Université du Québec
Institut National de la Recherche Scientifique
Centre Énergie Matériaux Télécommunications
(INRS-EMT)

In Vitro Selections of Mammaglobin B and Mammaglobin A Aptamers for detecting Metastatic Breast Cancer Cells Using Terahertz Chemical Microscopy

Par

Eman Hassan

Une thèse présenté en application partielle
des exigences pour le degré de
Doctorat en philosophie (Ph.D.)
en sciences de l'énergie et des Matériaux

Jury d'évaluation

Président du jury et examinateur interne	Marc A. Gauthier INRS - Énergie, matériaux et télécommunications
Examinateur externe	Juewen Liu Département de chimie, Waterloo Institute for Nanotechnology
Directeur de recherché	Tsuneyuki Ozaki INRS - Énergie, matériaux et télécommunications
Codirecteur de recherché	Maria DeRosa Département de chimie. Université Carleton
Codirecteur de recherche	Bill Willmore Institut de biochimie. Université Carleton

ACKNOWLEDGMENTS

I would like to sincerely acknowledge my supervisor Prof. Ozaki for giving me the chance to do my experimental part of my P.hD. at Carleton University in Ottawa. I want to thank him also for his support and supervision throughout my PhD program.

I want to send very special thanks to my co-supervisor Dr. Maria DeRosa for her great supervision and support through the years I was doing my Ph.D. research at her lab. Moreover, I want to thank Dr. Bill Willmore (my co-supervisor) to giving me a chance to do part of my experimental work at his lab and for his supervision.

I would like to express my deep appreciation to my friends from the DeRosa lab starting from Emily, Nadine, Phepa, Annamaria, Erin, Chris, and McKenzie for all their support and the great friendship that support and help me through my PhD program. I want to thank also Willmore lab members for their support especially Mary and Haiyun.

I would like to thank my undergraduate students James Podrebarac, Usman Khan, Andries Selst, Anisha Ghelani, and Jacqueline Addo that help me in my Ph.D. research.

Many thanks also goes to Dr. Bruce McKay from the Biology department at Carleton University to give me a chance to use his flow cytometry instrument and for his help in the data analysis as well. Also I want to thank Andrew Robinette from the Willmore lab for his help in running the flow cytometry.

Special thanks goes to Dr. Kiwa (and his student Yuki Hanaoka) for doing the TCM experiments at his lab at Okayama University (Japan) and for the data analysis as well.

Finally I would like to express my gratitude to INRS for their financial support during the years of my Ph.D. program.

ABSTRACT

Breast cancer is the most common cancer diagnosed in Canadian women, accounting for 26% of all newly diagnosed cancers in women. Metastatic breast cancer (MBC) is the stage of cancer where the disease has spread to distant parts of the body, further than the axillary lymph nodes. Breast cancer cells that leave the primary tumor to the blood are called circulating tumor cells (CTCs). Studying CTCs allows for better understanding of the metastases process. The detection of CTCs is a challenging task due to the heterogeneous nature of CTCs and lack of biomarkers. There is an urgent need to identify more biomarkers to help in the detection of CTCs from breast cancer.

Mammaglobin B (MGB2) and mammaglobin A (MGB1) proteins are small secretory proteins that are believed to have roles in cancer development, immune system regulation, and androgen binding. MGB2 is found at high levels in many secretions, including those from uterine, prostatic, pulmonary, and lacrimal and salivary glands, whereas MGB1 expression is exclusive to the breast tissue. Since both proteins are overexpressed in most cases of metastatic breast cancer,, MGB2 and MGB1 are very attractive biomarkers that could be used in the detection of breast cancer and breast CTCs.

Aptamers are powerful recognition elements. They are single-stranded ssDNA or RNA oligonucleotides that bind to their targets with high affinity and selectivity. Aptamers are chemically produced via a process called Systematic Evolution of Ligands by EXponential enrichment (SELEX). SELEX is an iterative *in vitro* processwhich enriches target-binding sequencesfrom randomized large ssDNA or RNA libraries. In each round of SELEX, target–binding sequences are eluted from target molecules, amplified, and used as the library for subsequent rounds. This process results in an aptamer with high affinity and specificity for the target.

In this work, a group of aptamers against MGB2 and MGB1 proteins were developed for the first time using a unique type of SELEX called hybrid SELEX, in which the ssDNA library was alternated between two targets that are the recombinant form of both MGB2, and MGB1 proteins and breast cancer cells MCF7 and MDA-MB-415 respectively.

All selected aptamers were then tested for their affinity by determining the dissociation constant (K_d) using flow cytometry. Aptamers were tested for their selectivity by testing their affinities to other normal and cancer cell lines. Aptamers for both targets that had the highest affinity and selectivity were further tested for their binding affinity to their target cancer cell lines in plasma and whole blood lysate environments using flow cytometry. Aptamers for both targets were tested for their binding affinity in blood containing normal blood cells and spiked with different percentages of target breast cancer cells. The binding affinity of the selected aptamers was tested against their target proteins using Electrophoretic Mobility Shift Assay (EMSA). Finally, and for the first time, the selected aptamers against the two targets were tested to detect breast cancer cells using Terahertz Chemical Microscopy (TCM).

Results obtained by this study revealed the successful selection of a group of aptamers against MGB2 and MGB1 proteins. Among all the selected aptamers, mammaglobin B1 (MAMB1) against MGB2 (MCF7 cell line), and mammaglobin A2 (MAMA2) against MGB1 (MDA-MB-415 cell line) showed the highest affinities to their targets (as indicated by their K_d values). Moreover, these aptamers showed minimal binding to normal and other cancer cell lines, indicating the high selectivity to their target breast cancer cells. Plasma and whole blood lysate results showed that MAMB1 and MAMA2 aptamers bound specifically to MCF7 and MDA-MB-415 cells respectively. MAMB1 showed higher affinity to MCF7 than MAMA2 showed to MDA-MB-415 cells. Moreover, the study with spiked breast cancer cells showed an increase in the fluorescence intensity with increasing percentages of spiked cells for both aptamers, indicating the great potential to use both aptamers in the detection of breast CTCs.

TCM results showed that terahertz (THz) amplitude was higher when MAMB1 was interacting with MCF7 cells and MAMA2 with MDA-MB-415, compared to the control on the surface of the sensing plate. Moreover, TCM could detect the interaction of a low number of cells (10 cells) binding to their aptamers. MAMB1 aptamer was more selective than MAMA2 aptamer, as shown by TCM.

Overall results obtained by flow cytometry and TCM were very promising and indicate the possibility of using MAMB1 and MAMA2 aptamers in the detection of breast cancer and breast CTCs.

Table of Contents

List of Tables

List of Figures

List of Schemes

List of Abbreviations

ACS	American Cancer Society
AML	Acute Myeloid Leukemia
APTES	(3-Aminopropyl) triethoxysilane
ATCC	American Type Culture Collection
BPE	Bovine Pituitary Extract
BRCA1	Breast Cancer 1
BRCA2	Breast Cancer 2
BSA	Bovine Serum Albumin
CCS	Canadian Cancer Society
CD45-APC	CD45-Allophycocyan
CE-SELEX	Capillary Electrophoresis SELEX
CEUS	Contrast-Enhanced Ultrasound
CKs	Cytokeratins
CK-PE	Cytokeratin Phycoerythrin
CPG	Controlled Pore Glass
CRP	C-Reactive Protein
CSCs	Cancer Stem Cells
CTCs	Circulating Tumor Cells
DAPI	4’,6-Diamidino-2-Phenylindole
DCIS	Ductal Carcinoma *In Situ*
DI	Deionized Water
DMEM	Dulbecco's modified Eagle's Medium,
Ds	double-stranded
ECM	Extracellular Matrix
EDTA	Ethylenediaminetetraacetic acid
EGFR	Epidermal Growth Factor Receptor
EGTA	Ethylene Glycol Tetraacetic Acid
EMEM	Eagle's Minimum Essential Medium
EpCAM	Epithelial Cell Adhesion Molecule
ER	Estrogen-Receptor
EMSA	Electrophoretic Mobility Shift Assay
EMT	Epithelial to Mesenchymal Transition
FACS	Fluorescence-Activated Cell Sorter
FBS	Fetal Bovine Serum
FCS	Fetal Calf Serum
FDA	Food and Drug Administration,
FISH	Fluorescence *In Situ* Hybridization
FSC	Forward Scatter
GA	Gentamicin-Amphotericin
GFP	Green Fluorescence Protein

GNR	Gold Nano Rods
GST	Glutathione S transferase
HBSS	Hank's Balanced Salt Solution
HCAEC	Human Primary Coronary Artery Endothelial Cells
hEGF	human Epidermal Growth Factor
HER-2	Human Epidermal Growth Factor Receptor 2
hMAM	human Mammaglobin
HTS	High Throughput Sequencing,
IDC-NST	Invasive Ductal Carcinoma of No Special Type
IGF-1	Insulin-like Growth Factor-1
IHC	Immunohistochemistry
K_d	Dissociation Constant
LCIS	Lobular Carcinoma *In Situ*
LOD	Limit of Detection
LOQ	Limit of Quantification
MBC	Metastatic Breast Cancer
MEGM	Mammary Epithelial Cell Growth Medium
MET	Mesenchymal to Epithelial Transition
MFI	Mean of Fluorescence Intensity
µFFE	Micro- Free Flow Electrophoresis
MGB1	Mammaglobin A
MGB2	Mammaglobin B
MRI	Magnetic Resonance Imaging
MS	Mass Spectrometry
NAT	Ni- Nitrilotriacetic
NGS	Next- Generation Sequencing
OPN	Osteopontin
PAGE	Polyacrylamide Gel Electrophoresis
PBMC	Primary Peripheral Blood Mononuclear Cells
PCR	Polymerase Chain Reaction
PN	Periostin
PNDA-3	Benzyl-d(U)TP-Modified DNA Aptamers-3
PR	Progesterone-Receptor
PVDF	Polyvinylidene Difluoride
RAN	Random
RBCs	Red Blood Cells
RT	Room Temperature
RT-PCR	Real Time PCR
SELEX	Systematic Evolution of Ligands by EXponential enrichment
SDS	Sodium Dodecyl Sulfate
SLS	Sodium Lauroyl Sarcosinate
SPR	Surface Plasmon Resonance
Ss	single-stranded
SSC	Side Scatter
TBS-T	Tris Buffer Saline-Tween
TCM	Terahertz Chemical Microscopy

TECS-SELEX	Target Expressed on Cell Surface-SELEX
THz	Terahertz
VEGF	Vascular Endothelial Growth Factor

CHAPTER 1

Introduction

Some parts of this chapter were reviewed in: *Hassan EM, Willmore WG, and DeRosa, MC Aptamers: Promising Tools for the Detection of Circulating Tumor Cells. Nucleic Acid Therapeutics. 2016*

1.1 Breast Cancer

Cancer is a complex disease that originates as a result of multiple genomic mutations leading to a disruption of normal cellular homeostasis. Mutant cells gain unlimited proliferative capacity, the ability to invade surrounding tissues and metastasize, insensitivity to growth signals, sustained angiogenesis, and resistance to apoptosis [1,2]. The main factor behind malignant transformation is genetic mutations by inheritance or acquisition over time due to loss of the DNA repair system [1,2].

Breast cancer is the most common cancer diagnosed in women, with an estimated 1.67 million new cases diagnosed worldwide in 2012 [3]. The incidence of breast cancer is highest in developed countries with a rate of 80 per 100,000 in the European Union and 92 per 100,000 in North America [4]. Breast cancer is the leading incident cancer and second-leading (following lung cancer) cause of cancer death among Canadian women [5]. In 2015, an estimated 25,000 Canadian women were diagnosed with breast cancer. Breast cancer remains to be the most commonly diagnosed cancer in Canadian women, accounting for 26% of all newly diagnosed cancers in women [5].

One of the most important risk factors for developing breast cancer is having a family history of breast cancer [6]. Women with inherited mutations in the Breast Cancer 1 (BRCA1) and Breast Cancer 2 (BRCA2) genes have a lifetime risk of breast cancer of 50%–85% [7,8].

Other factors related to the reproductive system such as early menarche, late menopause, and having children have been demonstrated by American Cancer Society (ACS) and Canadian Cancer Society (CCS) to increase the risk of having breast cancer. The increase in risk may be due to a longer lifetime exposure to the hormones estrogen and progesterone [9,10]. Women who were exposed or treated with radiation in the chest, neck, and arm pit area are more likely to develop breast cancer in their lifetime [11,12]. Lifestyle-related factors have been shown to increase the risk of having breast cancer, these factors include: drinking alcohol, being overweight or obese, smoking [13-15], using birth control [16,17], and hormone replacement therapy [18,19].

1.2 Classification of breast cancer

Breast cancer is a disease in which malignant cells are found either lining the milk ducts or in the milk-producing lobules of the breast. Lobules and ducts are formed from three lineages of cells in two germ layers: the myoepithelial layer is common to both structures and forms the basal layer, while ductal epithelial cells line the ducts, and alveolar epithelial cells synthesize the milk within the lobules [20]. Almost all breast cancers (>95%) are epithelial in origin, and are classified as adenocarcinomas [21]. Cancers of the breast are classified according to their histological type (growth pattern) as *in situ* (cancers which are contained within the breast lobule or duct) or invasive breast cancers (cancers which are not contained by the lobular or ductal walls and have invaded the breast stroma) [21]. Ductal Carcinoma *In Situ* (DCIS) or Lobular Carcinoma *In Situ* (LCIS) are the two most common types of *in situ* breast cancer. Both are non-invasive cancerous lesions of the breast and are classified as stage 0. They rarely produce symptoms or breast lumps [22]. Invasive Ductal Carcinoma of No Special Type (IDC-NST) is the most common carcinoma of the breast accounting for 40% to 75% of invasive breast cancers [23]. The cancerous cells invade the surrounding normal tissue and replace them. Special types of breast cancer account for up to 25% of all breast cancers and the latest edition of the World Health Organization classification recognizes the presence of at least 17 different histological special types of breast cancer [24]. Some of the cancers, including tubular, invasive cribriform, medullary, and infiltrating lobular breast cancers have confirmed more favorable prognoses compared with IDC-NST [22,25]. These cancers typically demonstrate rates of 10-year survival that generally exceed 80% [26].

On the other hand, other special types of cancer, such as mixed or solid lobular, mixed ductal and lobular, or ductal/NST have shown considerably lower rates of 10-year survival [26].

Breast cancers are further subtyped based on the stage of development, which includes tumor size, where it first started in the body, and the locations to which it has travelled. TNM staging is determined by the collective score on three components, including: tumor size (T), involvement of the regional lymph nodes (armpits, neck, and the chest) (N) and presence of distant metastasis (M) [26]. Breast cancers diagnosed at a later stage (large size and metastasis to lymph nodes and other parts of the body) are associated with poorer prognoses. The four main stages of breast cancers are: stage 0 which describe the *in situ* types of breast cancers (DCIS and LCIS), stage 1-3: breast cancer spread within the breast or lymph nodes, stage 4: breast cancers spread to different parts of the body [26].

Histologic grade is also strongly linked with breast cancer survival and recurrence. Grading assesses the morphological features of a breast cancer by measuring: a) the degree of differentiation (which reflects how closely they resemble normal breast epithelial cells), b) the tubule formation and nuclear pleomorphism parameters, and c) the proliferative activity by the mitotic index parameter [27]. The Elston-Ellis modification to the Scarff-Bloom-Richardson grading system (Nottingham grading system) is the most commonly used histologic grading system, in which each of the parameters is assigned a score from 1 to 3 [27]. Tumors are considered well-differentiated if the overall score is 3 to 5 (histologic grade I), moderately-differentiated if the score was 6 to 7 (histologic grade II) or poorly-differentiated if the score was 8 to 9 (histologic grade III). It has been reported that the survival of breast cancer patients was significantly decreased for those with a histologic grade of II vs. I, and III vs. I [28]. Also, rates of disease-free survival were also less for breast cancer patients with a higher histological grade. However this association only reached statistical significance for grade III vs. I [27]. A study by van Diest *et al*. demonstrated a poorer prognosis for breast cancer patients with a higher mitotic index compared to those with a lower index [29].

Breast cancers are further classified by their expression levels of steroid hormone receptors (Estrogen-Receptors (ER) and Progesterone-Receptors (PR)) measured with Immunohistochemistry (IHC), as well as expression of Human Epidermal Growth Factor Receptor 2/neu (HER-2), measured with IHC and Fluorescence *In Situ* Hybridization (FISH) [30,31]. ER is expressed in 60-80% of invasive breast tumors [22] and over half of ER+ tumors are also PR+ [32]. Less than 10% of breast cancers are ER-/PR+ [32]. The HER-2 oncogene is overexpressed in an estimated 20% of invasive breast tumors [22]. Compared with ER+/PR+/HER-2- tumors, ER-/PR-/HER-2+ and ER-/PR-/HER-2- (triple negative) tumors are more likely to be diagnosed at a later stage and at higher histologic grade [33]. Moreover, triple negative tumors have low rates of overall survival [33,34]. The expression profile of ER, PR and HER-2 receptors direct which molecular pathway affects a tumor, and thus are used to determine the type of endocrine therapy. ER+ tumors depend mainly on ER levels for their growth; therefore they can be treated with drugs to reduce either the effect of ER (e.g. tamoxifen) or the actual level of estrogen (e.g. aromatase inhibitors), and have a response rate of 60 to 70% [35,36]. ER- tumors however have a response rate of less than 10% to hormone therapy [37]. HER-2+ breast cancer patients are treated with anti-HER-2 therapies such as trastuzumab (Herceptin) and pertuzumab which are used in combination with chemotherapy and had been reported to improve the response in HER2+ breast cancers [38,39]. Triple negative breast cancer lacks the expression of the three receptors which make it more difficult to treat because most chemotherapies target at least one of the three receptors. Due to this lack of receptors, triple-negative cancers often involve combination therapies. No specific drug therapy has been identified for triple negative tumors, therefore chemotherapy remains the mainstay in treatment of these cancers [38,40,41].

1.3 Metastatic breast cancer

Metastatic Breast Cancer (MBC) (also known as advanced or stage IV breast cancer) is a stage of breast cancer where the disease has spread from the axillary lymph nodes to distant parts of the body (e.g. bones, liver, lung, or brain) [42]. In the US and Canada, approximately 5 and 10% of all breast cancer patients have distant metastases at initial diagnosis, respectively [43,44]. Approximately 30% of women who are first diagnosed with early stage breast cancer will develop MBC later in their lives [43,44]. Metastasis is the main cause of death for the majority of breast cancer patients [45]. In Canada, the five year survival rate of metastatic breast cancer patients is only 20%, as indicated by CCS [5].

In order to metastasize, breast cancer cells must complete a cascade of events from invasion of local tissues, entry into the blood and colony establishment at distance sites, resulting in the development of a secondary tumor (Figure 1.1). During the initial stages, some of the cancer cells gain invasive properties and leave the primary tumor to intravasate into the blood stream or lymphatic system. Subsequently, the cells circulate through the body in the blood, adhere at the secondary site and extravasate from the circulatory system into the secondary tissue. Cancer cells must then maintain growth in the secondary site to form micro-metastases, which then form macro-metastases (along with the correspondingly associated angiogenesis) at the secondary site [46-48]. The process of metastasis is complex and the localization in a secondary tumor is non-random, usually depending upon the primary tumor characteristics [49] and certain genes that drive the metastatic process [50]. In breast cancer, more than 70 gene signatures have been associated with distant metastasis [51]. Bone metastasis was found in up to 70% of advanced breast cancer patients with osteolytic sites being the most dominant type of metastasis [52].

1.3.1 Circulating tumor cells

It is well known that cancer cells are heterogeneous, and thus not all cancer cells are capable of initiating metastasis [53]. Circulating Tumor Cells (CTCs) are cancer cells that detach from a primary tumor, circulate within the blood and then invade other distant organs in the body to form a secondary tumor [53]. CTCs have stem cell-like properties, and thus are capable of initiating the metastatic process (Figure 1.1). In 1994, John Dick and his colleagues were the first ones to isolate and characterize Cancer Stem Cells (CSCs) in Acute Myeloid Leukemia (AML) [54]. They discovered that only a certain population of cells (CD34+CD38−) had the ability to initiate and sustain the leukemia. After isolating CD34+CD38−cells and CD34-CD38-cells and injecting them into NOD/SCID mice, it was observed that only the CD34+CD38− cells developed into leukemia cells [55]. This study was the first spark for the CSC research. Subsequently, there is now evidence supporting the existence of CSC populations in brain, prostate, colon, and pancreatic tumors, as well as in melanomas [56-58]. In breast cancer, CSCs were identified based on high CD44 and low CD24 expression ($CD44^{high}/CD24^{low}$) on the cell surface. These cells showed CSC properties (self-renewal and the initiation and sustainment of breast cancer growth) when injected into immunocompromised mice [59]. Based on their ability to give rise to tumors, a small subset of CTCs from metastatic breast cancer patients have been reported to have CSC characteristics, and thus all CSCs may be considered CTCs [60-63].

1.3.2 The metastatic cascade: a closer look

Epithelial to Mesenchymal Transition (EMT) was first identified in embryonic development and is the conversion of epithelial cells to a mesenchymal phenotype [64]. CTCs have the ability to initiate the metastatic cascade by converting differentiated epithelial cancer cells into dedifferentiated cells that have more mesenchymal characteristics. EMT is considered a crucial event in the metastatic process (Figure 1.1) [65]. The most important events that take place in EMT are: 1) down-regulation of E-cadherin, which is considered the most dominant cell adhesion protein on the surface of the cancer cell, 2) the decline of the expression of epithelial markers, such as cytokeratin [93], 3) increased expression of vimentin, a mesenchymal marker, and 4) change of the expression of integrin, an extracellular matrix molecule [66,67].

During EMT, several signaling pathways control the expression of various factors on the molecular level. The growth factor transforming growth factor-β (TGF-β) has been shown to induce reversible EMT, along with the Wnt pathway proteins (in particular β-catenin), Notch, and Hedgehog signaling pathways. These pathways often act in a sequential manner to induce EMT [68-71]. Furthermore, transcription factors such as Snail and NF-κB have been shown to induce apoptotic resistance to tumor cells undergoing EMT [72,73]. The EMT cells break through the basement membrane, blood vessels, or interstitial spaces and start entering the circulation in a process called intravasation (Figure 1.1), allowing them to initiate distant metastasis. It is worth mentioning that intravasation could take place early in cancer development, even before the diagnosis of the primary tumor [74]. In breast cancer, it has been reported that tumor cells start to disseminate out of the primary tumor early in the disease (initial diagnosis), but it is not sufficient for metastatic outgrowth [75].

A solid tumor of 1 cm size (approximately 10^9 cancer cells) could shed one million cancer cells into the circulation per day [76]. These cancer cells will be carried either via the blood (CTCs) or via the lymph. CTCs do not spend long circulating in the bloodstream. Indeed, most of them are collected by the first capillary bed they encounter within minutes of entering the circulation due their large size and inability to pass through capillaries [74]. It is well known that the majority of secondary tumors are formed in organs that are not connected to lymph nodes (e.g. brain, liver, lung, and bone) [77].

The phenotypic transition of CTCs is reversible and it has been reported that once CTCs reach their destination, they will convert back into an epithelial phenotype in a process called Mesenchymal to Epithelial Transition (MET). This transition is crucial for tumor growth at the secondary site (Figure 1.1) [78]. Little is known about the role that MET plays in the formation of a secondary tumor, but recently researchers have begun to examine MET as one of many potential therapeutic targets in the prevention of metastases [79,80].

Metastasis is a complex process and it is unlikely that all cancer cells successfully complete all the steps necessary to form clinically relevant metastases. Therefore, metastasis is described as a highly inefficient process. Few cancer cells pass through all the metastasis steps and finally make it to their destination. From a million cancer cells that escape the primary tumor, only a small number (~0.1%) remain alive in the circulation after 24 h, and an even smaller number of cells (0.01%) could start metastasis [53,81]. Indeed, the majority of CTCs (85%) get washed from the circulations afew hours of intravasation

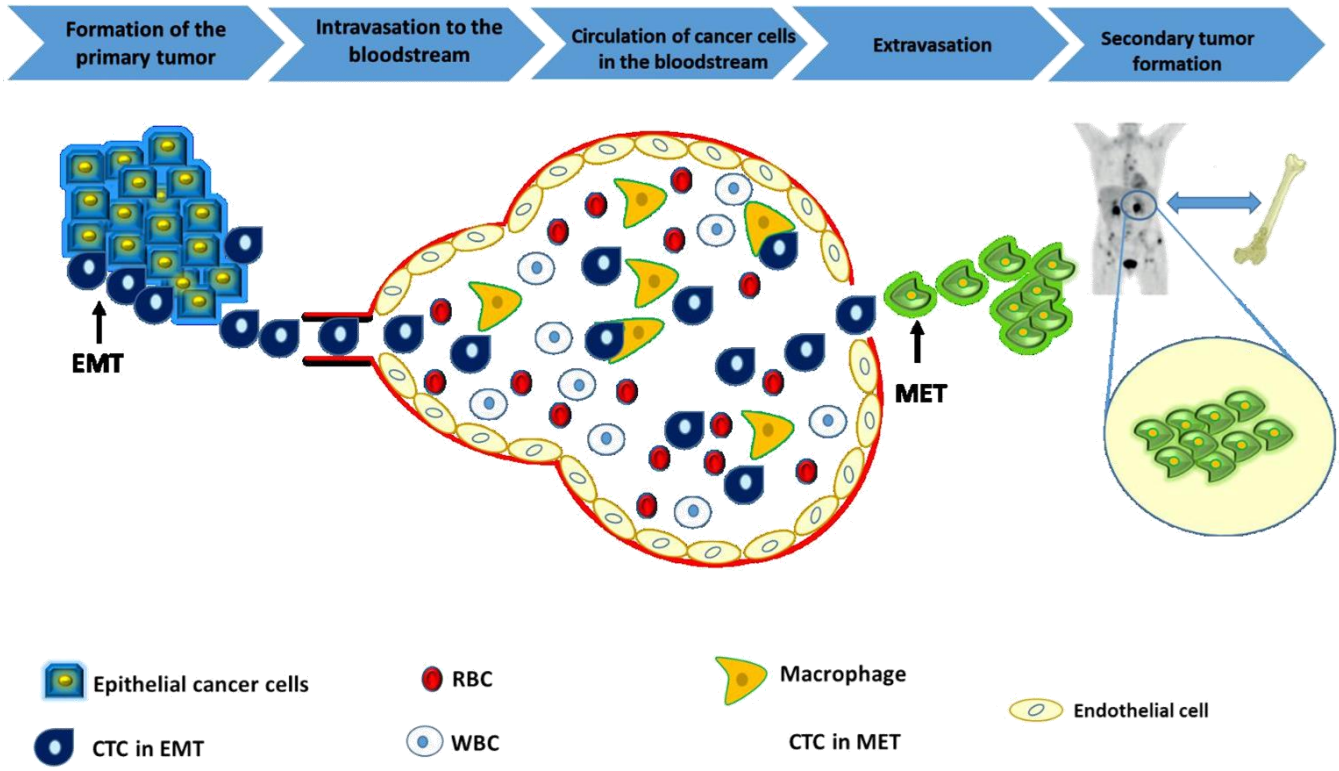


Figure 1.1: Metastasis cascade: In the first stage, which includes the formation of the primary tumor, cancer cells start to invade a local epithelial tissue and divide there. Some of them (CSCs) have the ability to initiate metastasis and undergo EMT to be considered as CTCs 2: CTCs, which have the properties of CSCs, enter the bloodstream and circulate with the blood components, some of which are eliminated by macrophage cells (part of the immune system). 3: The rest of CTCs will exit the blood in a process called extravasation. 4: After they exit the blood, CTCs undergo MET and start to invade distant organs (in the case of breast cancer, it is most often the bone) forming secondary tumors. (Reprinted with permission from Hassan *et al*. (2016).

CTCs are very rare, with a frequency of typically 1 per 10^{6-7} leukocytes [83]. CTCs exhibit highly pleomorphic characteristics, and therefore it is difficult to distinguish them from other blood cells [84,85]. Breast cancer CTCs have a diameter of 13.1 µm [86] vs 10 µm for blood leukocytes [87]. To be considered as CTCs, cancer cells must have a nucleus, visible cytoplasm, expression of cytokeratin and absence of CD45 expression [88]. Detection of CTCs in breast cancer patients is commonly associated with an increased risk of metastases and poor prognosis [89]. However, variation of the detection rate of CTCs and its correlation with prognosis has been reported [90]. Although, to date, the use of CTCs as a routine prognostic clinical tool is not applied [91]. Many clinical studies have demonstrated the importance of CTCs in monitoring disease progression or treatment failure. For instance, MBC patients who had a reduction in their CTC count after initial treatment showed identical prognosis to CTC-negative patients [92]. Moreover, patients who had higher initial counts of CTCs and reduced CTC counts following treatment showed better prognoses than those whose CTC count did not change [93]. Other studies failed to observe whether a change in chemotherapy would improve the overall survival rate in breast cancer patients with increasing CTC levels [94].

In summary, there is solid evidence that a sub-population of cells that have stem-cell like properties (CTCs) can contribute to tumor initiation. The functional role of such cells in tumor metastasis such as breast metastasis and treatment remains poorly understood.

1.4 Current methods of the detection and quantification of CTCs

The study of CTCs allows for a better understanding of the metastasis process. In comparison to traditional tissue biopsy, the isolation of CTCs (which has been called “liquid biopsy”) offers many advantages: collection of peripheral blood is easy to perform, the technique is relatively inexpensive with less risk to patients, and it is easily repeatable over time [95,96]. The idea of monitoring CTCs over time (real time) is very attractive to researchers as it allows for the tracking of the changes which occur in cancer cells during metastasis. Isolation of CTCs also allows for better understanding of the metastatic progression, as well as the reasons for the resistance of CTCs to current therapies [95,96].

Many methods have been developed to detect CTCs in the blood of cancer patients. Some of these methods are based on the physical properties of CTCs, including density and size, as well as immune and cell-surface electrical properties [97]. Other methods detect CTCs as a result of the interaction of a certain biological probe (*e.g.* antibody) to a protein (bio-marker) on the surface of CTCs in an instrument such as the CellSearch® system.

It has been demonstrated that tumor cells exhibit epithelial properties and thus express on their surface molecules of epithelial origin [98]. Epithelial Cell Adhesion Molecule (EpCAM) is expressed in tumors of epithelial origin and allows distinguishing CTCs from blood cells [99]. EpCAM is a transmembrane glycoprotein first identified as a tumor antigen, and is highly expressed in epithelial cancers in comparison to normal epithelial tissue [100]. EpCAM has a role in cell signaling, migration, proliferation, and differentiation [101-103]. It has been reported that EpCAM plays a role in tumorgenesis and metastasis, so it can be used as a prognostic marker and a target for therapy in patients with epithelial cancers [104].

The only system that has been approved by the U.S. Food and Drug Administration (FDA) and uses EpCAM to detect CTCs is the CellSearch system. CellSearch detects and quantifies CTCs in human blood from epithelial carcinomas (breast, prostate or colorectal cancers) using fluorescence microscopy and immunology-based techniques [105-107]. CTCs are enriched and separated from a 7.5 ml blood sample in the CellTracks AutoPrep system by antibody-mediated ferrofluid-based magnetic separation in which monoclonal antibodies against EpCAM are conjugated to iron magnetic nanoparticles to specifically target CTCs. Anti-Cytokeratin Phycoerythrin (CK-PE), a second monoclonal antibody, is used to stain CTCs, followed by the DNA stain 4',6-Diamidino-2-Phenylindole (DAPI). Blood leukocytes are then separated from epithelial cells by using anti-CD45-Allophycocyan (CD45-APC) [108]. The CellTracks Magnet analyzes the enriched CTCs using the CellTracks Analyzer II, a semi-automated fluorescence-based microscopy system that obtains images using a 10X objective lens, and filters corresponding to the different DAPI, PE, and APC stains [109]. A gallery of CTC images is then created. Qualitative analysis software selects CTCs based on the differential staining method mentioned earlier.

EpCAM+, CK+, and CD45- cells are counted as CTCs [108]. Quantification of the actual number of CTCs, as well as assessment of the morphology of CTCs can be accomplished using the CellSearch system. It is difficult to set criteria of what counts as a CTC due to the diversity in their viability or apoptotic stage, which makes their population very heterogeneous [110].

The percentage of CTCs not detected by the CellSearch system is unknown. During the EMT process, the expression of epithelial markers decreases and therefore the expression of EpCAM is not the same amongst all CTCs. For example, some studies showed a change in the level of EpCAM expression in mammary carcinomas [111]. It has been reported that in patients with MBC, CTCs tend not to express EpCAM on their surface, but instead exhibit stem cell features [112]. Moreover some breast cancer subtypes lack the expression of Cytokeratins (CKs) 8/18/19 [113]. A study investigating EpCAM expression in 50 breast cancer cell lines showed that 20% of cell lines had low levels of EpCAM expression [114]. These cells were reported to have a basal-like phenotype, high level of vimentin (a mesenchymal marker) and low levels of epithelial markers CKs 8/18/19 and E-cadherin, indicating a mesenchymal phenotype. Therefore, the CellSearch system would not be efficient in detecting CTCs in those patients. Studies have shown the successful clinical usage of the CellSearch system to monitor the prognosis by counting CTCs in metastatic breast cancer patients, treated for long periods.. The results of these studies confirmed the prognostic value of CTCs, but they failed to evaluate the effectiveness of chemotherapy in those patients [115,116].

In conclusion, the detection of CTCs specifically in MBC, is a challenge due to a) their low abundance in the blood stream and b) the lack of reliable markers present to recognize these cells. Therefore, new markers are still required for the detection and isolation of different sub-types of tumor cells in MBC with high specificity and sensitivity. Moreover, increasing the sensitivity and specificity by investigating new ligands other than antibodies is required and this represents the central topic of this thesis.

1.5 Aptamers as powerful ligands in the detection of CTCs in metastatic breast cancer

1.5.1 Aptamers

Aptamers were described for the first time in 1990 [117]. They are non-natural, short (~ 15–100 nucleotides in length) single-stranded DNA or RNA oligonucleotides that bind to their targets with high affinity and selectivity, with a unique three- dimensional interaction [118].

Aptamers can replace their antibody counterparts due to their similar recognition properties [119]. Aptamers are chemically synthesized at a relatively low cost, without the use of animals, which are required for antibody production [119]. Because of their chemical and thermal stability, aptamers can be used under a wide range of experimental conditions; unlike antibodies which are restricted to their host conditions, rapidly losing their tertiary structure at higher temperatures. Antibodies have batch-to-batch variations, whereas aptamers do not. In addition, the modification (labeling for proteins) can affect the structure and the function of antibodies while aptamers, when synthesized, allow for the introduction of functional groups on specific sites without the disruption of their structure or function [120]. As mentioned earlier, aptamers can be selected against a wide range of targets with different sizes and structures, whereas the affinity of antibodies is limited to those molecules that can trigger immune responses [121]. The first RNA sequence that could discriminate between different targets was the one that binds to theophylline, with 10^4 times higher binding affinity to theophylline than to caffeine, which differ by only a methyl group [122].

Recently, slow off-rate modified aptamers (SOMAmers) were established by Brody *et al*. These new aptamers were capable of binding to their protein targets with very high affinity (in the very low nanomolar range and with picomolar dissociation constants) and have been tested in clinical samples [123]. Table 1.1 summarizes the different properties that aptamers have versus antibodies [118, 120].

Table 1.1: Characteristics of aptamers versus antibodies

Properties	Aptamers	Antibodies
Stability	Stable at wide range of temperatures (reversible denaturation)	Stable only at certain temperatures
Target Potential	Wide range of targets from ions to bacteria or whole mammalian cells, also non-immunogenic or toxic targets	The targets are limited to mostly proteins/peptides
Immunogenicity	Not recognized by the immune system due to their small size	The immune system reacts against them due to their size
Production	Produced via *in vitro* process	Produced *in vivo* with the need of animals or cell culture
Shelf life	Long shelf life due to the stability of aptamers (when dried they can stay at room temperature)	Short shelf life due to in-stability, since they are limited to their host conditions
Ability to be Modified	Easy to chemically modify at relatively low cost	Difficult to chemically modify and expensive
Nuclease Degradation	Subject to nuclease degradation but can be overcome with modification	Nucleases cannot degrade antibodies, since they made of amino acids

1.5.2 Systematic Evolution of Ligands by EXponential enrichment

Aptamers are produced in vitro in a process called Systematic Evolution of Ligands by EXponential enrichment (SELEX). Aptamers can be selected for a wide range of targets, starting with ions and ending with whole mammalian cells [124-127]. Therefore, SELEX can be performed against different targets, with different conditions according to the researcher's needs. Many improvements to SELEX have taken place since the original SELEX was established in 1990 to select aptamers with higher affinity and selectivity to their targets [128].

In general, SELEX has multiple steps (Figure 1.2). The first step is the selection step, where the target, which is either immobilized on a solid support or is free in solution (according to the final application of the selected aptamers), is incubated with a random library of DNA or RNA (depends on whether the application needed DNA or RNA aptamers). This library is flanked with two conserved primer binding sites to be used later for amplification. The number of random sequences in the library can be 10^{13}-10^{16} (Figure 1.2). In thesecond step, partitioning of the nonbinding sequences from those that bound the target takes place. Nonbinding sequences are washed away, and the sequences which bound the target are eluted in the third step. The amplification of the eluted sequences takes place using the Polymerase Chain Reaction (PCR), using the conserved regions of the randomized library for primer-binding (Figure 1.2). The selection process can be repeated several times (rounds) until the percentage recovery of the DNA (monitored during the selection) reaches a plateau. Usually a counter selection of a similar target takes place after or before the main selection to ensure the specificity of all the selected aptamers for the target of interest [129]. Throughout the selection process, it is recommended to increase the stringency to improve binder enrichment in the library thus generating higher affinity aptamers. This increased stringency can be achieved by decreasing the target concentration to encourage competition, changing the ionic strength of the binding and washing buffer or/ and increasing the number of washing steps for the unbound sequences [129].

1.5.3 Cell- SELEX

The idea of using complex targets for the generation of specific aptamers was first suggested in 1998 [130]. This idea was confirmed when an aptamer against red blood cell membranes was identified in 2000 [131].

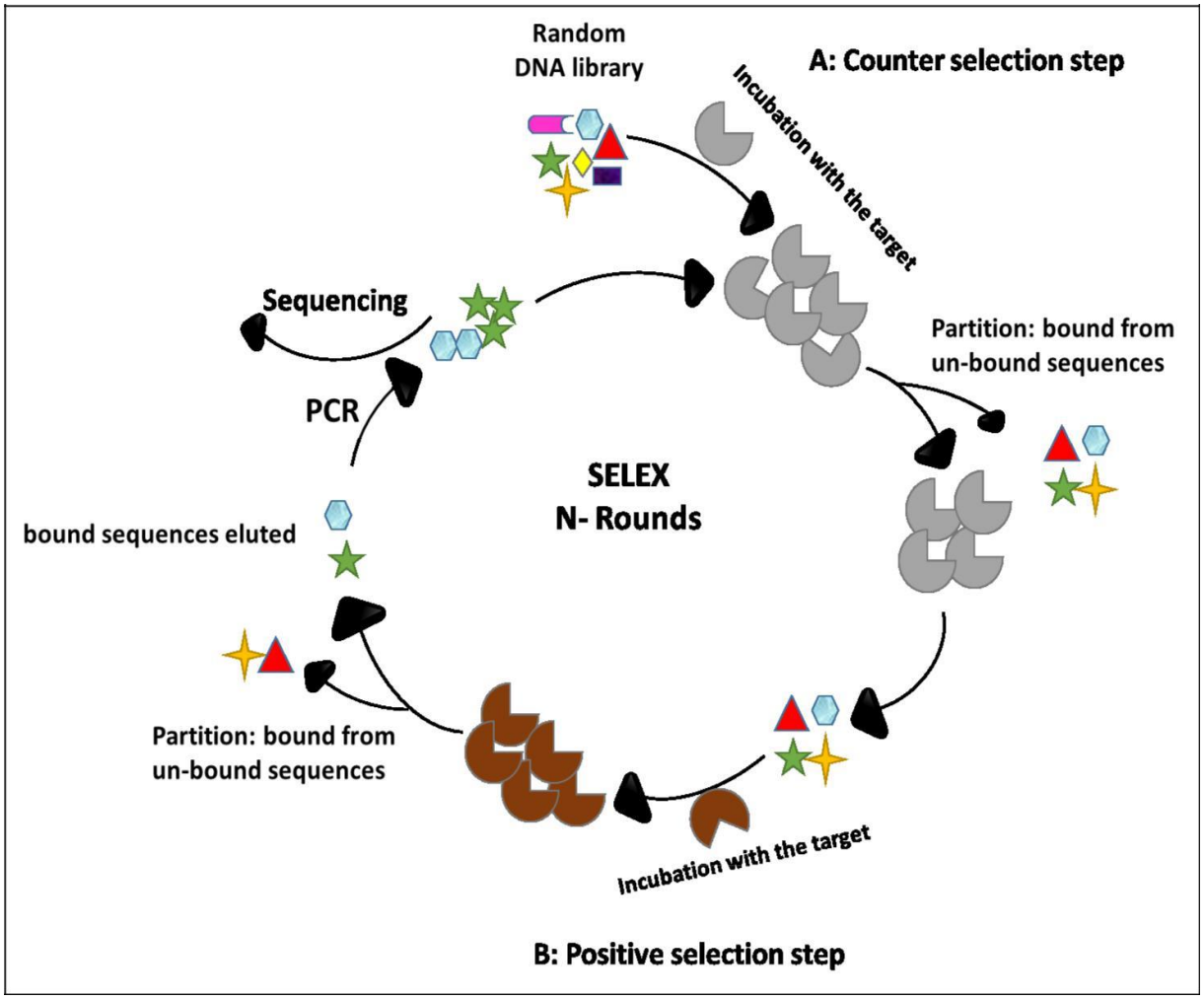


Figure 1.2: General Scheme of SELEX: SELEX starts with counter selection step (A) in which a random DNA library is incubated with a target similar to the target of interest. The unbound sequences (from washing step) are then incubated with the target of interest for the positive selection (B). Unbound sequences are washed off and binding sequences are eluted off the target. The eluted sequences are amplified by PCR, and the rounds are then repeated.

At that time, the application of aptamers for cancer detection was limited due to the absence of aptamers able to bind target cancer cells, or target proteins on the cell

surface [132]. Cell- SELEX technology offers many advantages over other types of SELEX: 1. A prior knowledge of the protein receptors on the surface is unnecessary since the generated aptamers will recognize the molecular features only on the abnormal cells. Moreover, the aptamers bind to the native state of the targets. In this manner it is important to mention that a counter selection is an important step in cell-SELEX. In the case of cancer cell-SELEX, the ideal counter selection should be normal cells from the same type of tissue *i.e.* breast cancer cells vs healthy normal cells surrounding the cancer cells [132]. 2. Multiple aptamers can be selected for many different targets and thus provide accurate diagnosis of the disease (*e.g.* cancer) [132]. 3. The generation of new aptamers to unknown targets on the surface of the cells help to identify new biomarkers. Cell-SELEX for cancer may serve to improve the understanding of many complicated processes that occur during the development of cancer including metastasis [132].

The cell-SELEX procedure is similar to that of conventional SELEX, but differs in the partitioning step, in which washing or centrifugation is used to separate the bound from unbound sequences. In cell-SELEX, heating is usually used for isolating aptamer-target cell complexes [133]. Despite the short history of cell-SELEX, it has become a promising platform for generating large numbers of aptamers against a wide range of cell lines, especially cancer cells [132]. Although considerable progress has been made, cell-SELEX is still facing many challenges. One challenge is the complexity of the cell-SELEX process, as selection for whole cells takes many rounds (up to 35) to identify the aptamer candidates, which is time consuming and labor intensive. It is hard to see a significant enrichment for the starting library [134]. Another challenge for cell-SELEX is the diversity of the ligands on the cell surface.Althoughthis diversity increases the chance of identifying new bio-markers, it affects the progression of the SELEX process as aptamers prefer to bind the more abundant targets (common targets on all cells types) versus the low abundance ones. If the more abundant targets were not the desired ones, this might affect the enrichment of the aptamers [132]. The abundance of such targets (*e.g.* membrane proteins) may change during the cell growth and passages, therefore cell culture maintenance is important for cell-SELEX [132]. Target identification after cell-SELEX is necessary in order to use these aptamers in further applications, especially the ones selected against cancer cell lines [133].

This step is a serious limitation due to the difficulties in the separation and purification of aptamer targets, especially the membrane protein targets [133]. For this reason, a limited number of aptamer targets have been identified [135-138].

In order to overcome these challenges, many modified cell-SELEX platforms have been reported [139-142]. One of the reasons behind the variations in enrichment in cell-SELEX is the contamination from dead cells. Dead cells tend to adsorb non-specifically to single-stranded nucleic acids and thus reduce the enrichment factor of the positive selection step [143]. Fluorescence-Activated Cell Sorting (FACS) has been reported to eliminate the dead cells from a mixture of cells [144]. Raddatz *et al.* used FACS for the positive selection step to eliminate dead cells and successfully identified DNA aptamers against Burkitt lymphoma cells [144]. FACS separates distinct subpopulations of cells from other subpopulations within composite cell mixtures. For example, live and dead cells could be separated on the basis of their different light-scattering characteristics, or their ability to convert cell-permeable dyes into a fluorescent form enzymatically [145]. It is hard to maintain the activity of cells over long periods of time, and thus different cells may be used for selection in different rounds in cell SELEX. In cancer cells for example, variations between cells could affect the SELEX process, as cancer cells are heterogeneous. Microfluidic systems have been integrated into cell-SELEX to shorten the selection process and make it more efficient [146]. An automated on chip cell-SELEX has been demonstrated. In this system, the entire cell-SELEX is performed on a single chip. In order to achieve this, three modules, including a microfluidic control module, a magnetic bead-based aptamer extraction module, and a temperature control module for cooling reagents and rapid nucleic acid amplification have been used [147]. It has been reported that on chip cell–SELEX has been successfully used in the identification of aptamers against ovarian cancer; these aptamers binds to target cells specifically with high affinity [148].

Targeting of specific cell surface proteins in their native forms using cell-SELEX helps to take the application of cell-SELEX even further, especially in cancer research where identifying new biomarkers is a great benefit to improve the diagnosis and treatment of cancer patients. This part of cell SELEX will be discussed in detail in chapter two as a core topic.

1.6 High throughput sequencing and aptamer characterization

Integration of High Throughput Sequencing (HTS) or Next- Generation Sequencing (NGS) methods in SELEX is of great benefit, due to the many advantages that it provides compared with traditional cloning/sequencing methods. Using traditional methods, sequences from enriched nucleic acid libraries are cloned into a plasmid and a few hundred individual clones, at most, would be sequenced to identify individual high affinity aptamers [149,150]. HTS allows for sequencing of millions of DNA strands in parallel, yielding substantially more throughput and reducing the need for the cloning method [150]. Cloning is typically applied only after the last selection round, where as HTS can be applied to the nucleic acid libraries after each selection round, thus avoiding the need for a high number of selection rounds, reducing time, PCR bias and artefacts [150]. Comprehensive characterization of obtained aptamers, identification of their functional and rare motifs, and a comparison of functional motifs in each oligonucleotide population along with quantification of their abundance can be done by combining HTS to bio-informatics tools and powerful software [151]. HTS has contributed noticeably to *in vitro* selection techniques as in many other research fields, including personalized medicine [152]. HTS was first combined with SELEX in 2002 in a SELEX-SAGE method in which CTF/NFI transcription factor (TF) ligands were identified in genomic DNA [153]. Many improvements have been made on HTS in the following years to identify many targets [154-156]. Recently HTS was combined with whole cell SELEX and applied to identify DNA aptamers targeting the HER2 receptor in breast cancer patients [157]. In this work, they applied HTS after each round of selection, reducing the number of rounds to only five and increasing the specificity and affinity of the identified aptamers [157]. Many software tools are now available to analyze HTS data, including AptaCluster [158] and FASTAptamer [159].

The Dissociation Constant (K_d) describes how tightly a ligand binds to its target under equilibrium conditions. For aptamers, the K_d is the concentration of aptamer at which the concentration of target with aptamer bound equals the concentration of target with no aptamer bound [160,161]. K_d reflects the affinity of the aptamers to their targets and has a molar unit when the binding stoichiometry is 1:1. The smaller the K_d, the higher the affinity [160,161]. The K_d of aptamer-target complexes typically range from the low micromolar to high picomolar levels

[162]. Several methods have been established to measure the K_d of aptamers [163]. The methods range from simple, low cost methods such as dialysis and filter binding assays, to complicated ones such as Fluorescence anisotropy and Surface Plasmon Resonance (SPR) [164,165]. Figure 1.3 shows the most popular methods used for aptamer affinity determination. These methods vary in their sensitivities and the sample volume needed for analysis, as well as the analysis time required. Moreover, the post-selection labelling or immobilization of the aptamers in these methods could affect the value of K_d [161]. Thus a particular aptamer could have many values of K_d considering the factors above when applied to different methods of K_d determination. Therefore, it is recommended to use the method that most closely simulates the circumstances in which the aptamer is planned to be used [161].

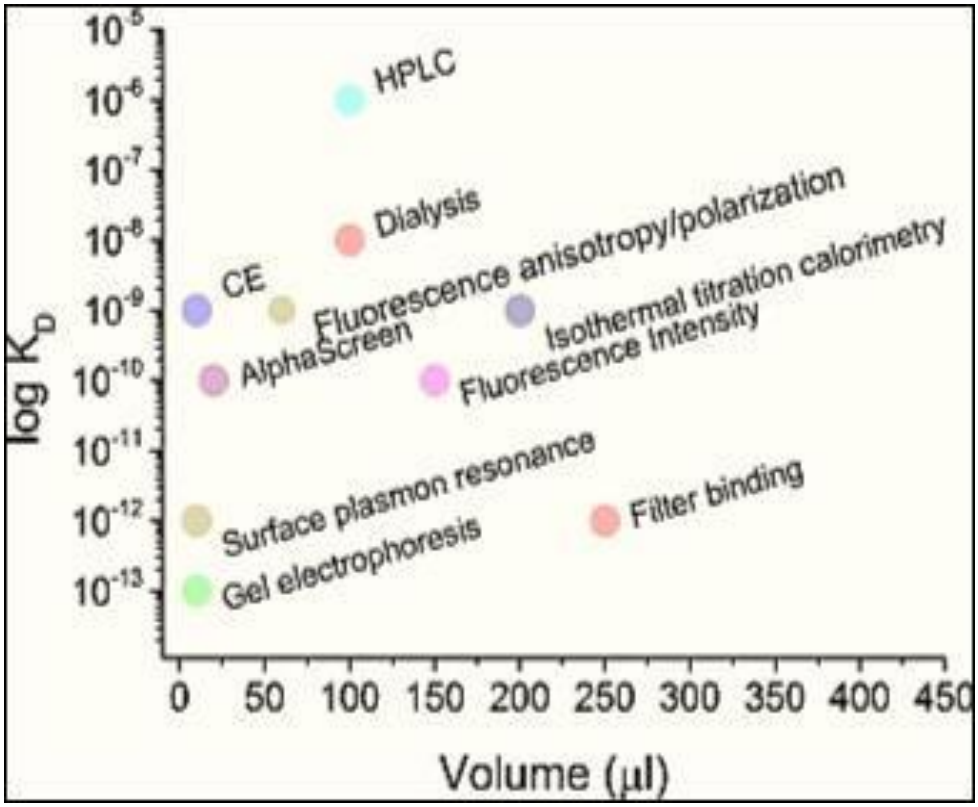


Figure 1.3: The different methods used to measure the dissociation constant (K_d) of aptamer–protein complexes. The graph shows the K_d values with the respective methods as a function of the volume required for the analysis. Reprinted with permission from Szeitner *et al*.[161].

1.7 Aptamers generated for the detection of metastatic breast cancer and breast CTCs

As mentioned earlier, aptamers can rival antibodies in their specificity and affinity for the target [166], therefore, they are good candidate probes for capturing CTCs in the bloodstream [167] The overexpression of surface biomarkers is often related to tumorigenesis [167], therefore cell surface biomarkers are attractive targets for cancer diagnosis including CTCs detection [166,167]. As mentioned earlier, EpCAM is the most abundant protein reported to be overexpressed in patients with epithelial carcinomas [99]. Moreover, it has been reported that EpCAM is a biomarker for CTCs in these patients as it can be used to distinguish between cancer cells in the blood and other blood cells [99]. Although it has been reported that EpCAM expression changes during EMT process (as mentioned earlier in this chapter) and might be absent in some breast cancer patients [111,112,114], many studies have been developed using aptamers against EpCAM to detect CTCs in a mixture of cancer cells [168-170]. In a study by Song Y *et al*. a DNA aptamer (SYL3C) against EpCAM protein was successfully selected and used to capture CTCs [171]. SYL3C has a high affinity to the metastatic form of breast cancer cells (MDA-MB-231) *in vitro* [171]. Moreover, SYL3C was able to capture the target cancer cells from a mixture of cancer cells with up to 63% cancer cell capture efficiency and about 80% purity [171].

Human Periostin (PN) is a protein produced and secreted by fibroblasts as a component of the Extracellular Matrix (ECM) [172]. PN is a very interesting protein, as it has a role in the interaction with multiple cell-surface receptors, particularly integrins [173,174] or with the PI3-K/Akt pathway which when activated, stimulates cell survival, angiogenesis, invasion, metastasis, and more importantly, epithelial-mesenchymal transition of carcinoma cells [175,176]. In breast cancer, the overexpression of PN is correlated with the poorest outcomes [172]. Characterizing PN functional roles in breast cancer progression could improve the outcome of breast cancer. Benzyl-d(U)TP-Modified DNA Aptamers-3 (PNDA-3) were selected against PN [177] PNDA-3 inhibits periostin-mediated breast cancer cell invasion *in vitro* and metastasis *in vivo* [177]. PNDA-3 aptamer could be used to target metastatic breast cancer cells [177].

Another secretory protein which is similar in function to PN and can induce breast cancer metastasis in almost the same way is Osteopontin (OPN) [178]. OPN is a secreted phosphoprotein cell attachment protein and cytokine that signals through two cell adhesion molecules: αvβ3-integrin and CD44 [178]. It has been reported that OPN increased cellular migratory and invasive behavior, and increased metastasis in many cancers including breast cancer [178]. Like PN, OPN is overexpressed in breast cancer patients and is correlated with poor outcomes [179]. An RNA aptamer (OPN-R3) against OPN was developed in 2009 [179]. This study showed that OPN-R3 could inhibit breast cancer metastasis *in vivo* [179]. Since OPN is a secretory protein and based on the high affinity of OPN-R3 aptamer to OPN, an electrochemical aptasensor for the detection of human OPN was developed in 2015 by Meirinho *et al*. [180]. The detection limit of the aptasensor was 3.7 ± 0.6 nM in standard solution. This limit corresponds with OPN levels reported in the literature for patients with metastatic breast cancer [180]. This study demonstrated a non-invasive method to detect metastasis in breast cancer patients [180].

Regarding the use of aptamers as imaging probes in metastatic breast cancer, a DNA aptamer (LXL-1) was selected against metastatic MDA-MB-231 cells [181]. The truncated aptamer sequence LXL-1-A showed a very high affinity and specificity when used as a probe in lymph node tissues from metastatic breast cancer patients, with 76% detection rate [181]. This suggests the potential use of LXL-1-A aptamer as an imaging probe for the detection of breast cancer metastasis [181].

The expression of biomarkers on the surface of CTCs is heterogeneous, therefore a single CTC marker is insufficient to capture all CTCs from the blood. A study by Nellore *et al*. established aptamer-modified porous graphene oxide membranes for the capture and identification of multiple types of CTCs spiked in rabbit blood [182]. Three aptamers S6, A9, and YJ-1 for breast, prostate, and colon cancer cells, respectively, were immobilized to porous graphene oxide membranes. The three aptamers were capable of capturing SKBR3 breast cancer cells, LNCaP prostate cancer cells, and SW-948 colon cancer cells selectively and simultaneously when 100 cells were spiked in one milliliter of rabbit blood, with a capture efficiency as high as 98% [182].

In summary, aptamers generated for the detection of metastatic breast cancer and breast CTCs are few in the literature, therefore more studies involving aptamers in breast CTCs detection are needed to improve the outcome of patients with metastatic breast cancer. Moreover, more clinical data is needed for those already published to be able to use these aptamer in clinics.

1.8 Flow cytometry as a tool to detect circulating tumor cells from breast cancer

As mentioned earlier, CTCs are heterogeneous in nature, making their detection a challenging task, therefore methods that are more complex are needed to detect CTCs in an enriched environment. Flow cytometry is an analytical technique that became more common in biomedical fields, due to its many advantages, including: the high statistical value of the method, analysis of a very large number of events, and the short analysis time [183]. Flow cytometry has many functions including cell counting and sorting, biomarker detection, and cell cycle and protein studies [183,184]. Flow cytometry recognizes cells according to their size and level of internal complexity (granularity) [183,184]. The forward Scatter (FSC) parameter allows for the determinination of the size of a cell (using known size standard such as beads) by measuring the amount of the laser beam that passes around the cell. The Side Scatter (SSC) parameter allows for the determination of the granularity level of a cell by measuring the amount of laser light that bounces off particulates inside of the cell (Figure 1.4) [185]. Because of its ability to differentiate cells according to their size and complexity, flow cytometry has been used in the detection of cancer cells in the blood [186-188].

Most flow cytometry methods isolate and count breast CTCS by using monoclonal antibodies directed against different antigens on the surface of CTCs. The most popular surface antigens are CKs and EpCAM. These two biomarkers, as already discussed, can be down regulated in malignant cells during EMT and this leads to the generation of false negative results. Since flow cytometry has the advantage of using multiple labelling (multi-color analysis) of different antigens on the cell surface, this limitation can overcome by using additional antibodies against CD45, making it possible to detect leukocytes [189]. Multiple parameters can be measured

simultaneously (i.e. DNA content, cell size, cell viability, and intra- and extracellular markers) using this method, therefore high specificity is obtained. Despite this, some studies have reported low sensitivity of flow cytometric techniques compared to RT-PCR [190].

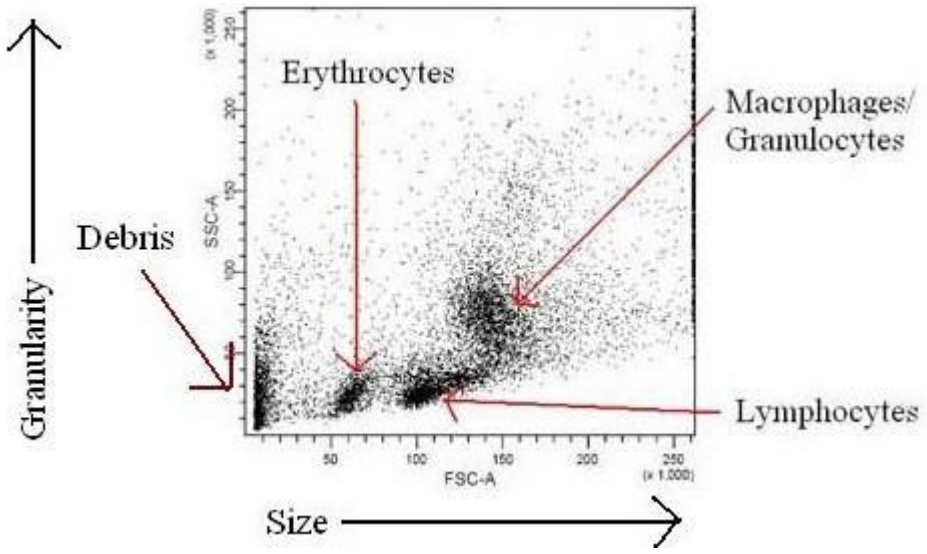


Figure 1.4: A dot plot obtained after running a mixture of blood cells is suspended through a flow cytometer: the plot shows different populations of cells according to their size (the x-axis (FSC)) and granularity (the y-axis (SSC)). The cells to the far left of the plot represent the debris (dead cells). Going further right on the FSC axis the cells are bigger and more complex. Examples of these cells are macrophages that have granules. Reprinted with the permission of Sandra Hope (Ph.D) at Brigham Young University [185].

Flow cytometry has been used as a tool to detect CTCs from breast cancer in many studies [187,188, 191, 192]. In a study by Wang L *et al*. CTCs from breast cancer patients were detected using flow cytometry in which the expression of CK19 was quantified [191]. In this study, the expression of CK19 was quantified (using anti-CK19 antibody) for patients in different stages of breast cancer (primary and metastatic stages) as well as patients who underwent chemotherapy for three months. The results obtained by this study showed that the number of CTCS detected increased gradually along with the disease stages. Moreover, the expression of CK19 declined for those who went for chemotherapy as obtained by flow cytometry [191].

The limit of the detection of the assay was one cancer cell among 10^4 human white blood cells [191]. In a recent study by Warawdekar *et al*. flow cytometry was used in combination with RT-PCR to detect CTCs in spiked breast cancer cells in lymphocytes as well as blood samples from breast cancer patients [192]. The results revealed a sensitivity of one breast cancer cell detected in 10^5 lymphocytes [192]. In breast cancer patients, flow cytometry showed a range of 1-85 CTCs per 10 mL of blood [192]. Monoclonal antibodies against EpCAM, CKs as well as CD45 were used in this study [192].

Song *et al*. has reported the detection of breast CTCs using flow cytometry and aptamers as the probe [193]. In this study, an aptamer (SYL3C) against EpCAM was developed and tested to capture CTCs derived from breast, colorectal, and gastric cancers in mixed media as well as in blood [193]. The capture efficiency was 63% with 80% purity [193].

In this context, aptamers could be used as probes for flow cytometry, and offer certain advantages over antibodies. First, aptamers can be chemically synthesized and functionalized with a wide variety of fluorophores and modifications. Second, unlike antibodies, aptamers are more stable which makes them less susceptible to degradation by heat and pH shifts. These two main advantages allow easy handling and storage for both manufacturers and operators [194].

Newly developed aptamers (against new breast cancer biomarkers) for the detection of metastatic breast cancer using flow cytometry were used in this work and will be discussed in details in Chapter 3.

1.9 Terahertz radiation

Terahertz (THz) radiation, also known as THz waves, lies between the microwave and infrared regions of the electromagnetic spectrum (Figure 1.5) [195]. THz has unique characteristics with potential applications for research in physics, chemistry, biology, materials science and medicine (Figure 1.4) [195]. Since THz radiation has a long wavelength (300μm) it can penetrate a wide variety of non-conducting materials, e.g. clothing, paper, cardboard, wood, and plastic. For this reason, THz radiation is used for security purposes in airports and for weapon

detection. In most cases, the THz radiation used in such applications is called passive or incoherent [196]. THz radiation has relatively low photon energy, which is insufficient to cause chemical damage to tissues and DNA. Some frequencies of THz radiation can penetrate several millimeters of tissue with low water content (e.g., fatty tissue) [196]. THz can be used to differentiate between different materials or even isomers, since the energy of rotational and vibrational transitions of molecules lies in the THz region and intermolecular vibrations such as hydrogen bonds show different spectral features in the THz region [196]. THz waves can provide a better contrast for soft tissues than X-rays, due to the fact that THz radiation is very sensitive to polar substances, such as water. The type of THz used in imaging is called active [196].

1.9.1 Biomedical applications of THz radiation

1.9.1.1 THz applications and DNA

The four nucleobases (adenine (A), cytosine (C), guanine (G), and thymine (T)) have different absorption coefficients in the THz range due to the fact that the low-frequency molecular motions that originate from the hydrogen bonds of DNA base pairs and non-bonded interactions are sensitive to the base composition and conformational state of nucleic acids [197]. Furthermore, some studies demonstrated the difference in absorption between single-stranded (ss) and double-stranded (ds) DNA [198]. This unique feature of DNA allows for label-free detection of mutations of unknown DNA molecules and this is critical in many biomedical applications, especially those involved in cancer research [199]. On the other hand, it has been reported that intense THz pulses caused DNA damage and increased the concentration of protein that helps the body to fight against cancer (tumor suppressor proteins) [200]. This has been suggested in a study by Titova *et al.* in which a lab-grown human skin model was exposed to 0.1 –1.0 μJ of THz pulse for 10 mins [200]. Intense THz increased the amount of phosphorylated H2AX; a protein that is formed when there is a double strand break in the DNA, suggesting that intense THz could cause double strand breaks, and at the same time induced tumor suppressor proteins, indicating that DNA damage repair mechanisms are rapidly activated [200].

1.9.1.2 THz and biological imaging

As a result of its unique features, THz research is growing exponentially [201]. One of the most interesting applications is using THz in imaging biological tissues [195]. As mentioned earlier, THz radiation has low photon energy (1 THz= 4.1 meV), and thus it can be applied for *in vivo* real-time diagnosis without causing ionization damage to the tissue or DNA [195,196]. Furthermore, THz has a longer wavelength than infrared and visible light, and scattering losses in biological tissues are negligible [195,196]. However, there are limitations for using THz in imaging biological tissues, which include absorption of THz radiation to water content, which limit the penetration into the human body in areas where water content is high [195,196]. The development of suitable THz sources is another challenge for using THz in imaging [195,196]. THz uses femtosecond lasers that are bulky systems and expensive to use [196]. Due to its long wavelength, THz imaging has limited spatial resolution [196]. To overcome the challenges of using THz in imaging, nanotechnology has been integrated to THz. This will be explained in the following section [196].

1.9.1.3 THz and cancer imaging

Since THz radiation is strongly absorbed by water, normal tissues and cancer tissues can be differentiated, because their water contents are different [196]. It has been reported that cancer tissues show higher absorption coefficients and refractive indexes than normal tissues [195,196]. However, in this regard, THz radiation with low frequencies (0.5 THz) should be used, to give the highest contrast between normal cells and the cancerous cells, as a result, the resolution of the image will be affected [196]. Gold nanorods (GNR) were integrated to THz imaging to improve image contrast [202].

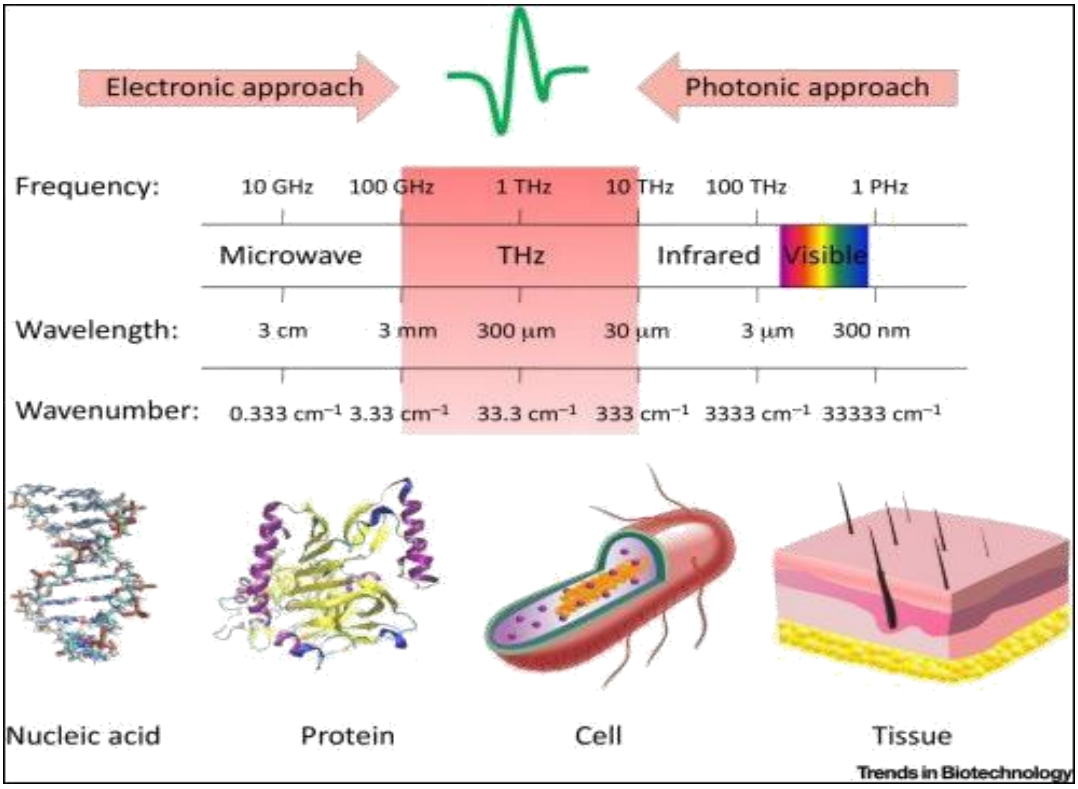


Figure 1.5: THz spectrum and the potential applications of THz radiation. THz radiation lies between the microwave and infrared regions of the electromagnetic spectrum. THz frequencies start from 100 GHz to 10 THz on the electromagnetic spectrum which equals to 30 mm–3000 mm wavelength or 3.33 cm^{-1}–333 cm^{-1} wavenumber. THz has several applications including imaging biological tissues and proteins, and nucleic acids applications. Reprinted with permission from Yang *et al*. [195].

GNR were made so that they can only be absorbed by cells with high concentrations of Epidermal Growth Factor Receptor (EGFR), which is overexpressed only in cancer cells [203]. In order for theses nano-rods to work, they would need to be stimulated by IR illumination. The IR heats up the nanorods with a surface polaritons causing changing of THz properties of water inside the cancer cells, which in turn affects the reflectivity of THz [203].

THz was used in the detection of early breast cancer. In a study by Chen *et al.*, *in vivo* THz transmission imaging was performed on a subcutaneous xenograft mouse model [204]. The results showed that THz can monitor the growth of breast tumor in vivo by distinguishing cancer from the surrounding fatty tissue [204]. This study revealed the potential of using a THz imaging system in the early detection of breast cancer [204].

In another interesting study, Yngvesson *et al.* showed the possibility to use a THz imaging system to determine the margins of malignant tissues in breast cancer by differentiating between breast carcinoma, normal fibro-glandular tissue, and adipose tissue [205]. Usually surgeons tend to remove the healthy tissue surrounding the malignant ones, and after at least one day examination of the excised tissue, some of the patients may undergo a second surgery. Using THz imaging in defining the margins of the malignant tissues is of great benefit for the patients, since it decreases patient injury, emotional stress and health care [205]. Recently, THz was used in the early detection of gastric cancer *in vitro* [206]. In this study cancerous tissues were successfully differentiated based on their higher THz reflection intensities compared with normal mucosa [206].

In summary, THz radiation has been used in many fields due to its low energy and non-ionizing characters. Imaging using THz emerges as a novel tool in the biological medicine field. However, methods to overcome the disadvantages of THz imaging and improve the image quality of Terahertz Chemical Microscopy (TCM) have been established, which will be discussed in details in Chapter 5.

1.10 Thesis objectives

The main objectives of the thesis are first, the development of aptamers against mammaglobin B and mammaglobin A (breast cancer biomarkers), and their characterization, which includes testing for their affinity and selectivity to their targets. Second, the selected aptamers are used to detect metastatic breast cancer cells in biological media (blood and plasma) using flow cytometry. Third, the selected aptamers are integrated into a new detection platform called Terahertz Chemical Microscopy (TCM) to detect metastatic breast cancer cells.

The following is the breakdown of the main objectives.

Objective 1: Screening of novel aptamers against breast cancer biomarkers (mammaglobin B and mammaglobin A)

1. Screening new aptamers by applying hybrid SELEX (cross-over) for the recombinant form of the targets and breast cancer cell lines, along with applying the correct counter selections

2. Using high throughput sequencing to reveal the aptamer sequences and their families

3. Using specialized software and bioinformatics tools to analyze the sequences and choose the best candidates

4. Testing the binding of the aptamer candidates to all cell lines used in the selection

Objective 2: Characterization of the aptamer candidates and determination of the specificity for the highest affinity aptamers

1. Determine the dissociation constant (K_d) for the aptamer candidates

2. Determine the specificity of the aptamers with the highest affinity to the target cell line by testing the binding of these aptamers to other cancer and normal cell lines using flow cytometry and fluorescence microscopy

Objective 3: Testing the potential of the highest affinity aptamers to detect cancer cells in blood and plasma samples and in cell mixtures (whole blood lysate spiked with cancer cells)

Objective 4: Verify the targets of the aptamers on the cell surface and in solution

1. Applying a pull-down assay to concentrate and determine the aptamer's target
2. Using transfection as a tool to identify the targets of the aptamers
3. Testing the binding of the selected aptamers to their recombinant proteins in solution

Objective 5: Testing the binding of the highest affinity aptamers to their respective cell lines using Terahertz Chemical Microscopy (TCM) technology.

References:

1. Hanahan D and RA Weinberg. The hallmarks of cancer. Cell. 2000, 100:57–70.
2. Hanahan D and RA Weinberg. Hallmarks of cancer: the next generation. Cell. 2011, 144:646–674.
3. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]. International Agency for Research on Cancer. 2013, Available on: http://globocan.iarc.fr, accessed on 10/06/2016.
4. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 2015, 136(5): E359–E386.
5. Canadian Cancer Society's Advisory Committee on Cancer Statistics [CCS], 2015.
6. Collaborative Group on Hormonal Factors in Breast Cancer. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. Lancet. 2001, 358 (9291):1389-1399.
7. King MC, Marks JH, Mandell JB, New York Breast Cancer Study Group. Breast and ovarian cancer risks due to inherited mutations in BRCA1 and BRCA2. Science. 2003, 302 (5645): 643–646.
8. Lindor NM, McMaster ML, Lindor CJ, Greene MH; National Cancer Institute, Division of Cancer Prevention, Community Oncology and Prevention Trials Research Group. Concise handbook of familial cancer susceptibility syndromes—second edition. J Natl Cancer Inst Monogr. 2008, 1–93.
9. American Cancer Society, 2016. http://www.cancer.org
10. Canadian Cancer Society, 2016. http://www.cancer.ca
11. John EM, Phipps AI, Knight JA, Milne RL, Dite GS, Hopper JL, Andrulis IL, Southey M, Giles GG, West DW, Whittemore AS. Medical radiation exposure and breast cancer risk: findings from the Breast Cancer Family Registry. Int J Cancer. 2007, 121(2):386-394.
12. De Bruin ML, Sparidans J, van't Veer MB, Noordijk EM, Louwman MW, *et al*. Breast cancer risk in female survivors of Hodgkin's lymphoma: lower risk after smaller radiation volumes. J Clin Oncol. 2009, 27(26):4239-4246.
13. Harvie M, Howell A, Evans DG. Can diet and lifestyle prevent breast cancer: what is the evidence? Am Soc Clin Oncol Educ Book. 2015, e66-73.
14. Key TJ, Allen NE, Spencer EA, Travis RC. Nutrition and breast cancer. Breast. 2003, 12(6):412-416.
15. Johnson KC, Miller AB, Collishaw NE, Palmer JR, Hammond SK, Salmon AG, *et al*. Active smoking and secondhand smoke increase breast cancer risk: the report of the Canadian Expert Panel on Tobacco Smoke and Breast Cancer Risk (2009). Tob Control. 2011, 20(1):e2.

16. Ursin G, Henderson BE, Haile RW, Pike MC, Zhou N, Diep A, *et al*. Does oral contraceptive use increase the risk of breast cancer in women with BRCA1/BRCA2 mutations more than in other women? Cancer Res. 1997, 57(17):3678-3681.
17. Urban M, Banks E, Egger S, Canfell K, O'Connell D, Beral V, *et al*. Injectable and oral contraceptive use and cancers of the breast, cervix, ovary, and endometrium in black South African women: case-control study. PLoS Med. 2012, 9(3):e1001182.
18. Eden JA. Menopausal status, adipose tissue, and breast cancer risk: impact of estrogen replacement therapy. Horm Mol Biol Clin Investig. 2013, 14(2):57-63.
19. Beral V; Million Women Study Collaborators. Breast cancer and hormone-replacement therapy in the Million Women Study. Lancet. 2003, 362(9382):419-427.
20. Kalirai H, Clarke RB. Human breast epithelial stem cells and their regulation. J Pathol. 2006, 208(1):7-16.
21. Colditz GA, Baer HJ, Tamimi RM. Breast cancer. In: Schottenfeld D, Fraumeni JF (eds), Cancer epidemiology and prevention, 3rd ed. Cary, NC: Oxford University Press, 2006.
22. Vanio H, Bianchini F. IARC Handbooks of Cancer Prevention, Volume 7: Breast Cancer Screening. Lyon, France: IARC Press, 2002.
23. B. Weigelt, J.S. Reis-Filho. Histological and molecular types of breast cancer: is there a unifying taxonomy? Nat. Rev. Clin Oncol. 2009, 6(12): 718–730
24. Ellis IO, Schnitt SJ, Sastre-Garau X, Bussolati G, Tavassoli Fa, Eusebi V, J.L. Peterse, *et al*. Invasive breast carcinoma. In: Tavassoli FA, Devilee P (eds). WHO Classification of Tumours Pathology and Genetics of Tumours of the Breast and Female Genital Organs. Lyon, 2003.
25. Pereira H1, Pinder SE, Sibbering DM, Galea MH, Elston CW, Blamey RW, Robertson JF, Ellis IO. Pathological prognostic factors in breast cancer. IV: Should you be a type or a grader? A comparative study of two histological prognostic features in operable breast carcinoma. Histopathology. 1995, 27(3):219-226.
26. Edge S, Byrd DR, Compton CC, Fritz AG, Greene FL, Trotti A (eds). AJCC Cancer Staging Manual, 7th edn. New York, NY: Springer, 2010.
27. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow up. Histopathology. 1991, 19(5): 403-410.
28. Rakha EA, El-Sayed ME, Lee AH, Elston CW, Grainge MJ, Hodi Z, Blamey RW, Ellis IO. Prognostic significance of Nottingham histologic grade in invasive breast carcinoma. J Clin Oncol. 2008, 26(19):3153-3158
29. Van Diest PJ, van der Wall E, Baak JP. Prognostic value of proliferation in invasive breast cancer: a review. J Clin Pathol. 2004, 57(7):675-681.
30. Colditz GA, Baer HJ, Tamimi RM. Breast cancer. In: Schottenfeld D, Fraumeni JF (eds), Cancer epidemiology and prevention, 3rd ed. Cary, NC: Oxford University Press, 2006.
31. Sims AH, Clarke RB, Howell A, Howell SJ. The cellular origins of breast cancer subtypes. In: Pasqualini JR (ed). Breast cancer: Prognosis, treatment, and prevention, 2nd ed. Informa Healthcare, New York, 2008.

