools respectively. Number 9 anti-VEGF and number 10 anti-HER2 sequences where the chosen ones. Number 9 anti-VEGF sequence was 58 nucleotides long and the longest sequence among the others. It had a free energy which relatively low when compared to other sequences and it forms secondary structure motifs. From the sequence information it was easy to distinguish the possible formation of a G-quartet, which corresponds to the random region. Such kind of quadruplex structures are encountered in promoter regions, the structure of some aptamers (eg. Thrombin aptamer), etc. The presence of the G-quartet does not make this candidate automatically the best one, but it is an advantage over other candidates. Whenever G-quartet is found in a structure (to my knowledge) it appears to be functional and increase the stability. These structures are more stable that other structures like hairpins, loops, etc. which are commonly seen in the aptamer structure. However it is not possible to see this structure from the bioinformatic tool used which shows only stem and loop formation. In this version also a loop is formed by the sequences in the random region. This sequence appeared only once.

Number 10 anti-HER2 sequence was also the longest sequence in the respective pool and 55 nucleotides long. It appeared twice and also has a free energy which is relatively low when compared to others. The secondary structure is composed of two

stem-loops and the random regions lies to one side of one of the stem-loops. This structure contained a GTG motif which appeared in 40% of the candidates at the side of stem-loops.

The choice of these two aptamers for further analysis was more an assumption for high affinity binding than rational explanation. It is usually hard to give a decision before performing binding assays. It should be mentioned that there were some motifs in the secondary structure most candidates that were common. These common motifs were formed in the constant regions, but this cannot be used as rationale for choosing the best candidate. One reason is that the same constant regions were used for aptamer selection of both targets and these motifs appear in both pools. Even if these secondary structures were the ones that confer affinity, it would be unlikely that they show affinity for two completely different targets. Even if this too is the case, the resulting aptamers would not be specific at all, which does not serve the final goal of this project. One explanation could be that these constant regions support the real binding structures. The similar free energy values can support this assumption. This is also supported by the observation that constant region participate more in the structure of aptamers with shorter random region and affect their thermostability more when compared to other aptamers. On the other hand from the observations also tell that in most of the cases there are the random regions that participate in the functional structure, not the constant regions (39).

The chosen candidates were the longest sequences with the random regions forming a complete secondary structure motif. It was previously mentioned that constant regions participate more in the secondary structures of the aptamers with the shortest random regions. For the reasons mentioned above it was a better idea to select the aptamers with the lowest participation of the constant regions. However it is not known how much difference makes a random region of 12 nucleotides from 3 nucleotides. Secondly, in order to decrease the production cost and to find the sequences that confer affinity to the target, truncation is applied. Since the random region of these aptamers forms a secondary structure motif, after truncation it will be easier to understand whether it is the random region or the constant region that plays role in the functional structure of the selected aptamers. To conclude with the best candidate selection part, it is worth saying that all the candidates that are sequences must confer

affinity to the respective targets. It would be highly improbable for the bad binders to pass the stringency test.

Binding assays of the selected aptamers with target protein were performed using change in fluorescence intensity after incubation of the two. The fluorescence property of the aptamers makes it easier to detect the unbound ssDNA. The same conditions as during the selection were used for binding assay measurements. The concentration of the ssDNA was kept constant while the concentration of the protein was increased up to 860 nM. As the aptamers bind to the target protein, the fluorescence intensity in the supernatant decreased. In order not to disturb the equilibrium reaction between the aptamer and target protein, the target coated magnetic beads were gently attracted to the walls of the tube and the fluorescence of the supernatant was measured. For binding curve construction, the data given were in terms of aptamer-target binding. In other words, the y-axis values were increasing until the plateau was reached. So, the decreasing fluorescence data was reconstructed by subtracting the ratio of fluorescence reading of any protein concentration to the fluorescence reading of the reference by 1. There was seen an increase and then no change in the fluorescence signal for anti-VEGF, but the change in fluorescence signal of anti-HER2 was not significant. Moreover anti-HER2 failed the selectivity test, but anti-VEGF showed no binding to BSA.

