DNA APTAMERS FOR BIOMARKER DISCOVERY IN OVARIAN CANCER

By

DIMITRI VAN SIMAEYS

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2012

To my Family

ACKNOWLEDGMENTS

I would like to give special thanks to my advisor Dr. Weihong Tan of the department of Chemistry, University of Florida for the support and faith he has shown me and the chances he has provided me throughout my studies at the University of Florida that lead to this dissertation and the development of my career. I also owe my gratitude to Dr. Rebecca Sutphen from the College of Medicine, University of South Florida for her support and knowledge in our collaboration. I would like to thank my committee members Dr. Gail Fanucci, Dr. Nicole Horenstein, Dr. Ben Smith and Dr. Brad Fletcher for their input and constructive feedback that lead to the overall improvement of this dissertation.

I would like to thank the members from the Tan group, with special thanks to Dr. Kwame Sefah and Dr. Dalia López-Colón for teaching me the ropes and SELEX to the best of their ability and for their patience. Dalia's help cannot be underestimated, as she was a defining factor in many aspects of the work I have done on ovarian cancer. I would also like to thank the current and past members of the Tan group, for the many meaningful discussions, friendship and support by Xiangling Xiong, Dr. Jennifer Martin, Elizabeth Jimenez, Dr. Meghan O'Donoghue, Dr. Karen Martinez, Dr. M.-Carmen Estevez, Jun Lui, Kejing Zhang, Sena Cansiz, Guizhi Zhu, Diane Turek, Danny Chang, Dr. Tahir Bayrac, Jing Zheng, Carole Champanhac and Julia Vaizer. I would also like to give thanks to Dr. Pamela Havre for her friendship and guidance.

I would like to thank my parents Filip and Nadia Van Simaeys and my brother Mathieu Van Simaeys, for their support and encouragement and having me in their prayers. Finally I would like to thank all my friends in both Belgium and the US for their support and encouragement.

TABLE OF CONTENTS

page

LIST OF TABLES

LIST OF FIGURES

LIST OF SCHEMES

Scheme page

LIST OF ABBREVIATIONS

17AAG	17-N-allylamino-17-demethoxygeldanamycin
2D LC-MS/MS	2 Dimensional liquid chromatography- mass spectrometry/ mass spectrometry
AJCC	American joint committee on cancer
ATCC	American type cell culture
ATP	Adenosine tri-phosphate
BB	Binding buffer
BRCA	Breast cancer
BSA	Bovine serum albumin
CA125	Carbohydrate antigen 125
CD	Cluster of differentiation
ChIP	Chromatin immuno-precipitation
CPG	Controlled porous glass
CSC	Cancer stem cell
EDTA	Ethylenediaminetetraacetic acid
FACS	Fluorescence assisted cell sorting
FBS	Fetal bovine serum
FDA	Food and drug administration
FITC	Fluorescein isothiocyanate
Freq.	Frequency
HBSS	Hank’s balanced salt solution
HE4	Human epididymis 4
HOP	Hsp70-hsp90 organizing protein
HSP	Heat shock protein

IC50	Inhibition concentration were 50% of the maximum activity is reached
LOD	Limit of detection
M-CSF	Macrophage colony-stimulating factor
MMP	Matrix metalo protease
MTT	3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
NHS	N-Hydroxysuccinimide
OCCA	Ovarian clear cell adenocarcinoma
OCSI	Ovarian cancer symptom index
PBS	Phosphate buffer solution
PCR	Polymerase chain reaction
PE-Cy5.5	Phycoerythrin-cyanine derivative 5.5
PERL	Practical extraction and reporting language
PTK7	Protein tyrosin kinase 7
qPCR	Quantitative polymerase chain reaction
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SELEX	Systematic evolution of ligands by exponential enrichment
siRNA	Small interfering RNA
SOAC	Serous ovarian adenocarcinoma
ssDNA	Single stranded deoxyribonucleic acid
ssRNA	Single stranded ribonucleic acid
StIP1	Stress induced protein 1
Surfactant	Surface-active agents
TCA	Trichloroacetic acid
TVU	Trans vaginal ultrasound
USPSTF	U.S. preventive services task force

WB	Washing buffer

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

DNA APTAMERS FOR BIOMARKER DISCOVERY IN OVARIAN CANCER

By

Dimitri Van Simaeys

August 2012

Chair: Weihong Tan
Major: Chemistry

Ovarian cancer is the most deadly gynecological malignancy. In particular, ovarian clear cell adenocarcinoma shows an especially poor response to the standard treatment. In general, the five-year survival rate of patients diagnosed in Stage I ovarian cancer is about 90%; however, only 2% of all ovarian cancer is diagnosed at this stage. The underlying reasons for this low number are the vague symptoms of this disease and the poor diagnostic tests available. Furthermore, there is a need to differentiate the two most prevalent types of ovarian cancer, clear cell adenocarcinoma and serous adenocarcinoma, which behave different clinically.

To enable physicians to diagnose ovarian cancer at an earlier stage and differentiate clear cell carcinoma from serous carcinoma, in this research a panel of aptamers was selected against clear cell adenocarcinoma and ovarian cancer stem cells. Aptamers are ssDNA or ssRNA oligonucleotides with selectivities and binding affinities similar to antibodies, but with the benefit of *in vitro* generation. A total of nine aptamers have been selected for TOV-21G, an ovarian clear cell adenocarcinoma cell line, and 5 aptamers for ovarian cancer stem cells were selected. The aptamers all show subnanomolar to low nanomolar affinities. While the aptamers selected for TOV-

21G show no binding to HeLa or the serous adenocarcinoma cell line CAOV3, there was binding to CEM, A172, HCT-116 and HL60 cells, suggesting that these unrelated cell lines have the same surface molecular markers.

Furthermore, a strategy for aptamer target identification was outlined and used to determine that the molecular target of the aptamer TOV6 is Stress-induced Protein 1, a co-chaperone that is known to interact with heat shock protein 90. The target was validated by siRNA silencing, a protein blot on the recombinant protein and antibody binding. Stress-Induced Protein 1 is expressed in several other ovarian cancer cell lines, including SKOV3, OVCAR3, TOV112D and C13.

Studies of TOV-21G and aptamer TOV6 indicate that TOV6 is a potent inhibitor of invasion of the basement membrane by binding to Stress-induced protein 1. The aptamer also has cytostatic properties.

CHAPTER 1
BACKGROUND

Cancer

Cancer is a group of diseases in which abnormal cells are growing and spreading in an abnormal fashion. Carcinogenesis, the initiation of cancer, can be caused by external factors, internal factors or a combination of both. Predictions in 2011 projected that 1,596,670 people would to be diagnosed with cancer and 571,950 people would die of cancer. In 2011, one in four deaths will be caused by cancer[1].

Over the past century, remarkable progress has been made in the fight against various diseases, as can be seen in Figure 1-1. The leading cause of deaths in the US, heart disease, has seen over a 50% reduction from 1950 to 2001. Similarly, deaths due to cerebrovascular disease were reduced by two thirds in the same period, and deaths caused by the flu and related diseases were reduced by about one-half. However, there has been no progress towards improving the survival rate for cancer. Progress has been made in the last ten years, but with no more then a few percent reduction per year, dependent on the type of cancer[2]. There is a dire need for better treatment, as well as better diagnostic methods[3].

Cancers in which considerable progress has been made towards reducing mortality rates include lung and bronchial cancer (20% reduction 1990-2000), and prostate cancer (45% reduction 1990-2005) in men. Similar figures have been observed in women with regards to stomach and breast cancers. However pancreatic cancer, liver cancer, ovarian cancer and leukemia have not experienced reductions in mortality rates over the past decades[1]. The absence of improvement in the statistics from ovarian cancer indicates that the current methods of fighting this deadly disease

are inadequate, and that more insight in the molecular biology of this type of cancer is needed.

Ovarian Cancer

Ovarian cancer is cancer that begins in the ovaries, with ninety percent attributed to epithelial origin. Epithelial ovarian cancer is classified by its histological subtypes, with serous and clear cell carcinoma being the most predominant subtypes of ovarian cancer[4].

As mentioned above, there has been little progress in the treatment and survival rate of ovarian cancer, with 15,460 people expected to die from this disease in 2012[1]. Symptoms that could indicate the presence of ovarian cancer are vague: pelvic or abdominal pain, increased urinary urgency or frequency, bloating and difficulty eating, all symptoms that can be misinterpreted with other common (non-) gynecological diseases. Symptoms of this type arise slowly and sometimes go unnoticed until the disease has made severe progress. Factors that increase the risk of ovarian disease include obesity, smoking, the use of estrogen as a postmenopausal hormone and certain mutations in the BRCA1 and BRCA2 gene[1].

Ovarian cancer is usually treated with a combination of surgery and chemotherapy. Surgical removal usually results in the complete removal of the ovaries and the fallopian tubes. As can be seen in Figure 1-2, the success of treatment is highly dependent on the stage in which the patient is diagnosed. Given the same treatment, the 5-year survival rate of patients below 65 is nearly double the rate of patients over 65 (57% compared to 29%[4]). Table 1-1 tabulates how ovarian cancer is staged by the American Joint Committee on Cancer (AJCC)[5]. The five-year survival

rate of patients diagnosed at stage I is 90%, but only 2% of the patients with ovarian cancer are diagnosed at that stage[3].

Diagnosis of Ovarian Cancer

Early diagnosis of ovarian cancer is imperative towards the effective treatment of this disease, as illustrated in Figure 1-2. The vague symptoms of the disease make early diagnosis very difficult without additional aid, such as a transvaginal ultrasound (TVU) or serum levels of biomarkers assigned to the presence of ovarian cancer.

Currently, patients complete a 27-item form over a period of several weeks as an initial diagnostic tool, called the Ovarian Cancer Symptom Index (OCSI)[6,7]. If symptoms persist, a TVU is then performed. An OSCI is considered to be positive when the listed symptoms appear more then 12 times in a single month and these symptoms first appeared within the previous year. In addition, the use of longitudinal serum levels of various markers in combination with TVUs is currently being investigated[8,9].

As previously above, a transvaginal ultrasound is an important tool for early detection of tumors or benign swellings[10]. However, in 1996 the U.S. Preventive Services Task Force (USPSTF) recommended against the use of TVU for preventive screening as the test is not sufficiently specific[11] and the cost paid ($550, national average[12]) is not justified, considering the marginal reduction in mortality.

Biomarkers

Biomarkers in Ovarian Cancer

It is thought that the genetic changes causing carcinogenesis can be translated in the release of several proteins, called biomarkers, for example: the sensitivity of the OCSI (combined with transvaginal ultrasonography) can be enhanced when cross-referenced with the patient's CA125 blood levels. CA125 is the only FDA approved

biomarker for ovarian cancer, but it is heterogeneously expressed: Some types of ovarian cancer do not express this glycoprotein, while in some other non-cancerous situations like pregnancy or cirrhosis, elevated levels of CA125 are known to exist in the blood of the patient[13,14]. The use of CA125 alone gives a sensitivity (measures the proportion of truly positive test results) of 48% at 98% specificity (measures the proportion of truly negative test results). When the test for early diagnosis was expanded with more markers, sensitivities of 75% were reached[15]. Other possible ovarian cancer biomarkers include HE4, mesothelin, M-CSF, osteopontin, kallikrein(s), and soluble EGF receptor[9,15].

An Introduction to Biomarkers

"A biomarker is a measurable indicator of a specific biological state, particularly one relevant to the risk of contraction, the presence or the stage of disease"[16]. From a historical point of view, a biomarker could be a physical trait, a measurable biomolecule or a process. In recent years, the word biomarker has become synonymous with a molecular marker that indicates biological state. Biomarkers can take many forms, e.g., karyotype in a cell, methylation profile, the number of circulating cancer cells or copy number on a gene[17-20]. Biomarker can also be a metabolite related to the disease[21].

Protein detection has become the gold standard in the biomarker field[16,22]. Also, protein levels or presence can be easily measured in the blood, enabling researchers to directly correlate the efficacy of novel drugs to the biomarker level[24]. Therefore, considerable effort in the field of proteomics has been focused on discovering novel protein biomarkers that are up- or down regulated in cancer. However, diagnostic protein measurements also face many challenges, such as heterogeneous protein expression, the wide possible concentration range of the putative biomarker in the

blood, or the low abundance of the biomarker in the blood[16]. These problems make protein biomarker discovery difficult and uncertain, and many putative biomarkers have been discarded due to poor analytical precision.

Membrane Proteins

One class of proteins that can prove outstanding as biomarkers are those associated with the membrane[25]. This class of proteins is one of the most understudied and underestimated class studied to day.

Membrane proteins can be divided in two main classes: 1) integral or transmembrane proteins that pass through the membrane multiple times and which are often found as porins or signaling proteins; and 2) peripheral or anchored proteins extend from the membrane, often extracellularly. The majority of this discussion will be on integral proteins, which can be further subdivided in two groups: the ß-barrels and the α-helices. Gene segments coding for α-helices are the most prevalent group, estimated to comprise 20-30% of most genomes[26], with the ß-barrel group representing about 2-3%[27]. Membrane proteins have a wide range of functions: 1) transportation of ions or metabolites across the membrane; 2) transduction of chemical signals from their environment to initiate intracellular responses, 3) propagation of electrical signals, 4) cell attachment and 5) the control of membrane of lipid composition, vesicular transport, and organizing the shape of organelles within the cell itself[28]. To become cancerous, cells often must change their membrane protein footprint to recruit growth factors, invade foreign tissues or evade apoptosis[29-31].

Membrane proteins are in their native states only when intercalated in the membrane, and they are thus often insoluble in aqueous buffers due to sections of non-polar lipid associating amino acids. Thus, methods such as two-dimensional gel-

electrophoresis or 2D LC-MS/MS are required to search for potential biomarkers[32,33]. As these proteins are predominantly found in the membrane, limitations of these analytical techniques make it difficult to obtain a good overview of the exact protein content of a membrane.

There are three characteristic problems associated with membrane protein analysis. First, membrane proteins are generally in low abundance, making it difficult to detect these proteins in standard gels with current extraction methods. This problem can be circumvented by various enrichment techniques[34]. Another problem is the limited aqueous solubility of membrane proteins. In 2D gel electrophoresis, a technique in which a protein sample is separated by iso-electrical focusing, followed by SDS-PAGE, membrane proteins commonly precipitate at their iso-electrical points. Also, due to the limited solubility, surfactants are often used to increase the amount of protein in solution, but these chemicals are not compatible with mass spectrometry, because some ionic surfactants or salts can suppress the ionization of peptides. The final is the alkalinity of membrane proteins and the absence of ionic peptides, in general[35]. Since most 2D gels focus on analyzing proteins with a pI<8, proteins with higher pI's are simply not detected. The LOD of mass spectrometry lays a thousand fold higher for unknown protein detection[40]. Putting all these factors together, membrane proteins are underrepresented in global large-scale proteomics studies[35].

Other Biomarker Discovery Techniques

Alternatives to high-throughput mass spectrometry techniques can also lead to the discovery of novel biomarker leads such as immuno-affinity capture techniques[36] (protein microarrays[37]) or transcriptional profiling methods (e.g., the analysis of transcriptomes in blood for pharmacological analysis[38]).

Immuno-affinity capture is currently considered as the most effective method for the detection and quantification of putative protein biomarkers. These techniques can be used for the detection of proteins at the ng/mL level commonly observed for biomarkers (with the help of signal amplification techniques through enzymatic amplification)[39]. New or unknown biomarkers are difficult to be detected with this approach, as the detection of proteins through immune-affinity require knowledge about the analyte and as the production of antibodies is a lengthy process, many new promising proteins (which have a low incidence due to their novelty) do not always have (good) antibodies available[41]. Therefore, the search for alternative capturing agents has been an important research area. Some alternatives for antibodies can be found in peptoids[42], phage display[43], and in aptamers[44,45].

Aptamers

An Introduction to Aptamers

Aptamers are oligo nucleotides, usually ssDNA or ssRNA. Although researchers have developed some unusual aptamers with a peptide backbone[46], or with an extra unnatural base[47]. The typical length of an aptamer is 20-100 nucleotides, which assume very distinctive and unique tertiary structure[48]. Aptamers were first described by Gold and Tuerk[49] and were also independently developed by Stozak and Ellington[50], who coined the term ”aptamer”. The word is composed of two Greek roots ‘‘apta”, which means “to fit”, and “mer”, which means “more then one”. Aptamers are developed from large random libraries of oligo nucleotides, each containing 20 to 60 random nucleotides flanked by primers sequences (which are needed for amplification by PCR). Target sequences are selected for a target of choice through a process called Systematic Evolution of Ligands by EXponential enrichment (SELEX). Their tertiary

structures allow them to bind to a wide variety of targets, from metal ions[51], to organic molecules[52], peptides[53] and proteins[54]. Often, aptamers are selected for the proteins on the surface of viruses[55], bacteria[56] and eukaryotic cells[57]. Aptamers bind to their targets with specificities and affinities comparable to antibodies.

There are several other differences between antibodies and aptamers that change the ways in which they are used. For *in vivo* applications for example, antibodies have a relatively long half-life, while aptamers, when unaltered, are relatively quickly removed by DNases or RNases and the kidneys[58,59]. However, because of the ease of chemical modification, the half-life can by extended and fine tuned[59,60]. By adding polyethylene glycol segments, the size can also be fine tuned, ultimately affecting how the aptamer is cleared from the system by the first pass effect (renal filtration and liver filtration)[59]. Other possible modifications are to thiolation or methylation of certain hydroxides on the ribose sugars to hamper the hydrolytic action of DNases or RNases[48,61].

Another difference between antibodies and aptamers are in the way they are generated, which allows selection of aptamer against i.e. toxins. In order to make any antibody, the antigen of interest needs to be injected in an animal, after which the T-cell that makes the antibody can be transformed into a hybridoma[62]. SELEX is a procedure that is completely *in vitro* and automatable, which can reduce the development time of an aptamer in the hands of well-trained selectors. This also allows easy scale-up to the commercial level. The affinity of antibodies is also something that is hard to optimize (especially for low incidence antigens[41], which is why it may be challenging sometimes to find the right vendor for an antibody that can bind to the antigen of interest. Aptamers themselves can be designed to bind with a high dissociation constant, and recently it is

even possible to precisely design the selection process to match the affinity constant, in so called “smart aptamers”[63]. It is also assumed that aptamers are less prone to give immunologic responses in time as proteins like antibodies are known to be doing[61].

Many research groups and corporations in the world are actively investigating aptamers. Currently, there is one aptamer on the market, Pegaptanib sodium, for the treatment of age-related macular degeneration[64], and several other aptamers are now being in phase two and three clinical trails[45]. It is estimated that the current aptamer therapeutic market was $10 million in 2009. It is predicted that by this growth rate this number will grow to $1.2 billion in 2014[65].

SELEX

As stated in the introduction, aptamers are generated from large libraries that are prepared by solid phase chemistry[66]. Because each library is consisted of a segment containing a string of 20-40 random nucleotides, flanked by primers for PCR amplification, the library contains in theory 4^{20}-4^{40} (10^{10}-10^{24}) different possible sequences or femto- to millimoles of unique molecules. Since each sequence forms its own unique tertiary structure, one library can be used to identify aptamer candidates for almost any target of interest.

The basic scheme of SELEX is explained in Figure 1-3. In the initial round, usually 20nmol of naïve library is melted at 95ºC and allowed to refold at the binding temperature (usually 4ºC). This ensures that all strands fold into their optimal tertiary structures. This mix is incubated with the target of interest. Non-binding sequences are removed by washing, after which the binding sequences are recovered by denaturing the target. In cell-SELEX, one may choose to do a counter or negative selection against a non-target cell line for additional stringency. However, usually this step is

omitted in the first round in the SELEX process. When negative selection is introduced, only those sequences that do not bind to the negative target are gathered and prepared for amplification. The pool that is retrieved after incubation with the target or counter target is amplified by PCR and prepared for the next round, leading towards the next generation's library. After every couple of rounds, the enrichment is monitored by flow cytometry and when the desired enrichment is reached, the selection is stopped and one or a couple of pools are analyzed for sequence homology. The largest families obtained will be the primary aptamer candidates, which are further tested for binding with the target.

Aptamer Target Recognition Interactions

Aptamers bind to their targets via a combination of possible molecular interactions, resulting in affinities higher than their natural counterparts (riboswitches[68]) and comparable to that of antibodies[45]. There are in principle two kinds of oligonucleotides that can perform target recognition: naturally raised oligonucleotides or ribozymes and *in vitro* selected aptamers. The comparison between these two classes of ligands has shed light on the principles that underlie aptamer-target recognition interactions[67]. Since natural nucleic acids need to perform a certain action as part of a larger network, it is sometimes necessary for different structural motifs to be recognized. However, aptamers are selected to bind to one specific target though. For this reason, aptamers tend to have lower dissociation constants (i.e., higher affinities) than their naturally raised counterparts. It is assumed that when aptamers are freely flowing in solution, they often exist as unpaired disordered loop regions[68]. When the aptamer binds to its target, it assumes a very distinctive structure that wraps in a very specific way around

its target. In some oligonucleotide-peptide binding studies, it has been observed that unpaired nucleotides become flaps that cover the ligand[69].

Thus, an aptamer can bind to its target with high affinity, and also with high specificity. A classic example of this is the inability of the theophylline aptamer to bind to caffeine, because the methyl group on caffeine induces steric hindrance that prevents the aptamer from binding[70]. With this example it is apparent that the binding of aptamers is not merely dependent on the π-π stacking of flat ligands. If that were true, caffeine could very easily bind to the aptamer by intercalation (or to any piece of oligonucleotide for that matter), as do common DNA imagining dyes do; e.g. ethidium bromide[71]. In the example of theophylline and caffeine, the binding of the aptamer with theophylline is governed by hydrogen bonds in a specific S-loop, with further stabilization by π-π stacking[70].

Aptamers fold into controlled conformations that allow binding to their selection targets via a variety of possible intermolecular interactions. Hydrogen bonding is an important factor toward the selectivity of target specific aptamers, and has been illustrated in several examples of aptamer/ small molecules interactions (e.g., arginine, adenosine mono phosphate, flavin mononucleotide)[72-74]. The increased complexity of proteins allows for multiple types of interactions, including hydrogen bonding, π-π stacking, shape conformations and electrostatic interactions[75-77]. It is especially important to consider electrostatic interactions when selecting for aptamers, as oligonucleotides have a phosphate backbone: it is important to introduce the right amount of competition from non-specific poly-anions during the selection process to increase the selective evolutionary pressure towards specific interactions[78].

Cell-SELEX

Principle

In cell-SELEX, a library is enriched for a cell line[79], xenograft[80] or tissue[81]. As described above, it is very common for a counter selection to be performed with healthy tissue or a non-target cell line. In the end, the hypothesis is that the cells for positive selection have markers expressed that are not found on the negative (counter) cell line. Therefore, aptamers are also useful tools for the study of membrane protein expression in cancer, as many cancers have membrane proteins that are uniquely expressed. Cell-SELEX is performed in the same way as normal SELEX is performed, but the main difference is that libraries are enriched for a complex target. Because of this, an enriched pool can have monoclonal aptamers for a wide array of membrane-associated molecules, including proteins and their posttranslational modifications (lipids, poly- or oligo- saccharides). Similar studies have been performed on the proteome of human plasma[82]. The enriched pool will thus contain a wide array of aptamers that can bind to different proteins. By analyzing the pools by cloning, or next generation sequencing, it is possible to find the aptamers that bind to the most significantly different target markers, resulting in a panel of aptamers that can distinguish the positive cell line from the negative cell line.

Applications of aptamers derived from cell-SELEX

There are several possible applications, many of which correspond to antibody-based applications, for aptamers selected for cells. One obvious application is the use of this technique on cells that overexpress a certain type of membrane protein for drug delivery purposes and tissue profiling. The classic example of such an aptamer is the tenascin-c aptamer. In this example, the parallel selection of aptamers in compared

between cell lines that are known to express tenascin-c and a cell line that has been made to express the protein. The tenascin-c aptamer that this study yielded could be used for cancer tissue targeting, as tenascin-c is an extracellular matrix protein often overexpressed in cancer[83].

Although Cell-SELEX is usually performed *in vitro*, *in vivo* selection is also possible. In the work conducted by Mi *et al.*, a mouse xenograft was targeted with a nuclease-resistant RNA pool, which yielded an aptamer for p68, an RNA helicase that allows for xenograft homing[81]. The interesting observation is the unusual location of the selected protein for this aptamer. Although p68 has a known function in the nucleus of the cell, the elucidated protein was also found in the cell membrane. Cell-SELEX will become a very potent technique for the study of cell membranes, as current methodologies would discard the unexpected protein based on previous assumptions about membrane proteins, cytosolic proteins and their sub-cellular location[84]. A similar example can be found in nucleolin, another RNA helicase that is found on the surface of many cancer cell types[85].

Another application of cell-SELEX is: the identification of protein associated at the membrane of cells. Researchers have demonstrated that it is possible to perform subtractive SELEX against virus-infected versus uninfected cells[86], and different lineages of hematopoietic cells or solid tumors[79,87,88]. For some of these selections, the target identification of these aptamers can lead to the identification of proteins that play a role in the differentiation state of cells, or in the infected state of cells. The identification usually occurs through a method where the aptamer is used as the ligand

in an affinity column, in the cell lysate or on the living cell[81,89-91]. The captured protein is then analyzed by standard mass spectrometric protein identification.

The identification of proteins that these aptamers bind can have several possible applications. The classic example is the aptamer that is FDA-approved for the treatment of age-related macular degeneration, branded Macugen by Gilead Sciences. This aptamer has anti-angiogenic properties, as it binds to vascular endothelial growth factor (VEGF). It is an excellent example of what the possibilities are for aptamers towards changing their pharmacological profile by post-SELEX modifications, as the aptamer is heavily modified by PEGylation[64]. Several companies are developing aptamer-based drugs that are now in clinical trails. Archemix is on the forefront by having several aptamer in phase II and phase III clinical trails. With their flagship aptamer, ARC1779, a phase II clinical trail was performed in patients with three types of Von Willebrand Factor-Related Platelet Function Disorders[92]. Another phase II aptamer is ARC 19499, for the treatment of hemophilia[93]. Archemix is also aiding large pharmaceutical companies like GlaxoSmithKline (AS1411), Arca (anticoagulant aptamers) and Regalo Biosciences (anticoagulant aptamers)in developing aptamers for the treatment of several pathologies[94].

Overview of the Dissertation

In this dissertation a panel of aptamers have been developed to bind to clear cell adenocarcinoma and ovarian cancer stem cells. Furthermore, a detailed description of each selection process, including aptamer characterization, is presented. The identity of the binding protein to one of the selected aptamers has been determined and the methodology for the identification process outlined. Finally, the oncological function of

this protein is being investigated, along with the action of the aptamer on viability, proliferation, migration and invasiveness of the cell line.

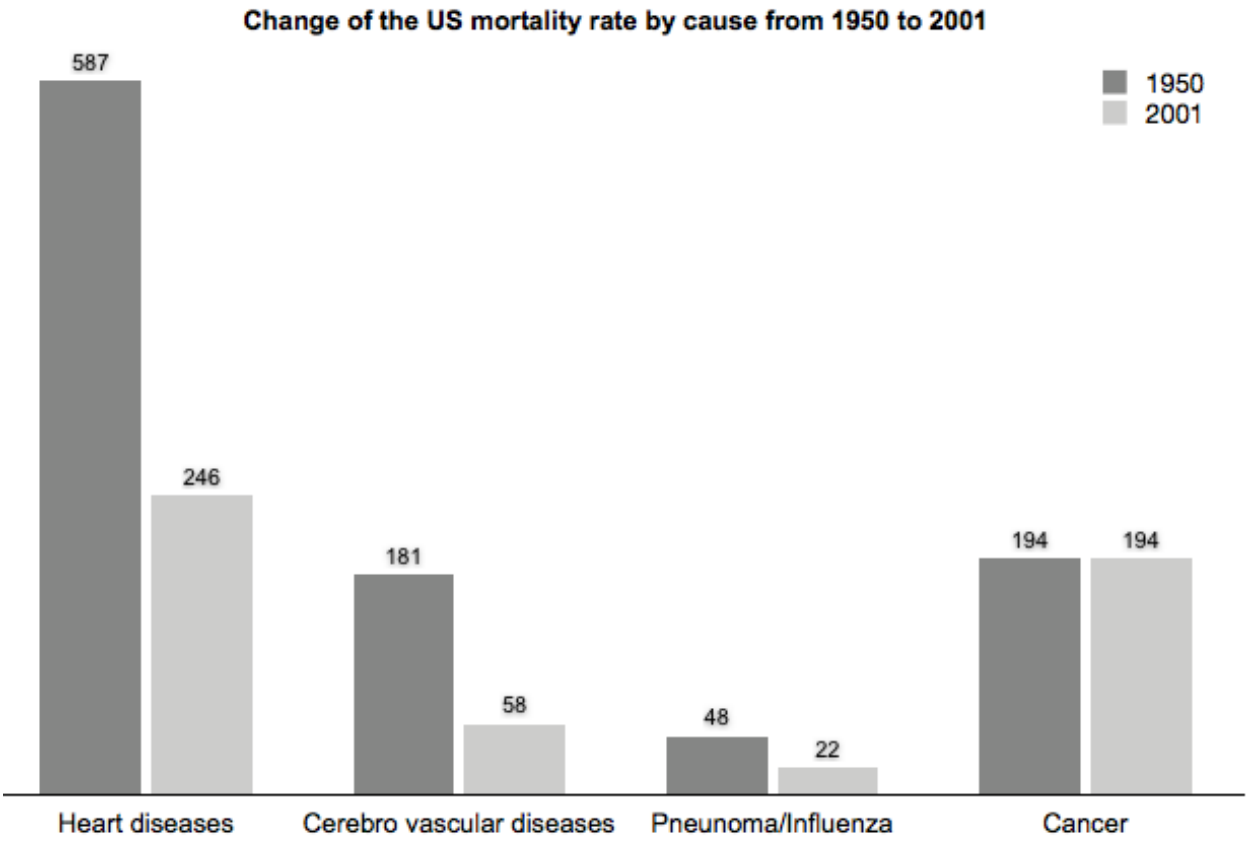


Figure 1-1. Change of the US mortality rate by cause from 1950 to 2001. The numbers represent the rate per 100,000. (Figure based on data found in Vital Statistics of the United States, annual, Vol. I and Vol II I; 1971-2001)

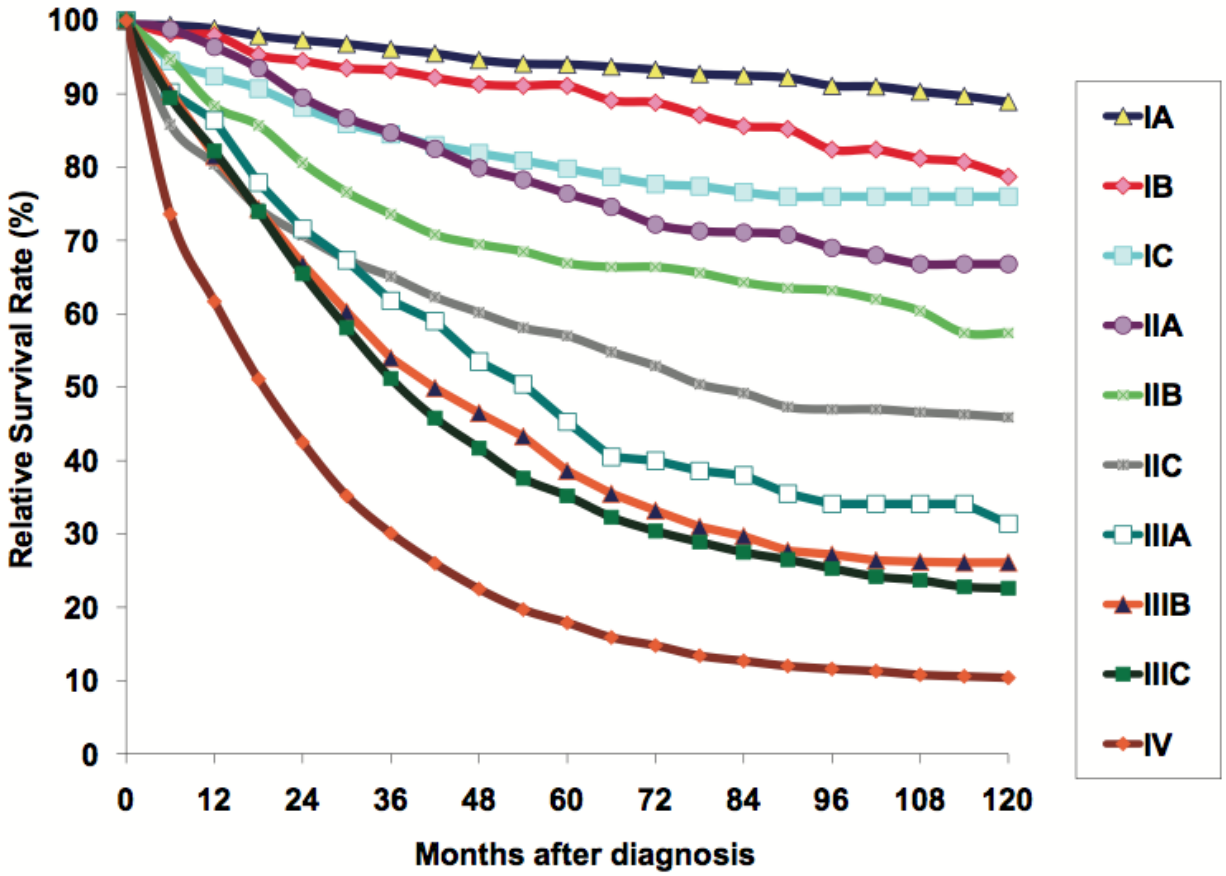


Figure 1-2. Adenocarcinoma of the ovary (excluding borderline tumors): relative survival rates (%) by AJCC Stage, Ages 20+, 12 SEER Areas, 1988-2001. See Table 1.1 for the description of the stages (Adapted from Kosary C. Cancer of the ovary. In: Ries L.A.G., Young J.L., Keel G.E., Eisner M.P., Lin Y.D., et al. (editors) SEER survival monograph: cancer survival among adults: US SEER Program, 1988–2001, patient and tumor characteristics. Bethesda (Maryland): National Cancer Institute, SEER Program (2007))

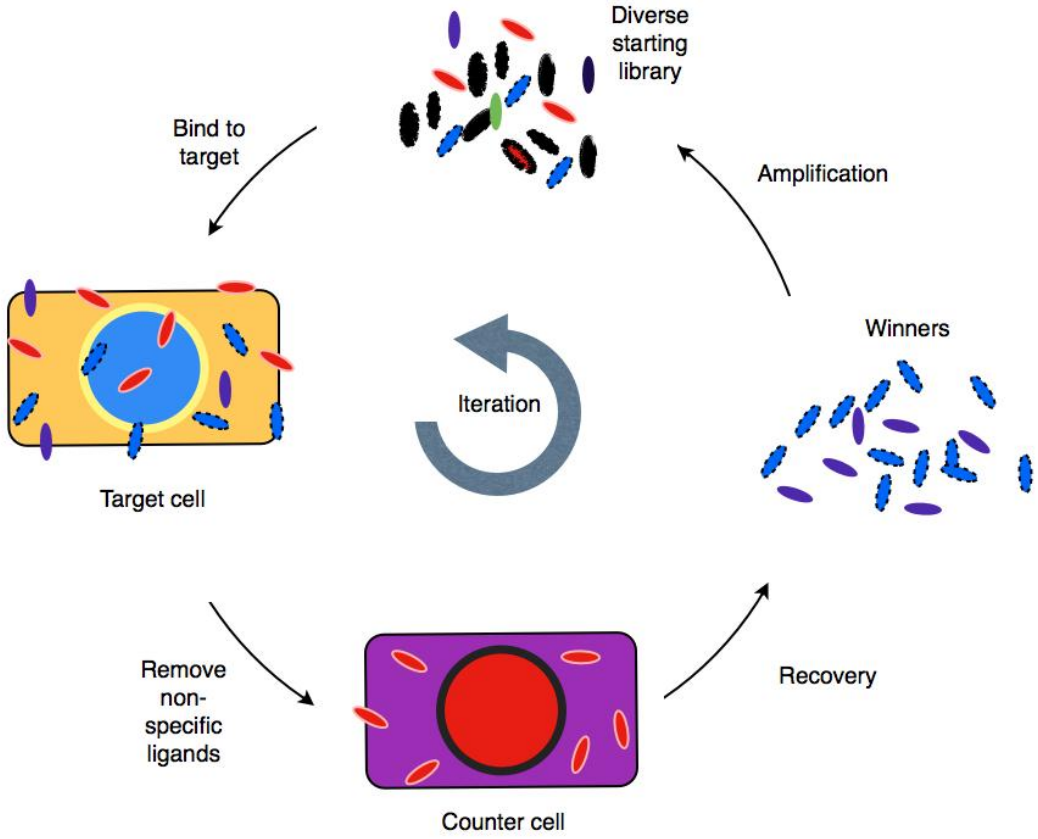


Figure 1-3. The basic scheme for cell-SELEX

Table 1-1. The description of the different stages of ovarian cancer according to the American joint committee on cancer (AJCC)

AJCC stage	Description
Stage I	Ovarian cancer is growth limited to the ovaries.
Stage IA	Growth limited to one ovary; no ascites. No tumor on the external surface; capsule intact.
Stage IB	Growth limited to both ovaries; no ascites. No tumor on the external surfaces; capsules intact.
Stage IC	Tumor either stage IA or IB, but with tumor on the surface of one or both ovaries; or with capsule ruptured; or with ascites present containing malignant cells or with positive peritoneal washings.
Stage II	Ovarian cancer is growth involving one or both ovaries with pelvic extension.
Stage IIA	Extension and/or metastases to the uterus and/or tubes. Stage IIB extension to other pelvic tissues. Stage IIC tumor either stage IIA or stage IIB, but with tumor on the surface of one or both ovaries; or with capsule(s) ruptured; or with ascites present containing malignant cells or with positive peritoneal washings.
Stage III	Ovarian cancer is tumor involving one or both ovaries with peritoneal implants outside the pelvis and/or positive retroperitoneal or inguinal nodes. Superficial liver metastasis equals stage III. Tumor is limited to the true pelvis but with histologically verified malignant extension to small bowel or omentum.
Stage IIIA	Tumor grossly limited to the true pelvis with negative nodes but with histologically confirmed microscopic seeding of abdominal peritoneal surfaces.
Stage IIIB	Tumor of one or both ovaries with histologically confirmed implants of abdominal peritoneal surfaces, none exceeding 2 centimeters in diameter. Nodes negative.
Stage IIIC	Abdominal implants greater than 2 centimeters in diameter and/or positive retroperitoneal or inguinal nodes.
Stage IV	Ovarian cancer is growth involving one or both ovaries with distant metastasis. If pleural effusion is present, there must be positive cytologic test results to allot a case to stage IV. Parenchymal liver metastasis equals stage IV.