32. Ghayad S, Cohen PA. Steroid receptors and associated transcriptional cofactors in predicting the response to endocrine therapy. In: Walker RA, Thompson AM (eds). Prognostic and predictive factors in breast cancer, 2nd ed. London, UK: Informa UK Ltd., 2008.
33. O'Brien KM, Cole SR, Tse CK, Perou CM, Carey LA, Foulkes WD, Dressler LG, Geradts J, Millikan RC. Intrinsic breast tumor subtypes, race, and long-term survival in the Carolina Breast Cancer Study. Clin Cancer Res. 2010, 16(24):6100-6110.
34. Bauer KR, Brown M, Cress RD, Parise CA, Caggiano V. Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype: a population-based study from the California cancer Registry. Cancer. 2007, 109(9):1721-8.
35. De Marchi T, Foekens JA, Umar A, Martens JW. Endocrine therapy resistance in estrogen receptor (ER)-positive breast cancer. Drug Discov Today. 2016, pii: S1359-6446(16)30163-5.
36. Early Breast Cancer Trialists' Collaborative Group (EBCTCG), Dowsett M, Forbes JF, Bradley R, Ingle J, Aihara T, Bliss J, Boccardo F, Coates A, Coombes RC, *et al*. Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. Lancet. 2015, 386(10001):1341-1352.
37. Richard G. Margolese,, MD, Gabriel N. Hortobagyi,, MD, and Thomas A. Buchholz,, MD. 6th ed Management of Metastatic Breast Cancer. Kufe DW, Pollock RE, Weichselbaum RR, *et al*., editors. Hamilton (ON): BC Decker, 2003.
38. Kelly CM and Buzdar AU. Using multiple targeted therapies in oncology: considerations for use, and progress to date in breast cancer. Drugs. 2013, 73(6):505-515.
39. Nitta H, Kelly BD, Allred C, Jewell S, Banks P, Dennis E, Grogan TM. The assessment of HER2 status in breast cancer: the past, the present, and the future. Pathol Int. 2016, 66(6):313-24.
40. Zeichner SB, Terawaki H, Gogineni. A Review of Systemic Treatment in Metastatic Triple-Negative Breast Cancer. Breast Cancer (Auckl). 2016, 10:25-36.
41. Jiao Q, Wu A, Shao G, Peng H, Wang M, Ji S, Liu P, Zhang J. The latest progress in research on triple negative breast cancer (TNBC): risk factors, possible therapeutic targets and prognostic markers. J Thorac Dis. 2014 6(9):1329-1335.
42. Lu J, Steeg PS, Price JE, Krishnamurthy S, Mani SA, Reuben J, Cristofanilli M, Dontu G, Bidaut L, *et al*. Breast cancer metastasis: challenges and opportunities. Cancer Res. 2009, 69(12):4951-4953.
43. SEER cancer statistics review, 1975–2013. National Cancer Institute, Bethesda, MD. (Released April 15, 2016), available on http://seer.cancer.gov/csr/1975_2013/ accessed on 11/06/2016.
44. Canadian Breast Cancer Foundation (CBCF), available on http://www.cbcf.org/central/AboutBreastCancerMain/MetastaticBreastCancer accessed on 11/06/2016

45. Redig AJ, McAllister SS. Breast cancer as a systemic disease: a view of metastasis. J Intern Med. 2013, 274(2):113-126.
46. Chambers AF, Groom AC, MacDonald IC. Dissemination and growth of cancer cells in metastatic sites. Nat Rev Cancer. 2002, 2(8):563-572.
47. Langley RR, Fidler IJ. Tumor cell-organ microenvironment interactions in the pathogenesis of cancer metastasis. Endocr Rev. 2007, 28(3):297-321.
48. Pantel K, Brakenhoff RH. Dissecting the metastatic cascade. Nat Rev Cancer. 2004, 4(6):448-456.
49. Nguyen DX, Bos PD, Massagué J. Metastasis: from dissemination to organ-specific colonization. Nat Rev Cancer. 2009, 9(4):274-784.
50. Minn AJ, Gupta GP, Siegel PM, Bos PD, Shu W, Giri DD, Viale A, Olshen AB, Gerald WL, Massagué J. Genes that mediate breast cancer metastasis to lung. Nature. 2005, 436(7050):518-524.
51. Sotiriou C, Pusztai L. Gene-expression signatures in breast cancer. N Engl J Med. 2009, 360(8):790-800.
52. Martins D, Beca F, Schmitt F. Metastatic breast cancer: mechanisms and opportunities for cytology. Cytopathology. 2014, 25(4):225-230.
53. Paterlini-Brechot P, Benali NL. Circulating tumor cells (CTC) detection: clinical impact and future directions. Cancer Lett. 2007, 253(2):180-204.
54. Lapidot T, Sirard C, Vormoor J, Murdoch B, Hoang T, Caceres-Cortes J, Minden M, Paterson B, Caligiuri MA, Dick JE. A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. Nature. 1994, 367(6464):645-648.
55. Bonnet D, Dick JE. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Nat Med. 1997, 3(7):730-737.
56. Singh SK, Hawkins C, Clarke ID, Squire JA, Bayani J, Hide T, Henkelman RM, Cusimano MD, Dirks PB. Identification of human brain tumour initiating cells. Nature. 2004, 432(7015):396-401.
57. O'Brien CA, Pollett A, Gallinger S, Dick JE. A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. Nature. 2007, 445(7123):106-110.
58. Li C, Heidt DG, Dalerba P, Burant CF, Zhang L, Adsay V, Wicha M, Clarke MF, Simeone DM. Identification of pancreatic cancer stem cells. Cancer Res. 2007, 67(3):1030-1037.
59. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci U S A. 2003, 100(7):3983-3988.
60. Aktas B, Tewes M, Fehm T, Hauch S, Kimmig R, Kasimir-Bauer S. Stem cell and epithelial-mesenchymal transition markers are frequently overexpressed in circulating tumor cells of metastatic breast cancer patients. Breast Cancer Res. 2009, 11(4):R46.
61. Theodoropoulos PA, Polioudaki H, Agelaki S, Kallergi G, Saridaki Z, Mavroudis D, Georgoulias V. Circulating tumor cells with a putative stem cell phenotype in peripheral blood of patients with breast cancer. Cancer Lett. 2010, 288(1):99-106.

62. Kasimir-Bauer S, Hoffmann O, Wallwiener D, Kimmig R, Fehm T. Expression of stem cell and epithelial-mesenchymal transition markers in primary breast cancer patients with circulating tumor cells. Breast Cancer Res. 2012, 14(1):R15.

63. Scatena R, Bottoni P, Giardina B. Circulating tumour cells and cancer stem cells: a role for proteomics in defining the interrelationships between function, phenotype and differentiation with potential clinical applications. Biochim Biophys Acta. 2013, 1835(2):129-143.

64. Thiery JP. Epithelial-mesenchymal transitions in tumour progression. Nat Rev Cancer. 2002, 2(6):442-454.

65. Gradilone A, Raimondi C, Nicolazzo C, Petracca A, Gandini O, Vincenzi B, Naso G, Aglianò AM, Cortesi E, Gazzaniga P. Circulating tumour cells lacking cytokeratin in breast cancer: the importance of being mesenchymal. J Cell Mol Med. 2011, 15(5):1066-1070.

66. Dejuan Kong, Yiwei Li, Zhiwei Wang, and Fazlul H. Sarkar. Cancer Stem Cells and Epithelial-to-Mesenchymal Transition (EMT)-Phenotypic Cells: Are They Cousins or Twins? Cancers (Basel). 2011, 3(1): 716–729.

67. Kang Y, Massagué J. Epithelial-mesenchymal transitions: twist in development and metastasis. Cell. 2004, 118(3):277-279.

68. Stover DG, Bierie B, Moses HL. A delicate balance: TGF-beta and the tumor microenvironment. J Cell Biochem. 2007, 101(4):851-861.

69. Dreesen O, Brivanlou AH. Signaling pathways in cancer and embryonic stem cells. Stem Cell Rev. 2007, 3(1):7-17.

70. Moustakas A, Heldin CH. Signaling networks guiding epithelial-mesenchymal transitions during embryogenesis and cancer progression. Cancer Sci. 2007, 98(10):1512-1520.

71. Zavadil J, Cermak L, Soto-Nieves N, Böttinger EP. Integration of TGF-beta/Smad and Jagged1/Notch signalling in epithelial-to-mesenchymal transition. EMBO J. 2004, 23(5):1155-1165.

72. Derksen PW, Liu X, Saridin F, van der Gulden H, Zevenhoven J, Evers B, *et al*. Somatic inactivation of E-cadherin and p53 in mice leads to metastatic lobular mammary carcinoma through induction of anoikis resistance and angiogenesis. Cancer Cell. 2006, 10(5):437-449.

73. Rennebeck G, Martelli M, Kyprianou N. Anoikis and survival connections in the tumor microenvironment: is there a role in prostate cancer metastasis? Cancer Res. 2005, 65(24):11230-11235.

74. Rhim AD1, Mirek ET, Aiello NM, Maitra A, Bailey JM, McAllister F, Reichert M, Beatty GL, Rustgi AK, *et al*. EMT and dissemination precede pancreatic tumor formation. Cell. 2012, 148(1-2):349-361.

75. Hüsemann Y, Geigl JB, Schubert F, Musiani P, Meyer M, Burghart E, Forni G, Eils R, Fehm T, at al. Systemic spread is an early step in breast cancer. Cancer Cell. 2008, 13(1):58-68.

76. Fidler IJ. Cancer biology is the foundation for therapy. Cancer Biol Ther. 2005, 4(9):1036-1039.

77. Esmaeilsabzali H, Beischlag TV, Cox ME, Parameswaran AM, Park EJ. Detection and isolation of circulating tumor cells: principles and methods. Biotechnol Adv. 2013, 31(7):1063-1084.
78. Stover DG, Bierie B, Moses HL. A delicate balance: TGF-beta and the tumor microenvironment. J Cell Biochem. 2007, 101(4):851-861.
79. Pattabiraman DR, Bierie B, Kober KI, Thiru P, Krall JA, Zill C, Reinhardt F, Tam WL, Weinberg RA. Activation of PKA leads to mesenchymal-to-epithelial transition and loss of tumor-initiating ability. Science. 2016, 351(6277):aad3680.
80. Yao D, Dai C, Peng S. Mechanism of the mesenchymal-epithelial transition and its relationship with metastatic tumor formation. Mol Cancer Res. 2011, 9(12):1608-1620.
81. Mocellin S, Keilholz U, Rossi CR, Nitti D. Circulating tumor cells: the 'leukemic phase' of solid cancers. Trends Mol Med. 2006, 12(3):130-139.
82. Wong CW, Lee A, Shientag L, Yu J, Dong Y, Kao G, Al-Mehdi AB, Bernhard EJ, Muschel RJ. Apoptosis: an early event in metastatic inefficiency. Cancer Res. 2001, 61(1):333-338.
83. Sun YF, Yang XR, Zhou J, Qiu SJ, Fan J, Xu Y. Circulating tumor cells: advances in detection methods, biological issues, and clinical relevance. J Cancer Res Clin Oncol. 2011, 137(8):1151-1173.
84. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci U S A. 2003, 100(7):3983-3988.
85. Marrinucci D, Bethel K, Bruce RH, Curry DN, Hsieh B, Humphrey M, Krivacic RT, Kroener J, et al. Case study of the morphologic variation of circulating tumor cells. Hum Pathol. 2007, 38(3):514-519.
86. Coumans FA, van Dalum G, Beck M, Terstappen LW. Filter characteristics influencing circulating tumor cell enrichment from whole blood. PLoS One. 2013, 8(4):e61770.
87. Zheng S, Tai YC, Kasdan H. A micro device for separation of erythrocytes and leukocytes in human blood. Conf Proc IEEE Eng Med Biol Soc. 2005, 1:1024-1027.
88. Allard WJ, Matera J, Miller MC, Repollet M, Connelly MC, Rao C, Tibbe AG, Uhr JW, Terstappen LW. Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases. Clin Cancer Res. 2004, 10(20):6897-6904.
89. Zhang L, Riethdorf S, Wu G, Wang T, Yang K, Peng G, Liu J, Pantel K. Meta-analysis of the prognostic value of circulating tumor cells in breast cancer. Clin Cancer Res. 2012, 18(20):5701-5710.
90. Van der Auwera I, Peeters D, Benoy IH, Elst HJ, Van Laere SJ, Prové A, Maes H, Huget P, *et al*. Circulating tumour cell detection: a direct comparison between the CellSearch System, the AdnaTest and CK-19/mammaglobin RT-PCR in patients with metastatic breast cancer. Br J Cancer. 2010, 102(2):276-284.
91. Harris L, Fritsche H, Mennel R, Norton L, Ravdin P, Taube S, Somerfield MR, Hayes DF, *et al*. American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. J Clin Oncol. 2007, 25(33):5287-5312.

92. Hayes DF, Cristofanilli M, Budd GT, Ellis MJ, Stopeck A, Miller MC, Matera J, Allard WJ, *et al*. Circulating tumor cells at each follow-up time point during therapy of metastatic breast cancer patients predict progression-free and overall survival. Clin Cancer Res. 2006, 12(14 Pt 1):4218-4224.
93. Pachmann K, Camara O, Kavallaris A, Krauspe S, Malarski N, Gajda M, Kroll T, Jörke C, Hammer U, *et al*. Monitoring the response of circulating epithelial tumor cells to adjuvant chemotherapy in breast cancer allows detection of patients at risk of early relapse. J Clin Oncol. 2008, 26(8):1208-1215.
94. Smerage JB, Barlow WE, Hortobagyi GN, Winer EP, Leyland-Jones B, Srkalovic G, Tejwani S, Schott AF, et al. Circulating tumor cells and response to chemotherapy in metastatic breast cancer: SWOG S0500. J Clin Oncol. 2014, 32(31):3483-3489.
95. Murtaza M, Dawson SJ, Tsui DW, Gale D, Forshew T, Piskorz AM, Parkinson C, Chin SF, Kingsbury Z, *et al*. Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. Nature. 2013, 497(7447):108-112.
96. Alix-Panabières C, Pantel K. Circulating tumor cells: liquid biopsy of cancer. Clin Chem. 2013, 59(1):110-118.
97. Harouaka RA, Nisic M, Zheng SY. Circulating tumor cell enrichment based on physical properties. J Lab Autom. 2013, 18(6):455-468.
98. Trzpis M, McLaughlin PM, de Leij LM, Harmsen MC. Epithelial cell adhesion molecule: more than a carcinoma marker and adhesion molecule. Am J Pathol. 2007, 171(2):386-395.
99. Raimondi C, Nicolazzo C, Gradilone A. Circulating tumor cells isolation: the "post-EpCAM era". Chin J Cancer Res. 2015, 27(5):461-470.
100. Herlyn M, Steplewski Z, Herlyn D, Koprowski H. Colorectal carcinoma-specific antigen: detection by means of monoclonal antibodies. Proc Natl Acad Sci U S A. 1979, 76(3):1438-1442.
101. Maetzel D, Denzel S, Mack B, Canis M, Went P, Benk M, Kieu C, Papior P, *et al*. Nuclear signalling by tumour-associated antigen EpCAM. Nat Cell Biol. 2009, 11(2):162-171.
102. Osta WA, Chen Y, Mikhitarian K, Mitas M, Salem M, Hannun YA, Cole DJ, Gillanders WE. EpCAM is overexpressed in breast cancer and is a potential target for breast cancer gene therapy. Cancer Res. 2004, 64(16):5818-5824.
103. Litvinov SV, van Driel W, van Rhijn CM, Bakker HA, van Krieken H, Fleuren GJ, Warnaar SO. Expression of Ep-CAM in cervical squamous epithelia correlates with an increased proliferation and the disappearance of markers for terminal differentiation. Am J Pathol. 1996, 148(3):865-875.
104. Armstrong A, Eck SL. EpCAM: A new therapeutic target for an old cancer antigen. Cancer Biol Ther. 2003, 2(4):320-326.
105. Cristofanilli M, Budd GT, Ellis MJ, Stopeck A, Matera J, Miller MC, Reuben JM, Doyle GV, Allard WJ, *et al*. Circulating tumor cells, disease progression, and survival in metastatic breast cancer. N Engl J Med. 2004, 351(8):781-791.

106. de Bono JS, Scher HI, Montgomery RB, Parker C, Miller MC, Tissing H, Doyle GV, Terstappen LW, *et al*. Circulating tumor cells predict survival benefit from treatment in metastatic castration-resistant prostate cancer. Clin Cancer Res. 2008, 14(19):6302-6309.

107. Cohen SJ, Punt CJ, Iannotti N, Saidman BH, Sabbath KD, Gabrail NY, Picus J, Morse M, Mitchell E, *et al*. Relationship of circulating tumor cells to tumor response, progression-free survival, and overall survival in patients with metastatic colorectal cancer. J Clin Oncol. 2008, 26(19):3213-3221.

108. Lowes LE, Hedley BD, Keeney M, Allan AL. User-defined protein marker assay development for characterization of circulating tumor cells using the CellSearch system. Cytometry A. 2012, 81(11):983-995.

109. de Wit S, van Dalum G, Terstappen LW. Detection of circulating tumor cells. Scientifica (Cairo). 2014, 2014:819362.

110. Larson CJ, Moreno JG, Pienta KJ, Gross S, Repollet M, O'hara SM, Russell T, Terstappen LW. Apoptosis of circulating tumor cells in prostate cancer patients. Cytometry A. 2004, 62(1):46-53.

111. Thurm H, Ebel S, Kentenich C, Hemsen A, Riethdorf S, Coith C, Wallwiener D, Braun S, *et al*. Rare expression of epithelial cell adhesion molecule on residual micrometastatic breast cancer cells after adjuvant chemotherapy. Clin Cancer Res. 2003, 9(7):2598-2604.

112. Sieuwerts AM, Kraan J, Bolt J, van der Spoel P, Elstrodt F, Schutte M, *et al*. Anti-epithelial cell adhesion molecule antibodies and the detection of circulating normal-like breast tumor cells. J Natl Cancer Inst. 2009, 101(1):61-66.

113. Mostert B, Kraan J, Sieuwerts AM, van der Spoel P, Bolt-de Vries J, Prager-van der Smissen WJ, Smid M, et al. CD49f-based selection of circulating tumor cells (CTCs) improves detection across breast cancer subtypes. Cancer Lett. 2012, 319(1):49-55.

114. Punnoose EA, Atwal SK, Spoerke JM, Savage H, Pandita A, Yeh RF, Pirzkall A, Fine BM, Amler LC, *et al*. Molecular biomarker analyses using circulating tumor cells. PLoS One. 2010, 5(9):e12517.

115. Bidard FC, Peeters DJ, Fehm T, Nolé F, Gisbert-Criado R, Mavroudis D, Grisanti S, Generali D, Garcia-Saenz JA, *et al*. Clinical validity of circulating tumour cells in patients with metastatic breast cancer: a pooled analysis of individual patient data. Lancet Oncol. 2014, 15(4):406-414.

116. Raimondi C, Nicolazzo C, Gradilone A. Circulating tumor cells isolation: the "post-EpCAM era" Chin J Cancer Res. 2015, 27(5):461-470.

117. Ellington AD, Szostak JW. In vitro selection of RNA molecules that bind specific ligands. Nature. 1990, 346(6287):818-822.

118. Chen A, Yang S. Replacing antibodies with aptamers in lateral flow immunoassay. Biosens Bioelectron. 2015, 71:230-242.

119. Kong HY, Byun J. Nucleic Acid aptamers: new methods for selection, stabilization, and application in biomedical science. Biomol Ther (Seoul). 2013, 21(6):423-434.

120. Toh SY, Citartan M, Gopinath SC, Tang TH. Aptamers as a replacement for antibodies in enzyme-linked immunosorbent assay. Biosens Bioelectron. 2015, 64:392-403.

121. Jenison RD, Gill SC, Pardi A, Polisky B. High-resolution molecular discrimination by RNA. Science. 1994, 263(5152):1425-1429.

122. Zimmermann GR, Wick CL, Shields TP, Jenison RD, and Pardi A. Molecular interactions and metal binding in the theophylline-binding core of an RNA aptamer. RNA. 2000, 6(5):659-667.

123. Kawakami J, Imanaka H, Yokota Y, Sugimoto N. *In vitro* selection of aptamers that act with Zn2+. J Inorg Biochem. 2000, 82(1-4):197-206.

124. Mann D, Reinemann C, Stoltenburg R, Strehlitz B. In vitro selection of DNA aptamers binding ethanolamine. Biochem Biophys Res Commun. 2005, 338(4):1928-34.

125. Hamula CL, Le XC, Li XF. DNA aptamers binding to multiple prevalent M-types of Streptococcus pyogenes. Anal Chem. 2011, 83(10):3640-3647.

126. Chen F, Hu Y, Li D, Chen H, Zhang XL. CS-SELEX generates high-affinity ssDNA aptamers as molecular probes for hepatitis C virus envelope glycoprotein E2. PLoS One. 2009, 4(12):e8142.

127. Ruscito A, DeRosa MC. Small-molecule binding aptamers: selection strategies, characterization, and applications. Front Chem. 2016, 4:14.

128. Jijakli K, Khraiwesh B, Fu W, Luo L, Alzahmi A, Koussa J, Chaiboonchoe A, *et al*. The In Vitro Selection World. Methods. 2016, 106:3-13.

129. Yüce M, Ullah N, Budak H. Trends in aptamer selection methods and applications. Analyst. 2015, 140(16):5379-5399.

130. Vant-Hull B, Payano-Baez A, Davis RH, Gold L. The mathematics of SELEX against complex targets. J Mol Biol. 1998, 278(3):579-597.

131. Morris KN, Jensen KB, Julin CM, Weil M, Gold L. High affinity ligands from in vitro selection: complex targets. Proc Natl Acad Sci U S A. 1998, 95(6):2902-2907.

132. Dihua Shangguan, Tao Bing and Nan Zhang. Cell-SELEX: Aptamer Selection against Whole Cells. In: Weihong Tan, Xiaohong Fang. Aptamers Selected by Cell-SELEX for Theranostics (eds). 1st ed. Springer Heidelberg New York Dordrecht London, 2015.

133. Sefah K, Shangguan D, Xiong X, O'Donoghue MB, Tan W. Development of DNA aptamers using Cell-SELEX. Nat Protoc. 2010, 5(6):1169-1185.

134. Ohuchi S. Cell-SELEX Technology. Biores Open Access. 2012, 1(6):265-272.

135. Tang Z, Shangguan D, Wang K, Shi H, Sefah K, Mallikratchy P, and Chen HW, *et al*. Selection of aptamers for molecular recognition and characterization of cancer cells. Anal Chem. 2007, 79(13):4900-4907.

136. Van Simaeys D, Turek D, Champanhac C, Vaizer J, Sefah K, Zhen J, *et al*. Identification of cell membrane protein stress-induced phosphoprotein 1 as a potential ovarian cancer biomarker using aptamers selected by cell systematic evolution of ligands by exponential enrichment. Anal Chem. 2014, 86(9):4521-4527.

137. Mallikaratchy P, Tang Z, Kwame S, Meng L, Shangguan D, Tan W. Aptamer directly evolved from live cells recognizes membrane bound immunoglobin heavy mu chain in Burkitt's lymphoma cells. Mol Cell Proteomics. 2007, 6(12):2230-2238.

138. Yang M, Jiang G, Li W, Qiu K, Zhang M, Carter CM, *et al*. Developing aptamer probes for acute myelogenous leukemia detection and surface protein biomarker discovery. J Hematol Oncol. 2014, 7:5.

139. Stoltenburg R, Reinemann C, Strehlitz B. FluMag-SELEX as an advantageous method for DNA aptamer selection. Anal Bioanal Chem. 2005, 383(1):83-91.

140. Wang J, Rudzinski JF, Gong Q, Soh HT, Atzberger PJ. Influence of target concentration and background binding on in vitro selection of affinity reagents. PLoS One. 2012, 7(8):e43940.

141. Ahn JY, Jo M, Dua P, Lee DK, Kim S. A sol-gel-based microfluidics system enhances the efficiency of RNA aptamer selection. Oligonucleotides. 2011, 21(2):93-100.

142. Darmostuk M, Rimpelova S, Gbelcova H, Ruml T. Current approaches in SELEX: An update to aptamer selection technology. Biotechnol Adv. 2015, 33(6 Pt 2):1141-1161.

143. Avci-Adali M, Metzger M, Perle N, Ziemer G, Wendel HP. Pitfalls of cell-systematic evolution of ligands by exponential enrichment (SELEX): existing dead cells during in vitro selection anticipate the enrichment of specific aptamers. Oligonucleotides. 2010, 20(6):317-323.

144. Raddatz MS, Dolf A, Endl E, Knolle P, Famulok M, Mayer G. Enrichment of cell-targeting and population-specific aptamers by fluorescence-activated cell sorting. Angew Chem Int Ed Engl. 2008, 47(28):5190-5193.

145. McGann LE, Walterson ML, Hogg LM. Light scattering and cell volumes in osmotically stressed and frozen-thawed cells. Cytometry. 1988, 9(1):33-38.

146. Weng CH, Hung LY, Lin HI, Hsieh IS, Shieh SC, Chen YL, Lee GB. An integrated microfluidic system for automating on-chip selex process to screen tumor cell-specific aptamers. Proceeding of The 15th International Conference on Miniaturized Systems for Chemistry and Life Sciences. 2011, 2086–2088.

147. Hung LY, Wang CH, Hsu KF, Chou CY, Lee GB. An on-chip Cell-SELEX process for automatic selection of high-affinity aptamers specific to different histologically classified ovarian cancer cells. Lab Chip. 2014, 14(20):4017-4028.

148. Hasegawa H, Savory N, Abe K, Ikebukuro K. Methods for improving aptamer binding affinity. Molecules. 2016, 21(4). pii: E421.

149. Blind M and Blank M. Aptamer Selection Technology and Recent Advances. Molecular Therapy Nucleic Acids. 2015, 4: e223.

150. Dupont DM, Larsen N, Jensen JK, Andreasen PA, Kjems J. Characterisation of aptamer-target interactions by branched selection and high-throughput sequencing of SELEX pools. Nucleic Acids Res. 2015, 43(21):e139.

151. Roulet E, Busso S, Camargo AA, Simpson AJ, Mermod N, Bucher P. High-throughput SELEX SAGE method for quantitative modeling of transcription-factor binding sites. Nat Biotechnol. 2002, 20(8):831-835.

152. Ohashi H, Miyamoto-Sato E. Towards personalized medicine mediated by *in vitro* virus-based interactome approaches. Int J Mol Sci. 2014, 15(4):6717-6724.

153. Jolma A, Kivioja T, Toivonen J, Cheng L, Wei G, Enge M, *et al*. Multiplexed massively parallel SELEX for characterization of human transcription factor binding specificities. Genome Res. 2010, 20(6):861-873.

154. Slattery M, Riley T, Liu P, Abe N, Gomez-Alcala P, Dror I. Cofactor binding evokes latent differences in DNA binding specificity between Hox proteins. Cell. 2011, 147(6):1270-1282.

155. Dittmar KA, Jiang P, Park JW, Amirikian K, Wan J, Shen S, *et al*. Genome-wide determination of a broad ESRP-regulated posttranscriptional network by high-throughput sequencing. Mol Cell Biol. 2012, 32(8):1468-1482.

156. Riley KR, Gagliano J, Xiao J, Libby K, Saito S, Yu G, *et al*. Combining capillary electrophoresis and next-generation sequencing for aptamer selection. Anal Bioanal Chem. 2015, 407(6):1527-3152.

157. Gijs M, Penner G, Blackler GB, Impens NR, Baatout S, Luxen A, *et al*. Improved aptamers for the diagnosis and potential treatment of HER2-positive cancer. Pharmaceuticals (Basel). 2016, 9(2). pii: E29.

158. Hoinka J, Berezhnoy A, Sauna ZE, Gilboa E, Przytycka TM. AptaCluster-a method to cluster HT-SELEX aptamer pools and lessons from its application. Res Comput Mol Biol. 2014, 8394:115-128.

159. Alam KK, Chang JL, Burke DH. FASTAptamer: A bioinformatic tool kit for high-throughput sequence analysis of combinatorial selections. Mol Ther Nucleic Acids. 2015, 4:e230.

160. Tombelli S, Minunni M, Mascini M. Analytical applications of aptamers. Biosens Bioelectron. 2005, 20(12):2424-2434.

161. Szeitner Z, András J, Gyurcsányi RE, Mészáros T. Is less more? Lessons from aptamer selection strategies. J Pharm Biomed Anal. 2014, 101:58-65.

162. Jing M and Bowser MT. Isolation of DNA aptamers using micro free flow electrophoresis. Lab Chip. 2011, 11(21):3703-3709.

163. Jing M, Bowser MT. Methods for measuring aptamer-protein equilibria: a review. Anal Chim Acta. 2011, 686(1-2):9-18.

164. Gokulrangan G1, Unruh JR, Holub DF, Ingram B, Johnson CK, Wilson GSDNA aptamer-based bioanalysis of IgE by fluorescence anisotropy. Anal Chem. 2005, 77(7):1963-1970.

165. Chang AL, McKeague M, Smolke CD. Facile characterization of aptamer kinetic and equilibrium binding properties using surface plasmon resonance. Methods Enzymol. 2014, 549:451-466.

166. Dickey DD, Giangrande PH. Oligonucleotide aptamers: A next-generation technology for the capture and detection of circulating tumor cells. Methods. 2016, 97:94-103.

167. Shangguan D, Li Y, Tang Z, Cao ZC, Chen HW, Mallikaratchy P, *et al*. Aptamers evolved from live cells as effective molecular probes for cancer study. Proc Natl Acad Sci U S A. 2006, 103(32):11838-11843.

168. Song Y, Zhu Z, An Y, Zhang W, Zhang H, Liu D, *et al*. Selection of DNA aptamers against epithelial cell adhesion molecule for cancer cell imaging and circulating tumor cell capture. Anal Chem. 2013, 85(8):4141-4149.

169. Shigdar S, Lin J, Yu Y, Pastuovic M, Wei M, Duan W. RNA aptamer against a cancer stem cell marker epithelial cell adhesion molecule. Cancer Sci. 2011, 102(5):991-998.

170. Kim JW, Kim EY, Kim SY, Byun SK, Lee D, Oh KJ, *et al*. Identification of DNA aptamers toward epithelial cell adhesion molecule via cell-SELEX. Mol Cells. 2014, 37(10):742-746.

171. Nuzzo PV, Rubagotti A, Zinoli L, Salvi S, Boccardo S, Boccardo F. The prognostic value of stromal and epithelial periostin expression in human breast cancer: correlation with clinical pathological features and mortality outcome. BMC Cancer. 2016, 16:95.

172. Gillan L, Matei D, Fishman DA, Gerbin CS, Karlan BY, Chang DD. Periostin secreted by epithelial ovarian carcinoma is a ligand for alpha(V)beta(3) and alpha(V)beta(5) integrins and promotes cell motility. Cancer Res. 2002, 62(18):5358-5364.

173. Baril P, Gangeswaran R, Mahon PC, Caulee K, Kocher HM, Harada T, *et al*. Periostin promotes invasiveness and resistance of pancreatic cancer cells to hypoxia-induced cell death: role of the beta4 integrin and the PI3k pathway. Oncogene. 2007, 26(14):2082-2094.

174. Bao S, Ouyang G, Bai X, Huang Z, Ma C, Liu M, *et al*. Periostin potently promotes metastatic growth of colon cancer by augmenting cell survival via the Akt/PKB pathway. Cancer Cell. 2004, 5(4):329-339.

175. Morra L and Moch H. Periostin expression and epithelial-mesenchymal transition in cancer: a review and an update. Virchows Arch. 2011, 459(5):465-475.

176. Lee YJ, Kim IS, Park SA, Kim Y, Lee JE, Noh DY, *et al*. Periostin-binding DNA aptamer inhibits breast cancer growth and metastasis. Mol Ther. 2013, 21(5):1004-1013.

177. Rodrigues LR, Teixeira JA, Schmitt FL, Paulsson M, Lindmark-Mänsson H. The role of osteopontin in tumor progression and metastasis in breast cancer. Cancer Epidemiol Biomarkers Prev. 2007, 16(6):1087-1097.

178. Denhardt DT, Lopez CA, Rollo EE, Hwang SM, An XR, Walther SE. Osteopontin-induced modifications of cellular functions. Ann N Y Acad Sci. 1995, 760:127-142.

179. Mi Z, Guo H, Russell MB, Liu Y, Sullenger BA, Kuo PC. RNA aptamer blockade of osteopontin inhibits growth and metastasis of MDA-MB231 breast cancer cells. Mol Ther. 2009, 17(1):153-161.

180. Meirinho SG, Dias LG2, Peres AM3, Rodrigues LR4. Development of an electrochemical RNA-aptasensor to detect human osteopontin Biosens Bioelectron. 2015, 71:332-341.

181. Li X, Zhang W, Liu L, Zhu Z, Ouyang G, An Y, *et al*. In vitro selection of DNA aptamers for metastatic breast cancer cell recognition and tissue imaging. Anal Chem. 2014, 86(13):6596-603.

182. Viraka Nellore BP, Kanchanapally R, Pramanik A, Sinha SS, Chavva SR, Hamme A 2nd, Ray PC. Aptamer-conjugated graphene oxide membranes for highly efficient capture and accurate identification of multiple types of circulating tumor cells. Bioconjug Chem. 2015, 26(2):235-242.

183. Brown M, Wittwer C. Flow cytometry: principles and clinical applications in hematology. Clin Chem. 2000, 46(8 Pt 2):1221-9.

184. Aebisher D, Bartusik D, Tabarkiewicz J. Laser flow cytometry as a tool for the advancement of clinical medicine. Biomed Pharmacother. 2017, 85:434-443.
185. Sandra Hope, Ph.D., https://ricfacility.byu.edu/Research/CellSizeGranularity
186. Racila E, Euhus D, Weiss AJ, Rao C, McConnell J, Terstappen LW, *et al*. Detection and characterization of carcinoma cells in the blood. Proc Natl Acad Sci U S A. 1998. 95(8):4589-4594.
187. Hu Y, Fan L, Zheng J, Cui R, Liu W, He Y, *et al*. Detection of circulating tumor cells in breast cancer patients utilizing multiparameter flow cytometry and assessment of the prognosis of patients in different CTCs levels. Cytometry A. 2010, 77(3):213-219.
188. Beitsch PD, Clifford E. Detection of carcinoma cells in the blood of breast cancer patients. Am J Surg. 2000, 180(6):446-448.
189. Takao M, Takeda K. Enumeration, characterization, and collection of intact circulating tumor cells by cross contamination-free flow cytometry. Cytometry A. 2011, 79(2):107-117.
190. Sun YF, Yang XR, Zhou J, Qiu SJ, Fan J, Xu Y. Circulating tumor cells: advances in detection methods, biological issues, and clinical relevance. J Cancer Res Clin Oncol. 2011, 137(8):1151-1173.
191. Wang L, Wang Y, Liu Y, Cheng M, Wu X, Wei H. Flow cytometric analysis of CK19 expression in the peripheral blood of breast carcinoma patients: relevance for circulating tumor cell detection. J Exp Clin Cancer Res. 2009, 28;28:57.
192. Warawdekar UM, Parmar V, Prabhu A, Kulkarni A, Chaudhari M, Badwe RA. A versatile methodology for the enumeration and characterization of circulating tumor cells from patients with breast cancer. J Cancer Metastasis Treat 2017, 3:23-33.
193. Song Y, Zhu Z, An Y, Zhang W, Zhang H, Liu D, *et al*. Selection of DNA aptamers against epithelial cell adhesion molecule for cancer cell imaging and circulating tumor cell capture. Anal Chem. 2013, 85(8):4141-4149.
194. Meyer M, Scheper T, Walter JG. Aptamers: versatile probes for flow cytometry. Appl Microbiol Biotechnol. 2013, 97(16):7097-7109.
195. Yang X, Zhao X, Yang K, Liu Y, Liu Y, Fu W, Luo Y. Biomedical Applications of Terahertz Spectroscopy and Imaging. Trends Biotechnol. 2016, pii: S0167-7799(16)30027-0.
196. Stylianou A, Talias MA. Nanotechnology-supported THz medical imaging. F1000Res. 2013, 2:100.
197. Fischer BM, Walther M, Uhd Jepsen P. Far-infrared vibrational modes of DNA components studied by terahertz time-domain spectroscopy. Phys Med Biol. 2002, 47(21):3807-3814.
198. Globus T, Woolard D, Crowe TW, Khromova T, Gelmont B, Hesler J. Terahertz Fourier transform characterization of biological materials in a liquid phase. J. Phys. D: Appl. Phys. 2006, 39: 3405–3413.
199. Tang M, Zhang M, Yan S, Yang Z, Du C, *et al*. Detection of oligonucleotides in aqueous solution based on terahertz spectroscopy. In The 8th International Symposium on

Ultrafast Phenomena and Terahertz Waves, OSA Technical Digest (online) (Optical Society of America, 2016), paper IT2A.33.

200. Titova LV, Ayesheshim AK, Golubov A, Fogen D, Rodriguez-Juarez R, *et al*. Intense THz pulses cause H2AX phosphorylation and activate DNA damage response in human skin tissue Biomed Opt Express. 2013, 4(4):559-568.

201. Fan S, He Y, Ung BS, Pickwell-MacPherson E. The growth of biomedical terahertz research J. Phys. D: Appl. Phys. 2014, 47: 374009.

202. Oh SJ, Kang J, Maeng I, Suh JS, Huh YM, Haam S, Son JH. Nanoparticle-enabled terahertz imaging for cancer diagnosis. Opt Express. 2009, 17(5):3469-3475.

203. Oh SJ, Choi J, Maeng I, Leec J, Kangc J, Haamc S, *et al*. High-sensitivity terahertz imaging technique using nanoparticle probes for medical applications. IEEE photonics society winter topicals meeting. 2010, 10.1109/PHOTWTM.5421967.

204. Chen H, Chen TH, Tseng TF, Lu JT, Kuo CC, Fu SC, *et al*. High-sensitivity *in vivo* THz transmission imaging of early human breast cancer in a subcutaneous xenograft mouse model. Opt Express. 2011, 19(22):21552-21562.

205. Yngvesson SK, St Peter B, Siqueira P, Kelly P, Glick S, Karellas A, Khan A. Feasibility demonstration of frequency domain terahertz imaging in breast cancer margin determination. Proc SPIE Int Soc Opt Eng. 2012, 8221.

206. Ji YB, Park CH, Kim H, Kim SH, Lee GM, Noh SK. Feasibility of terahertz reflectometry for discrimination of human early gastric cancers. Biomed Opt Express. 2015, 6(4):1398-406.

CHAPTER 2

Selection of aptamers against mammaglobin B and mammaglobin A proteins

(The work in this chapter and the following ones (chapters 3 and 4) was accepted by Scientific Report Journal)

2.1 Introduction

2.1.1 Mammaglobin B and mammaglobin A proteins

The identification and characterization of new biomarkers for the detection of breast cancer and CTCs from breast cancer is crucial for successful diagnosis and therapy. Mammaglobin B (MGB2) was first identified by Becker *et al*. in 1998 as a secretory protein and a member of the uteroglobin gene family [1]. MGB2 can be referred to as secretoglobin family 2A member 1 (SCGB2A1), lipophilin C or lacryglobin [2]. MGB2 is a small protein consisting of 95 amino acids for the precursor protein, and 77 amino acids for the mature protein (the first 18 amino acids form a signal peptide which is cleaved off by proteases; only the mature protein is secreted from the cell), with a mass of 8.9 kDa and an isoelectric point of 4.94 for the mature protein [2]. Mammaglobin A (MGB1), another member of the uteroglobin gene family, was first identified by Watson and Fleming in 1996 [3]. MGB1 is also referred to as secretoglobin family 2A member 2 (SCGB2A2), mammaglobin 1, or human Mammaglobin (hMAM), and consists of 93 amino acids for the precursor protein, and 75 amino acids for the mature protein. As with MGB2, the first 18 amino acids correspond to a signal peptide which is cleaved for the secretion of the mature protein that has a mass of 8.5 kDa and an isoelectic point of 3.91.

All secretoglobins, including MGB2 and MGB1, contain alpha-helices in their structure and are often found as dimers [5]. MGB2 is mostly expressed in mucosal tissues and is found at high levels in many secretions including those from uterine, prostatic, pulmonary, and lacrimal and salivary

glands [1,2,6,7]. Little is known about MGB2 and MGB1 function, but it has been reported that they have a role in cancer development, immune system regulation, chemotaxis, and the transport of aromatic molecules like steroid hormones and biphenyls [8-10]. MGB2 and MGB1 often form dimers with other lipophilin proteins (MGB1 forms a dimer with lipophilin B (Figure 2.1) [11]. A study by Zuo *et al*., proposed that MGB1 is a transmembrane protein due to the presence of an N-terminal domain that may act as a transmembrane domain, therefore, MGB1 could be used as a marker of breast cancer [12]. On the other hand no similar studies were found for MGB2.

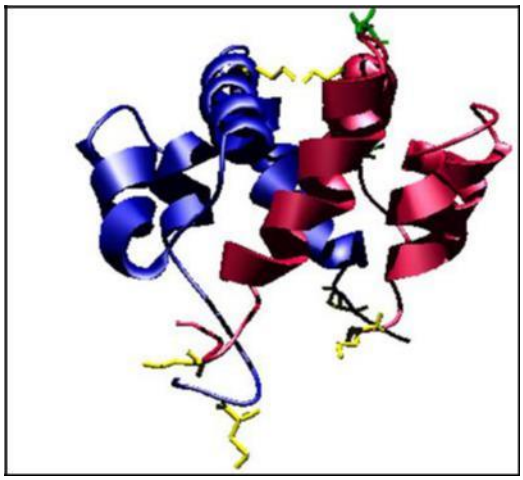


Figure 2.1: MGB1 dimer structure.
MGB1 (in red) consists of two alpha helices and binds to lipophilin B (in blue). The four alpha helices interact together via six cysteines residues (in yellow as sticks) to form the complex. The two N linked-glycosylation sites are shown at the beginning and the end of the helix (green sticks) [11].

MGB2 and MGB1 are homologous (58% homology) [1], but they differ in their patterns of expression. MGB2 is expressed in the breast, uterus, salivary gland, tear fluid, thyroid, testis, and ovary [1,13]. On the other hand, MGB1 expression is exclusive to the adult mammary gland [14]. MGB2 is over expressed in several types of cancers, including breast [14], ovarian [16-18], and endometrium cancers [19]. It has been reported that MGB2 is a marker of CTCs found in the blood of breast cancer patients [20], whereas MGB1 is over expressed only in breast cancer [3,21,22].

Both proteins have been reported as markers of breast cancer micro-metastases to lymph nodes [14,23,24]. MGB1 has been investigated as a CTC marker in the blood of breast cancer patients; Real Time PCR (RT-PCR) is the method used to detect MGB1 mRNA in the blood [25-28]. MGB1 showed higher specificity than any other marker used in CTC detection (e.g. EGFR and CK- 19) in the blood of breast cancer patients measured by the RT- PCR method [29]. Since MGB1 is limited to breast cancer, it is a very attractive target and has higher clinical values than many other markers [30,31]. Recently, a MGB1 DNA vaccine has successfully passed phase 1 clinical trial demonstrating the safety to use this vaccine in patients with metastatic breast cancer [32]. On the other hand, no further investigations of the role of MGB2 as a marker of CTCs in the blood of breast cancer patients have been done.

No reports have been found in the literature that investigate the detection of CTCs in the blood using MGB1 or MGB2, or both, as markers of CTCs from breast cancer using new methods (other than RT-PCR). Moreover, no reports have been published which establish new probes (other than antibodies) against these two biomarkers and investigate their binding to CTCs using new methods *in vitro* or *in vivo*. Herein, the development and characterization of aptamers against MGB1 and MGB2 were established for the first time.

2.1.2 Hybrid (Cross-over) SELEX

As mentioned in chapter one, cell-SELEX has many advantages over other types of SELEX. Moreover, cell-SELEX has several sub-types depending on the target used which include: whole cell culture or tissue as the target [33], recombinant protein and cells both as target [34], cells as target and antibodies as epitope pre-determination [35]. In this section, only the type that uses protein and cells both as target will be discussed as the core subject of the thesis.

Purified recombinant proteins can be used as targets for SELEX [36]. Proteins are subject to posttranslational modifications when synthesized in the cell naturally, and thus the purified form of these proteins might not represent the real protein, especially in the case of highly glycosylated proteins [37].

Moreover, recombinant proteins or peptides often can not fold into the right three-dimensional structure that is formed under physiological conditions [37], therefore this will affect the binding affinity of those aptamers to their corresponding targets on the surface of the cell [37]. To overcome this problem, cell-SELEX based technology was introduced to use whole live cells as targets in the selection process [38]. The cell has a large number of surface molecules and potential target biomarkers, and thus many unwanted aptamers will be generated [38,39]. Counter cell lines (control cells) that do not express the target biomarker of interest, or express it in very low levels, are used in a counter selection step to eliminate the unwanted aptamers at the end of the selection [37,39]. It has been reported that the target of interest could be over expressed in a particular cell line by gene transfection. The infected cell line could be used in cell SELEX [40], however many off-target surface biomarkers/molecules are co-expressed on the cells of interest leading to low enrichment of aptamers [39].

Hybrid or crossover SELEX was first established in 2001 by Hicke *et al*. In this study they developed aptamers with high affinity against Tenascin-C protein by alternating the selection process between the recombinant form of the protein and tumor cells [34]. In a subsequent study in 2006, Pestourie *et al*. developed aptamers against the mutated form of the membrane receptor tyrosine kinase RET [41]. Moreover, Parekh *et al*. screened a new aptamer against CD30 (a protein over expressed in lymphoma) by a method called hybrid SELEX [42]. The C2NP aptamer has a high affinity to CD30 on the surface of lymphoma cell lines, and was proposed to be a potential target for cancer immunotherapy [42].

Hybrid SELEX leads to a higher enrichment of aptamers against the target of interest, since the DNA library is exposed to both the purified form of the protein, and the native form on the cell surface. Moreover, the aptamers resulting from hybrid SELEX are more specific to the tumor cell line versus the control cell line compared to aptamers developed using only whole cells as target [41]. Some limitations of hybrid SELEX include the requirement for more equipment, and a higher cost compared to a normal cell SELEX.

2.1.3 High throughput sequencing

As mentioned in in Chapter one, HTS is a powerful tool that has improved the selection process in many ways. Many types of HTS platforms are commercially available [43]. Illumina's platforms are the most dominant in the market of DNA sequencing [44]. Three major factors affect the use of a certain platform: the amount of output data that the machine can produce, the cost, and the errors produced during sequencing [43]. So far, Illumina's platforms have invaded the market with a series of models with improved output, and reduced cost and error [43,44]. The first Illumina platform was the Genome Analyzer II in 2006, followed by a set of advanced models [43]. The Illumina sequencing process starts with the proper preparation of the DNA library, in which specific adaptors are ligated (by regular PCR) to the DNA library. The DNA library is then fragmented to shorter sequences using enzymes. Complementary sequences for those adaptors are found on a glass flow cell to bind the adaptor-ligated DNA fragments. Bridge amplification then takes place on the surface of the flow cell, creating small clusters of identical sequences to be analyzed. One strand of DNA is removed, and fluorescently labeled nucleotides are passed by each cluster. For the first cycle, an image of the flow cell is recorded. The fluorescent label is cleaved and a second round starts again, passing fluorescently labeled nucleotides by each cluster. Another image is taken, and the whole DNA sequence is read by this process. All the sequences are then aligned to what they were before the fragmentation process [45,46]. Across all Illumina sequencers (HiSeq, MiSeq, and NextSeq 500), the overall error rates are below 1%, and the most common type of error is substitution [47].

Many algorithms have been developed to analyze HTS sequencing data [48-50]. AptaCluster is an algorithm which allows for an efficient clustering of HTS-SELEX aptamer pools, a task that traditional clustering algorithms could not accomplish due to the massive size of such datasets [51,52]. AptaCluster can cluster millions of sequences in the context of whole-cell HTS-SELEX, without loss of noticeable performance [51,52]. AptaCluster is a good match in the case of HTS hybrid SELEX due to the large dataset obtained after HTS sequencing. The AptaCluster algorithm analyzes the HTS data according to sequence similarity [51].

AptaGUI is the visual face of AptaCluster, in which all the results are visualized and data are monitored in terms of sequence-based analysis of pools, including their round-to-round enrichments, and secondary structure predictions [53].

This chapter will describe the selection process of MGB2 and MGB1 aptamers in detail and their initial screening.

2.2 Materials and methods

2.2.1 Materials and cell culture supplements

All amidites including, dA-CE Phosphoramidite, Ac-dC-CE Phosphoramidite, dmf-dG-5'-CE Phosphoramidite, dT-CE Phosphoramidite, 5’-fluorescein phosphoramidite (6-FAM), and reagents for DNA synthesis (Cap Mix A, Cap Mix B, Oxidizing Solution, Deblocking Mix) were purchased from Glen Research (USA). 1000Å and 500 Å DNA synthesis columns (Controlled Pore Glass (CPG)) were ordered from Bioautomation (USA).

Glutathione S transferase (GST), MGB1, and MGB2 recombinant fusion proteins were purchased from Novus Biologicals (Canada), GST recombinant protein and Human Mammaglobin-B(SCGB2A1) ELISA kit were purchased from Cedarlane (Canada). Glutathione-agarose beads were purchased from Sigma (Canada). Yeast tRNA and all tissue culture supplements were obtained from Thermo Fisher Scientific (Canada). Salmon sperm DNA solution was purchased from Invitrogen (Canada). Mammary Epithelial Cell Growth Medium (MEGM) and MEGM kits were purchased from Lonza (USA), Endothelial Cell Basal Medium MV 2 and MV2 kits were purchased from Promo-cell (Canada), Dulbecco's modified Eagle's Medium (DMEM), Eagle's Minimum Essential Medium (EMEM), and Fetal Bovine Serum (FBS) were ordered from Gibco™ / Thermo Fisher Scientific (Canada). Recombinant human insulin was ordered from Santa Cruz Biotechnology Inc. (USA). All chemicals and SELEX reagents were purchased from Sigma and Bishop (Canada) unless otherwise indicated. All buffers were prepared with Milli-Q water and filtered with Corning 0.22 µm cellulose acetate low protein binding filter units (Corning, USA).

2.2.2 Cell lines and cell culture

MCF-7 (breast adenocarcinoma, ATCC: HTB-22), MDA-MB-415 (breast adenocarcinoma, ATCC: HTB-128), MCF-10A (fibrocystic disease, ATCC: CRL-10317), HCAEC (Human Primary Coronary Artery Endothelial Cells; Normal, (ATCC: PCS-100-020)), were purchased from American Type Culture Collection (ATCC). MCF7 and MDA-MB-415 were cultured in DMEM, 10% FBS and 0.01 mg/ml human recombinant insulin as supplement. MCF10A cells were cultured in MEGM containing Bovine Pituitary Extract (BPE) (2 mL); human Epidermal Growth Factor (hEGF), 0.5 mL; hydrocortisone, 0.5, (0.5 mL); Gentamicin-Amphotericin (GA)-1000, 0.5 mL; insulin, 0.5 mL for every 500 mL of medium. HCAEC were cultured in MV2 basal medium phenol red- free containing Fetal Calf Serum (FCS) (25 mL); hEGF 0.5mL; Insulin-like Growth Factor (IGF-1) (0.5mL); Hydrocortisone, 0.5, (0.5 mL); Vascular Endothelial Growth Factor 165 (recombinant human) (0.5mL); Basic Fibroblast Growth Factor (recombinant human) 0.5mL; ascorbic acid 0.5mL for every 500 mL of MV2 medium. All cells were cultured in a 5% CO2-humidified chamber at 37ºC to reach 90% confluence before further sub culture or being used in experiments.

2.2.3 DNA synthesis

A MerMade 6 automated DNA synthesizer was used to synthesize all DNA sequences and the library. Phosphoramidite chemistry using solid phase CPG columns was used as the standard method of synthesis [54]. The following DNA sequences were used: ssDNA library: 5'-CAT GCT TAC CTA TAG TGA AC (N24) CTT TGA GAA CTG ACT CAT AC-3', forward primer: 5'- (6-FAM)-CAT GCT TAC CTA TAG TGA AC-3', reverse primer: 5'- (A20) GTA TGA GTC AGT TCT CAA AG-3'. All amidites and CPG columns were stored at -20 degrees C and were warmed to Room Temperature (RT) in a dessicator for 20 minutes, then amidites were dissolved in acetonitrile (ACN) under argon gas, and loaded onto the DNA synthesizer. Upon the completion of the synthesis process, all ssDNA was cleaved off the CPG solid support using ammonium hydroxide at high temperature (55^{o}C for 18 hours). After that, ssDNA was dried using a Speed Vac (Savant, USA) to be ready for purification using 12%-19% Polyacrylamide Gel Electrophoresis (PAGE).

Gel bands were heated in DI water overnight at 37^{o}C to release the DNA out of the gel, then freeze dried using a lyophilizer (LABCONCO, USA). Amicon ultra 0.5mL centrifugal filters (cut off 3KDa) (Millipore, USA) were used to desalt all ssDNA. Then Cary 300 Bio UV-Visible Spectrophotometer (Agilent: Santa Clara, CA, USA) running the Cary Win UV Scan application software (Version 3.00 (1.82)) was used to characterize all ssDNA, and the molecular weight of all sequences was verified by ESI- LC-MS (McGill University, Canada). To make it easier to track the DNA during the selection, the ssDNA library was tagged with 6-FAM (fluorescein) after amplification by PCR (explained below) and was performed for the first time on the library with the primers before the start of the SELEX process.

2.2.4 Hybrid (Cross-over) SELEX

2.2.4.1 MGB2 and MGB1 Recombinant protein SELEX

Glutathione-Agarose beads were used as the solid support in the protein SELEX. The beads were prepared by incubating them with DI overnight at 4˚C at a concentration of 200 mL/g for 100% swelling. The beads were then spun down at 1000 g for 1 min, washed with DI followed by binding buffer (150 mM Tris, 150 mM NaCl, 5 mM KCl, 5 mM MgCl2, pH 7.4). MGB2 or GST proteins were immobilized on the beads by incubating 16µg of GST protein and 9 µg of MGB2 with 100 µL of the beads in binding buffer at RT for 30 min. The beads were spun down in the same way as above, and washed with binding buffer three times. To verify the binding of the protein to the beads, the washes were scanned at 280nm using the UV-Vis. Round 1 of the SELEX process started with heating the tagged DNA library (pool) at 95^{o}C for 5 min and gradually cooled for 10 min at 4^{o}C then 10 min at RT (to allow the DNA to fold into the right 3D structure). 5 µL of the starting DNA library were transferred to a separate tube (to be used later in calculating the percentage recovery of the DNA). A counter selection round was then performed on GST-beads by incubating 5 nmoles of ssDNA library with 100 µL of the beads in binding buffer for 1 hour at RT. The pass through solution (containing sequences that did not bind to GST-beads) was collected to be used as the starting pool for MGB2 selection (Positive selection).

MGB2 first-round was done in the same way as GST, but the incubation time of the pool with MGB2 was 30 min. A partitioning step to separate the bound from un-bound sequences took place by washing the beads with binding buffer two times. The protein-DNA complex was eluted using elution buffer (7 M urea, 5 mM EDTA, pH 5.5) heated at 95^{o}C five times each in 100 µL. In order to separate the DNA from the protein-DNA complex, a phenol-chloroform extraction followed by ethanol precipitation was performed (Figure 2.2) [55]. The percentage recovery of the first- round DNA was then calculated by a Fluorolog (HORIBA, using HORIBA FluoroEssence V3.5 software, USA) in which the fluorescence intensity of 6-FAM tagged to the ssDNA pool was recorded by observing the emission peak of 6-FAM at 520 nm. The percentage recovery of the bound DNA was then calculated as percentage of DNA bound = the intensity of the total elutions the intensity the initial pool. Elutions were then recovered, and a desalting step was done before the amplification (PCR) to ensure no salts were present which could interfere with the amplification.

DNA was then prepared for PCR as follow, for 100 µL of PCR reaction: 50 µL of 2X Flu-Mag buffer (100 mM KCl, 200 mM Tris, 2% Triton X100 pH 9.00), 8µL of 25mM $MgCl_2$, 2µL of 10mM DNTP, 0.5 µL of 0.2mM of forward primer, 0.5 µL of 0.2 mM of reverse primer, 1 µL Taq polymerase (5 units) were mixed with the DNA template. PCR was done using a thermal cycler (Eppendorf, USA) with the following conditions: denaturation of DNA at 95^{o}C for 30 s, annealing at 58^{o}C for 30 s, and elongation at 72^{o}C for 30 s for 25 cycles, then a final elongation step at 72^{o}C was done for 10 min. A purification of the first-round DNA for downstream processing was done as mentioned above, starting with PAGE, heat-shock, freeze drying, followed by desalting and quantification. Round 2 was then repeated the same way, and so on. Rounds of SELEX using MGB2 protein as the positive target, and GST as the counter selection target were repeated 15 times.

MGB1 was used as a counter selection instead of GST starting from round 16 until round 18 of the MGB2 selection, resulting in the separation of the original DNA library into two halves.

Rounds 16-18 were applied same way as mentioned above, but the pass through from MGB1 in round 17 was kept as the starting DNA library for MGB1 selections. Rounds 19-21 of MGB2 SELEX were performed with GST as a counter selection. On the other hand, a SELEX for MGB1 was started using GST as a counter selection for 4 rounds, since MGB1 was already used as a counter selection for 3 rounds in the MGB2 SELEX (a total of 7 rounds for MGB1). All the downstream processing of both libraries was done as mentioned above. In order to increase the stringency throughout the selection for both libraries, the number of washes was increased (up to 5 times), the amount of the initial DNA was decreased to 100pmol, the amount of both the beads and proteins was decreased to 25 µl, and the incubation time for the counter selection step was decreased to 25 min and 15 min for the positive selection step. The SELEX selection for both MGB1 and MGB2 ended when the % recovery of both libraries reached a plateau.

In summary, by the end of both SELEX selection processes, we had two enriched pools for both MGB1 and MGB2 with 21 and 7 rounds of SELEX for MGB2 and MGB1, respectively (Figure 2.2). The same primers were used for PCR throughout all SELEX steps.

2.2.4.2 Cell-SELEX on breast cancer cell lines using protein SELEX libraries

Both final enriched DNA libraries stemming from protein SELEX were used as the initial pool for cell SELEX. Both MCF7 and MDA-MB-415 are breast cancer cell lines derived from metastatic sites. Both of them were used as the positive target in the selection of MGB2 and MGB1, respectively. MCF10A is an immortal breast cell line and was used as a counter selection target for both MGB2 and MGB1 SELEX. Another counter selection target used was HCAEC, a normal endothelial cell line. MGB2 and MGB1 ELISA kits were used to investigate the expression of MGB2 and MGB1 in all cell lines used in cell-SELEX.

MCF10A and HCAEC ($2X10^6$ cells) were maintained in 100 mm tissue culture plates, and cells were washed with washing buffer (4.5g/L glucose, 5mM MgCl2 in PBS, pH 7.4) to remove media, using a cell scraper to scrape the cells for collection in a 1.5 mL Eppendorf tube.

Cells were then centrifuged at 500 xg for 5 min and re-suspended in binding buffer (4.5g/L glucose, 5mM MgCl2, 2mg/mL BSA, 0.2 mg/mL yeast tRNA, 200µL of FBS in PBS, pH 7.4). At the same time, the DNA libraries of either MGB2 or MGB1 were prepared by heating and cooling down the same way as in the protein SELEX. The DNA library was then incubated with the HCAEC cell mixture for 1 hour at 4^{o}C. After the incubation, the cells were centrifuged as above. The supernatant, which contained the non-binding sequences to HCAEC, was incubated with MCF10A for 1 hour at 4^{o}C, and then cells were centrifuged as above, using the supernatant as the starting pool for MCF7 cells or MDA-MB-415 (Figure 2.2). In either SELEX, 1X10^{6} of cells were maintained, scraped and collected in the same way as was described for the counter selection cells. The supernatant from the final counter selection cells was then incubated with the positive selection cells in binding buffer for 30 min at 4^{o}C. The cells were then centrifuged as described above. A partitioning step was then performed in the same way as in the protein SELEX by washing the cells three times with washing buffer (Figure 2.2). To obtain the target-bound DNA sequences, cells were re-suspended with 1 mL of washing buffer and boiled at 95^{o}C for 5min to pellet the cell fragments, and to obtain the bound DNA. The cells were then centrifuged at 12000 xg for 10 min. and the supernatant was collected and dried overnight using a speed Vac. A PCR was then done on either pool following the same conditions used in the protein SELEX, but the number of cycles decreased to 20 PCR cycles instead of 25. Then downstream processing after PCR was done in the same way as described for the protein SELEX. The final ssDNA library was used as the starting pool for the second round of cell-SELEX, which was repeated 7 times for both MCF7 and MDA-MB-415.

Cells have many receptors (potential targets) on their surface with which the DNA library has a chance to interact, therefore the final libraries after the 7 rounds of each MCF7 and MDA-MB-415 selection were transferred back to protein SELEX for 2 rounds for each MGB2 and MGB1 proteins to minimize the non-specific binders, and to ensure that the final aptamers would bind to MGB2 or MGB1 on the surface of breast cancer cells (Figure 2.2).

Unlike the protein SELEX in which the percentage recovery of the DNA library was monitored by a fluorolog, flow cytometry (BD Accuri™ C6 Cytometer, using BD Accuri™ C6 software, Canada) was used to monitor the binding of the final libraries to their cell line.

The random pool (before the SELEX), the enriched pools (last pool of protein SELEX) and the last pools of the whole cell SELEX were incubated at a concentration of 200nM with $2X10^5$ cells of MCF7, MDA-MB-415, HCAE, and MCF10A and 0.3 mg/mL of ssDNA from salmon sperm, 0.5 mg/mL bovine serum albumin (BSA) in 200µL of binding buffer for 30 min at 4°C. The cells were then centrifuged as mentioned above, washed with washing buffer, centrifuged again, re-suspended with 200 µL washing buffer, and loaded in a 96-well plate read using flow cytometry.

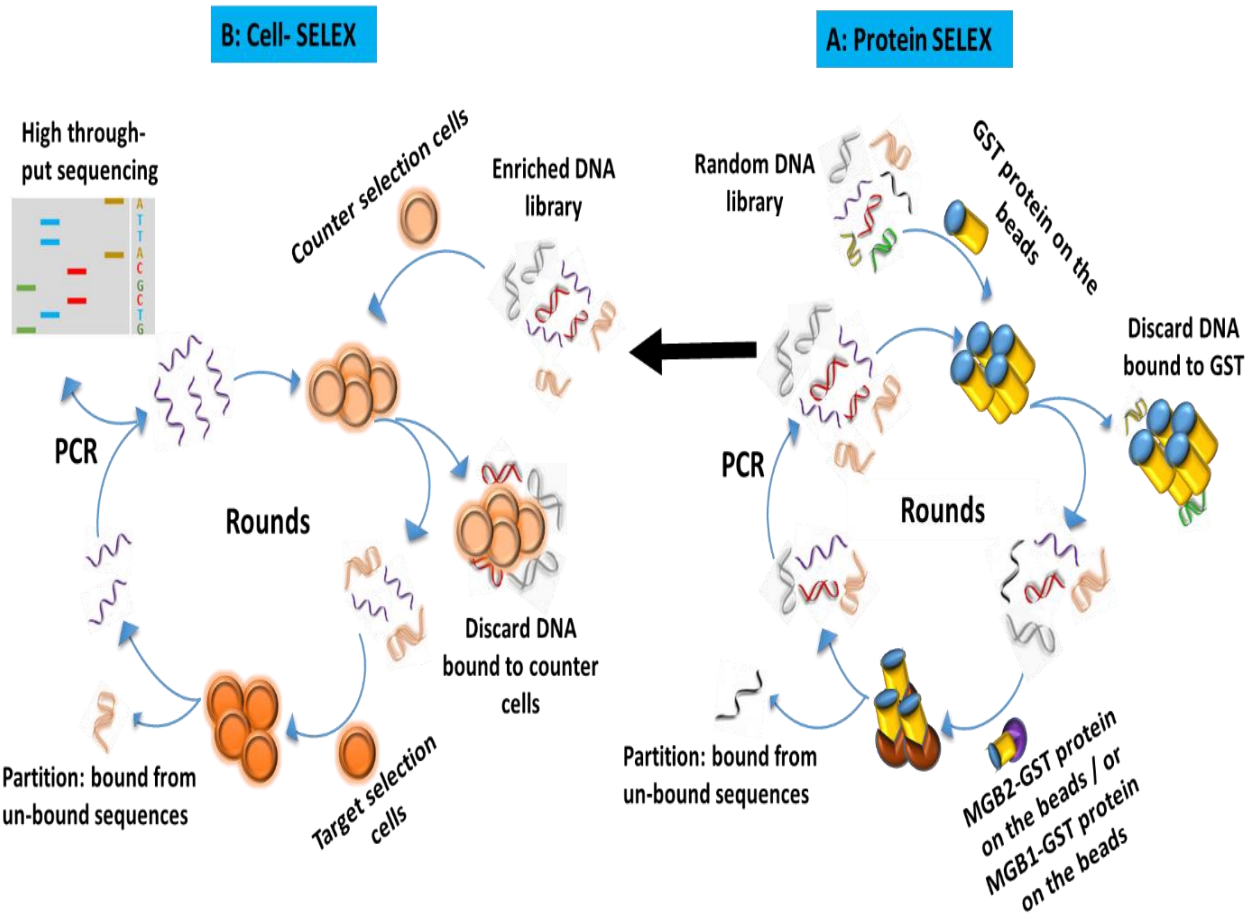


Figure 2.2: Schematic diagram of the hybrid SELEX method for the selection of MGB2 and MGB1-specific DNA aptamers. The hybrid SELEX in our experiment was divided into A: a protein-based SELEX, and B: a cell based SELEX. First GST-MGB2 and GST-MGB1 proteins were used as targets for a random ssDNA library. After 21 rounds of selection, the enriched pool for both proteins were transferred to be used in cell SELEX in which HCAEC and MCF10A were used as counter selection cell lines. MCF7 was used as the positive selection cell lines for MGB2, and MDA-MB-415 was used as the positive selection cell line for MGB1. After 7 rounds of selection, the final pools for both targets were transferred back to the protein for another 2 additional rounds to make the library more specific to MGB2 and MGB1. The final two pools after that were prepared to be sent to HTS.

2.2.5 High throughput sequencing

Several SELEX-rounds from each MGB2 and MGB1 hybrid SELEX were prepared for HTS sequencing. Special adaptors (DNA primers which contained a distinct barcode inserted within the primer sequence) were ligated to the desired libraries by PCR, using the same conditions used in the SELEX but with a minimum number of cycles (17 cycles). The final product was purified using 8% PAGE instead of 12% PAGE. In the case of MGB2, five libraries were prepared for HTS: R0, R16, R16 counter selection, R30, and R30 counter selection. From the MGB1 selection, eight libraries were prepared for HTS: R17 counter selection of MGB2, R4, R5, R7, R10, R11, R13, and R13 counter selection. All prepared libraries were combined in which the five pools for MGB2 were sent as one sample at 75 ng total DNA and the eight pools for MGB1 as one sample at 75ng as well. Each separate round was distinguished using the barcodes inserted in the ligated adaptors. All samples were sent to Genome Quebec facility (Canada) to be sequenced using Illumina HiSeq sequencing platform. Upon receiving the HTS data, AptaCluster was used as the software to analyze the data, and RNAstructure software (Mathews Lab RNAstructure) was used to predict the secondary structure of the chosen sequences. Note that RNAstructure cannot predict G-guadruplexes structures.

2.2.6 Screening of MGB2 and MGB1 aptamers

The binding of the aptamer candidates from both MGB2 and MGB1 was tested against HCAEC, MCF10A, MCF7, and MDA-MB-415 cell lines by using flow cytometry. Each aptamer for each target was incubated at a concentration of 200nM with $2X10^5$ cells of each cell line mentioned in 0.3 mg/mL of ssDNA (DNA for blocking), and 0.5 mg/mL BSA in 200µL of binding buffer for 30 min at 4^oC. The cells were then centrifuged in the same way as in SELEX, washed with washing buffer, centrifuged again, re-suspended with 200µL washing buffer, and loaded in a 96-well plate read using a flow cytometry. Each aptamer trial was repeated three times in duplicate for each sample.

2.3 Results and Discussion

2.3.1 Monitoring the percentage recovery of MGB2 and MGB1 DNA libraries during the selection

2.3.1.1 Monitoring the percentage recovery of DNA libraries during protein SELEX

As Figure 2.3 illustrates, the percentage recovery of MGB2-binding DNA reached a plateau at round 15 (A). Due to the fact that MGB2 is 58% similar to MGB1 (as mentioned in the introduction of this chapter), MGB1 was used as a counter selection to enrich the library with aptamers more specific to MGB2, and the same applied for MGB1. It was noticeable that the percentage recovery for MGB2 library dropped to a little over half (3.38-1.35%) when MGB1 was used as a counter selection, which matched our expectations and the literature since both are similar proteins [1]. The percentage recovery of the MGB2 library increased to reach 4.8% after 21 rounds of SELEX. This indicated that the library was now more enriched toward MGB2 protein. As we can see in Figure 2.4A, the percentage recovery for MGB1 was higher than the percentage recovery for MGB2 in round 17. The pool in round 17 was more specific to MGB1 protein, and this pool was used as the initial pool to select aptamers against MGB1. In Figure 2.3B, the percentage recovery against MGB1 increased to reach 4.8% in the last round of SELEX, which was similar to that of MGB2.

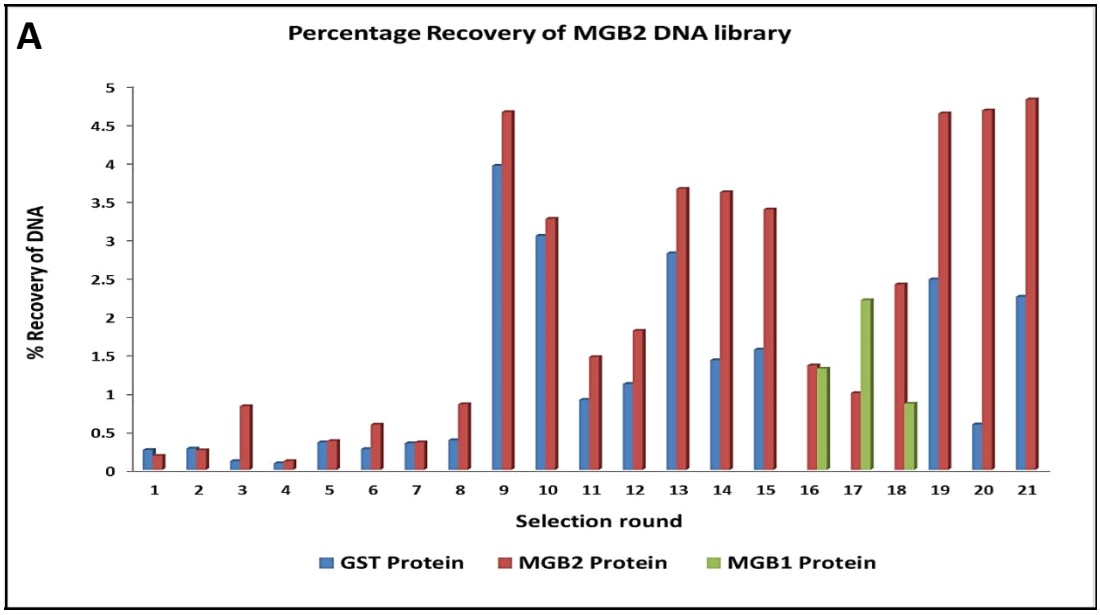
MGB2 and MGB1 are found as dimers with lipophilin proteins (as discussed in the introduction of this chapter). Counter selection against lipophilin proteins did not take place because of the fact that the potential selected aptamers for both MGB2 and MGB1 might bind to the site that connects the two proteins together. Therefore, the binding could not be achieved if the sequences against lipophilin were excluded by a counter selection step.

2.3.1.2 MGB2 and MGB1 are exclusively expressed in positive selection cell lines and not in the counter selection cell lines

The final pools of both protein SELEX procedures were transferred to breast cancer cell lines MCF7 and MDA-MB-415, in order to enrich the DNA pools against the native form of MGB2 and MGB1 on the surface of their cells. As mentioned earlier, the expression of MGB2 and MGB1 was investigated in all cell lines, using MGB2 and MGB1 ELISA kits before cell-SELEX was initiated. As Figure 2.4 showed, the highest expression of MGB2 and MGB1 proteins was observed in MCF7 and MDA-MB-415 cells (16.19 and 21.98 ng/mL respectively).

The lowest level of MGB2 and MGB1 was seen in HCAEC cells (around 0.25 ng/mL for both proteins). MCF10A cells showed low levels of MGB2 and MGB1 (2.4 and 3.7 ng/mL respectively). MDA-MB-415 and MCF7 expressed low levels of their counter proteins (Figure 2.4). Our ELISA results confirmed the results found in previous studies [1,3].

The final pools from the protein SELEX for each target, were transferred to cell lines to perform cell–SELEX. It was expected that the MGB2 pool of selected aptamers would bind preferentially to MCF7 cells, while the MGB1 pool of selected aptamers would bind preferentially to MDA-MB-415 cells.



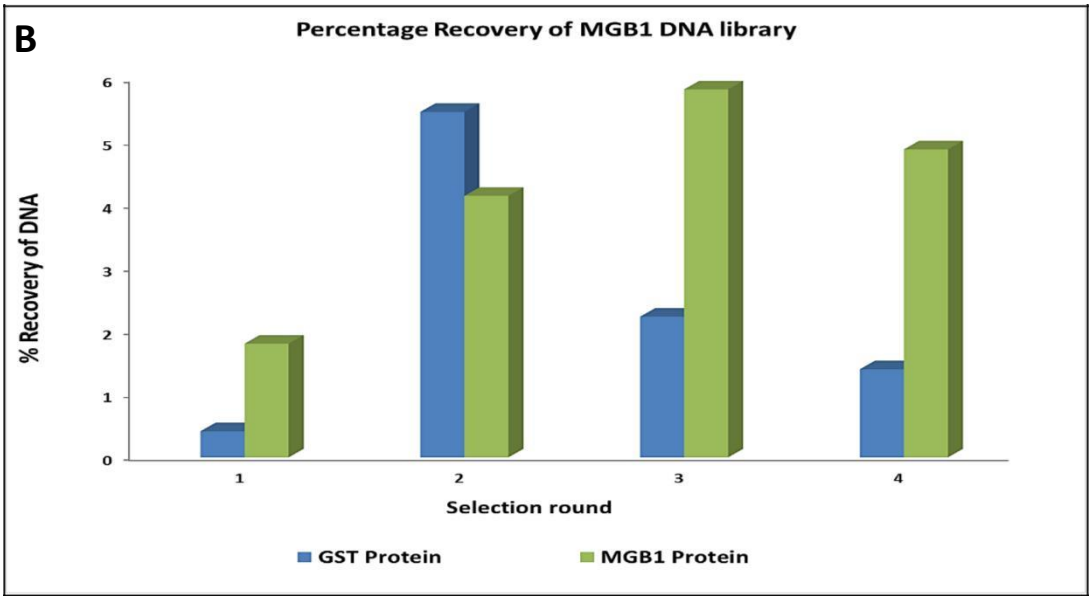


Figure 2.3: Percentage recovery of MGB2 (A) and MGB1 (B) ssDNA libraries monitored using a fluorolog. The emission of the tagged ssDNA libraries (with 6-FAM) was monitored at 520nm using a fluorolog. The percentage of ssDNA that bound to the targets (percentage recovery of DNA) was calculated as percentage of DNA bound = the fluorescence intensity of the total elutions / the fluorescence intensity the initial pool.