The results of the anti-VEGF binding to VEGF were fitted according to one to one binding. The dissociation constant was calculated to be 315 nM. This is still not the best binding affinity of the aptamer-protein binding but the affinity can be increased by structure modification of the aptamer. The G-quartet structure of the anti-VEGF is promising in this aspect. Regarding the anti-HER2 aptamer it is not completely clear why it does not specificity for the target protein.

## 5.1. Future Perspectives

The objective of this study was to select high affinity and specificity aptamers that bind to VEGF and HER2 target proteins. Introducing a new pool and other stringent conditions there exists the possibility to obtain such aptamers. However several challenges were faced during the aptamer selection process. Work with short aptamers (around 55 bases), gel extraction, separation of ssDNA and lack of resources were some

of the difficulties encountered during the process. On the other hand use of magnetic beads for separation of the bound to unbound ssDNA and implementation of negative selection were two strategies that made the work easier to select more successful candidates. As a result a number of successful candidates were obtained. They were successful because the selected aptamers were enriched for at least 7 cycles and passed through stringency conditions for binding. The chosen anti-VEGF aptamer showed a relatively good binding and was selective with respect to BSA. On the other hand the chosen anti-HER2 was not a good candidate with respect to affinity and selectivity, but still these measurements should be compared with results from other methods of binding affinity measurements (e.g. SPR). The same thing can be said for anti-VEGF.

Next it would be a good idea to perform the binding analysis for at least one additional aptamer and compare. Either the candidates with the shortest random region or the candidates that have secondary structure from constant regions only could be chosen. Regarding further analysis with the chosen aptamers, after verification with at least one other method truncation and/ site mutagenesis can be applied to determine the functional sites in the aptamer structure. Modifications can be applied to first increase the affinity and add other properties to the aptamer. Meanwhile binding tests with other non-target proteins can be performed. Then NMR analysis of the best aptamer – target protein complex would provide valuable information.

## 6. BIBLIOGRAPHY

1. Cancer types. (2012). Retrieved July 25, 2012, from
   http://www.cancer.org/Cancer/ShowAllCancerTypes/index

2. Marusyk, A. and K. Polyak (2010). "Tumor heterogeneity: Causes and consequences." Biochimica Et Biophysica Acta-Reviews on Cancer **1805**(1): 105-117.

3. Hanahan, D. and R. A. Weinberg (2000). "The hallmarks of cancer." Cell **100**(1): 57-70.

4. Hanahan, D. and R. A. Weinberg (2011). "Hallmarks of Cancer: The Next Generation." Cell **144**(5): 646-674.

5. Pang, R. W. and R. T. Poon (2006). "Clinical implications of angiogenesis in cancers." Vasc Health Risk Manag **2**(2): 97-108.

6. Gerstner, E. R., D. G. Duda, et al. (2009). "VEGF inhibitors in the treatment of cerebral edema in patients with brain cancer." Nature Reviews Clinical Oncology **6**(4): 229-236.

7. Bergers, G. and L. E. Benjamin (2003). "Tumorigenesis and the angiogenic switch." Nature Reviews Cancer **3**(6): 401-410.

8. Meric-Bernstam, F. and M. C. Hung (2006). "Advances in targeting human epidermal growth factor receptor-2 signaling for cancer therapy." Clinical Cancer Research **12**(21): 6326-6330.

9. Baselga, J. (2002). "A new anti-ErbB2 strategy in the treatment of cancer: prevention of ligand-dependent ErbB2 receptor heterodimerization." Cancer Cell **2**(2): 93-95.

10. Aertgeerts, K., R. Skene, et al. (2011). "Structural analysis of the mechanism of inhibition and allosteric activation of the kinase domain of HER2 protein." Journal of Biological Chemistry **286**(21): 18756-18765.