CHAPTER 2
SELECTION AND CHARACTERIZATION OF OVARIAN CLEAR CELL ADENOCARCINOMA CELL APTAMERS

Introduction: Ovarian Clear Cell Adenocarcinoma

Ovarian cancer is the most lethal gynecological malignancy, and the ovarian clear cell adenocarcinoma subtype (OCCA) demonstrates a particularly poor response to standard treatment[95]. OCCA is the second most prevalent histological subtype of ovarian adenocarcinoma (4-25% of ovarian adenocarcinomas), which accounts for almost 90% of all ovarian cancers. The largest histological group is serous ovarian adenocarcinoma (SOAC) (±32,2% total)[4]. Although ovarian cancer shows generally good response to standard treatment with platinum-based drugs, OCCA does not respond to these drugs, resulting in poor prognosis for OCCA patients[95].

Clinically, OCCA and SOAC behave distinctly different. Some notible features of OCCA include its rare bilateral occurrence, its large pelvic mass presence, and its association with endometriosis, thromboembolic complications and hyper-calcemia. Like serous adenocarcinoma, early-staged OCCA (e.g., stages I and II) exhibits a better survival rate then stage III or IV OCCA. Of particular concern is the survival rate of OCCA, which is much lower than that of serous adenocarcinoma (median survival rate in serous adenocarcinoma is 3 to 4 years versus about a year in OCCA at stage I/II).

Since OCCA hardly responds to platinum-based therapy, different therapeutic targets are needed for OCCA in comparison to SOCA[95-97]. Some potential candidates for the treatment of OCCA are tyrosine kinases, and inhibitors for the PI3K–AKT–mTOR pathway[98]. Studies to make these inhibitors approved drugs for cancer therapy are underway, but still more drug candidates are needed, as these approaches need to be proven safe[95].

With the current system of FDA approval through clinical trails, a new and faster way to discover targets for cancer therapy in OCCA (and other diseases) is needed. It has been proposed that aptamers, or disease cell enriched libraries can lead towards the identification of new disease markers[99]. In Cell-SELEX, the majority of the possible targets are proteins, and when a library pool is generated against a disease cell, more information on the proteins expressed on the surface of that disease cell can be generated. Improvements in ovarian cancer outcomes, especially for OCCA, can be expected with a clearer understanding of the molecular pathology to guide strategies for earlier diagnosis and more effective treatment.

This chapter describes the SELEX on the OCCA cell line TOV-21G[100], using the cervical cancer cell line HeLa[101] for counter selection. Biochemical and biophysical properties of the selected aptamers are also studied.

Materials and Methods

Instrumentation and Reagents

All oligonucleotides were synthesized by standard phosphoroamidite chemistry using a 3400 DNA synthesizer (Applied Biosystems) and were purified by reversed-phase HPLC(Varian Prostar), using a C18 column in 0,1M Triethylamonium acetate in water/acetonitrile gradient. All PCR mixtures contained 50 mM KCl, 10 mM Tris- HCl (pH 8.3), 2.0 mM $MgCl_2$, dNTPs (each at 2.5 mM), 0.5 mM of each primer, and Hot start Taq DNA polymerase (5 U/mL) (TaKaRa). PCRs were performed on a Biorad Thermocycler. The monitoring of pool enrichment, characterization of the selected aptamers, and study of the target protein assays were performed by flow cytometry using a FACScan cytometer (BD Immunocytometry Systems). Trypsin and Proteinase K were purchased from Fisher Biotech. The DNA sequences in the pools were determined

by the Genome Sequencing Services Laboratory at the University of Florida with the use of 454 sequencing (Roche). Clustal X was used for analysis of sequence homology for sequence selection.

Cell Culture and Buffers

All cell lines were obtained from the American Type Cell Culture (ATCC). The CAOV-3 and TOV-21G ovarian cancer cell lines where maintained in culture with MCBD 105: Medium 199 (1:1); the HeLa cell line was cultured in RPMI-1640; and the Hs832(C)T cell line was cultured in Dulbecco's Modified Eagle's Medium (DMEM). All media where supplemented with 10% FBS and 100 IU/mL Penicillin-Streptomycin. Other cell lines used for selectivity assays included CEM (T cell acute lymphoblastic leukemia), Ramos (Burkitt's lymphoma), HCT-116, DLD-1, HT-29 (colorectal adenocarcinoma), NCI H23 (non-small cell lung cancer) and A172 (glioblastoma), all of which were cultured according to ATCC recommendations. All cell lines where incubated at 37ºC in a 5% CO_2 atmosphere. During the selection, cells were washed before and after incubation with washing buffer (WB), containing 4.5 g/L glucose and 5 mM $MgCl_2$ in Dulbecco's phosphate buffered saline with $CaCl_2$ and $MgCl_2$ (Sigma). Binding buffer (BB) used for selection was prepared by adding yeast tRNA (0.1 mg/mL) (Sigma) and BSA (1 mg/mL - Fisher) to the washing buffer to reduce non-specific binding.

SELEX Library and Primer Design

A typical DNA library for SELEX constitutes a random nucleotide region, flanked with primers that allow for amplification of the pool after incubation with the target. The primers are designed so, that so-called molecular parasites are reduced. Molecular parasites are found in unoptimized libraries, where the primers can interact with

themselves instead of their complement in the template[48]. The primers in this thesis were optimized so that the T_m differ less then 1ºC, and, to minimize unwanted elongation, the annealing temperature maximum was 80% of the elongation temperature (55 - 60ºC). The primers were not allowed to form hairpins, and only three Watson-Crick base pairs were permitted between the different primers. Finally, the complement of the reverse primer and the forward primer were not allowed to form hairpins. The optimized primers were chosen using mfold software[102]. The chosen primer pair was incorporated into a library, with each component having a ssDNA strand that is synthesized as the forward primer, followed by the random sequences, and ended by the complement of the reverse primer. In this selection 5'- ATC CAG AGT GAC GCA GCA $(N)_{40}$ TGG ACA CGG TGG CTT AGT-3' was used, which was synthesized using standard solid phase synthesis and HPLC purification, as described above. The forward and reverse primer were labeled with 5'-FITC and with 5'-biotin, respectively.

***In Vitro* Cell-SELEX**

In this study, TOV-21G was used as the target cell line and HeLa was used for counter-selection. For the first round, the cells were incubated with 20 nmol of naïve ssDNA library dissolved in BB. For later rounds, 50 pmol of enriched pool were used for incubation, also dissolved in BB. Before incubation, the ssDNA pool was denatured by heating at 95ºC for 5 minutes and was placed on ice for 5 minutes, allowing each sequence to form the most stable structure. The cells were washed twice (2 minutes) with WB and incubated with the DNA pool on ice in an orbital shaker for 30 minutes. In later selection rounds, the washing stringency was increased to remove weakly binding sequences (a larger number of washes and increased washing time, up to 5 minutes).

The bound sequences were eluted in 500 mL WB by heating at 95ºC for 10-15 minutes, cooling on ice for 5 minutes and centrifuging at 14,000 rpm for 2 minutes. The supernatant, which contained the binding sequences, was then incubated with a negative cell line to remove general sequences, as described in Figure 1.4. The remaining sequences were amplified by PCR using the FITC- and biotin-labeled primers. Amplifications were carried out at 95ºC for 30s, 60ºC for 30s, and 72ºC for 30s, followed by final extension for 3 minutes at 72ºC. The selected sense ssDNA was separated from the biotinylated antisense ssDNA by streptavidin-coated Sepharose beads (Amersham Bioscience). The ssDNA was eluted from the streptavidin beads by melting the dsDNA in a 0.2M NaOH solution. The ssDNA is then desalted with the use of a G25 Sephadex size exclusion column (GE healthcare). The enrichment of specific sequences was assayed using flow cytometry, as explained below. When the level of enrichment was satisfactory, pools of interest were submitted for sequencing.

Sequencing and Selection of Putative Aptamers

In this study, next generation sequencing was used for the analysis of enriched pools. Several pools were labeled with a specific sequence 'coding' each pool with the use of PCR, so enrichment of sequences could be monitored later on. More specifically, 454 sequencing was used, which yielded about 6000 sequences per run. Since it was foreseen that the field of next-gen sequencing will make drastic steps in the future[103], a script was written in PERL[104] (see appendix A) for the high throughput trimming of the primers prior to the processing of actual aptamer sequences for sequence homology. Sequences that fulfilled the requirements (data flanked with the primer pair and minimum 70 nt long) were analyzed for homology in ClustalX[105].

Affinity Studies: Flow Cytofluorometric Analysis for the Determination of Binding Affinity

To determine the binding affinities of the aptamers and the enrichment of the pools, the target cells (5×10^5) were incubated with various concentrations of 5'- biotin labeled aptamers or FITC-labeled pools on ice for 30 minutes in 100 mL of BB. Cells were then washed twice with 500 mL of BB, and suspended in 100 mL of BB containing streptavidin PE-Cy5.5 at appropriate dilution. Cells were then washed twice with 500 mL of WB, and then suspended in 200 mL of BB for flow cytometry analysis, using a 5'-biotin labeled random sequence as the negative control. All the experiments for binding assays were repeated at least twice. The specific binding intensity was calculated by subtraction of the mean fluorescence intensity of the background binding from the mean fluorescence intensity of the aptamers. The apparent equilibrium dissociation constant (K_d) of the fluorescent ligand was obtained by non-linear regression of the specific binding intensity (Y) and the aptamer concentration (X) to the equation $Y=B_{max}X/(K_d+X)$ using SigmaPlot. (Jandel, San Rafael, CA).

Selectivity and Specificity

To determine the cell specificity of the selected aptamers, cell lines including HeLa, K562, H23, H69, A172, HL-60. HT-29, Ramos and CEM were used in binding assays by flow cytometry, as described above.

Effect of Temperature on Aptamer Binding

The aptamer selection process and all of the binding assays were performed on ice. It has been observed that some (one out of dozens selected to date from the Tan lab) of the aptamers selected at lower temperatures may not bind well at 37ºC, leading to poor performance under physiological conditions[89]. In order to verify binding stability,

aptamers were incubated with the target at 37ºC, and fluorescence intensity was determined by flow cytometry. Aptamers incubated at 4ºC were used as the positive control.

Protease Digestion Assay

Target cells ($5x10^5$) were detached using a non-enzymatic cell dissociation solution (Sigma). After suspending the cells, they were washed with 3 mL of PBS and then incubated with 1 mL of 0.05% trypsin/0.53 mM EDTA in HBSS or 0.1 mg/mL proteinase K in PBS at 37ºC for 1, 5, 15, 30 and 60 minutes. Pure FBS was added to quench the protease reactions. After washing with 2 mL of BB, the treated cells were used for binding assays as described above.

Results

Monitoring of Pool Enrichment for TOV-21G against HeLa

To start the selection process, 20 pmol of naïve library was enriched by sequential binding to TOV-21G cell monolayers. Sequences showing non-specific binding to general cell surface markers were removed by incubating the enriched pool with HeLa cells (rounds 2, 4, 5, 7, 8, 9, 12, 20, 21, 22). The eluted pool for each round of SELEX was amplified through semi-quantitative PCR. Special care was given to prevent unwanted amplification of molecular parasites. An example can be found in Figure 2-1, in the bands for 18 and 20 PCR cycles. As the figure shows, after 16 rounds of PCR, unwanted amplification occured. Therefore, a larger scale (1mL vs 50µL) PCR was repeated at 16 rounds in order to maximize yields. This procedure was followed throughout the entire SELEX procedure. When counter-selection in rounds 13 to 19 was omitted, re-enrichment for some HeLa-binding sequences occurred (Figure 2-2). The sequences binding to HeLa cells were successfully removed by reintroducing

counter selection in subsequent rounds, while the enrichment towards the target cell line was maintained. After 22 rounds of SELEX, an enriched pool that specifically bound to the model OCCA cell line, but marginally to HeLa cells, was obtained (Figure 2-2). This procedure produced a pool enriched for sequences binding to surface markers expressed by the model OCCA cell line, but not by the cervical cancer cells.

Aptamer Characterization

After 22 rounds a pool with satisfactory (i.e. specific binding to target) enrichment was obtained and pools were selected for sequencing, after which aptamers were chosen from the sequencing data based on the homology and sizes of the families. As observed in a SELEX experiment on plasma (another complex selection target), observable enrichment is not always necessary to obtain sufficiently large homologous families for aptamer choice[82,106].

It was hypothesized that 454 sequencing would give the resolution needed to find aptamers at an early stage. In other words, the information given by 454-sequencing of a pool was hypothesized to give a good representation of the actual pool under investigation. The frequency of selected sequence families is summarized in Table 2-1. Pool 13 was submitted for sequencing, as it was the first pool with a slight fluorescence intensity increase in flow cytometry. Pool 22 was also submitted for analysis, because it was the final pool, and pool 21 was submitted to study the possible changes that can occur in one round. Pool 21 had some binding to HeLa, while pool 22's binding to HeLa was only marginal. As can be seen in the pool, only a small change was observed in the frequency of the presented aptamers between pool 13 and 22, and the frequency seemed to be independent of the binding (i.e. aptTOV6 does decrease, which could

suggest that HeLa may have AptTOV6-binding motifs; however the aptamer binds specifically to TOV-21G, but not to HeLa).

From the alignment data, 9 sequences (and one extra sequence for aptTOV2 that showed an interesting 'mutation') were chosen and characterized by protease digestion assays and affinity determination, as described in the Methods and Materials section. Table 2-2 contains the sequences of the aptamers with their respective dissociation constants (K_d) and frequencies in pool 22. The K_ds for the selected aptamers were all in the picomolar to the nanomolar range. Especially aptTOV1 showed tight binding with the cells, with a K_d of 250 ± 80pM (Figure 2-12). It is interesting to observe that the aptamers with sub-nano molar affinities were enriched relatively rapidly (Table 2-1). All the selected aptamers were found in all the pools and it was observed that they could bind to TOV-21G but not to HeLa, i.e., with the same binding profile as their respective pool (Figure 2-13 illustrates this for aptTOV1).

More binding studies were performed on an array of different cell lines from different types of cancer, and the results are tabulated in Table 2-3. Some basic characterization was performed by protease treatment studies, where it was found that all the targets for the aptamers were cleaved of by proteases (Figure 2-3 to Figure 2-11). The protease-mediated removal of the fluorescence implies that the target of the aptamer is associated with a protein. However, it does not specify the exact epitope of the target, which could be a peptide chain or aterminal oligosaccharide[89,91,107].

Concluding Remarks

A series of aptamers with high affinity for OCCA (TOV-21G) have been selected that can distinguish ovarian cancer from cervical cancer (Figure 2-13). In particular, AptTOV 1 showed very high affinity towards TOV-21G, with a K_d of 250 pM. Given the

limited number of biomarkers for ovarian cancer currently available, the aptamers obtained from these selections have potential for improving diagnosis and treatment of this deadly disease.

Because the aptamers also bind benign cysts (Table 2-3), the aptamers cannot be used to identify ovarian cancer *per se* (more binding studies are warranted). However, since the aptTOV aptamers do not bind to a cancer of similar etiology (CAOV3) and also not to HeLa, they still have the potential to provide more insight into the pathology of ovarian cancer, and to become a valuable tool for distinguishing OCCA from other ovarian cancer subtypes. It has been observed that there are significant differences in the proteome of serous and clear cell ovarian cancer[108], and the need for oncologists to distinguish clear cell adenocarcinoma from serous ovarian adenocarcinoma was recently unequivocally established to be a genuine need[95]. The targets for these aptamers are most likely down-regulated or silenced in these two cell models.

Additionally, the aptTOV aptamers show binding to cancer cell lines from different non-related cancers (Table 2-3) and some aptTOV aptamers also bind CEM cells. This result suggests that the aptamers obtained from this SELEX can be used for profiling the expression of membrane proteins of different cancers. Identifying the targets of the selected aptamers is expected to shed light on the underlying mechanisms and pathways involved of these deadly diseases[109].

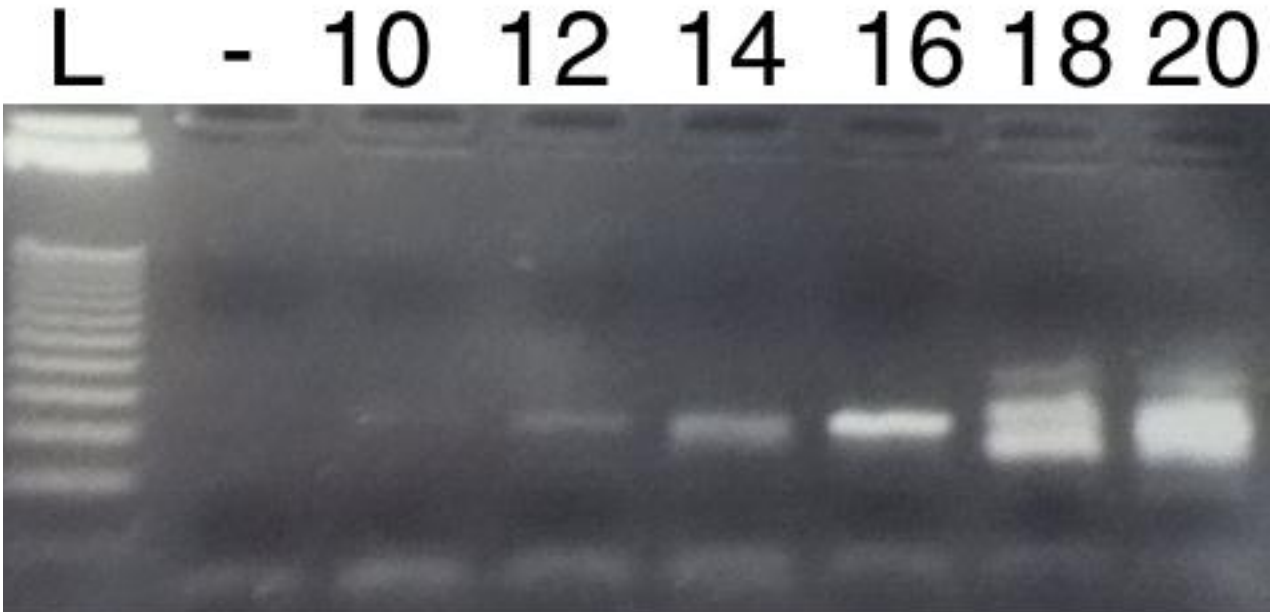


Figure 2-1. Sample agarose gel of a semiquantitative evaluation of pool enrichment. L: 25bp ladder, - : no template, the numbers stand for the number of PCR rounds

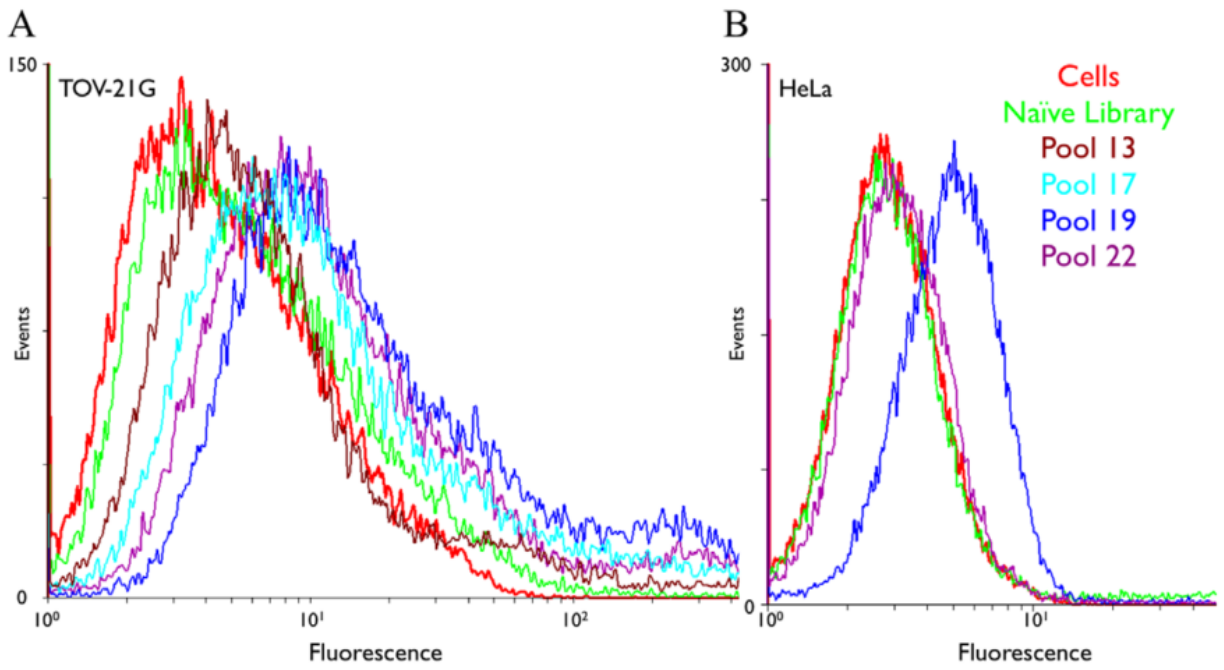


Figure 2-2. Binding assay of the enriched pools with TOV-21G and HeLa cells. A) Enrichment with TOV-21G cells. B) Marginal binding of the respective pools to HeLa cells. By doing counter selection, sequences binding to HeLa were removed

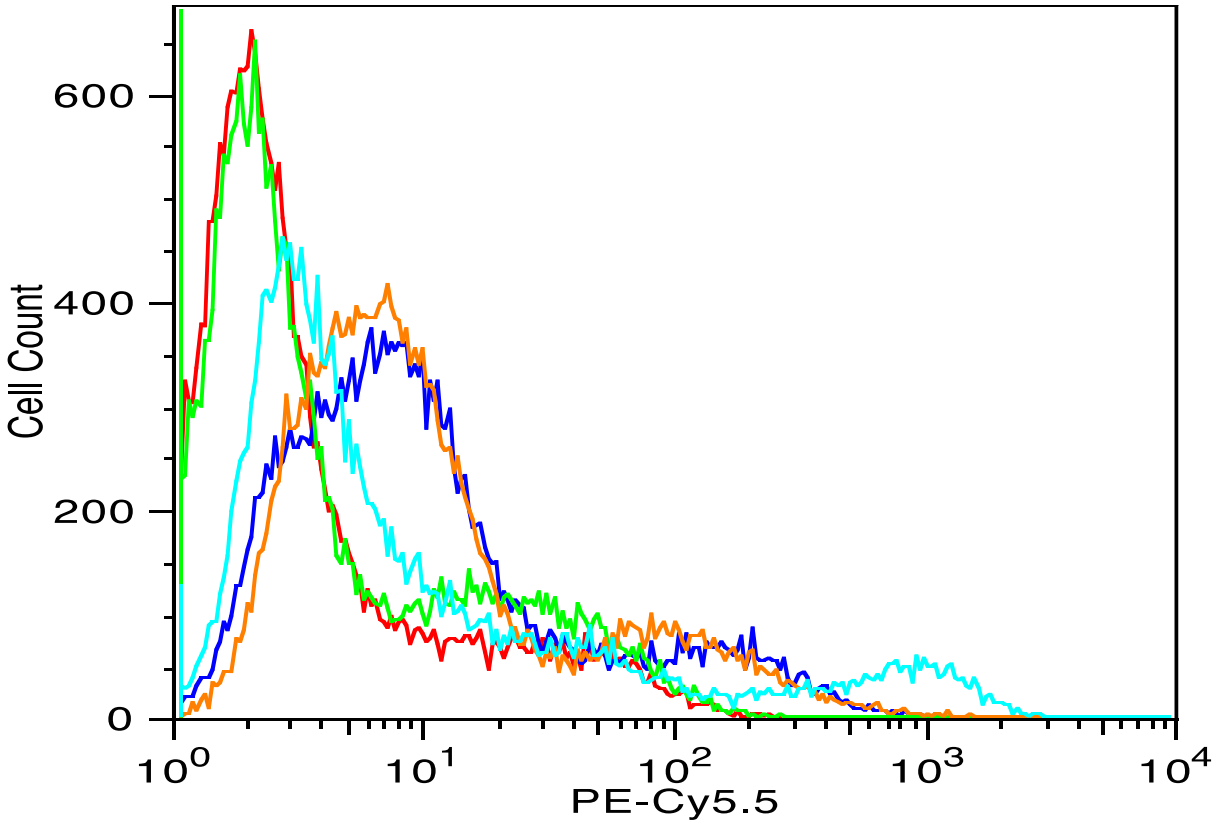


Figure 2-3. Binding profile of aptTOV1 (Red: TOV21G; Green: Library; Dark blue: 4ºC for 30min; Orange: 37ºC; Light blue: 37ºC after 30 minutes 0.05% trypsin treatment)

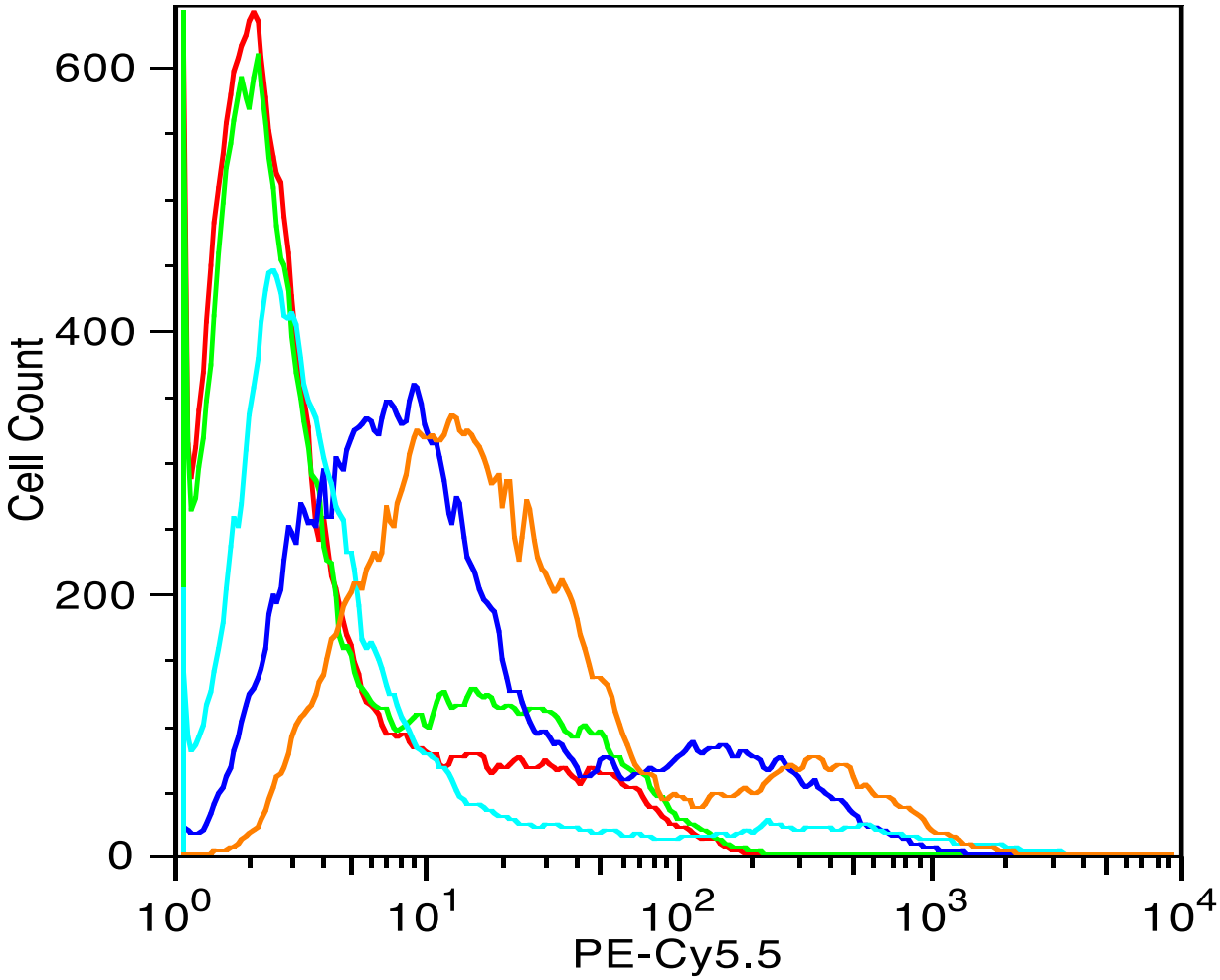


Figure 2-4. Binding profile of aptTOV2 (Red: TOV21G; Green: Library; Dark blue: 4ºC for 30min; Orange: 37ºC; Light blue: 37ºC after 30 minutes 0.05% trypsin treatment)

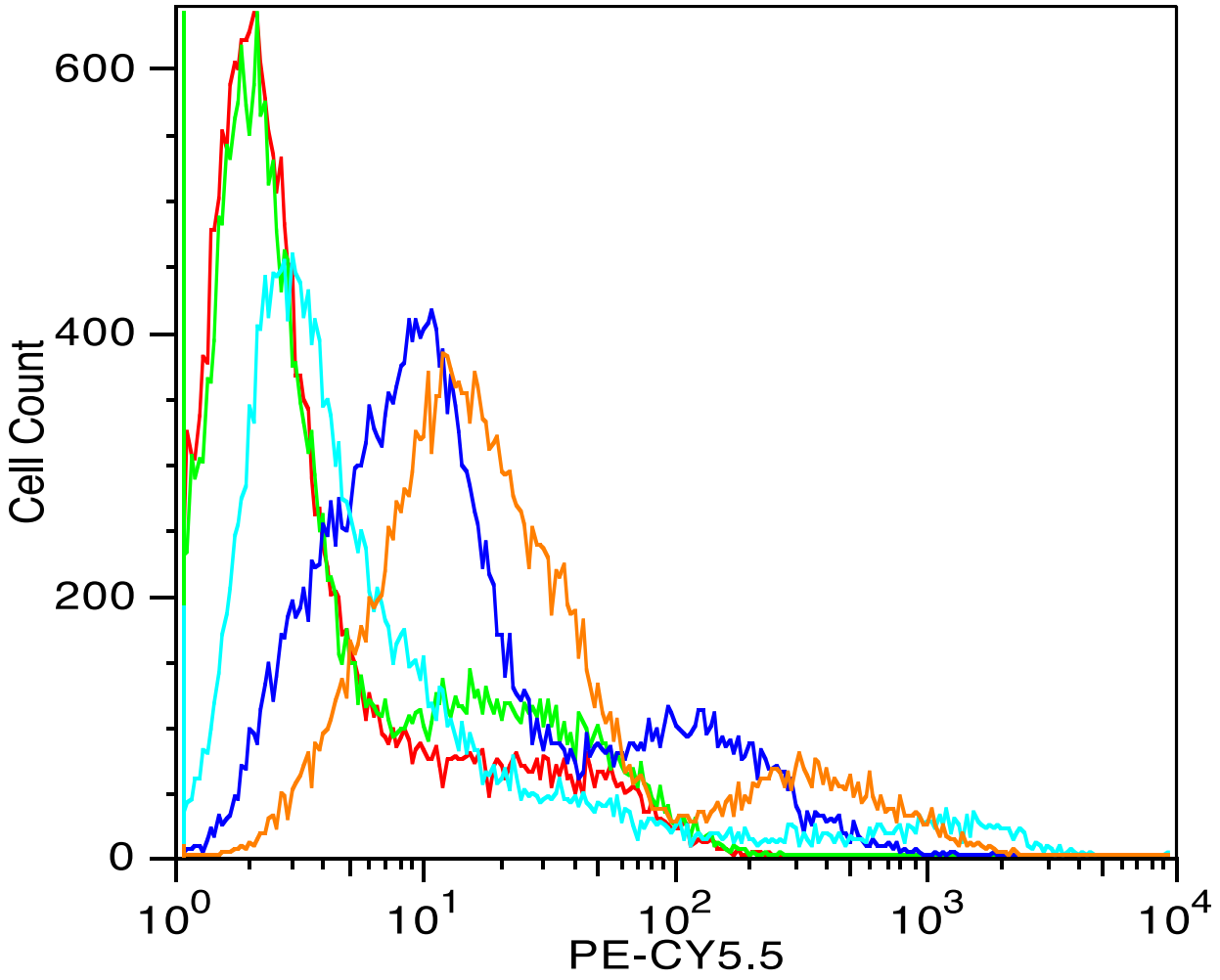


Figure 2-5. Binding profile of aptTOV3 (Red: TOV21G; Green: Library; Dark blue: 4ºC for 30min; Orange: 37ºC; Light blue: 37ºC after 30 minutes 0.05% trypsin treatment)

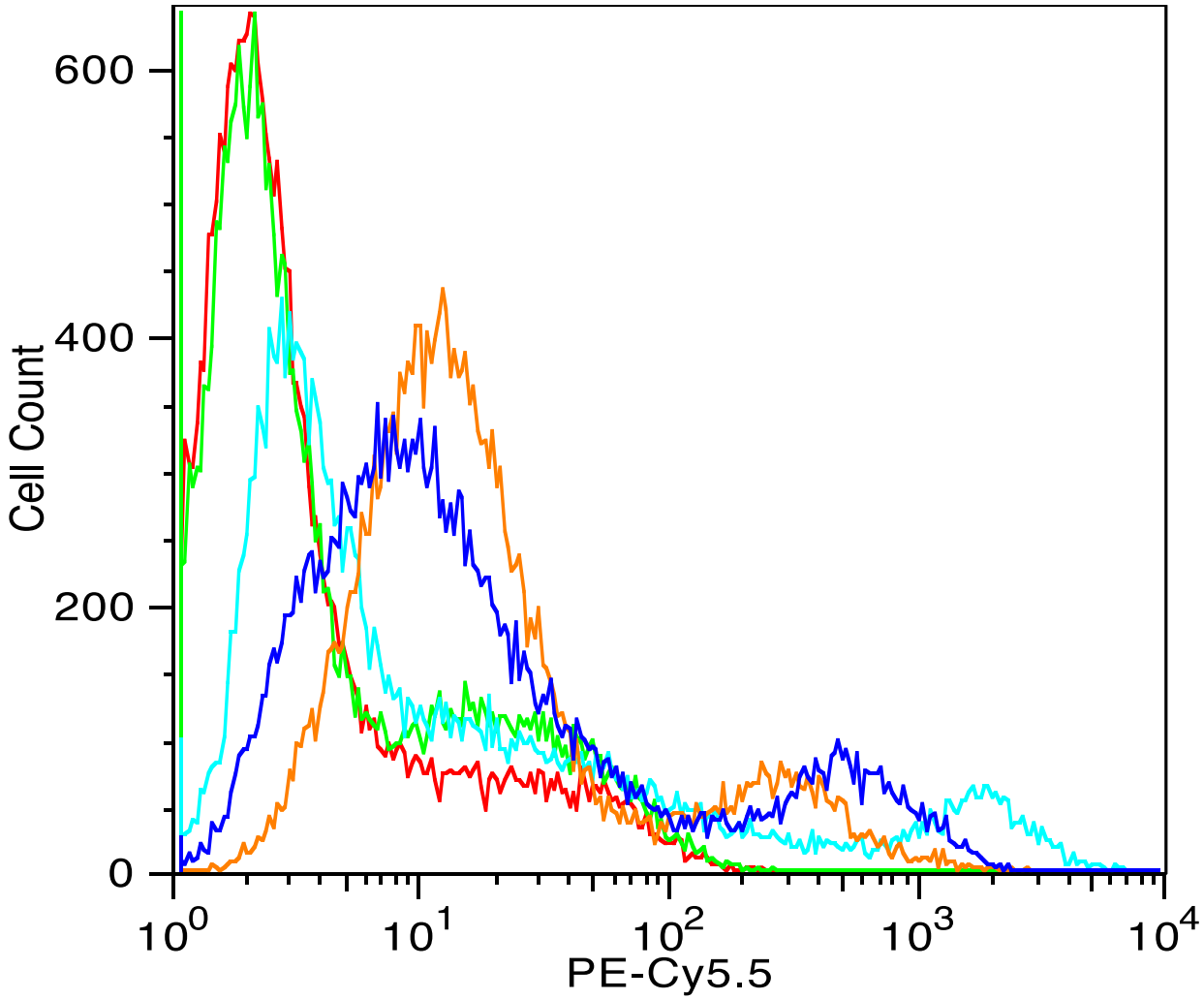


Figure 2-6. Binding profile of aptTOV4 (Red: TOV21G; Green: Library; Dark blue: 4ºC for 30min; Orange: 37ºC; Light blue: 37ºC after 30 minutes 0.05% trypsin treatment)

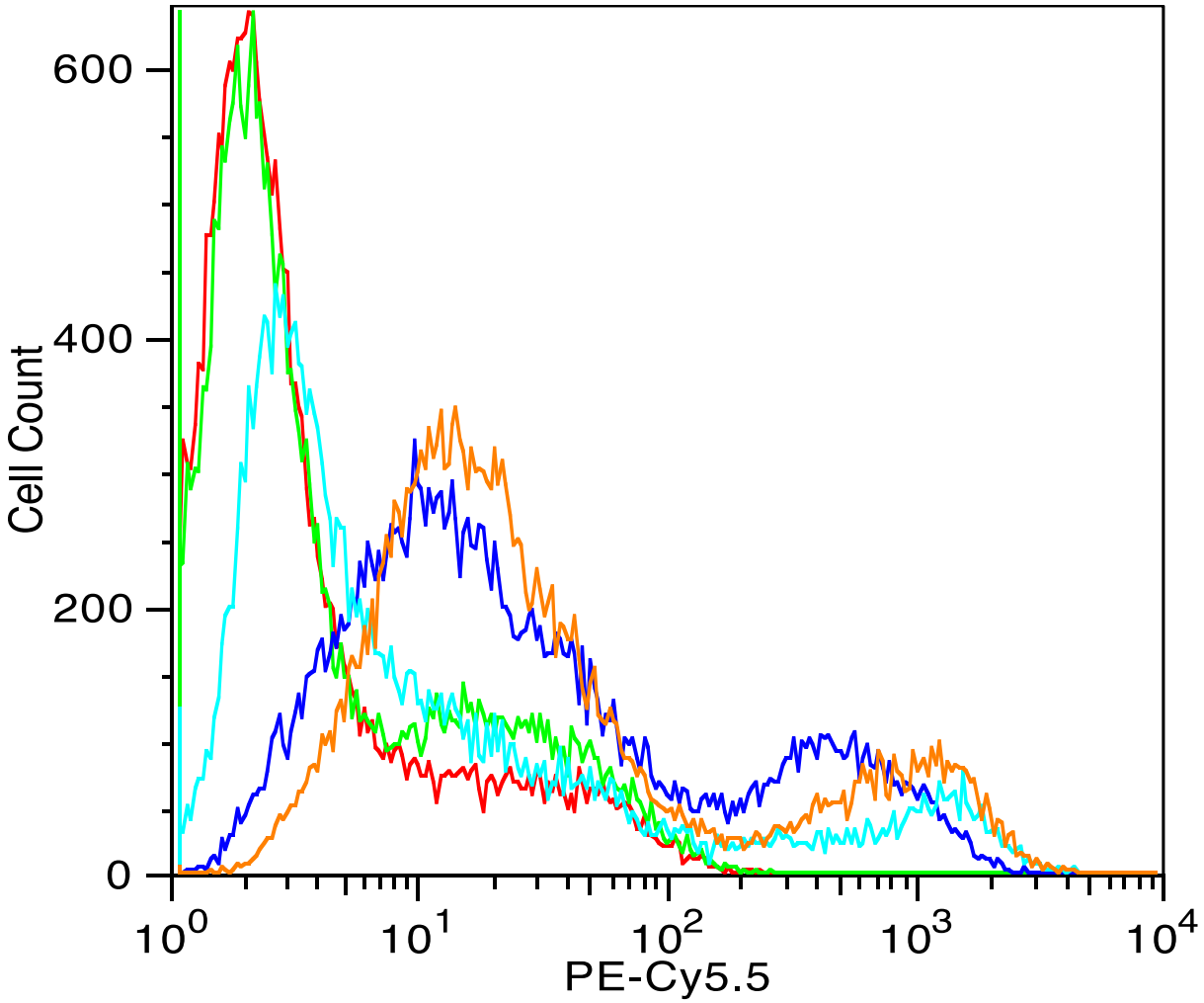


Figure 2-7. Binding profile of aptTOV5 (Red: TOV21G; Green: Library; Dark blue: 4ºC for 30min; Orange: 37ºC; Light blue: 37ºC after 30 minutes 0.05% trypsin treatment)