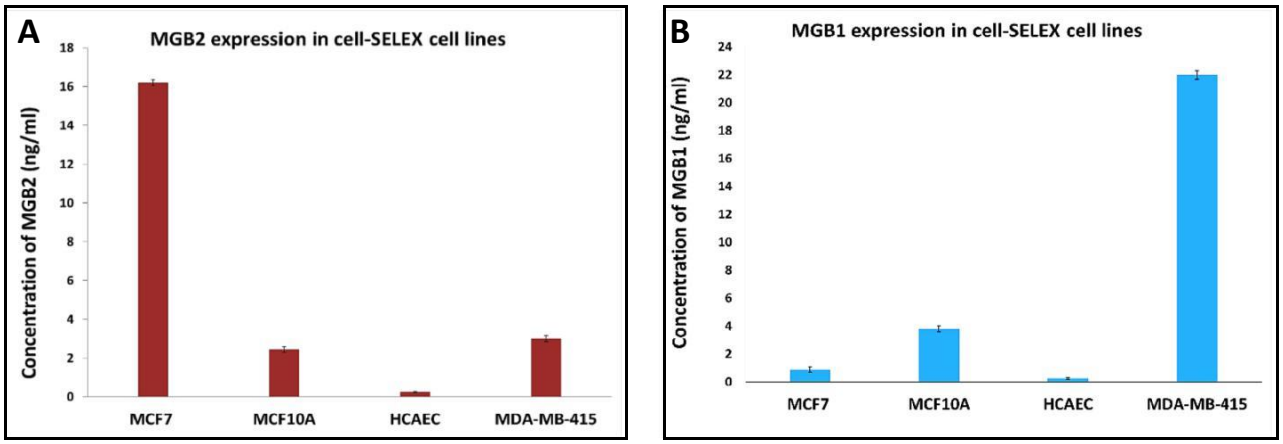


Figure 2.4: MGB2 and MGB1 expression of the lysate of all cell lines used in cell-SELEX. The highest expression of MGB2 and MGB1 proteins was observed in MCF7 (A) and MDA-MB-415 cells (B) (16.19 and 21.98 ng/mL respectively). The lowest expression of the same proteins was seen in HCAEC cells (around 0.25 ng/mL for both proteins). MCF10A showed low expression levels of both proteins, MDA-MB-415 and MCF7 expressed low level of their counter proteins. The expression of MGB2 and MGB1 was measured using ELISA kits for both proteins.

2.3.1.3 Monitoring the percentage recovery of DNA libraries during cell SELEX

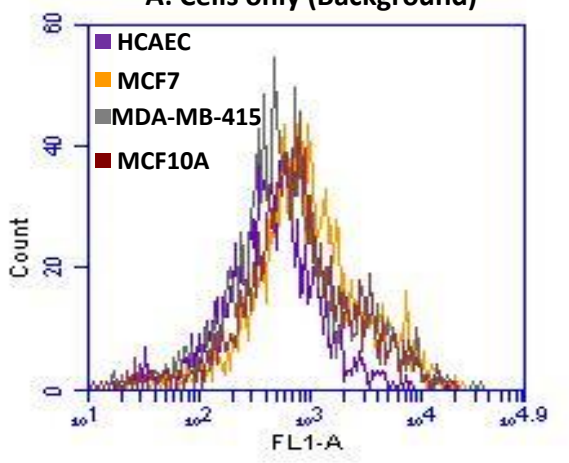
R0, R21 MGB2, R4 MGB1, R30 MCF7 and R13 MDA-MB-415 were monitored during SELEX against the cell lines (live cells were used) used in the selection by flow cytometry, in which pools of aptamers, tagged with 6-FAM, were mixed with cells and monitored by counting 10000 events.

The results showed that the last selection round (R13) against MDA-MB-415 cells had the highest fluorescence intensity (40409.41) (Figure 2.5 D and G) compared to cells only (background) and counter selection cell lines (Figure 2.5 A-C). Similar results were shown when monitoring R30 of MCF7 which had a fluorescence intensity of 32199.84 (Figure 2.5 E and H). The enriched pools for both targets bound to their respective positive cell lines, but R13 for MGB1 bound more to MDA-MB-415 (9513.33) than to R21 MGB2 (7076.7). Low non-specific binding of all pools was observed when testing them against the counter selection cell lines, HCAEC and MCF10A, compared to cells only (Figure 2.5 A, B, C, F, G and H).

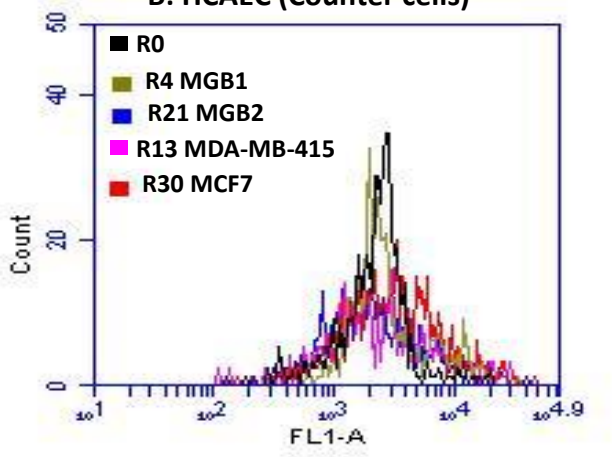
The binding of the last pools to both targets (R30 MCF7 and R13 MDA-MB-415) was significant ($P<0.05$) compared to counter cells. Moreover, the binding of rounds 21 of MCF7 and

4 of MDA-MB-415 was also significant ($P<0.05$) compared to counter cells. The binding of round 0 (the initial pool) was not significant ($p>0.05$) compared to counter cell lines. Finally, the binding of counter cells was not significant ($P>0.05$) compared to cells only.

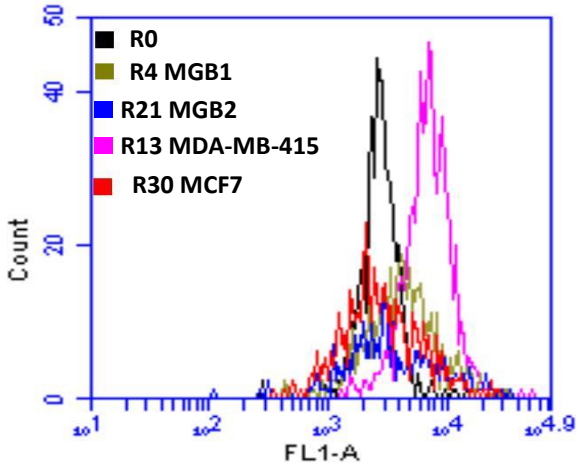
A: Cells only (Background)



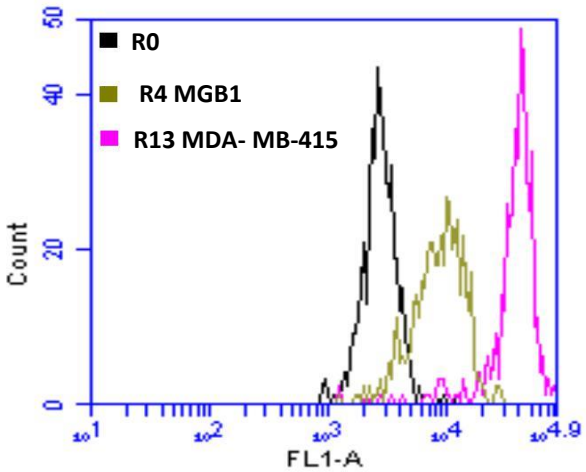
B: HCAEC (Counter cells)



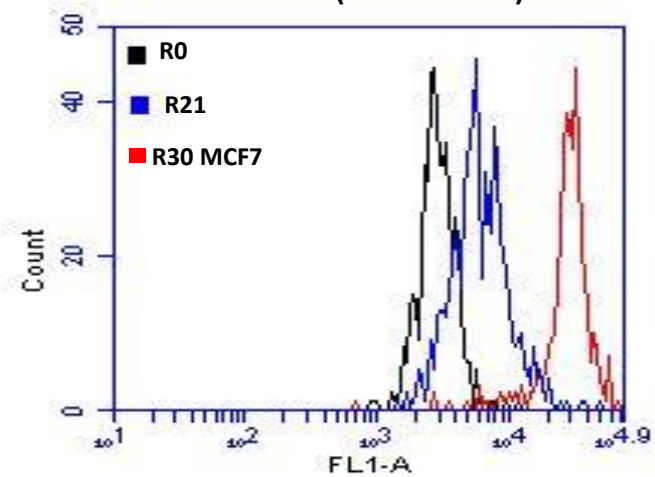
C: MCF10A (Counter cells)

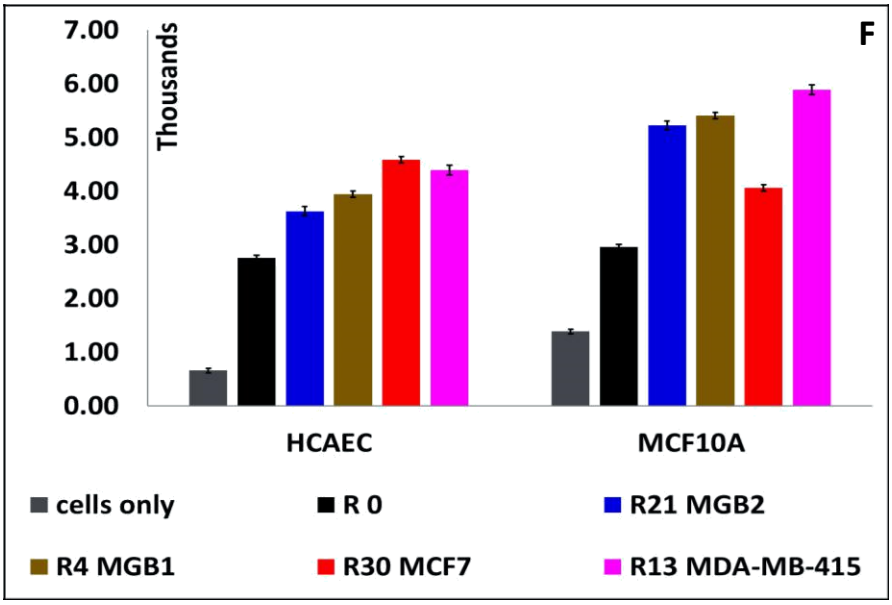


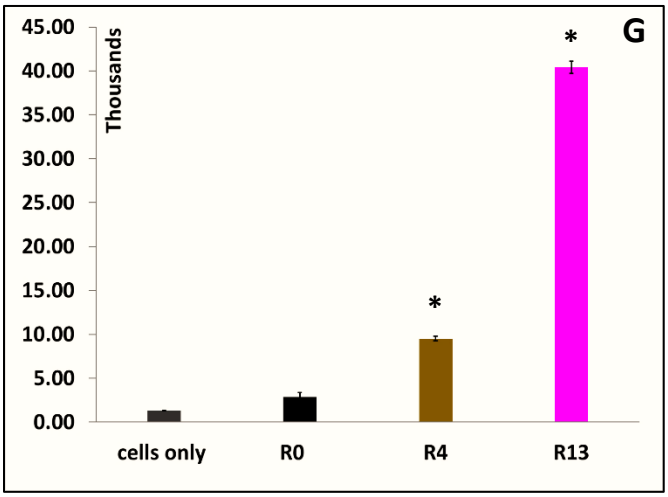
D: MDA-MB-415 (Positive cells)



E: MCF7 (Positive cells)







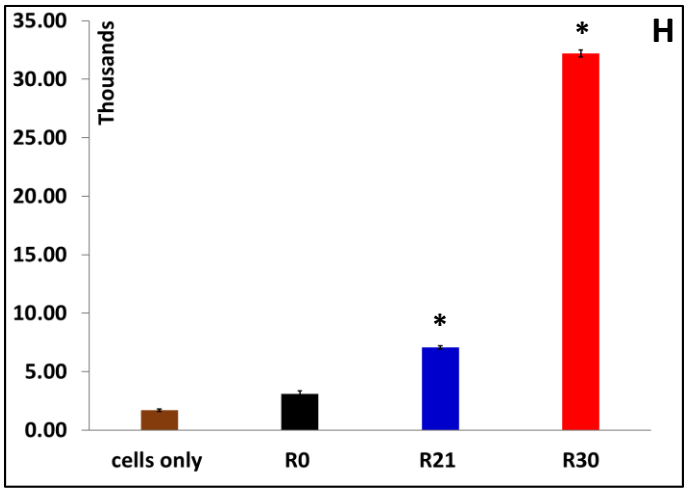


Figure 2.5: Flow cytometry monitoring results of R0, R21 MGB2, R4 MGB1, R30 MCF7 and R13 MDA-MB-415 against cell SELEX cell lines. A: fluorescence intensity of all cell lines alone without any treatment (cells only) B: fluorescence intensity of HCAEC cells (counter selection cells) when binding to all pools from different rounds. C: fluorescence intensity of MCF10A cells (counter selection cells) when bound to all pools from different rounds. D: fluorescence intensity of MDA-MB-415 cells (positive selection cells) when bound to R0, R4, and R13 of MGB1. E: fluorescence intensity of MCF7 cells (positive section cells) when bound to R0, R21, and R30 of MGB2. F-H: Histograms with the exact values of fluorescence intensities (from A-E) of all pools tested against all cell lines. Error bars represent the S.E.M. of three trials. The statistical significance was determined by one-way ANOVA, followed by the Fisher LSD test using SPSS software (SPSS, version 23). The p values less than 0.05 were considered to be significant.

2.3.2 Computational and bioinformatics analysis using AptaCluster software

Several DNA pools from both selections were sent to HTS. AptaCluster software was obtained from collaborators (Teresa M. Przytycka, and Jan Hoinka) at the National Institute of Health (Washington, D.C., USA). AptaCluster grouped both MGB2 and MGB1 target datasets into clusters, according to their sequence similarity (in our case, sequences which had 6 or less base differences were grouped in the same cluster). In each cluster, sequences were organized according to their count (frequency). The sequence with the highest count in each cluster is called a seed sequence. AptaCluster allows for the screening of all sequences (from all the different clusters) of analyzed SELEX rounds according to their count or enrichment, thus making it easier to organize all the sequences to choose from by exporting them into a separate file.

The bioinformatics analysis by AptaCluster revealed information about many aspects of the SELEX process by providing specific data for each round analyzed. The pool size (total number of sequences) in each selection round (the analyzed rounds) for both MGB2 and MGB1 selections was found to be at least 10 million for each target. The random region size analysis for MGB2 showed that the majority of the sequences in the last round had 25 bases in their randomized region, compared to 24 bases in the starting pool (Figure 2.6). This one base increase was a result of a mutation during the SELEX process, as a result of PCR. The random region size of the MGB1 pool was 24 bases. Many studies have shown the generation of new sequences as a result of mutation during the SELEX process [56-58]. These mutations could increase the binding affinity of the aptamers to their targets and can show vital structural information regarding the effects of specific nucleotides on binding affinity [56]. In our case, the increase in the random region size for MGB2-binding sequences might be a great benefit towards the binding affinity of MGB2 aptamers to their target.

The bioinformatics analysis also showed the distribution of the four bases in the sequences generated for each target. For MGB2, the starting library had almost an equal random distribution of the four bases (20-27%, Table 2.1).

This distribution noticeably changed during the selection process to favour more “C”s and ”G”s (Table 2.1). Conversely, the MGB1 starting library had noticeably higher percentage of the bases “C” and “G” regards to the other bases (28-30%). The distribution changed toward “T” and “G” bases in the final round (Table 2.1). The non-even base distribution of the MGB1 starting library was expected since the starting library was obtained from R17 of the MGB2 pool, and not synthesized as a separate pool. It has been reported that a good starting pool (ssDNA library) should have a random region consisting of an equal distribution of all the four bases [59]. Therefore, our starting library (pool) distribution matched the literature and matched our expectations. Moreover, the high percentage of “G” bases in the final rounds (as obtained from AptaCluster analysis) of both targets is a normal output and it might indicate the presence of aptamers that have G-quadruplex structures, considered to be the most common ssDNA structures identified by SELEX [60].

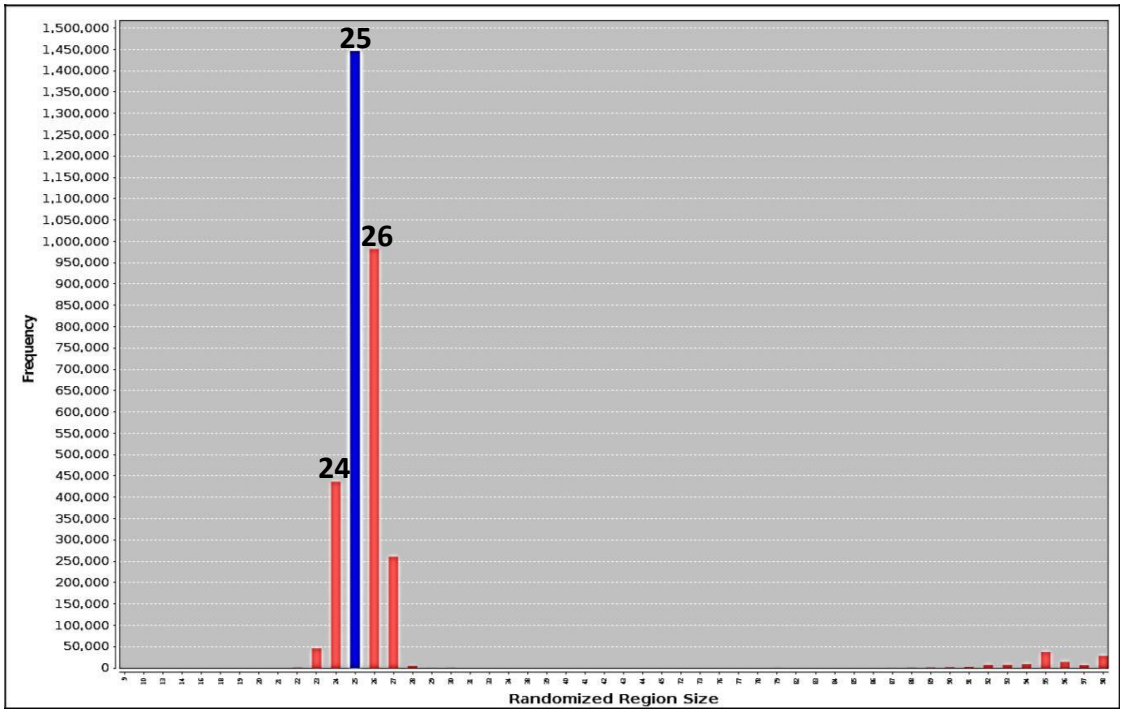


Figure 2.6: Analysis of the random region size of MGB2 last round pool using AptaCluster. Last round of MCF7 (MGB2) showed the dominant size of random region is 25 bases (blue) and other sizes 24 and 26 (red).

Table 2.1: Distribution of the four bases in the random region of the starting library (R0) and the final pools (R30, R13) for MGB2 and MGB1 targets as obtained from bioinformatics analysis using AptaCluster software.

Base distribution (%)	A	T	C	G
MGB2				
Initial round (R0)	20.1	26.4	27.6	26.0
Final round (R30)	22.1	17.7	30.7	29.6
MGB1				
Initial round (R0)	22.8	19.7	28.2	29.3
Final round (R13)	17.6	21.5	26.6	34.3

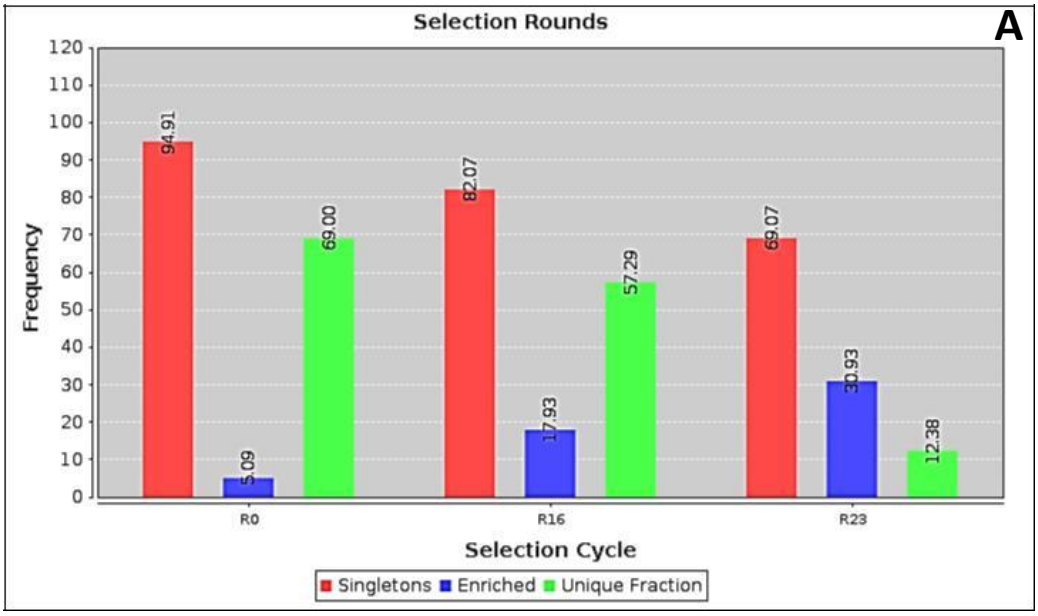
Moreover, the HiSeq data analysis showed that the percentage of the unique sequences (expressed here by singleton) for both targets decreased in the last round of SELEX compared to R0 (Figure 2.7 A and B (red columns)). If the pool size was taken into account, corresponding to the unique sequences in every round, the result would be a unique fraction which also decreased for specific nucleotides as the SELEX progressed for both targets (69.00 to 12.38 in MGB2 and 54.11 to 7.34 for MGB1, Figure 2.7 A and B (green columns). The reason behind the decline in either the percentage of unique sequences or the unique fraction is that the starting libraries for both targets were not enriched for a specific target. Thus, they contained a large number of unique sequences that decreased as the SELEX advanced, reaching the lowest limit in the final round when the library was more enriched for specific targets. Another important finding the bioinformatics analysis revealed was that the ssDNA libraries of both targets were enriched as the SELEX advanced; for MGB2 the enrichment values increased from 5.09 to 30.93 (Figure 2.7A), and for MBG1 from 20.61 to 27.12 (Figure 2.7B). Once again, the starting enrichment value of MGB1 pool was higher compared to MGB2 pool (as indicated by the values shown in Figure 2.7). This was due to the fact that the starting pool of MGB1 was already enriched (21 compared to 5 for MGB2) as MGB1 was used as a counter selection in MGB2 protein SELEX.

In addition, AptaCluster analysis showed the number of copies of sequences (count) in each round. It was assumed that the number of copies for a certain sequence was a reflection of higher binding affinity, and that it would indicate that the SELEX process was free of artifacts as the high count sequences were present in the initial pool as well [51]. Artifacts are revealed as a large fraction of sequences in the final pool being absent from the initial pool, indicating that they were introduced in a later stage of the SELEX process [51]. This introduction could be the result of artifacts introduced or the enrichment of a high affinity sequence, if the sequence was found in the initial pool and enriched as the SELEX advanced. The only way to verify this is by testing the highest count sequences and this was applied to our SELEX.

It was clear that changing the conditions that were applied to the SELEX of both MGB2 and MGB1 (number and volume of washes, pool and target concentrations, and other conditions as mentioned earlier in the method section of this chapter) increased stringency, leading to increased enrichment of common sequences, and decreased enrichment of unique sequences as the SELEX progressed. Moreover, applying the counter selection in every round, simultaneously with the positive selection, facilitated the elimination of the sequences that could bind to a structurally similar target (in our case MGB1 in the protein SELEX and MCF10A in the cell SELEX), thus further enhancing the selectivity of the aptamers. Another important factor, which led to increased enrichment of the ssDNA libraries, is the number of selection rounds. In our case, the enrichment was increasing with increasing number of rounds, as indicated by the percentage binding of ssDNA libraries to both targets (Figure2.3 and Figure 2.5). This indicated that a high number of selection rounds was vital to fully enrich the ssDNA libraries for sequences with desired affinity.

Upon the completion of HiSeq data analysis, 12 sequences for MGB2 aptamers and 8 sequences for MGB1 aptamers were chosen for further investigation (binding and affinity studies) from the final rounds of SELEX. Aptamers were chosen according to their count (seed sequences) and enrichment values. Aptamers with the highest count in all clusters were chosen (MAMB0 (34108 copies), MAMA1 (223690 copies)), as well as the seed sequences from candidate clusters for both targets. Aptamers with higher enrichment values in positive selection rounds compared to thethe counter selections of the same round (e.g. MAMB6, MAMA5) were also chosen.

To investigate the role of enrichment values provided by AptaCluster in the binding of these aptamers to their targets, some of the aptamers that were chosen have the same or less enrichment than the counter-selection pools (e.g.MAMB1, MAMB4, MAMA2). Table 2.2 shows the exact sequences for both targets and their enrichment values.



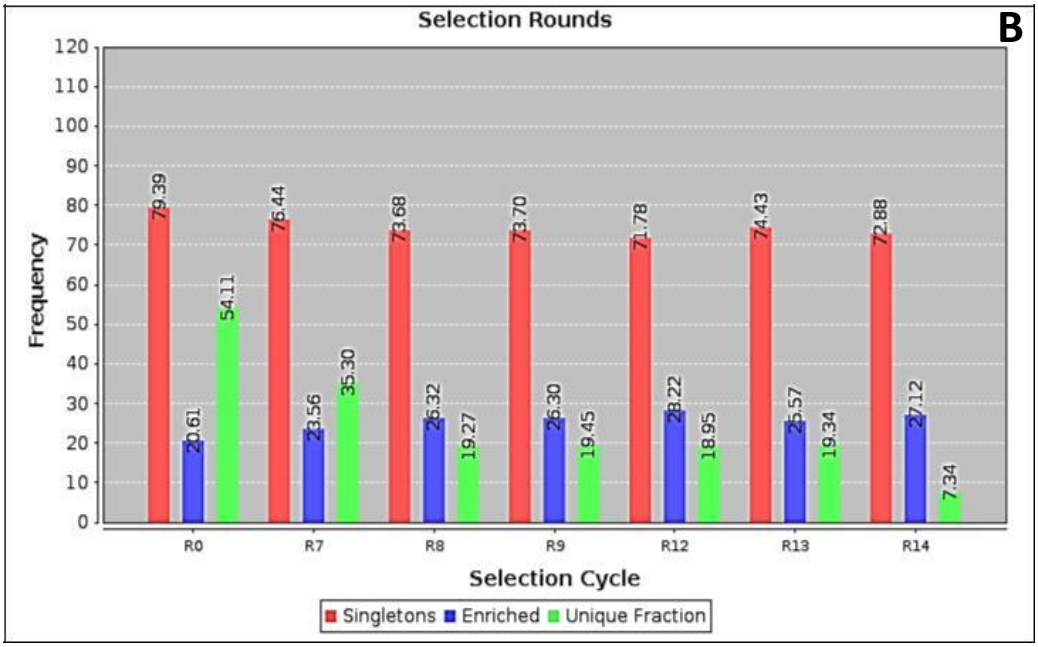


Figure 2.7: The experimental details of both A: MGB2 and B: MGB1 as shown by AptaCluster software analysis. The unique sequences decreased in the selections for both targets and the unique fraction is shown. The percentage of the unique fraction is represented by singleton, which also decreased for both targets. The enrichment values, however, increased during the selection for both targets as well.

Table 2.2: MGB2 and MGB1 sequences and enrichment values

Name/ cluster	Aptamer sequence (5'-3') 5' 20 bases forward primer (N 25 or 24) 20 bases reverse primer'3	Enrichment of Positive /Counter	
MAMB0	CATGCTTACCTATAGTGAACCCCAACACACGTGTAGATCCTGCGGCTTTGAGAACTGACTCATAC	241.990	175.373
MAMB1	CATGCTTACCTATAGTGAACCCAACGTCGAACTGAATCCCGTGTCCTTTGAGAACTGACTCATAC	122.649	123.569
MAMB4	CATGCTTACCTATAGTGAACGCGGCATGTTGGCATCTTGGTCCTGCTTTGAGAACTGACTCATAC	11.636	11.079
MAMB6	CATGCTTACCTATAGTGAACCCACACGATATGGCGCTACACCGTGCTTTGAGAACTGACTCATAC	143.579	96.389
MAMB8	CATGCTTACCTATAGTGAACCACGACACGCGCGATCGTCTCACTGCTTTGAGAACTGACTCATAC	60.063	51.133
MAMB12	CATGCTTACCTATAGTGAACCCACCACACAGCGGATACACCATGGCTTTGAGAACTGACTCATAC	920.639	680.997
MAMB19	CATGCTTACCTATAGTGAACCCACCACGATCGATCGTACCACGTCCTTTGAGAACTGACTCATAC	157.224	121.557
MAMB21	CATGCTTACCTATAGTGAACCGTGGCGTGATGGACGTGGGGATGGCTTTGAGAACTGACTCATAC	68.241	40.644
MAMB29	CATGCTTACCTATAGTGAACCCCCAACGACACCGGATTGCCCTTGCTTTGAGAACTGACTCATAC	100.34	83.660
MAMB50	CATGCTTACCTATAGTGAACCCCGTAGCGATGACCGATCGATGTGCTTTGAGAACTGACTCATAC	162.864	125.615
MAMB57	CATGCTTACCTATAGTGAACCCGAAGAGGATGTGCGGTCCCATTGCTTTGAGAACTGACTCATAC	80.200	61.521
MAMB198	CATGCTTACCTATAGTGAACCCGGGGGATGTGGACAGAACGTGCGCTTTGAGAACTGACTCATAC	305.155	182.854
MAMA1	CATGCTTACCTATAGTGAACGCGGCATGTTGGCATCTTGGTCCTCTTTGAGAACTGACTCATAC	25.80	15.93
MAMA2	CATGCTTACCTATAGTGAACCCGGGACAGAACGTGCGCTTTGAGCTTTGAGAACTGACTCATAC	5836.12	5277.48
MAMA4	CATGCTTACCTATAGTGAACGCGGCATGTTGGCCATCTTGGTCCCTTTGAGAACTGACTCATAC	1833.64	690.30
MAMA5	CATGCTTACCTATAGTGAACGGTTGGCATCTTGGTCCTGCTTTGCTTTGAGAACTGACTCATAC	1666.05	236.97

MAMA6	CATGCTTACCTATAGTGAACTGTTGGCATCTTGGTCCTGCTTTGCTTTGAGAACTGACTCATAC	1242.12	191.94
MAMA7	CATGCTTACCTATAGTGAACGCGCATCTTGGTCCTGCTTTGAGACTTTGAGAACTGACTCATAC	1084.41	180.49
MAMA12	CATGCTTACCTATAGTGAACGTTGGCATCTTGGTCCTGCTTTGACTTTGAGAACTGACTCATAC	1409.74	242.50
MAMA13	CATGCTTACCTATAGTGAACGCGGCATGCATCTTGGTCCTGCTTCTTTGAGAACTGACTCATAC	1330.87	286.83

2.3.3 Secondary structure prediction of the aptamer candidates

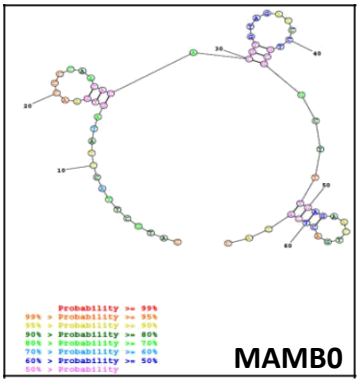
The structural analysis of many aptamers show that they form pockets when interacting with their targets by folding into distinct 3D structures (including stems and loops) [61]. The folding occurs only when they are interacting with their targets, therefore they are unstructured when they are free in solution [61,62]. The knowledge of the aptamer structure is important to know the binding mechanism and therefore might affect the binding affinity.

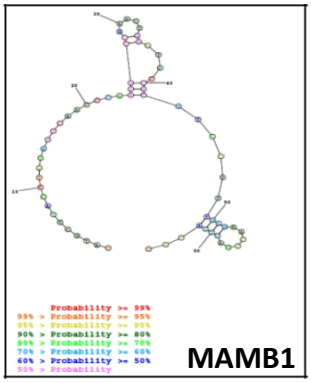
In our results, the secondary structures of the chosen aptamers from the selections for MGB2 and MGB1 (12 for MGB2 and 8 for MGB1) were predicted using RNAstructure software. Figure 2.8 showed most of MGB2 and MGB1-binding sequences had simple structures (e.g. MAMB6, MAMB19) and some of them had more complex structures (e.g. MAMB 198, MAMA 13). The structure of these aptamers, however, may become more complex when interacting with their targets. Moreover, some of the software used to predict the secondary structure of the aptamers does not recognize special structures such as pseudoknots and G- quadruplexes that could be found in our selected aptamers.

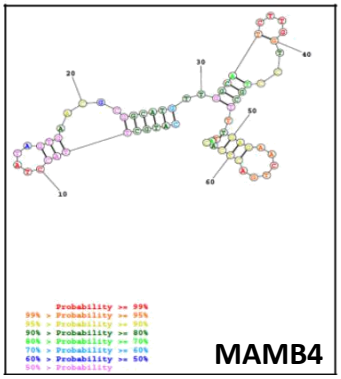
It has been reported that most of the existing aptamers generated from SELEX have simple structures with low degrees of complexity and that highly complex structures are less abundant [63,64]. For example, aptamers that bind to adenosine triphosphate (ATP) and streptomycin have simple structures (linear stem loops) [65,66]. The reason behind this is that the starting ssDNA libraries have limited structural diversity (due to the fact that DNA has only

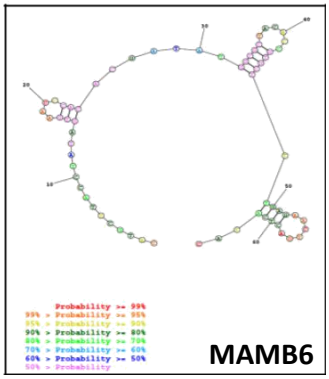
four bases (A,T,G,C)) [63]. Conversely, it has been shown that the structural complexity of a sequence depends on the length of that sequence. As the length increases, there is a higher probability of having structures that are more complex [67]. However, applying this to the starting libraries of SELEX is a challenge due to the lack of efficiency when synthesizing more than 100 nucleotides [68]. In addition, a conserved region (primer binding site) of 15-20 bases must be included in the length of the sequence (on the 3' and 5' ends of the sequence). These sequences may participate in the binding of the aptamer to its target, but they also reduce the variable region size to 70 bases as a maximum number to generate new structures.

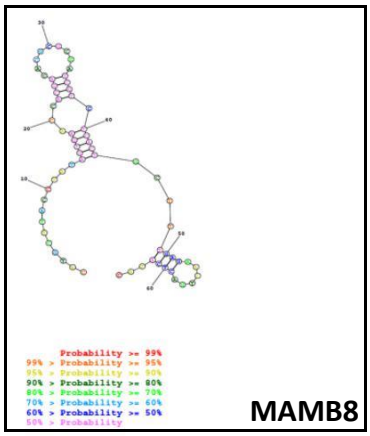
In this context, the predicted secondary structures of our selected aptamers are simple, as are most of the selected aptamer sequences found in the literature. Moreover, the length of our aptamers was 65 bases for MGB2, and 64 bases for MGB1 (as showed by Table 2.2), which included the primer binding sites. Our starting library consisted of only 24 bases of variable region and therefore the selected aptamers had simple predicted structures. The reason behind choosing a short variable region, in our case, was the size of the targets. The goal of this thesis was to develop aptamers against two proteins, MGB2 and MGB1, using the recombinant human protein. The other targets in the hybrid SELEX were cells. Since proteins and cells are big in size, the binding of the aptamers to these proteins/cells may not require complex aptamer structure. This is unlike the binding of aptamers to small molecules, which requires aptamers with more complex structures [69].

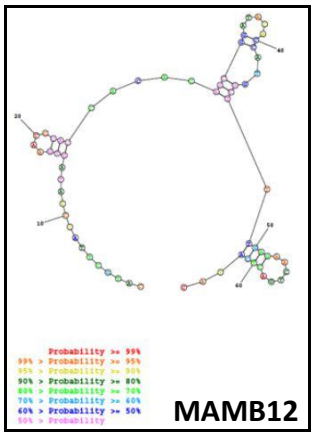

MAMB0

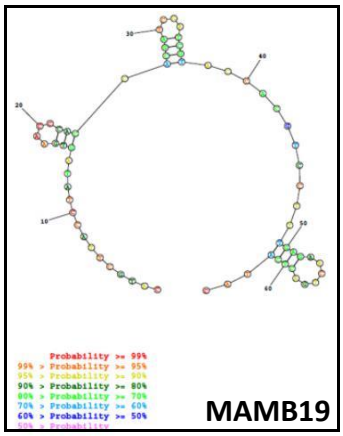

MAMB1

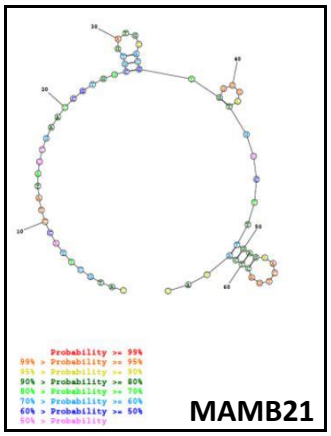

MAMB4

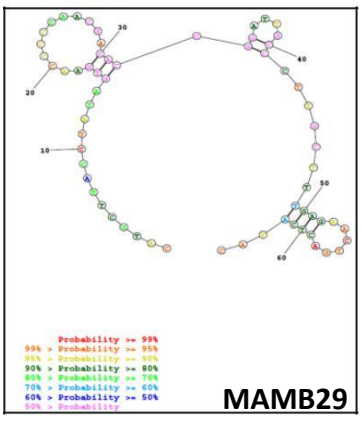

MAMB6

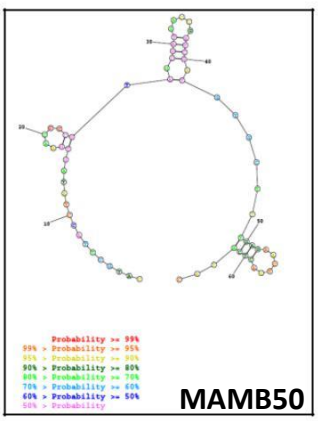

MAMB8

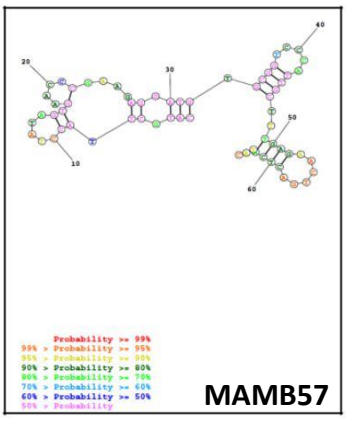

MAMB12

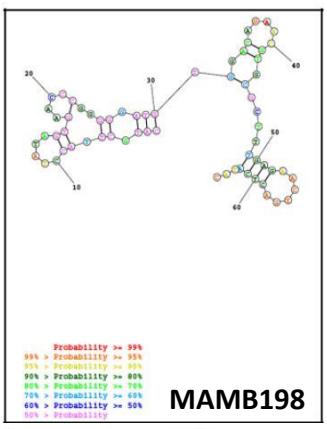

MAMB19

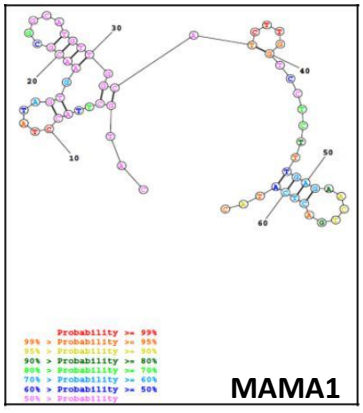

MAMB21

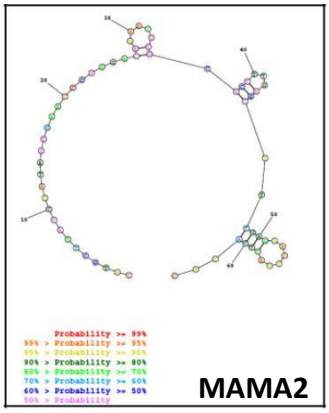

MAMB29

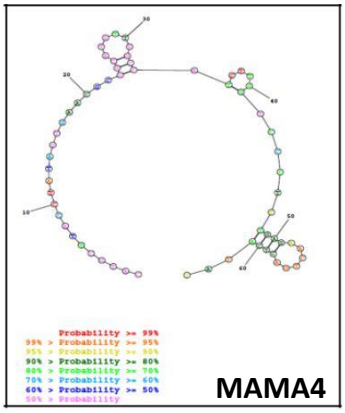

MAMB50

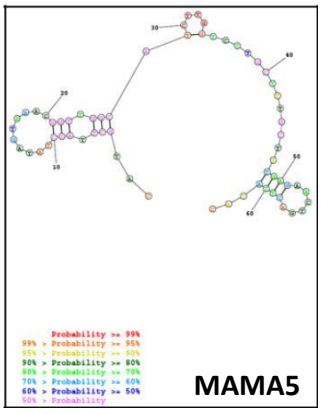

MAMB57


MAMB198


MAMA1


MAMA2


MAMA4


MAMA5

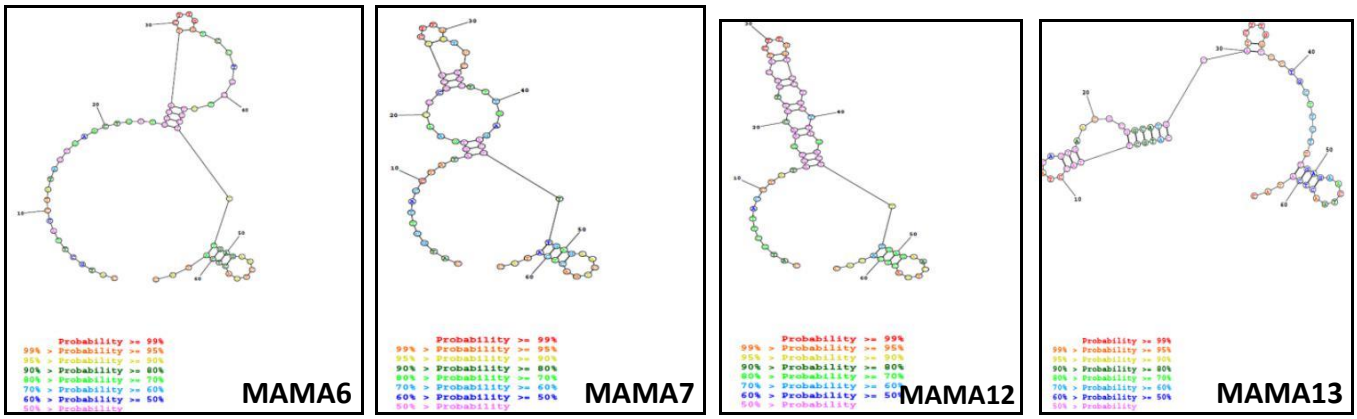


Figure 2.8 : The predicted secondary structure of the chosen aptamers of MGB2 and MGB1 using RNAstructure software.

2.3.4 Screening of MGB2 and MGB1 aptamers

The goal of this test was to perform initial screening (assessment) on the selected aptamers (12 for MGB2 and 8 for MGB1) to have a rapid yes or no answer to further test them for their affinity and selectivity.

The binding of the chosen aptamer candidates to both targets was initially screened against the cell lines used in cell-SELEX, using flow cytometry. The screening results are shown in Figure 2.9, where the mean of fluorescence intensity represented the binding of the aptamer-cell complex. The highest mean of fluorescence intensity of the aptamers binding to the target cell lines, compared to the counter selection cell lines (MCF10A and HCAEC), was 15647 for MAMA2 binding to MDA-MB-415 cells (Figure 2.9B). MAMB1 showed similar binding to MCF7 cells (10586, Figure 2.9A). MAMB1 and MAMB12, MAMA2, MAMA5, MAMA6, MAMA7, and MAMA12 showed similar results (Figure 2.9 A and B). MAMB0, MAMA4, and MAMB57 showed low fluorescence intensity values (ranging from 3000 to 1600) binding to their target cancer cells MCF7 but still had higher binding to their target than to the counter selection cell lines (Figure 2.9A). MAMA1, MAMA4, and MAMA13 showed non-specific binding to MCF10A cells (as indicating by their fluorescence intensity values), but this binding was still lower than that of these aptamers to MDA-MB-415 cells (Figure 2.9B).

According to the results mentioned above, six aptamers out of twelve for MGB2 or MCF7 binding aptamers (MAMB0, MAMB1, MAMB4, MAMB8, MAMB12, and MAMB57) were chosen to further investigate their affinity and selectivity. Four aptamers out of eight (MAMA2, MAMA5, MAMA6, and MAMA12) were chosen for MGB1 or MDA-MB-415 for further investigation of their affinity and selectivity.

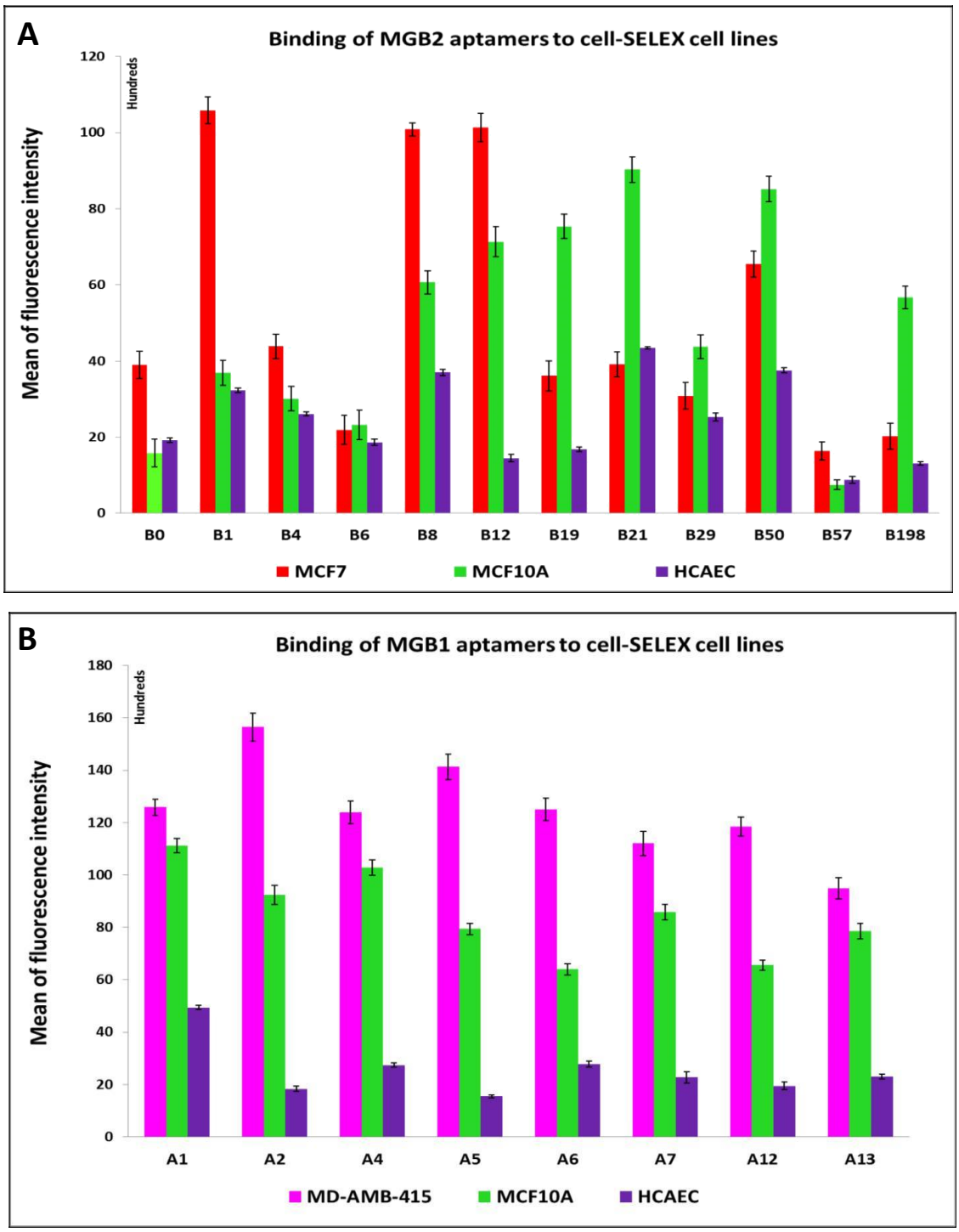


Figure 2.8: Screening results of the aptamer candidates against cell-SELEX cell lines. A: Screening results of MGB2 aptamers showed six aptamers out of twelve (MAMB0, MAMB1, MAMB4, MAMB8, MAMB12, MAMB57) with higher fluorescence intensity in the positive selection cell lines than counter selection ones. B: Screening results of MGB1 aptamers showed high non-specific binding to the MCF10A cell line. However, some aptamers showed less non-specific binding (MAMA2, MAMA5, MAMA6, and MAMA12). Promising aptamers were chosen for further studies. The experiments were repeated twice. Results shown were means ± the standard error of the means (SEM).

2.4 Conclusion

The first in vitro selections (SELEX) against MGB2 and MGB1 proteins were reported here. MGB2 and MGB1 aptamers were selected successfully from a random DNA library containing 24 nucleotides as a random (variable) region, using hybrid SELEX as the main method of selection. Hybrid SELEX targeted specific proteins on the surface of the cell, in our case MGB2 on the surface of MCF7 cells and MGB1 on the surface of MDA-MB-415 cells. The percentage recovery of DNA for the last pools (R30 and R13) of the SELEX showed significant ($P<0.05$) and specific binding to target cancer cells lines (MCF7 and MDA-MB-415) but not to counter selection cell lines (HCAEC and MCF10A; Figure 2.5).

The HTS sequencing technique, unlike the conventional cloning, opened the door to explore and monitor the dynamics that took place every round in the SELEX process (such as total number of sequences, count (frequency) of sequences, percentage of unique sequences, the distribution of the four bases, enrichment, and background binders) [70]. This allows the researcher to change the conditions of the SELEX to serve their needs, which was not possible before. In our study, HTS sequencing (using Illumina HiSeq as a platform and AptaCluster as software) was done on different chosen pools of both aptamers. Data analysis was performed on both targets to get aptamers arranged in different clusters with high enrichment and lower unique sequences (Figure 2.7). Several aptamers were selected from many of the clusters according to their count and enrichment (Table 2.2).

The aptamer candidates were analyzed to predict their secondary structures, and were then screened against their target cancer cell lines, as well as counter selection cell lines (Figure 2.9). To pass the screening test, a good aptamer should bind to the target cancer cell line, with minimal non-specific binding to the counter cell line. MAMB0, MAMB1, MAMB4, MAMB8, MAMB12, MAMB57, MAMA2, MAMA5, MAMA6, and MAMA12 aptamers were chosen to test their affinity towards their target cancer cell lines (K_d test) and their specificity by testing the binding of the high affinity aptamers to other cancer and normal cells.

References

1. Becker RM, Darrow C, Zimonjic DB, Popescu NC, Watson MA, Fleming TP. Identification of mammaglobin B, a novel member of the uteroglobin gene family. Genomics. 1998, 54(1):70-78.
2. Lehrer RI, Xu G, Abduragimov A, Dinh NN, Qu XD, Martin D, Glasgow BJ. Lipophilin, a novel heterodimeric protein of human tears. FEBS Lett. 1998, 432(3):163-167.
3. Watson MA, Fleming TP. Mammaglobin, a mammary-specific member of the uteroglobin gene family, is overexpressed in human breast cancer. Cancer Res. 1996, 56(4):860-865.
4. Colpitts TL, Billing-Medel P, Friedman P, Granados EN, Hayden M, Hodges S *et al*. Mammaglobin is found in breast tissue as a complex with BU101. Biochemistry. 2001, 40(37):11048-11059.
5. Klug J, Beier HM, Bernard A, Chilton BS, Fleming TP, Lehrer RI *et al*. Uteroglobin/Clara cell 10-kDa family of proteins: nomenclature committee report. Ann N Y Acad Sci. 2000, 923:348-354.
6. Tassi RA, Bignotti E, Falchetti M, Calza S, Ravaggi A, Rossi E, *et al*. Mammaglobin B expression in human endometrial cancer. Int J Gynecol Cancer. 2008, 18(5):1090-1096.
7. Jackson BC, Thompson DC, Wright MW, McAndrews M, Bernard A, Nebert DW, Vasiliou V. Update of the human secretoglobin (SCGB) gene superfamily and an example of 'evolutionary bloom' of androgen-binding protein genes within the mouse Scgb gene superfamily. Hum Genomics. 2011, 5(6):691-702.
8. Lacroix M. Significance, detection and markers of disseminated breast cancer cells. Endocr Relat Cancer. 2006, 13(4):1033-1067.
9. Carter D, Douglass JF, Cornellison CD, Retter MW, Johnson JC, Bennington AA, *et al*. Purification and characterization of the mammaglobin/lipophilin B complex, a promising diagnostic marker for breast cancer. Biochemistry. 2002, 41(21):6714-6722.
10. Brown NM, Stenzel TT, Friedman PN, Henslee J, Huper G, Marks JR. Evaluation of expression based markers for the detection of breast cancer cells. Breast Cancer Res Treat. 2006, 97(1):41-47.

11. Zehentner BK, Carter D. Mammaglobin: a candidate diagnostic marker for breast cancer. Clin Biochem. 2004, 37(4):249-257.

12. Zuo L, Li L, Wang Q, Fleming TP, You S. Mammaglobin as a potential molecular target for breast cancer drug delivery. Cancer Cell Int. 2009, 9:8.

13. Zhao C, Nguyen T, Yusifov T, Glasgow BJ, Lehrer RI. Lipophilins: human peptides homologous to rat prostatein. Biochem Biophys Res Commun. 1999, 256(1):147-155.

14. Watson MA, Dintzis S, Darrow CM, Voss LE, DiPersio J, Jensen R, Fleming TP. Mammaglobin expression in primary, metastatic, and occult breast cancer. Cancer Res. 1999, 59(13):3028-3031.

15. Aihara T, Fujiwara Y, Ooka M, Sakita I, Tamaki Y, Monden M. Mammaglobin B as a novel marker for detection of breast cancer micrometastases in axillary lymph nodes by reverse transcription-polymerase chain reaction. Breast Cancer Res Treat. 1999, 58(2):137-140.

16. Bellone S, Tassi R, Betti M, English D, Cocco E, Gasparrini S, *et al*. Mammaglobin B (SCGB2A1) is a novel tumour antigen highly differentially expressed in all major histological types of ovarian cancer: implications for ovarian cancer immunotherapy. Br J Cancer. 2013, 109(2):462-471.

17. Tassi RA, Calza S, Ravaggi A, Bignotti E, Odicino FE, Tognon G, *et al*. Mammaglobin B is an independent prognostic marker in epithelial ovarian cancer and its expression is associated with reduced risk of disease recurrence. BMC Cancer. 2009, 9:253.

18. Tassi RA, Bignotti E, Rossi E, Falchetti M, Donzelli C, Calza S *et al*. Overexpression of mammaglobin B in epithelial ovarian carcinomas. Gynecol Oncol. 2007, 105(3):578-585.

19. Tassi RA, Bignotti E, Falchetti M, Calza S, Ravaggi A, Rossi E *et al*. Mammaglobin B expression in human endometrial cancer. Int J Gynecol Cancer. 2008, 18(5):1090-1096.

20. Mercatali L, Valenti V, Calistri D, Calpona S, Rosti G, Folli S, *et al*. RT-PCR determination of maspin and mammaglobin B in peripheral blood of healthy donors and breast cancer patients. Ann Oncol. 2006, 17(3):424-428.

21. Jiang Y, Harlocker SL, Molesh DA, Dillon DC, Stolk JA, Houghton RL, *et al*. Discovery of differentially expressed genes in human breast cancer using subtracted cDNA libraries and cDNA microarrays. Oncogene. 2002, 21(14):2270-2282.

22. Janku F, Kleibl Z, Novotny J, Tesarova P, Petruzelka L, Matous B. Mammaglobin A, a novel marker of minimal residual disease in early stages breast cancer. Neoplasma. 2004, 51(3):204-208.

23. Ouellette RJ, Richard D, Maïcas E. RT-PCR for mammaglobin genes, MGB1 and MGB2, identifies breast cancer micrometastases in sentinel lymph nodes. Am J Clin Pathol. 2004, 121(5):637-643.

24. Ooka M, Sakita I, Fujiwara Y, Tamaki Y, Yamamoto H, Aihara T, *et al*. Selection of mRNA markers for detection of lymph node micrometastases in breast cancer patients. Oncol Rep. 2000, 7(3):561-566.

25. Aristizábal-Pachón AF, de Carvalho TI, Carrara HH, de Andrade JM, Takahashi CS. Detection of human mammaglobin A mRNA in peripheral blood of breast cancer patients before treatment and association with metastasis. J Egypt Natl Canc Inst. 2015, 27(4):217-222.

26. Ceballos MP, Zumoffen C, Massa E, Cipulli G, Funes CC, Gil AB, *et al*. Detection of mammaglogin A in blood from breast cancer patients, before and after treatment, using a one-tube nested PCR protocol. Association with the absence of tumor estrogen receptors. Clin Biochem. 2011, 44(17-18):1429-1433.

27. Ballestrero A, Garuti A, Bertolotto M, Rocco I, Boy D, Nencioni A *et al*. Effect of different cytokines on mammaglobin and maspin gene expression in normal leukocytes: possible relevance to the assays for the detection of micrometastatic breast cancer. Br J Cancer. 2005, 92(10):1948-1952.

28. Zach O, Kasparu H, Krieger O, Hehenwarter W, Girschikofsky M, Lutz D. Detection of circulating mammary carcinoma cells in the peripheral blood of breast cancer patients via a nested reverse transcriptase polymerase chain reaction assay for mammaglobin mRNA. J Clin Oncol. 1999, 17(7):2015-2019.

29. Grünewald K, Haun M, Urbanek M, Fiegl M, Müller-Holzner E, Gunsilius E, *et al*. Mammaglobin gene expression: a superior marker of breast cancer cells in peripheral blood in comparison to epidermal-growth-factor receptor and cytokeratin-19. Lab Invest. 2000, 80(7):1071-1077.

30. Ghersevich S, Ceballos MP. Mammaglobin A: review and clinical utility. Adv Clin Chem. 2014, 64:241-268.

31. Al Joudi FS. Human mammaglobin in breast cancer: a brief review of its clinical utility. Indian J Med Res. 2014, 139(5):675-685.

32. Kim SW, Goedegebuure P, Gillanders WE. Mammaglobin-A is a target for breast cancer vaccination. Oncoimmunology. 2016, 5(2):e1069940.

33. Li S, Xu H, Ding H, Huang Y, Cao X, Yang G, *et al*. Identification of an aptamer targeting hnRNP A1 by tissue slide-based SELEX. J Pathol. 2009, 218(3):327-336.

34. Hicke BJ, Marion C, Chang YF, Gould T, Lynott CK, Parma D, *et al*. Tenascin-C aptamers are generated using tumor cells and purified protein. J Biol Chem. 2001, 276(52):48644-48654.

35. Zumrut HE, Ara MN, Fraile M, Maio G, Mallikaratchy P. Ligand-guided selection of target-specific aptamers: A screening technology for identifying specific aptamers against cell-surface proteins. Nucleic Acid Ther. 2016, 26(3):190-198.

36. Nimjee SM, Rusconi CP, Sullenger BA. Aptamers: an emerging class of therapeutics. Annu Rev Med. 2005, 56:555-583.

37. Ray P, Sullenger BA, White RR. Further characterization of the target of a potential aptamer biomarker for pancreatic cancer: cyclophilin B and its posttranslational modifications. Nucleic Acid Ther. 2013, 23(6):435-442.

38. Sefah K, Shangguan D, Xiong X, O'Donoghue MB, Tan W. Development of DNA aptamers using Cell-SELEX. Nat Protoc. 2010, 5(6):1169-1185.

39. Larry H. Bernstein. Nanotechnology: aptamers for specific & better delivery systems of existing drugs. Molecular Therapy Nucleic Acids, 2014 3: e182.

40. Shigdar S, Qiao L, Zhou SF, Xiang D, Wang T, Li Y, *et al*. RNA aptamers targeting cancer stem cell marker CD133. Cancer Lett. 2013, 330(1):84-95.

41. Pestourie C, Cerchia L, Gombert K, Aissouni Y, Boulay J, De Franciscis V, *et al*. Comparison of different strategies to select aptamers against a transmembrane protein target. Oligonucleotides. 2006, 16(4):323-335.

42. Parekh P, Kamble S, Zhao N, Zeng Z, Portier BP, Zu Y. Immunotherapy of CD30-expressing lymphoma using a highly stable ssDNA aptamer. Biomaterials. 2013, 34(35):8909-89017.

43. Reuter JA, Spacek DV, Snyder MP. High-throughput sequencing technologies. Mol Cell. 2015, 58(4):586-597.

44. Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, *et al*. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. BMC Genomics. 2012, 13:341.

45. Minoche AE, Dohm JC, Himmelbauer H. Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and genome analyzer systems. Genome Biol. 2011, 12(11):R112.

46. Kircher M, Kelso J. High-throughput DNA sequencing--concepts and limitations. Bioessays. 2010, 32(6):524-536.

47. Dohm JC, Lottaz C, Borodina T, Himmelbauer H. Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. Nucleic Acids Res. 2008, 36(16):e105.

48. Mielczarek M, Szyda J. Review of alignment and SNP calling algorithms for next-generation sequencing data. J Appl Genet. 2016, 57(1):71-79.

49. Caboche S, Audebert C, Lemoine Y, Hot D. Comparison of mapping algorithms used in high-throughput sequencing: application to Ion Torrent data. BMC Genomics. 2014, 15:264.

50. Mane SP, Modise T, Sobral BW. Analysis of high-throughput sequencing data. Methods Mol Biol. 2011, 678:1-11.

51. Hoinka J, Berezhnoy A, Sauna ZE, Gilboa E, Przytycka TM. AptaCluster-a method to cluster HT-SELEX aptamer pools and lessons from itsaApplication. Res Comput Mol Biol. 2014, 8394:115-128.

52. Hoinka J, Berezhnoy A, Dao P, Sauna ZE, Gilboa E, Przytycka TM. Large scale analysis of the mutational landscape in HT-SELEX improves aptamer discovery. Nucleic Acids Res. 2015, 43(12):5699-5707.

53. Hoinka J, Dao P, Przytycka TM. AptaGUI-A graphical user interface for the efficient analysis of HT-SELEX data. Mol Ther Nucleic Acids. 2015, 4:e257.

54. Kosuri S, Church GM. Large-scale de novo DNA synthesis: technologies and applications. Nat Methods. 2014 May, 11(5):499-507.

55. Sambrook J, Russell DW. Purification of nucleic acids by extraction with phenol: chloroform. CSH Protoc. 2006, 2006(1). pii: pdb.prot4455.

56. Levay A, Brenneman R, Hoinka J, Sant D, Cardone M, Trinchieri G, *et al*. Identifying high-affinity aptamer ligands with defined cross-reactivity using high-throughput guided systematic evolution of ligands by exponential enrichment. Nucleic Acids Res. 2015, 43(12):e82.

57. Hoinka J, Berezhnoy A2, Dao P, Sauna ZE, Gilboa E, Przytycka TM. Large scale analysis of the mutational landscape in HT-SELEX improves aptamer discovery. Nucleic Acids Res. 2015, 43(12):5699-5707.

58. McKeague M, Foster A, Miguel Y, Giamberardino A, Verdin C, Chan J. Y. S, *et al*. Development of a DNA aptamer for direct and selective homocysteine detection in human serum. RSC Advances. 2013, 3i, 24415-24422.

59. Blind M, Blank M. Aptamer selection technology and recent advances. Mol Ther Nucleic Acids. 2015, 4:e223.

60. Tucker WO, Shum KT, Tanner JA. G-quadruplex DNA aptamers and their ligands: structure, function and application. Curr Pharm Des. 2012 18(14):2014-2026.

61. Schwalbe H, Buck J, Fürtig B, Noeske J, Wöhnert J. Structures of RNA switches: insight into molecular recognition and tertiary structure. Angew Chem Int Ed Engl. 2007, 46(8):1212-1219.

62. Hasegawa H, Savory N, Abe K, Ikebukuro K. Methods for improving aptamer binding affinity. Molecules. 2016, 21(4). pii: E421.

63. Pobanz K, Lupták A2. Improving the odds: Influence of starting pools on *in vitro* selection outcomes. Methods. 2016, 106:14-20.

64. Zinnen SP, Domenico K, Wilson M, Dickinson BA, Beaudry A, Mokler V, *et al*. Selection, design, and characterization of a new potentially therapeutic ribozyme. RNA. 2002, 8(2):214-228.

65. Sassanfar M, Szostak JW. An RNA motif that binds ATP. Nature. 1993, 364(6437):550-3.

66. Laserson U, Gan HH, Schlick T. Searching 2D RNA geometries in bacterial genomes. 2004, in proceedings of the 12th Annual Symposium on Computational geometry: 373-377.

67. Sabeti PC, Unrau PJ, Bartel DP. Accessing rare activities from random RNA sequences: the importance of the length of molecules in the starting pool. Chem Biol. 1997, 4(10):767-774.

68. Beattie KL, Logsdon NJ, Anderson RS, Espinosa-Lara JM, Maldonado-Rodriguez R, Frost JD 3rd. Gene synthesis technology: recent developments and future prospects. Biotechnol Appl Biochem. 1988, 10(6):510-521.

69. Ruscito A, DeRosa MC. Small-Molecule binding aptamers: selection strategies, characterization, and applications. Front Chem. 2016, 4:14.

70. Schütze T, Wilhelm B, Greiner N, Braun H, Peter F, Mörl M, *et al*. Probing the SELEX process with next-generation sequencing. PLoS One. 2011, 6(12):e29604.

CHAPTER 3

Characterization of MGB2 and MGB1 aptamers and testing their binding to breast cancer cells in plasma and blood

3.1 Introduction

3.1.1 Methods of aptamer binding affinity

As mentioned in Chapter 1, the dissociation constant (K_d) can be defined as Equation 1 when the stoichiometry of binding of target to aptamer is 1:1, (Equation 1) [1].

$$K_d = \frac{[\mathbf{Aptamer}][\mathbf{Target}]}{[\mathbf{complex}]} \qquad \text{Equation 1}$$

The K_d can be measured by different methods [1,2]. The simplest method is dialysis, which separates the free aptamers from the aptamer-target complex according to their size [1]. In dialysis, the aptmer-target complex is placed in one compartment which is separated by a semi-permeable membrane from another one [1]. After reaching equilibrium, only the free aptamers can make it to the second compartment. The K_d is determined from the free aptamer concentrations, measured at different target concentrations, in the second compartment [1]. Measuring K_d using dialysis can be challenging, due to the experimental complications of this method. These challenges include the increased osmotic pressure as a result of impermeability of the target and the complex [3], and long incubation times which may affect the target and the ligand [3]. Other methods have been reported to measure aptamer K_ds, such as surface plasmon resonance (SPR) and gel electrophoresis, that are typically limited to large targets such as proteins [4,5].

In the case of aptamers selected by cell-SELEX, the most common method to determine the K_d is to measure the mean of fluorescence intensity by flow cytometry [6]. In this method, the aptamers are labelled and allowed to bind to their target cells without immobilization, thus increasing the chance of interaction. The K_d value is calculated by the following equation

$$Y = Bmax \frac{X}{Kd+X} \qquad \text{Equation 2}$$

where Y, B_{max}, and X are fluorescence intensity, maximum fluorescence intensity, and aptamer concentration respectively [7]. Flow cytometry has many advantages including high sensitivity, very high statistical accuracy due to the large numbers of cells used, and the outcome data can be analyzed easily to provide quantitative results [6].

3.1.2 Aptamer interaction to their targets

Aptamers resulting from cell-SELEX have the ability to interact with different cellular receptors (biomarkers), and thus provide valuable information on cellular activities and their response to the environment, particularly when used as a therapeutic agent [7]. Cell-SELEX aptamers are considered better than their antibody counterparts when binding to their targets, due to their high binding affinity (low K_d) and specificity. For example, the K_d of the HER2 (breast cancer biomarker) aptamer is 0.56 nM, compared to 18.9 nM for the antibody [8,9]. The small size of aptamers allows them to have a higher tumor penetration rate compared to antibodies, as well as reduced immunogenicity [10,11].

Specific binding of aptamers to their target is a requirement when using them in diagnosis or therapy. Aptamers can discriminate their targets on the basis of small structural differences such as the presence or absence of a methyl [12,13] or a hydroxyl group [14,15]. This degree of specificity is a result of the SELEX process, which allows for the elimination of sequences that bind to closely related analogs of the target in the counter selection step [12].

In the case of cell-SELEX against cancer cells, normal healthy cells from the surrounding tissue are the ideal counter selection targets in order to generate aptamers that discriminate cancer cells from normal healthy ones.

This has great benefits when using these aptamers as therapeutic agents, or for targeted delivery [16].

Aptamers bind to their target via a unique three-dimensional structure through the formation of non-covalent bonds such as van der Waals and hydrogen bonding [17]. The three-dimensional structures of aptamer complexes form highly optimized scaffolds for specific ligand recognition [18]. The stability in the interaction (bond formation) affects the specificity of the aptamers. For example, the specificity of the AMP aptamer was attributed to a specific hydrogen bonding arrangement that involves pseudo-base pair formation with the ligand [19]. In a different study of the flavin mononucleotide-RNA aptamer complex, a similar aptamer ligand recognition conformation, based on hydrogen bonding and stacking interactions, was reported [20].

This Chapter describes the characterization of the affinity (Kd values) of the MGB2 (MCF7) and the MGB1 (MDA-MB-415) aptamers to their target cancer cells and their selectivity against other cancer and normal cells. It then describes the binding of the highest affinity aptamers to their target cancer cells in the complex environment of biological media (plasma and blood).

3.2 Materials and Methods

3.2.1 Materials and Cell Lines

Plasma from human, human blood (Pb and Cd) (BCR® certified reference material) and Hank's Balanced Salt Solution (HBSS) without Ca^{2+} or Mg^{2+} were purchased from Sigma (Canada), EMEM was ordered from Gibco™ / Thermo Fisher Scientific (Canada). Human Primary Peripheral Blood Mononuclear Cells (PBMC), (normal, ATCC: PCS-800-011), HepG2 (hepatocellular carcinoma, ATCC: HB-8065), HeLa (human cervical carcinoma, ATCC: CCL-2), HEK-293 (human embryonic kidney, ATCC: CRL-1573), WI-38 (human lung normal fibroblast ATCC: CCL-75) were purchased from ATCC. Frozen stocks of MDA-MB-231 (breast adenocarcinoma) were obtained from Dr. Christine Pratt, Department of Cellular and Molecular Medicine, University of Ottawa.

PBMC cells were cultured in HBSS, MDA-MB-231, HepG2, HeLa and HEK-293 cells were cultured in DMEM, 10% fetal bovine serum (FBS). WI-38 cells were cultured in EMEM medium with 10% FBS. All cells were cultured in a 5% CO_2-humidified chamber at 37^{o}C and 95% humidity to reach 90% confluence before further sub-culturing, with the exception of PBMC that were harvested after 48 hours of culture.

3.2.2 Determination of K_d for MGB2 and MGB1 aptamers using flow cytometry

All tested DNA aptamers were 6-FAM modified on their 5' ends. MCF7 and MDA-MB-415 cells (2.5×10^5) were fixed in 4% formaldehyde solution in PBS for 15 min at RT, then centrifuged at 1000 xg for 5 min, cells then were re-suspended in 250 µL of binding buffer (5 mM $MgCl_2$ in phosphate-buffered saline (PBS; pH 7.4), 0.3 mg/mL of ssDNA from salmon sperm, and 0.5 mg/mL BSA). Aptamers were heated at 95^{o}C for 5 min and cooled down gradually as mentioned earlier in Chapter 2 to preserve their 3D structure. Fixed MCF7 and MDA-MB-415 cells were incubated with MGB2 and MGB1 aptamers, respectively, at different concentrations (0, 5, 10, 40, 60, 100, 140, 160, 180, and 200 nM) at 4^{o}C for 30 min. Cells then were centrifuged as above and washed with washing buffer (5 mM $MgCl_2$ in PBS (pH 7.4), then re-suspended in 250 µL of binding buffer. Pool 0 was used as a control in K_d determinations in which cells were incubated with the same concentrations as aptamers and tested against other Pools of aptamers and cells using flow cytometry. After flow cytometry, the mean of fluorescence intensity of each aptamer-cell complex Pool was subtracted from the mean of fluorescence intensity of Pool 0 aptamer-cell complexes at the same corresponding concentrations of cells and aptamers.

3.2.3 Selectivity of MGB2 and MGB1 aptamers

The highest affinity aptamers (lowest K_d values) were tested for their selectivity by evaluating their binding to cancer cells other than breast cancer cells and normal cell lines. MCF7, MDA-MB-415, MDA-MB-231, HepG2, HeLa, HEK-293 and WI-38 cells (2.5×10^5) were fixed in 4% formaldehyde solution in PBS for 15 min at RT, then centrifuged at 1000 x g for 5 min. Cells then were resuspended in 250 µL of binding buffer. MAMB1, MAMB12, MAMA2, MAMA12 aptamers (200 nM each; heated and cooled down prior to use as mentioned above) were incubated with cells from each cell line at 4^{o}C for 30 min. Cells were then centrifuged as above, washed with washing buffer, and then re-suspended in 250 µL of binding buffer. Cells alone were used as the controls in which the mean of fluorescence intensity of each aptamer-cell complex of each cell line was subtracted from the mean of fluorescence intensity of cells alone of the same cell line.

3.2.4 Binding of MGB2 and MGB1 aptamers to their target breast cancer cells by fluorescence microscopy

MAMB1 and MAMA2 binding aptamers were selected to further study their binding to breast cancer cells, other cancer cells and normal cells by fluorescence microscopy. Briefly, cells (5×10^5) were plated on coverslips in 35 mm tissue culture dishes and allowed to grow for 24 hours in a tissue culture incubator (37°C at 5% CO_2 and 95% humidity). Cells were then washed with PBS, fixed in 4% formaldehyde solution in PBS for 15 min at RT, washed again with PBS, and aptamers were added for 30 min at RT at a concentration of 300 nM in binding buffer (aptamers were heated and cooled down prior to use). Cells were then washed three times with cold PBS, and 4',6-diamidino-2-phenylindole (DAPI) (diluted 1:20,000 in PBS from a stock solution of 10 mg/mL) was added for 5 min. Cells were then washed three times, for 5 min each, using deionized water. Coverslips were then mounted on microscope slides using VECTASHIELD mounting medium (Vector Laboratories, USA), and imaged using ZOE™ fluorescent cell Imager (Bio-Rad, Canada) using the green and blue channels.

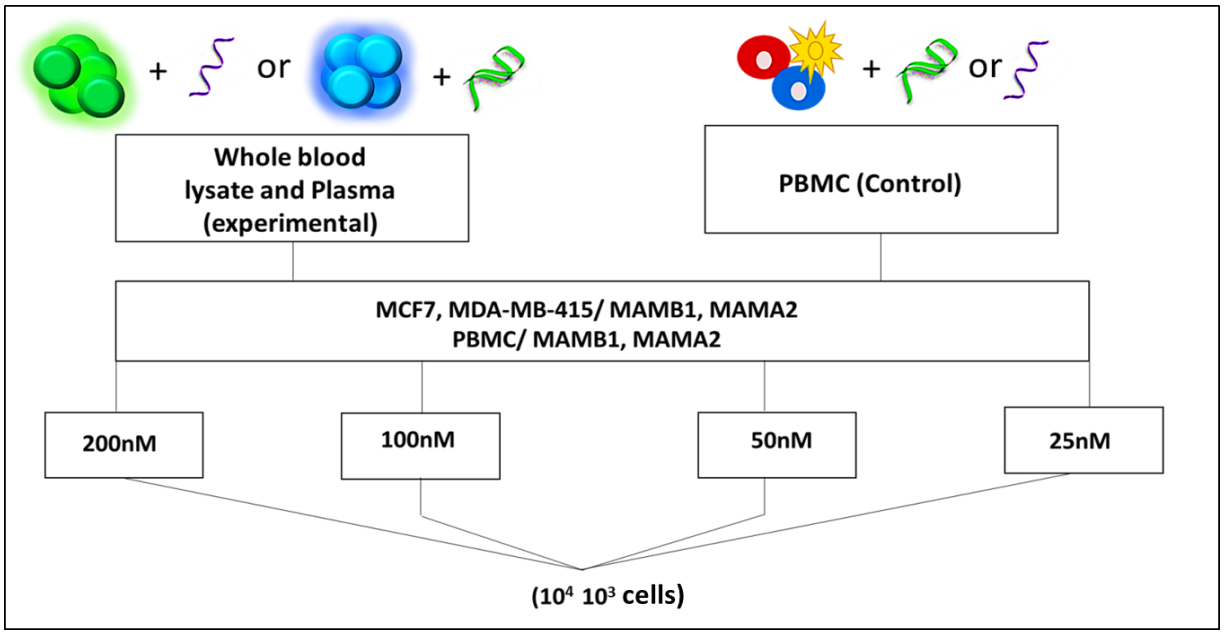
MCF7 and MDA-MB-415 cells were used as the target breast cancer cell lines for testing MAMB1 and MAMA2 aptamers respectively. HCAEC, MCF10A, HEK293, WI38, MDA-MB-231 and HepG2 cell lines were used as negative controls to test the selectivity of both aptamers. A random sequence was used as an aptamer control and was incubated with each cell line in the same way as the test aptamers, and imaged using fluorescence microscopy.