11. Arteaga, C. L., M. X. Sliwkowski, et al. (2012). "Treatment of HER2-positive breast cancer: current status and future perspectives." Nature Reviews Clinical Oncology **9**(1): 16-32.

12. Rubin, I. and Y. Yarden (2001). "The basic biology of HER2." Annals of Oncology **12 Suppl 1**: S3-8.

13. Baselga, J. and S. M. Swain (2009). "Novel anticancer targets: revisiting ERBB2 and discovering ERBB3." Nature Reviews Cancer **9**(7): 463-475.

14. Prenzel, N., O. M. Fischer, et al. (2001). "The epidermal growth factor receptor family as a central element for cellular signal transduction and diversification." Endocr Relat Cancer **8**(1): 11-31.

15. Menard, S., E. Tagliabue, et al. (2000). "Role of HER2 gene overexpression in breast carcinoma." J Cell Physiol **182**(2): 150-162.

16. Baselga, J. and S. M. Swain (2009). "Novel anticancer targets: revisiting ERBB2 and discovering ERBB3." Nature Reviews Cancer **9**(7): 463-475.

17. Ghosh, R., A. Narasanna, et al. (2011). "Trastuzumab Has Preferential Activity against Breast Cancers Driven by HER2 Homodimers." Cancer Research **71**(5): 1871-1882.

18. Engel, R. H. and V. G. Kaklamani (2007). "HER2-positive breast cancer: current and future treatment strategies." Drugs **67**(9): 1329-1341.

19. Cserni, G., M. Francz, et al. (2011). "Estrogen Receptor Negative and Progesterone Receptor Positive Breast Carcinomas-How Frequent are they?" Pathology & Oncology Research **17**(3): 663-668.

20. de Ruijter, T. C., J. Veeck, et al. (2011). "Characteristics of triple-negative breast cancer." J Cancer Res Clin Oncol **137**(2): 183-192.

21. Konecny, G. E., Y. G. Meng, et al. (2004). "Association between HER-2/neu and vascular endothelial growth factor expression predicts clinical outcome in primary breast cancer patients." Clinical Cancer Research **10**(5): 1706-1716.

22. Stimpfl, M., D. Tong, et al. (2002). "Vascular endothelial growth factor splice variants and their prognostic value in breast and ovarian cancer." Clinical Cancer Research **8**(7): 2253-2259.

23. Ho, Q. T. and C. J. Kuo (2007). "Vascular endothelial growth factor: biology and therapeutic applications." Int J Biochem Cell Biol 39(7-8): 1349-1357.

24. Sweeney, C. J., K. D. Miller, et al. (2003). "Resistance in the anti-angiogenic era: nay-saying or a word of caution?" Trends in Molecular Medicine **9**(1): 24-29.

25. Folkman J. In: DeVita VT Jr, Hellman S, Rosenberg SA, eds. Cancer: Principles & Practice of Oncology. Vol 2. 8th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008: ch 8 103-116.

**26.** Understanding Angiogenesis and the VEGF Ligand. (2012). Retrieved July 28, 2012, from
http://www.biooncology.com/research-education/vegf/ligand/index.html

27. Ferrara, N. (2004). "Vascular endothelial growth factor: basic science and clinical progress." Endocr Rev **25**(4): 581-611.

28. Gerber, H. P. and N. Ferrara (2005). "Pharmacology and pharmacodynamics of bevacizumab as monotherapy or in combination with cytotoxic therapy in preclinical studies." Cancer Research **65**(3): 671-680.

29. Holmes, K., O. L. Roberts, et al. (2007). "Vascular endothelial growth factor receptor-2: Structure, function, intracellular signalling and therapeutic inhibition." Cellular Signalling **19**(10): 2003-2012.

30. Ferrara, N., H. P. Gerber, et al. (2003). "The biology of VEGF and its receptors." Nature Medicine **9**(6): 669-676.