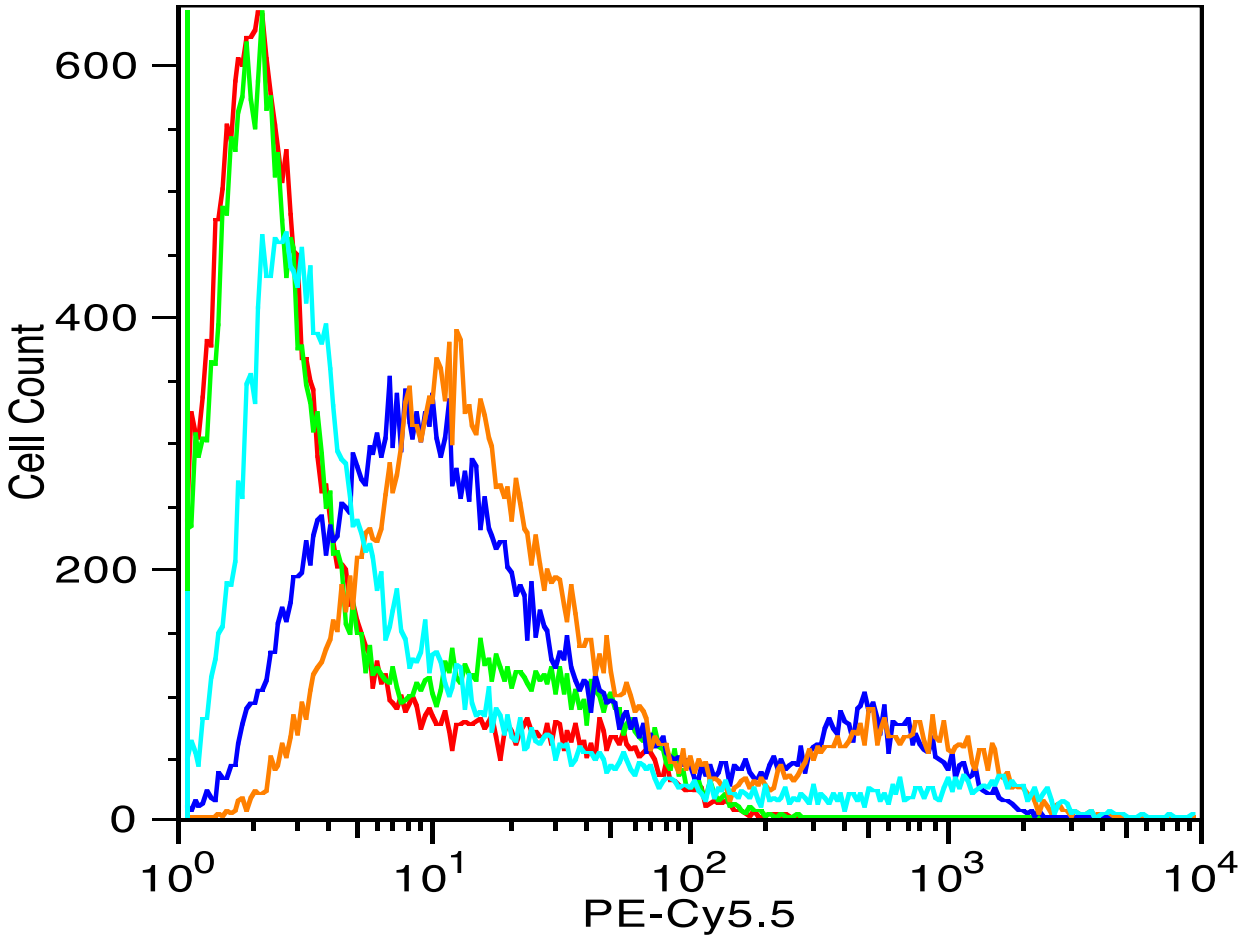


Figure 2-8. Binding profile of aptTOV6 (Red: TOV21G; Green: Library; Dark blue: 4ºC for 30min; Orange: 37ºC; Light blue: 37ºC after 30 minutes 0.05% trypsin treatment)

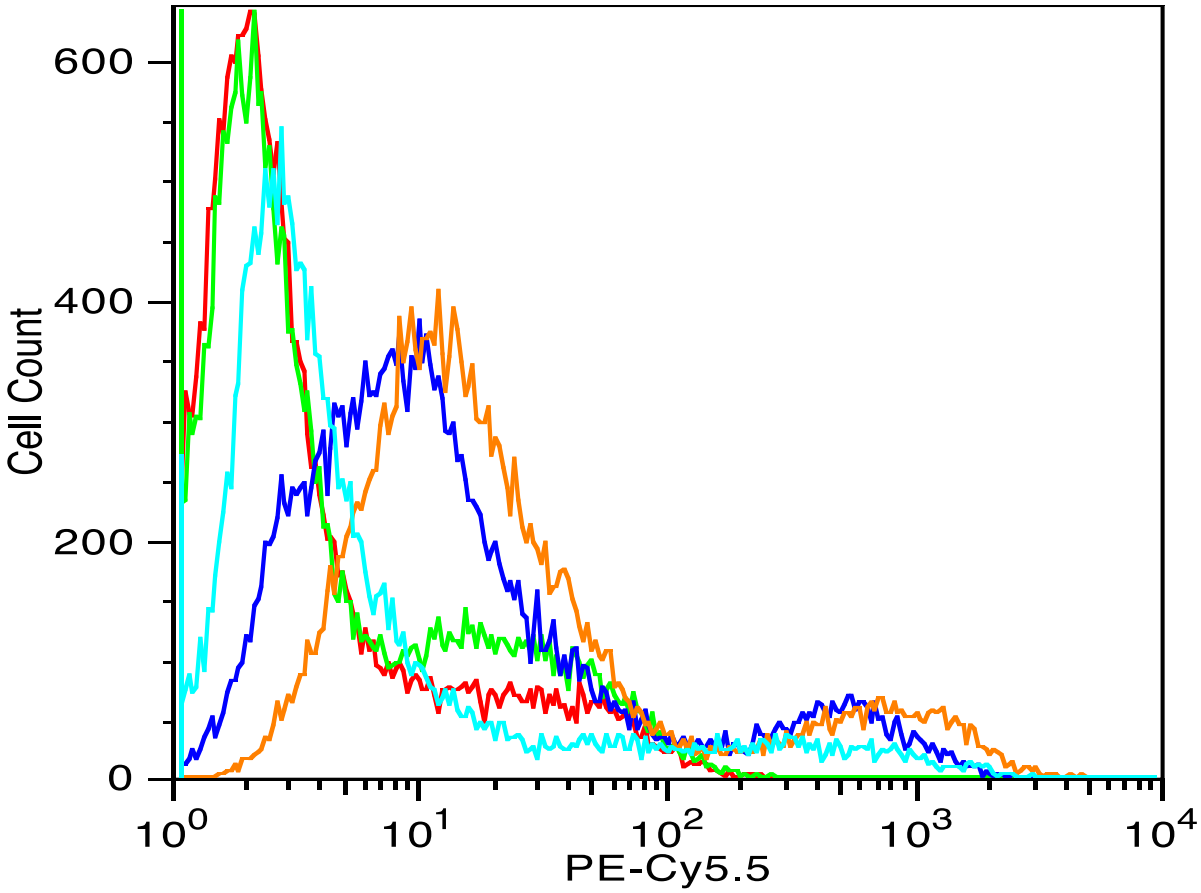


Figure 2-9. Binding profile of aptTOV7 (Red: TOV21G; Green: Library; Dark blue: 4ºC for 30min; Orange: 37ºC; Light blue: 37ºC after 30 minutes 0.05% trypsin treatment)

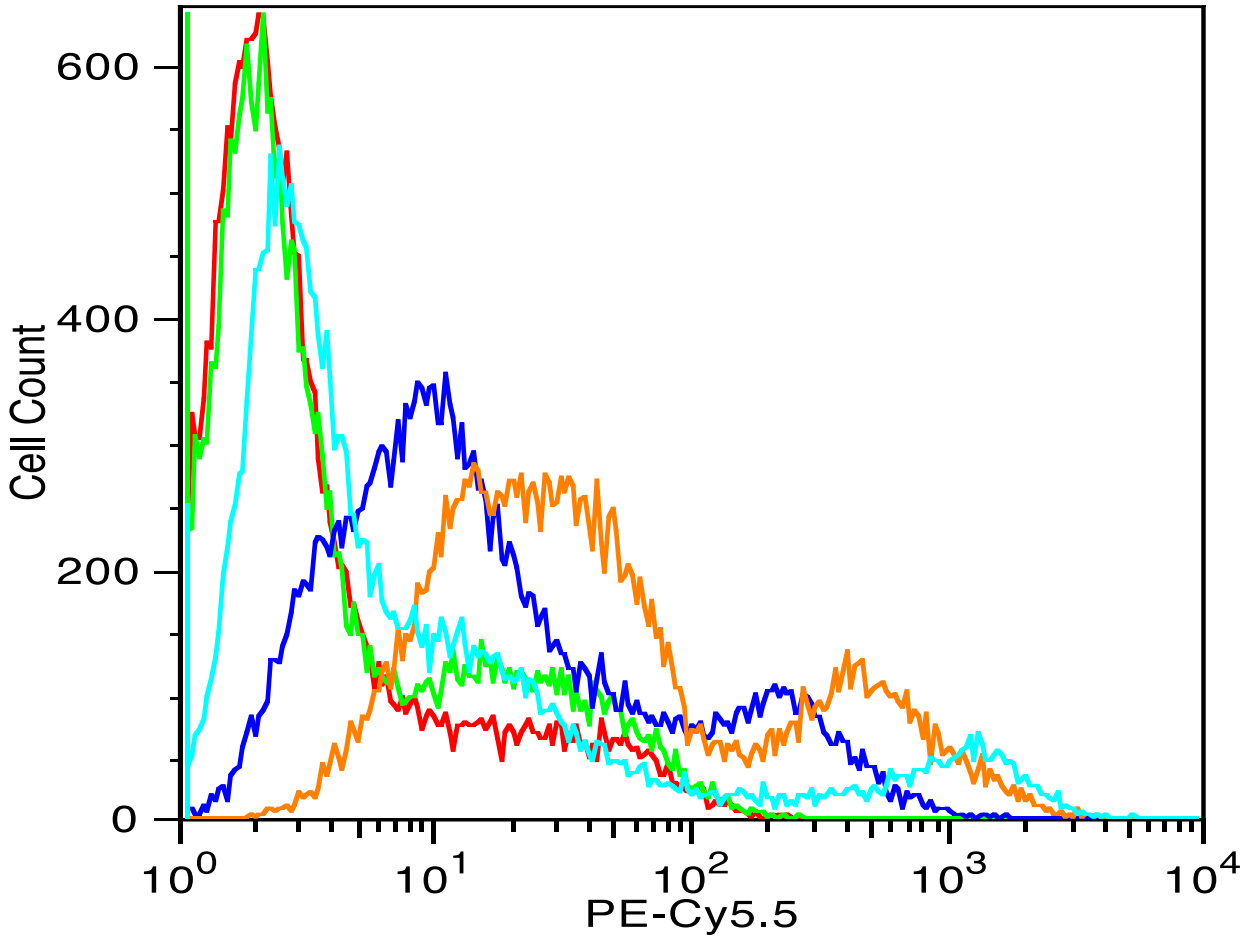


Figure 2-10. Binding profile of aptTOV8 (Red: TOV21G; Green: Library; Dark blue: 4ºC for 30min; Orange: 37ºC; Light blue: 37ºC after 30 minutes 0.05% trypsin treatment)

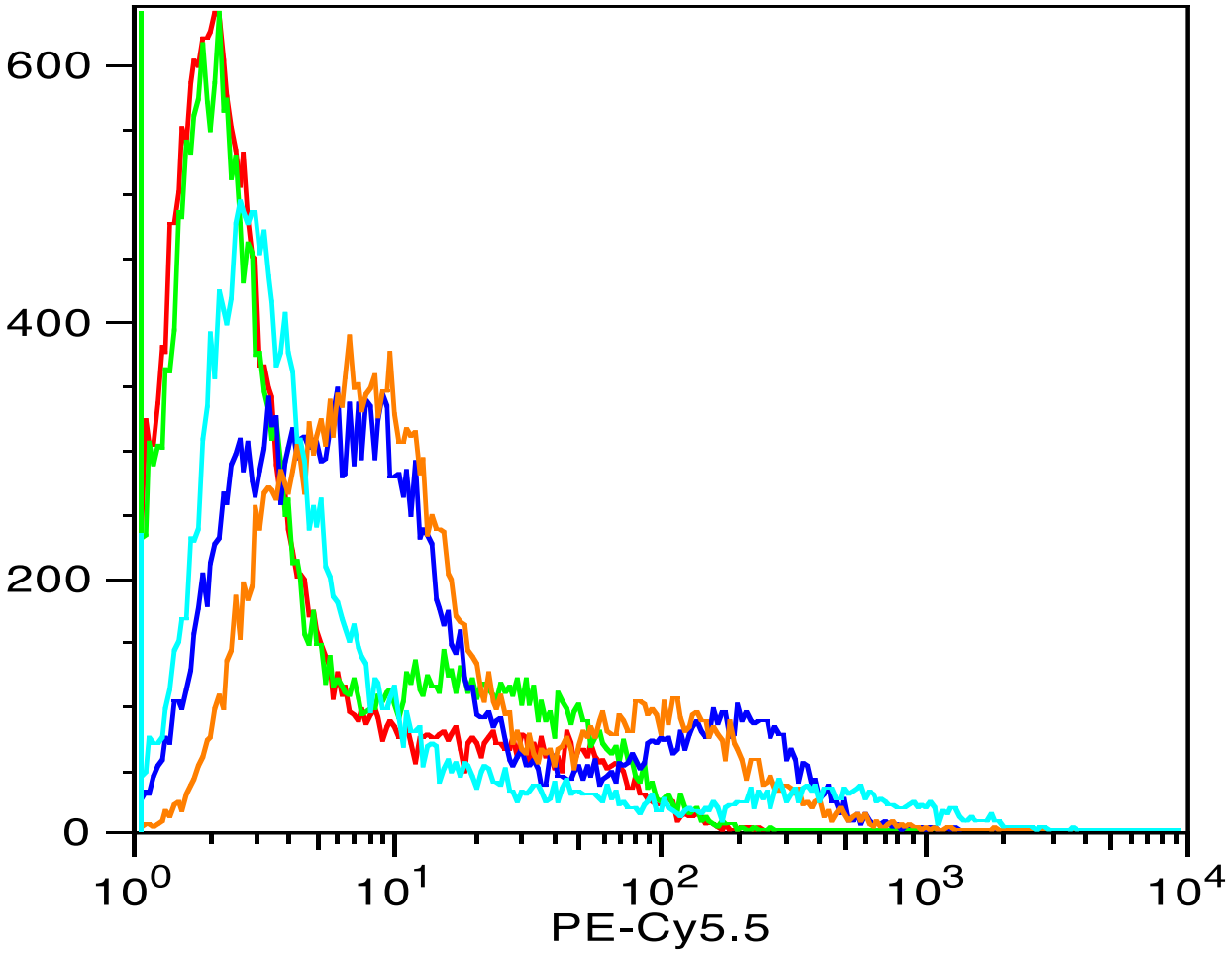


Figure 2-11. Binding profile of aptTOV9 (Red: TOV21G; Green: Library; Dark blue: 4ºC for 30min; Orange: 37ºC; Light blue: 37ºC after 30 minutes 0.05% trypsin treatment)

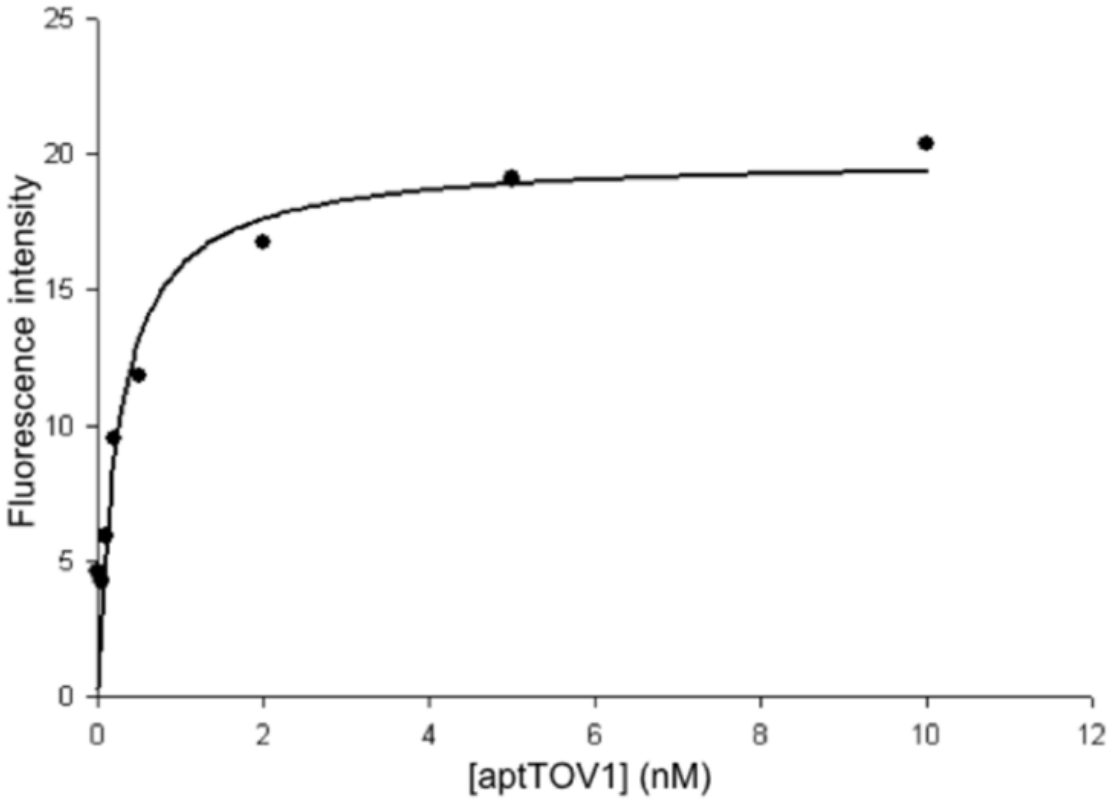


Figure 2-12. K_d determination for aptTOV1. Cells were incubated with varying concentrations of PE-Cy5-labeled aptamer in duplicate. The fluorescence intensity originating from background binding at each concentration was subtracted from the mean fluorescence intensity of the corresponding aptamer

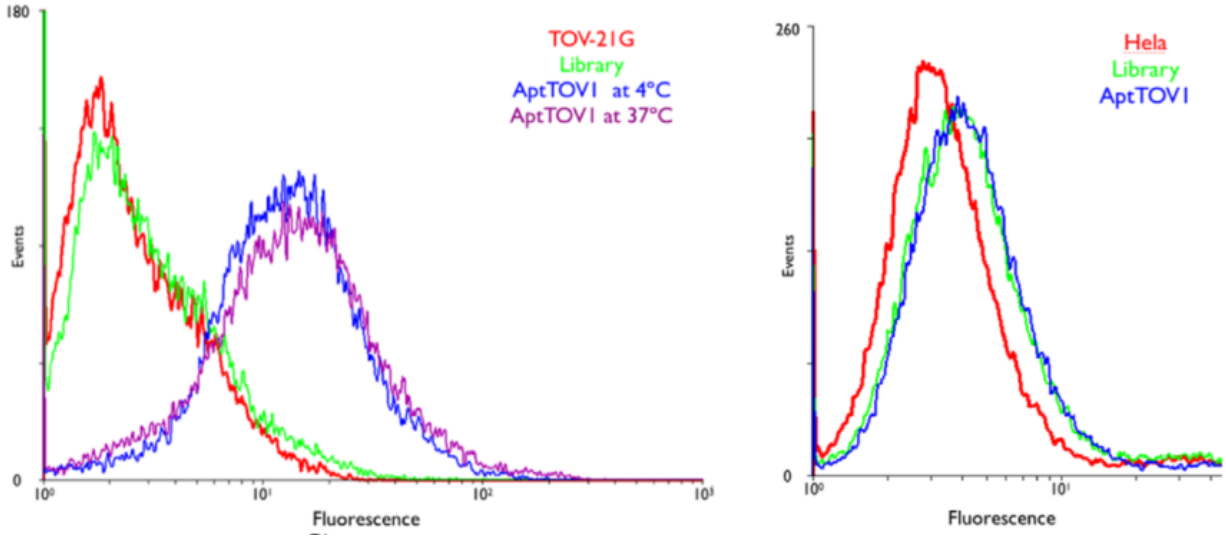


Figure 2-13. Specificity assay for aptTOV1 on TOV-21G and HeLa. The aptamer binds well to TOV-21G, but poorly to the negative SELEX cell line HeLa. Aptamers were labeled with PE-cy5.5 [119]

Table 2-1. Selected aptamer frequencies in the analyzed pools

	Pool 13 (%)	Pool 21 (%)	Pool 22 (%)
aptTOV1	7.56	1.91	2.53
aptTOV2	6.43	10.84	18.62
aptTOV2-a	2.96	4.92	7.65
aptTOV2 total	9.30	15.76	26.27
aptTOV3	4.05	7.97	8.74
aptTOV6	0.30	0.92	0.58

Table 2-2. Compendium of the aptamers obtained by selection against TOV-21G

Name	Sequence	K_d (nM)	Freq. (%)
aptTOV1	5’ - ATC CAG AGT GAC GCA GCA GAT CTG TGT AGG ATC GCA GTG TAG TGG ACA TTT GAT ACG ACT GGC TCG ACA CGG TGG CTT A - 3'	0.25 ± 0.08	2.52
aptTOV2	5’ - ATC CAG AGT GAC GCA GCA TAA TCT CTA CAG GCG CAT GTA ATA TAA TGA AGC CCA TCC ACC TGG ACA CGG TGG CTT A- 3’	0.90 ± 0.25	18.62
aptTOV3	5’ - ATC CAG AGT GAC GCA GCA CTC ACT CTG ACC TTG GAT CGT CAC ATT ACA TGG GAT CAT CAG TCG ACA CGG TGG CTT A- 3’	30 ± 9	8.73
aptTOV4	5’ - ATC CAG AGT GAC GCA GCA GGC ACT CTT CAC AAC ACG ACA TTT CAC TAC TCA CAA TCA CTC TCG ACA CGG TGG CTT A- 3’	20 ± 5	0.52
aptTOV5	5’ - ATC CAG AGT GAC GCA GCA CAA CAT CCA CTC ATA ACT TCA ATA CAT ATC TGT CAC TCT TTC TCG ACA CGG TGG CTT A- 3’	4.5 ± 1.2	0.82
aptTOV6	5’ - ATC CAG AGT GAC GCA GCA CGG CAC TCA CTC TTT GTT AAG TGG TCT GCT TCT TAA CCT TCA TCG ACA CGG TGG CTT A- 3’	29 ± 7	0.58
aptTOV7	5’ - ATC CAG AGT GAC GCA GCA CCA ACT CGT ACA TCC TTC ACT TAA TCC GTC AAT CTA CCA CTC TCG ACA CGG TGG CTT A- 3’	6.6 ± 2.3	0.19
aptTOV8	5’ - ATC CAG AGT GAC GCA GCA CCA GTC CAT CCC AAA ATC TGT CGT CAC ATA CCC TGC TGC GCC TCG ACA CGG TGG CTT A- 3’	17 ± 3	0.76
aptTOV9	5’ - ATC CAG AGT GAC GCA GCA GCA ACA CAA ACC CAA CTT CTT ATC TTT TCG TTC ACT CTT CTC TCG ACA CGG TGG CTT A- 3’	26 ± 10	0.06

Table 2-3. Overview of the binding studies performed on multiple types of cell lines. (- : no binding; + : less then a 5-fold signal increase; ++ : between 5 to 10 fold signal increase; +++ : between 10 to 100 fold signal increase; ++++ : more then 100 fold signal increase)

Aptamer	TOV-21G	CAOV3	HeLa	BCC	H23	HT-29	HCT-116	A172	Ramos	CEM	HL60	DLD-1
aptTOV1	+++	-	-	++	++	-	++	+	-	-	-	-
aptTOV2	+++	-	-	++	-	-	++	++	-	-	-	-
aptTOV3	+++	-	-	++	+	-	++	++	-	-	-	-
aptTOV4	+++	+	-	++	+	-	++	+++	-	+	-	-
aptTOV5	+++	+	+	+++	++	+	++	+++	-	+++	-	-
aptTOV6	++++	-	-	++	++	-	++	+++	-	++	-	-
aptTOV7	+++	-	-	++	++	-	++	+++	-	++	-	-
aptTOV8	+++	-	-	++	++	-	++	+++	-	+	-	-
aptTOV9	+++	-	-	++	++	-	++	++	-	-	-	-

CHAPTER 3
METHOD DEVELOPMENT FOR APTAMER TARGET IDENTIFICATION

Introduction

Ovarian cancer is the most lethal gynecological malignancy, with an annual mortality rate of around 140,000 per year globally[110]. The average relative survival rate for ovarian cancer is 43.9%, a number that can dramatically change depending on the stage (Stage IA: 94.0%; Stage 4: 17.9%)[4]. Unfortunately, only 2% of all ovarian cancer is detected at stage I and better methods are needed for early detection[3]. A method that is showing great promise to facilitate early cancer detection is biomarker detection in the peripheral fluids[16]. A cancer biomarker is understood to be a protein that is circulating in the blood, allowing simple blood tests to determine whether a patient has cancer by its presence or absense[22]. However, there is a lack of good biomarkers that can be used for the early detection of ovarian cancer[111]. So far, only CA125 is recognized by the FDA as a marker that may indicate the presence of ovarian cancer, but with limited specificity[10]. In ovarian cancer, there is therefore a need for more biomarkers, in order to increase specificity and sensitivity towards early detection of this malignancy.

As described in Chapter 1, Systematic Enhancement of Ligands through EXponential enrichment of aptamers on living cells (Cell-SELEX) is touted as a potentially important tool for finding ligands that bind to biomarker leads[44,89-91] and is becoming more and more a routine procedure[79,86-88].

The most important advantage of cell-SELEX is that no prior knowledge about the molecular profile of the cell is required[60,112]. The selected aptamers can be used in applications without further knowledge when a good target/non-target model is chosen.

When an interesting aptamer is found, the question about the identity of the aptamer target presents itself. Since approximately 50% of the mass of a eukaryotic cell membrane consists of protein, proteins are the prime candidates as aptamer-selected cell antigens[113]. Several research groups have attempted to solve the problem of aptamer target identification, but progress has been hampered, as the identification of individual aptamers' target is finicky and intricate. Thus, there is still a clear need for an easy and robust method for elucidating the protein targets of an aptamer generated from Cell-SELEX. Successful ventures towards such a technique have been performed on aptamers[89,91] and whole pools[90]. These methods cannot be applied to all aptamers[91], and are laborious and expensive, as they require many optimization steps. Some of these optimization steps include the insertion of nucleotides that can crosslink with the target without loss of binding, which requires extensive optimization to retain the aptamer's binding ability[89].

Various biochemical techniques are used to study the interactions between DNA and proteins, including Chromatin immuno-precipitation (ChIP). In this technique, formaldehyde fixes the protein and DNA within 2Å of their respective binding groups at the primary amines of both moieties. The protein is then immuno-precipitated, and the DNA is further analyzed by PCR[114].

The reaction mechanism for the protein-DNA crosslinking is described in Scheme 3-1[115]. The advantages of formaldehyde crosslinking include: the small size of the resulting crosslink, which allows only interacting DNA and protein to react with each other and the ability to reverse the crosslink once the molecule of interest (usually a short piece of DNA) is extracted[116]. This allows for easy analysis after the extraction

and precipitation step. ChIP is a valuable tool for various genetic studies, as transcription factors can only be retrieved from the pieces of DNA (e.g., promotor regions) with which they interact[117].

An innovative recent technique has used the hybridization of genomic DNA to elucidate the proteins involved in the binding of specific loci. Several proteins that play a role in telomere interactions were identified, leading to the identification of proteins involved in the alternative lengthening of telomeres[118]. This study led us to the hypothesis that studies of interactions between aptamers and their targets are feasible with formaldehyde crosslinking, and the aptamer could serve as a valuable capturing agent for its target in this manner.

Thus, implementing the basic principles of ChIP on aptamer-protein interactions can be a very attractive strategy for the elucidation of the target of an aptamer. As described in Chapter 2, a series of aptamers have been selected against ovarian cancer, and they bind to an ovarian clear cell carcinoma cell line (TOV-21G), but not to cervical cancer cells (HeLa)[119]. In order to acquire a deeper understanding about the underlying molecular differences between these two types of malignancies, it is important that the targets for the aptamers be elucidated. This chapter describes the method that was used to elucidate the target for aptamer TOV6. The target determined to be StIP1 by mass spectrometry, and the assignment was further validated bysiRNA knockdown and antibody binding studies with the help of flow cytometry.

Materials and Methods

Instrumentation and Reagents

AptTOV6 (5’ - ATC CAG AGT GAC GCA GCA CGG CAC TCA CTC TTT GTT AAG TGG TCT GCT TCT TAA CCT TCA TCG ACA CGG TGG CTT A- 3’) and library

(N_{76}) were synthesized by standard phosphoroamidite chemistry using a 3400 DNA synthesizer (Applied Biosystems) on biotin or desthiobiotin-CPG for protein capture (Glen Research) and were purified by reversed-phase HPLC (Varian Prostar) as described in Chapter 2. StIP1 antibody M33 (clone 2E1) was purchased from Abnova (Taiwan), and biotinylated with the EZ-link® Sulfo-NHS-LC-Biotin kit (Pierce). The capturing of the aptamer hybrids was done using Dynabeads® M-280 Streptavidin (Invitrogen).

Cell Culture and Buffers

The TOV-21G cell line was purchased from the American Type Cell Culture (ATCC) and was maintained in culture with MCBD 105: Medium 199 (1:1), supplemented with 10% FBS and 100 IU/mL Penicillin-Streptomycin. TOV-21G was cultured at 37ºC in a 5% CO_2 atmosphere. For aptamer binding, cells were washed after non-enzymatic dissociation buffer treatment (Sigma) and after incubation with washing buffer (WB) containing 4.5 g/L glucose and 5 mM $MgCl_2$ in Dulbecco's phosphate buffered saline with $CaCl_2$ and $MgCl_2$ (PBS - Sigma). Binding buffer (BB) used for aptamer binding was prepared by adding yeast tRNA (0.1 mg/mL, Sigma) and BSA (1 mg/mL, Fisher) to the washing buffer to reduce non-specific binding. All chemicals used in the buffers were purchased from Sigma, unless otherwise specified. For crosslinking, a 1% formaldehyde (Fisher) in PBS solution was used. The cell lysis buffer contained 2% Triton-100X (Fisher), 1.5% Nonidet (Fisher) and 0.5% cholate in dideionized water. For the washing of the magnetic beads, a 10mM HEPES NaOH buffer (pH 7.8) was used with 100mM NaCl, 2mM ethylenediaminetetraacetic acid (EDTA), 1mM ethylene glycol tetraacetic acid (EGTA) , 0.2% sodium dodecyl sulfate(SDS) and 0.1% sodium lauroyl sarcosinate (SLS)(All from Sigma) (PI washing buffer). As an elution buffer,

12.5mM biotin in 7.5 mM HEPES NaOH in 75mM NaCl, 1.5% EDTA, 0.5%EGTA , 0.15% SDS and 0.075% SLS was used. As a crosslink reversal solution, 250mM Tris buffer (pH 8.8) with 2% SDS and 0.2M mercaptoethanol was used. All solutions, except for WB and BB contain 0.1mM phenylmethanesulfonylfluoride (PMSF). For SDS-PAGE, the SilverQuest staining kit was used to visualize the bands (Invitrogen).

Aptamer Target Purification for Protein Identification

TOV-21G cells (10^8) were incubated with 200pmol of desthiobiotin-oligonucleotide according to the scheme in Figure 3-1. The cells were washed and the aptamer was cross-linked to the cells by incubating for one minute in 1% formaldehyde PBS. Washing the cells three times at 4ºC in PBS diluted the formaldehyde, as quenching the reaction with lysine (the usual method for quenching formaldehyde) can make MS analysis nearly impossible[118]. Subsequently, the cells were lysed in a dounce homogenizer for two minutes (75 strokes per minute) in lysis buffer. The lysate in lysis buffer was incubated overnight at room temperature in the presence of 200µg of magnetic beads. The beads were then washed with PI washing buffer on a magnetic stand until any remaining membrane was washed from the beads. Once the beads were clean, the protein aptamer hybrid was eluted by incubation for 1h with elution buffer. TCA precipitation and acetone washing were used for further purification of the eluate protein fraction. The pellet was dissolved in crosslink reversal buffer and boiled for 1h, after which the sample was loaded on an SDS-polyacrylamide gel. Bands of interest were sent for MS analysis.

StIP1 siRNA Knockdown

Hs_StIP1 5, 6, 10 and 11 (QAIGEN) was used with the HiPerfect starter kit (QAIGEN), as directed, on 0.8×10^5 TOV-21G cells. Aptamer binding was verified 72

hours past transfection. The transfection efficiency was tested by a cell death positive control and a scrambled siRNA negative control to ensure proper cell viability and delivery efficiency, both provided by the supplier.

Antibody Biotinylation

StIP1 antibody (50mg) was incubated with a 20 molar excess of sulfo-NHS-biotin reagent, as described in the supplier's guidelines (Pierce).

Aptamer Blotting

rhStIP1 and BSA was blotted as described[131]. In brief, 20µg of the respective protein is blotted in a nitrocellulose membrane and blocked in 4% non-fat milk in PBS containing 0.05% (v/v) Tween 20 and 1 mM EDTA. 250nM aptamer solution incubated on the membrane, after which streptavidin-horseradish peroxidase is added. The complex can then be visualized with the ECL plus Western blot system (GE lifesciences).

Results

Outline of the Aptamer Mediated Protein Identification Procedure

In order to determine the identity of the protein binding to aptTOV6, it was hypothesized that the interaction between the aptamer and the protein had to be fixed in order to ensure easy extraction with detergents without loss of the binding between the aptamer and its target. Formaldehyde is an easy-to-use and proven cross-linker, which is resistant to surface-active agents (surfactants) and thus maintains the interaction between DNA and protein[120]. To ensure efficient capture of the protein with the aptamer, the exposure time of the formaldehyde with the aptamer bound cells was optimized. This led to the procedure schematically shown in Figure 3.2.

In essence, aptamers were bound to the cells of interest and fixed in formaldehyde. After dilution of the cross linker, the membranes were dissolved in lysis buffer and the protein-aptamer hybrid was recovered using streptavidin covered magnetic beads. These beads were washed 5 times with PI washing buffer to remove any remaining membranes. The aptamer-protein hybrid was then eluted and identified using mass spectrometry (MS). Elution of the hybrid was facilitated by desthiobiotin conjugated at the 3' end of the aptamer. Since the K_d of streptavidin is about 350 times higher (i.e., lower affinity) towards desthiobiotin ($K_d = 3.5 \times 10^{-13}$) than towards biotin($K_d \sim 1 \times 10^{-15}$)[121], the desthiobiotin conjugated protein-aptamer hybrid elutes more rapidly then other biotin-containing proteins in the cell lysate. To ensure that the sample would be MS compatible, surfactants and salts that might cause ionization supression[122] were removed by TCA precipitation. The resulting pellet was dissolved and boiled for 2 hours to reverse the crosslink between the aptamer and the protein[115]. This sample is then separated by SDS-PAGE, where the differential band was identified.

Aptamer TOV6, which was selected from the ovarian clear cell adenocarcinoma cell line TOV-21G, was used for target identification. The aptamer does not bind to the ovarian serous carcinoma cell line CAOV3 or to the cervical cancer cell line HeLa. It was first necessary to verify the binding of the desthiobiotin aptamer due to possible formaldehyde-induced denaturation of the target protein or potential desthiobiotin reactions with formaldehyde. Therefore, an elution study with streptavidin labeled PE-cy5.5 was performed using flow cytometry. In this study, the labeling efficiency was studied by comparing the signal from desthiobiotinylated aptamer with the signal from aptamer bound cells after formaldehyde treatment and after incubation with biotin. If the

biotin is able to compete with desthiobiotin, the cells lose their fluorescent signal. As can be found in Figure 3.2, the desthiobiotin can easily be eluted by biotin and formaldehyde does not prevent aptamer binding.

SDS-PAGE of the Bead-Binding Fraction of Whole Cell Lysate

The protein fractions that bound to aptamer TOV6 with or without formaldehyde treatment were studied by SDS-PAGE. As can be seen from the SDS-polyacrylamide gel in Figure 3-3, a clear band at around 78kDa appears in the crosslinked sample, and no bands are detected in the non-crosslinked section. This indicates that formaldehyde is effective in maintaining the aptamer-protein interaction, which would otherwise be lost in the extraction process. Optimization of the crosslink time was imperative however, as excessive crosslinking can lead to significant background, while insufficient crosslinking does not yield any protein fraction (Figure 3-4). Because of the lower affinity of desthiobiotin towards streptavidin compared to biotin, the aptamer-protein hybrid can be eluted prior to other biotin-containing protein fractions, significantly reducing the background that originates from biotin containing enzymes or other non-specific streptavidin interactions[91]. As shown in Figure 3-3, a protein band (band 8) was effectively eluted of the beads, as a result of the replacement of biotin by desthiobiotin. After incubating longer at 65ºC (to increase the release of biotin containing protein), only trace amounts of protein can be observed (band 9). This indicates that the aptamer-protein hybrid can easily be removed from of the cell lysate and specifically recovered from the beads.