3.2.5 Testing the binding of MGB2 and MGB1 aptamers to their target cell lines in human plasma and blood

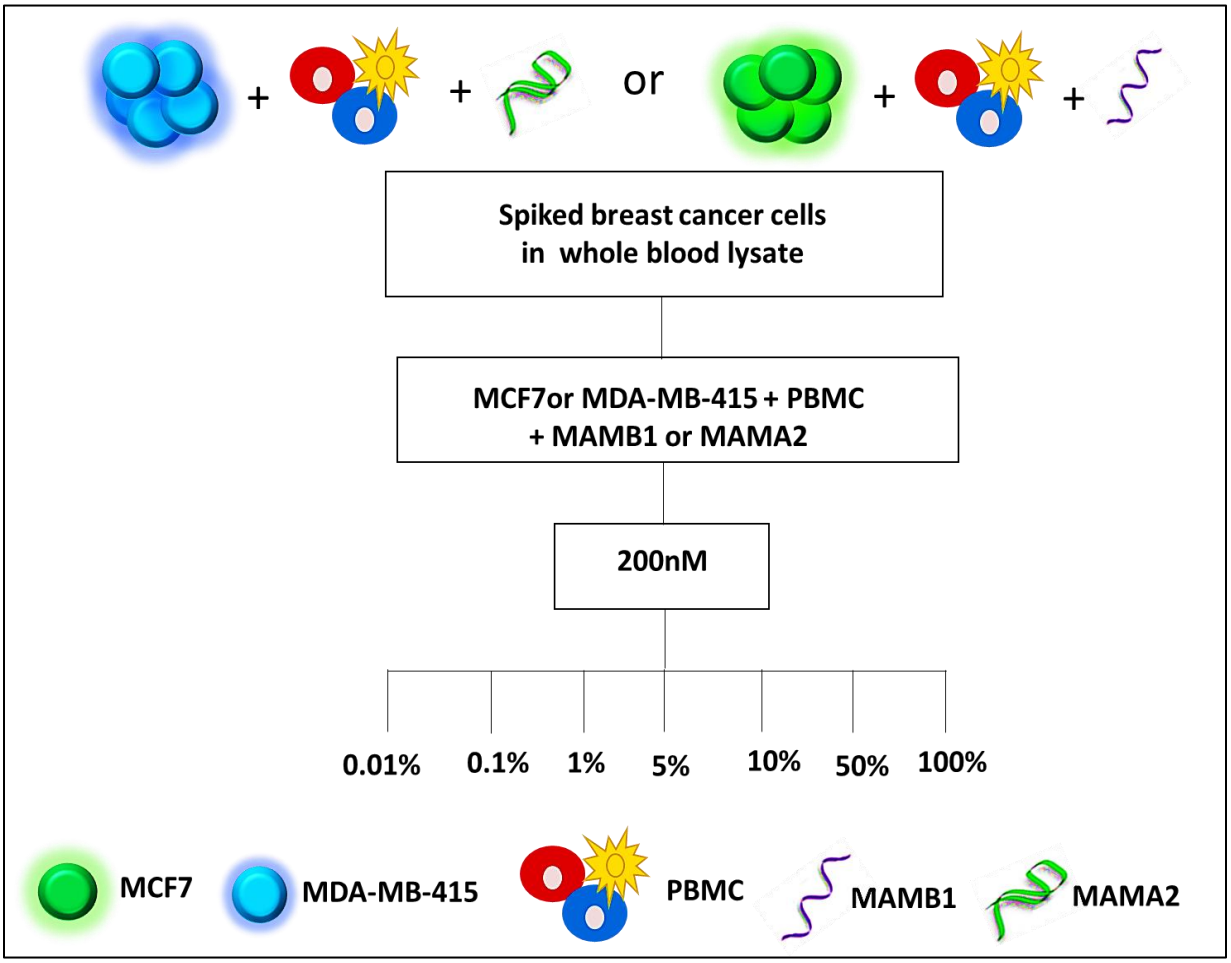
Human plasma was dissolved in deionized water for 2 hours on a shaker at RT. Human lyophilized blood was dissolved in deionized water for 3 hours on a shaker at RT. After that, 50 µL of blood was prepared and observed on a microscope slide. Red Blood Cells (RBCs) were then lysed by RBCs lysis solution (154 mM NH_4Cl, 5.7 mM K_2HPO_4, 0.1 mM EDTA) for 15 min at RT. The blood was then centrifuged at 4,000 xg for 10 min and re-suspended with deionized water. PBMC cells represent normal peripheral blood cells and include mainly lymphocytes (T cells, B cells, natural killer cells) and monocytes. These cells were used as controls (i.e. non-cancerous). MCF7 and MDA-MB-415 were used as breast cancer target cells.

MCF7, MDA-MB-415, and PBMC cells (1×10^7) were fixed in 4% formaldehyde solution in PBS for 15 min at RT. Cells were then centrifuged at 1,000 xg for 5 min, and resuspended each in 1 mL of human whole blood lysate and plasma to be used as the stocks. Two dilutions were prepared from the stocks (1,000 and 10,000 cells) from each cell line. MAMB1 and MAMA2 aptamers were then incubated (heated and cooled down prior to use as mentioned above) with the dilutions prepared above, from each corresponding cell line at different concentrations (200, 50, 25 nM for the plasma and 200, 100, and 50 nM for the whole blood lysate) to a final volume of 250 µL at 4^oC for only 15 min. to avoid any endonuclease activity from whole blood lysate and plasma components. Samples were then analyzed by flow cytometry by counting 1,000 events for both dilutions (Scheme 3.1).

A


+
or
+
+
or
Whole blood
lysate and Plasma
(experimental)
PBMC (Control)
MCF7, MDA-MB-415/ MAMB1, MAMA2
PBMC/ MAMB1, MAMA2
200nM
100nM
50nM
25nM
(10⁴ 10³ cells)

B


+
+
or
+
+
Spiked breast cancer cells
in whole blood lysate
MCF7or MDA-MB-415 + PBMC
+ MAMB1 or MAMA2
200nM
0.01%
0.1%
1%
5%
10%
50%
100%
MCF7
MDA-MB-415
PBMC
MAMB1
MAMA2

Scheme 3.1. Schematic diagram of the experimental design of plasma and blood studies. A. Whole blood lysate and plasma experimental design: MCF7 and MDA-MB-415 were used as the target breast cancer cells, PBMC cells were used as the control (non-cancerous), target cells were incubated with their aptamers (MCF7 with MAMB1 and MDA-MB-415 with MAMA2). PBMC cells were incubated with the same aptamers. Different cell densities ($1X10^4$ and $1X10^3$ cells) were used for each target cell with their aptamer, as well as PMBC cells. Cell-aptamer complexes were then analyzed by flow cytometry. Different concentrations of the aptamers (50, 100, and 200 nM) were used in whole blood lysate experiments. In plasma experiments, 25, 50 and 200 nM of aptamers were used. B. Whole blood lysate spiking experiments: MCF7 breast cancer cells were incubated at different percentages with PBMC cells in blood to a total number of $1X10^4$ cells at 200 nM of MAMB1. The same procedure was applied for MDA-MB-415 cells with their aptamer (MAMA2). After that, both samples were subjected to flow cytometry analysis.

3.2.6 Testing the binding of MAMB1 and MAMA2 aptamers to their cancer cell lines in spiked blood

MCF7, MDA-MB-415, and PBMC cells stocks ($1X10^7$ cells each) were used in this study. For each cell line cancer cells were spiked in different percentages in PBMC cells-whole blood lysate to a total number of $1X10^4$ cells. The percentages of spiked cancer cells from each cell line were 0.01, 0.1, 1, 5, 10, 50, and 100%. MAMB1 and MAMA2 (200 nM each) were incubated with their corresponding spiked cancer cells in whole blood lysate at a total volume of 300 µL for 15 min. at RT. The binding of aptamers to breast cancer cells was then analyzed by flow cytometry.

3.3 Results and discussion

3.3.1 Determination of K_d for MGB2 and MGB1 aptamers

Aptamers that showed the best binding results from cell-SELEX for both MGB2 and MGB1 selections were further tested for their affinity (K_d) to their corresponding cell lines using flow cytometry. Although live cells were used in the selection experiments, breast cancer cells as well as normal cells were fixed in all the characterization assays as well as in the complex media and TCM studies (chapter 3, 4 and 5). The reason behind the fixation was to obtain less variable results and to test the reproducibility of the aptamers. Live cells tend to give more variable results. When fixing the cells our target proteins (MGB2 and MGB1) would be crossed linked, as a result they might not be secreted outside the cells.

Therefore, this might affect our results in a way that these proteins would be more available on the surface of the cells and thus the mean of fluorescence intensity would increased. However, the proteases enzymes that cleave the signal peptide part of our proteins and release them outside the cells are not available in the cell medium. Thus, fixing the cells was a good approach to achieve our goals.

MAMB0, MAMB1, MAMB4, MAMB8, MAMB12, MAMB57 were tested against MCF7 cells as described in the Materials and Methods section of this chapter. Two out of the six aptamers showed K_d values in the low nanomolar range (Figure 3.1A). MAMB1 had a K_d of 14 ± 4 nM against MCF7 cell line. MAMB12 had a K_d of 20 ± 6 nM against the same cell line (Figure 3.1B). MAMB0, MAMB4, MAMB8, MAMB57, showed K_d values of 80 ± 14 nM, 105 ± 20 nM, 78 ± 15 nM, and 72 ± 9 nM respectively. MAMA2, MAMA5, MAMA6, MAMA12 were tested against MDA-MB-415 cell line. MAMA2 and MAMA12 showed K_d values of 3 ± 1 nM and 8 ± 3 nM respectively, indicating that MAMA2 has higher affinity to MDA-MB-415 than MAMA12 (Figure 3.2). MAMA5 showed a K_d value of 55 ± 14 nM and MAMA6 had a K_d value of 50 ± 20 nM. Flow cytometry has been used by many to measure the K_d values of protein-binding aptamers and aptamers generated by cell-SELEX, and these were reported to be in the nanomolar range [21-23] and our K_d values for both targets were within this expected range.

It has been previously reported that determination of binding sites of a sequence through the predicted secondary structure and sequence truncation can affect the overall binding affinity of a sequence [24]. In our case, MGB2 and MGB1 aptamers were both selected from a 24 base variable region library. However, the length of tested aptamers is 65 bases for MGB2 and 64 bases for MGB1, due to the addition of the 20 bases primers on the 5’ and 3’ ends, which might have helped in the interaction with the target molecule. Overall, the predicated secondary structure of all aptamers (presented in Figures 3.1 and 3.2) were relatively simple. MAMA2 (which had the lowest K_d value) had few loops on the 3’ end. The 5’ end of MAMA2 had no loops, as predicted by its secondary structure (Figure 3.2A), indicating that the removal of the 5’ 20 base primer might impact its binding affinity. The same was applicable to the MAMB1 aptamer (Figure 3.1 A). The secondary structure provided here was only a prediction of the most stable structure,

with different probabilities, as indicated by the colors (red had the highest stability and pink had the lowest stability).

The enrichment of the sequences in the ssDNA libraries increased with the increasing number of selection rounds [25]. Therefore, enrichment could affect the binding affinity of the aptamers to their targets. In our case, and according to the HiSeq data analysis by AptaCluster (Section 2.4 (Table 2.2)), the enrichment value of MAMB1 was lower than MAMB12 of the same pool (122.649 and 920.639 respectively).

It was almost the same for its corresponding counter selection cell line for MAMB1 (122.649 vs 123.569) and higher in MAMB12 (920.639 vs 680.997).
This might have been due to the fact that the MGB2 data was analyzed differently from the MGB1 data (paired-end read vs single-end reads). The quality of MGB2 data received from the high-throughput sequencing center could have affected the overall analysis, including the enrichment values.

In contrast, the enrichment value of MAMA2 in the last pool of the positive selection cell line (MDA-MB-415) was higher than the one for MAMA12 of the same pool (5836.12 and 1409.74 respectively). The enrichment of both was higher than that for their corresponding counter selection of the same pool (5836.12 vs 5277.48 for MAMA2 and 1409.74 vs 242.50 for MAMA12). This might explain the higher K_d value of MAMA2, compared to MAMA12 (3 vs 8 nM respectively).

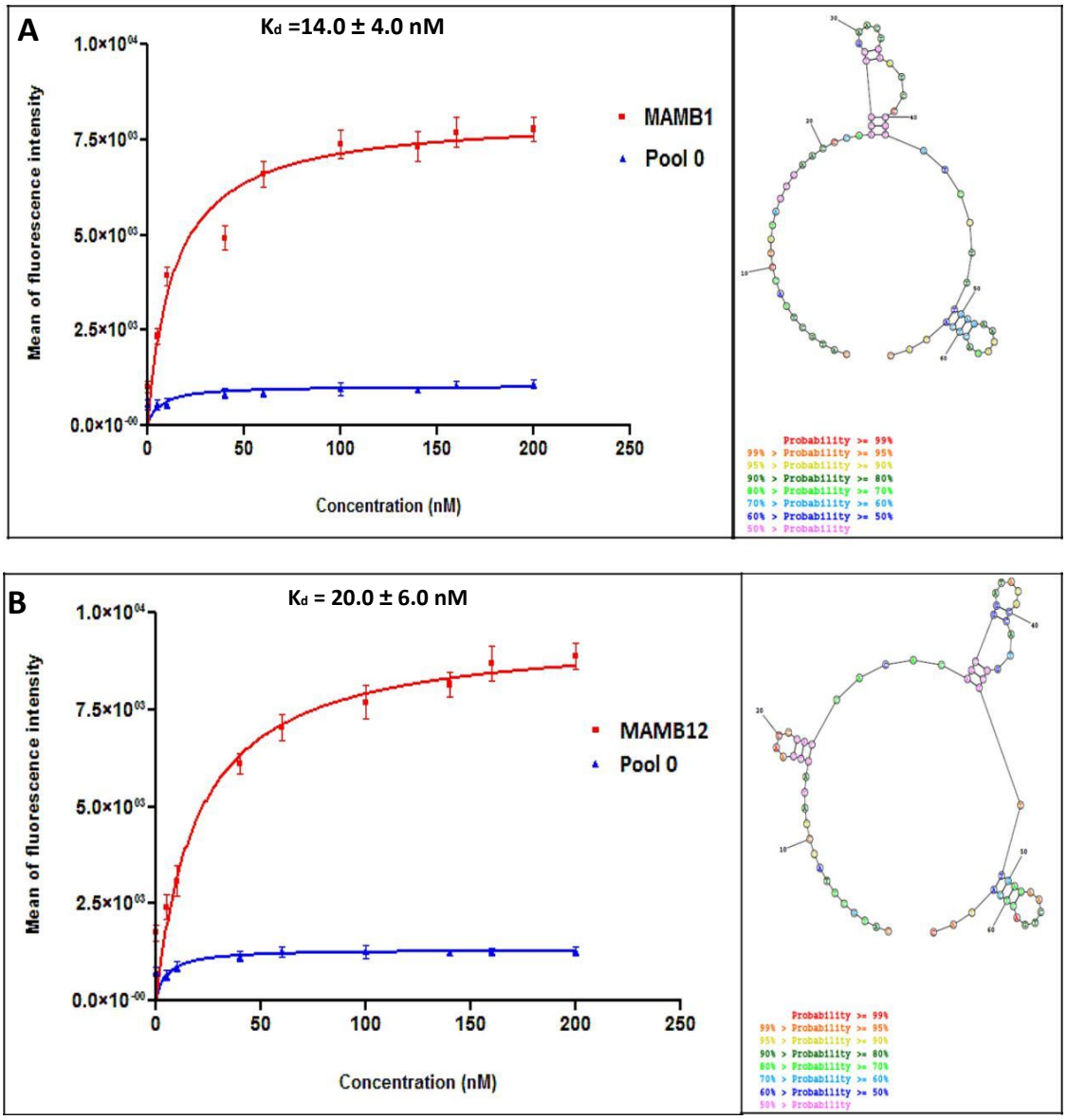


Figure 3.1 Binding curves of 6-FAM-labeled aptamer sequences to MCF7 cells (MAMB1 (A) and MAMB12 (B), (left) and the predicted secondary structure of MAMB1 and MAMB12 (right). The concentrations of DNA were 0, 5, 10, 40, 60, 100, 140, 160, 180, 200 nM and 6-FAM labelled Pool 0. The K_d values were analyzed using GraphPad Prism 3.0 by the classical B_{max} equation. The secondary structure (right) of MAMB1 and MAMB12 was taken from Figure 2.7 and analyzed by RNAstructure software. Error bars represent the S.E.M. of three trails

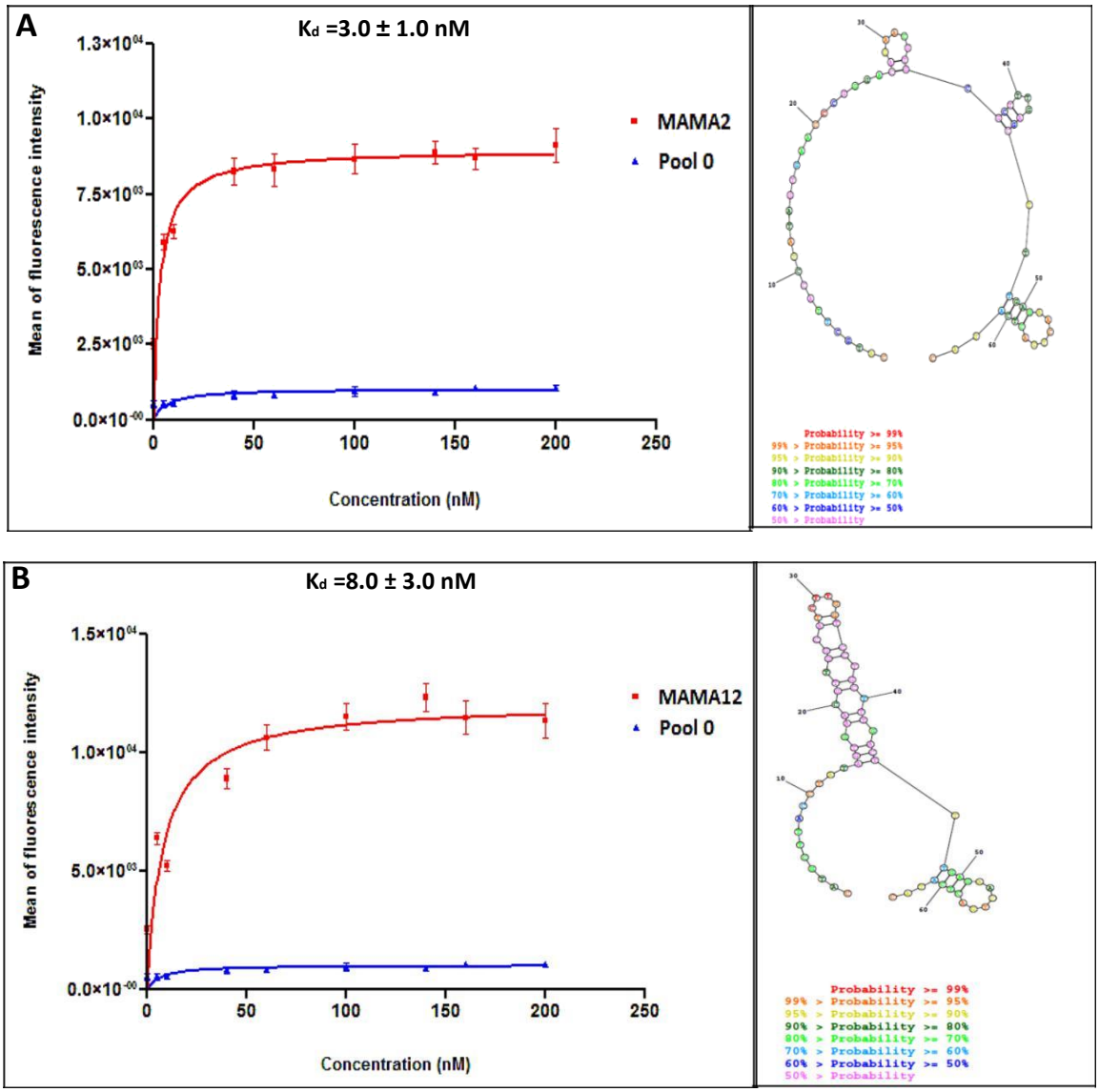


Figure 3.2: Binding curves of 6-FAM-labeled aptamer sequences to MCF7 cells (MAMA2 (A) and MAMA12 (B), (left) and the predicted secondary structure of MAMA2 and MAMA12 (right). The concentrations of DNA are 0, 5, 10, 40, 60, 100, 140, 160, 180, 200 nM and 6-FAM labelled Pool 0. The k_d values were analyzed using GraphPad Prism 3.0 by the classical B_{max} equation mentioned in the introduction. The secondary structure (right) of MAMA2 and MAMA12 was taken from Figure 2.7 and analyzed by RNAstructure software. Error bars represent the S.E.M. of three trails.

3.3.2 Selectivity of MGB2 and MGB1 aptamers to other cancer and normal cell lines

Many cell-SELEX aptamers have shown high selectivity when tested against non-target cancer and normal cell lines [21,23]. In our case, MAMB1, MAMB12, MAMA2, MAMA12 were tested for their selectivity to other cancer and normal cell lines by flow cytometry.

Among all cell lines tested in this study, neither MGB2 nor MGB1 aptamers bound with high affinity to non-target cell lines (Figure 3.3). MAMB1 and MAMB12 did not recognize MDA-MB-415 cells, which was used in the MGB1 cell-SELEX, which indicated the high selectivity of both aptamers. Moreover, none of the aptamers bound to MDA-MB-231 breast cancer cell lines (Figure 3.3A). MAMA2 and MAMA12 had the same binding pattern as the MGB2 aptamers (Figure 3.3B). No binding was observed to non-target cell lines or to normal cell lines. This study indicated the specificity of MGB2 and MGB1 aptamers to their corresponding target cell lines. These results matched our expectations to develop aptamers with high affinity and selectivity to their targets.

3.3.3 Fluorescence microscopy results showed high affinity of MAMB1 and MAMA2 to their breast cancer cells

Using fluorescence microscopy to test the binding of aptamers to their target cells is a common method in the literature [26, 27]. Therefore, MAMB1 and MAMA2 were further tested by this method. MAMB1 and MAMA2 aptamers showed high specific binding to their target breast cancer cells lines MCF7 and MDA-MB-415, respectively (Figure 3.4A and D). MAMB1 did not show any binding to MDA-MB-415 cells (Figure 3.4F), and the same was observed with MAMA2 against MCF7 (Figure 3.4C). This matched our expectations, since the pool was divided into two parts when MGB1 was used as a counter selection target in the protein-SELEX (as mentioned in Chapter 2). No binding was observed when both aptamers were tested against the counter cell lines HCAEC and MCF10A (Figure 3.4G, I, J and L). No affinity was seen of the random sequence when incubated with the target cell lines MCF7 and MDA-MB-415 (Figure 3.4B and E) and the counter cell lines HCAEC and MCF10A (Figure 3.4H and K).

Testing other cancer and normal cell lines with MAMB1 and MAMA2 aptamers, as well as the random sequence, showed no binding (Appendix A). The fluorescence microscopy results matched the selectivity results obtained by flow cytometry, indicating that both aptamers were selective to their breast cancer cell targets, using these two methods.

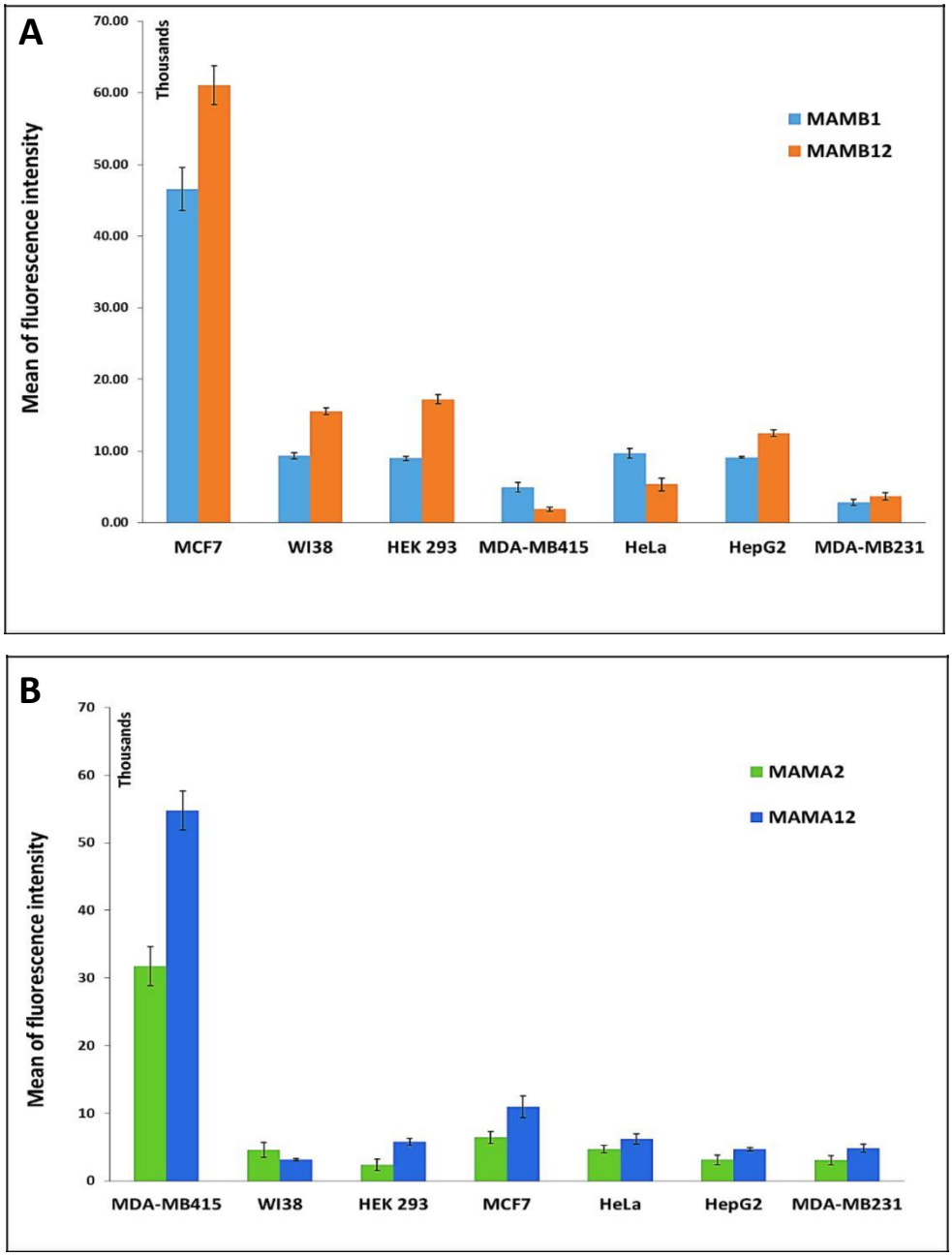
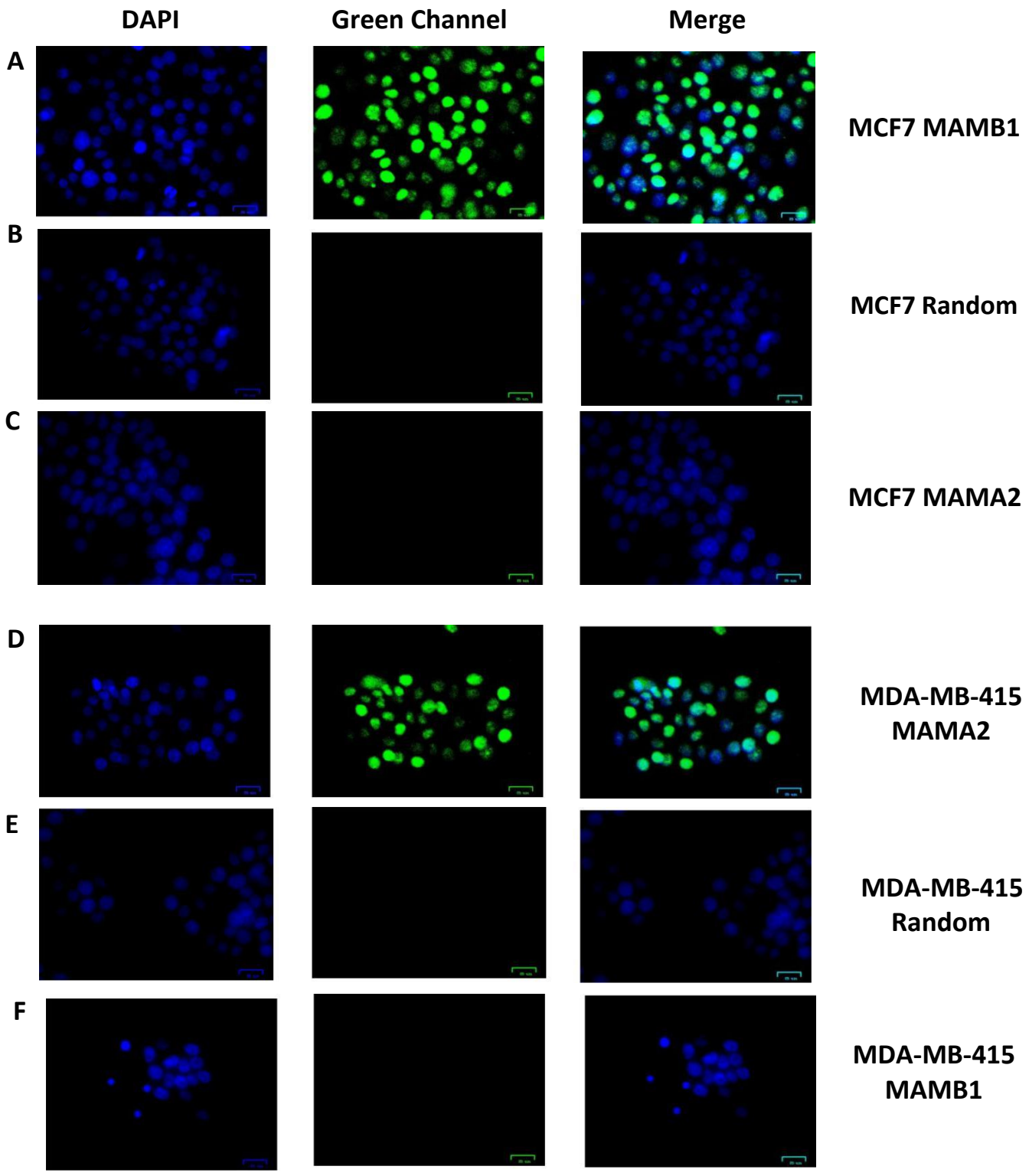


Figure 3.3: The specificity of MGB2 (A) and MGB1 (B) aptamers to different cell lines. The cell lines used were MCF7 and MDA-MB-415, as the target cell lines, for both MGB2 and MGB1 respectively. WI-38 is a human normal lung fibroblast cell line. HEK 293 is a transformed human embryonic kidney cell line. HeLa is a human cervical cancer cell line. HepG2 is a human liver cancer cell line. MDA-MB-231 is a breast cancer cell line. All were used to test the binding of MGB2 and MGB1 aptamers. Error bars represent the S.E.M. of three trails.


DAPI
Green Channel
Merge
A
MCF7 MAMB1
B
MCF7 Random
C
MCF7 MAMA2
D
MDA-MB-415 MAMA2
E
MDA-MB-415 Random
F
MDA-MB-415 MAMB1

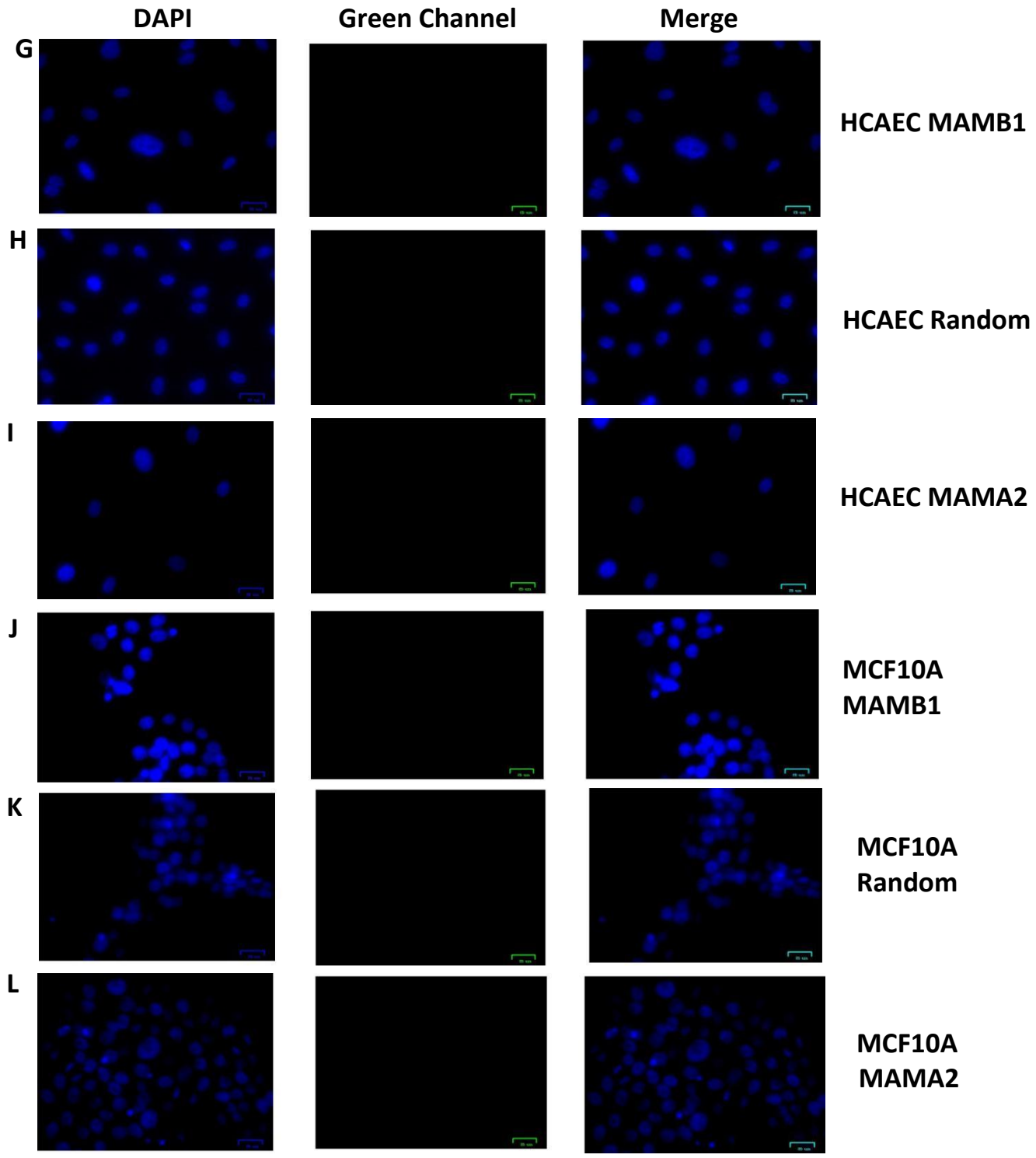


Figure 3.4: Fluorescence microscopy images of aptamers MAMB1 and MAMA2 binding to target cancer and counter selection cell lines. A: MAMB1 binding to MCF7 cells, B: random sequence binding to MCF7, C: MAMA2 binding to MCF7. D: MAMA2 binding to MDA-MB-415 cells, E: random sequence binding to MDA-MB-415, F: MAMB1 binding to MDA-MB-415, G-L: binding of both MAMB1 and MAMA2 as well as random sequence to counter selection cell lines HCAEC (G-I) and MCF10A (J-L). Both aptamers and the random sequence were incubated with the cells for 30 min and then imaged. Scale bars correspond to 25 µm in all images.

3.3.4 Binding of MAMB1 and MAMA2 to their target cell lines in plasma and whole blood lysate

As mentioned in Chapter 1 of this thesis, flow cytometry is a popular method to detect cancer cells in solution, and since it has many advantages (including analyzing large numbers of cells), it was chosen to be used in our study. The ability of MAMB1 and MAMA2 aptamers to bind to breast cancer cells in plasma and whole blood lysate was investigated using PBMC cells as control. MCF7 and MDA-MB-415 were used as breast cancer cell line targets for MAMB1 and MAMA2 aptamers, respectively. The binding of the aptamers to their target breast cancer cells, as well as to control cells PBMC, was indicated by the mean of fluorescence intensity obtained from the flow cytometry instrument after counting aptamers that bind to the cells.

The human plasma used in this study was prepared from pooled human blood (according to the source) and contained active clotting factors. The plasma results showed that both aptamers (MAMB1 and MAMA2) bound specifically ($P<0.05$) to their breast cancer cell target compared to non-cancerous cells (PBMC) (Figure 3.5). Moreover, the results revealed that the mean of fluorescence intensity for both aptamers was proportional to the concentration of the aptamer and the number of cells used in the study (Figure 3.5). MAMA2-MDA-MB-415 cells showed increased fluorescence intensity with increasing concentration of the aptamer from 25 to 200 nM and from 1×10^3 to 1×10^4 cells (Figure 3.5A and D). The same was noticed in MAMB1-MCF7 cells (Figure 3.5 A and D). However, MAMB1 aptamer showed a higher affinity for MCF7 compared to MAMA2 aptamer affinity to MDA-MB-415 at all aptamer concentrations, and both 1×10^4 and 1×10^3 cells. This could be seen in the values of the fluorescence intensity for both aptamers at all concentrations and number of cells used (Figure 3.5). The mean of fluorescence intensity of all aptamer-cell complexes was above that obtained from cells alone in plasma (without the aptamer) (Table 3.1).

Further testing of the aptamers in more complex environments such as whole blood lysate was performed using flow cytometry. The lyophilized human blood samples that were used in this study contained all the blood components found in normal human blood (according to the source), which included lysed cells and plasma proteins (such as albumins, globulins, fibrinogen).

Cell lysis occurred due to the way the blood was handled (freeze-dried), and this was confirmed under the microscope, as the cells were not intact.

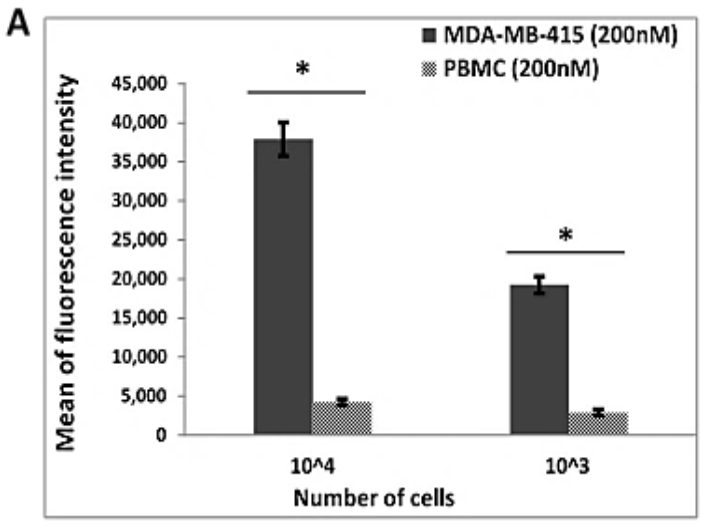
The results using whole blood lysate showed that MAMB1 and MAMA2 aptamers significantly ($P<0.05$) distinguished MCF7 and MDA-MB-415 breast cancer cells from PBMC cells (Figure 3.6). The fluorescence intensity increased with increasing aptamer concentration, and number of cells, for both cell lines. MAMB1 had a higher affinity for MCF7 cells than MAMA2 had to MDA-MB-415 cells (Figure 3.6). The mean of fluorescence intensity of all aptamer-cell complexes was above that for cells alone in whole blood lysate (without the aptamer) (Table 3.1).

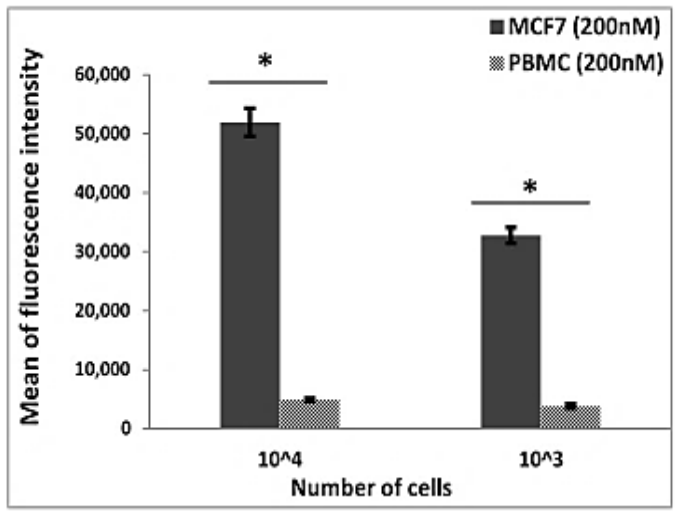
When comparing plasma and whole blood lysate results, the mean of fluorescence intensity of the aptamer-cell complexes was higher in whole blood lysate for both aptamers at all concentrations and number of cells used. For example, the fluorescence intensity of MAMB1-MCF7 cells at 200 nM and $1X10^4$ cells was around 51,000 for plasma and 83,000 for whole blood lysate (Figures 3.4D and 3.5D respectively). This is could be due to the fact that the background for blood was higher than plasma (as indicated by the values of cells only in whole blood lysate;, Table 3.1), since it is a more complex mixture and contained more components (enzymes and proteins). However, the results indicated that MAMB1 aptamer is more specific and has higher affinity to its target cancer cells than MAMA2. This could be as a result of the PCR mutation during the SELEX process (as mentioned in chapter 2), that led to increasing the number of random region to 25 bases instead of 24 as it's the case in MAMA2 aptamer. As discussed in chapter 2 this mutation could increase the binding affinity of an aptamer as reported in the literature.

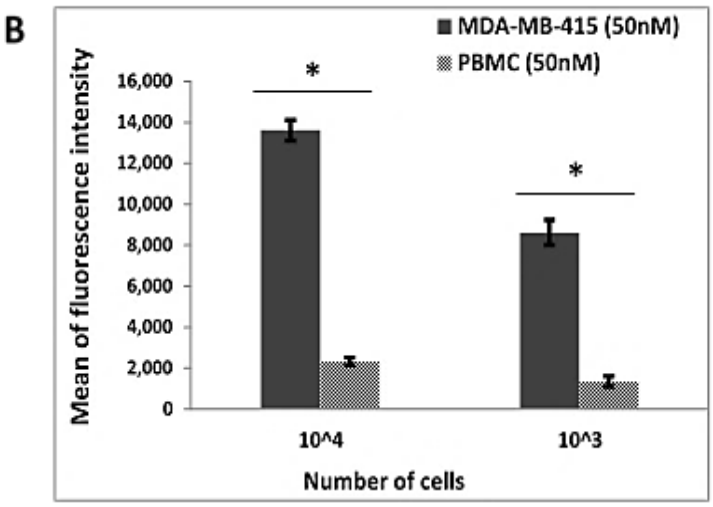
In summary, the binding of MAMB1 and MAMA2 was studied in plasma and whole blood lysate-like environments. The main goal was to investigate the behavior of both aptamers in complex environments. It worth noting that both aptamers were unmodified in terms of adding functional groups, the only modification they had was 6-FAM labelling on the 5' end for both aptamers. Blood and plasma are complex environments and contain an abundance of proteins and enzymes (nuclease) that could affect the binding of the aptamers (in this case MAMA2 aptamer) to their targets.

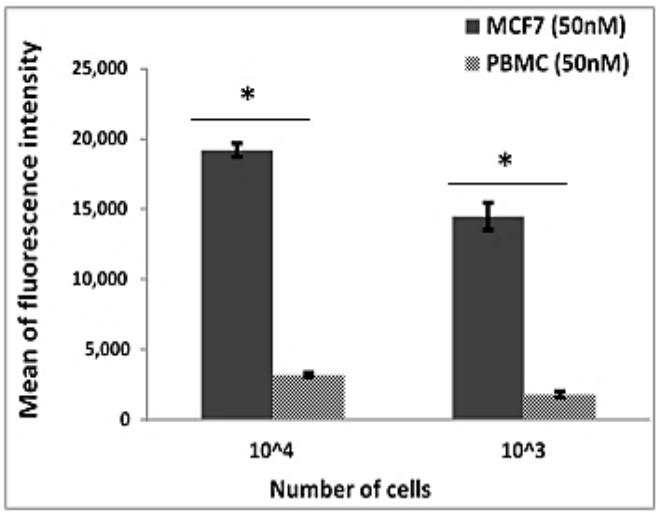
Since both aptamers were unmodified, the aptamer incubation time with their target cells and normal cells (controls) was decreased from 30 to 15 min. Also, it is important to note is that, in the plasma experiments, the investigated concentrations of aptamers were lower than that used in the blood experiments due to the fact that plasma is a less complex environment. Lower concentrations of aptamer (25 nM) were tested in whole blood lysate but both aptamers did not show significant binding at this concentration ($P>0.05$), indicating that components of the blood interfered with aptamer-target binding (Figure 3.6G and H).

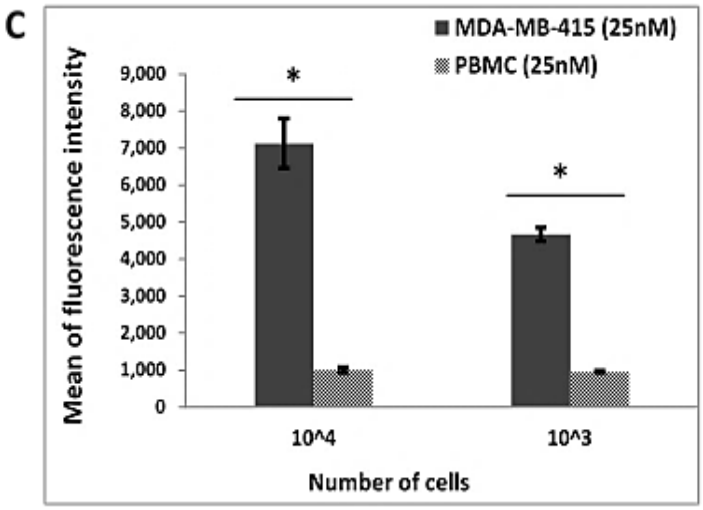
The results showed here are promising to further test the aptamers in more complex environments that simulate the environment that CTCs circulate within, in which normal cells and cancer cells occur together in the whole blood. This will be discussed in the following section.


A
MDA-MB-415 (200nM)
PBMC (200nM)
Mean of fluorescence intensity
45,000
40,000
35,000
30,000
25,000
20,000
15,000
10,000
5,000
0
*
*
10^4
10^3
Number of cells


D
MCF7 (200nM)
PBMC (200nM)
Mean of fluorescence intensity
60,000
50,000
40,000
30,000
20,000
10,000
0
*
*
10^4
10^3
Number of cells


B
MDA-MB-415 (50nM)
PBMC (50nM)
Mean of fluorescence intensity
16,000
14,000
12,000
10,000
8,000
6,000
4,000
2,000
0
*
*
10^4
10^3
Number of cells


E
MCF7 (50nM)
PBMC (50nM)
Mean of fluorescence intensity
25,000
20,000
15,000
10,000
5,000
0
*
*
10^4
10^3
Number of cells


C
MDA-MB-415 (25nM)
PBMC (25nM)
Mean of fluorescence intensity
9,000
8,000
7,000
6,000
5,000
4,000
3,000
2,000
1,000
0
*
*
10^4
10^3
Number of cells

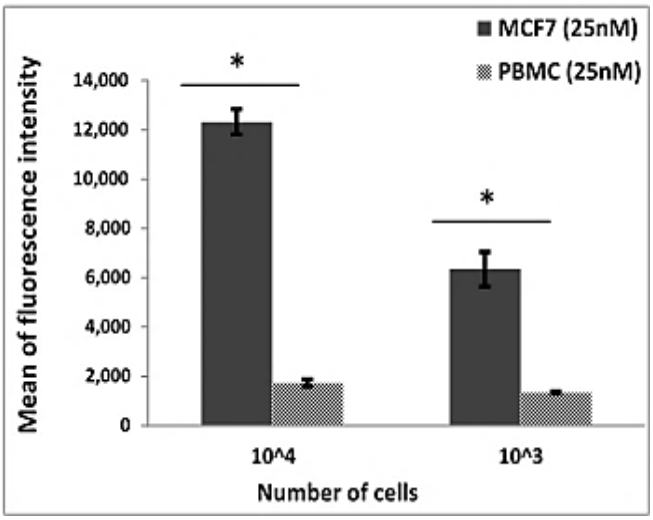
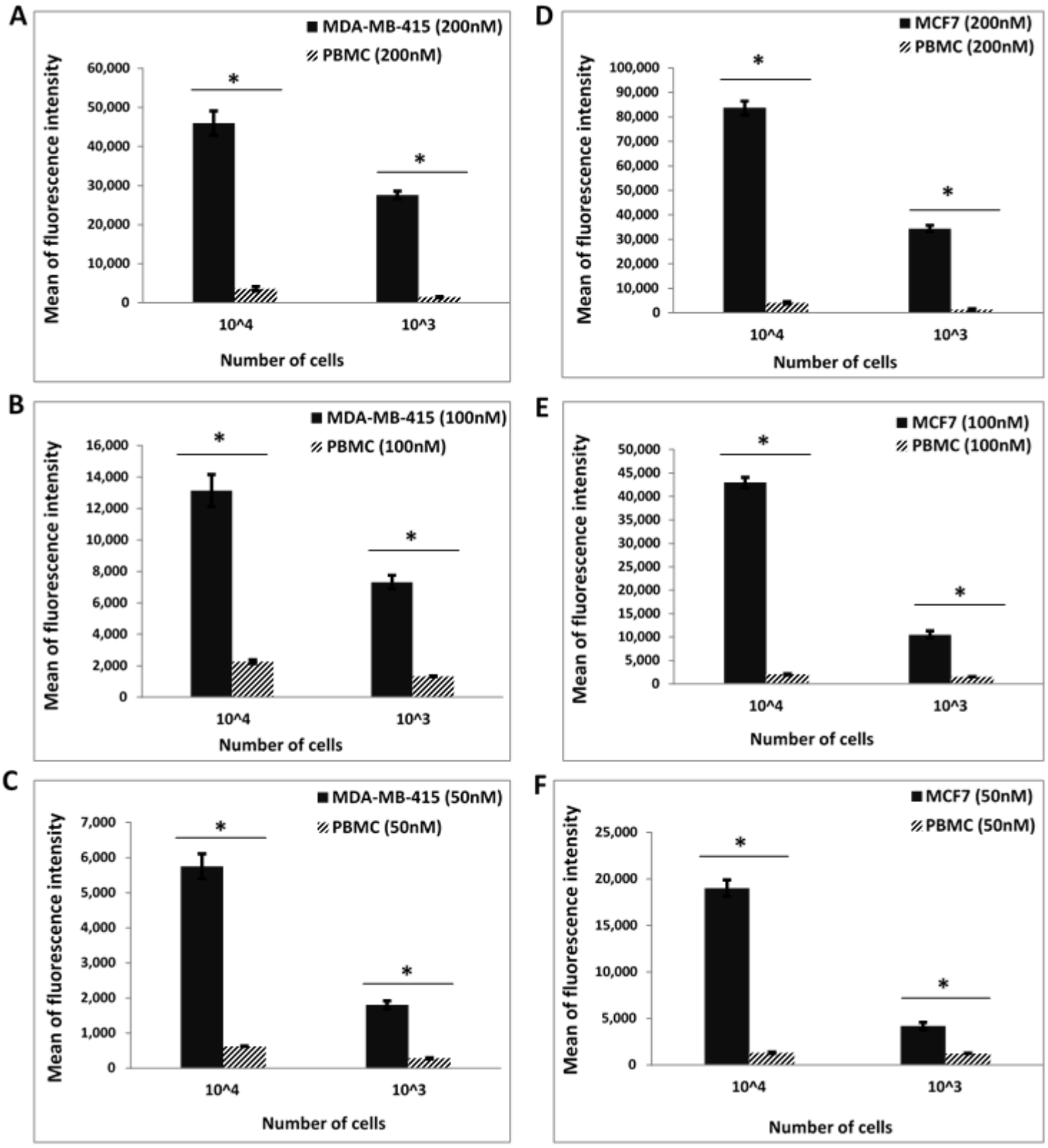

F
MCF7 (25nM)
PBMC (25nM)
Mean of fluorescence intensity
14,000
12,000
10,000
8,000
6,000
4,000
2,000
0
*
*
10^4
10^3
Number of cells

Figure 3.5: Mean of fluorescence intensity of cancer cell lines (MCF7, MDA-MB-415) versus normal cells (PBMC) for MAMB1 and MAMA2 aptamers in plasma. A-C: Mean of fluorescence intensity of MAMA2- MDA-MB-415 cancer cells at 200 nM (A), 50 nM (B) and 25 nM (C) of the aptamer at $1X10^4$ and $1X\ 10^3$ cells. D-F: Mean of fluorescence intensity of MAMB1-MCF7 cell complexes at 200 nM (D), 50 nM (E) and 25 nM (F) of the aptamer at $1X10^4$ and $1X\ 10^3$ cells. Error bars represent the S.E.M. of three trials. The statistical significance was determined by one-way ANOVA followed by the Fisher LSD test using SPSS software (SPSS, version 23). P Values less than 0.05 were considered to be statistically significant.


A
MDA-MB-415 (200nM)
PBMC (200nM)
Mean of fluorescence intensity
60,000
50,000
40,000
30,000
20,000
10,000
0
*
*
10^4
10^3
Number of cells
D
MCF7 (200nM)
PBMC (200nM)
Mean of fluorescence intensity
100,000
90,000
80,000
70,000
60,000
50,000
40,000
30,000
20,000
10,000
0
*
*
10^4
10^3
Number of cells
B
MDA-MB-415 (100nM)
PBMC (100nM)
Mean of fluorescence intensity
16,000
14,000
12,000
10,000
8,000
6,000
4,000
2,000
0
*
*
10^4
10^3
Number of cells
E
MCF7 (100nM)
PBMC (100nM)
Mean of fluorescence intensity
50,000
45,000
40,000
35,000
30,000
25,000
20,000
15,000
10,000
5,000
0
*
*
10^4
10^3
Number of cells
C
MDA-MB-415 (50nM)
PBMC (50nM)
Mean of fluorescence intensity
7,000
6,000
5,000
4,000
3,000
2,000
1,000
0
*
*
10^4
10^3
Number of cells
F
MCF7 (50nM)
PBMC (50nM)
Mean of fluorescence intensity
25,000
20,000
15,000
10,000
5,000
0
*
*
10^4
10^3
Number of cells

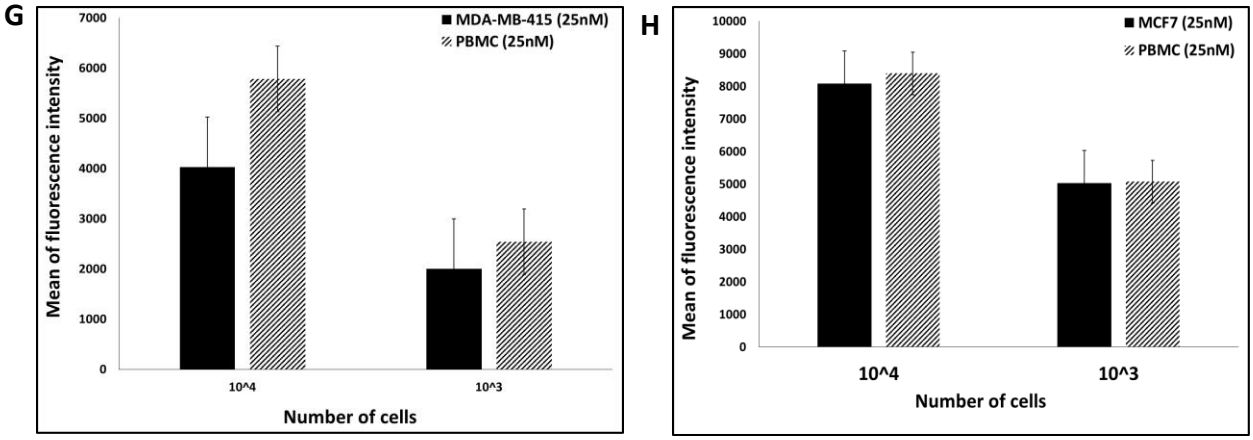


Figure 3.6: Mean of fluorescence intensity of cancer cell lines vs normal cells (PBMC) for MAMB1 and MAMA2 aptamers in whole blood lysate. A-C: Mean of fluorescence intensity of MAMA2-MDA-MB-415 cell complexes at 200 nM (A), 100 nM (B) and 50 nM (C) of the aptamer at $1X10^4$ and $1X\ 10^3$ cells. D-F: Mean of fluorescence intensity of MAMB1-MCF7 cell complexes at 200 nM (D), 100 nM (E) and 50 nM (F) of the aptamer at $1X10^4$ and $1X\ 10^3$ cells. G: Mean of fluorescence intensity of MDA-MB-415 cell-aptamer complexes at 25 nM of MAMA2 at $1X10^4$ and $1X\ 10^3$ cells. H: Mean of fluorescence intensity of MCF7 cell aptamer complex at 25 nM of MAMB1 at $1X10^4$ and $1X\ 10^3$ cells. Error bars represent the S.E.M. of three trials. The statistical significance was determined by one-way ANOVA, followed by the Fisher LSD test using SPSS software (SPSS, version 23). The p values less than 0.05 were considered to be statistically significant.

Table 3.1: Mean of fluorescence intensity of cells only (without aptamer) in plasma and whole blood lysates.

Type of cells (Cells only)	Mean of fluorescence intensity in plasma	Mean of fluorescence intensity in whole blood lysate
MCF7	2,104	2,624
MDA-MB-415	1,598	1,945
PBMC	854	1,168

3.3.5 Binding of MAMB1 and MAMA2 to breast cancer cells in mixture of cells in whole blood lysates

In this Section, the binding of MAMB1 and MAMA2 aptamers to their target breast cancer cells was investigated, in a mixture of cells, using flow cytometry. The binding of MAMAB1 and MAMA2 was investigated in a more complex environment in which cancer cells were spiked in whole blood lysate containing PBMC cells.

It has been reported that only a small portion of CTCs remain viable in the circulation (0.1%) and even fewer CTCs could initiate the metastasis (0.01%) [29]. Therefore, MCF7 and MDA-MB-415 were spiked separately at different percentages (as mentioned in the Material and Method section) in PBMC cells-blood suspension. The incubation time was 15 min to avoid any degradation of aptamers that could happen as a result of endonuclease activity of the blood components.

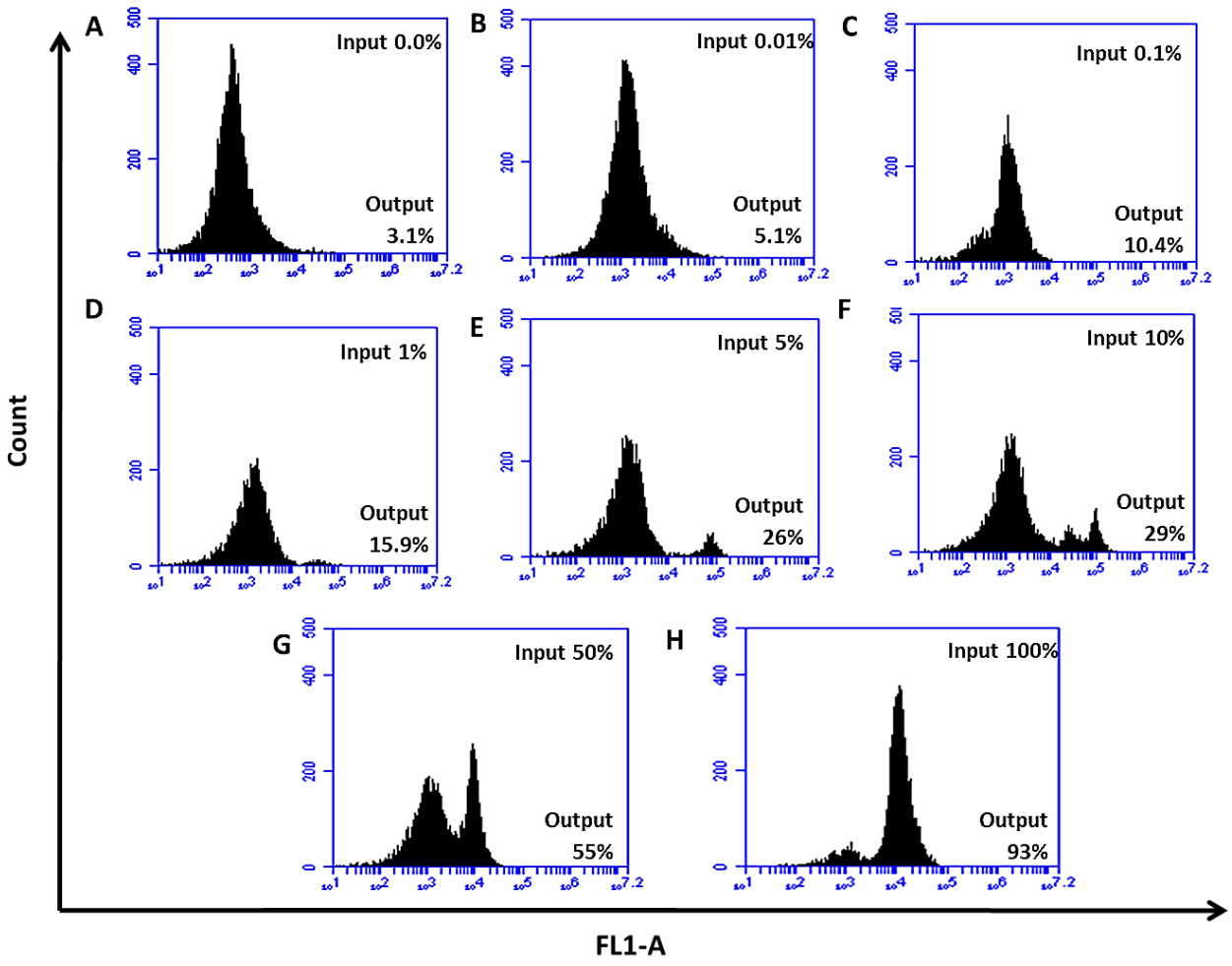
The results showed an increase of the percentage of positive cells recognized by 6-FAM labeled MAMB1 and MAMA2 (output) with increasing the percentage of spiked breast cancer cells (input cells) from 3.1 (Figure 3.7A) to 93 (Figure 3.7H) for MAMB1 aptamer, and from 2.9 (Figure 3.8A) to 93 (Figure 3.8H) for the MAMA2 aptamer. The increase of the percentage of positive cancer cells spiked (X-axis) and the percentage of positive cells identified by flow (y-axis) is linear for both aptamers as indicated by the $R^2 > 0.92$ (Figures 3.7I and 3.8I). Moreover, we can see the appearance of the new population of detected cells at higher fluorescent intensities with increasing percentages of spiked cancer cells for both aptamers (Figures 3.7 and 3.8).

MAMB1 and MAMA2 aptamers showed significant binding ($P<0.05$) to spiked MCF7 and MDA-MB-415 cancer cells respectively, in whole blood lysate containing PBMC cells at 5, 10, 50, and 100% compared to 0% (only PBMC cells). The detection at 0.01, 0.1, and 1% was approaching significance ($P=0.07$). Overall, the results of both aptamers were similar. Therefore, MAMB1 and MAMA2 have great potential to be used as breast CTCs biomarkers. Further testing of both aptamers in real clinical blood samples is required.

The limit of detection (LOD) of MAMB1 and MAMA2 in the spiked experiment was determined using the following formula: Limit of detection = 3* the standard deviation of the blank / by the slope of the calibration curve. LOD value was 4 cells in 10000 for both MAMB1 and MAMA2 aptamers using flow cytometry. Moreover, limit of quantification (LOQ) was determined according t the following formula: Limit of quantification = 10* LOD. The value of LOQ for both aptamers was 40 cells. These results showed that MAMB1 and MAMA2 could detect 4 cells in 10000 of metastatic breast cancer cells using flow cytometry. Other studies showed similar LOD values [30]. Our results indicated that both aptamers could be used in the detection of metastatic breast cancer cells using flow cytometry.

Many methods have been reported to detect different types of metastatic cancer cells in complex media using aptamers [31,32]. These methods includes microfluidic systems [33-35]. These devices are portable and sample and reagent volumes are small, allowing for fast capture of cancer cells. However, flow cytometry allows analysis of large number of cells besides sorting cells on a one-to-one basis [33]. The combination of both microchips and flow cytometry systems could be of great benefit in the field of cancer as separation and high throughput analysis of rare cell populations (such as CTCs) are needed [36]. Other method used aptamer-conjugated nanoparticles (streptavidin-conjugated fluorescent silica nanoparticles) to detect cancer cells from the liver using flow cytometry [37]. This method prevents the photobleaching of the fluorescent labels attached to the aptamer, and thus, increases the specificity of the aptamers [37].

Song *et al.* studied spiked metastatic breast cancer cells into cell media and detected them using the EpCAM aptamer and flow cytometry as the method of detection [38]. The results of this study showed an increase of the mean of fluorescence intensity of the aptamer binding to the cells, as the percentages of spiked cells increased in the cell media [38]. In our case, the flow cytometry method recognized as low as 5% spiked cancer cells (lower limit of detection; Figures 3.7 and 3.8) indicating the potential use of our aptamers to detect CTCs from breast cancer.


A
Input 0.0%
Output
3.1%
B
Input 0.01%
Output
5.1%
C
Input 0.1%
Output
10.4%
D
Input 1%
Output
15.9%
E
Input 5%
Output
26%
F
Input 10%
Output
29%
G
Input 50%
Output
55%
H
Input 100%
Output
93%
Count
FL1-A

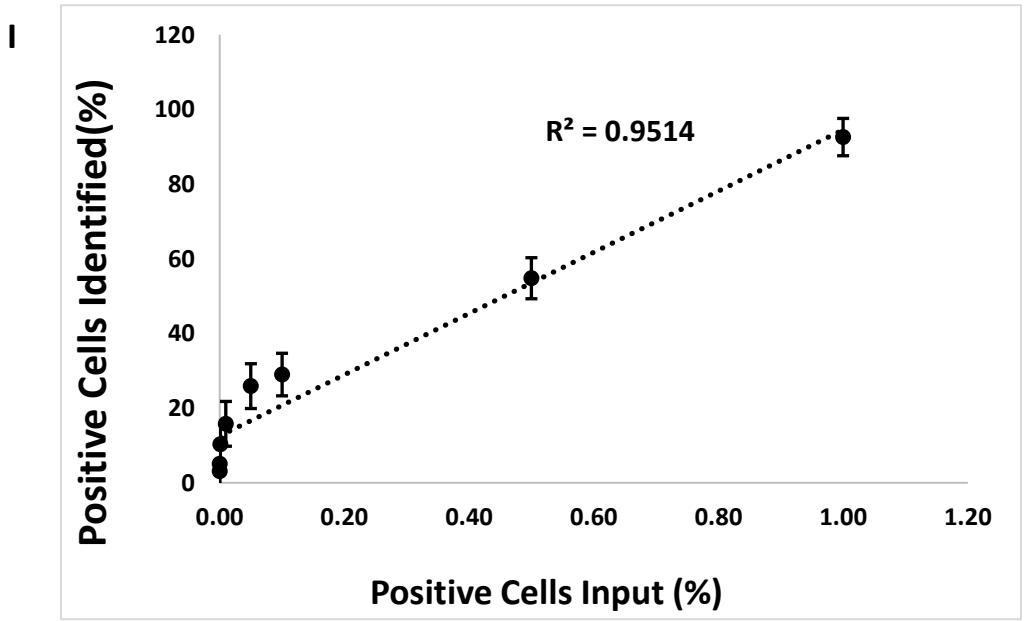
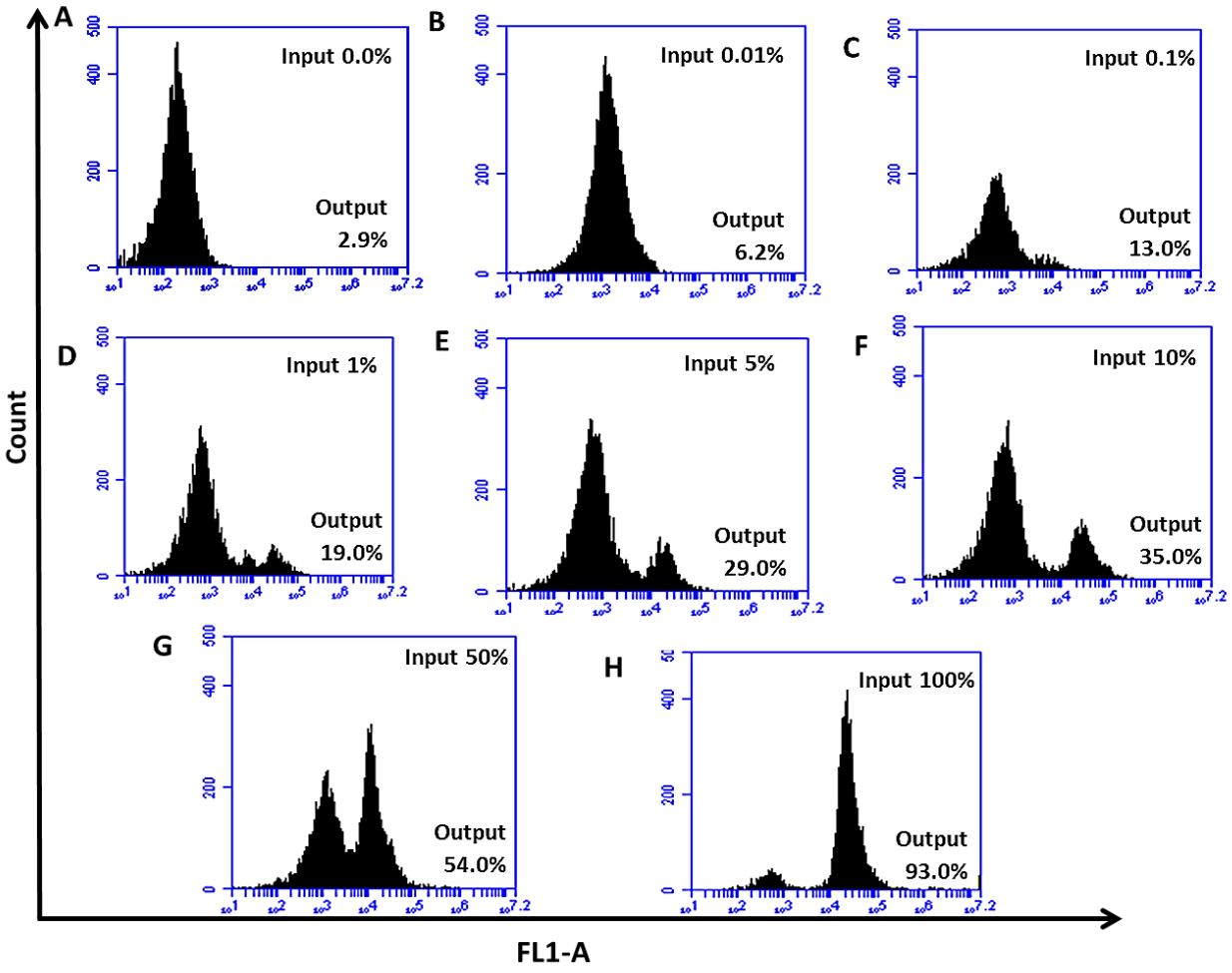

I
R² = 0.9514
Positive Cells Identified(%)
Positive Cells Input (%)
120
100
80
60
40
20
0
0.00
0.20
0.40
0.60
0.80
1.00
1.20

Figure 3.7: Flow cytometric analysis of the recognition of MCF7 cells in spiked whole blood lysate by 6-FAM-labelled MAMB1 aptamer. MCF7 cells were spiked in whole blood lysate containing PBMC in different percentages (A-G: 0%, 0.01%, 0.1%, 1%, 5%, 10%, 50%, and 100%). I: A plot of the positive cells (input) (X-axis) and the positive cells identified (Y-axis). Values are shown as means ± S.E.M. of three trials. The statistical significance was determined by one-way ANOVA followed by Fisher LSD multiple tests using SPSS software (SPSS, version 23). P Values less than 0.05 were considered to be statistically significantly different.


A
Input 0.0%
Output
2.9%
B
Input 0.01%
Output
6.2%
C
Input 0.1%
Output
13.0%
D
Input 1%
Output
19.0%
E
Input 5%
Output
29.0%
F
Input 10%
Output
35.0%
G
Input 50%
Output
54.0%
H
Input 100%
Output
93.0%
Count
FL1-A

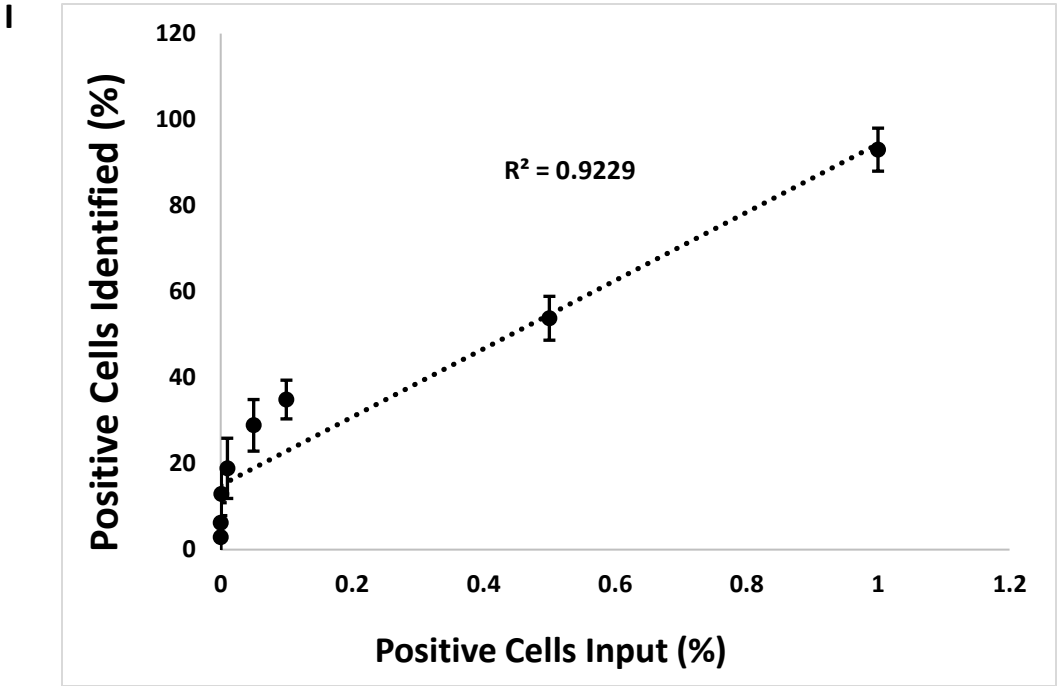

I
Positive Cells Identified (%)
Positive Cells Input (%)
R² = 0.9229
0
20
40
60
80
100
120
0
0.2
0.4
0.6
0.8
1
1.2

Figure 3.8: Flow cytometric analysis of the recognition of MDA-MB-415 cells in spiked whole blood lysate by 6-FAM-labelled MAMA2 aptamer. MDA-MB-415 cells were spiked in whole blood lysate containing PBMC in different percentages (A-G: 0%, 0.01%, 0.1%, 1%, 5%, 10%, 50%, and 100%). I: A plot of positive cells (input) (X-axis) and the positive cells identified (Y-axis). Values are shown as means ± S.E.M. of three trials. The statistical significance was determined by one-way ANOVA followed by Fisher LSD multiple tests using SPSS software (SPSS, version 23). P Values less than 0.05 were considered to be statistically significantly different.

3.4 Conclusion and future work

MAMB1, MAMB12, MAMA2, and MAMA12 showed the highest affinity to their target cancer cells lines, as shown by their K_d values (Figures 3.1 and 3.2). Moreover, all of these aptamers showed specific binding to their target cancer cell lines and not to other cancer cell lines or normal cell lines, as indicated by flow cytometry (Figure 3.3). MAMB1 and MAMA2 showed higher affinity (according to their K_d values) than MAMB12 and MAMA12 and therefore were chosen for further test their binding to breast cancer cell lines, other cancer cell lines and normal cell lines by fluorescence microscopy. Both aptamers showed high affinity to their target cells, and no binding was observed in other cancer cell lines and normal cell lines (Figure 3.3 and 3.4).