31. Goel, S., D. G. Duda, et al. (2011). "Normalization of the vasculature for treatment of cancer and other diseases." Physiol Rev **91**(3): 1071-1121.

32. Hoeben, A., B. Landuyt, et al. (2004). "Vascular endothelial growth factor and angiogenesis." Pharmacol Rev **56**(4): 549-580.

33. Vosseler, S., N. Mirancea, et al. (2005). "Angiogenesis inhibition by vascular endothelial growth factor receptor-2 blockade reduces stromal matrix metal loproteinase expression, normalizes stromal tissue, and reverts epithelial tumor phenotype in surface heterotransplants." Cancer Research **65**(4): 1294-1305.

34. Hermann, T. and D. J. Patel (2000). "Adaptive recognition by nucleic acid aptamers." Science 287(5454): 820-825.

35. Patel, D. J. and A. K. Suri (2000). "Structure, recognition and discrimination in RNA aptamer complexes with cofactors, amino acids, drugs and aminoglycoside antibiotics." Journal of Biotechnology 74(1): 39-60.

36. Mascini, M., I. Palchetti, et al. (2012). "Nucleic Acid and Peptide Aptamers: Fundamentals and Bioanalytical Aspects." Angewandte Chemie-International Edition **51**(6): 1316-1332.

37. Kulbachinskiy, A. V. (2007). "Methods for selection of aptamers to protein targets." Biochemistry (Mosc) **72**(13): 1505-1518.

38. Stoltenburg, R., C. Reinemann, et al. (2007). "SELEX-A (r)evolutionary method to generate high-affinity nucleic acid ligands." <u>Biomolecular Engineering</u> **24**(4): 381-403.

39. Cowperthwaite, M. C. and A. D. Ellington (2008). "Bioinformatic analysis of the contribution of primer sequences to aptamer structures." Journal of Molecular Evolution 67(1): 95-102.

40. Patel, D. J., A. K. Suri, et al. (1997). "Structure, recognition and adaptive binding in RNA aptamer complexes." <u>Journal of Molecular Biology</u> **272**(5): 645-664.

41. Peeters, E., C. Wartel, et al. (2007). "Analysis of the DNA-binding sequence specificity of the archaeal transcriptional regulator Ss-LrpB from Sulfolobus solfataricus by systematic mutagenesis and high resolution contact probing." <u>Nucleic Acids Research</u> **35**(2): 623-633.

42. Goodrich, James A.Kugel, Jennifer F. (2007) *Binding and kinetics for molecular biologists /*Cold Spring Harbor, N.Y. : Cold Spring Harbor Laboratory Press, ch 4 70-95.

43. Jing, M. and M. T. Bowser (2011). "Methods for measuring aptamer-protein equilibria: a review." <u>Analytica Chimica Acta</u> **686**(1-2): 9-18.

44. Goodrich, James A.Kugel, Jennifer F. (2007) *Binding and kinetics for molecular biologists /*Cold Spring Harbor, N.Y. : Cold Spring Harbor Laboratory Press, ch 1 1-17.

45. Pavski, V. and X. C. Le (2003). "Ultrasensitive protein-DNA binding assays." <u>Curr Opin Biotechnol</u> **14**(1): 65-73.

46. Dynabeads® M-270 Carboxylic Acid, 2012, Retreived Nov 10, 2011, from
http://products.invitrogen.com/ivgn/product/14305D

47. Niazi, J. H., S. J. Lee, et al. (2008). "ssDNA aptamers that selectively bind oxytetracycline." <u>Bioorganic & Medicinal Chemistry</u> **16**(3): 1254-1261.

48. TOPO® TA Cloning® Kit for Sequencing *with One Shot® MAX Efficiency® DH5α-T1$^R$ E. Coli*, 2012, Retrieved March 15, 2012, from
http://products.invitrogen.com/ivgn/product/K459501