There was one distinct band that could be eluted, but couldn't be detected in the non-crosslinked sample. This band was excised and sent for MS analysis. Stress Induced Protein 1 (StIP1)[123] was the top result for both samples sent for MS analysis.

The MS results from the service lab can be found in Figure 3-5. StIP1's alias is Hsp70-hsp90 Organizing Protein (HOP), a protein that is known to play a regulative role as a co-chaperone in the heat shock protein (hsp) 90 chaperone complex[124]. Recently, the protein was found in the membrane of various cell types, including ovarian cancer cell lines (TOV-21G, ES2, SKOV3, OVCAR3)[125], the pancreatic cell line Panc-1[126] and the glioblastoma cell line A172[127].

Confirmation of the Binding of aptTOV6 to StIP1

In order to confirm the binding of AptTOV6 to the StIP1, TOV-21G was incubated with biotinylated StIP1 antibody and analyzed with flow cytometry. The effect of AptTOV6 and StIP1 antibody binding to STIP1 siRNA treated TOV-21G cells was investigated with the help of flow cytometry.

The target of aptTOV6 was confirmed with the help of StIP1 siRNA knock down, the effect of the binding of aptTOV6 has been investigated by using predesigned and experimentally validated siRNAs for StIP1[128]. StIP1 is a known membrane protein in TOV-21G and also in A172[125,127,129]. Figure 3-6 demonstrates that the binding of aptTOV6 decreases when incubated with StIP1 siRNA-treated cells, while the binding is unaffected against scrambled siRNA-treated TOV-21G cells. Figure 3-7 shows similar results for the glioblastoma cell line A172. In both experiments, the expression of PTK7[91,130] was also tested to demonstrate that the general expression of membrane protein was unaffected (PTK7 is a protein that is generally expressed in many cell lines, including TOV21G) (Figures 3.8 and 3.9). Finally, the binding of the StIP1 antibody M33 was abrogated by siRNA-induced knockdown of StIP1 expression (Figure 3.10B). All these results prove that the aptamer TOV6 binds to the membrane protein StIP1. Furthermore, rhStIP1 was immobilized on a nitrocellulose membrane, and the TOV6-

biotin blot[131] was able to stain the membrane, while it could not stain immobilized BSA (Figure 3.11).

Discussion

An efficient method for the identification of an aptamer's binding protein has been developed. Formaldehyde in combination with a desthiobiotin-conjugated aptamer was proven to be efficient and specific (i.e. only aptamer-interacting proteins were extracted) in order to identify the proteins that aptamers target. The crosslinking of the DNA aptamer protein interaction proved to be an efficient and straightforward way for the efficient extraction of membrane proteins. Without crosslinking, proteins that are imbedded in the lipid bilayer of the cell lose their tertiary structures when the lysis buffer is employed to remove the proteins from of the bilayer (Figure 3-3)[28,113]. Therefore, in order to be able to extract an unknown protein with an aptamer and keep the tertiary structure intact, a crosslinker is warranted. The use of surfactants in the lysis buffer can also disrupt the interaction between the aptamer and the target, as the interaction between protein and target are governed by Van der Waals interactions and hydrogen bonds. Thus, in order to prevent loss of binding between the aptamer and the unknown target, it is imperative that the interaction be fixed.

Our lab has previously developed a method that is based on a similar concept, but its application requires the use of specialized nucleotides, which, when incorporated into the aptamer at the key positions, destroy the binding of the aptamer with the target. At the other extreme, if the crosslinking modality is bonded at the wrong location, no protein can be extracted, as the necessary fixation between aptamer and protein would be absent. Thus, by using formaldehyde, there is no need for separate optimization of

the probe for target identification. The only optimization that is needed is the reaction time between the aptamer and the protein (Figure 3-4).

Formaldehyde only crosslinks amines that are in close proximity of each other. In this manner, only molecules that truly interact with each other can be extracted with the aptamer[120]. Furthermore, the crosslinked protein-aptamer hybrid is considerably more soluble then the membrane protein in itself, as the phosphate backbone of the aptamer provides extra anionic groups that increase the solubility of the protein. One of the biggest advantages of SELEX, the selection of ligands for membrane proteins in their native forms, could have been problematic for aptamer-target elucidation, as the lysis of the cells potentially leads to the refolding of the target, that may lead to the loss of interaction. The use of a crosslinker like formaldehyde has solved this problem. Once the aptamer is fixed to its target, the maintenance of the tertiary structure of the membrane protein becomes irrelevant, as identification solely depends on the individual peptides analyzed. Furthermore, the use of formaldehyde does not prevent the analysis of the protein, as the formaldehyde-mediated crosslink is unstable above 72ºC, and therefore reversible, allowing the release of the aptamer at the point where the protein needs to be analyzed[132]. Furthermore, because the conventional biotin conjugation is replaced with a desthiobiotin conjugation, the protein-aptamer hybrid can easily be removed from the streptavidin by free biotin competition (Figure 3-2).

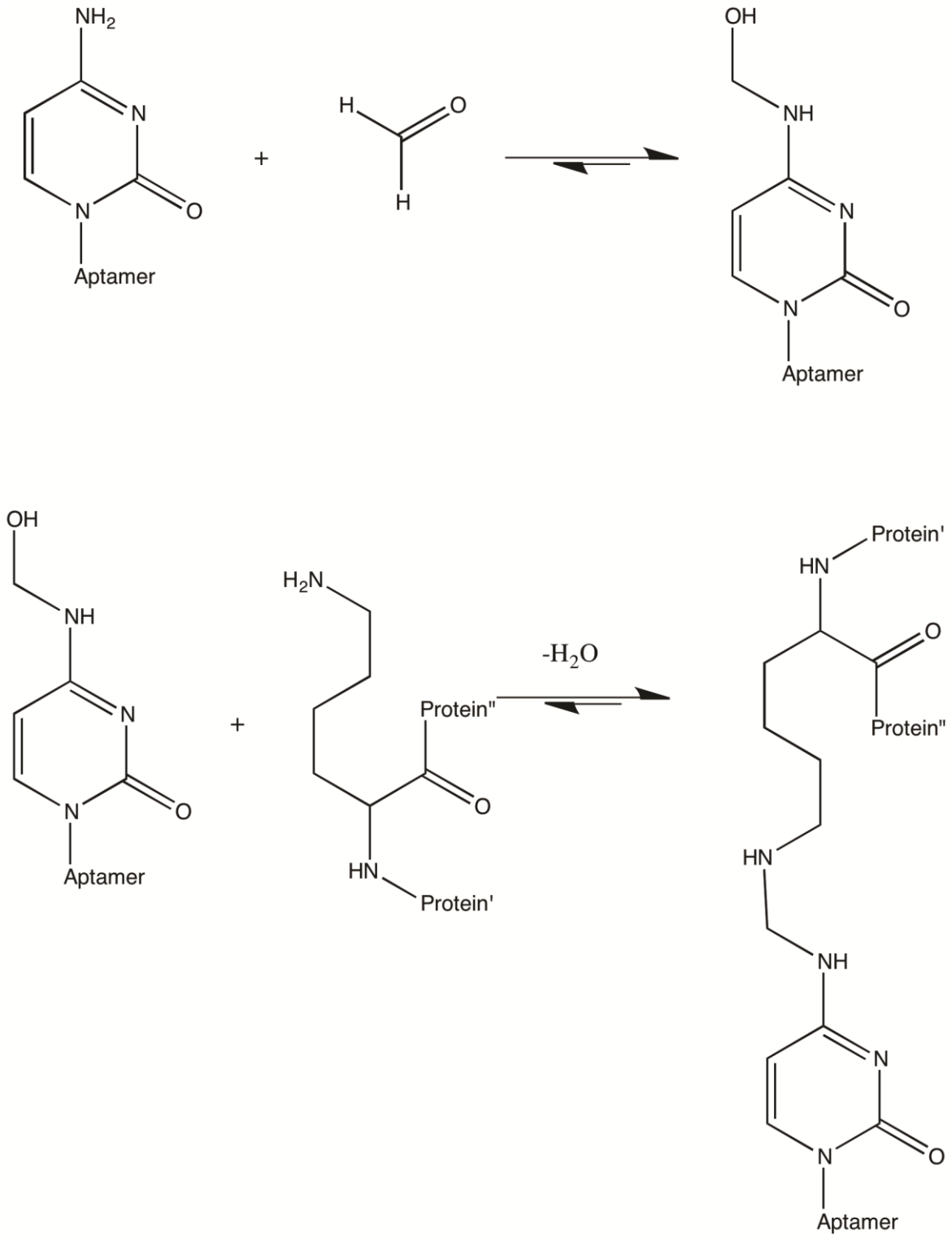
That StIP1, a membrane protein, is in cancerous cells, including TOV-21G, is supported in the literature. Wang *et al.* found StIP1 in a 2D-gel that was run as a comparison between tumor interstitial fluids and non-tumor interstitial fluids from ovarian cancer patients. They confirmed the presence of StIP1 in the cell membrane of several

ovarian cancer cell lines, including SKOV3 and TOV-21G. They also performed a pilot study, and the results suggested that StIP1 can be used in combination with CA125 for the early detection of ovarian cancer in patients[125]. Shin *et al.* showed that chaperone and co-chaperone proteins can commonly be found in the membrane of cancerous cells[133]. Although StIP1 was not found in their findings, many other (co-) chaperone proteins were extracted from the membrane of ovarian cancer patient tissue, demonstrating that chaperone proteins can be retrieved from the membrane of cancerous cells. In ovarian cancer, StIP1 is found in the cytoplasmic membrane, where it interacts with ERK and induces proliferation[125]. In pancreatic cancer it is found to play a crucial role in metastasis[126,134].

Conclusions

This work demonstrates that protein identification of aptamer targets can become a powerful tool in the development of new cancer biomarkers. SELEX can be employed for the selection of aptamers that are specific for diseased cells. The target identification of these proteins can lead to deeper insight of the mechanism that drives the pathology of a cancer. Due to the nature of the selection procedure (i.e., target cell is cancerous, negative cell is healthy), it is very likely that the targets binding these disease-specific aptamers, can lead to new biomarkers or at least to deeper insight in the oncopathology of the cancer cells studied.

These biomarkers can not only be used for early diagnosis, but can also be useful for the development of new patient care strategies, as new pathways can be identified that can be targeted with small molecule drugs, for a more efficient form of therapy and drug discovery[135].



Scheme 3-1. The chemistry of formaldehyde mediated DNA-Protein crosslinking. In this example cytidine crosslinks to a lysine

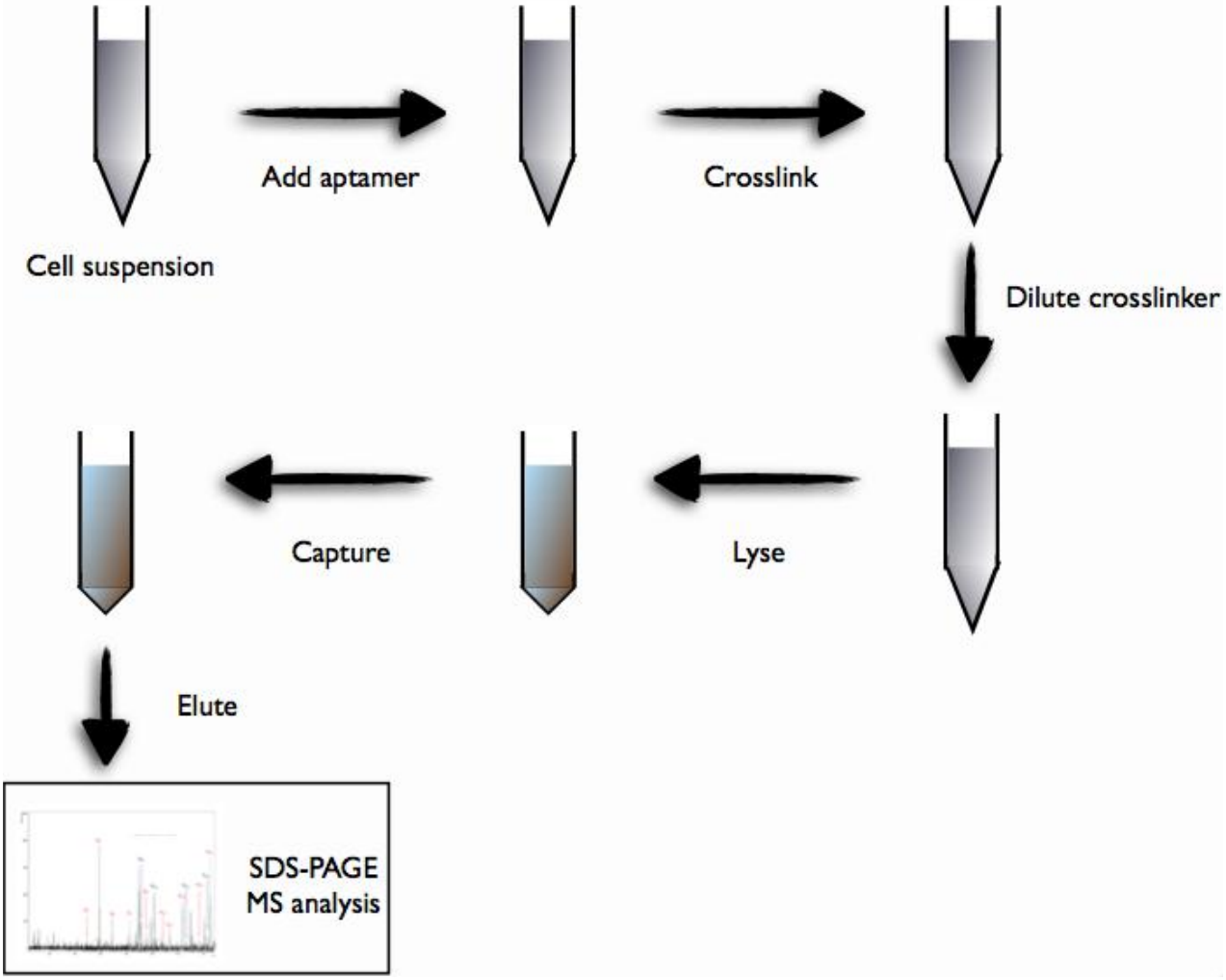


Figure 3-1. General procedure for protein identification with the use of aptamers. The aptamer is bound to the cells. Washing the cells in WB removes non-binding aptamer. Then the 1% formaldehyde solution is added to the cells and allowed to crosslink for 2 minutes, after which the formaldehyde is diluted by washing in WB. The cross-linked cells are homogenized in lysis buffer, magnetic beads are added, and the aptamer-hybrid is captured on the beads. The beads are further washed, after which the hybrid is eluted by biotin elution. The crosslink is reversed and the protein fraction is dissolved and separated by SDS-PAGE, after which the resulting band is analyzed

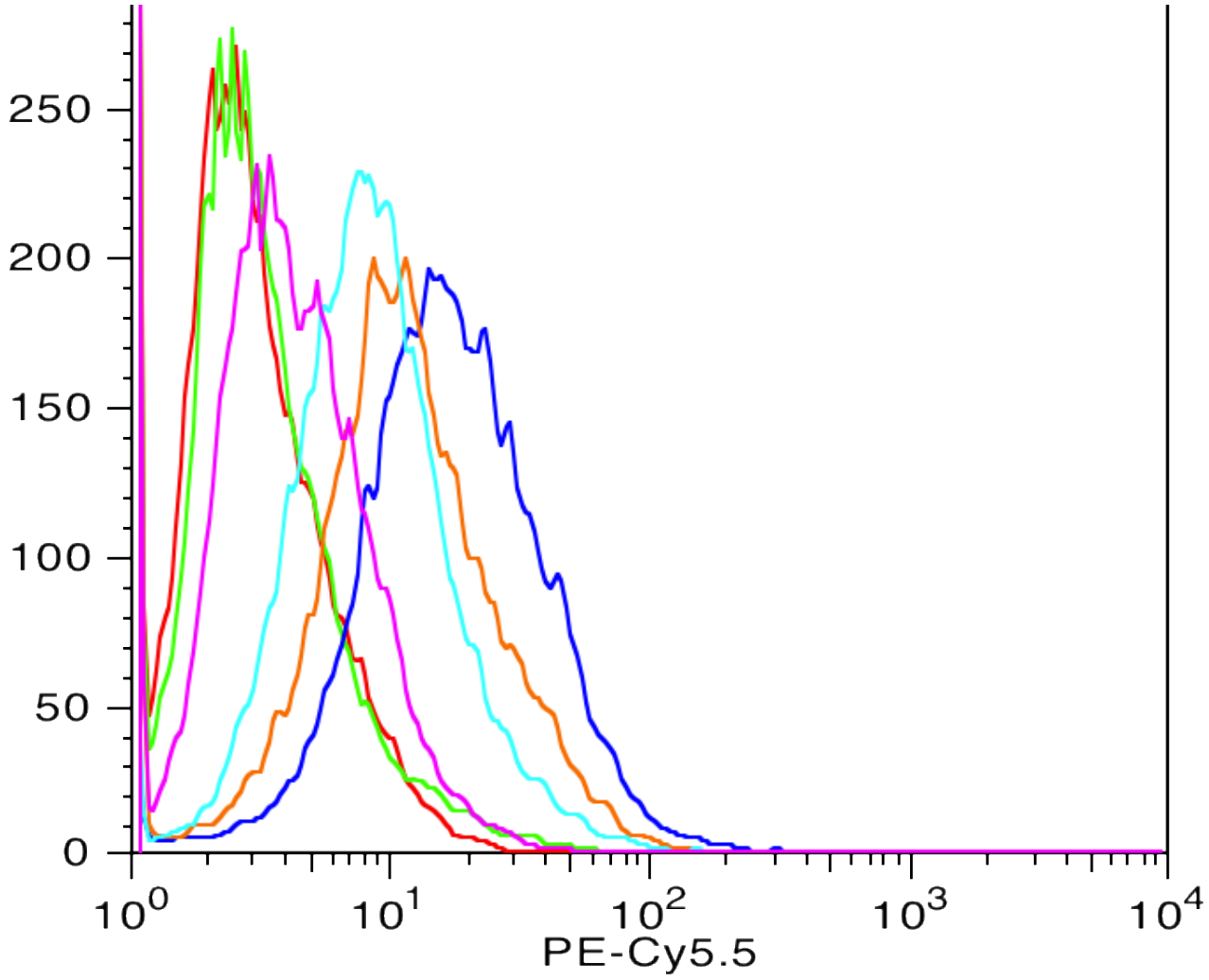


Figure 3-2. Study of the effect of formaldehyde on streptavidin binding and biotin elution of the desthiobiotin conjugated aptamer. Red: TOV-21G; Green: Library; Dark blue: TOV6; Orange: TOV6 in 1% CH_2O; Light blue: TOV6 after 30 min in 5mM biotin solution; Magenta: TOV6 in 1% CH_2O after 30 min in 5mM biotin solution

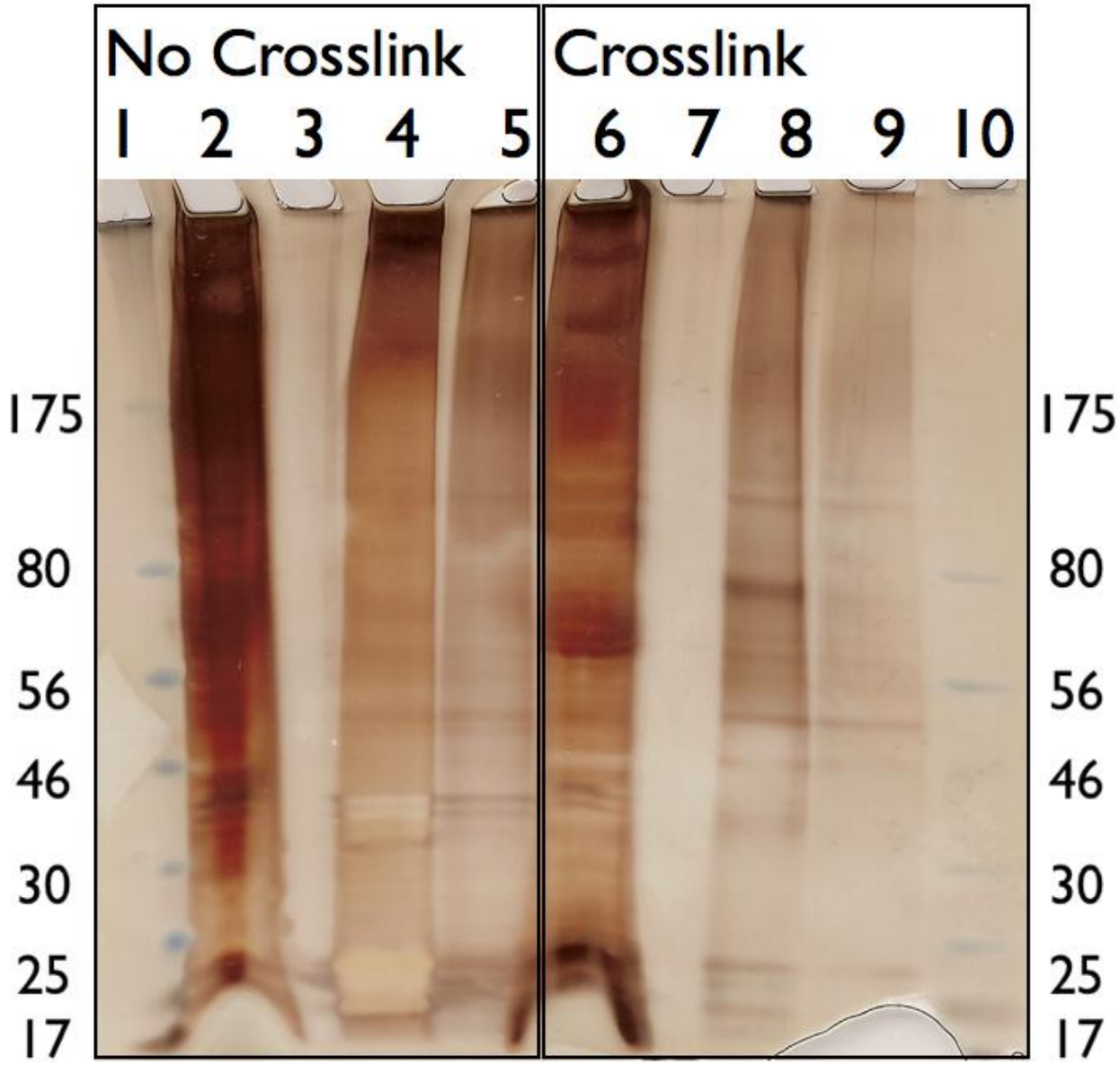


Figure 3-3. Silver staining of the material obtained from the aptamer mediated protein purification for aptTOV6. Stains for cells in presence or absence of formaldehyde are shown. 1 & 10: bands from ladder; 2: unbound fraction from non-crosslinked sample; 3: Fraction that bound to naïve library from non-crosslinked sample; 4: fraction that bound to aptTOV6, eluted with 1 h of 5mM biotin in PBS at 37ºC from non-crosslinked sample; 5: fraction that remained on the beads after elution at 37ºC, obtained by eluting at 65ºC from non-crosslinked sample; 6: unbound fraction from crosslinked sample; 7: fraction that bound to naïve library from crosslinked sample; 8: fraction that bound to aptTOV6, eluted with 1 h of 5mM biotin in PBS at 37ºC from crosslinked sample; 9: Fraction that remained on the beads after elution at 37ºC, obtained by eluting at 65ºC. The crosslink time here was set at 2 minutes

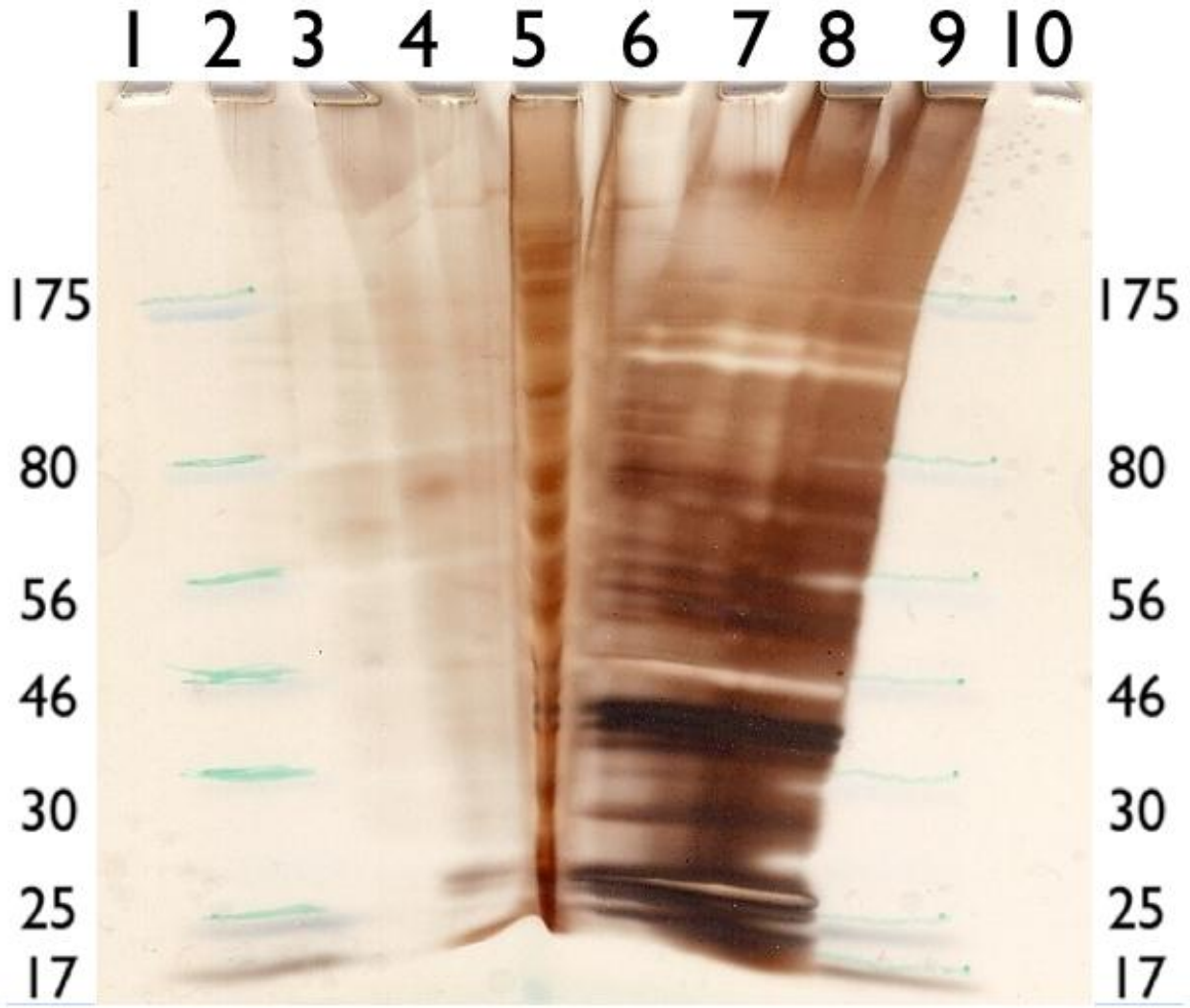


Figure 3-4. Comparison of sample runs with insufficient or excessive crosslink times. Bands 2-4 represent samples from 1 minute incubation with formaldehyde. Bands 5-9 are from samples after 10 minutes incubation. Bands 1 and 10 represent ladder; 2: the eluted sample after 1h of incubation at 65ºC after 37ºC incubation for TOV6; 3: the eluted sample at 37ºC after 1h for TOV6; 3: the eluted sample at 37ºC after 1h for Library; 4: bead fraction after elution at 37ºC and 65ºC 5: Unbound fraction; 6: bead fraction after elution at 37ºC and 65ºC; 7: elution fraction after 1h at 37ºC for TOV6, 8 and 9: the eluted sample after 1h of incubation at 65ºC after 37ºC incubation for TOV6. The samples in bands 3 and 7 show the presence of the protein of interest which was observed in Figure 3-3

Protein matches			
34	62	3.516	LMNA_IPI:IPI00021405.3
32	35	3.351	FLNA_IPI:IPI00302592.2
23	28	3.292	HNRNPM_IPI:IPI00171903.2
21	25	2.820	STIP1_IPI:IPI00013894.1
20	25	3.708	CKAP4_IPI:IPI00141318.2
19	22	3.274	RPN1_IPI:IPI00025874.2
18	27	3.700	TKT_IPI:IPI00643920.3
16	22	4.042	NOP56_IPI:IPI00411937.4
16	18	2.848	DDX5_IPI:IPI00017617.1
13	21	3.426	XRCC5_IPI:IPI00220834.8
13	15	2.708	LMNB1_IPI:IPI00217975.4
12	14	3.345	CLTC_IPI:IPI00024067.4
12	14	3.246	RPA1_IPI:IPI00020127.1
12	13	2.994	MATR3_IPI:IPI00017297.1
10	10	3.684	HSPA8_IPI:IPI00003865.1
10	18	3.306	NCL_IPI:IPI00183526.6
10	10	3.064	MYEF2_IPI:IPI00386418.4
9	9	3.299	XRCC6_IPI:IPI00644712.4
9	9	3.266	DDX17_IPI:IPI00023785.7
9	9	2.802	COPG_IPI:IPI00783982.1
8	8	2.770	LMNB2_IPI:IPI00009771.6
8	9	2.634	EEF2_IPI:IPI00186290.6
7	11	3.953	NCL_IPI:IPI00444262.3
7	7	3.711	STXBP2_IPI:IPI00019971.4

Protein matches				
5	5	2.774	0.00	STIP1_IPI:IPI00013894.1
4	4	3.198	0.00	ALB_IPI:IPI00022434.4
4	4	3.151	0.00	TKT_IPI:IPI00643920.3
3	3	4.030	0.00	HSPA1A_IPI:IPI00304925.5
3	3	3.677	0.00	HSPA8_IPI:IPI00003865.1
3	3	3.663	0.00	RPN1_IPI:IPI00025874.2
3	3	3.184	0.00	NCL_IPI:IPI00444262.3
3	3	3.161	0.00	ACTA2_IPI:IPI00008603.1
2	2	5.055	0.00	IPI:IPI00550731.2_IPI:IPI00550731.2
2	2	3.716	0.00	HSPA5_IPI:IPI00003362.2
2	2	3.655	0.00	TUBA1C_IPI:IPI00166768.3
2	2	3.392	0.00	ACTBL2_IPI:IPI00003269.1
2	2	3.374	0.00	ANXA2P2_IPI:IPI00334627.3
2	2	3.161	0.00	GAPDH_IPI:IPI00219018.7
2	2	2.927	0.00	PKM2_IPI:IPI00220644.8
1	1	4.008	0.00	ENO1_IPI:IPI00465248.5
1	1	3.871	0.00	FUS_IPI:IPI00221354.1
1	1	3.761	0.00	HSPA1L_IPI:IPI00301277.1
1	1	3.643	0.00	ECM1_IPI:IPI00003351.2
1	1	3.600	0.00	HSPA9_IPI:IPI00007765.5
1	1	3.471	0.00	HSPD1_IPI:IPI00784154.1
1	1	3.417	0.00	LMNA_IPI:IPI00021405.3
1	1	3.241	0.00	EEF1A2_IPI:IPI00014424.1
1	1	3.228	0.00	SDHA_IPI:IPI00217143.3

Figure 3-5. Proteins found in the aptamer TOV6 binding fractions. StIP1 was the only top protein found in both samples sent for analysis. Each column represents the protein hits for each sample

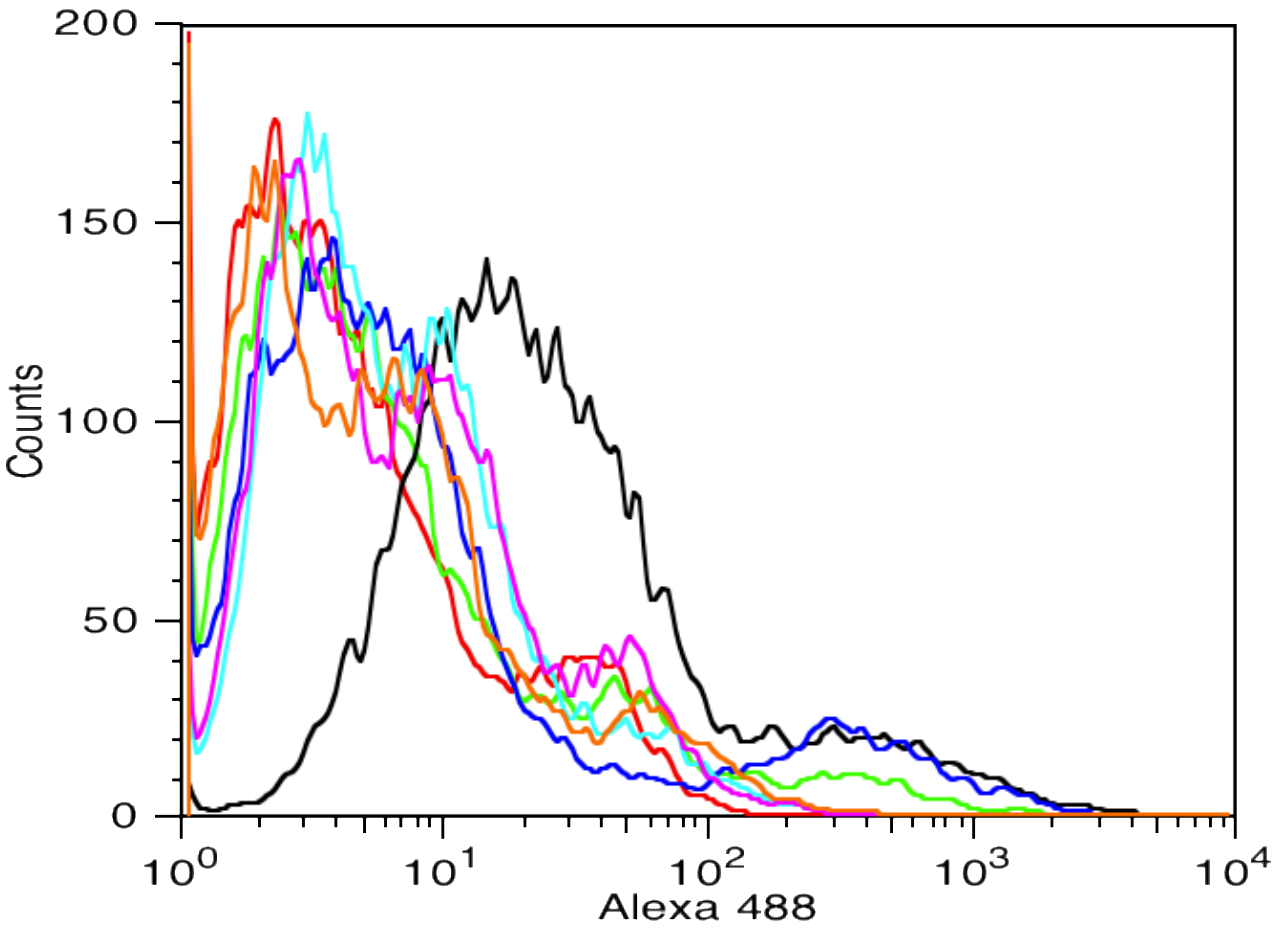


Figure 3-6. Silencing of StIP1 in TOV-21G cells. Red: Str. Alexa 488-only treated TOV-21G; Green: Library treated cells; Black: mock siRNA treated cells; Dark blue: StIP1 siRNA 5 Orange: StIP1 siRNA 6;Light blue: StIP1 siRNA 10;Magenta: StIP1 siRNA 11. The cells were tested for TOV6 binding after 72 h of siRNA treatment

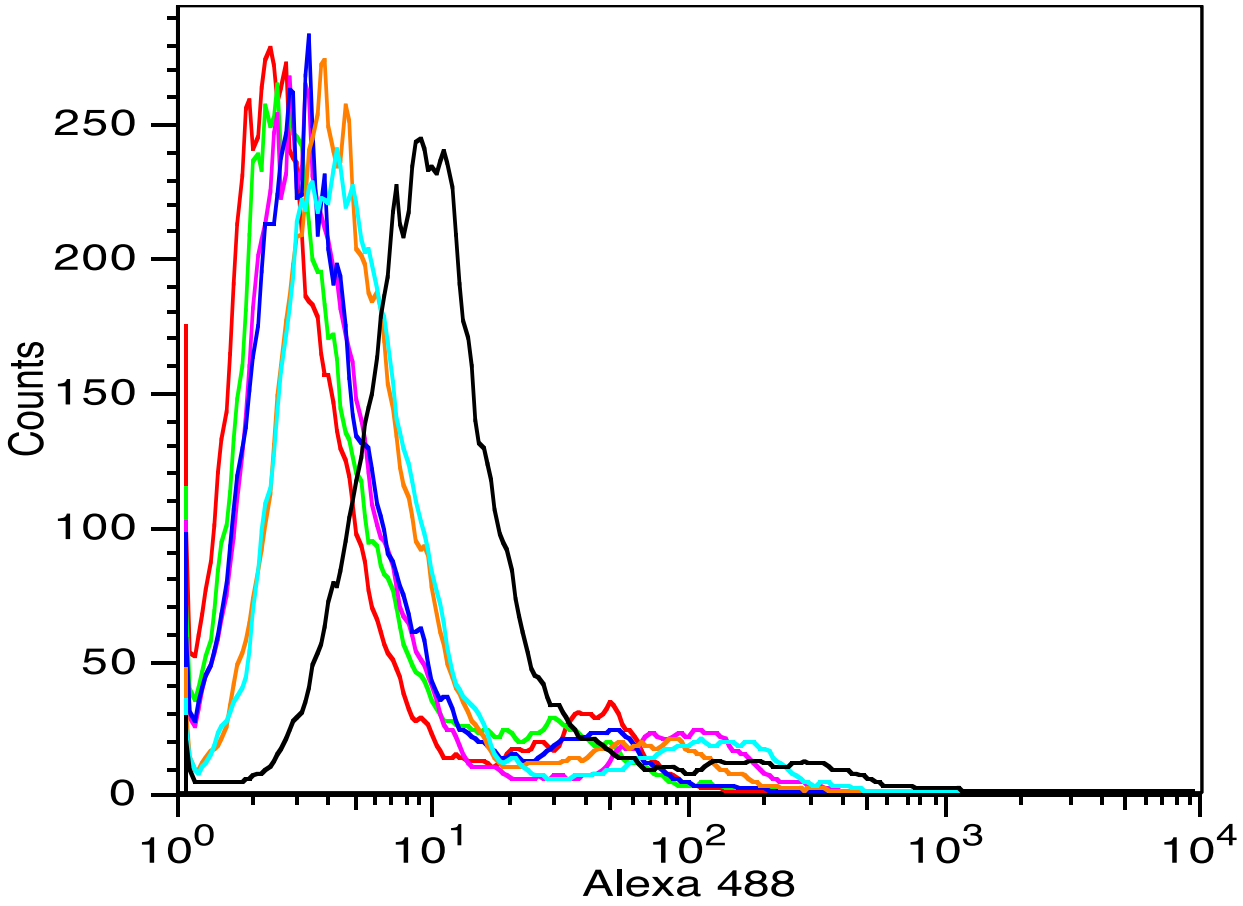


Figure 3-7. Silencing of StIP1 in A172 cells. Red: Str. Alexa 488-only treated A172; Green: Library treated cells; Black: mock siRNA treated cells; Dark blue: StIP1 siRNA 5 Orange: StIP1 siRNA 6; Light blue: StIP1 siRNA 10; Magenta: StIP1 siRNA 11. The cells were tested for TOV6 binding after 72 h of siRNA treatment