Both aptamers bound specifically to their target breast cancer cells in plasma and whole blood lysates (Figure 3.5 and 3.6). In general, MAMB1 is more specific in recognizing its target cancer cells in plasma and whole blood lysate (Figure 3.5D, E, F and 3.6D, E, F). Both aptamers showed great potential to be used in breast cancer CTC detection, as shown by the spiked cancer cells results (Figure 3.7 and Figure 3.8).

Several factors could be further investigated in the future in order to improve binding of both aptamers to cancer cells in complex environments, including modification of the aptamers by adding functional groups at certain nucleotides. It has been reported that modifications (the most common modification is caps at the 3'- and 5'-ends) can increase aptamer resistance to nucleases without affecting their binding to target molecules [38]. Therefore, modification of both aptamers might improve their performance in complex environments such as whole blood. Moreover, studying different fluorophores (such as R-phycoerythrin and rhodamine) and their effect on the overall fluorescence intensity might be useful.

In the spiked cancer cells study (Figures 3.7 and 3.8), investigation of different concentrations of both aptamers showed that both aptamers could detect 4 cells in 10000 cancer cells as indicated by their LOD values.

Investigating the binding of both aptamers *in situ* using tissue samples from breast cancer patients is necessary to verify their binding affinity to breast cancer cells in surrounding tissues. Both aptamers could be labeled to work in immunohistochemistry-like assays to test their specificity to their targets in tissue samples. MAMA2 is only expressed in breast tissue (as mentioned in Chapter 2) so it might show more promising results when used *in situ* compared to whe used *in solution* (*e.g.* whole blood lysate).

References

1. Jing M, Bowser MT. Methods for measuring aptamer-protein equilibria: a review. Anal Chim Acta. 2011, 686(1-2):9-18.
2. Zhang Q, Landgraf R. Selecting molecular recognition. What can existing aptamers tell us about their inherent recognition capabilities and modes of interaction? Pharmaceuticals (Basel). 2012, 5(5):493-513.
3. Bowers WF, Fulton S, Thompson J. Ultrafiltration vs equilibrium dialysis for determination of free fraction. Clin Pharmacokinet. 1984, 9 Suppl 1:49-60.
4. Chang AL, McKeague M, Smolke CD. Facile characterization of aptamer kinetic and equilibrium binding properties using surface plasmon resonance. Methods Enzymol. 2014, 549:451-466.
5. Gokulrangan G1, Unruh JR, Holub DF, Ingram B, Johnson CK, Wilson GSDNA aptamer-based bioanalysis of IgE by fluorescence anisotropy. Anal Chem. 2005, 77(7):1963-1970.
6. Dihua Shangguan, Tao Bing and Nan Zhang. Cell-SELEX: Aptamer selection against whole cells.

   in: Weihong Tan, Xiaohong Fang. Aptamers selected by Cell-SELEX for theranostics (eds). 1st ed. Springer Heidelberg New York Dordrecht London, 2015.
7. Tao Chen, Cuichen Wu, Weihong Tan. Cell-specific aptamer characterization. In: Weihong Tan, Xiaohong Fang. Aptamers selected by cell-SELEX for theranostics (eds). 1st ed. Springer Heidelberg New York Dordrecht London, 2015.
8. Keefe AD, Pai S, Ellington A. Aptamers as therapeutics. Nat Rev Drug Discov. 2010, 9(7):537-550.
9. Rudnick SI, Lou J, Shaller CC, Tang Y, Klein-Szanto AJ, Weiner LM, *et al*. Influence of affinity and antigen internalization on the uptake and penetration of anti-HER2 antibodies in solid tumors. Cancer Res. 2011, 71(6):2250-2259.
10. Liu Z, Duan JH, Song YM, Ma J, Wang FD, Lu X, Yang XD. Novel HER2 aptamer selectively delivers cytotoxic drug to HER2-positive breast cancer cells in vitro. J Transl Med. 2012, 10:148.
11. Shangguan D, Li Y, Tang Z, Cao ZC, Chen HW, Mallikaratchy P, Sefah K, *et al*. Aptamers evolved from live cells as effective molecular probes for cancer study. Proc Natl Acad Sci U S A. 2006, 103(32):11838-11843.
12. Chen Y, Munteanu AC, Huang YF, Phillips J, Zhu Z, Mavros M, Tan W. Mapping receptor density on live cells by using fluorescence correlation spectroscopy. Chemistry. 2009, 15(21):5327-36.
13. Jenison RD, Gill SC, Pardi A, Polisky B. High-resolution molecular discrimination by RNA. Science. 1994, 263(5152):1425-1429.
14. Haller AA, Sarnow P. *In vitro* selection of a 7-methyl-guanosine binding RNA that inhibits translation of capped mRNA molecules. Proc Natl Acad Sci U S A. 1997, 94(16):8521-85216.
15. Sassanfar M, Szostak JW. An RNA motif that binds ATP. Nature. 1993, 364(6437):550-553.
16. Mannironi C, Di Nardo A, Fruscoloni P, Tocchini-Valentini GP. *In vitro* selection of dopamine RNA ligands. Biochemistry. 1997, 36(32):9726-9734.
17. Yüce M, Ullah N, Budak H. Trends in aptamer selection methods and applications. Analyst. 2015, 140(16):5379-5399.
18. Hermann T, Patel DJ. Adaptive recognition by nucleic acid aptamers. Science. 2000, 287(5454):820-825.

19. Batey RT, Gilbert SD, Montange RK. Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine. Nature. 2004, 432(7015):411-415.
20. Jiang F, Kumar RA, Jones RA, Patel DJ. Structural basis of RNA folding and recognition in an AMP-RNA aptamer complex. Nature. 1996, 382(6587):183-186.
21. Davis KA, Lin Y, Abrams B, Jayasena SD. Staining of cell surface human CD4 with 2'-F-pyrimidine-containing RNA aptamers for flow cytometry. Nucleic Acids Res. 1998, 26(17):3915-3924.
22. Ninomiya K, Kaneda K, Kawashima S, Miyachi Y, Ogino C, Shimizu N. Cell-SELEX based selection and characterization of DNA aptamer recognizing human hepatocarcinoma. Bioorg Med Chem Lett. 2013, 23(6):1797-1802.
23. Almasi F, Mousavi Gargari SL, Bitaraf F, Rasoulinejad S. Development of a single stranded DNA aptamer as a molecular probe for LNCap cells using cell-SELEX. Avicenna J Med Biotechnol. 2016, 8(3):104-111.
24. Fan P, Suri AK, Fiala R, Live D, Patel DJ. Molecular recognition in the FMN-RNA aptamer complex. J Mol Biol. 1996, 258(3):480-500.
25. McKeague M, McConnell EM, Cruz-Toledo J, Bernard ED, Pach A, Mastronardi E, *et al*. Analysis of *in vitro* aptamer selection parameters. J Mol Evol. 2015, 81(5-6):150-161.
26. Jurek PM, Zabłocki K, Waśko U, Mazurek MP, Otlewski J, Jeleń F. Anti-FGFR1 aptamer-tagged superparamagnetic conjugates for anticancer hyperthermia therapy. Int J Nanomedicine. 2017, 12:2941-2950.
27. Yoon JW, Jang IH, Heo SC, Kwon YW, Choi EJ, Bae KH, *et al*. solation of foreign material-free endothelial progenitor cells using CD31 aptamer and therapeutic application for ischemic injury. PLoS One. 2015, 10(7):e0131785
28. Holahan MR, Madularu D, McConnell EM, Walsh R, DeRosa MC. Intra-accumbens injection of a dopamine aptamer abates MK-801-induced cognitive dysfunction in a model of schizophrenia. PLoS One. 2011, 6(7):e22239.
29. Yang J and Weinberg RA. Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. Dev Cell. 2008, 14(6):818-829.
30. Warawdekar UM, Parmar V, Prabhu A, Kulkarni A, Chaudhari M, Badwe RA. A versatile methodology for the enumeration and characterization of circulating tumor cells from patients with breast cancer. J Cancer Metastasis Treat 2017, 3:23-33.
31. Hassan EM, Willmore WG, DeRosa MC. Aptamers: promising tools for the detection of circulating tumor cells. Nucleic Acid Ther. 2016, 26(6):335-347.
32. Dickey DD, Giangrande PH. Oligonucleotide aptamers: A next-generation technology for the capture and detection of circulating tumor cells. Methods. 2016, 97:94-103.
33. Murthy SK, A Sin, RG Tompkins and TM Langmuir. Effect of flow and surface conditions on human lymphocyte isolation using microfluidic chambers. Langmuir. 20:11649–11655.
34. Du Z, N Colls, KH Cheng, MW Vaughn and L Gollahon. Microfluidic-based diagnostics for cervical cancer cells. Biosens Bioelectron. 2006, 21:1991–1995.
35. Nagrath S, LV Sequist, S Maheswaran, DW Bell, D Irimia, L Ulkus, MR Smith, EL Kwak, S Digumarthy, *et al*. Isolation of rare circulating tumor cells in cancer patients by microchip technology. Nature. 2007, 450:1235–1239.

35. Piyasena ME, Graves SW. The intersection of flow cytometry with microfluidics and microfabrication. Lab Chip. 2014, 14(6):1044-1059.
36. Hu Z, Tan J1, Lai Z, Zheng R, Zhong J, Wang Y, *et al*. Aptamer combined with fluorescent silica nanoparticles for detection of hepatoma cells. Nanoscale Res Lett. 2017, 12(1):96.
37. Song Y, Zhu Z, An Y, Zhang W, Zhang H, Liu D, et al. Selection of DNA aptamers against epithelial cell adhesion molecule for cancer cell imaging and circulating tumor cell capture. Anal Chem. 2013, 85(8):4141-4149.
38. Hasegawa H, Savory N, Abe K, Ikebukuro K. Methods for improving aptamer binding affinity. molecules. 2016, 21(4). pii: E421.

CHAPTER 4

Identification of MAMB1 and MAMA2 targets on the surface of breast cancer cells and protein studies

4.1 Introduction

4.1.1 Aptamer target Identification after cell-SELEX

Cell-SELEX has been mentioned in the details presented in Chapter 1. Moreover, cross-over SELEX or hybrid SELEX has been outlined in Chapter 2. In this Chapter, target identification after cell-SELEX will be discussed in detail. It is known that cell-SELEX is a powerful tool to identify new biomarkers, especially in cancer research where identifying new biomarkers helps to improve the diagnosis and therapy of cancer patients. In order to be used in cancer diagnosis or therapy, aptamer-target identification is an essential step.

Affinity pull-down of targets by the aptamers, followed by mass spectrometry (MS), is the most popular and widely used method to identify aptamer targets [1,2]. Only a few targets have been identified via whole cell-SELEX against cancer cell lines. Among them, none were identified for breast cancer (Table 4.1). One of the first aptamer targets identified by this method was tenascin-C (a 250 kDa protein) after whole cell-SELEX on glioblastoma cell line (Table 4.1) [3]. Target identification by pull-down-MS approach faces many challenges including the nature of the target proteins on the surface of cells. These proteins are extremely difficult to work with due to the fact that most of them are plasma membrane proteins, which are hydrophobic and therefore have low solubility [1].

Another obstacle is that the yield of the protein pulled by the aptamers is usually very low, which makes MS a challenging task [1].

Target Expressed on Cell Surface-SELEX (TECS-SELEX) is a modified whole cell-SELEX approach which has been established to develop aptamers against a specific target on the surface of the cell that is overexpressed [4]. This approach allows for the selection of aptamers against a specific target without the need to identify the aptamer target at the end of the SELEX process. An RNA aptamer S6 against HER-2 protein was identified by this method through transfection of its gene in a SK-BR-3 breast cancer cell line [5]. In another study by Song *et al*. EpCAM was transfected into HEK-293T cells and used as the target cell line to develop an aptamer (SYL3) that specifically binds to EpCAM in multiple cancers including breast cancer [6]. Moreover, the SYL3 aptamer captured breast cancer cells in a mixture of cells with 63% capture efficiency and 80% purity [6].

Table 4.1: Target identification via whole cell-SELEX against cancer cell lines

Aptamer name	Aptamer type	Target cancer cell line	Cell-SELEX type	Aptamer target	Ref.
GBI-10	DNA	Glioblastoma cell line U251	Whole cell-SELEX	Tenascin-C	[3]
Sgc8	DNA	T-cell acute lymphoblastic leukemia CCRF-CEM	Whole cell-SELEX	Tyrosine-protein kinase-like 7 (PTK7)	[7]
TD05	DNA	Burkitt's lymphoma cell line Ramos	Whole cell-SELEX	Ig μ heavy chain	[8]
TOV6	DNA	Ovarian clear cell adenocarcinoma TOV-21G	Whole cell-SELEX	STIP1 stress induced phosphoprotein 1 (STIP1)	[9]
K19	DNA	Acute myelogenous leukemia (AML) cell line NB4 AML	Whole cell-SELEX	Siglec-5	[10]

4.1.2 MGB2 and MGB1 as secretory proteins

[14]. MGB2 and MGB1 are both secretory proteins [11,12]. Therefore, they might be detected in the plasma or serum of breast cancer patients. MGB1 has been reported as a breast cancer-specific protein [13] and thus many studies have tried to measure its levels in breast cancer patients. The first study reported that MGB1 was detected in the serum of breast cancer patients by ELISA (0.07-9.6 ng/mL in breast cancer patients compared to 0-0.07 ng/mL in cancer free individuals) [14]. In a second study measuring the levels of MGB1 in serum of breast cancer patients using ELISA it was confirmed that MGB1 levels were affected by the cancer stage and tumor size. Stages I-III had 0.9 to 1.4 ng/mL of MGB1 and 2.3 ng/mL had been reported for stage IV patients. 1.71 ng/mL was set as the cut off value for that study [15]. In a third study, MGB1 levels were the highest in metastatic breast cancer patients (9.38 ng/mL vs 7.9 ng/mL control). The cut-off value for was 8.8 ng/mL [16]. The sensitivity and specificity for that study was 68 and 88.8% respectively [16]. The differences in the cut off values of both studies depend on the antibodies used to measure the levels of MGB1 in serum and whether the antibodies are monoclonal [15] or polyclonal [16]. In a more recent study, which used peptides instead of antibodies as the immunogen (four polyclonal antisera), it was shown that better discrimination of MGB1 levels was achieved between control and breast cancer patients with a sensitivity of 86.3% and a specificity of 96% [17]. In the last study, however, the levels of MGB1 were not correlated to cancer stage [17]. Moreover, no correlation was found between levels of MGB1 and age nor menopausal age [17].

In contrast, no studies have been reported on the measurement of MGB2 levels in the serum of human breast cancer patients using ELISA or other methods. MGB2 mRNA has been detected in the peripheral blood of breast cancer patients using RT-PCR [18]. The lack of such studies for MGB2 protein could be related to the fact that MGB2 is overexpressed in many of the epithelial origin cancers (as mentioned in Chapter 2).

Aptamers have the potential to replace antibodies and could be used for the detection of MGB1 in the serum of breast cancer patients resulting in a more specific and more reliable method of detection. This chapter will seek to confirm whether MGB2 and MGB1 are the actual targets that the aptamers interact with on the surface of breast cancer cells.

4.2 Materials and Methods

4.2.1 Materials and Cell Lines

MGB1 and MGB2 recombinant proteins, as well as MGB2 and MGB1 monoclonal antibodies were purchased from Novus Biologicals (Canada). Dynabeads® M-280 Streptavidin magnetic beads were purchased from Thermo-Fisher Scientific (Canada). Nonenzymatic cell dissociation buffer and Dulbecco's Phosphate-Buffered Saline (DPBS) with calcium and magnesium were purchased from Sigma (Canada). Biotin (B20656) was purchased from Invitrogen. MGB1 and MGB2 plasmids (10µg) were purchased from Cedarlane (Canada). All other chemicals were ordered from Bioshop (Canada) unless otherwise indicated. All buffers were prepared using Milli-Q water and filtered using 0.2µm filters. MCF-7, MDA-MB-415, and HEK293 cell lines were used in this part of study and obtained from the American Type Culture Collection (ATCC).

4.2.2 Purification of MAMB1 and MAMA2 targets in cell surface proteins extract of MCF7 and MDA-MB-415 cells by pull-down assay followed by SDS and mass spectroscopy (MS) analysis

4.2.2.1 Cell lysis and surface proteins extraction

Cells lysis and protein extraction protocol was modified from Daniels *et al.* [3]. Briefly, MCF7 and MDA-MB-415 cells were plated ($1X10^7$ cells from each cell line in 15 cm tissue culture plates), then cells were washed with DPBS. Then, 2 mL of hypotonic solution (10 mM Tris·HCl (pH 7.5) containing 5 mM KCl and 0.5 mM MgCl2 with cOmplete ^{TM} protease inhibitors) was added to the cells and incubated for 30min at 4°C. Scrapers were used to harvest the cells, and cells were homogenized in 10 mL of 250 nM sucrose solution using a Dounce tissue grinder. To remove whole cells and nuclei, the mixture was centrifuged at 1,000 x g for 10 min at 4˚C (this step was repeated if needed). After that, the membrane fractions were obtained by centrifuging the supernatant at 105,000 × g for 1 h at 4°C using an ultra-centrifuge. The membrane fractions were allowed to solubilize overnight on a shaker at 4˚C in extraction buffer (10 mM Tris·HCl (pH 7.5) containing 200 mM NaCl, 0.1% Triton X-100, 1 mM MgCl2, 1 mM CaCl2, and cOmplete^{TM} protease inhibitors) The pellet now was stored until use for pull-down assay (Scheme 4.1A). In a second trial, the salt concentrations were changed in the hypotonic extraction buffers to be as the following: hypotonic buffer, 20 mM Tris·HCl (pH 7.5) containing 10 mM KCl and 1.5 mM $MgCl_2$

with cOmplete protease inhibitors. For the extraction buffer, 20 mM Tris·HCl (pH 7.5) containing 250 mM NaCl, 0.1% Triton X-100, 5 mM MgCl2, 3 mM CaCl2, and cOmplete protease inhibitors.

4.2.2.2 Pull-down assay using MAMB1 and MAMA2 biotinylated aptamers followed by SDS-PAGE and MS analysis

MAMB1 and MAMA2 biotinylated aptamers were synthesized using the DNA synthesizer, as described in Chapter 2. Streptavidin magnetic beads were washed according to the manufacture instructions. Then, 200 pmols of each aptamer was incubated with 1 mg of the beads in 1 X binding and washing buffer (according to manufacturer's instructions) for 15 min at RT with gentle agitation. After that, the membrane fractions of each cell line were incubated with their corresponding aptamer-beads in affinity purification buffer (extraction buffer containing 0.5 µg/µL BSA and 4% glycerol) and 300 µg/mL of ssDNA from salmon sperm to a total volume of 1 mL. The mixture was then incubated at 4˚C for 1 hr with agitation. The beads then were washed with extraction buffer 3 times each for 10 min at 4°C. Then the bound proteins from each aptamer-beads complex was eluted in 10 mM Tris·HCl (pH 7.5) containing 1 M NaCl, 5 mM EDTA, and 0.1% Triton X-100 elution buffer (Scheme 4.1B).

Aptamer purified proteins (the pellet) were loaded onto 19% SDS-PAGE, and the gel was stained with R-250 Coomassie blue (Bio-Rad, Canada) for 1 hr at RT with shaking. The gel was allowed to destain overnight in DI water, and protein bands at low molecular weight (10-15 KDa) were cut and tryptic digestion was applied. Then protein samples were sent to the MS center at Carleton University (Scheme 4.1B).

In a third trial of cell-lysis followed by pull down assay, a new method was applied which was modified from Simaeys *et al*. [9]. Briefly, MCF7 and MDA-MB-415 cells were plated (1 X 10^8 cells from each cell line in 15 cm tissue culture plates) and washed with nonenzymatic cell dissociation buffer (Sigma, Canada) to remove them from the culture plates. Each aptamer (MAMB1 and MAMA2) was then cross-linked to its corresponding target cells by incubating them for 2 min in a 1% formaldehyde/PBS solution. The cells were then immediately washed (three times) at 4°C in PBS to remove the formaldehyde. Cells then were lysed in cell lysis buffer contained 2% Triton-X100 (Fisher), 1.5% Nonidet, and 0.5% cholate using a Dounce tissue grinder.

The lysate was then incubated overnight at RT in the presence of 200 µg of streptavidin magnetic beads. After that, the lysate aptamer complex was washed with washing buffer (10 mM HEPES/NaOH (pH 7.8), 100 mM NaCl, 2 mM ethylenediaminetetraacetic acid (EDTA), 1 mM ethylene glycol tetraacetic acid (EGTA), 0.2% sodium dodecyl sulfate (SDS), and 0.1% sodium lauroyl sarcosinate (SLS)) until the beads were clear (about 3-4 times with 1 mL of washing buffer each). Then the aptamer-lysate (proteins) complex was eluted by adding 50 µL of elution buffer (12.5 mM biotin in 7.5 mM HEPES/NaOH (pH 7.8), 75 mM NaCl, 1.5% EDTA, 0.5% EGTA, 0.15% SDS, and 0.075% SLS). Proteins were precipitated by adding cold acetone and incubated for 3 hours at -20°C. Proteins were then centrifugated at 12,000 x g for 10 min at 4°C. The pellet was then dried to be used in SDS-PAGE.

4.2.3 Purification of MAMB1 target in cell surface protein extract of MCF7 cells by pull-down assay followed by Western blot analysis

The cell lysis and protein extraction was done the same as above. For the pull-down assay it was only applied for MAMB1 aptamer on the beads. The aptamer purified protein was loaded onto 19% SDS-PAGE, and then the gels were transferred onto a Polyvinylidene Difluoride (PVDF) membrane for a standard Western blot analysis. An MGB2 monoclonal antibody (1:500) was diluted in 3% BSA-0.1%TBST and incubated with the first membrane overnight at 4°C. Then the membrane was washed with Tris Buffer Saline Tween-20 (0.05%) (TBST) for 5 times each 5 min. A secondary HRP antibody (1:2,000) diluted in the same diluent was incubated with the membrane for 1 hr at RT then the membrane was washed as above then an HRP chemiluminescence substrate was added (Bio-Rad) and the gel was subjected for imaging using Alpha Innotech gel imager documentation system (USA). For the second membrane, MAMB1 biotinylated aptamer (1µg/mL) diluted with 3% BSA-0.1% Phosphate=Buffered Saline Tween-20 (PBST) was incubated at 4˚C for 2 hrs. Then the membrane wascwashed with PBST for 5 times each 5 min. Streptavidin-HRP (1:5,000) diluted in the same diluent was incubated with the membrane for 1 hr at RT then the membrane was processed for imaging same way as the first membrane (Scheme 4.1B).

4.2.4 Over expression of MGB2 and MGB1 Proteins in HEK293 cells by transfection

In order to determine whether MGB2 and MGB1 proteins are the targets of MAMB1 and MAMA2 respectively, human embryonic kidney (HEK293) cells were transfected with MGB2 and MGB1 overexpression plasmids. First, the expression levels of both proteins was determined in HEK293 cells using MGB2 and MGB1 ELISA kits (mentioned in Chapter 2). Second, MGB1 and MGB2 plasmids were transformed into DH5α competent cells [19] and then purified using standard midiprep system. Third, MGB1 and MGB2 purified plasmids were transfected into HEK293 cells according to Longo et al. and others [21-23]. Briefly, HEK293 cells (5 X 10^5 cells) were plated in 60 mm tissue culture dishes and left to grow for 24 hours in the incubator (37°C, 5% CO_2, 95% humidity). The next day, the cell culture media was replaced and HEK293 cells were left in the incubator for another 30 min. The DNA (plasmids) were prepared for transfection; 10 µg of each purified MGB1 and MGB2 plasmids were added to 250 µL of Opti-MEM media, then 50 µL (final concentration of 5 µg/mL) of polyethylenimine (PEI) was added to 250 µL of Opti-MEM media, in separate tubes. Then the PEI mix was added to the DNA (plasmids) mix and incubated for 15 min at RT. The PEI-DNA was mixed by vortex every 3 min for 5 sec. each, enabling the uptake of the DNA by PEI to form small insoluble aggregates (polyplexes), facilitating its delivery into cells [24]. After that, the PEI-DNA complex was added to the cells gradually drop-by-drop, and HEK293 cells were left to grow for 48 hours. PUC19 empty vector was used as a control. Green Fluorescence Protein (GFP) was used to validate the transfection efficiency. Both pUC19 and GFP plasmids were infected using the same amount of DNA as MGB2 and MGB1 plasmids under the same conditions and time. After 48 hours, HEK293 cells were harvested and fixed with 4% formaldehyde for 15 min at RT. Then washed and resuspend in PBS to be used in flow cytometry analysis.

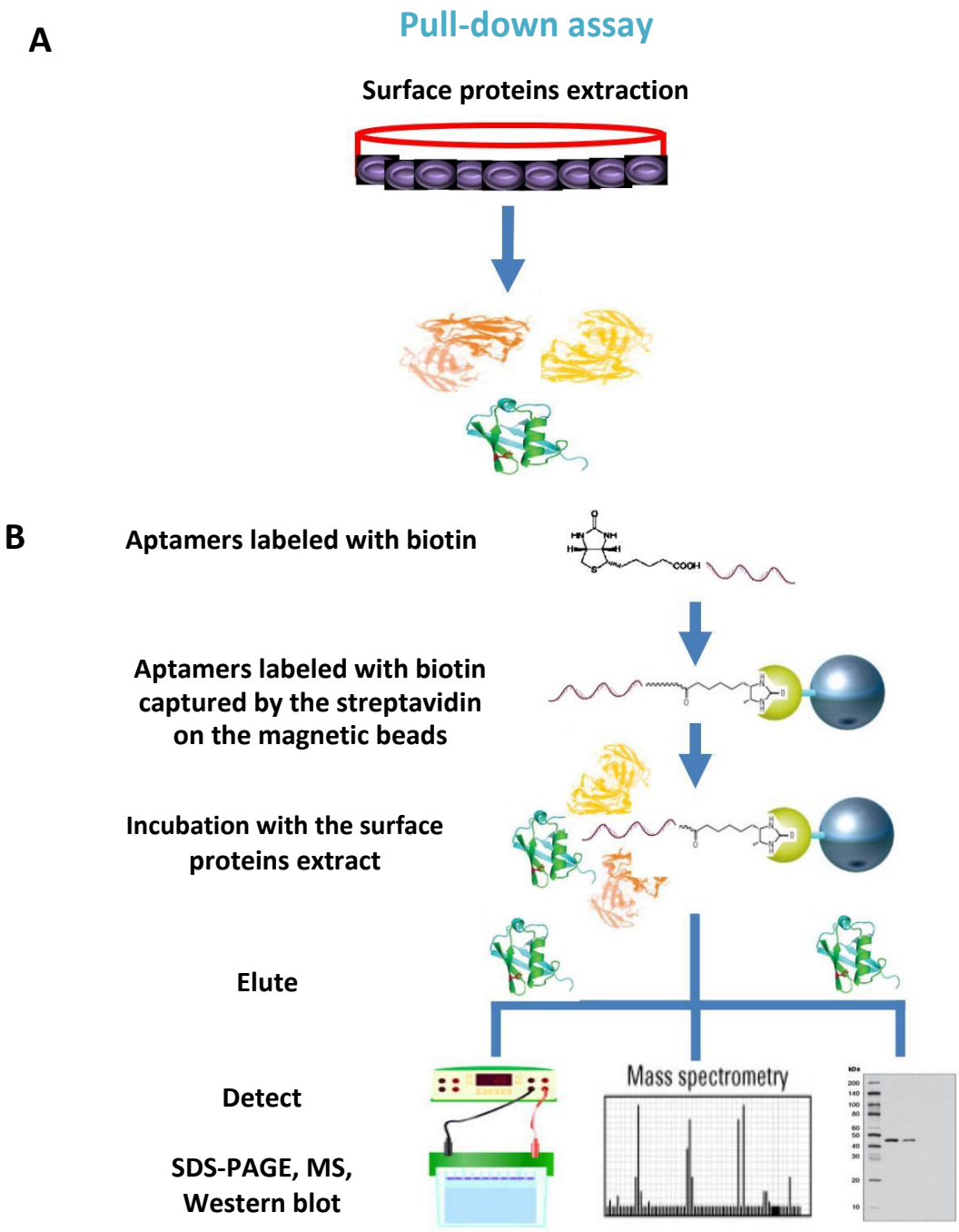
4.2.5 Studies of aptamer-antibody competition by flow cytometry

Anti-MGB2 and anti-MGB1 monoclonal antibodies, goat anti-mouse FITC-labelled secondary antibody, 6-FAM-labelled and unlabeled MAMB1 and MAMA2 aptamers, non-transfected and transfected HEK293 cells, as well as MCF7 and MDA-MB-415 cells, were used in a competition study. Briefly, 5 X 10^5 fixed cells of each cell line mentioned above was added to 500 µL of 5 mM $MgCl_2$ in PBS, then competition studies were performed as follows. 1) non-transfected HEK293, MCF7 and MDA-MB-415 cells were incubated with 300 nM of 6-FAM-labelled aptamers (MAMB1 with MCF7 cells, MAMA2 with MDA-MB-415 cells, and each of them with HEK293 cells) or a control random sequence for 30 min at 4°C, then centrifuged at 1,000 x g for 5 min. Cells then were re-suspended with 500 µL of 5 mM $MgCl_2$ in PBS. 2) Non-transfected HEK293, MCF7 and MDA-MB-415 cells were incubated with anti-MGB2 and anti-MGB1 primary monoclonal antibodies (1:400 of 100 µg/0.5 mL, anti-MGB2 with MCF7 cells, anti-MGB1 with MDA-MB-415 cells, and each of them with HEK293 cells) for 30 min at RT. Cells then were centrifuged at 1,000 x g for 5 min and washed three times with PBS. Goat anti-mouse FITC labelled secondary antibody (1:500 of 100 µg/ 0.5 mL) was added for 20 min at RT. Finally, cells were centrifuged and washed as above, then re-suspended with 500 µL of PBS, 3) Transfected HEK293 cells were incubated with either 300 nM 6-FAM-labelled aptamers, or unlabeled aptamers, or random sequence for 30 min at 4°C. Transfected cells were then incubated with anti-MGB2 or anti-MGB1 primary monoclonal antibodies with goat anti-mouse-FITC labelled secondary antibody. All cells were studied using flow cytometry by counting 10,000 events and the data was analyzed using the flow cytometry (with C6 sampler software). The fluorescence intensity of the goat anti-mouse secondary antibody alone was subtracted from the samples where primary and secondary antibodies were used together to obtain the fluorescence intensity.

4.2.6 Protein-aptamer study using electrophoretic mobility shift assay

Electrophoretic Mobility Shift Assay (EMSA) was used to investigate the binding of MAMB1 and MAMA2 aptamers to their target recombinant proteins MGB2 and MGB1 respectively. The recombinant proteins used in this study were the same proteins used in the protein SELEX part (Chapter 2).

Briefly, different concentrations of target proteins (0-220 nM) were incubated with fixed concentrations of 6-FAM-MAMB1 and 6-FAM-MAMA2 aptamers (600 nM) in 50 mM Tris-HCl, 10 mM reduced glutathione, pH 8.0 for 1 hr at 37°C. Protein-aptamer complexes were then loaded onto 8% non-denaturing PAGE and allowed to run for 3 hours at 160 V at RT. The gels were then visualized using an Alpha Innotech gel imager documentation system. For the control experiments, a random sequence (different DNA aptamer, 6-FAM-60 bases), and Bovine Serum Albumin (BSA) were used as a control aptamer and a control protein respectively. They were treated in the same manner as experimental samples (MAMB1 and MAMA2 aptamers with their target proteins). Control aptamer-protein complexes were loaded to PAGE and visualized at the same time as the experimental samples.



Scheme 4.1: Schematic diagram of the experimental design of pull-down assay using MAMB1 and MAMA2 aptamers on cell surface protein extracts of MCF7 and MDA-MB-415 cells. A: Surface protein extraction on MCF7 and MDA-MB-415 cells by applying different buffers to lysis the cell component. B: pull-down assay steps using Biotinylated MAMB1 and MAMA2 and streptavidin magnetic beads to pull- down the protein (proteins) aptamers binding then analyze it using SDS-PAGE, MS, and Western blot.

4.3 Results and Discussion

4.3.1 SDS-PAGE and mass-spectroscopy analysis of MAMB1 and MAMA2 after applying pull-down assay on MCF7 and MDA-MB-415 surface proteins extract

In our study, a hybrid SELEX approach was applied against MGB2 and MGB1 proteins, therefore the intended target was known before cell-SELEX was performed. Moreover, the ssDNA libraries were previously enriched to MGB2 and MGB1 proteins before cell-SELEX was applied. Therefore, it would be expected that the targets which the aptamers are binding to on the surface of the cells are MGB2 for MCF7 cells and MGB1 for MDA-MB-415 cells. Additionally, MAMB1 and MAMA2 aptamers would be expected to bind to their recombinant protein targets in solution. Herein, a series of experiments were done to verify the identity of the targets on the surface of the breast cancer cell lines MCF7 and MDA-MB-415. Moreover, a protein study was performed to test the selective binding of both aptamers to their target proteins in solution.

The most common way to identify a target protein on the surface of the cell, after pulling down the proteins using the candidate aptamers and cell-SELEX, is MS (as mentioned earlier in this Chapter). None of the examples mentioned in Table 1 were hybrid SELEX selections and thus applying this type of SELEX does not typically required the identification of the target on the cell surface because the ssDNA library alternates between the cell and the recombinant protein form of the target of interest. Nevertheless, protein extraction of the cell surface proteins and MAMB1 and MAMA2 aptamers were used to pull-down their target proteins, then an SDS-PAGE analysis was performed. As Figure 4.1A shows, MAMB1 was successfully able to bind to low molecular weight proteins. The cell has many proteins that could be similar to MGB2 protein and may bind MAMB1 aptamer non-specifically (Figure 4.1A Elution 1 and 2), but when the pull-down was repeated a second time most of the non-specific binding was eliminated (Figure 4.1A Elution 3). MAMA2 was able to bind to the low molecular weight protein (its target), but the bands are not clear indicating that MAMA2 bound to a very small amount of the protein. Another factor that could affect the results in Figure 4.1B is SDS as it consider being harsh detergent. Perhaps MAMA2 bound to its target but SDS could degrade the native form of the protein leading to a lack of a clear signal at the right molecular weight.

After that, a tryptic digestion was done to extract the protein peptides that both aptamers bound to, then samples were sent to MS. MGB2 and MGB1 proteins, unfortunately, were not on the list of proteins we obtained after MS analysis was performed. One of the reasons behind the fact that MGB2 and MGB1 proteins are very small proteins (around 10 kDa) and need special techniques to isolate them from other proteins that have the same molecular weight such as histones (which are more abundant and bind to DNA). Moreover, it could be that the expression level of our proteins is low compared to other similar proteins and therefore these proteins masked our proteins of interest. In addition, the presence of histones and ribosomal protein in MS analysis results revealed some kind of contamination of cytoplasmic and nuclear proteins within the gel. Therefore, several other experiments were done in which new conditions for cell lysis and extraction of membrane fractions were applied, but the MS results obtained did not show MGB2 or MGB1. The last trial was performed using a cross linking technique for the aptamer-target while the cells are intact then a cell-lysis was performed and the cross-linking was recovered [9]. Unfortunately, the MS results did not show our target proteins (Table 4.2 showed the list of proteins obtained after mass analysis form the three different trials). However, all the SDS-PAGE images obtained from the different protocols showed a protein band at approximately 10 kDa for both aptamers same as the one in Figure 4.1.

4.3.2 Western blot analysis of MAMB1 after applying pull-down assay on MCF7 surface proteins extract

A Western blot analysis comparing the binding of MGB2 monoclonal antibody and MAMB1 aptamer to the pull- down elutions from MAMB1-protein complex on MCF7 was performed. The results showed that MAMB1 aptamer is more specific in binding to the pull-down extract than MGB2 antibody (Figure 4.2A and B). Moreover, MAMB1 could recognize the amount of MGB2 protein in E2 (which was obtained after re-purification) while the antibody could not (Figure 4.2B, lane E2) but recognized MGB2 in the total protein as there is higher amount of it there (Figure 4.2B lane total protein). The results obtained from Western blot suggest that MGB2 protein is present in the pull-down extract and that aptamers specifically recognized this protein.

Table 4.2: List of proteins identified from mass-spectroscopy after pull-down assay of MCF7 and MDA-MB-415 cells using MAMB1 and MAMA2 aptamers respectively.

Cell- lysis and pull-down assay for both MCF7 and MDA-MB-415 using MAMB1 and MAMA2 aptamers	**List of proteins name obtained from mass spectroscopy analysis**
Trial 1: Unmodified method as mentioned in the Materials and Methods section and Reference 3	Histone H2B type 1-M (*Homo sapiens*) Histone H4 (*Homo sapiens*) 40S ribosomal protein S19 (*Homo sapiens*)
Trial 2: Salt modifications as mentioned in the Materials and Methods section	60S ribosomal protein L23 (*Homo sapiens*) 60S ribosomal protein L35 (*Homo sapiens*) cytokeratin 9 (*Homo sapiens*)
Trial 3: Cross-linking method as mentioned in the Materials and Method section and Reference 9	unnamed protein product synsav (synthetic construct)

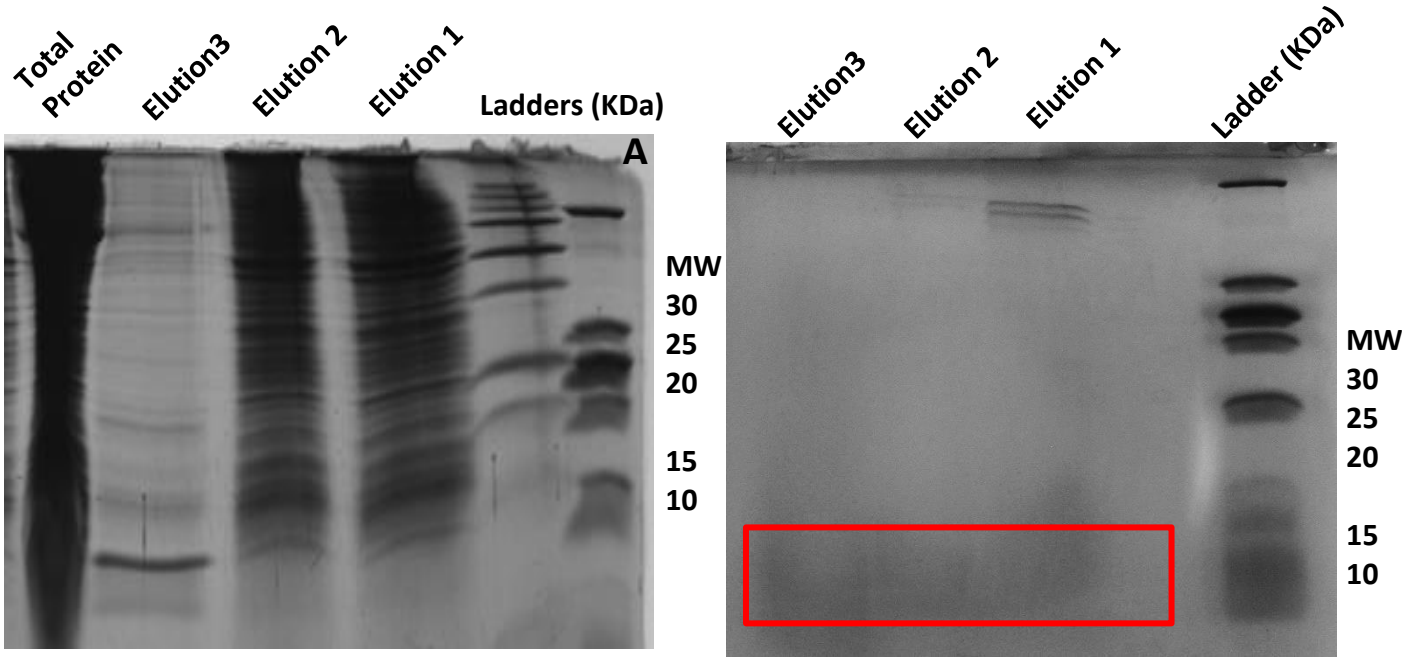


Figure 4.1: SDS-PAGE image for MAMB1 and MAMA2 aptamer-purified proteins. Both aptamers were used to purify proteins from cell surface extract proteins of MCF7 and MDA-MB-415 breast cancer cells. A: MAMB1-protein aptamer SDS-image. B: MAMA2-protein aptamer SDS-image. Elution 1 and elution 2 were obtained after applying the elution buffer, elution 3 obtained after repeating the pull-down on the same extract (re-purification) then the beads were washed and boiled. Sample of the protein extract before applying the purification was loaded only in the case of MAMB1. Images were taken using Alpha Innotech gel imager documentation system.

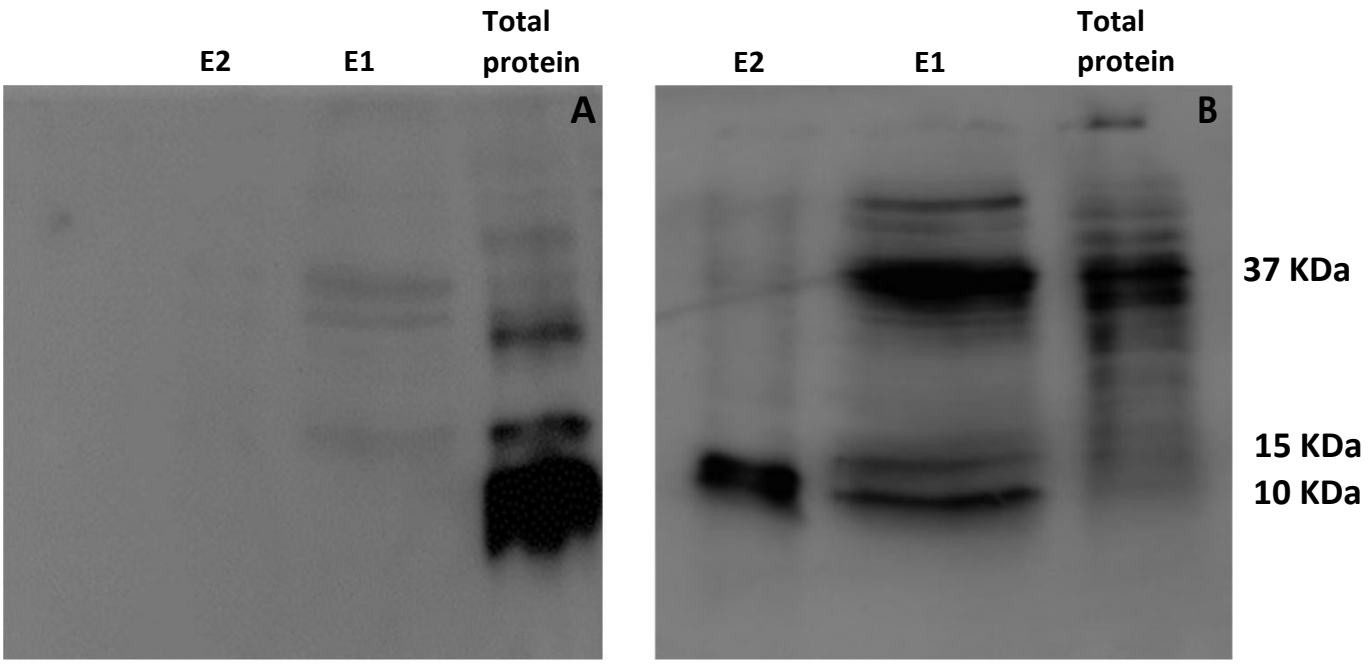


Figure 4.2: Western blot analysis of MAMB1 aptamer-protein complex after performing the pull-down assay on MCF7 surface protein extracts. A: probing with MGB2 monoclonal antibody. The PVDF membrane was incubated with MGB2 monoclonal antibody then incubated with HRP-secondary antibody and visualized using a gel documentation system. B: probing with biotinylated MAMB1 aptamer. The PVDF membrane was incubated with biotinylated MAMB1 aptamer, incubated with streptavidin-HRP secondary antibody and visualized using a gel documentation system. The bands that both the monoclonal antibody and MAMB1 aptamer bound to were at 37, 15, and 10 kDa. The lanes from right to left are: total protein which is the protein extract before applying the pull-down assay, E1: elution after first time purification using the beads, E2: elution after second time using the same beads.

4.3.3 MGB2 and MGB1 ELISA and transfection results

Human embryonic kidney cells (HEK293) are widely used as a host for gene expression [25]. Therefore, HEK293 cells were transfected with human MGB2- and MGB1-expressing plasmids in this study. To investigate the expression levels of MGB2 and MGB1 proteins in HEK293 cells before transfection, ELISA was performed on HEK293 cell lysates. As Figure 4.3 shows, HEK293 cells express very low levels of MGB2 and MGB1 proteins (1.78 and 1.56 ng/mL respectively) compared to breast cancer cell lines MCF7 and MDA-MB-415 (16.19 and 21.98 ng/mL respectively) (Figure 4.3). These results matched our hypothesis, which suggested that HEK293 cells did not express breast cancer biomarkers.

In order overexpress MGB2 and MGB1 protein in HEK293 cells, plasmids were transformed into DH5α competent bacterial cells to propagate and then purified using standard midiprep system [19,20]. The bacterial cells gave many clones of MGB2 and MGB1 plasmids after 24 hours of transformation and a high amount of DNA (3μg/μL) was obtained after purification (Appendix A).

MGB2 and MGB1 plasmids were transfected to HEK293 cells using polyethylenimine (PEI), which is widely used as cationic transfection reagent [26]. PEI condenses DNA (due to the high concentration of positively charged nitrogen atoms) into cationic particles (polyplexes), which bind to the anionic cell surface and transported into the cell via endocytosis [21]. Once inside the cells protonation of the amine groups of PEI result in lowering the osmotic potential of the cell, releasing polyplexes into the cytoplasm. DNA is then free to diffuse to the nucleus [27]. In our study, the transfection of MGB2 and MGB1 plasmids was successful as indicated by GFP transfection (Figure 4.4). Fluorescence images showed that GFP was successfully delivered to HEK293 cells using PEI.

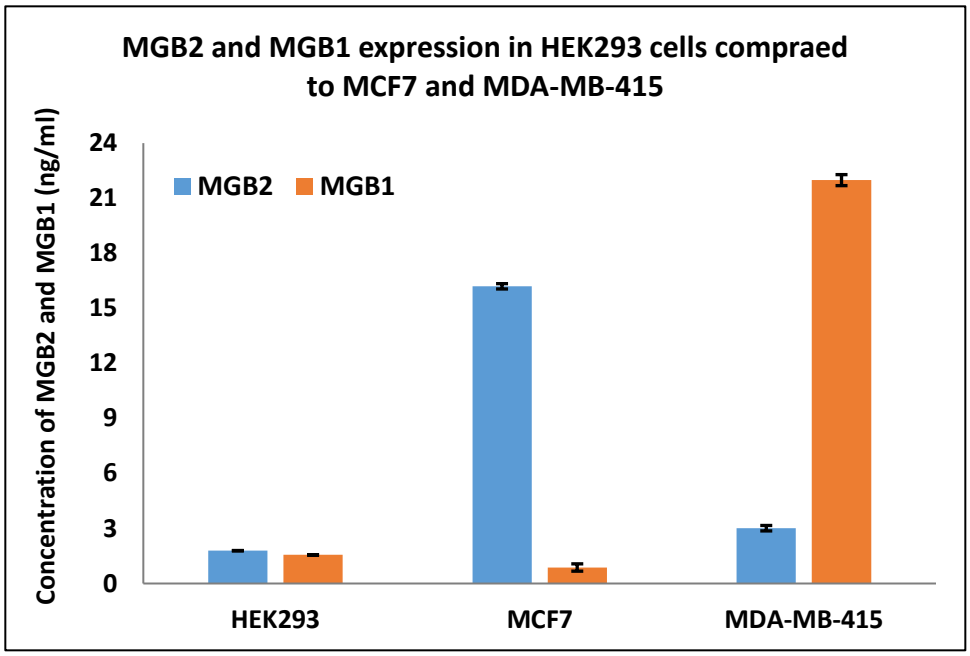


Figure 4.3: MGB2 and MGB1 expression of HEK293, MCF7, and MDA-MB-415 lysates. HEK293 cells showed low level of MGB2 and MGB1 (1.78 and 1.56 ng/mL respectively) compared to MCF7 (16.19 and 0.87ng/ml for MGB2 and MGB1 respectively) and MDA-MB-415 (3.01 and 21.98ng/ml for MGB2 and MGB1 respectively). The expression of MGB2 and MGB1 was measured using an ELISA kits for both proteins.

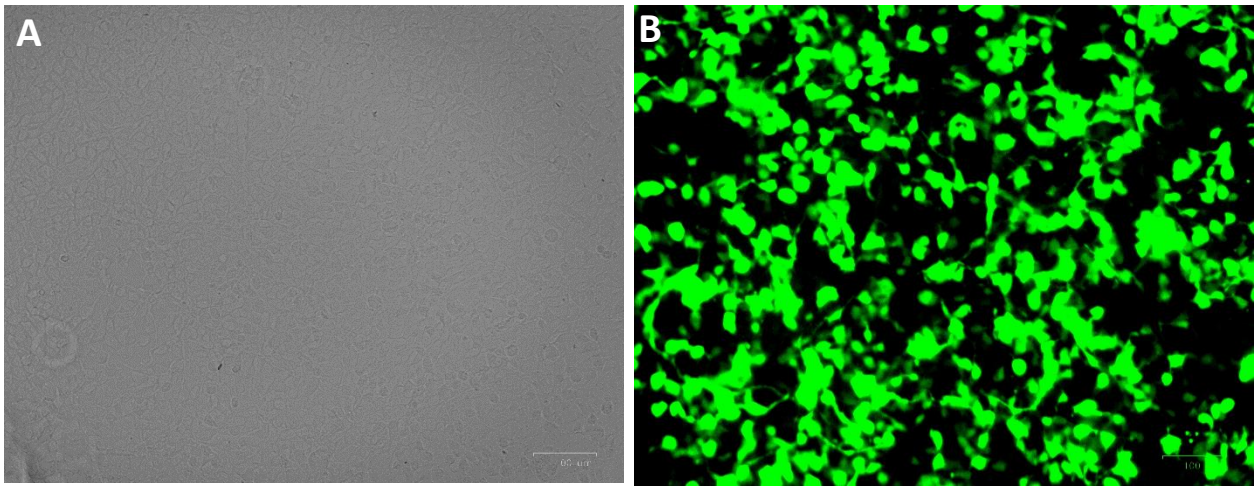


Figure 4.4: Microscopy images of HEK293 cells after transfection with GFP plasmid. A: bright field image of HEK293 cells. B: fluorescence image of HEK 293 cells after the transfection of GFP. Images were taken by ZOE™ Fluorescent Cell Imager. The scale was 100µm for all images.

4.3.4 Aptamer- antibody competition studies by flow cytometry

To confirm that MGB2 and MGB1 proteins are the targets for MAMB1 and MAMA2 aptamers on the surface of the breast cancer cells respectively, competition experiments were carried out using antibodies against MGB2 and MGB1 proteins, as well as 6-FAM-labelled and unlabeled MAMB1 and MAMA2 aptamers on non-transfected and transfected HEK293 cells. Figure 4.5 showed the results of binding the non-transfected HEK293, MCF7 and MDA-MB-415 to 6-FAM-labelled MAMB1 and MAMA2 aptamers as well as to MGB2 and MGB1 antibodies. HEK293 did not bind to MAMB1 and MAMA2 aptamers (Figure 4.5B; black and green respectively) compared to MCF7 and MDA-MB-415 (Figure 4.5B; red and blue respectively), as the mean of fluorescence intensity for HEK293, binding to MAMB1 and MAMA2 aptamers, was close to that for the control cells (Figure 4.5A black). The same results were shown when non-transfected HEK293 cells were incubated with MGB2 and MGB1 monoclonal antibodies (Figure 4.5C black and green respectively), whereas a shift in the mean of fluorescence intensity was seen when MCF7 and MDA-MB-415 cells were incubated with MGB2 and MGB1 antibodies (Figure 4.5C red and blue respectively). A random sequence (same sequence used in the fluorescence microscopy) was used as a control and incubated with non-transfected HEK293, MCF7, and MDA-MB-415 cells. The results showed no binding of the random sequence to any of the cells mentioned above, demonstrating the specificity of MAMB1 and MAMA2 aptamers (Figure 4.5D).

These results were expected as HEK293 cells showed no expression of MGB2 nor MGB1 proteins on their surface, compared to MCF7 and MDA-MB-415 cells, as confirmed by ELISA (Figure 4.3). Moreover, the results above indicated that MAMB1 and MAMA2 aptamers might be binding to MGB2 and MGB1 on the surface of their target cells. Therefore, transfection of both MGB2 and MGB1 plasmids were carried out to confirm this hypothesis. The shift in the fluorescence intensity of MAMB1 and MAMA2 aptamers binding to MCF7 and MDA-MB-415 cells respectively (Figure 4.5B) was higher than that for MGB2 and MGB1 monoclonal antibodies binding to MCF7 and MDA-MB-415 respectively (Figure 4.5C).

This indicated that both aptamers are more specific for their proteins than their corresponding antibodies, but this needed to be confirmed. Therefore, transfection and overexpression of the mammaglobin proteins was performed in HEK293 cells.

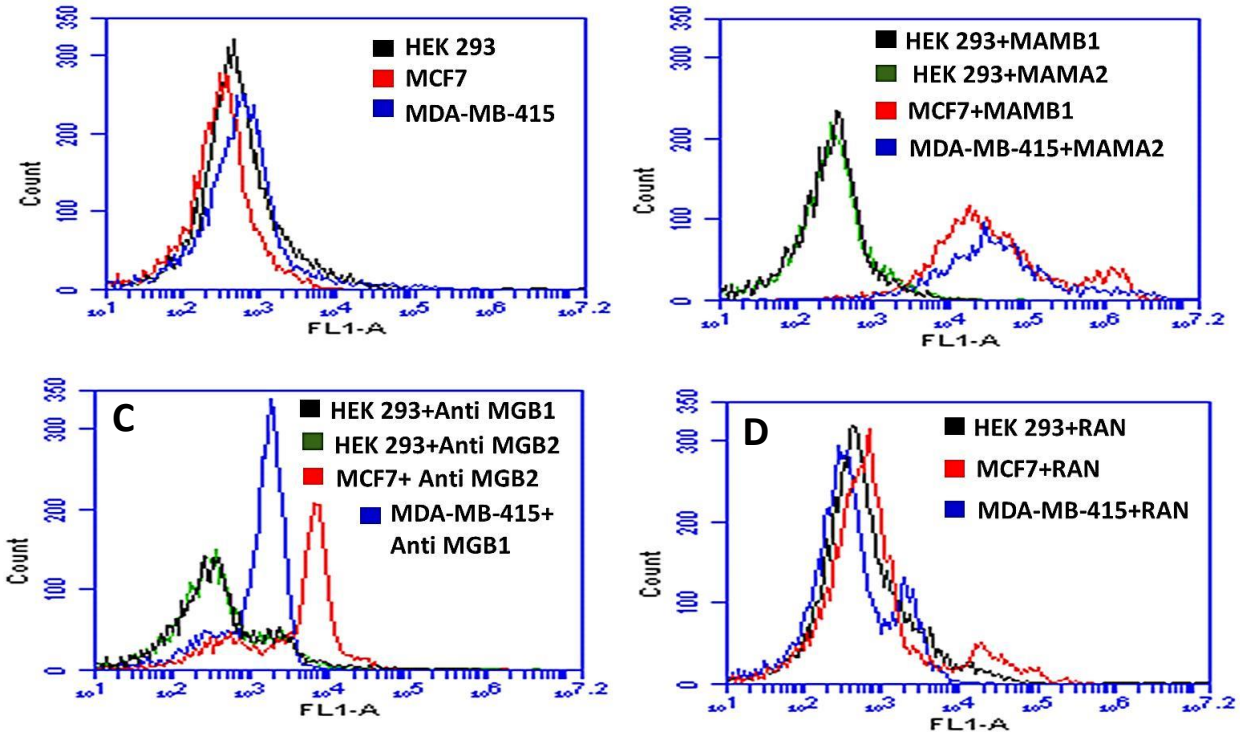


Figure 4.5: Binding of MAMB1 and MAMA2 aptamers and MGB2 and MGB1 antibodies to non-transfected HEK293, MCF7, and MDA-MB-415 cells. A: Mean of fluorescence intensity of non-transfected HEK293 cells (black), MCF7 (red), and MDA-MB-415 (blue) alone. B: Mean of fluorescence intensity of non-transfected HEK293 cells binding to MAMB1 (black), MAMA2 (green), MCF7 binding to MAMB1 (red), and MDA-MB-415 binding to MAMA2 (blue). C: Mean of fluorescence intensity of non-transfected HEK293 cells binding to anti-MGB1 antibody (black), anti MGB2 antibody (green), MCF7 binding to anti MGB2 antibody (red), and MDA-MB-415 binding to MGB1 antibody (blue). D: Mean of fluorescence intensity of non-transfected HEK293 cells binding to a random sequence (black), MCF7 binding to a random sequence, and MDA-MB-415 binding to a random sequence. Cells were incubated with their corresponding probes and 10,000 events was counted by flow cytometry. Then data was analyzed using flow cytometry C6 sampler software.

Competition binding assays results of transfected HEK293 cells with either MGB2 or MGB1 plasmids showed that both aptamers and antibodies could compete against each other for the binding sites of MGB2 and MGB1 overexpressed proteins on the surface of transfected HEK293 cells (Figure 4.6). MAMB1 and MAMA2 aptamers bounds 20-fold higher to HEK293 cells transfected with MGB2 (Trans B) and MGB1 (Trans A) respectively (Figure 4.6A; brown and purple) than to MCF7 and MDA-MB-415 cells (Figure 4.5B; red and blue).

Similar results were observed for the antibodies; anti-MGB2 and anti-MGB1 bound 10-fold higher to Trans B and Trans A HEK293 cells respectively (Figure 4.6A; orange and green) than to MCF7 and MDA-MB-415 cells (Figure 4.5B; red and blue). Both aptamers bound with higher affinity (around 20-fold more) to Trans B and Trans A HEK293 cells than the antibodies (Figure 4.6A), indicating that both aptamers are more specific than their corresponding antibodies to their target proteins.

When Trans B and Trans A HEK293 cells were blocked with unlabeled MAMB1 and MAMA2 aptamers, and then probed with anti-MGB2 and anti-MGB1 respectively, the fluorescence intensity decreased by almost 10-fold (Figure 4.6B; brown and red). Blocking with a random sequence restored the binding of anti-MGB2 to Trans B HEK 293 cells (Figure 4.6C; dark yellow), similar to the binding seen in Figure 4.6B (red), indicating that both MAMB1 and anti-MGB2 were competing for the same binding sites. This binding is specific, as it was restored when a random sequence was used for blocking. A small shift (2-fold less than the binding in Figure 4.6 (brown)) in the mean of fluorescence intensity of anti-MGB1 binding to Trans A HEK 293 cells was seen when the same random sequence was used to block the binding sites of MGB1 on Trans A HEK293 cells (Figure 4.6C; dark yellow and light blue). This could be non-specific binding of the random sequence to MGB1 sites on Trans A HEK293 cells. Other random sequences could be used in the future to restore the original binding seen in Figure 4.6A.

PUC19 empty plasmid was used as a control for the competition assays. PUC19 was transfected same way as MGB2 and MGB1 plasmids simultaneously, and treated with the same probes as Trans B and Trans A HEK293 cells. Figure 4.7A showed no binding of MAMB1, MAMA2, anti MGB2, or anti-MGB1 probes to pUC19-transfected HEK293 cells (Figure 4.7A). Moreover, no binding was seen when pUC19-transfected HEK293 cells were blocked with unlabeled MAMB1 and MAMB1 aptamers, nor when they were blocked with a random sequence (same sequence that was used to block MGB2 and MGB1 sites on transfected HEK293 cells in Figure 4.6C (Figure 4.7B)). PUC19 did not contain either MGB2 or MGB1 genes, and therefore MGB2 and MGB1 were not expressed on the surface of HEK293 cells. The mean of fluorescence intensity of pUC19 transfected HEK293 cells was close to the one for non-transfected HEK293 cells (Figure 4.5).

A selectivity test was performed to investigate the binding of MAMB1 and MAMA2 to Trans A and Trans B HEK293 respectively (Figure 4.8). The results showed that the MAMB1 aptamer did not bind to Trans A HEK293 cells (Figure 4.8; black). The same was observed for the MAMA2 aptamer, which did not bind to Trans B HEK293 cells (Figure 4.8; pink). This was expected as the main pool was divided into two halves (as mentioned in Chapter 2; Figure 2.3A) to eliminate sequences that could bind to both MGB2 and MGB1 proteins together. On the other hand, the binding of anti-MGB2 and anti-MGB1 to Trans A (Figure 4.8; dark yellow) and Tran B HEK293 cells (Figure 4.8; purple) respectively, showed a small shift in the mean of fluorescence intensity. This could be a result of cross-reactivity of both antibodies to their counter proteins. These results indicate that both MAMB1 and MAMA2 aptamers are more selective than their antibodies counterparts.

In summary, our aptamer–antibody competition binding assay results confirmed that MGB2 and MGB1 proteins are the targets of the MAMB1 and MAMA2 aptamers respectively. Moreover, our results were similar to the ones in the literature [28,29]. Many studies have been reported that used similar approaches to identify aptamer targets on the surface of cells, after cell-SELEX [28,29]. More controls were used in our study, when compared to others, such as the use of random sequence and pUC19 empty plasmid transfections.

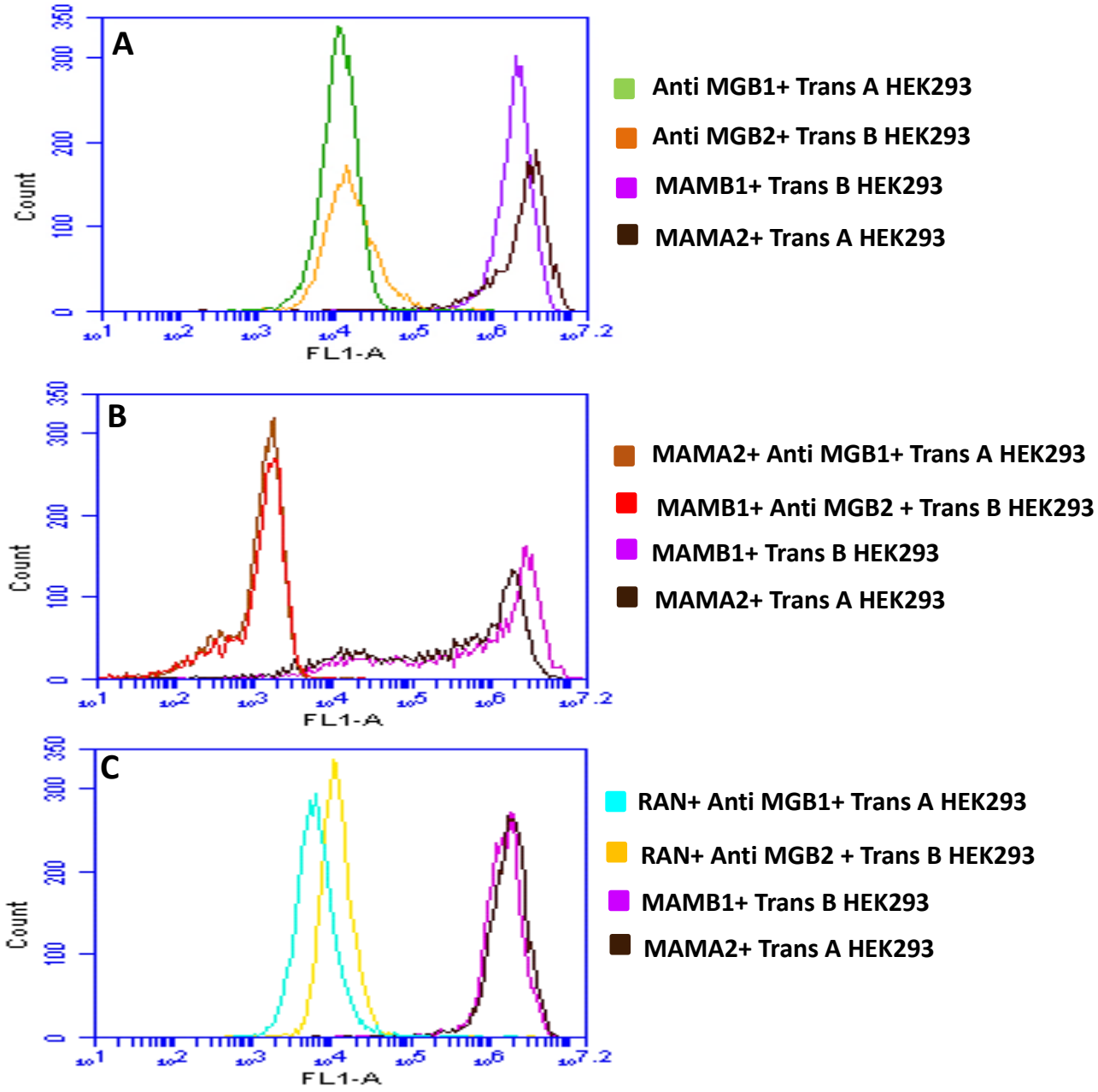


Figure 4.6: Competition binding assays of MAMB1, MAMA2 aptamers and anti-MGB, anti-MGB1 antibodies on transfected HEK293 cells with MGB2 and MGB1 plasmids. A: mean of fluorescence intensity of transfected HEK293 with MGB2 plasmid (Trans B) binding to anti-MGB2 (orange) and MAMB1 (purple), and transfected HEK293 with MGB1 plasmid (Trans A) binding to anti-MGB1 (green) and MAMA2 (brown). B: mean of fluorescence intensity of Trans B and Trans A HEK293 blocked with unlabeled MAMB1 and MAMA2 aptamers respectively, and probed with anti-MGB2 (red) in the case of Trans B, and anti-MGB1 (brown) in the case of Trans A. Mean of fluorescence intensity of Trans B and Trans A binding to MAMB1 (purple), and MAMA2 (brown) same as in A. C: mean of fluorescence intensity of Trans B and Trans A HEK293 blocked with random sequence (RAN) and probed with anti-MGB2 (dark yellow) in the case of Trans B and anti-MGB1 (light blue) in the case of Trans A. Mean of fluorescence intensity of Trans B and Trans A binding to MAMB1 (purple) and MAMA2 (brown) same as in A. Cells were incubated with their corresponding probes and 10,000 events were counted by flow cytometry. Then data was analyzed using flow cytometry C6 sampler software.

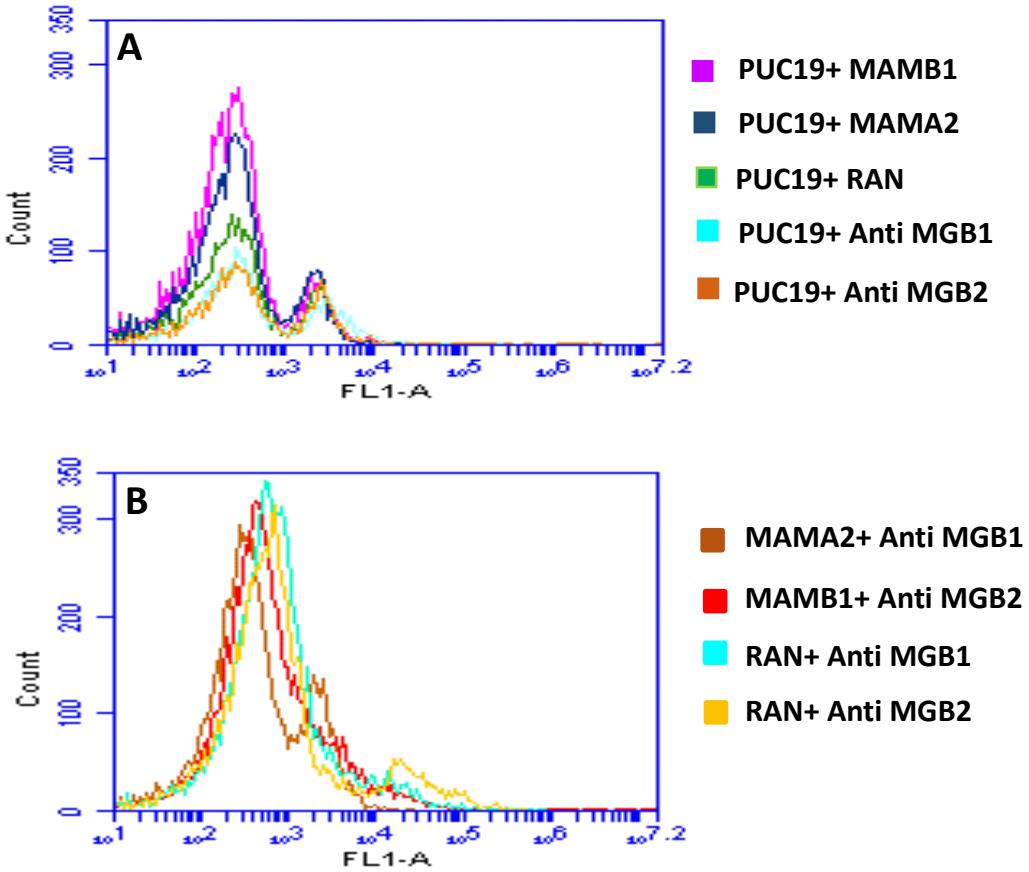


Figure 4.7: Competition binding assays of MAMB1, MAMA2 aptamers and anti-MGB, anti-MGB1 antibodies on transfected HEK293 cells with pUC19 empty plasmid. A: Mean of fluorescence intensity of transfected HEK293 with pUC19 plasmid binding to MAMB1 (Pink), MAMA2 (blue), random sequence (green), anti-MGB2 (orange), and anti-MGB1 (light blue). B: Mean of fluorescence intensity of transfected HEK293 with pUC19 plasmid and blocked with unlabeled MAMB1 and MAMA2 aptamers, then probed with anti-MGB2 (red), and anti-MGB1 (brown) respectively. Yellow Light blue and histograms showed transfected HEK293 with pUC19 blocked with random sequence and then probed with anti-MGB2 and anti-MGB1 respectively. Cells were incubated with their corresponding probes and 10,000 events were counted by flow cytometry. Then data was analyzed using flow cytometry C6 sampler software.

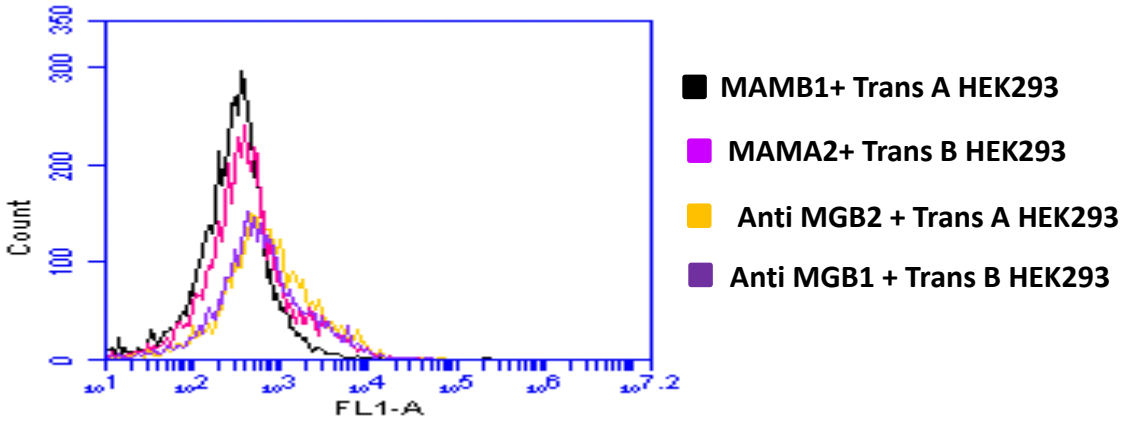


Figure 4.8 Selectivity test performed on transfected HEK293 cells with MGB2 and MGB1 plasmids. The figure shows the mean of fluorescence intensity of MAMB1 (black), MAMA2 (pink), anti MGB2 (dark yellow), and anti-MGB1 (purple) binding to Trans A and Trans B HEK293 cells respectively. Cells were incubated with their counter probes transfected HEK293, and 10,000 events were counted by flow cytometry. Then data was analyzed using flow cytometry C6 sampler software.

4.3.5 Aptamer complex analysis using EMSA

A protein- aptamer study was applied to test the selective binding of both aptamers to their targets in solution using EMSA. Since our target proteins are secretory proteins, testing their binding affinity to their recombinant protein targets free in solution would indicate if these aptamers would bind to their targets in serum. Therefore, EMSA assay was performed.

Both aptamers successfully bound to their target proteins as shown in the increasing of the mobility shift for both MAMB1 and MAMA2 aptamers (Figure 4.9A and C). The shift started at 60 nM for MAMB1 (Figure 4.9A) and at 40 nM for MAMA2 (Figure 4.9E). However, the increasing in the mobility shift was higher for MAMA2 than MAMB1. The random sequence (control aptamer) did not bind to MGB2 protein (Figure 4.9B and F). Yet it showed low level of non-specific binding to MGB2 and MGB1 proteins at the last three concentrations (140, 180, and 220 nM, Figure 4.9B and F). A low level of non-specific binding was seen when BSA was incubated with MAMB1 aptamer at 180 and 220 nM (Figure 4.9C), whereas no binding was shown when BSA was incubated with MAMA2 aptamer (Figure 4.9G). The results indicated that MAMA2 is more specific to its target protein than MAMB1 aptamer and thus indicates the potential use of this aptamer to detect its target free in solution. The different concentrations used in EMSA assay for both proteins were plotted against the shift distance (cm) corresponding to each concertation compared to the control (0 nM) creating a K_d curve for both aptamers (Figure 4.9D and H), the K_d values were 95±7.0 and 88±6.0 nM for MAMB1 and MAMA2 respectively. As expected MAMA2 has lower K_d value which reflect higher affinity to it target protein than MAMB1 aptamer (Figure 4.9D and H).

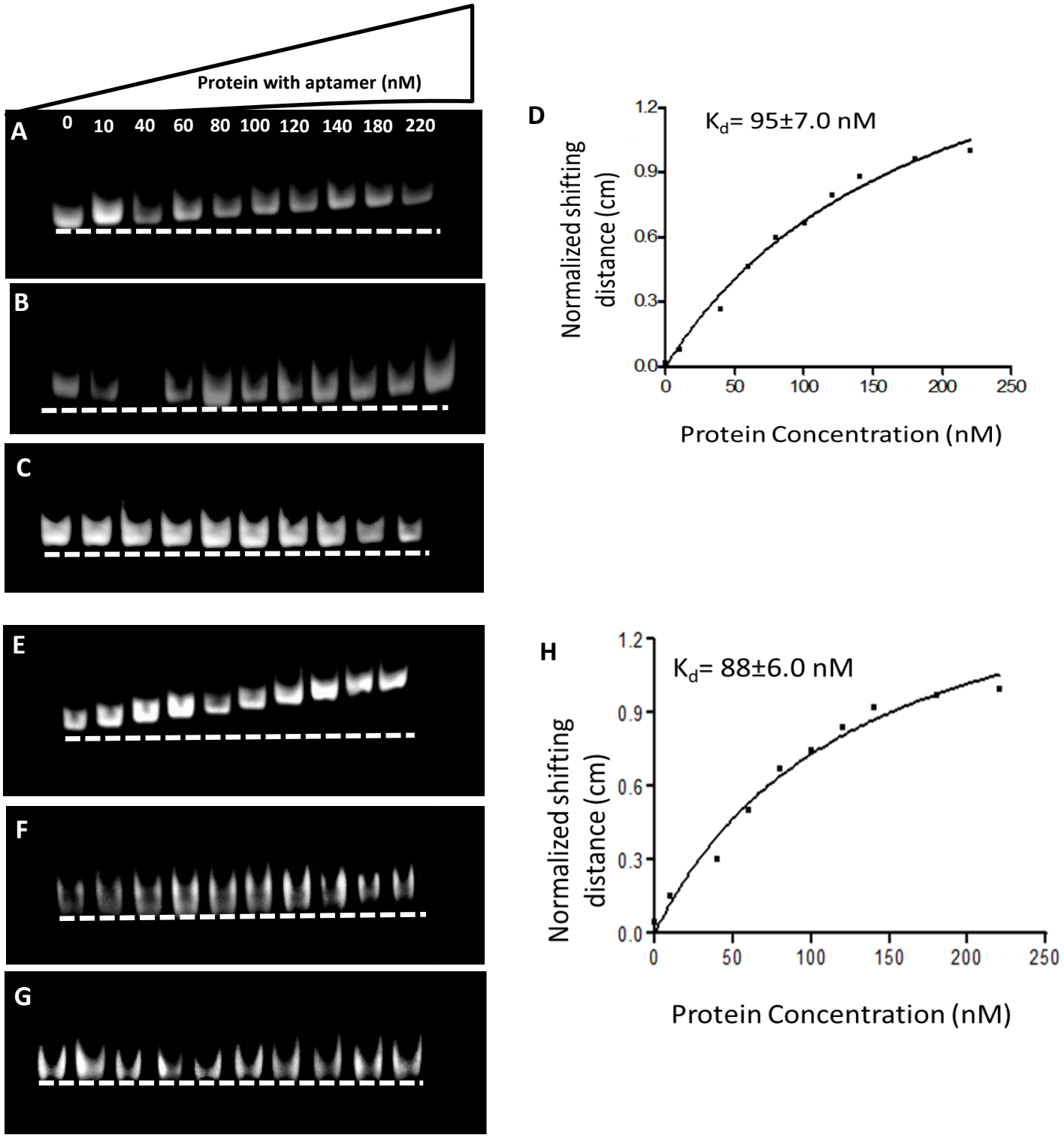


Figure 4.9: EMSA gel after electrophoresis: (A) EMSA for MGB2 protein with MAMB1 aptamer. (B) EMSA for MGB2 protein with DNA control aptamer (random sequence). (C) EMSA for BSA protein with MAMB1 (D) plotting curve of MAMB1 (0-220 nM) against the shifting distance of MAMB1-MGB2 complex compared to the control (0 nM). The K_d value is 95±7.0 as obtained from GraphPad Prism software. (E) EMSA for MGB1 protein with MAMA2 aptamer. (F) EMSA for MGB1 protein with DNA control aptamer. (G) EMSA for BSA protein with MAMA2 (H) plotting curve of (E) MAMA2 (0-220 nM) against the shifting distance of MAMA2-MGB1 complex compared to the control (0nM) The K_d value is 88±6.0 as obtained from GraphPad Prism software. Aptamers (600 nM) and different concentrations of proteins were incubated at 37C for one hour then EMSA was applied using 8% native PAGE. The distance of the complex compared to the control (0 nM) was measured using the analysis tool in the Alpha-imager software. The brightness of gel images was improved by Image J software.