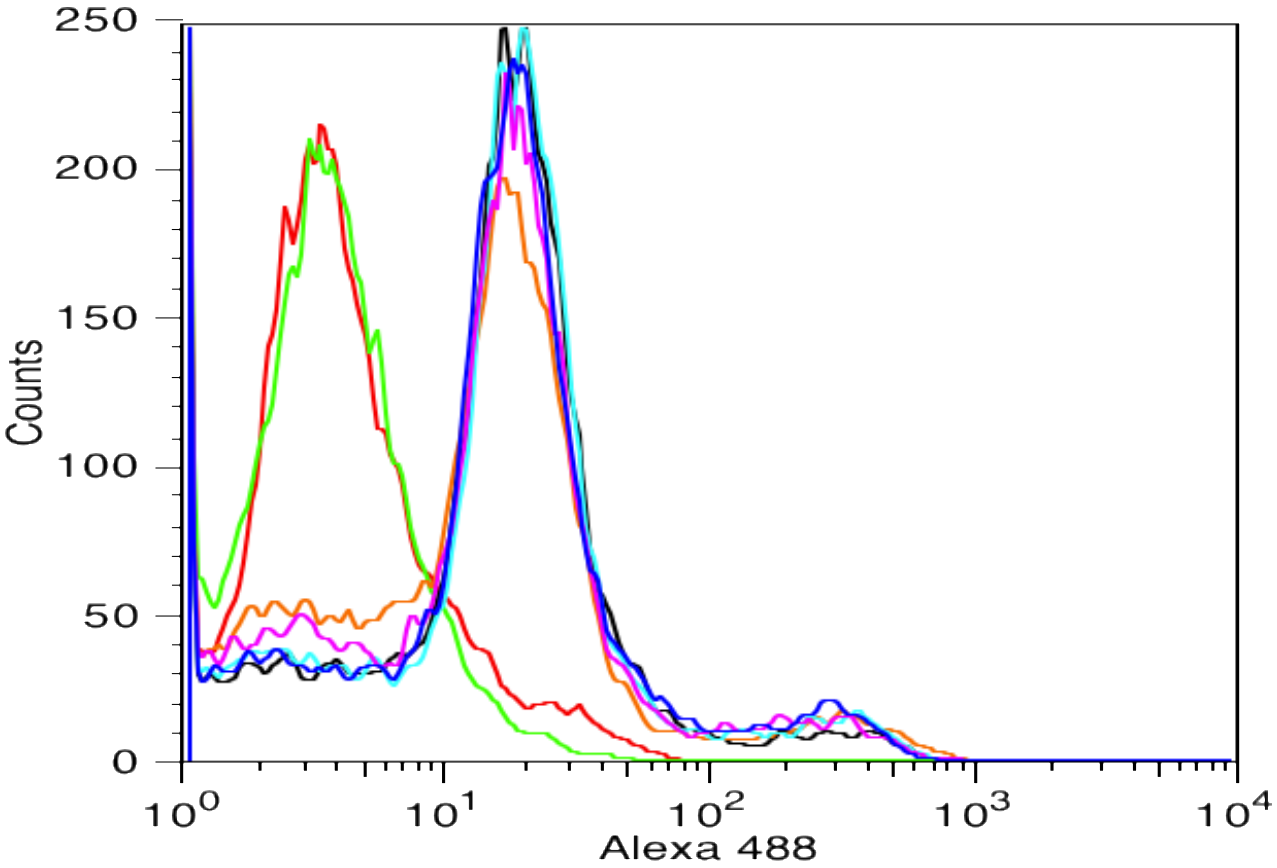


Figure 3-8. Absence of PTK7 silencing with StIP1 siRNA treatment in A172 cells. Red: Str. Alexa 488-only treated A172; Green: Library treated cells; Black: mock siRNA treated cells; Dark blue: StIP1 siRNA 5; Orange: StIP1 siRNA 6; Light blue: StIP1 siRNA 10; Magenta: StIP1 siRNA 11. The cells were tested for sgc8 binding after 72 h of siRNA treatment

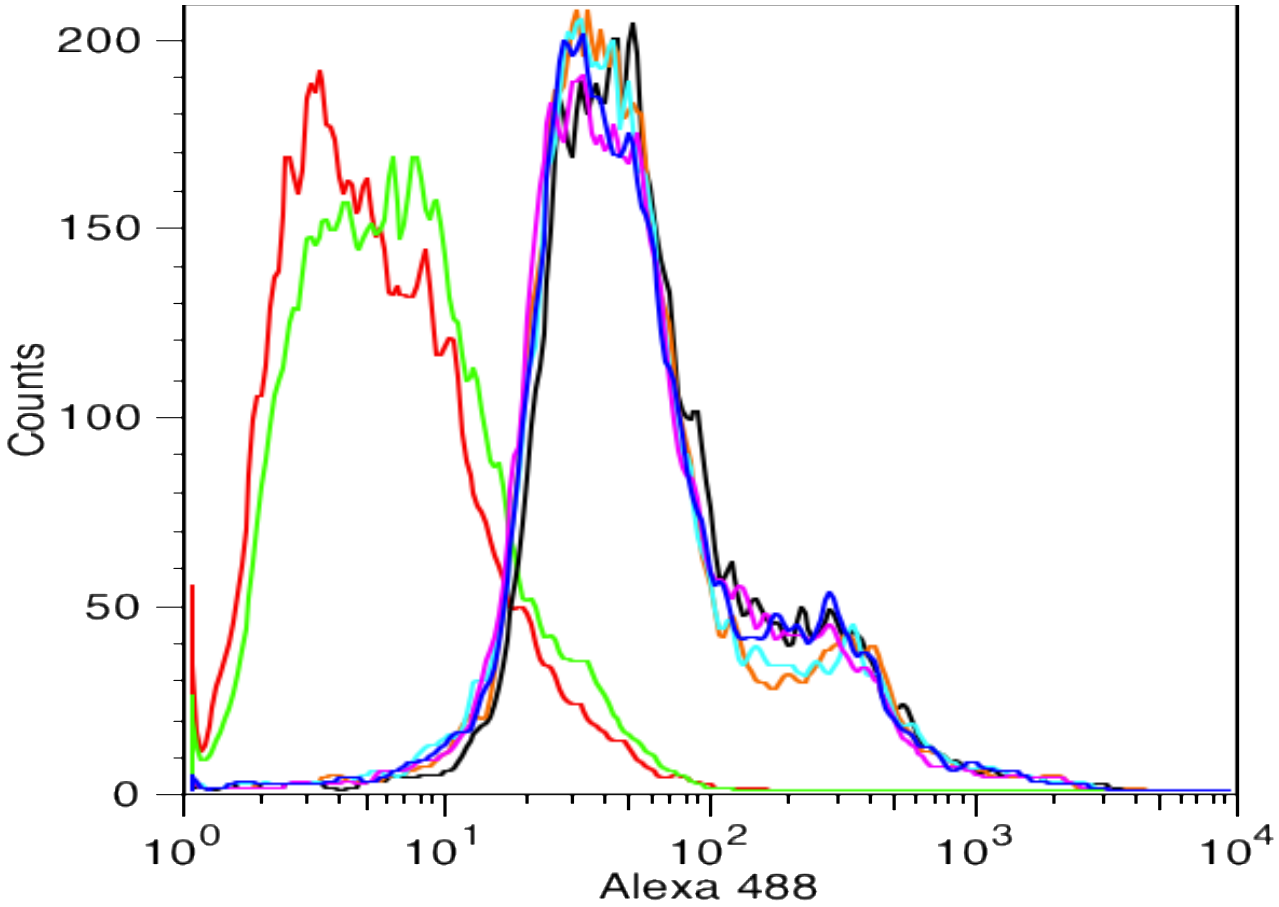
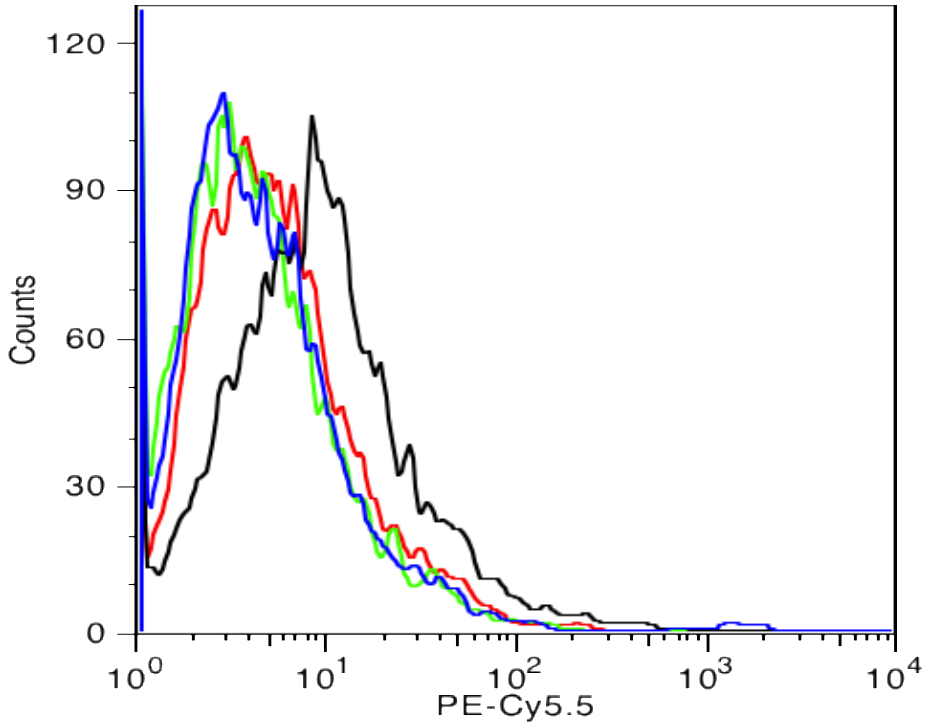
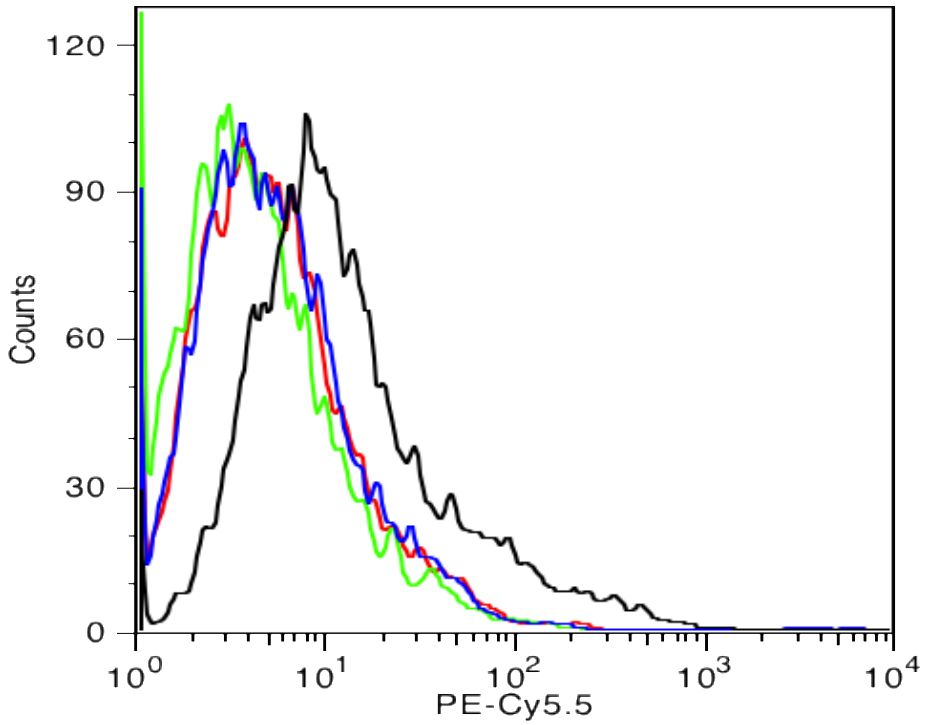


Figure 3-9. Absence of PTK7 silencing with StIP1 siRNA treatment in TOV-21G cells. Red: Str. Alexa 488-only treated TOV-21G; Green: Library treated cells; Black mock siRNA treated cells; Dark blue: StIP1 siRNA 5; Orange: StIP1 siRNA 6; Light blue: StIP1 siRNA 10; Magenta: StIP1 siRNA 11. The cells were tested for sgc8 binding after 72 h of siRNA treatment



(A)



(B)

Figure 3-10. A: The binding of AptTOV6 with (blue) and without (black) StIP1 knock down; B: The binding of StIP1 antibody M33 with (blue) or without (black) StIP1 silencing

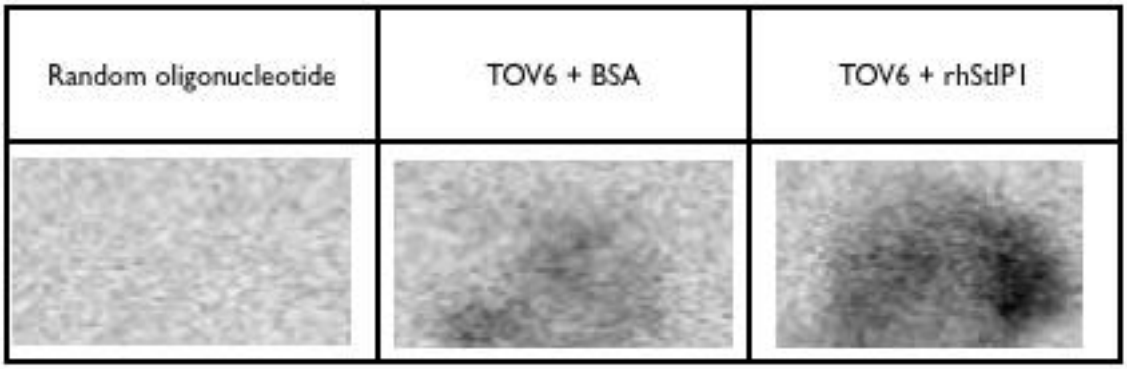


Figure 3-11. Chemi-luminescent blot of rhStIP1. The aptamer is able to induce a strong chemiluminescent signal with rhSTIP1, but not with BSA. When staining the protein with library, no luminescence can be observed

CHAPTER 4
THE FUNCTION OF STIP1 IN TOV-21G

Introduction

As shown in Chapter 3, the TOV6 aptamer binds to the membrane protein StIP1. StIP1 or heat shock protein organizing protein (hop) is known to interact with heat shock protein 70 (HSP70) and heat shock protein 90 (HSP90)[136]. StIP1 found in the cytosol is a component of the chaperone complex, HSP70, StIP1 and HSP90 which plays a key role in retaining the stability and proper folding of numerous proteins involved in cell viability[137]. HSP90 acts as a foldase and that needs ATP to adopt a conformation that folds client proteins back to their active forms. A recent study on the structure of the HSP90-StIP1-HSP70 complex provided information about the function of StIP1 in this complex. StIP1 plays an important role in stabilizing the client loading conformation of HSP90, and allows for easy interactions between HSP90 and client proteins such as HSP70. The interaction of StIP1 with HSP90 forces conformational changes that mimic the ATP-bound HSP90 form, and allows HSP90 to interact more easy with client proteins by bringing the hydrophobic surfaces of HSP90 together in an ATP-independent manner[138].

Chaperone and co-chaperone proteins have been identified in the cell membrane of several cancer cell lines by global profiling techniques[133], and the presence of StIP1 has been documented in several of these[125,133]. By performing Western blots with cell surface isolation kits, StIP1 has been found to be overexpressed in the membranes of TOV-21G, OVCAR3, SKOV3 and ES2. This study also performed blood sample tests on several ovarian cancer patients (both at early and late stages of cancer progression), which indicate that StIP1 shows promise as a potential biomarker for ovarian cancer[133].

However, the test was too limited for conclusive results. To determine the extent to which in what extent StIP1 is expressed, more ovarian cancer cell lines were tested.

StIP1 is also found in the membrane of several other cancer cell lines, including A172, a glioblastoma cell line [127,139,140], and Panc-1, a pancreatic cancer cell line[126,134]. In pancreatic cancer, silencing of StIP1 results in a significant decrease in invasiveness. Walsh and collegues observed that StIP1 was secreted along with HSP90 and matrix metallo-protease 2 (MMP2) to facilitate invasion. This chapter describes studies of the role of StIP1 in TOV-21G cells, as well the effect of TOV6 binding on the cell line's proliferation or viability. We hypothesize that StIP1 has a similar function in TOV-21G as it has in pancreatic cancer and thus plays a role in the metastatic properties of TOV-21G[163].

Materials and Methods

Instrumentation, Cell Culture and Reagents

The TOV-21G cell line was purchased from the American Type Cell Culture (ATCC). Dr. Patricia Kruk from the University of South Florida kindly donated the cell lines OVCAR3, OVCAR8, TOV112D, SKOV3, A2780s, A2780cp and C13. All cell lines were maintained in culture with MCBD 105: Medium 199 (1:1), supplemented with 10% FBS and 100 IU/mL Penicillin-Streptomycin. The cells were cultured at 37ºC in a 5% CO_2 atmosphere. Aptamer TOV6 and library were synthesized without any 5'- or 3'-modifications or with 3'-biotin as needed. The HSP90 specific inhibitor 17-N-Allylamino-17-demethoxygeldanamycin (17AAG) was purchased from Sigma.

Proliferation Assay

For proliferation studies, 100,000 cells were seeded in 24 well plates and allowed to grow for 3 days. The cell viability was tested in 0.05% trypan blue and MTT. The cells were counted in a hemocytometer.

Tumor Invasion Assay

The invasion assays were performed as described elsewhere[141]. In brief, the direct invasiveness of the cells was evaluated with the BD Falcon FluoroBlok 24-Multiwell Insert System (BD biosciences), precoated with matrigel (Figure 4-1). A migration control was run on the same system, using a BD Falcon FluoroBlok 24-Multiwell without matrigel coating (BD biosciences). The top compartment was loaded with 60,000 cells/ well in minimal media (RPMI), the lower compartment was filled with RPMI with a chemo-attractant (10% FBS) added. The 17AAG or aptamer TOV6 that was added in the minimal media was pre-filtered with 0.2µM syringe filters (Fisher). Percent invasion is calculated by #invading cells/#migrating cells *100%.

Membrane Expression of StIP1 in Ovarian Cancer Cell Lines

To determine the binding of the aptamers with different ovarian cell lines, the target cells (3×10^5) were incubated with varying concentrations of 5'- biotin labeled TOV6 on ice for 30 minutes in 100 mL of BB. Cells were then washed twice with 500 mL of BB, and suspended in 100 mL of BB containing streptavidin PE-Cy5.5 at an appropriate dilution. Cells were then washed twice with 500 mL of WB, and were suspended in 200 mL of BB for flow cytometry analysis, using a 5'-biotin labeled random sequence as the negative control.

Statistical Analysis

All errors reported are the standard deviation obtained from three replicates unless otherwise reported. The statistical significance between different invasion samples was determined with the student-t test, treating p-values below 0.05 as significant. The EC50 or IC50 of the tested drugs were calculated with JMP (SAS).

Results

Growth Inhibitory Effects of TOV6

The aptamer and the library were added to monolayers of cells, and the cell viability was tested with MTT and Trypan Blue. TOV6 did not induce any cell death, as shown in Figures 4-2 and 4-3. However, the cell proliferation was hampered at higher levels of TOV6, as can be seen in Figure 4-4. The IC50 of TOV6 in TOV-21G cells is 0.2μM. The effects of the TOV6 are cytostatic, as the viability of the cells was not hampered (Figure 4-3) and no cell death was observed visually (Figure 4-2), but the proliferation of the cells was affected (Figure 4-4).

The Non-Proliferative Effects of 17AAG

Compound 17AAG is an efficient HSP90-specific inhibitor[142]. StIP1 binds with HSP90 to form a complex that allows proteins to bind and be refolded[138]. Since HSP90's activity also regulates other crucial cellular mechanisms[142], an IC50 study was performed on TOV-21G cells *in vitro*, in order to determine the inhibition of HSP90 before a decrease in proliferation can be observed. If the concentration of 17AAG can be minimized in order to maintain proliferation, a more correct view about the role of HSP90 in its interaction with StIP1 can possibly be observed. Figure 4-5 shows the effect of 17AAG on TOV-21G proliferation. The IC50 of 17AAG in TOV-21G was 0.060μM, which is close the value found in the literature, 0.100μM[167]. Figure 4-6

represents the visual control for cells that remained attached on the dish after 17AAG treatment.

The Effect of StIP1 siRNA Silencing on Invasion

Walsh *et al.* showed that StIP1 is involved in the invasion process *in vitro* in pancreatic cancer[134]. Since StIP1 is expressed on the cell membrane of TOV-21G, an invasive cell line[143], it was hypothesized that the protein's location inside the membrane is playing a role in metastasis. The true invasive potential of a cell line is measured by the propensity to migration through a microporous membrane and the ability of the cell line to cross the basement membrane (modeled by a microporous membrane covered with matrigel). TOV-21G that had StIP1 silenced through siRNA showed a reduction in its propensity to migrate (Figure 4-7, A versus B histograms). As shown in Figure 4-7 neither 17AAG (2A and 3A) nor saturation concentrations of TOV6 (4) by themselves (without StIP1 silencing) inhibits migration. In Figure 4-7, histogram 3B, 0.10 µM 17AAG slowed the migration, but at this concentration, the observed decrease in migration rate could be attributed to cell death triggered by the drug 17AAG. In a duplicate experiment with matrigel-coated membranes (Figure 4-8), it can be observed that HSP90 inhibition strongly reduces the ability of the cells to invade, even at a concentration where no loss in proliferation is observed over the tested time span. The StIP1 silenced cells also had a diminished ability to cross the membrane. As can be seen, the effect of the inhibition of invasion works in a cumulative fashion. Remarkably, TOV6 strongly inhibits the invasion of the cells, indicating that the aptamer could be used in further studies to fight metastasis. Figure 4-9 (%-invasion) shows that that treatment with TOV6 provides invasion inhibitory effects to the same degree as StIP1 silencing when cells are treated treated with 0.10 µM 17AAG.

The Expression of StIP1 in Ovarian Cancer Cell Lines' Membrane

StIP1, the target for aptamer TOV6, is expressed in TOV21G (OCCA), but not in CAOV3 (SOAC). Some binding tests have been performed for other cell lines in Chapter 2, where expression of StIP1 was found in the membrane of A172, CEM, HCT-116, H23. To determine if the aptamer could be useful clinically for possible early detection, and to what extent the marker could be used for cancer different than ovarian clear cell adenocarcinoma, a panel of cell lines was selected that included several examples of serous adenocarcinoma (OVCAR3, OVCAR8, SKOV3). Other types of ovarian histological subtypes included ovarian endometrioid carcinoma (TOV-112D), and in the ovarian cancer cell lines A2780s, A2780cp and C13. Both A2780cp and C13 are cisplatin resistant serous adenocarcinoma cell lines[100,162] . The flow cytometry experiments (Figure 4-10) on these cells show that there is some heterogeneity in serous adenocarcinoma (binding to OVCAR3 and SKOV3), but that most resistant cell lines do not express StIP1, with the exception of C13 (Figure 4-10). The rate of expression in StIP1 positive cells is variable, as some cell line show high expression (SKOV3, OVCAR3, TOV112D, TOV21G) of StIP1, while some signals from TOV6 binding to StIP1 is but marginal (A2780s, C13). The results suggest that StIP1 is a protein that can be found in other histological types than OCCA.

Discussion

This chapter describes a study that tests the hypothesis that StIP1 plays a role in cell invasion, as an adaptor protein that helps in the activation mechanism of HSP90 has with of matrix-metallo proteases (MMPs). Recent structural studies have shown that StIP1 is binds in the N-terminal ATP binding pocket of HSP90[138,165]. That HSP90 is

involved in the invasion mechanism in TOV-21G has been demonstrated by the inhibition of invasion with the HSP90 specific inhibitor, 17AAG[142,144]. An interesting result was observed when TOV6 was added to the media of TOV-21G. The invasion dropped down to almost the same levels that occurred with StIP1 silencing, combined with 17AAG in the media. Although to more clearly elucidate this phenomenon, structural studies are needed, it is likely that the aptamer inhibits the activation of pro-MMP2. For example, HSP90 is a known activator of MMP2 (constitutively expressed, but needs to be activated for function), and the actions of HSP90 have been subject to extensive study[165,166]. It may be possible to expand the current understanding of MMP activation by HSP90 by considering that StIP1 plays a critical role in the activation of these important proteins involved in cell invasion, perhaps by maintaining the folding of HSP90 that is required for its activation function. It can be hypothesized that the aptamer sterically hinders the binding of proMMP2 to StIP1 in a complex with HSP90, thereby preventing proMMP2 from being cleaved to MMP2. It is known, for example, that MMP2 cannot be found in the conditioned media of pancreatic cells in which StIP1 has been knocked down[134]. Furthermore, in TOV-21G have the levels of active MMP-2 been directly correlated with HSP90 inhibition by 17AAG[165]. A proposed mechanism of activation is represented in Figure 4-11. Figure 4-11 described how proMMP2 binds to the StIP1-HSP90 complex to form active MMP2. It is hypothesized that the proMMP2 has to bind to both proteins in complex to be activated, as StIP1 silencing and HSP90 inhibition strongly reduce invasion. Since the aptamer TOV6 is as affective as the combination of StIP1 silencing and HSP90 inhibition, it is very likely that the aptamer blocks the epitope for proMMP2 that requires both proteins. To gain more detailed

insight in this phenomenon, more study is definitely warranted. The question can be asked, for example, whether the aptamer is inhibiting the complex on the cell surface, or the secreted complex of StIP1 and HSP90. The role of trans membrane (TM) MMP may also be needed to be investigated.

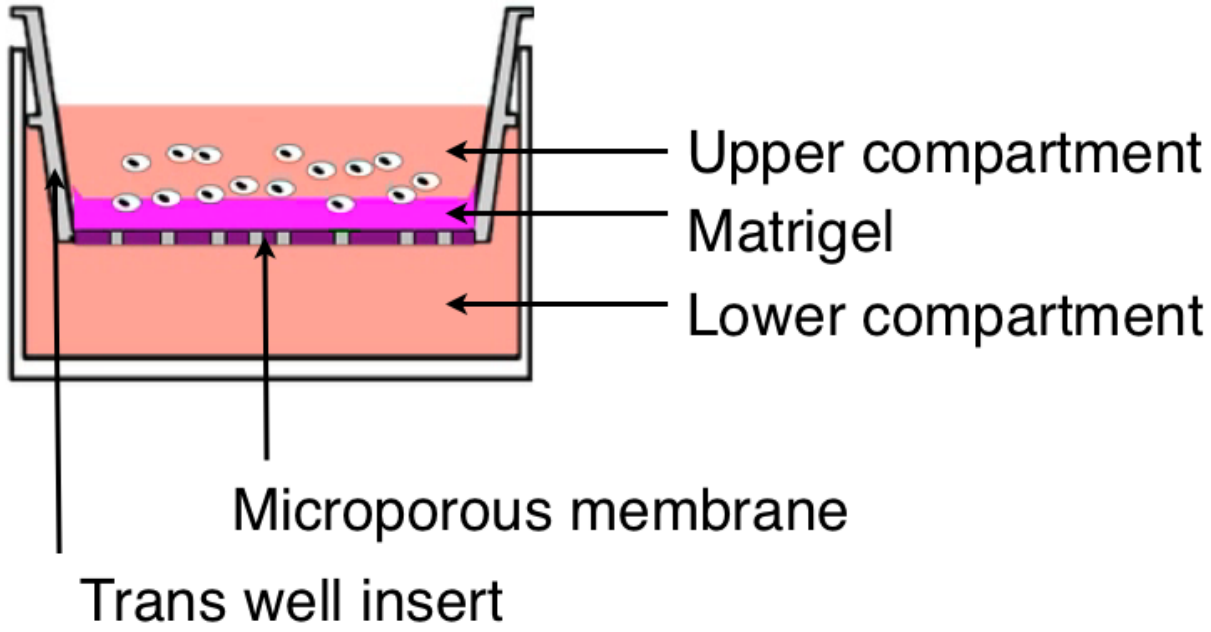


Figure 4-1. Figure of a Boyden chamber for migration and invasion studies. For these studies, the top chamber was filled with a minimal media and the cells of interest. The lower compartment contained a chemo-attractant (10% FBS), by which migrating (no matrigel) or invading cells (matrigel) can travel through the microporous membrane. The cells can be stained with a fluorophore by transferring the transwell insert in a fluorophore containing solution. The fluorescense of the cells was measured with a bottom plate reader. Non migrated or invaded cells were not detected due to an optical filter in the trans well insert

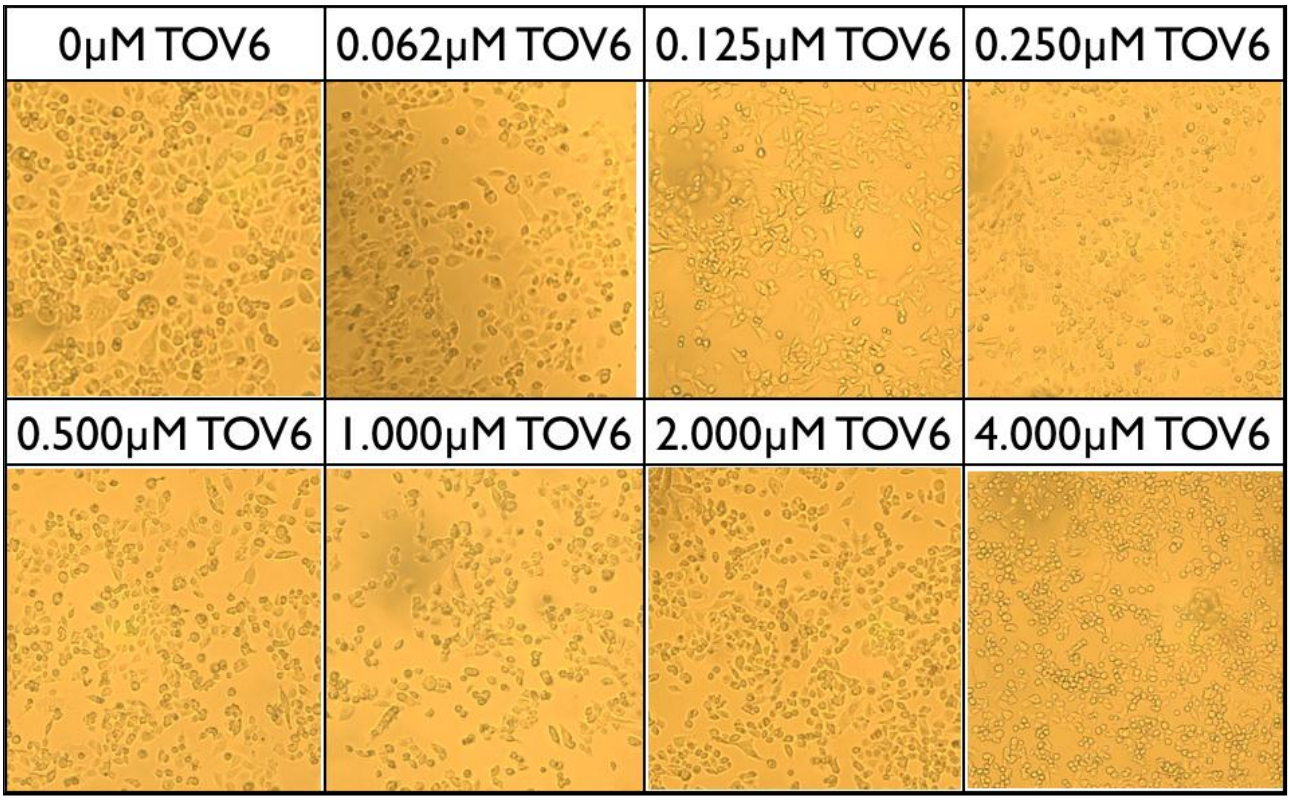


Figure 4-2. Microscopic images of TOV-21G cells treated with TOV6 for 3 days. The cells look more round with increasing levels of TOV6. In order to detect dead cells, the cells were stained with 0.05% Trypan Blue. (Lack of blue cells indicates that TOV6 is not cytotoxic)

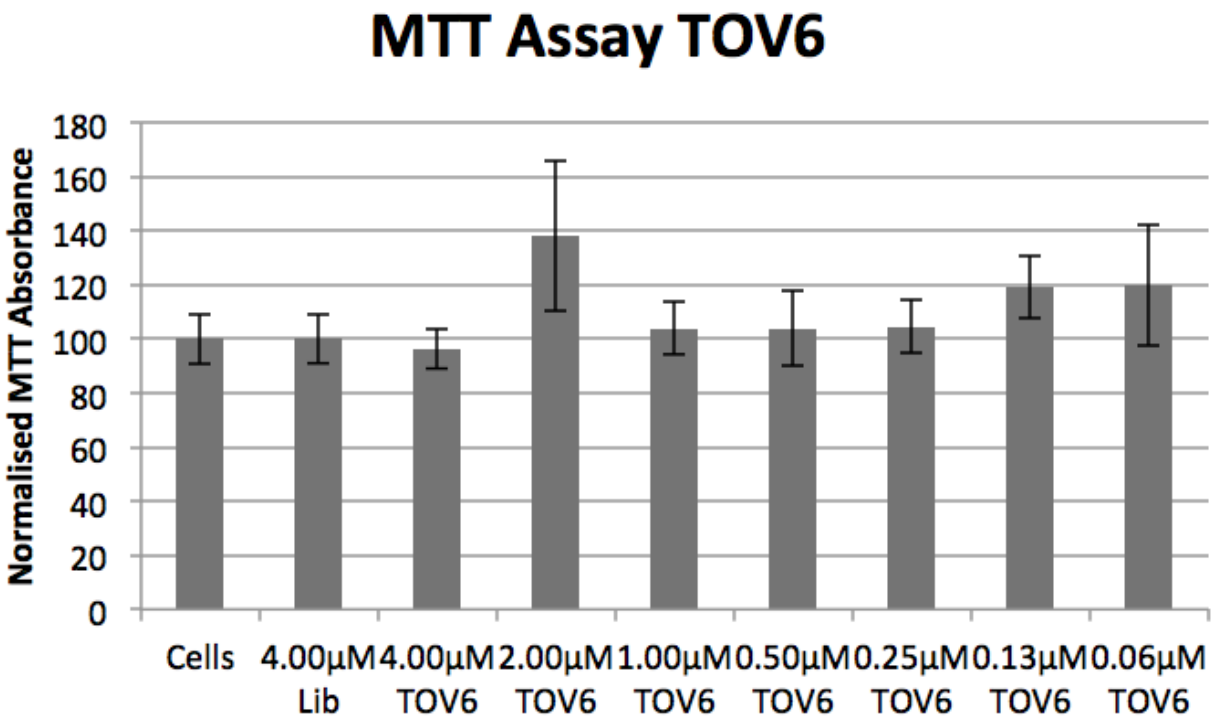


Figure 4-3. MTT assay of TOV-21G cells incubated with library and TOV6. There was no decrease of viability of the cells at the concentrations of the aptamer tested. Error bars represent the standard deviation (n=3)

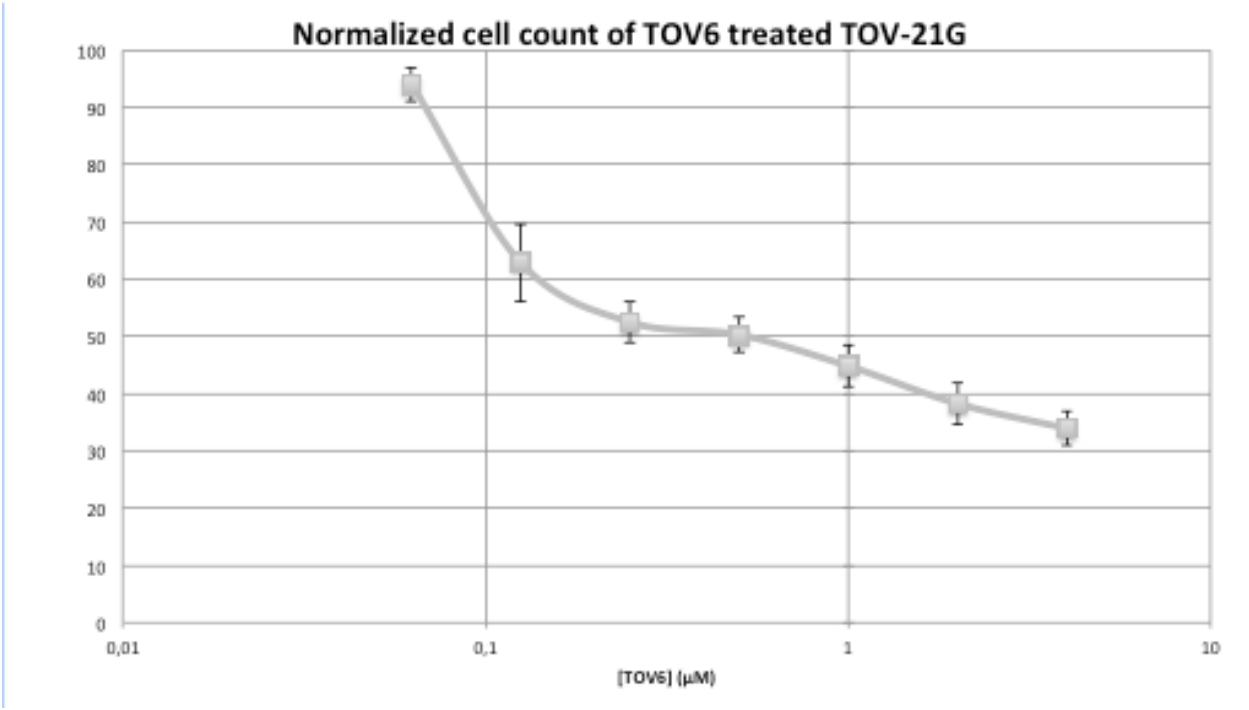


Figure 4-4. Normalized cell count of TOV-21G after 3 days of incubation with TOV6. The cells were normalized against untreated cells after 3 days of growth. The IC50 was determined to be 0.20µM by non-linear regression in JMP. The error flags are the standard deviation (n = 3)

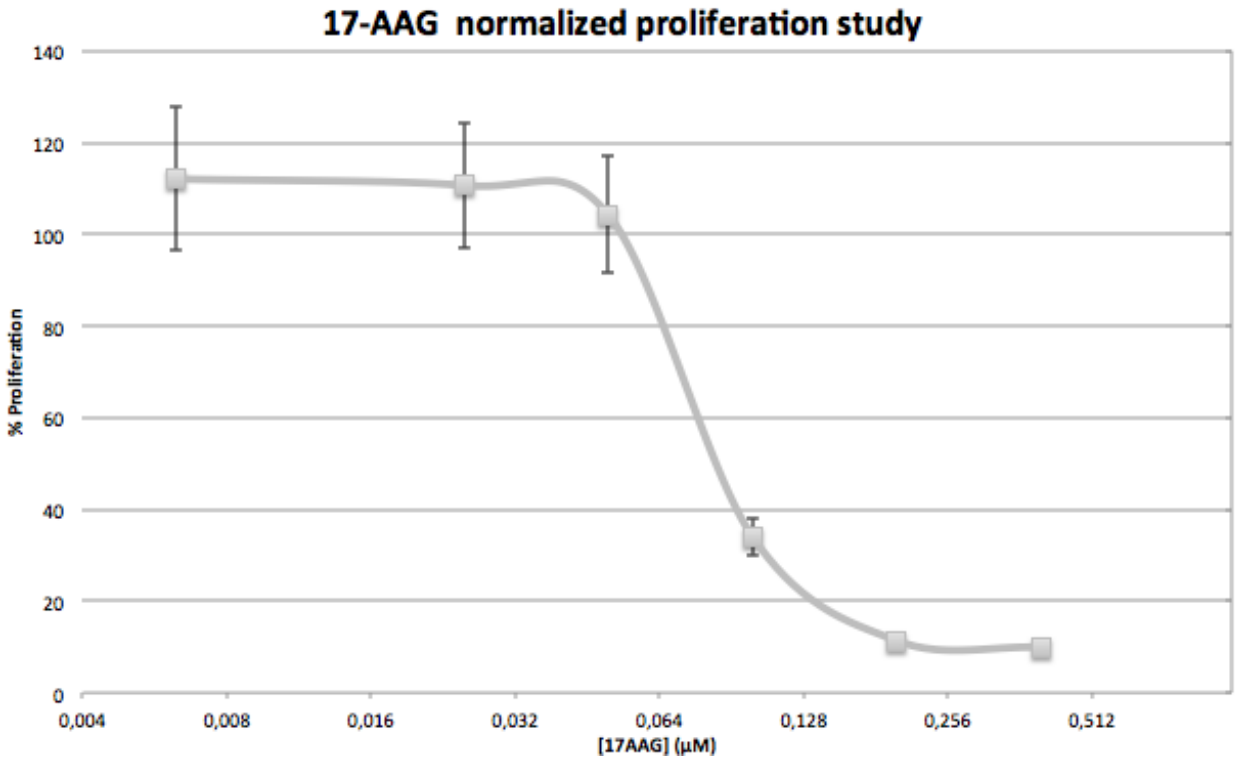


Figure 4-5. Normalized proliferation study of TOV-21G, after three days of incubation with 17AAG, the IC50 was determined to be 0.060µM by non-linear regression in JMP. The error bars give the standard deviation (n = 3)

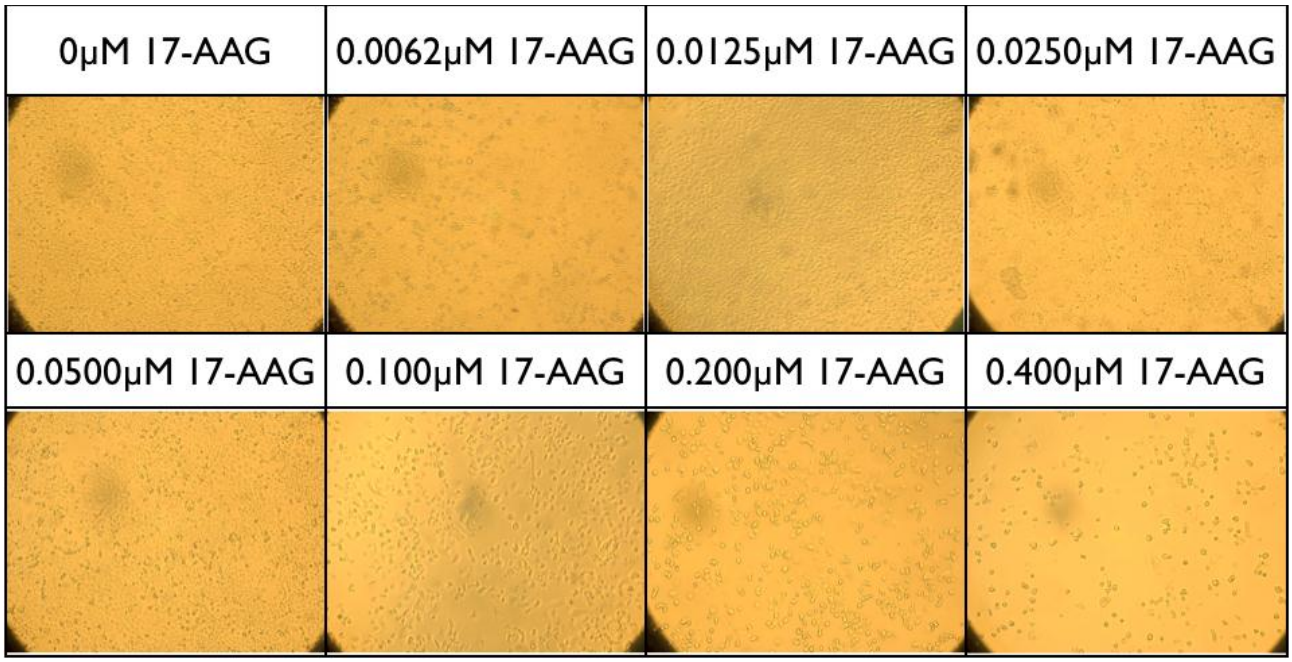


Figure 4-6. Microscopic images of TOV-21G cells treated with different levels of 17AAG in full media after 2 days incubation. The cells were stained with 0.05% Trypan Blue. The overall cell density decreases with increasing 17AAG concentrations.

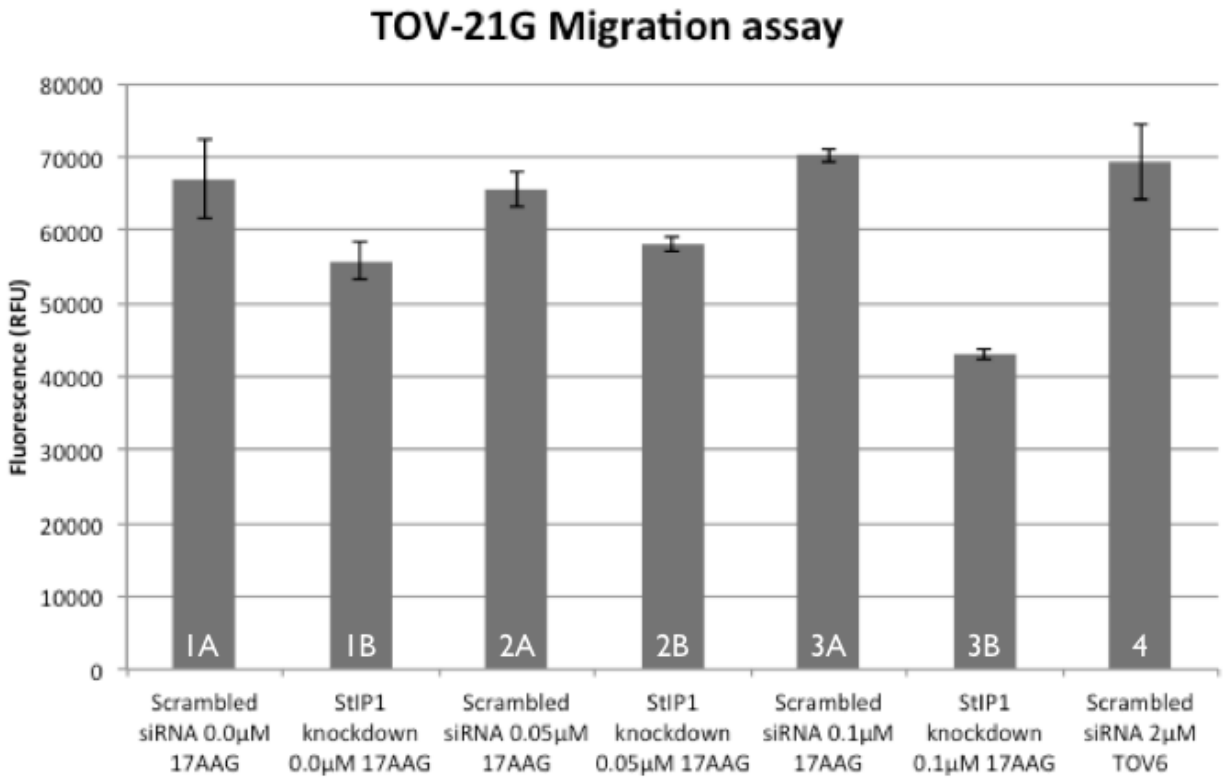


Figure 4-7. Migration of TOV-21G across a microporous membrane. StIP1 siRNA treated cells migrate slower (histograms labeled with B) then scrambled siRNA treated cells (histograms labeled with A). The reduction of migration with 0.1µM 17AAG can be explained by the reduced health of TOV-21G cells at this concentration. TOV6 and 17AAG do not affect migration by themselves. Error flags represent the standard deviation (n = 3)

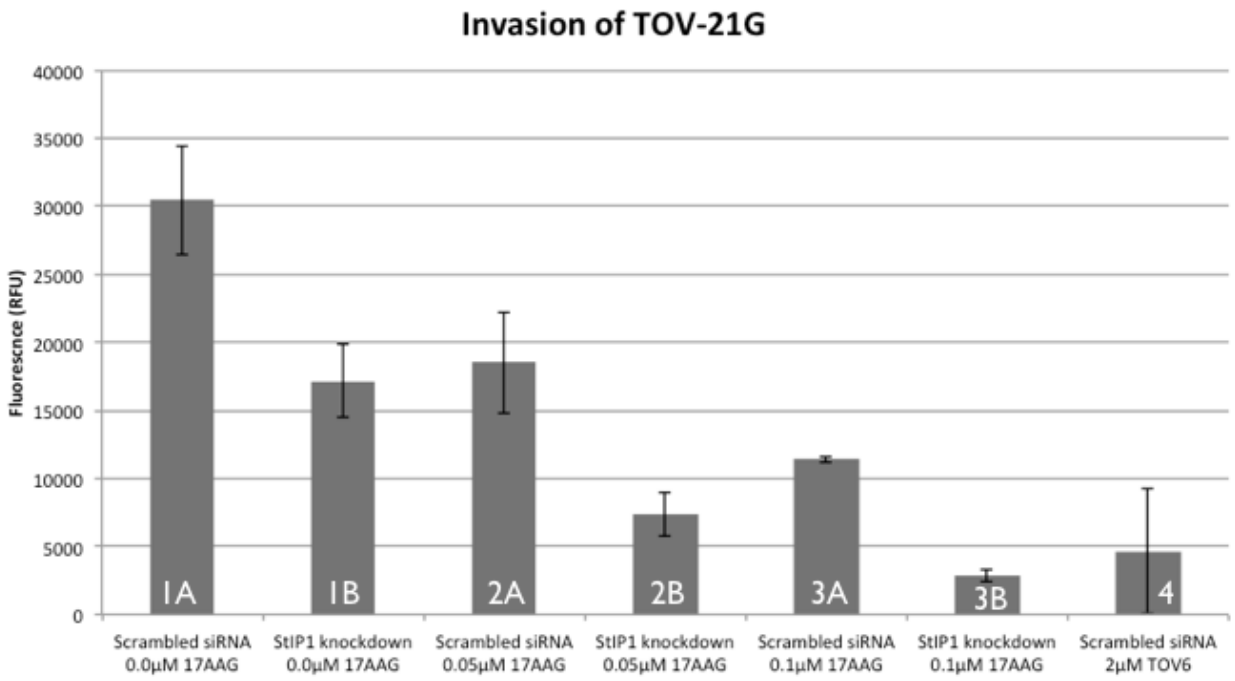


Figure 4-8. Invasion assay of TOV-21G determining the effect of TOV6 on the ability of TOV-21G to cross a matrigel layer. StIP1 siRNA treated cells are hampered in their ability to digest the matrigel layer (all B histograms), an effect that is amplified by 17-AAG inhibition (1A - 2A - 3A). TOV6 (histogram 4), without any siRNA treatment is slows down the invasion at similar rates at StIP1 knocked down 17AAG treated samples (3B). Error flags represent the standard deviation (n = 3)

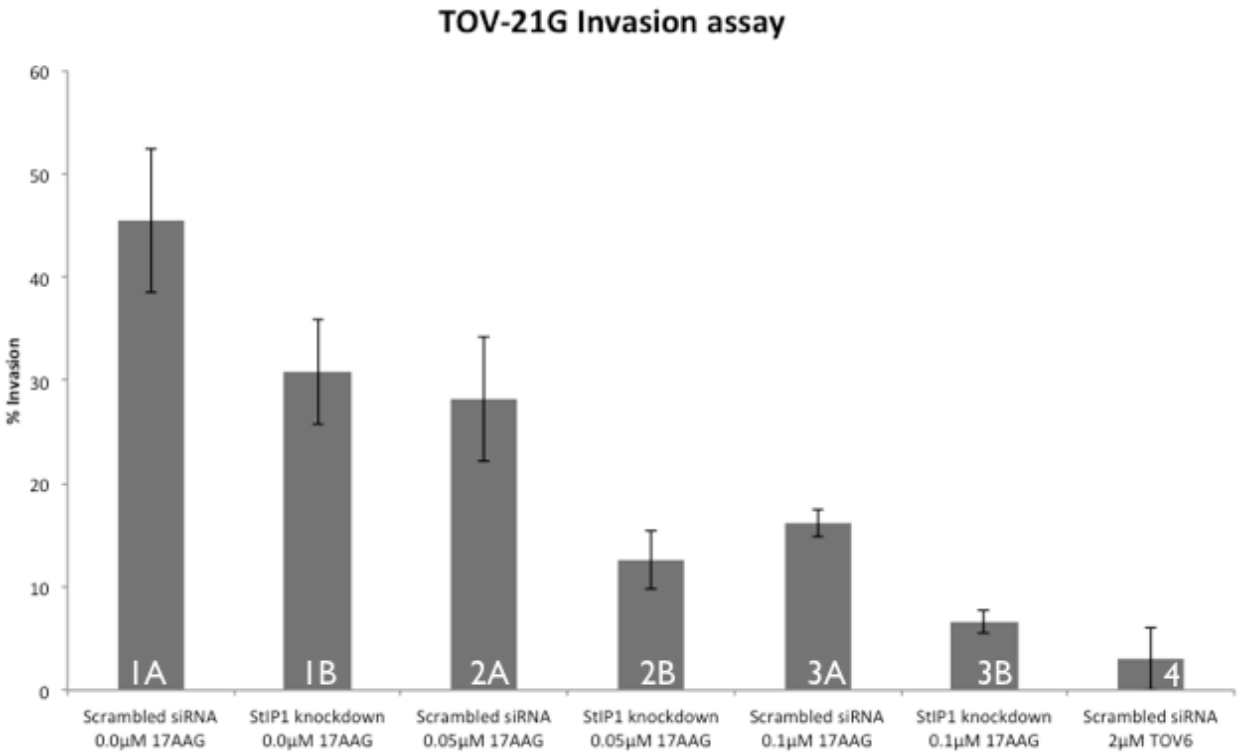


Figure 4-9. Combined TOV-21G cell invasion assay. The data suggests that the invasion is facilitated through a mechanism where StIP1 and HSP90 are involved. The effect of HSP90 inhibition and StIP1 knockdown seems to be cumulative. Furthermore, TOV6 is able to inhibit invasion as strong as StIP1 knockdown in combination with 17AAG

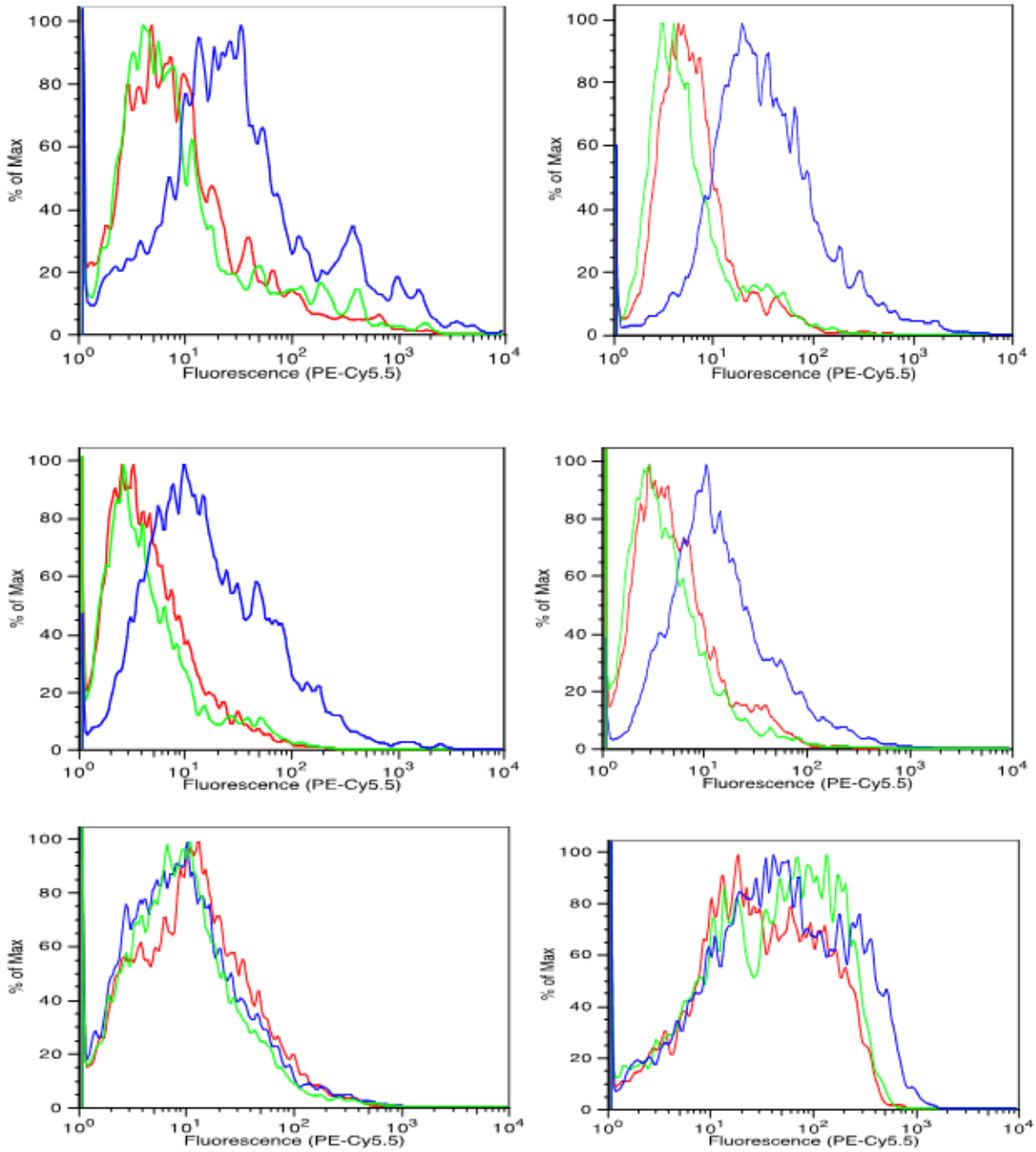


Figure 4-10. Binding assay of several ovarian cancer cell lines with TOV6First column, fro top to bottom: a) OVCAR3; b) TOV112D; c) OVCAR8; d) A2780cp; second column, a) SKOV3; b) TOV-21G; c)A2780s; d)C13. Red: Unlabeled cells; Green Library; Dark blue: TOV6. d) is found on the next page

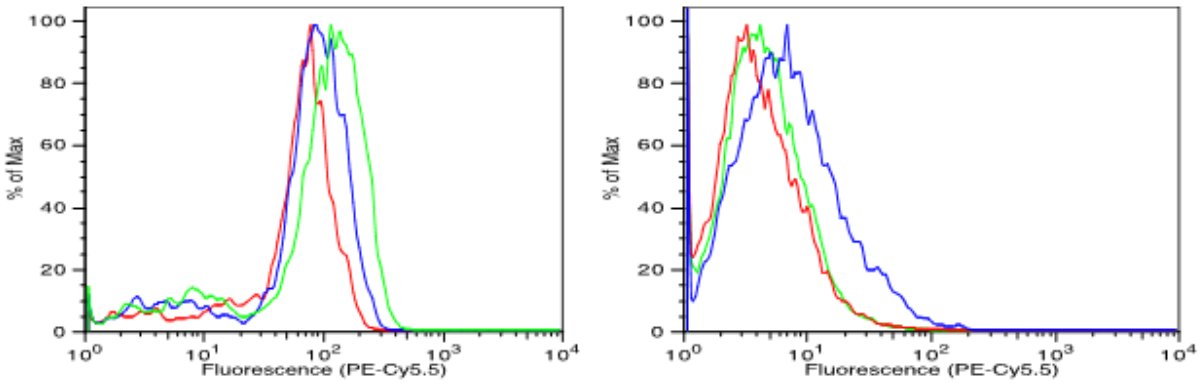


Figure 4-10. Continued

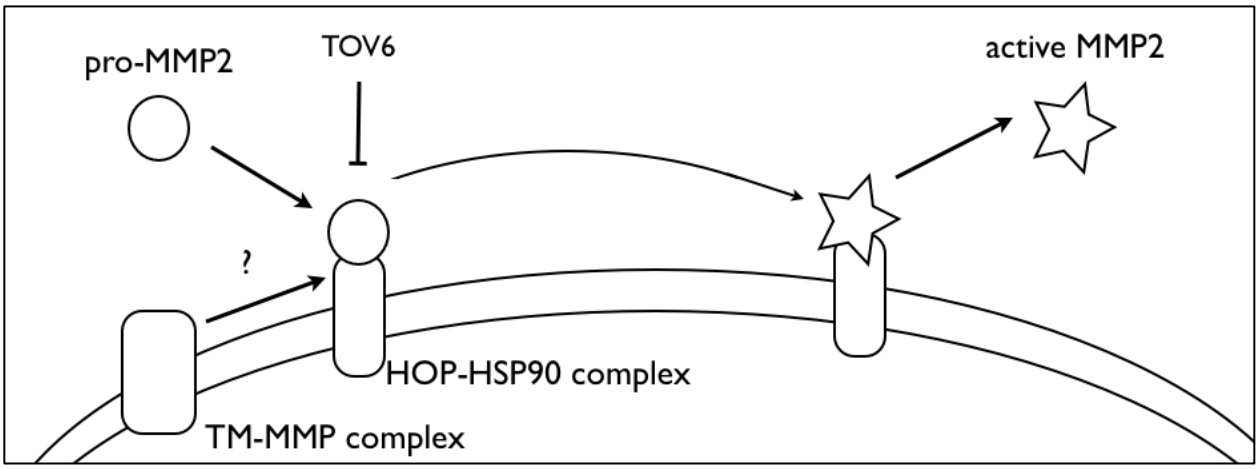


Figure 4-11. A proposed mechanism for TOV6 caused inhibition of cell invasion. MMP2 needs to be activated by cell membrane protein interactions, possibly either by direct proMMP2/HOP-HSP90 interaction or via complexation with MT-MMP and HOP-HSP90. Either of these possibilities could be blocked by direct competition with TOV6

CHAPTER 5
OVARIAN CANCER STEM CELL SELEX

Introduction

Ovarian cancer tumors, like all cancer types, exist as heterogeneous cell clusters. Not all cells in the tumor are the same, but there is a hierarchical distribution of cells, making tumors caricatures to the normal tissues of their origin. The hierarchy in normal tissue is attributed to stem cells, a type of cell that can differentiate into several different tissues (e.g., skin, bone marrow). The similarities between normal tissue and tumors have led to the attractive model of cancer stem cells. This model considers a tumor to be an agglomerate of differentiated, non-mitotic cells that also contains another type of cells capable of self-renewal, which can stay in the body for a life-time and are resistant to chemical or electromagnetic (i.e. radiation) attacks. These properties go hand-in-hand with the observation that stem cells can be non-dividing for a long time and that they have the ability to colonize other parts of the body.

These properties explain what is commonly observed in ovarian cancer: the patient shows an excellent response to chemo- or radiation therapy, but cannot be considered cured[145]. It is not rare for a patient who has been declared healthy to suffer a reoccurring cancer, that is even more aggressive and shows strong drug resistance, and these properties become more heinous as the length of the treatment free time interval increase[146]. Basically, it is assumed that most of the cells die from a successful chemical or irradiative assault, but the cells that show stem cell-like properties can survive the assault and sustain in the tissue for a long time. This results in the relative enrichment of a chemical resistant cell. Metastatic relapse can occur many years after a patient has been declared healthy[147]. Since these types of cancer, including ovarian

cancer, fit the stem cell model remarkably well, the cells responsible for these relapses have been coined “Cancer Stem Cells” (CSC)[148,149].

In ovarian cancer, cancer stem cells have been acknowledged to be involved in cancer aggressiveness[150]. CSC’s extracted from tumors and have shown to be very effective in the initiation of tumors in the xenograft models; as few as a 100 cells are needed to initiate the growth of a tumor[151]. CSC’s can be enriched from ovarian cancer by Fluorescence Assisted Cell Sorting (FACS) based on the efflux of an organic dye, like Hoecht 33324[152], or on the expression of a combination of specific cell surface markers[151]. The use of these properties is somewhat controversial, as many of the proposed markers have been shown not to be not absolute markers for CSC biology[153,154]. In other words, the use of cell surface markers is highly useful to sort out cells that show stemness properties, but the used markers are not defining the stemness properties of the extracted cells[155]. An example of the limitations of the currently known stem cell surface markers can be found in CD133, or prominin, which was thought for many years to be a stem cell specific surface marker, but is expressed in many normal epithelial tissues as well[156]. Due to the limited understanding of these CSC markers, it is highly desirable to have more tools to extract CSC out of tumors and to gain deeper knowledge about the molecular mechanisms that underlie CSC emergence.

In addition, many stem cell markers are expressed in a plastic fashion, meaning that the expression of markers doe not remain constant. An illustrative example can be found again with prominin; cells sorted for CD133 will return to the pre-sorting levels of CD133[157,159]. A recent publication has reported this problem for all 15 commonly used

cell surface biomarkers[158]. In order to get deeper insight into the behavior that CSCs exhibit, more knowledge about their cell surface markers would be highly beneficial. Also, a stable, specific cancer stem cell line could be of great benefit for the treatment of cancer. Currently, companies and labs around the world are able to provide cell lines that can be passaged sufficient amounts for application of SELEX[168, 169]. Selecting probes that bind to cells that show stemness, without the knowledge of the actual molecular target may be very advantageous. By doing so, novel markers can be identified that play an important role in the biology of cancer stem cells. Such an approach has been demonstrated to be plausible for mammalian membrane proteins. The system that is currently used for phenotyping cells by their membrane proteins (the Cluster of Differentiation system), contains 350 proteins. However it is estimated that around one-third of the protein in the genome is found in the membrane[28,113].

As described in previous chapters, Cell-SELEX has been touted as a new method for the elucidation of cell surface markers[44]. This chapter describes the application of cell-SELEX to identify aptamers against ovarian CSCs.

Cell-SELEX on an Ovarian CSC Cell Line

Materials and Methods

Instrumentation and reagents

All oligonucleotides were synthesized by standard phosphoroamidite chemistry using a 3400 DNA synthesizer (Applied Biosystems) and were purified by reversed-phase HPLC (Varian Prostar). All PCR mixtures contained 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 2.0 mM $MgCl_2$, dNTPs (each at 2.5 mM), 0.5 mM of each primer, and Hot start Taq DNA polymerase (5units/mL) (TaKaRa). PCR was performed on a Biorad Thermocycler. Monitoring of pool enrichment, characterization of the selected

aptamers, and quantitiation of the target protein assays were performed by flow cytometry using a FACScan cytometer (BD Immunocytometry Systems). The DNA sequences were determined by the Genome Sequencing Services Laboratory at the University of Florida with the use of iontorrent sequencing, analysis on sequence homology was performed with MAFFT[160].

Cell culture and buffers

The ovarian cancer stem cell line was purchased from Celprogen (San Pedro). The cells were maintained in an undifferentiated state according to the supplier's specifications. In short, the cells were maintained beneath 50-60% confluence, using the supplier's optimized non-differentiation buffer. The stem cells were differentiated with specifically designed media, provided by the supplier. The TOV112D, CAOV-3 and TOV-21G ovarian cancer cell lines where maintained in culture with MCBD 105: Medium 199 (1:1), supplemented with 10% FBS and antibiotics. All the cell lines were maintained in a humidified incubator in a 5% CO_2 atmosphere.

During the selection, cells were washed before and after incubation with wash buffer (WB), containing 4.5 g/L glucose and 5 mM $MgCl_2$ in Dulbecco's phosphate buffered saline with calcium chloride and magnesium chloride (Sigma). Binding buffer (BB) used for selection was prepared by adding yeast tRNA (0.1 mg/mL) (Sigma) and BSA (1 mg/mL) (Fisher) to the wash buffer to reduce background binding. A non-enzymatic cell dissociation buffer was used for cell detachment (Sigma).

Cell SELEX library

A library was designed as described in Chapter 2 with the following primer sequences: ATC CAG AGT GAC GCA GCA $(N)_{40}$TGG ACA CGG TGG CTT AGT. The forward primer was labeled with FITC and the reverse primer was labeled with biotin for

efficient separation of the forward sequence for ssDNA elution from streptavidin columns[88].

In vitro cell-SELEX on ovarian CSCs

The commercially available ovarian cancer stem cell (OCSC) line was chosen as the target cell line. Differentiated ovarian cancer stem cells were used for the negative selection. Differentiation of cells was monitored under the microscope: undifferentiated cells were round, while differentiated cells were larger, spindle shaped and had a more complex cellular structure. The selection occurred in similar fashion as described in Chapter 2, except that: all the cells that were used during the SELEX procedure were detached from the flask with non-enzymatic dissociation buffer in order to maximize the possible contact surface between the cells and the library. This procedure also resulted in increased the washing stringency, as the supernatant was easier to remove compared to a selection for adherent cells on the plate[119]. PCR amplifications were carried out at 95ºC for 30s, 60ºC for 30s, and 72ºC for 30s, followed by a final extension step for 3 minutes at 72ºC. Enrichment was monitored with flow cytometry. Due to the relative instability of the cell line, SELEX was stopped and submitted for sequencing as soon as specific a pool was selected.

Specificity and affinity studies

Flow cytometry was used for the determination of the affinity constants of the found aptamers. The ovarian CSCs (2×10^5) were incubated with various concentrations of 5'- biotin labeled aptamers or FITC labeled pools (for enrichment assessment) on ice for 30 minutes in 100 mL of BB. Cells were then suspended in 100 mL of BB containing streptavidin-PE-Cy5.5 for the labeling of the cells. After labeling, the cells were washed twice with 500 mL of WB, and were suspended in 200 mL of WB

for flow cytometric analysis, using a 5'-biotin labeled random sequence as the negative control. All the experiments for binding assays were repeated three times. The specific binding intensity was calculated by subtracting the mean fluorescence intensity of the background binding from the mean fluorescence intensity of the aptamers. The equilibrium apparent dissociation constant (K_d') of the fluorescent ligand was obtained by non-linear regression analysis of the specific binding intensity (Y) versus the aptamer concentration (X) fitted to the equation $Y=B_{max}X/(K_d+X)$ using SigmaPlot (Jandel, San Rafael, CA).

Aptamer mediated cell sorting with magnetic beads

The streptavidin covered magnetic beads (Invitrogen) were incubated with aptamer solution for 10 minutes, after which they were washed twice with WB, resulting in aptamer functionalization of the beads. The beads were then incubated in non-enzymatic dissociation buffer treated cells for 30 minutes. The beads were used to extract cells that bound to the aptamer. The cells were released from the beads by a 30 minute incubation with DNase I (Invitrogen) and cultured for one day before analysis.

Results

Monitoring the pool enrichment for undifferentiated OCSC vs differentiated ovarian stem cells

At the start of the selection, 20 pmol of library was used. See Figure 5-1 for the progressive enrichment. Sequences that bound generally expressed surface markers for the ovarian cell line were removed for the most part by counter-selecting with the differentiated cell line (Figure 5-2). The selection was concluded after 15 rounds, as some amplification for differentiated cells was observed (blue in Figures 5-1 Figure 5-2), while a good pool was obtained after 14 rounds (light blue in Figures 5-1 and 5-2).

Furthermore, due to the instability of the cells (the cells are only undifferentiated for 6 to 50 passages[169] (Celprogen, personal correspondence)), it was decided that the enriched pools were to be analyzed by next-gen sequencing. In round 15, there was an increase in the binding with OCSCs observed. However, this pool also bound a little to the differentiated cells (Figure 5-2). Since pool 14 shows a good binding profile with the cells (better specificity), pool 14 was also included in the iontorrent analysis. As can be seen in Figure 5.1, the cells that bound to the pools were split in two distinct populations, which could reflect the fashion by which stems cells split into differentiated daughter cells and in undifferentiated mother cells[161].

Binding assay of putative aptamers and determination of K_d'

In total, 5 aptamers chosen from the alignment data showed specificity for undifferentiated ovarian cancer stem cells; they showed binding to undifferentiated cells, while they did not bind to the differentiated cells. Flow cytometry plots for binding assays are shown in Figures 5-3 to 5-7 and the dissociation constants are summarized in Table 5-1. As can be seen, the apparent K_ds of the selected aptamers all lie in the lower nanomolar range, which is the usual value for aptamers selected through cell-SELEX. Figure 5-8 shows an example of the K_d' determination of DOCSC-3. The aptamers were tested against several ovarian cancer cell lines. As shown in Table 5-2, DOCSC-2 is binding to the OCCA cell line TOV-21G, ovarian endometrial adenocarcinoma TOV112D and not to the serous adeno-carcinoma cell line CAOV3. Also DOCSC-4 and DOCSC-5 also bind to TOV-21G.

DOCSC2 and DOCSC5 were used to enrich for cells that specifically bound to the aptamers (Figure 5-9). For this, the aptamers were immobilized on magnetic microbeads. The beads were successful in the enrichment of cells expressing the

marker binding to the aptamers. The phenotypes of the two cells showed some differences from each other. As shown in Figure 5-10, the cells enriched from DOCSC-2 are small and round, while the cells from DOSCS-5 are larger and are spindle-shaped. The cells were also tested with flow cytometry for DOSCS-5 in Figure 5-9, and here it can be observed that after one day of culturing the cells that bound to the beads there was some re-enrichment of cells that showed strong binding to DOSCS-5.

Discussion

Five aptamers for OCSC were selected from the sequencing data. The aptamers selected show good binding properties, with K_ds in the low to middle nanomolar range. Two aptamers (DOCSC-2 and DOCSC-5) were tested for the enrichment purposes. Figure 5-9 shows flow cytometry results for DOCSC-5 sorting of cells that did not show clear binding to DOCSC-5 (passage number 58). The histogram in orange represents the unsorted cells, while the black histogram clearly shows two populations of cells, similar to results obtained for the enriched pools (Figure 5-1). This experiment shows that the aptamers can be used for the extraction of the cells out of the population of undifferentiated cells. When comparing the sorted cells with each other under the microscope, there is a difference observable in the cell shape (Figure 5-10). For example, sorting with DOCSC-5 leads to the growth of large spindle-shaped cells, compared to small round cells sorted with DOCSC-3. This implies that the cell surface markers binding to the aptamers are related to a defined phenotype.

Future Work

Molecular profiling is required to verify if these samples are truly binding to markers related to stem cells. As has been demonstrated with the flow data and microscope images, it is possible to re-enrich cells, which can be further analyzed for

stemness. There are several experiments that can be conducted to verify if the aptamers bind to a marker for a CSC. The experiment with the aptamer functionalized microbeads shows that this is possible. A classic experiment is the analysis of the transcriptome of the cells according to the different aptamer sorted cells. There are a few genes that are known to be associated with the stemness properties of CSCs[164]. For example, two the membrane markers for cells in the undifferentiated state are CD133 and CD44. Other stemness markers are telomerase and c-kit. If the cells indeed show stemness, the expression of these genes can be verified with qPCR. The true experiment to verify for stemness however lies in the use of the xenograft seeding model[150,151], in which fewer than 1000 cells are subcutaneously injected in nude mice. By verifying the tumor size every 3 days for two months, the tumor initiating ability of the sorted cells could be determined. At the moment, the current model is the only accepted method to verify stemness, or tumor initiating ability[145].