4.4 Conclusions and Future Work

MAMB1 and MAMA2 targets were verified by pull-down assay followed by SDS-PAGE, MS, and Western blot analysis (Figures 4.1 and 4.2). MS results did not show MGB2 and MGB1 proteins (Table 4.2). However, Western blot analysis showed that MAMB1 binds to MGB2 protein and is more specific than the MGB2 monoclonal antibody (Figure 4.2). Moreover, low molecular weight bands (around 10 kDa) were seen on SDS-PAGE after the pulldown of each aptamer from a mixture of cell surface protein extractions of MCF7 and MDA-MB-415 breast cancer cells (Figure 4.1). For more investigation of MAMB1 and MAMA2 targets on the surface of their target breast cancer cells, MGB2 and MGB1 plasmids were transfected to HEK293 cells and competition assay between aptamers and antibodies were performed. The results confirmed that MGB2 and MGB1 proteins are the targets for MAMB1 and MAMA2 respectively (Figures 4.5-4.8). Both aptamers bind to their recombinant proteins and they did not show binding to the random sequence as well as to BSA as a control protein as shown by EMSA (Figure 4.9).

Overall, the results suggest that MAMB1 and MAMA2 aptamers bind to their targets on the surface of the cells as well as the free protein form indicating the possibility of their use in detecting MGB2 and MGB1 proteins in solution (and perhaps in serum).

To identify the targets that both aptamer bind to by mass-spectrometry, two-dimensional (2D) gel electrophoresis would help in separating proteins of similar molecular weight, but different charges which may mask MGB2 and MGB1 proteins on one-dimensional (1D) gel electrophoresis. This would include the histone proteins, which are small and bind non-specifically to any DNA molecule, such as aptamers. Future studies will focus on 2D separation of MGB targets from histone proteins.

References

1. Hou Z, Meyer S, Propson NE, Nie J, Jiang P, Stewart R, Thomson JA. Characterization and target identification of a DNA aptamer that labels pluripotent stem cells. Cell Res. 2015, 25(3):390-393.
2. Wu CC, Yates JR 3rd. The application of mass spectrometry to membrane proteomics. Nat Biotechnol. 2003, 21(3):262-267.
3. Daniels DA, Chen H, Hicke BJ, Swiderek KM, Gold L. A tenascin-C aptamer identified by tumor cell SELEX: systematic evolution of ligands by exponential enrichment. Proc Natl Acad Sci U S A. 2003, 100(26):15416-15421.
4. Ohuchi SP, Ohtsu T, Nakamura Y. Selection of RNA aptamers against recombinant transforming growth factor-beta type III receptor displayed on cell surface. Biochimie. 2006, 88(7):897-904.
5. Kan H, Yong K, So L, Dong L. Isolation of RNA aptamers targeting HER-2-overexpressing breast cancer cells using cell-SELEX. Bull. Korean Chem. Soc. 2009, 30(8): 1827-1831.
6. Song Y, Zhu Z, An Y, Zhang W, Zhang H, Liu D, *et al*. Selection of DNA aptamers against epithelial cell adhesion molecule for cancer cell imaging and circulating tumor cell capture. Anal Chem. 2013, 85(8):4141-4149.
7. Shangguan D, Cao Z, Meng L, Mallikaratchy P, Sefah K, Wang H, *et al*. ell-specific aptamer probes for membrane protein elucidation in cancer cells. J Proteome Res. 2008, 7(5):2133-2139.
8. Mallikaratchy P, Tang Z, Kwame S, Meng L, Shangguan D, Tan W. Aptamer directly evolved from live cells recognizes membrane bound immunoglobin heavy mu chain in Burkitt's lymphoma cells. Mol Cell Proteomics. 2007, 6(12):2230-2238.
9. Van Simaeys D, Turek D, Champanhac C, Vaizer J, Sefah K, Zhen J, *et al*. Identification of cell membrane protein stress-induced phosphoprotein 1 as a potential ovarian cancer biomarker using aptamers selected by cell systematic evolution of ligands by exponential enrichment. Anal Chem. 2014, 86(9):4521-4527.

10. Yang M, Jiang G, Li W, Qiu K, Zhang M, Carter CM, *et al*. Developing aptamer probes for acute myelogenous leukemia detection and surface protein biomarker discovery. J Hematol Oncol. 2014, 7:5.
11. Becker RM, Darrow C, Zimonjic DB, Popescu NC, Watson MA, Fleming TP. Identification of mammaglobin B, a novel member of the uteroglobin gene family. Genomics. 1998, 54(1):70-78.
12. Watson MA, Fleming TP. Mammaglobin, a mammary-specific member of the uteroglobin gene family, is overexpressed in human breast cancer. Cancer Res. 1996, 56(4):860-865.
13. Watson MA, Dintzis S, Darrow CM, Voss LE, DiPersio J, Jensen R, Fleming TP. Mammaglobin expression in primary, metastatic, and occult breast cancer. Cancer Res. 1999, 59(13):3028-3031
14. Fanger GR, Houghton RL, Retter MW, Hendrickson RC, Babcook J, Dillon DC, et al. Detection of mammaglobin in the sera of patients with breast cancer. Tumour Biol. 2002, 23(4):212-421.
15. Zehentner BK, Persing DH, Deme A, Toure P, Hawes SE, Brooks L, *et al*. Mammaglobin as a novel breast cancer biomarker: multigene reverse transcription-PCR assay and sandwich ELISA. Clin Chem. 2004, 50(11):2069-2076.
16. Bernstein JL, Godbold JH, Raptis G, Watson MA, Levinson B, Aaronson SA, Fleming TP. Identification of mammaglobin as a novel serum marker for breast cancer. Clin Cancer Res. 2005, 11(18):6528-6235.
17. Galvis-Jiménez JM, Curtidor H, Patarroyo MA, Monterrey P, Ramírez-Clavijo SR. Mammaglobin peptide as a novel biomarker for breast cancer detection. Cancer Biol Ther. 2013, 14(4):327-332.
18. Mercatali L, Valenti V, Calistri D, Calpona S, Rosti G, Folli S, *et al*. RT-PCR determination of maspin and mammaglobin B in peripheral blood of healthy donors and breast cancer patients. Ann Oncol. 2006, 17(3):424-428.
19. Invitrogen protocol:

    http://www.openwetware.org/images/d/df/Subcloning_Efficiency_DH5a_Chemically_competent_E._Coli.pdf

20. https://www.promega.ca/-/media/files/resources/protocols/technicalmanuals/0/pureyield-plasmid-midiprep-system-protocol.pdf?la=en

21. Longo PA, Kavran JM, Kim MS, Leahy DJ. Transient mammalian cell transfection with polyethylenimine (PEI). Methods Enzymol. 2013, 529:227-240.

22. Hsu CY, Uludağ H. A simple and rapid nonviral approach to efficiently transfect primary tissue-derived cells using polyethylenimine. Nat Protoc. 2012, 7(5):935-945.

23. Reed SE, Staley EM, Mayginnes JP, Pintel DJ, Tullis GE. Transfection of mammalian cells using linear polyethylenimine is a simple and effective means of producing recombinant adeno-associated virus vectors. J Virol Methods. 2006, 138(1-2):85-98.

24. Dai Z and Wu C. How does DNA complex with polyethylenimine with different chain lengths and topologies in their aqueous solution mixtures? Macromolecules. 2012, (10): 4346–4353.

25. Graham FL, Smiley J, Russell WC, Nairn R. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. J Gen Virol. 1977, 36(1):59-74.

26. Boussif O, Lezoualc'h F, Zanta MA, Mergny MD, Scherman D, Demeneix B, *et al*. A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. Proc Natl Acad Sci U S A. 1995, 92(16):7297-7301.

27. Akinc A, Thomas M, Klibanov AM, Langer R. Exploring polyethylenimine-mediated DNA transfection and the proton sponge hypothesis. J Gene Med. 2005, 7(5):657-663.

28. Yang M, Jiang G, Li W, Qiu K, Zhang M, Carter CM, *et al*. Developing aptamer probes for acute myelogenous leukemia detection and surface protein biomarker discovery. J Hematol Oncol. 2014, 7(5): 1-14.

29. Sefah K, Tang ZW, Shangguan DH, Chen H, Lopez-Colon D, Li Y, *et al*. Molecular recognition of acute myeloid leukemia using aptamers. Leukemia. 2009, 23(2):235-244.

CHAPTER 5

Testing the binding of MAMB1 and MAMA2 aptamers to their target breast cancer cells using terahertz chemical microscopy

5.0 Statement of contribution:

Part of the TCM experiments were done at Dr. Toshihiko Kiwa's lab at Okayama University (Japan). Mr. Yuki Hanaoka (a student in Dr. Kiwa's lab) performed part the TCM scan experiments. This work is a collaboration between the INRS, Carleton University (Department of Chemistry and Department of Biology) and Okayama University (Japan).

The second part of the TCM experiments was done at the INRS, by me and Ahmed Mohammed (a Ph.D. student at Prof. Ozaki's lab).

5.1 Introduction

5.1.1 Terahertz chemical microscopy

The terahertz chemical microscope (TCM) was first developed to image chemical reactions [1]. Rather than using THz absorption as contrast, the TCM is a type of a THz emission microscope, which uses THz emission to monitor changes in the chemical potential [2]. A key component of the TCM consists of a sensing plate made of thin silicon dioxide (SiO_2), and silicon (Si) films on a sapphire substrate (Figure 5.1) [2]. When a femtosecond laser pulse irradiates the sensing plate (specifically the Si film), it generates carriers. These carriers are then accelerated by the local electric field gradient within the depletion layer between the Si and SiO_2 films, leading to the generation of THz radiation.

When a chemical reaction occurs on the surface of the sensing plate, the electric field (chemical potential) generated on the surface of the sensing plate will shift (change) the electric field gradient in the depletion layer, which will result in the variation of the generated THz pulse's amplitude [1]. An advantage of the TCM is that it can overcome the limited spatial resolution of THz imaging systems [1]. TCM has higher spatial resolution than standard THz imaging systems, because the spatial resolution is not determined by the wavelength of the THz pulse, but by that of the femtosecond laser, which is typically around 790 nm [1]. Moreover, TCM could monitor the concentration of chemicals in media that are highly absorptive to THz radiation, such as water. TCM allows one to extend THz technology to biology, which with conventional techniques had been difficult.

The TCM was first used to detect potassium (K^+) and sodium (Na^+) ions in aqueous solution. A difference in electrical potential is seen between membrane surfaces when the sensing membrane (sensing membranes immobilized with sodium (Na^+) and potassium (K^+) ions) comes into contact with the aqueous solution containing the specific ions.

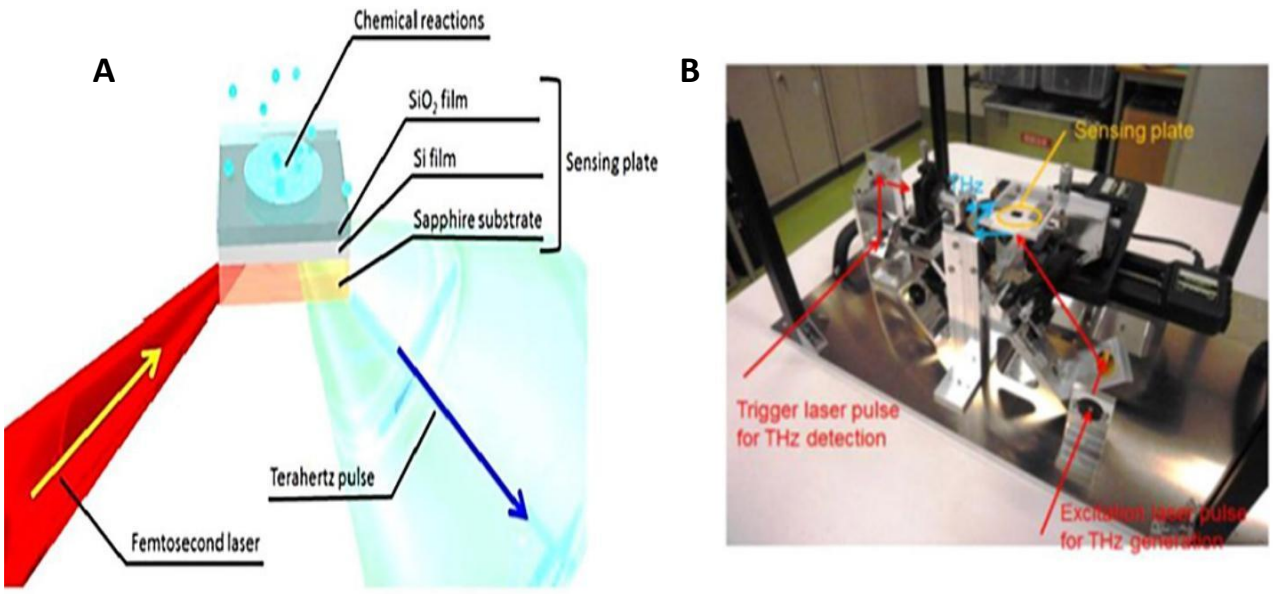


Figure 5.1: A: Schematic diagram of the terahertz (THz) chemical microscope (TCM) sensing plate. A femtosecond laser is focused on the thin silicon (Si) film from the substrate side of the plate, and a THz pulse radiates into the free space. B: Photograph of the prototype TCM system. Reprinted with the permission of Kiwa *et al*. [1]

Another application of the TCM is to study the interaction between an analyte and its receptor, by providing a label free assay [3]. Kiwa *et al.* has shown that the TCM could measure the change in the chemical potential on the surface of the sensing plate as a result of the binding between an antigen and its antibody [3]. Kiwa *et al*. studied the change of the THz amplitude when different concentrations of anti-IgG antibodies interacted with the antigen (IgG) (Figure 5.2) [3]. This study revealed that the TCM could be used to measure the binding between a ligand and its target. Moreover, the TCM can be used for the detection of a protein complex [4]. In a study by Kiwa *et al*., biotin-avidin complexes were detected as a change in the amplitude of the THz peak when biotin interacted with avidin on the surface of the sensing plate (Figure 5.3) [4].

In this study, aptamers were used in the TCM for the first time as a probe to study the interaction results after binding with their targets (breast cancer cells). The details of how aptamers could be used in the TCM to detect breast cancer cells by interacting with their targets is described in this chapter.

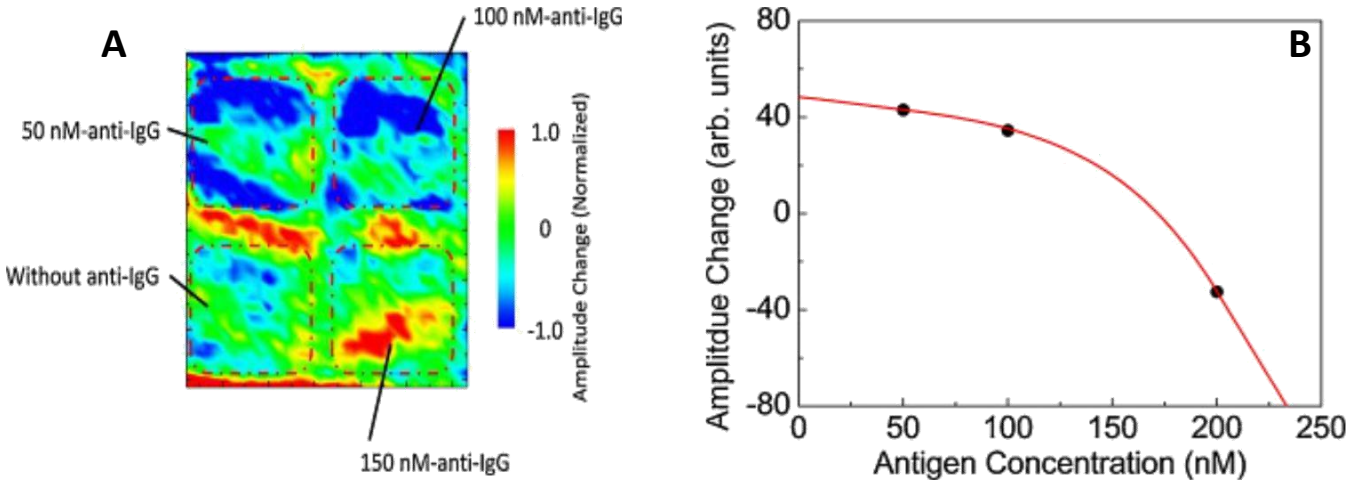


Figure 5.2: Using TCM to study antigen-antibody interactions. A: A differential image showing the changes of THz amplitude upon the interaction of different concentrations of anti-IgG antibody to IgG antigen. In this study 50, 100, and 150nM of anti-IgG antibody was added to the surface of the sensing plate which was immobilized with IgG antigen then, divided into 4 parts to study the different concentrations. One of the parts was a control (without IgG antibody) to show the THz amplitude before the interaction of the antibody to its antigen. B: the changes of the THz amplitude upon the interaction of the different concentrations of the antibody to the antigen plotted in a graph. Reprinted with permission of Kiwa *et al*. [3].

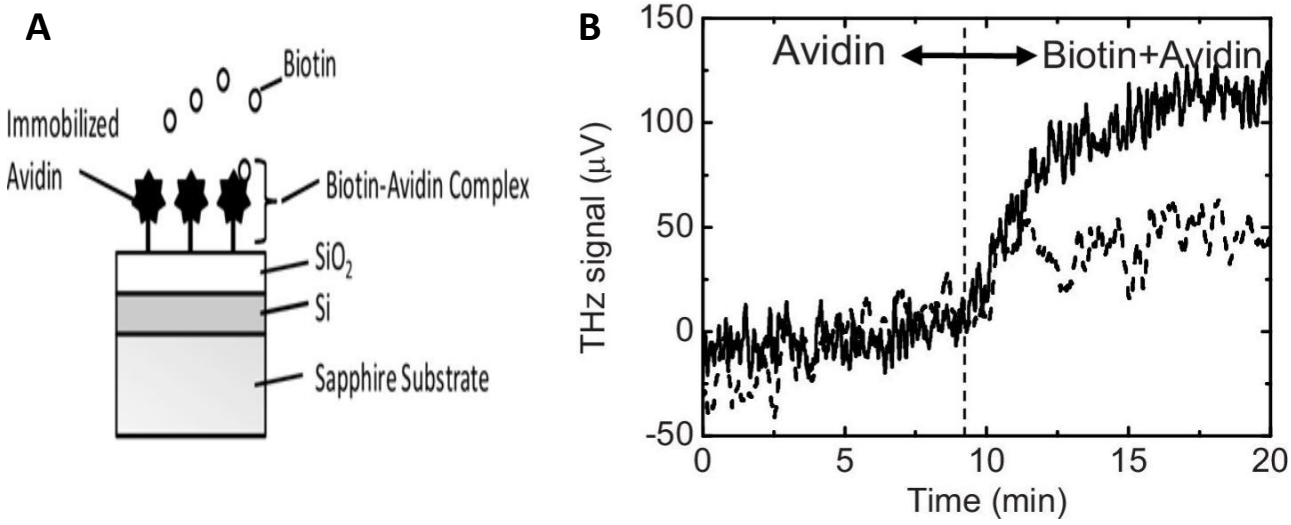


Figure 5.3: Using TCM to detect protein complex. A: the image shows the TCM sensing plate with avidin protein immobilized on its surface, and biotin is dropped (dissolved in solution) on the surface to interact with avidin. B: the graph represents the change of the THz amplitude peak before and after the interaction of biotin and avidin. Reprinted with permission of Kiwa *et al*. [4].

5.2 Materials and methods

The same materials and methods were used for both the TCM studies done in Japan as well as those done at the INRS, unless otherwise indicated.

5.2.1 Materials and DNA synthesis

MAMB1 and MAMA2 aptamers were synthesized to obtain dual conjugate aptamers (NH_2 on the 3' end and 6-FAM (Fluorescein) on the 5' end) using a DNA synthesizer. MCF10A was used as the control cells ("normal" breast cells). MCF7 and MDA-MB-415 were used as target breast cancer cells for MAMB1 and MAMA2 aptamers, respectively. Sensing plates (Wafer chips) contain a silicon dioxide (SiO_2) layer (2 nm thick) deposited on a thin film of silicon were obtained from MTI Corporation (USA).

5.2.2 Preparation of the TCM wafer chips

5.2.2.1 Cleaning of wafer chips and surface activation

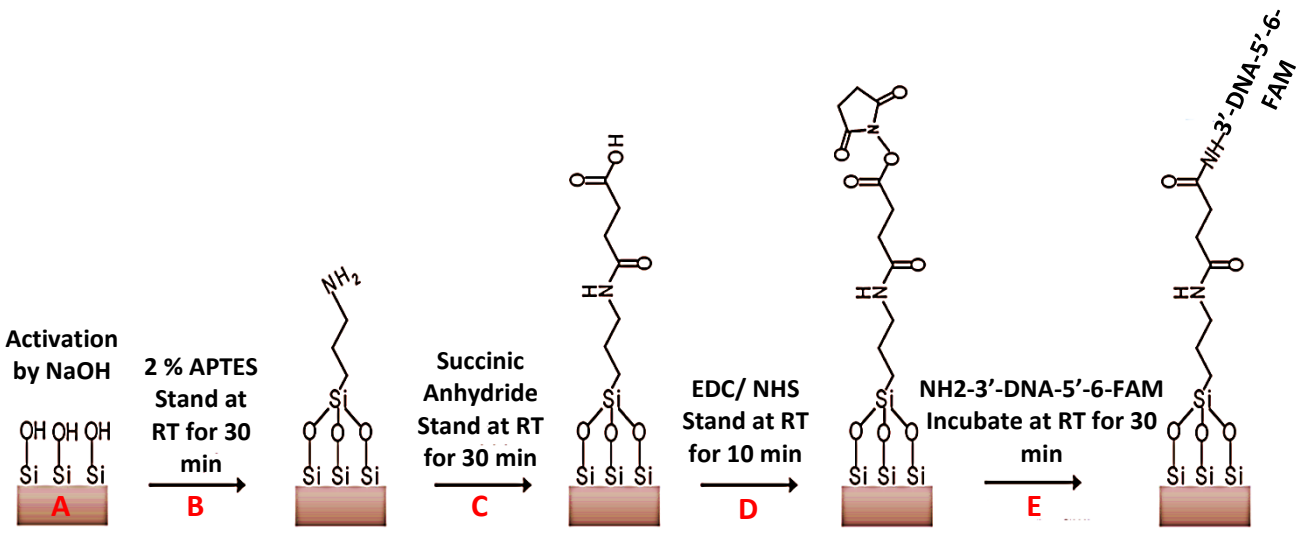
Wafer pieces were incubated in warm acetone for 10 min, then transferred to absolute ethanol and incubated for another 10min. The wafers were then washed with absolute ethanol and dried using argon. For surface hydroxylation, the wafers were then incubated in 1M NaOH for 45 min at RT [5,6].

5.2.2.3 Silanization and immobilization of aptamers on the wafer surface

Wafers were salanized by using (3-Aminopropyl) triethoxysilane (APTES) (Scheme 5.1). 2 % saline solution (APTES) in 95:5 (vol/vol) absolute ethanol/water mixture was prepared and let to stand at RT for 30 min. Wafers were then incubated in this mixture for 2 hrs at RT. Wafers were then washed with absolute ethanol 3 times with an incubation of 5 min each time, dried by argon, and then were let to stand in an oven at 120^{o}C for 20 min. The amide bond was formed by adding Succinic Anhydride (SA). Briefly, 140 mg of SA was dissolved in 2 mL of DMSO, at the same time that the wafers were placed in a Petri dish, and immersed with borate buffer (200 mM boric acid pH8.00). SA solution was injected into the Petri dish with quick mixing, and let to incubate for 30 min at RT.

The wafer was then sonicated in Deionized Water (DI) for 3 min and allowed to dry using argon. The cross linker EDC/NHS was then used by dissolving 100mM of EDC and 50mM of NHS in DI (1 mL each) and adding them with quick mixing to a 10mM PBS solution, where the wafers were immersed. The reaction was let to stand at RT for 10 min. The wafers were then washed with DI and dried with argon [7].

MAMB1 and MAMA2 aptamers were prepared in PBS to a final concentration of 10µM, heated at 95^{o}C for 5min, cooled gradually to 4^{o}C for 10 min, and let to stand at RT for another 10min before being incubated with the wafers (by immersion).. The incubation time of each aptamer with the wafer was 30 min. The non-binding aptamers were washed away with 10 mM Tris and 20mM $MgCl_2$ three times. After that, the wafers were dried by argon and were ready for TCM scan [7].



Scheme 5.1: The process of silanization and immobilization of aptamers (MGB2 and MGB1) on the surface of the wafer chip. A: hydroxylation of the surface took place by incubating the wafer with NaOH, then APTES amino- silane agent was added to create an amino terminated surface (B). Amide bond formation took place by adding SA, followed by EDC/NHS cross linker for activation of the carboxylic group to mediate the formation of the covalent bond between the amine on the 3' end of the DNA and the carboxylic group on the surface of the wafer (C,D,E). Adapted from [8].

5.2.2.3 Validation of the immobilization of both aptamers on the surface of the sensing plates

MAMB1 and MAMA2 wafer chips (sensing plates) were subjected to fluorescence microscope imaging to confirm the immobilization step. Each chip (immobilized with MAMB1 or MAMA2) was loaded on a ZOE™ fluorescent cell Imager (Bio-Rad, Canada) which has a bright field channel, and three fluorescence channels. The green-channel (417nm excitation, 517 emission) was used to detect 6-FAM (fluorescein) on the surface of the chips. Samples were viewed on the screen, and images were taken (within the first 5 min) by an integrated digital camera at 100µm scale. A control chip (without aptamer on the surface) was used as a reference to compare the intensity of the green color on the same scale (100µm).

5.2.3 TCM scan for the chips

5.2.3.1 Cell preparation

MCF10A, MCF7, and MDA-MB-415 ($1X10^6$/mL) were fixed using 4% formaldehyde for 15 min at RT, and then cells were washed and re-suspended in 10mM Tris buffer pH 7.4.

5.2.3.2 TCM protocol

Wafer chips (two) immobilized with aptamers (MAMB1 and MAMA2) on their surfaces were scanned by TCM following the protocol on how to prepare the samples for TCM.

5.2.3.2.1 Binding of MAMB1 and MAMA2 aptamers to their target breast cancer cells

For both MAMB1 and MAMA2 chips, each chip was placed on the holder of the TCM (as indicated by Figure 5.1 B). PBS buffer (blank) was pipetted (100 µL) on the surface of the chip (same for both chips). Then a scan was run using the TCM system to measure the amplttiude of the THz. After that the buffer was washed away gently using acetone. The control cells (MCF10A, same for both chips) $1x10^5$ were then pipetted (100 µL) same was as the buffer and the amplttiude of the THz was recoreded.

Acetone washing was perfomed after this step for each chip. For MAMB1 chip, MCF7 cells (100 µL) were added to the surface of MAMB1 sensing plate. Same number of cells as the control cells was scanned to measure the amplttiude of the THz. For MAMA2 chip MDA-MB-415 (100 µL) was pipetted on the surface of the sensing plate. Then different number of MDA-MB-415 was tested using the TCM system same as in MAMB1 chip. After that all the TCM scan data was saved.

5.2.3.2.2 Studying different number of target cells on MAMB1 and MAMA2 sensing plates using TCM

To further test the minimum number of target cells each aptamer-chip could bind, different dilutions of cells were prepared and then applied to TCM for scanning as shown in table 5.1.

5.2.3.2.3 Selectivity test of MAMB1 and MAMA2 aptamers

Both aptamers were further tested for their selectivity using TCM by applying the same protocol above, the cell lines. In this study, MAMB1 aptamer was tested against the MDA-MB-415 cells, and the MAMA2 aptamer was tested against the MCF7 cells.

Table 5.1: preparation of stocks of different number of cells

Stocks (cells)	Buffer (µl)	MCF7 (µl) original stock	MDA-MB-415 (µl) original stock	MCF10A (µl) original stock	Total volume (µl)
$1X10^4$	198	2	0	0	200
	198	0	2	0	200
	297	0	0	3	300
$1X10^3$	180	20 from 10^4 stock	0	0	200
	180	0	20 from 10^4 stock	0	200
	270	0	0	30 from 10^4 stock	300
$1X10^2$	180	20 from 10^3 stock	0	0	200
	180	0	20 from 10^3 stock	0	200
	270	0	0	30 from 10^3 stock	300
10	180	20 from 10^2 stock	0	0	200
	180	0	20 from 10^2 stock	0	200
	270	0	0	30 from 10^2 stock	300

5.2.3.3 TCM sensing plates scanning (imaging) details (Okayama University-Japan)

For all TCM scanned chips, the central wavelength of the femtosecond laser was 780 nm. The average laser power that was focused onto the sensing plate was 120 mW, and the pulse repetition rate was 80 MHz. MAMB1 and MAMA2 chips (sensing plates) were loaded on the TCM holder, which is integrated into the TCM platform (as shown in Fig. 5.1B). Femtosecond laser was warmed warm up and stabilize (if the laser was completely off) before the scanning started (to stabilize the laser power). Each chip was then scanned for about 90 min. TCM measurement was done by scanning the chips in two dimensions (2D).

THz peak amplitude maps for the initial scan and for selectivity scans were developed as colored contour images (using the scanned data points) for control scans (buffer only (Blank), and MCF10A cells) as well as experimental scans (MCF7 and MDA-MB-415 breast cancer cells) using Origin software (Origin pro-USA). Differential images were then developed by subtracting the experimental images (breast cancer cells) from the control images (buffer and MCF10A) to identify the shift in the THz amplitude (delta THz amplitude, (Δ THz mV)).

For experiments with varying cell numbers, the numerous scanned data points (in the differential images) were averaged to give a single point, which represents the value of delta THz amplitude (Δ THz mV). The number of cells scanned (10^5, 10^4, 10^3, 10^2, 10) from each cell line (MCF7 and MDA-MB-415) on its corresponding chip (MAMB1 for MCF7 and MAMA2 for MDA-MB-415) were plotted against the difference in the THz amplitude (Δ THz [mV]) to determine the lowest number of cells that each aptamer could bind to using TCM.

The initial scan for both MAMB1 and MAMA2 chips were performed once using 10^5 cells from both MCF7 and MDA-MB-415 cells.

This scan was repeated (using different chips) when the varying cell number experiments were done in duplicate (the first time 10^4, 10^3, 10^2, 10 cells were used, then the second time 10^5, 10^4, 10^3, 10^2, 10 cells were used). The selectivity test was performed twice on two different chips.

5.2.3.4 TCM sensing plates scanning (imaging) details (INRS-Canada)

The scanning method of the sensing plates used at the INRS was the same as that used in Japan, with a few differences. The central wavelength of the femtosecond laser was between 790 and 800 nm. The average laser power that was focused onto the sensing plate was 130 mW, and the pulse repetition rate was 80 MHz. An optical (mechanical) chopper was used to improve the signal-to-noise ratio. Before starting the scan, the laser was warmed up for up to 30 min. The scanning process for each sensing plate started by first scanning the entire plate in 2D, searching for a region with relatively homogenous THz signal. For both plates used, an area of 2.5 × 2 mm was found to have a homogenous signal. The spatial resolution was 0.3 mm. The scanning time for each sensing plate was 10 min.

Scanning was performed three times for each number of cells (10-10^6 of each MCF7, MCF10A, and MDA-MB-415) as well as for the selectivity test (MCF7 cells (10^6) on MAMA2 chip and MDA-MB-415 cells (10^6) on MAMB1 chip). The data was then analyzed in the same way as the first set of experiments in Japan, using the same software (Origin 2016 graphing and analysis). Three trials of each sensing plate (MAMB1 and MAMA2) were performed on the same sensing plate.

5.3 Results and discussion

5.3.1 MAMB1 and MAMA2 immobilization on the surface of the wafer chips (Okayama and INRS)

MAMB1 and MAMA2 aptamers were successfully immobilized on the surface of the wafer chips (for Okayama and INRS experiments), and were compared to the control chip, as shown in Figure 5.4. The wafer chip alone without the aptamer showed some fluorescence (as seen when examined under the fluorescence microscope) (Figure 5.4A), which may be from sapphire or from the combination of the sapphire with silicon and silicon oxide.

This fluorescence is negligible (as we see in Figure 5.4A) compared to the fluorescence of the aptamers. The fluorescence shown by MAMB1 (Figure 5.4B) and MAMA2 (Figure 5.4C) binding aptamers is much more (as the image indicated (Figure 5.4B)), indicating that the DNA aptamers were immobilized on the surface of the chip.

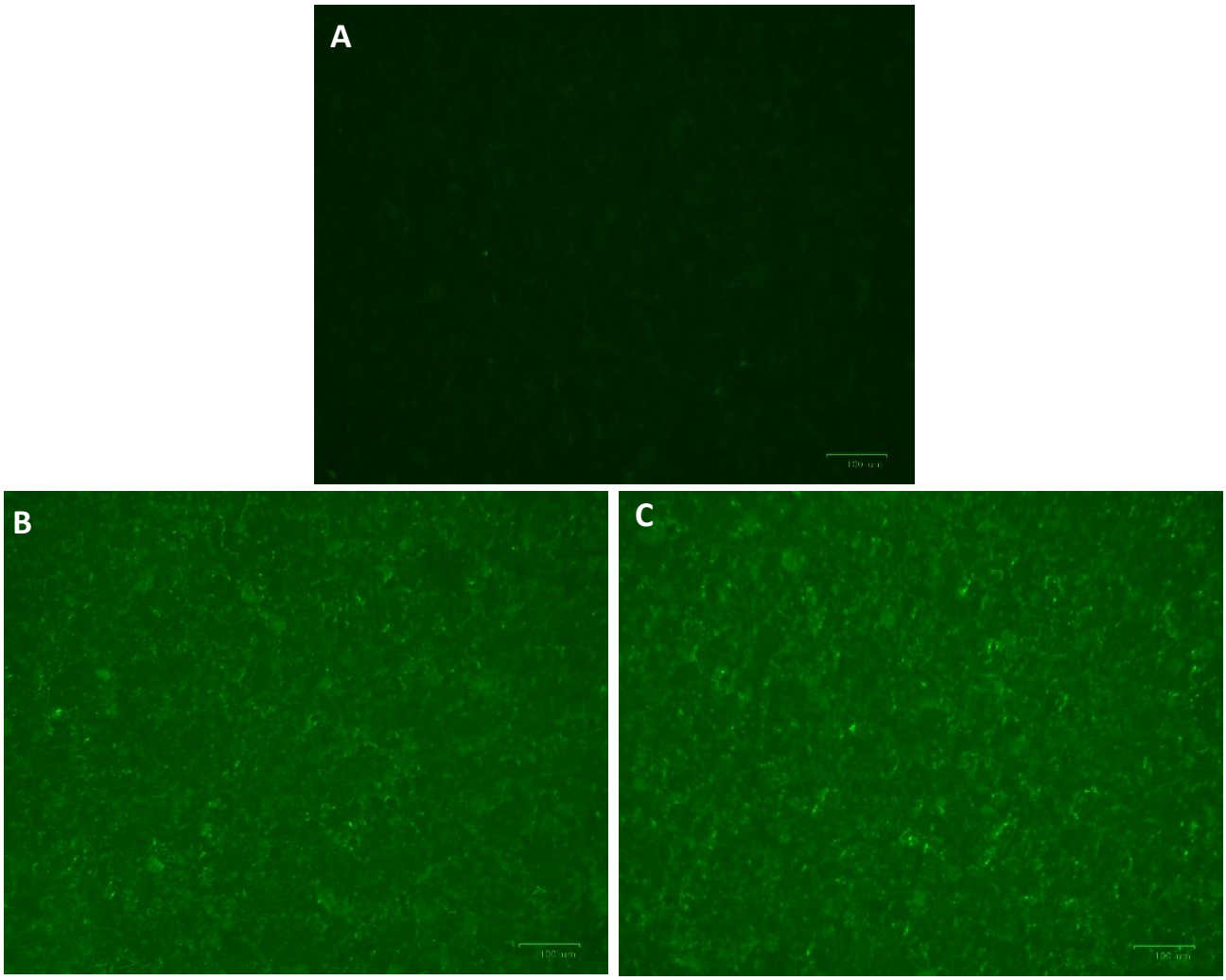


Figure 5.4: Fluorescence images of the wafer chips after aptamer immobilization taken by ZOE™ Fluorescent Cell Imager. A: wafer chip alone without aptamer. B: MGB2 chip after immobilization of the aptamer (MAMB1) showing a proof that the 6-FAM labeled aptamer was covalently bound to the wafer chip. C: MGB1 chip after immobilization of the aptamer (MAMA2) showing a proof that the 6-FAM labeled aptamer was covalently bound to the wafer chip. Images were taken on the green channel (417nm excitation, 517 emission). The scale was 100µm for all images.

5.3.2 Okayama University results

5.3.2.1 Binding of MAMB1 and MAMA2 aptamers to their target breast cancer cells

Aptamers are similar to antibodies in the way they bind to their targets, and it has been reported that the TCM was used to detect the binding of an antibody to its antigen (target) (as discussed in the introduction of this chapter) [3]. Therefore, we tested MAMB1 and MAMA2 aptamers using TCM to detect breast cancer cell lines. MCF7 and MDA-MB-415 (target cells) as well as MCF10A (control cells, (1×10^5 cells from each)) were added (pipetting (100µL)) to their corresponding sensing plates and scanned by TCM (as mentioned in the TCM protocol of the method section). The results showed more interaction (binding) between the aptamers and their target breast cancer cells (MCF7 and MDA-MB415) compared to the interaction between aptamers and control cells (MCF10A), as well as the buffer, as indicated by the intensity of the color in contour images (Figure 5.5 and Figure 5.6). The interaction of the aptamers with their targets on the surface of each sensing plate led to a change in the chemical potential on the surface and therefore resulted in a THz amplitude shift (change). This shift was higher when using MCF7 cells on MAMB1 sensing plate and MDA-MB-415 cells on MAMA2 plate, when compared to the control cells (MCF10A) (Figure 5.5C and B and 5.6C and B).

Different THz peak intensities can be seen in Figures 5.5C and 5.6C, especially at the centers of the images. The intensity in Figure 5.5C is higher (orange) than that of Figure 5.6C (green to orange), as indicated by the color scale on the right of both images, which represents the THz amplitude (mV). This could explain the interaction of both aptamers to their targets. MAMAB1 aptamer seemed to have a higher affinity (more interaction) to MCF7 cells compared to MAMA2 aptamer and its target (MDA-MB- 415) and thus resulted in more change in the THz amplitude on the sensing plate than MAMA2 aptamer. The change observed with both aptamers is indeed higher than the shifting resulting from the buffer (Blank, (dark green)) or control cells ((dark green) MCF10A) when interacting with either the MAMB1 chip (Figure 5.5A and B) or MAMA2 chip (Figure 5.6A and B). The interactions of the buffer and the control with the aptamers were considered as the baseline.

In order to evaluate the THz amplitude change due to the actual aptamer-target interaction (binding), the THz amplitude for the target cells should be subtracted from that of the control cells (MCF10A), and generating differential images (Figure 5.5 F and Figure 5.6 F). Again, by looking at Figures 5.5F and 5.6F and comparing the signal intensities, Figure 5.5F shows a higher intensity (dark orange) than Figure 5.6F (light orange). This emphasized that the binding of MAMB1 aptamer developed more shifting on the surface of the MAMAB1 chip, compared to the MAMA2 aptamer on the surface of MAMA2 chip.

The interaction of the aptamers with their targets depends on the conditions in which the aptamers are present [9]. For example, it has been reported that the presence of positively charged ions such as Mg^{2+} could dramatically affect the interaction of aptamers to their targets [10-12]. In TCM experiments done at Okayama, the targets (breast cancer cells) were present in 10mM Tris buffer, which might not have been a favorable environment for the aptamers, since the ionic strength was not enough for the aptamer-target interaction. As such, the aptamers may not have formed the necessary three dimensional structures for the optimal target interaction. This could explain the behavior of MAMA2 aptamers compared to MAMAB1 aptamers, in which MAMA2 might need extra ionic strength to fold properly on the target and generate a higher shift in THz amplitude (Figure 5.5F and 5.6F). On the other hand, TCM is sensitive to salts, since it can be used to study chemical reactions, such as the interaction of potassium and sodium ions with water (as mentioned in the introduction section of this chapter) [1], and the use of extra salts could develop false negative results (the buffer will cause high shift of the THz amplitude on the surface of the sensing plate and thus, masking the real interaction of aptamer and the target). It is worth mentioning here that these results match the results in chapter 3, when both aptamers were used to detect breast cancer cells in complex biological media using flow cytometry (Figures 3.4 and 3.5). MAMA2 showed a lower mean of fluorescence intensity (binding) than MAMB1.

Figure 5.5D and E and Figure 5.6D and E represent the differential images for both the control cells (MACF10A) and the cancer cells (MCF7 and MDA-MB-415 respectively) subtracted from the buffer.

The shift of the THz amplitude in control cells was small (baseline) as indicated by the color scale (dark green) (Figure 5.5D and 4.6D) compared to the shift as a result of the aptamers interaction with their targets (dark orange) (Figure 5.5E and 5.6E).

Overall results of the initial testing of MAMB1 and MAMA2 using TCM indicate high binding to their breast cancer cell targets compared to control cells. These results matched the results obtained by Kiwa *et al.* when they used an antibody (Ig-G) to detect its antigen (Ig-G) [3], as they also noticed an increase in the THz amplitude compared to the control [3]. Aptamers could therefore be used in this manner, and more experiments were done to determine the lowest number of cells that could be detected using TCM. These results are explained in the following section.

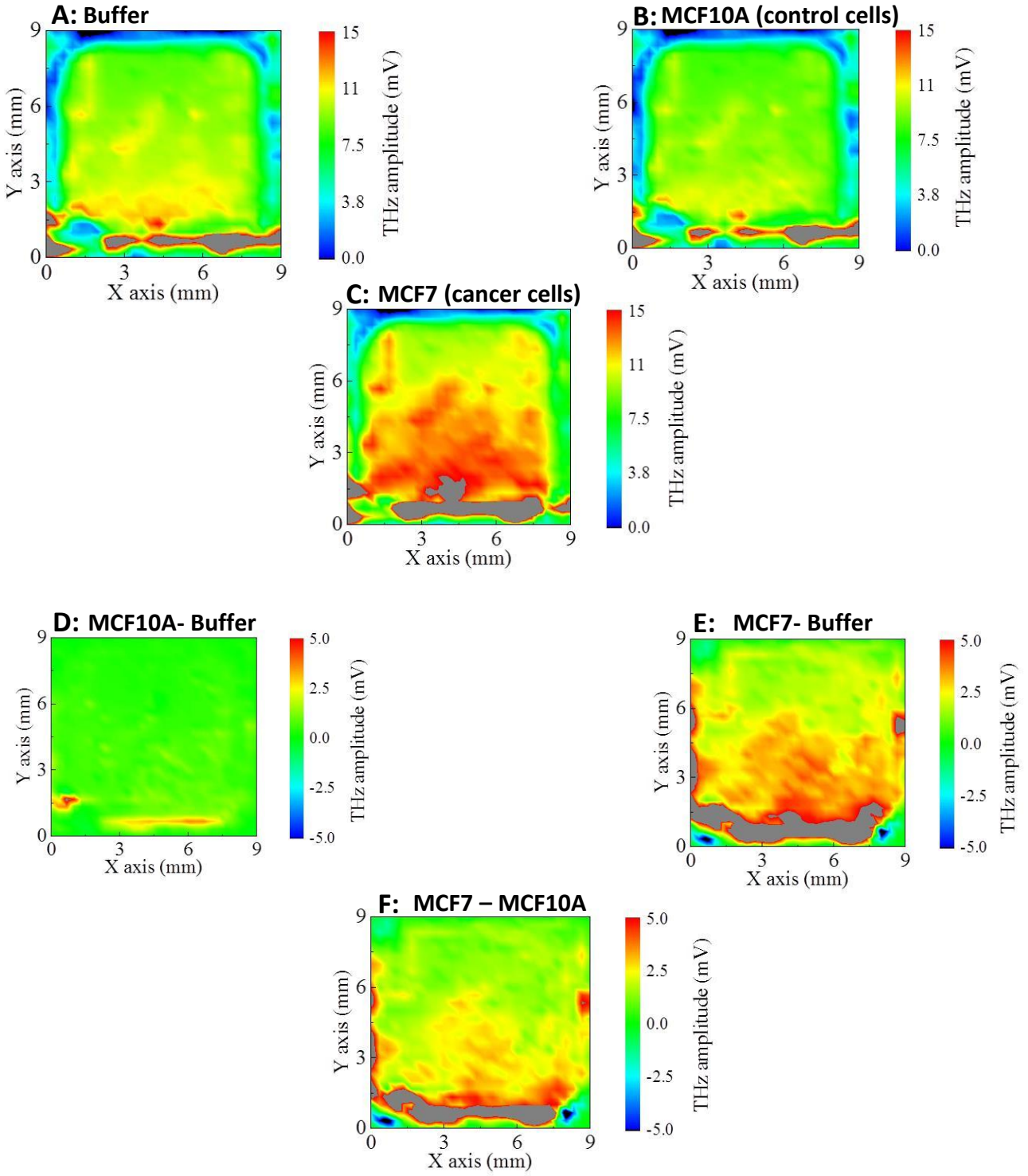


Figure 5.5: THz peak amplitude mapping for MAMAB1 chip. A: buffer alone B: MCF10A (normal control cells), C: MCF7 cells (breast cancer cells), D,E, and F: Differential images after subtraction from buffer and control cells. Images were created using Origin software. MCF7 breast cancer cells (1×10^5) were placed on the surface of the sensing plate by pipetting a 100µl. then the sensing plate was scanned in 2D for 90 min. many data points were measured during the scanning. These data points were converted to contour images using Origin software.

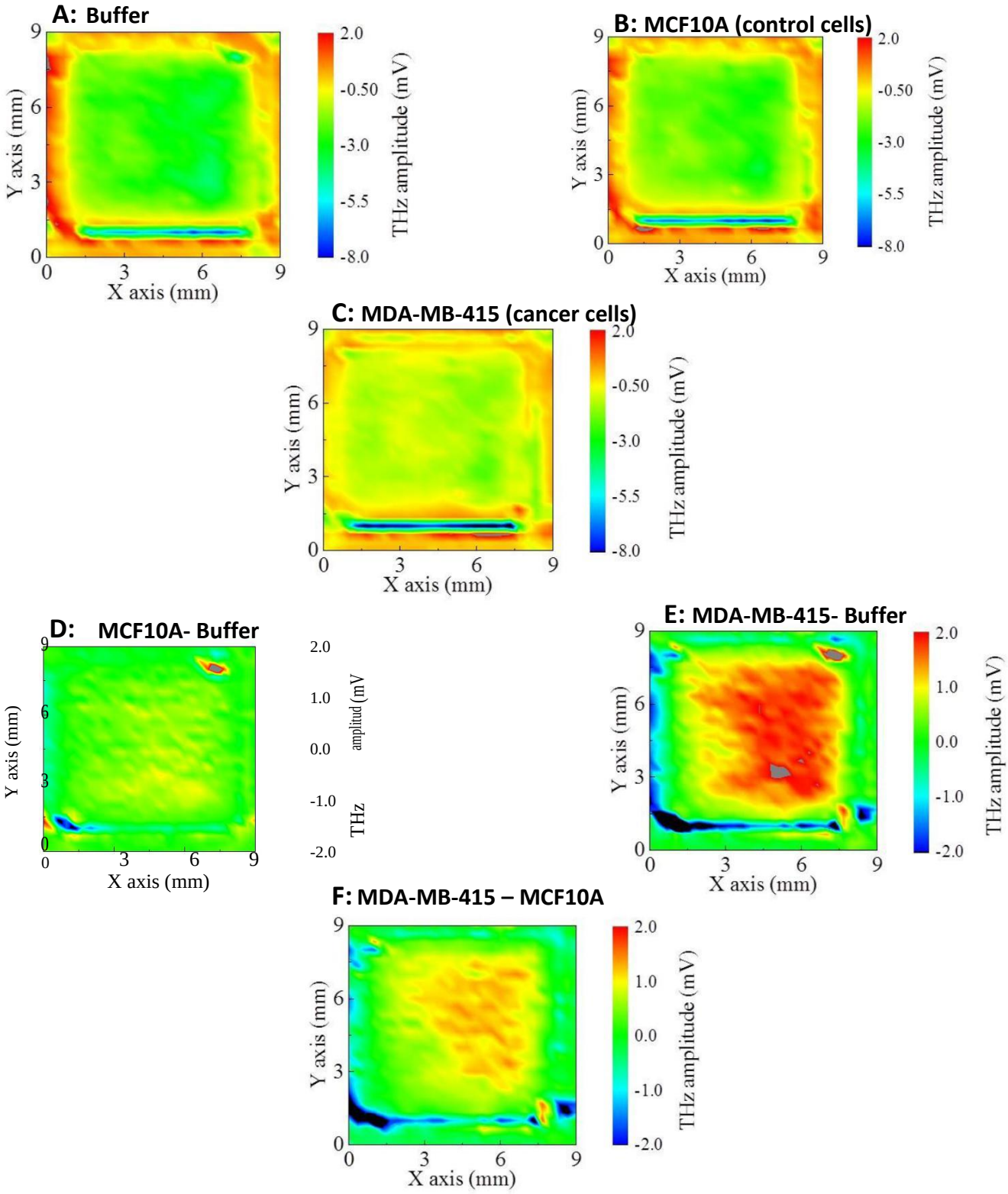


Figure 5.6: THz peak amplitude mapping for MAMA2 chip. A: buffer alone B: MCF10A (normal control cells), C: MDA-MB415 cells (breast cancer cells), D,E, and F: Differential images after subtraction from buffer and control cells. Images were created using Origin software. MDA-MB-415 breast cancer cells (1×10^5) were placed on the surface of the sensing plate by pipetting a 100µl. The sensing plate was then scanned in 2D for 90 min. Many data points were measured during the scanning. These data points were converted to contour images using Origin software.

5.3.2.2 Studying different numbers of target cells on MAMB1 and MAMA2 sensing plates using TCM

Different numbers of breast cancer cells (MCF7 and MDA-MB-415), as well as control cells (MCF10A) (10^5, 10^4, 10^3, 10^2, 10) were used in this experiment. Cells were pipetted onto the sensing plate (100µL) in the same way as described in the initial test experiment (control cells first and then cancer cells). The chip was then scanned in 2D for each set of cells for around 90 min, using a femtosecond laser as the laser source for the TCM. The experiment was done twice on two different TCM chips. The 10^5 set of cells was only used in the second trial.

As mentioned earlier (in the TCM protocol method section), many data points were scanned on the surface of the TCM chips (MAMB1 and MAMA2 chips). This data was averaged (the area represented by the dashed black lines on the differential images (Figure 5.7A- 5.7D and 5.8A-5.8E, and figure 5.9A-5.9E and Figure 5.10A-5.10F)) to give a single point on the graph (Figure 5.7E and 5.8F and figure 5.9E and 5.10F). Due to the large number of measurements, the error in each measurement was small.

The first trial using the MAMB1 chip showed increasing delta THz amplitude with increasing number of scanned breast cancer cells (MCF7), from 0.08mV at 10 cells to 0.90mV at 10^4 cells, as shown by the differential images of breast cancer cells (Figure 5.7). In contrast, the second trial results showed decreasing of the THz amplitude from 0.65mV at 10 cells to - 0.94 mV at 10^5 cells, as shown by the differential images of MCF7 (Figure 5.8). The results obtained from MAMA2 chip scanned by TCM showed that THz amplitude was decreasing with increasing number of cells in the two trials, from 0.50 mV at 10 cells to -3.30mV at 10^4 cells in the first trial (Figure 5.9) and from -0.59mV at 10 cells to -0.99mV in the second trial (Figure 5.10).

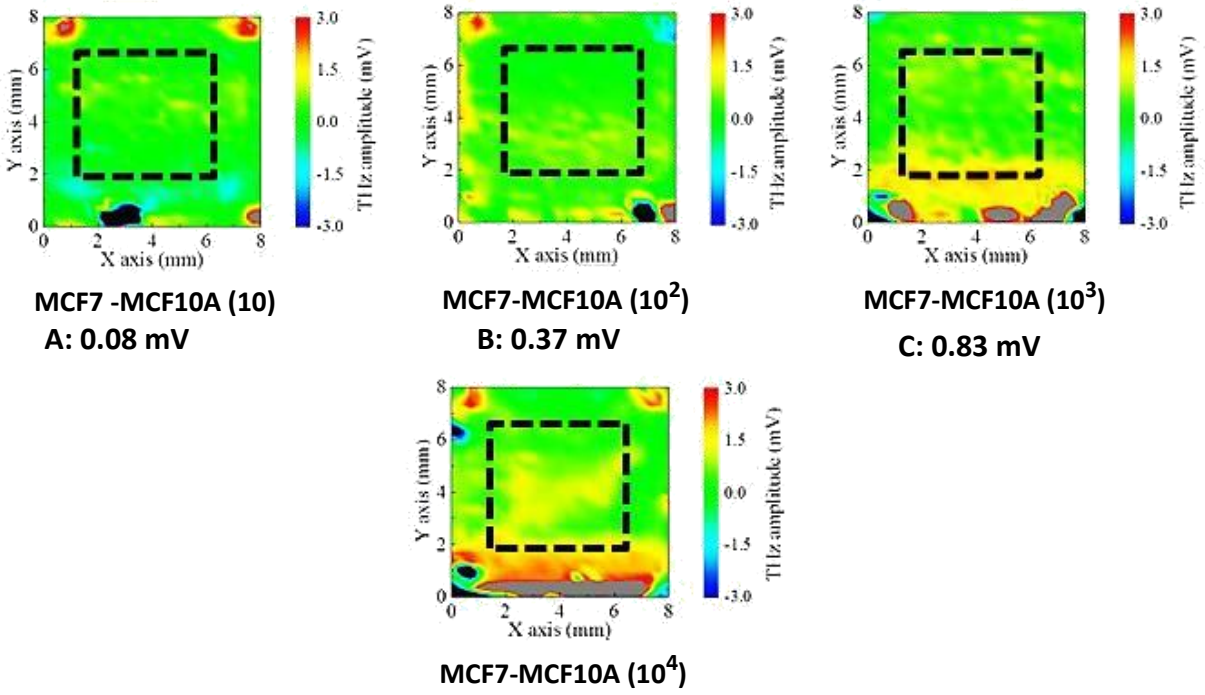
In Figure 5.7, the THz amplitude changed at positions where the MAMB1 aptamer interacted with MCF7. The error bars (on figure 5.7E) that represent the standard deviation of each single image at each concentration of the cells (calculated as [13]), Figure 5.7 A-D) indicated high variability in the shift of the THz amplitude, as a result of either the inhomogeneity in the immobilization of MAMB1 or the THz signal due to the nature of the sensing plate.

This non-uniformity was also reported in the study by Kiwa *et al*. for antibody-antigen interaction [3]. On the other hand, the change in the THz amplitude for the second trial of MAMB1 was more uniform than the first one, as indicated by the standard deviation values (Figure 5.8F). The interaction of MAMA2 with MDA-MB-415 cells in the first and second trials showed more uniformity in the THz amplitude change, as shown in figures 5.9E and 5.10F. These results suggest that the TCM can evaluate aptamer-breast cancer cell complex formation. However, more trials were needed to confirm this point, and thus more sensing plates (MAMB1) were scanned at the INRS (Canada).

The first trial with the MAMA2 chip showed an increase in the THz amplitude at 10 and 10^2 cells, and then a drop to – 0.35 mV at 10^3 cells (Figure 5.9C). The electric field turned the opposite direction at the same number of cells in the second trial of the MAMB1 chip (Figure 5.8C). The relationship between the THz amplitude change and the number of cells is not quite linear (Figure 5.9E). Results of the second trial showed that the THz amplitude was very similar for 10^3 and 10^4 cells (Figure 5.10C and D). No correlation was found in THz amplitude change and the number of cells (Figure 5.10F). In this case, more trials were also needed to reach a conclusion, and thus more sensing plates (MAMA2) were scanned at the INRS (Canada).

To summarize, the results obtained by studying different number of cells on MAMB1 and MAMA2 chips revealed that the TCM could be used to detect the interaction of aptamers with breast cancer cells. There were however, many factors that could affect the changes in the THz amplitude (that is, the binding of aptamers to their target of MAMB1 and MAMA2 chips) on the surface of the sensing plate, such as: the initial charges (initial chemical potential) on the surface of the sensing plate (chip), the quality of the TCM sensing plate, as good chips give rise to more uniform and less variable results and, the nature of the target, in our case the targets were normal cells (in case of the control) and cancer cells (in case of the targets). Cells are large in size, holding many charges on their surfaces that could change the values of the THz amplitude. Finally, the environment surrounding the aptamers could affect the results. For example, the buffer is one of the conditions that could affect the interaction of the aptamers with their targets, as mentioned earlier in this section, as well as the temperature and the pH [12].

All of these factors should be investigated. It is worth mentioning here that the above experiments were done using different sensing plates (not the same sensing plate), and this could play a role in the output as well. Dr. Kiwa (Okayama lab) propsed that in order to study the reproducibility of MAMB1 and MAMA2 aptamers, the trials should be done on the same sensing plate up to four times as soon as the bond between the aptamers and their targets could be broken by mild solvents. Therefore, three more trials were performed at the INRS (Canada) for each MAMB1 and MAMA2 sensing plate, and the results are discussed in the following sections.



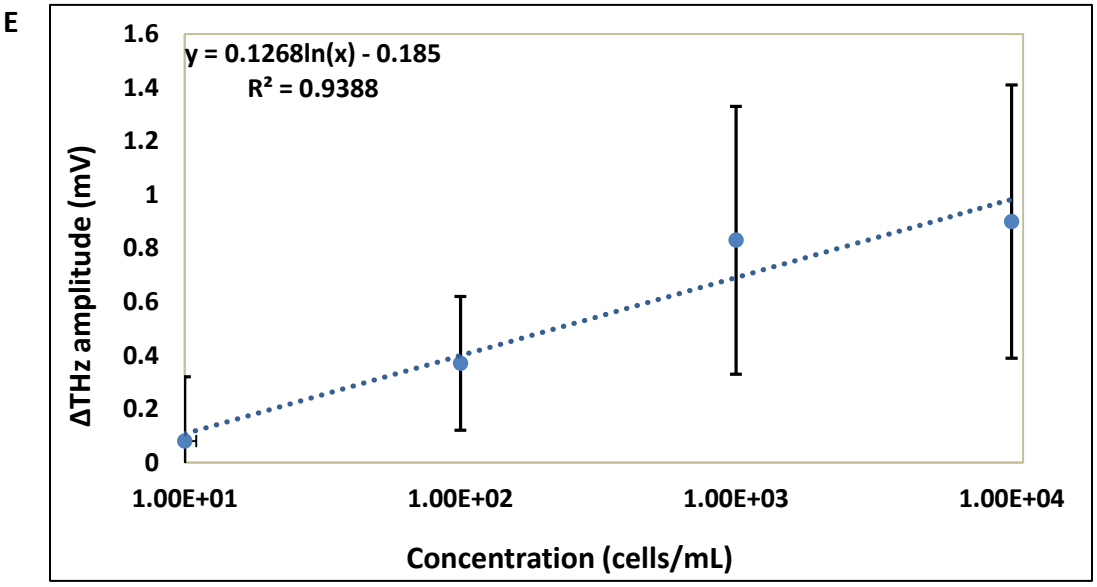


Figure 5.7: THz peak amplitude mapping of different numbers of MCF7 cancer cells on the surface of MAMB1 sensing plate (first trial). A-D: differential images (MCF7-MCF10A) of MAMB1 sensing plate when different numbers of MCF7 cancer cells were used (10, 10^2, 10^3 and 10^4 respectively). E: THz amplitude (mV) curve against different number of cells. To calculate the THz amplitude (mV) an average area of 6.7 mm was created on both the X axis and the Y axis. Images were created using Origin software.

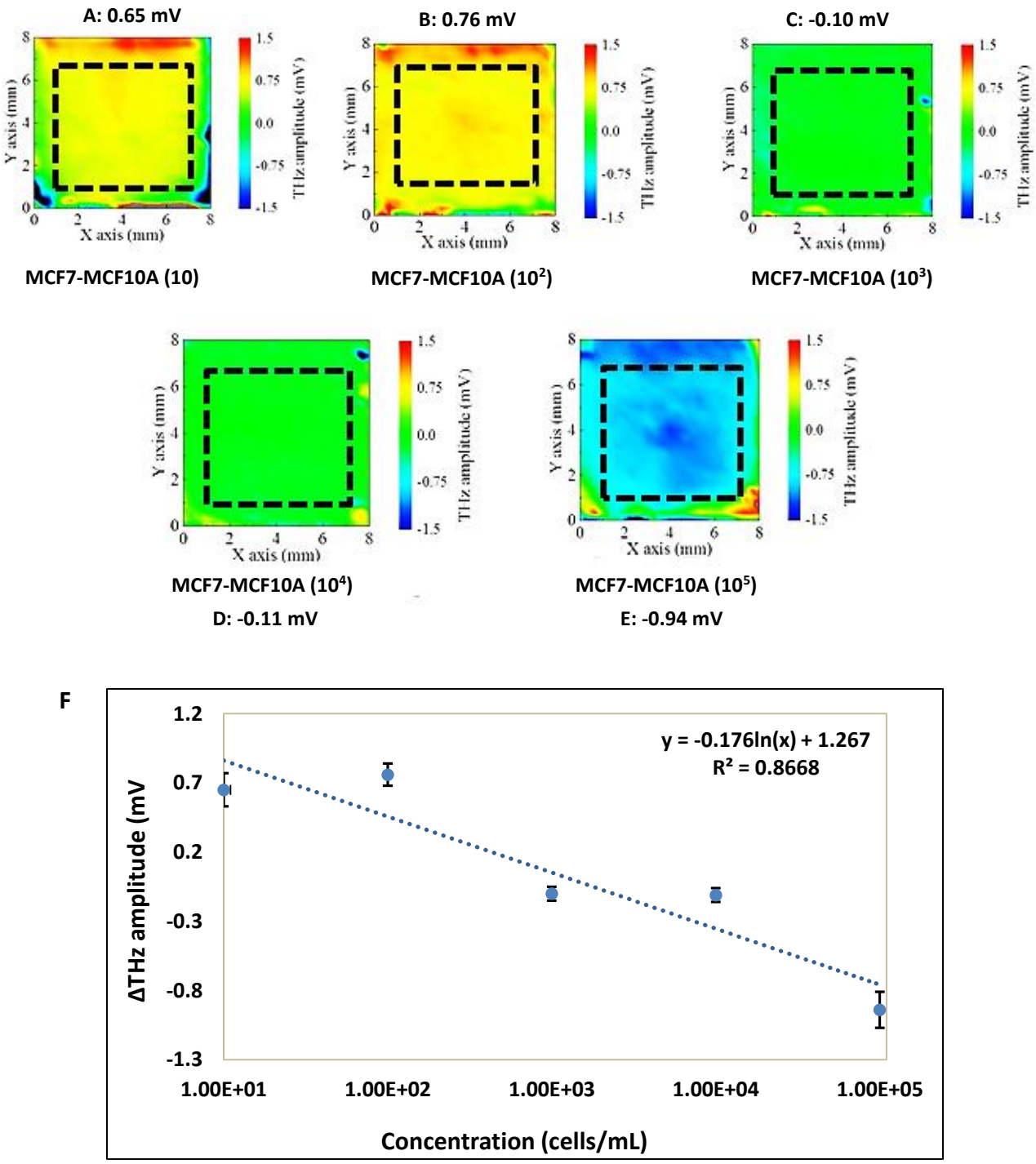
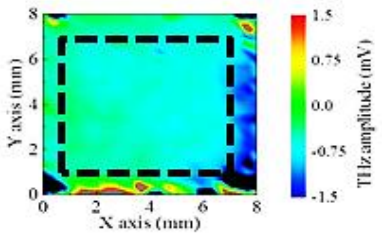
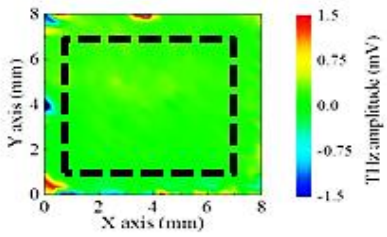


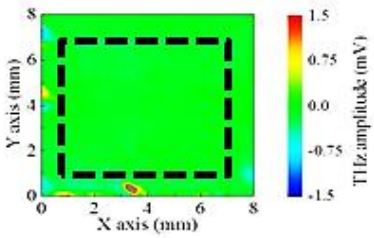
Figure 5.8: THz peak amplitude mapping of different numbers of MCF7 cancer cells on the surface of MAMB1 sensing plate (second trial). A-E: differential images (MCF7-MCF10A) of MAMB1 sensing plate when different numbers of MCF7 cancer cells were used (10, 10^2, 10^3, 10^4 and 10^5 respectively). F: THz amplitude (mV) curve against different number of cells. To calculate the THz amplitude (mV) an average area of 7 mm was created on both the X axis and the Y axis. Images were created using Origin software.



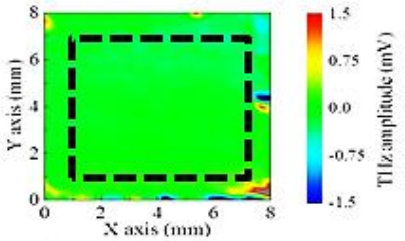
MDA-MB-415-MCF10A (10)
A: -0.59 mV



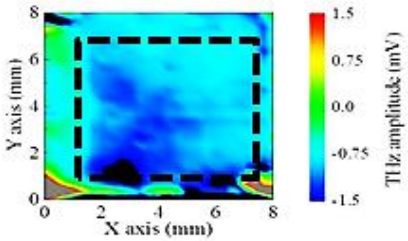
MDA-MB-415-MCF10A (10^2)
B: 0.16 mV



MDA-MB-415-MCF10A (10^3)
C: -0.13 mV



MDA-MB-415-MCF10A (10^4)
D: -0.11 mV



MDA-MB-415-MCF10A (10^5)
E: -0.99 mV

F

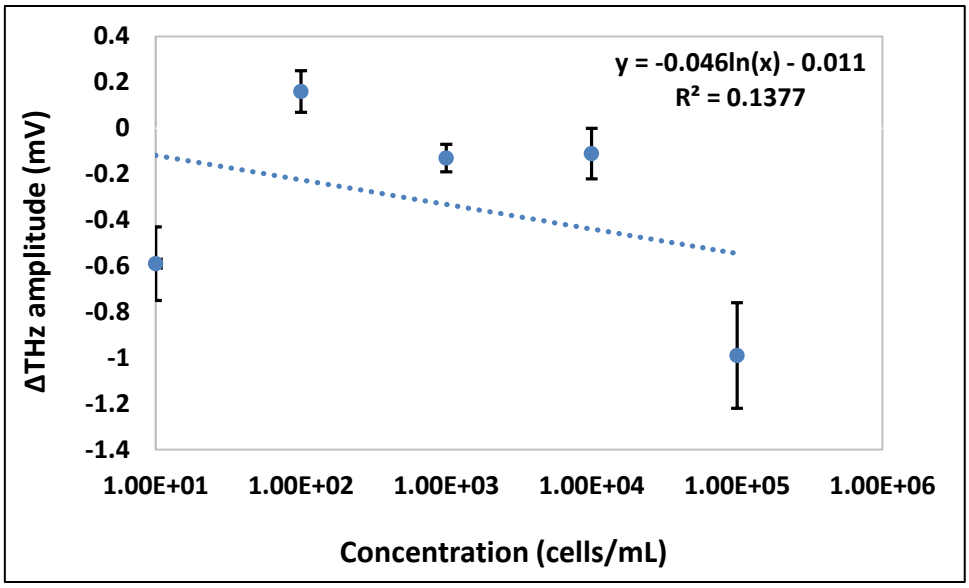
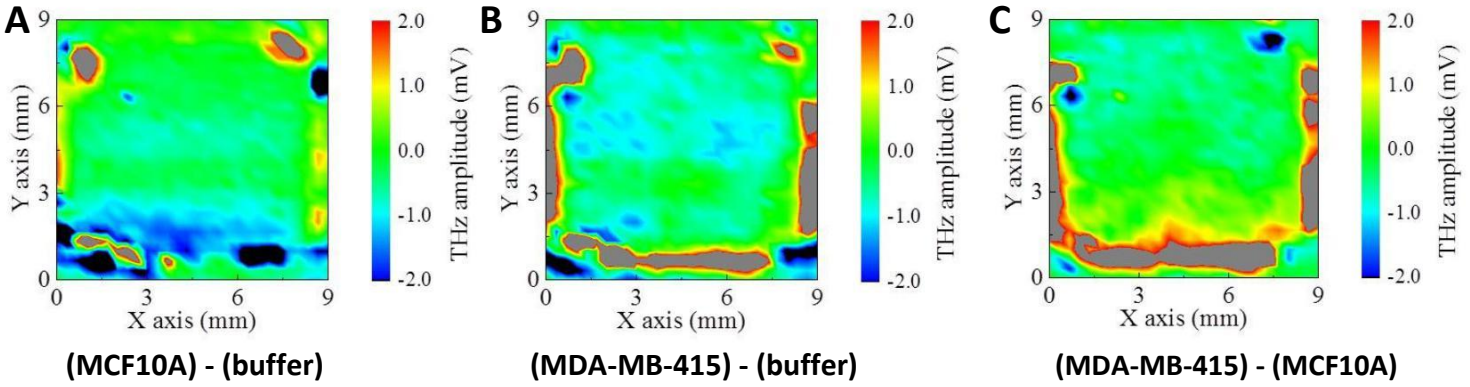


Figure 5.10: THz peak amplitude mapping of different numbers of MDA-MB-415 cancer cells on the surface of MAMA2 sensing plate (second trial). A-E: differential images (MDA-MB-415-MCF10A) of MAMA2 sensing plate when different numbers of MDA-MB-415 cancer cells were used (10, 10^2, 10^3, 10^4 and 10^5 respectively). F: THz amplitude (mV) curve against different number of cells. To calculate the THz amplitude (mV) an average area of 7 mm was created on both the X axis and the Y axis. Images were created using Origin software.

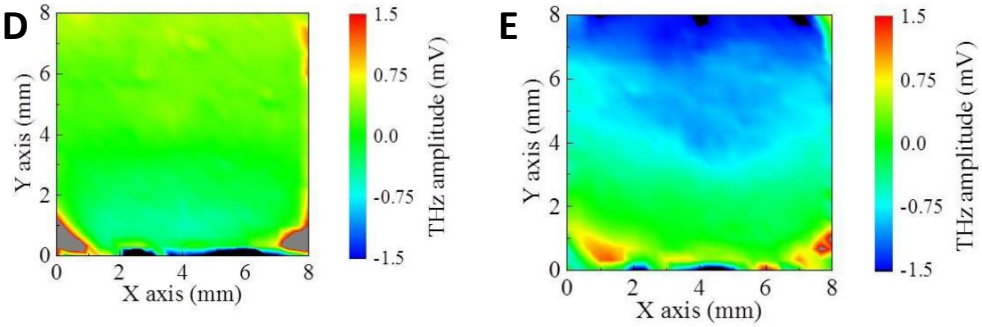
5.3.2.3 Selectivity of MAMB1 and MAMA2 aptamers

The selectivity of both aptamers was investigated by adding MDA-MB-415 cells (1×10^5 cells) on the MAMB1 sensing plate, and adding MCF7 cells (1×10^5 cells) to the MAMA2 sensing plate, and scanning both plates using the TCM. The results showed that the MAMB1 aptamer was very selective to MCF7 cancer cells, as shown by the TCM differential images in both first and second trials (Figure 5.11C, E and K). The change in THz amplitude was higher when MAMB1 interacted with MCF7 cells compared to MCF10A and MDA-MB-415 cells (Figure 5.11K). On the other hand, the interaction of MAMA2 with MCF7 was higher than MDA-MB-415 in both the first and second trials (Figure 5.11 H, J and K). Furthermore, the binding of the MAMA2 aptamer to its target breast cancer cells MDA-MB-415 was low, as indicated by the value of the THz amplitude (0.69 mV) in both trials compared to 2.025 mV for MAMB1 to MCF7 (Figure 5.11 K). This might have been because the two cell lines are different (in terms of size and gene expression) and therefore they could have exhibited different interaction behavior to their targets, and that the TCM method was sensitive enough to detect the changes of these cells. Moreover, MAMA2 could have been more sensitive to the conditions of the surrounding environment than the MAMB1 aptamer, as we mentioned earlier, since buffer, temperature, and pH could affect the interaction of aptamers to their targets [12]. The high non-specific binding of MAMA2 to MCF7 could be reduced by using a blocking reagent such as ethanol amine to block the extra sites on the sensing plate that could bind non-specifically to MAMA2 aptamer on the surface of the sensing plate.

To summarize, the selective study results were done twice on two different chips and revealed that TCM could be used as a tool to test the selectivity of the aptamers to their targets. However, more trials needed to be done and thus, selectivity tests were repeated three times on new chips for each aptamer at the INRS (Canada).