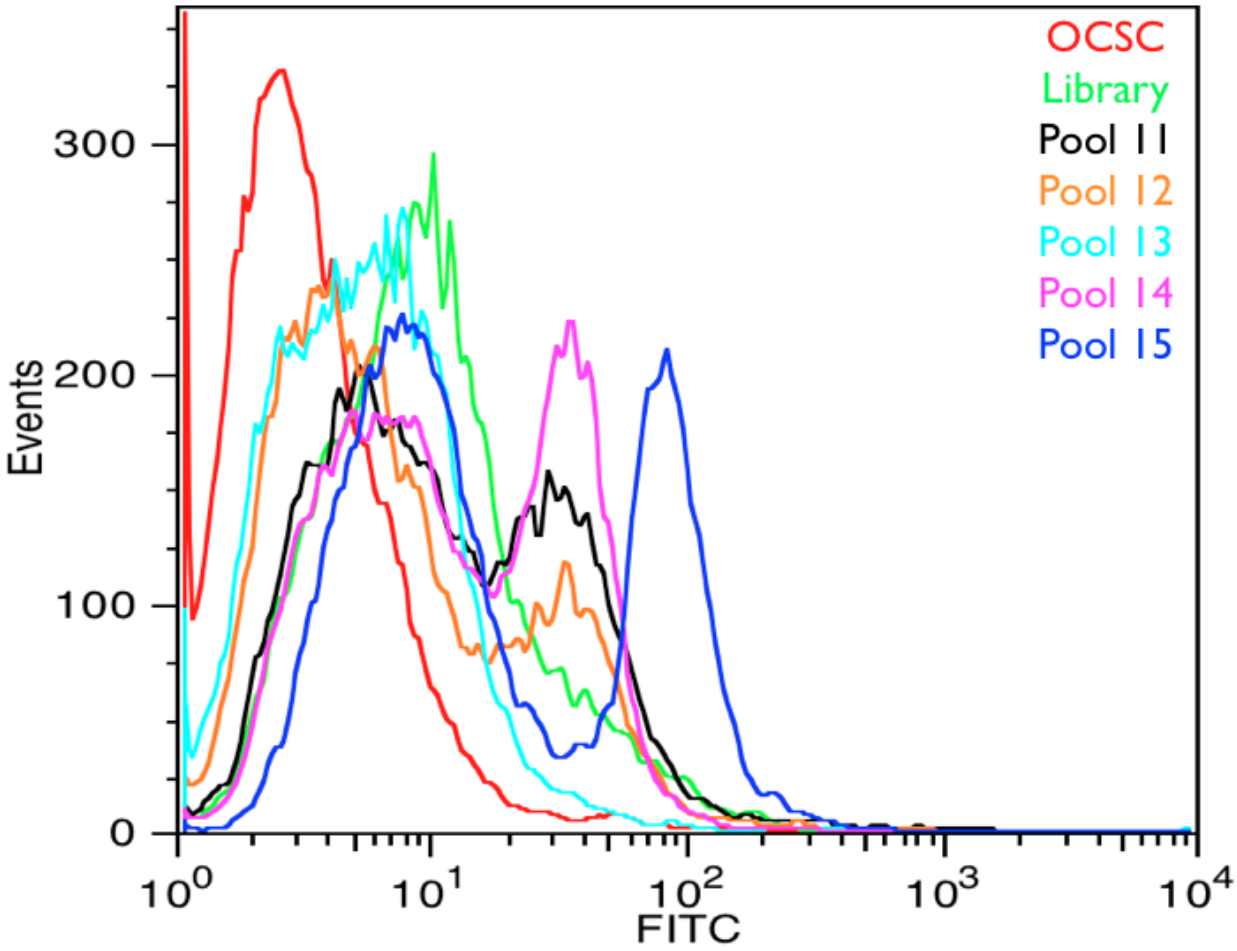


Figure 5-1. Enrichment for ovarian cancer stem cells. The pools show binding to the undifferentiated cells. The fluorescence profile shows a separation of the enrichment in to two distinct populations

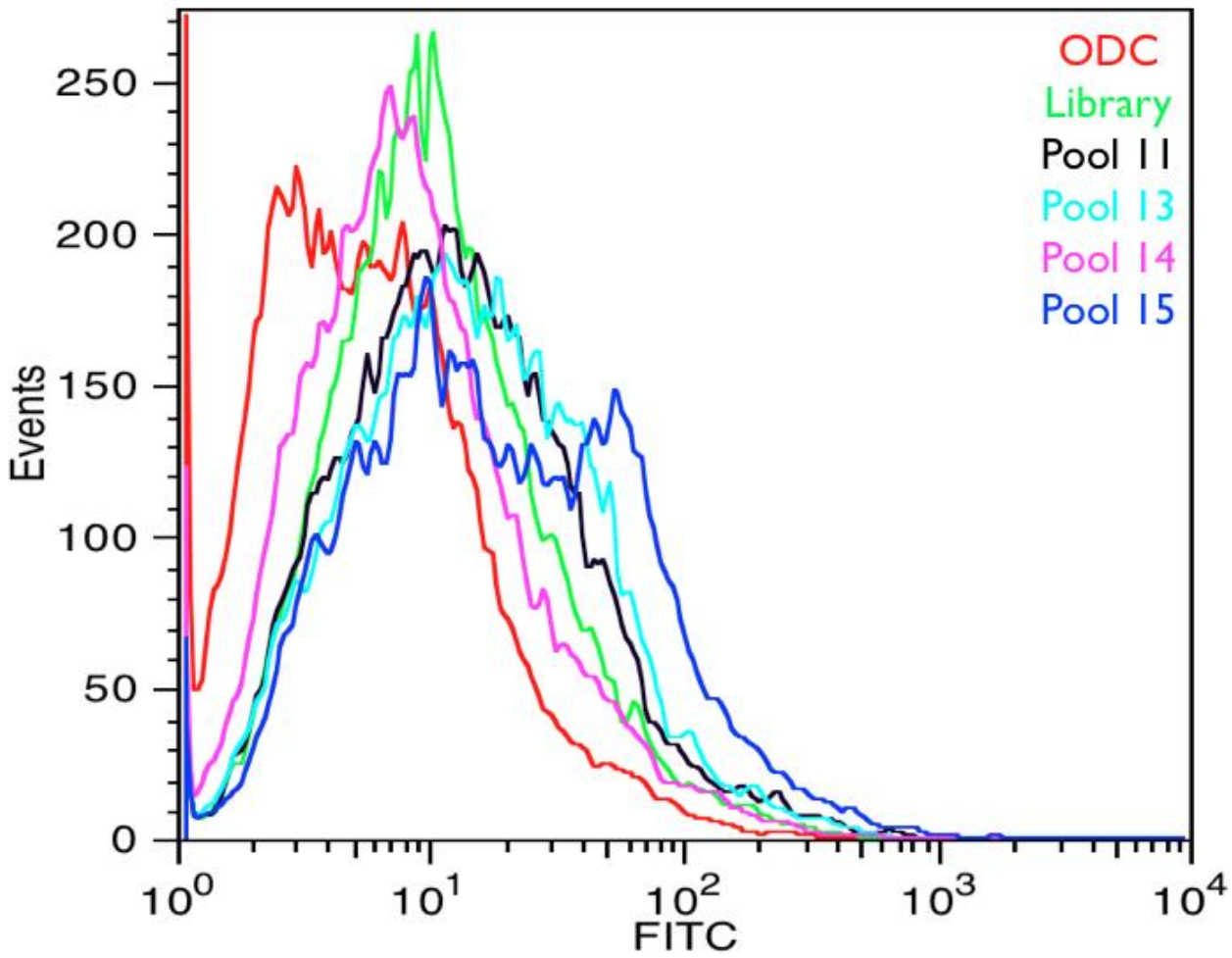


Figure 5-2. Binding profile of the selected pools on the negative cell line, ovarian differentiated cancer cells (ODC). Pool 14 shows less binding than the library; pool 15 shows some binding to ODC, but relatively less than the binding with OCSC

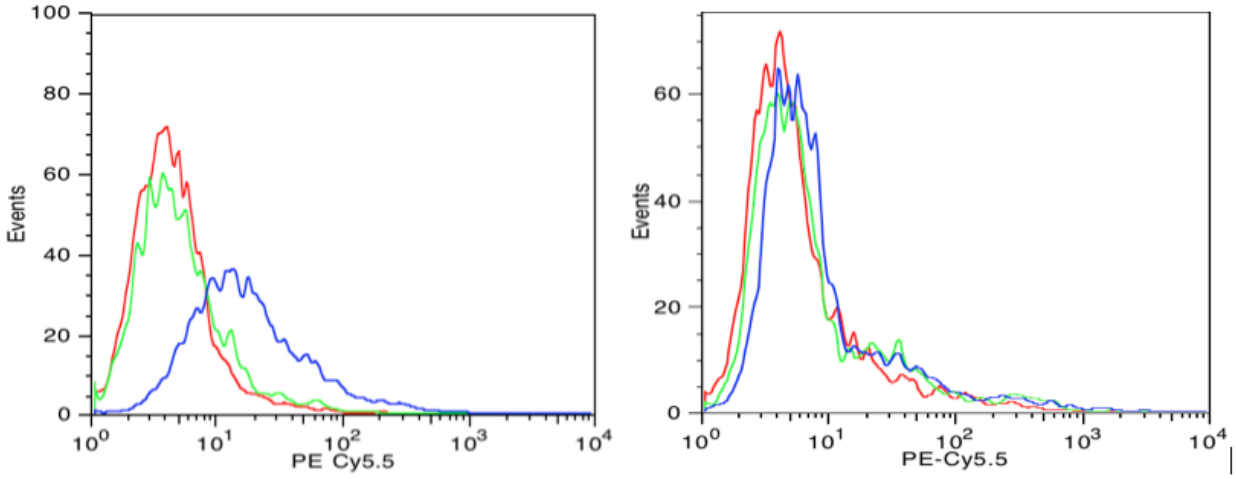


Figure 5-3. Binding assay of DOCSC-1 on undifferentiated OCSC vs differentiated cells

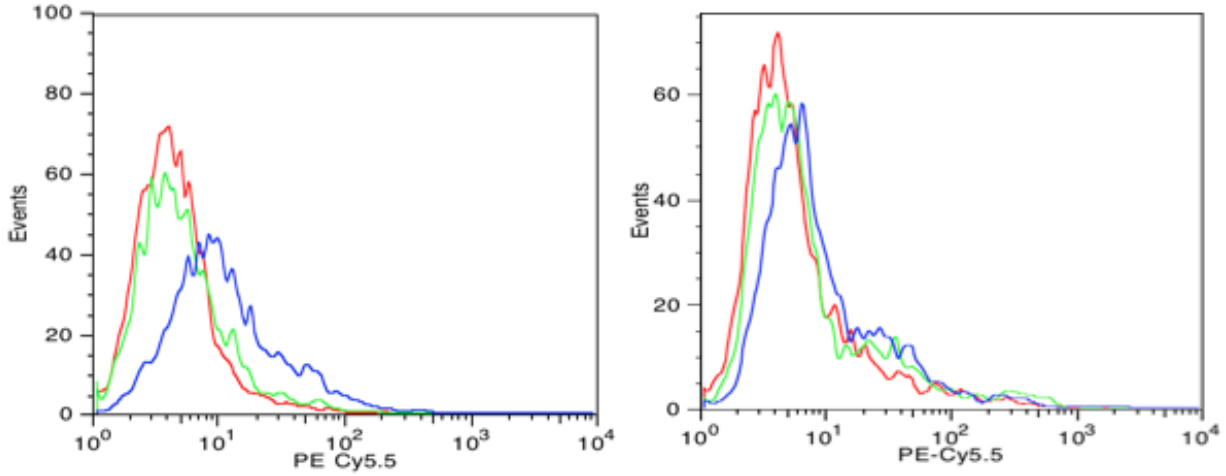


Figure 5-4. Binding assay of DOCSC-2 on undifferentiated OCSC vs differentiated cells

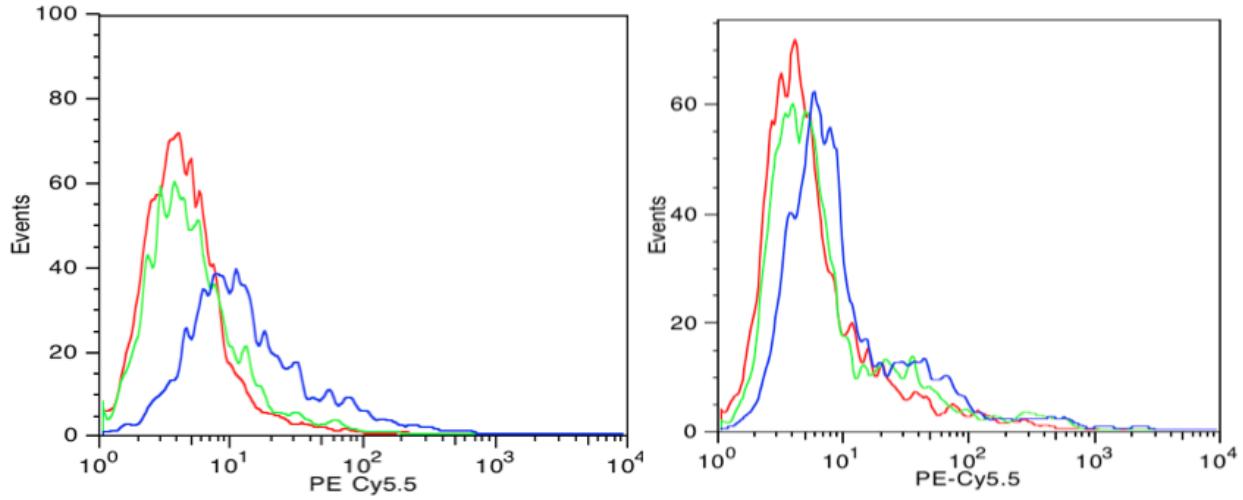


Figure 5-5. Binding assay of DOCSC-3 on undifferentiated OCSC vs differentiated cells

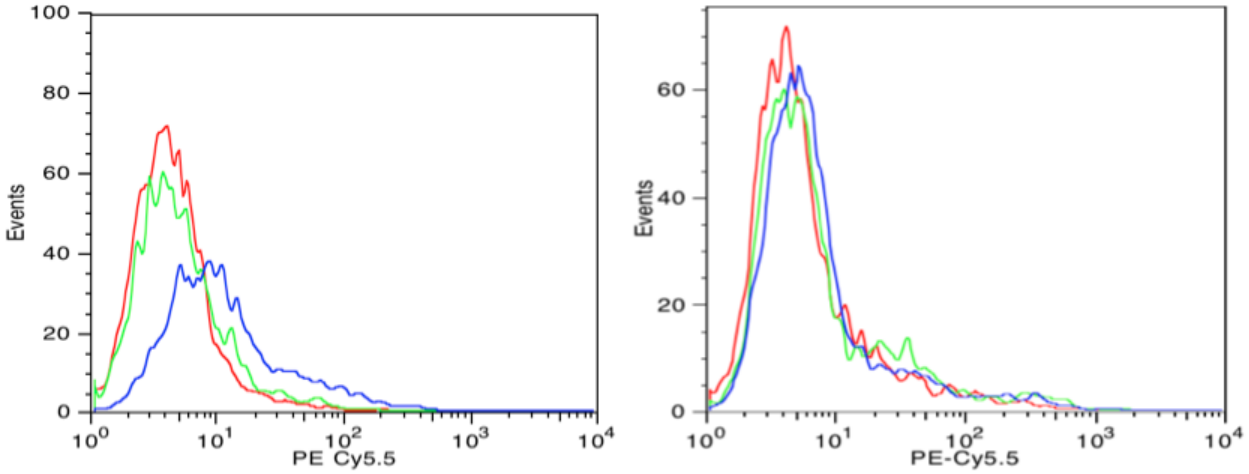


Figure 5-6. Binding assay of DOCSC-4 on undifferentiated OCSC vs differentiated cells

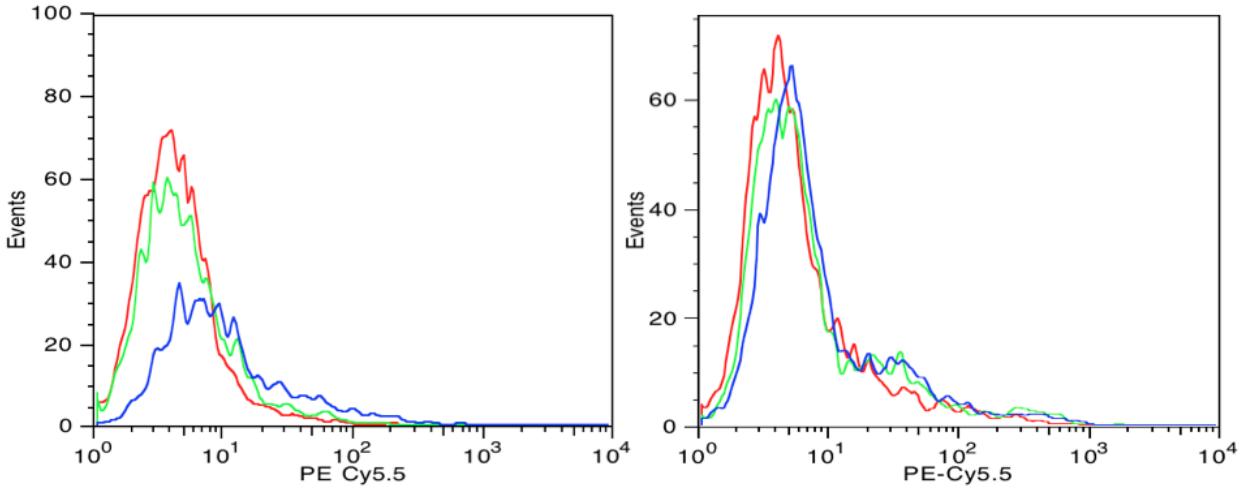


Figure 5-7. Binding assay of DOCSC-5 on undifferentiated OCSC vs differentiated cells

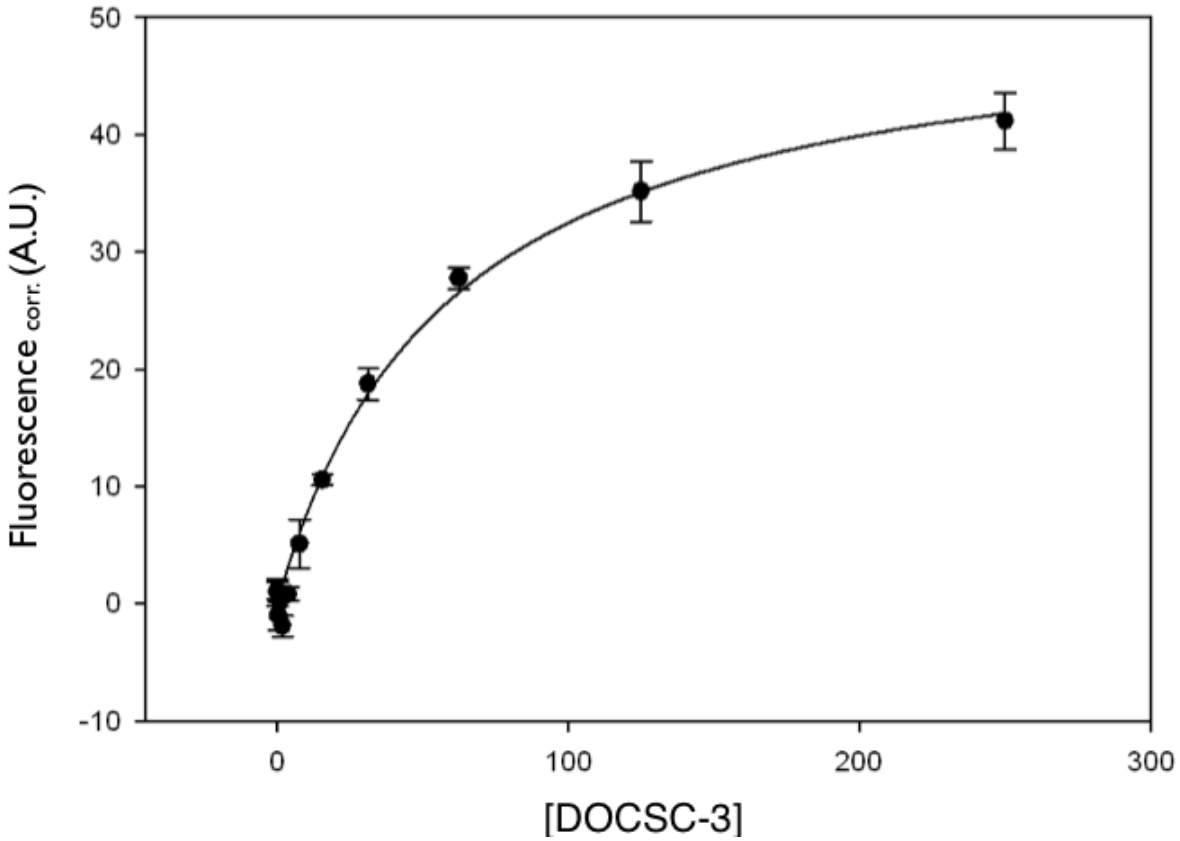


Figure 5-8. Example of a K_d determination for the aptamer DOCSC-3. The fluorescence is corrected against the signal for the random library. The error bars are the standard deviation for n = 3

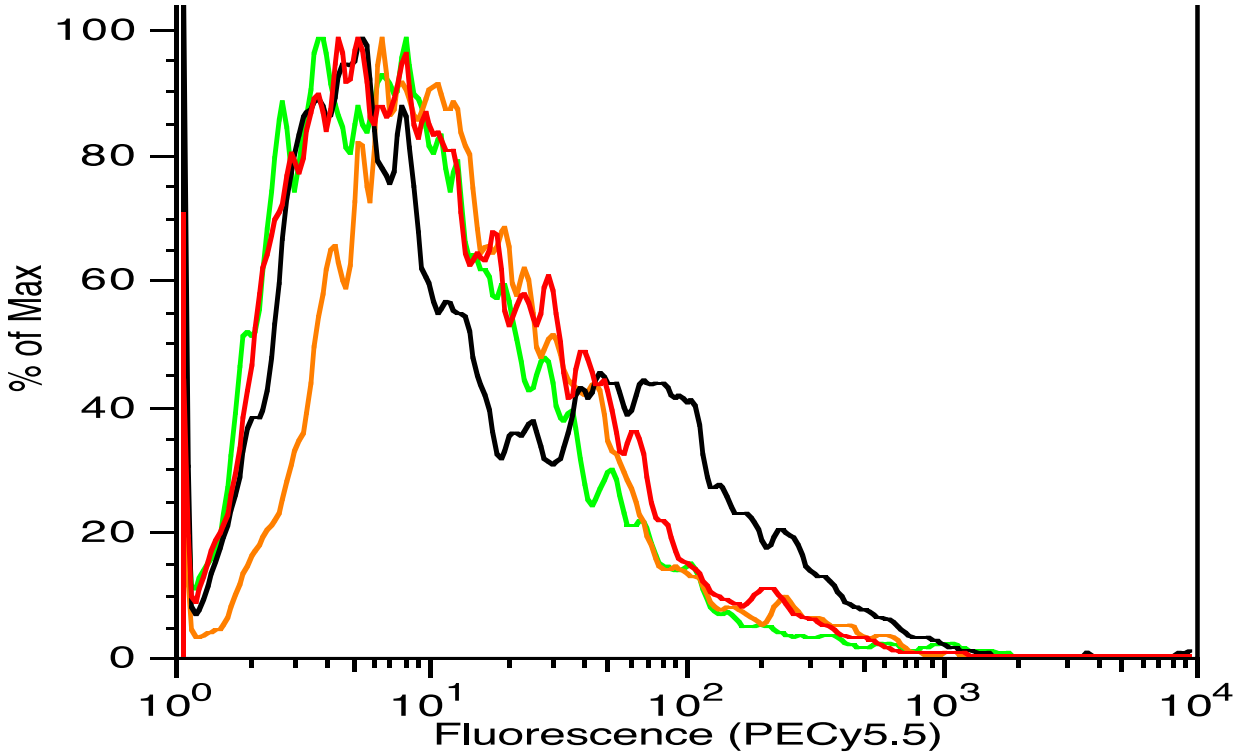


Figure 5-9. Cell sorting experiment with DOCSC-5. Red: Cells; Green: Library; Orange: DOCSC-5 unsorted; Black: DOCSC-5 after sorting

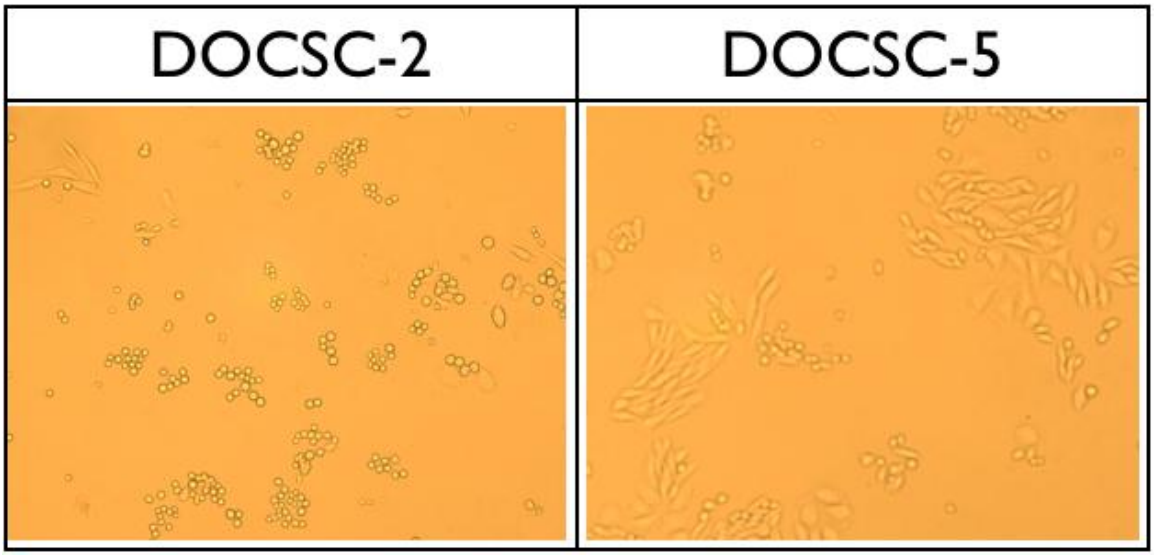


Figure 5-10. Microscope images of one day old cells that have been sorted with aptamer functionalized magnetic beads. Left: DOCSC-2, Right: DOCSC-5

Table 5-1. The aptamers that show specificity towards OCSCs

Name	Sequence	Kd' (nM)
DOCSC 1	5'- ATC CAG AGT GAC GCA GCA TCA TAC CCG AGA TTC ATC ACC CTT ACC TGT CGC TCT GCC TGG ACA CGG TGG CTT AGT A-3'	ND
DOCSC 2	5'- ATC CAG AGT GAC GCA GCA CCC GAC ACA TCT CAT TCA ATT TCG CCT CTC TGG ACA CGG TGG CTT A -3'	3.8±1.6
DOCSC 3	5'- ATC CAG AGT GAC GCA GCA TCA CCA CAC TAC ACA AAT GAT ATT CTC CAA TCC CCC GGC TGG ACA CGG TGG CTT AGT A -3'	59.0 ± 5.9
DOCSC 4	5'- ATC CAG AGT GAC GCA GCA CCA AAC ACA ACT CCG GAA ACG TCA CTA ATC TGC GCA CCT GGA CAC GGT GGC TTA -3'	11.6 ± 2.2
DOCSC 5	5'- ATC CAG AGT GAC GCA GCA CAC CAC CTG ACT ACA TAC CGA ACA TTC GAC TGC TGC GCC TGG ACA CGG TGG CTT AGT A -3'	1.84 ± 0.47

Table 5-2. Binding assay of aptamers from OSCS with other ovarian cancer cell lines

Aptamer	OCSC	dOCSC	CAOV3	TOV-112D	TOV-21G
DOCSC1	+	-	-	-	-
DOCSC2	+	-	-	+	+
DOCSC3	+	-	-	-	-
DOCSC4	+	-	-	-	+
DOCSC5	+	-	-	ND	+

CHAPTER 6
SUMMARY AND FUTURE DIRECTIONS

Summary

In this thesis, the use of cell-SELEX was employed as an innovative way for the determination of a cell surface marker (StIP1) that shows excellent potential for its use as a biomarker for ovarian clear cell carcinoma. This protein also seems to play an important role in the invasive properties of the model cell line TOV-21G. This thesis has outlined the current difficulties that exist in the field of biomarker discovery and explains how cell-SELEX can be used as a complementary technology for rapid, comprehensive determination of biomarker leads. Since the primairy goal of cell-SELEX is selection of aptamers, the method developed in this research has an important advantage, as new ligands with high binding affinities are developed towards these markers.

A pool of ssDNA was enriched against TOV-21G, an ovarian clear cell carcinoma cell line. From this enriched pool of ssDNA sequences, aptamers were selected that show nM to pM range dissociation contants, and did not bind to the negative selection cancer cell line HeLa. More importantly, the aptamers also bound to other cell lines, that were not used at any point during the selection, including CEM, A172, H23, HCT-116. This behavior implies that the cell surface markers of these different types of cancers have commonalities, or more importantly, that the molecular mechanisms of carcinogenesis could be shared between these different types of cancer. More insight toward these mechanisms can prove to be invaluable towards curing or treating the cancer.

After the aptamers were selected, the project turned to the identification of the cell surface markers that bind these aptamers. An important advantage of cell-SELEX is

that aptamer selection does not require any knowledge about the cell surface profile of the cancerous cell studied. However, the selected aptamers can be key players in the identification process. Previous methods toward elucidating an aptamers target have proven to be difficult and are not generally applicable. Markers of interest come with their intrinsic problems, such as low abundance or high hydrophobicity. Problems can also occur because the interaction of the aptamer with the target of interest can be lost when the target is extracted from the cell. The need for a method that reduced these problems therefore presented itself.

In this research the use of formaldehyde as a crosslinker between the protein and the aptamer proved to be an elegant solution for target elucidation for one aptamer of interest, TOV6, which has excellent binding affinity to TOV-21G. Because TOV6 also binds to cancer cells like CEM and A172, identificating the target protein can have significance for leukemic leukemic lymphoblasts and glioblastoma.

After crosslinking the aptamer with the target protein the band to be resected was easily retrieved due to the specific elution from the cell lysate was facilitated by desthiobiotin, which was conjugated to the aptamer. Because of the specific and easy elution using biotin, a clear protein band could be resected from SDS-PAGE and submitted for mass spectral analysis. The MS results indicated that the target protein was StIP1 (also known as HOP).

The identity of the protein target for TOV6 was confirmed to be StIP1 by several methods. The expression of StIP1 was hampered by the use of StIP1 siRNA silencing. This silencing experiment was repeated on A172 cells (which also bound TOV6), where A172 showed the same behavior as TOV-21G, the binding with TOV6 was reduced. All

the knockdown experiments hampered the expression of other non-related membrane proteins as proven by the binding of aptamer SGC8 to its target (targeting PTK7, a protein involved in planar cell polarity). The experiment was repeated with the StIP1 antibody, giving confirmation of the expression of StIP1 in TOV-21G. Finally, the binding of TOV6 was confirmed by an aptamer blot on rhStIP1.

The literature suggested that StIP1 may play an important role in cancer biology, more specifically in the stabilization of certain proteins as HSP90's co-chaperone and also in the pericellular activation of MMPs. The expression of StIP1 has been confirmed in the work of Walsh *et al.* and Wang *et al.*. Wang and co-workers showed the potential of using StIP1 in combination with other ovarian cancer biomarkers for early detection of ovarian cancer. Walsh *et al.* proved that StIP1 plays a crucial role in the invasive capability of Panc-1, a pancreatic cancer line. It has been shown in this thesis that the same mechanism governs TOV-21G invasion. Experiments for this showed that aptamer TOV6 is cytostatic to TOV-21G and that the aptamer inhibits TOV-21G invasion. Further study is needed to elucidate the mechanisms of these phenomena. Characterization of other TOV aptamers will most likely result in the identification of HSP90 as the target, as this protein is always associated with StIP1 and also plays a role in metastasis.

In this work, aptamers were also selected for ovarian cancer stem cells. Even though there are no known stable cell surface markers on cancer stem cells, cell-SELEX was still able to select a pool of aptamers that show binding to undifferentiated cells, but not to differentiated cells. Binding studies on five aptamers from the pool yielded Kd values in the lower to middle nano-molar range. The results indicate good

promise that some of these aptamers bind to cancer stem cells, but further experimental validation is required.

Future Work

This thesis provides an example of the additional dimension that cell-SELEX can add to the field of biomarker discovery. Today, the search for biomarker discovery is hampered by the massive amounts of data generated in proteomic studies, while some of the more interesting biomarkers may never be detected considering the limitations of current analytical methods. Although the limit of detection for proteins can reach close to the attomolar region (in MS), this is only the case for known proteins. When it comes to the analysis of unknown samples (e.g., search for a biomarker in blood), it is common that the ionization of important peptides can be suppressed by more abundant peptides of lesser significance. Membrane proteins are especially victim of this, because they are usually comprised of low polarity amino acids that ionize poorly and have a relatively low solubility. This research has demonstrated the value of cell-SELEX in locating interesting surface proteins with the added benefits of mass spectroscopy to determine their identities. Furthermore, cell-SELEX provides new ligands for the very targets that are identified. The field of cell-SELEX will gain in efficiency andcapability as methodologies develop: for example selection of aptamers without the primer regions; use of next-gen sequencing with fewer rounds and the use of cell sorting to make the cell-SELEX process faster.

The insight that can be gained from further study of cancers through cell-SELEX is tremendous. The study of ovarian cancer, as outlined in this thesis, has provided insight not only in the pathogenesis of this particular cancer, but also other malignancies, such as glioblastoma, leukemia and colon cancer. As this field gains

ground and becomes more developed, the insight in how a cancer interacts with its environment will increase to provide the required knowledge for advances in the treatment of this collection of deadly diseases.

The aptamers selected from cell-Selex can be used for drug discovery via a number of approaches. With the right selection model, aptamers can be selected that bring drugs to the cancer cells, but not to healthy tissue. Also, the identification of new proteins that are expressed in these cancers will lead to improvements in drug discovery and development. Aptamers can be used as a tool for identifying unknown pathways, which could then be targeted with novel small molecule drugs, or by the aptamer itself. The aptamer could also be designed to act as a precursor to a highly toxic drug. The aptamer drug would be specific in the sense that the aptamer would only bind to the cancer, avoiding toxic effects the surrounding healthy tissue. This would also allow the effective dosage for the patient to be lowered, resulting in more cost effective treatment with fewer side-effects. Aptamers themselves can also have therapeutic efficacy. This has been demonstrated in the Macugen aptamer (blocks angiogenesis in the eye) and also in the nucleolin aptamer, that has shown to induce cell death.

It should be mentioned that the selection of aptamers for cells should be undertaken with a word of caution. The work described in this dissertation used *in vitro* cultures, but today's understanding of cancer is suggests that not all surface molecules are the same between *in vitro* cultures and xenografts or real tumors. There is more and more evidence that the expression of these molecules may be partially defined by their biological niches. The causes for a cell to become cancerous are rarely found in the

environmental changes, but due to genetic change within the cells. However, because genetic mutations allow cells to adapt and grow uninhibited; when selecting aptamers for therapy it is important to consider the cells in balance with the niche where they grow. On the other hand selecting aptamers for cancer biomarkers may be to find the markers that define the mechanisms of the cancerous cell, and that do not respond to the niche environment. This leads to the question of the the difference is between the membrane proteome of a xenograft derived from a cell-line and the membrane proteome from an *in vitro* culture, a study that could also be conducted with cell-SELEX.

Aptamer TOV6 has identified StIP1 as a protein that plays a role in metastasis. More importantly, the aptamer itself can significantly inhibit metastasis and potentially help prevent metastasis from occuring. But, the aptamer itself must be further optimized if it is to be used as a drug. For this, a structural optimization step will be needed, as 76 nucleotide long aptamer drugs are not economically feasible at the moment of writing. Alternatively, the answer may also lie in the development of small molecules that directly inhibit the formation of the complex in which StIP1 is involved for blockage of the tumor's ability for metastasis.

Wang *et al.* have demonstrated that StIP1 can be used for diagnosis of ovarian cancer in combination with other biomarkers. An important improvement would be use of TOV6 for Enzyme Linked Aptamer Sorbent Assays (ELASAs) or a hybrid form of this with an antibody, especially since most rare proteins do not have the antibodies with the properties required to work in ELISAs.

Furthermore, the mechanism by which TOV6 inhibits the invasion of TOV-21G needs to be investigated. Is invasion prohibited by the inhibition of the protein at the

membrane level, or by the inhibition of secreted StIP1? A deeper understanding about this mechanism will prove important towards the elucidation on how invasion is regulated in the complex biology of cancer. A hypothesis that can be made here is that StIP1 is plays an important role in the stabilization of HSP90 in the cancer cell membrane.

APPENDIX
THE ANALYSIS OF NEXT-GEN SEQUENCING DATA FOR CELL-SELEX

Introduction

Next generation sequencing or high throughput sequencing are techniques which that allow massive throughput and which generates usually thousands to millions of DNA sequences. This thesis research has demonstrated the need for this information, as well as the advantage of having sequence data. In this thesis, the methods developed by 454 Life Technologies and IonTorrent technology were used.

These technologies are is based on the release of either light, or a proton as nucleotides are added to beads in an emulsion. Each of these beads contains one sequence, and with the use of a sensitive charge-coupled device (CCD) camera, the synthesis is followed as triphopho-nucleotides are added. Then sulfurylase converts pyrophosphate to ATP, which in turn activates luciferin, which emits a light that can be detected by the CCD[170]. In IonTorrent, this light based system is replaced by the detection of the pH change that accompanies the addition of a nucleotide to the strand[171].

Pools from the selection can be analyzed by the elongation of the primer sequences with adapter sequences. These sequences allow the pool to be oriented in the right direction, in order to have the forward sequence of the dsDNA that needs to be analyzed. The actual run and analysis is done by the 454 or IonTorrent instrument, and in the end, the sequences in the library are returned. If more pools need to be sequenced, there is the option to include additional sequences between the template specific primer and the bead-specific primer (Figure A-1).

The Trimming of the Sequencing Data With PERL

Sequencing data are provided as a FASTA file, that contains the sequences of the actual pools, which are considered to be representative of the entire pool. However, if the sequences need to be ordered according to their homology, it is imperative that the template specific primer be trimmed fromdata, with the help of bioinformatics. A PERL script was written that can easily remove the sequences of the primers, remove sequences that are too short, and order sequences according to words they may contain.

In the first script ‘prep’, the sequences are prepared for primer removal. The script removes all the data in the title line of the fasta line and links all the sequence data in one long string.

```
#!/usr/bin/perl -w

open(FASTA, <STDIN>) || die "Sorry, I couldn't find that file... \n";
open(OUT, ">OUTPUT.fna");
my $seq;
my $prim1;
my $prim2;

while (my $line = <FASTA>){
chomp ($line);
if ($line =~ /^>/) {
@fasta = split(/ /, $line);
$string = join(" ", @fasta);
print "\n", $string, "\n";
print OUT "\n", $string, "\n";
}
else{
$seq = $line;
print $seq;
print OUT $seq;
}
}
print "\n";
print OUT "\n";
close(FASTA);
```

‘Prep’ is then followed by ’newgen’:

```
#!/usr/bin/perl -w

open(FASTA1, <STDIN>) || die "Sorry, I couldn't find that file... \n";
```

```
open(OUT1, ">MINING.fna");
open(RY, ">RY.fna");
open(SW, ">SW.fna");

print "Enter your first primer sequence:\n";
$prim1 = <STDIN>;
print "Now, please enter the length of the random sequence:\n";
$random = <STDIN>;

$begin = length($prim1);

while (my $line = <FASTA1>){
chomp ($line);
if ($line =~ /^>/) {
@fasta = split(/ /, $line);
print $fasta[0], "\n";
print OUT1 $fasta[0], "\n";
print RY $fasta[0], "\n";
print SW $fasta[0], "\n";
}

else{
$seq = $line;
$total = length($line);
$subtot = $random + $begin;
if ($total > $subtot){
$offset = $total - $begin;
$seq = substr($seq, $begin, $random);

$ry = $seq;
$sw = $seq;
$ry =~ tr/ACGTacgt/RYRYryry/;
$sw =~ tr/ACGTacgt/WSSWwssw/;

print $seq, "\n";
print OUT1 $seq, "\n";
}
else{
print OUT1 FALSE,"\n";
}
}
}
print "\n";
print OUT1 "\n";
print RY "\n";
print SW "\n";
close(FASTA1);
```

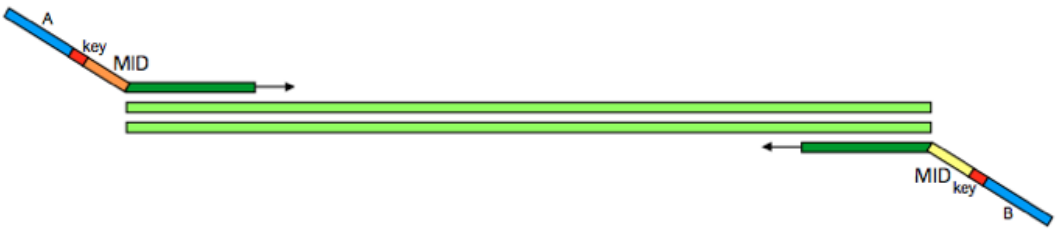
In this ‘newgen’, the file that needs to be trimmed is prompted. Following that, the program asks for the primer and the length of the random sequence. The output of this script is the sequence of the input data, after the primer has been removed. ‘Newgen’ also removes the sequences after the length specified by the user. If the sequence of this output is too short, the sequence is named ‘FALSE’. ‘Aptacount’ allows for the

sorting of interesting families after a preliminary multiple sequence alignment[160]. The sequences named ‘FALSE’ can be removed in similar fashion.

```
#!/usr/bin/perl -w
# use strict;

open(FASTA1, <STDIN>) || die "Sorry, I couldn't find that file... \n";
open(OUT, ">aptamer2.fna");
open(OUT1,">newpool.fna");

print "Enter your putative aptamer sequence:\n";
my $apt = <STDIN>;
chomp ($apt);
@pool = "";
my $n = 0;
my $total = 0;
my $m = 0;

while (my $line = <FASTA1>) {
chomp ($line);
$total++;
# @pool = split(/  /, $line);

if ($line =~ /^>/) {
$pool[0] = $line;
}
if ($line =~ m/($apt)/i){
$n++;
$pool[1] = $line;
if (defined ($pool[1])){
print OUT $pool[0], "\n", $pool[1], "\n";
$pool[1] = "";
}
}
elsif (($line !~ m/($apt)/i) && ($line !~ m/^>/)){
$m++;
$rest[1] = $line;
if (defined ($rest[1])){
$rest[1] =~ tr/AGCT/agct/;
print OUT1 $pool[0], "\n", $rest[1], "\n";
$rest[1] = "";
}
}
}
$n *= 100;
$total /= 2;
# $remain = $total - $m;
$m *= 100;
print $total, "\n";
print "The number of aptamers of this type is: ", $n/100, ".";
print "This is ", $n/$total, "%.\n";
print $m/100, "sequences remaining.\n\n";
print "That's ",$m/$total, "%.\n\n";
exit;
```

Future Work

Improvements in the data processing would be an incorporation of the three scripts in one maser script. Additionally, the software could be expanded with MAFFT software, so sequence data can directly be generated by simply inserting the input file.



Forward primer (Primer A-Key):

5’-CGTATCGCCTCCCTCGCGCCATCAG-{MID}-{template-specific-sequence}-3’

Reverse primer (Primer B-Key):

5’-CTATGCGCCTTGCCAGCCCGCTCAG-{MID}-{template-specific-sequence}-3’

Figure A-1. The scheme of a pool’s primers submitted for sequencing. The blue portions of the DNA are the sections of the primers that allow proper orientation with the beads. The MID sequences are sequences that can be inserted in case more then one sample is analysed. This is attached to the pools-to-analyse by PCR with the help of the template specific sequences

LIST OF REFERENCES

1. American Cancer Society. Cancer Facts & Figures 2011. Atlanta: American Cancer Society (2011).

2. U.S. Public Health Service, Vital Statistics of the United States, annual, Vol. I and Vol II; (1900-1970), U.S. National Center for Health Statistics, Vital Statistics of the United States, annual; National Vital Statistics Report (NVSR) (formerly Monthly Vital Statistics Report)(1971-2001).

3. Lucy Gilbert, Olga Basso, John Sampalis, Igor Karp, Claudia Martins, Jing Feng, Sabrina Piedimonte, Louise Quintal, Agnihotram V Ramanakumar, Janet Takefman, Maria S Grigorie, Giovanni Artho, Srinivasan Krishnamurthy. Assessment of symptomatic women for early diagnosis of ovarian cancer: results from the prospective DOvE pilot project - The Lancet Oncology - Medical Journal. *Lancet Oncol; 13:* 285-91 (2012).

4. Kosary C. Cancer of the ovary. In: Ries LAG, Young JL, Keel GE, Eisner MP, Lin YD, et al., editors. SEER survival monograph: cancer survival among adults: US SEER Program, 1988–2001, patient and tumor characteristics. Bethesda (Maryland): National Cancer Institute, SEER Program (2007).

5. Beahrs, OH, Henson DE, Hutter RVP, Myers MH (eds). AJCC Cancer Staging Manual, Third edition. American Joint Committee on Cancer. Philadelphia: Lippincott (1988).

6. Goff, B., Mandel, L., Melancon, C., & Muntz, H. (2004). Frequency of symptoms of ovarian cancer in women presenting to primary care clinics. Journal of the American Medical Association, 291(22), 2705–2712.

7. Goff, B., Mandel, L., Drescher, C., Urban, N., Gough, S., Schur- man, K., *et al.* (2007). Development of an Ovarian Cancer Symptom Index: Possibilities for earlier detection. Cancer, 109(2), 221–227.

8. Yurkovetsky, Z., Skates, S., Lomakin, A., Nolen, B., *et al.* Development of a multimarker assay for early detection of ovarian cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **28**, 2159-66 (2010).

9. Bast, R., Brewer, M., Zou, C., Hernandez, M. A., Daley, M., Ozols, R., Lu, K., *et al*. Prevention and Early Detection of Ovarian Cancer : Mission Impossible ? Cancer. (2007).

10. Cesario, S. Advances in the early detection of ovarian cancer: How to hear the whispers early. *Nursing for women's health* **14**, 222-34 (2010).

11. http://www.uspreventiveservicestaskforce.org/3rduspstf/ovariancan/ovcanrs.htm. Last accessed June 14, 2012.

12. http://www.newchoicehealth.com/Directory/Procedure/60/Transvaginal%20Ultrasound. Last accessed June 14, 2012.

13. http://www.medicinenet.com/ca_125/page2.htm Last accessed June 14, 2012.

14. Haga Y., Sakamoto K., Egami H., Yoshimura R., Akagi M. Evaluation of serum CA125 values in healthy individuals and pregnant women. *The American journal of the medical sciences*, **292**, 25-9 (1986).

15. Bast, R.C., Badgwell, D., Lu Z., Marquez, R. *et al.* New tumor markers: CA125 and beyond. *International journal of gynecological cancer : official journal of the International Gynecological Cancer Society* **15 Suppl 3**, 274-81

16. Rifai, N., Gillette, M. a & Carr, S. A Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nature biotechnology* **24**, 971-83 (2006).

17. Verhoef, G., Meeus, P., Stul, M., Mecucci, C., *et al.* Cytogenetic and molecular studies of the Philadelphia translocation in myelodysplastic syndromes. *Cancer Genetics and Cytogenetics* **59**, 161-166 (1992).

18. Sandoval, J. & Esteller, M. Cancer epigenomics: beyond genomics. *Current opinion in genetics & development* **22**, 55-50 (2012).

19. Mikolajczyk, S.D. Millar, L. S., Tsinberg, P., Coutts, S. M. *et al.* Detection of EpCAM-Negative and Cytokeratin-Negative Circulating Tumor Cells in Peripheral Blood. Journal of Oncology **2011**, 1-10 (2011).

20. Nielsen, K.V., Ejlertsen, B., Møller, S., Jørgensen, J. T. *et al.* The value of TOP2A gene copy number variation as a biomarker in breast cancer: Update of DBCG trial 89D. *Acta oncologica (Stockholm, Sweden)* **47**, 725-34 (2008).

21. *Huang, Z., Lin, L., Gao, Y., Chen, Y. et al.* Bladder cancer determination via two urinary metabolites: a biomarker pattern approach. *Molecular & cellular proteomics : MCP* **10**, M111.007922 (2011).

22. Hanash, S.M., Pitteri, S.J. & Faca, V.M. Mining the plasma proteome for cancer biomarkers. *Nature* **452**, 571-9 (2008).

23. Haga, Y., Sakamoto, K., Egami, H., Yoshimura, R. & Akagi, M. Evaluation of serum CA125 values in healthy individuals and pregnant women. *The American journal of the medical sciences* **292**, 25-9 (1986).

24. Frank, R. & Hargreaves, R. Clinical biomarkers in drug discovery and development. *Nature reviews. Drug discovery* **2**, 566-80 (2003).

25. Adachi, J., Kumar, C., Zhang, Y., Olsen, J.V. & Mann, M. The human urinary proteome contains more than 1500 proteins, including a large proportion of membrane proteins. *Genome biology* **7**, R80 (2006).

26. Almen, M.; Nordstrom, K.; Fredriksson, R.; Schioth, H. Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin. *BMC Biol* **7**, 50 (2009).

27. Garrow, A., Agnew, A.; Westhead, D. TMB-Hunt: an amino acid composition based method to screen proteomes for beta-barrel transmembrane proteins. *BMC Bioinform* **6**, 56 (2005).

28. Luckey, M. Membrane Structural Biology - With Biochemical and Biophysical Foundations; Cambridge University Press: New York (2008).

29. Klijn, J.G.M., Berns, P.M.J.J., Schmitz, P.I.M. & Foekens, J.A. The Clinical Significance of Epidermal Growth Factor Receptor (EGF-R) in Human Breast Cancer: A Review on 5232 Patients. *Endocrine Reviews* **13**, 3-17 (1992).

30. Sato, H., Takino, T., Okada, Y., Cao, J. *et al.* A matrix metalloproteinase expressed on the surface of invasive tumour cells. *Nature* **370**, 61-5 (1994).

31. Igney, F.H. & Krammer, P.H. Death and anti-death: tumour resistance to apoptosis. *Nature reviews. Cancer* **2**, 277-88 (2002).

32. Srinivas, P.R., Verma, M., Zhao, Y. & Srivastava, S. Proteomics for Cancer Biomarker Discovery. *Clin. Chem.* **48**, 1160-1169 (2002).

33. Delahunty, C. & Yates, J.R. Protein identification using 2D-LC-MS/MS. *Methods (San Diego, Calif.)* **35**, 248-55 (2005).

34. Shin, B. K., Wang, H., Yim, A. M., Le Naour, F. *et al.* Global profiling of the cell surface proteome of cancer cells uncovers an abundance of proteins with chaperone function. *The Journal of biological chemistry* **278**, 7607-16 (2003).

35. Barrera, N.P. & Robinson, C.V. Advances in the Mass Spectrometry of Membrane Proteins: From Individual Proteins to Intact Complexes. *Annu. Rev. Biochem.* **80**, 247–71 (2011).

36. Tang, N., Tornatore, P. & Weinberger, S.R. Current developments in SELDI affinity technology. *Mass spectrometry reviews* **23**, 34-44

37. *Targeted Therapies*. (Humana Press: Totowa, NJ, 2011).

38. Burczynski, M.E. & Dorner, A.J. Transcriptional profiling of peripheral blood cells in clinical pharmacogenomic studies. *Pharmacogenomics* **7**, 187-202 (2006).

39. Elshal, M.F. & McCoy, J.P. Multiplex bead array assays: performance evaluation and comparison of sensitivity to ELISA. *Methods (San Diego, Calif.)* **38**, 317-23 (2006).

40. Domon, B. & Aebersold, R. Mass spectrometry and protein analysis. *Science (New York, N.Y.)* **312**, 212-7 (2006).

41. Garratty, G. How concerned should we be about missing antibodies to low incidence antigens? *Transfusion* **43**, 844-847 (2003).

42. Udugamasooriya, D.G. & Kodadek, T. On-Bead Two-Color (OBTC) Cell Screen for Direct Identification of Highly Selective Cell Surface Receptor Ligands. *Current protocols in chemical biology* **4**, 35-48 (2012).

43. Wang, X., Yu, J., Sreekumar, A., Varambally, S. *et al.* Autoantibody Signatures in Prostate Cancer. *New England Journal of Medicine* **353**, 1224-1235 (2005).

44. Kim, Y., Liu, C. & Tan, W. Aptamers generated by Cell SELEX for biomarker discovery. *Future Medicine* **3**, 193-202 (2009).

45. Keefe, A.D., Pai, S. & Ellington, A. Aptamers as therapeutics. *Nature reviews. Drug discovery* **9**, 537-50 (2010).

46. Felix Hoppe-Seyler, Irena Crnkovic-Mertens, Evangelia Tomai & Karin Butz Peptide Aptamers: Specific Inhibitors of Protein Function. *Current Molecular Medicine* **4**, 10 (2004).

47. Vaught, J. D.,Bock, C., Carter, J.,Fitzwater, T. *et al.* Expanding the chemistry of DNA for in vitro selection. *Journal of the American Chemical Society* **132**, 4141-51 (2010).

48. Mayer, G. The chemical biology of aptamers. *Angewandte Chemie (International ed. in English)* **48**, 2672-89 (2009).

49. Tuerk, C. & Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505-510 (1990).

50. Ellington, A.D. & Szostak, J.W. In vitro selection of RNA molecules that bind specific ligands. *Nature* **346**, 818-822 (1990).

51. Ciesiolka, J. & Yarus, M. Small RNA-divalent domains. RNA 2, 785-793 (1996).

52. Jiang, L., Majumdar, A., Hu, W., Jaishree T.J. *et al.*, Saccharide–RNA recognition in a complex formed between neomycin B and an RNA aptamer. *Structure* **7**, 817-827 (1999).

53. Jiang, F., Gorin, A., Hu, W., Majumdar A. *et al.*, Anchoring an extended HTLV-1 Rex peptide within an RNA major groove containing junctional base triples *Structure* **7**, 1461 (1999).

54. Orito, N. *et al.* High-affinity RNA aptamers to C-reactive protein (CRP): newly developed pre-elution methods for aptamer selection. *Journal of Physics: Conference Series* **352**, 012042 (2012).

55. Zhang, G. & Simon, A.E. A Multifunctional Turnip Crinkle Virus Replication Enhancer Revealed by in vivo Functional SELEX. *Journal of Molecular Biology* **326**, 35-48 (2003).

56. Chen, F., Zhou, J., Luo, F., Mohammed, A.-B. & Zhang, X.-L. Aptamer from whole-bacterium SELEX as new therapeutic reagent against virulent Mycobacterium tuberculosis. *Biochemical and biophysical research communications* **357**, 743-8 (2007).

57. Graham, J.C. & Zarbl, H. Use of Cell-SELEX to Generate DNA Aptamers as Molecular Probes of HPV-Associated Cervical Cancer Cells. *PLoS ONE* **7**, e36103 (2012).

58. Lee, J.F., Stovall, G.M. & Ellington, A.D. Aptamer therapeutics advance. *Current opinion in chemical biology* **10**, 282-9 (2006).

59. Healy, J.M. *et al.* Pharmacokinetics and Biodistribution of Novel Aptamer Compositions. *Pharmaceutical Research* **21**, 2234-2246 (2004).

60. Pestourie, C., Tavitian, B. & Duconge, F. Aptamers against extracellular targets for in vivo applications. *Biochimie* **87**, 921-30 (2005).

61. Pendergrast, P.S., Marsh, H.N., Grate, D., Healy, J.M. & Stanton, M. Nucleic acid aptamers for target validation and therapeutic applications. *Journal of biomolecular techniques : JBT* **16**, 224-34 (2005).

62. Cole S.P., Campling B.G., Atlaw T., Kozbor D., Roder J.C. Human monoclonal antibodies . *Mol. Cell. Biochem* **62**, 109–20(1984).

63. Drabovich, A.P., Berezovski, M., Okhonin, V. & Krylov, S.N. Selection of smart aptamers by methods of kinetic capillary electrophoresis. *Analytical chemistry* **78**, 3171-8 (2006).

64. Gragoudas, E. S., Adamis, A. P., Cunningham, E. T., Feinsod, M., & Guyer, D. R. Pegaptanib for Neovascular Age-Related Macular Degeneration. *New England Journal of Medicine* **351**, 2805-2816 (2004).

65. http://www.bccresearch.com/report/BIO071A.html?utm_campaign=Global%20market%20for%20aptamers%20worth%20%25241%252E8%20billion%20by%202014&utm_content=tom@aptagen.com&utm_medium=Email&utm_source=VerticalResponse&utm_term=Global%20market%20for%20aptamerscampaign. Last accessed June 14, 2012.

66. US Patent 5,141,813

67. Breaker, R.R. Complex riboswitches. *Science* **319**, 1795-7 (2008).

68. Hermann, T. Adaptive Recognition by Nucleic Acid Aptamers. *Science* **287**, 820-825 (2000).

69. Baskerville, S., Zapp, M. & Ellington, A.D. Anti-Rex Aptamers as Mimics of the Rex-Binding Element. *J. Virol.* **73**, 4962-4971 (1999).

70. Zimmermann, G.R., Jenison, R.D., Wick, C.L., Simorre, J.-P. & Pardi, A. Interlocking structural motifs mediate molecular discrimination by a theophylline-binding RNA. *Nature Structural Biology* **4**, 644-649 (1997).

71. Waring, M.J. Complex formation between ethidium bromide and nucleic acids. *Journal of Molecular Biology* **13**, 269-282 (1965).

72. Fan P., Suri A.K., Fiala R., Live, D. Patel D.J. (1996). Molecular recognition in the FMN-RNA aptamer complex. *J Mol Biol*, **258**, 480-500.

73. Jiang, F., Kumar, R.A., Jones, R.A. & Patel, D.J. Structural basis of RNA folding and recognition in an AMP-RNA aptamer complex. *Nature* **382**, 183-6 (1996).

74. Lin, C.H. & Patei, D.J. Structural basis of DNA folding and recognition in an AMP-DNA aptamer complex: distinct architectures but common recognition motifs for DNA and RNA aptamers complexed to AMP. *Chemistry & Biology* **4**, 817-832 (1997).

75. Horn, W. T., Convery, M. A., Stonehouse, N. J., Adams, C. J., *et al.* The crystal structure of a high affinity RNA stem-loop complexed with the bacteriophage MS2 capsid: further challenges in the modeling of ligand-RNA interactions. *RNA (New York, N.Y.)* **10**, 1776-82 (2004).

76. Long, S.B., Long, M.B., White, R.R. & Sullenger, B.A. Crystal structure of an RNA aptamer bound to thrombin. *RNA (New York, N.Y.)* **14**, 2504-12 (2008).

77. Huang, D-B., Vu, D., Cassiday, L. A., Zimmerman, J. M. *et al.* Crystal structure of NF-kappaB (p50)2 complexed to a high-affinity RNA aptamer. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 9268-73 (2003).

78. Hopfield, J.J. Kinetic Proofreading: A New Mechanism for Reducing Errors in Biosynthetic Processes Requiring High Specificity. *Proceedings of the National Academy of Sciences* **71**, 4135-4139 (1974).

79. Shangguan, D., Li, Y., Tang, Z., Cao, Z. C. *et al.* Aptamers evolved from live cells as effective molecular probes for cancer study. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 11838-43 (2006).

80. Mi, J., Liu, Y., Rabbani, Z. N., Yang, Z. *et al.* In vivo selection of tumor-targeting RNA motifs. *Nature chemical biology* **6**, 22-4 (2010).

81. Noma T., Ikebukuro K., Sode K., Ohkubo T., *et al.* A screening method for DNA aptamers that bind to a specific, unidentified protein in tissue samples. *Biotechnology letters* **28**, 1377-81 (2006).

82. Fitter, S. & James, R. Deconvolution of a complex target using DNA aptamers. *The Journal of biological chemistry* **280**, 34193-201 (2005).

83. Hicke B.J., Marion C., Chang Y.F., Gould T., *et al.* Tenascin-C aptamers are generated using tumor cells and purified protein. *The Journal of biological chemistry* **276**, 48644-54 (2001).

84. Washburn, M.P., Wolters, D. & Yates, J.R. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature Biotechnology* **19**, 242-247 (2001).

85. Bates, P.J. Antiproliferative Activity of G-rich Oligonucleotides Correlates with Protein Binding. *Journal of Biological Chemistry* **274**, 26369-26377 (1999).

86. Tang, Z., Parekh, P., Turner, P., Moyer, R.W. & Tan, W. Generating aptamers for recognition of virus-infected cells. *Clinical chemistry* **55**, 813-22 (2009).

87. Bayrac A.T., Sefah K., Parekh P., Bayrac C.,. *et al.* In Vitro Selection of DNA Aptamers to Glioblastoma Multiforme. *ACS Chemical Neuroscience* **2**, 175-181 (2011).

88. Sefah, K., Shangguan, D., Xiong, X., O'Donoghue, M.B. & Tan, W. Development of DNA aptamers using Cell-SELEX. *Nature protocols* **5**, 1169-85 (2010).

89. Mallikaratchy P., Tang Z., Kwame S., Meng L., *et al.* Aptamer directly evolved from live cells recognizes membrane bound immunoglobin heavy mu chain in Burkitt's lymphoma cells. *Molecular & cellular proteomics : MCP* **6**, 2230-8 (2007).

90. Berezovski, M.V., Lechmann, M., Musheev, M.U., Mak, T.W. & Krylov, S.N. Aptamer-facilitated biomarker discovery (AptaBiD). *Journal of the American Chemical Society* **130**, 9137-43 (2008).

91. Shangguan D., Cao Z., Meng L., Mallikaratchy P., *et al.* Cell-Specific Aptamer Probes for Membrane Protein Elucidation in Cancer Cells research articles. *Journal of Proteome Research* 2133-2139 (2008).

92. http://clinicaltrials.gov/ct2/show/NCT00632242?term=ARC1779&rank=1. Last accessed June 14, 2012.

93. http://clinicaltrials.gov/ct2/show/NCT01191372?term=archemix&rank=6. Last accessed June 14, 2012.

94. http://www.thepharmaletter.com/file/100055/baxter-buys-all-hemophilia-related-assets-of-archemix-for-30-million-and-285-million-in-potential-milestones.html. Last accessed June 14, 2012.

95. Anglesio, M.S., Carey, M.S., Köbel, M., Mackay, H. & Huntsman, D.G. Clear cell carcinoma of the ovary: a report from the first Ovarian Clear Cell Symposium, June 24th, 2010. *Gynecologic oncology* **121**, 407-15 (2011).

96. Sugiyama T., Kamura T., Kigawa J., Terakawa N., *et al.* Clinical characteristics of clear cell carcinoma of the ovary: a distinct histologic type with poor prognosis and resistance to platinum-based chemotherapy. *Cancer* **88**, 2584-9 (2000).

97. Takano M., Kikuchi Y., Yaegashi N., Kuzuya K., *et al.* Clear cell carcinoma of the ovary: a retrospective multicentre experience of 254 patients with complete surgical staging. *British journal of cancer* **94**, 1369-74 (2006).

98. LoPiccolo, J., Blumenthal, G.M., Bernstein, W.B. & Dennis, P.A. Targeting the PI3K/Akt/mTOR pathway: effective combinations and clinical considerations. *Drug resistance updates : reviews and commentaries in antimicrobial and anticancer chemotherapy* **11**, 32-50

99. Kerschgens, J., Egener-Kuhn, T. & Mermod, N. Protein-binding microarrays: probing disease markers at the interface of proteomics and genomics. *Trends in molecular medicine* **15**, 352-8 (2009).

100. Provencher D.M., Lounis H., Champoux L., Tétrault M., *et al.* Characterization of Four Novel Epithelial Ovarian Cancer Cell Lines. *In Vitro Cellular & Developmental Biology. Animal* **36**, 357-36 (2000).

101. Scherer, W.F., Syverton, J.T. & Gey, G.O. Studies on the propagation in vitro of poliomyelitis viruses. IV. Viral multiplication in a stable strain of human malignant epithelial cells (strain HeLa) derived from an epidermoid carcinoma of the cervix. *The Journal of experimental medicine* **97**, 695-710 (1953).

102. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* **31**, 3406-3415 (2003).

103. Kahn, S.D. On the future of genomic data. *Science (New York, N.Y.)* **331**, 728-9 (2011).

104. Wall L., Programming in Perl **3**, (O'Reilly Media, 2000)

105. Larkin M.A., Blackshields G., Brown N.P., Chenna R.,*et al.*, Clustal W and Clustal X version 2.0., *Bioinformatics* **23**, 2947-2948 (2007).

106. Shamah, S.M., Healy, J.M. & Cload, S.T. Complex target SELEX. *Accounts of chemical research* **41**, 130-8 (2008).

107. Svarovsky S.A. and Joshi L. Biocombinatorial Selection of Carbohydrate Binding Agents of Therapeutic Significance. *Current Drug Discovery Technologies* **5**, 20-28 (2008).

108. Faca V. M., Ventura A. P., Fitzgibbon M. P., Pereira-Faca S. R., *et al.* Proteomic Analysis of ovarian cancer cells reveals dynamic processes of protein secretion and shedding of extra-cellular domains. *PLoS ONE* **3**: e2425(2008).

109. Vogelstein, B. & Kinzler, K.W. Cancer genes and the pathways they control. *Nature medicine* **10**, 789-99 (2004).

110. Garcia, M., Jemal, A., Ward, E., Center, M., Hao, Y., *et al.*, Global cancer facts and figures. Atlanta, GA: American Cancer Society (2007).

111. Lawrenson, K. & Gayther, S.A. Ovarian cancer: a clinical challenge that needs some basic answers. *PLoS medicine* **6**, e25 (2009).

112. Cibiel, A., Dupont, D.M. & Ducongé, F. Methods To Identify Aptamers against Cell Surface Biomarkers. *Pharmaceuticals* **4**, 1216-1235 (2011).

113. Hedin, L.E., Illerg, K. & Elofsson, A. An Introduction to Membrane Proteins. *Journal of Proteome Research* **10**, 3324-3331 (2011).

114. Collas, P. Chromatin Immunoprecipitation Assays. *Methods in Molecular Biology* **567**, (2009). And

115. Solomon, M.J. Formaldehyde-Mediated DNA-Protein Crosslinking: A Probe for in vivo Chromatin Structures. *Proceedings of the National Academy of Sciences* **82**, 6470-6474 (1985).

116. Von Hippel, P.H. & McGhee, J.D. DNA-protein interactions. *Annual review of biochemistry* **41**, 231-300 (1972).

117. Schmidt D., Wilson M.D., Spyrou C., Brown G.D., *et al.* ChIP-seq: using high-throughput sequencing to discover protein-DNA interactions. *Methods (San Diego, Calif.)* **48**, 240-8 (2009).

118. Déjardin, J. & Kingston, R.E. Purification of proteins associated with specific genomic Loci. *Cell* **136**, 175-86 (2009).

119. Van Simaeys D., López-Colón D., Sefah K., Sutphen R., *et al.* Study of the molecular recognition of aptamers selected through ovarian cancer cell-SELEX. *PloS one* **5**, e13770 (2010).

120. Orlando, V., Strutt, H. & Paro, R. Analysis of chromatin structure by in vivo formaldehyde cross-linking. *Methods (San Diego, Calif.)* **11**, 205-14 (1997).

121. Wortmann, A., Jecklin, M.C., Touboul, D., Badertscher, M. & Zenobi, R. Binding constant determination of high-affinity protein-ligand complexes by electrospray ionization mass spectrometry and ligand competition. *Journal of mass spectrometry : JMS* **43**, 600-8 (2008).

122. Annesley, T.M. Ion Suppression in Mass Spectrometry. *Clinical Chemistry* **49**, 1041-1044 (2003).

123. da Silva, V.C.H. & Ramos, C.H.I. The network interaction of the human cytosolic 90kDa heat shock protein Hsp90: A target for cancer therapeutics. *Journal of proteomics* 1-13 (2012).doi:10.1016/j.jprot.2011.12.028

124. Onuoha, S.C., Coulstock, E.T., Grossmann, J.G. & Jackson, S.E. Structural studies on the co-chaperone Hop and its complexes with Hsp90. *Journal of molecular biology* **379**, 732-44 (2008).

125. Wang T.-H., Chao A., Tsai C.L., Chang C.L., *et al.* Stress-induced phosphoprotein 1 as a secreted biomarker for human ovarian cancer promotes cancer cell proliferation. *Molecular & cellular proteomics : MCP* **9**, 1873-84 (2010).

126. Walsh N., O'Donovan N., Kennedy S., Henry M., *et al.* Identification of pancreatic cancer invasion-related proteins by proteomic analysis. *Proteome science* **7**, 3 (2009).

127. Zanata, S.M. *et al.* Stress-inducible protein 1 is a cell surface ligand for cellular prion that triggers neuroprotection. *The EMBO journal* **21**, 3307-16 (2002).

128. https://www.qiagen.com/geneglobe/genesolutionview.aspx?ID=GS10963. Last accessed June 14, 2012.

129. Erlich R.B., Kahn S.A., Lima F.R., Muras A.G., *et al.* STI1 promotes glioma proliferation through MAPK and PI3K pathways. *Glia* **55**, 1690-8 (2007).

130. Lu X., Borchers A.G., Jolicoeur C., Rayburn H., *et al.* PTK7/CCK-4 is a novel regulator of planar cell polarity in vertebrates. *Nature* **430**, 93-8 (2004).

131. Noma T., Ikebukuro K., Sode K., Ohkubo T., *et al.* A screening method for DNA aptamers that bind to a specific, unidentified protein in tissue samples. *Biotechnology letters* **28**, 1377-81 (2006).

132. Jackson, V. Studies on histone organization in the nucleosome using formaldehyde as a reversible cross-linking agent. *Cell* **15**, 945-954 (1978).

133. Shin B.K., Wang H., Yim A.M., Le Naour F., *et al.* Global profiling of the cell surface proteome of cancer cells uncovers an abundance of proteins with chaperone function. *The Journal of biological chemistry* **278**, 7607-16 (2003).

134. Walsh N., Larkin A., Swan N., Conlon K., *et al.* RNAi knockdown of Hop (Hsp70/Hsp90 organising protein) decreases invasion via MMP-2 down regulation. *Cancer letters* **306**, 180-9 (2011).

135. Frank, R. & Hargreaves, R. Clinical biomarkers in drug discovery and development. *Nature reviews. Drug discovery* **2**, 566-80 (2003).

136. Carrigan, P.E. Domain:domain interactions within Hop, the Hsp70/Hsp90 organizing protein, are required for protein stability and structure. *Protein Science* **15**, 522-532 (2006).

137. Buchner, H.W.·.L.M.·.J. *Reviews of Physiology, Biochemistry and Pharmacology*. **151**, 1–44 (Springer Berlin Heidelberg: Berlin, Heidelberg, 2004).

138. Southworth, D.R. & Agard, D.A. Client-loading conformation of the Hsp90 molecular chaperone revealed in the cryo-EM structure of the human Hsp90:Hop complex. *Molecular cell* **42**, 771-81 (2011).

139. Americo, T. a, Chiarini, L.B. & Linden, R. Signaling induced by hop/STI-1 depends on endocytosis. *Biochemical and biophysical research communications* **358**, 620-5 (2007).

140. Erlich R.B., Kahn S.A., Lima F.R., Muras A.G., *et al.* STI1 promotes glioma proliferation through MAPK and PI3K pathways. *Glia* **55**, 1690-8 (2007).

141. Partridge, J. & Flaherty, P. An in vitro FluoroBlok tumor invasion assay. *Journal of visualized experiments : JoVE* (2009).doi:10.3791/1475

142. Neckers, L. Hsp90 inhibitors as novel cancer chemotherapeutic agents. *Trends in molecular medicine* **8**, S55-61 (2002).

143. Hagemann, T. & Lawrence, T. Inflammation and Cancer. *Business* **512**, 325-332 (2009).

144. Trepel, J., Mollapour, M., Giaccone, G. & Neckers, L. Targeting the dynamic HSP90 complex in cancer. *Nature reviews. Cancer* **10**, 537-49 (2010).

145. Clevers, H. The cancer stem cell: premises, promises and challenges. *Nature medicine* **17**, 313-9 (2011).

146. Agarwal, R. & Kaye, S.B. Ovarian cancer: strategies for overcoming resistance to chemotherapy. *Nature reviews. Cancer* **3**, 502-16 (2003).

147. Aguirre-Ghiso, J.A. Models, mechanisms and clinical evidence for cancer dormancy. *Nat. Rev. Cancer* **7**, 834–846 (2007).

148. Durante F. Nesso fisio-pathologico tra la struttura dei nei materni e la genesi di alcuni tumori maligni. *Arch Memor Observ Chir Pract* **11**, 217–26 (1874).

149. Sell, S. Stem cell origin of cancer and differentiation therapy. *Critical reviews in oncology/hematology* **51**, 1-28 (2004).

150. Bapat, S.A., Mali, A.M., Koppikar, C.B. & Kurrey, N.K. Stem and progenitor-like cells contribute to the aggressive behavior of human epithelial ovarian cancer. *Cancer research* **65**, 3025-9 (2005).

151. Zhang S., Balch C., Chan M.W., Lai H.C., *et al.* Identification and characterization of ovarian cancer-initiating cells from primary human tumors. *Cancer research* **68**, 4311-20 (2008).

152. Szotek P.P., Pieretti-Vanmarcke R., Masiakos P.T., Dinulescu D.M., *et al.* Ovarian cancer side population defines cells with stem cell-like characteristics and Mullerian Inhibiting Substance responsiveness. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 11154-9 (2006).

153. Lichtenauer U.D., Shapiro I., Geiger K., Quinkler M., *et al.* Side population does not define stem cell-like cancer cells in the adrenocortical carcinoma cell line NCI h295R. *Endocrinology* **149**, 1314-22 (2008).

154. Gambelli F., Sasdelli F., Manini I., Gambarana C., *et al.* Identification of cancer stem cells from human glioblastomas: growth and differentiation capabilities and CD133/prominin-1 expression. *Cell biology international* **36**, 29-38 (2012).

155. Pan, Y. & Huang, X. Epithelial ovarian cancer stem cells-a review. *International journal of clinical and experimental medicine* **1**, 260-6 (2008).

156. Karbanová J., Missol-Kolka E., Fonseca A.V., Lorra C., *et al.* The stem cell marker CD133 (Prominin-1) is expressed in various human glandular epithelia. *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society* **56**, 977-93 (2008).

157. Quintana E., Shackleton M., Sabel M.S., Fullen D.R., *et al.* Efficient tumour formation by single human melanoma cells. *Nature* **456**, 593–598 (2008). Shackleton, M., Quintana, E., Fearon, E.R. & Morrison, S.J. Heterogeneity in cancer: cancer stem cells versus clonal evolution. *Cell* **138**, 822–829 (2009).

158. Quintana E., Shackleton M., Foster H.R., Fullen D.R., *et al.* Phenotypic heterogeneity among tumorigenic melanoma cells from patients that is reversible and not hierarchically organized. *Cancer Cell* **18**, 510–523 (2010).

159. Bidlingmaier, S., Zhu, X. & Liu, B. The utility and limitations of glycosylated human CD133 epitopes in defining cancer stem cells. *Journal of molecular medicine (Berlin, Germany)* **86**, 1025-32 (2008).

160. Katoh, K., Asimenos, G. & Toh, H. Multiple alignment of DNA sequences with MAFFT. *Methods in molecular biology (Clifton, N.J.)* **537**, 39-64 (2009).

161. Scadden, D.T. Cancer stem cells refined. *Nature immunology* **5**, 701-3 (2004).

162. http://www.cellbankaustralia.com/estore/productdetail.aspx?productid=245&categoryid=133. Last accessed June 14, 2012.

163. Wang, Y., Sheng, Q., Spillman, M. a, Behbakht, K. & Gu, H. Gab2 regulates the migratory behaviors and E-cadherin expression via activation of the PI3K pathway in ovarian cancer cells. *Oncogene* 1-9 (2011). doi:10.1038/onc.2011.435

164. http://www.celprogen.com/DS/36113-40%20-%20DS.pdf. Last accessed June 14, 2012.

165. Eustace B.K., Sakurai T., Stewart J.K., Yimlamai D., *et al.* Functional proteomic screens reveal an essential extracellular role for hsp90 alpha in cancer cell invasiveness. *Nature cell biology* **6**, 507-14 (2004).

166. Koga F., Kihara K., Neckers L. Inhibition of Cancer Invasion and Metastasis by Targeting the Molecular Chaperone Heat-shock Protein 90. *Anticancer Res* **29**, 797-807 (2009).

167. Banerji U., O'Donnell A., Scurr M., Pacey S., *et al*. Phase I pharmacokinetic and pharmacodynamic study of 17-allylamino, 17-demethoxygeldanamycin in patients with advanced malignancies. *J Clin Oncol* **23**: 4152-4161, (2005).

168. Baba T., Convery P.A., Matsumura N., Whitaker R.S., *et al.* Epigenetic regulation of CD133 and tumorigenicity of CD133+ ovarian cancer cells. *Oncogene* **28**, 209-18 (2009).

169. http://www.celprogen.com/DS/M36113-40%20-%20DS%20-%20Human%20Ovarian%20Cancer%20Stem%20Cell%20Culture%20Serum%20Free%20Media.pdf. Last accessed June 14, 2012.

170. England, R. & Pettersson, M. Pyro Q-CpG™: quantitative analysis of methylation in multiple CpG sites by Pyrosequencing®. *Nature Methods* **2**, (2005).

BIOGRAPHICAL SKETCH

Dimitri Van Simaeys was born in Tielt, Belgium. He earned his Industrial Engineering in chemistry, option biochemistry degree with distinction at the Proviciale Industriële Hogeschool West-Vlaanderen, Kortrijk, Belgium in 2004. He subsequently earned a master's degree at the University of Ghent in 2005, also with distinction in molecular biotechnology. After this, he joined Procter and Gamble, where he was working on global formulations for hand dish products, with a focus on strategic use of raw materials. In 2007, he joined the chemistry graduate program at the University of Florida, where he obtained his PhD in Chemistry in 2012, under the tutelage of Dr. Weihong Tan.