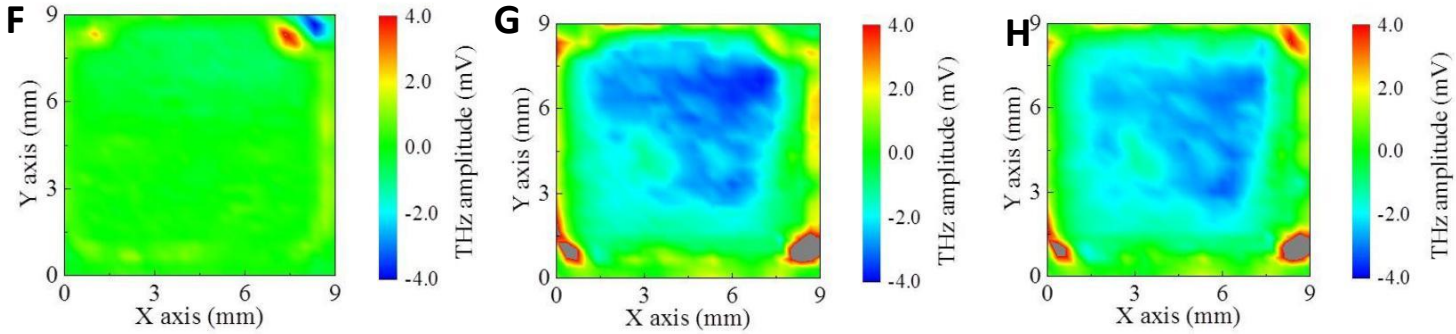


(MCF10A) - (buffer) (MDA-MB-415) - (buffer) (MDA-MB-415) - (MCF10A)

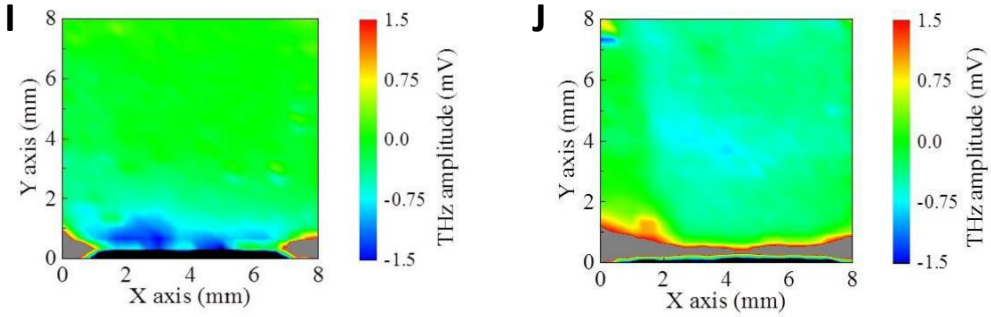


(MCF10A) - (buffer (Second trial)

(MDA-MB-415) - (MCF10A) (Second trial)



(MCF10A) - (buffer) (MCF7) - (buffer) (MCF7) - (MCF10A)



(MCF10A) - (buffer) (Second trial)

(MCF7) - (MCF10A) (Second trial)

K

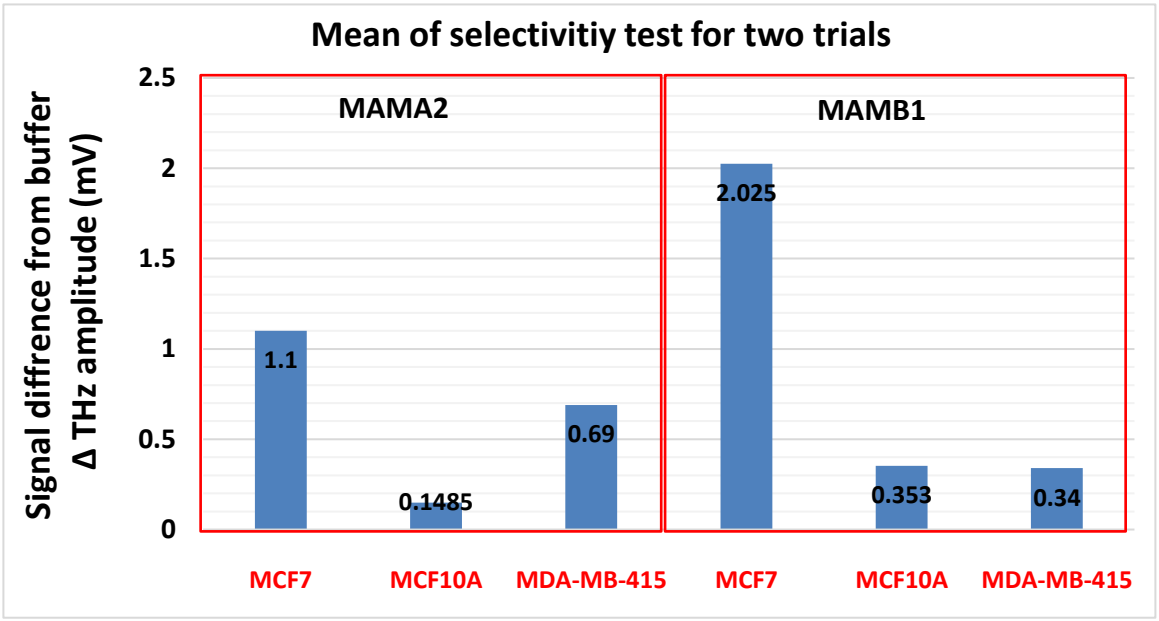


Figure 5.11: THz peak amplitude mapping obtained for MAMB1 and MAMA2 sensing plates after the selectivity test. A-C: differential images of MAMB1 sensing plate after adding MDA-MB-415 cells for the first trial. D,E: differential images of MAMB1 sensing plate after adding MDA-MB-415 cells for the second trial. F-H: differential images of MAMA2 sensing plate after adding MCF7 cells for the first trial. I, J: differential images of MAMA2 sensing plate after adding MCF7 cells for the second trial. K: A graph represent the mean values of THz amplitude for MAMB1 and MAMA2 sensing plate after adding MCF7, MDA-MB-415 and MCF10A cells on each of them. Images were created using Origin software, and the graph was created using excel 2013.

5.3.3 INRS results

5.3.3.1 MAMB1 and MAMA2 aptamers bind to their target breast cancer cells using TCM

Both MAMB1 and MAMA2 sensing plates were scanned using the TCM at the INRS, as mentioned earlier in this chapter. The scanning was done three times on the same sensing plate for each number of cells (10-10^6) on each chip, and for the selectivity tests (MDA-MB-415 (1×10^6) on MAMB1 chip, and MCF7 (1×10^6) on MAMA2 chip).

The results obtained from MAMB1 chip showed that the THz amplitude for MCF7 (cancer cells) was higher than the one obtained from the buffer and MCF10A (normal cells) at all number of cells (Figures 12- 17 A-C for each). The differential image results indicated that the MAMB1 aptamer was specific for cancer cells only, regardless of the number of cells measured, as the THz amplitude indicated (Figures 12-17 D for each). Plotting the number of cells against the THz amplitude change (differential images) showed increasing (toward the positive) of the THz amplitude with increasing number of cells from 0.08 mV at 10 cells to 0.42 mV at 10^6 cells (Figure 5.18 A-F for each), with the relationship being linear with $R^2 = 0.991$ (Figure 5.18 G). The binding of the MCF7 cancer cells to the MAMB1 aptamer (as ΔTHz after subtracting from normal cells) was significant at all numbers of cells ($P<0.05$) compared to the buffer, indicating that the TCM could detect as low as 10 cancer cells on the sensing plate (MAMB1) (Figure 5.18). These results were similar to the first trial using the MAMB1 chip that was done at Okayama (Figure 5.7). However, the variation here was very small because the three trials were done on the same chip, reducing the variability coming from using different sensing plates, which was the major factor for the large variability seen in the results obtained at Okayama. Moreover, the chip was scanned for a homogenous THz signal before the scanning of the samples started, which also contributed to reducing the variability.

The MAMA2 chip showed results that were similar to those of MAMB1. The THz amplitude for MDA-MB-415 (cancer cells) was higher than that obtained from the buffer and MCF10A (normal cells) at all number of cells (Figures 12-17 E-G for each). Subtracting the THz signal of normal cells from cancer cells (differential images) indicated the specificity of MAMA2 to cancer cells vs non-cancer cells (Figures 12-17 H for each). The slope of the THz amplitude

(THz) as a function of the number of cells was negative (this could be the result of the initial charges on the sensing plate or the charges of the cells, as discussed earlier in the Okayama results). The absolute THz amplitude increases with increasing number of cells from -0.076 mV at 10 cells to -0.33 mV at 10^6 cells of MDA-MB-415 (Figure 5.19 A-F). The relation was linear with R^2 = 0.992 (Figure 5.19G). The THz amplitude was significant at all number of cells ($P<0.05$) compared to the buffer, indicating that the TCM could detect as small as 10 cancer cells on the sensing plate (MAMA2) (Figure 5.19). The results here were similar to the first trial on MAMA2 chip obtained from Okayama (Figure 5.8). However, the THz signal from the INRS results were more uniform because the three trials were performed on the same chip, similar to the case for MAMB1.

Another important factor that have contributed to reducing the variability of the INRS compared with those at Okayama results is the scanning time of the sensing plates. The time was 90 min for the Okayama experiments vs 10 min for the INRS experiments. The interaction of aptamers with their targets takes place in the first 5 min (when the target is introduced to the aptamer), until all the binding sites on the target are occupied with the aptamer (saturation) The interaction will reach equilibrium, and the aptamers might go into cycles of binding and unbinding the target (in the case of the 90 min scan). This could affect the THz signal and create more variability.

The limit of detection and limit of quantification were 1 and 10 cells in a million of cancer cells respectively for both MAMB1 and MAMA2 aptamers. These values indicated the high sensitivity of our TCM detection method.

In summary, the results obtained at the INRS indicate that the TCM could be used to detect a small number of breast cancer cells (1 cell in a million cells of breast cancer cells). However, more optimization needs to be done in the future as this was the first time the TCM has been used for detecting breast cancer cells. Moreover, testing the binding of MAMB1 and MAMA2 to their breast cancer cells in a mixture of normal cells and spiked cancer cells would reveal the sensitivity of TCM in detecting metastatic breast cancer cells.

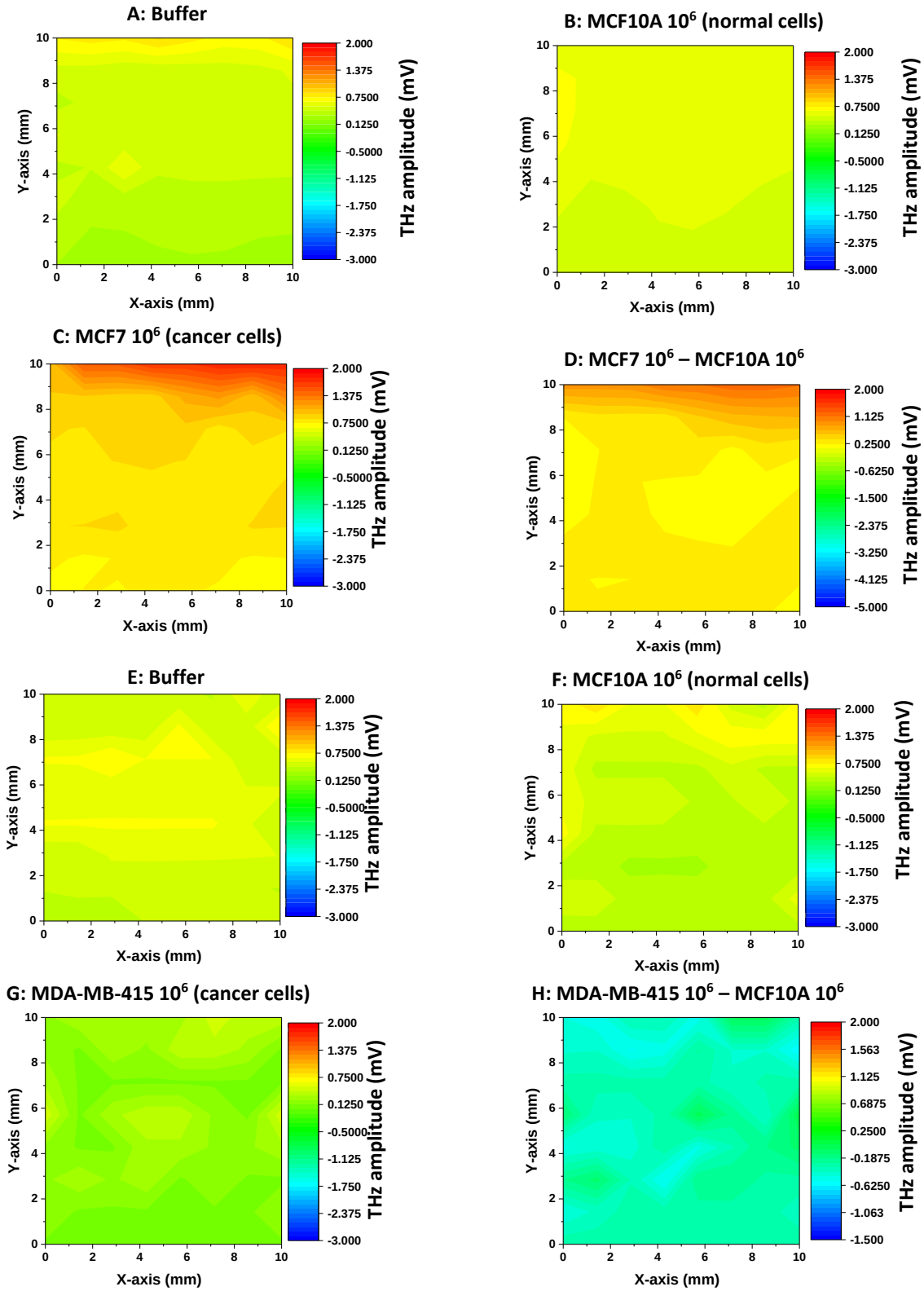


Figure 5.12: THz peak amplitude mapping for MAMB1 and MAMA2 chips at 10^6 of breast cancer cells. A,E: buffer alone B,F: MCF10A (normal control cells), C,G: MCF7 cells and MDA-MB-415 (breast cancer cells) respectively D,H : Differential images after subtraction from control cells. Trials were repeated three times (for each chip) on the same sensing plate. Images were created using Origin software. MCF7 and MDA-MB-415 breast cancer cells (1×10^6) were placed on the surface of their sensing plates by pipetting a 120µL. then each sensing plate was scanned in 2D for 10 min. many data points were measured during the scanning. These data points were converted to contour images using Origin software.

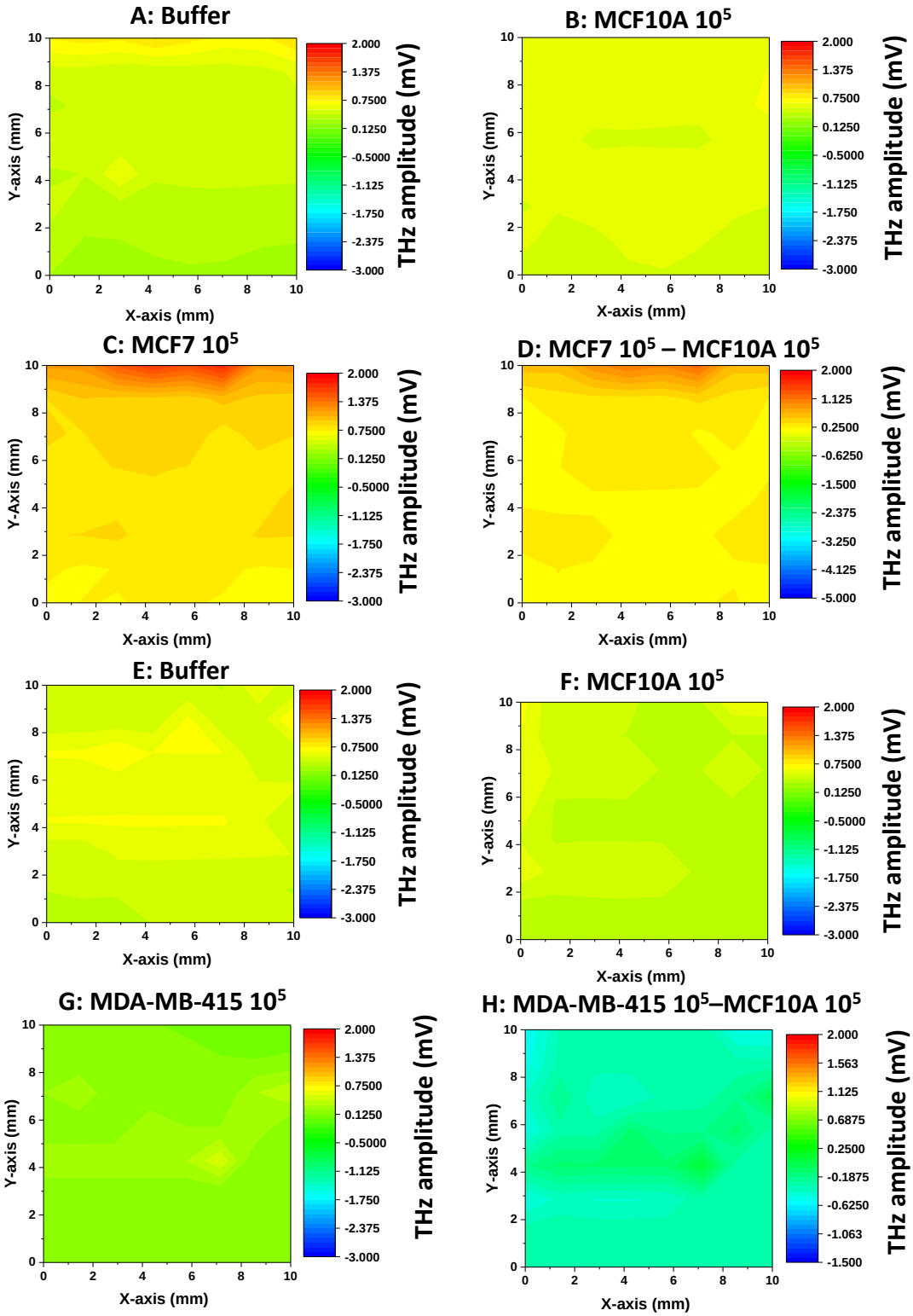


Figure 5.13: THz peak amplitude mapping for MAMB1 and MAMA2 chips at 10^5 of breast cancer cells. A,E: buffer alone B,F: MCF10A (normal control cells), C,G: MCF7 cells and MDA-MB-415 (breast cancer cells) respectively D,H : Differential images after subtraction from control cells. Trials were repeated three times (for each chip) on the same sensing plate. Images were created using Origin software. MCF7 and MDA-MB-415 breast cancer cells (1×10^5) were placed on the surface of their sensing plates by pipetting a 120µL. then each sensing plate was scanned in 2D for 10 min. many data points were measured during the scanning. These data points were converted to contour images using Origin software.

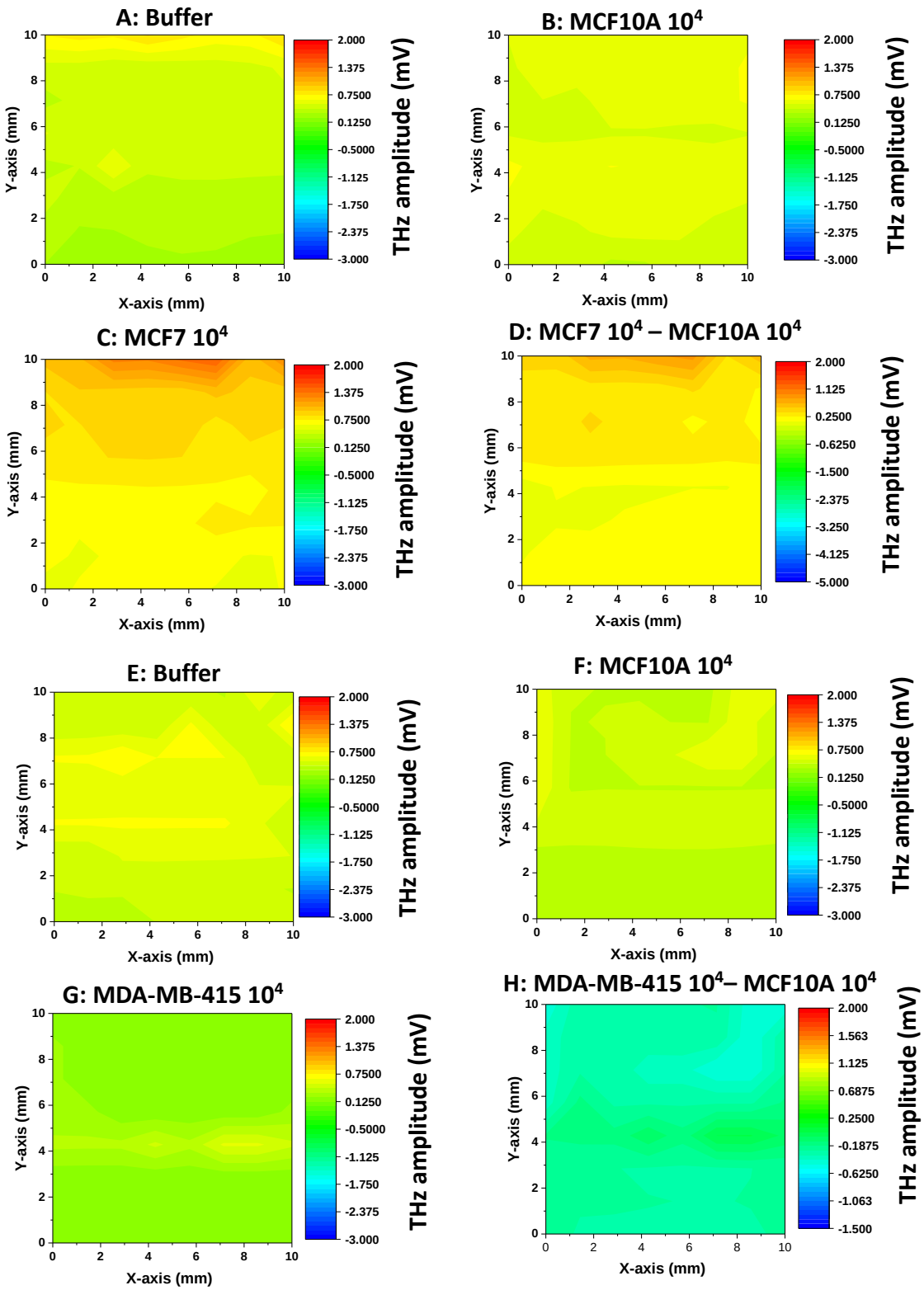


Figure 5.14: THz peak amplitude mapping for MAMB1 and MAMA2 chips at 10^4 of breast cancer cells. A,E: buffer alone B,F: MCF10A (normal control cells), C,G: MCF7 cells and MDA-MB-415 (breast cancer cells) respectively D,H : Differential images after subtraction from control cells. Trials were repeated three times (for each chip) on the same sensing plate. Images were created using Origin software. MCF7 and MDA-MB-415 breast cancer cells (1×10^4) were placed on the surface of their sensing plates by pipetting a 120µL. then each sensing plate was scanned in 2D for 10 min. many data points were measured during the scanning. These data points were converted to contour images using Origin software.

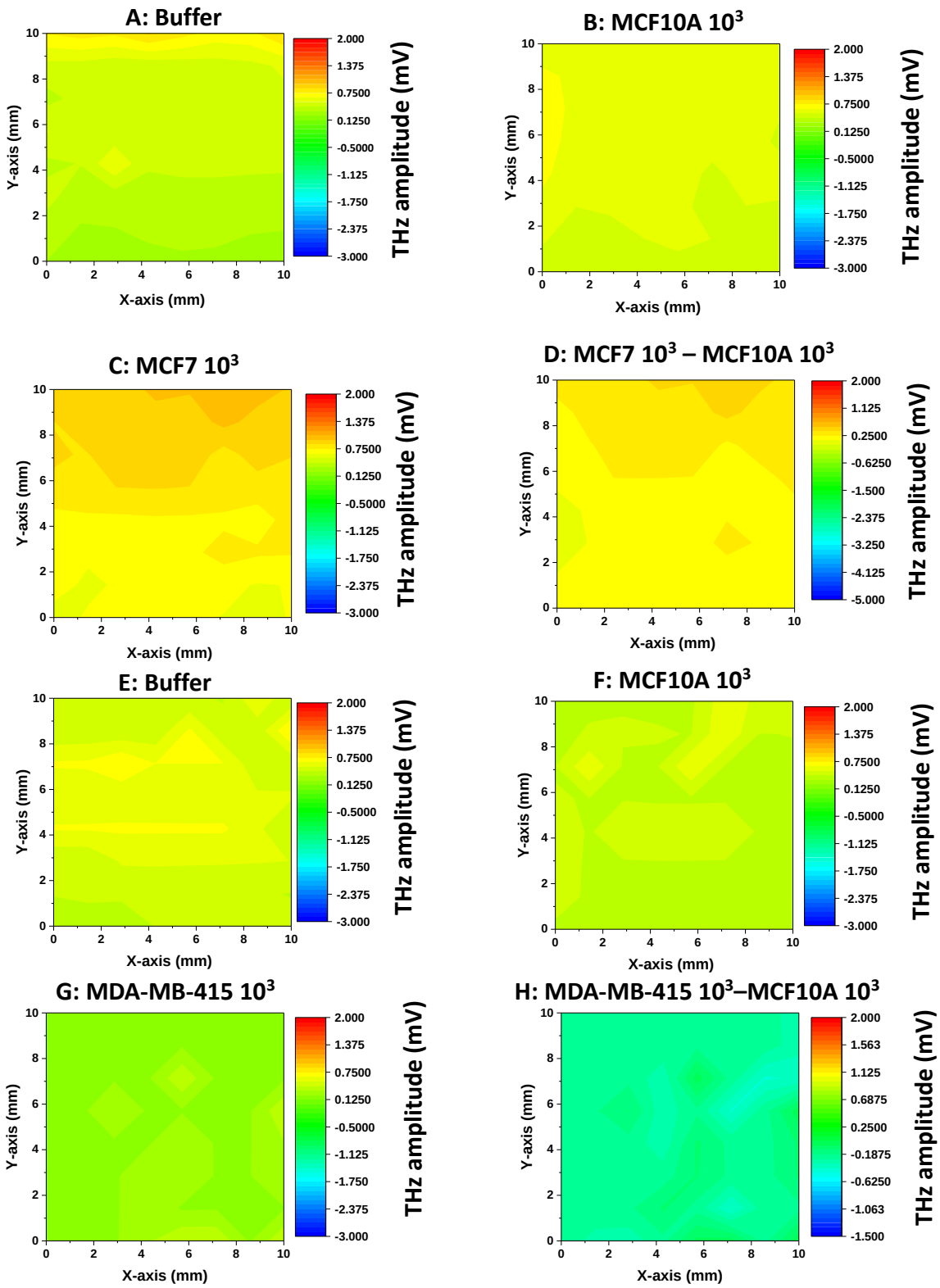


Figure 5.15: THz peak amplitude mapping for MAMB1 and MAMA2 chips at 10^3 of breast cancer cells. A,E: buffer alone B,F: MCF10A (normal control cells), C,G: MCF7 cells and MDA-MB-415 (breast cancer cells) respectively D,H : Differential images after subtraction from control cells. Trials were repeated three times (for each chip) on the same sensing plate. Images were created using Origin software. MCF7 and MDA-MB-415 breast cancer cells (1×10^3) were placed on the surface of their sensing plates by pipetting a 120µL. then each sensing plate was scanned in 2D for 10 min. many data points were measured during the scanning. These data points were converted to contour images using Origin software.

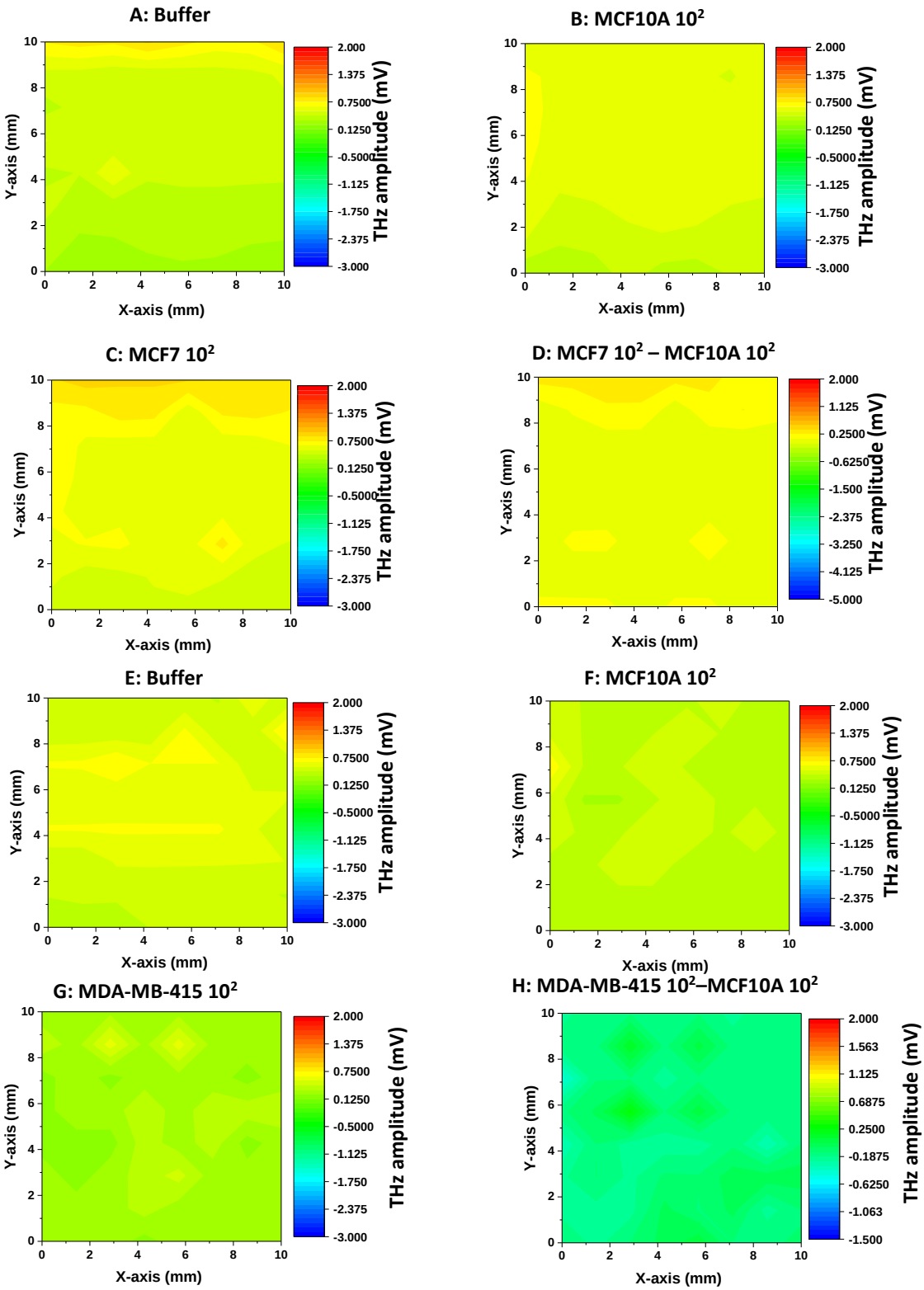


Figure 5.16: THz peak amplitude mapping for MAMB1 and MAMA2 chips at 10^2 of breast cancer cells. A,E: buffer alone B,F: MCF10A (normal control cells), C,G: MCF7 cells and MDA-MB-415 (breast cancer cells) respectively D,H : Differential images after subtraction from control cells. Trials were repeated three times (for each chip) on the same sensing plate. Images were created using Origin software. MCF7 and MDA-MB-415 breast cancer cells (1×10^2) were placed on the surface of their sensing plates by pipetting a 120µL. then each sensing plate was scanned in 2D for 10 min. many data points were measured during the scanning. These data points were converted to contour images using Origin software.

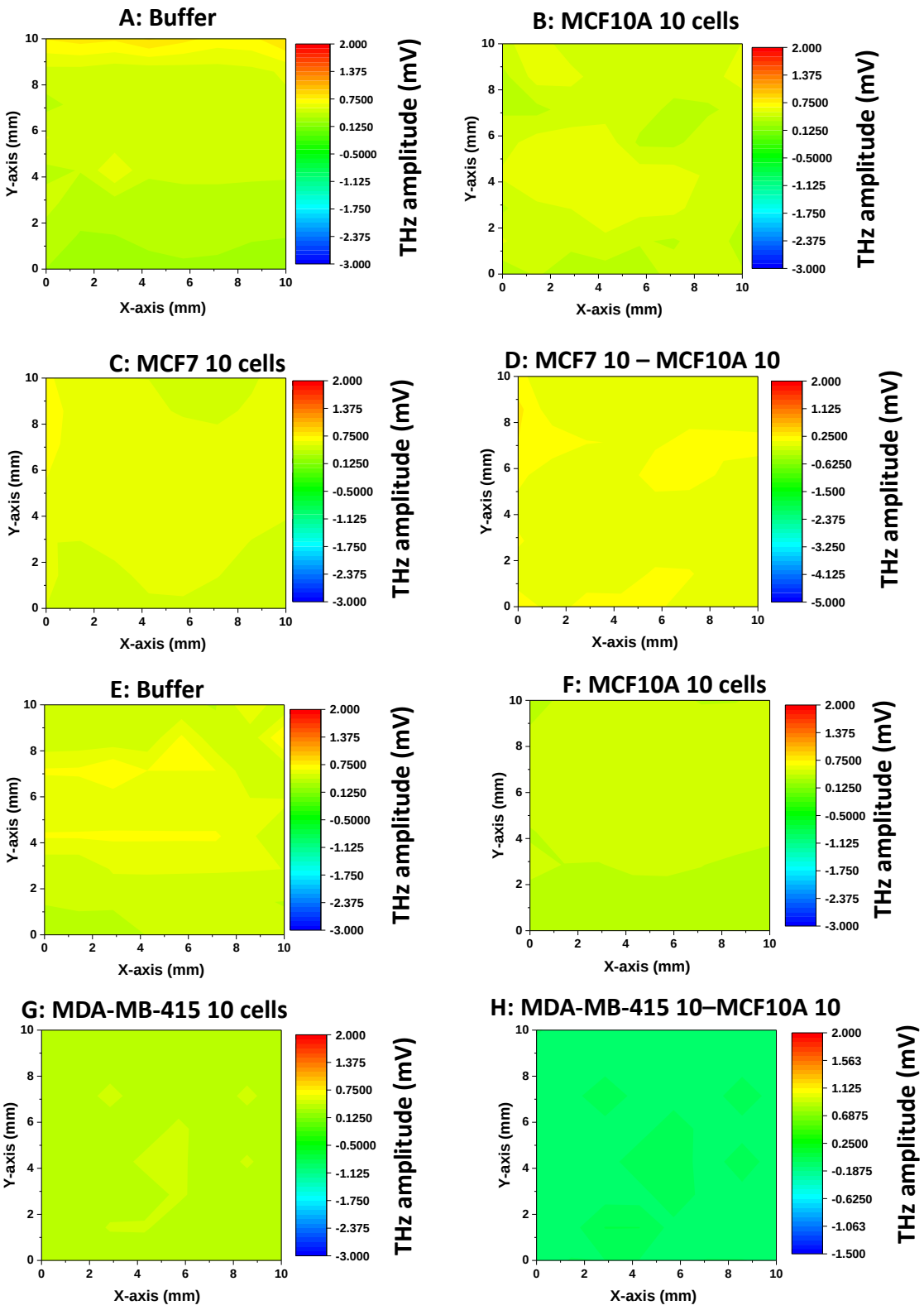


Figure 5.17: THz peak amplitude mapping for MAMB1 and MAMA2 chips at 10 cells of breast cancer cells. A,E: buffer alone B,F: MCF10A (normal control cells), C,G: MCF7 cells and MDA-MB-415 (breast cancer cells) respectively D,H : Differential images after subtraction from control cells. Trials were repeated three times (for each chip) on the same sensing plate. Images were created using Origin software. MCF7 and MDA-MB-415 breast cancer cells (1×10) were placed on the surface of their sensing plates by pipetting a 120µL. then each sensing plate was scanned in 2D for 10 min. many data points were measured during the scanning. These data points were converted to contour images using Origin software.

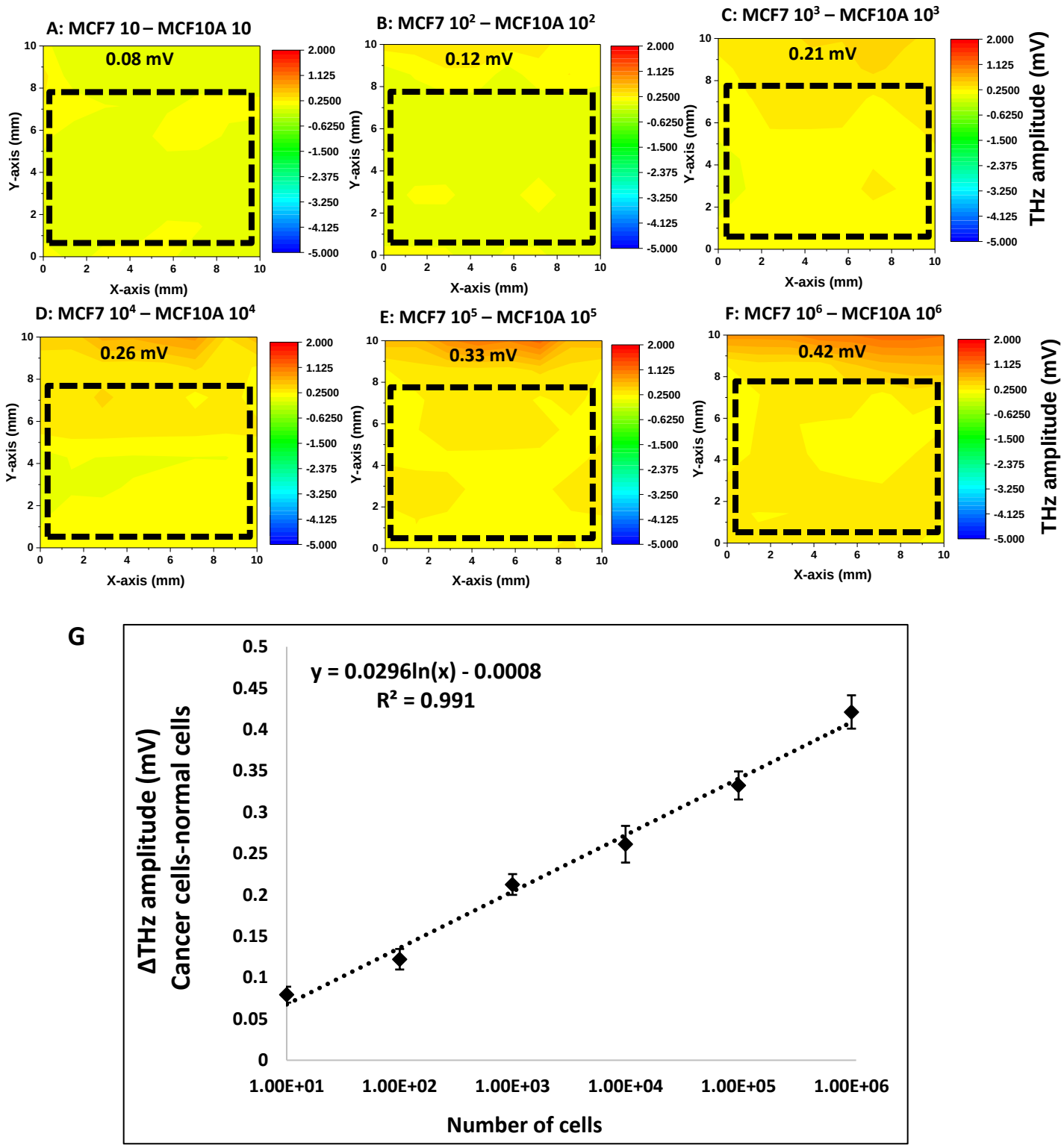


Figure 5.18: THz peak amplitude mapping of different numbers of MCF7 cancer cells on the surface of MAMB1 sensing plate. A-F: differential images (MCF7-MCF10A) of MAMB1 sensing plate when different numbers of MCF7 cancer cells were used (10, 10^2, 10^3, 10^4, 10^5, and 10^6 respectively). G: THz amplitude (mV) against different number of MCF7 cells. To calculate the THz amplitude (mV) an average area of 2.0 mm was created on the X axis and 1.3 mm on the Y axis. Images were created using Origin software. Values are shown as means ± S.E.M. of three trials. The statistical significance was determined by one way ANOVA followed by Fisher LSD multiple tests using SPSS software (SPSS, version 23). P Values less than 0.05 were considered to be significantly different.

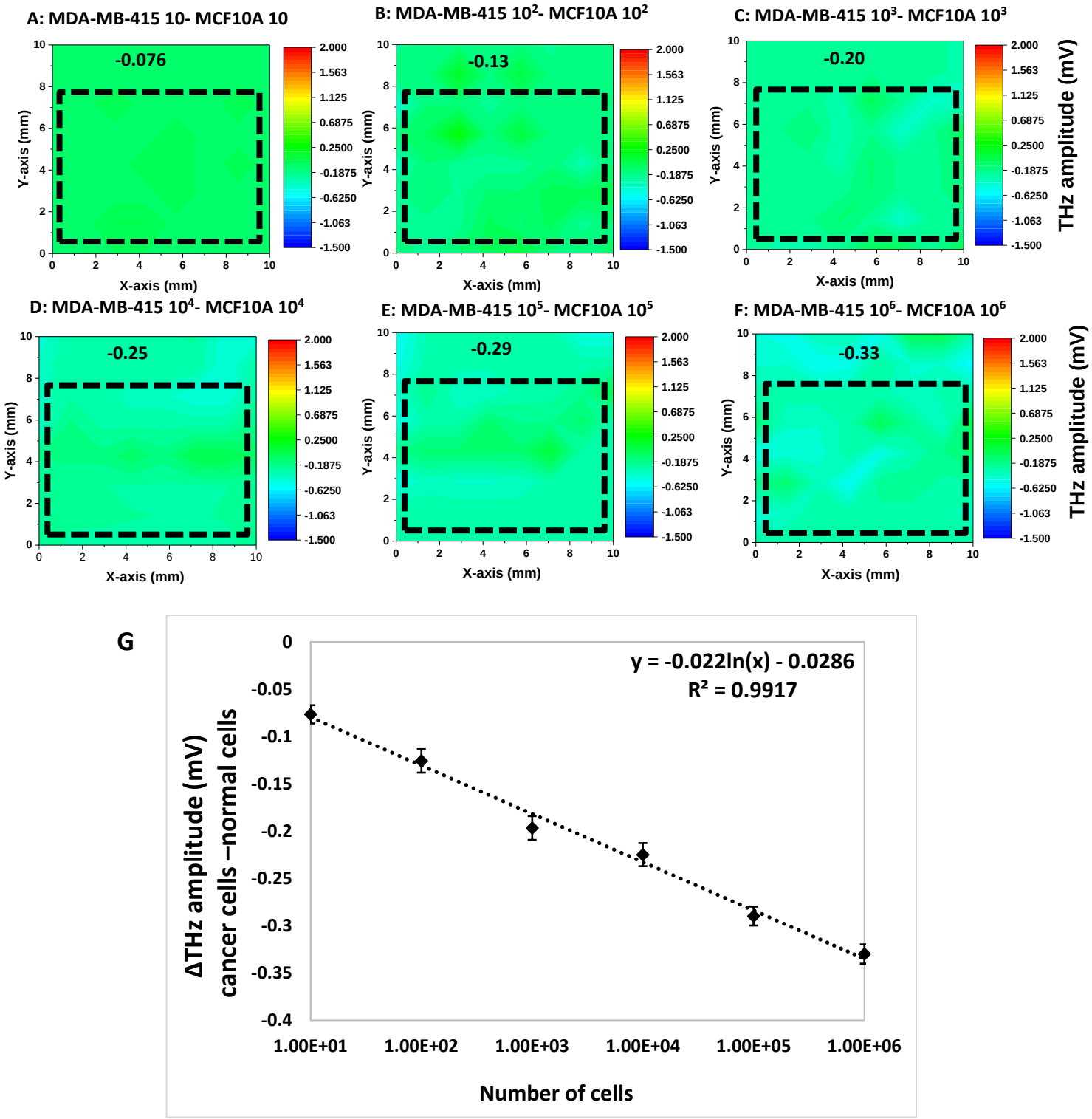


Figure 5.19: THz peak amplitude mapping of different numbers of MDA-MB-415 cancer cells on the surface of MAMA2 sensing plate. A-F: differential images (MDA-MB-415-MCF10A) of MAMA2 sensing plate when different numbers of MDA-MB-415 cancer cells were used (10, 10^2, 10^3, 10^4, 10^5, and 10^6 respectively). G: THz amplitude (mV) curve against different number of cells. To calculate the THz amplitude (mV) an average area of 2.0 mm was created on the X axis and 1.3 mm on the Y axis. Images were created using Origin software. Values are shown as means ± S.E.M. of three trials. The statistical significance was determined by one way ANOVA followed by Fisher LSD multiple tests using SPSS software (SPSS, version 23). P Values less than 0.05 were considered to be significantly different.

5.3.3.2 MAMB1 and MAMA2 are selective for their target cancer cells

The selectivity test of both MAMB1 and MAMA2 aptamers immobilized on the surface of the sensing plates were performed by adding MDA-MB-415 cells (1×10^6 cells) to the MAMB1 sensing plate, and MCF7 (1×10^6 cells) to the MAMA2 sensing plate. The results showed that MAMB1 and MAMA2 aptamers were selective to their breast cancer cell targets (MCF7 and MDA-MB-415, respectively) (Figure 5.20 and 21) as shown by the differential images (subtracting from the buffer) (Figure 5.20 D and E, Figure 5.201 D and E).

MAMA2 showed high nonspecific binding when tested at Okayama (Figure 5.11). However, the number of cells added to the sensing plate was different (1×10^5 at Okayama versus 1×10^6 cells at the INRS, for both sensing plates). 10^6 cells could be a better choice to perform the selectivity test. Applying a different number of cells to the selectivity test could answer more questions.

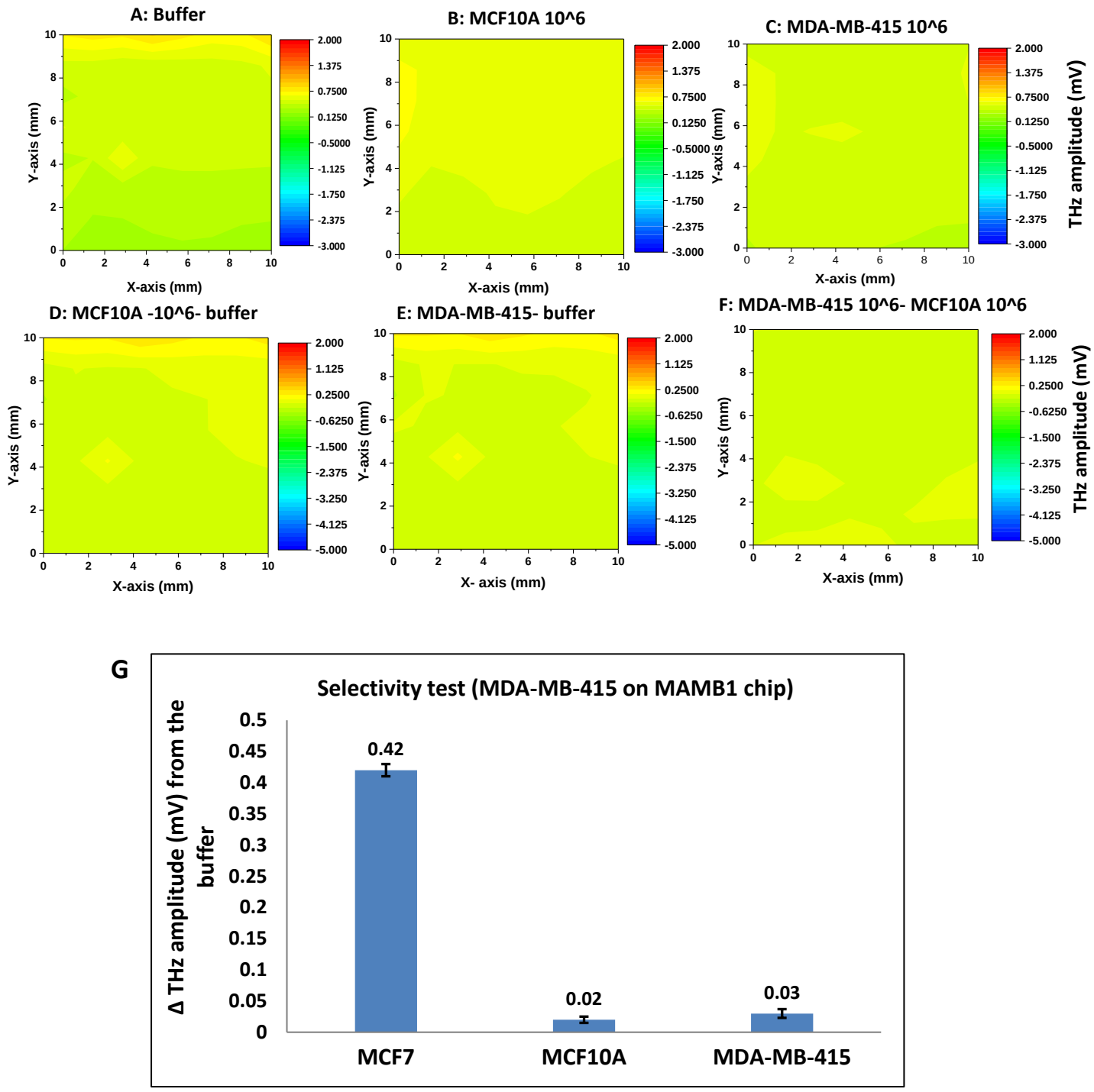
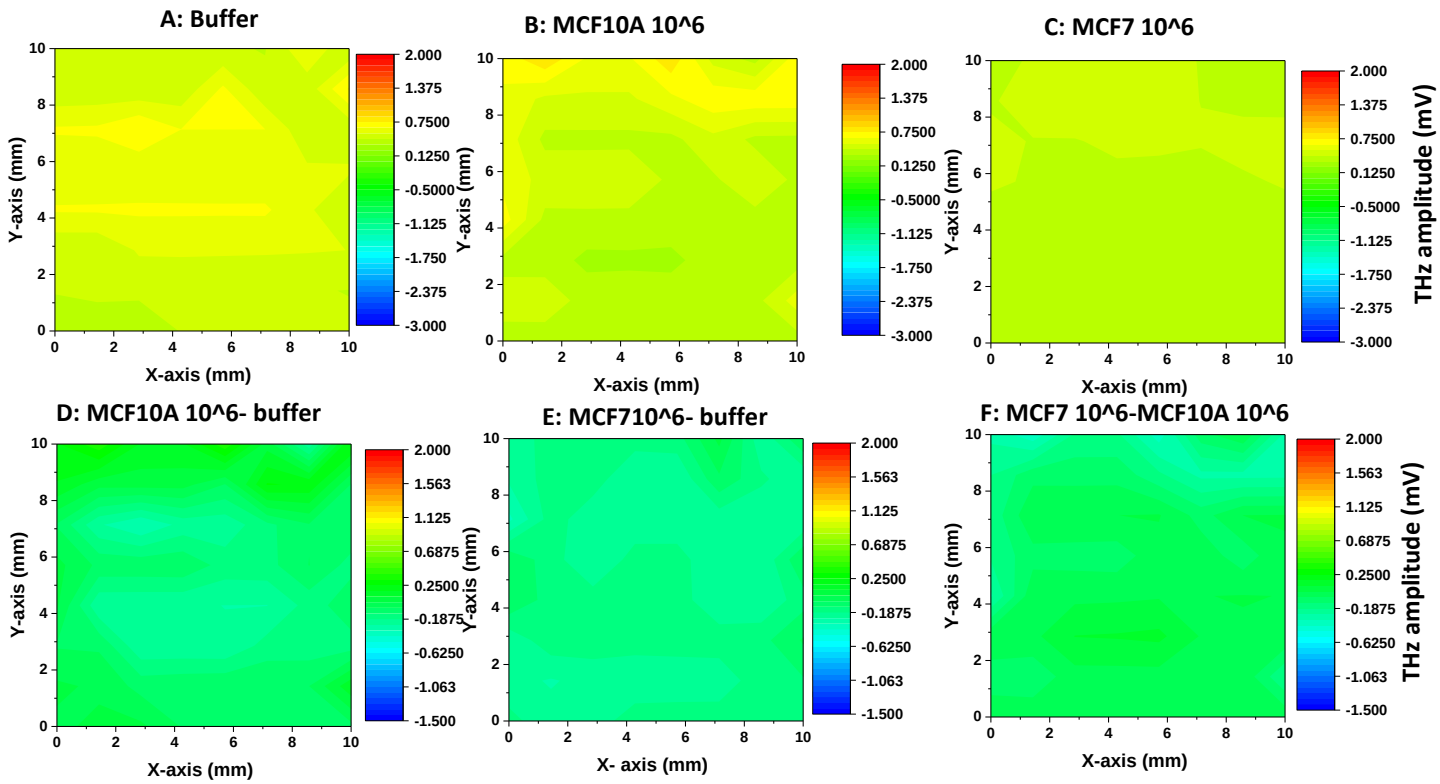


Figure 5.20: THz peak amplitude mapping obtained for MAMB1 sensing plate after the selectivity test. A-C: images of MAMB1 sensing plate pf buffer (A), MCF10A (B), and MDA-MB-415 cells (c). D-F: differential images of MCF10A and MDA-MB-415 subtracting from the buffer (D,E) and MDA-MB-415 subtracting from MCF10A (F). G: a graph represent the mean values (of three trials) of THz amplitude of MCF, MCF10A and, MDA-MB-415 subtracting from the buffer. Images were created using Origin software, and the graph was created using excel 2013.



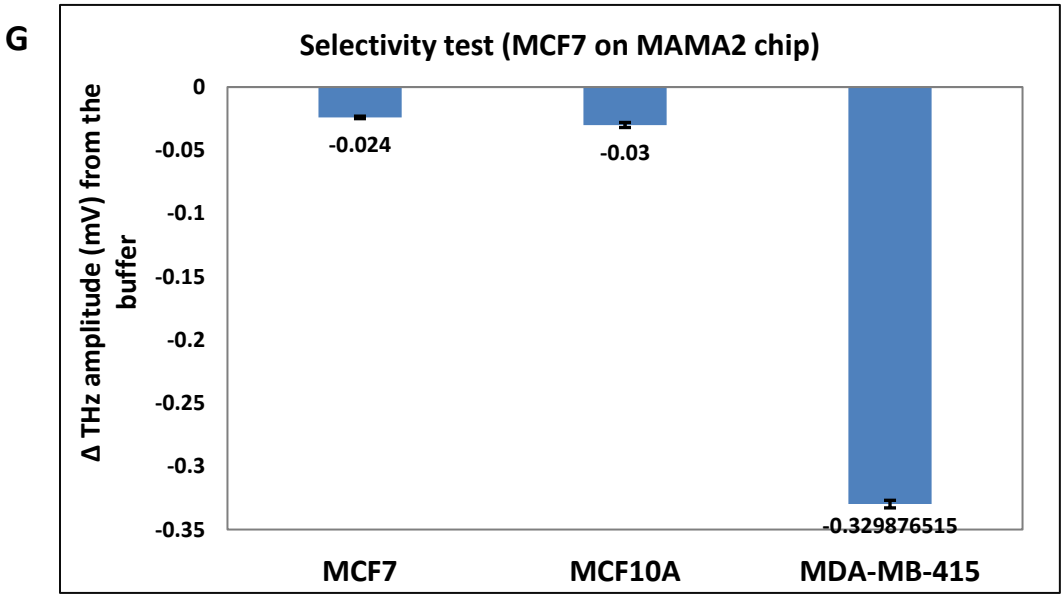


Figure 5.21: THz peak amplitude mapping obtained for MAMA2 sensing plate after the selectivity test. A-C: images of MAMB1 sensing plate of buffer (A), MCF10A (B), and MCF7 cells (c). D-F: differential images of MCF10A, MCF7 subtracting from the buffer (D,E) and MCF7 subtracting from MCF10A (F). G: a graph represent the mean values (of three trials) of THz amplitude of MCF, MCF10A and, MDA-MB-415 subtracting from the buffer. Images were created using Origin software, and the graph was created using excel 2013.

5.4 Conclusion

For the first time, the interaction of aptamers with their targets on the surface of the sensing plate was performed in this study. Aptamers are small in size and are easily modified to be immobilized on the surface of the sensing plate. Moreover, aptamers are more stable than antibodies and thus, can handle variable conditions (e.g. scanning at RT).

The experiments were performed at Okayama University (Dr. Kiwa's lab) and the INRS (Dr. Ozaki's lab). The conditions of TCM scanning were slightly different at both labs. The first step of this study was to run initial testing to study the interaction of MAMB1 and MAMA2 aptamers that were immobilized on the surface of the sensing plates, allowing scanning times of 90 min when MCF7 cells were pipetted on the surface of the MAMB1 plate, and MDA-MB-415 on the surface of the MAMA2 plate. The results of the initial testing showed that the TCM could be used to detect breast cancer cells, as the TCM signal from the aptamers binding with breast cancer cells was higher than that binding to normal breast cells (control) (Figure 5.5). The results using different numbers of cells showed more consistent results for MAMB1 aptamer interacting with its target MCF7 cells, compared to MAMA2 interacting with its target MDA-MB-415 cells (Figures 5.6-5.8). These results were similar to the ones obtained from the initial testing. A third set of experiments were done to investigate the selectivity of MAMB1 and MAMA2 aptamers to their target breast cancer cells by TCM. The results showed high selectivity of MAMB1 to its target breast cancer cells, and high non-specific binding of MAMA2 to MCF7 (Figure 5.11). However, the variability of the results obtained at Okayama was high and thus, more trials were performed at the INRS (three trials on the same sensing plate).

The results obtained at the INRS were less variable, because we used one sensing plate of each aptamer for all the trials to minimize the variability and test the reproducibility of the aptamers. ΔTHz amplitude increased with increasing number of both cell types on both sensing plates (Figures 12-18). Moreover, the TCM could detect ($P<0.05$) as low as 10 cancer cells for both sensing plates. In addition, the result showed that both aptamers were selective at 10^6 cancer cells (Figures 5.19-20).

In conclusion, TCM technology could use aptamers as recognition elements to detect breast cancer cells. This proof-of-concept study could be the spark for a new aptamer detection method for cancer and maybe in other fields in the future.

References

1. Kiwa T, Sakai K, Tsukada K. A novel laser-terahertz emission microscope is used to visualize changes in chemical and electrical potential. Biomedical Optics & Medical Imaging. 2013.

2. Kiwa T, Hagiwara T, Shinomiya M, Sakai K, and Tsukada K. Work function shifts of catalytic metals under hydrogen gas visualized by terahertz chemical microscopy. Opt Express. 2012, 20, 11637.

3. Kiwa T, Tenma A, Takahashi S, Sakai K, Tsukada K. Label free immune assay using terahertz chemical microscope. Sensors and Actuators Sensors and Actuators. 2013, B 187: 8–11.

4. Kiwa T, Kondo Y, Minami Y, Kawayama I, Tonoushi M, Tsukada K. Terahertz chemical microscope for label-free detection of protein complex. Appl Phys Lett. 2010, 96, 211114.

5. Glass NR, Tjeung R, Chan P, Yeo LY, Friend JR. Organosilane deposition for microfluidic applications. Biomicrofluidics. 2011, 5(3):36501-365017.

6. Janissen R, Oberbarnscheidt L, Oesterhelt F. Optimized straight forward procedure for covalent surface immobilization of different biomolecules for single molecule applications. Colloids Surf B Biointerfaces. 2009, 71(2):200-2007.

7. Römhildt L. Biochemical functionalization of silicon dioxide surfaces for sensing applications. PhD thesis in material sciences and nanotechnology. Dresden University of Technology. Germany, 2014. Obtained from the following data base http://www.qucosa.de/recherche/frontdoor/?tx_slubopus4frontend%5bid%5d=urn:nbn:de:bsz:14-qucosa-148666

8. Kim J, Cho J, Seidler PM, Kurland NE, Yadavalli VK. Investigations of chemical modifications of amino-terminated organic films on silicon substrates and controlled protein immobilization. Langmuir. 2010, 26(4):2599-2608.

9. Hianik T, Ostatná V, Sonlajtnerova M, Grman I. Influence of ionic strength, pH and aptamer configuration for binding affinity to thrombin. Bioelectrochemistry. 2007, 70(1):127-33.

10. Yang X, Bing T, Mei H, Fang C, Cao Z, *et al*. Characterization and application of a DNA aptamer binding to L-tryptophan. Analyst. 2011, 136(3):577-585.

11. Tan SY, Acquah C, Sidhu A, Ongkudon CM, Yon L, *et al*. SELEX Modifications and Bioanalytical Techniques for Aptamer-Target Binding Characterization. Crit Rev Anal Chem. 2016, 46(6):521-537.

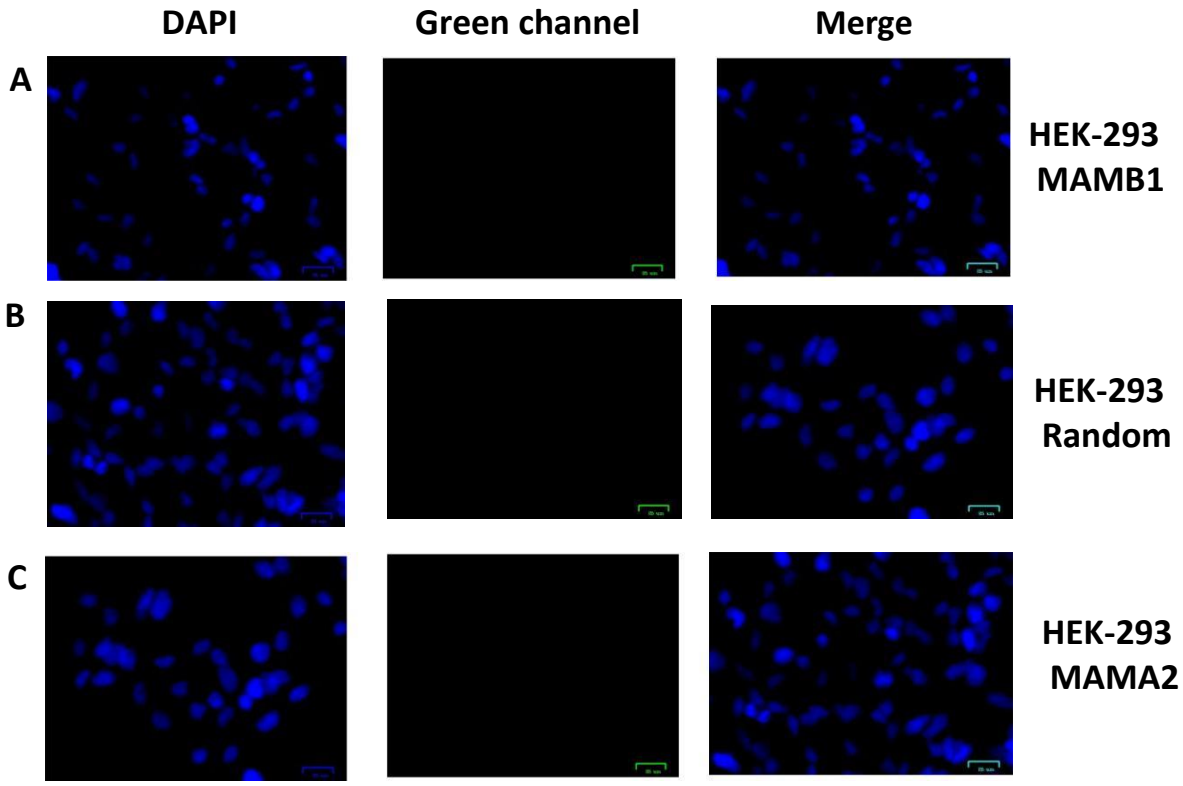
12. McKeague M, McConnell EM, Cruz-Toledo J, Bernard ED, Pach A, Mastronardi E, *et al*. Analysis of In Vitro Aptamer Selection Parameters. J Mol Evol. 2015, 81(5-6):150-161.

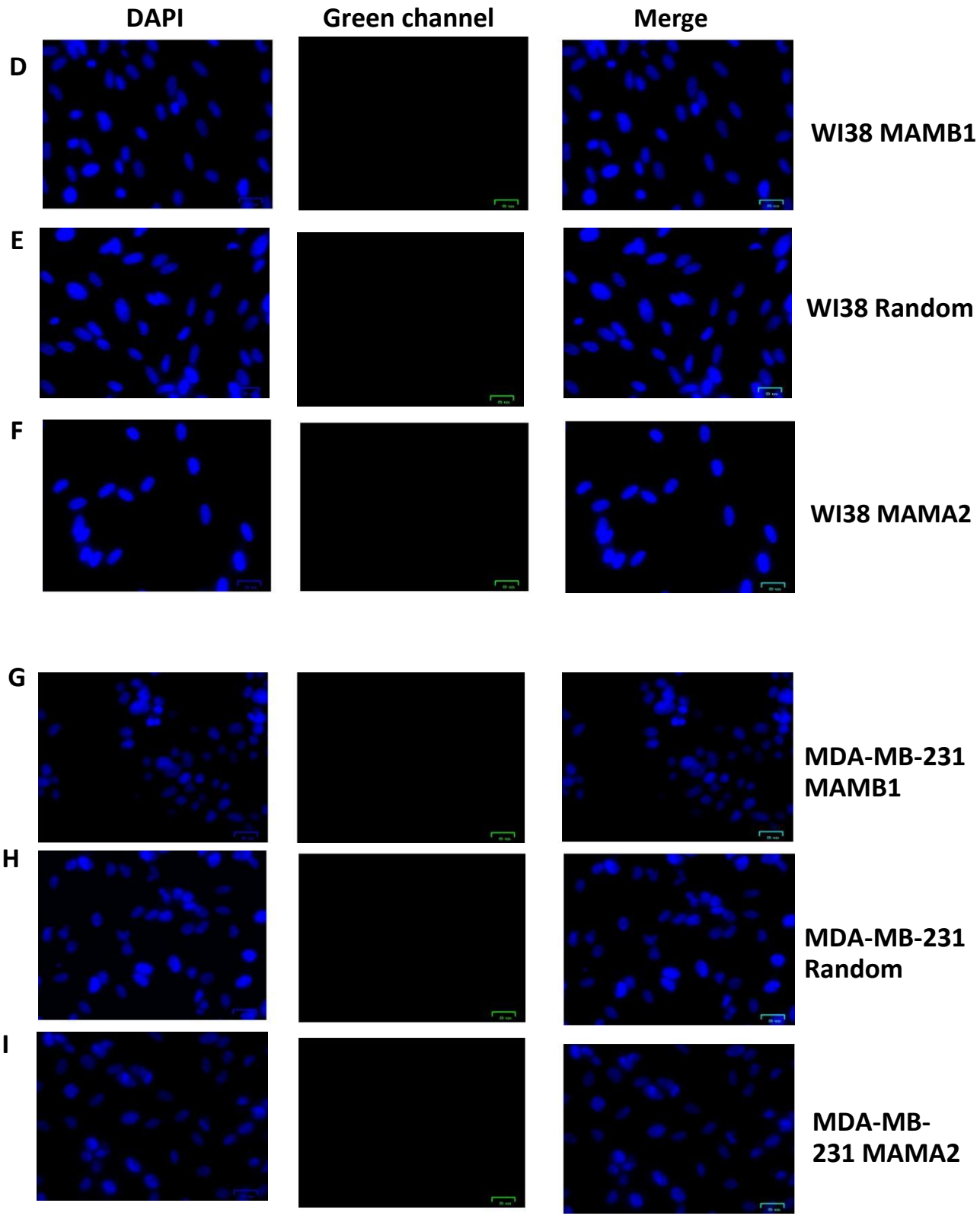
13. Kuwana T, Ogawa M, Sakai K, Kiwa T, and Tsukada K. Label-free detection of low-molecular-weight samples using a terahertz chemical microscope. Applied Physics Express 2016, 9, 042401.

Appendices

Appendix A

A.1 Fluorescence microscopy images of MAMB1 and MAMA2 binding to other cancer and normal cell lines




DAPI
Green channel
Merge
D
WI38 MAMB1
E
WI38 Random
F
WI38 MAMA2
G
MDA-MB-231 MAMB1
H
MDA-MB-231 Random
I
MDA-MB-231 MAMA2

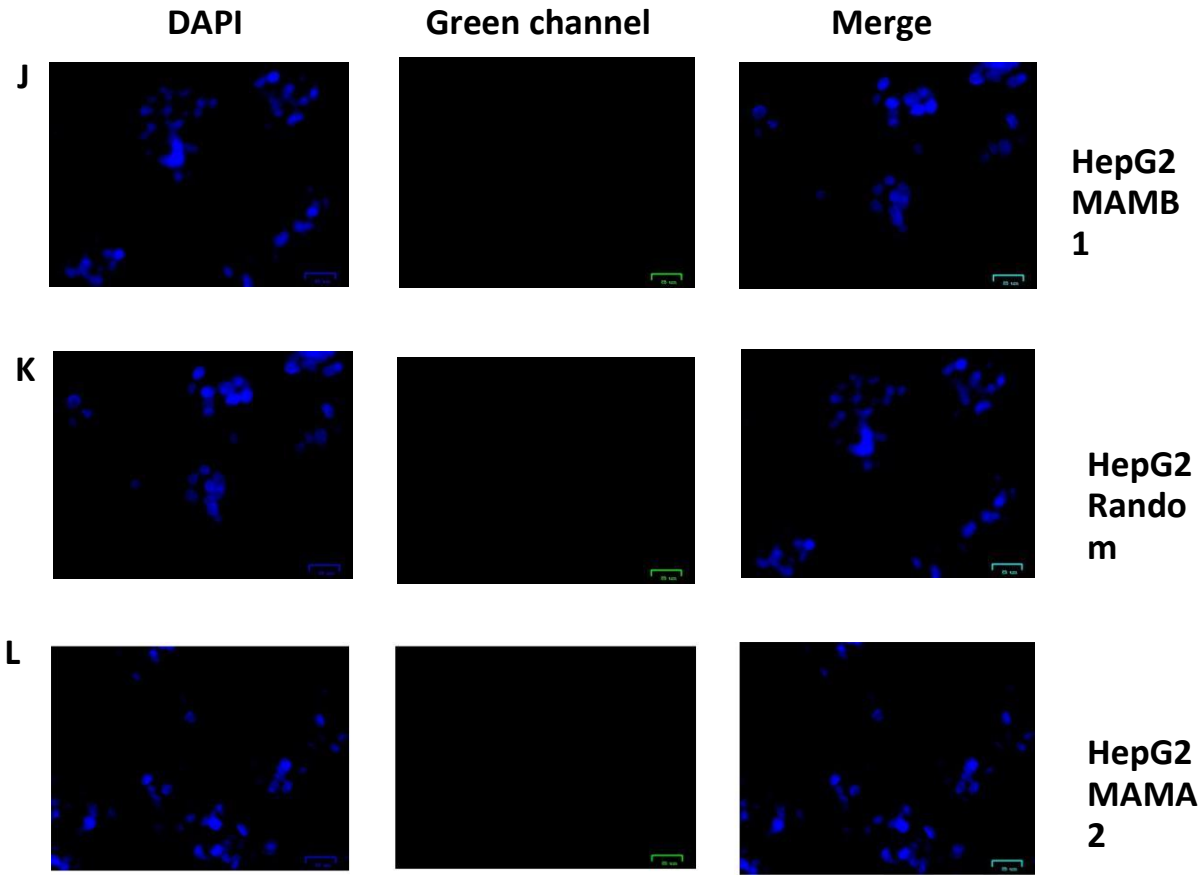


Figure A.1: Fluorescence microscopy images of aptamers MAMB1 and MAMA2 binding to different cancer and normal cell lines. A: MAMB1 binding to HEK293 cells, B: random sequence binding to HEK293 cells, C: MAMA2 binding to HEK293 cells. D: MAMB1 binding to WI38 cells, E: random sequence binding to WI38 cells, F: MAMA2 binding to WI38 cells, G: binding of both MAMB1 to MDA-MB-231 cells, H: random sequence binding to MDA-MB-231 cells, I: binding MAMA2 to MDA-MB-231 cells, J: MAMB1 binding to HepG2 cells, K: random sequence binding to HepG2 cells, L: MAMA2 binding to HepG2 cells. Both aptamers and the random sequence were incubated with the cells for 30 min and then imaged. Scale bars correspond to 25µm in all images.

A.2 Transformation and purification results of MGB2 and MGB1 plasmids

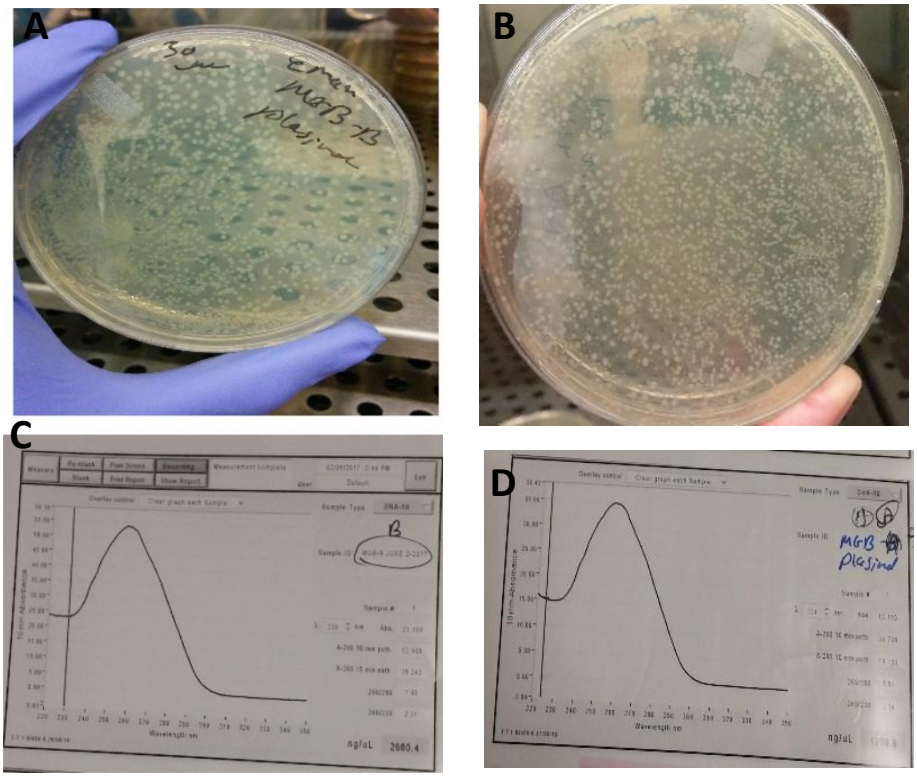


Figure A.2: Transformation and purification results of MGB2 and MGB1 plasmids. A-B: photos of agar plates with growing colonies after transformation of MGB2 (A) and MGB1 (B) plasmids to alpha DH5 bacteria. C-D: nano-drop reports for MGB2 (C) and MGB1 (D) indicating high amount of DNA after using midiprep system for purification of plasmids.

Résumé en français

1. Introduction

Cancer du sein

Le cancer du sein est le cancer le plus fréquemment diagnostiqué chez les femmes canadiennes. Il correspond à 26% de tous les cancers nouvellement diagnostiqués chez les femmes. L'incidence du cancer du sein est plus élevée dans les pays développés avec un taux de 80 pour 100 000 dans l'Union européenne et 92 pour 100 000 en Amérique du Nord. Le cancer du sein est le principal cancer diagnostiqué et la deuxième cause (après le cancer du poumon) de décès par cancer chez les femmes canadiennes. En 2015, environ 25 000 Canadiennes ont été diagnostiquées avec un cancer du sein. L'un des facteurs de risque les plus importants pour le développement du cancer du sein est d'avoir des antécédents familiaux de cancer du sein. De plus, les femmes qui ont été exposées ou traitées par rayonnement dans la zone de la poitrine, du cou et des bras sont plus susceptibles de développer un cancer du sein au cours de leur vie. On a montré que les facteurs de risque de cancer du sein liés au mode de vie augmentaient le risque de cancer du sein, ces facteurs incluent l'alcool, le surpoids ou l'obésité, le tabagisme, le contrôle des naissances et la thérapie de remplacement hormonal. La mammographie a prouvé son efficacité pour réduire le taux de mortalité du cancer du sein. D'autres méthodes telles que l'échographie mammaire et l'imagerie par résonance magnétique mammaire (IRM) ont été utilisées dans le diagnostic du cancer du sein.

Le cancer du sein métastatique (MBC) et les cellules tumorales circulant (CTC)

Le cancer du sein métastatique (MBC) est une étape du cancer du sein où la maladie s'est propagée à différentes parties du corps en plus des ganglions axillaires, telles que les os, le foie, les poumons ou le cerveau. La métastase est la principale cause de décès chez la majorité des patientes atteintes de cancer du sein. Pour qu'il y ait des métastases, les cellules cancéreuses du

sein doivent compléter une cascade d'événements, de l'invasion des tissus locaux au sang et de la croissance à distance du premier site provoquant une tumeur secondaire (figure 1). Au cours des premiers stades, certaines cellules cancéreuses gagnent des propriétés invasives et laissent la tumeur primaire s'insinuer dans la circulation sanguine ou le système lymphatique. Par la suite, les cellules circulent dans le sang à travers le corps, adhèrent à leur site secondaire préféré et extravasent de la circulation dans le tissu secondaire. Les cellules cancéreuses doivent ensuite maintenir la croissance dans le site secondaire pour former des micro-métastases, puis la formation ultérieure de macro-métastases par angiogenèse sur le site secondaire. Les cellules qui se détachent de la tumeur primaire et circulent dans le sang sont appelées CTC.

Les CTC ont la capacité d'initier des métastases en convertissant des cellules cancéreuses épithéliales différenciées en cellules dédifférenciées qui ont plus de caractéristiques mésenchymateuses par transition épithéliale à mésenchymateuse (EMT). Ce qui est considéré comme un événement crucial dans le processus métastatique (figure 1). La transition phénotypique des CTC est réversible et il a été signalé qu'une fois que les cellules CTC atteignent leur destination, elles se retrouveront dans un phénotype épithélial dans un processus appelé transition mésenchymateuse à épithéliale (MET), cette transition est déterminante pour la croissance tumorale dans le secondaire Site (figure 1).

Les CTC sont des événements très rares dans le sang, par conséquent, la détection de ces cellules dans le sang est une tâche difficile. Le système CellSearchTM est le seul instrument approuvé par la FDA (Food and Drug Administration) pour détecter et quantifier les CTC dans le sang humain à partir de carcinomes épithéliaux, Cela inclus le cancer du sein basé sur l'interaction de la molécule d’adhésion cellulaire (EpCAM) à la surface des CTC et de son anticorps monoclonal. EpCAM est considéré comme le biomarqueur de choix pour la détection de CTC dans le sang des patients atteints de cancer du sein. Cependant, l'expression de protéines EpCAM n'est pas la même chez tous les CTC. En outre, certains des CTC perdent l'expression d'EpCAM pendant la circulation dans le sang. Par conséquent, plus de biomarqueurs doivent être identifiés pour aider à détecter les CTC du cancer du sein.

Les protéines Mammaglobin B et mammaglobin A

La mammaglobine B (MGB2) a d'abord été identifiée par Becker et al. En 1998 en tant que protéine sécrétoire et membre d'une famille de gènes d'utéroglobine. MGB2 est une petite protéine composée de 95 acides aminés pour l'ensemble de la protéine et 77 acides aminés pour la protéine mature (les 18 premiers acides aminés est un peptide signal qui est clivé lorsque la protéine mature est sécrétée hors de la cellule), avec une masse de 8854,1 KDa pour la protéine mature et le point isoélectrique 4.94.

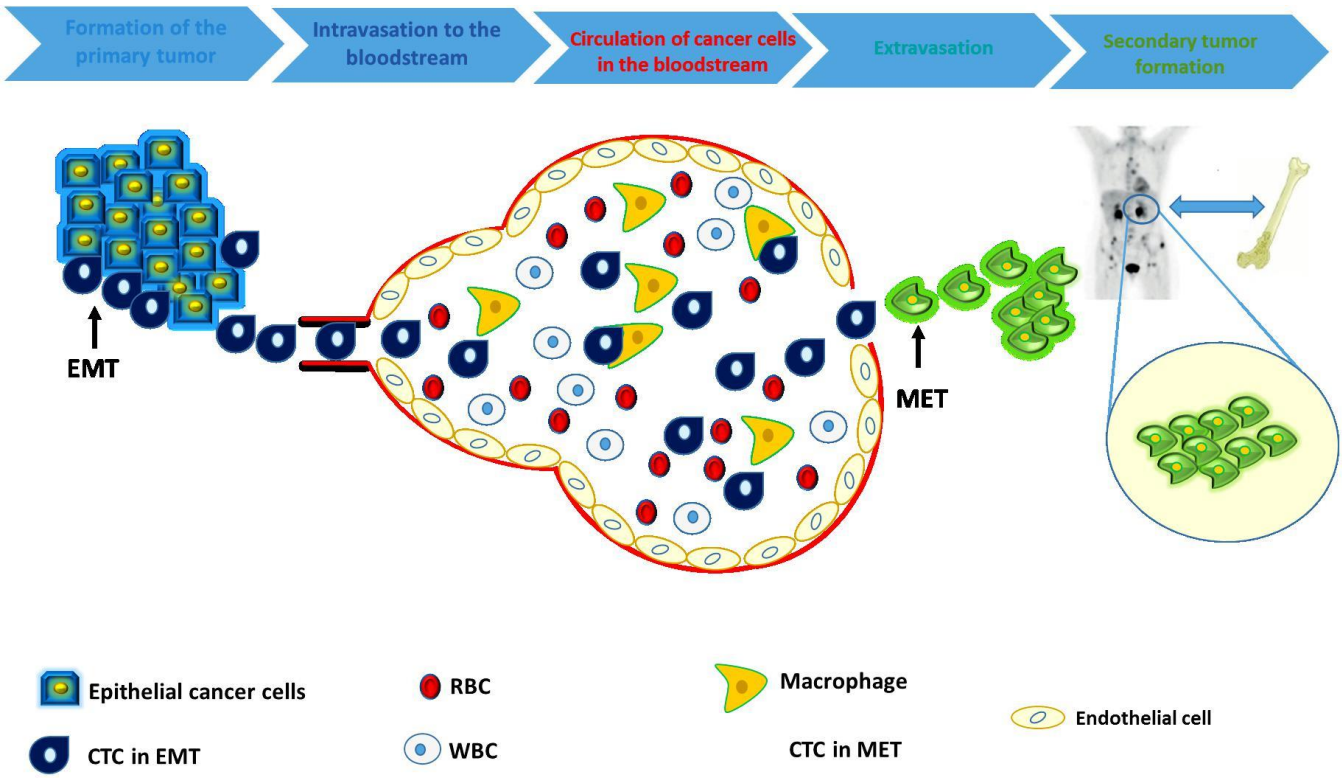


Figure 1 : cascade de métastases, dans la première étape qui inclut la formation de la tumeur primaire, les cellules cancéreuses commencent à envahir un tissu épithélial local et à se diviser, certaines d'entre elles (CSC) ont la capacité d'initier une métastase et de subir une EMT pour être considérées comme CTC, 2 : Les CTC qui ont les propriétés du CSC entrent dans la circulation sanguine et circulent avec les composants sanguins, certains ont été éliminés par des cellules de macrophages (partie du système immunitaire), 3 : Le reste des CTC quitteront le sang dans un processus appelé Extravasation, 4 : après leur sortie du sang, les cellules cancéreuses circulantes subissent un TEM et commencent à envahir un organe éloigné (en cas de cancer du sein est l'os) formant une tumeur secondaire.

L'association Mammaglobin A (MGB1) d'un autre membre de la famille de gènes d'utéroglobine a d'abord été identifiée par Watson et Fleming en 1996. Le MGB1 consiste en 93 acides aminés avec un poids moléculaire de 10,5 kDa. Comme MGB2, les 18 premiers résidus correspondent à un peptide signal qui est clivé lors de la sécrétion de la protéine mature de 8,5 kDa. Le MGB2 est surtout exprimé dans les tissus muqueux et se trouve à des niveaux élevés dans de nombreuses sécrétions, y compris celles des glandes utérines, prostatiques, pulmonaires et lacrymales et salivaires. D'autre part, l'expression MGB1 est exclusive à la glande mammaire adulte. On sait peu de choses au sujet des fonctions de MGB2 et MGB1, mais il a été rapporté qu'ils ont un rôle dans le développement du cancer, la régulation du système immunitaire et les récepteurs de la liaison des androgènes. Les deux sont surexprimés dans la plupart des cancers primaires et métastatiques du sein. Ils ont tous deux été signalés comme biomarqueurs de CTC chez les patients atteints de cancer du sein. Par conséquent, MGB2 et MGB1 sont des biomarqueurs très intéressants pour la détection de cancer du sein et de cancer des seins.

Les aptamères sont des éléments de reconnaissance puissants

Les aptamères sont des oligonucleotides d'ADN ou d'ARN monocaténaire (~ 15-100 nucleotides), qui se lient à leurs cibles avec une affinité et une sélectivité élevées et avec une interaction tridimensionnelle unique. Les aptamères peuvent remplacer les anticorps en raison de leurs propriétés de reconnaissance similaires. Les aptamères sont synthétisés chimiquement sans les problèmes associés à l'utilisation d'animaux pour la génération d'anticorps. En raison de leur stabilité chimique et thermique, les aptamères peuvent être utilisés dans un large éventail de conditions expérimentales. Contrairement aux anticorps, ils sont limités à leurs conditions d'hôte et perdent rapidement leur structure tertiaire à des températures élevées. Les aptamères sont produits par un processus appelé « évolution systématique des ligands par enrichissement exponentie » (SELEX). SELEX sont des cycles d'enrichissante répétitif in vitro à partir d'une large bibliothèque aléatoire de ssDNA ou d'ARN. Chaque cycle du procédé SELEX est effectué en liant

et en éluant à partir de molécules cibles qui aboutissent à un aptamateur avec une grande affinité et une spécificité pour la cible.

Une large gamme de molécules pourrait être la cible dans le processus SELEX pour développer un aptamère contre eux. En partant d'un ion et en finissant par des cellules de mammifères entières. Cell-SELEX est un outil puissant pour identifier de nouveaux biomarqueurs. La procédure Cell-SELEX est semblable à celle générale, mais elle diffère dans l'étape de séparation dans laquelle le lavage ou la centrifugation est utilisé pour séparer les séquences liées des non-liées et le chauffage est généralement utilisé pour l'isolement du complexe de cellules aptamères-cibles (figure 2).

Les cellules ont de nombreux récepteurs à leur surface. Par conséquent, le développement d'aptamères pour des cibles inactives est une chance élevée. Le SELEX hybride est une cellule modifiée SELEX où la banque aléatoire alterne entre la forme purifiée de la protéine et les cellules. Mais dans ce cas, la cible devrait être identifiée avant le début du processus SELEX. Le SELEX hybride conduit à un enrichissement plus élevé des aptamères contre la cible d'intérêt à la surface de la cellule contrairement au SELEX de cellules entières, car la banque d'ADN se transforme en la forme purifiée de la protéine. De plus, les aptamères résultant du SELEX hybride sont plus spécifiques à la lignée cellulaire tumorale par rapport au contrôle qu'un des aptamères développés par des cellules entières comme cible.

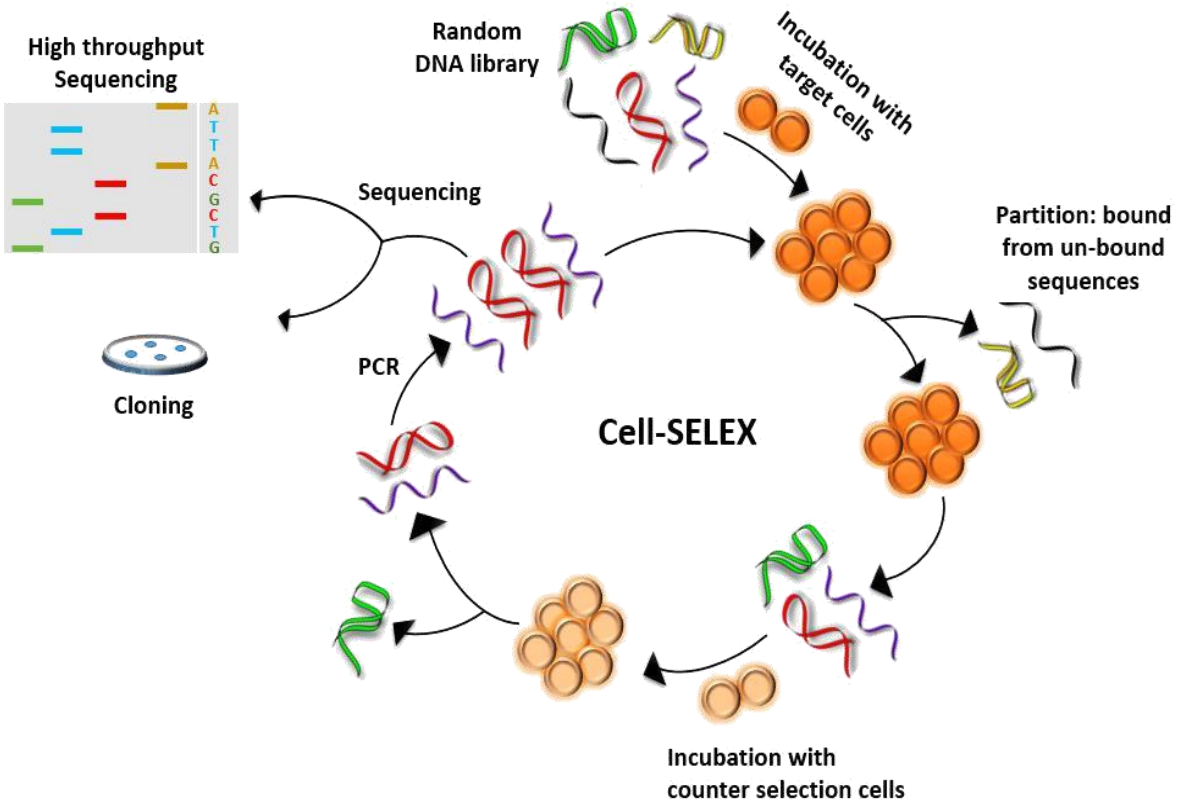


Figure 2 : Cell-SELEX, Cell-SELEX commence par l'incubation d'une certaine cellule cible avec une banque d'ADN aléatoire, cette étape est une sélection positive. Ensuite, les séquences de faible liant et non-liées sont éliminées par lavage et les restes sont élués hors des cellules par ébullition. Pour éliminer toute liaison des séquences à des cellules similaires, une étape de comptage sélectif se déroule pour les séquences éluées. Les séquences non consolidées des cellules comptabilisés sont amplifiées par PCR et le cycle se répète. La bibliothèque finale est séquencée pour identifier les candidats des aptamères soit par clonage traditionnel, soit par séquençage à haut débit. PCR: réaction en chaîne par polymérase.

THz radiation

Le rayonnement Terahertz (THz) également connu sous le nom d'ondes THz se situe entre les régions infrarouges et hyperfréquences du spectre électromagnétique (figure 3). Les THz présentent des caractéristiques uniques qui lui permettent d'être un domaine potentiel de recherche en physique, chimie, biologie, sciences des matériaux et médecine. Puisque le

rayonnement THz a une longue longueur d'onde (300 µm), il peut pénétrer une grande variété de matériaux non conducteurs, par exemple Les vêtements, le papier, le carton, le bois et le plastique. Pour cette raison, le rayonnement THz est utilisé à des fins de sécurité dans les aéroports et pour la détection d'armes. Ce type de rayonnement THz appelé passif ou incohérent. Le rayonnement de THz a une énergie de photon relativement faible qui est insuffisante pour causer des dommages chimiques aux tissus et à l'ADN. Certaines fréquences du rayonnement de THz peuvent pénétrer plusieurs millimètres de tissu à faible teneur en eau (par exemple, tissu adipeux). THz peut être utilisé pour différencier différents matériaux ou même des isomères puisque l'énergie des transitions rotatives et vibratoires des molécules se situe dans la région de THz et les vibrations intermoléculaires telles que les liaisons hydrogène présentent différentes caractéristiques spectrales dans la région de THz. Les ondes de THz peuvent offrir un meilleur contraste pour les tissus mous que les rayons X en raison du fait que le rayonnement THz est très sensible aux substances polaires, comme l'eau et l'état d'hydratation. Le type de THz utilisé dans l'imagerie s'appelle actif ou cohérent.

La microscopie chimique Terahertz comme outil de détection du cancer du sein

La microscopie chimique Terahertz (TCM) a une résolution spatiale plus élevée que les systèmes d'imagerie THz standard car la résolution spatiale n'est pas déterminée par la longueur d'onde de l'impulsion THz mais par celle du laser femtoseconde, qui est généralement d'environ 790 nm. TCM se compose d'une plaque de détection de films minces de dioxyde de silicium (SiO_2) et de silicium (Si) sur un substrat en saphir (figure 4). Lorsqu'une impulsion laser femtoseconde frappe la plaque de détection, elle accélère les porteurs de charge libres de Si et ces porteurs commencent à se déplacer entre Si et SiO_2 (la couche de déplétion). Par conséquent, le potentiel chimique à la surface de la plaque de détection changera, ce qui entraînera une variation de l'amplitude de l'impulsion THz générée. Le TCM peut être utilisé dans des analyses immunitaires sans étiquettes. De cette manière, le TCM mesure le changement de potentiel chimique sur la surface de la plaque de détection en raison de la liaison entre l'antigène et l'anticorps. L'intensité du THz augmente lors de l'interaction de l'anticorps et de son antigène. Par conséquent, le TCM pourrait être utilisé pour mesurer la liaison entre un ligand et sa cible.

À cet égard, les aptamères sont potentiellement utilisables avec la technologie TCM comme une sonde au lieu d'anticorps pour tester l'interaction entre les aptamères et leurs cibles. De plus, les aptamères pourraient être testés pour la liaison aux cellules cancéreuses à la surface des plaques de détection. Le TCM pourrait donc être utilisée dans la détection de cellules cancéreuses et ceci est réalisé dans le cadre de ce projet de recherche pour la première fois.

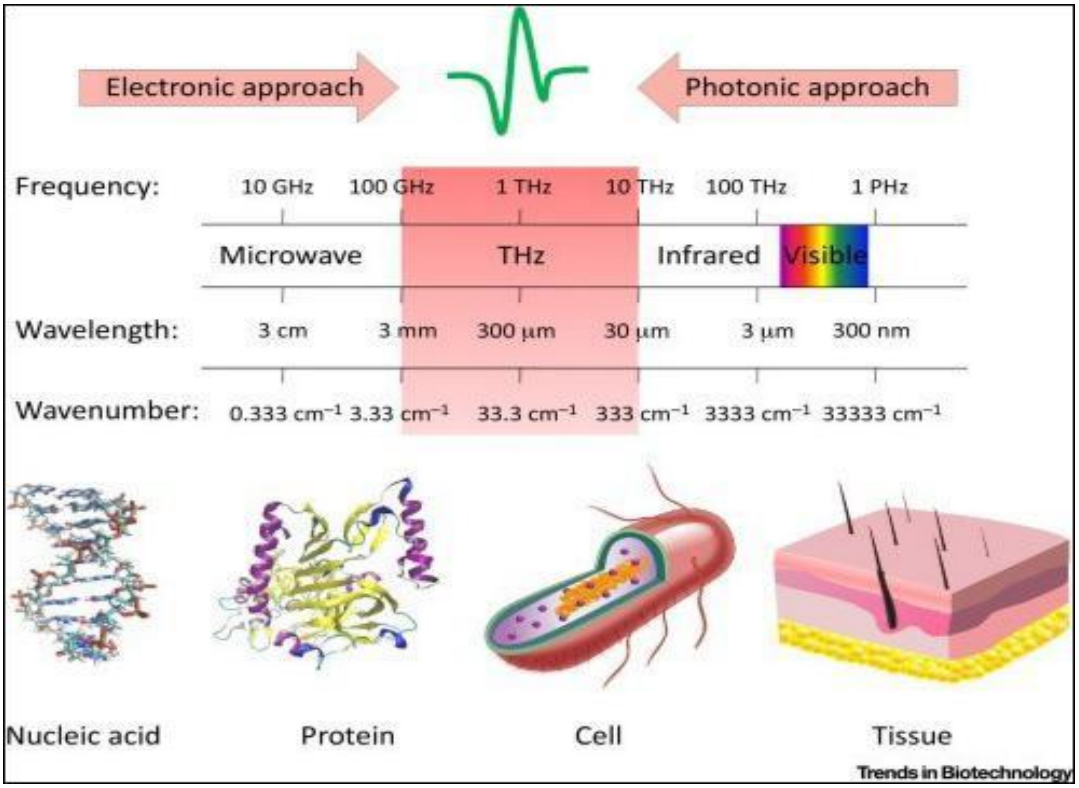


Figure 3: spectre THz et les applications potentielles du rayonnement THz. Le rayonnement THz se situe entre le micro-ondes et les régions infrarouges du spectre électromagnétique. Les fréquences THz s'étalent de 100 GHz à 10 THz sur le spectre électromagnétique, ce qui équivaut une longueur d'onde de 30 mm à 3000 mm ou à un nombre d'onde de 3,33 cm^{-1} à 333 cm^{-1}. Le THz a plusieurs applications incluant l'imagerie sur les tissus biologiques, les protéines et les acides nucléiques.

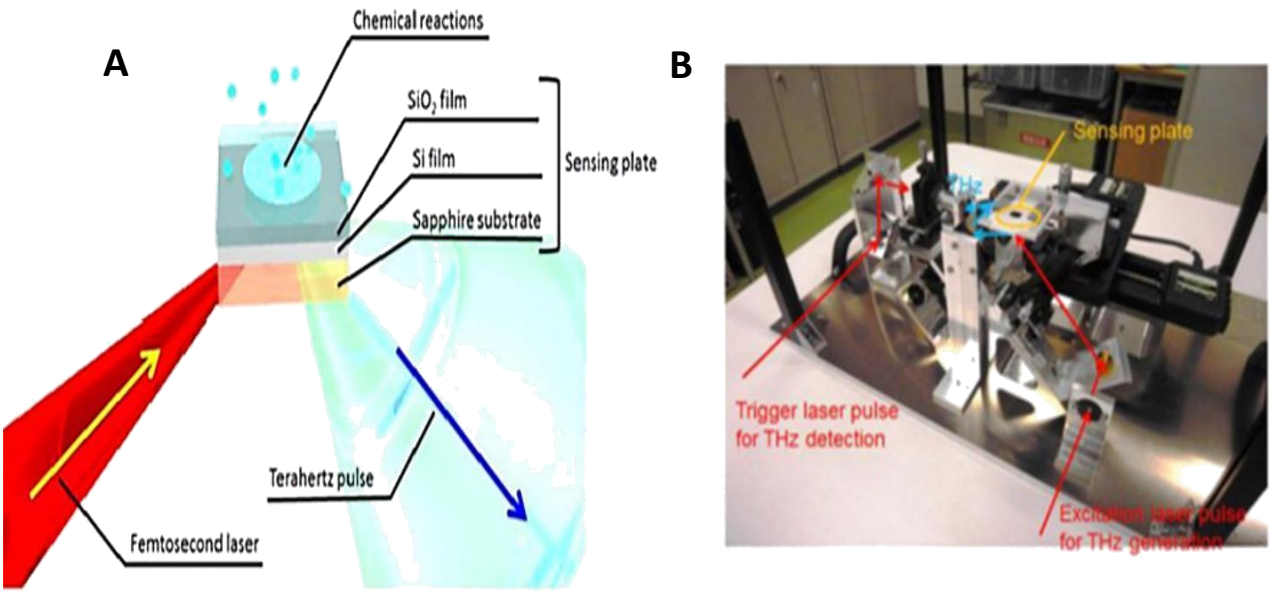


Figure 4 : A, Diagramme schématique de la plaque de détection du microscope chimique terahertz (THz). Un laser femtoseconde se concentre sur le film mince en silicium (Si) du côté substrat de la plaque et une impulsion THz rayonne dans l'espace libre. B, photographie du prototype du système TCM.

Objectifs de thèse

1 Sélection de nouveaux aptamères contre les biomarqueurs mammaires (mammaglobine B et mammaglobine A)

2 Caractérisation des candidats aptamères et détermination de leurs spécificités pour les meilleurs

3 Test du potentiel des aptamères ayant une affinité la plus élevée pour détecter les cellules cancéreuses dans les échantillons de sang et de plasma et dans le mélange de cellules (cellules cancéreuses à pointes)

4 Vérifier les cibles des aptamères à la surface de la cellule et en solution

5 Test de la liaison des aptamères de haute affinité à leurs lignées cellulaires à l'aide de la technologie TCM.

2. Conception expérimentale

SELEX hybride et séquençage à haut débit (HTP)

Dans cette étude, les aptamères MGB2 et MGB1 ont été développés en utilisant l'approche hybride SELEX. La banque de ssDNA marquée en fluorescéine était altérée entre les formes recombinantes des deux cibles et des cellules MCF7 comme cible pour MGB2 et MDA-MB-415 comme cible pour MGB1. Les lignées cellulaires MCF10A et HCAEC ont été utilisées comme lignes cellulaires de contre sélection (Figure 5).

Après un total de 30 cycles pour MGB2 et 13 cycles pour MGB1, plusieurs rondes SELEX ont été envoyées pour un séquençage à haut débit pour identifier les séquences de tous les aptamères dans ces rondes. Une analyse de bioinformatique a été effectuée pour sélectionner les candidats aptamères.

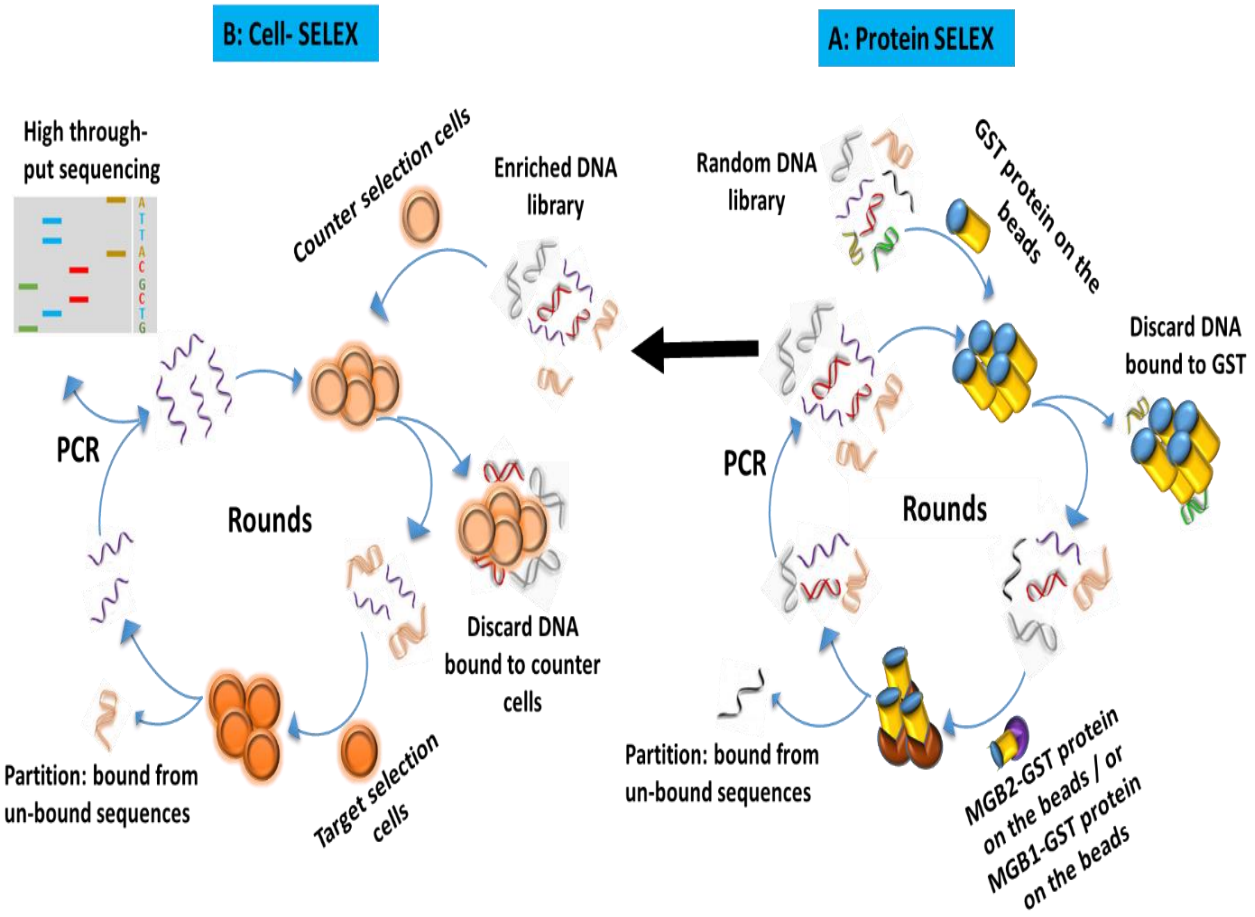


Figure 5 : diagramme schématique de la méthode SELEX hybride pour la sélection des aptamères d'ADN spécifiques à MGB2 et MGB1. Le SELEX croisé dans notre expérience a été divisé en A, une protéine SELEX et B, une SELEX à base de cellules. D'abord, une GST-MGB2, des protéines GST-MGB1 ont été utilisées pour sélectionner contre une bibliothèque de ssDNA aléatoire. Après 21 cycles de sélection, le pool enrichi des deux protéines a été transféré pour être utilisé dans la cellule SELEX dans laquelle HCAEC et MCF10A ont été utilisés comme lignes de sélection de contre-sélection. MCF7 a été utilisé comme ligne de sélection positive pour MGB2 et MDA-MB-415 a été utilisé comme ligne de sélection positive pour MGB1. Après 7 cycles de sélection, les pools finaux pour les deux cibles ont été transférés à la protéine pendant encore 2 tours supplémentaires pour rendre la bibliothèque plus spécifique au MGB2 et au MGB1.

Détermination de l'affinité et de la sélectivité des aptamères sélectionnés

Les aptamères sélectionnés ont été testés comme suit: Tout d'abord, l'affinité de tous les aptamères sélectionnés a été testée contre leurs lignes cellulaires positives et opposées. Deuxièmement, les aptamères qui ont montré une affinité élevée à leurs lignées cellulaires cibles comparées aux opposés ont été testés pour déterminer leur affinité (constante de dissociation (Kd)) à MCF7 dans le cas de

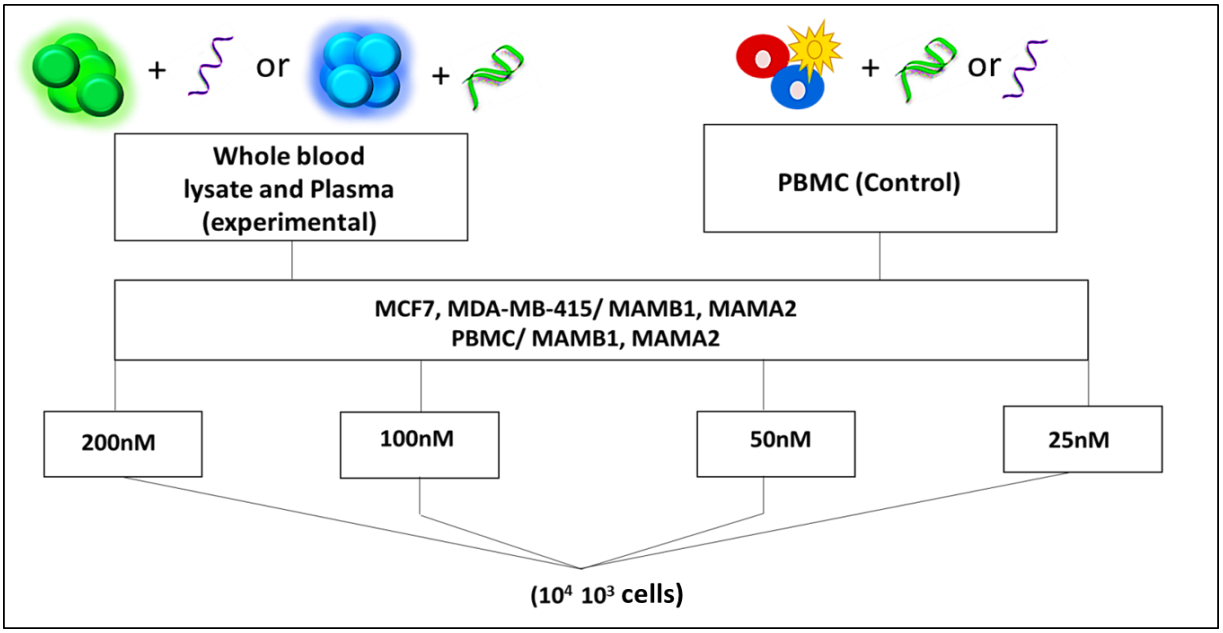
MGB2 et MDA-MB-415 dans le cas de MGB1. Troisièmement, les aptamères qui présentaient une affinité maximale (valeurs kd les plus faibles) ont été testés pour

leur sélectivité en testant leur affinité à d'autres cellules normales et cancéreuses. La cytométrie de flux a été utilisée comme outil d'analyse dans tous les tests effectués ci-dessus, où les cellules ont été fixées dans du formaldéhyde et incubées avec chaque aptamère sélectionné (marqué par fluorescéine) dans un tampon de liaison (PBS et MgCl2 5 mM) pendant 30 minutes à 4 ° C. Ensuite, les cellules-aptamères étaient prêts à être analysées. Dans le cas de la détermination de kd, les cellules ont été incubées avec différentes concentrations de l'aptamère (0 nM-200 nM).

Des tests supplémentaires dans le lysat de plasma et du sang total

Les aptamères qui présentaient une affinité et une sélectivité élevées contre leurs lignées cellulaires cibles ont encore été testés pour leur liaison dans le plasma et le lysat sanguin entier. Différentes concentrations des aptamères candidats pour les deux ont été utilisées et un nombre différent de cellules. Le schéma 1 explique la conception de cette partie.

A


+
or
+
+
or
Whole blood
lysate and Plasma
(experimental)
PBMC (Control)
MCF7, MDA-MB-415/ MAMB1, MAMA2
PBMC/ MAMB1, MAMA2
200nM
100nM
50nM
25nM
(10⁴ 10³ cells)

B

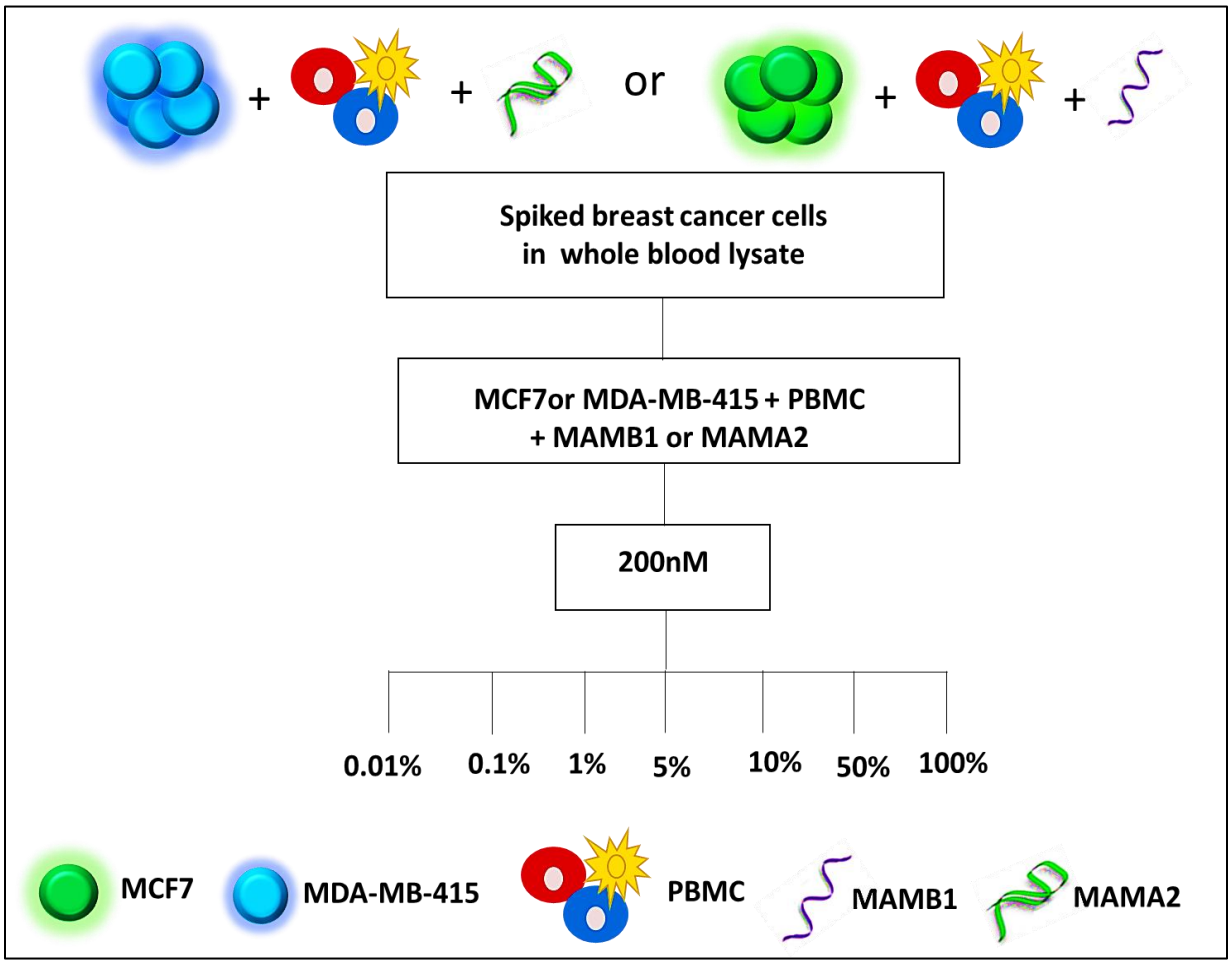

+
+
or
+
+
Spiked breast cancer cells
in whole blood lysate
MCF7or MDA-MB-415 + PBMC
+ MAMB1 or MAMA2
200nM
0.01%
0.1%
1%
5%
10%
50%
100%
MCF7
MDA-MB-415
PBMC
MAMB1
MAMA2

Schéma 1. Diagramme schématique de la conception expérimentale des études plasmatiques et sanguines. A. Lysat sanguin entier et conception expérimentale du plasma: MCF7 et MDA-MB-415 ont été utilisés comme cellules témoins de cancer du sein, les cellules PBMC ont été utilisées comme témoins (non cancéreuses), les cellules cibles ont été incubées avec leurs aptamères (MCF7 avec MAMB1 Et MDA-MB-415 avec MAMA2). Les cellules PBMC ont été incubées avec les mêmes aptamères. Les densités cellulaires 1X104 et 1X103 ont été utilisées pour chaque cellule cible avec leur aptamère ainsi que des cellules PMBC. Ensuite, ils ont été analysés par cytométrie en flux. Des concentrations différentes des aptamères ont été utilisées. 200 nM, 100 nM et 50 nM ont été utilisés dans des expériences de lysat de sang total. 200 nM, 50 nM et 25 nM ont été utilisés dans les expériences plasmatiques. B. Expériences d'épissage du lysat sanguin entier: les cellules de cancer du sein MCF7 ont été incubées à différents pourcentages avec des cellules de PBMC dans le sang à un nombre total de cellules 1X104 à 200 nM de MAMB1. La même procédure a été appliquée aux cellules MDA-MB-415 avec leur aptamère (MAMA2). Ensuite, les deux ont été soumis à une analyse de cytométrie en flux.

Aptamer se liant aux cellules cancéreuses du sein en utilisant TCM

Les meilleurs aptamers d'affinité (MAMB1, MAMA2) ont été testés pour se lier à leurs cellules cancéreuses cibles en utilisant la technologie TCM. En bref, les aptamères ont été immobilisés sur la surface de la puce de plaquette après modification, comme indiqué dans le schéma 2. Ensuite, la numérisation TCM a été effectuée pour les deux aptamères et des images différentielles ont été créées.

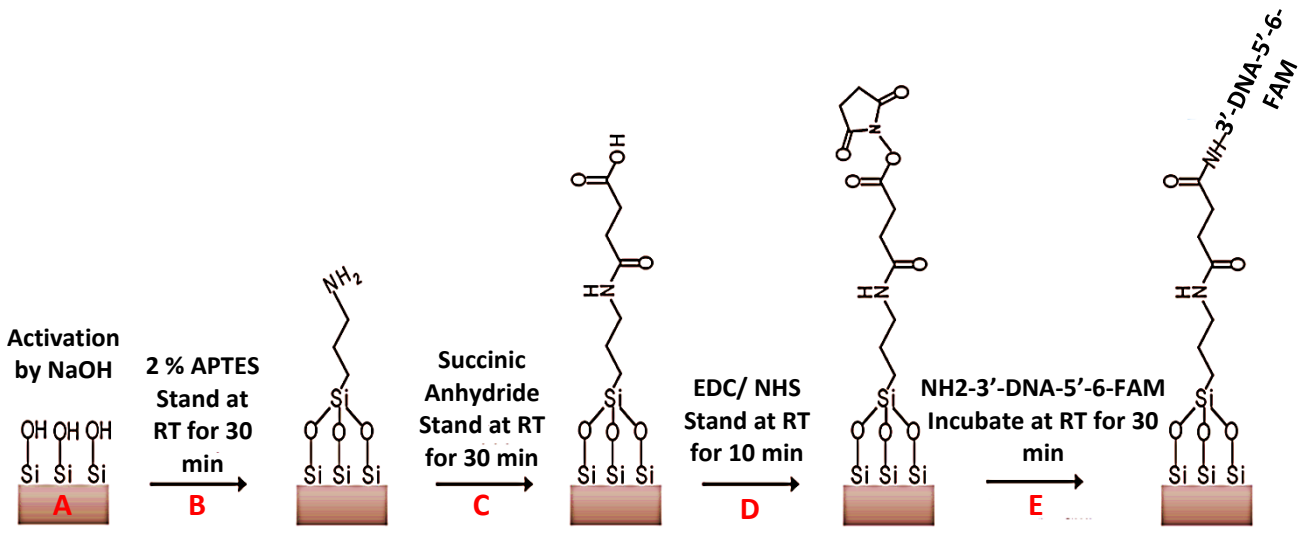


Schéma 2: Procédé d'immobilisation de la silanisation et des aptamères (MAMB1 et MAMA2) sur la surface de la puce.

A: l'hydroxylation de la surface a lieu en incubant la plaquette avec du NaOH, puis on ajoute de l'agent APTES amino-silane pour créer une surface à terminaison amino (B). La formation de liaison amide s'effectue par addition de SA suivie d'une liaison croisée EDC / NHS pour l'activation du groupe carboxylique afin de mesurer la formation de la liaison covalente entre l'amine sur l'extrémité 3 'de l'ADN et le groupe carboxylique sur la surface de la plaquette (C, D, E).

3. Résultats

SELEX hybride et séquençage à haut débit

Les bibliothèques ssDNA ont été surveillées pour s'assurer que la liaison aux cibles (protéines et cellules) augmente, l'enrichissement de l'aptamère augmente également, comparé aux contrôles. Le SELEX se termine lorsque l'enrichissement des bibliothèques d'ADN atteint le maximum. Après avoir envoyé les bibliothèques à HTP MGB2 et les aptamères MGB1 ont été sélectionnés en fonction de leur enrichissement. Douze séquences ont été sélectionnées pour MGB2 et huit pour MGB1 (tableau 1).

Nom / grappe	Séquence aptamère (5'-3 ') 5 '20 bases avant amorce (N 25 ou 24) 20 bases amorce inverse '3	Enrichissement de Positif / Ccontraire	
MAMB0	CATGCTTACCTATAGTGAACCCCAACACACGTGTAGATCCTGCGGCTTTGAGAACTGACTCATAC	241.990	175.373
MAMB1	CATGCTTACCTATAGTGAACCCAACGTCGAACTGAATCCCGTGTCCTTTGAGAACTGACTCATAC	122.649	123.569
MAMB4	CATGCTTACCTATAGTGAACGCGGCATGTTGGCATCTTGGTCCTGCTTTGAGAACTGACTCATAC	11.636	11.079
MAMB6	CATGCTTACCTATAGTGAACCCACACGATATGGCGCTACACCGTGCTTTGAGAACTGACTCATAC	143.579	96.389
MAMB8	CATGCTTACCTATAGTGAACCACGACACGCGCGATCGTCTCACTGCTTTGAGAACTGACTCATAC	60.063	51.133
MAMB12	CATGCTTACCTATAGTGAACCCACCACACAGCGGATACACCATGGCTTTGAGAACTGACTCATAC	920.639	680.997
MAMB19	CATGCTTACCTATAGTGAACCCACCACGATCGATCGTACCACGTCCTTTGAGAACTGACTCATAC	157.224	121.557
MAMB21	CATGCTTACCTATAGTGAACCGTGGCGTGATGGACGTGGGGATGGCTTTGAGAACTGACTCATAC	68.241	40.644
MAMB29	CATGCTTACCTATAGTGAACCCCCAACGACACCGGATTGCCCTTGCTTTGAGAACTGACTCATAC	100.34	83.660
MAMB50	CATGCTTACCTATAGTGAACCCCGTAGCGATGACCGATCGATGTGCTTTGAGAACTGACTCATAC	162.864	125.615
MAMB57	CATGCTTACCTATAGTGAACCCGAAGAGGATGTGCGGTCCCATTGCTTTGAGAACTGACTCATAC	80.200	61.521
MAMB198	CATGCTTACCTATAGTGAACCCGGGGGATGTGGACAGAACGTGCGCTTTGAGAACTGACTCATAC	305.155	182.854
MAMA1	CATGCTTACCTATAGTGAACGCGGCATGTTGGCATCTTGGTCCTCTTTGAGAACTGACTCATAC	25.80	15.93
MAMA2	CATGCTTACCTATAGTGAACCCGGGACAGAACGTGCGCTTTGAGCTTTGAGAACTGACTCATAC	5836.12	5277.48
MAMA4	CATGCTTACCTATAGTGAACGCGGCATGTTGGCCATCTTGGTCCCTTTGAGAACTGACTCATAC	1833.64	690.30
MAMA5	CATGCTTACCTATAGTGAACGGTTGGCATCTTGGTCCTGCTTTGCTTTGAGAACTGACTCATAC	1666.05	236.97

MAMA6	CATGCTTACCTATAGTGAACTGTTGGCATCTTGGTCCTGCTTTGCTTTGAGAACTGACTCATAC	1242.12	191.94
MAMA7	CATGCTTACCTATAGTGAACGCGCATCTTGGTCCTGCTTTGAGACTTTGAGAACTGACTCATAC	1084.41	180.49
MAMA12	CATGCTTACCTATAGTGAACGTTGGCATCTTGGTCCTGCTTTGACTTTGAGAACTGACTCATAC	1409.74	242.50
MAMA13	CATGCTTACCTATAGTGAACGCGGCATGCATCTTGGTCCTGCTTCTTTGAGAACTGACTCATAC	1330.87	286.83

Détermination de l'affinité et de la sélectivité des aptamères sélectionnés

Les résultats du dépistage ont montré des résultats prometteurs avec une plus grande affinité de liaison pour les lignées cellulaires de sélection positive que les contre-sélection pour les deux cibles. Dans le cas de MGB2, six aptamères sur douze ont montré des résultats prometteurs (MAMB0, MAMB1, MAMB4, MAMB8, MAMB12 et MAMB57) et quatre sur huit pour MGB1 (MAMA2, MAMA5, MAMA6 et MAMA12 (figures 6A et 6B). Le tableau 2 présente les valeurs de kd pour ces aptamères. MAMB1, MAMB12, MAMA2 et MAMA2 ont été testés pour leur sélectivité. Les résultats ont montré que tous ces aptamères ne présentent pas de réelles affinité avec des cellules cancéreuses normal ou d'autre type (figures 7A et 7B).

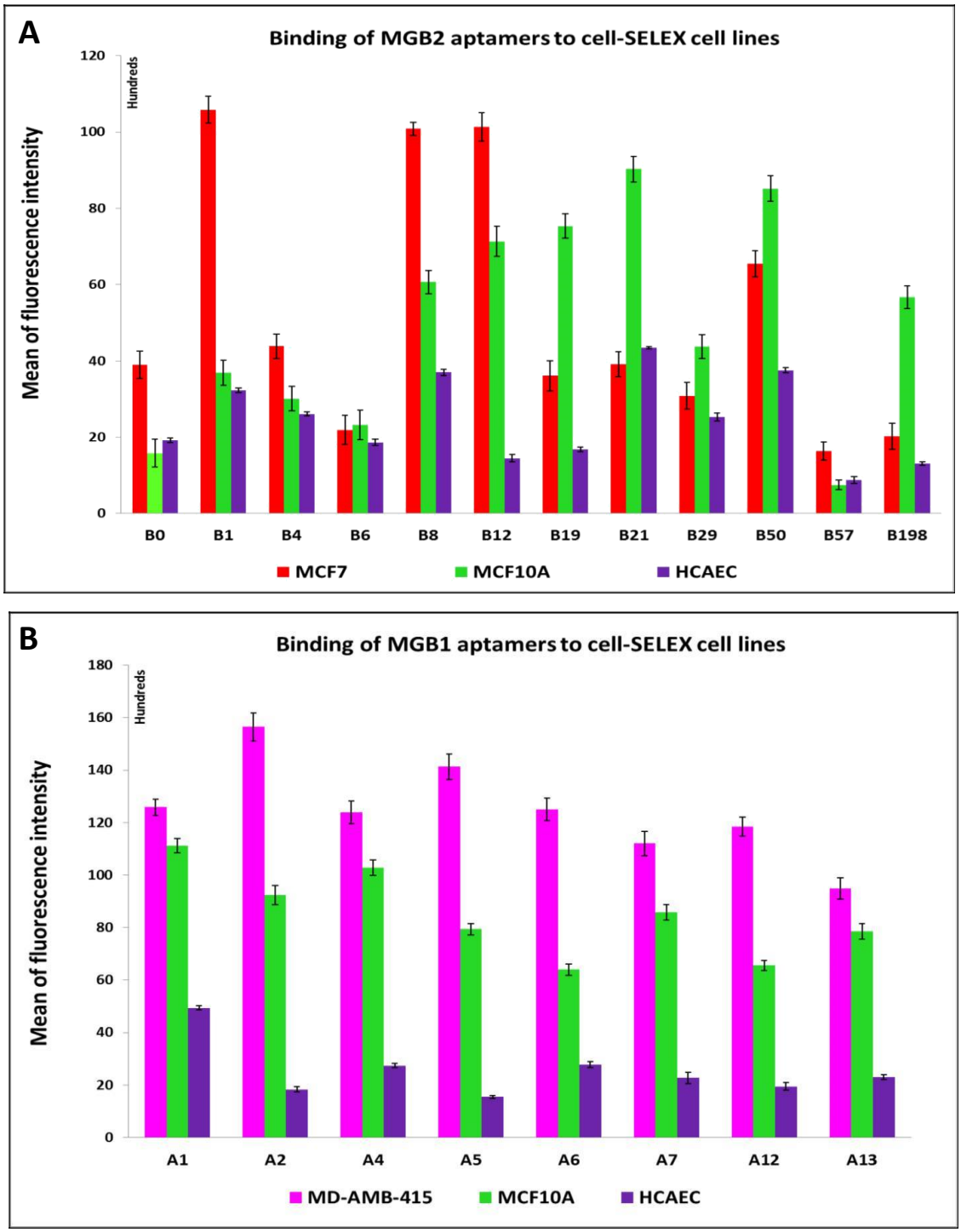


Figure 6: Résultats de dépistage des candidats aptamères en focntion des lignées cellulaires SELEX. A : Les résultats de dépistage des aptamères MGB2 montrent six aptamères sur douze (MAMB0, MAMB1, MAMB4, MAMB8, MAMB12, MAMB57) avec une intensité de fluorescence plus élevée dans les lignées cellulaires de sélection positive que celles de sélection opposée. B : Les résultats de dépistage des aptamères MGB1 montrent une liaison forte non spécifique à la lignée cellulaire MCF10A. Cependant, certains aptamères montrent moins de liaison non spécifique (MAMA2, MAMA5, MAMA6 et MAMA12). Des aptamères prometteurs ont été choisis pour d'autres études. Les expériences ont été répétées deux fois. Les barres d'erreur représentent le S.E.M de deux essais.

Tableau 2: séquences aptamères des cibles MGB2 et MGB1 et leurs valeurs kd correspondantes. A: les sites d'amorces sont soulignés

Aptamer	Séquence aptamer (5'-3 ')	K_d (nM)
MAMB0	CATGCTTACCTATAGTGAACCCCAACACACGTGTAGATCCTGCGGCTTTGAGAACTGACTCATAC	80.0 ± 14.65
MAMB1	CATGCTTACCTATAGTGAACCCAACGTCGAACTGAATCCCGTGTCCTTTGAGAACTGACTCATAC	13.74 ± 3.74
MAMB4	CATGCTTACCTATAGTGAACGCGGCATGTTGGCATCTTGGTCCTGCTTTGAGAACTGACTCATAC	105.23 ± 20.32
MAMB8	CATGCTTACCTATAGTGAACCACGACACGCGCGATCGTCTCACTGCTTTGAGAACTGACTCATAC	78.0 ± 15.35
MAMB12	CATGCTTACCTATAGTGAACCCACCACACAGCGGATACACCATGGCTTTGAGAACTGACTCATAC	19.93 ± 5.80
MAMB57	CATGCTTACCTATAGTGAACCCGAAGAGGATGTGCGGTCCCATTGCTTTGAGAACTGACTCATAC	72.0 ± 9.00
MAMA2	CATGCTTACCTATAGTGAACCCGGGACAGAACGTGCGCTTTGAGCTTTGAGAACTGACTCATAC	3.17 ± 1.41
MAMA5	CATGCTTACCTATAGTGAACGGTTGGCATCTTGGTCCTGCTTTGCTTTGAGAACTGACTCATAC	54.90 ±13.57
MAMA6	CATGCTTACCTATAGTGAACTGTTGGCATCTTGGTCCTGCTTTGCTTTGAGAACTGACTCATAC	49.68 ± 20.36
MAMA12	CATGCTTACCTATAGTGAACGTTGGCATCTTGGTCCTGCTTTGACTTTGAGAACTGACTCATAC	8.23 ± 3.27

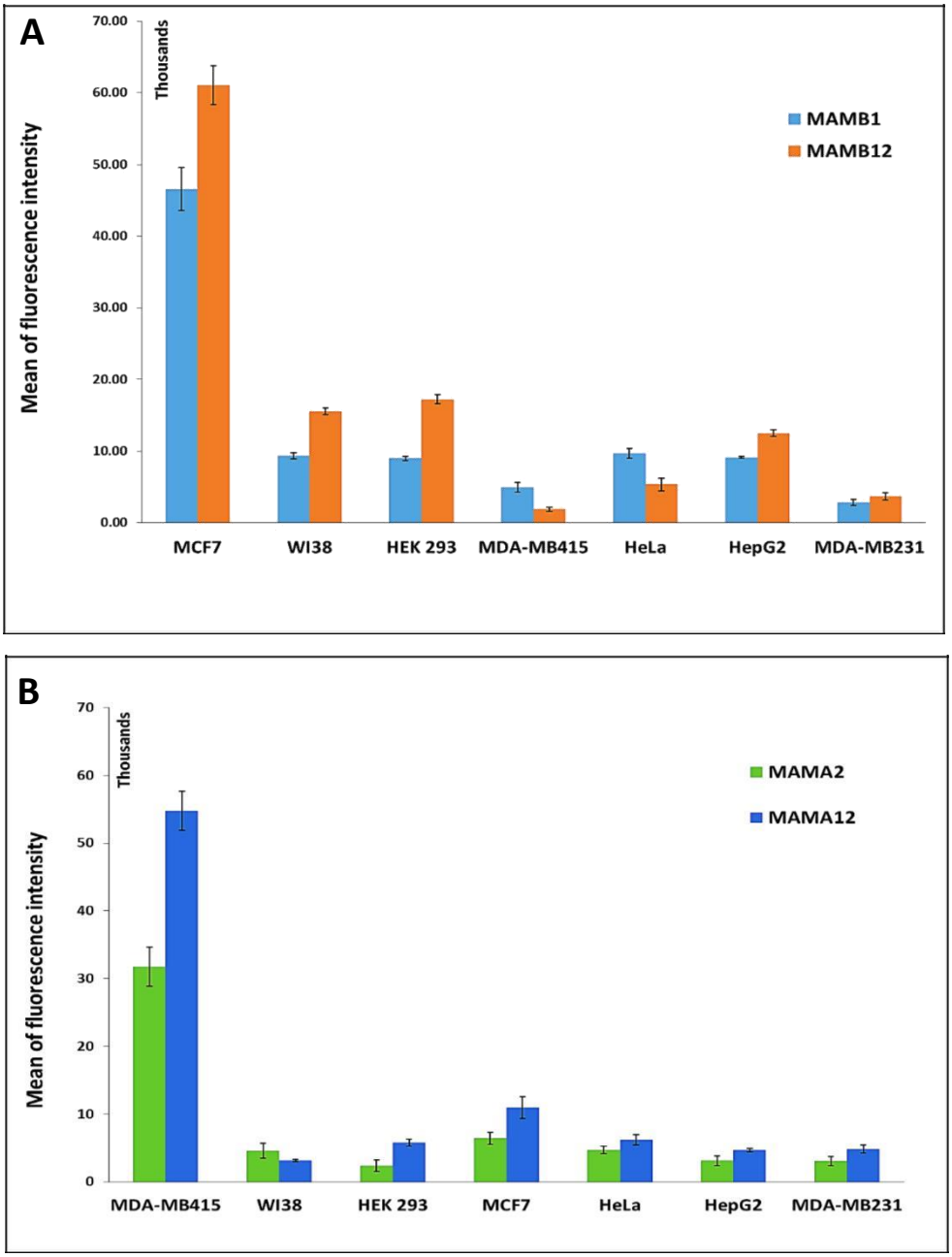


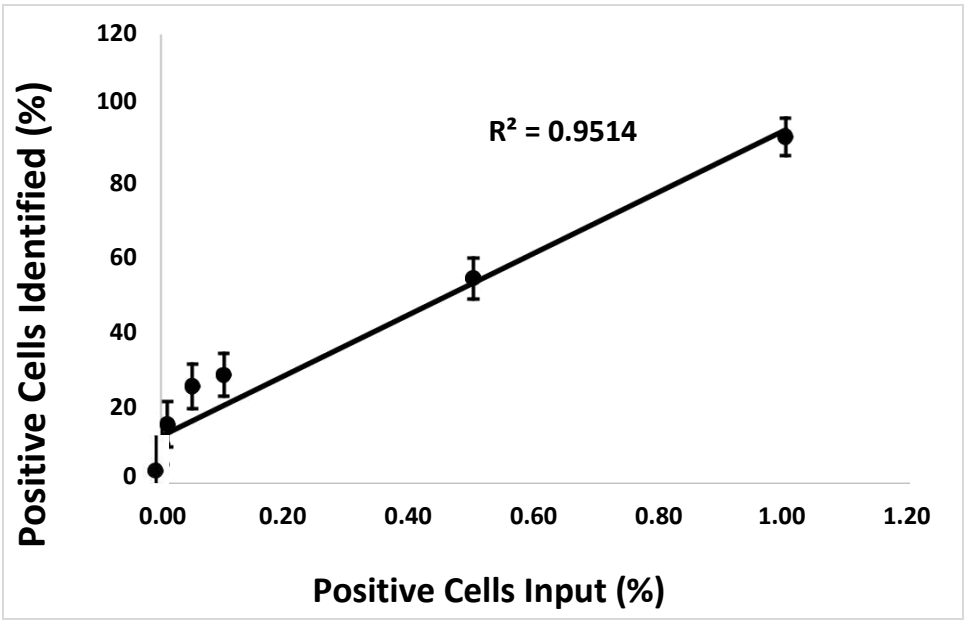
Figure 7: Spécificité des aptamères MGB2 (A) et MGB1 (B) sur différentes lignes de cellules. Les lignes cellulaires utilisées sont MCF7 et MDA-MB-415, en tant que lignes cellulaires cibles pour MGB2 et MGB1 respectivement. WI38 est une cellule de fibroblastes de poumon normale, HEK 293 est une cellule de rein embryonnaire humain transformée, HeLa est une cellule de cancer du col de l'utérus, HepG2 est une cellule de cancer du foie et le MDA-MB-231 est utilisé pour tester la liaison des aptamères MGB2 et MGB1. Les barres d'erreur représentent le S.E.M. De trois essais.

Résultats du lysat de plasma et du sang total

La capacité de MAMB1 et MAMA2 à se lier à leurs cellules de cancer du sein cibles dans le plasma et le lysat de sang complet a été étudiée. Les résultats globaux ont montré que les deux aptamères se lient spécifiquement à leurs cellules cancéreuses cibles par rapport aux cellules témoins à différentes concentrations des aptamères et de différentes densités cellulaires (cellules 1X104 et 1X103). Cependant, MAMB1 a tendance à se lier plus spécifiquement que MAMA2 comme le montrent les valeurs de l'intensité de fluorescence pour les deux aptamères.

Les deux aptamères ont été testés pour leur liaison dans le lysat de sang complet contenant des cellules sanguines normales et des cellules de cancer du sein à pointes en différents pourcentages (0%, 0.01%, 0.1%, 1%, 5%, 10%, 50% et 100%). . Les résultats ont montré que les deux aptamères pouvaient reconnaître leurs cellules cibles à partir d'un mélange de cellules dans le lysat. Le pourcentage de cellules positives identifiées a augmenté avec l'augmentation des pourcentages de cellules tamponnées pour les deux aptamères (figure 8).

A



B

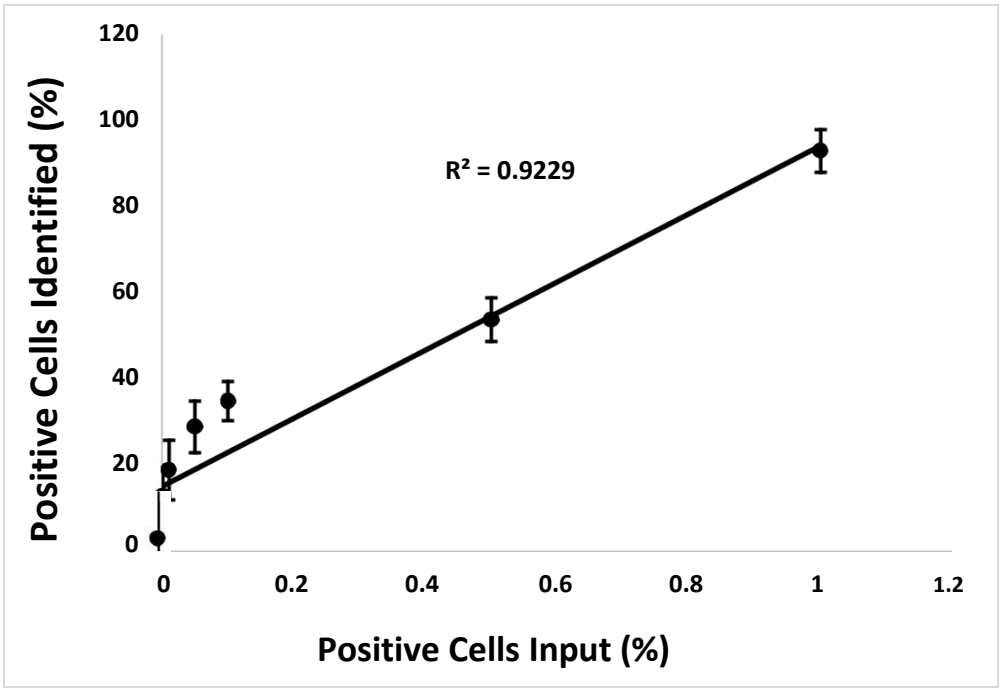


Figure 8 : Analyse cytométrique de flux de la reconnaissance des cellules cancéreuses du sein dans des échantillons de lyses de sang entier spiked par des aptamères MAMB1 et MAMA2 marqués par 6-FAM. A : Un graphique du pourcentage de cellules positives (MCF7) spiked (axe X) par rapport au pourcentage de cellules positives identifiées (axe Y). B : Un graphique du pourcentage de cellules positives (MDA-MB-415) pointu (axe X) par rapport au pourcentage de cellules positives identifiées (axe Y). MCF7 et MDA-MB-415 ont été additionnés de lysat de sang total contenant des PBMC en différents pourcentages (0%, 0.01%, 0.1%, 1%, 5%, 10%, 50% et 100%) et lus par cytométrie en flux.

Liaison MAMB1 et MAMA2 aux cellules cancéreuses du sein utilisant une MTC

MAMB1 et MAMA2 ont montré une grande affinité pour leurs cellules de cancer du sein (MCF7 et MDA-MB-415 respectivement) par rapport aux cellules de contrôle (MCF10A) en utilisant le TCM (Figure 9 et Figure 10). En outre, les deux aptamères pourraient reconnaître aussi peu que 10 cellules cancéreuses comme le montrent les valeurs de l'amplitude THz (Figure 11). Cependant, MAMB1 est plus sélectif pour ses cellules de cancer du sein cible MCF7 que MAMA2 à MDA-MB-415.

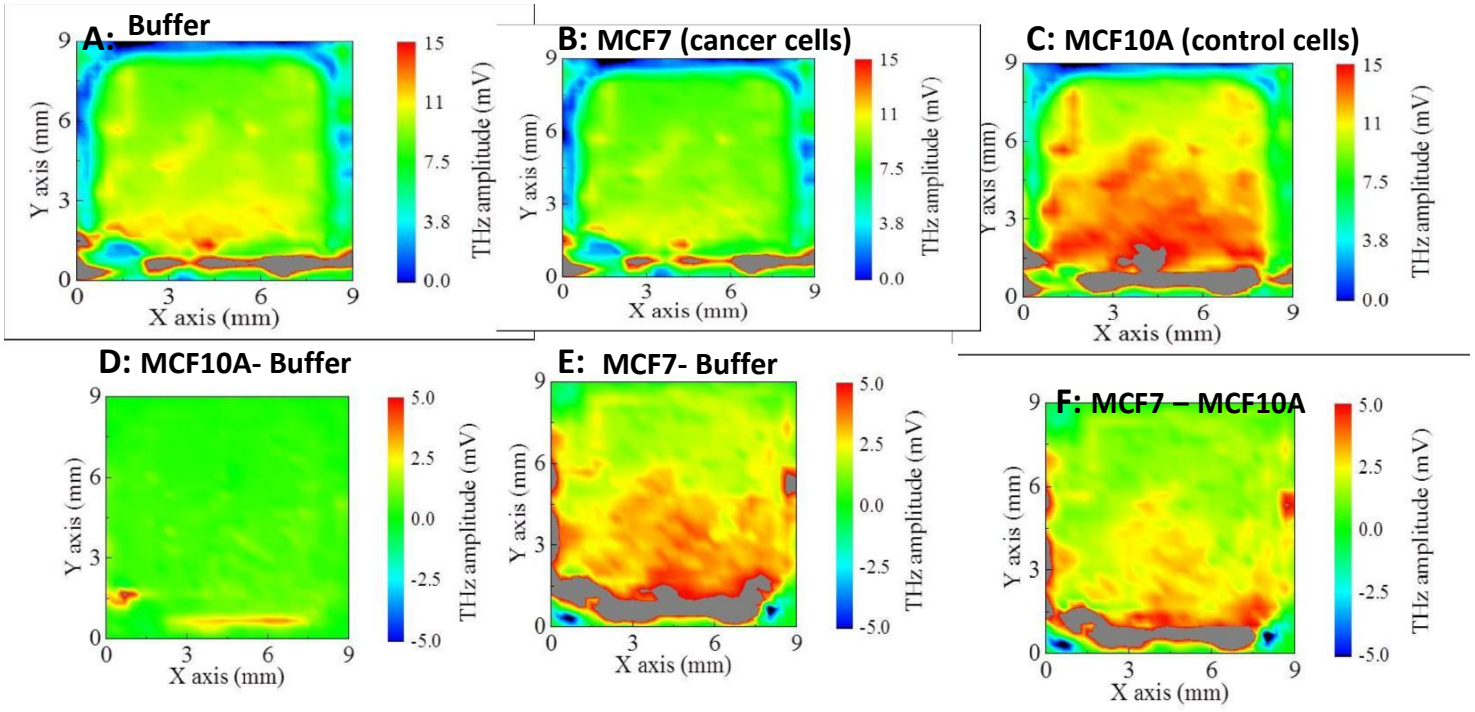


Figure 9: Mappage d'amplitude crête de l'impulsion THz pour A : tampon seul, B : MCF10A (cellules de contrôle normales), C : cellules MCf7 (cellules de cancer du sein), D, E, F: images différentielles après soustraction des cellules tampon et témoins. Les images ont été créées à l'aide du logiciel Origin.

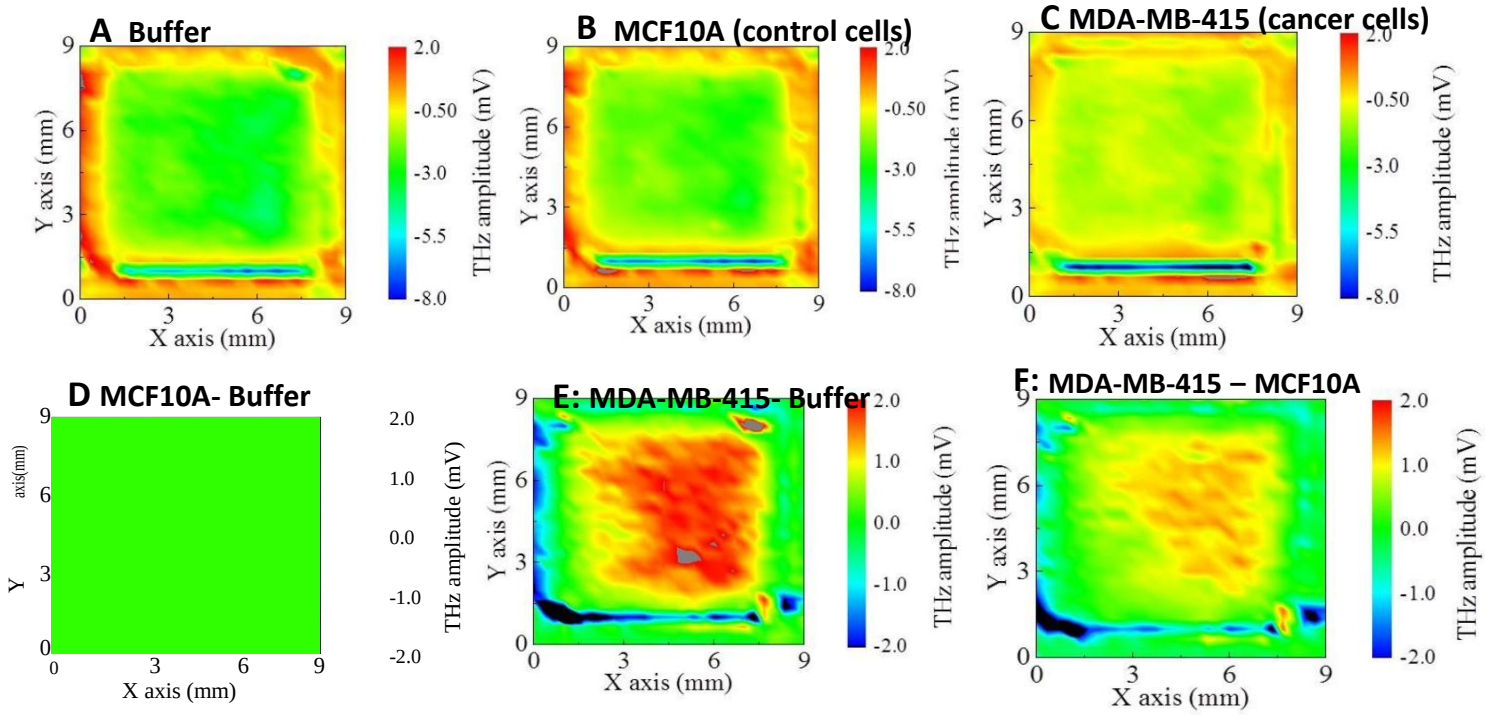


Figure 10: Mappage d'amplitude crête de l'impulsion THz pour A : tampon seu,l B : MCF10A (cellules de contrôle normales), C: cellules MDA-MB415 (cellules de cancer du sein), D, E, F: images différentielles après soustraction des cellules tampon et témoins. Les images ont été créées à l'aide du logiciel Origin

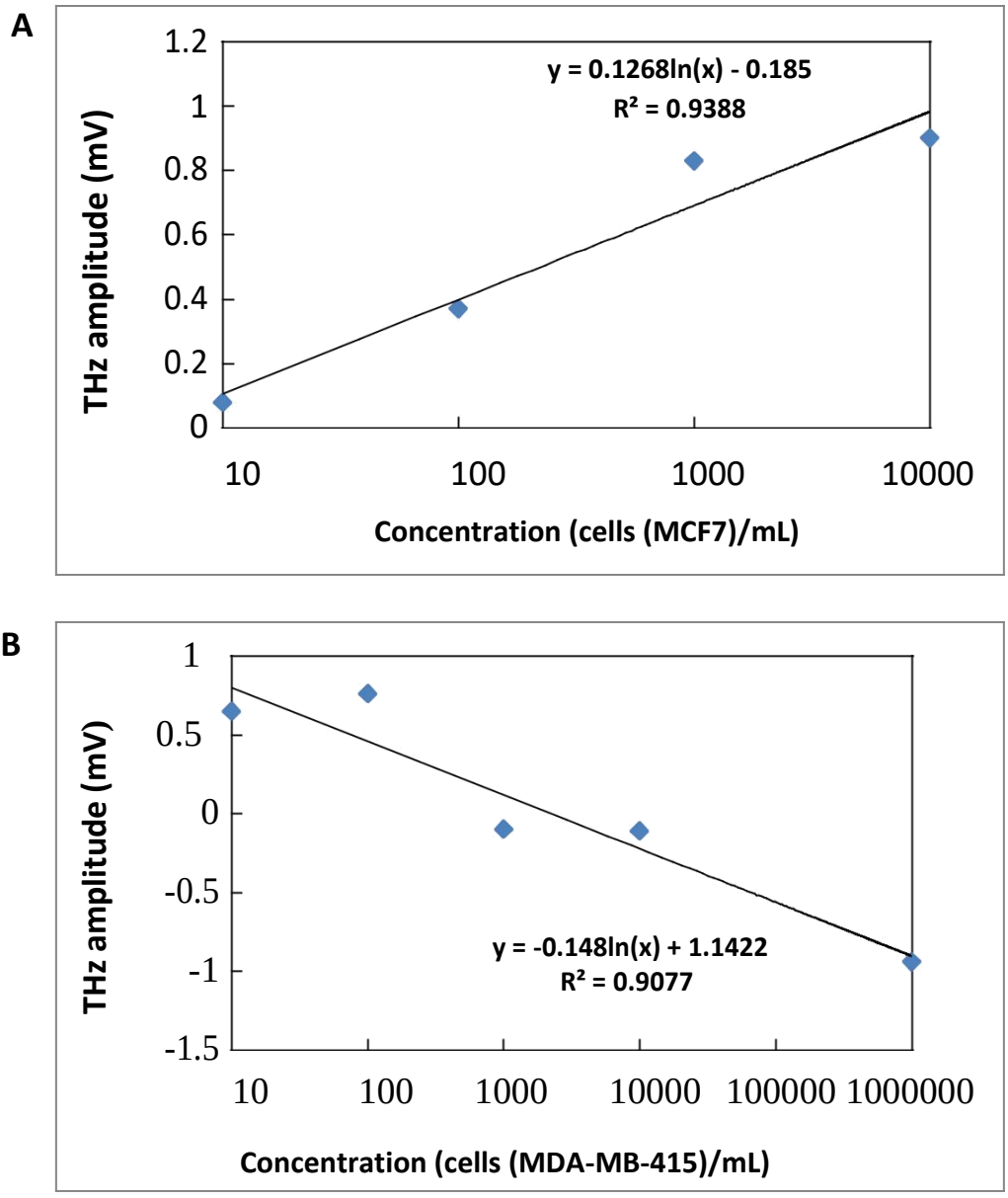


Figure 11: Mappage de l'amplitude du pic THz de différents nombres de cellules cancéreuses à la surface des plaques de détection. A-B : courbe de THz amplitude (mV) contre différents nombres de cellules cancéreuses MCF7 (A) et de cellules cancéreuses MDA-MB-415 (cellules 10, 1X102, 1X103, 1X104). Pour calculer l'amplitude THz (mV), une zone moyenne de 7 mm a été créée à la fois sur l'axe X et sur l'axe des Y. Les images ont été créées à l'aide du logiciel Origin

4. Discussion

Dans cette étude, MAMB1 et MAMA2, deux aptamères contre les protéines MGB2 et MGB1, respectivement, ont été sélectionnés avec succès. Les deux aptamères ont montré une affinité élevée à leurs lignées de cancer du sein cible MCF7 et MDA-MB-415 respectivement comme indiquées par leurs valeurs de kd (tableau 2). En outre, les deux aptamères ne présentaient aucune affinité aux lignées de cellules cancéreuses normales ou autres, ce qui indique une sélectivité élevée pour leurs cellules cancéreuses cérébrales (Figure 7). Les tests supplémentaires des aptamères dans le plasma et le lysat sanguin complet ont montré une spécificité élevée de l'aptamère à leurs cellules cibles par rapport aux cellules normales (témoins). Afin de simuler l'environnement de CTC, chaque cellule de cancer du sein MCF7 et MDA-MB-415 a été enrichie avec des pourcentages différents dans le sang contenant des cellules sanguines normales et les deux aptamères ont été testés pour leur liaison aux cellules cancéreuses du sein. Les résultats ont montré une augmentation de l'intensité de fluorescence des cellules détectées en augmentant le pourcentage de cellules poinçonnées pour les deux aptamères. La MAMA2 a montré moins d'affinité avec ses cellules cibles du cancer du sein comparées à MAMA2 indiquant que MAMA2 pourrait nécessiter certaines modifications (chimiquement) pour améliorer son affinité de liaison. La même tendance des résultats a été observée lorsque le TCM a été utilisée pour tester la liaison des deux aptamères à leurs cellules cibles. MAMA2 a montré moins de sélectivité et d'affinité pour MDA-MB415. Il se pourrait que MAMA2 fonctionne mieux en dehors des environnements en solution.

5. Conclusion et perspectives futures

Les aptamères pour la mammaglobine B et la mammaglobine A ont été sélectionnés pour la première fois dans cette étude. Les deux aptamères ont montré un grand potentiel d'utilisation pour leur liaison aux cellules cancéreuses du sein dans un mélange de cellules dans le lysat du sang complet. En outre, ils ont montré une liaison aux cellules de cancer du sein en utilisant la technologie TCM qui a été appliquée ici pour la première fois. Cependant, des investigation supplémentaires sont nécessaires, comme la limite de détection des deux aptamères par rapport
à la liaison aux cellules cancéreuses du sein dans le lysat du sang entier. En outre, étudier la

stabilité des deux aptamères dans le sang serait très avantageux si ces aptamères étaient utilisés dans le diagnostic de cancer du sein ou de CTC. En plus de cela en utilisant les deux aptamères dans un essai immunohistochimique comme un test pour détecter l'origine du cancer, en particulier pour la cible MGB2, comme étant seulement surexprimé dans le tissu mammaire. D'autres travaux pourraient également être réalisés en utilisant le TCM. Par exemple, d'autres cancers et des cellules normales pourraient être utilisées pour déterminer la spécificité des deux aptamères. Une modification des deux aptamères pour améliorer leur affinité de liaison est très avantageuse.